Quasi-likelihood analysis and Bayes-type estimators of an ergodic diffusion plus noise

Shogo H. Nakakita[1] · Yusuke Kaino[1] · Masayuki Uchida[2]

Abstract

We consider adaptive maximum-likelihood-type estimators and adaptive Bayes-type ones for discretely observed ergodic diffusion processes with observation noise whose variance is constant. The quasi-likelihood functions for the diffusion and drift parameters are introduced and the polynomial-type large deviation inequalities for those quasi-likelihoods are shown to see the asymptotic properties of the adaptive Bayes-type estimators and the convergence of moments for both adaptive maximum-likelihood-type estimators and adaptive Bayes-type ones.

Keywords Bayes-type estimation · Convergence of moments · Diffusion processes · Observation noise · Quasi-likelihood analysis · Stochastic differential equations

1 Introduction

We consider a d-dimensional ergodic diffusion process defined by the following stochastic differential equation such that

$$\mathrm{d}X_t = b(X_t, \beta)\,\mathrm{d}t + a(X_t, \alpha)\,\mathrm{d}w_t, \ X_0 = x_0,$$

where $\{w_t\}_{t \geq 0}$ is an r-dimensional Wiener process, x_0 is a random variable independent of $\{w_t\}_{t \geq 0}$, $\alpha \in \Theta_1$ and $\beta \in \Theta_2$ are unknown parameters, $\Theta_1 \subset \mathbf{R}^{m_1}$ and $\Theta_2 \subset \mathbf{R}^{m_2}$ are bounded, open, and convex sets in $\mathbf{R}^{m_i}$ admitting Sobolev's inequalities for embedding $W^{1,p}(\Theta_i) \hookrightarrow C(\overline{\Theta_i})$ for $i = 1, 2$, $\theta^\star = (\alpha^\star, \beta^\star)$ is the true value of the parameter, and $a : \mathbf{R}^d \times \Theta_1 \to \mathbf{R}^d \otimes \mathbf{R}^r$ and $b : \mathbf{R}^d \times \Theta_2 \to \mathbf{R}^d$ are known functions.

✉ Shogo H. Nakakita
  nakakita@sigmath.es.osaka-u.ac.jp

[1] Graduate School of Engineering Science, Osaka University, 1-3, Machikaneyama, Toyonaka, Osaka 560-8531, Japan

[2] Graduate School of Engineering Science, Center for Mathematical Modeling and Data Science, Osaka University, 1-3, Machikaneyama, Toyonaka, Osaka 560-8531, Japan

A matter of interest is to estimate the parameter $\theta = (\alpha, \beta)$ with partial and indirect observation of $\{X_t\}_{t \geq 0}$: the observation is discretised and contaminated by exogenous noise. The sequence of observation $\{Y_{ih_n}\}_{i=0,\ldots,n}$, which our parametric estimation is based on, is defined as

$$Y_{ih_n} = X_{ih_n} + \Lambda^{1/2} \varepsilon_{ih_n}, \quad i = 0, \ldots, n,$$

where $h_n > 0$ is the discretisation step such that $h_n \to 0$ and $T_n = nh_n \to \infty$, $\{\varepsilon_{ih_n}\}_{i=0,\ldots,n}$ is an i.i.d. sequence of random variables independent of $\{w_t\}_{t \geq 0}$ and x_0 such that $\mathbf{E}_{\theta^\star}[\varepsilon_{ih_n}] = 0$ and $\mathrm{Var}_{\theta^\star}(\varepsilon_{ih_n}) = I_d$ where I_m is the identity matrix in $\mathbf{R}^m \otimes \mathbf{R}^m$ for every $m \in \mathbf{N}$, and $\Lambda \in \mathbf{R}^d \otimes \mathbf{R}^d$ is a positive semi-definite matrix which is the variance of noise term. We also assume that the half vectorisation of Λ has bounded, open, and convex parameter space Θ_ε, and let us denote $\Xi := \Theta_\varepsilon \times \Theta_1 \times \Theta_2$. We also notate the true parameter of Λ as $\Lambda_\star$, its half vectorisation as $\theta_\varepsilon^\star = \mathrm{vech}\Lambda_\star$, and $\vartheta^\star = (\theta_\varepsilon^\star, \alpha^\star, \beta^\star)$. That is to say, our interest is parametric inference for an ergodic diffusion with long-term and high-frequency noised observation. One of the concrete examples is the wind velocity data provided by NWTC Information Portal (2018) whose observation is contaminated by exogenous noise with statistical significance according to the test for noise detection (Nakakita and Uchida 2019a) (Fig. 1).

As the existent discussion, Nakakita and Uchida (2019a) propose the following estimators $\hat{\Lambda}_n$, $\hat{\alpha}_n$, and $\hat{\beta}_n$ such that

$$\hat{\Lambda}_n = \frac{1}{2n} \sum_{i=0}^{n-1} \left(Y_{(i+1)h_n} - Y_{ih_n} \right)^{\otimes 2},$$

$$\mathbb{H}_{1,n}^\tau \left(\hat{\alpha}_n; \hat{\Lambda}_n \right) = \sup_{\alpha \in \Theta_1} \mathbb{H}_{1,n}^\tau \left(\alpha; \hat{\Lambda}_n \right),$$

$$\mathbb{H}_{2,n} \left(\hat{\beta}_n; \hat{\alpha}_n \right) = \sup_{\beta \in \Theta_2} \mathbb{H}_{2,n} \left(\beta; \hat{\alpha}_n \right),$$

where for every matrix A, A^{T} is the transpose of A and $A^{\otimes 2} = AA^{\mathrm{T}}$, $\mathbb{H}_{1,n}^\tau$ and $\mathbb{H}_{2,n}$ are the adaptive quasi-likelihood functions of α and β, respectively, defined in Sect. 3, and $\tau \in (1, 2]$ is a tuning parameter, and Nakakita and Uchida (2019a) show that these estimators are asymptotically normal and particularly the drift one is asymptotically efficient. To obtain the convergence rates of the estimators, it is necessary to see the asymptotic properties of the quasi-likelihood functions. Both of them are functions of local means of observation defined as

$$\bar{Y}_j = \frac{1}{p_n} \sum_{i=0}^{p_n-1} Y_{j\Delta_n + ih_n}, \quad j = 0, \ldots, k_n - 1,$$

where k_n is the number of partitions given for observation, p_n is that of observation in each partition, and $\Delta_n = p_n h_n$ is the time interval which each partition has, and note that these parameters have the properties $k_n \to \infty$, $p_n \to \infty$ and $\Delta_n \to 0$. Intuitively speaking, k_n and Δ_n correspond to n and h_n in the observation scheme
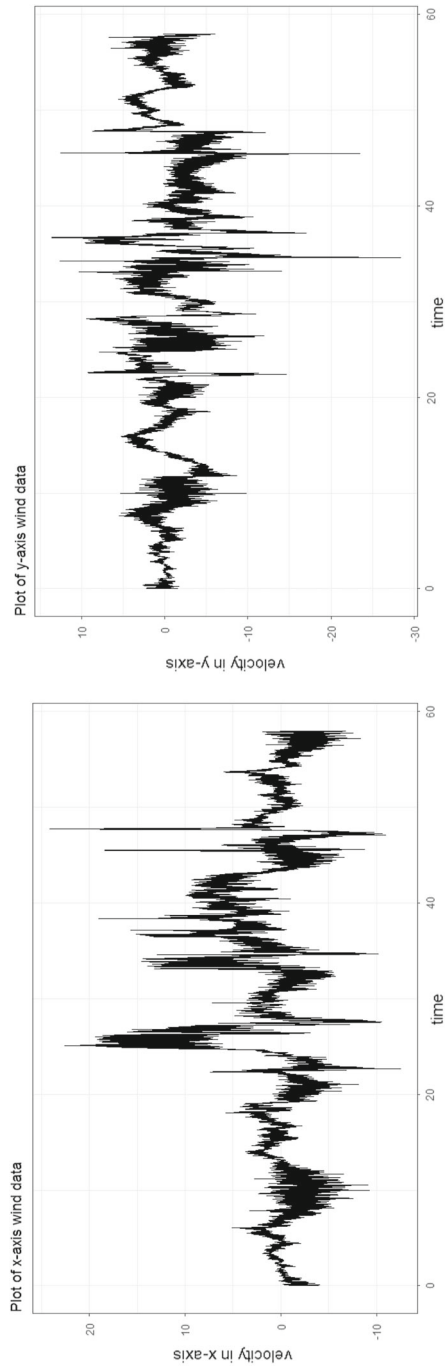
Fig. 1 Plot of wind velocity labelled Sonic x (left) and y (right) (119M) at the M5 tower from 00:00:00 on 1 July 2017 to 20:00:00 on 5 July 2017 with 0.05-s resolution (NWTC Information Portal 2018)

without exogenous noise, and divergence of p_n works to eliminate the influence of noise by law of large numbers. Hence, it should be also easy to understand that we have the asymptotic normality with the convergence rates $\sqrt{k_n}$ and $\sqrt{T_n}$ for α and β; that is,

$$\left[\sqrt{k_n}\left(\hat{\alpha}_n - \alpha^\star\right), \sqrt{T_n}\left(\hat{\beta}_n - \beta^\star\right)\right] \to^d \xi,$$

where ξ is an $(m_1 + m_2)$-dimensional Gaussian distribution with zero-mean.

The statistical inference for diffusion processes with discretised observation has been investigated for several decades, see Florens-Zmirou (1989), Yoshida (1992), Bibby and Sørensen (1995), and Kessler (1995, 1997). In practice, it is necessary to argue whether exogenous noise exists in observation, and it has been pointed out that the observational noise, known as microstructure noise, certainly exists in high-frequency financial data which is one of the major disciplines where statistics for diffusion processes is applied. Inference for diffusions under the noisy and discretised observation in fixed time interval [0, 1] is discussed by Jacod et al. (2009). Favetto (2014, 2016) examines the same model as our study and shows that simultaneous maximum-likelihood-type (ML-type) estimation has consistency under the situation where the variance of noise is unknown and asymptotic normality under the situation where the variance is known. As mentioned above, Nakakita and Uchida (2019a) propose adaptive ML-type estimation which has asymptotic normality even if we do not know the variance of noise, and test for noise detection which succeeds in showing the real example of the data (NWTC Information Portal 2018) contaminated by observational noise.

Our study aims to obtain polynomial-type large deviation inequalities for statistical random fields and to construct the adaptive Bayes-type estimators of both drift and diffusion parameters. Moreover, it is shown that both the adaptive ML-type estimators proposed in Nakakita and Uchida (2019a) and the adaptive Bayes-type estimators have not only asymptotic normality but also a certain type of convergence of moments. It is well known that asymptotic normality is one of the hopeful properties that estimators are expected to have; for instance, Nakakita and Uchida (2019b) utilise this result to compose likelihood-ratio-type statistics and related ones for parametric test and prove the convergence in distribution to a χ^2-distribution under null hypothesis and consistency of the test under alternative one. However, it is also known that asymptotic normality is not sufficient to develop some discussion requiring convergence of moments such as information criterion. In concrete terms, it is necessary to show the convergence of moments such that for every $f \in C(\mathbf{R}^{m_1} \times \mathbf{R}^{m_2})$ with at most polynomial growth and adaptive ML-type estimators $\hat{\alpha}_n$ and $\hat{\beta}_n$,

$$\mathbf{E}_{\vartheta^\star}\left[f\left(\sqrt{k_n}\left(\hat{\alpha}_n - \alpha^\star\right), \sqrt{T_n}\left(\hat{\beta}_n - \beta^\star\right)\right)\right] \to \mathbf{E}_{\theta^\star}\left[f\left(\xi\right)\right].$$

This is a stronger property than mere asymptotic normality since if we take f as a bounded and continuous function, then indeed asymptotic normality follows.

To see the asymptotic properties of the adaptive Bayes-type estimators and the convergence of moments for adaptive ML-type estimator, we can utilise polynomial-type

large deviation inequalities (PLDI) and quasi-likelihood analysis (QLA) proposed by Yoshida (2011) which have been widely used to discuss the asymptotic properties of the Bayes-type estimators and the convergence of moments of both the ML-type estimation and the Bayes-type one in statistical inference for continuous-time stochastic processes. This approach is developed from the exponential-type large deviation and likelihood analysis introduced by Ibragimov and Has'minskii (1972, 1973, 1981) and Kutoyants (1984, 1994, 2004) for continuously observed stochastic processes. Yoshida (2011) also discusses the asymptotic properties of the simultaneous Bayes-type estimators and adaptive Bayes-type ones and convergence of moments for both the adaptive ML-type estimation and the Bayes-type estimation for ergodic diffusions with $nh_n \to \infty$ and $nh_n^2 \to 0$. Uchida and Yoshida (2012, 2014) examine the same problem for adaptive ML-type and adaptive Bayes-type estimation of ergodic diffusions with more relaxed condition: $nh_n \to \infty$ and $nh_n^p \to 0$ for some $p \geq 2$. Ogihara and Yoshida (2011) study the asymptotic properties of the adaptive Bayes-type estimators and the convergence of moments for both the ML-type estimators and the Bayes-type estimators of ergodic jump-diffusion processes in the scheme of $nh_n \to \infty$ and $nh_n^2 \to 0$. For the Bayes-type estimation for diffusion-type processes in a general setting, see Ogihara (2018, 2019). Clinet and Yoshida (2017) show PLDI for the quasi-likelihood function for ergodic point processes and the convergence of moments for the corresponding ML-type and Bayes-type estimators. From the viewpoint of computational statistics, it is crucial to get the Bayes-type estimation for diffusion-type processes from high-frequency data. In particular, Bayesian statistics work well for the estimation of unknown nonlinear parameters based on the multimodal quasi-likelihood functions for diffusion-type processes. In order to obtain the ML-type estimators for diffusion-type processes, the hybrid-type estimators with initial Bayes-type estimators are studied by Kamatani and Uchida (2015) and Kaino and Uchida (2018a, b). For the convergence of moments for Z-estimators of ergodic diffusion processes, see Negri and Nishiyama (2017). As the applications of the convergence of moments for estimators of stochastic differential equations, Uchida (2010) composes AIC-type information criterion for ergodic diffusion processes, and Eguchi and Masuda (2018) propose BIC-type one for local asymptotic quadratic statistical experiments including some schemes for diffusion processes. For the convergence of moments of regularized estimators for a discretely observed ergodic diffusion process, see Masuda and Shimizu (2017).

The composition of the paper is as follows: in Sect. 2, we set some notations and assumptions; Sect. 3 proposes the adaptive Bayes-type estimators of both drift and diffusion parameters and gives the main results of our study such as QLA for our ergodic diffusion plus noise model and the convergence of moments of both the adaptive ML-type estimators and the adaptive Bayes-type estimators for the model; Sect. 4 studies a concrete example of adaptive Bayes-type estimators for ergodic diffusion plus noise model and shows the result of computational simulation; Sect. 5 provides the technical proofs for the main results shown in Sect. 3.

2 Notation and assumption

We set the following notations.

– For every matrix A, A^{T} is the transpose of A, and $A^{\otimes 2} := AA^{\mathrm{T}}$.
– For every set of matrices A and B whose dimensions coincide, $A[B] := \mathrm{tr}\left(AB^{\mathrm{T}}\right)$. Moreover, for any $m \in \mathbf{N}$, $A \in \mathbf{R}^m \otimes \mathbf{R}^m$ and $u, v \in \mathbf{R}^m$, $A[u, v] := v^{\mathrm{T}} A u$.
– Let us denote the ℓth element of any vector v as $v^{(\ell)}$ and (ℓ_1, ℓ_2)th one of any matrix A as $A^{(\ell_1, \ell_2)}$.
– For any vector v and any matrix A, $|v| := \sqrt{\mathrm{tr}\left(v^{\mathrm{T}} v\right)}$ and $\|A\| := \sqrt{\mathrm{tr}\left(A^{\mathrm{T}} A\right)}$.
– For every $p > 0$, $\|\cdot\|_p$ is the L^P ($P_{\theta^\star}$)-norm.
– $A(x, \alpha) := a(x, \alpha)^{\otimes 2}$, $a(x) := a(x, \alpha^\star)$, $A(x) := A(x, \alpha^\star)$ and $b(x) := b(x, \beta^\star)$.
– For given $\tau \in (1, 2]$, $p_n := h_n^{-1/\tau}$, $\Delta_n := p_n h_n$, and $k_n := n/p_n$, and we define the sequence of local means such that

$$\bar{Z}_j = \frac{1}{p_n} \sum_{i=0}^{p_n - 1} Z_{j\Delta_n + ih_n}, \; j = 0, \ldots, k_n - 1,$$

where $\left\{Z_{ih_n}\right\}_{i=0,\ldots,n}$ indicates an arbitrary sequence defined on the mesh $\{ih_n\}_{i=0,\ldots,n}$ such as $\left\{Y_{ih_n}\right\}_{i=0,\ldots,n}$, $\left\{X_{ih_n}\right\}_{i=0,\ldots,n}$ and $\left\{\varepsilon_{ih_n}\right\}_{i=0,\ldots,n}$.

Remark 1 Since the observation is masked by the exogenous noise, it should be transformed to obtain the undermined process $\{X_t\}_{t\geq 0}$. As illustrated by Nakakita and Uchida (2019a), the sequence $\left\{\bar{Y}_j\right\}_{j=0,\ldots,k_n-1}$ can extract the state of the latent process $\{X_t\}_{t\geq 0}$ in the sense of the statement of Lemma 2.

– $\mathcal{G}_t := \sigma(x_0, w_s : s \leq t)$, $\mathcal{G}_{j,i}^n := \mathcal{G}_{j\Delta_n + ih_n}$, $\mathcal{G}_j^n := \mathcal{G}_{j,0}^n$, $\mathcal{A}_{j,i}^n := \sigma(\varepsilon_{\ell h_n} : \ell \leq j p_n + i - 1)$, $\mathcal{A}_j^n := \mathcal{A}_{j,0}^n$, $\mathcal{H}_{j,i}^n := \mathcal{G}_{j,i}^n \vee \mathcal{A}_{j,i}^n$ and $\mathcal{H}_j^n := \mathcal{H}_{j,0}^n$.
– We define the real-valued function as for $l_1, l_2, l_3, l_4 = 1, \ldots, d$:

$$V((l_1, l_2), (l_3, l_4))$$
$$:= \sum_{k=1}^d \left(\Lambda_\star^{1/2}\right)^{(l_1, k)} \left(\Lambda_\star^{1/2}\right)^{(l_2, k)} \left(\Lambda_\star^{1/2}\right)^{(l_3, k)} \left(\Lambda_\star^{1/2}\right)^{(l_4, k)} \left(\mathbf{E}_{\theta^\star}\left[\left|\epsilon_0^{(k)}\right|^4\right] - 3\right)$$
$$+ \frac{3}{2}\left(\Lambda_\star^{(l_1, l_3)}\Lambda_\star^{(l_2, l_4)} + \Lambda_\star^{(l_1, l_4)}\Lambda_\star^{(l_2, l_3)}\right),$$

and with the function σ as for $i = 1, \ldots, d$ and $j = i, \ldots, d$,

$$\sigma(i, j) := \begin{cases} j & \text{if } i = 1, \\ \sum_{\ell=1}^{i-1}(d - \ell + 1) + j - i + 1 & \text{if } i > 1, \end{cases}$$

we define the matrix W_1 as for $i_1, i_2 = 1, \ldots, d(d+1)/2$,

$$W_1^{(i_1, i_2)} := V\left(\sigma^{-1}(i_1), \sigma^{-1}(i_2)\right).$$

– Let

$$\left\{ B_\kappa(x) \,\middle|\, \kappa = 1, \ldots, m_1, \; B_\kappa = (B_\kappa^{(j_1, j_2)})_{j_1, j_2} \right\},$$

$$\left\{ f_\lambda(x) \,\middle|\, \lambda = 1, \ldots, m_2, \; f_\lambda = (f_\lambda^{(1)}, \ldots, f_\lambda^{(d)}) \right\}$$

be sequences of $\mathbf{R}^d \otimes \mathbf{R}^d$-valued functions and $\mathbf{R}^d$-valued ones, respectively, such that the components of themselves and their derivative with respect to x are polynomial growth functions for all κ and λ. Then, we define the following matrix-valued functionals, for $\bar{B}_\kappa := \frac{1}{2}\left(B_\kappa + B_\kappa^{\mathrm{T}}\right)$,

$$\left(W_2^{(\tau)}\left(\{B_\kappa : \kappa = 1, \ldots, m_1\}\right)\right)^{(\kappa_1, \kappa_2)}$$

$$:= \begin{cases} \nu\left(\mathrm{tr}\left\{\left(\bar{B}_{\kappa_1} A \bar{B}_{\kappa_2} A\right)(\cdot)\right\}\right) & \text{if } \tau \in (1, 2), \\ \nu\left(\mathrm{tr}\left\{\left(\bar{B}_{\kappa_1} A \bar{B}_{\kappa_2} A + 4\bar{B}_{\kappa_1} A \bar{B}_{\kappa_2} \Lambda_\star + 12\bar{B}_{\kappa_1} \Lambda_\star \bar{B}_{\kappa_2} \Lambda_\star\right)(\cdot)\right\}\right) & \text{if } \tau = 2, \end{cases}$$

$$\left(W_3\left(\{f_\lambda : \lambda = 1, \ldots, m_2\}\right)\right)^{(\lambda_1, \lambda_2)}$$

$$:= \nu\left(\left(f_{\lambda_1} A \left(f_{\lambda_2}\right)^{\mathrm{T}}\right)(\cdot)\right),$$

where $\nu = \nu_{\theta^\star}$ is the invariant measure of X_t discussed in assumption [A1]—(iv), and for all function f on $\mathbf{R}^d$, $\nu\left(f\left(\cdot\right)\right) := \int_{\mathbf{R}^d} f(x)\, \nu(\mathrm{d}x)$.

With respect to X_t, we assume the following conditions.

[A1] (i) $\inf_{x, \alpha} \det A(x, \alpha) > 0$.

(ii) For some constant C, for all $x_1, x_2 \in \mathbf{R}^d$,

$$\sup_{\alpha \in \Theta_1} \|a(x_1, \alpha) - a(x_2, \alpha)\| + \sup_{\beta \in \Theta_2} |b(x_1, \beta) - b(x_2, \beta)| \le C |x_1 - x_2|$$

(iii) For all $p \ge 0$, $\sup_{t \ge 0} \mathbf{E}_{\theta^\star}\left[|X_t|^p\right] < \infty$.

(iv) There exists an unique invariant measure $\nu = \nu_0$ on $\left(\mathbf{R}^d, \mathcal{B}\left(\mathbf{R}^d\right)\right)$ and for all $p \ge 1$ and $f \in L^p(\nu)$ with polynomial growth,

$$\frac{1}{T} \int_0^T f(X_t)\, \mathrm{d}t \to^P \int_{\mathbf{R}^d} f(x)\, \nu(\mathrm{d}x).$$

(v) For any polynomial growth function $g : \mathbf{R}^d \to \mathbf{R}$ satisfying $\int_{\mathbf{R}^d} g(x)\, \nu(\mathrm{d}x) = 0$, there exist $G(x)$, $\partial_{x^{(i)}} G(x)$ with at most polynomial growth for $i = 1, \ldots, d$ such that for all $x \in \mathbf{R}^d$,

$$L_{\theta^\star} G(x) = -g(x),$$

where $L_{\theta^\star}$ is the infinitesimal generator of X_t.

Remark 2 Pardoux and Veretennikov (2001) show a sufficient condition for [A1]-(v). Uchida and Yoshida (2012) also introduce the sufficient condition for [A1]—(iii)–(v) assuming [A1]—(i)-(ii), $\sup_{x,\alpha} A(x,\alpha) < \infty$ and $\exists c_0 > 0$, $M_0 > 0$ and $\gamma \geq 0$ such that for all $\beta \in \Theta_2$ and $x \in \mathbf{R}^d$ satisfying $|x| \geq M_0$,

$$\frac{1}{|x|} x^\mathsf{T} b(x, \beta) \leq -c_0 |x|^\gamma .$$

[A2] There exists $C > 0$ such that $a : \mathbf{R}^d \times \Theta_1 \to \mathbf{R}^d \otimes \mathbf{R}^r$ and $b : \mathbf{R}^d \times \Theta_2 \to \mathbf{R}^d$ have continuous derivatives satisfying

$$\sup_{\alpha \in \Theta_1} \left| \partial_x^j \partial_\alpha^i a(x,\alpha) \right| \leq C(1 + |x|)^C , \ 0 \leq i \leq 4, \ 0 \leq j \leq 2,$$

$$\sup_{\beta \in \Theta_2} \left| \partial_x^j \partial_\beta^i b(x,\beta) \right| \leq C(1 + |x|)^C , \ 0 \leq i \leq 4, \ 0 \leq j \leq 2.$$

With the invariant measure ν, we define

$$\mathbb{Y}_1^\tau (\alpha; \vartheta^\star) := -\frac{1}{2} \int \left\{ \mathrm{tr} \left(A^\tau (x, \alpha, \Lambda_\star)^{-1} A^\tau (x, \alpha^\star, \Lambda_\star) - I_d \right) + \log \frac{\det A^\tau (x, \alpha, \Lambda_\star)}{\det A^\tau (x, \alpha^\star, \Lambda_\star)} \right\} \nu (\mathrm{d}x),$$

$$\mathbb{Y}_2 (\beta; \vartheta^\star) := -\frac{1}{2} \int A(x, \alpha^\star)^{-1} \left[(b(x, \beta) - b(x, \beta^\star))^{\otimes 2} \right] \nu (\mathrm{d}x),$$

where $A^\tau (x, \alpha, \Lambda) := A(x, \alpha) + 3\Lambda \mathbf{1}_{\{2\}}(\tau)$. For these functions, let us assume the following identifiability conditions hold.

[A3] For all $\tau \in (1, 2]$, there exists a constant $\chi(\alpha^\star) > 0$ such that $\mathbb{Y}_1^\tau (\alpha; \theta^\star) \leq -\chi(\theta^\star) |\alpha - \alpha^\star|^2$ for all $\alpha \in \Theta_1$.

[A4] For all $\tau \in (1, 2]$, there exists a constant $\chi'(\beta^\star) > 0$ such that $\mathbb{Y}_2 (\beta; \theta^\star) \leq -\chi'(\theta^\star) |\beta - \beta^\star|^2$ for all $\beta \in \Theta_2$.

The next assumption is with respect to the moments of noise.

[A5] For any $k > 0$, ε_{ih_n} has kth moment and the components of ε_{ih_n} are independent of the other components for all i, $\{w_t\}_{t \geq 0}$ and x_0. In addition, for all odd integer k, $i = 0, \ldots, n$, $n \in \mathbf{N}$, and $\ell = 1, \ldots, d$, $\mathbf{E}_{\theta^\star} \left[\left(\varepsilon_{ih_n}^{(\ell)} \right)^k \right] = 0$, and $\mathbf{E}_{\theta^\star} \left[\varepsilon_{ih_n}^{\otimes 2} \right] = I_d$.

The following assumption determines the balance of convergence or divergence of several parameters. Note that τ is a tuning parameter and hence we can control it arbitrarily in its space $(1, 2)$.

[A6] $p_n = h_n^{-1/\tau}$, $\tau \in (1, 2]$, $h_n \to 0$, $T_n = nh_n \to \infty$, $k_n = n/p_n \to \infty$, $k_n \Delta_n^2 \to 0$ for $\Delta_n := p_n h_n$. Furthermore, there exists $\epsilon_0 > 0$ such that $nh_n \geq k_n^{\epsilon_0}$ for sufficiently large n.

Remark 3 The tuning parameter τ controls the number of local means and that of samples used for one local mean. If we set τ to be close to 2, we can make $\sqrt{k_n}$ larger which is the convergence rate of the adaptive maximum-likelihood-type estimator $\hat{\alpha}_n$ and the adaptive Bayes-type one $\tilde{\alpha}_n$ defined later on. Note that the estimation with larger values of τ does not necessarily outperform that with smaller value of τ if the noise variance is large (e.g. see the simulation study in Nakakita and Uchida 2019a).

Remark 4 Let us denote $\epsilon_1 = \epsilon_0/2$ and $f \in \mathcal{C}^{1,1}\left(\mathbf{R}^d \times \Xi\right)$ where f and the components of their derivatives are polynomial growth with respect to x uniformly in $\vartheta \in \Xi$. Then, the discussion in Uchida (2010) verifies under [A1] and [A6], for all $M > 0$,

$$\sup_{n \in \mathbf{N}} \mathbf{E}_{\theta^\star}\left[\sup_{\vartheta \in \Xi}\left(k_n^{\epsilon_1}\left|\frac{1}{k_n}\sum_{j=1}^{k_n-2} f\left(X_{j\Delta_n}, \vartheta\right) - \int_{R^d} f\left(x, \vartheta\right)\nu\left(dx\right)\right|\right)^M\right] < \infty.$$

3 Quasi-likelihood analysis

First of all, we introduce and analyse some quasi-likelihood functions and estimators which are defined in Nakakita and Uchida (2019a). The quasi-likelihood functions for the diffusion parameter α and the drift one β using this sequence are as follows:

$$\mathbb{H}_{1,n}^{\tau}\left(\alpha; \Lambda\right) := -\frac{1}{2}\sum_{j=1}^{k_n-2}\left(\left(\frac{2}{3}\Delta_n A_n^{\tau}\left(\bar{Y}_{j-1}, \alpha, \Lambda\right)\right)^{-1}\left[\left(\bar{Y}_{j+1} - \bar{Y}_j\right)^{\otimes 2}\right]\right.$$

$$\left. + \log\det A_n^{\tau}\left(\bar{Y}_{j-1}, \alpha, \Lambda\right)\right),$$

$$\mathbb{H}_{2,n}\left(\beta; \alpha\right) := -\frac{1}{2}\sum_{j=1}^{k_n-2}\left(\left(\Delta_n A\left(\bar{Y}_{j-1}, \alpha\right)\right)^{-1}\left[\left(\bar{Y}_{j+1} - \bar{Y}_j - \Delta_n b\left(\bar{Y}_{j-1}, \beta\right)\right)^{\otimes 2}\right]\right),$$

where $A_n^{\tau}\left(x, \alpha, \Lambda\right) := A\left(x, \alpha\right) + 3\Delta_n^{\frac{2-\tau}{\tau-1}}\Lambda$. We set the adaptive ML-type estimators $\hat{\Lambda}_n$, $\hat{\alpha}_n$, and $\hat{\beta}_n$ such that

$$\hat{\Lambda}_n := \frac{1}{2n}\sum_{i=0}^{n-1}\left(Y_{(i+1)h_n} - Y_{ih_n}\right)^{\otimes 2},$$

$$\mathbb{H}_{1,n}^{\tau}\left(\hat{\alpha}_n; \hat{\Lambda}_n\right) = \sup_{\alpha \in \Theta_1}\mathbb{H}_{1,n}^{\tau}\left(\alpha; \hat{\Lambda}_n\right),$$

$$\mathbb{H}_{2,n}\left(\hat{\beta}_n; \hat{\alpha}_n\right) = \sup_{\beta \in \Theta_2}\mathbb{H}_{2,n}\left(\beta; \hat{\alpha}_n\right).$$

Assume that $\pi_\ell, \ell = 1, 2$ are continuous and $0 < \inf_{\theta_\ell \in \Theta_\ell}\pi_\ell\left(\theta_\ell\right) < \sup_{\theta_\ell \in \Theta_\ell}\pi_\ell\left(\theta_\ell\right) < \infty$ and denote the adaptive Bayes-type estimators

$$
\tilde{\alpha}_n := \left\{ \int_{\Theta_1} \exp\left(\mathbb{H}_{1,n}^{\tau}\left(\alpha; \hat{\Lambda}_n \right) \right) \pi_1(\alpha) \, d\alpha \right\}^{-1} \int_{\Theta_1} \alpha \exp\left(\mathbb{H}_{1,n}^{\tau}\left(\alpha; \hat{\Lambda}_n \right) \right) \pi_1(\alpha) \, d\alpha,
$$

$$
\tilde{\beta}_n := \left\{ \int_{\Theta_2} \exp\left(\mathbb{H}_{2,n}\left(\beta; \tilde{\alpha}_n \right) \right) \pi_2(\beta) \, d\beta \right\}^{-1} \int_{\Theta_2} \beta \exp\left(\mathbb{H}_{2,n}\left(\beta; \tilde{\alpha}_n \right) \right) \pi_2(\beta) \, d\beta.
$$

Our purpose is to show the polynomial-type large deviation inequalities for the quasi-likelihood functions defined above in the framework introduced by Yoshida (2011), and the convergences of moments for the adaptive ML-type estimators and the Bayes-type estimators as the application of them. Note that the asymptotic properties of the Bayes-type estimators are shown by using the polynomial-type large deviation inequality. Let us denote the following statistical random fields for $u_1 \in \mathbf{R}^{m_1}$ and $u_2 \in \mathbf{R}^{m_2}$

$$
\mathbb{Z}_{1,n}^{\tau}\left(u_1; \hat{\Lambda}_n, \alpha^{\star} \right) := \exp\left(\mathbb{H}_{1,n}^{\tau}\left(\alpha^{\star} + k_n^{-1/2} u_1; \hat{\Lambda}_n \right) - \mathbb{H}_{1,n}^{\tau}\left(\alpha^{\star}; \hat{\Lambda}_n \right) \right),
$$

$$
\mathbb{Z}_{2,n}^{\mathrm{ML}}\left(u_2; \hat{\alpha}_n, \beta^{\star} \right) := \exp\left(\mathbb{H}_{2,n}\left(\beta^{\star} + T_n^{-1/2} u_2; \hat{\alpha}_n \right) - \mathbb{H}_{2,n}\left(\beta^{\star}; \hat{\alpha}_n \right) \right),
$$

$$
\mathbb{Z}_{2,n}^{\mathrm{Bayes}}\left(u_2; \tilde{\alpha}_n, \beta^{\star} \right) := \exp\left(\mathbb{H}_{2,n}\left(\beta^{\star} + T_n^{-1/2} u_2; \tilde{\alpha}_n \right) - \mathbb{H}_{2,n}\left(\beta^{\star}; \tilde{\alpha}_n \right) \right),
$$

and some sets

$$
\mathbb{U}_{1,n}^{\tau}\left(\alpha^{\star} \right) := \left\{ u_1 \in \mathbf{R}^{m_1}; \alpha^{\star} + k_n^{-1/2} u_1 \in \Theta_1 \right\},
$$

$$
\mathbb{U}_{2,n}\left(\beta^{\star} \right) := \left\{ u_2 \in \mathbf{R}^{m_2}; \beta^{\star} + T_n^{-1/2} u_2 \in \Theta_2 \right\},
$$

and for $r \geq 0$,

$$
V_{1,n}^{\tau}\left(r, \alpha^{\star} \right) := \left\{ u_1 \in \mathbb{U}_{1,n}^{\tau}\left(\alpha^{\star} \right); r \leq |u_1| \right\},
$$

$$
V_{2,n}\left(r, \beta^{\star} \right) := \left\{ u_2 \in \mathbb{U}_{2,n}\left(\beta^{\star} \right); r \leq |u_2| \right\}.
$$

We use the notation as Nakakita and Uchida (2019a) for the information matrices

$$
\mathcal{I}^{\tau}\left(\vartheta^{\star} \right) := \mathrm{diag}\left\{ W_1, \mathcal{I}^{(2,2),\tau}, \mathcal{I}^{(3,3)} \right\}\left(\vartheta^{\star} \right),
$$

$$
\mathcal{J}^{\tau}\left(\vartheta^{\star} \right) := \mathrm{diag}\left\{ I_{d(d+1)/2}, \mathcal{J}^{(2,2),\tau}, \mathcal{J}^{(3,3)} \right\}\left(\vartheta^{\star} \right),
$$

where for $i_1, i_2 \in \{1, \ldots, m_1\}$,

$$
\mathcal{I}^{(2,2),\tau}(\vartheta^{\star}) := W_2^{(\tau)}\left(\left\{ \frac{3}{4}\left(A^{\tau} \right)^{-1}\left(\partial_{\alpha^{(k_1)}} A \right)\left(A^{\tau} \right)^{-1}\left(\cdot, \vartheta^{\star} \right) : k_1 = 1, \ldots, m_1 \right\} \right),
$$

$$
\mathcal{J}^{(2,2),\tau}(\vartheta^{\star}) := \left[\frac{1}{2} \nu\left(\mathrm{tr}\left\{ \left(A^{\tau} \right)^{-1}\left(\partial_{\alpha^{(i_1)}} A \right)\left(A^{\tau} \right)^{-1}\left(\partial_{\alpha^{(i_2)}} A \right) \right\}\left(\cdot, \vartheta^{\star} \right) \right) \right]_{i_1, i_2},
$$

and for $j_1, j_2 \in \{1, \ldots, m_2\}$,

$$\mathcal{I}^{(3,3)}(\theta^\star) = \mathcal{J}^{(3,3)}(\theta^\star) := \left[\nu \left((A)^{-1} \left[\partial_{\beta^{(j_1)}} b, \partial_{\beta^{(j_2)}} b \right] (\cdot, \theta^\star) \right) \right]_{j_1, j_2}.$$

We also denote $\hat{\theta}_{\varepsilon,n} := \mathrm{vech} \hat{\Lambda}_n$ and $\theta_\varepsilon^\star := \mathrm{vech} \Lambda_\star$.

Theorem 1 *Under [A1]–[A6], we have the following results.*

1. *The polynomial-type large deviation inequalities hold: for all $L > 0$, there exists a constant $C(L)$ such that for all $r > 0$,*

$$P_{\theta^\star} \left[\sup_{u_1 \in V_{1,n}^\tau(r, \alpha^\star)} \mathbb{Z}_{1,n}^\tau \left(u_1; \hat{\Lambda}_n, \alpha^\star \right) \geq e^{-r} \right] \leq \frac{C(L)}{r^L},$$

$$P_{\theta^\star} \left[\sup_{u_2 \in V_{2,n}(r, \beta^\star)} \mathbb{Z}_{2,n}^{\mathrm{ML}} \left(u_2; \hat{\alpha}_n, \beta^\star \right) \geq e^{-r} \right] \leq \frac{C(L)}{r^L},$$

$$P_{\theta^\star} \left[\sup_{u_2 \in V_{2,n}(r, \beta^\star)} \mathbb{Z}_{2,n}^{\mathrm{Bayes}} \left(u_2; \tilde{\alpha}_n, \beta^\star \right) \geq e^{-r} \right] \leq \frac{C(L)}{r^L}.$$

2. *The convergences of moment hold:*

$$\mathbf{E}_{\theta^\star} \left[f \left(\sqrt{n} \left(\hat{\theta}_{\varepsilon,n} - \theta_\varepsilon^\star \right), \sqrt{k_n} \left(\hat{\alpha}_n - \alpha^\star \right), \sqrt{T_n} \left(\hat{\beta}_n - \beta^\star \right) \right) \right] \to \mathbb{E} \left[f \left(\zeta_0, \zeta_1, \zeta_2 \right) \right],$$

$$\mathbf{E}_{\theta^\star} \left[f \left(\sqrt{n} \left(\hat{\theta}_{\varepsilon,n} - \theta_\varepsilon^\star \right), \sqrt{k_n} \left(\tilde{\alpha}_n - \alpha^\star \right), \sqrt{T_n} \left(\tilde{\beta}_n - \beta^\star \right) \right) \right] \to \mathbb{E} \left[f \left(\zeta_0, \zeta_1, \zeta_2 \right) \right],$$

where

$$(\zeta_0, \zeta_1, \zeta_2) \sim N_{d(d+1)/2+m_1+m_2} \left(\mathbf{0}, \left(\mathcal{J}^\tau (\vartheta^\star) \right)^{-1} \left(\mathcal{I}^\tau (\vartheta^\star) \right) \left(\mathcal{J}^\tau (\vartheta^\star) \right)^{-1} \right)$$

and f is an arbitrary continuous functions of at most polynomial growth.

Remark 5 It is worth noting that the adaptive Bayes-type estimators are newly proposed for the ergodic diffusion plus noise models in this study and their properties, which are not only asymptotic normality but also mere convergences of moments, are shown. As an application of the Bayes-type estimation, Kaino et al. (2018) study the hybrid estimators with initial Bayes-type estimators for our ergodic diffusion plus noise model and give an example and simulation results of the hybrid estimators.

4 Example and simulation results

We consider the two-dimensional diffusion process.

$$\mathrm{d}X_t = b(X_t, \beta)\mathrm{d}t + a(X_t, \alpha)\mathrm{d}w_t, \quad t \geq 0, \quad X_0 = (1, 1)^\star,$$

where

$$b(X_t, \beta) = \begin{pmatrix} -\beta_1 X_{t,1} - \beta_3 X_{t,2} + \beta_5 \\ -\beta_2 X_{t,1} - \beta_4 X_{t,2} + \beta_6 \end{pmatrix},$$

$$a(X_t, \alpha) = \begin{pmatrix} \alpha_1 & \alpha_2 \\ \alpha_2 & \alpha_3 \end{pmatrix}.$$

Moreover, the true parameter values are $(\alpha_1^*, \alpha_2^*, \alpha_3^*) = (1, 0, 2)$ and $(\beta_1^*, \beta_2^*, \beta_3^*, \beta_4^*, \beta_5^*, \beta_6^*) = (1, 0.1, 0.1, 1, 1, 1)$. The parameter space of α is $\Theta_1 = \{(\alpha_1, \alpha_2, \alpha_3) \in [0.1, 50] \times [-25, 25] \times [0.1, 50] | \alpha_1 \alpha_3 - \alpha_2^2 \geq 10^{-3}\}$, and the parameter space of β is $\Theta_2 = [-25, 25]^6$.

The noisy data $\{Y_{ih_n}\}_{i=0,\ldots,n}$ are defined as for all $i = 0, \ldots, n$,

$$Y_{ih_n} := X_{ih_n} + \Lambda^{1/2} \varepsilon_{ih_n},$$

where $n = 10^6$, $h_n = 6.309573 \times 10^{-5}$, $T = nh_n = 63.09573$, $\Lambda = 10^{-3} I_2$, I_2 is the 2×2-identity matrix, $\{\varepsilon_{ih_n}\}_{i=0,\ldots,n}$ is the i.i.d. sequence of two-dimensional normal random vectors with $\mathbf{E}[\varepsilon_0] = \mathbf{0}$ and $\mathrm{Var}(\varepsilon_0) = I_2$.

For the true model, 1000 independent sample paths are generated by the Euler–Maruyama scheme, and the mean, the standard deviation (SD), the theoretical standard deviation and the mean squared error for the estimators in Theorem 1 are computed and shown in Tables 1, 2, 3, 4, and 5. The personal computer with Intel i7-6950X (3.00GHz) was used for the simulations.

Table 1 shows the simulation results of the estimator $\hat{\Lambda}_n = (\hat{\Lambda}_{n,i,j})_{i,j=1,2}$ of $\Lambda = (\Lambda_{ij})_{i,j=1,2}$.

Tables 2 and 3 show the simulation results of the adaptive ML-type estimator $(\hat{\alpha}_{A,n}, \hat{\beta}_{A,n})$ with the initial value being the true value, where

$$\hat{\alpha}_{A,n} = \arg\sup_{\alpha \in \Theta_1} \mathbb{H}_{1,n}^{\tau} \left(\alpha; \hat{\Lambda}_n\right),$$

$$\hat{\beta}_{A,n} = \arg\sup_{\beta \in \Theta_2} \mathbb{H}_{2,n} \left(\beta; \hat{\alpha}_{A,n}\right),$$

the quasi-likelihood functions $\mathbb{H}_{1,n}^{\tau}$ and $\mathbb{H}_{2,n}$ are that

$$\mathbb{H}_{1,n}^{\tau}(\alpha; \Lambda) = -\frac{1}{2} \sum_{j=1}^{k_n-2} \left(\left(\frac{2}{3} \Delta_n A_n^{\tau} \left(\bar{Y}_{j-1}, \alpha, \Lambda\right) \right)^{-1} \left[(\bar{Y}_{j+1} - \bar{Y}_j)^{\otimes 2} \right] + \log \det A_n^{\tau} \left(\bar{Y}_{j-1}, \alpha, \Lambda\right) \right),$$

$$\mathbb{H}_{2,n}(\beta; \alpha) = -\frac{1}{2} \sum_{j=1}^{k_n-2} \left((\Delta_n A \left(\bar{Y}_{j-1}, \alpha\right))^{-1} \left[(\bar{Y}_{j+1} - \bar{Y}_j - \Delta_n b \left(\bar{Y}_{j-1}, \beta\right))^{\otimes 2} \right] \right),$$

the local mean $\{\bar{Y}_j\}_{j=0,\ldots,k_n-1}$ is defined as

$$\bar{Y}_j = \frac{1}{p_n} \sum_{i=0}^{p_n-1} Y_{j\Delta_n + ih_n}.$$

Table 1 Estimator of Λ

	$\hat{\Lambda}_{11}(0.001)$	$\hat{\Lambda}_{12}(0)$	$\hat{\Lambda}_{22}(0.001)$
Mean	0.001	0.000	0.001
SD	(1.795×10^{-6})	(1.318×10^{-6})	(1.837×10^{-6})
Theoretical SD	(1.732×10^{-6})	(1.225×10^{-6})	(1.732×10^{-6})
Mean squared error	9.951×10^{-10}	1.750×10^{-12}	1.595×10^{-8}

Table 2 Adaptive ML-type estimator of α with the initial value being the true value

	$\hat{\alpha}_1(1)$	$\hat{\alpha}_2(0)$	$\hat{\alpha}_3(2)$
Mean	0.997	0.000	1.994
SD	(0.008)	(0.008)	(0.016)
Theoretical SD	(0.008)	(0.008)	(0.017)
Mean squared error	7.667×10^{-5}	6.473×10^{-5}	3.044×10^{-4}

Table 3 Adaptive ML-type estimator of β with the initial value being the true value

	$\hat{\beta}_1(1)$	$\hat{\beta}_2(0.1)$	$\hat{\beta}_3(0.1)$	$\hat{\beta}_4(1)$	$\hat{\beta}_5(1)$	$\hat{\beta}_6(1)$
Mean	1.079	0.097	0.105	1.074	1.075	1.049
SD	(0.197)	(0.387)	(0.094)	(0.198)	(0.251)	(0.492)
Theoretical SD	(0.205)	(0.409)	(0.087)	(0.174)	(0.250)	(0.500)
Mean squared error	0.045	0.150	0.088	0.045	0.069	0.244

Here, $\tau = 2.0$, $k_n = 8000$, $p_n = 125$, $\Delta_n = 0.007886967$, $T_n = k_n \Delta_n = 63.09573$, $A_n^\tau(x, \alpha, \Lambda) = A(x, \alpha) + 3\Delta_n^{\frac{2-\tau}{\tau-1}} \Lambda$, and $A(x, \alpha) = aa^{\mathrm{T}}(x, \alpha)$. The adaptive ML-type estimator $(\hat{\alpha}_{A,n}, \hat{\beta}_{A,n})$ are obtained by means of **optim()** based on the "L-BFGS-B" method in the R Language.

From Tables 1, 2, and 3, we see that all the estimators have good behaviour.

Tables 4 and 5 show the simulation results of Bayes-type estimators with uniform priors defined as

$$
\tilde{\alpha}_n = \frac{\int_{\Theta_1} \alpha \exp\left(\mathbb{H}_{1,n}^\tau\left(\alpha; \hat{\Lambda}_n\right)\right) d\alpha}{\int_{\Theta_1} \exp\left(\mathbb{H}_{1,n}^\tau\left(\alpha; \hat{\Lambda}_n\right)\right) d\alpha},
$$

$$
\tilde{\beta}_n = \frac{\int_{\Theta_2} \beta \exp\left(\mathbb{H}_{2,n}\left(\beta; \tilde{\alpha}\right)\right) d\beta}{\int_{\Theta_2} \exp\left(\mathbb{H}_{2,n}\left(\beta : \tilde{\alpha}\right)\right) d\beta}.
$$

The Bayes-type estimators of α and β are calculated with MpCN method proposed by Kamatani (2018) for 10^4 Markov chains and 10^3 burn-in iterations.

Table 4 Bayes-type estimator of α

	$\hat{\alpha}_1(1)$	$\hat{\alpha}_2(0)$	$\hat{\alpha}_3(2)$
Mean	0.997	0.000	1.995
SD	(0.009)	(0.009)	(0.018)
Theoretical SD	(0.008)	(0.008)	(0.017)
Mean squared error	9.627×10^{-5}	7.993×10^{-5}	3.537×10^{-4}

Table 5 Bayes-type estimator of β

	$\hat{\beta}_1(1)$	$\hat{\beta}_2(0.1)$	$\hat{\beta}_3(0.1)$	$\hat{\beta}_4(1)$	$\hat{\beta}_5(1)$	$\hat{\beta}_6(1)$
Mean	1.079	0.094	0.105	1.073	1.076	1.045
SD	(0.198)	(0.392)	(0.094)	(0.198)	(0.253)	(0.499)
Theoretical SD	(0.205)	(0.409)	(0.087)	(0.174)	(0.250)	(0.500)
Mean squared error	0.045	0.154	0.088	0.044	0.070	0.250

From Tables 4 and 5, we can see that the Bayes-type estimators have good behaviour. Furthermore, the performance of the Bayes-type estimators is almost the same as that of the estimators in Tables 2 and 3.

Figures 2, 4, and 6 show the plots of the empirical distribution functions, the Q–Q plots and the histograms of the adaptive ML-type estimators $\hat{\beta}_{A,n}^{(i)}(i = 2, 3, 6)$. Figures 3, 5, and 7 show those of the the Bayes-type estimators $\tilde{\beta}_n^{(i)}(i = 2, 3, 6)$. For Figs. 2, 3, 4, 5, 6, and 7, the left side is the empirical distribution function (solid line) and the theoretical cumulative distribution function (dotted line). Moreover, the middle is the Q–Q plot and the right side is the histogram and the theoretical probability density function (dotted line). From Figs. 2, 3, 4, 5, 6 and 7, we can see that both the adaptive ML-type estimators and the Bayes-type estimators have good behaviour.

5 Proofs

5.1 Evaluation for local means

In the first place, we give some evaluations related to local means. Some of the instruments are inherited from the previous studies by Nakakita and Uchida (2017, 2019a). We define the following random variables:

$$
\zeta_{j+1,n} := \frac{1}{p_n} \sum_{i=0}^{p_n-1} \int_{j\Delta_n + ih_n}^{(j+1)\Delta_n} dw_s, \quad \zeta'_{j+2,n} := \frac{1}{p_n} \sum_{i=0}^{p_n-1} \int_{(j+1)\Delta_n}^{(j+1)\Delta_n + ih_n} dw_s.
$$

The next lemma is Lemma 11 in Nakakita and Uchida (2019a).

Fig. 2 Simulation results of the adaptive ML-type estimator $\hat{\beta}_{A,n}^{(2)}$

Fig. 3 Simulation results of the Bayes-type estimator $\tilde{\beta}_n^{(2)}$

Fig. 4 Simulation results of the adaptive ML-type estimator $\hat{\beta}_{A,n}^{(3)}$

Fig. 5 Simulation results of the Bayes-type estimator $\tilde{\beta}_n^{(3)}$

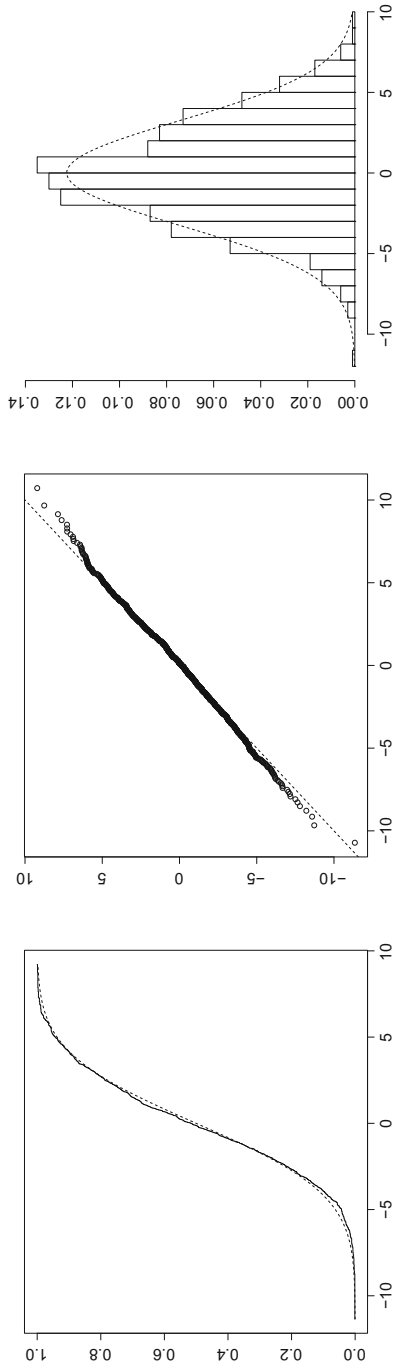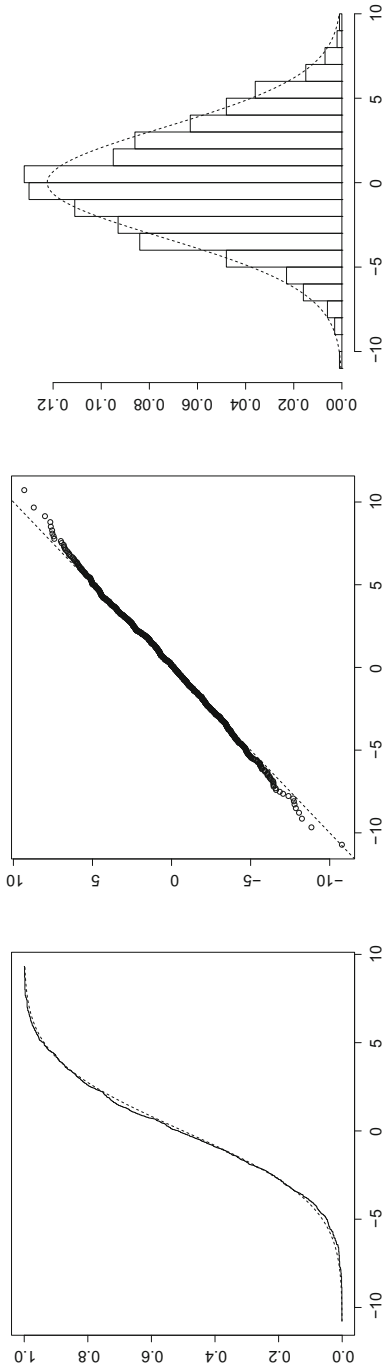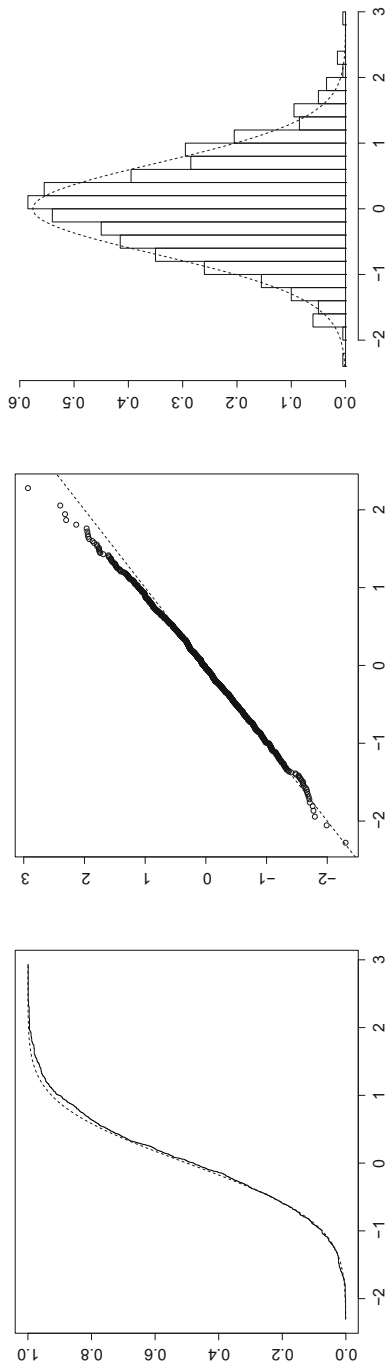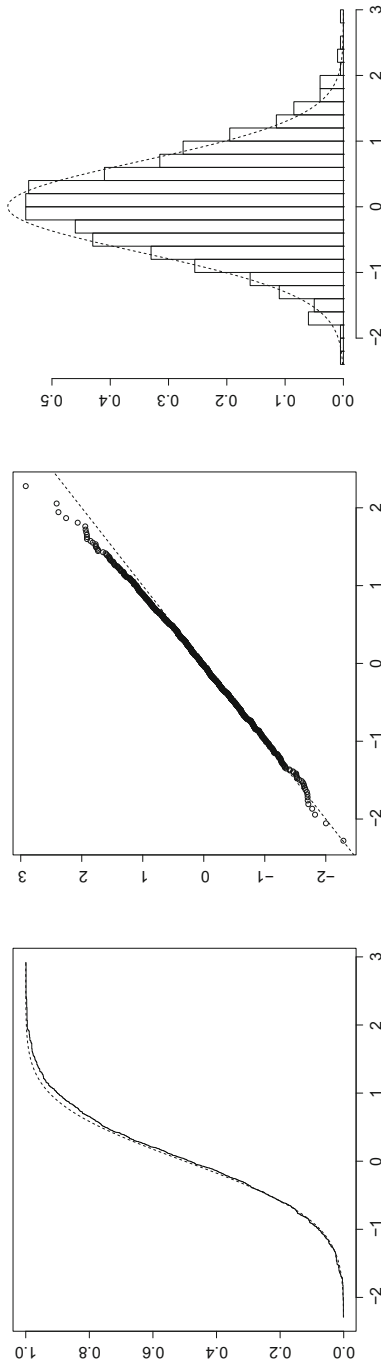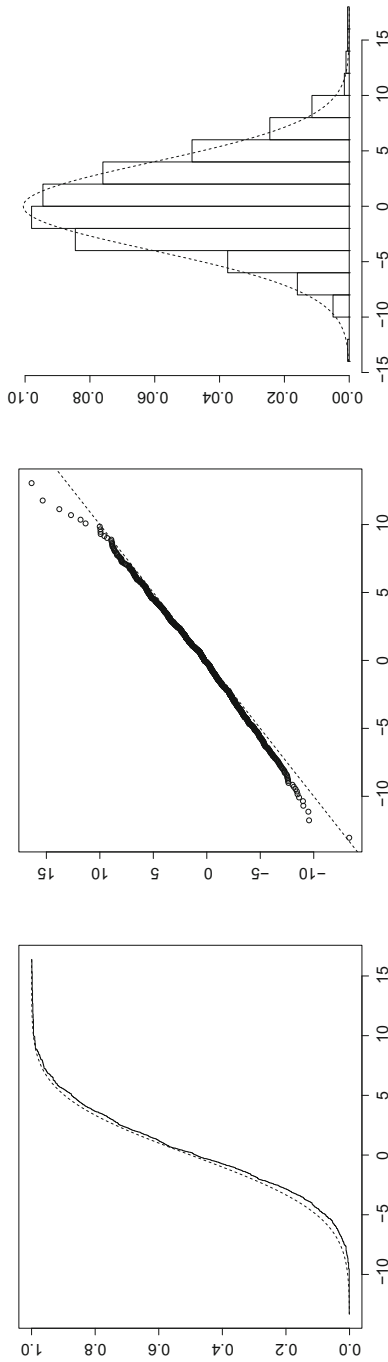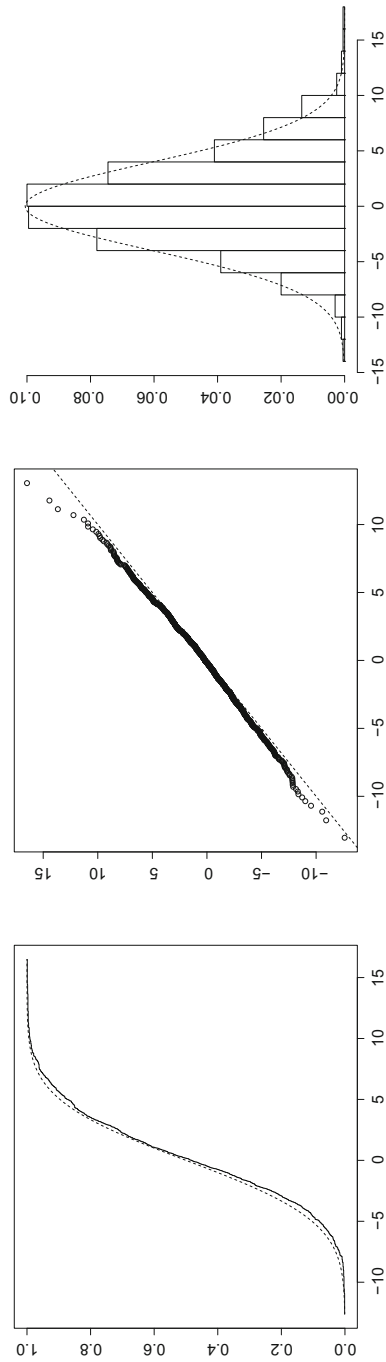**Fig. 6** Simulation results of the adaptive ML-type estimator $\hat{\beta}_{A,n}^{(6)}$

Fig. 7 Simulation results of the Bayes-type estimator $\tilde{\beta}_n^{(6)}$

Lemma 1 $\zeta_{j+1,n}$ and $\zeta'_{j+1,n}$ are $\mathcal{G}^n_{j+1}$-measurable, independent of $\mathcal{G}^n_j$ and Gaussian. These variables have the next decompositions:

$$\zeta_{j+1,n} = \frac{1}{p_n} \sum_{k=0}^{p_n-1} (k+1) \int_{j\Delta_n+kh_n}^{j\Delta_n+(k+1)h_n} \mathrm{d}w_s,$$

$$\zeta'_{j+1,n} = \frac{1}{p_n} \sum_{k=0}^{p_n-1} (p_n-k-1) \int_{j\Delta_n+kh_n}^{j\Delta_n+(k+1)h_n} \mathrm{d}w_s.$$

The evaluation of the following conditional expectations holds:

$$\mathbf{E}_{\theta^\star}\left[\zeta_{j,n}|\mathcal{G}^n_j\right] = \mathbf{E}_{\theta^\star}\left[\zeta'_{j+1,n}|\mathcal{G}^n_j\right] = \mathbf{0},$$

$$\mathbf{E}_{\theta^\star}\left[\zeta_{j+1,n}\left(\zeta_{j+1,n}\right)^{\mathrm{T}}|\mathcal{G}^n_j\right] = m_n \Delta_n I_r,$$

$$\mathbf{E}_{\theta^\star}\left[\zeta'_{j+1,n}\left(\zeta'_{j+1,n}\right)^{\mathrm{T}}|\mathcal{G}^n_j\right] = m'_n \Delta_n I_r,$$

$$\mathbf{E}_{\theta^\star}\left[\zeta_{j+1,n}\left(\zeta'_{j+1,n}\right)^{\mathrm{T}}|\mathcal{G}^n_j\right] = \chi_n \Delta_n I_r,$$

where $m_n = \left(\frac{1}{3} + \frac{1}{2p_n} + \frac{1}{6p_n^2}\right)$, $m'_n = \left(\frac{1}{3} - \frac{1}{2p_n} + \frac{1}{6p_n^2}\right)$, and $\chi_n = \frac{1}{6}\left(1 - \frac{1}{p_n^2}\right)$.

The next lemma can be obtained with same discussion as Proposition 12 in Nakakita and Uchida (2019a).

Lemma 2 Assume the component of the function $f \in C^1\left(\mathbf{R}^d \times \Xi; \mathbf{R}\right)$ and $\partial_x f$ are polynomial growth functions uniformly in $\vartheta \in \Xi$. For all $p \geq 1$, there exists $C(p) > 0$ such that for all $n \in \mathbf{N}$,

$$\sup_{j=0,\ldots,k_n-1} \left\| \sup_{\vartheta \in \Xi} \left| f\left(\bar{Y}_j, \vartheta\right) - f\left(X_{j\Delta_n}, \vartheta\right)\right| \right\|_p \leq C(p) \Delta_n^{1/2}.$$

Lemma 3 Assume the component of the function $f \in C^1\left(\mathbf{R}^d \times \Xi; \mathbf{R}\right)$ and $\partial_x f$ are polynomial growth functions uniformly in $\vartheta \in \Xi$. For all $p \geq 1$, there exists $C(p) > 0$ such that for all $n \in \mathbf{N}$

$$\left\| \sup_{\vartheta \in \Xi} \left| \frac{1}{k_n} \sum_{j=1}^{k_n-2} f\left(\bar{Y}_j, \vartheta\right) - \frac{1}{k_n} \sum_{j=1}^{k_n-2} f\left(X_{j\Delta_n}, \vartheta\right)\right| \right\|_p \leq C(p) \Delta_n^{1/2}.$$

Proof By Lemma 2,

$$\left\| \sup_{\vartheta \in \Xi} \left| \frac{1}{k_n} \sum_{j=1}^{k_n-2} f\left(\bar{Y}_j, \vartheta\right) - \frac{1}{k_n} \sum_{j=1}^{k_n-2} f\left(X_{j\Delta_n}, \vartheta\right) \right| \right\|_p$$

$$\leq \left\| \frac{1}{k_n} \sum_{j=1}^{k_n-2} \sup_{\vartheta \in \Xi} \left| f\left(\bar{Y}_j, \vartheta\right) - f\left(X_{j\Delta_n}, \vartheta\right) \right| \right\|_p$$

$$\leq \frac{1}{k_n} \sum_{j=1}^{k_n-2} \left\| \sup_{\vartheta \in \Xi} \left| f\left(\bar{Y}_j, \vartheta\right) - f\left(X_{j\Delta_n}, \vartheta\right) \right| \right\|_p$$

$$\leq C\left(p\right) \Delta_n^{1/2}.$$

□

Lemma 4 *Assume the components of the functions* $f, g \in C^2\left(\mathbf{R}^d; \mathbf{R}\right)$, $\partial_x f$, $\partial_x g$, $\partial_x^2 f$ $\partial_x^2 g$ *are polynomial growth functions. Then, we have*

$$\left| \mathbf{E}\left[f\left(\bar{Y}_j\right) g\left(X_{(j+1)\Delta_n}\right) - f\left(X_{j\Delta_n}\right) g\left(X_{j\Delta_n}\right) | \mathcal{H}_j^n \right] \right| \leq C\Delta_n \left(1 + \left|X_{j\Delta_n}\right|\right)^C.$$

Proof For Taylor's expansion, we have

$$f\left(\bar{Y}_j\right) g\left(X_{(j+1)\Delta_n}\right)$$
$$= f\left(X_{j\Delta_n}\right) g\left(X_{(j+1)\Delta_n}\right) + \partial_x f\left(X_{j\Delta_n}\right) \left[\bar{Y}_j - X_{j\Delta_n}\right] g\left(X_{(j+1)\Delta_n}\right)$$
$$+ \int_0^1 (1-s)\, \partial_x^2 f\left(X_{j\Delta_n} + s\left(\bar{Y}_j - X_{j\Delta_n}\right) \mathrm{d}s\right) \left[\left(\bar{Y}_j - X_{j\Delta_n}\right)^{\otimes 2}\right] g\left(X_{(j+1)\Delta_n}\right),$$

and Ito–Taylor expansion and Proposition 3.2 in Favetto (2014) verify

$$\left| \mathbf{E}_{\theta^\star}\left[f\left(X_{j\Delta_n}\right) g\left(X_{(j+1)\Delta_n}\right) - f\left(X_{j\Delta_n}\right) g\left(X_{j\Delta_n}\right) | \mathcal{H}_j^n \right] \right| \leq C\Delta_n \left(1 + \left|X_{j\Delta_n}\right|\right)^C,$$

$$\left| \mathbf{E}_{\theta^\star}\left[\int_0^1 (1-s)\, \partial_x^2 f\left(X_{j\Delta_n} + s\left(\bar{Y}_j - X_{j\Delta_n}\right) \mathrm{d}s\right) \left[\left(\bar{Y}_j - X_{j\Delta_n}\right)^{\otimes 2}\right] g\left(X_{(j+1)\Delta_n}\right) | \mathcal{H}_j^n \right] \right|$$

$$\leq \mathbf{E}_{\theta^\star}\left[C\left(1 + \sup_{t \in [j\Delta_n, (j+1)\Delta_n]} |X_t| + |\bar{\varepsilon}_j|\right)^C | \mathcal{H}_j^n \right] \mathbf{E}_{\theta^\star}\left[\left|\bar{Y}_j - X_{j\Delta_n}\right|^2 | \mathcal{H}_j^n \right]$$

$$\leq C\Delta_n \left(1 + \left|X_{j\Delta_n}\right|\right)^C.$$

It holds that

$$\partial_x f\left(X_{j\Delta_n}\right) \left[\bar{Y}_j - X_{j\Delta_n}\right] g\left(X_{(j+1)\Delta_n}\right) = \partial_x f\left(X_{j\Delta_n}\right) \left[\bar{Y}_j - X_{j\Delta_n}\right] g\left(X_{j\Delta_n}\right)$$
$$+ \partial_x f\left(X_{j\Delta_n}\right) \left[\bar{Y}_j - X_{j\Delta_n}\right] \int_0^1 \partial_x g\left(X_{j\Delta_n} + s\left(X_{(j+1)\Delta_n} - X_{j\Delta_n}\right)\right)$$
$$\mathrm{d}s \left[X_{(j+1)\Delta_n} - X_{j\Delta_n}\right]$$

and Proposition 3.2 in Favetto (2014) leads to

$$
\left| \mathbf{E}_{\theta^\star} \left[\partial_x f \left(X_{j\Delta_n} \right) \left[\bar{Y}_j - X_{j\Delta_n} \right] g \left(X_{j\Delta_n} \right) | \mathcal{H}_j^n \right] \right| \leq C \Delta_n \left(1 + \left| X_{j\Delta_n} \right| \right)^C ,
$$

$$
\left| \mathbf{E}_{\theta^\star} \left[\partial_x f \left(X_{j\Delta_n} \right) \left[\bar{Y}_j - X_{j\Delta_n} \right] \right. \right.
$$

$$
\left. \left. \times \int_0^1 \partial_x g \left(X_{j\Delta_n} + s \left(X_{(j+1)\Delta_n} - X_{j\Delta_n} \right) \right) \mathrm{d}s \left[X_{(j+1)\Delta_n} - X_{j\Delta_n} \right] | \mathcal{H}_j^n \right] \right|
$$

$$
\leq C \left(1 + \left| X_{j\Delta_n} \right| \right)^C \mathbf{E}_{\theta^\star} \left[\left| \bar{Y}_j - X_{j\Delta_n} \right|^2 | \mathcal{H}_j^n \right]^{1/2} \mathbf{E}_{\theta^\star} \left[\left| X_{(j+1)\Delta_n} - X_{j\Delta_n} \right|^2 | \mathcal{H}_j^n \right]^{1/2}
$$

$$
\leq C \Delta_n \left(1 + \left| X_{j\Delta_n} \right| \right)^C .
$$

Hence, we obtain the result. □

Lemma 5 (i) *The next expansion holds:*

$$
\bar{Y}_{j+1} - \bar{Y}_j = \Delta_n b \left(X_{j\Delta_n} \right) + a \left(X_{j\Delta_n} \right) \left(\zeta_{j+1,n} + \zeta'_{j+2,n} \right) + e_{j,n} + \left(\Lambda_\star \right)^{1/2} \left(\bar{\varepsilon}_{j+1} - \bar{\varepsilon}_j \right)
$$

where $e_{j,n}$ is a $\mathcal{H}_{j+2}^n$-measurable random variable such that $\left\| e_{j,n} \right\|_p \leq C \left(p \right) \Delta_n$, for $j = 1, \ldots, k_n - 2$, $n \in \mathbf{N}$ and $p \geq 1$.

(ii) *For any $p \geq 1$ and $\mathcal{H}_j^n$-measurable $\mathbf{R}^d \otimes \mathbf{R}^r$-valued random variable $\mathbb{B}_j^n$ such that $\sup_{j,n} \mathbf{E} \left[\left\| \mathbb{B}_j^n \right\|^m \right] < \infty$ for all $m \in \mathbf{N}$, we have the next L^p-boundedness:*

$$
\mathbf{E}_{\theta^\star} \left[\left| \sum_{j=1}^{k_n-2} \mathbb{B}_j^n \left[e_{j,n} \left(\zeta_{j+1,n} + \zeta'_{j+2,n} \right)^{\mathrm{T}} \right] \right|^p \right]^{1/p} \leq C \left(p \right) k_n \Delta_n^2 .
$$

(iii) *For any $p \geq 1$ and $\mathcal{H}_j^n$-measurable $\mathbf{R}^d$-valued random variable $\mathbb{C}_j^n$ such that $\sup_{j,n} \mathbf{E} \left[\left| \mathbb{C}_j^n \right|^m \right] < \infty$ for all $m \in \mathbf{N}$, we have the next L^p-boundedness:*

$$
\mathbf{E}_{\theta^\star} \left[\left| \sum_{j=1}^{k_n-2} \mathbb{C}_j^n \left[e_{j,n} \right] \right|^p \right]^{1/p} \leq C \left(p \right) k_n \Delta_n^{3/2} .
$$

Proof Firstly, we prove (i). Without loss of generality, assume p is an even number. It holds

$$
\bar{Y}_{j+1} - \bar{Y}_j = \bar{X}_{j+1} - \bar{X}_j + \left(\Lambda_\star \right)^{1/2} \left(\bar{\varepsilon}_{j+1} - \bar{\varepsilon}_j \right) ,
$$

and

$$\bar{X}_{j+1} - \bar{X}_j$$

$$= \frac{1}{p_n} \sum_{i=0}^{p_n-1} \left(X_{(j+1)\Delta_n+ih_n} - X_{j\Delta_n+ih_n} \right)$$

$$= \frac{1}{p_n} \sum_{i=0}^{p_n-1} \left(X_{(j+1)\Delta_n+ih_n} - X_{(j+1)\Delta_n+(i-1)h_n} + X_{(j+1)\Delta_n+(i-1)h_n} - \cdots - X_{j\Delta_n+ih_n} \right)$$

$$= \frac{1}{p_n} \sum_{i=0}^{p_n-1} \left(\int_{j\Delta_n+(p_n+i-1)h_n}^{j\Delta_n+(p_n+i)h_n} dX_s + \cdots + \int_{j\Delta_n+ih_n}^{j\Delta_n+(i+1)h_n} dX_s \right)$$

$$= \frac{1}{p_n} \sum_{i=0}^{p_n-1} \sum_{l=0}^{p_n-1} \int_{j\Delta_n+(i+l)h_n}^{j\Delta_n+(i+l+1)h_n} dX_s$$

$$= \frac{1}{p_n} \sum_{i=0}^{p_n-1} (i+1) \int_{j\Delta_n+ih_n}^{j\Delta_n+(i+1)h_n} dX_s + \frac{1}{p_n} \sum_{i=0}^{p_n-1} (p_n-i-1) \int_{(j+1)\Delta_n+ih_n}^{(j+1)\Delta_n+(i+1)h_n} dX_s$$

$$= \frac{1}{p_n} \sum_{i=0}^{p_n-1} (i+1) \int_{j\Delta_n+ih_n}^{j\Delta_n+(i+1)h_n} dX_s + \frac{1}{p_n} \sum_{i=0}^{p_n-1} (p_n-i-1) \int_{(j+1)\Delta_n+ih_n}^{(j+1)\Delta_n+(i+1)h_n} dX_s$$

$$+ \Delta_n b \left(X_{j\Delta_n} \right) - \frac{1}{p_n} \sum_{i=0}^{p_n-1} \left((i+1) h_n + (p_n-i-1) h_n \right) b \left(X_{j\Delta_n} \right)$$

$$= \Delta_n b \left(X_{j\Delta_n} \right) + a \left(X_{j\Delta_n} \right) \left(\zeta_{j+1,n} + \zeta'_{j+2,n} \right) + e_{j,n},$$

where $e_{j,n} = \sum_{l=1}^{3} \left(r_{j,n}^{(l)} + s_{j,n}^{(l)} \right)$,

$$r_{j,n}^{(1)} = \frac{1}{p_n} \sum_{i=0}^{p_n-1} (i+1) \int_{j\Delta_n+ih_n}^{j\Delta_n+(i+1)h_n} \left(a \left(X_{j\Delta_n+ih_n} \right) - a \left(X_{j\Delta_n} \right) \right) dw_s,$$

$$r_{j,n}^{(2)} = \frac{1}{p_n} \sum_{i=0}^{p_n-1} (i+1) \int_{j\Delta_n+ih_n}^{j\Delta_n+(i+1)h_n} \left(a \left(X_s \right) - a \left(X_{j\Delta_n+ih_n} \right) \right) dw_s,$$

$$r_{j,n}^{(3)} = \frac{1}{p_n} \sum_{i=0}^{p_n-1} (i+1) \int_{j\Delta_n+ih_n}^{j\Delta_n+(i+1)h_n} \left(b \left(X_s \right) - b \left(X_{j\Delta_n} \right) \right) ds,$$

$$s_{j,n}^{(1)} = \frac{1}{p_n} \sum_{i=0}^{p_n-1} (p_n-i-1) \int_{(j+1)\Delta_n+ih_n}^{(j+1)\Delta_n+(i+1)h_n} \left(a \left(X_{(j+1)\Delta_n+ih_n} \right) - a \left(X_{j\Delta_n} \right) \right) dw_s,$$

$$s_{j,n}^{(2)} = \frac{1}{p_n} \sum_{i=0}^{p_n-1} (p_n-i-1) \int_{(j+1)\Delta_n+ih_n}^{(j+1)\Delta_n+(i+1)h_n} \left(a \left(X_s \right) - a \left(X_{(j+1)\Delta_n+ih_n} \right) \right) dw_s,$$

$$s_{j,n}^{(3)} = \frac{1}{p_n} \sum_{i=0}^{p_n-1} (p_n-i-1) \int_{(j+1)\Delta_n+ih_n}^{(j+1)\Delta_n+(i+1)h_n} \left(b \left(X_s \right) - b \left(X_{j\Delta_n} \right) \right) ds,$$

using Lemma 1. By BDG inequality, Hölder's inequality, and triangular inequality for $L^{p/2}$-norm, we have

$$
\left\| r_{j,n}^{(1)} \right\|_p = \mathbf{E}_{\theta^\star} \left[\left| \frac{1}{p_n} \sum_{i=0}^{p_n-1} (i+1) \int_{j\Delta_n+ih_n}^{j\Delta_n+(i+1)h_n} \left(a\left(X_{j\Delta_n+ih_n}\right) - a\left(X_{j\Delta_n}\right) \right) \mathrm{d}w_s \right|^p \right]^{1/p}
$$

$$
\leq C(p) \mathbf{E}_{\theta^\star} \left[\left| \frac{1}{p_n^2} \sum_{i=0}^{p_n-1} (i+1)^2 \int_{j\Delta_n+ih_n}^{j\Delta_n+(i+1)h_n} \left\| a\left(X_{j\Delta_n+ih_n}\right) - a\left(X_{j\Delta_n}\right) \right\|^2 \mathrm{d}s \right|^{p/2} \right]^{1/p}
$$

$$
\leq C(p) \mathbf{E}_{\theta^\star} \left[\left| \sum_{i=0}^{p_n-1} \int_{j\Delta_n+ih_n}^{j\Delta_n+(i+1)h_n} \left\| a\left(X_{j\Delta_n+ih_n}\right) - a\left(X_{j\Delta_n}\right) \right\|^2 \mathrm{d}s \right|^{p/2} \right]^{1/p}
$$

$$
= C(p) \left(\mathbf{E}_{\theta^\star} \left[\left| \sum_{i=0}^{p_n-1} \int_{j\Delta_n+ih_n}^{j\Delta_n+(i+1)h_n} \left\| a\left(X_{j\Delta_n+ih_n}\right) - a\left(X_{j\Delta_n}\right) \right\|^2 \mathrm{d}s \right|^{p/2} \right]^{2/p} \right)^{1/2}
$$

$$
\leq C(p) \left(\sum_{i=0}^{p_n-1} \mathbf{E}_{\theta^\star} \left[\left| \int_{j\Delta_n+ih_n}^{j\Delta_n+(i+1)h_n} \left\| a\left(X_{j\Delta_n+ih_n}\right) - a\left(X_{j\Delta_n}\right) \right\|^2 \mathrm{d}s \right|^{p/2} \right]^{2/p} \right)^{1/2}
$$

$$
\leq C(p) \left(\sum_{i=0}^{p_n-1} h_n^{1-2/p} \mathbf{E}_{\theta^\star} \left[\int_{j\Delta_n+ih_n}^{j\Delta_n+(i+1)h_n} \left\| a\left(X_{j\Delta_n+ih_n}\right) - a\left(X_{j\Delta_n}\right) \right\|^p \mathrm{d}s \right]^{2/p} \right)^{1/2}
$$

$$
= C(p) \left(\sum_{i=0}^{p_n-1} h_n \sup_{s\in[j\Delta_n,(j+1)\Delta_n]} \mathbf{E}_{\theta^\star} \left[\left\| a\left(X_s\right) - a\left(X_{j\Delta_n}\right) \right\|^p \right]^{2/p} \right)^{1/2}
$$

$$
\leq C(p) \left(\sum_{i=0}^{p_n-1} h_n \left(C(p) \Delta_n^{p/2} \mathbf{E}_{\theta^\star} \left[\left(1+\left|X_{j\Delta_n}\right|\right)^{C(p)} \right] \right)^{2/p} \right)^{1/2}
$$

$$
\leq C(p) \left(C(p) \Delta_n^2 \right)^{1/2} \leq C(p) \Delta_n
$$

and we also have $\left\| s_{j,n}^{(1)} \right\|_p \leq C(p) \Delta_n$ which can be obtained in the analogous manner. For $r_{j,n}^{(2)}$, we obtain

$$
\left\| r_{j,n}^{(2)} \right\|_p = \mathbf{E}_{\theta^\star} \left[\left| \frac{1}{p_n} \sum_{i=0}^{p_n-1} (i+1) \int_{j\Delta_n+ih_n}^{j\Delta_n+(i+1)h_n} \left(a\left(X_s\right) - a\left(X_{j\Delta_n+ih_n}\right) \right) \mathrm{d}w_s \right|^p \right]^{1/p}
$$

$$
\leq C(p) \mathbf{E}_{\theta^\star} \left[\left| \frac{1}{p_n^2} \sum_{i=0}^{p_n-1} (i+1)^2 \int_{j\Delta_n+ih_n}^{j\Delta_n+(i+1)h_n} \left\| a\left(X_s\right) - a\left(X_{j\Delta_n+ih_n}\right) \right\|^2 \mathrm{d}s \right|^{p/2} \right]^{1/p}
$$

$$
\leq C(p) \left(\sum_{i=0}^{p_n-1} \mathbf{E}_{\theta^\star} \left[\left| \int_{j\Delta_n+ih_n}^{j\Delta_n+(i+1)h_n} \left\| a\left(X_s\right) - a\left(X_{j\Delta_n+ih_n}\right) \right\|^2 \mathrm{d}s \right|^{p/2} \right]^{2/p} \right)^{1/2}
$$

$$\leq C(p)\left(\sum_{i=0}^{p_n-1}h_n^{1-2/p}\mathbf{E}_{\theta^\star}\left[\int_{j\Delta_n+ih_n}^{j\Delta_n+(i+1)h_n}\left\|a\left(X_s\right)-a\left(X_{j\Delta_n+ih_n}\right)\right\|^p\mathrm{d}s\right]^{2/p}\right)^{1/2}$$

$$\leq C(p)\left(\sum_{i=0}^{p_n-1}h_n^{1-2/p}\left(\int_{j\Delta_n+ih_n}^{j\Delta_n+(i+1)h_n}\mathbf{E}_{\theta^\star}\left[\left\|a\left(X_s\right)-a\left(X_{j\Delta_n+ih_n}\right)\right\|^p\right]\mathrm{d}s\right)^{2/p}\right)^{1/2}$$

$$\leq C(p)\left(\sum_{i=0}^{p_n-1}h_n\left(\sup_{s\in[j\Delta_n+ih_n,j\Delta_n+(i+1)h_n]}\mathbf{E}_{\theta^\star}\left[\left\|a\left(X_s\right)-a\left(X_{j\Delta_n+ih_n}\right)\right\|^p\right]\mathrm{d}s\right)^{2/p}\right)^{1/2}$$

$$\leq C(p)\left(p_nh_n^2\right)^{1/2}\leq C(p)\Delta_n^{3/2}$$

because of BDG inequality, Hölder's inequality, Fubini's theorem, and the fact that $h_n=\Delta_n/p_n\leq\Delta_n^2$, and the same evaluation can be proved for $s_{j,n}^{(2)}$. It also holds

$$\left\|r_{j,n}^{(3)}\right\|_p=\frac{1}{p_n}\sum_{k=0}^{p_n-1}(k+1)\mathbf{E}_{\theta^\star}\left[\left|\int_{j\Delta_n+kh_n}^{j\Delta_n+(k+1)h_n}\left(b\left(X_s\right)-b\left(X_{j\Delta_n}\right)\right)\mathrm{d}s\right|^p\right]^{1/p}$$

$$\leq\frac{C(p)}{p_n}\sum_{k=0}^{p_n-1}(k+1)\mathbf{E}_{\theta^\star}\left[\left(\int_{j\Delta_n+kh_n}^{j\Delta_n+(k+1)h_n}\left|b\left(X_s\right)-b\left(X_{j\Delta_n}\right)\right|\mathrm{d}s\right)^p\right]^{1/p}$$

$$\leq\frac{C(p)}{p_n}\sum_{k=0}^{p_n-1}(k+1)h_n^{1-1/p}\mathbf{E}_{\theta^\star}\left[\int_{j\Delta_n+kh_n}^{j\Delta_n+(k+1)h_n}\left|b\left(X_s\right)-b\left(X_{j\Delta_n}\right)\right|^p\mathrm{d}s\right]^{1/p}$$

$$\leq\frac{C(p)}{p_n}\sum_{k=0}^{p_n-1}(k+1)h_n^{1-1/p}\left(\int_{j\Delta_n+kh_n}^{j\Delta_n+(k+1)h_n}\mathbf{E}_{\theta^\star}\left[\left|b\left(X_s\right)-b\left(X_{j\Delta_n}\right)\right|^p\right]\mathrm{d}s\right)^{1/p}$$

$$\leq\frac{C(p)}{p_n}\sum_{k=0}^{p_n-1}(k+1)h_n\left(\sup_{s\in[j\Delta_n,(j+1)\Delta_n]}\mathbf{E}_{\theta^\star}\left[\left|b\left(X_s\right)-b\left(X_{j\Delta_n}\right)\right|^p\right]\right)^{1/p}$$

$$\leq\frac{C(p)\Delta_n^{1/2}h_n}{p_n}\sum_{k=0}^{p_n-1}(k+1)\leq C(p)\Delta_n^{3/2}$$

by Hölder's inequality and Fubini's theorem, and same evaluation holds for $s_{j,n}^{(3)}$: $\left\|s_{j,n}^{(3)}\right\|_p\leq C(p)\Delta_n^{3/2}$. Hence, we obtain the evaluation for $\left\|e_{j,n}\right\|_p$.

In the next place, we show (ii) holds. Note that it is sufficient to see only the moments for $r_{j,n}^{(1)}\zeta_{j+1,n}^{\mathrm{T}}$ and $s_{j,n}^{(1)}\left(\zeta'_{j+2,n}\right)^{\mathrm{T}}$ because Hölder's inequality and orthogonality are applicable for the others. We have the following expression for $r_{j,n}^{(1)}$ and $s_{j,n}^{(1)}$:

$$r_{j,n}^{(1)}=\frac{1}{p_n}\sum_{i=0}^{p_n-1}(i+1)\left(a\left(X_{j\Delta_n+ih_n}\right)-a\left(X_{j\Delta_n}\right)\right)\int_{j\Delta_n+ih_n}^{j\Delta_n+(i+1)h_n}\mathrm{d}w_s,$$

$$s_{j,n}^{(1)} = \frac{1}{p_n} \sum_{i=0}^{p_n-1} (p_n - i - 1) \left(a \left(X_{(j+1)\Delta_n+ih_n} \right) - a \left(X_{j\Delta_n} \right) \right) \int_{(j+1)\Delta_n+ih_n}^{(j+1)\Delta_n+(i+1)h_n} \mathrm{d}w_s.$$

Let us define for all $\ell = p_n, \ldots, (k_n - 2) p_n + p_n - 1$, $\ell_1(\ell) = \lfloor \ell/p_n \rfloor$, and $\ell_2(\ell) = \ell - \ell_1(\ell)$,

$$\mathbb{D}_\ell^n = \sum_{j=1}^{\ell_1(\ell)} \sum_{i=0}^{\ell_2(\ell)} \frac{i+1}{p_n} \mathbb{B}_j^n \left(a \left(X_{j\Delta_n+ih_n} \right) - a \left(X_{j\Delta_n} \right) \right) \left[\left(\int_{j\Delta_n+ih_n}^{j\Delta_n+(i+1)h_n} \mathrm{d}w_s \right)^{\otimes 2} \right],$$

$$\mathbf{D}_\ell^n = \sum_{j=1}^{\ell_1(\ell)} \sum_{i=0}^{\ell_2(\ell)} \frac{i+1}{p_n} \mathbb{B}_j^n \left(a \left(X_{j\Delta_n+ih_n} \right) - a \left(X_{j\Delta_n} \right) \right) [h_n I_r],$$

and then we have $\sum_{j=1}^{k_n-2} \mathbb{B}_j^n \left[r_{j,n}^{(1)} \left(\zeta_{j+1,n} \right)^{\mathrm{T}} \right] = \mathbb{D}_{(k_n-2)p_n+p_n-1}^n$. We can easily observe that $\mathbb{D}_\ell^n - \mathbf{D}_\ell^n$ is a martingale with respect to $\left\{ \mathcal{H}_{\ell_1(\ell),\ell_2(\ell)}^n \right\}$. Then, Burkholder's inequality is applicable and it follows that

$$\mathbf{E}_{\theta^\star} \left[\left| \mathbb{D}_{(k_n-2)p_n+p_n-1}^n - \mathbf{D}_{(k_n-2)p_n+p_n-1}^n \right|^p \right]$$

$$\leq C(p) \mathbf{E}_{\theta^\star} \left[\left| \sum_{j=1}^{k_n-2} \sum_{i=0}^{p_n-1} \left(\frac{i+1}{p_n} \right)^2 \left\| \mathbb{B}_j^n \right\|^2 \left\| a \left(X_{j\Delta_n+ih_n} \right) - a \left(X_{j\Delta_n} \right) \right\|^2 \right. \right.$$

$$\times \left. \left. \left(\left| \int_{j\Delta_n+ih_n}^{j\Delta_n+(i+1)h_n} \mathrm{d}w_s \right|^4 + r^2 h_n^2 \right) \right|^{p/2} \right]$$

$$\leq C(p) \frac{n^{p/2}}{n} \sum_{j=1}^{k_n-2} \sum_{i=0}^{p_n-1} \mathbf{E}_{\theta^\star} \left[\left| \left\| \mathbb{B}_j^n \right\|^2 \left\| a \left(X_{j\Delta_n+ih_n} \right) - a \left(X_{j\Delta_n} \right) \right\|^2 \right. \right.$$

$$\times \left. \left. \left(\left| \int_{j\Delta_n+ih_n}^{j\Delta_n+(i+1)h_n} \mathrm{d}w_s \right|^4 + r^2 h_n^2 \right) \right|^{p/2} \right]$$

$$\leq C(p) \frac{n^{p/2}}{n} \sum_{j=1}^{k_n-2} \sum_{i=0}^{p_n-1} \mathbf{E}_{\theta^\star} \left[\left\| a \left(X_{j\Delta_n+ih_n} \right) - a \left(X_{j\Delta_n} \right) \right\|^{2p} \right.$$

$$\left. \left(\left| \int_{j\Delta_n+ih_n}^{j\Delta_n+(i+1)h_n} \mathrm{d}w_s \right|^4 - r^2 h_n^2 \right)^p \right]^{1/2}$$

$$\leq C(p) \frac{n^{p/2}}{n} \sum_{j=1}^{k_n-2} \sum_{i=0}^{p_n-1} \mathbf{E}_{\theta^\star} \left[\left\| a \left(X_{j\Delta_n+ih_n} \right) - a \left(X_{j\Delta_n} \right) \right\|^{4p} \right]^{1/4}$$

$$\times \, \mathbf{E}_{\theta^\star} \left[\left(\left| \int_{j\Delta_n+ih_n}^{j\Delta_n+(i+1)h_n} \mathrm{d}w_s \right|^4 + r^2 h_n^2 \right)^{2p} \right]^{1/4}$$

$$\leq C\left(p\right) n^{p/2} \Delta_n^{p/2} h_n^p = C\left(p\right) k_n^{p/2} \Delta_n^p h_n^{p/2} \leq C\left(p\right) k_n^{p/2} \Delta_n^{3p}.$$

Hence, we have $\left\| \mathbf{D}_{(k_n-2)p_n+p_n-1}^n - \mathbf{D}_{(k_n-2)p_n+p_n-1}^n \right\|_p \leq C\left(p\right) k_n^{1/2} \Delta_n$. Furthermore, let us define

$$\mathbf{D}_\ell^{1,n} = \sum_{j=1}^{\ell_1(\ell)} \sum_{i=0}^{\ell_2(\ell)} \frac{i+1}{p_n} \mathbb{B}_j^n \left(a\left(X_{j\Delta_n+ih_n}\right) - \mathbf{E}_{\theta^\star}\left[a\left(X_{j\Delta_n+ih_n}\right) | \mathcal{H}_j^n \right] \right) [h_n I_r] \,,$$

$$\mathbf{D}_\ell^{2,n} = \sum_{j=1}^{\ell_1(\ell)} \sum_{i=0}^{\ell_2(\ell)} \frac{i+1}{p_n} \mathbb{B}_j^n \left(\mathbf{E}_{\theta^\star}\left[a\left(X_{j\Delta_n+ih_n}\right) | \mathcal{H}_j^n \right] - a\left(X_{j\Delta_n}\right) \right) [h_n I_r] \,,$$

and clearly, we have $\mathbf{D}_\ell^n = \mathbf{D}_\ell^{1,n} + \mathbf{D}_\ell^{2,n}$. In addition, we see $\left\{ \mathbf{D}_{jp_n+p_n-1}^{1,n} \right\}_{j=1,\ldots,k_n-2}$ is a martingale with respect to $\left\{ \mathcal{H}_j^n \right\}_{j=1,\ldots,k_n-2}$, and then, Burkholder's inequality leads to

$$\mathbf{E}_{\theta^\star} \left[\left| \mathbf{D}_{(k_n-2)p_n+p_n-1}^{1,n} \right|^p \right]$$

$$\leq \mathcal{C}\left(p\right) \mathbf{E}_{\theta^\star} \left[\left| \sum_{j=1}^{k_n-2} p_n \sum_{i=0}^{p_n-1} \left(\frac{i+1}{p_n} \right)^2 \left\| \mathbb{B}_j^n \right\|^2 \left\| a\left(X_{j\Delta_n+ih_n}\right) - \mathbf{E}_{\theta^\star}\left[a\left(X_{j\Delta_n+ih_n}\right) | \mathcal{H}_j^n \right] \right\|^2 h_n^2 \right|^{p/2} \right]$$

$$\leq \mathcal{C}\left(p\right) n^{p/2} p_n^{p/2} \Delta_n^{p/2} h_n^p \leq \mathcal{C}\left(p\right) k_n^{p/2} \Delta_n^{3p/2} \leq \mathcal{C}\left(p\right) k_n^p \Delta_n^{2p}.$$

Regarding $\mathbf{D}_\ell^{2,n}$, we have

$$\mathbf{E}_{\theta^\star} \left[\left| \mathbf{D}_{(k_n-2)p_n+p_n-1}^{2,n} \right|^p \right]$$

$$= \mathbf{E}_{\theta^\star} \left[\left| \sum_{j=1}^{k_n-2} \sum_{i=0}^{p_n-1} \frac{i+1}{p_n} \mathbb{B}_j^n \left(\mathbf{E}_{\theta^\star}\left[a\left(X_{j\Delta_n+ih_n}\right) | \mathcal{H}_j^n \right] - a\left(X_{j\Delta_n}\right) \right) [h_n I_r] \right|^p \right]$$

$$\leq \mathbf{E}_{\theta^\star} \left[\left| \sum_{j=1}^{\ell_1} \sum_{i=0}^{\ell_2} \left\| \mathbb{B}_j^n \right\| \left\| \mathbf{E}_{\theta^\star}\left[a\left(X_{j\Delta_n+ih_n}\right) | \mathcal{H}_j^n \right] - a\left(X_{j\Delta_n}\right) \right\| \right|^p \right] h_n^p$$

$$\leq C\left(p\right) n^p h_n^p \Delta_n^p = C\left(p\right) k_n^p \Delta_n^{2p},$$

since $\left\| \mathbf{E}_{\theta^\star}\left[a\left(X_{j\Delta_n+ih_n}\right) | \mathcal{H}_j^n \right] - a\left(X_{j\Delta_n}\right) \right\| \leq C\Delta_n \left(1 + \left| X_{j\Delta_n} \right| \right)^C$. The same evaluation holds for $s_{j,n}^{(1)}$, and hence, we obtain the result.

Finally, we check that (iii) holds. It is only necessary to verify it for $r_{j,n}^{(1)}$ and $s_{j,n}^{(1)}$, and we show with respect to $r_{j,n}^{(1)}$. Since $\left\{\sum_{k=1}^{j} \mathbb{C}_{k,n}\left[r_{k,n}^{(1)}\right]\right\}$ for $\ell \leq k_n - 2$ is a martingale with respect to $\left\{\mathcal{H}_j^n\right\}$, we can utilise Burkholder's inequality and then

$$\mathbf{E}_{\theta^\star}\left[\left|\sum_{j=1}^{k_n-2} \mathbb{C}_{j,n}\left[r_{j,n}^{(1)}\right]\right|^p\right] \leq C\left(p\right) k_n^{p/2-1} \sum_{j=1}^{k_n-2} \mathbf{E}_{\theta^\star}\left[\left|r_{j,n}^{(1)}\right|^{2p}\right]^{1/2} \leq C\left(p\right) k_n^{p/2} \Delta_n^p$$

and we can have the same evaluation for $s_{j,n}^{(1)}$. $\qquad\square$

Remark 6 When the evaluation $\left\|e_{j,n}\right\|_p \leq C\left(p\right) \Delta_n$ is sufficient, then we can abbreviate $\Delta_n b\left(X_{j\Delta_n}\right)$ in the right-hand side.

Lemma 6 (a) For all $p \geq 1$, there exists $C\left(p\right) > 0$ such that for all $j = 0, \ldots, k_n - 1$ and $n \in \mathbf{N}$,

$$\left\|\bar{\varepsilon}_j\right\|_p \leq C\left(p\right) p_n^{-1/2}.$$

(b) For all $p \geq 1$, there exists $C\left(p\right) > 0$ such that for all $n \in \mathbf{N}$

$$\left\|\hat{\Lambda}_n - \Lambda_\star\right\|_p \leq C\left(p\right)\left(h_n + \frac{1}{\sqrt{n}}\right).$$

Proof (a) Because of Hölder's inequality, it is enough to evaluate it in the case where p is an even integer. We easily obtain

$$\mathbf{E}_{\theta^\star}\left[\left|\bar{\varepsilon}_j\right|^p\right] \leq \sum_{\ell=1}^{d} \mathbf{E}_{\theta^\star}\left[\left|\bar{\varepsilon}_j^{(\ell)}\right|^p\right] = \frac{1}{p_n^p} \sum_{\ell=1}^{d} \sum_{i_1=0}^{p_n-1} \cdots \sum_{i_{p/2}=0}^{p_n-1} \mathbf{E}_{\theta^\star}\left[\left|\varepsilon_{j\Delta_n+i_1 h_n}^{(\ell)}\right|^2 \cdots \left|\varepsilon_{j\Delta_n+i_{p/2} h_n}^{(\ell)}\right|^2\right]$$

$$\leq C\left(p\right) p_n^{-p/2}$$

for [A5].

(b) As (a), it is enough to evaluate in the case where p is an even integer. Then, we have

$$\left\|\hat{\Lambda}_n - \Lambda_\star\right\|_p = \mathbf{E}_{\theta^\star}\left[\left\|\frac{1}{2n}\sum_{i=1}^{n}\left(Y_{ih_n} - Y_{(i-1)h_n}\right)^{\otimes 2} - \Lambda_\star\right\|^p\right]^{1/p}$$

$$\leq \mathbf{E}_{\theta^\star}\left[\left\|\frac{1}{2n}\sum_{i=1}^{n}\left(X_{ih_n} - X_{(i-1)h_n}\right)^{\otimes 2}\right\|^p\right]^{1/p}$$

$$+ \mathbf{E}_{\theta^\star}\left[\left\|\frac{1}{2n}\sum_{i=1}^{n}\left(X_{ih_n} - X_{(i-1)h_n}\right)\left(\varepsilon_{ih_n} - \varepsilon_{(i-1)h_n}\right)^{\mathrm{T}}\left(\Lambda_\star\right)^{1/2}\right\|^p\right]^{1/p}$$

$$+ \mathbf{E}_{\theta^\star}\left[\left\|\frac{1}{2n}\sum_{i=1}^{n}\left[\left(\Lambda_\star\right)^{1/2}\left(\varepsilon_{ih_n} - \varepsilon_{(i-1)h_n}\right)\right]^{\otimes 2} - \Lambda_\star\right\|^p\right]^{1/p}$$

$$\leq \frac{1}{2n} \sum_{i=1}^{n} \mathbf{E}_{\theta^\star} \left[\left| X_{ih_n} - X_{(i-1)h_n} \right|^{2p} \right]^{1/p}$$

$$+ \frac{C(p)}{2n} \mathbf{E}_{\theta^\star} \left[\left\| \sum_{i=1}^{n} \left(X_{ih_n} - X_{(i-1)h_n} \right) \varepsilon_{ih_n}^{\mathrm{T}} \right\|^{p} \right]^{1/p}$$

$$+ \frac{C(p)}{2n} \mathbf{E}_{\theta^\star} \left[\left\| \sum_{i=1}^{n} \left(X_{ih_n} - X_{(i-1)h_n} \right) \varepsilon_{(i-1)h_n}^{\mathrm{T}} \right\|^{p} \right]^{1/p}$$

$$+ C(p) \mathbf{E}_{\theta^\star} \left[\left\| \frac{1}{2n} \sum_{i=1}^{n} \varepsilon_{ih_n}^{\otimes 2} - \frac{1}{2} I_d \right\|^{p} \right]^{1/p}$$

$$+ C(p) \mathbf{E}_{\theta^\star} \left[\left\| \frac{1}{2n} \sum_{i=1}^{n} \varepsilon_{(i-1)h_n}^{\otimes 2} - \frac{1}{2} I_d \right\|^{p} \right]^{1/p}$$

$$+ \frac{C(p)}{n} \mathbf{E}_{\theta^\star} \left[\left\| \sum_{i=1}^{n} \varepsilon_{ih_n} \varepsilon_{(i-1)h_n}^{\mathrm{T}} \right\|^{p} \right]^{1/p}.$$

The first term of the right-hand side has the evaluation

$$\frac{1}{2n} \sum_{i=1}^{n} \mathbf{E}_{\theta^\star} \left[\left| X_{ih_n} - X_{(i-1)h_n} \right|^{2p} \right]^{1/p} \leq C(p) h_n.$$

We can evaluate the second term of the right-hand side

$$\mathbf{E}_{\theta^\star} \left[\left\| \sum_{i=1}^{n} \left(X_{ih_n} - X_{(i-1)h_n} \right) \varepsilon_{ih_n}^{\mathrm{T}} \right\|^{p} \right]$$

$$= \mathbf{E}_{\theta^\star} \left[\sum_{i_1} \cdots \sum_{i_{p/2}} \left\| \left(X_{i_1 h_n} - X_{(i_1-1)h_n} \right) \varepsilon_{i_1 h_n}^{\mathrm{T}} \right\|^2 \cdots \left\| \left(X_{i_{p/2} h_n} - X_{(i_{p/2}-1)h_n} \right) \varepsilon_{i_{p/2} h_n}^{\mathrm{T}} \right\|^2 \right]$$

$$\leq \sum_{i_1} \cdots \sum_{i_{p/2}} \mathbf{E}_{\theta^\star} \left[\left\| \left(X_{i_1 h_n} - X_{(i_1-1)h_n} \right) \varepsilon_{i_1 h_n}^{\mathrm{T}} \right\|^2 \cdots \left\| \left(X_{i_{p/2} h_n} - X_{(i_{p/2}-1)h_n} \right) \varepsilon_{i_{p/2} h_n}^{\mathrm{T}} \right\|^2 \right]$$

$$\leq \sum_{i_1} \cdots \sum_{i_{p/2}} C(p) h_n^{p/2} \leq C(p) (nh_n)^{p/2}$$

and hence

$$\frac{C(p)}{2n} \mathbf{E}_{\theta^\star} \left[\left\| \sum_{i=1}^{n} \left(X_{ih_n} - X_{(i-1)h_n} \right) \varepsilon_{ih_n}^{\mathrm{T}} \right\|^{p} \right]^{1/p} \leq C(p) \sqrt{\frac{h_n}{n}}.$$

The evaluation for the third term can be obtained in the same manner. For the fourth term, we have

$$C\left(p\right)\mathbf{E}_{\theta^{\star}}\left[\left\|\frac{1}{2n}\sum_{i=1}^{n}\varepsilon_{ih_n}^{\otimes2}-\frac{1}{2}I_d\right\|^{p}\right]^{1/p}$$

$$=C\left(p\right)\mathbf{E}_{\theta^{\star}}\left[\left\|\frac{1}{2n}\sum_{i=1}^{n}\left(\varepsilon_{ih_n}^{\otimes2}-I_d\right)\right\|^{p}\right]^{1/p}$$

$$=\frac{C\left(p\right)}{n}\mathbf{E}_{\theta^{\star}}\left[\sum_{i_1}\cdots\sum_{i_{p/2}}\left\|\varepsilon_{i_1h_n}^{\otimes2}-I_d\right\|^{2}\cdots\left\|\varepsilon_{i_{p/2}h_n}^{\otimes2}-I_d\right\|^{2}\right]^{1/p}$$

$$\leq\frac{C\left(p\right)}{\sqrt{n}},$$

and the same evaluation holds for the fifth term. Finally, we obtain

$$\frac{C\left(p\right)}{n}\mathbf{E}_{\theta^{\star}}\left[\left\|\sum_{i=1}^{n}\varepsilon_{ih_n}\varepsilon_{(i-1)h_n}^{\mathrm{T}}\right\|^{p}\right]^{1/p}$$

$$=\frac{C\left(p\right)}{n}\mathbf{E}_{\theta^{\star}}\left[\sum_{i_1}\cdots\sum_{i_{p/2}}\left\|\varepsilon_{i_1h_n}\varepsilon_{(i_1-1)h_n}^{\mathrm{T}}\right\|^{2}\cdots\left\|\varepsilon_{i_{p/2}h_n}\varepsilon_{(i_{p/2}-1)h_n}^{\mathrm{T}}\right\|^{2}\right]^{1/p}$$

$$\leq\frac{C\left(p\right)}{\sqrt{n}}.$$

Hence, the evaluation for L^p-norm stated above holds. □

Lemma 7 *For every function f such that $f\in C^{1}\left(\mathbf{R}^{d}\times\Xi;\ \mathbf{R}\right)$ and all the elements of f and the derivatives are polynomial growth with respect to x uniformly in ϑ,*

$$\mathbf{E}_{\theta^{\star}}\left[k_n^{\epsilon_1}\sup_{\vartheta\in\Xi}\left|\frac{1}{k_n}\sum_{j=1}^{k_n-2}\left(f\left(\bar{Y}_{j-1},\vartheta,\hat{\Lambda}_n\right)-f\left(\bar{Y}_{j-1},\vartheta,\Lambda_{\star}\right)\right)\right|^{p}\right]^{1/p}$$

$$\leq C\left(p\right)\left(n^{1/2}h_n^{3/2}+\frac{1}{\sqrt{p_n}}\right).$$

Proof We have

$$\mathbf{E}_{\theta^{\star}}\left[k_n^{\epsilon_1}\sup_{\vartheta\in\Xi}\left|\frac{1}{k_n}\sum_{j=1}^{k_n-2}\left(f\left(\bar{Y}_{j-1},\vartheta,\hat{\Lambda}_n\right)-f\left(\bar{Y}_{j-1},\vartheta,\Lambda_{\star}\right)\right)\right|^{p}\right]^{1/p}$$

$$\leq k_n^{\epsilon_1}\left(\frac{1}{k_n}\sum_{j=1}^{k_n-2}\mathbf{E}_{\theta^{\star}}\left[\sup_{\vartheta\in\Xi}\left|f\left(\bar{Y}_{j-1},\vartheta,\hat{\Lambda}_n\right)-f\left(\bar{Y}_{j-1},\vartheta,\Lambda_{\star}\right)\right|^{p}\right]\right)^{1/p}$$

$$\leq k_n^{\epsilon_1} \left(\frac{1}{k_n} \sum_{j=1}^{k_n-2} \mathbf{E}_{\theta^\star} \left[C \left(1 + \left| \bar{Y}_{j-1} \right| \right)^C \left\| \hat{\Lambda}_n - \Lambda_\star \right\|^p \right] \right)^{1/p}$$

$$\leq C\left(p\right) k_n^{\epsilon_1} \mathbf{E}_{\theta^\star} \left[\left\| \hat{\Lambda}_n - \Lambda_\star \right\|^{2p} \right]^{1/2p}$$

$$\leq C\left(p\right) \left(k_n^{\epsilon_1} h_n + \frac{k_n^{\epsilon_1}}{\sqrt{n}} \right) \leq C\left(p\right) \left(n^{1/2} h_n^{3/2} + \frac{1}{\sqrt{p_n}} \right).$$

$\square$

5.2 LAN for the quasi-likelihoods and proof for the main theorem

To prove the main theorem, we set some additional preliminary lemmas. Before the discussion, let us define the statistical random fields:

$$\mathbb{Y}_{1,n}^\tau \left(\alpha; \vartheta^\star \right) = \frac{1}{k_n} \left(\mathbb{H}_{1,n}^\tau \left(\alpha; \hat{\Lambda}_n \right) - \mathbb{H}_{1,n}^\tau \left(\alpha^\star; \hat{\Lambda}_n \right) \right)$$

$$= -\frac{1}{2k_n} \sum_{j=1}^{k_n-2} \left(\left(A_n^\tau \left(\bar{Y}_{j-1}, \alpha, \hat{\Lambda}_n \right)^{-1} \right. \right.$$

$$\left. - A_n^\tau \left(\bar{Y}_{j-1}, \alpha^\star, \hat{\Lambda}_n \right)^{-1} \right) \left[\left(\bar{Y}_{j+1} - \bar{Y}_j \right)^{\otimes 2} \right]$$

$$\left. \times \left(\frac{2}{3} \Delta_n \right)^{-1} + \log \frac{\det A_n^\tau \left(\bar{Y}_{j-1}, \alpha, \hat{\Lambda}_n \right)}{\det A_n^\tau \left(\bar{Y}_{j-1}, \alpha^\star, \hat{\Lambda}_n \right)} \right),$$

$$\mathbb{Y}_{2,n}^{\mathrm{ML}} \left(\beta; \vartheta^\star \right) = \frac{1}{k_n \Delta_n} \left(\mathbb{H}_{2,n} \left(\beta; \hat{\alpha}_n \right) - \mathbb{H}_{2,n} \left(\beta^\star; \hat{\alpha}_n \right) \right)$$

$$= \frac{1}{k_n \Delta_n} \left(\sum_{j=1}^{k_n-2} A \left(\bar{Y}_{j-1}, \hat{\alpha}_n \right)^{-1} \left[b \left(\bar{Y}_{j-1}, \beta \right) - b \left(\bar{Y}_{j-1}, \beta^\star \right), \bar{Y}_{j+1} - \bar{Y}_j \right] \right.$$

$$\left. - \frac{\Delta_n}{2} \sum_{j=1}^{k_n-2} A \left(\bar{Y}_{j-1}, \hat{\alpha}_n \right)^{-1} \left[b \left(\bar{Y}_{j-1}, \beta \right)^{\otimes 2} - b \left(\bar{Y}_{j-1}, \beta^\star \right)^{\otimes 2} \right] \right),$$

$$\mathbb{Y}_{2,n}^{\mathrm{Bayes}} \left(\beta; \vartheta^\star \right) = \frac{1}{k_n \Delta_n} \left(\mathbb{H}_{2,n} \left(\beta; \tilde{\alpha}_n \right) - \mathbb{H}_{2,n} \left(\beta^\star; \tilde{\alpha}_n \right) \right)$$

$$= \frac{1}{k_n \Delta_n} \left(\sum_{j=1}^{k_n-2} A \left(\bar{Y}_{j-1}, \tilde{\alpha}_n \right)^{-1} \left[b \left(\bar{Y}_{j-1}, \beta \right) - b \left(\bar{Y}_{j-1}, \beta^\star \right), \bar{Y}_{j+1} - \bar{Y}_j \right] \right.$$

$$\left. - \frac{\Delta_n}{2} \sum_{j=1}^{k_n-2} A \left(\bar{Y}_{j-1}, \tilde{\alpha}_n \right)^{-1} \left[b \left(\bar{Y}_{j-1}, \beta \right)^{\otimes 2} - b \left(\bar{Y}_{j-1}, \beta^\star \right)^{\otimes 2} \right] \right).$$

We give the locally asymptotic quadratic at $\vartheta^\star \in \Xi$ for $u_1 \in \mathbf{R}^{m_1}$ and $u_2 \in \mathbf{R}^{m_2}$,

$$\mathbb{Z}_{1,n}^{\tau}\left(u_1; \hat{\Lambda}_n, \alpha^\star\right) := \exp\left(\Delta_{1,n}^{\tau}\left(\vartheta^\star\right)[u_1] - \frac{1}{2}\Gamma_1^{\tau}\left(\vartheta^\star\right)\left[u_1^{\otimes 2}\right] + r_{1,n}^{\tau}\left(u; \vartheta^\star\right)\right),$$

$$\mathbb{Z}_{2,n}^{\mathrm{ML}}\left(u_2; \hat{\alpha}_n, \beta^\star\right) := \exp\left(\Delta_{2,n}^{\mathrm{ML}}\left(\vartheta^\star\right)[u_2] - \frac{1}{2}\Gamma_2^{\mathrm{ML}}\left(\vartheta^\star\right)\left[u_2^{\otimes 2}\right] + r_{2,n}^{\mathrm{ML}}\left(u; \vartheta^\star\right)\right),$$

$$\mathbb{Z}_{2,n}^{\mathrm{Bayes}}\left(u_2; \tilde{\alpha}_n, \beta^\star\right) := \exp\left(\Delta_{2,n}^{\mathrm{Bayes}}\left(\vartheta^\star\right)[u_2] - \frac{1}{2}\Gamma_2^{\mathrm{Bayes}}\left(\vartheta^\star\right)\left[u_2^{\otimes 2}\right] + r_{2,n}^{\mathrm{Bayes}}\left(u; \vartheta^\star\right)\right),$$

where

$$\Delta_{1,n}^{\tau}\left(\vartheta^\star\right)[u_1] := -\frac{1}{2k_n^{1/2}}\sum_{j=1}^{k_n-2}\left(\partial_\alpha A_n^{\tau}\left(\bar{Y}_{j-1}, \alpha^\star, \hat{\Lambda}_n\right)^{-1}\left[u_1, \left(\bar{Y}_{j+1} - \bar{Y}_j\right)^{\otimes 2}\right]\left(\frac{2\Delta_n}{3}\right)^{-1}\right.$$

$$\left. + \partial_\alpha \log \frac{\det A_n^{\tau}\left(\bar{Y}_{j-1}, \alpha^\star, \hat{\Lambda}_n\right)}{\det A_n^{\tau}\left(\bar{Y}_{j-1}, \alpha^\star, \hat{\Lambda}_n\right)}[u_1]\right),$$

$$\Delta_{2,n}^{\mathrm{ML}}\left(\vartheta^\star\right)[u_2] := \frac{1}{(k_n\Delta_n)^{1/2}}\sum_{j=1}^{k_n-2} A\left(\bar{Y}_{j-1}, \hat{\alpha}_n\right)^{-1}$$

$$\left[\partial_\beta b\left(\bar{Y}_{j-1}, \beta^\star\right)u_2, \bar{Y}_{j+1} - \bar{Y}_j - \Delta_n b\left(\bar{Y}_{j-1}, \beta^\star\right)\right],$$

$$\Delta_{2,n}^{\mathrm{Bayes}}\left(\vartheta^\star\right)[u_2] := \frac{1}{(k_n\Delta_n)^{1/2}}\sum_{j=1}^{k_n-2} A\left(\bar{Y}_{j-1}, \tilde{\alpha}_n\right)^{-1}$$

$$\left[\partial_\beta b\left(\bar{Y}_{j-1}, \beta^\star\right)u_2, \bar{Y}_{j+1} - \bar{Y}_j - \Delta_n b\left(\bar{Y}_{j-1}, \beta^\star\right)\right]$$

and

$$\Gamma_{1,n}^{\tau}\left(\alpha; \vartheta^\star\right)\left[u_1^{\otimes 2}\right]$$

$$:= \frac{1}{2k_n}\sum_{j=1}^{k_n-2}\left(\partial_\alpha^2 A_n^{\tau}\left(\bar{Y}_{j-1}, \alpha, \hat{\Lambda}_n\right)^{-1}\left[u_1^{\otimes 2}, \left(\bar{Y}_{j+1} - \bar{Y}_j\right)^{\otimes 2}\right]\left(\frac{2\Delta_n}{3}\right)^{-1}\right.$$

$$\left. + \partial_\alpha^2 \log \frac{\det A_n^{\tau}\left(\bar{Y}_{j-1}, \alpha, \hat{\Lambda}_n\right)}{\det A_n^{\tau}\left(\bar{Y}_{j-1}, \alpha^\star, \hat{\Lambda}_n\right)}\left[u_1^{\otimes 2}\right]\right),$$

$$\Gamma_{2,n}^{\mathrm{ML}}\left(\beta; \vartheta^\star\right)\left[u_2^{\otimes 2}\right]$$

$$:= \frac{1}{k_n\Delta_n}\sum_{j=1}^{k_n-2} A\left(\bar{Y}_{j-1}, \hat{\alpha}_n\right)^{-1}\left[\partial_\beta b\left(\bar{Y}_{j-1}, \beta\right)[u_2], \Delta_n b\left(\bar{Y}_{j-1}, \beta\right)[u_2]\right]$$

$$- \frac{1}{k_n\Delta_n}\sum_{j=1}^{k_n-2} A\left(\bar{Y}_{j-1}, \hat{\alpha}_n\right)^{-1}\left[\partial_\beta^2 b\left(\bar{Y}_{j-1}, \beta\right)\left[u_2^{\otimes 2}\right], \bar{Y}_{j+1} - \bar{Y}_j - \Delta_n b\left(\bar{Y}_{j-1}, \beta\right)\right],$$

$$\Gamma_{2,n}^{\mathrm{Bayes}}\left(\beta; \vartheta^\star\right)\left[u_2^{\otimes 2}\right]$$

$$
:= \frac{1}{k_n \Delta_n} \sum_{j=1}^{k_n-2} A\left(\bar{Y}_{j-1}, \tilde{\alpha}_n\right)^{-1} \left[\partial_\beta b\left(\bar{Y}_{j-1}, \beta\right)[u_2], \Delta_n b\left(\bar{Y}_{j-1}, \beta\right)[u_2]\right]
$$

$$
- \frac{1}{k_n \Delta_n} \sum_{j=1}^{k_n-2} A\left(\bar{Y}_{j-1}, \tilde{\alpha}_n\right)^{-1} \left[\partial_\beta^2 b\left(\bar{Y}_{j-1}, \beta\right)\left[u_2^{\otimes 2}\right], \bar{Y}_{j+1} - \bar{Y}_j - \Delta_n b\left(\bar{Y}_{j-1}, \beta\right)\right],
$$

and

$$
\Gamma_1^\tau\left(\vartheta^\star\right)\left[u_1^{\otimes 2}\right]
$$
$$
:= \frac{1}{2} \int_{\mathbf{R}^d} \left(\partial_\alpha^2 A^\tau\left(x, \alpha, \Lambda_\star\right)^{-1}\left[u_1^{\otimes 2}, A^\tau\left(x, \alpha^\star, \Lambda_\star\right)\right] + \partial_\alpha^2 \log \frac{\det A^\tau\left(x, \alpha, \Lambda_\star\right)}{\det A^\tau\left(x, \alpha^\star, \Lambda_\star\right)}\left[u_1^{\otimes 2}\right]\right)\Bigg|_{\alpha=\alpha^\star}
$$
$$
\times \nu\left(\mathrm{d}x\right), \Gamma_2\left(\vartheta^\star\right)\left[u_2^{\otimes 2}\right]
$$
$$
:= \frac{1}{2} \int_{\mathbf{R}^d} \left(A\left(x, \alpha\right)^{-1}\left[\partial_\beta b\left(x, \beta^\star\right)[u_2], \partial_\beta b\left(x, \beta^\star\right)[u_2]\right]\right) \nu\left(\mathrm{d}x\right),
$$

and

$$
r_{1,n}^\tau\left(u; \vartheta^\star\right) := \int_0^1 (1-s) \left\{\Gamma_1^\tau\left(\vartheta^\star\right)\left[u_1^{\otimes 2}\right] - \Gamma_{1,n}^\tau\left(\alpha^\star + sk_n^{-1/2}u_1; \vartheta^\star\right)\left[u_1^{\otimes 2}\right]\right\} \mathrm{d}s,
$$

$$
r_{2,n}^{\mathrm{ML}}\left(u; \vartheta^\star\right) := \int_0^1 (1-s) \left\{\Gamma_2^{\mathrm{ML}}\left(\vartheta^\star\right)\left[u_2^{\otimes 2}\right] - \Gamma_{2,n}^{\mathrm{ML}}\left(\beta^\star + sT_n^{-1/2}u_2; \vartheta^\star\right)\left[u_2^{\otimes 2}\right]\right\} \mathrm{d}s,
$$

$$
r_{2,n}^{\mathrm{Bayes}}\left(u; \vartheta^\star\right) := \int_0^1 (1-s) \left\{\Gamma_2^{\mathrm{Bayes}}\left(\vartheta^\star\right)\left[u_2^{\otimes 2}\right] - \Gamma_{2,n}^{\mathrm{Bayes}}\left(\beta^\star + sT_n^{-1/2}u_2; \vartheta^\star\right)\left[u_2^{\otimes 2}\right]\right\} \mathrm{d}s.
$$

We evaluate the moments of these random variables and fields in the following lemmas.

Lemma 8 (a) *For every $p > 1$,*

$$
\sup_{n \in \mathbf{N}} \mathbf{E}_{\theta^\star}\left[\left|\Delta_{1,n}^\tau\left(\vartheta^\star\right)\right|^p\right] < \infty.
$$

(b) *Let $\epsilon_1 = \epsilon_0/2$. Then, for every $p > 0$,*

$$
\sup_{n \in \mathbf{N}} \mathbf{E}_{\theta^\star}\left[\left(\sup_{\alpha \in \Theta_1} k_n^{\epsilon_1}\left|\mathbb{Y}_{1,n}^\tau\left(\alpha; \vartheta^\star\right) - \mathbb{Y}_1^\tau\left(\alpha; \vartheta^\star\right)\right|\right)^p\right] < \infty.
$$

Proof We start with the proof for (a). By Lemma 5, we obtain a decomposition

$$
\Delta_{1,n}^\tau\left(\vartheta^\star\right)[u_1] = M_{1,n}^\tau + R_{1,n}^{\tau(1)} + R_{1,n}^{\tau(2)} + R_{1,n}^{\tau(3)}
$$

for

$$M_{1,n}^{\tau} := -\frac{1}{2k_n^{1/2}} \sum_{j=1}^{k_n-2} \left(\partial_\alpha A_n^{\tau} \left(X_{j\Delta_n}, \alpha^{\star}, \Lambda_{\star} \right)^{-1} \left[u_1, \widehat{A_{j,n}^{\tau}} \right] + \partial_\alpha \log \det A_n^{\tau} \left(X_{j\Delta_n}, \alpha^{\star}, \Lambda_{\star} \right) [u_1] \right),$$

$$R_{1,n}^{\tau(1)} := \frac{1}{2k_n^{1/2}} \sum_{j=1}^{k_n-2} \left(\partial_\alpha A_n^{\tau} \left(X_{j\Delta_n}, \alpha^{\star}, \Lambda_{\star} \right)^{-1} \left[u_1, \widehat{A_{j,n}^{\tau}} \right] + \partial_\alpha \log \det A_n^{\tau} \left(X_{j\Delta_n}, \alpha^{\star}, \Lambda_{\star} \right) [u_1] \right)$$

$$- \frac{1}{2k_n^{1/2}} \sum_{j=1}^{k_n-2} \left(\partial_\alpha A_n^{\tau} \left(\bar{Y}_{j-1}, \alpha^{\star}, \Lambda_{\star} \right)^{-1} \left[u_1, \widehat{A_{j,n}^{\tau}} \right] + \partial_\alpha \log \det A_n^{\tau} \left(\bar{Y}_{j-1}, \alpha^{\star}, \Lambda_{\star} \right) [u_1] \right),$$

$$R_{1,n}^{\tau(2)} := \frac{1}{2k_n^{1/2}} \sum_{j=1}^{k_n-2} \left(\partial_\alpha A_n^{\tau} \left(\bar{Y}_{j-1}, \alpha^{\star}, \Lambda_{\star} \right)^{-1} \left[u_1, \widehat{A_{j,n}^{\tau}} \right] + \partial_\alpha \log \det A_n^{\tau} \left(\bar{Y}_{j-1}, \alpha^{\star}, \Lambda_{\star} \right) [u_1] \right)$$

$$- \frac{1}{2k_n^{1/2}} \sum_{j=1}^{k_n-2} \left(\partial_\alpha A_n^{\tau} \left(\bar{Y}_{j-1}, \alpha^{\star}, \hat{\Lambda}_n \right)^{-1} \left[u_1, \widehat{A_{j,n}^{\tau}} \right] + \partial_\alpha \log \det A_n^{\tau} \left(\bar{Y}_{j-1}, \alpha^{\star}, \hat{\Lambda}_n \right) [u_1] \right),$$

$$R_{1,n}^{\tau(3)} := -\frac{1}{2k_n^{1/2}} \sum_{j=1}^{k_n-2} \partial_\alpha A_n^{\tau} \left(\bar{Y}_{3j+i-1}, \alpha^{\star}, \hat{\Lambda}_n \right)^{-1} \left[u_1, \frac{3}{2\Delta_n} \left(\bar{Y}_{3j+i+1} - \bar{Y}_{3j+i} \right)^{\otimes 2} - \widehat{A_{3j+i,n}^{\tau}} \right],$$

where

$$\widehat{A_{j,n}^{\tau}} := \frac{1}{\Delta_n} \left[\frac{1}{\sqrt{m_n + m_n'}} a \left(X_{j\Delta_n}, \alpha^{\star} \right) \left(\zeta_{j+1,n} + \zeta_{j+2,n}' \right) + \sqrt{\frac{3}{2}} \left(\Lambda_{\star} \right)^{1/2} \left(\bar{\varepsilon}_{j+1} - \bar{\varepsilon}_j \right) \right]^{\otimes 2}$$

with the following property

$$\mathbf{E}_{\theta^{\star}} \left[\widehat{A_{j,n}^{\tau}} | \mathcal{H}_j^n \right] = A \left(X_{j\Delta_n}, \alpha^{\star} \right) + \frac{3}{p_n \Delta_n} \Lambda_{\star} = A \left(X_{j\Delta_n}, \alpha^{\star} \right) + 3\Delta_n^{\frac{2-\tau}{\tau-1}} \Lambda_{\star} = A_n^{\tau} \left(X_{j\Delta_n}, \alpha^{\star}, \Lambda_{\star} \right)$$

because of Lemma 1, $\Delta_n = p_n^{1-\tau}$, $\Delta_n^{\frac{1}{1-\tau}} = p_n$ and $(\Delta_n p_n)^{-1} = \Delta_n^{\frac{2-\tau}{\tau-1}}$. Furthermore, we have the L^p-boundedness such that

$$\mathbf{E}_{\theta^{\star}} \left[\left\| \widehat{A_{j,n}^{\tau}} - \frac{3}{2\Delta_n} \left[a \left(X_{j\Delta_n}, \alpha^{\star} \right) \left(\zeta_{j+1,n} + \zeta_{j+2,n}' \right) + \left(\Lambda_{\star} \right)^{1/2} \left(\bar{\varepsilon}_{j+1} - \bar{\varepsilon}_j \right) \right]^{\otimes 2} \right\|^p \right]^{1/p}$$

$$\leq \frac{1}{\Delta_n} \left| \frac{3}{2} - \frac{1}{m_n + m_n'} \right| \mathbf{E}_{\theta^{\star}} \left[\left\| a \left(X_{j\Delta_n}, \alpha^{\star} \right) \left(\zeta_{j+1,n} + \zeta_{j+2,n}' \right) \right\|^{2p} \right]^{1/p}$$

$$+ \frac{2}{\Delta_n} \left| \frac{3}{2} - \sqrt{\frac{3}{2}} \sqrt{\frac{1}{m_n + m_n'}} \right| \mathbf{E}_{\theta^{\star}} \left[\left\| a \left(X_{j\Delta_n}, \alpha^{\star} \right) \left(\zeta_{j+1,n} + \zeta_{j+2,n}' \right) \left(\bar{\varepsilon}_{j+1} - \bar{\varepsilon}_j \right)^{\mathrm{T}} \left(\Lambda_{\star} \right)^{1/2} \right\|^p \right]^{1/p}$$

$$\leq C(p) \left(\frac{3}{2} - \frac{1}{2/3 + 1/ \left(3p_n^2 \right)} \right) + \frac{C(p)}{\Delta_n^{1/2}} \left(1 - \sqrt{\frac{2/3}{2/3 + 1/ \left(3p_n^2 \right)}} \right)$$

$$\leq \frac{C(p)}{p_n^2} + \frac{C(p)}{\Delta_n^{1/2} p_n^{1/2}} \left(1 - \sqrt{1 - \frac{1}{1 + 2p_n^2}} \right) \leq \frac{C(p)}{p_n^2} + \frac{C(p)}{\Delta_n^{1/2} p_n^{5/2}} \leq C(p) \Delta_n^2$$

because of $\left\| \zeta_{j+1,n} + \zeta_{j+2,n} \right\|_p \le C\left(p\right) \Delta_n^{1/2}$ and $\left\| \bar{\varepsilon}_j \right\|_p = C\left(p\right) p_n^{-1/2}$ for all $j = 0, \ldots, k_n - 1$ and $n \in \mathbf{N}$, and the Taylor expansion for $f\left(x\right) = \sqrt{1+x}$ around $x = 0$. The L^p-boundedness of $R_{1,n}^{\tau(1)}$ is led by Lemma 4 and Burkholder's inequality for martingale, and that of $R_{1,n}^{\tau(2)}$ can be easily obtained by Lemma 6. With respect to $R_{1,n}^{\tau(3)}$, we decompose as $R_{1,n}^{\tau(3)} = \sum_{i=0}^{2} R_{i,1,n}^{\tau(3)}$ where

$$
R_{i,1,n}^{\tau(3)} = -\frac{1}{2k_n^{1/2}} \sum_{1 \le 3j+i \le k_n-2} \partial_\alpha A_n^\tau \left(\bar{Y}_{3j+i-1}, \alpha^\star, \hat{\Lambda}_n \right)^{-1}
$$
$$
\left[u_1, \left(\frac{2\Delta_n}{3} \right)^{-1} \left(\bar{Y}_{3j+i+1} - \bar{Y}_{3j+i} \right)^{\otimes 2} - \widehat{A_{3j+i,n}^\tau} \right].
$$

We only evaluate $R_{0,1,n}^{\tau(3)}$ and for the case p is an even number. The next inequality holds because of the L^p-boundedness shown above:

$$
\mathbf{E}_{\theta^\star} \left[\left| R_{0,1,n}^{\tau(3)} \right|^p \right]^{1/p}
$$
$$
= \mathbf{E}_{\theta^\star} \left[\left| \frac{1}{2k_n^{1/2}} \sum_{1 \le 3j \le k_n-2} \partial_\alpha A_n^\tau \left(\bar{Y}_{3j-1}, \alpha^\star, \hat{\Lambda}_n \right)^{-1} \left[u_1, \left(\frac{2\Delta_n}{3} \right)^{-1} \left(\bar{Y}_{3j+1} - \bar{Y}_{3j} \right)^{\otimes 2} - \widehat{A_{3j,n}^\tau} \right] \right|^p \right]^{1/p}
$$
$$
= \frac{3}{4k_n^{1/2}\Delta_n} \mathbf{E}_{\theta^\star} \left[\left| \sum_{1 \le 3j \le k_n-2} \partial_\alpha A_n^\tau \left(\bar{Y}_{3j-1}, \alpha^\star, \hat{\Lambda}_n \right)^{-1} \left[u_1, \left(\bar{Y}_{3j+1} - \bar{Y}_{3j} \right)^{\otimes 2} - \frac{2\Delta_n}{3} \widehat{A_{3j,n}^\tau} \right] \right|^p \right]^{1/p}
$$
$$
\le \frac{3}{4k_n^{1/2}\Delta_n} \mathbf{E}_{\theta^\star} \left[\left| \sum_{1 \le 3j \le k_n-2} \partial_\alpha A_n^\tau \left(\bar{Y}_{3j-1}, \alpha^\star, \hat{\Lambda}_n \right)^{-1} \left[u_1, \left(e_{3j,n} + \Delta_n b\left(X_{3j\Delta_n} \right) \right)^{\otimes 2} \right] \right|^p \right]^{1/p}
$$
$$
+ \frac{3}{2k_n^{1/2}\Delta_n} \mathbf{E}_{\theta^\star} \left[\left| \sum_{1 \le 3j \le k_n-2} \partial_\alpha A_n^\tau \left(\bar{Y}_{3j-1}, \alpha^\star, \hat{\Lambda}_n \right)^{-1} \right. \right.
$$
$$
\left. \left. \left[u_1, e_{3j,n} \left(a\left(X_{3j\Delta_n}, \alpha^\star \right) \left(\zeta_{3j+1,n} + \zeta'_{3j+2,n} \right) \right)^{\mathrm{T}} \right] \right|^p \right]^{1/p}
$$
$$
+ \frac{3}{2k_n^{1/2}\Delta_n} \mathbf{E}_{\theta^\star} \left[\left| \sum_{1 \le 3j \le k_n-2} \partial_\alpha A_n^\tau \left(\bar{Y}_{3j-1}, \alpha^\star, \hat{\Lambda}_n \right)^{-1} \right. \right.
$$
$$
\left. \left. \left[u_1, \Delta_n b\left(X_{3j\Delta_n} \right) \left(a\left(X_{3j\Delta_n}, \alpha^\star \right) \left(\zeta_{3j+1,n} + \zeta'_{3j+2,n} \right) \right)^{\mathrm{T}} \right] \right|^p \right]^{1/p}
$$
$$
+ \frac{3}{2k_n^{1/2}\Delta_n} \mathbf{E}_{\theta^\star} \left[\left| \sum_{1 \le 3j \le k_n-2} \partial_\alpha A_n^\tau \left(\bar{Y}_{3j-1}, \alpha^\star, \hat{\Lambda}_n \right)^{-1} \right. \right.
$$
$$
\left. \left. \left[u_1, \left(e_{3j,n} + \Delta_n b\left(X_{3j\Delta_n} \right) \right) \left(\left(\Lambda_\star \right)^{1/2} \left(\bar{\varepsilon}_{3j+1} - \bar{\varepsilon}_{3j} \right) \right)^{\mathrm{T}} \right] \right|^p \right]^{1/p}
$$
$$
+ o\left(1\right).
$$

We easily obtain the evaluation for the first term in the right-hand side

$$\frac{3}{4k_n^{1/2}\Delta_n}\mathbf{E}_{\theta^\star}\left[\left\|\sum_{1\le 3j\le k_n-2}\partial_\alpha A_n^\tau\left(\bar{Y}_{3j-1},\alpha^\star,\hat{\Lambda}_n\right)^{-1}\left[u_1,(e_{3j,n}+\Delta_n b\left(X_{3j\Delta_n}\right))^{\otimes 2}\right]\right\|^p\right]^{1/p}$$
$$\le C\left(p\right)|u|\,k_n^{1/2}\Delta_n\to 0,$$

and that for the second term

$$\frac{3}{2k_n^{1/2}\Delta_n}\mathbf{E}_{\theta^\star}\left[\left\|\sum_{1\le 3j\le k_n-2}\partial_\alpha A_n^\tau\left(\bar{Y}_{3j-1},\alpha^\star,\hat{\Lambda}_n\right)^{-1}\right.\right.$$
$$\left.\left.\left[u_1,e_{3j,n}\left(a\left(X_{3j\Delta_n},\alpha^\star\right)\left(\zeta_{3j+1,n}+\zeta'_{3j+2,n}\right)\right)^\mathrm{T}\right]\right\|^p\right]^{1/p}$$
$$=\frac{3}{2k_n^{1/2}\Delta_n}\mathbf{E}_{\theta^\star}\left[\left\|\sum_{1\le 3j\le k_n-2}\partial_\alpha A_n^\tau\left(\bar{Y}_{3j-1},\alpha^\star,\Lambda_\star\right)^{-1}\right.\right.$$
$$\left.\left.\left[u_1,e_{3j,n}\left(a\left(X_{3j\Delta_n},\alpha^\star\right)\left(\zeta_{3j+1,n}+\zeta'_{3j+2,n}\right)\right)^\mathrm{T}\right]\right\|^p\right]^{1/p}+o\left(1\right)$$
$$\le C\left(p\right)|u|\,k_n^{1/2}\Delta_n\to 0,$$

because of Lemmas 5 and 6. For the third term, we can replace $\hat{\Lambda}_n$ with $\Lambda_\star$ and $\bar{Y}_{3j-1}$ with $X_{3j\Delta_n}$ because of Lemma 6 and the result from combining Lemma 1 and Proposition 12 in Nakakita and Uchida (2019a), we denote $\eta_{3j,n}\left(u_1\right)=\left(a\left(X_{3j\Delta_n}\right)\right)^\mathrm{T}\left(\partial_\alpha A_n^\tau\left(X_{3j\Delta_n},\alpha^\star,\Lambda_\star\right)[u_1]\right)b\left(X_{3j\Delta_n}\right)$ which is a $\mathcal{H}_{3j}^n$-measurable random variable. Because of Lemma 1 and BDG inequality, we have

$$\frac{3}{2k_n^{1/2}}\mathbf{E}_{\theta^\star}\left[\left\|\sum_{1\le 3j\le k_n-2}\eta_{3j,n}\left(u_1\right)\left[\zeta_{3j+1,n}+\zeta'_{3j+2,n}\right]\right\|^p\right]^{1/p}$$
$$\le\frac{C\left(p\right)}{k_n^{1/2}}\mathbf{E}_{\theta^\star}\left[\left(\int_0^{k_n\Delta_n}\sum_{1\le 3j\le k_n-2}\left\|\eta_{3j,n}\left(u_1\right)\right\|^2\mathbf{1}_{[3j\Delta_n,(3j+1)\Delta_n]}\left(s\right)\mathrm{d}s\right)^{p/2}\right]^{1/p}$$
$$\le\frac{C\left(p\right)}{k_n^{1/2}}\mathbf{E}_{\theta^\star}\left[\left(\int_0^{k_n\Delta_n}\sum_{1\le 3j\le k_n-2}\left\|\eta_{3j,n}\left(u_1\right)\right\|^p\mathbf{1}_{[3j\Delta_n,(3j+1)\Delta_n]}\left(s\right)\mathrm{d}s\right)\left(\int_0^{k_n\Delta_n}\mathrm{d}s\right)^{p/2-1}\right]^{1/p}$$
$$=\frac{C\left(p\right)\left(k_n\Delta_n\right)^{1/2-1/p}}{k_n^{1/2}}\left(\int_0^{k_n\Delta_n}\sum_{1\le 3j\le k_n-2}\mathbf{E}_{\theta^\star}\left[\left\|\eta_{3j,n}\left(u_1\right)\right\|^p\right]\mathbf{1}_{[3j\Delta_n,(3j+1)\Delta_n]}\left(s\right)\mathrm{d}s\right)^{1/p}$$
$$\le\frac{C\left(p\right)\left(k_n\Delta_n\right)^{1/2}}{k_n^{1/2}}|u|\le C\left(p\right)\Delta_n^{1/2}\to 0.$$

It is obvious that the fourth term can be evaluated as bounded because $\{\varepsilon_{ih_n}\}$ is independent of X and i.i.d. Therefore, we obtain $\left\|R_{0,1,n}^{\tau(3)}\right\|_p<\infty$ and $\left\|R_{1,n}^{\tau(3)}\right\|_p<\infty$.

With respect to $M_{1,n}^\tau$, we utilise Burkholder's inequality for martingale: let us define $M_{i,1,n}^\tau$ for $i = 0, 1, 2$ as same as $R_{i,1,n}^{\tau(3)}$ and then

$$
\mathbf{E}_{\theta^\star}\left[\left|M_{0,1,n}^\tau\right|^p\right]
$$

$$
\leq C(p)\,\mathbf{E}_{\theta^\star}\left[\left|\frac{1}{4k_n}\sum_{1\leq 3j\leq k_n-2}\left|\partial_\alpha A_n^\tau\left(X_{j\Delta_n},\alpha^\star,\Lambda_\star\right)^{-1}\left[u_1, \widehat{A_{j,n}^\tau}\right]\right.\right.\right.
$$

$$
\left.\left.\left. + \partial_\alpha \log\det A_n^\tau\left(X_{j\Delta_n},\alpha^\star,\Lambda_\star\right)[u_1]\right|^2\right|^{p/2}\right]
$$

$$
\leq \frac{C(p)}{k_n}\sum_{1\leq 3j\leq k_n-2}\mathbf{E}_{\theta^\star}\left[\left|\partial_\alpha A_n^\tau\left(X_{j\Delta_n},\alpha^\star,\Lambda_\star\right)^{-1}\left[u_1,\widehat{A_{j,n}^\tau}\right]+\partial_\alpha\log\det A_n^\tau\left(X_{j\Delta_n},\alpha^\star,\Lambda_\star\right)[u_1]\right|^p\right]
$$

$$
< \infty
$$

because of the integrability.

In the next place, we give the proof for (b). Let us denote

$$
\mathbb{Y}_{1,n}^{\tau(\dagger)}(\alpha;\vartheta^\star) = -\frac{1}{2k_n}\sum_{j=1}^{k_n-2}\left(\left(A_n^\tau\left(\bar{Y}_{j-1},\alpha,\hat{\Lambda}_n\right)^{-1} - A_n^\tau\left(\bar{Y}_{j-1},\alpha^\star,\hat{\Lambda}_n\right)^{-1}\right)\right.
$$

$$
\left[A_n^\tau\left(X_{j\Delta_n},\alpha^\star,\Lambda_\star\right)\right]
$$

$$
\left.+ \log\frac{\det A_n^\tau\left(\bar{Y}_{j-1},\alpha,\hat{\Lambda}_n\right)}{\det A_n^\tau\left(\bar{Y}_{j-1},\alpha^\star,\hat{\Lambda}_n\right)}\right).
$$

Define $R_{1,n}^{\tau(\dagger)}$ by $R_{1,n}^{\tau(\dagger)} = \mathbb{Y}_{1,n}^\tau(\alpha;\vartheta^\star) - \mathbb{Y}_{1,n}^{\tau(\dagger)}(\alpha;\vartheta^\star) - M_{1,n}^{\tau(\dagger)}$ for

$$
M_{1,n}^{\tau(\dagger)} = -\frac{1}{2k_n}\sum_{j=1}^{k_n-2}\left(\left(A_n^\tau\left(\bar{Y}_{j-1},\alpha,\hat{\Lambda}_n\right)^{-1} - A_n^\tau\left(\bar{Y}_{j-1},\alpha^\star,\hat{\Lambda}_n\right)^{-1}\right)\right.
$$

$$
\left.\left[\widehat{A_{j,n}^\tau} - A_n^\tau\left(X_{j\Delta_n},\alpha^\star,\Lambda_\star\right)\right]\right).
$$

Firstly, we show L^p-boundedness of $k_n^{\epsilon_1}R_{1,n}^{\tau(\dagger)}$ uniformly for n and α for every p. We have the representation such that

$$
R_{1,n}^{\tau(\dagger)} = -\frac{1}{2k_n}\sum_{j=1}^{k_n-2}\left(\left(A_n^\tau\left(\bar{Y}_{j-1},\alpha,\hat{\Lambda}_n\right)^{-1} - A_n^\tau\left(\bar{Y}_{j-1},\alpha^\star,\hat{\Lambda}_n\right)^{-1}\right)\right.
$$

$$
\left[\left(\bar{Y}_{j+1}-\bar{Y}_j\right)^{\otimes 2}\right]\left(\frac{2}{3}\Delta_n\right)^{-1}
$$

$$
\left.+ \log\frac{\det A_n^\tau\left(\bar{Y}_{j-1},\alpha,\hat{\Lambda}_n\right)}{\det A_n^\tau\left(\bar{Y}_{j-1},\alpha^\star,\hat{\Lambda}_n\right)}\right)
$$

$$+ \frac{1}{2k_n} \sum_{j=1}^{k_n-2} \left(\left(A_n^\tau \left(\bar{Y}_{j-1}, \alpha, \hat{\Lambda}_n \right)^{-1} - A_n^\tau \left(\bar{Y}_{j-1}, \alpha^\star, \hat{\Lambda}_n \right)^{-1} \right) \right.$$

$$\left[A_n^\tau \left(X_{j\Delta_n}, \alpha^\star, \Lambda_\star \right) \right]$$

$$\left. + \log \frac{\det A_n^\tau \left(\bar{Y}_{j-1}, \alpha, \hat{\Lambda}_n \right)}{\det A_n^\tau \left(\bar{Y}_{j-1}, \alpha^\star, \hat{\Lambda}_n \right)} \right)$$

$$+ \frac{1}{2k_n} \sum_{j=1}^{k_n-2} \left(\left(A_n^\tau \left(\bar{Y}_{j-1}, \alpha, \hat{\Lambda}_n \right)^{-1} - A_n^\tau \left(\bar{Y}_{j-1}, \alpha^\star, \hat{\Lambda}_n \right)^{-1} \right) \right.$$

$$\left. \left[\widehat{A_{j,n}^\tau} - A_n^\tau \left(X_{j\Delta_n}, \alpha^\star, \Lambda_\star \right) \right] \right)$$

$$= - \frac{1}{2k_n} \sum_{j=1}^{k_n-2} \left(\left(A_n^\tau \left(\bar{Y}_{j-1}, \alpha, \hat{\Lambda}_n \right)^{-1} - A_n^\tau \left(\bar{Y}_{j-1}, \alpha^\star, \hat{\Lambda}_n \right)^{-1} \right) \right.$$

$$\left. \left[(\bar{Y}_{j+1} - \bar{Y}_j)^{\otimes 2} \right] \left(\frac{2}{3} \Delta_n \right)^{-1} \right)$$

$$+ \frac{1}{2k_n} \sum_{j=1}^{k_n-2} \left(\left(A_n^\tau \left(\bar{Y}_{j-1}, \alpha, \hat{\Lambda}_n \right)^{-1} - A_n^\tau \left(\bar{Y}_{j-1}, \alpha^\star, \hat{\Lambda}_n \right)^{-1} \right) \right.$$

$$\left. \left[A_n^\tau \left(X_{j\Delta_n}, \alpha^\star, \Lambda_\star \right) \right] \right)$$

$$+ \frac{1}{2k_n} \sum_{j=1}^{k_n-2} \left(\left(A_n^\tau \left(\bar{Y}_{j-1}, \alpha, \hat{\Lambda}_n \right)^{-1} - A_n^\tau \left(\bar{Y}_{j-1}, \alpha^\star, \hat{\Lambda}_n \right)^{-1} \right) \right.$$

$$\left. \left[\widehat{A_{j,n}^\tau} - A_n^\tau \left(X_{j\Delta_n}, \alpha^\star, \Lambda_\star \right) \right] \right)$$

$$= - \frac{1}{2k_n} \sum_{j=1}^{k_n-2} \left(A_n^\tau \left(\bar{Y}_{j-1}, \alpha, \hat{\Lambda}_n \right)^{-1} - A_n^\tau \left(\bar{Y}_{j-1}, \alpha^\star, \hat{\Lambda}_n \right)^{-1} \right)$$

$$\times \left[\left(\frac{2}{3} \Delta_n \right)^{-1} (\bar{Y}_{j+1} - \bar{Y}_j)^{\otimes 2} - \widehat{A_{j,n}^\tau} \right].$$

Because of Lemma 5, the following evaluation holds:

$$\left\| \left(\frac{2}{3} \Delta_n \right)^{-1} (\bar{Y}_{j+1} - \bar{Y}_j)^{\otimes 2} - \widehat{A_{j,n}^\tau} \right\|_p$$

$$\leq \left\| \left(\frac{2}{3} \Delta_n \right)^{-1} \left[(\bar{Y}_{j+1} - \bar{Y}_j)^{\otimes 2} - \left(a \left(X_{j\Delta_n}, \alpha^\star \right) \left(\zeta_{j+1,n} + \zeta'_{j+2,n} \right) + (\Lambda_\star)^{1/2} (\bar{\varepsilon}_{j+1} - \bar{\varepsilon}_j) \right)^{\otimes 2} \right] \right\|_p$$

$$+ C(p) \Delta_n$$

$$= \left(\frac{2}{3}\Delta_n\right)^{-1} \left\| e_{j,n}^{\otimes 2} - e_{j,n}\left(a\left(X_{j\Delta_n},\alpha^\star\right)\left(\zeta_{j+1,n}+\zeta_{j+2,n}'\right)+(\Lambda_\star)^{1/2}\left(\bar{\varepsilon}_{j+1}-\bar{\varepsilon}_j\right)\right)^{\mathsf{T}} \right.$$
$$\left. - \left(a\left(X_{j\Delta_n},\alpha^\star\right)\left(\zeta_{j+1,n}+\zeta_{j+2,n}'\right)+(\Lambda_\star)^{1/2}\left(\bar{\varepsilon}_{j+1}-\bar{\varepsilon}_j\right)\right)e_{j,n}^{\mathsf{T}} \right\|_p$$
$$+ C\left(p\right)\Delta_n$$
$$\leq \left(\frac{2}{3}\Delta_n\right)^{-1}\left(\left\|e_{j,n}\right\|_p^2 + 2\left\|a\left(X_{j\Delta_n},\alpha^\star\right)\left(\zeta_{j+1,n}+\zeta_{j+2,n}'\right)+(\Lambda_\star)^{1/2}\left(\bar{\varepsilon}_{j+1}-\bar{\varepsilon}_j\right)\right\|_{2p}\left\|e_{j,n}\right\|_{2p}\right)$$
$$+ C\left(p\right)\Delta_n$$
$$\leq C\left(p\right)\left(\Delta_n + \Delta_n^{1/2}\right).$$

Hence, we have the evaluation $\sup_{\alpha\in\Theta_1}\sup_{n\in\mathbf{N}}\left\|R_{1,n}^{\tau(\dagger)}\right\|_p \leq C\left(p\right)\Delta_n + C\left(p\right)\Delta_n^{1/2} \leq C\Delta_n^{1/2}$, and hence,

$$\sup_{\alpha\in\Theta_1}\sup_{n\in\mathbf{N}}\left\|k_n^{\epsilon_1}R_{1,n}^{\tau(\dagger)}\right\|_p \leq C\left(p\right)k_n^{\epsilon_1}\Delta_n^{1/2} = C\left(p\right)\left(k_n^{\epsilon_0}\Delta_n\right)^{1/2} \to 0.$$

In the next place, we see the same uniform L^p-boundedness of $k_n^{\epsilon_1}M_{1,n}^{\tau(\dagger)}$ for every p. As the approximation, we set $M_{1,n}^{\tau(\ddagger)} := \sum_{i=0}^2 M_{i,1,n}^{\tau(\ddagger)}$ where for $i=0,1,2$, $M_{i,1,n}^{\tau(\ddagger)} := -\frac{1}{2k_n}\sum_{1\leq 3j+i\leq k_n-2}\mu_{3j+i,n}$,
where

$$\mu_{3j+i,n} = \left(A_n^\tau\left(\bar{Y}_{3j+i-1},\alpha,\Lambda_\star\right)^{-1} - A_n^\tau\left(\bar{Y}_{3j+i-1},\alpha^\star,\Lambda_\star\right)^{-1}\right)$$
$$\left[\widehat{A_{3j+i,n}^\tau} - A_n^\tau\left(X_{(3j+i)\Delta_n},\alpha^\star,\Lambda_\star\right)\right].$$

It is easy to show $\mathbf{E}\left[\sup_{\alpha\in\Theta_1}k_n^{\epsilon_1}\left|M_{1,n}^{\tau(\dagger)}-M_{1,n}^{\tau(\ddagger)}\right|^p\right]^{1/p} \leq C\left(p\right)k_n^{\epsilon_1}n^{-1/2} \leq C\left(p\right)h_n^{1/2} \to 0$ for Lemma 6. For simplicity, we only evaluate $k_n^{\epsilon_1}M_{0,1,n}^{\tau(\ddagger)}$. We have for all p,

$$\mathbf{E}_{\theta^\star}\left[\left|\mu_{3j,n}\right|^p\right]$$
$$= \mathbf{E}_{\theta^\star}\left[\left|\left(A_n^\tau\left(\bar{Y}_{3j-1},\alpha,\Lambda_\star\right)^{-1} - A_n^\tau\left(\bar{Y}_{3j-1},\alpha^\star,\Lambda_\star\right)^{-1}\right)\left[\widehat{A_{3j,n}^\tau} - A_n^\tau\left(X_{3j\Delta_n},\alpha^\star,\Lambda_\star\right)\right]\right|^p\right]$$
$$\leq \mathbf{E}_{\theta^\star}\left[\left\|A_n^\tau\left(\bar{Y}_{3j-1},\alpha,\Lambda_\star\right)^{-1} - A_n^\tau\left(\bar{Y}_{3j-1},\alpha^\star,\Lambda_\star\right)^{-1}\right\|^p\left\|\widehat{A_{3j,n}^\tau} - A_n^\tau\left(X_{3j\Delta_n},\alpha^\star,\Lambda_\star\right)\right\|^p\right]$$
$$\leq C\left(p\right)\mathbf{E}_{\theta^\star}\left[\left\|\widehat{A_{3j,n}^\tau} - A_n^\tau\left(X_{3j\Delta_n},\alpha^\star,\Lambda_\star\right)\right\|^{2p}\right]^{1/2}$$
$$\leq C\left(p\right).$$

Hence, by Burkholder's inequality, for all p,

$$
\begin{aligned}
\mathbf{E}_{\theta^\star}\left[\left|k_n^{\epsilon_1} M_{0,1,n}^{\tau(\ddagger)}\right|^p\right] &\leq C\left(p\right) k_n^{\epsilon_1 p} \mathbf{E}_{\theta^\star}\left[\left|\frac{1}{k_n^2} \sum_{1\leq 3j\leq k_n-2} \mu_{3j,n}^2\right|^{p/2}\right] \\
&\leq C\left(p\right) k_n^{\epsilon_1 p} k_n^{-p/2} \frac{1}{k_n} \sum_{1\leq 3j\leq k_n-2} \mathbf{E}_{\theta^\star}\left[\left|\mu_{3j,n}^2\right|^{p/2}\right] \\
&\leq C\left(p\right) k_n^{(\epsilon_1-1/2)p} \frac{1}{k_n} \sum_{1\leq 3j\leq k_n-2} \mathbf{E}_{\theta^\star}\left[\left|\mu_{3j,n}\right|^p\right] \\
&\leq C\left(p\right) k_n^{(\epsilon_1-1/2)p}
\end{aligned}
$$

and then $\sup_{n,\theta^\star}\left\|k_n^{\epsilon_1} M_{1,n}^{\tau(\ddagger)}\right\|_p < \infty$. With the same procedure, we obtain the uniform L^p-boundedness of $k_n^{\epsilon_1} \partial_\alpha R_{1,n}^{\tau(\dagger)}$ and $k_n^{\epsilon_1} \partial_\alpha M_{1,n}^{\tau(\ddagger)}$. Sobolev's inequality leads to $\sup_{n\in\mathbf{N}}\left\|\sup_{\alpha\in\Theta_1}\left|k_n^{\epsilon_1} R_{1,n}^{\tau(\dagger)}\right|\right\|_p < \infty$ and $\sup_{n\in\mathbf{N}}\left\|\sup_{\alpha\in\Theta_1}\left|k_n^{\epsilon_1} M_{1,n}^{\tau(\ddagger)}\right|\right\|_p < \infty$, and then $\sup_{n\in\mathbf{N}}\left\|\sup_{\alpha\in\Theta_1}\left|k_n^{\epsilon_1} M_{1,n}^{\tau(\dagger)}\right|\right\|_p < \infty$. Note that for

$$
\begin{aligned}
\mathbb{Y}_{1,n}^{\tau(\ddagger)}\left(\alpha;\vartheta^\star\right) = &-\frac{1}{2k_n} \sum_{j=1}^{k_n-2}\left(\left(A_n^\tau\left(X_{j\Delta_n},\alpha,\Lambda_\star\right)^{-1} - A_n^\tau\left(X_{j\Delta_n},\alpha^\star,\Lambda_\star\right)^{-1}\right)\left[A_n^\tau\left(X_{j\Delta_n},\alpha^\star,\Lambda_\star\right)\right]\right. \\
&\left. + \log\frac{\det A_n^\tau\left(X_{j\Delta_n},\alpha,\Lambda_\star\right)}{\det A_n^\tau\left(X_{j\Delta_n},\alpha^\star,\Lambda_\star\right)}\right),
\end{aligned}
$$

we can evaluate $\sup_{n\in\mathbf{N}}$ $\left\|\sup_{\alpha\in\Theta_1}\left|k_n^{\epsilon_1}\left(\mathbb{Y}_{1,n}^{\tau(\ddagger)}\left(\alpha;\vartheta^\star\right)\left(\alpha;\vartheta^\star\right) - \mathbb{Y}_{1,n}^{\tau(\dagger)}\left(\alpha;\vartheta^\star\right)\left(\alpha;\vartheta^\star\right)\right)\right|\right\|_p < \infty$ because of Lemmas 3 and 7. Hence, the discussion of Remark 4 leads to the proof. □

Lemma 9 (a) *For any $M_3 > 0$,*

$$
\sup_{n\in\mathbf{N}} \mathbf{E}_{\theta^\star}\left[\left(k_n^{-1} \sup_{\vartheta\in\Xi}\left|\partial_\alpha^3 \mathbb{H}_{1,n}^\tau\left(\alpha;\Lambda\right)\right|\right)^{M_3}\right] < \infty.
$$

(b) *Let $\epsilon_1 = \epsilon_0/2$. Then for $M_4 > 0$,*

$$
\sup_{n\in\mathbf{N}} \mathbf{E}_{\theta^\star}\left[\left(k_n^{\epsilon_1}\left|\Gamma_{1,n}^\tau\left(\alpha^\star;\vartheta^\star\right) - \Gamma_1^\tau\left(\vartheta^\star\right)\right|\right)^{M_4}\right] < \infty.
$$

Proof With respect to (a), we have

$$\sup_{\vartheta \in \Xi} \left| \partial_\alpha^3 \mathbb{H}_{1,n}^\tau (\alpha; \Lambda) \right|$$

$$= \sup_{\vartheta \in \Xi} \left| \frac{1}{2} \partial_\alpha^3 \sum_{j=1}^{k_n-2} \left(\left(\frac{2}{3} \Delta_n A_n^\tau \left(\bar{Y}_{j-1}, \alpha, \Lambda \right) \right)^{-1} \left[(\bar{Y}_{j+1} - \bar{Y}_j)^{\otimes 2} \right] + \log \det A_n^\tau \left(\bar{Y}_{j-1}, \alpha, \Lambda \right) \right) \right|$$

$$\leq \sup_{\vartheta \in \Xi} \left| \sum_{j=1}^{k_n-2} \partial_\alpha^3 \left(A_n^\tau \left(\bar{Y}_{j-1}, \alpha, \Lambda \right) \right)^{-1} \left[\frac{3}{4\Delta_n} (\bar{Y}_{j+1} - \bar{Y}_j)^{\otimes 2} \right] \right|$$

$$+ \sup_{\vartheta \in \Xi} \frac{1}{2} \sum_{j=1}^{k_n-2} \left| \partial_\alpha^3 \log \det A_n^\tau \left(\bar{Y}_{j-1}, \alpha, \Lambda \right) \right|$$

$$\leq C \sum_{j=1}^{k_n-2} \left(1 + |\bar{Y}_{j-1}| \right)^C \Delta_n^{-1} \left| \bar{Y}_{j+1} - \bar{Y}_j \right|^2 + C \sum_{j=1}^{k_n-2} \left(1 + |\bar{Y}_{j-1}| \right)^C$$

and hence

$$\sup_{n \in \mathbf{N}} \mathbf{E}_{\theta^\star} \left[\left(k_n^{-1} \sup_{\vartheta \in \Xi} \left| \partial_\alpha^3 \mathbb{H}_{1,n}^\tau (\alpha; \Lambda) \right| \right)^{M_3} \right]$$

$$\leq C \sup_{n \in \mathbf{N}} \mathbf{E}_{\theta^\star} \left[\left(k_n^{-1} \sum_{j=1}^{k_n-2} \left(1 + |\bar{Y}_{j-1}| \right)^C \Delta_n^{-1} \left| \bar{Y}_{j+1} - \bar{Y}_j \right|^2 \right)^{M_3} \right]$$

$$+ C \sup_{n \in \mathbf{N}} \mathbf{E}_{\theta^\star} \left[\left(k_n^{-1} \sum_{j=1}^{k_n-2} \left(1 + |\bar{Y}_{j-1}| \right)^C \right)^{M_3} \right]$$

$$\leq C \sup_{n \in \mathbf{N}} k_n^{-1} \sum_{j=1}^{k_n-2} \mathbf{E}_{\theta^\star} \left[1 + \left| X_{j\Delta_n} \right|^C \right] < \infty.$$

For (b), the discussion same as Lemma 8 leads to the result. $\qquad \square$

Proposition 1 *For any $p > 0$,*

$$\sup_{n \in \mathbf{N}} \mathbf{E}_{\theta^\star} \left[\left| \sqrt{k_n} \left(\hat{\alpha}_n - \alpha^\star \right) \right|^p \right] < \infty, \quad \sup_{n \in \mathbf{N}} \mathbf{E}_{\theta^\star} \left[\left| \sqrt{k_n} \left(\tilde{\alpha}_n - \alpha^\star \right) \right|^p \right] < \infty.$$

Proof Theorem 3 in Yoshida (2011), Lemmas 8 and 9 lead to the following polynomial large deviation inequality $P_{\theta^\star} \left[\sup_{u_1 \in V_{1,n}^\tau (r, \alpha^\star)} \mathbb{Z}_{1,n}^\tau \left(u_1; \hat{\Lambda}_n, \alpha^\star \right) \geq e^{-r} \right] \leq \frac{C(L)}{r^L}$ for all $r > 0$ and $n \in \mathbf{N}$. The L^p-boundedness of $\sqrt{k_n} \left(\hat{\alpha}_n - \alpha^\star \right)$ is then obtained with the discussion parallel to Yoshida (2011).

With respect to the Bayes-type estimator, we need to verify the next boundedness: there exists $\delta_1 > 0$ and $C > 0$ such that $\sup_{n \in \mathbf{N}} \mathbf{E}_{\theta^\star} \left[\left(\int_{u_1 : |u_1| \leq \delta_1} \mathbb{Z}_{1,n}^\tau \left(u_1; \hat{\Lambda}_n, \alpha^\star \right) du_1 \right)^{-1} \right] < \infty$. Because of the Lemma 2 in Yoshida (2011), it is sufficient to show that

for some $p > d, \delta > 0$ and $C > 0$, $\sup_{n \in \mathbf{N}} \mathbf{E}_{\theta^\star} \left[\left| \log \mathbb{Z}_{1,n}^\tau \left(u_1; \hat{\Lambda}_n, \alpha^\star \right) \right|^p \right] \leq C |u_1|^p$ for all u_1 such that $|u_1| \leq \delta$ and actually it is easy to obtain by Lemmas 8 and 9. $\quad \square$

Lemma 10 (a) *For every $p > 0$,*

$$\sup_{n \in \mathbf{N}} \mathbf{E}_{\theta^\star} \left[\left| \Delta_{2,n}^{\mathrm{ML}} \left(\vartheta^\star \right) \right|^p \right] < \infty, \quad \sup_{n \in \mathbf{N}} \mathbf{E}_{\theta^\star} \left[\left| \Delta_{2,n}^{\mathrm{Bayes}} \left(\vartheta^\star \right) \right|^p \right] < \infty.$$

(b) *Let $\epsilon_1 = \epsilon_0 / 2$. Then, for every $p > 0$,*

$$\sup_{n \in \mathbf{N}} \left\| \sup_{\beta \in \Theta_2} (k_n \Delta_n)^{\epsilon_1} \left| \mathbb{Y}_{2,n}^{\mathrm{ML}} \left(\beta; \vartheta^\star \right) - \mathbb{Y}_2 \left(\beta; \vartheta^\star \right) \right| \right\|_p < \infty,$$

$$\sup_{n \in \mathbf{N}} \left\| \sup_{\beta \in \Theta_2} (k_n \Delta_n)^{\epsilon_1} \left| \mathbb{Y}_{2,n}^{\mathrm{Bayes}} \left(\beta; \vartheta^\star \right) - \mathbb{Y}_2 \left(\beta; \vartheta^\star \right) \right| \right\|_p < \infty.$$

Proof We only show the proof for $\Delta_{2,n}^{\mathrm{ML}}$ and $\mathbb{Y}_{2,n}^{\mathrm{ML}}$ since the proof for $\Delta_{2,n}^{\mathrm{Bayes}}$ and $\mathbb{Y}_{2,n}^{\mathrm{Bayes}}$ is quite parallel. For (a), we decompose
$\Delta_{2,n}^{\mathrm{ML}} \left(\vartheta^\star \right) [u_2] = M_{2,n}^{\mathrm{ML}} + R_{2,n}^{\mathrm{ML}}$,
where

$$M_{2,n}^{\mathrm{ML}} = \frac{1}{(k_n \Delta_n)^{1/2}} \sum_{j=1}^{k_n-2} A \left(\bar{Y}_{j-1}, \hat{\alpha}_n \right)^{-1} \left[\partial_\beta b \left(\bar{Y}_{j-1}, \beta^\star \right) u_2, a \left(X_{j\Delta_n} \right) \left(\zeta_{j+1,n} + \zeta'_{j+2,n} \right) \right]$$

$$+ \frac{1}{(k_n \Delta_n)^{1/2}} \sum_{j=1}^{k_n-2} A \left(\bar{Y}_{j-1}, \hat{\alpha}_n \right)^{-1} \left[\partial_\beta b \left(\bar{Y}_{j-1}, \beta^\star \right) u_2, (\Lambda_\star)^{1/2} \left(\bar{\varepsilon}_{j+1} - \bar{\varepsilon}_j \right) \right],$$

$$R_{2,n}^{\mathrm{ML}} = \frac{\Delta_n}{(k_n \Delta_n)^{1/2}} \sum_{j=1}^{k_n-2} A \left(\bar{Y}_{j-1}, \hat{\alpha}_n \right)^{-1} \left[\partial_\beta b \left(\bar{Y}_{j-1}, \beta^\star \right) u_2, b \left(X_{j\Delta_n} \right) - b \left(\bar{Y}_{j-1} \right) \right]$$

$$+ \frac{1}{(k_n \Delta_n)^{1/2}} \sum_{j=1}^{k_n-2} A \left(\bar{Y}_{j-1}, \hat{\alpha}_n \right)^{-1} \left[\partial_\beta b \left(\bar{Y}_{j-1}, \beta^\star \right) u_2, e_{j,n} \right].$$

We can use L^p-boundedness of $\sqrt{k_n} \left(\hat{\alpha}_n - \alpha^\star \right)$ and Burkholder's inequality; then, we obtain $\sup_{n \in \mathbf{N}} \mathbf{E}_{\theta^\star} \left[\left| M_{2,n}^{\mathrm{ML}} \right|^p \right]^{1/p} \leq C(p)$, and for the residuals, Lemmas 3 and 5 lead to $\mathbf{E}_{\theta^\star} \left[\left| R_{2,n}^{\mathrm{ML}} \right|^p \right]^{1/p} \leq C(p) \sqrt{k_n} \Delta_n \to 0$. Then, we obtain (a). We prove (b) in the second place. We decompose $\mathbb{Y}_{2,n}^{\mathrm{ML}} \left(\beta; \vartheta^\star \right)$ as $\mathbb{Y}_{2,n}^{\mathrm{ML}} \left(\beta; \vartheta^\star \right) = M_{2,n}^{\mathrm{ML}(\dagger)} \left(\hat{\alpha}_n, \beta \right) + R_{2,n}^{\mathrm{ML}(\dagger)} \left(\hat{\alpha}_n, \beta \right) + \mathbb{Y}_{2,n}^{\mathrm{ML}(\dagger)} \left(\beta; \vartheta^\star \right)$, where

$$M_{2,n}^{\mathrm{ML}(\dagger)}(\alpha, \beta) = \frac{1}{k_n \Delta_n} \sum_{j=1}^{k_n-2} A\left(\bar{Y}_{j-1}, \alpha\right)^{-1} \left[b\left(\bar{Y}_{j-1}, \beta\right), a\left(X_{j\Delta_n}\right)\left(\zeta_{j+1,n} + \zeta_{j+2,n}\right)\right]$$

$$- \frac{1}{k_n \Delta_n} \sum_{j=1}^{k_n-2} A\left(\bar{Y}_{j-1}, \alpha\right)^{-1} \left[b\left(\bar{Y}_{j-1}, \beta^\star\right), a\left(X_{j\Delta_n}\right)\left(\zeta_{j+1,n} + \zeta_{j+2,n}\right)\right],$$

$$+ \frac{1}{k_n \Delta_n} \sum_{j=1}^{k_n-2} A\left(\bar{Y}_{j-1}, \alpha\right)^{-1} \left[b\left(\bar{Y}_{j-1}, \beta\right), (\Lambda_\star)^{1/2} \left(\bar{\varepsilon}_{j+1} - \bar{\varepsilon}_j\right)\right]$$

$$- \frac{1}{k_n \Delta_n} \sum_{j=1}^{k_n-2} A\left(\bar{Y}_{j-1}, \alpha\right)^{-1} \left[b\left(\bar{Y}_{j-1}, \beta^\star\right), (\Lambda_\star)^{1/2} \left(\bar{\varepsilon}_{j+1} - \bar{\varepsilon}_j\right)\right],$$

$$R_{2,n}^{\mathrm{ML}(\dagger)}(\alpha, \beta) = \frac{1}{k_n \Delta_n} \sum_{j=1}^{k_n-2} A\left(\bar{Y}_{j-1}, \alpha\right)^{-1} \left[b\left(\bar{Y}_{j-1}, \beta\right), e_{j,n}\right]$$

$$- \frac{1}{k_n \Delta_n} \sum_{j=1}^{k_n-2} A\left(\bar{Y}_{j-1}, \alpha\right)^{-1} \left[b\left(\bar{Y}_{j-1}, \beta^\star\right), e_{j,n}\right]$$

$$+ \frac{1}{k_n} \sum_{j=1}^{k_n-2} A\left(\bar{Y}_{j-1}, \alpha\right)^{-1} \left[b\left(\bar{Y}_{j-1}, \beta\right), b\left(X_{j\Delta_n}, \alpha^\star\right) - b\left(\bar{Y}_{j-1}, \alpha^\star\right)\right]$$

$$- \frac{1}{k_n} \sum_{j=1}^{k_n-2} A\left(\bar{Y}_{j-1}, \alpha\right)^{-1} \left[b\left(\bar{Y}_{j-1}, \beta^\star\right), b\left(X_{j\Delta_n}, \alpha^\star\right) - b\left(\bar{Y}_{j-1}, \alpha^\star\right)\right],$$

$$\mathbb{Y}_{2,n}^{\mathrm{ML}(\dagger)}(\beta; \vartheta^\star) = -\frac{1}{2k_n} \sum_{j=1}^{k_n-2} A\left(\bar{Y}_{j-1}, \hat{\alpha}_n\right)^{-1} \left[\left(b\left(\bar{Y}_{j-1}, \beta\right) - b\left(\bar{Y}_{j-1}, \beta^\star\right)\right)^{\otimes 2}\right].$$

It is easy to obtain $\sup_{n \in \mathbf{N}} \mathbf{E}_{\theta^\star}\left[\sup_{\theta \in \Theta}\left|M_{2,n}^{\mathrm{ML}(\dagger)}\right|^p\right] \le C(p)\left(k_n \Delta_n\right)^{-p/2}$ using L^p-boundedness of $\sqrt{k_n}\left(\hat{\alpha}_n - \alpha^\star\right)$, Burkholder's inequality, and Sobolev's one, and $\sup_{n \in \mathbf{N}} \mathbf{E}_{\theta^\star}\left[\sup_{\theta \in \Theta}\left|R_{2,n}^{\mathrm{ML}(\dagger)}\right|^p\right] \le C(p)\Delta_n^{p/2}$ because of Lemma 5. Let us define

$$\mathbb{Y}_{2,n}^{\mathrm{ML}(\ddagger)}(\beta; \vartheta^\star) = -\frac{1}{2k_n} \sum_{j=1}^{k_n-2} A\left(X_{j\Delta_n}, \alpha^\star\right)^{-1} \left[\left(b\left(X_{j\Delta_n}, \beta\right) - b\left(X_{j\Delta_n}, \beta^\star\right)\right)^{\otimes 2}\right],$$

and then because of L^p-boundedness of $\sqrt{k_n}\left(\hat{\alpha}_n - \alpha^\star\right)$, and Lemma 3, we obtain

$$k_n^{\epsilon_1} \left\| \sup_{\beta \in \Theta_2} \left|\mathbb{Y}_{2,n}^{\mathrm{ML}(\dagger)}(\beta; \vartheta^\star) - \mathbb{Y}_{2,n}^{\mathrm{ML}(\ddagger)}(\beta; \vartheta^\star)\right| \right\|_p \to 0.$$

Then, L^p-boundedness of $\sup_{\beta \in \Theta_2}\left(k_n \Delta_n\right)^{\epsilon_1}\left|\mathbb{Y}_{2,n}^{\mathrm{ML}(\ddagger)}(\beta; \vartheta^\star) - \mathbb{Y}_2(\beta; \vartheta^\star)\right|$ is obtained by the discussion in Remark 4 and it verifies (b). □

Lemma 11 (a) *For every $M_3 > 0$,*

$$\sup_{n\in\mathbf{N}} \mathbf{E}_{\theta^\star}\left[\left((k_n\Delta_n)^{-1} \sup_{\beta\in\Theta_2} \left|\partial_\beta^3\mathbb{H}_{2,n}\left(\hat{\alpha}_n, \beta\right)\right|\right)^{M_3}\right] < \infty,$$

$$\sup_{n\in\mathbf{N}} \mathbf{E}_{\theta^\star}\left[\left((k_n\Delta_n)^{-1} \sup_{\beta\in\Theta_2} \left|\partial_\beta^3\mathbb{H}_{2,n}\left(\tilde{\alpha}_n, \beta\right)\right|\right)^{M_3}\right] < \infty.$$

(b) *Let $\epsilon_1 = \epsilon_0/2$. Then, for every $M_4 > 0$,*

$$\sup_{n\in\mathbf{N}} \mathbf{E}_{\theta^\star}\left[\left((k_n\Delta_n)^{\epsilon_1} \left|\Gamma_{2,n}^{\mathrm{ML}}\left(\beta^\star; \vartheta^\star\right) - \Gamma_2\left(\vartheta^\star\right)\right|\right)^{M_4}\right] < \infty,$$

$$\sup_{n\in\mathbf{N}} \mathbf{E}_{\theta^\star}\left[\left((k_n\Delta_n)^{\epsilon_1} \left|\Gamma_{2,n}^{\mathrm{Bayes}}\left(\beta^\star; \vartheta^\star\right) - \Gamma_2\left(\vartheta^\star\right)\right|\right)^{M_4}\right] < \infty.$$

Proof With respect to (a), we have for all $\alpha \in \Theta_1$ and $\beta \in \Theta_2$,

$$\frac{1}{k_n\Delta_n}\left|\partial_\beta^3\mathbb{H}_{2,n}\left(\alpha, \beta\right)\right|$$

$$= \frac{1}{k_n\Delta_n}\sum_{j=1}^{k_n-2}\left|\partial_\beta^2\left(A\left(\bar{Y}_{j-1}, \alpha\right)\left[\bar{Y}_{j+1}-\bar{Y}_j-\Delta_n b\left(\bar{Y}_{j-1}, \beta\right), \Delta_n\partial_\beta b\left(\bar{Y}_{j-1}, \beta\right)^\mathrm{T}\right]\right)\right|$$

$$= \frac{1}{k_n}\sum_{j=1}^{k_n-2}\left|\partial_\beta^2\left(A\left(\bar{Y}_{j-1}, \alpha\right)\left[\bar{Y}_{j+1}-\bar{Y}_j-\Delta_n b\left(\bar{Y}_{j-1}, \beta\right), \partial_\beta b\left(\bar{Y}_{j-1}, \beta\right)^\mathrm{T}\right]\right)\right|$$

$$\leq \frac{1}{k_n}\sum_{j=1}^{k_n-2} C\left(1 + \left|\bar{Y}_{j-1}\right| + \left|\bar{Y}_j\right| + \left|\bar{Y}_{j+1}\right|\right)^C.$$

Hence, the evaluation of (a) can be obtained because of the integrability of $\left\{\bar{Y}_j\right\}_{j=0,\ldots,k_n-1}$.

For (b), it is quite analogous to the (b) in Lemma 10. □

Proof *(Proof of Theorem 1)* The first polynomial-type large deviation inequality is shown in Proposition 1, and the second and third ones are also the consequence of Lemmas 10 and 11 and Theorem 3 in Yoshida (2011). This result, Lemma 6, and convergence in distribution shown by Nakakita and Uchida (2019a) complete the proof for convergence of moments with respect to the adaptive ML-type estimator.

Let us define the following statistical random fields, for all $u_0 \in \mathbf{R}^{d(d+1)/2}$ and $n \in \mathbf{N}$ such that $\theta_\varepsilon^\star + n^{-1/2} u_0 \in \Theta_\varepsilon$,

$$\mathbb{H}_{0,n} (\theta_\varepsilon) := -\frac{1}{2} \sum_{i=1}^{n-1} \left| \frac{1}{2} Z_{i+1} - \theta_\varepsilon \right|^2,$$

$$\mathbb{Z}_{0,n} (u_0; \theta_\varepsilon^\star) := \exp \left(\mathbb{H}_{0,n} \left(\theta_\varepsilon^\star + n^{-1/2} u_0 \right) - \mathbb{H}_{0,n} \left(\theta_\varepsilon^\star \right) \right),$$

where $\theta_\varepsilon = \mathrm{vech}\Lambda$ and $Z_{i+1} = \mathrm{vech} \left\{ \left(Y_{(i+1)h_n} - Y_{ih_n} \right)^{\otimes 2} \right\}$. Note that $\hat{\theta}_{\varepsilon,n}$ maximises $\mathbb{H}_{0,n}$. Now we prove the convergence in distribution such that for all $R > 0$,

$$\left[\mathbb{Z}_{0,n} \left(u_0; \theta_\varepsilon^\star \right), \mathbb{Z}_{1,n}^\tau \left(u_1; \hat{\Lambda}_n, \alpha^\star \right), \mathbb{Z}_{2,n} (u_2; \tilde{\alpha}_n, \beta^\star) \right]$$
$$\xrightarrow{d} \left[\mathbb{Z}_0 \left(u_0; \theta_\varepsilon^\star \right), \mathbb{Z}_1^\tau (u_1; \Lambda_\star, \alpha^\star), \mathbb{Z}_2 (u_2; \alpha^\star, \beta^\star) \right] \text{ in } \mathcal{C} \left(B \left(R; \mathbf{R}^{d(d+1)/2+m_1+m_2} \right) \right),$$

where for $\Delta_0 \sim N_{d(d+1)/2} \left(\mathbf{0}, \mathcal{I}^{(1,1)} (\vartheta^\star) \right)$, $\Delta_1^\tau \sim N_{m_1} \left(\mathbf{0}, \mathcal{I}^{(2,2),\tau} (\vartheta^\star) \right)$, $\Delta_2 \sim N_{m_2} \left(\mathbf{0}, \mathcal{I}^{(3,3)} (\vartheta^\star) \right)$ such that Δ_0, Δ_1^τ and Δ_2 are diagonal,

$$\mathbb{Z}_0 \left(u_0; \vartheta^\star \right) := \exp \left(\Delta_0 [u_0] - |u_0|^2 \right),$$

$$\mathbb{Z}_1^\tau \left(u_1; \Lambda_\star, \alpha^\star \right) := \exp \left(\Delta_1^\tau [u_1] - \Gamma_1^\tau (\vartheta^\star) \left[u_1^{\otimes 2} \right] \right),$$

$$\mathbb{Z}_2 \left(u_2; \alpha^\star, \beta^\star \right) := \exp \left(\Delta_2 [u_2] - \Gamma_2 (\vartheta^\star) \left[u_2^{\otimes 2} \right] \right),$$

and $\mathcal{C} (B (R; \mathbf{R}^m))$ is a metric space of continuous functions on the closed ball such that $B (R; \mathbf{R}^m) = \{u \in \mathbf{R}^m; |u| \le R\}$, whose norm is defined as the supreme one. To prove it, it is sufficient to show the finite-dimensional convergence of

$$\left[\log \mathbb{Z}_{0,n} \left(u_0; \theta_\varepsilon^\star \right), \log \mathbb{Z}_{1,n}^\tau \left(u_1; \hat{\Lambda}_n, \alpha^\star \right), \log \mathbb{Z}_{2,n} (u_2; \tilde{\alpha}_n, \beta^\star) \right]$$
$$\xrightarrow{d} \left[\log \mathbb{Z}_0 \left(u_0; \theta_\varepsilon^\star \right), \log \mathbb{Z}_1^\tau (u_1; \Lambda_\star, \alpha^\star), \log \mathbb{Z}_2 (u_2; \alpha^\star, \beta^\star) \right],$$

and the tightness of $\left\{ \log \mathbb{Z}_{0,n} (u_0) |_{C(B(R))}; n \in \mathbf{N} \right\}$, $\left\{ \log \mathbb{Z}_{1,n}^\tau (u_1) |_{C(B(R))}; n \in \mathbf{N} \right\}$, and $\left\{ \log \mathbb{Z}_{2,n} (u_3) |_{C(B(R))}; n \in \mathbf{N} \right\}$. The finite-dimensional convergence is a simple consequence of Nakakita and Uchida (2019a), and the tightness can be obtained if we can show

$$\sup_{n\in\mathbf{N}} \mathbf{E}_{\theta^\star} \left[\sup_{u_0\in B\left(R;\mathbf{R}^{d(d+1)/2}\right)} \left|\partial_{u_0} \log \mathbb{Z}_{0,n}\left(u_0;\theta_\varepsilon^\star\right)\right| \right] < \infty,$$

$$\sup_{n\in\mathbf{N}} \mathbf{E}_{\theta^\star} \left[\sup_{u_1\in B\left(R;\mathbf{R}^{m_1}\right)} \left|\partial_{u_1} \log \mathbb{Z}_{1,n}^\tau\left(u_1;\hat{\Lambda}_n,\alpha^\star\right)\right| \right] < \infty,$$

$$\sup_{n\in\mathbf{N}} \mathbf{E}_{\theta^\star} \left[\sup_{u_2\in B\left(R;\mathbf{R}^{m_2}\right)} \left|\partial_{u_2} \log \mathbb{Z}_{2,n}\left(u_2;\tilde{\alpha}_n,\beta^\star\right)\right| \right] < \infty,$$

as Ogihara and Yoshida (2011) or Yoshida (2011). We have the first evaluation for the simple computation and the remaining ones by Lemmas 8, 9, 10, and 11. Hence, we obtain the convergences in distribution in $\mathcal{C}\left(B\left(R;\mathbf{R}^{d(d+1)/2+m_1+m_2}\right)\right)$.

Finally, it is necessary to show the following evaluations for the proof utilising Theorem 10 in Yoshida (2011): there exists $\delta_1 > 0$ and $\delta_2 > 0$ such that

$$\sup_{n\in\mathbf{N}} \mathbf{E}_{\theta^\star} \left[\left(\int_{u_1:|u_1|\leq\delta_1} \mathbb{Z}_{1,n}^\tau\left(u_1;\hat{\Lambda}_n,\alpha^\star\right) du_1 \right)^{-1} \right] < \infty,$$

$$\sup_{n\in\mathbf{N}} \mathbf{E}_{\theta^\star} \left[\left(\int_{u_2:|u_2|\leq\delta_2} \mathbb{Z}_{2,n}\left(u_2;\tilde{\alpha}_n,\beta^\star\right) du_2 \right)^{-1} \right] < \infty.$$

Because of the Lemma 2 in Yoshida (2011), it is sufficient to show that for some $p > d$, $\delta > 0$ and $C > 0$,

$$\sup_{n\in\mathbf{N}} \mathbf{E}_{\theta^\star} \left[\left|\log \mathbb{Z}_{1,n}^\tau\left(u_1;\hat{\Lambda}_n,\alpha^\star\right)\right|^p \right] \leq C |u_1|^p ,$$

$$\sup_{n\in\mathbf{N}} \mathbf{E}_{\theta^\star} \left[\left|\log \mathbb{Z}_{2,n}\left(u_2;\tilde{\alpha}_n,\beta^\star\right)\right|^p \right] \leq C |u_2|^p ,$$

for all u_1, u_2 satisfying $|u_1|+|u_2| \leq \delta$, and actually, it is easily obtained by Lemmas 8, 9, 10, and 11. These results above lead to the following convergences because of Theorem 10 in Yoshida (2011):

$$\left[\mathbb{Z}_{0,n}\left(u_0;\theta_\varepsilon^\star\right), \int f_1\left(u_1\right) \mathbb{Z}_{1,n}^\tau\left(u_1;\hat{\Lambda}_n,\alpha^\star\right) du_1, \int f_2\left(u_2\right) \mathbb{Z}_{2,n}\left(u_2;\tilde{\alpha}_n,\beta^\star\right) du_2 \right]$$
$$\xrightarrow{d} \left[\mathbb{Z}_0\left(u_0;\theta_\varepsilon^\star\right), \int f_1\left(u_1\right) \mathbb{Z}_1^\tau\left(u_1;\Lambda_\star,\alpha^\star\right) du_1, \int f_2\left(u_2\right) \mathbb{Z}_2\left(u_2;\alpha^\star,\beta^\star\right) du_2 \right]$$

in $\mathcal{C}\left(B\left(R;\mathbf{R}^{d(d+1)/2}\right)\right)$ for the functions f_1 and f_2 of at most polynomial growth, and the continuous mapping theorem verifies

$$\left[\sqrt{n}\left(\hat{\theta}_{\varepsilon,n}-\theta_\varepsilon^\star\right), \sqrt{k_n}\left(\tilde{\alpha}_n-\alpha^\star\right), \sqrt{T_n}\left(\tilde{\beta}_n-\beta^\star\right) \right] \xrightarrow{d} \left[\zeta_0, \zeta_1^\tau, \zeta_2 \right].$$

Moreover, in a similar way as in the proof of Theorem 8 in Yoshida (2011), one has that for every $p > 0$, $\sup_{n\in\mathbf{N}} \mathbf{E}_{\theta^\star}\left[\left|\sqrt{T_n}(\tilde{\beta}_n-\beta^\star)\right|^p\right] < \infty$, which completes the proof. $\square$

References

Bibby, B. M., Sørensen, M. (1995). Martingale estimating functions for discretely observed diffusion processes. *Bernoulli*, *1*, 17–39.

Clinet, S., Yoshida, N. (2017). Statistical inference for ergodic point processes and application to limit order book. *Stochastic Processes and Their Applications*, *127*(6), 1800–1839.

Eguchi, S., Masuda, H. (2018). Schwarz-type model comparison for LAQ models. *Bernoulli*, *24*(3), 2278–2327.

Favetto, B. (2014). Parameter estimation by contrast minimization for noisy observations of a diffusion process. *Statistics*, *48*(6), 1344–1370.

Favetto, B. (2016). Estimating functions for noisy observations of ergodic diffusions. *Statistical Inference for Stochastic Processes*, *19*, 1–28.

Florens-Zmirou, D. (1989). Approximate discrete time schemes for statistics of diffusion processes. *Statistics*, *20*(4), 547–557.

Ibragimov, I. A., Has'minskii, R. Z. (1972). The asymptotic behaviour of certain statistical estimates in the smooth case. I. Investigation of the likelihood ratio. *Teorija Verojatnostei i ee Primenenija*, *17*, 469–486. (**Russian**).

Ibragimov, I. A., Has'minskii, R. Z. (1973). Asymptotic behaviour of certain statistical estimates. II. Limit theorems for a posteriori density and for Bayesian estimates. *Teorija Verojatnostei i ee Primenenija*, *18*, 78–93. (**Russian**).

Ibragimov, I. A., Has'minskii, R. Z. (1981). *Statistical estimation*. New York: Springer.

Jacod, J., Li, Y., Mykland, P. A., Podolskij, M., Vetter, M. (2009). Microstructure noise in the continuous case: The pre-averaging approach. *Stochastic Processes and Their Applications*, *119*(7), 2249–2276.

Kaino, Y., Uchida, M. (2018a). Hybrid estimators for small diffusion processes based on reduced data. *Metrika*, *81*(7), 745–773.

Kaino, Y., Uchida, M. (2018b). Hybrid estimators for stochastic differential equations from reduced data. *Statistical Inference for Stochastic Processes*, *21*(2), 435–454.

Kaino, Y., Nakakita, S. H., Uchida, M. (2018). Hybrid estimation for ergodic diffusion processes based on noisy discrete observations. arXiv:1812.07497.

Kamatani, K. (2018). Efficient strategy for the Markov chain Monte Carlo in high-dimension with heavy-tailed target probability distribution. *Bernoulli*, *24*(4B), 3711–3750.

Kamatani, K., Uchida, M. (2015). Hybrid multi-step estimators for stochastic differential equations based on sampled data. *Statistical Inference for Stochastic Processes*, *18*(2), 177–204.

Kessler, M. (1995). Estimation des parametres d'une diffusion par des contrastes corriges. *Comptes rendus de l'Académie des sciences. Série 1, Mathématique*, *320*(3), 359–362.

Kessler, M. (1997). Estimation of an ergodic diffusion from discrete observations. *Scandinavian Journal of Statistics*, *24*, 211–229.

Kutoyants, Y. A. (1984). *Parameter estimation for stochastic processes* (B. L. S. Prakasa Rao, Ed., Trans.). Berlin: Herdermann.

Kutoyants, Y. A. (1994). *Identification of dynamical systems with small noise*. Dordrecht: Kluwer.

Kutoyants, Y. A. (2004). *Statistical inference for ergodic diffusion processes*. London: Springer.

Masuda, H., Shimizu, Y. (2017). Moment convergence in regularized estimation under multiple and mixed-rates asymptotics. *Mathematical Methods of Statistics*, *26*(2), 81–110.

Nakakita, S. H., Uchida, M. (2017). Adaptive estimation and noise detection for an ergodic diffusion with observation noises. arXiv:1711.04462.

Nakakita, S. H., Uchida, M. (2019a). Inference for ergodic diffusions plus noise. *Scandinavian Journal of Statistics*, *46*(2), 470–516.

Nakakita, S. H., Uchida, M. (2019b). Adaptive test for ergodic diffusions plus noise. *Journal of Statistical Planning and Inference*, *203*, 131–150.

Negri, I., Nishiyama, Y. (2017). Moment convergence of *Z*-estimators. *Statistical Inference for Stochastic Processes*, *20*(3), 387–397.

NWTC Information Portal. (2018). *NWTC 135-m meteorological towers data repository*. Retrieved from https://nwtc.nrel.gov/135mdata. Accessed 15 Feb 2018.

Ogihara, T. (2018). Parametric inference for nonsynchronously observed diffusion processes in the presence of market microstructure noise. *Bernoulli*, *24*(4B), 3318–3383.

Ogihara, T. (2019). On the asymptotic properties of Bayes-type estimators with general loss functions. *Journal of Statistical Planning and Inference*, *199*, 136–150.

Ogihara, T., Yoshida, N. (2011). Quasi-likelihood analysis for the stochastic differential equation with jumps. *Statistical inference for stochastic processes*, *14*(3), 189–229.

Pardoux, E., Veretennikov, A. Y. (2001). On the Poisson equation and diffusion approximation. I. *The Annals of Probability*, *29*(3), 1061–1085.

Uchida, M. (2010). Contrast-based information criterion for ergodic diffusion processes from discrete observations. *Annals of the Institute of Statistical Mathematics*, *62*(1), 161–187.

Uchida, M., Yoshida, N. (2012). Adaptive estimation of an ergodic diffusion process based on sampled data. *Stochastic Processes and their Applications*, *122*(8), 2885–2924.

Uchida, M., Yoshida, N. (2014). Adaptive Bayes-type estimators of ergodic diffusion processes from discrete observations. *Statistical Inference for Stochastic Processes*, *17*(2), 181–219.

Yoshida, N. (1992). Estimation for diffusion processes from discrete observation. *Journal of Multivariate Analysis*, *41*(2), 220–242.

Yoshida, N. (2011). Polynomial-type large deviation inequalities and quasi-likelihood analysis for stochastic differential equations. *Annals of the Institute of Statistical Mathematics*, *63*, 431–479.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.