**Electronic Journal of Statistics**
Vol. 16 (2022) 4275–4331
ISSN: 1935-7524
https://doi.org/10.1214/22-EJS2049

# LAMN property for multivariate inhomogeneous diffusions with discrete observations

**Ngoc Khue TRAN**

*School of Applied Mathematics and Informatics, Hanoi University of Science and Technology, 1 Dai Co Viet, Hai Ba Trung, Hanoi, Vietnam*
*e-mail:* khue.tranngoc@hust.edu.vn

**Hoang-Long NGO**

*Hanoi National University of Education, 136 Xuan Thuy, Cau Giay, Hanoi, Vietnam*
*e-mail:* ngolong@hnue.edu.vn

**Abstract:** We consider a class of multidimensional inhomogeneous diffusions whose drift coefficient depends on a multidimensional unknown parameter. Under some appropriate assumptions, we prove the local asymptotic mixed normality property for the drift parameter from high-frequency observations when the length of the observation window tends to infinity. To obtain the result, we use the Malliavin calculus techniques and the change of measures. Our approach is applicable for both ergodic and non-ergodic diffusions.

**MSC2020 subject classifications:** Primary 60H07, 60J60; secondary 62F12, 62M05.
**Keywords and phrases:** Asymptotic efficiency, multivariate inhomogeneous diffusion, local asymptotic mixed normality property, Malliavin calculus, non-ergodic diffusion.

Received November 2021.

## Contents

## 1. Introduction

We consider on a complete probability space $(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{\mathrm{P}})$ a $d$-dimensional process $X^\theta = (X_t^\theta)_{t\geq 0}$ solution to the following inhomogeneous stochastic differential equation (SDE)

$$dX_t^\theta = b(\theta, t, X_t^\theta)dt + \sigma(t, X_t^\theta)dB_t, \tag{1.1}$$

where $X_0^\theta = x_0 \in \mathbb{R}^d$, $B = (B_t)_{t\geq 0}$ is a $d$-dimensional Brownian motion. The unknown parameter $\theta = (\theta_1, \ldots, \theta_m)$ belongs to $\Theta$, an open subset of $\mathbb{R}^m$ for some integer $m \geq 1$. Let $\{\widehat{\mathcal{F}}_t\}_{t\geq 0}$ denote the natural filtration generated by $B$. We always suppose that the coefficients $b = (b_1, \ldots, b_d) : \Theta \times \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ are measurable functions satisfying the Lipschitz continuity and linear growth condition **(A1)** below under which equation (1.1) has a unique $\{\widehat{\mathcal{F}}_t\}_{t\geq 0}$-adapted solution $X^\theta$ possessing the strong Markov property. Let $\widehat{\mathrm{P}}^\theta$ denote the probability measure induced by $X^\theta$ on the canonical space $(C(\mathbb{R}_+, \mathbb{R}^d), \mathcal{B}(C(\mathbb{R}_+, \mathbb{R}^d)))$ endowed with $\{\widehat{\mathcal{F}}_t\}_{t\geq 0}$, where $C(\mathbb{R}_+, \mathbb{R}^d)$ is the set of $\mathbb{R}^d$-valued continuous functions defined on $\mathbb{R}_+$, and $\mathcal{B}(C(\mathbb{R}_+, \mathbb{R}^d))$ is its Borel $\sigma$-algebra. Let $\widehat{\mathrm{E}}^\theta$ denote the expectation with respect to (w.r.t.) $\widehat{\mathrm{P}}^\theta$. Let $\xrightarrow{\widehat{\mathrm{P}}^\theta}$, $\xrightarrow{\mathcal{L}(\widehat{\mathrm{P}}^\theta)}$, $\widehat{\mathrm{P}}^\theta$-a.s., $\xrightarrow{\mathrm{P}}$, and $\xrightarrow{\mathcal{L}(\mathrm{P})}$ denote the convergence in $\widehat{\mathrm{P}}^\theta$-probability, in $\widehat{\mathrm{P}}^\theta$-law, in $\widehat{\mathrm{P}}^\theta$-almost surely, in P-probability, and in P-law, respectively. For $x \in \mathbb{R}^d$, $|x|$ denotes the Euclidean norm. $|A|$ denotes the Frobenius norm of the square matrix $A$, $\mathrm{tr}(A)$ denotes the trace, and $^*$ denotes the transpose.

For $n \geq 1$, we consider a discrete observation $X^{n,\theta} = (X_{t_0}^\theta, X_{t_1}^\theta, \ldots, X_{t_n}^\theta)$ at deterministic and equidistant times $t_k = k\Delta_n$, $k \in \{0, \ldots, n\}$ of the process $X^\theta$ solution to (1.1) under the high-frequency and infinite horizon conditions. That is, $\Delta_n \to 0$ and $n\Delta_n \to \infty$ as $n \to \infty$. Let $\mathrm{P}_n^\theta$ denote the probability law of the random vector $X^{n,\theta}$. We say that the local asymptotic mixed normality (LAMN) property holds at $\theta^0 \in \Theta$ with rate of convergence $\varphi_{n\Delta_n}(\theta^0)$ and asymptotic random Fisher information matrix $\Gamma(\theta^0)$ if for any $u \in \mathbb{R}^m$,

$$\log \frac{d\mathrm{P}_n^{\theta^0+\varphi_{n\Delta_n}(\theta^0)u}}{d\mathrm{P}_n^{\theta^0}}(X^{n,\theta^0}) \overset{\mathcal{L}(\widehat{\mathrm{P}}^{\theta^0})}{\longrightarrow} u^*\Gamma(\theta^0)^{1/2}\mathcal{N}(0,I_m) - \frac{1}{2}u^*\Gamma(\theta^0)u,$$

as $n \to \infty$, where $\mathcal{N}(0,I_m)$ is a centered $\mathbb{R}^m$-valued Gaussian random vector independent of $\Gamma(\theta^0)$ with identity covariance matrix $I_m$. Here, $\Gamma(\theta^0)$ is a symmetric positive definite random matrix in $\mathbb{R}^{m\times m}$, and $\varphi_{n\Delta_n}(\theta^0)$ is a diagonal matrix in $\mathbb{R}^{m\times m}$ whose diagonal entries tend to zero as $n$ goes to infinity. If $\Gamma(\theta^0)$ is non-random, we say that the local asymptotic normality (LAN) property holds at $\theta^0$. The LAMN property plays a fundamental role in the asymptotic theory of statistics. This property developed by Jeganathan [19] extends the LAN property which was introduced by Le Cam [23] and Hájek [13] in the situations where the asymptotic Fisher information matrix is deterministic. These properties allow giving the notion of asymptotically efficient estimators in the sense of Hájek-Le Cam convolution theorem as well as the lower bounds for the variance of estimators (see Jeganathan [19]). More precisely, a sequence of estimators $(\widehat{\theta}_n)_{n\geq 1}$ of the parameter $\theta^0$ is called regular at $\theta^0$ if for any $u \in \mathbb{R}^m$, as $n \to \infty$,

$$\varphi_{n\Delta_n}^{-1}(\theta^0)\left(\widehat{\theta}_n - \left(\theta^0 + \varphi_{n\Delta_n}(\theta^0)u\right)\right) \overset{\mathcal{L}(\widehat{\mathrm{P}}^{\theta^0+\varphi_{n\Delta_n}(\theta^0)u})}{\longrightarrow} V(\theta^0), \tag{1.2}$$

for some $\mathbb{R}^m$-valued random variable $V(\theta^0)$, independent of $u$. When the LAMN property holds at point $\theta^0$, the law of $V(\theta^0)$ conditionally on $\Gamma(\theta^0)$ is a convolution between the Gaussian law $\mathcal{N}(0,\Gamma(\theta^0)^{-1})$ and some other law $G_{\Gamma(\theta^0)}$ on $\mathbb{R}^m$, i.e.,

$$\mathcal{L}\left(V(\theta^0)|\Gamma(\theta^0)\right) = \mathcal{N}\left(0,\Gamma(\theta^0)^{-1}\right) \star G_{\Gamma(\theta^0)}.$$

Hence, $V(\theta^0)$ can be written as a sum of two independent random variables

$$V(\theta^0) \overset{\text{law}}{=} \Gamma(\theta^0)^{-1/2}\mathcal{N}(0,I_m) + R,$$

where $R$ is a random variable with distribution $G_{\Gamma(\theta^0)}$, independent of $\mathcal{N}(0,I_m)$ (see [19, Corollary 1]). Consequently, a sequence of regular estimators $(\widehat{\theta}_n)_{n\geq 1}$ of the parameter $\theta^0$ is called asymptotically efficient at $\theta^0$ in the sense of Hájek-Le Cam convolution theorem if as $n \to \infty$,

$$\varphi_{n\Delta_n}^{-1}(\theta^0)\big(\widehat{\theta}_n - \theta^0\big) \overset{\mathcal{L}(\widehat{\mathrm{P}}^{\theta^0})}{\longrightarrow} \Gamma(\theta^0)^{-1/2}\mathcal{N}(0,I_m),$$

where $\Gamma(\theta^0)$ and $\mathcal{N}(0,I_m)$ are independent (i.e., take $u=0$ in (1.2) and $R=0$). We refer the reader to Section 7.1 of Höpfner [14] or Le Cam and Lo Yang [24] for further details.

On the basis of continuous observations with increasing observation window, the LAMN property was established by Luschgy [26] for semimartingale, Kutoyants [21] for ergodic diffusions (see Proposition 2.2), null-recurrent process (see [21, Remark 3.42]) and for Ornstein-Uhlenbeck process (see [21, Remark

3.47]), Bishwal [7] for inhomogeneous diffusions (see [7, Chapter 4]), Overbeck [30] for Cox-Ingersoll-Ross process, and Benke *et al.* [6] for Heston model. The asymptotic likelihood theory for multidimensional inhomogeneous diffusion processes (1.1) whose drift coefficient depends linearly on the parameter can be found in Section 5 of [2, Chapter 9]. Besides, the asymptotic properties of the maximum likelihood estimator and Bayes estimator for the nonlinear drift parameter of one-dimensional inhomogeneous and homogeneous diffusions were studied in [7, Chapter 4], [28] and [38].

On the basis of discrete observations at high frequency, Gobet [11] proved the LAMN property for inhomogeneous diffusion on a fixed time interval. In the case of increasing observation window $n\Delta_n$, Gobet [12] obtained the LAN property for homogeneous ergodic diffusions using Malliavin calculus. Later on, Shimizu [33] showed the LAMN property for the non-recurrent Ornstein-Uhlenbeck process using the explicit expression of the transition density. More recently, Ben Alaya *et al.* [5] have proved the LAN, LAMN, and LAQ (local asymptotic quadraticity) properties for the Cox-Ingersoll-Ross process. Recall also that results on parameter estimation for discretely observed non-ergodic diffusions can be found in Jacod [16] where the rate is $(\sqrt{n\Delta_n}, \sqrt{n})$ for the drift and diffusion parameters, and in Shimizu [34] where the rate varies depending on the observed Fisher information. Indeed, in [16], the author constructed estimators from a moment type contrast function for the drift and diffusion parameters of multidimensional homogeneous and non-ergodic diffusions and established the consistency of the estimators in the sense of tightness under some suitable smoothness and identifiability conditions. These estimators converge at rate $\sqrt{n\Delta_n}$ for the drift parameter and at rate $\sqrt{n}$ for the diffusion parameter. In [34], the author constructed $M$-estimators from a quadratic-type contrast function for the drift and diffusion parameters of one-dimensional homogeneous diffusions without ergodicity assumption and established the consistency of the $M$-estimators in the sense of tightness. These $M$-estimators converge with a variety of rates of convergence for the drift and diffusion parameters. Besides, the parameter estimation for discretely observed multidimensional ergodic diffusions can be found in Yoshida [39].

On the basis of discrete observations at low frequency, Aït-Sahalia [1] studied the LAN, LAMN, LAQ properties, and the asymptotic properties of maximum likelihood estimator (MLE) for one-dimensional homogeneous diffusions. For this, the author constructs a closed-form sequence of approximations to the transition density via the Hermite polynomials.

To summarize, the following table contains the aforementioned known results on the LAMN and LAN properties for homogeneous and inhomogeneous diffusions on the basis of continuous or discrete observations on the time interval

$[0, T]$ with $T \to \infty$.

<table>
<tr><th colspan="2">Diffusions</th><th>Continuous observations</th><th>Discrete observations</th></tr>
<tr><td rowspan="2">Homogeneous</td><td>Ergodic<br>(LAN)</td><td>• Kutoyants [21]:<br>diffusions (Pro. 2.2)<br>• Overbeck [30]:<br>Cox-Ingersoll-Ross<br>• Benke *et al.* [6]:<br>Heston model</td><td>• Gobet [12]:<br>diffusions<br>• Ben Alaya *et al.* [5]:<br>Cox-Ingersoll-Ross</td></tr>
<tr><td>Non-ergodic<br>(LAMN)</td><td>• Kutoyants [21]:<br>null-recurrent process<br>(Rem. 3.42)<br>Ornstein-Uhlenbeck<br>(Rem. 3.47)<br>• Overbeck [30]:<br>Cox-Ingersoll-Ross<br>• Benke *et al.* [6]:<br>Heston model</td><td>• Shimizu [33]:<br>Ornstein-Uhlenbeck<br>• Ben Alaya *et al.* [5]:<br>Cox-Ingersoll-Ross</td></tr>
<tr><td rowspan="2">Inhomogeneous</td><td>Ergodic<br>(LAN)</td><td>• Luschgy [26]:<br>semimartingale</td><td></td></tr>
<tr><td>Non-ergodic<br>(LAMN)</td><td>• Luschgy [26]:<br>semimartingale<br>• Bishwal [7]:<br>1-d diffusions (Cha. 4)</td><td></td></tr>
</table>

The validity of the LAN or LAMN property based on discrete observations at a high frequency of solution to a general inhomogeneous and ergodic or non-ergodic SDE when the length of the observation window tends to infinity has not been investigated yet. In addition, one of the motivations for the current work is to understand the problem in [1] for the case of high frequency.

In this paper, we prove the LAMN property for a general class of inhomogeneous diffusions observed at discrete times without assuming ergodicity. The validity of the LAQ property will be considered in future work. Unlike the Ornstein-Uhlenbeck process, the transition density of the solution to the general equation (1.1) is not explicit. Our strategy is to use the Malliavin calculus approach initiated by Gobet [11] in order to derive an explicit expression for the logarithm derivative of the transition density with respect to the parameter (see Lemma 3.3). With the help of this explicit expression, we derive an appropriate expansion of the log-likelihood ratio (see (3.11), (3.12) of Section 3). To treat the main contributions, we need to use the asymptotic behavior of the observed Fisher information process based on the continuous observation (see condition **(A4)**) and the multivariate central limit theorem for continuous local martingales. Thanks to conditions **(A5)**-**(A7)**, the negligible contribution is shown by using four technical Lemmas 4.1-4.5. This technique is not the same as the one of Gobet [12]. Indeed, Gobet [12] used a change of transition densities and the upper and lower bounds of the Gaussian type of the densities. In our situation, it is not clear if one could use that technique. Instead, we need to use a change of measures. Two approaches will lead to the squared exponential moment. To deal with this moment, Gobet [12] used the ergodic property whereas we need here condition **(A7)** that can be verified in several models, which may not possess the ergodicity, by using the explicit expression of the density or moment estimates. It should be noted that our new strategy allows us to get, for the first time, the LAN and LAMN properties from high-frequency observations for diffusions whose both drift and diffusion coefficients are time-dependent (see Section 5.2.2).

The paper is organized as follows. In Section 2, we formulate assumptions on equation (1.1) and state our main result in Theorem 2.1. In Section 3, an explicit expression for the score function is first presented, which allows transforming the log-likelihood ratio. The proof of the main result is then given by following the aforementioned strategy. The convergence of the remainder terms is given in Section 4. Several illustrated examples will be given in Section 5 which discusses homogeneous ergodic diffusion processes, homogeneous Ornstein-Uhlenbeck process, two-dimensional Gaussian diffusion process, null-recurrent diffusion process, and a generalized exponential growth process, inhomogeneous Ornstein-Uhlenbeck process and a special inhomogeneous diffusion process. Finally, the proofs of some technical results are given in Section 6.

## 2. Assumptions and main result

We first recall a few concepts on the statistical inference for the experiment based on continuous observations. For details, we refer the reader to Barndorff-Nielsen and Sørensen [4]. For any $T \geq 0$ and $\theta \in \Theta$, we let $\widehat{\mathrm{P}}_T^\theta$ denote the probability measure generated by the process $X^{T,\theta} := (X_t^\theta)_{t\in[0,T]}$ solving (1.1) on the measurable space $(C([0,T],\mathbb{R}^d), \mathcal{B}(C([0,T],\mathbb{R}^d)))$. Here $C([0,T],\mathbb{R}^d)$ denotes the set of $\mathbb{R}^d$-valued continuous functions defined on $[0,T]$, and $\mathcal{B}(C([0,T],\mathbb{R}^d))$ is its Borel $\sigma$-algebra. Therefore, $\widehat{\mathrm{P}}_T^\theta$ is the restriction of $\widehat{\mathrm{P}}^\theta$ to $\widehat{\mathcal{F}}_T$. We define the log-likelihood function of the family of probability measures $(\widehat{\mathrm{P}}_T^\theta)_{\theta\in\Theta}$ as $\ell_T(\theta) = \log \frac{d\widehat{\mathrm{P}}_T^\theta}{d\widehat{\widehat{\mathrm{P}}}_T}$, where $\widehat{\widehat{\mathrm{P}}}_T$ is a probability measure on $(C([0,T],\mathbb{R}^d), \mathcal{B}(C([0,T],\mathbb{R}^d)))$ which is supposed to satisfy that $\widehat{\mathrm{P}}_T^\theta$ is absolutely continuous w.r.t. $\widehat{\widehat{\mathrm{P}}}_T$, for all $T \geq 0$ and $\theta \in \Theta$. In fact, by [18, Chapter III] and [25, Chapter 7], for all $\theta, \theta^1 \in \Theta$, the probability measures $\widehat{\mathrm{P}}_T^\theta$ and $\widehat{\mathrm{P}}_T^{\theta^1}$ are absolutely continuous w.r.t. each other and its Radon-Nikodym derivative is given by

$$\frac{d\widehat{\mathrm{P}}_T^\theta}{d\widehat{\mathrm{P}}_T^{\theta^1}}((X_t^{\theta^1})_{t\in[0,T]}) = \exp\left\{ \int_0^T \left(b(\theta,t,X_t^{\theta^1}) - b(\theta^1,t,X_t^{\theta^1})\right)^* \sigma^{-1}(t,X_t^{\theta^1}) dB_t \right.$$
$$\left. - \frac{1}{2}\int_0^T \left| \left(b(\theta,t,X_t^{\theta^1}) - b(\theta^1,t,X_t^{\theta^1})\right)^* \sigma^{-1}(t,X_t^{\theta^1}) \right|^2 dt \right\}.$$

Therefore, the log-likelihood function is given by

$$\begin{aligned}\ell_T(\theta) &= \log \frac{d\widehat{\mathrm{P}}_T^{\theta}}{d\widehat{\mathrm{P}}_T^{\theta^1}}((X_t^{\theta^1})_{t\in[0,T]})\\ &= \int_0^T \big(b(\theta,t,X_t^{\theta^1}) - b(\theta^1,t,X_t^{\theta^1})\big)^* \sigma^{-1}(t,X_t^{\theta^1})dB_t\\ &\quad - \frac{1}{2}\int_0^T \Big|\big(b(\theta,t,X_t^{\theta^1}) - b(\theta^1,t,X_t^{\theta^1})\big)^* \sigma^{-1}(t,X_t^{\theta^1})\Big|^2 dt,\end{aligned}$$

where $\widehat{\mathrm{P}}_T^{\theta^1}$ is considered as the dominating probability measure $\widehat{\widehat{\mathrm{P}}}_T$ of the family of probability measures $(\widehat{\mathrm{P}}_T^{\theta})_{\theta\in\Theta}$. The score vector which is defined as the vector of the first derivatives of the log-likelihood function is given under the measure $\widehat{\mathrm{P}}$ by the gradient

$$\nabla_\theta \ell_T(\theta) = \int_0^T (\nabla_\theta b(\theta,t,X_t^\theta))^* \sigma^{-1}(t,X_t^\theta)dB_t, \tag{2.1}$$

which is a martingale w.r.t. the filtration $\{\widehat{\mathcal{F}}_t\}_{t\in[0,T]}$. The quadratic variation of the score vector martingale which is also the bracket process is given by

$$[\nabla_\theta \ell(\theta)]_T = \langle \nabla_\theta \ell(\theta)\rangle_T = \int_0^T (\nabla_\theta b(\theta,t,X_t^\theta))^* \sigma^{-2}(t,X_t^\theta)\nabla_\theta b(\theta,t,X_t^\theta)dt, \tag{2.2}$$

which can be interpreted as the observed Fisher information process at $\theta$ based on $(X_t^\theta)_{t\in[0,T]}$.

We next impose the following assumptions on equation (1.1).

**(A1)** For any $\theta\in\Theta$, there exists a constant $L>0$ such that for all $x,y\in\mathbb{R}^d$ and $t\geq 0$,

$$\begin{aligned}&|b(\theta,t,x)-b(\theta,t,y)| + |\sigma(t,x)-\sigma(t,y)| \leq L|x-y|,\\ &|b(\theta,t,x)| + |\sigma(t,x)| \leq L\left(1+|x|\right).\end{aligned}$$

Moreover, the Lipschitz constant $L$ is uniformly bounded on $\Theta$.

**(A2)** The diffusion matrix $\sigma$ is symmetric, positive and satisfies a uniform ellipticity condition. That is, there exists a constant $c\geq 1$ such that for all $x,\xi\in\mathbb{R}^d$ and $t\geq 0$,

$$\frac{1}{c}|\xi|^2 \leq |\sigma(t,x)\xi|^2 \leq c|\xi|^2.$$

**(A3)** The functions $b$ and $\sigma$ are of class $C^1$ w.r.t. $\theta$, $t$ and $x$. Each partial derivative $\partial_{x_i} b$ and $\partial_{x_i}\sigma$ is of class $C^1$ w.r.t. $x$, $\partial_{\theta_i} b$ is of class $C^1$ w.r.t. $t$, $\partial_{\theta_i} b$ is of class $C^2$ w.r.t. $x$. Moreover, for any $(\theta,\theta^1,\theta^2,x)\in\Theta^3\times\mathbb{R}^d$ and $t\geq 0$, there exist positive constants $C$ and $\gamma\in(0,1]$, independent of $(\theta,\theta^1,\theta^2,x,t)$, such that

(a) $|g(\cdot,t,x)| \le C$ for $g(\cdot,t,x) = \partial_{x_i} b(\theta,t,x), \partial_{x_i}\sigma(t,x), \partial_t \sigma(t,x)$, $\partial^2_{\theta_i x_j} b(\theta,t,x)$, $\partial^2_{x_i x_j} b(\theta,t,x), \partial^2_{x_i x_j}\sigma(t,x)$, $\partial^3_{\theta_i x_j x_k} b(\theta,t,x)$;

(b) $|h(\cdot,t,x)| \le C\,(1+|x|)$
for $h(\cdot,x) = \partial_{\theta_i} b(\theta,t,x), \partial_t b(\theta,t,x), \partial^2_{\theta_i t} b(\theta,t,x)$;

(c) $|\partial_{\theta_i} b(\theta^1,t,x) - \partial_{\theta_i} b(\theta^2,t,x)| \le C|\theta^1-\theta^2|^\gamma\,(1+|x|)$.

**(A4)** For $\theta \in \Theta$, there exist a $m \times m$ non-random diagonal matrix

$$\varphi_T(\theta) = \mathrm{diag}(\varphi_T^1(\theta), \dots, \varphi_T^m(\theta))$$

whose diagonal entries $\varphi_T^1(\theta), \dots, \varphi_T^m(\theta)$ are strictly positive and tend to zero as $T \to \infty$, and an $m \times m$ symmetric positive definite random matrix $\Gamma(\theta)$ such that $\langle \nabla_\theta \ell(\theta)\rangle_T$ converges to $\Gamma(\theta)$ at rate $\varphi_T(\theta)\varphi_T(\theta)$ in $\widehat{\mathrm{P}}^\theta$-probability as $T \to \infty$. That is, as $T \to \infty$,

$$\varphi_T(\theta)\,\langle \nabla_\theta \ell(\theta)\rangle_T\,\varphi_T(\theta) \xrightarrow{\widehat{\mathrm{P}}^\theta} \Gamma(\theta).$$

**(A5)** For any $\theta \in \Theta$, $p \ge 1$, there exists a function $\psi_t(\theta)$ which is strictly positive and $\psi_t^{-1}(\theta) \ge 1$ for any $t \ge 0$ such that

$$\sup_{t\ge 0} \widehat{\mathrm{E}}^\theta\left[\left|\psi_t(\theta) X_t^\theta\right|^p\right] < \infty.$$

**(A6)** For any $\theta \in \Theta$, $i,j,p \in \{1,\dots,m\}$ with $i \le j$, $i \le p$, and $p_0 > 1$ close to 1, as $n \to \infty$,

$$\Delta_n \varphi^i_{n\Delta_n}(\theta)\varphi^j_{n\Delta_n}(\theta)(\varphi^p_{n\Delta_n}(\theta))^\gamma \sum_{k=0}^{n-1}\left(1+\psi_{t_k}^{-2}(\theta)\right) \to 0,$$

$$\Delta_n^{\frac{3}{2}} \varphi^i_{n\Delta_n}(\theta)\varphi^j_{n\Delta_n}(\theta) \sum_{k=0}^{n-1}\left(1+\psi_{t_k}^{-6p_0}(\theta)\right) \to 0.$$

**(A7)** For all $\theta \in \Theta$, $i \in \{1,\dots,m\}$, $k \in \{0,...,n-1\}$ and constant $C > 0$, there exists a constant $c > 0$ such that for $n$ large enough,

$$\widehat{\mathrm{E}}^\theta\left[e^{C(\varphi^i_{n\Delta_n}(\theta))^2\Delta_n |X^\theta_{t_k}|^2}\right] \le c.$$

In order to apply the Malliavin calculus, the uniform ellipticity condition **(A2)** and regularity condition **(A3)** on the coefficients are required. Conditions **(A4)**, **(A5)**, **(A6)** and **(A7)** are needed in this setting where the diffusions can be ergodic or non-ergodic. On the one hand, condition **(A4)** ensures the asymptotic behavior for the main terms in the expansion of the log-likelihood ratio which are determined by the score vector and its quadratic variation. Let us recall that condition **(A4)** is analogous to the general condition (3.3) of Barndorff-Nielsen and Sørensen [4] which is given for general asymptotic likelihood theory for stochastic processes. This condition **(A4)** is also similar to the condition (2.12) of Luschgy [26] which is established for semimartingales. On the other hand, conditions **(A5)**, **(A6)** and **(A7)** guarantee the convergence of the negligible terms in the expansion of the log-likelihood ratio.

Now, for fixed $\theta^0 \in \Theta$, recall that a discrete observation of the process $X^{\theta^0}$ is given by $X^{n,\theta^0} = (X_{t_0}^{\theta^0}, X_{t_1}^{\theta^0}, \ldots, X_{t_n}^{\theta^0})$. The main result of this paper is the following LAMN property.

**Theorem 2.1.** *Assume conditions **(A1)**-**(A7)**. Then, the LAMN property holds for the likelihood at $\theta^0$ with rate of convergence*

$$\varphi_{n\Delta_n}(\theta^0) = \operatorname{diag}(\varphi_{n\Delta_n}^1(\theta^0), \ldots, \varphi_{n\Delta_n}^m(\theta^0)),$$

*and asymptotic random Fisher information matrix $\Gamma(\theta^0)$. That is, for all $u \in \mathbb{R}^m$, as $n \to \infty$,*

$$\log \frac{d\mathrm{P}_n^{\theta^0 + \varphi_{n\Delta_n}(\theta^0)u}}{d\mathrm{P}_n^{\theta^0}}(X^{n,\theta^0}) \xrightarrow{\mathcal{L}(\widehat{\mathrm{P}}^{\theta^0})} u^*\Gamma(\theta^0)^{1/2}\mathcal{N}(0, I_m) - \frac{1}{2}u^*\Gamma(\theta^0)u,$$

*where $\mathcal{N}(0, I_m)$ is a centered $\mathbb{R}^m$-valued Gaussian random variable independent of $\Gamma(\theta^0)$ with identity covariance matrix $I_m$.*

**Remark 2.2.** The exact maximum likelihood estimation is essentially based on the transition densities which are available in some very special cases. Thus, the MLE cannot be written explicitly for general diffusions. Instead, the parameter estimation for diffusion processes observed discretely at high frequency has been investigated mainly in the ergodic and homogeneous cases by using different approximate schemes. More precisely, in the one-dimensional setting, the approximate discrete-time scheme known as Euler-Maruyama's is used in [9] for diffusion processes with constant diffusion coefficient. Later on, in [20] the author studies diffusion processes with nonlinear coefficients and constructs a minimum contrast estimator via a contrast function by approximating the transition density by the density of a Gaussian law. In the multidimensional setting, in [39] the author proposes an adaptive maximum likelihood type estimator for diffusion processes with multiplicative diffusion coefficient by using a discretization of the continuous-time likelihood function. Later on, the authors propose respectively in [35, 36] two kinds of adaptive maximum likelihood type estimators and three kinds of adaptive Bayes type estimators based on the quasi log-likelihood functions for diffusion processes with nonlinear coefficients. For the case of the Ornstein-Uhlenbeck process, the asymptotic behavior of MLE-type approximation in the ergodic case is established in Theorem 1 of [32] whereas the asymptotic behavior of a trajectory-fitting estimator (TFE) in the non-ergodic case is obtained in Theorem 2 of [33].

As a consequence of Theorem 2.1, all of the aforementioned estimators for the drift parameters are asymptotically efficient. There are very few results on the estimation for non-ergodic homogeneous diffusions, and for ergodic and non-ergodic inhomogeneous diffusions (see Chapter 7 of [7]). However, Theorem 2.1 could serve as a benchmark to verify the asymptotic efficiency of any estimators in these cases.

As usual, constants are denoted by $C$ and may change value from one line to the next.

## 3. Proof of the main result

### 3.1. Representation of the score function using Malliavin calculus

To simplify the exposition, for $i \in \{1, \ldots, m\}$ we use the notations

$$\begin{aligned}
&\theta^0 = (\theta_1^0, \ldots, \theta_m^0), u = (u_1, u_2, \ldots, u_m),\\
&\theta^{0+} := \theta^0 + \varphi_{n\Delta_n}(\theta^0)u = (\theta_1^0 + \varphi^1_{n\Delta_n}(\theta^0)u_1, \ldots, \theta_m^0 + \varphi^m_{n\Delta_n}(\theta^0)u_m),\\
&\theta_i^{0+} := (\theta_1^0, \ldots, \theta_{i-1}^0, \theta_i^0 + \varphi^i_{n\Delta_n}(\theta^0)u_i, \theta_{i+1}^0 + \varphi^{i+1}_{n\Delta_n}(\theta^0)u_{i+1},\\
&\qquad \ldots, \theta_m^0 + \varphi^m_{n\Delta_n}(\theta^0)u_m),\\
&\theta_i^{0+}(\ell) := (\theta_1^0, \ldots, \theta_{i-1}^0, \theta_i^0 + \ell\varphi^i_{n\Delta_n}(\theta^0)u_i, \theta_{i+1}^0 + \varphi^{i+1}_{n\Delta_n}(\theta^0)u_{i+1},\\
&\qquad \ldots, \theta_m^0 + \varphi^m_{n\Delta_n}(\theta^0)u_m).
\end{aligned}$$

Under **(A1)**, **(A2)** and **(A3)**(a), for any $t > s$ the law of $X_t^\theta$ conditioned on $X_s^\theta = x$ possesses a positive transition density $p^\theta(s, t, x, y)$ which is differentiable w.r.t. $\theta$. The density of $X^{n,\theta} = (X^\theta_{t_0}, X^\theta_{t_1}, \ldots, X^\theta_{t_n})$ is denoted by $p_n(\cdot;\theta)$. To transform the log-likelihood ratio, we first use the Markov property to rewrite the global likelihood function in terms of a product of transition densities and then apply the mean value theorem to get the following decomposition

$$\begin{aligned}
&\log \frac{d\mathrm{P}_n^{\theta^0+\varphi_{n\Delta_n}(\theta^0)u}}{d\mathrm{P}_n^{\theta^0}}(X^{n,\theta^0}) = \log \frac{p_n(X^{n,\theta^0};\theta^{0+})}{p_n(X^{n,\theta^0};\theta^0)}\\
&= \sum_{k=0}^{n-1} \log \frac{p^{\theta_1^{0+}}}{p^{\theta^0}}(t_k, t_{k+1}, X^{\theta^0}_{t_k}, X^{\theta^0}_{t_{k+1}})\\
&= \sum_{k=0}^{n-1} \log \left( \frac{p^{\theta_1^{0+}}}{p^{\theta_2^{0+}}} \frac{p^{\theta_2^{0+}}}{p^{\theta_3^{0+}}} \cdots \frac{p^{\theta_i^{0+}}}{p^{\theta_{i+1}^{0+}}} \cdots \frac{p^{\theta_{m-1}^{0+}}}{p^{\theta_m^{0+}}} \frac{p^{\theta_m^{0+}}}{p^{\theta^0}} \right)(t_k, t_{k+1}, X^{\theta^0}_{t_k}, X^{\theta^0}_{t_{k+1}})\\
&= \sum_{k=0}^{n-1} \log \frac{p^{\theta_1^{0+}}}{p^{\theta_2^{0+}}}(t_k, t_{k+1}, X^{\theta^0}_{t_k}, X^{\theta^0}_{t_{k+1}}) + \sum_{k=0}^{n-1} \log \frac{p^{\theta_2^{0+}}}{p^{\theta_3^{0+}}}(t_k, t_{k+1}, X^{\theta^0}_{t_k}, X^{\theta^0}_{t_{k+1}}) + \cdots\\
&\quad + \sum_{k=0}^{n-1} \log \frac{p^{\theta_m^{0+}}}{p^{\theta^0}}(t_k, t_{k+1}, X^{\theta^0}_{t_k}, X^{\theta^0}_{t_{k+1}})\\
&= \sum_{k=0}^{n-1} \varphi^1_{n\Delta_n}(\theta^0)u_1 \int_0^1 \frac{\partial_{\theta_1} p^{\theta_1^{0+}(\ell)}}{p^{\theta_1^{0+}(\ell)}}(t_k, t_{k+1}, X^{\theta^0}_{t_k}, X^{\theta^0}_{t_{k+1}})d\ell\\
&\quad + \cdots + \sum_{k=0}^{n-1} \varphi^m_{n\Delta_n}(\theta^0)u_m \int_0^1 \frac{\partial_{\theta_m} p^{\theta_m^{0+}(\ell)}}{p^{\theta_m^{0+}(\ell)}}(t_k, t_{k+1}, X^{\theta^0}_{t_k}, X^{\theta^0}_{t_{k+1}})d\ell. \qquad (3.1)
\end{aligned}$$

We next apply the Malliavin calculus to get an explicit expression for the logarithm derivative of the transition density appearing in (3.1). For this, to avoid confusion with the observed process $X^\theta$ generated by the Brownian motion $B$, we introduce a new independent $d$-dimensional standard Brownian motion $W=(W_t)_{t\geq 0}$ by which an independent copy $Y^\theta=(Y_t^\theta)_{t\geq 0}$ of $X^\theta$ is generated. The Malliavin calculus on the Wiener space induced by $W$ will be applied. Let $(\widehat{\Omega},\widehat{\mathcal{F}},\{\widehat{\mathcal{F}}_t\}_{t\geq 0},\widehat{\mathrm{P}})$ and $(\widetilde{\Omega},\widetilde{\mathcal{F}},\{\widetilde{\mathcal{F}}_t\}_{t\geq 0},\widetilde{\mathrm{P}})$ be two canonical filtered probability spaces associated respectively to each of two Brownian motions $B$ and $W$. The product filtered probability space of these two canonical spaces is $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geq 0},\mathrm{P})$, that is, $\Omega=\widehat{\Omega}\times\widetilde{\Omega}$, $\mathcal{F}=\widehat{\mathcal{F}}\otimes\widetilde{\mathcal{F}}$, $\mathrm{P}=\widehat{\mathrm{P}}\otimes\widetilde{\mathrm{P}}$, $\mathcal{F}_t=\widehat{\mathcal{F}}_t\otimes\widetilde{\mathcal{F}}_t$. Let $\mathrm{E}$, $\widehat{\mathrm{E}}$, $\widetilde{\mathrm{E}}$ denote the expectation w.r.t. $\mathrm{P}$, $\widehat{\mathrm{P}}$ and $\widetilde{\mathrm{P}}$, respectively.

On the probability space $(\widetilde{\Omega},\widetilde{\mathcal{F}},\widetilde{\mathrm{P}})$, we consider the flow process $Y^\theta(s,x)=(Y_t^\theta(s,x),t\geq s)$, $x\in\mathbb{R}^d$ on the time interval $[s,\infty)$ and with initial condition $Y_s^\theta(s,x)=x$ satisfying

$$Y_t^\theta(s,x)=x+\int_s^t b(\theta,u,Y_u^\theta(s,x))du+\int_s^t \sigma(u,Y_u^\theta(s,x))dW_u. \tag{3.2}$$

The independent copy $Y^\theta=(Y_t^\theta)_{t\geq 0}$ of $X^\theta$ is defined by $Y_t^\theta\equiv Y_t^\theta(0,x_0)$ which thus satisfies

$$Y_t^\theta=x_0+\int_0^t b(\theta,u,Y_u^\theta)du+\int_0^t \sigma(u,Y_u^\theta)dW_u. \tag{3.3}$$

The Malliavin derivative and the Skorohod integral w.r.t. $W$ are respectively denoted by $D$ and $\delta$. The space of random variables which are differentiable in the sense of Malliavin and the domain of $\delta$ are respectively denoted by $\mathbb{D}^{1,2}$ and Dom $\delta$. The Malliavin derivative of a differentiable random variable $F\in\mathbb{D}^{1,2}$ is $DF=(D^1F,\ldots,D^dF)$ where $D^i$ denotes the Malliavin derivative in the $i$th direction $W^i$ of the Brownian motion $W=(W^1,\ldots,W^d)$ for $i\in\{1,\ldots,d\}$. The Skorohod integral of a $\mathbb{R}^d$-valued process $U=(U^1,\ldots,U^d)\in$ Dom $\delta$ is defined as $\delta(U)=\sum_{i=1}^d\delta^i(U^i)$ where $\delta^i$ is the Skorohod integral w.r.t. $W^i$. For a more detailed presentation on the Malliavin calculus, we refer to the book [29] by Nualart.

For any $k\in\{0,\ldots,n-1\}$, under **(A1)**, **(A2)** and **(A3)**(a)-(b), the process $(Y_t^\theta(t_k,x),t\in[t_k,t_{k+1}])$ is differentiable w.r.t. $x$ and $\theta$ (see Kunita [22]). Moreover, the corresponding Jacobian matrix and the derivative w.r.t. $\theta_i$ are respectively denoted by $(\nabla_x Y_t^\theta(t_k,x),t\in[t_k,t_{k+1}])$ and $(\partial_{\theta_i}Y_t^\theta(t_k,x),t\in[t_k,t_{k+1}])$ which are solutions to the equations

$$\begin{aligned}\nabla_x Y_t^\theta(t_k,x)&=\mathrm{I}_d+\int_{t_k}^t \nabla_x b(\theta,s,Y_s^\theta(t_k,x))\nabla_x Y_s^\theta(t_k,x)ds\\&\quad+\sum_{j=1}^d\int_{t_k}^t \nabla_x\sigma_j(s,Y_s^\theta(t_k,x))\nabla_x Y_s^\theta(t_k,x)dW_s^j,\end{aligned} \tag{3.4}$$

$$\partial_{\theta_i}Y_t^\theta(t_k,x)=\int_{t_k}^t\left(\partial_{\theta_i}b(\theta,s,Y_s^\theta(t_k,x))+\nabla_x b(\theta,s,Y_s^\theta(t_k,x))\partial_{\theta_i}Y_s^\theta(t_k,x)\right)ds$$

$$+\sum_{j=1}^{d}\int_{t_k}^{t}\nabla_x\sigma_j(s,Y_s^\theta(t_k,x))\partial_{\theta_i}Y_s^\theta(t_k,x)dW_s^j, \tag{3.5}$$

for $i\in\{1,\dots,m\}$, where $\sigma_1,...,\sigma_d:\mathbb{R}_+\times\mathbb{R}^d\to\mathbb{R}^d$ are the columns of the matrix $\sigma$. Furthermore, the random variables $Y_t^\theta(t_k,x)$, $\nabla_xY_t^\theta(t_k,x)$, $(\nabla_xY_t^\theta(t_k,x))^{-1}$ and $\partial_{\theta_i}Y_t^\theta(t_k,x)$ belong to $\mathbb{D}^{1,2}$ for any $t\in[t_k,t_{k+1}]$ (see Nualart [29, Section 2.2]). The Malliavin derivative of $Y_t^\theta(t_k,x)$ satisfies the following equation

$$D_sY_t^\theta(t_k,x)=\sigma(s,Y_s^\theta(t_k,x))+\int_s^t\nabla_xb(\theta,u,Y_u^\theta(t_k,x))D_sY_u^\theta(t_k,x)du$$
$$+\sum_{j=1}^{d}\int_s^t\nabla_x\sigma_j(u,Y_u^\theta(t_k,x))D_sY_u^\theta(t_k,x)dW_u^j,$$

for $s\le t$ a.e., and $D_sY_t^\theta(t_k,x)=0$ for $s>t$ a.e. By (2.59) of Nualart [29], we have

$$D_sY_t^\theta(t_k,x)=\nabla_xY_t^\theta(t_k,x)(\nabla_xY_s^\theta(t_k,x))^{-1}\sigma(s,Y_s^\theta(t_k,x))\mathbf{1}_{[t_k,t]}(s). \tag{3.6}$$

For all $k\in\{0,...,n-1\}$, $x\in\mathbb{R}^d$, the probability law of $Y^\theta$ starting at $x$ at time $t_k$ is denoted by $\widetilde{\mathrm{P}}_{t_k,x}^\theta$, i.e., $\widetilde{\mathrm{P}}_{t_k,x}^\theta(A)=\widetilde{\mathrm{E}}[\mathbf{1}_A|Y_{t_k}^\theta=x]$ for all $A\in\widetilde{\mathcal{F}}$, and the expectation w.r.t. $\widetilde{\mathrm{P}}_{t_k,x}^\theta$ is denoted by $\widetilde{\mathrm{E}}_{t_k,x}^\theta$, i.e., $\widetilde{\mathrm{E}}_{t_k,x}^\theta[V]=\widetilde{\mathrm{E}}[V|Y_{t_k}^\theta=x]$ for all $\widetilde{\mathcal{F}}$-measurable random variables $V$.

Following the approach developed by Gobet in [11, Proposition 4.1], the score function is represented in terms of a conditional expectation involving the Skorohod integral.

**Lemma 3.1.** *Under **(A1)**, **(A2)** and **(A3)**(a)-(b), for all $i\in\{1,\dots,m\}$, $k\in\{0,...,n-1\}$, $\theta\in\Theta$, and $x,y\in\mathbb{R}^d$,*

$$\frac{\partial_{\theta_i}p^\theta}{p^\theta}(t_k,t_{k+1},x,y)=\frac{1}{\Delta_n}\widetilde{\mathrm{E}}_{t_k,x}^\theta\left[\delta\left(U^\theta(t_k,x)\partial_{\theta_i}Y_{t_{k+1}}^\theta(t_k,x)\right)\big|Y_{t_{k+1}}^\theta=y\right],$$

*where $U^\theta(t_k,x)=(U_t^\theta(t_k,x),t\in[t_k,t_{k+1}])$ with $U_t^\theta(t_k,x)=(D_tY_{t_{k+1}}^\theta(t_k,x))^{-1}$.*

Next, the Skorohod integral appearing in Lemma 3.1 can be decomposed as follows.

**Lemma 3.2.** *Under **(A1)**, **(A2)** and **(A3)**(a)-(b), for all $i\in\{1,\dots,m\}$, $k\in\{0,...,n-1\}$, $\theta\in\Theta$, and $x\in\mathbb{R}^d$,*

$$\delta\left(U^\theta(t_k,x)\partial_{\theta_i}Y_{t_{k+1}}^\theta(t_k,x)\right)$$
$$=\Delta_n(\partial_{\theta_i}b(\theta,t_k,x))^*\sigma^{-2}(t_k,x)\left(Y_{t_{k+1}}^\theta-Y_{t_k}^\theta-b(\theta,t_k,Y_{t_k}^\theta)\Delta_n\right)$$
$$-R_1^{\theta,k}+R_2^{\theta,k}+R_3^{\theta,k}-R_4^{\theta,k}-R_5^{\theta,k},$$

*where*

$$R_1^{\theta,k}=\int_{t_k}^{t_{k+1}}\int_s^{t_{k+1}}\mathrm{tr}\Big(D_s\Big(((\nabla_xY_u^\theta(t_k,x))^{-1}\partial_{\theta_i}b(\theta,u,Y_u^\theta(t_k,x)))^*\Big)$$

$$\cdot \sigma^{-1}(s, Y_s^\theta(t_k, x)) \nabla_x Y_s^\theta(t_k, x)\Big) duds,$$

$$R_2^{\theta,k} = \int_{t_k}^{t_{k+1}} \left((\nabla_x Y_s^\theta(t_k, x))^{-1} \partial_{\theta_i} b(\theta, s, Y_s^\theta(t_k, x))\right)^* ds$$
$$\cdot \int_{t_k}^{t_{k+1}} \big((\nabla_x Y_s^\theta(t_k, x))^* \sigma^{-1}(s, Y_s^\theta(t_k, x))$$
$$-(\nabla_x Y_{t_k}^\theta(t_k, x))^* \sigma^{-1}(t_k, Y_{t_k}^\theta(t_k, x))\big) dW_s,$$

$$R_3^{\theta,k} = \int_{t_k}^{t_{k+1}} \big(((\nabla_x Y_s^\theta(t_k, x))^{-1} \partial_{\theta_i} b(\theta, s, Y_s^\theta(t_k, x)))^*$$
$$-((\nabla_x Y_{t_k}^\theta(t_k, x))^{-1} \partial_{\theta_i} b(\theta, t_k, Y_{t_k}^\theta(t_k, x)))^*\big) ds$$
$$\cdot \int_{t_k}^{t_{k+1}} (\nabla_x Y_{t_k}^\theta(t_k, x))^* \sigma^{-1}(t_k, Y_{t_k}^\theta(t_k, x)) dW_s,$$

$$R_4^{\theta,k} = \Delta_n (\partial_{\theta_i} b(\theta, t_k, Y_{t_k}^\theta))^* \sigma^{-2}(t_k, Y_{t_k}^\theta) \int_{t_k}^{t_{k+1}} \left(b(\theta, s, Y_s^\theta) - b(\theta, t_k, Y_{t_k}^\theta)\right) ds,$$

$$R_5^{\theta,k} = \Delta_n (\partial_{\theta_i} b(\theta, t_k, Y_{t_k}^\theta))^* \sigma^{-2}(t_k, Y_{t_k}^\theta) \int_{t_k}^{t_{k+1}} \left(\sigma(s, Y_s^\theta) - \sigma(t_k, Y_{t_k}^\theta)\right) dW_s.$$

The following explicit expression of the score function is due to Lemmas 3.1 and 3.2.

**Lemma 3.3.** *Under **(A1)**, **(A2)** and **(A3)**(a)-(b), for all $i \in \{1, \ldots, m\}$, $k \in \{0, ..., n-1\}$, $\theta \in \Theta$, and $x, y \in \mathbb{R}^d$,*

$$\frac{\partial_{\theta_i} p^\theta}{p^\theta} (t_k, t_{k+1}, x, y)$$
$$= (\partial_{\theta_i} b(\theta, t_k, x))^* \sigma^{-2}(t_k, x) \left(y - x - b(\theta, t_k, x)\Delta_n\right)$$
$$+ \frac{1}{\Delta_n} \widetilde{\mathrm{E}}_{t_k,x}^\theta \left[-R_1^{\theta,k} + R_2^{\theta,k} + R_3^{\theta,k} - R_4^{\theta,k} - R_5^{\theta,k} \middle| Y_{t_{k+1}}^\theta = y\right].$$

Using **(A1)**, **(A2)**, **(A3)**(a)-(b) and Gronwall's inequality, it is straightforward that for any $k \in \{0, ..., n-1\}$, $\theta \in \Theta$, $x \in \mathbb{R}^d$ and $p \geq 2$, there exists a constant $C > 0$ such that

$$\widetilde{\mathrm{E}}_{t_k,x}^\theta \left[\left|\nabla_x Y_t^\theta(t_k, x)\right|^p\right] + \widetilde{\mathrm{E}}_{t_k,x}^\theta \left[\left|(\nabla_x Y_t^\theta(t_k, x))^{-1}\right|^p\right]$$
$$+ \sup_{s \in [t_k, t_{k+1}]} \widetilde{\mathrm{E}}_{t_k,x}^\theta \left[\left|D_s Y_t^\theta(t_k, x)\right|^p\right]$$
$$+ \sup_{s \in [t_k, t_{k+1}]} \widetilde{\mathrm{E}}_{t_k,x}^\theta \left[\left|D_s(\nabla_x Y_t^\theta(t_k, x))\right|^p\right] \leq C, \tag{3.7}$$

$$\widetilde{\mathrm{E}}_{t_k,x}^\theta \left[\left|\partial_{\theta_i} Y_t^\theta(t_k, x)\right|^p\right] \leq C\left(1 + |x|^p\right), \tag{3.8}$$

for $t \in [t_k, t_{k+1}]$, where $C$ is uniform in $\theta$. Consequently, we have the following estimates.

**Lemma 3.4.** *Under conditions **(A1)**, **(A2)**, and **(A3)**(a)-(b), for any $k \in \{0, ..., n-1\}$, $\theta \in \Theta$, $x \in \mathbb{R}^d$ and $p \geq 2$, there exists a constant $C > 0$ such that*

$$\widetilde{\mathrm{E}}^{\theta}_{t_k,x}\left[-R_1^{\theta,k} + R_2^{\theta,k} + R_3^{\theta,k}\right] = 0, \tag{3.9}$$

$$\widetilde{\mathrm{E}}^{\theta}_{t_k,x}\left[\left|-R_1^{\theta,k} + R_2^{\theta,k} + R_3^{\theta,k}\right|^p\right] \leq C\Delta_n^{2p}\left(1+|x|^p\right). \tag{3.10}$$

The equality (3.9) follows from the proof of Lemma 3.2, $\widetilde{\mathrm{E}}^{\theta}_{t_k,x}[W_{t_{k+1}} - W_{t_k}] = 0$, and $\widetilde{\mathrm{E}}^{\theta}_{t_k,x}[\delta(U^{\theta}(t_k,x)\partial_{\theta_i}Y^{\theta}_{t_{k+1}}(t_k,x))] = 0$. The estimate (3.10) can be obtained by using Itô's formula, Burkholder-Davis-Gundy's inequality, conditions **(A1)**, **(A2)**, **(A3)**(a)-(b) and (3.7).

### *3.2. Proof of Theorem 2.1*

*Proof.* First, using the decomposition (3.1), Lemma 3.3 and equation (1.1), we obtain

$$\begin{aligned}
&\log \frac{d\mathrm{P}_n^{\theta^0+\varphi_{n\Delta_n}(\theta^0)u}}{d\mathrm{P}_n^{\theta^0}}(X^{n,\theta^0})\\
&= \sum_{k=0}^{n-1}\sum_{i=1}^{m}\varphi^i_{n\Delta_n}(\theta^0)u_i\int_0^1\bigg((\partial_{\theta_i}b(\theta_i^{0+}(\ell),t_k,X^{\theta^0}_{t_k}))^*\sigma^{-2}(t_k,X^{\theta^0}_{t_k})\\
&\qquad\cdot\Big(X^{\theta^0}_{t_{k+1}} - X^{\theta^0}_{t_k} - b(\theta_i^{0+}(\ell),t_k,X^{\theta^0}_{t_k})\Delta_n\Big)\\
&\qquad+\frac{1}{\Delta_n}\widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}\Big[R^{\theta_i^{0+}(\ell),k} - R_4^{\theta_i^{0+}(\ell),k} - R_5^{\theta_i^{0+}(\ell),k}|Y^{\theta_i^{0+}(\ell)}_{t_{k+1}} = X^{\theta^0}_{t_{k+1}}\Big]\bigg)d\ell\\
&= \sum_{k=0}^{n-1}\sum_{i=1}^{m}\xi_{i,k,n} + \sum_{k=0}^{n-1}\sum_{i=1}^{m}\frac{\varphi^i_{n\Delta_n}(\theta^0)u_i}{\Delta_n}\int_0^1\bigg\{Z^{4,\ell}_{i,k,n} + Z^{5,\ell}_{i,k,n}\\
&\quad+\widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}\Big[R^{\theta_i^{0+}(\ell),k} - R_4^{\theta_i^{0+}(\ell),k} - R_5^{\theta_i^{0+}(\ell),k}\Big|Y^{\theta_i^{0+}(\ell)}_{t_{k+1}} = X^{\theta^0}_{t_{k+1}}\Big]\bigg\}d\ell,
\end{aligned} \tag{3.11}$$

where $R^{\theta_i^{0+}(\ell),k} = -R_1^{\theta_i^{0+}(\ell),k} + R_2^{\theta_i^{0+}(\ell),k} + R_3^{\theta_i^{0+}(\ell),k}$ and

$$\begin{aligned}
\xi_{i,k,n} &= \varphi^i_{n\Delta_n}(\theta^0)u_i\int_0^1(\partial_{\theta_i}b(\theta_i^{0+}(\ell),t_k,X^{\theta^0}_{t_k}))^*\sigma^{-2}(t_k,X^{\theta^0}_{t_k})\\
&\quad\cdot\Big(\sigma(t_k,X^{\theta^0}_{t_k})(B_{t_{k+1}}-B_{t_k}) + \big(b(\theta^0,t_k,X^{\theta^0}_{t_k}) - b(\theta_i^{0+}(\ell),t_k,X^{\theta^0}_{t_k})\big)\Delta_n\Big)d\ell,\\
Z^{4,\ell}_{i,k,n} &= \Delta_n(\partial_{\theta_i}b(\theta_i^{0+}(\ell),t_k,X^{\theta^0}_{t_k}))^*\sigma^{-2}(t_k,X^{\theta^0}_{t_k})\\
&\qquad\cdot\int_{t_k}^{t_{k+1}}\Big(b(\theta^0,s,X^{\theta^0}_s) - b(\theta^0,t_k,X^{\theta^0}_{t_k})\Big)ds,\\
Z^{5,\ell}_{i,k,n} &= \Delta_n(\partial_{\theta_i}b(\theta_i^{0+}(\ell),t_k,X^{\theta^0}_{t_k}))^*\sigma^{-2}(t_k,X^{\theta^0}_{t_k})\\
&\qquad\cdot\int_{t_k}^{t_{k+1}}\Big(\sigma(s,X^{\theta^0}_s) - \sigma(t_k,X^{\theta^0}_{t_k})\Big)dB_s.
\end{aligned}$$

Next, we decompose $\sum_{k=0}^{n-1}\sum_{i=1}^{m}\xi_{i,k,n}$. For this, observe that $\xi_{i,k,n}=\xi_{1,i,k,n}+\xi_{2,i,k,n}$, where

$$\begin{aligned}\xi_{1,i,k,n}&=\varphi^i_{n\Delta_n}(\theta^0)u_i\int_0^1(\partial_{\theta_i}b(\theta^0,t_k,X^{\theta^0}_{t_k}))^*\sigma^{-2}(t_k,X^{\theta^0}_{t_k})\\&\quad\cdot\Big(\sigma(t_k,X^{\theta^0}_{t_k})(B_{t_{k+1}}-B_{t_k})+\big(b(\theta^0,t_k,X^{\theta^0}_{t_k})-b(\theta^{0+}_i(\ell),t_k,X^{\theta^0}_{t_k})\big)\Delta_n\Big)\,d\ell,\\\xi_{2,i,k,n}&=\varphi^i_{n\Delta_n}(\theta^0)u_i\int_0^1\Big(\partial_{\theta_i}b(\theta^{0+}_i(\ell),t_k,X^{\theta^0}_{t_k})-\partial_{\theta_i}b(\theta^0,t_k,X^{\theta^0}_{t_k})\Big)^*\\&\qquad\cdot\sigma^{-2}(t_k,X^{\theta^0}_{t_k})\Big(\sigma(t_k,X^{\theta^0}_{t_k})(B_{t_{k+1}}-B_{t_k})\\&\qquad+\big(b(\theta^0,t_k,X^{\theta^0}_{t_k})-b(\theta^{0+}_i(\ell),t_k,X^{\theta^0}_{t_k})\big)\Delta_n\Big)d\ell.\end{aligned}$$

Then, we write $\xi_{1,i,k,n}=\xi_{1,1,i,k,n}+\xi_{1,2,i,k,n}$, where

$$\begin{aligned}\xi_{1,1,i,k,n}&=\varphi^i_{n\Delta_n}(\theta^0)u_i(\partial_{\theta_i}b(\theta^0,t_k,X^{\theta^0}_{t_k}))^*\sigma^{-1}(t_k,X^{\theta^0}_{t_k})(B_{t_{k+1}}-B_{t_k}),\\\xi_{1,2,i,k,n}&=\varphi^i_{n\Delta_n}(\theta^0)u_i\int_0^1(\partial_{\theta_i}b(\theta^0,t_k,X^{\theta^0}_{t_k}))^*\sigma^{-2}(t_k,X^{\theta^0}_{t_k})\\&\qquad\cdot\Big(b(\theta^0,t_k,X^{\theta^0}_{t_k})-b(\theta^{0+}_i(\ell),t_k,X^{\theta^0}_{t_k})\Big)\,\Delta_n d\ell.\end{aligned}$$

Notice that

$$\sum_{k=0}^{n-1}\sum_{i=1}^{m}\xi_{1,1,i,k,n}=u^*\varphi_{n\Delta_n}(\theta^0)\nabla_\theta\ell_{n\Delta_n}(\theta^0)-\sum_{k=0}^{n-1}H_{1,k,n},$$

where $\nabla_\theta\ell_T(\theta)$ is given by (2.1) and

$$\begin{aligned}H_{1,k,n}=u^*\varphi_{n\Delta_n}(\theta^0)\int_{t_k}^{t_{k+1}}&\Big(\sigma^{-1}(t,X^{\theta^0}_t)\nabla_\theta b(\theta^0,t,X^{\theta^0}_t)\\&-\sigma^{-1}(t_k,X^{\theta^0}_{t_k})\nabla_\theta b(\theta^0,t_k,X^{\theta^0}_{t_k})\Big)\,dB_t.\end{aligned}$$

Now, we treat $\xi_{1,2,i,k,n}$. For this, using the mean value theorem, we write

$$\begin{aligned}&b(\theta^0,t_k,X^{\theta^0}_{t_k})-b(\theta^{0+}_i(\ell),t_k,X^{\theta^0}_{t_k})\\&=-\Big(b(\theta^{0+}_i(\ell),t_k,X^{\theta^0}_{t_k})-b(\theta^{0+}_{i+1},t_k,X^{\theta^0}_{t_k})+b(\theta^{0+}_{i+1},t_k,X^{\theta^0}_{t_k})\\&\qquad-b(\theta^{0+}_{i+2},t_k,X^{\theta^0}_{t_k})+\cdots+b(\theta^{0+}_{m-1},t_k,X^{\theta^0}_{t_k})-b(\theta^{0+}_{m},t_k,X^{\theta^0}_{t_k})\\&\qquad+b(\theta^{0+}_{m},t_k,X^{\theta^0}_{t_k})-b(\theta^0,t_k,X^{\theta^0}_{t_k})\Big)\\&=-\Big(\ell\varphi^i_{n\Delta_n}(\theta^0)u_i\int_0^1\partial_{\theta_i}b(\theta^{0+}_i(\alpha\ell),t_k,X^{\theta^0}_{t_k})d\alpha\\&\qquad+\varphi^{i+1}_{n\Delta_n}(\theta^0)u_{i+1}\int_0^1\partial_{\theta_{i+1}}b(\theta^{0+}_{i+1}(\alpha),t_k,X^{\theta^0}_{t_k})d\alpha\end{aligned}$$

$$
\begin{aligned}
&+\cdots+\varphi^m_{n\Delta_n}(\theta^0)u_m\int_0^1\partial_{\theta_m}b(\theta^{0+}_m(\alpha),t_k,X^{\theta^0}_{t_k})d\alpha\Big)\\
=&-\Big(\ell\varphi^i_{n\Delta_n}(\theta^0)u_i\partial_{\theta_i}b(\theta^0,t_k,X^{\theta^0}_{t_k})\\
&+\varphi^{i+1}_{n\Delta_n}(\theta^0)u_{i+1}\partial_{\theta_{i+1}}b(\theta^0,t_k,X^{\theta^0}_{t_k})+\cdots+\varphi^m_{n\Delta_n}(\theta^0)u_m\partial_{\theta_m}b(\theta^0,t_k,X^{\theta^0}_{t_k})\Big)\\
&-\Big(\ell\varphi^i_{n\Delta_n}(\theta^0)u_i\int_0^1\big(\partial_{\theta_i}b(\theta^{0+}_i(\alpha\ell),t_k,X^{\theta^0}_{t_k})-\partial_{\theta_i}b(\theta^0,t_k,X^{\theta^0}_{t_k})\big)d\alpha\\
&+\varphi^{i+1}_{n\Delta_n}(\theta^0)u_{i+1}\int_0^1\big(\partial_{\theta_{i+1}}b(\theta^{0+}_{i+1}(\alpha),t_k,X^{\theta^0}_{t_k})-\partial_{\theta_{i+1}}b(\theta^0,t_k,X^{\theta^0}_{t_k})\big)d\alpha\\
&+\cdots+\varphi^m_{n\Delta_n}(\theta^0)u_m\int_0^1\big(\partial_{\theta_m}b(\theta^{0+}_m(\alpha),t_k,X^{\theta^0}_{t_k})-\partial_{\theta_m}b(\theta^0,t_k,X^{\theta^0}_{t_k})\big)d\alpha\Big),
\end{aligned}
$$

where, to simplify the exposition, we have set for $j\in\{i+1,\dots,m\}$,

$$
\begin{aligned}
\theta^{0+}_i(\alpha\ell):=&(\theta^0_1,\dots,\theta^0_{i-1},\theta^0_i+\alpha\ell\varphi^i_{n\Delta_n}(\theta^0)u_i,\theta^0_{i+1}+\varphi^{i+1}_{n\Delta_n}(\theta^0)u_{i+1},\\
&\dots,\theta^0_m+\varphi^m_{n\Delta_n}(\theta^0)u_m),\\
\theta^{0+}_j(\alpha):=&(\theta^0_1,\dots,\theta^0_{j-1},\theta^0_j+\alpha\varphi^j_{n\Delta_n}(\theta^0)u_j,\theta^0_{j+1}+\varphi^{j+1}_{n\Delta_n}(\theta^0)u_{j+1},\\
&\dots,\theta^0_m+\varphi^m_{n\Delta_n}(\theta^0)u_m).
\end{aligned}
$$

Thus,

$$
\begin{aligned}
\sum_{k=0}^{n-1}\sum_{i=1}^m\xi_{1,2,i,k,n}=&-\frac12u^*\varphi_{n\Delta_n}(\theta^0)\left\langle\nabla_\theta\ell(\theta^0)\right\rangle_{n\Delta_n}\varphi_{n\Delta_n}(\theta^0)u+\frac12\sum_{k=0}^{n-1}H_{2,k,n}\\
&-\sum_{k=0}^{n-1}\sum_{i=1}^m\Big(K_{i,k,n}+K_{i+1,k,n}+\cdots+K_{m,k,n}\Big),
\end{aligned}
$$

where $\langle\nabla_\theta\ell(\theta)\rangle_T$ is given by (2.2), and for $j\in\{i+1,\dots,m\}$,

$$
\begin{aligned}
K_{i,k,n}=&\varphi^i_{n\Delta_n}(\theta^0)u_i\int_0^1\int_0^1(\partial_{\theta_i}b(\theta^0,t_k,X^{\theta^0}_{t_k}))^*\sigma^{-2}(t_k,X^{\theta^0}_{t_k})\ell\varphi^i_{n\Delta_n}(\theta^0)u_i\\
&\cdot\Big(\partial_{\theta_i}b(\theta^{0+}_i(\alpha\ell),t_k,X^{\theta^0}_{t_k})-\partial_{\theta_i}b(\theta^0,t_k,X^{\theta^0}_{t_k})\Big)\Delta_n d\alpha d\ell,\\
K_{j,k,n}=&\varphi^i_{n\Delta_n}(\theta^0)u_i\int_0^1(\partial_{\theta_i}b(\theta^0,t_k,X^{\theta^0}_{t_k}))^*\sigma^{-2}(t_k,X^{\theta^0}_{t_k})\varphi^j_{n\Delta_n}(\theta^0)u_j\\
&\cdot\Big(\partial_{\theta_j}b(\theta^{0+}_j(\alpha),t_k,X^{\theta^0}_{t_k})-\partial_{\theta_j}b(\theta^0,t_k,X^{\theta^0}_{t_k})\Big)\Delta_n d\alpha,\\
H_{2,k,n}=&u^*\varphi_{n\Delta_n}(\theta^0)\int_{t_k}^{t_{k+1}}\Big((\nabla_\theta b(\theta^0,t,X^{\theta^0}_t))^*\sigma^{-2}(t,X^{\theta^0}_t)\nabla_\theta b(\theta^0,t,X^{\theta^0}_t)\\
&-(\nabla_\theta b(\theta^0,t_k,X^{\theta^0}_{t_k}))^*\sigma^{-2}(t_k,X^{\theta^0}_{t_k})\nabla_\theta b(\theta^0,t_k,X^{\theta^0}_{t_k})\Big)dt\,\varphi_{n\Delta_n}(\theta^0)u.
\end{aligned}
$$

Therefore, we have shown that

$$
\sum_{k=0}^{n-1}\sum_{i=1}^m\xi_{i,k,n}=u^*\varphi_{n\Delta_n}(\theta^0)\nabla_\theta\ell_{n\Delta_n}(\theta^0)
$$

$$
\begin{aligned}
&- \frac{1}{2} u^* \varphi_{n\Delta_n}(\theta^0) \left\langle \nabla_\theta \ell(\theta^0) \right\rangle_{n\Delta_n} \varphi_{n\Delta_n}(\theta^0) u - \sum_{k=0}^{n-1} H_{1,k,n} \\
&+ \frac{1}{2} \sum_{k=0}^{n-1} H_{2,k,n} - \sum_{k=0}^{n-1} \sum_{i=1}^{m} (K_{i,k,n} + K_{i+1,k,n} + \cdots + K_{m,k,n}) \\
&+ \sum_{k=0}^{n-1} \sum_{i=1}^{m} \xi_{2,i,k,n}.
\end{aligned} \tag{3.12}
$$

Next, using condition **(A4)** and the multivariate central limit theorem for continuous local martingales (see [37, Theorem 4.1]), we obtain that as $n \to \infty$,

$$
\begin{aligned}
&\Big( \varphi_{n\Delta_n}(\theta^0) \nabla_\theta \ell_{n\Delta_n}(\theta^0), \varphi_{n\Delta_n}(\theta^0) \left\langle \nabla_\theta \ell(\theta^0) \right\rangle_{n\Delta_n} \varphi_{n\Delta_n}(\theta^0) \Big) \\
&\xrightarrow{\mathcal{L}(\widehat{\mathrm{P}}^{\theta^0})} \big( \Gamma(\theta^0)^{1/2} \mathcal{N}(0, I_m), \Gamma(\theta^0) \big),
\end{aligned} \tag{3.13}
$$

where $\mathcal{N}(0, I_m)$ is independent of $\Gamma(\theta^0)$. Thus, from (3.11), (3.12), (3.13) and Lemmas 4.8-4.11 below, we complete the proof of the desired result. □

## 4. Negligible contributions

This section aims to prove the convergence of the remainder terms in the expansion. For this, we first need some preliminary results.

**Lemma 4.1.** *Let $n \in \mathbb{N}$ and $(\zeta_{k,n})_{k\geq 1}$ be a sequence of random variables defined on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathrm{P})$ such that $\zeta_{k,n}$ is $\mathcal{F}_{t_{k+1}}$-measurable for all $k$.*

a) [10, Lemma 9] *Assume that as $n \to \infty$,*

$$
\text{(i)} \sum_{k=0}^{n-1} \mathrm{E}\left[\zeta_{k,n} | \mathcal{F}_{t_k}\right] \xrightarrow{\mathrm{P}} 0, \quad \text{and} \quad \text{(ii)} \sum_{k=0}^{n-1} \mathrm{E}\left[\zeta_{k,n}^2 | \mathcal{F}_{t_k}\right] \xrightarrow{\mathrm{P}} 0.
$$

*Then as $n \to \infty$, $\sum_{k=0}^{n-1} \zeta_{k,n} \xrightarrow{\mathrm{P}} 0$.*

b) [17, Lemma 3.4] *Assume that $\sum_{k=0}^{n-1} \mathrm{E}\left[|\zeta_{k,n}| | \mathcal{F}_{t_k}\right] \xrightarrow{\mathrm{P}} 0$ as $n \to \infty$. Then $\sum_{k=0}^{n-1} \zeta_{k,n} \xrightarrow{\mathrm{P}} 0$ as $n \to \infty$.*

Using (1.1) and Burkholder-Davis-Gundy's inequality, we get the following estimates.

**Lemma 4.2.** *Assume conditions **(A1)**-**(A2)**.*

(i) *For any $p \geq 2$, $k \in \{0, ..., n-1\}$, $\theta \in \Theta$ and $x \in \mathbb{R}^d$, there exists a constant $C > 0$ such that for all $t \in [t_k, t_{k+1}]$,*

$$
\widehat{\mathrm{E}}^\theta \left[ \left| X_t^\theta - X_{t_k}^\theta \right|^p \big| \widehat{\mathcal{F}}_{t_k} \right] \leq C \left| t - t_k \right|^{\frac{p}{2}} \left( 1 + |X_{t_k}^\theta|^p \right).
$$

(ii) *For any $k \in \{0,...,n-1\}$ and function $g$ defined on $\Theta\times\mathbb{R}^d$ with polynomial growth in $x$ uniformly in $\theta \in \Theta$, that is, $|g(\theta,x)| \leq c(1+|x|^p)$ for some constants $c,p>0$, there exists a constant $C>0$ such that for all $t \in [t_k,t_{k+1}]$,*
$$\widehat{\mathrm{E}}^{\theta}\left[\left|g(\theta,X_t^{\theta})\right|\middle|\widehat{\mathcal{F}}_{t_k}\right] \leq C\left(1+|X_{t_k}^{\theta}|^p\right).$$

Thanks to condition **(A5)**, the following estimate will be useful in the sequel.

**Lemma 4.3.** *Assume **(A1)**, **(A5)**. For any $p \geq 1$, there exists a constant $C>0$ such that*
$$\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}\left[\left|X_{t_k}^{\theta^0}\right|^p\right] \leq C\sum_{k=0}^{n-1}\psi_{t_k}^{-p}(\theta^0).$$

Now the change of measures on $I_k=[t_k,t_{k+1}]$ is applied. The probability law of $X^{\theta}$ starting at $x$ at time $t_k$ is denoted by $\widehat{\mathrm{P}}_{t_k,x}^{\theta}$, i.e., $\widehat{\mathrm{P}}_{t_k,x}^{\theta}(A)=\widehat{\mathrm{E}}[\mathbf{1}_A|X_{t_k}^{\theta}=x]$ for all $A \in \widehat{\mathcal{F}}$, and the expectation w.r.t. $\widehat{\mathrm{P}}_{t_k,x}^{\theta}$ is denoted by $\widehat{\mathrm{E}}_{t_k,x}^{\theta}$, i.e., $\widehat{\mathrm{E}}_{t_k,x}^{\theta}[V]=\widehat{\mathrm{E}}[V|X_{t_k}^{\theta}=x]$ for all $\widehat{\mathcal{F}}$-measurable random variables $V$. From [18, Chapter III] and [25, Chapter 7], under **(A1)** and **(A2)**, for all $\theta,\theta^1 \in \Theta$, $x \in \mathbb{R}^d$ and $k \in \{0,...,n-1\}$, the probability measures $\widehat{\mathrm{P}}_{t_k,x}^{\theta}$ and $\widehat{\mathrm{P}}_{t_k,x}^{\theta^1}$ are absolutely continuous w.r.t. each other and its Radon-Nikodym derivative is given by
$$\begin{aligned}\frac{d\widehat{\mathrm{P}}_{t_k,x}^{\theta}}{d\widehat{\mathrm{P}}_{t_k,x}^{\theta^1}}((X_t^{\theta^1})_{t\in I_k})&=\exp\bigg\{\int_{t_k}^{t_{k+1}}(b(\theta,t,X_t^{\theta^1})-b(\theta^1,t,X_t^{\theta^1}))^*\sigma^{-1}(t,X_t^{\theta^1})dB_t\\&-\frac{1}{2}\int_{t_k}^{t_{k+1}}\left|(b(\theta,t,X_t^{\theta^1})-b(\theta^1,t,X_t^{\theta^1}))^*\sigma^{-1}(t,X_t^{\theta^1})\right|^2dt\bigg\}.\end{aligned}\tag{4.1}$$

To treat the conditional expectations in the remainder terms, two following lemmas will be used. For this, we fix some $k \in \{0,\dots,n-1\}$, and let $V^{\theta}$ be a $\widetilde{\mathcal{F}}_{t_{k+1}}$-measurable random variable which will be $R^{\theta,k}$, $(R^{\theta,k})^2$, $R_5^{\theta,k}$, $(R_5^{\theta,k})^2$. Moreover, let $\widehat{V}^{\theta}$ be a $\widehat{\mathcal{F}}_{t_{k+1}}$-measurable random variable which will be defined in the proof of Lemma 4.11.

**Lemma 4.4.** *Assume **(A1)** and **(A2)**. Then, for any $k \in \{0,...,n-1\}$, $\theta \in \Theta$, $x \in \mathbb{R}^d$,*
$$\begin{aligned}&\widehat{\mathrm{E}}_{t_k,x}^{\theta^0}\left[\widetilde{\mathrm{E}}_{t_k,x}^{\theta}\left[V^{\theta}|Y_{t_{k+1}}^{\theta}=X_{t_{k+1}}^{\theta^0}\right]\right]\\&=\widetilde{\mathrm{E}}_{t_k,x}^{\theta}\left[V^{\theta}\right]+\widetilde{\mathrm{E}}_{t_k,x}^{\theta}\left[\widetilde{\mathrm{E}}_{t_k,x}^{\theta}[V^{\theta}|Y_{t_{k+1}}^{\theta}]\left(\frac{d\widetilde{\mathrm{P}}_{t_k,x}^{\theta^0}}{d\widetilde{\mathrm{P}}_{t_k,x}^{\theta}}((Y_t^{\theta})_{t\in I_k})-1\right)\right].\end{aligned}\tag{4.2}$$

*Similarly, we have*
$$\widehat{\mathrm{E}}_{t_k,x}^{\theta^0}\left[\widehat{V}^{\theta^0}\right]=\widehat{\mathrm{E}}_{t_k,x}^{\theta}\left[\widehat{V}^{\theta}\right]+\widehat{\mathrm{E}}_{t_k,x}^{\theta}\left[\widehat{V}^{\theta}\left(\frac{d\widehat{\mathrm{P}}_{t_k,x}^{\theta^0}}{d\widehat{\mathrm{P}}_{t_k,x}^{\theta}}((X_t^{\theta})_{t\in I_k})-1\right)\right].\tag{4.3}$$

The following lemma is used to estimate two second terms appearing in (4.2) and (4.3) when $\theta = \theta_i^{0+}(\ell)$. For this, for $j \in \{1, \dots, m\}$ we set

$$\theta_j(0+) := (\theta_1^0, \dots, \theta_{j-1}^0, \theta_j, \theta_{j+1}^0 + \varphi_{n\Delta_n}^{j+1}(\theta^0)u_{j+1}, \dots, \theta_m^0 + \varphi_{n\Delta_n}^m(\theta^0)u_m).$$

**Lemma 4.5.** *Assume* ***(A1)***, ***(A2)***, ***(A3)***(b) *and $q > 1$. Then, there exist constants $C, C_i, \dots, C_m > 0$ such that for any $i \in \{1, \dots, m\}$, $k \in \{0, ..., n-1\}$, $x \in \mathbb{R}^d$ and $n$ large enough,*

$$\begin{aligned}
&\left| \widetilde{\mathrm{E}}_{t_k,x}^{\theta_i^{0+}(\ell)} \left[ \widetilde{\mathrm{E}}_{t_k,x}^{\theta_i^{0+}(\ell)} \left[ V^{\theta_i^{0+}(\ell)} | Y_{t_{k+1}}^{\theta_i^{0+}(\ell)} \right] \Big( \frac{d\widetilde{\mathrm{P}}_{t_k,x}^{\theta^0}}{d\widetilde{\mathrm{P}}_{t_k,x}^{\theta_i^{0+}(\ell)}} ((Y_t^{\theta_i^{0+}(\ell)})_{t\in I_k}) - 1 \Big) \right] \right| \\
&\le C \Big( \widetilde{\mathrm{E}}_{t_k,x}^{\theta_i^{0+}(\ell)} [|V^{\theta_i^{0+}(\ell)}|^q] \Big)^{\frac{1}{q}} \sqrt{\Delta_n} (1+|x|) \\
&\quad \times \bigg( \varphi_{n\Delta_n}^i(\theta^0) \Big( 1 + \sqrt{\Delta_n} \varphi_{n\Delta_n}^i(\theta^0)|x| \Big) e^{C_i(\varphi_{n\Delta_n}^i(\theta^0))^2 \Delta_n |x|^2} \\
&\quad + \varphi_{n\Delta_n}^{i+1}(\theta^0) \Big( 1 + \sqrt{\Delta_n} \big( \varphi_{n\Delta_n}^i(\theta^0) + \varphi_{n\Delta_n}^{i+1}(\theta^0) \big) |x| \Big) \\
&\quad \times e^{C_{i+1}((\varphi_{n\Delta_n}^i(\theta^0))^2 + (\varphi_{n\Delta_n}^{i+1}(\theta^0))^2)\Delta_n |x|^2} + \cdots \\
&\quad + \varphi_{n\Delta_n}^m(\theta^0) \Big( 1 + \sqrt{\Delta_n} \big( \varphi_{n\Delta_n}^i(\theta^0) + \cdots + \varphi_{n\Delta_n}^m(\theta^0) \big) |x| \Big) \\
&\quad \times e^{C_m((\varphi_{n\Delta_n}^i(\theta^0))^2 + \cdots + (\varphi_{n\Delta_n}^m(\theta^0))^2)\Delta_n |x|^2} \bigg).
\end{aligned} \tag{4.4}$$

*Similarly, let $q > 1$. Then, there exists a constant $C > 0$ such that for any $i \in \{1, \dots, m\}$, $k \in \{0, ..., n-1\}$ and $x \in \mathbb{R}^d$,*

$$\begin{aligned}
&\left| \widehat{\mathrm{E}}_{t_k,x}^{\theta_i^{0+}(\ell)} \left[ \widehat{V}^{\theta_i^{0+}(\ell)} \Big( \frac{d\widehat{\mathrm{P}}_{t_k,x}^{\theta^0}}{d\widehat{\mathrm{P}}_{t_k,x}^{\theta_i^{0+}(\ell)}} ((X_t^{\theta_i^{0+}(\ell)})_{t\in I_k}) - 1 \Big) \right] \right| \\
&\le C\sqrt{\Delta_n}\,(1+|x|) \bigg( \Big| \int_{\theta_i^0 + \ell\varphi_{n\Delta_n}^i(\theta^0)u_i}^{\theta_i^0} \Big( \widehat{\mathrm{E}}_{t_k,x}^{\theta_i(0+)} \Big[ \big| \widehat{V}^{\theta_i(0+)} \big|^q \Big] \Big)^{\frac{1}{q}} d\theta_i \Big| \\
&\quad + \Big| \int_{\theta_{i+1}^0 + \varphi_{n\Delta_n}^{i+1}(\theta^0)u_{i+1}}^{\theta_{i+1}^0} \Big( \widehat{\mathrm{E}}_{t_k,x}^{\theta_{i+1}(0+)} \Big[ \big| \widehat{V}^{\theta_{i+1}(0+)} \big|^q \Big] \Big)^{\frac{1}{q}} d\theta_{i+1} \Big| \\
&\quad + \cdots + \Big| \int_{\theta_m^0 + \varphi_{n\Delta_n}^m(\theta^0)u_m}^{\theta_m^0} \Big( \widehat{\mathrm{E}}_{t_k,x}^{\theta_m(0+)} \Big[ \big| \widehat{V}^{\theta_m(0+)} \big|^q \Big] \Big)^{\frac{1}{q}} d\theta_m \Big| \bigg).
\end{aligned} \tag{4.5}$$

**Remark 4.6.** The expression (4.2) in Lemma 4.4 is analogous to the last equality on page 911 of [11]. From (4.2), the estimate (4.4) in Lemma 4.5 is analogous to (4.19) of [11, Proposition 4.2]. In [11], the author used the change of transition densities and the Gaussian lower and upper bounds for the densities. There is no exponential term in (4.19) of [11, Proposition 4.2] since the coefficients are bounded. The arguments we use here are the change of measures and the estimate of the squared exponential moment. The exponential terms appear in (4.4) due to the fact that the drift coefficient and its derivative w.r.t the parameter are unbounded.

**Remark 4.7.** As will be seen in the proof of (4.4), when $|\partial_{\theta_i} b(\theta,t,x)|$ is bounded, the exponential term $e^{C_j(\varphi^i_{n\Delta_n}(\theta^0))^2\Delta_n|x|^2}$ does not appear in the estimate (4.4). As a consequence, condition **(A7)** is not required for $i$.

In all what follows, without further mention, conditions **(A1)**, **(A2)**, **(A3)** and Lemma 4.2 (ii) will be applied repeatedly.

**Lemma 4.8.** *Under conditions **(A1)**-**(A3)** and **(A5)**-**(A6)**, as $n\to\infty$,*

$$-\sum_{k=0}^{n-1} H_{1,k,n} + \frac{1}{2}\sum_{k=0}^{n-1} H_{2,k,n} - \sum_{k=0}^{n-1}\sum_{i=1}^{m}(K_{i,k,n}+K_{i+1,k,n}+\cdots+K_{m,k,n}) + \sum_{k=0}^{n-1}\sum_{i=1}^{m}\xi_{2,i,k,n} \overset{\widehat{\mathrm{P}}^{\theta^0}}{\longrightarrow} 0.$$

**Lemma 4.9.** *Under conditions **(A1)**-**(A3)** and **(A5)**-**(A7)**, as $n\to\infty$,*

$$\sum_{k=0}^{n-1}\sum_{i=1}^{m}\frac{\varphi^i_{n\Delta_n}(\theta^0)u_i}{\Delta_n}\int_0^1 \widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}\Big[R^{\theta_i^{0+}(\ell),k}\big|Y^{\theta_i^{0+}(\ell)}_{t_{k+1}}=X^{\theta^0}_{t_{k+1}}\Big]d\ell \overset{\widehat{\mathrm{P}}^{\theta^0}}{\longrightarrow} 0.$$

*Proof.* We apply Lemma 4.1 a) with

$$\zeta_{k,n}=\zeta_{i,k,n}:=\frac{\varphi^i_{n\Delta_n}(\theta^0)u_i}{\Delta_n}\int_0^1 \widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}\Big[R^{\theta_i^{0+}(\ell),k}\big|Y^{\theta_i^{0+}(\ell)}_{t_{k+1}}=X^{\theta^0}_{t_{k+1}}\Big]d\ell,$$

for $i\in\{1,\ldots,m\}$. Applying (4.2) of Lemma 4.4 to $V^{\theta_i^{0+}(\ell)}=R^{\theta_i^{0+}(\ell),k}$, using the fact that $\widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}[R^{\theta_i^{0+}(\ell),k}]=0$ by (3.9), (4.4) of Lemma 4.5 with $q=2$, and (3.10) with $p=2$,

$$\begin{aligned}
&\Big|\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}\Big[\zeta_{i,k,n}|\widehat{\mathcal{F}}_{t_k}\Big]\Big|\\
&=\Big|\sum_{k=0}^{n-1}\frac{\varphi^i_{n\Delta_n}(\theta^0)u_i}{\Delta_n}\int_0^1\widehat{\mathrm{E}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}\Big[\widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}\Big[R^{\theta_i^{0+}(\ell),k}\big|Y^{\theta_i^{0+}(\ell)}_{t_{k+1}}=X^{\theta^0}_{t_{k+1}}\Big]\Big]d\ell\Big|\\
&\le C\sum_{k=0}^{n-1}\frac{\varphi^i_{n\Delta_n}(\theta^0)}{\sqrt{\Delta_n}}(1+|X^{\theta^0}_{t_k}|)\\
&\quad\times\int_0^1\big(\widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}[|R^{\theta_i^{0+}(\ell),k}|^2]\big)^{\frac12}\Big(\varphi^i_{n\Delta_n}(\theta^0)\big(1+\sqrt{\Delta_n}\varphi^i_{n\Delta_n}(\theta^0)|X^{\theta^0}_{t_k}|\big)\\
&\quad\times e^{C_i(\varphi^i_{n\Delta_n}(\theta^0))^2\Delta_n|X^{\theta^0}_{t_k}|^2}+\varphi^{i+1}_{n\Delta_n}(\theta^0)\big(1+\sqrt{\Delta_n}(\varphi^i_{n\Delta_n}(\theta^0)+\varphi^{i+1}_{n\Delta_n}(\theta^0))|X^{\theta^0}_{t_k}|\big)\\
&\quad\times e^{C_{i+1}((\varphi^i_{n\Delta_n}(\theta^0))^2+(\varphi^{i+1}_{n\Delta_n}(\theta^0))^2)\Delta_n|X^{\theta^0}_{t_k}|^2}+\cdots
\end{aligned}$$

$$
\begin{aligned}
&+\varphi^m_{n\Delta_n}(\theta^0)\big(1+\sqrt{\Delta_n}(\varphi^i_{n\Delta_n}(\theta^0)+\cdots+\varphi^m_{n\Delta_n}(\theta^0))|X^{\theta^0}_{t_k}|\big)\\
&\times e^{C_m((\varphi^i_{n\Delta_n}(\theta^0))^2+\cdots+(\varphi^m_{n\Delta_n}(\theta^0))^2)\Delta_n|X^{\theta^0}_{t_k}|^2}\Big)d\ell\\
\le{}& C(\varphi^i_{n\Delta_n}(\theta^0))^2\Delta_n^{\frac{3}{2}}\sum_{k=0}^{n-1}\Big(1+|X^{\theta^0}_{t_k}|^2+\sqrt{\Delta_n}\varphi^i_{n\Delta_n}(\theta^0)|X^{\theta^0}_{t_k}|^3\Big)\\
&\times e^{C_i(\varphi^i_{n\Delta_n}(\theta^0))^2\Delta_n|X^{\theta^0}_{t_k}|^2}+C\varphi^i_{n\Delta_n}(\theta^0)\varphi^{i+1}_{n\Delta_n}(\theta^0)\Delta_n^{\frac{3}{2}}\\
&\times\sum_{k=0}^{n-1}\Big(1+|X^{\theta^0}_{t_k}|^2+\sqrt{\Delta_n}(\varphi^i_{n\Delta_n}(\theta^0)+\varphi^{i+1}_{n\Delta_n}(\theta^0))|X^{\theta^0}_{t_k}|^3\Big)\\
&\times e^{C_{i+1}((\varphi^i_{n\Delta_n}(\theta^0))^2+(\varphi^{i+1}_{n\Delta_n}(\theta^0))^2)\Delta_n|X^{\theta^0}_{t_k}|^2}+\cdots+C\varphi^i_{n\Delta_n}(\theta^0)\varphi^m_{n\Delta_n}(\theta^0)\Delta_n^{\frac{3}{2}}\\
&\times\sum_{k=0}^{n-1}\Big(1+|X^{\theta^0}_{t_k}|^2+\sqrt{\Delta_n}(\varphi^i_{n\Delta_n}(\theta^0)+\cdots+\varphi^m_{n\Delta_n}(\theta^0))|X^{\theta^0}_{t_k}|^3\Big)\\
&\times e^{C_m((\varphi^i_{n\Delta_n}(\theta^0))^2+\cdots+(\varphi^m_{n\Delta_n}(\theta^0))^2)\Delta_n|X^{\theta^0}_{t_k}|^2}.
\end{aligned}
$$

Thus, using Young's inequality for products with $\frac{1}{p_0}+\frac{1}{q_0}=1$ and $p_0$ close to 1, we get

$$
\begin{aligned}
&\Big|\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}\Big[\zeta_{i,k,n}|\widehat{\mathcal{F}}_{t_k}\Big]\Big|\le C(\varphi^i_{n\Delta_n}(\theta^0))^2\Delta_n^{\frac{3}{2}}\\
&\times\sum_{k=0}^{n-1}\Big\{\frac{1}{p_0}(1+|X^{\theta^0}_{t_k}|^2+\sqrt{\Delta_n}\varphi^i_{n\Delta_n}(\theta^0)|X^{\theta^0}_{t_k}|^3)^{p_0}+\frac{1}{q_0}e^{q_0C_i(\varphi^i_{n\Delta_n}(\theta^0))^2\Delta_n|X^{\theta^0}_{t_k}|^2}\Big\}\\
&+C\varphi^i_{n\Delta_n}(\theta^0)\varphi^{i+1}_{n\Delta_n}(\theta^0)\Delta_n^{\frac{3}{2}}\sum_{k=0}^{n-1}\Big\{\frac{1}{p_0}(1+|X^{\theta^0}_{t_k}|^2+\sqrt{\Delta_n}(\varphi^i_{n\Delta_n}(\theta^0)\\
&+\varphi^{i+1}_{n\Delta_n}(\theta^0))|X^{\theta^0}_{t_k}|^3)^{p_0}+\frac{1}{q_0}e^{q_0C_{i+1}((\varphi^i_{n\Delta_n}(\theta^0))^2+(\varphi^{i+1}_{n\Delta_n}(\theta^0))^2)\Delta_n|X^{\theta^0}_{t_k}|^2}\Big\}\\
&+\cdots+C\varphi^i_{n\Delta_n}(\theta^0)\varphi^m_{n\Delta_n}(\theta^0)\Delta_n^{\frac{3}{2}}\sum_{k=0}^{n-1}\Big\{\frac{1}{p_0}(1+|X^{\theta^0}_{t_k}|^2+\sqrt{\Delta_n}(\varphi^i_{n\Delta_n}(\theta^0)\\
&+\cdots+\varphi^m_{n\Delta_n}(\theta^0))|X^{\theta^0}_{t_k}|^3)^{p_0}+\frac{1}{q_0}e^{q_0C_m((\varphi^i_{n\Delta_n}(\theta^0))^2+\cdots+(\varphi^m_{n\Delta_n}(\theta^0))^2)\Delta_n|X^{\theta^0}_{t_k}|^2}\Big\}.
\end{aligned}
$$

Then, taking expectation in both sides and using Lemma 4.3 and **(A7)**, we obtain

$$
\begin{aligned}
&\widehat{\mathrm{E}}^{\theta^0}\Big[\Big|\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}\Big[\zeta_{i,k,n}|\widehat{\mathcal{F}}_{t_k}\Big]\Big|\Big]\\
&\le C\Delta_n^{\frac{3}{2}}(\varphi^i_{n\Delta_n}(\theta^0))^2\Big(n+\sum_{k=0}^{n-1}\psi_{t_k}^{-2p_0}(\theta^0)+(\sqrt{\Delta_n}\varphi^i_{n\Delta_n}(\theta^0))^{p_0}\sum_{k=0}^{n-1}\psi_{t_k}^{-3p_0}(\theta^0)\Big)
\end{aligned}
$$

$$
\begin{aligned}
&+C\Delta_n^{\frac{3}{2}}\varphi^i_{n\Delta_n}(\theta^0)\varphi^{i+1}_{n\Delta_n}(\theta^0)\Big(n+\sum_{k=0}^{n-1}\psi_{t_k}^{-2p_0}(\theta^0)+(\sqrt{\Delta_n}(\varphi^i_{n\Delta_n}(\theta^0)\\
&+\varphi^{i+1}_{n\Delta_n}(\theta^0))^{p_0}\sum_{k=0}^{n-1}\psi_{t_k}^{-3p_0}(\theta^0)\Big)+\cdots+C\Delta_n^{\frac{3}{2}}\varphi^i_{n\Delta_n}(\theta^0)\varphi^m_{n\Delta_n}(\theta^0)\\
&\times\Big(n+\sum_{k=0}^{n-1}\psi_{t_k}^{-2p_0}(\theta^0)+(\sqrt{\Delta_n}(\varphi^i_{n\Delta_n}(\theta^0)+\cdots+\varphi^m_{n\Delta_n}(\theta^0))^{p_0}\sum_{k=0}^{n-1}\psi_{t_k}^{-3p_0}(\theta^0)\Big),
\end{aligned}
$$

which, by condition **(A6)**, tends to zero. Thus, $\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}[\zeta_{i,k,n}|\widehat{\mathcal{F}}_{t_k}]\xrightarrow{\widehat{\mathrm{P}}^{\theta^0}}0$.

Next, applying Jensen's inequality, (4.2) of Lemma 4.4 to the random variable $V^{\theta_i^{0+}(\ell)}=(R^{\theta_i^{0+}(\ell),k})^2$, (4.4) of Lemma 4.5, and (3.10) with $p\in\{2,4\}$, we obtain

$$
\begin{aligned}
&\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}\big[\zeta_{i,k,n}^2|\widehat{\mathcal{F}}_{t_k}\big]\\
&\leq\sum_{k=0}^{n-1}\frac{(\varphi^i_{n\Delta_n}(\theta^0))^2u_i^2}{\Delta_n^2}\int_0^1\widehat{\mathrm{E}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}\Big[\widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}\Big[(R^{\theta_i^{0+}(\ell),k})^2|Y^{\theta_i^{0+}(\ell)}_{t_{k+1}}=X^{\theta^0}_{t_{k+1}}\Big]\Big]d\ell\\
&\leq\sum_{k=0}^{n-1}\frac{(\varphi^i_{n\Delta_n}(\theta^0))^2u_i^2}{\Delta_n^2}\int_0^1\Big\{\widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}\Big[(R^{\theta_i^{0+}(\ell),k})^2\Big]\\
&\quad+C\sqrt{\Delta_n}(1+|X^{\theta^0}_{t_k}|)\big(\widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}[|R^{\theta_i^{0+}(\ell),k}|^4]\big)^{\frac{1}{2}}\\
&\quad\times\Big(\varphi^i_{n\Delta_n}(\theta^0)\big(1+\sqrt{\Delta_n}\varphi^i_{n\Delta_n}(\theta^0)|X^{\theta^0}_{t_k}|\big)e^{C_i(\varphi^i_{n\Delta_n}(\theta^0))^2\Delta_n|X^{\theta^0}_{t_k}|^2}\\
&\quad+\varphi^{i+1}_{n\Delta_n}(\theta^0)\big(1+\sqrt{\Delta_n}(\varphi^i_{n\Delta_n}(\theta^0)+\varphi^{i+1}_{n\Delta_n}(\theta^0))|X^{\theta^0}_{t_k}|\big)\\
&\quad\times e^{C_{i+1}((\varphi^i_{n\Delta_n}(\theta^0))^2+(\varphi^{i+1}_{n\Delta_n}(\theta^0))^2)\Delta_n|X^{\theta^0}_{t_k}|^2}+\cdots\\
&\quad+\varphi^m_{n\Delta_n}(\theta^0)\big(1+\sqrt{\Delta_n}(\varphi^i_{n\Delta_n}(\theta^0)+\cdots+\varphi^m_{n\Delta_n}(\theta^0))|X^{\theta^0}_{t_k}|\big)\\
&\quad\times e^{C_m((\varphi^i_{n\Delta_n}(\theta^0))^2+\cdots+(\varphi^m_{n\Delta_n}(\theta^0))^2)\Delta_n|X^{\theta^0}_{t_k}|^2}\Big)\Big\}d\ell\\
&\leq C\Delta_n^2(\varphi^i_{n\Delta_n}(\theta^0))^2\sum_{k=0}^{n-1}(1+|X^{\theta^0}_{t_k}|^2)+C(\varphi^i_{n\Delta_n}(\theta^0))^3\Delta_n^{\frac{5}{2}}\\
&\quad\times\sum_{k=0}^{n-1}\Big(1+|X^{\theta^0}_{t_k}|^3+\sqrt{\Delta_n}\varphi^i_{n\Delta_n}(\theta^0)|X^{\theta^0}_{t_k}|^4\Big)e^{C_i(\varphi^i_{n\Delta_n}(\theta^0))^2\Delta_n|X^{\theta^0}_{t_k}|^2}\\
&\quad+C(\varphi^i_{n\Delta_n}(\theta^0))^2\varphi^{i+1}_{n\Delta_n}(\theta^0)\Delta_n^{\frac{5}{2}}\sum_{k=0}^{n-1}\Big(1+|X^{\theta^0}_{t_k}|^3+\sqrt{\Delta_n}(\varphi^i_{n\Delta_n}(\theta^0)\\
&\quad+\varphi^{i+1}_{n\Delta_n}(\theta^0))|X^{\theta^0}_{t_k}|^4\Big)e^{C_{i+1}((\varphi^i_{n\Delta_n}(\theta^0))^2+(\varphi^{i+1}_{n\Delta_n}(\theta^0))^2)\Delta_n|X^{\theta^0}_{t_k}|^2}+\cdots\\
&\quad+C(\varphi^i_{n\Delta_n}(\theta^0))^2\varphi^m_{n\Delta_n}(\theta^0)\Delta_n^{\frac{5}{2}}\sum_{k=0}^{n-1}\Big(1+|X^{\theta^0}_{t_k}|^3+\sqrt{\Delta_n}(\varphi^i_{n\Delta_n}(\theta^0)
\end{aligned}
$$

$$+\cdots+\varphi^{m}_{n\Delta_n}(\theta^0))|X^{\theta^0}_{t_k}|^4\Big)e^{C_m((\varphi^{i}_{n\Delta_n}(\theta^0))^2+\cdots+(\varphi^{m}_{n\Delta_n}(\theta^0))^2)\Delta_n|X^{\theta^0}_{t_k}|^2}.$$

Thus, using Young's inequality with $\frac{1}{p_0}+\frac{1}{q_0}=1$ and $p_0$ close to 1, Lemma 4.3 and **(A7)**,

$$\begin{aligned}&\widehat{\mathrm{E}}^{\theta^0}\Big[\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}\Big[\zeta^2_{i,k,n}|\widehat{\mathcal{F}}_{t_k}\Big]\Big]\leq C\Delta_n^2(\varphi^{i}_{n\Delta_n}(\theta^0))^2\Big(n+\sum_{k=0}^{n-1}\psi^{-2}_{t_k}(\theta^0)\Big)\\
&+C\Delta_n^{\frac{5}{2}}(\varphi^{i}_{n\Delta_n}(\theta^0))^3\Big(n+\sum_{k=0}^{n-1}\psi^{-3p_0}_{t_k}(\theta^0)+(\sqrt{\Delta_n}\varphi^{i}_{n\Delta_n}(\theta^0))^{p_0}\sum_{k=0}^{n-1}\psi^{-4p_0}_{t_k}(\theta^0)\Big)\\
&+C\Delta_n^{\frac{5}{2}}(\varphi^{i}_{n\Delta_n}(\theta^0))^2\varphi^{i+1}_{n\Delta_n}(\theta^0)\Big(n+\sum_{k=0}^{n-1}\psi^{-3p_0}_{t_k}(\theta^0)+(\sqrt{\Delta_n}(\varphi^{i}_{n\Delta_n}(\theta^0)\\
&+\varphi^{i+1}_{n\Delta_n}(\theta^0))^{p_0}\sum_{k=0}^{n-1}\psi^{-4p_0}_{t_k}(\theta^0)\Big)+\cdots+C\Delta_n^{\frac{5}{2}}(\varphi^{i}_{n\Delta_n}(\theta^0))^2\varphi^{m}_{n\Delta_n}(\theta^0)\\
&\times\Big(n+\sum_{k=0}^{n-1}\psi^{-3p_0}_{t_k}(\theta^0)+(\sqrt{\Delta_n}(\varphi^{i}_{n\Delta_n}(\theta^0)+\cdots+\varphi^{m}_{n\Delta_n}(\theta^0))^{p_0}\sum_{k=0}^{n-1}\psi^{-4p_0}_{t_k}(\theta^0)\Big),\end{aligned}$$

which, by condition **(A6)**, tends to zero. Thus, $\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}[\zeta^2_{i,k,n}|\widehat{\mathcal{F}}_{t_k}]\overset{\widehat{\mathrm{P}}^{\theta^0}}{\longrightarrow}0$ for any $i\in\{1,\ldots,m\}$. Thus, by Lemma 4.1 a), the result follows. □

To prove the two following lemmas, we will proceed as in the proof of Lemma 4.9.

**Lemma 4.10.** *Under conditions **(A1)-(A3)** and **(A5)-(A7)**, as $n\to\infty$,*

$$\sum_{k=0}^{n-1}\sum_{i=1}^{m}\frac{\varphi^{i}_{n\Delta_n}(\theta^0)u_i}{\Delta_n}\int_0^1\Big\{Z^{5,\ell}_{i,k,n}-\widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}\Big[R^{\theta_i^{0+}(\ell),k}_5|Y^{\theta_i^{0+}(\ell)}_{t_{k+1}}=X^{\theta^0}_{t_{k+1}}\Big]\Big\}d\ell\overset{\widehat{\mathrm{P}}^{\theta^0}}{\longrightarrow}0.$$

**Lemma 4.11.** *Under conditions **(A1)-(A3)** and **(A5)-(A7)**, as $n\to\infty$,*

$$\sum_{k=0}^{n-1}\sum_{i=1}^{m}\frac{\varphi^{i}_{n\Delta_n}(\theta^0)u_i}{\Delta_n}\int_0^1\Big\{Z^{4,\ell}_{i,k,n}-\widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}\Big[R^{\theta_i^{0+}(\ell),k}_4|Y^{\theta_i^{0+}(\ell)}_{t_{k+1}}=X^{\theta^0}_{t_{k+1}}\Big]\Big\}d\ell\overset{\widehat{\mathrm{P}}^{\theta^0}}{\longrightarrow}0.$$

## 5. Examples

### *5.1. Homogeneous diffusions*

#### *5.1.1. Homogeneous ergodic diffusion processes*

Let $X^\theta=(X^\theta_t)_{t\geq0}$ be the unique strong solution of the following $d$-dimensional SDE

$$dX^\theta_t=b(\theta,X^\theta_t)dt+\sigma(X^\theta_t)dB_t, \tag{5.1}$$

where $X_0^\theta = x_0 \in \mathbb{R}^d$. This is a particular case of the model discussed in [12]. To guarantee the ergodicity of $X^\theta$, we impose the following assumption on the drift coefficient.

**(A4')** There exist constants $c_0, K > 0$ such that for all $(\theta, x) \in \Theta \times \mathbb{R}^d$,

$$b(\theta, x)x \leq -c_0|x|^2 + K.$$

Under **(A1)**-**(A3)** and **(A4')**, the process $X^\theta$ admits an unique invariant probability measure $\pi_\theta(dx)$ and the ergodic theorem holds. That is, for any $\pi_\theta$-integrable function $g$, as $T \to \infty$,

$$\frac{1}{T}\int_0^T g(X_t^\theta)dt \xrightarrow{\widehat{\mathrm{P}}^\theta} \int_{\mathbb{R}^d} g(x)\pi_\theta(dx).$$

Moreover, $\sup_{t\geq 0} \widehat{\mathrm{E}}^\theta[|X_t^\theta|^p] < \infty$ for all $\theta \in \Theta$ and $p \geq 1$. Therefore, **(A4)** is satisfied with $m \times m$ diagonal matrix $\varphi_T(\theta) = \mathrm{diag}(\frac{1}{\sqrt{T}}, \ldots, \frac{1}{\sqrt{T}})$ with $\varphi_T^1(\theta) = \cdots = \varphi_T^m(\theta) = \frac{1}{\sqrt{T}}$ and

$$\Gamma(\theta) = \int_{\mathbb{R}^d} \left(\nabla_\theta b(\theta, x)\right)^* \sigma^{-2}(x) \nabla_\theta b(\theta, x)\pi_\theta(dx).$$

Moreover, **(A5)** is valid with $\psi_t(\theta) = 1$ for $t \geq 0$, and **(A6)** holds since $\psi_t(\theta) = 1$ and $\varphi_{n\Delta_n}^1(\theta) = \cdots = \varphi_{n\Delta_n}^m(\theta) = \frac{1}{\sqrt{n\Delta_n}}$. By [12, Proposition 1.1], **(A7)** is valid. As a consequence of Theorem 2.1, under **(A1)**-**(A3)** and **(A4')**, the LAN property holds at $\theta^0$ with $\varphi_{n\Delta_n}(\theta^0) = \mathrm{diag}(\frac{1}{\sqrt{n\Delta_n}}, \ldots, \frac{1}{\sqrt{n\Delta_n}})$ and asymptotic Fisher information matrix $\Gamma(\theta^0)$.

**Remark 5.1.** Theorem 2.1 can be seen as an extension of [12, Theorem 4.1] when the unknown parameter appears only in the drift coefficient and when equation (1.1) is inhomogeneous and can be ergodic or non-ergodic.

#### *5.1.2. Homogeneous Ornstein-Uhlenbeck process*

Let $X^{a,b} = (X_t^{a,b})_{t\geq 0}$ be the unique strong solution of the one-dimensional SDE

$$dX_t^{a,b} = (b - aX_t^{a,b})dt + \sigma dB_t, \tag{5.2}$$

with given initial condition $X_0^{a,b} = x_0$, $\theta = (a, b) \in \mathbb{R}^2$ are unknown parameters and $\sigma > 0$. By Itô's formula, the solution process is given by

$$X_t^{a,b} = X_0^{a,b} e^{-at} + \frac{b}{a}(1 - e^{-at}) + \sigma \int_0^t e^{-a(t-s)} dB_s. \tag{5.3}$$

For $t > 0$, the transition density $p^{a,b}(t, x_0, y)$ of $X_t^{a,b}$ is given by

$$p^{a,b}(t, x_0, y) = \sqrt{\frac{a}{\pi\sigma^2(1 - e^{-2at})}} \exp\left\{\frac{-(y - \frac{b}{a} - (x_0 - \frac{b}{a})e^{-at})^2}{\frac{\sigma^2}{a}(1 - e^{-2at})}\right\}. \tag{5.4}$$

Based on $(X_t^{a,b})_{t\in[0,T]}$, the observed Fisher information process at $(a,b)$ is given by

$$\frac{1}{\sigma^2}\begin{pmatrix}\int_0^T (X_t^{a,b})^2dt & -\int_0^T X_t^{a,b}dt \\ -\int_0^T X_t^{a,b}dt & T\end{pmatrix}.$$

**Case 1:** $a>0$. The solution $X^{a,b}$ is ergodic with invariant Gaussian distribution $\mathcal{N}(\frac{b}{a},\frac{\sigma^2}{2a})$ (see [21, Example 1.26]). That is,

$$\pi_{a,b}(dx)=f(a,b,x)dx=\sqrt{\frac{a}{\pi\sigma^2}}\exp\Big\{-\frac{(ax-b)^2}{a\sigma^2}\Big\}dx.$$

By [21, Example 1.35], as $T\to\infty$,

$$\frac{1}{T}\int_0^T (X_t^{a,b})^2dt \xrightarrow{\widehat{\mathrm{P}}^{a,b}} \int_{\mathbb{R}} x^2\pi_{a,b}(dx)=\frac{b^2}{a^2}+\frac{\sigma^2}{2a},$$
$$\frac{1}{T}\int_0^T X_t^{a,b}dt \xrightarrow{\widehat{\mathrm{P}}^{a,b}} \int_{\mathbb{R}} x\pi_{a,b}(dx)=\frac{b}{a}.$$

Thus, **(A4)** is satisfied with $\varphi_T(a,b)=\mathrm{diag}(\frac{1}{\sqrt{T}},\frac{1}{\sqrt{T}})$ and

$$\Gamma(a,b)=\frac{1}{a^2\sigma^2}\begin{pmatrix} b^2+\frac{a\sigma^2}{2} & -ab \\ -ab & a^2\end{pmatrix}.$$

Moreover, $\sup_{t\geq 0}\widehat{\mathrm{E}}^{a,b}[|X_t^{a,b}|^p]<\infty$ for all $(a,b)$ and $p\geq 1$. Thus, condition **(A5)** is valid with $\psi_t(a,b)=1$ for $t\geq 0$. Condition **(A6)** holds since $\psi_t(a,b)=1$ and $\varphi^1_{n\Delta_n}(a,b)=\varphi^2_{n\Delta_n}(a,b)=\frac{1}{\sqrt{n\Delta_n}}$. Using (5.4) and the fact that $1-e^{-2at_k}\leq 2at_k$, $\frac{t_k}{n\Delta_n}\leq 1$ and $e^{-at_k}\leq 1$, we get

$$\widehat{\mathrm{E}}^{a,b}\Big[e^{C(\varphi^1_{n\Delta_n}(a,b))^2\Delta_n|X_{t_k}^{a,b}|^2}\Big]=\widehat{\mathrm{E}}^{a,b}\Big[e^{C\frac{\Delta_n}{n\Delta_n}|X_{t_k}^{a,b}|^2}\Big]\leq c,$$

for $n$ large enough and some constant $c>0$. Thus, **(A7)** holds. As a consequence of Theorem 2.1, the LAN property holds at $(a_0,b_0)$ with $\varphi_{n\Delta_n}(a_0,b_0)=\mathrm{diag}(\frac{1}{\sqrt{n\Delta_n}},\frac{1}{\sqrt{n\Delta_n}})$ and $\Gamma(a_0,b_0)$.

**Case 2:** $a<0$. From (5.3), $e^{at}X_t^{a,b}-X_0^{a,b}-\frac{b}{a}(e^{at}-1)=\sigma\int_0^t e^{as}dB_s$, $t\geq 0$ is a square integrable martingale. Thus, the martingale convergence theorem implies that as $t\to\infty$,

$$e^{at}X_t^{a,b}\to X_0^{a,b}-\frac{b}{a}+Z^a,\quad \widehat{\mathrm{P}}^{a,b}\text{-a.s.},$$

where $Z^a:=\sigma\int_0^\infty e^{as}dB_s$ has Gaussian law $\mathcal{N}(0,-\frac{\sigma^2}{2a})$. Then, the integral version of the Toeplitz lemma implies that as $t\to\infty$,

$$\frac{\int_0^t X_s^{a,b}ds}{\int_0^t e^{-as}ds}\to X_0^{a,b}-\frac{b}{a}+Z^a,\quad \widehat{\mathrm{P}}^{a,b}\text{-a.s.},$$

$$\frac{\int_0^t (X_s^{a,b})^2 ds}{\int_0^t e^{-2as} ds} \to \left(X_0^{a,b} - \frac{b}{a} + Z^a\right)^2, \quad \widehat{\mathrm{P}}^{a,b}\text{-a.s.}$$

which deduces that as $t \to \infty$,

$$\frac{1}{\sqrt{t}} e^{at} \int_0^t X_s^{a,b} ds \to 0, \quad \widehat{\mathrm{P}}^{a,b}\text{-a.s.},$$
$$e^{2at} \int_0^t (X_s^{a,b})^2 ds \to -\frac{1}{2a}\left(X_0^{a,b} - \frac{b}{a} + Z^a\right)^2, \quad \widehat{\mathrm{P}}^{a,b}\text{-a.s.}$$

Thus, **(A4)** is satisfied with $\varphi_T(a,b) = \operatorname{diag}(e^{aT}, \frac{1}{\sqrt{T}})$ and

$$\Gamma(a,b) = \begin{pmatrix} -\frac{1}{2a\sigma^2}(x_0 - \frac{b}{a} + Z^a)^2 & 0 \\ 0 & \frac{1}{\sigma^2} \end{pmatrix}.$$

Moreover, $\sup_{t\geq 0} \widehat{\mathrm{E}}^{a,b}[|e^{at} X_t^{a,b}|^p] < \infty$ for $p \geq 1$. Thus, **(A5)** is satisfied with $\psi_t(a,b) = e^{at}$. Assume that $\frac{\sqrt{\Delta_n}}{n\Delta_n} e^{-6p_0 an\Delta_n} \to 0$ where $p_0 > 1$ and $p_0$ close to 1, then **(A6)** is valid. Using (5.4) and $e^{2a(n\Delta_n - t_k)} \leq 1$ and $e^{at_k} \leq 1$, we get for $n$ large enough and some constant $c > 0$,

$$\widehat{\mathrm{E}}^{a,b}\left[e^{C(\varphi^1_{n\Delta_n}(a,b))^2 \Delta_n |X_{t_k}^{a,b}|^2}\right] = \widehat{\mathrm{E}}^{a,b}\left[e^{Ce^{2an\Delta_n}\Delta_n |X_{t_k}^{a,b}|^2}\right] \leq c.$$

Thus, **(A7)** holds.

As a consequence of Theorem 2.1, under condition $\frac{\sqrt{\Delta_n}}{n\Delta_n} e^{-6p_0 a_0 n\Delta_n} \to 0$, the LAMN property holds at $(a_0, b_0)$ with $\varphi_{n\Delta_n}(a_0,b_0) = \operatorname{diag}(e^{a_0 n\Delta_n}, \frac{1}{\sqrt{n\Delta_n}})$ and $\Gamma(a_0, b_0)$.

When $\Delta_n = -\frac{\log n}{6\alpha p_0 a_0 n}$ with $\alpha \geq 2$, we have $\Delta_n \to 0$, $n\Delta_n \to \infty$ and $\frac{\sqrt{\Delta_n}}{n\Delta_n} e^{-6p_0 a_0 n\Delta_n} \to 0$.

### *5.1.3. Two-dimensional Gaussian diffusion process*

Let $X^\theta = (X_1^\theta, X_2^\theta)^* = (X_t^\theta)_{t\geq 0}$ be the unique strong solution of the following 2-dimensional SDE (see [26, Section 4.1])

$$dX_t^\theta = A(\theta) X_t^\theta dt + dB_t, \tag{5.5}$$

with $X_0^\theta = 0$, where

$$A(\theta) = \begin{pmatrix} -\theta_1 & -\theta_2 \\ \theta_2 & -\theta_1 \end{pmatrix},$$

$B = (B_t)_{t\geq 0}$ is a 2-dimensional Brownian motion, $\theta = (\theta_1, \theta_2)$ and $\Theta = \mathbb{R}^2$. By Itô's formula,

$$X_t^\theta = e^{A(\theta)t} \int_0^t e^{-A(\theta)s} dB_s, \qquad \text{where} \quad e^{A(\theta)t} = e^{-\theta_1 t} \begin{pmatrix} \cos\theta_2 t & -\sin\theta_2 t \\ \sin\theta_2 t & \cos\theta_2 t \end{pmatrix}.$$

For $t > 0$, the density $p^\theta(t, y)$ of the marginal $X_t^\theta$ is given by

$$p^\theta(t, y) = \frac{1}{\pi\sqrt{2}}\sqrt{\frac{\theta_1}{1 - e^{-2\theta_1 t}}}\exp\left\{\frac{-\theta_1|y|^2}{1 - e^{-2\theta_1 t}}\right\}. \tag{5.6}$$

The observed Fisher information process at $\theta = (\theta_1, \theta_2)$ based on $(X_s^\theta)_{s\in[0,t]}$ is $\int_0^t |X_s^\theta|^2 ds\, I_2$.

**Case 1:** $\theta_1 > 0$. By ergodicity, as $t \to \infty$,

$$\frac{1}{t}\int_0^t |X_s^\theta|^2 ds \longrightarrow \lim_{s\to\infty} \widehat{\mathrm{E}}^\theta[|X_s^\theta|^2] = \int_{\mathbb{R}^2} |x|^2 \pi_\theta(dx) = \frac{1}{\theta_1}, \quad \widehat{\mathrm{P}}^\theta\text{-a.s.}$$

Thus, **(A4)** is satisfied with $\varphi_t^1(\theta) = \varphi_t^2(\theta) = \frac{1}{\sqrt{t}}$ and $\Gamma(\theta) = \frac{1}{\theta_1} I_2$. Furthermore, for any $p \geq 1$, $\sup_{t\geq 0} \widehat{\mathrm{E}}^\theta[|X_t^\theta|^p] < \infty$. Thus, **(A5)** holds with $\psi_t(\theta) = 1$ for $t \geq 0$. Condition **(A6)** holds since $\psi_t(\theta) = 1$ and $\varphi_{n\Delta_n}^1(\theta) = \varphi_{n\Delta_n}^2(\theta) = \frac{1}{\sqrt{n\Delta_n}}$. Using (5.6) and the fact that $1 - e^{-2\theta_1 t_k} \leq 2\theta_1 t_k$, $\frac{t_k}{n\Delta_n} \leq 1$ and $e^{-\theta_1 t_k} \leq 1$, we get that for $n$ large enough

$$\widehat{\mathrm{E}}^\theta\left[e^{C(\varphi_{n\Delta_n}^1(\theta))^2\Delta_n|X_{t_k}^\theta|^2}\right] = \widehat{\mathrm{E}}^\theta\left[e^{C(\varphi_{n\Delta_n}^2(\theta))^2\Delta_n|X_{t_k}^\theta|^2}\right] = \widehat{\mathrm{E}}^\theta\left[e^{C\frac{\Delta_n}{n\Delta_n}|X_{t_k}^\theta|^2}\right] \leq c,$$

for some constant $c > 0$. Thus, **(A7)** holds. As a consequence of Theorem 2.1, the LAN property holds at $\theta^0 = (\theta_1^0, \theta_2^0)$ with $\varphi_{n\Delta_n}(\theta^0) = \mathrm{diag}(\frac{1}{\sqrt{n\Delta_n}}, \frac{1}{\sqrt{n\Delta_n}})$ and $\Gamma(\theta^0) = \frac{1}{\theta_1^0} I_2$.

**Case 2:** $\theta_1 < 0$. From [26, Section 4.1], as $t \to \infty$,

$$\begin{aligned}
e^{-A(\theta)t}X_t^\theta &\longrightarrow \sqrt{-\frac{1}{2\theta_1}}V(\theta), \quad \widehat{\mathrm{P}}^\theta\text{-a.s.},\\
-\theta_1 e^{2\theta_1 t}|X_t^\theta|^2 &\longrightarrow \frac{1}{2}|V(\theta)|^2, \quad \widehat{\mathrm{P}}^\theta\text{-a.s.},\\
2\theta_1^2 e^{2\theta_1 t}\int_0^t |X_s^\theta|^2 ds &\longrightarrow \frac{1}{2}|V(\theta)|^2, \quad \widehat{\mathrm{P}}^\theta\text{-a.s.},
\end{aligned}$$

where $V(\theta) \sim \mathcal{N}(0, I_2)$. Thus, **(A4)** is satisfied with $\varphi_t^1(\theta) = \varphi_t^2(\theta) = -\sqrt{2}\theta_1 e^{\theta_1 t}$ and $\Gamma(\theta) = \frac{1}{2}|V(\theta)|^2 I_2$. Moreover, $\sup_{t\geq 0} \widehat{\mathrm{E}}^\theta[|e^{\theta_1 t}X_t^\theta|^p] < \infty$ for any $p \geq 1$. Thus, **(A5)** holds with $\psi_t(\theta) = e^{\theta_1 t}$. Assume that $\sqrt{\Delta_n}e^{-(6p_0-2)\theta_1 n\Delta_n} \to 0$ where $p_0 > 1$ and $p_0$ close to 1, then **(A6)** is valid. Using (5.6), we get that for $n$ large enough

$$\begin{aligned}
&\widehat{\mathrm{E}}^\theta\left[e^{C(\varphi_{n\Delta_n}^1(\theta))^2\Delta_n|X_{t_k}^\theta|^2}\right] = \widehat{\mathrm{E}}^\theta\left[e^{C(\varphi_{n\Delta_n}^2(\theta))^2\Delta_n|X_{t_k}^\theta|^2}\right]\\
&= \widehat{\mathrm{E}}^\theta\left[e^{2C\theta_1^2 e^{2\theta_1 n\Delta_n}\Delta_n|X_{t_k}^\theta|^2}\right] \leq c,
\end{aligned}$$

for some constant $c > 0$. Thus, **(A7)** holds. As a consequence of Theorem 2.1, under condition $\sqrt{\Delta_n}e^{-(6p_0-2)\theta_1^0 n\Delta_n} \to 0$, the LAMN property holds at $\theta^0 =$

$(\theta_1^0, \theta_2^0)$ with $\varphi_{n\Delta_n}(\theta^0) = \operatorname{diag}(-\sqrt{2}\theta_1^0 e^{\theta_1^0 n\Delta_n}, -\sqrt{2}\theta_1^0 e^{\theta_1^0 n\Delta_n})$ and $\Gamma(\theta^0) = \frac{1}{2}|V(\theta^0)|^2 I_2$.

Notice that when choosing $\Delta_n = -\frac{\log n}{(6p_0-2)\theta_1^0 \alpha n}$ for some $\alpha > 2$, we have $\Delta_n \to 0$, $n\Delta_n \to \infty$ and $\sqrt{\Delta_n} e^{-(6p_0-2)\theta_1^0 n\Delta_n} \to 0$ as $n \to \infty$ provided that $n\Delta_n^{\frac{\alpha}{2}} \to 0$ since $\theta_1^0 < 0$.

### *5.1.4. Null-recurrent diffusion process*

Let $X^\theta = (X_t^\theta)_{t\geq 0}$ be the unique strong solution of the one-dimensional SDE (see [21, Section 3.5.1])

$$dX_t^\theta = -\theta \frac{X_t^\theta}{1+(X_t^\theta)^2} dt + \sigma dB_t, \tag{5.7}$$

where $X_0^\theta = x_0$ and $\sigma > 0$. Notice that $X_t^\theta$ does not follow the Gaussian law. Observe that $b(\theta, x)$ and $\partial_\theta b(\theta, x)$ are bounded. In this case, it can be checked that the sum of type $\sum_{k=0}^{n-1} |X_{t_k}^{\theta^0}|^p$ for $p \geq 1$ will not appear in the estimates of all the negligible terms and thus we do not need to use Lemma 4.3 to treat $\sum_{k=0}^{n-1} \widehat{\mathrm{E}}^{\theta^0}[|X_{t_k}^{\theta^0}|^p]$. As a result, **(A5)** is not required. Moreover, thanks to the fact that $\partial_\theta b(\theta, x)$ is bounded, **(A7)** is not required (see Remark 4.7). Condition **(A6)** now writes as $n\Delta_n^{\frac{3}{2}} (\varphi_{n\Delta_n}(\theta))^2 \to 0$ as $n \to \infty$. The observed Fisher information process at $\theta$ based on $(X_t^\theta)_{t\in[0,T]}$ is $\int_0^T \frac{(X_t^\theta)^2}{\sigma^2(1+(X_t^\theta)^2)^2} dt$.

**Case 1:** $\theta > \frac{\sigma^2}{2}$. The process $X^\theta$ is ergodic with invariant density

$$f(\theta, x) = \frac{1}{G(\theta)(1+x^2)^{\theta/\sigma^2}} \quad \text{with} \quad G(\theta) = \int_{-\infty}^{\infty} \frac{dx}{(1+x^2)^{\theta/\sigma^2}}.$$

By ergodicity, as $T \to \infty$,

$$\frac{1}{T}\int_0^T \frac{(X_t^\theta)^2}{\sigma^2(1+(X_t^\theta)^2)^2} dt \xrightarrow{\widehat{\mathrm{P}}^\theta} \Gamma(\theta) := \frac{1}{\sigma^2 G(\theta)} \int_{-\infty}^{\infty} \frac{x^2}{(1+x^2)^{2+\theta/\sigma^2}} dx.$$

Thus, **(A4)** holds with $\varphi_T(\theta) = \frac{1}{\sqrt{T}}$ and $\Gamma(\theta)$, and **(A6)** is valid. As a consequence of Theorem 2.1, the LAN property holds at $\theta^0$ with $\varphi_{n\Delta_n}(\theta^0) = \frac{1}{\sqrt{n\Delta_n}}$ and $\Gamma(\theta^0)$.

**Case 2:** $-\frac{\sigma^2}{2} < \theta < \frac{\sigma^2}{2}$. We set $\gamma(\theta) := \frac{1}{2} + \frac{\theta}{\sigma^2}$ and

$$K_*(B, \gamma(\theta)) = \frac{\Gamma(1+\gamma(\theta))}{2(\gamma(\theta)^2 B)^{\gamma(\theta)}\Gamma(1-\gamma(\theta))}, \qquad B = \frac{2}{\sigma^2}\Big(1 + \frac{2\theta}{\sigma^2}\Big)^{-\frac{4\theta}{\sigma^2+2\theta}},$$

where $\Gamma(\cdot)$ is the Gamma function. Let $\eta$ be a random variable with stable distribution function having the Laplace transform $\mathrm{E}[e^{-p\eta}] = e^{-p^{\gamma(\theta)}}$.

With $\varphi_T(\theta) = T^{-\frac{\gamma(\theta)}{2}}$, it follows from page 298 of [21], we have that as $T \to \infty$,

$$\frac{1}{T^{\gamma(\theta)}} \int_0^T \frac{(X_t^\theta)^2}{\sigma^2(1+(X_t^\theta)^2)^2} dt \xrightarrow{\mathcal{L}(\widehat{\mathrm{P}}^\theta)} \Gamma(\theta),$$

where

$$\Gamma(\theta) := K_*(B, \gamma(\theta)) \frac{2}{\sigma^2} \int_{-\infty}^{\infty} \frac{x^2}{(1+x^2)^{2+\theta/\sigma^2}} dx \; \eta^{-\gamma(\theta)}.$$

Thus, **(A4)** is not valid, and the central limit theorem for continuous local martingales cannot be applied in this case. Instead, from (3.113) of [21] or Proposition 1 of [15], we have that as $T \to \infty$,

$$\begin{aligned}
&\left( \frac{1}{T^{\frac{\gamma(\theta)}{2}}} \nabla_\theta \ell_T(\theta), \frac{1}{T^{\gamma(\theta)}} \langle \nabla_\theta \ell(\theta) \rangle_T \right) \\
&= \left( \frac{-1}{T^{\frac{\gamma(\theta)}{2}}} \int_0^T \frac{X_t^\theta}{\sigma(1+(X_t^\theta)^2)} dB_t, \frac{1}{T^{\gamma(\theta)}} \int_0^T \frac{(X_t^\theta)^2}{\sigma^2(1+(X_t^\theta)^2)^2} dt \right) \\
&\overset{\mathcal{L}(\widehat{\mathrm{P}}^\theta)}{\longrightarrow} \left( \sqrt{\Gamma(\theta)} \mathcal{N}(0,1), \Gamma(\theta) \right),
\end{aligned}$$

where $\mathcal{N}(0,1)$ is independent of $\Gamma(\theta)$, which shows the convergence (3.13) that does not require **(A4)**. Condition **(A6)** is valid provided that $n\Delta_n^{\frac{2(\sigma^2-\theta)}{\sigma^2-2\theta}} \to 0$ as $n \to \infty$. Consequently, from the proof of Theorem 2.1, under condition $n\Delta_n^{\frac{2(\sigma^2-\theta^0)}{\sigma^2-2\theta^0}} \to 0$, the LAMN property holds at $\theta^0$ with $\varphi_{n\Delta_n}(\theta^0) = (n\Delta_n)^{-\frac{\gamma(\theta^0)}{2}}$ and $\Gamma(\theta^0)$.

### *5.1.5. A generalized exponential growth process*

Let $X^\theta = (X_t^\theta)_{t\geq 0}$ be the unique strong solution of the one-dimensional SDE

$$dX_t^\theta = \theta a(X_t^\theta)dt + dB_t, \tag{5.8}$$

with given initial condition $X_0^\theta = x_0$. The unknown parameter $\theta$ is positive. For some constant $c > 0$, the known trend coefficient admits the representation

$$a(x) = cx + r(x), \quad x \in \mathbb{R},$$

such that the function $r$ satisfies the following Lipschitz and growth conditions with appropriate constants $K \geq 0$, $L \geq 0$ and $\gamma \in [0,1)$. That is, for all $(x,y) \in \mathbb{R}^2$,

$$\begin{aligned}
|r(x) - r(y)| &\leq L\,|x-y|\,, \\
|r(x)| &\leq K\,(1+|x|^\gamma)\,.
\end{aligned}$$

See Dietz and Kutoyants [8]. Suppose further that $r$ is of class $C^2$ and its first and second derivatives $r'$ and $r''$ are bounded. When $r(x) = 0$ for all $x \in \mathbb{R}$, $X^\theta$ is an Ornstein-Uhlenbeck process with exponential rate in infinity. When taking a large value, $X^\theta$ behaves like an Ornstein-Uhlenbeck process. Notice that $X_t^\theta$ does not follow the Gaussian law generally. The observed Fisher information process at $\theta$ based on the continuous observation $(X_t^\theta)_{t\in[0,T]}$ is given by $\int_0^T a^2(X_t^\theta)dt$.

By [8, Lemma 2.1 and Corollary 2.2], as $t \to \infty$ and $T \to \infty$,

$$e^{-\theta ct} X_t^\theta \longrightarrow x_0 + \xi_\infty^\theta + \rho_\infty^\theta, \quad \widehat{\mathrm{P}}^\theta\text{-a.s.},$$
$$e^{-2\theta cT} \int_0^T a^2(X_t^\theta) dt \longrightarrow \frac{c}{2\theta}(x_0 + \xi_\infty^\theta + \rho_\infty^\theta)^2, \quad \widehat{\mathrm{P}}^\theta\text{-a.s.},$$

where $\xi_\infty^\theta = \int_0^\infty e^{-\theta cs} dB_s$ and $\rho_\infty^\theta = \int_0^\infty e^{-\theta cs}\theta r(X_s^\theta) ds$. Thus, condition **(A4)** satisfies with $\varphi_T(\theta) = e^{-\theta cT}$ and $\Gamma(\theta) = \frac{c}{2\theta}(x_0 + \xi_\infty^\theta + \rho_\infty^\theta)^2$.

Moreover, $\sup_{t\geq 0} \widehat{\mathrm{E}}^\theta[|e^{-\theta ct} X_t^\theta|^p] < \infty$ for $p \geq 1$, see Lemma 2.1 of [8]. Hence, condition **(A5)** holds with $\psi_t(\theta) = e^{-\theta ct}$. Assume that $\sqrt{\Delta_n} e^{(6p_0-2)\theta cn\Delta_n} \to 0$ where $p_0 > 1$ and $p_0$ close to 1, then condition **(A6)** is valid. Now, it remains to check condition **(A7)**. For this, using the Maclaurin series of the exponential function, we write

$$\begin{aligned}\widehat{\mathrm{E}}^\theta \left[ e^{C(\varphi_{n\Delta_n}(\theta))^2 \Delta_n |X_{t_k}^\theta|^2} \right] &= 1 + \sum_{i=1}^\infty \frac{(C\Delta_n)^i}{i!} \widehat{\mathrm{E}}^\theta \left[ \left( e^{-\theta cn\Delta_n} |X_{t_k}^\theta| \right)^{2i} \right] \\ &\leq 1 + \sum_{i=1}^\infty \frac{(C\Delta_n)^i}{i!} \widehat{\mathrm{E}}^\theta \left[ \left| e^{-\theta ct_k} X_{t_k}^\theta \right|^{2i} \right]. \end{aligned} \tag{5.9}$$

From Lemma 2.1 of [8], we have the following decomposition

$$e^{-\theta ct} X_t^\theta =: \eta_t^\theta = x_0 + \xi_t^\theta + \rho_t^\theta,$$

for any $t \geq 0$, where

$$\xi_t^\theta = \int_0^t e^{-\theta cs} dB_s, \qquad \rho_t^\theta = \int_0^t e^{-\theta cs}\theta r(X_s^\theta) ds.$$

Then, for any $t > 0$ and even number $p \geq 2$ and $p \in \mathbb{N}$, we get that

$$\widehat{\mathrm{E}}^\theta \left[ \left| e^{-\theta ct} X_t^\theta \right| \right] \leq 3^{p-1} \left( |x_0|^p + \widehat{\mathrm{E}}^\theta \left[ |\xi_t^\theta|^p \right] + \widehat{\mathrm{E}}^\theta \left[ |\rho_t^\theta|^p \right] \right). \tag{5.10}$$

First, since $\xi_t^\theta$ follows the Gaussian distribution, we have

$$\begin{aligned}\widehat{\mathrm{E}}^\theta \left[ |\xi_t^\theta|^p \right] &= (p-1)!! \left( \widehat{\mathrm{E}}^\theta \left[ |\xi_t^\theta|^2 \right] \right)^{\frac{p}{2}} \\ &= (p-1)!! \left( \frac{1}{2\theta c}(1 - e^{-2\theta ct}) \right)^{\frac{p}{2}} \leq (p-1)!! \left( \frac{1}{2\theta c} \right)^{\frac{p}{2}}, \end{aligned} \tag{5.11}$$

where $(p-1)!!$ denotes the double factorial of $p-1$.

Next, using (2.15) and (2.16) of [8], we have that $\widehat{\mathrm{P}}$-a.s.

$$|\rho_t^\theta| \leq \frac{\theta K}{\mu} \left( 2 + \eta_*^\theta \right) \leq \frac{\theta K}{\mu} \left( 2 + \left( |x_0| + \xi_*^\theta + \frac{2\theta K}{\mu} \right) e^{\frac{\theta K}{\mu}} \right),$$

where $\mu = \theta c(1-\gamma)$, $\eta_*^\theta = \sup_{t\geq 0} |\eta_t^\theta|$ and $\xi_*^\theta = \sup_{t\geq 0} |\xi_t^\theta|$.

Using Doob's maximal inequalities, the fact that $1 < \frac{p}{p-1} \leq 2$ and (5.11), we have that for any $t \geq 0$,

$$\widehat{\mathrm{E}}^\theta \left[ \left( \sup_{0 \leq s \leq t} |\xi_s^\theta| \right)^p \right] \leq \left( \frac{p}{p-1} \right)^p \widehat{\mathrm{E}}^\theta \left[ |\xi_t^\theta|^p \right] \leq 2^p (p-1)!! \Big( \frac{1}{2\theta c} \Big)^{\frac{p}{2}}.$$

This implies that

$$\widehat{\mathrm{E}}^\theta \left[ |\xi_*^\theta|^p \right] = \widehat{\mathrm{E}}^\theta \left[ \left( \sup_{s \geq 0} |\xi_s^\theta| \right)^p \right] = \lim_{t \to \infty} \widehat{\mathrm{E}}^\theta \left[ \left( \sup_{0 \leq s \leq t} |\xi_s^\theta| \right)^p \right] \leq 2^p (p-1)!! \Big( \frac{1}{2\theta c} \Big)^{\frac{p}{2}}.$$

Thus, we obtain

$$\begin{aligned} \widehat{\mathrm{E}}^\theta \left[ |\rho_t^\theta|^p \right] &\leq \left( \frac{\theta K}{\mu} \right)^p 2^{p-1} \bigg( \Big( 2 + \big( |x_0| + \frac{2\theta K}{\mu} \big) e^{\frac{\theta K}{\mu}} \Big)^p + e^{p \frac{\theta K}{\mu}} \widehat{\mathrm{E}}^\theta \left[ |\xi_*^\theta|^p \right] \bigg) \\ &\leq \left( \frac{\theta K}{\mu} \right)^p 2^{p-1} \bigg( \Big( 2 + \big( |x_0| + \frac{2\theta K}{\mu} \big) e^{\frac{\theta K}{\mu}} \Big)^p + e^{p \frac{\theta K}{\mu}} 2^p (p-1)!! \Big( \frac{1}{2\theta c} \Big)^{\frac{p}{2}} \bigg). \end{aligned} \tag{5.12}$$

From (5.10), (5.11) and (5.12), we obtain that for any $t > 0$ and even number $p \geq 2$ and $p \in \mathbb{N}$,

$$\begin{aligned} \widehat{\mathrm{E}}^\theta \left[ \left| e^{-\theta c t} X_t^\theta \right|^p \right] &\leq 3^{p-1} \bigg\{ |x_0|^p + (p-1)!! \Big( \frac{1}{2\theta c} \Big)^{\frac{p}{2}} + \left( \frac{\theta K}{\mu} \right)^p 2^{p-1} \\ &\times \bigg( \Big( 2 + \big( |x_0| + \frac{2\theta K}{\mu} \big) e^{\frac{\theta K}{\mu}} \Big)^p + e^{p \frac{\theta K}{\mu}} 2^p (p-1)!! \Big( \frac{1}{2\theta c} \Big)^{\frac{p}{2}} \bigg) \bigg\}. \end{aligned} \tag{5.13}$$

Inserting (5.13) into (5.9) and using again Maclaurin series of the exponential function, we get that

$$\begin{aligned} &\widehat{\mathrm{E}}^\theta \left[ e^{C (\varphi_{n\Delta_n}^i(\theta))^2 \Delta_n |X_{t_k}^\theta|^2} \right] \leq 1 + \frac{1}{3} \sum_{i=1}^\infty \frac{(C\Delta_n)^i}{i!} 9^i \bigg\{ |x_0|^{2i} + (2i-1)!! \Big( \frac{1}{2\theta c} \Big)^i \\ &+ \frac{1}{2} \left( \frac{\theta K}{\mu} \right)^{2i} 4^i \bigg( \Big( 2 + \big( |x_0| + \frac{2\theta K}{\mu} \big) e^{\frac{\theta K}{\mu}} \Big)^{2i} + e^{2i \frac{\theta K}{\mu}} 4^i (2i-1)!! \Big( \frac{1}{2\theta c} \Big)^i \bigg) \bigg\} \\ &\leq 1 + S_1 + S_2 \\ &+ \frac{1}{3} \left( \exp\left\{ 9C |x_0|^2 \Delta_n \right\} + \frac{1}{2} \exp\left\{ 36 C \big( \frac{\theta K}{\mu} \big)^2 \Big( 2 + \big( |x_0| + \frac{2\theta K}{\mu} \big) e^{\frac{\theta K}{\mu}} \Big)^2 \Delta_n \right\} \right), \end{aligned}$$

where

$$\begin{aligned} S_1 &= \frac{1}{3} \sum_{i=1}^\infty \frac{1}{i!} \left( \frac{9C\Delta_n}{2\theta c} \right)^i (2i-1)!!, \\ S_2 &= \frac{1}{6} \sum_{i=1}^\infty \frac{1}{i!} \left( \frac{144 C \Delta_n}{2\theta c} \left( \frac{\theta K}{\mu} \right)^2 e^{2 \frac{\theta K}{\mu}} \right)^i (2i-1)!!. \end{aligned}$$

Using Stirling's approximation $n! \sim \sqrt{2\pi n}(\frac{n}{e})^n$ for $n$ large enough, there exists a natural number $i_0$ such that for all $i > i_0$,

$$(2i-1)!! = \frac{(2i)!}{2^i i!} \sim \frac{\sqrt{2\pi 2i}(\frac{2i}{e})^{2i}}{\sqrt{2\pi i}(\frac{i}{e})^i} = \sqrt{2}\Big(\frac{2i}{e}\Big)^i.$$

This implies that

$$\begin{aligned} S_1 &\approx \frac{1}{3}\sum_{i=1}^{i_0}\frac{1}{i!}\left(\frac{9C\Delta_n}{2\theta c}\right)^i(2i-1)!! + \frac{1}{3}\sum_{i=i_0+1}^{\infty}\frac{1}{\sqrt{2\pi i}(\frac{i}{e})^i}\left(\frac{9C\Delta_n}{2\theta c}\right)^i\sqrt{2}\Big(\frac{2i}{e}\Big)^i \\ &= \frac{1}{3}\sum_{i=1}^{i_0}\frac{1}{i!}\left(\frac{9C\Delta_n}{2\theta c}\right)^i(2i-1)!! + \frac{1}{3\sqrt{\pi}}\sum_{i=i_0+1}^{\infty}\frac{1}{\sqrt{i}}\left(\frac{9C\Delta_n}{\theta c}\right)^i, \end{aligned}$$

which, by the D'Alembert criterion, converges when $n$ is large enough. Similarly, the series $S_2$ converges for $n$ large enough by using the same arguments. Thus, we have shown that there exists a constant $C_0 > 0$ such that for $n$ large enough

$$\widehat{\mathrm{E}}^\theta\left[e^{C(\varphi^i_{n\Delta_n}(\theta))^2\Delta_n|X^\theta_{t_k}|^2}\right] \le C_0.$$

Hence, **(A7)** is valid.

As a consequence of Theorem 2.1, under condition $\sqrt{\Delta_n}e^{(6p_0-2)\theta^0 cn\Delta_n} \to 0$, the LAMN property holds for the likelihood at $\theta^0$ with rate of convergence $\varphi_{n\Delta_n}(\theta^0) = e^{-\theta^0 cn\Delta_n}$ and asymptotic random Fisher information

$$\Gamma(\theta^0) = \frac{c}{2\theta^0}(x_0 + \xi^{\theta^0}_\infty + \rho^{\theta^0}_\infty)^2.$$

Notice that when choosing $\Delta_n = \frac{\log n}{(6p_0-2)\theta^0 c\alpha n}$ for some $\alpha > 2$, we have $\Delta_n \to 0$, $n\Delta_n \to \infty$ and $\sqrt{\Delta_n}e^{(6p_0-2)\theta^0 cn\Delta_n} \to 0$ as $n \to \infty$ provided that $n\Delta_n^{\frac{\alpha}{2}} \to 0$ since $\theta^0 > 0$.

## 5.2. Inhomogeneous diffusions

### 5.2.1. Inhomogeneous Ornstein-Uhlenbeck process

Let $X^\theta = (X^\theta_t)_{t\ge 0}$ be the unique strong solution of the one-dimensional SDE

$$dX^\theta_t = -\theta A(t)X^\theta_t dt + dB_t, \tag{5.14}$$

where $X^\theta_0 = 0$, $A : \mathbb{R}_+ \to \mathbb{R}$ is measurable with $\int_0^t A^2(s)ds < \infty$ for every $t$ (see [26, Section 4.2]). By Itô's formula, $X^\theta_t = f(\theta,t)\int_0^t f(\theta,s)^{-1}dB_s$ where $f(\theta,t) = \exp\{-\theta\int_0^t A(s)ds\}$. The observed Fisher information process at $\theta$ based on $(X^\theta_s)_{s\in[0,t]}$ is given by $\int_0^t A^2(s)(X^\theta_s)^2 ds$. The expected Fisher information at $\theta$ and $t$ based on $(X^\theta_s)_{s\in[0,t]}$ is given by

$$I_{X^\theta}(t) = \int_0^t A^2(s)f(\theta,s)^2\int_0^s f(\theta,u)^{-2}duds.$$

**Case 1:** Consider the set of explosive parameters

$$\Theta_0 := \Big\{\theta \in \mathbb{R} : -\theta \int_0^t A(s)ds \to \infty \ \text{ as } \ t \to \infty \ \text{ and } \ \int_0^\infty f(\theta, s)^{-2} ds < \infty\Big\}.$$

For any $\theta \in \Theta_0$, we have $\sup_{t\geq 0} f(\theta,t)^{-2}\widehat{\mathrm{E}}^\theta[(X_t^\theta)^2] = \int_0^\infty f(\theta,s)^{-2}ds < \infty$ and as $t \to \infty$,

$$f(\theta,t)^{-1}X_t^\theta \longrightarrow \big(\int_0^\infty f(\theta,s)^{-2}ds\big)^{\frac{1}{2}}\xi, \quad \widehat{\mathrm{P}}^\theta\text{-a.s.},$$
$$\varphi_t(\theta)^2\int_0^t A^2(s)(X_s^\theta)^2 ds \longrightarrow \xi^2, \quad \widehat{\mathrm{P}}^\theta\text{-a.s.},$$

where $\xi \sim \mathcal{N}(0,1)$ and

$$\varphi_t(\theta) = \Big(\int_0^\infty f(\theta,s)^{-2}ds \int_0^t A^2(s)f(\theta,s)^2 ds\Big)^{-\frac{1}{2}}.$$

Thus, **(A4)** is satisfied with $\varphi_t(\theta)$ and $\Gamma(\theta) = \xi^2$.

Moreover, $\sup_{t\geq 0}\widehat{\mathrm{E}}^\theta[|f(\theta,t)^{-1}X_t^\theta|^p] < \infty$ for any $p \geq 1$. Thus, **(A5)** holds with $\psi_t(\theta) = f(\theta,t)^{-1}$. As a consequence of Theorem 2.1, under **(A6)-(A7)**, the LAMN property holds at $\theta^0 \in \Theta_0$ with $\varphi_{n\Delta_n}(\theta^0)$ and $\Gamma(\theta^0) = \xi^2$.

**Case 2:** Consider the set of parameters

$$\Theta_1 := \Big\{\theta \in \mathbb{R} : \lim_{t\to\infty} I_{X^\theta}(t) = \infty,$$
$$\text{and } \lim_{t\to\infty} \frac{1}{\sqrt{I_{X^\theta}(t)}} A(t)f(\theta,t)^2 \int_0^t f(\theta,s)^{-2}ds = 0\Big\},$$

where assume that $A$ is continuous. For any $\theta \in \Theta_1$, as $t \to \infty$,

$$\varphi_t(\theta)^2 \int_0^t A^2(s)(X_s^\theta)^2 ds \longrightarrow 1, \quad \text{in } \ L^2(\widehat{\mathrm{P}}^\theta),$$

where $\varphi_t(\theta) = I_{X^\theta}(t)^{-\frac{1}{2}}$. Thus, **(A4)** is satisfied with $\Gamma(\theta) = 1$. As a consequence of Theorem 2.1, under **(A5)-(A7)**, the LAN property holds at $\theta^0 \in \Theta_1$ with $\varphi_{n\Delta_n}(\theta^0)$ and $\Gamma(\theta^0) = 1$.

When $A(t) = 1$, $X^\theta$ becomes the classical homogeneous Ornstein-Uhlenbeck process which has been addressed in Section 5.1.2.

When $A(t) = \frac{1}{1+t}$, then $\Theta_0 = (-\infty, -\frac{1}{2})$ and $\Theta_1 = (-\frac{1}{2}, \infty)$. For $t > 0$, the density $p^\theta(t,y)$ of the marginal $X_t^\theta$ is given by

$$p^\theta(t,y) = \frac{1}{\sqrt{\frac{2\pi}{2\theta+1}\big(1+t-\frac{1}{(1+t)^{2\theta}}\big)}} \exp\left\{\frac{-y^2}{\frac{2}{2\theta+1}\big(1+t-\frac{1}{(1+t)^{2\theta}}\big)}\right\}. \tag{5.15}$$

For $\theta \in \Theta_0$, we choose $\varphi_t(\theta) = -(2\theta+1)t^{\theta+\frac{1}{2}}$, $\psi_t(\theta) = (1+t)^\theta$, and **(A6)** is valid provided that $n\Delta_n^{1+\frac{1}{2((-6p_0+2)\theta+2)}} \to 0$ where $p_0 > 1$ and $p_0$ close to

1. Using (5.15), it can be checked that **(A7)** holds. For $\theta \in \Theta_1$, we choose $\varphi_t(\theta) = \sqrt{\frac{2\theta+1}{\log(1+t)}}$, **(A5)** is valid with $\psi_t(\theta) = \frac{1}{\sqrt{1+t}}$, **(A6)** is valid provided that $\frac{\sqrt{\Delta_n}(n\Delta_n)^{3p_0+1}}{\log(1+n\Delta_n)} \to 0$ where $p_0 > 1$ and $p_0$ close to 1. Using (5.15), it can be checked that **(A7)** holds.

### 5.2.2. *A special inhomogeneous diffusion process*

Let $X^\theta = (X_t^\theta)_{t\geq 0}$ be the unique strong solution of the one-dimensional SDE, which is a special case of the Hull-White model,

$$dX_t^\theta = \theta b(t) X_t^\theta dt + \sigma(t) dB_t, \tag{5.16}$$

with given initial condition $X_0^\theta = 0$, where $b : \mathbb{R}_+ \to \mathbb{R}$ and $\sigma : \mathbb{R}_+ \to (0, \infty)$ are known Borel-measurable functions. Here, $\theta \in \mathbb{R}$ is an unknown parameter. See Barczy and Pap [3].

By Itô's formula, the SDE (5.16) has a unique strong solution given by

$$X_t^\theta = \int_0^t \sigma(s) \exp\Big\{\theta \int_s^t b(u)du\Big\} dB_s = f(\theta, t) \int_0^t \sigma(s) f(\theta, s)^{-1} dB_s,$$

where $f(\theta, t) = \exp\{\theta \int_0^t b(u)du\}$. The observed Fisher information process at $\theta$ based on $(X_s^\theta)_{s\in[0,t]}$ is given by $\int_0^t \frac{b^2(s)(X_s^\theta)^2}{\sigma^2(s)} ds$. The expected Fisher information at $\theta$ and $t$ is

$$I_{X^\theta}(t) = \int_0^t \frac{b^2(s)}{\sigma^2(s)} \widehat{\mathrm{E}}^\theta \Big[(X_s^\theta)^2\Big] ds = \int_0^t \frac{b^2(s)}{\sigma^2(s)} f(\theta, s)^2 \int_0^s \sigma^2(u) f(\theta, u)^{-2} du ds.$$

**Case 1:** Consider the set of parameters

$$\Theta_0 := \Big\{\theta \in \mathbb{R} : \lim_{t\to\infty} I_{X^\theta}(t) = \infty \quad \text{and} \quad \int_0^\infty \sigma^2(s) f(\theta, s)^{-2} ds < \infty\Big\}.$$

Then, for any $\theta \in \Theta_0$, we have

$$\sup_{t\geq 0} f(\theta, t)^{-2} \widehat{\mathrm{E}}^\theta[(X_t^\theta)^2] = \int_0^\infty \sigma^2(s) f(\theta, s)^{-2} ds < \infty$$

and

$$f(\theta, t)^{-1} X_t^\theta \longrightarrow \Big(\int_0^\infty \sigma^2(s) f(\theta, s)^{-2} ds\Big)^{\frac{1}{2}} \xi, \ \widehat{\mathrm{P}}^\theta\text{-a.s.},$$
$$\frac{1}{I_{X^\theta}(t)} \int_0^t \frac{b^2(s)(X_s^\theta)^2}{\sigma^2(s)} ds \longrightarrow \xi^2, \ \widehat{\mathrm{P}}^\theta\text{-a.s.},$$

as $t \to \infty$, where $\xi \sim \mathcal{N}(0, 1)$ (see the proof of [3, Theorem 7]). Thus, **(A4)** is satisfied with $\varphi_t(\theta) = I_{X^\theta}(t)^{-\frac{1}{2}}$ and $\Gamma(\theta) = \xi^2$.

Moreover, $\sup_{t\geq 0}\widehat{\mathrm{E}}^{\theta}[|f(\theta,t)^{-1}X_t^{\theta}|^p]<\infty$ for any $p\geq 1$. Thus, **(A5)** holds with $\psi_t(\theta)=f(\theta,t)^{-1}$. As a consequence of Theorem 2.1, under **(A6)-(A7)**, the LAMN property holds at $\theta^0$ with $\varphi_{n\Delta_n}(\theta^0)=I_{X^{\theta^0}}(n\Delta_n)^{-\frac{1}{2}}$ and $\Gamma(\theta^0)=\xi^2$.

**Case 2:** Consider the set of parameters

$$\Theta_1:=\Big\{\theta\in\mathbb{R}:\lim_{t\to\infty}I_{X^{\theta}}(t)=\infty,\\ \text{and }\lim_{t\to\infty}\frac{1}{\sqrt{I_{X^{\theta}}(t)}}\frac{b(t)}{\sigma^2(t)}f(\theta,t)^2\int_0^t\sigma^2(s)f(\theta,s)^{-2}ds=0\Big\}.$$

Then, for any $\theta\in\Theta_1$, as $t\to\infty$, (see [3, Theorem 10])

$$\frac{1}{I_{X^{\theta}}(t)}\int_0^t\frac{b^2(s)(X_s^{\theta})^2}{\sigma^2(s)}ds\longrightarrow 1,\quad\text{in }L^2(\widehat{\mathrm{P}}^{\theta}).$$

Thus, **(A4)** is satisfied with $\varphi_t(\theta)=I_{X^{\theta}}(t)^{-\frac{1}{2}}$ and $\Gamma(\theta)=1$. By Theorem 2.1, under **(A5)-(A7)**, the LAN property holds at $\theta^0$ with $\varphi_{n\Delta_n}(\theta^0)=I_{X^{\theta^0}}(n\Delta_n)^{-\frac{1}{2}}$ and $\Gamma(\theta^0)=1$.

When $\sigma(t)=\sigma>0$, $X^{\theta}$ becomes the inhomogeneous Ornstein-Uhlenbeck process which has been considered in Section 5.2.1.

When $b(t)=\frac{1}{1+t}$, and $\sigma(t)=\frac{2+t}{1+t}$ then $\Theta_0=(\frac{1}{2},\infty)$ and $\Theta_1=(-\infty,\frac{1}{2})\setminus\{-\frac{1}{2}\}$. For $t>0$, the density $p^{\theta}(t,y)$ of the marginal $X_t^{\theta}$ is given by

$$\begin{aligned}p^{\theta}(t,y)&=\frac{1}{\sqrt{2\pi v^2(\theta,t)}}\exp\left\{\frac{-y^2}{2v^2(\theta,t)}\right\},\\ v^2(\theta,t)&=\frac{1}{-2\theta+1}(t+1)-\frac{1}{\theta}-\frac{1}{2\theta+1}\frac{1}{t+1}+\frac{8\theta^2-1}{\theta(4\theta^2-1)}(t+1)^{2\theta}.\end{aligned}\tag{5.17}$$

For any $\theta\in\Theta_0$, we choose $\varphi_t(\theta)=\frac{2\theta-1}{t^{\theta-\frac{1}{2}}}$, $\psi_t(\theta)=\frac{1}{(1+t)^{\theta}}$, and **(A6)** is valid provided that $n\Delta_n^{1+\frac{1}{2((6p_0-2)\theta+2)}}\to 0$ where $p_0>1$ and $p_0$ close to 1. Using (5.17), **(A7)** is valid. For any $\theta\in\Theta_1$, we choose $\varphi_t(\theta)=\sqrt{\frac{-2\theta+1}{\log(2+t)}}$, condition **(A5)** is valid with $\psi_t(\theta)=\frac{1}{\sqrt{1+t}}$, and **(A6)** is valid provided that $\frac{\sqrt{\Delta_n}(n\Delta_n)^{3p_0+1}}{\log(2+n\Delta_n)}\to 0$ where $p_0>1$ and $p_0$ close to 1. Using (5.17), **(A7)** is valid. To the best of our knowledge, the results obtained for this concrete inhomogeneous diffusion process are new in the context where both the drift and diffusion coefficients depend on the time variable $t$.

**Remark 5.2.** In models (5.14) and (5.16) we consider a particular case that the initial condition $X_0^{\theta}=0$ for simplicity. The case that $X_0^{\theta}\neq 0$ can be treated by a similar argument.

## 6. Proof of technical results

### 6.1. Proof of Lemma 3.1

*Proof.* For all $t \in [t_k, t_{k+1}]$, using the chain rule of the Malliavin calculus, we have

$$\left(D_t(f(Y^\theta_{t_{k+1}}(t_k,x)))\right)^* = \left(\nabla f(Y^\theta_{t_{k+1}}(t_k,x))\right)^* D_t Y^\theta_{t_{k+1}}(t_k,x),$$

where $f : \mathbb{R}^d \to \mathbb{R}$ is a continuously differentiable function with compact support. Thus, we have $(\nabla f(Y^\theta_{t_{k+1}}(t_k,x)))^* = (D_t(f(Y^\theta_{t_{k+1}}(t_k,x))))^* U^\theta_t(t_k,x)$ with $U^\theta_t(t_k,x) = (D_t Y^\theta_{t_{k+1}}(t_k,x))^{-1}$ since $D_t Y^\theta_{t_{k+1}}(t_k,x)$ is an invertible matrix a.s. Then for any $i \in \{1,\dots,m\}$, using the duality relationship of the Malliavin calculus on $[t_k,t_{k+1}]$, we get

$$\begin{aligned}
\partial_{\theta_i}\widetilde{\mathrm{E}}\left[f(Y^\theta_{t_{k+1}}(t_k,x))\right] &= \widetilde{\mathrm{E}}\left[\left(\nabla f(Y^\theta_{t_{k+1}}(t_k,x))\right)^* \partial_{\theta_i} Y^\theta_{t_{k+1}}(t_k,x)\right]\\
&= \frac{1}{\Delta_n}\widetilde{\mathrm{E}}\left[\int_{t_k}^{t_{k+1}} \left(D_t(f(Y^\theta_{t_{k+1}}(t_k,x)))\right)^* U^\theta_t(t_k,x)\,\partial_{\theta_i} Y^\theta_{t_{k+1}}(t_k,x)dt\right]\\
&= \frac{1}{\Delta_n}\widetilde{\mathrm{E}}\left[f(Y^\theta_{t_{k+1}}(t_k,x))\delta\left(U^\theta(t_k,x)\partial_{\theta_i} Y^\theta_{t_{k+1}}(t_k,x)\right)\right]. \qquad (6.1)
\end{aligned}$$

By (3.8), the family $((\nabla f(Y^\theta_{t_{k+1}}(t_k,x)))^* \partial_{\theta_i} Y^\theta_{t_{k+1}}(t_k,x), \theta \in \Theta)$ is uniformly integrable. This is the reason why we can interchange $\partial_{\theta_i}$ and $\widetilde{\mathrm{E}}$. Next, using the stochastic flow property,

$$\partial_{\theta_i}\widetilde{\mathrm{E}}\left[f(Y^\theta_{t_{k+1}}(t_k,x))\right] = \int_{\mathbb{R}^d} f(y)\partial_{\theta_i} p^\theta(t_k,t_{k+1},x,y)dy, \qquad (6.2)$$

and

$$\begin{aligned}
&\widetilde{\mathrm{E}}\Big[f(Y^\theta_{t_{k+1}}(t_k,x))\delta\big(U^\theta(t_k,x)\partial_{\theta_i} Y^\theta_{t_{k+1}}(t_k,x)\big)\Big]\\
&= \widetilde{\mathrm{E}}\Big[f(Y^\theta_{t_{k+1}})\delta\big(U^\theta(t_k,x)\partial_{\theta_i} Y^\theta_{t_{k+1}}(t_k,x)\big)|Y^\theta_{t_k} = x\Big]\\
&= \int_{\mathbb{R}^d} f(y)\widetilde{\mathrm{E}}\left[\delta\left(U^\theta(t_k,x)\partial_{\theta_i} Y^\theta_{t_{k+1}}(t_k,x)\right)|Y^\theta_{t_k} = x, Y^\theta_{t_{k+1}} = y\right]\\
&\qquad \times p^\theta(t_k,t_{k+1},x,y)dy. \qquad (6.3)
\end{aligned}$$

Thus, the desired result follows from (6.1)-(6.3). □

### 6.2. Proof of Lemma 3.2

*Proof.* First, using (3.4) and Itô's formula, we obtain

$$(\nabla_x Y^\theta_t(t_k,x))^{-1} = \mathrm{I}_d - \sum_{j=1}^d \int_{t_k}^t (\nabla_x Y^\theta_s(t_k,x))^{-1}\nabla_x\sigma_j(s,Y^\theta_s(t_k,x))dW^j_s$$

$$- \int_{t_k}^{t} (\nabla_x Y_s^\theta(t_k, x))^{-1} \big(\nabla_x b(\theta, s, Y_s^\theta(t_k, x)) - \sum_{j=1}^{d} (\nabla_x \sigma_j(s, Y_s^\theta(t_k, x)))^2\big) ds.$$

This combined with (3.5) and Itô's formula implies that

$$(\nabla_x Y_t^\theta(t_k, x))^{-1} \partial_{\theta_i} Y_t^\theta(t_k, x) = \int_{t_k}^{t} (\nabla_x Y_s^\theta(t_k, x))^{-1} \partial_{\theta_i} b(\theta, s, Y_s^\theta(t_k, x)) ds. \tag{6.4}$$

Then, using (3.6), the product rule (1.48) of Nualart [29], the fact that the Skorohod integral and the Itô integral of an adapted process coincide, and (6.4), we obtain

$$\begin{aligned}
&\delta\big(U^\theta(t_k, x)\partial_{\theta_i} Y_{t_{k+1}}^\theta(t_k, x)\big) \\
&= \delta\Big(\sigma^{-1}(\cdot, Y_\cdot^\theta(t_k, x)) \nabla_x Y_\cdot^\theta(t_k, x)(\nabla_x Y_{t_{k+1}}^\theta(t_k, x))^{-1} \partial_{\theta_i} Y_{t_{k+1}}^\theta(t_k, x)\Big) \\
&= (\partial_{\theta_i} Y_{t_{k+1}}^\theta(t_k, x))^* ((\nabla_x Y_{t_{k+1}}^\theta(t_k, x))^{-1})^* \\
&\qquad \cdot \int_{t_k}^{t_{k+1}} (\nabla_x Y_s^\theta(t_k, x))^* \sigma^{-1}(s, Y_s^\theta(t_k, x)) dW_s \\
&\qquad - \int_{t_k}^{t_{k+1}} \mathrm{tr}\Big( D_s\big((\partial_{\theta_i} Y_{t_{k+1}}^\theta(t_k, x))^* ((\nabla_x Y_{t_{k+1}}^\theta(t_k, x))^{-1})^*\big) \\
&\qquad \cdot \sigma^{-1}(s, Y_s^\theta(t_k, x)) \nabla_x Y_s^\theta(t_k, x)\Big) ds \\
&= \int_{t_k}^{t_{k+1}} ((\nabla_x Y_s^\theta(t_k, x))^{-1} \partial_{\theta_i} b(\theta, s, Y_s^\theta(t_k, x)))^* ds \\
&\qquad \cdot \int_{t_k}^{t_{k+1}} (\nabla_x Y_s^\theta(t_k, x))^* \sigma^{-1}(s, Y_s^\theta(t_k, x)) dW_s \\
&\qquad - \int_{t_k}^{t_{k+1}} \int_{s}^{t_{k+1}} \mathrm{tr}\Big( D_s\big(((\nabla_x Y_u^\theta(t_k, x))^{-1} \partial_{\theta_i} b(\theta, u, Y_u^\theta(t_k, x)))^*\big) \\
&\qquad \cdot \sigma^{-1}(s, Y_s^\theta(t_k, x)) \nabla_x Y_s^\theta(t_k, x)\Big) du ds.
\end{aligned}$$

Next, adding and subtracting the vector $((\nabla_x Y_{t_k}^\theta(t_k, x))^{-1} \partial_{\theta_i} b(\theta, t_k, Y_{t_k}^\theta(t_k, x)))^*$ in the first integral and the matrix $(\nabla_x Y_{t_k}^\theta(t_k, x))^* \sigma^{-1}(t_k, Y_{t_k}^\theta(t_k, x))$ in the second integral together with $Y_{t_k}^\theta(t_k, x) = Y_{t_k}^\theta = x$, we obtain

$$\begin{aligned}
\delta\big(U^\theta(t_k, x)\partial_{\theta_i} Y_{t_{k+1}}^\theta(t_k, x)\big) = \Delta_n (\sigma^{-1}(t_k, x)\partial_{\theta_i} b(\theta, t_k, x))^* (W_{t_{k+1}} - W_{t_k}) \\
- R_1^{\theta,k} + R_2^{\theta,k} + R_3^{\theta,k}.
\end{aligned} \tag{6.5}$$

Finally, from equation (3.3) we write

$$\begin{aligned}
&W_{t_{k+1}} - W_{t_k} = \sigma^{-1}(t_k, Y_{t_k}^\theta)\Big(Y_{t_{k+1}}^\theta - Y_{t_k}^\theta - b(\theta, t_k, Y_{t_k}^\theta)\Delta_n \\
&\quad - \int_{t_k}^{t_{k+1}} \big(b(\theta, s, Y_s^\theta) - b(\theta, t_k, Y_{t_k}^\theta)\big) ds - \int_{t_k}^{t_{k+1}} \big(\sigma(s, Y_s^\theta) - \sigma(t_k, Y_{t_k}^\theta)\big) dW_s\Big).
\end{aligned}$$

This combined with (6.5) concludes the desired result. □

### 6.3. Proof of Lemma 4.4

*Proof.* For $y \in \mathbb{R}^d$, we denote $g(x,y) := \widetilde{\mathrm{E}}^{\theta}_{t_k,x}[V^\theta | Y^\theta_{t_{k+1}} = y]$. Then, we get

$$\begin{aligned}\widehat{\mathrm{E}}^{\theta^0}_{t_k,x}\left[\widetilde{\mathrm{E}}^{\theta}_{t_k,x}\left[V^\theta | Y^\theta_{t_{k+1}} = X^{\theta^0}_{t_{k+1}}\right]\right] &= \widehat{\mathrm{E}}^{\theta^0}_{t_k,x}\left[g(x, X^{\theta^0}_{t_{k+1}})\right] \\ &= \int_{\mathbb{R}^d} g(x,y) p^{\theta^0}(t_k, t_{k+1}, x, y) dy \\ &= \int_{\mathbb{R}^d} g(x,y) \frac{p^{\theta^0}}{p^\theta}(t_k, t_{k+1}, x, y) p^\theta(t_k, t_{k+1}, x, y) dy.\end{aligned}$$

Next, using (4.1) in [27] for homogeneous diffusions and (3.2) in [31] for inhomogeneous diffusions, together with the fact that $Y^\theta$ is the independent copy of $X^\theta$, we have

$$\frac{p^{\theta^0}}{p^\theta}(t_k, t_{k+1}, x, y) = \widetilde{\mathrm{E}}^{\theta}_{t_k,x}\left[\frac{d\widetilde{\mathrm{P}}^{\theta^0}_{t_k,x}}{d\widetilde{\mathrm{P}}^{\theta}_{t_k,x}}((Y^\theta_t)_{t \in I_k}) \Big| Y^\theta_{t_{k+1}} = y\right].$$

Thus, we write

$$\begin{aligned}&\widehat{\mathrm{E}}^{\theta^0}_{t_k,x}\left[\widetilde{\mathrm{E}}^{\theta}_{t_k,x}\left[V^\theta | Y^\theta_{t_{k+1}} = X^{\theta^0}_{t_{k+1}}\right]\right] \\ &= \int_{\mathbb{R}^d} g(x,y) \widetilde{\mathrm{E}}^{\theta}_{t_k,x}\left[\frac{d\widetilde{\mathrm{P}}^{\theta^0}_{t_k,x}}{d\widetilde{\mathrm{P}}^{\theta}_{t_k,x}}((Y^\theta_t)_{t \in I_k}) \Big| Y^\theta_{t_{k+1}} = y\right] p^\theta(t_k, t_{k+1}, x, y) dy \\ &= \widetilde{\mathrm{E}}^{\theta}_{t_k,x}\left[g(x, Y^\theta_{t_{k+1}}) \frac{d\widetilde{\mathrm{P}}^{\theta^0}_{t_k,x}}{d\widetilde{\mathrm{P}}^{\theta}_{t_k,x}}((Y^\theta_t)_{t \in I_k})\right] \\ &= \widetilde{\mathrm{E}}^{\theta}_{t_k,x}\left[g(x, Y^\theta_{t_{k+1}})\right] + \widetilde{\mathrm{E}}^{\theta}_{t_k,x}\left[g(x, Y^\theta_{t_{k+1}})\Big(\frac{d\widetilde{\mathrm{P}}^{\theta^0}_{t_k,x}}{d\widetilde{\mathrm{P}}^{\theta}_{t_k,x}}((Y^\theta_t)_{t \in I_k}) - 1\Big)\right] \\ &= \widetilde{\mathrm{E}}^{\theta}_{t_k,x}\left[V^\theta\right] + \widetilde{\mathrm{E}}^{\theta}_{t_k,x}\left[\widetilde{\mathrm{E}}^{\theta}_{t_k,x}\left[V^\theta | Y^\theta_{t_{k+1}}\right]\Big(\frac{d\widetilde{\mathrm{P}}^{\theta^0}_{t_k,x}}{d\widetilde{\mathrm{P}}^{\theta}_{t_k,x}}((Y^\theta_t)_{t \in I_k}) - 1\Big)\right],\end{aligned}$$

which concludes (4.2). The proof of (4.3) can be obtained by using change of measures. □

### 6.4. Proof of Lemma 4.5

We need to introduce the following auxiliary estimate.

**Lemma 6.1.** *Let $h : \mathbb{R} \to \mathbb{R}$ be a convex function and $\xi = (\xi_s)_{s \geq 0}$ be a stochastic process. Then, for any $\tau > 0$, we have*

$$\mathrm{E}\left[h\left(\int_0^\tau \xi_s ds\right)\right] \leq \sup_{s \in [0,\tau]} \mathrm{E}\left[h(\tau \xi_s)\right],$$

*provided that all the expectations are well-defined.*

*Proof.* Let $U$ be a continuous random variable which is independent of $\xi$, and has a uniform distribution over the interval $[0,\tau]$. Then, using Jensen's inequality, we get

$$\begin{aligned}&\mathrm{E}\Big[h\Big(\int_0^\tau \xi_s ds\Big)\Big] = \mathrm{E}\left[h\left(\mathrm{E}_U[\tau\xi_U]\right)\right]\\&\leq \mathrm{E}\left[\mathrm{E}_U\left[h\left(\tau\xi_U\right)\right]\right] = \frac{1}{\tau}\int_0^\tau \mathrm{E}\left[h(\tau\xi_s)\right]ds \leq \sup_{s\in[0,\tau]}\mathrm{E}\left[h(\tau\xi_s)\right],\end{aligned}$$

where $\mathrm{E}_U$ is the expectation w.r.t. the uniform distribution over $[0,\tau]$. The result follows. □

*Proof.* We start proving (4.4). Using (4.1) for $Y$, we get

$$\begin{aligned}&\frac{d\widetilde{\mathrm{P}}_{t_k,x}^{\theta^0}}{d\widetilde{\mathrm{P}}_{t_k,x}^{\theta_i^{0+}(\ell)}}((\mathrm{Y}_t^{\theta_i^{0+}(\ell)})_{t\in I_k})-1\\&=\frac{(d\widetilde{\mathrm{P}}_{t_k,x}^{\theta_{i+1}^{0+}}-d\widetilde{\mathrm{P}}_{t_k,x}^{\theta_i^{0+}(\ell)})+(d\widetilde{\mathrm{P}}_{t_k,x}^{\theta_{i+2}^{0+}}-d\widetilde{\mathrm{P}}_{t_k,x}^{\theta_{i+1}^{0+}})}{d\widetilde{\mathrm{P}}_{t_k,x}^{\theta_i^{0+}(\ell)}}((\mathrm{Y}_t^{\theta_i^{0+}(\ell)})_{t\in I_k})\\&+\cdots+\frac{(d\widetilde{\mathrm{P}}_{t_k,x}^{\theta_m^{0+}}-d\widetilde{\mathrm{P}}_{t_k,x}^{\theta_{m-1}^{0+}})+(d\widetilde{\mathrm{P}}_{t_k,x}^{\theta^0}-d\widetilde{\mathrm{P}}_{t_k,x}^{\theta_m^{0+}})}{d\widetilde{\mathrm{P}}_{t_k,x}^{\theta_i^{0+}(\ell)}}((\mathrm{Y}_t^{\theta_i^{0+}(\ell)})_{t\in I_k})\\&=\int_{\theta_i^0+\ell\varphi_{n\Delta_n}^i(\theta^0)u_i}^{\theta_i^0}\partial_{\theta_i}\Big(\frac{d\widetilde{\mathrm{P}}_{t_k,x}^{\theta_i(0+)}}{d\widetilde{\mathrm{P}}_{t_k,x}^{\theta_i^{0+}(\ell)}}\Big)((\mathrm{Y}_t^{\theta_i^{0+}(\ell)})_{t\in I_k})d\theta_i\\&+\int_{\theta_{i+1}^0+\varphi_{n\Delta_n}^{i+1}(\theta^0)u_{i+1}}^{\theta_{i+1}^0}\partial_{\theta_{i+1}}\Big(\frac{d\widetilde{\mathrm{P}}_{t_k,x}^{\theta_{i+1}(0+)}}{d\widetilde{\mathrm{P}}_{t_k,x}^{\theta_i^{0+}(\ell)}}\Big)((\mathrm{Y}_t^{\theta_i^{0+}(\ell)})_{t\in I_k})d\theta_{i+1}\\&+\cdots+\int_{\theta_m^0+\varphi_{n\Delta_n}^m(\theta^0)u_m}^{\theta_m^0}\partial_{\theta_m}\Big(\frac{d\widetilde{\mathrm{P}}_{t_k,x}^{\theta_m(0+)}}{d\widetilde{\mathrm{P}}_{t_k,x}^{\theta_i^{0+}(\ell)}}\Big)((\mathrm{Y}_t^{\theta_i^{0+}(\ell)})_{t\in I_k})d\theta_m\\&=\int_{\theta_i^0+\ell\varphi_{n\Delta_n}^i(\theta^0)u_i}^{\theta_i^0}\int_{t_k}^{t_{k+1}}(\partial_{\theta_i}b(\theta_i(0+),t,Y_t^{\theta_i^{0+}(\ell)}))^*\sigma^{-1}(t,Y_t^{\theta_i^{0+}(\ell)})\\&\cdot\big(dW_t-\sigma^{-1}(t,Y_t^{\theta_i^{0+}(\ell)})\big(b(\theta_i(0+),t,Y_t^{\theta_i^{0+}(\ell)})-b(\theta_i^{0+}(\ell),t,Y_t^{\theta_i^{0+}(\ell)})\big)dt\big)\\&\cdot\frac{d\widetilde{\mathrm{P}}_{t_k,x}^{\theta_i(0+)}}{d\widetilde{\mathrm{P}}_{t_k,x}^{\theta_i^{0+}(\ell)}}((\mathrm{Y}_t^{\theta_i^{0+}(\ell)})_{t\in I_k})d\theta_i\\&+\int_{\theta_{i+1}^0+\varphi_{n\Delta_n}^{i+1}(\theta^0)u_{i+1}}^{\theta_{i+1}^0}\int_{t_k}^{t_{k+1}}(\partial_{\theta_{i+1}}b(\theta_{i+1}(0+),t,Y_t^{\theta_i^{0+}(\ell)}))^*\sigma^{-1}(t,Y_t^{\theta_i^{0+}(\ell)})\\&\cdot\big(dW_t-\sigma^{-1}(t,Y_t^{\theta_i^{0+}(\ell)})\big(b(\theta_{i+1}(0+),t,Y_t^{\theta_i^{0+}(\ell)})-b(\theta_i^{0+}(\ell),t,Y_t^{\theta_i^{0+}(\ell)})\big)dt\big)\end{aligned}$$

$$
\begin{aligned}
&\cdot \frac{d\widetilde{\mathrm{P}}_{t_k,x}^{\theta_{i+1}(0+)}}{d\widetilde{\mathrm{P}}_{t_k,x}^{\theta_i^{0+}(\ell)}}((Y_t^{\theta_i^{0+}(\ell)})_{t\in I_k})d\theta_{i+1}+\cdots\\
&+\int_{\theta_m^0+\varphi_{n\Delta_n}^m(\theta^0)u_m}^{\theta_m^0}\int_{t_k}^{t_{k+1}}(\partial_{\theta_m}b(\theta_m(0+),t,Y_t^{\theta_i^{0+}(\ell)}))^*\sigma^{-1}(t,Y_t^{\theta_i^{0+}(\ell)})\\
&\cdot\big(dW_t-\sigma^{-1}(t,Y_t^{\theta_i^{0+}(\ell)})\big(b(\theta_m(0+),t,Y_t^{\theta_i^{0+}(\ell)})-b(\theta_i^{0+}(\ell),t,Y_t^{\theta_i^{0+}(\ell)})\big)dt\big)\\
&\cdot\frac{d\widetilde{\mathrm{P}}_{t_k,x}^{\theta_m(0+)}}{d\widetilde{\mathrm{P}}_{t_k,x}^{\theta_i^{0+}(\ell)}}((Y_t^{\theta_i^{0+}(\ell)})_{t\in I_k})d\theta_m.
\end{aligned}\tag{6.6}
$$

Thus, we have

$$
\begin{aligned}
&\widetilde{\mathrm{E}}_{t_k,x}^{\theta_i^{0+}(\ell)}\Big[\widetilde{\mathrm{E}}_{t_k,x}^{\theta_i^{0+}(\ell)}[V^{\theta_i^{0+}(\ell)}|Y_{t_{k+1}}^{\theta_i^{0+}(\ell)}]\Big(\frac{d\widetilde{\mathrm{P}}_{t_k,x}^{\theta^0}}{d\widetilde{\mathrm{P}}_{t_k,x}^{\theta_i^{0+}(\ell)}}((Y_t^{\theta_i^{0+}(\ell)})_{t\in I_k})-1\Big)\Big]\\
&=\int_{\theta_i^0+\ell\varphi_{n\Delta_n}^i(\theta^0)u_i}^{\theta_i^0}\widetilde{\mathrm{E}}_{t_k,x}^{\theta_i^{0+}(\ell)}\Big[\widetilde{\mathrm{E}}_{t_k,x}^{\theta_i^{0+}(\ell)}[V^{\theta_i^{0+}(\ell)}|Y_{t_{k+1}}^{\theta_i^{0+}(\ell)}]\\
&\cdot\int_{t_k}^{t_{k+1}}(\partial_{\theta_i}b(\theta_i(0+),t,Y_t^{\theta_i^{0+}(\ell)}))^*\sigma^{-1}(t,Y_t^{\theta_i^{0+}(\ell)})\big(dW_t-\sigma^{-1}(t,Y_t^{\theta_i^{0+}(\ell)})\\
&\cdot\big(b(\theta_i(0+),t,Y_t^{\theta_i^{0+}(\ell)})-b(\theta_i^{0+}(\ell),t,Y_t^{\theta_i^{0+}(\ell)})\big)dt\big)\frac{d\widetilde{\mathrm{P}}_{t_k,x}^{\theta_i(0+)}}{d\widetilde{\mathrm{P}}_{t_k,x}^{\theta_i^{0+}(\ell)}}((Y_t^{\theta_i^{0+}(\ell)})_{t\in I_k})\Big]d\theta_i\\
&+\int_{\theta_{i+1}^0+\varphi_{n\Delta_n}^{i+1}(\theta^0)u_{i+1}}^{\theta_{i+1}^0}\widetilde{\mathrm{E}}_{t_k,x}^{\theta_i^{0+}(\ell)}\Big[\widetilde{\mathrm{E}}_{t_k,x}^{\theta_i^{0+}(\ell)}[V^{\theta_i^{0+}(\ell)}|Y_{t_{k+1}}^{\theta_i^{0+}(\ell)}]\\
&\cdot\int_{t_k}^{t_{k+1}}(\partial_{\theta_{i+1}}b(\theta_{i+1}(0+),t,Y_t^{\theta_i^{0+}(\ell)}))^*\sigma^{-1}(t,Y_t^{\theta_i^{0+}(\ell)})\big(dW_t-\sigma^{-1}(t,Y_t^{\theta_i^{0+}(\ell)})\\
&\cdot\big(b(\theta_{i+1}(0+),t,Y_t^{\theta_i^{0+}(\ell)})-b(\theta_i^{0+}(\ell),t,Y_t^{\theta_i^{0+}(\ell)})\big)dt\big)\\
&\cdot\frac{d\widetilde{\mathrm{P}}_{t_k,x}^{\theta_{i+1}(0+)}}{d\widetilde{\mathrm{P}}_{t_k,x}^{\theta_i^{0+}(\ell)}}((Y_t^{\theta_i^{0+}(\ell)})_{t\in I_k})\Big]d\theta_{i+1}\\
&+\cdots+\int_{\theta_m^0+\varphi_{n\Delta_n}^m(\theta^0)u_m}^{\theta_m^0}\widetilde{\mathrm{E}}_{t_k,x}^{\theta_i^{0+}(\ell)}\Big[\widetilde{\mathrm{E}}_{t_k,x}^{\theta_i^{0+}(\ell)}[V^{\theta_i^{0+}(\ell)}|Y_{t_{k+1}}^{\theta_i^{0+}(\ell)}]\\
&\cdot\int_{t_k}^{t_{k+1}}(\partial_{\theta_m}b(\theta_m(0+),t,Y_t^{\theta_i^{0+}(\ell)}))^*\sigma^{-1}(t,Y_t^{\theta_i^{0+}(\ell)})\big(dW_t-\sigma^{-1}(t,Y_t^{\theta_i^{0+}(\ell)})\\
&\cdot\big(b(\theta_m(0+),t,Y_t^{\theta_i^{0+}(\ell)})-b(\theta_i^{0+}(\ell),t,Y_t^{\theta_i^{0+}(\ell)})\big)dt\big)\\
&\cdot\frac{d\widetilde{\mathrm{P}}_{t_k,x}^{\theta_m(0+)}}{d\widetilde{\mathrm{P}}_{t_k,x}^{\theta_i^{0+}(\ell)}}((Y_t^{\theta_i^{0+}(\ell)})_{t\in I_k})\Big]d\theta_m.
\end{aligned}
$$

Next, we treat the first term on the right-hand side. For this, using Hölder's inequality with $\frac{1}{q}+\frac{1}{p}+\frac{1}{r}=1$ and Jensen's inequality, we get

$$\begin{aligned}
D_i &:= \Big| \int_{\theta_i^0+\ell\varphi^i_{n\Delta_n}(\theta^0)u_i}^{\theta_i^0} \widetilde{\mathrm{E}}_{t_k,x}^{\theta_i^{0+}(\ell)} \Big[ \widetilde{\mathrm{E}}_{t_k,x}^{\theta_i^{0+}(\ell)}[V^{\theta_i^{0+}(\ell)} | Y_{t_{k+1}}^{\theta_i^{0+}(\ell)}] \\
&\cdot \int_{t_k}^{t_{k+1}} (\partial_{\theta_i} b(\theta_i(0+),t,Y_t^{\theta_i^{0+}(\ell)}))^* \sigma^{-1}(t,Y_t^{\theta_i^{0+}(\ell)}) \big(dW_t - \sigma^{-1}(t,Y_t^{\theta_i^{0+}(\ell)}) \\
&\cdot \big(b(\theta_i(0+),t,Y_t^{\theta_i^{0+}(\ell)}) - b(\theta_i^{0+}(\ell),t,Y_t^{\theta_i^{0+}(\ell)})\big)dt\big) \\
&\cdot \frac{d\widetilde{\mathrm{P}}_{t_k,x}^{\theta_i(0+)}}{d\widetilde{\mathrm{P}}_{t_k,x}^{\theta_i^{0+}(\ell)}} ((Y_t^{\theta_i^{0+}(\ell)})_{t\in I_k}) \Big] d\theta_i \Big| \\
&\leq \Big| \int_{\theta_i^0+\ell\varphi^i_{n\Delta_n}(\theta^0)u_i}^{\theta_i^0} \big( \widetilde{\mathrm{E}}_{t_k,x}^{\theta_i^{0+}(\ell)} \big[ |V^{\theta_i^{0+}(\ell)}|^q \big] \big)^{\frac{1}{q}} \\
&\cdot \Big( \widetilde{\mathrm{E}}_{t_k,x}^{\theta_i^{0+}(\ell)} \Big[ \Big| \int_{t_k}^{t_{k+1}} (\partial_{\theta_i} b(\theta_i(0+),t,Y_t^{\theta_i^{0+}(\ell)}))^* \sigma^{-1}(t,Y_t^{\theta_i^{0+}(\ell)}) \\
&\cdot \big(dW_t - \sigma^{-1}(t,Y_t^{\theta_i^{0+}(\ell)})(b(\theta_i(0+),t,Y_t^{\theta_i^{0+}(\ell)}) - b(\theta_i^{0+}(\ell),t,Y_t^{\theta_i^{0+}(\ell)}))dt\big) \Big|^p \Big] \Big)^{\frac{1}{p}} \\
&\cdot \Big( \widetilde{\mathrm{E}}_{t_k,x}^{\theta_i^{0+}(\ell)} \Big[ \Big( \frac{d\widetilde{\mathrm{P}}_{t_k,x}^{\theta_i(0+)}}{d\widetilde{\mathrm{P}}_{t_k,x}^{\theta_i^{0+}(\ell)}} ((Y_t^{\theta_i^{0+}(\ell)})_{t\in I_k}) \Big)^r \Big] \Big)^{\frac{1}{r}} d\theta_i \Big|.
\end{aligned}$$

Now, using the mean value theorem, Burkholder-Davis-Gundy's inequality, conditions **(A2)** and **(A3)**(b), and Lemma 4.2 (ii) applied to $Y^{\theta_i^{0+}(\ell)}$, we get for some constant $C>0$,

$$\begin{aligned}
&\widetilde{\mathrm{E}}_{t_k,x}^{\theta_i^{0+}(\ell)} \Big[ \Big| \int_{t_k}^{t_{k+1}} (\partial_{\theta_i} b(\theta_i(0+),t,Y_t^{\theta_i^{0+}(\ell)}))^* \sigma^{-1}(t,Y_t^{\theta_i^{0+}(\ell)}) \big(dW_t - \sigma^{-1}(t,Y_t^{\theta_i^{0+}(\ell)}) \\
&\cdot (b(\theta_i(0+),t,Y_t^{\theta_i^{0+}(\ell)}) - b(\theta_i^{0+}(\ell),t,Y_t^{\theta_i^{0+}(\ell)}))dt\big) \Big|^p \Big] \\
&= \widetilde{\mathrm{E}}_{t_k,x}^{\theta_i^{0+}(\ell)} \Big[ \Big| \int_{t_k}^{t_{k+1}} (\partial_{\theta_i} b(\theta_i(0+),t,Y_t^{\theta_i^{0+}(\ell)}))^* \sigma^{-1}(t,Y_t^{\theta_i^{0+}(\ell)}) \\
&\cdot \big(dW_t - \sigma^{-1}(t,Y_t^{\theta_i^{0+}(\ell)})(-\ell\varphi^i_{n\Delta_n}(\theta^0)u_i \int_0^1 \partial_{\theta_i} b(\theta_i^{0+}(\alpha\ell),t,Y_t^{\theta_i^{0+}(\ell)})d\alpha)dt\big) \Big|^p \Big] \\
&\leq C(\sqrt{\Delta_n})^p(1+|x|^p)\big(1+(\sqrt{\Delta_n}\varphi^i_{n\Delta_n}(\theta^0))^p|x|^p\big).
\end{aligned}$$

Next, using again (4.1) for $Y$, Cauchy-Schwarz inequality, and the fact that the expectation of an exponential martingale is equal to 1, we get

$$\begin{aligned}
&\widetilde{\mathrm{E}}_{t_k,x}^{\theta_i^{0+}(\ell)} \Big[ \Big( \frac{d\widetilde{\mathrm{P}}_{t_k,x}^{\theta_i(0+)}}{d\widetilde{\mathrm{P}}_{t_k,x}^{\theta_i^{0+}(\ell)}} ((Y_t^{\theta_i^{0+}(\ell)})_{t\in I_k}) \Big)^r \Big] \\
&\leq \Big( \widetilde{\mathrm{E}}_{t_k,x}^{\theta_i^{0+}(\ell)} \Big[ \exp\Big\{ 2r \int_{t_k}^{t_{k+1}} \big(b(\theta_i(0+),t,Y_t^{\theta_i^{0+}(\ell)}) - b(\theta_i^{0+}(\ell),t,Y_t^{\theta_i^{0+}(\ell)})\big)^*
\end{aligned}$$

$$
\begin{aligned}
&\cdot \sigma^{-1}(t, Y_t^{\theta_i^{0+}(\ell)})dW_t - 2r^2 \int_{t_k}^{t_{k+1}} \big|\big(b(\theta_i(0+), t, Y_t^{\theta_i^{0+}(\ell)}) \\
&- b(\theta_i^{0+}(\ell), t, Y_t^{\theta_i^{0+}(\ell)})\big)^* \sigma^{-1}(t, Y_t^{\theta_i^{0+}(\ell)})\big|^2 dt\Big\}\Big]\Big)^{\frac{1}{2}} \\
&\times \Big(\widetilde{\mathrm{E}}_{t_k,x}^{\theta_i^{0+}(\ell)}\Big[\exp\Big\{2(r^2 - \frac{r}{2})\int_{t_k}^{t_{k+1}} \big|\big(b(\theta_i(0+), t, Y_t^{\theta_i^{0+}(\ell)}) \\
&- b(\theta_i^{0+}(\ell), t, Y_t^{\theta_i^{0+}(\ell)})\big)^* \sigma^{-1}(t, Y_t^{\theta_i^{0+}(\ell)})\big|^2 dt\Big\}\Big]\Big)^{\frac{1}{2}} \\
= &\Big(\widetilde{\mathrm{E}}_{t_k,x}^{\theta_i^{0+}(\ell)}\Big[\exp\Big\{2(r^2 - \frac{r}{2})\int_{t_k}^{t_{k+1}} \big|\big(b(\theta_i(0+), t, Y_t^{\theta_i^{0+}(\ell)}) \\
&- b(\theta_i^{0+}(\ell), t, Y_t^{\theta_i^{0+}(\ell)})\big)^* \sigma^{-1}(t, Y_t^{\theta_i^{0+}(\ell)})\big|^2 dt\Big\}\Big]\Big)^{\frac{1}{2}}.
\end{aligned}
$$

Then, using the mean value theorem, conditions **(A2)** and **(A3)**(b), Lemma 6.1, and the argument in the proof of [12, Lemma A.1], we get for $n$ large enough,

$$
\begin{aligned}
&\widetilde{\mathrm{E}}_{t_k,x}^{\theta_i^{0+}(\ell)}\Big[\exp\Big\{2(r^2 - \frac{r}{2})\int_{t_k}^{t_{k+1}} \big|\big(b(\theta_i(0+), t, Y_t^{\theta_i^{0+}(\ell)}) \\
&\quad - b(\theta_i^{0+}(\ell), t, Y_t^{\theta_i^{0+}(\ell)})\big)^* \sigma^{-1}(t, Y_t^{\theta_i^{0+}(\ell)})\big|^2 dt\Big\}\Big] \\
&= \widetilde{\mathrm{E}}_{t_k,x}^{\theta_i^{0+}(\ell)}\Big[\exp\Big\{2(r^2 - \frac{r}{2})\ell^2 u_i^2 (\varphi_{n\Delta_n}^i(\theta^0))^2 \\
&\quad \times \int_{t_k}^{t_{k+1}} \int_0^1 \big|\big(\partial_{\theta_i} b(\theta_i^{0+}(\alpha\ell), t, Y_t^{\theta_i^{0+}(\ell)})\big)^* \sigma^{-1}(t, Y_t^{\theta_i^{0+}(\ell)})\big|^2 d\alpha dt\Big\}\Big] \\
&\leq \widetilde{\mathrm{E}}_{t_k,x}^{\theta_i^{0+}(\ell)}\Big[\exp\Big\{C(\varphi_{n\Delta_n}^i(\theta^0))^2 \int_{t_k}^{t_{k+1}} (1 + |Y_t^{\theta_i^{0+}(\ell)}|^2) dt\Big\}\Big] \\
&\leq \sup_{t\in[t_k,t_{k+1}]} \widetilde{\mathrm{E}}_{t_k,x}^{\theta_i^{0+}(\ell)}\Big[\exp\Big\{C(\varphi_{n\Delta_n}^i(\theta^0))^2 \Delta_n (1 + |Y_t^{\theta_i^{0+}(\ell)}|^2)\Big\}\Big] \\
&= \sup_{t\in[0,\Delta_n]} \widetilde{\mathrm{E}}_{0,x}^{\theta_i^{0+}(\ell)}\Big[\exp\Big\{C(\varphi_{n\Delta_n}^i(\theta^0))^2 \Delta_n (1 + |Y_t^{\theta_i^{0+}(\ell)}|^2)\Big\}\Big] \\
&\leq K \sup_{t\in[0,\Delta_n]} \widetilde{\mathrm{E}}_{0,x}^{\theta_i^{0+}(\ell)}\Big[\exp\Big\{C(\varphi_{n\Delta_n}^i(\theta^0))^2 \Delta_n |Y_t^{\theta_i^{0+}(\ell)}|^2\Big\}\Big] \\
&\leq K e^{C(\varphi_{n\Delta_n}^i(\theta^0))^2 \Delta_n |x|^2},
\end{aligned}
$$

for some constants $K, C > 0$. Thus, we have shown that

$$
\begin{aligned}
D_i \leq\ & C(\widetilde{\mathrm{E}}_{t_k,x}^{\theta_i^{0+}(\ell)}[|V^{\theta_i^{0+}(\ell)}|^q])^{\frac{1}{q}} \sqrt{\Delta_n}(1 + |x|)\varphi_{n\Delta_n}^i(\theta^0)(1 + \sqrt{\Delta_n}\varphi_{n\Delta_n}^i(\theta^0)|x|) \\
&\times e^{C_i(\varphi_{n\Delta_n}^i(\theta^0))^2 \Delta_n |x|^2},
\end{aligned}
$$

for some constants $C, C_i > 0$. The same arguments also apply to the other terms. Therefore,

$$
\Big|\widetilde{\mathrm{E}}_{t_k,x}^{\theta_i^{0+}(\ell)}\Big[\widetilde{\mathrm{E}}_{t_k,x}^{\theta_i^{0+}(\ell)}[V^{\theta_i^{0+}(\ell)}|Y_{t_{k+1}}^{\theta_i^{0+}(\ell)}]\Big(\frac{d\widetilde{\mathrm{P}}_{t_k,x}^{\theta^0}}{d\widetilde{\mathrm{P}}_{t_k,x}^{\theta_i^{0+}(\ell)}}((Y_t^{\theta_i^{0+}(\ell)})_{t\in I_k}) - 1\Big)\Big]\Big|
$$

$$
\begin{aligned}
&\leq C\Big(\widetilde{\mathrm{E}}_{t_k,x}^{\theta_i^{0+}(\ell)}[|V^{\theta_i^{0+}(\ell)}|^q]\Big)^{\frac{1}{q}}\sqrt{\Delta_n}(1+|x|)\\
&\quad\times\bigg(\varphi_{n\Delta_n}^i(\theta^0)\big(1+\sqrt{\Delta_n}\varphi_{n\Delta_n}^i(\theta^0)|x|\big)e^{C_i(\varphi_{n\Delta_n}^i(\theta^0))^2\Delta_n|x|^2}\\
&\quad+\varphi_{n\Delta_n}^{i+1}(\theta^0)\big(1+\sqrt{\Delta_n}(\varphi_{n\Delta_n}^i(\theta^0)+\varphi_{n\Delta_n}^{i+1}(\theta^0))|x|\big)\\
&\quad\times e^{C_{i+1}((\varphi_{n\Delta_n}^i(\theta^0))^2+(\varphi_{n\Delta_n}^{i+1}(\theta^0))^2)\Delta_n|x|^2}+\cdots\\
&\quad+\varphi_{n\Delta_n}^m(\theta^0)\big(1+\sqrt{\Delta_n}(\varphi_{n\Delta_n}^i(\theta^0)+\cdots+\varphi_{n\Delta_n}^m(\theta^0))|x|\big)\\
&\quad\times e^{C_m((\varphi_{n\Delta_n}^i(\theta^0))^2+\cdots+(\varphi_{n\Delta_n}^m(\theta^0))^2)\Delta_n|x|^2}\bigg),
\end{aligned}
$$

for any $q>1$, some constants $C,C_i,\dots,C_m>0$ and $n$ large enough. This concludes (4.4).

We next prove (4.5). Proceeding as (6.6) and using (1.1), we get

$$
\begin{aligned}
&\frac{d\widehat{\mathrm{P}}_{t_k,x}^{\theta^0}}{d\widehat{\mathrm{P}}_{t_k,x}^{\theta_i^{0+}(\ell)}}((X_t^{\theta_i^{0+}(\ell)})_{t\in I_k})-1\\
&=\int_{\theta_i^0+\ell\varphi_{n\Delta_n}^i(\theta^0)u_i}^{\theta_i^0}\int_{t_k}^{t_{k+1}}(\partial_{\theta_i}b(\theta_i(0+),t,X_t^{\theta_i^{0+}(\ell)}))^*\sigma^{-2}(t,X_t^{\theta_i^{0+}(\ell)})\\
&\quad\cdot\big(dX_t^{\theta_i^{0+}(\ell)}-b(\theta_i(0+),t,X_t^{\theta_i^{0+}(\ell)})dt\big)\frac{d\widehat{\mathrm{P}}_{t_k,x}^{\theta_i(0+)}}{d\widehat{\mathrm{P}}_{t_k,x}^{\theta_i^{0+}(\ell)}}((X_t^{\theta_i^{0+}(\ell)})_{t\in I_k})d\theta_i\\
&\quad+\int_{\theta_{i+1}^0+\varphi_{n\Delta_n}^{i+1}(\theta^0)u_{i+1}}^{\theta_{i+1}^0}\int_{t_k}^{t_{k+1}}(\partial_{\theta_{i+1}}b(\theta_{i+1}(0+),t,X_t^{\theta_i^{0+}(\ell)}))^*\sigma^{-2}(t,X_t^{\theta_i^{0+}(\ell)})\\
&\quad\cdot\big(dX_t^{\theta_i^{0+}(\ell)}-b(\theta_{i+1}(0+),t,X_t^{\theta_i^{0+}(\ell)})dt\big)\frac{d\widehat{\mathrm{P}}_{t_k,x}^{\theta_{i+1}(0+)}}{d\widehat{\mathrm{P}}_{t_k,x}^{\theta_i^{0+}(\ell)}}((X_t^{\theta_i^{0+}(\ell)})_{t\in I_k})d\theta_{i+1}\\
&\quad+\cdots+\int_{\theta_m^0+\varphi_{n\Delta_n}^m(\theta^0)u_m}^{\theta_m^0}\int_{t_k}^{t_{k+1}}(\partial_{\theta_m}b(\theta_m(0+),t,X_t^{\theta_i^{0+}(\ell)}))^*\sigma^{-2}(t,X_t^{\theta_i^{0+}(\ell)})\\
&\quad\cdot\big(dX_t^{\theta_i^{0+}(\ell)}-b(\theta_m(0+),t,X_t^{\theta_i^{0+}(\ell)})dt\big)\frac{d\widehat{\mathrm{P}}_{t_k,x}^{\theta_m(0+)}}{d\widehat{\mathrm{P}}_{t_k,x}^{\theta_i^{0+}(\ell)}}((X_t^{\theta_i^{0+}(\ell)})_{t\in I_k})d\theta_m.
\end{aligned}
$$

Then, using change of measures and (1.1), we get

$$
\begin{aligned}
&\widehat{\mathrm{E}}_{t_k,x}^{\theta_i^{0+}(\ell)}\Big[\widehat{V}^{\theta_i^{0+}(\ell)}\Big(\frac{d\widehat{\mathrm{P}}_{t_k,x}^{\theta^0}}{d\widehat{\mathrm{P}}_{t_k,x}^{\theta_i^{0+}(\ell)}}((X_t^{\theta_i^{0+}(\ell)})_{t\in I_k})-1\Big)\Big]\\
&=\int_{\theta_i^0+\ell\varphi_{n\Delta_n}^i(\theta^0)u_i}^{\theta_i^0}\widehat{\mathrm{E}}_{t_k,x}^{\theta_i^{0+}(\ell)}\Big[\widehat{V}^{\theta_i^{0+}(\ell)}\int_{t_k}^{t_{k+1}}(\partial_{\theta_i}b(\theta_i(0+),t,X_t^{\theta_i^{0+}(\ell)}))^*\\
&\cdot\sigma^{-2}(t,X_t^{\theta_i^{0+}(\ell)})\big(dX_t^{\theta_i^{0+}(\ell)}-b(\theta_i(0+),t,X_t^{\theta_i^{0+}(\ell)})dt\big)\frac{d\widehat{\mathrm{P}}_{t_k,x}^{\theta_i(0+)}}{d\widehat{\mathrm{P}}_{t_k,x}^{\theta_i^{0+}(\ell)}}((X_t^{\theta_i^{0+}(\ell)})_{t\in I_k})\Big]d\theta_i
\end{aligned}
$$

$$
\begin{aligned}
&+ \int_{\theta^0_{i+1}+\varphi^{i+1}_{n\Delta_n}(\theta^0)u_{i+1}}^{\theta^0_{i+1}} \widehat{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,x}\Big[\widehat{V}^{\theta_i^{0+}(\ell)} \int_{t_k}^{t_{k+1}} (\partial_{\theta_{i+1}} b(\theta_{i+1}(0+),t,X_t^{\theta_i^{0+}(\ell)}))^* \\
&\cdot \sigma^{-2}(t,X_t^{\theta_i^{0+}(\ell)})\big(dX_t^{\theta_i^{0+}(\ell)} - b(\theta_{i+1}(0+),t,X_t^{\theta_i^{0+}(\ell)})dt\big) \\
&\times \frac{d\widehat{\mathrm{P}}^{\theta_{i+1}(0+)}_{t_k,x}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,x}}((X_t^{\theta_i^{0+}(\ell)})_{t\in I_k})\Big]d\theta_{i+1} \\
&+ \cdots + \int_{\theta^0_m+\varphi^m_{n\Delta_n}(\theta^0)u_m}^{\theta^0_m} \widehat{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,x}\Big[\widehat{V}^{\theta_i^{0+}(\ell)} \int_{t_k}^{t_{k+1}} (\partial_{\theta_m} b(\theta_m(0+),t,X_t^{\theta_i^{0+}(\ell)}))^* \\
&\cdot \sigma^{-2}(t,X_t^{\theta_i^{0+}(\ell)})\big(dX_t^{\theta_i^{0+}(\ell)} - b(\theta_m(0+),t,X_t^{\theta_i^{0+}(\ell)})dt\big) \\
&\times \frac{d\widehat{\mathrm{P}}^{\theta_m(0+)}_{t_k,x}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,x}}((X_t^{\theta_i^{0+}(\ell)})_{t\in I_k})\Big]d\theta_m \\
&= \int_{\theta^0_i+\ell\varphi^i_{n\Delta_n}(\theta^0)u_i}^{\theta^0_i} \widehat{\mathrm{E}}^{\theta_i(0+)}_{t_k,x}\Big[\widehat{V}^{\theta_i(0+)} \int_{t_k}^{t_{k+1}} (\partial_{\theta_i} b(\theta_i(0+),t,X_t^{\theta_i(0+)}))^* \\
&\cdot \sigma^{-2}(t,X_t^{\theta_i(0+)})\big(dX_t^{\theta_i(0+)} - b(\theta_i(0+),t,X_t^{\theta_i(0+)})dt\big)\Big]d\theta_i \\
&+ \int_{\theta^0_{i+1}+\varphi^{i+1}_{n\Delta_n}(\theta^0)u_{i+1}}^{\theta^0_{i+1}} \widehat{\mathrm{E}}^{\theta_{i+1}(0+)}_{t_k,x}\Big[\widehat{V}^{\theta_{i+1}(0+)} \int_{t_k}^{t_{k+1}} (\partial_{\theta_{i+1}} b(\theta_{i+1}(0+),t,X_t^{\theta_{i+1}(0+)}))^* \\
&\cdot \sigma^{-2}(t,X_t^{\theta_{i+1}(0+)})\big(dX_t^{\theta_{i+1}(0+)} - b(\theta_{i+1}(0+),t,X_t^{\theta_{i+1}(0+)})dt\big)\Big]d\theta_{i+1} \\
&+ \cdots + \int_{\theta^0_m+\varphi^m_{n\Delta_n}(\theta^0)u_m}^{\theta^0_m} \widehat{\mathrm{E}}^{\theta_m(0+)}_{t_k,x}\Big[\widehat{V}^{\theta_m(0+)} \int_{t_k}^{t_{k+1}} (\partial_{\theta_m} b(\theta_m(0+),t,X_t^{\theta_m(0+)}))^* \\
&\cdot \sigma^{-2}(t,X_t^{\theta_m(0+)})\big(dX_t^{\theta_m(0+)} - b(\theta_m(0+),t,X_t^{\theta_m(0+)})dt\big)\Big]d\theta_m \\
&= \int_{\theta^0_i+\ell\varphi^i_{n\Delta_n}(\theta^0)u_i}^{\theta^0_i} \widehat{\mathrm{E}}^{\theta_i(0+)}_{t_k,x}\Big[\widehat{V}^{\theta_i(0+)} \int_{t_k}^{t_{k+1}} (\partial_{\theta_i} b(\theta_i(0+),t,X_t^{\theta_i(0+)}))^* \\
&\cdot \sigma^{-1}(t,X_t^{\theta_i(0+)})dB_t\Big]d\theta_i + \int_{\theta^0_{i+1}+\varphi^{i+1}_{n\Delta_n}(\theta^0)u_{i+1}}^{\theta^0_{i+1}} \widehat{\mathrm{E}}^{\theta_{i+1}(0+)}_{t_k,x}\Big[\widehat{V}^{\theta_{i+1}(0+)} \\
&\cdot \int_{t_k}^{t_{k+1}} (\partial_{\theta_{i+1}} b(\theta_{i+1}(0+),t,X_t^{\theta_{i+1}(0+)}))^* \sigma^{-1}(t,X_t^{\theta_{i+1}(0+)})dB_t\Big]d\theta_{i+1} \\
&+ \cdots + \int_{\theta^0_m+\varphi^m_{n\Delta_n}(\theta^0)u_m}^{\theta^0_m} \widehat{\mathrm{E}}^{\theta_m(0+)}_{t_k,x}\Big[\widehat{V}^{\theta_m(0+)} \\
&\cdot \int_{t_k}^{t_{k+1}} (\partial_{\theta_m} b(\theta_m(0+),t,X_t^{\theta_m(0+)}))^* \sigma^{-1}(t,X_t^{\theta_m(0+)})dB_t\Big]d\theta_m.
\end{aligned}
$$

Next, using Hölder's inequality with $\frac{1}{q}+\frac{1}{p}=1$, Burkholder-Davis-Gundy's inequality, conditions **(A2)** and **(A3)**(b), and Lemma 4.2 (ii), we get

$$\begin{aligned}&\Big|\widehat{\mathrm{E}}_{t_k,x}^{\theta_i^{0+}(\ell)}\Big[\widehat{V}^{\theta_i^{0+}(\ell)}\Big(\frac{d\widehat{\mathrm{P}}_{t_k,x}^{\theta^0}}{d\widehat{\mathrm{P}}_{t_k,x}^{\theta_i^{0+}(\ell)}}((X_t^{\theta_i^{0+}(\ell)})_{t\in I_k})-1\Big)\Big]\Big|\\&\leq C\sqrt{\Delta_n}(1+|x|)\Big(\Big|\int_{\theta_i^0+\ell\varphi_{n\Delta_n}^i(\theta^0)u_i}^{\theta_i^0}\big(\widehat{\mathrm{E}}_{t_k,x}^{\theta_i(0+)}[|\widehat{V}^{\theta_i(0+)}|^q]\big)^{\frac{1}{q}}d\theta_i\Big|\\&\qquad+\cdots+\Big|\int_{\theta_m^0+\varphi_{n\Delta_n}^m(\theta^0)u_m}^{\theta_m^0}\big(\widehat{\mathrm{E}}_{t_k,x}^{\theta_m(0+)}[|\widehat{V}^{\theta_m(0+)}|^q]\big)^{\frac{1}{q}}d\theta_m\Big|\Big),\end{aligned}$$

which concludes (4.5). Thus, the result follows. □

### *6.5. Proof of Lemma 4.8*

*Proof.* First, observe that $\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}[H_{1,k,n}|\widehat{\mathcal{F}}_{t_k}]=0$. Next, using Itô's formula, we get

$$\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}\big[H_{1,k,n}^2|\widehat{\mathcal{F}}_{t_k}\big]\leq C\big((\varphi_{n\Delta_n}^i(\theta^0))^2+\ldots+(\varphi_{n\Delta_n}^m(\theta^0))^2\big)\Delta_n^2\sum_{k=0}^{n-1}(1+|X_{t_k}^{\theta^0}|^4).$$

Then, taking expectations on both sides and using Lemma 4.3, we obtain

$$\begin{aligned}\widehat{\mathrm{E}}^{\theta^0}\Big[\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}\big[H_{1,k,n}^2|\widehat{\mathcal{F}}_{t_k}\big]\Big]&\leq C\Delta_n^2\big((\varphi_{n\Delta_n}^i(\theta^0))^2+\ldots+(\varphi_{n\Delta_n}^m(\theta^0))^2\big)\\&\qquad\times\Big(n+\sum_{k=0}^{n-1}\psi_{t_k}^{-4}(\theta^0)\Big),\end{aligned}$$

which, by condition **(A6)**, tends to zero. Thus, $\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}[H_{1,k,n}^2|\widehat{\mathcal{F}}_{t_k}]\overset{\widehat{\mathrm{P}}^{\theta^0}}{\longrightarrow}0$. Therefore, by Lemma 4.1 a), $H_{1,k,n}\overset{\widehat{\mathrm{P}}^{\theta^0}}{\longrightarrow}0$.

Next, using Itô's formula and Lemma 4.3,

$$\begin{aligned}\widehat{\mathrm{E}}^{\theta^0}\Big[\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}\big[|H_{2,k,n}||\widehat{\mathcal{F}}_{t_k}\big]\Big]&\leq C\sum_{i,j\in\{1,\ldots,m\}}\Delta_n^{\frac{3}{2}}\varphi_{n\Delta_n}^i(\theta^0)\varphi_{n\Delta_n}^j(\theta^0)\\&\qquad\times\Big(n+\sum_{k=0}^{n-1}\psi_{t_k}^{-3}(\theta^0)\Big),\end{aligned}$$

which, by condition **(A6)**, tends to zero. Thus, $\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}[|H_{2,k,n}||\widehat{\mathcal{F}}_{t_k}]\overset{\widehat{\mathrm{P}}^{\theta^0}}{\longrightarrow}0$. Therefore, by Lemma 4.1 b), $H_{2,k,n}\overset{\widehat{\mathrm{P}}^{\theta^0}}{\longrightarrow}0$.

Now, using **(A2)** and **(A3)** and Lemma 4.3, $\widehat{\mathrm{E}}^{\theta^0}[\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}[|K_{j,k,n}||\widehat{\mathcal{F}}_{t_k}]]$ is bounded by

$$C\Delta_n\varphi^i_{n\Delta_n}(\theta^0)\varphi^j_{n\Delta_n}(\theta^0)\big((\varphi^i_{n\Delta_n}(\theta^0))^\gamma+\ldots+(\varphi^m_{n\Delta_n}(\theta^0))^\gamma\big)\big(n+\sum_{k=0}^{n-1}\psi_{t_k}^{-2}(\theta^0)\big),$$

which, by **(A6)**, tends to zero. Thus, by Lemma 4.1 b), $K_{j,k,n}\overset{\widehat{\mathrm{P}}^{\theta^0}}{\longrightarrow}0$ for $j\in\{i,i+1,\ldots,m\}$.

Finally, we treat $\xi_{2,i,k,n}$. For this, we write $\xi_{2,i,k,n}=\xi_{2,1,i,k,n}+\xi_{2,2,i,k,n}$, where

$$\begin{aligned}\xi_{2,1,i,k,n}&=\varphi^i_{n\Delta_n}(\theta^0)u_i\int_0^1\Big(\partial_{\theta_i}b(\theta_i^{0+}(\ell),t_k,X_{t_k}^{\theta^0})-\partial_{\theta_i}b(\theta^0,t_k,X_{t_k}^{\theta^0})\Big)^*\\&\qquad\cdot\sigma^{-2}(t_k,X_{t_k}^{\theta^0})\sigma(t_k,X_{t_k}^{\theta^0})(B_{t_{k+1}}-B_{t_k})d\ell,\\ \xi_{2,2,i,k,n}&=\varphi^i_{n\Delta_n}(\theta^0)u_i\int_0^1\Big(\partial_{\theta_i}b(\theta_i^{0+}(\ell),t_k,X_{t_k}^{\theta^0})-\partial_{\theta_i}b(\theta^0,t_k,X_{t_k}^{\theta^0})\Big)^*\\&\qquad\cdot\sigma^{-2}(t_k,X_{t_k}^{\theta^0})\Big(b(\theta^0,t_k,X_{t_k}^{\theta^0})-b(\theta_i^{0+}(\ell),t_k,X_{t_k}^{\theta^0})\Big)\Delta_n d\ell.\end{aligned}$$

Observe that $\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}[\xi_{2,1,i,k,n}|\widehat{\mathcal{F}}_{t_k}]=0$. Using **(A2)** and **(A3)** and Lemma 4.3,

$$\begin{aligned}&\widehat{\mathrm{E}}^{\theta^0}\Big[\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}\Big[\xi^2_{2,1,i,k,n}|\widehat{\mathcal{F}}_{t_k}\Big]\Big]\\&\leq C\Delta_n(\varphi^i_{n\Delta_n}(\theta^0))^2\big((\varphi^i_{n\Delta_n}(\theta^0))^{2\gamma}+\ldots+(\varphi^m_{n\Delta_n}(\theta^0))^{2\gamma}\big)\big(n+\sum_{k=0}^{n-1}\psi_{t_k}^{-2}(\theta^0)\big),\end{aligned}$$

which, by **(A6)**, tends to zero. Thus, by Lemma 4.1 a), $\xi_{2,1,i,k,n}\overset{\widehat{\mathrm{P}}^{\theta^0}}{\longrightarrow}0$.

Next, using **(A2)** and **(A3)** and Lemma 4.3,

$$\begin{aligned}&\widehat{\mathrm{E}}^{\theta^0}\Big[\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}\big[|\xi_{2,2,i,k,n}||\widehat{\mathcal{F}}_{t_k}\big]\Big]\\&\leq C\Delta_n\varphi^i_{n\Delta_n}(\theta^0)\big((\varphi^i_{n\Delta_n}(\theta^0))^\gamma+\ldots+(\varphi^m_{n\Delta_n}(\theta^0))^\gamma\big)\\&\qquad\times\big(\varphi^i_{n\Delta_n}(\theta^0)+\ldots+\varphi^m_{n\Delta_n}(\theta^0)\big)\big(n+\sum_{k=0}^{n-1}\psi_{t_k}^{-2}(\theta^0)\big),\end{aligned}$$

which, by **(A6)**, tends to zero. Thus, by Lemma 4.1 b), $\xi_{2,2,i,k,n}\overset{\widehat{\mathrm{P}}^{\theta^0}}{\longrightarrow}0$. The result follows. □

### 6.6. *Proof of Lemma 4.10*

*Proof.* We apply Lemma 4.1 a) with

$$\zeta_{k,n}=\zeta_{i,k,n}$$

$$= \frac{\varphi^i_{n\Delta_n}(\theta^0)u_i}{\Delta_n} \int_0^1 \Big\{ Z^{5,\ell}_{i,k,n} - \widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k, X^{\theta^0}_{t_k}} \big[ R_5^{\theta_i^{0+}(\ell),k} | Y^{\theta_i^{0+}(\ell)}_{t_{k+1}} = X^{\theta^0}_{t_{k+1}} \big] \Big\} d\ell.$$

Using (4.2) of Lemma 4.4 with $V^{\theta_i^{0+}(\ell)} = R_5^{\theta_i^{0+}(\ell),k}$, and (4.4) of Lemma 4.5 with $q=2$, the mean value theorem for vector-valued functions, and Lemma 4.2 (i) applied to $Y^{\theta_i^{0+}(\ell)}$, we get

$$\begin{aligned}
&\Big| \sum_{k=0}^{n-1} \widehat{\mathrm{E}}^{\theta^0} \big[ \zeta_{i,k,n} | \widehat{\mathcal{F}}_{t_k} \big] \Big| \\
&\leq C \sum_{k=0}^{n-1} \frac{\varphi^i_{n\Delta_n}(\theta^0)}{\Delta_n} \int_0^1 \sqrt{\Delta_n}(1+|X^{\theta^0}_{t_k}|) \big( \widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k, X^{\theta^0}_{t_k}} [|R_5^{\theta_i^{0+}(\ell),k}|^2] \big)^{\frac{1}{2}} \\
&\quad \times \Big( \varphi^i_{n\Delta_n}(\theta^0) \big(1+\sqrt{\Delta_n}\varphi^i_{n\Delta_n}(\theta^0)|X^{\theta^0}_{t_k}|\big) e^{C_i(\varphi^i_{n\Delta_n}(\theta^0))^2\Delta_n|X^{\theta^0}_{t_k}|^2} \\
&\quad + \varphi^{i+1}_{n\Delta_n}(\theta^0)\big(1+\sqrt{\Delta_n}(\varphi^i_{n\Delta_n}(\theta^0)+\varphi^{i+1}_{n\Delta_n}(\theta^0))|X^{\theta^0}_{t_k}|\big) \\
&\quad \times e^{C_{i+1}((\varphi^i_{n\Delta_n}(\theta^0))^2+(\varphi^{i+1}_{n\Delta_n}(\theta^0))^2)\Delta_n|X^{\theta^0}_{t_k}|^2} + \cdots \\
&\quad + \varphi^m_{n\Delta_n}(\theta^0)\big(1+\sqrt{\Delta_n}(\varphi^i_{n\Delta_n}(\theta^0)+\cdots+\varphi^m_{n\Delta_n}(\theta^0))|X^{\theta^0}_{t_k}|\big) \\
&\quad \times e^{C_m((\varphi^i_{n\Delta_n}(\theta^0))^2+\cdots+(\varphi^m_{n\Delta_n}(\theta^0))^2)\Delta_n|X^{\theta^0}_{t_k}|^2} \Big) d\ell \\
&\leq C(\varphi^i_{n\Delta_n}(\theta^0))^2\Delta_n^{\frac{3}{2}} \sum_{k=0}^{n-1} \Big( 1+|X^{\theta^0}_{t_k}|^3+\sqrt{\Delta_n}\varphi^i_{n\Delta_n}(\theta^0)|X^{\theta^0}_{t_k}|^4 \Big) \\
&\quad \times e^{C_i(\varphi^i_{n\Delta_n}(\theta^0))^2\Delta_n|X^{\theta^0}_{t_k}|^2} + C\varphi^i_{n\Delta_n}(\theta^0)\varphi^{i+1}_{n\Delta_n}(\theta^0)\Delta_n^{\frac{3}{2}} \sum_{k=0}^{n-1} \Big( 1+|X^{\theta^0}_{t_k}|^3 \\
&\quad + \sqrt{\Delta_n}(\varphi^i_{n\Delta_n}(\theta^0)+\varphi^{i+1}_{n\Delta_n}(\theta^0))|X^{\theta^0}_{t_k}|^4 \Big) e^{C_{i+1}((\varphi^i_{n\Delta_n}(\theta^0))^2+(\varphi^{i+1}_{n\Delta_n}(\theta^0))^2)\Delta_n|X^{\theta^0}_{t_k}|^2} \\
&\quad + \cdots + C\varphi^i_{n\Delta_n}(\theta^0)\varphi^m_{n\Delta_n}(\theta^0)\Delta_n^{\frac{3}{2}} \sum_{k=0}^{n-1} \Big( 1+|X^{\theta^0}_{t_k}|^3+\sqrt{\Delta_n}(\varphi^i_{n\Delta_n}(\theta^0)+\cdots \\
&\quad + \varphi^m_{n\Delta_n}(\theta^0))|X^{\theta^0}_{t_k}|^4 \Big) e^{C_m((\varphi^i_{n\Delta_n}(\theta^0))^2+\cdots+(\varphi^m_{n\Delta_n}(\theta^0))^2)\Delta_n|X^{\theta^0}_{t_k}|^2}.
\end{aligned}$$

Here, we have used the fact that $\widehat{\mathrm{E}}^{\theta^0}_{t_k, X^{\theta^0}_{t_k}}[Z^{5,\ell}_{i,k,n}] = \widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k, X^{\theta^0}_{t_k}}[R_5^{\theta_i^{0+}(\ell),k}] = 0$ and the mean value theorem for vector-valued functions to write

$$\begin{aligned}
&\sigma_i(s, Y_s^{\theta_i^{0+}(\ell)}) - \sigma_i(t_k, Y_{t_k}^{\theta_i^{0+}(\ell)}) \\
&= \Big( \int_0^1 J_{\sigma_i}(t_k+v(s-t_k), Y_{t_k}^{\theta_i^{0+}(\ell)} + v(Y_s^{\theta_i^{0+}(\ell)} - Y_{t_k}^{\theta_i^{0+}(\ell)}))dv \Big) \\
&\qquad \cdot \begin{pmatrix} s-t_k \\ Y_s^{\theta_i^{0+}(\ell)} - Y_{t_k}^{\theta_i^{0+}(\ell)} \end{pmatrix},
\end{aligned}$$

where the Jacobian matrix is given by

$$J_{\sigma_i} = \begin{pmatrix} \partial_t \sigma_{1i} & \partial_{x_1}\sigma_{1i} & \dots & \partial_{x_d}\sigma_{1i} \\ \vdots & \vdots & \ddots & \vdots \\ \partial_t \sigma_{di} & \partial_{x_1}\sigma_{di} & \dots & \partial_{x_d}\sigma_{di} \end{pmatrix}.$$

Thus, using Young's inequality with $\frac{1}{p_0} + \frac{1}{q_0} = 1$ and $p_0$ close to 1, Lemma 4.3 and **(A7)**,

$$\begin{aligned}
&\widehat{\mathrm{E}}^{\theta^0}\Big[\Big|\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}\Big[\zeta_{i,k,n}|\widehat{\mathcal{F}}_{t_k}\Big]\Big|\Big] \\
&\le C\Delta_n^{\frac{3}{2}}(\varphi^i_{n\Delta_n}(\theta^0))^2\Big(n + \sum_{k=0}^{n-1}\psi_{t_k}^{-3p_0}(\theta^0) + (\sqrt{\Delta_n}\varphi^i_{n\Delta_n}(\theta^0))^{p_0}\sum_{k=0}^{n-1}\psi_{t_k}^{-4p_0}(\theta^0)\Big) \\
&+ C\Delta_n^{\frac{3}{2}}\varphi^i_{n\Delta_n}(\theta^0)\varphi^{i+1}_{n\Delta_n}(\theta^0)\Big(n + \sum_{k=0}^{n-1}\psi_{t_k}^{-3p_0}(\theta^0) + (\sqrt{\Delta_n}(\varphi^i_{n\Delta_n}(\theta^0) \\
&+ \varphi^{i+1}_{n\Delta_n}(\theta^0))^{p_0}\sum_{k=0}^{n-1}\psi_{t_k}^{-4p_0}(\theta^0)\Big) + \dots + C\Delta_n^{\frac{3}{2}}\varphi^i_{n\Delta_n}(\theta^0)\varphi^m_{n\Delta_n}(\theta^0) \\
&\times\Big(n + \sum_{k=0}^{n-1}\psi_{t_k}^{-3p_0}(\theta^0) + (\sqrt{\Delta_n}(\varphi^i_{n\Delta_n}(\theta^0) + \dots + \varphi^m_{n\Delta_n}(\theta^0))^{p_0}\sum_{k=0}^{n-1}\psi_{t_k}^{-4p_0}(\theta^0)\Big),
\end{aligned}$$

which, by condition **(A6)**, tends to zero. Thus, $\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}[\zeta_{i,k,n}|\widehat{\mathcal{F}}_{t_k}] \xrightarrow{\widehat{\mathrm{P}}^{\theta^0}} 0$.

Next, applying Jensen's inequality, (4.2) of Lemma 4.4, (4.4) of Lemma 4.5, the mean value theorem for vector-valued functions, and Lemma 4.2 (i), we get

$$\begin{aligned}
\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}\Big[\zeta^2_{i,k,n}|\widehat{\mathcal{F}}_{t_k}\Big] &\le 2\frac{(\varphi^i_{n\Delta_n}(\theta^0)u_i)^2}{\Delta_n^2}\sum_{k=0}^{n-1}\int_0^1\Big\{\widehat{\mathrm{E}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}\Big[|Z^{5,\ell}_{i,k,n}|^2\Big] \\
&+ \widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}\Big[|R_5^{\theta_i^{0+}(\ell),k}|^2\Big] + C\sqrt{\Delta_n}(1+|X^{\theta^0}_{t_k}|)\Big(\widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}\Big[|R_5^{\theta_i^{0+}(\ell),k}|^4\Big]\Big)^{\frac{1}{2}} \\
&\times\Big(\varphi^i_{n\Delta_n}(\theta^0)\big(1+\sqrt{\Delta_n}\varphi^i_{n\Delta_n}(\theta^0)|X^{\theta^0}_{t_k}|\big)e^{C_i(\varphi^i_{n\Delta_n}(\theta^0))^2\Delta_n|X^{\theta^0}_{t_k}|^2} \\
&+ \varphi^{i+1}_{n\Delta_n}(\theta^0)\big(1+\sqrt{\Delta_n}(\varphi^i_{n\Delta_n}(\theta^0)+\varphi^{i+1}_{n\Delta_n}(\theta^0))|X^{\theta^0}_{t_k}|\big) \\
&\times e^{C_{i+1}((\varphi^i_{n\Delta_n}(\theta^0))^2+(\varphi^{i+1}_{n\Delta_n}(\theta^0))^2)\Delta_n|X^{\theta^0}_{t_k}|^2} + \cdots \\
&+ \varphi^m_{n\Delta_n}(\theta^0)\big(1+\sqrt{\Delta_n}(\varphi^i_{n\Delta_n}(\theta^0)+\dots+\varphi^m_{n\Delta_n}(\theta^0))|X^{\theta^0}_{t_k}|\big) \\
&\times e^{C_m((\varphi^i_{n\Delta_n}(\theta^0))^2+\dots+(\varphi^m_{n\Delta_n}(\theta^0))^2)\Delta_n|X^{\theta^0}_{t_k}|^2}\Big)\Big\}d\ell \\
\le C\Delta_n^2(\varphi^i_{n\Delta_n}(\theta^0))^2\sum_{k=0}^{n-1}(1+|X^{\theta^0}_{t_k}|^4) &+ C\Delta_n^{\frac{5}{2}}(\varphi^i_{n\Delta_n}(\theta^0))^3\sum_{k=0}^{n-1}\Big(1+|X^{\theta^0}_{t_k}|^5
\end{aligned}$$

$$
\begin{aligned}
&+\sqrt{\Delta_n}\varphi^i_{n\Delta_n}(\theta^0)|X^{\theta^0}_{t_k}|^6\Big)e^{C_i(\varphi^i_{n\Delta_n}(\theta^0))^2\Delta_n|X^{\theta^0}_{t_k}|^2}+C\Delta_n^{\frac{5}{2}}(\varphi^i_{n\Delta_n}(\theta^0))^2\\
&\times\varphi^{i+1}_{n\Delta_n}(\theta^0)\sum_{k=0}^{n-1}\Big(1+|X^{\theta^0}_{t_k}|^5+\sqrt{\Delta_n}(\varphi^i_{n\Delta_n}(\theta^0)+\varphi^{i+1}_{n\Delta_n}(\theta^0))|X^{\theta^0}_{t_k}|^6\Big)\\
&\times e^{C_{i+1}((\varphi^i_{n\Delta_n}(\theta^0))^2+(\varphi^{i+1}_{n\Delta_n}(\theta^0))^2)\Delta_n|X^{\theta^0}_{t_k}|^2}+\cdots+C\Delta_n^{\frac{5}{2}}(\varphi^i_{n\Delta_n}(\theta^0))^2\\
&\times\varphi^{m}_{n\Delta_n}(\theta^0)\sum_{k=0}^{n-1}\Big(1+|X^{\theta^0}_{t_k}|^5+\sqrt{\Delta_n}(\varphi^i_{n\Delta_n}(\theta^0)+\cdots+\varphi^{m}_{n\Delta_n}(\theta^0))|X^{\theta^0}_{t_k}|^6\Big)\\
&\times e^{C_m((\varphi^i_{n\Delta_n}(\theta^0))^2+\cdots+(\varphi^{m}_{n\Delta_n}(\theta^0))^2)\Delta_n|X^{\theta^0}_{t_k}|^2}. \qquad (6.7)
\end{aligned}
$$

Thus, using Young's inequality with $\frac{1}{p_0}+\frac{1}{q_0}=1$ and $p_0$ close to 1, Lemma 4.3 and **(A7)**,

$$
\begin{aligned}
&\widehat{\mathrm{E}}^{\theta^0}\Big[\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}\Big[\zeta^2_{i,k,n}|\widehat{\mathcal{F}}_{t_k}\Big]\Big]\leq C\Delta_n^2(\varphi^i_{n\Delta_n}(\theta^0))^2\Big(n+\sum_{k=0}^{n-1}\psi^{-4}_{t_k}(\theta^0)\Big)\\
&+C\Delta_n^{\frac{5}{2}}(\varphi^i_{n\Delta_n}(\theta^0))^3\Big(n+\sum_{k=0}^{n-1}\psi^{-5p_0}_{t_k}(\theta^0)+(\sqrt{\Delta_n}\varphi^i_{n\Delta_n}(\theta^0))^{p_0}\sum_{k=0}^{n-1}\psi^{-6p_0}_{t_k}(\theta^0)\Big)\\
&+C\Delta_n^{\frac{5}{2}}(\varphi^i_{n\Delta_n}(\theta^0))^2\varphi^{i+1}_{n\Delta_n}(\theta^0)\Big(n+\sum_{k=0}^{n-1}\psi^{-5p_0}_{t_k}(\theta^0)\\
&+(\sqrt{\Delta_n}(\varphi^i_{n\Delta_n}(\theta^0)+\varphi^{i+1}_{n\Delta_n}(\theta^0))^{p_0}\sum_{k=0}^{n-1}\psi^{-6p_0}_{t_k}(\theta^0)\Big)\\
&+\cdots+C\Delta_n^{\frac{5}{2}}(\varphi^i_{n\Delta_n}(\theta^0))^2\varphi^{m}_{n\Delta_n}(\theta^0)\Big(n+\sum_{k=0}^{n-1}\psi^{-5p_0}_{t_k}(\theta^0)\\
&+(\sqrt{\Delta_n}(\varphi^i_{n\Delta_n}(\theta^0)+\cdots+\varphi^{m}_{n\Delta_n}(\theta^0))^{p_0}\sum_{k=0}^{n-1}\psi^{-6p_0}_{t_k}(\theta^0)\Big),
\end{aligned}
$$

which, by condition **(A6)**, tends to zero. Thus, $\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}[\zeta^2_{i,k,n}|\widehat{\mathcal{F}}_{t_k}]\xrightarrow{\widehat{\mathrm{P}}^{\theta^0}}0$. Therefore, by Lemma 4.1 a), $\sum_{k=0}^{n-1}\zeta_{i,k,n}\xrightarrow{\widehat{\mathrm{P}}^{\theta^0}}0$ for any $i\in\{1,\ldots,m\}$. Thus, the result follows. □

### 6.7. Proof of Lemma 4.11

*Proof.* We write

$$
\begin{aligned}
\zeta_{k,n}:&=\frac{\varphi^i_{n\Delta_n}(\theta^0)u_i}{\Delta_n}\int_0^1\Big\{Z^{4,\ell}_{i,k,n}-\widetilde{\mathrm{E}}^{\theta^{0+}_i(\ell)}_{t_k,X^{\theta^0}_{t_k}}\Big[R_4^{\theta^{0+}_i(\ell),k}\big|Y^{\theta^{0+}_i(\ell)}_{t_{k+1}}=X^{\theta^0}_{t_{k+1}}\Big]\Big\}d\ell\\
&=\zeta_{i,1,k,n}+\zeta_{i,2,k,n},
\end{aligned}
$$

where

$$\zeta_{i,1,k,n} = \varphi^i_{n\Delta_n}(\theta^0)u_i \int_0^1 (\partial_{\theta_i} b(\theta_i^{0+}(\ell), t_k, X_{t_k}^{\theta^0}))^* \sigma^{-2}(t_k, X_{t_k}^{\theta^0}) M_{i,1,k,n} d\ell,$$

$$\zeta_{i,2,k,n} = \varphi^i_{n\Delta_n}(\theta^0)u_i \int_0^1 (\partial_{\theta_i} b(\theta_i^{0+}(\ell), t_k, X_{t_k}^{\theta^0}))^* \sigma^{-2}(t_k, X_{t_k}^{\theta^0}) M_{i,2,k,n} d\ell,$$

$$\begin{aligned} M_{i,1,k,n} = \int_{t_k}^{t_{k+1}} \Big( & b(\theta^0, s, X_s^{\theta^0}) - b(\theta^0, t_k, X_{t_k}^{\theta^0}) \\ & - \big(b(\theta_i^{0+}(\ell), s, X_s^{\theta^0}) - b(\theta_i^{0+}(\ell), t_k, X_{t_k}^{\theta^0})\big)\Big) ds, \end{aligned}$$

$$\begin{aligned} M_{i,2,k,n} = \int_{t_k}^{t_{k+1}} & \Big( b(\theta_i^{0+}(\ell), s, X_s^{\theta^0}) - b(\theta_i^{0+}(\ell), t_k, X_{t_k}^{\theta^0})\Big) ds \\ & - \widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k, X_{t_k}^{\theta^0}} \Big[ \int_{t_k}^{t_{k+1}} \big( b(\theta_i^{0+}(\ell), s, Y_s^{\theta_i^{0+}(\ell)}) \\ & - b(\theta_i^{0+}(\ell), t_k, Y_{t_k}^{\theta_i^{0+}(\ell)})\big) ds \big| Y_{t_{k+1}}^{\theta_i^{0+}(\ell)} = X_{t_{k+1}}^{\theta^0} \Big]. \end{aligned}$$

First, we treat $\zeta_{i,1,k,n}$. For this, using the mean value theorem, we write

$$\begin{aligned} & b(\theta^0, s, X_s^{\theta^0}) - b(\theta_i^{0+}(\ell), s, X_s^{\theta^0}) = b(\theta_{i+1}^{0+}, s, X_s^{\theta^0}) - b(\theta_i^{0+}(\ell), s, X_s^{\theta^0}) \\ & \quad + b(\theta_{i+2}^{0+}, s, X_s^{\theta^0}) - b(\theta_{i+1}^{0+}, s, X_s^{\theta^0}) + \cdots + b(\theta_m^{0+}, s, X_s^{\theta^0}) - b(\theta_{m-1}^{0+}, s, X_s^{\theta^0}) \\ & \quad + b(\theta^0, s, X_s^{\theta^0}) - b(\theta_m^{0+}, s, X_s^{\theta^0}) \\ & = -\ell \varphi^i_{n\Delta_n}(\theta^0) u_i \int_0^1 \partial_{\theta_i} b(\theta_i^{0+}(\alpha\ell), s, X_s^{\theta^0}) d\alpha \\ & \quad - \varphi^{i+1}_{n\Delta_n}(\theta^0) u_{i+1} \int_0^1 \partial_{\theta_{i+1}} b(\theta_{i+1}^{0+}(\alpha), s, X_s^{\theta^0}) d\alpha \\ & \quad - \cdots - \varphi^m_{n\Delta_n}(\theta^0) u_m \int_0^1 \partial_{\theta_m} b(\theta_m^{0+}(\alpha), s, X_s^{\theta^0}) d\alpha. \end{aligned}$$

Therefore,

$$\begin{aligned} & b(\theta^0, s, X_s^{\theta^0}) - b(\theta^0, t_k, X_{t_k}^{\theta^0}) - \big(b(\theta_i^{0+}(\ell), s, X_s^{\theta^0}) - b(\theta_i^{0+}(\ell), t_k, X_{t_k}^{\theta^0})\big) \\ & = b(\theta^0, s, X_s^{\theta^0}) - b(\theta_i^{0+}(\ell), s, X_s^{\theta^0}) - \big(b(\theta^0, t_k, X_{t_k}^{\theta^0}) - b(\theta_i^{0+}(\ell), t_k, X_{t_k}^{\theta^0})\big) \\ & = -\ell \varphi^i_{n\Delta_n}(\theta^0) u_i \int_0^1 \big(\partial_{\theta_i} b(\theta_i^{0+}(\alpha\ell), s, X_s^{\theta^0}) - \partial_{\theta_i} b(\theta_i^{0+}(\alpha\ell), t_k, X_{t_k}^{\theta^0})\big) d\alpha \\ & - \varphi^{i+1}_{n\Delta_n}(\theta^0) u_{i+1} \int_0^1 \big(\partial_{\theta_{i+1}} b(\theta_{i+1}^{0+}(\alpha), s, X_s^{\theta^0}) - \partial_{\theta_{i+1}} b(\theta_{i+1}^{0+}(\alpha), t_k, X_{t_k}^{\theta^0})\big) d\alpha \\ & - \cdots - \varphi^m_{n\Delta_n}(\theta^0) u_m \int_0^1 \big(\partial_{\theta_m} b(\theta_m^{0+}(\alpha), s, X_s^{\theta^0}) - \partial_{\theta_m} b(\theta_m^{0+}(\alpha), t_k, X_{t_k}^{\theta^0})\big) d\alpha. \end{aligned}$$

Then, using the mean value theorem for vector-valued functions, we write

$$\partial_{\theta_j} b(\theta_j^{0+}(\alpha), s, X_s^{\theta^0}) - \partial_{\theta_j} b(\theta_j^{0+}(\alpha), t_k, X_{t_k}^{\theta^0})$$

$$= \Big( \int_0^1 J_{\partial_{\theta_j} b}(\theta_j^{0+}(\alpha), t_k + v(s-t_k), X_{t_k}^{\theta^0} + v(X_s^{\theta^0} - X_{t_k}^{\theta^0}))dv \Big) \cdot \begin{pmatrix} s - t_k \\ X_s^{\theta^0} - X_{t_k}^{\theta^0} \end{pmatrix},$$

for all $j \in \{i, \ldots, m\}$, where the Jacobian matrix is given by

$$J_{\partial_{\theta_j} b} = \begin{pmatrix} \partial^2_{\theta_j t} b_1 & \partial^2_{\theta_j x_1} b_1 & \ldots & \partial^2_{\theta_j x_d} b_1 \\ \vdots & \vdots & \ddots & \vdots \\ \partial^2_{\theta_j t} b_d & \partial^2_{\theta_j x_1} b_d & \ldots & \partial^2_{\theta_j x_d} b_d \end{pmatrix}.$$

This, combined with **(A2)-(A3)**, Lemma 4.2 (i) and Lemma 4.3, we obtain

$$\widehat{\mathrm{E}}^{\theta^0}\Big[ \sum_{k=0}^{n-1} \widehat{\mathrm{E}}^{\theta^0}\Big[ |\zeta_{i,1,k,n}| | \widehat{\mathcal{F}}_{t_k} \Big]\Big] \leq C\Delta_n^{\frac{3}{2}} \varphi^i_{n\Delta_n}(\theta^0) \big(\varphi^i_{n\Delta_n}(\theta^0) + \cdots + \varphi^m_{n\Delta_n}(\theta^0)\big)\big(n + \sum_{k=0}^{n-1} \psi_{t_k}^{-2}(\theta^0)\big),$$

which, by **(A6)**, tends to zero. Thus, by Lemma 4.1 b), $\sum_{k=0}^{n-1} \zeta_{i,1,k,n} \xrightarrow{\widehat{\mathrm{P}}^{\theta^0}} 0$.

Next, we apply Lemma 4.1 a) to $\zeta_{i,2,k,n}$. To simplify the exposition, we set

$$\widehat{V}^{\theta^0} = (\partial_{\theta_i} b(\theta_i^{0+}(\ell), t_k, X_{t_k}^{\theta^0}))^* \sigma^{-2}(t_k, X_{t_k}^{\theta^0}) \cdot \int_{t_k}^{t_{k+1}} \Big( b(\theta_i^{0+}(\ell), s, X_s^{\theta^0}) - b(\theta_i^{0+}(\ell), t_k, X_{t_k}^{\theta^0})\Big) ds,$$

$$V^{\theta_i^{0+}(\ell)} = (\partial_{\theta_i} b(\theta_i^{0+}(\ell), t_k, X_{t_k}^{\theta^0}))^* \sigma^{-2}(t_k, X_{t_k}^{\theta^0}) \cdot \int_{t_k}^{t_{k+1}} \Big( b(\theta_i^{0+}(\ell), s, Y_s^{\theta_i^{0+}(\ell)}) - b(\theta_i^{0+}(\ell), t_k, Y_{t_k}^{\theta_i^{0+}(\ell)})\Big) ds,$$

Then, using (4.3) and (4.2) of Lemma 4.4 with $\widehat{V}^{\theta^0}$ and $V^{\theta_i^{0+}(\ell)}$, (4.5) and (4.4) of Lemma 4.5 with $q = 2$, and Lemma 4.2 (i), $|\sum_{k=0}^{n-1} \widehat{\mathrm{E}}^{\theta^0}[\zeta_{i,2,k,n} | \widehat{\mathcal{F}}_{t_k}]|$ is bounded by

$$C\varphi^i_{n\Delta_n}(\theta^0)|u_i| \sum_{k=0}^{n-1} \int_0^1 \sqrt{\Delta_n}(1 + |X_{t_k}^{\theta^0}|) \times \Big\{ \Big| \int_{\theta_i^0 + \ell\varphi^i_{n\Delta_n}(\theta^0)u_i}^{\theta_i^0} \big(\widehat{\mathrm{E}}^{\theta_i(0+)}_{t_k, X_{t_k}^{\theta^0}}[|\widehat{V}^{\theta_i(0+)}|^2]\big)^{\frac{1}{2}} d\theta_i \Big| + \Big| \int_{\theta_{i+1}^0 + \varphi^{i+1}_{n\Delta_n}(\theta^0)u_{i+1}}^{\theta_{i+1}^0} \big(\widehat{\mathrm{E}}^{\theta_{i+1}(0+)}_{t_k, X_{t_k}^{\theta^0}}[|\widehat{V}^{\theta_{i+1}(0+)}|^2]\big)^{\frac{1}{2}} d\theta_{i+1} \Big| + \cdots + \Big| \int_{\theta_m^0 + \varphi^m_{n\Delta_n}(\theta^0)u_m}^{\theta_m^0} \big(\widehat{\mathrm{E}}^{\theta_m(0+)}_{t_k, X_{t_k}^{\theta^0}}[|\widehat{V}^{\theta_m(0+)}|^2]\big)^{\frac{1}{2}} d\theta_m \Big| + \big(\widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k, X_{t_k}^{\theta^0}}[|V^{\theta_i^{0+}(\ell)}|^2]\big)^{\frac{1}{2}}$$

$$
\begin{aligned}
&\times \Big(\varphi^i_{n\Delta_n}(\theta^0)\big(1+\sqrt{\Delta_n}\varphi^i_{n\Delta_n}(\theta^0)|X^{\theta^0}_{t_k}|\big)e^{C_i(\varphi^i_{n\Delta_n}(\theta^0))^2\Delta_n|X^{\theta^0}_{t_k}|^2}\\
&+\varphi^{i+1}_{n\Delta_n}(\theta^0)\big(1+\sqrt{\Delta_n}(\varphi^i_{n\Delta_n}(\theta^0)+\varphi^{i+1}_{n\Delta_n}(\theta^0))|X^{\theta^0}_{t_k}|\big)\\
&\times e^{C_{i+1}((\varphi^i_{n\Delta_n}(\theta^0))^2+(\varphi^{i+1}_{n\Delta_n}(\theta^0))^2)\Delta_n|X^{\theta^0}_{t_k}|^2}+\cdots\\
&+\varphi^{m}_{n\Delta_n}(\theta^0)\big(1+\sqrt{\Delta_n}(\varphi^i_{n\Delta_n}(\theta^0)+\cdots+\varphi^{m}_{n\Delta_n}(\theta^0))|X^{\theta^0}_{t_k}|\big)\\
&\times e^{C_m((\varphi^i_{n\Delta_n}(\theta^0))^2+\cdots+(\varphi^{m}_{n\Delta_n}(\theta^0))^2)\Delta_n|X^{\theta^0}_{t_k}|^2}\Big)\Big\}d\ell\\
&\leq C\Delta_n^2\varphi^i_{n\Delta_n}(\theta^0)\Big(\varphi^i_{n\Delta_n}(\theta^0)+\varphi^{i+1}_{n\Delta_n}(\theta^0)+\cdots+\varphi^{m}_{n\Delta_n}(\theta^0)\Big)\sum_{k=0}^{n-1}\big(1+|X^{\theta^0}_{t_k}|^3\big)\\
&+C\Delta_n^2(\varphi^i_{n\Delta_n}(\theta^0))^2\sum_{k=0}^{n-1}\Big(1+|X^{\theta^0}_{t_k}|^3+\sqrt{\Delta_n}\varphi^i_{n\Delta_n}(\theta^0)|X^{\theta^0}_{t_k}|^4\Big)\\
&\times e^{C_i(\varphi^i_{n\Delta_n}(\theta^0))^2\Delta_n|X^{\theta^0}_{t_k}|^2}+C\Delta_n^2\varphi^i_{n\Delta_n}(\theta^0)\varphi^{i+1}_{n\Delta_n}(\theta^0)\sum_{k=0}^{n-1}\Big(1+|X^{\theta^0}_{t_k}|^3\\
&+\sqrt{\Delta_n}(\varphi^i_{n\Delta_n}(\theta^0)+\varphi^{i+1}_{n\Delta_n}(\theta^0))|X^{\theta^0}_{t_k}|^4\Big)e^{C_{i+1}((\varphi^i_{n\Delta_n}(\theta^0))^2+(\varphi^{i+1}_{n\Delta_n}(\theta^0))^2)\Delta_n|X^{\theta^0}_{t_k}|^2}\\
&+\cdots+C\Delta_n^2\varphi^i_{n\Delta_n}(\theta^0)\varphi^{m}_{n\Delta_n}(\theta^0)\sum_{k=0}^{n-1}\Big(1+|X^{\theta^0}_{t_k}|^3+\sqrt{\Delta_n}(\varphi^i_{n\Delta_n}(\theta^0)+\cdots\\
&+\varphi^{m}_{n\Delta_n}(\theta^0))|X^{\theta^0}_{t_k}|^4\Big)e^{C_m((\varphi^i_{n\Delta_n}(\theta^0))^2+\cdots+(\varphi^{m}_{n\Delta_n}(\theta^0))^2)\Delta_n|X^{\theta^0}_{t_k}|^2}.
\end{aligned}
$$

Here, we have used $\widehat{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}[\widehat{V}^{\theta_i^{0+}(\ell)}]=\widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}[V^{\theta_i^{0+}(\ell)}]$ due to the fact that $Y^{\theta_i^{0+}(\ell)}$ is the independent copy of $X^{\theta_i^{0+}(\ell)}$ and the mean value theorem for vector-valued functions to write

$$
\begin{aligned}
&b(\theta_i^{0+}(\ell),s,Y_s^{\theta_i^{0+}(\ell)})-b(\theta_i^{0+}(\ell),t_k,Y_{t_k}^{\theta_i^{0+}(\ell)})\\
&=\Big(\int_0^1 J_b(\theta_i^{0+}(\ell),t_k+v(s-t_k),Y_{t_k}^{\theta_i^{0+}(\ell)}+v(Y_s^{\theta_i^{0+}(\ell)}-Y_{t_k}^{\theta_i^{0+}(\ell)}))dv\Big)\\
&\quad\cdot\begin{pmatrix}s-t_k\\ Y_s^{\theta_i^{0+}(\ell)}-Y_{t_k}^{\theta_i^{0+}(\ell)}\end{pmatrix},
\end{aligned}
$$

where the Jacobian matrix is given by

$$
J_b=\begin{pmatrix}\partial_t b_1 & \partial_{x_1}b_1 & \dots & \partial_{x_d}b_1\\ \vdots & \vdots & \ddots & \vdots\\ \partial_t b_d & \partial_{x_1}b_d & \dots & \partial_{x_d}b_d\end{pmatrix}.
$$

Then, using Young's inequality with $\frac{1}{p_0}+\frac{1}{q_0}=1$ and $p_0$ close to 1, Lemma 4.3

and **(A7)**,

$$\begin{aligned}
&\widehat{\mathrm{E}}^{\theta^0}\Big[\Big|\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}\Big[\zeta_{i,2,k,n}|\widehat{\mathcal{F}}_{t_k}\Big]\Big|\Big]\\
&\leq C\Delta_n^2\varphi^i_{n\Delta_n}(\theta^0)\big(\varphi^i_{n\Delta_n}(\theta^0)+\cdots+\varphi^m_{n\Delta_n}(\theta^0)\big)\Big(n+\sum_{k=0}^{n-1}\psi_{t_k}^{-3}(\theta^0)\Big)\\
&+C\Delta_n^2(\varphi^i_{n\Delta_n}(\theta^0))^2\Big(n+\sum_{k=0}^{n-1}\psi_{t_k}^{-3p_0}(\theta^0)+(\sqrt{\Delta_n}\varphi^i_{n\Delta_n}(\theta^0))^{p_0}\sum_{k=0}^{n-1}\psi_{t_k}^{-4p_0}(\theta^0)\Big)\\
&+C\Delta_n^2\varphi^i_{n\Delta_n}(\theta^0)\varphi^{i+1}_{n\Delta_n}(\theta^0)\Big(n+\sum_{k=0}^{n-1}\psi_{t_k}^{-3p_0}(\theta^0)+(\sqrt{\Delta_n}(\varphi^i_{n\Delta_n}(\theta^0)\\
&+\varphi^{i+1}_{n\Delta_n}(\theta^0))^{p_0}\sum_{k=0}^{n-1}\psi_{t_k}^{-4p_0}(\theta^0)\Big)+\cdots+C\Delta_n^2\varphi^i_{n\Delta_n}(\theta^0)\varphi^m_{n\Delta_n}(\theta^0)\\
&\times\Big(n+\sum_{k=0}^{n-1}\psi_{t_k}^{-3p_0}(\theta^0)+(\sqrt{\Delta_n}(\varphi^i_{n\Delta_n}(\theta^0)+\cdots+\varphi^m_{n\Delta_n}(\theta^0))^{p_0}\sum_{k=0}^{n-1}\psi_{t_k}^{-4p_0}(\theta^0)\Big),
\end{aligned}$$

which, by condition **(A6)**, tends to zero. Therefore, $\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}[\zeta_{i,2,k,n}|\widehat{\mathcal{F}}_{t_k}]\xrightarrow{\widehat{\mathrm{P}}^{\theta^0}}0$ as $n\to\infty$.

Next, applying Jensen's inequality, (4.2) of Lemma 4.4, (4.4) of Lemma 4.5, the mean value theorem for vector-valued functions, and Lemma 4.2 (i), $\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}[\zeta_{i,2,k,n}^2|\widehat{\mathcal{F}}_{t_k}]$ is bounded by

$$\begin{aligned}
&2(\varphi^i_{n\Delta_n}(\theta^0)u_i)^2\sum_{k=0}^{n-1}\int_0^1\Big\{\widehat{\mathrm{E}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}\Big[|\widehat{V}^{\theta^0}|^2\Big]+\widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}\Big[|V^{\theta_i^{0+}(\ell)}|^2\Big]\\
&\quad+C\sqrt{\Delta_n}(1+|X^{\theta^0}_{t_k}|)\Big(\widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}\Big[|V^{\theta_i^{0+}(\ell)}|^4\Big]\Big)^{\frac12}\\
&\quad\times\Big(\varphi^i_{n\Delta_n}(\theta^0)\big(1+\sqrt{\Delta_n}\varphi^i_{n\Delta_n}(\theta^0)|X^{\theta^0}_{t_k}|\big)e^{C_i(\varphi^i_{n\Delta_n}(\theta^0))^2\Delta_n|X^{\theta^0}_{t_k}|^2}\\
&\quad+\varphi^{i+1}_{n\Delta_n}(\theta^0)\big(1+\sqrt{\Delta_n}(\varphi^i_{n\Delta_n}(\theta^0)+\varphi^{i+1}_{n\Delta_n}(\theta^0))|X^{\theta^0}_{t_k}|\big)\\
&\quad\times e^{C_{i+1}((\varphi^i_{n\Delta_n}(\theta^0))^2+(\varphi^{i+1}_{n\Delta_n}(\theta^0))^2)\Delta_n|X^{\theta^0}_{t_k}|^2}+\cdots\\
&\quad+\varphi^m_{n\Delta_n}(\theta^0)\big(1+\sqrt{\Delta_n}(\varphi^i_{n\Delta_n}(\theta^0)+\cdots+\varphi^m_{n\Delta_n}(\theta^0))|X^{\theta^0}_{t_k}|\big)\\
&\quad\times e^{C_m((\varphi^i_{n\Delta_n}(\theta^0))^2+\cdots+(\varphi^m_{n\Delta_n}(\theta^0))^2)\Delta_n|X^{\theta^0}_{t_k}|^2}\Big)\Big\}d\ell\\
&\leq C\Delta_n^3(\varphi^i_{n\Delta_n}(\theta^0))^2\sum_{k=0}^{n-1}(1+|X^{\theta^0}_{t_k}|^4)+C\Delta_n^{\frac72}(\varphi^i_{n\Delta_n}(\theta^0))^3\sum_{k=0}^{n-1}\Big(1+|X^{\theta^0}_{t_k}|^5\\
&\quad+\sqrt{\Delta_n}\varphi^i_{n\Delta_n}(\theta^0)|X^{\theta^0}_{t_k}|^6\Big)e^{C_i(\varphi^i_{n\Delta_n}(\theta^0))^2\Delta_n|X^{\theta^0}_{t_k}|^2}+C\Delta_n^{\frac72}(\varphi^i_{n\Delta_n}(\theta^0))^2
\end{aligned}$$

$$\times\, \varphi^{i+1}_{n\Delta_n}(\theta^0) \sum_{k=0}^{n-1} \Big(1 + |X^{\theta^0}_{t_k}|^5 + \sqrt{\Delta_n}(\varphi^i_{n\Delta_n}(\theta^0) + \varphi^{i+1}_{n\Delta_n}(\theta^0))|X^{\theta^0}_{t_k}|^6\Big)$$
$$\times\, e^{C_{i+1}((\varphi^i_{n\Delta_n}(\theta^0))^2 + (\varphi^{i+1}_{n\Delta_n}(\theta^0))^2)\Delta_n |X^{\theta^0}_{t_k}|^2} + \cdots + C\Delta_n^{\frac{7}{2}}(\varphi^i_{n\Delta_n}(\theta^0))^2$$
$$\times\, \varphi^{m}_{n\Delta_n}(\theta^0) \sum_{k=0}^{n-1} \Big(1 + |X^{\theta^0}_{t_k}|^5 + \sqrt{\Delta_n}(\varphi^i_{n\Delta_n}(\theta^0) + \cdots + \varphi^{m}_{n\Delta_n}(\theta^0))|X^{\theta^0}_{t_k}|^6\Big)$$
$$\times\, e^{C_{m}((\varphi^i_{n\Delta_n}(\theta^0))^2 + \cdots + (\varphi^{m}_{n\Delta_n}(\theta^0))^2)\Delta_n |X^{\theta^0}_{t_k}|^2}.$$

Then, proceeding as for the term (6.7), we get

$$\widehat{\mathrm{E}}^{\theta^0}\Big[\sum_{k=0}^{n-1} \widehat{\mathrm{E}}^{\theta^0}\Big[\zeta^2_{i,2,k,n}|\widehat{\mathcal{F}}_{t_k}\Big]\Big] \le C\Delta_n^3(\varphi^i_{n\Delta_n}(\theta^0))^2\Big(n + \sum_{k=0}^{n-1}\psi^{-4}_{t_k}(\theta^0)\Big)$$
$$+ C\Delta_n^{\frac{7}{2}}(\varphi^i_{n\Delta_n}(\theta^0))^3\Big(n + \sum_{k=0}^{n-1}\psi^{-5p_0}_{t_k}(\theta^0) + (\sqrt{\Delta_n}\varphi^i_{n\Delta_n}(\theta^0))^{p_0}\sum_{k=0}^{n-1}\psi^{-6p_0}_{t_k}(\theta^0)\Big)$$
$$+ C\Delta_n^{\frac{7}{2}}(\varphi^i_{n\Delta_n}(\theta^0))^2\varphi^{i+1}_{n\Delta_n}(\theta^0)$$
$$\times\Big(n + \sum_{k=0}^{n-1}\psi^{-5p_0}_{t_k}(\theta^0) + (\sqrt{\Delta_n}(\varphi^i_{n\Delta_n}(\theta^0) + \varphi^{i+1}_{n\Delta_n}(\theta^0))^{p_0}\sum_{k=0}^{n-1}\psi^{-6p_0}_{t_k}(\theta^0)\Big)$$
$$+ \cdots + C\Delta_n^{\frac{7}{2}}(\varphi^i_{n\Delta_n}(\theta^0))^2\varphi^{m}_{n\Delta_n}(\theta^0)$$
$$\times\Big(n + \sum_{k=0}^{n-1}\psi^{-5p_0}_{t_k}(\theta^0) + (\sqrt{\Delta_n}(\varphi^i_{n\Delta_n}(\theta^0) + \cdots + \varphi^{m}_{n\Delta_n}(\theta^0))^{p_0}\sum_{k=0}^{n-1}\psi^{-6p_0}_{t_k}(\theta^0)\Big),$$

which, by condition **(A6)**, tends to zero. Thus, $\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}[\zeta^2_{i,2,k,n}|\widehat{\mathcal{F}}_{t_k}] \xrightarrow{\widehat{\mathrm{P}}^{\theta^0}} 0$. Therefore, by Lemma 4.1 a), $\sum_{k=0}^{n-1}\zeta_{i,2,k,n} \xrightarrow{\widehat{\mathrm{P}}^{\theta^0}} 0$ for any $i \in \{1,\dots,m\}$. Thus, the result follows. □

## Acknowledgments

The authors acknowledge support from the Vietnam Institute for Advanced Study in Mathematics (VIASM). H-L. Ngo would like to thank the Hanoi National University of Education for providing a fruitful working environment.

## References

[1] Aït-Sahalia, Y. (2002), Maximum likelihood estimation of discretely sampled diffusions: A closed-form approximation approach, *Econometrica*, **70**(1), 223–262. MR1926260
[2] Basawa, I.V. and Prakasa Rao, B.L.S. (1980), *Statistical Inference for Stochastic Processes*, London: Academic Press. MR0586053

[3] Barczy, M. and Pap, G. (2010), Asymptotic behavior of maximum likelihood estimator for time inhomogeneous diffusion processes, *J. Statist. Plann. Inference*, **140**, 1576–1593. MR2592235

[4] Barndorff-Nielsen, O. E. and Sørensen, M. (1994), A review of some aspects of asymptotic likelihood theory for stochastic processes, *Int. Stoch. Review*, **62**, 133–165.

[5] Ben Alaya, M., Kebaier, A. and Tran, N. K. (2020), Local asymptotic properties for Cox-Ingersoll-Ross process with discrete observations, *Scand. J. Statist.*, **47**(4), 1401–1464. MR4178198

[6] Benke, J.M. and Pap, G. (2017), Local asymptotic quadraticity of statistical experiments connected with a Heston model, *Acta Sci. Math. (Szeged)*, **83**, 313–344. MR3701048

[7] Bishwal, J.P.N. (2008), *Parameter Estimation in Stochastic Differential Equations*, Lecture Notes in Mathematics, Springer. MR2360279

[8] Dietz, H. M. and Kutoyants, Y.A. (2003), Parameter estimation for some non-recurrent solutions of SDE, *Statistics & Decisions*, **21**, 29–45. MR1985650

[9] Florens-Zmirou, D. (1989), Approximate discrete-time schemes for statistics of diffusion processes, *Statistics*, **20**(4), 547–557. MR1047222

[10] Genon-Catalot, V. and Jacod, J. (1993), On the estimation of the diffusion coefficient for multi-dimensional diffusion processes, *Ann. Inst. Henri Poincaré Probab. Stat.*, **29**, 119–151. MR1204521

[11] Gobet, E. (2001), Local asymptotic mixed normality property for elliptic diffusions: a Malliavin calculus approach, *Bernoulli*, **7**, 899–912. MR1873834

[12] Gobet, E. (2002), LAN property for ergodic diffusions with discrete observations, *Ann. Inst. Henri Poincaré Probab. Stat.*, **38**, 711–737. MR1931584

[13] Hájek, J. (1972), Local asymptotic minimax and admissibility in estimation, *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. I: Theory of statistics*, 175–194. MR0400513

[14] Höpfner, R. (2014), *Asymptotic statistics*, De Gruyter, Berlin. MR3185373

[15] Höpfner, R. and Kutoyants, Yu. (2003), On a Problem of Statistical Inference in Null Recurrent Diffusions, *Stat. Inference Stoch. Process.*, **6**(1), 25–42. MR1965183

[16] Jacod, J. (2006), Parametric inference for discretely observed non-ergodic diffusions, *Bernoulli*, **12**(3), 383–401. MR2232724

[17] Jacod, J. (2012), *Statistics and High Frequency Data*, In: M. Kessler, A. Lindner, and M. Sørensen (Eds.), Statistical Methods for Stochastic Differential Equations, Chapman & Hall/CRC Monographs on Statistics and Applied Probability, Volume 124. MR2976984

[18] Jacod, J. and Shiryaev, A. N. (2003), *Limit Theorems for Stochastic Processes*, Second Edition, Springer-Verlag, Berlin. MR1943877

[19] Jeganathan, P. (1982), On the asymptotic theory of estimation when the limit of the log-likelihood ratios is mixed normal, *Sankhyā Ser. A*, **44**(2),

173–212. MR0688800

[20] Kessler, M. (1997), Estimation of an ergodic diffusion from discrete observations, *Scandinavian J. Statist.*, **24**, 211–229. MR1455868

[21] Kutoyants, Y.A. (2004), *Statistical Inference for Ergodic Diffusion Processes*, Springer Series in Statistics, Springer-Verlag, London. MR2144185

[22] Kunita, H. (1997), *Stochastic Flows and Stochastic Differential Equations. Cambridge Studies in Advanced Mathematics*, Cambridge: Cambridge Univ. Press. MR1472487

[23] Le Cam, L. (1960), Locally asymptotically normal families of distributions, *Univ. California, Publ. Statist*, **3**, 37–98. MR0126903

[24] Le Cam, L. and Lo Yang, G. (1990), *Asymptotics in statistics: Some basic concepts*, Springer Series in Statistics. Springer-Verlag, New York. MR1066869

[25] Liptser, R.S. and Shiryaev, A. N. (1977), *Statistics of Random Processes. I, II*, Springer: New York. MR0474486

[26] Luschgy, H. (1992), Local asymptotic mixed normality for semimartingale experiments, *Probab. Theory Related Fields*, **92**(2), 151–176. MR1161184

[27] Lyons, T.J. and Zheng, W.A. (1990), On conditional diffusion processes, *Proc. Royal Soc. Edinburgh*, **115A**, 243–255. MR1069520

[28] Mishra, M. N. and Prakasa Rao, B.L.S. (1985), Asymptotic study of the maximum likelihood estimator for non-homogeneous diffusion processes, *Statistics & Decisions*, **3**, 193–203. MR0809360

[29] Nualart, D. (2006), *The Malliavin Calculus and Related Topics*, Second Edition, Springer, New York. MR2200233

[30] Overbeck, L. (1998), Estimation for continuous branching processes, *Scand. J. Statist.*, **25**(1), 111–126. MR1614256

[31] Qian, Z. and Wei, G. (1992), *A note on conditional diffusion processes*, In: Control Theory, Stochastic Analysis and Applications - Proceedings of Symposium on System Sciences and Control Theory by S P Chen (Editor) and J. Yong (Editor), World Scientific Publishing Company, Illustrated Edition, 185–192. MR1171674

[32] Shimizu, Y. (2009), Notes on drift estimation for certain non-recurrent diffusion processes from sampled data, *Statistics & Probability Letters*, **79**, 2200–2207. MR2572052

[33] Shimizu, Y. (2012), Local asymptotic mixed normality for discretely observed non-recurrent Ornstein-Uhlenbeck processes, *Ann. Inst. Stat. Math*, **64**, 193-211. MR2878923

[34] Shimizu, Y. (2012), Estimation of parameters for discretely observed diffusion processes with a variety of rates for information, *Ann. Inst. Stat. Math*, **64**, 545–575. MR2880869

[35] Uchida, M. and Yoshida, N. (2012), Adaptive estimation of an ergodic diffusion process based on sampled data, *Stochastic Processes and their Applications*, **122**(8), 2885–2924. MR2931346

[36] Uchida, M. and Yoshida, N. (2014), Adaptive Bayes type estimators of ergodic diffusion processes from discrete observations, *Stat. Inference Stoch. Process.*, **17**(2), 181–219. MR3219528

[37] Van Zanten, H. (2000), A multivariate central limit theorem for continuous local martingales, *Statistics & Probability Letters*, **50**(3), 229–235. MR1792301

[38] Wei, C. and Shu, H. (2016), Maximum likelihood estimation for the drift parameter in diffusion processes, *Stochastics*, **88**(5), 699–710. MR3484038

[39] Yoshida, N. (1992), Estimation for diffusion processes from discrete observation, *J. Multivar. Anal.*, **41**(2), 220–242. MR1172898