LAN PROPERTY FOR THE DRIFT PARAMETER OF ERGODIC DIFFUSIONS WITH JUMPS FROM DISCRETE OBSERVATIONS

NGOC KHUE TRAN

Abstract. We consider a multidimensional diffusion with jumps driven by a Brownian motion and a Poisson random measure associated with a Lévy process without Gaussian component, whose drift coefficient depends on a multidimensional unknown parameter. In continuity with the recent work by Kohatsu-Higa et *al.* [18] where only the case of finite jump activity is studied, in this paper the case of infinite jump activity is next investigated. We prove the local asymptotic normality property from high-frequency discrete observations with increasing observation window by assuming some hypotheses on the coefficients of the equation, the ergodicity of the solution and the integrability of the Lévy measure. To obtain the result, our approach is essentially based on Malliavin calculus techniques initiated by Gobet [7, 8] and a subtle analysis on the jump structure of the Lévy process developed recently by Ben Alaya et *al.* [2].

1. Introduction

On a complete probability space $(\Omega, \mathcal{F}, \mathrm{P})$ which will be specified later on, we consider the d-dimensional process $X^\theta = (X_t^\theta)_{t\geq 0}$ solution to the following stochastic differential equation (SDE) with jumps

$$dX_t^\theta = b(\theta, X_t^\theta)dt + \sigma(X_t^\theta)dB_t + \int_{\mathbb{R}_0^d} c(X_{t-}^\theta, z)\left(N(dt,dz) - \nu(dz)dt\right), \tag{1.1}$$

where $X_0^\theta = x_0 \in \mathbb{R}^d$ is fixed and known, $\mathbb{R}_0^d := \mathbb{R}^d \setminus \{0\}$, $B = (B_t)_{t\geq 0}$ is a d-dimensional Brownian motion, and $N(dt,dz)$ is a Poisson random measure in $(\mathbb{R}_+ \times \mathbb{R}_0^d, \mathcal{B}(\mathbb{R}_+ \times \mathbb{R}_0^d))$ independent of B with intensity measure $\nu(dz)dt$. The Lévy measure $\nu(dz)$ can be finite or infinite. The Poisson random measure $N(dt,dz)$ is associated to a centered Lévy process $Z = (Z_t)_{t\geq 0}$ without Gaussian component, that is, the Lévy-Itô decomposition of Z takes the form $Z_t = \int_0^t \int_{\mathbb{R}_0^d} z\widetilde{N}(ds,dz)$ for any $t \geq 0$, where $\widetilde{N}(dt,dz) := N(dt,dz) - \nu(dz)dt$ denotes the compensated Poisson random measure and $N(dt,dz) := \sum_{0\leq s\leq t} \mathbf{1}_{\{\Delta Z_s \neq 0\}}\delta_{(s,\Delta Z_s)}(ds,dz)$. Here, the jump amplitude of Z is defined as $\Delta Z_s := Z_s - Z_{s-}$ for any $s > 0$, $\Delta Z_0 := 0$, $\delta_{(s,z)}$ denotes the Dirac measure at the point $(s,z) \in \mathbb{R}_+ \times \mathbb{R}_0^d$, and $\mathcal{B}(\mathbb{R}_+ \times \mathbb{R}_0^d)$ denotes the Borel σ-algebra on $\mathbb{R}_+ \times \mathbb{R}_0^d$. The unknown parameter $\theta = (\theta_1, \ldots, \theta_m)$ belongs to Θ, an open

Date: April 15, 2019.

2010 *Mathematics Subject Classification.* 60H07; 60J75; 62F12; 62M05.

Key words and phrases. asymptotic efficiency, diffusions with jumps, local asymptotic normality property, Malliavin calculus.

The author wishes to thank the Vietnam Institute for Advanced Study in Mathematics (VIASM) for the intellectual environment and the hospitality where a part of this work was done during his stay at VIASM. This research is funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 101.03-2017.316.

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 and N. The coefficients $b = (b_1, \ldots, b_d) : \Theta \times \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ and $c : \mathbb{R}^d \times \mathbb{R}_0^d \to \mathbb{R}^d$ are measurable functions satisfying the Lipschitz continuity condition **(A1)** below under which equation (1.1) has a unique $\{\widehat{\mathcal{F}}_t\}_{t\geq 0}$-adapted càdlàg solution X^θ possessing the strong Markov property (see [13, Theorem III.2.32] or [1, Theorems 6.2.9. and 6.4.6.]). We denote by $\widehat{\mathrm{P}}^\theta$ the probability measure induced by the process X^θ on the canonical space $(D(\mathbb{R}_+, \mathbb{R}^d), \mathcal{B}(D(\mathbb{R}_+, \mathbb{R}^d)))$ endowed with the natural filtration $\{\widehat{\mathcal{F}}_t\}_{t\geq 0}$. Here $D(\mathbb{R}_+, \mathbb{R}^d)$ denotes the set of $\mathbb{R}^d$-valued càdlàg functions defined on $\mathbb{R}_+$, and $\mathcal{B}(D(\mathbb{R}_+, \mathbb{R}^d))$ is its Borel σ-algebra. We denote by $\widehat{\mathrm{E}}^\theta$ 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, and $\mathrm{tr}(A)$ denotes the trace. * denotes the transpose.

The class of Lévy-driven SDEs has recently received a lot of attention in various fields of applications such as physics, neurosciences, mathematical finance,.... The statistical study for these SDEs has become an active domain of research. Parameter estimation for diffusion processes from discrete observations can be found, for instance, in [4, 5, 16]. In the case of diffusions with jumps, see [23, 6, 33, 29, 32, 26].

For $\theta \in \Theta$ and $n \in \mathbb{N}^*$, a discrete observation at deterministic and equidistant times $t_k = k\Delta_n$, $k \in \{0, \ldots, n\}$ of the process X^θ solution to (1.1) is denoted by $X^{n,\theta} = (X^\theta_{t_0}, X^\theta_{t_1}, \ldots, X^\theta_{t_n})$. We assume that the sequence of sampling time-step sizes Δ_n satisfies the high-frequency and infinite horizon conditions: $\Delta_n \to 0$ and $n\Delta_n \to \infty$ as $n \to \infty$. Let P_n^θ denote the probability law of the random vector $X^{n,\theta}$.

The Local Asymptotic Normality (LAN) property is a fundamental concept in the asymptotic theory of statistics. This property was introduced by Le Cam [20] and Hájek [9] in the situations where the asymptotic Fisher information matrix is deterministic. In our setting, we say that the LAN property holds at $\theta^0 \in \Theta$ with rate of convergence $\varphi_{n\Delta_n}(\theta^0)$ and asymptotic Fisher information matrix $\Gamma(\theta^0)$ if for any $u \in \mathbb{R}^m$, as $n \to \infty$,

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

where $\mathcal{N}(0, \Gamma(\theta^0))$ is a centered $\mathbb{R}^m$-valued Gaussian random variable with covariance matrix $\Gamma(\theta^0)$. Here, $\Gamma(\theta^0)$ is a symmetric positive definite non-random matrix in $\mathbb{R}^{m\times m}$, $\varphi_{n\Delta_n}(\theta^0)$ is a diagonal matrix in $\mathbb{R}^{m\times m}$ whose diagonal entries tend to zero as $n \to \infty$. Later on, the concept of Local Asymptotic Mixed Normality (LAMN) property was developed by Jeganathan [14] when the asymptotic Fisher information matrix $\Gamma(\theta^0)$ is random. These properties allow to introduce the notion of asymptotically efficient estimators in the sense of Hájek-Le Cam convolution theorem and to give the lower bounds for the variance of estimators (see Jeganathan [14]). Assume that the LAN property holds at point θ^0, on the one hand, a sequence of estimators $(\widehat{\theta}_n)_{n\geq 1}$ of the parameter θ^0 is called asymptotically efficient at θ^0 in the sense of Hájek-Le Cam convolution theorem if as $n \to \infty$,

$$\varphi^{-1}_{n\Delta_n}(\theta^0) \left(\widehat{\theta}_n - \theta^0\right) \xrightarrow{\mathcal{L}(\widehat{\mathrm{P}}^{\theta^0})} \mathcal{N}(0, \Gamma(\theta^0)^{-1}).$$

On the other hand, the minimax theorem states that the lower bound for the asymptotic variance of estimators is given by the Cramer-rao lower bound $\Gamma(\theta^0)^{-1}$. We refer the reader to Subsection 7.1 of Höpfner [10] or Le Cam and Lo Yang [21] for further details.

On the basis of continuous observations, the LAMN property was studied by Luschgy [22] for semimartingale. In the case of discrete observations, the Malliavin calculus approach initiated by Gobet [7, 8] is used to obtain the LAMN and LAN properties. In [7], the author addressed the LAMN property for multidimensional elliptic diffusion processes. Later on, the LAN property was established in [8] for multidimensional ergodic diffusions. More recently, Ben Alaya et *al.* [3] have proved the LAN property in the subcritical case, the local asymptotic quadraticity (LAQ) in the critical case, and the LAMN property in the supercritical case for the Cox-Ingersoll-Ross process. In presence of jumps, several Lévy-driven SDEs have been investigated. More precisely, Kawai [15] studied the LAN property for the ergodic Ornstein-Uhlenbeck processes with jumps by using the fact that the solution and transition density are semi-explicit. See also [17, 34] in the case of a simple Lévy process and an ergodic Ornstein-Uhlenbeck process with Poisson jumps, respectively. Recently, Kohatsu-Higa et *al.* [18] have obtained the LAN property for the SDE with jumps (1.1) in a particular case where the driving Lévy process is a compound Poisson process with finite Lévy measure. More recently, in [2] Ben Alaya et *al.* have studied the local asymptotic properties for the growth rate of a jump-type CIR process driven by a subordinator with a possible infinite jump activity.

To our knowledge, the validity of the LAN property for SDEs (1.1) having a Brownian driver and a more general driving Lévy process with possible infinite Lévy measure has never been addressed in the literature. Thus, the purpose of this paper is to prove the LAN property for the drift parameter of diffusions with jumps (1.1) from discrete observations under some appropriate assumptions on the coefficients of the equation, the ergodicity of the solution and the integrability of the Lévy measure. This paper solves the open problem stated in page 933 of [15] and page 423 of [18] in the case where the unknown parameter appears only in the drift coefficient.

The first challenge is that the transition density of the solution to equation (1.1) is not explicit in general, which complicates the analysis of the log-likelihood of the discretized process $(X_{t_0}^{\theta}, X_{t_1}^{\theta}, \ldots, X_{t_n}^{\theta})$. To overcome this challenge, the Malliavin calculus approach initiated by Gobet [7, 8] is used to obtain an explicit expression for the logarithm derivative of the transition density w.r.t. the parameter (see Lemma 3.1 and 3.3). This allows us to derive an appropriate stochastic expansion of the log-likelihood ratio (see Lemma 4.1). Let us mention that this Malliavin calculus approach has been intensively developed in [15, 17, 18, 34, 3, 2]. In order to show the main contributions, we use a central limit theorem for triangular arrays of random variables and the ergodicity (see Lemma 4.2). These random variables are given by the terms which are determined by the Gaussian and drift components of equation (1.1).

The second challenge is to deal with the negligible contributions of the stochastic expansion of the log-likelihood ratio. As will be seen in Subsection 4.3, one difficulty comes from the fact that the conditional expectations are computed under the probability measure $\widetilde{\mathrm{P}}^{\theta_i^{0+}(\ell)}$ whereas the convergence is proven under the probability measure $\widehat{\mathrm{P}}^{\theta^0}$ with $\widehat{\mathrm{P}}^{\theta^0} \neq \widetilde{\mathrm{P}}^{\theta_i^{0+}(\ell)}$, where $\theta_i^{0+}(\ell)$ will be specified later on as a parameter value close to θ^0. To solve this problem, two technical Lemmas 3.6 and 3.7, which describe the Girsanov change of measures and the deviation of Girsanov change of measures, are mainly used. Recall that in [8] the author used a change of transition densities together with the upper and lower bounds of Gaussian type of the transition densities. This argument cannot be applied to our SDE with jumps where the upper and lower bounds of the transition densities may be of different characteristics due to the fact that the behavior of the transition density changes strongly with the presence of jumps.

The other difficulty is due to the jump components appearing in the stochastic expansion of the log-likelihood ratio (see Lemma 4.7). To resolve this difficulty, we apply a new approach developed recently by Ben Alaya et *al.* in [2] where a subtle analysis on the jump structure of the Lévy processes involving the amplitude of jumps and number of jumps is mainly used. More concretely, this approach consists in splitting the jumps of the Lévy processes into small jumps and big jumps, and then conditioning on the number of big jumps outside and inside the conditional expectation. When the number of big jumps of the Lévy processes outside and inside the conditional expectation is different, a large deviation principle in the estimate can be used (see (5.12) and (5.18)) and otherwise, an analysis on the complementary set is used (see (5.26)). All these arguments combined with the usual moment estimates allow to derive the exact large deviation type estimates given in Lemma 5.1 where the decreasing rate is determined by the intensity of the big jumps and the asymptotic behavior of the small jumps. This decreasing rate together with the help of condition **(A5)** on the integrability condition of the Lévy measure will show the negligible contributions of the jump components in the asymptotics. It is worth noticing that this new approach allows to include more general driving Lévy processes with possible infinite Lévy measure. Recall that the approach in [18, Lemma A.14] relies on conditioning on the jump structure involving number of jumps and amplitude of jumps, and using lower bounds for the transition density and upper bounds for the transition density conditioned on the jump structure in order to obtain the large deviation type estimates. Besides, only the case of compound Poisson process with finite Lévy measure is studied in [18]. Thus, the result derived in this paper can be seen as an improvement of the one obtained in [18].

The issue of parameter estimation for Lévy-driven SDEs from discrete observations usually requires an additional assumption on the decreasing rate of Δ_n, for instance, the rate $n\Delta_n^p \to 0$ for some $p > 1$ or the rate may depend on the behavior of the Lévy measure ν near zero, see [23, Theorem 3.5 and Theorem 4.6] and [6, Theorem 3.2]. On the one hand, our approach does not require neither additional assumption on the decreasing rate of Δ_n (see Remark 3.8) nor the tail behavior of the transition density. On the other hand, our approach keeps a wide class of Lévy processes in applications (see Example 2.1).

This paper is organized as follows. In Section 2, we formulate assumptions on equation (1.1) and provide a wide class of Lévy processes in applications which satisfies the assumption on the integrability of the Lévy measure. Furthermore, the main result is stated in Theorem 2.2. Section 3 is devoted to preliminary results which are needed for the proof of Theorem 2.2, such as an explicit expression for the logarithm derivative of the transition density, some crucial moment estimates, a conditional expectation formula, deviation of Girsanov change of measures, a discrete ergodic theorem. The proofs of these results are somewhat technical and are delayed to Appendix to maintain the flow of the exposition. We prove our main result in Section 4, which follows the aforementioned strategy. Finally, the proofs of some technical lemmas are presented in Section 5, where a result on the large deviation type estimates is also proven.

As usual, constants will be denoted by C which may change of value from one line to the next.

2. Assumptions and main result

We consider the following hypotheses on equation (1.1) we shall work with.

(A1) For any $\theta \in \Theta$, there exist a constant $L > 0$ and a function $\zeta : \mathbb{R}_0^d \to \mathbb{R}_+$ of polynomial growth in z satisfying $\zeta(z)\mathbf{1}_{|z|\leq 1} \leq C|z|$ with a constant $C > 0$ such that for all $x, y \in \mathbb{R}^d$, $z \in \mathbb{R}_0^d$,
$$|b(\theta, x) - b(\theta, y)| + |\sigma(x) - \sigma(y)| \leq L|x - y|,$$
$$|c(x, z) - c(y, z)| \leq \zeta(z)|x - y|, \quad |c(x, z)| \leq \zeta(z)(1 + |x|).$$
Moreover, the Lipschitz constant L is uniformly bounded on Θ.

(A2) The diffusion matrix σ satisfies an uniform ellipticity condition, that is, there exists a constant $c \geq 1$ such that for all $x, \xi \in \mathbb{R}^d$,
$$\frac{1}{c}|\xi|^2 \leq |\sigma(x)\xi|^2 \leq c|\xi|^2.$$

(A3) The functions b, σ and c are of class C^1 w.r.t. θ and x. Each partial derivative $\partial_{\theta_i} b$, $\partial_{x_i} b$, $\partial_{x_i}\sigma$ and $\partial_{x_i} c$ is of class C^1 w.r.t. x. Moreover, there exist positive constants C, q, $\gamma \in (0, 1]$, independent of $\left(\theta, \theta^1, \theta^2, x, y, z\right) \in \Theta^3 \times (\mathbb{R}^d)^2 \times \mathbb{R}_0^d$ such that
   (a) $|\partial_{x_i} b(\theta, x)| + |\partial_{x_i}\sigma(x)| \leq C$, and $|\partial_{x_i} c(x, z)| \leq \zeta(z)$;
   (b) $|h(\cdot, x)| \leq C\left(1 + |x|^q\right)$ for $h(\cdot, x) = \partial_{\theta_i} b(\theta, x), \partial^2_{x_i,x_j} b(\theta, x), \partial^2_{\theta_i,x_j} b(\theta, x)$ or $\partial^2_{x_i,x_j}\sigma(x)$;
   (c) $|\partial^2_{x_i,x_j} c(x, z)| \leq C\zeta(z)\left(1 + |x|\right)$;
   (d) $\partial_{\theta_i} b(\cdot, x)$ is γ- Hölder continuous w.r.t $\theta \in \Theta$:
$$|\partial_{\theta_i} b(\theta^1, x) - \partial_{\theta_i} b(\theta^2, x)| \leq C|\theta^1 - \theta^2|^\gamma \left(1 + |x|^q\right).$$

(A4) For all $\theta \in \Theta$, the process X^θ is ergodic in the sense that there exists a unique invariant probability measure $\pi_\theta(dx)$ such that the ergodic theorem holds, that is, as $T \to \infty$,
$$\frac{1}{T}\int_0^T g(X_t^\theta)dt \xrightarrow{\widehat{\mathrm{P}}^\theta} \int_{\mathbb{R}^d} g(x)\pi_\theta(dx),$$
for any π_θ-integrable function $g : \mathbb{R}^d \to \mathbb{R}$. Moreover, $\int_{\mathbb{R}^d} |x|^p \pi_\theta(dx) < \infty$, for any $p > 0$.

(A5) The Lévy measure ν satisfies $\int_{|z|>1} |z|^p \nu(dz) < \infty$ for any $p \geq 1$ and $\int_{0<|z|\leq 1} |z|\nu(dz) < \infty$.

The uniform ellipticity condition **(A2)** and regularity condition **(A3)**(a)-(c) on the coefficients are required in order to be able to apply the Malliavin calculus. Condition **(A3)**(d) is needed to show the main contributions of the stochastic expansion of the log-likelihood ratio (see Lemma 4.2). The ergodicity in the sense of **(A4)** was shown by Masuda in [24, Theorem 2.1] for a class of multidimensional diffusions with jumps. More recently, in [6, Lemma 2.1], conditions for the existence of an invariant measure π_θ and the ergodicity in the sense of **(A4)** are given for one-dimensional Lévy-driven SDEs with $c(x,z) = \gamma(x)z$. Several examples of ergodic diffusion processes with jumps are given in [32]. Moreover, results on ergodicity and exponential ergodicity which are understood in the sense of [27] and which are both stronger than the sense of **(A4)** have been established by Masuda [24] for diffusion processes with jumps.

The ergodicity also implies that for all $p > 0$,

$$\sup_{t\in\mathbb{R}_+} \widehat{\mathrm{E}}^{\theta}\left[|X_t^{\theta}|^p\right] < +\infty. \tag{2.1}$$

See [24, Theorem 2.2] and [6, Lemma 2.1].

The integrability condition **(A5)** of the Lévy measure controls the behavior of the small jumps and big jumps of the Lévy process, which is required in order to prove the negligible contribution of the jump component in the expansion (see Lemma 4.7). With the help of condition **(A5)**, the jump component is dominated over by the Gaussian component in a small time interval.

Example 2.1. *Condition* **(A5)** *can be verified for a wide class of Lévy measures: finite Lévy measure and infinite Lévy measure.*

1) *Lévy measure of a Poisson process $\nu(dz) = \lambda\delta_1(dz)$, where $\lambda > 0$ is the parameter of the Poisson process and $\delta_1(dz)$ is the Dirac measure supported on $\{1\}$.*

2) *Lévy measure of a compound Poisson process with exponentially distributed jump sizes $\nu(dz) = C\lambda e^{-\lambda z}\mathbf{1}_{(0,\infty)}(z)dz$, for some constants $C \in (0,\infty)$ and $\lambda \in (0,\infty)$.*

3) *Lévy measure of Gamma process $\nu(dz) = \gamma z^{-1}e^{-\lambda z}\mathbf{1}_{(0,\infty)}(z)dz$, where γ and λ are positive constants.*

4) *Lévy measure of inverse Gaussian process with $\nu(dz) = \frac{\delta}{\sqrt{2\pi z^3}}e^{-\frac{\gamma^2 z}{2}}\mathbf{1}_{(0,\infty)}(z)dz$, for a positive constant δ.*

5) *Lévy measure of a subordinator which is given by the gamma probability distribution. That is, $\nu(dz) = \frac{\lambda^{\alpha}}{\Gamma(\alpha)}z^{\alpha-1}e^{-\lambda z}\mathbf{1}_{(0,\infty)}(z)dz$ where $\alpha \in (-1,\infty)$ and λ is a positive constant.*

6) *Lévy measure of Variance gamma process*

$$\nu(dz) = \begin{cases} Ce^{Gz}|z|^{-1} & \text{if } z < 0 \\ Ce^{-Mz}z^{-1} & \text{if } z > 0, \end{cases}$$

where C, G, M are positive constants.

7) *Lévy measure of normal inverse Gaussian process*

$$\nu(dz) = \frac{\alpha\delta}{\pi}\frac{e^{\beta z}K_1(\alpha|z|)}{|z|}dz,$$

where $\delta > 0$, $\alpha > 0$, $-\alpha < \beta < \alpha$, $K_\lambda(z)$ is the modified Bessel function of the third kind

$$K_\lambda(z) = \frac{1}{2}\int_0^{\infty} u^{\lambda-1}\exp\left\{-\frac{1}{2}z(u+u^{-1})\right\}du.$$

8) *Lévy measure of some generalized tempered stable processes*

$$\nu(dz) = \frac{c_+e^{-\lambda_+ z}}{z^{1+\alpha_+}}\mathbf{1}_{(0,\infty)}(z)dz + \frac{c_-e^{-\lambda_-|z|}}{|z|^{1+\alpha_-}}\mathbf{1}_{(-\infty,0)}(z)dz,$$

with parameters satisfying $c_+ > 0$, $c_1 > 0$, $\lambda_+ > 0$, $\lambda_- > 0$, $\alpha_+ < 1$ *and* $\alpha_- < 1$.

For fixed $\theta^0 \in \Theta$, we consider a discrete observation $X^{n,\theta^0} = (X_{t_0}^{\theta^0}, X_{t_1}^{\theta^0}, \ldots, X_{t_n}^{\theta^0})$ of the process X^{θ^0}. The main result of this paper is the following LAN property.

Theorem 2.2. *Assume conditions* **(A1)-(A5)**. *Then, the LAN property holds for the likelihood at* θ^0 *with rate of convergence* $\varphi_{n\Delta_n}(\theta^0) = \mathrm{diag}(\frac{1}{\sqrt{n\Delta_n}}, \ldots, \frac{1}{\sqrt{n\Delta_n}})$ *where* $\varphi^1_{n\Delta_n}(\theta^0) = \cdots = \varphi^m_{n\Delta_n}(\theta^0) = \frac{1}{\sqrt{n\Delta_n}}$ *and asymptotic Fisher information matrix*

$$\Gamma(\theta^0) = \int_{\mathbb{R}^d} \left(\nabla_\theta b(\theta^0, x)\right)^* (\sigma\sigma^*)^{-1}(x) \nabla_\theta b(\theta^0, x) \pi_{\theta^0}(dx), \tag{2.2}$$

where the elements of matrix $\Gamma(\theta^0) = (\Gamma(\theta^0)_{i,j})_{1 \le i,j \le m} \in \mathbb{R}^m \otimes \mathbb{R}^m$ *are given by*

$$\Gamma(\theta^0)_{i,j} = \int_{\mathbb{R}^d} \left(\partial_{\theta_i} b(\theta^0, x)\right)^* (\sigma\sigma^*)^{-1}(x) \partial_{\theta_j} b(\theta^0, x) \pi_{\theta^0}(dx).$$

That is, for all $u \in \mathbb{R}^m$, *as* $n \to \infty$,

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

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

Remark 2.3. *Theorem 2.2 can be seen as an extension of the result obtained by Gobet in* [8, Theorem 4.1] *for ergodic diffusions in the case when the unknown parameter appears only in the drift coefficient and when a jump component is added to the solution process. When the jump component in* (1.1) *is degenerate, we recover the same formula for the asymptotic Fisher information matrix* $\Gamma(\theta^0)$ *of ergodic diffusions without jumps obtained in* [8, Theorem 4.1].

Remark 2.4. *Theorem 2.2 generalizes the result obtained by Kohatsu-Higa et al. in* [18, Theorem 2.2] *when the unknown parameter is multidimensional and when the jump component of driving Lévy processes are more general and of a possible infinite jump activity.*

Remark 2.5. *When the LAN property holds at* θ^0 *with rate of convergence* $\varphi_{n\Delta_n}(\theta^0) = \mathrm{diag}(\frac{1}{\sqrt{n\Delta_n}}, \ldots, \frac{1}{\sqrt{n\Delta_n}})$ *and asymptotic Fisher information matrix* $\Gamma(\theta^0)$, *in this case a sequence of estimators* $(\widehat{\theta}_n)_{n \ge 1}$ *of* θ^0 *is said to be asymptotically efficient at* θ^0 *in the sense of Hájek-Le Cam convolution theorem if as* $n \to \infty$,

$$\sqrt{n\Delta_n} \left(\widehat{\theta}_n - \theta^0\right) \overset{\mathcal{L}(\widehat{\mathrm{P}}^{\theta^0})}{\longrightarrow} \mathcal{N}(0, \Gamma(\theta^0)^{-1}).$$

Note that a sequence of estimators which is asymptotically efficient in the sense of Hájek-Le Cam convolution theorem achieves asymptotically the Cramér-Rao lower bound $\Gamma(\theta^0)^{-1}$ *for the estimation variance. For details, we refer the reader to e.g.* [21].

In [23], *the author constructs a discretized likelihood estimator with jump filtering from the time-continuous likelihood function, which is given by (6) in* [23], *for the drift parameter of an Ornstein-Uhlenbeck process driven by a Lévy process whose jump component is of finite jump activity or infinite jump activity. Combining our main result Theorem 2.2 and* [23, Theorem 3.5 and Theorem 4.6], *this estimator is asymptotically efficient in the sense of Hájek-Le Cam convolution theorem.*

More recently, in [6]*, a filtered maximum likelihood estimator (FMLE) for the drift parameter of the one-dimensional SDE* (1.1) *with* $c(x,z)=\gamma(x)z$*, driven by a Lévy process with a possible infinite jump activity, is constructed by applying a jump filter to the discretized likelihood function. This FMLE is given by (3.5) in* [6]*. As a consequence of our main result Theorem 2.2 and* [6, Theorem 3.2]*, this FMLE is asymptotically efficient in the sense of Hájek-Le Cam convolution theorem since its variance achieves the lower bound for the asymptotic variance of estimators with the optimal rate of convergence.*

Example 2.6. 1) *Consider the one-dimensional Ornstein-Uhlenbeck process driven by a Lévy process* $X^{\theta_1,\theta_2}=(X_t^{\theta_1,\theta_2})_{t\geq 0}$ *defined as*

$$\begin{aligned}X_t^{\theta_1,\theta_2}&=x_0+\int_0^t(\theta_2-\theta_1X_s^{\theta_1,\theta_2})ds+\sigma B_t+Z_t\\&=x_0+\int_0^t(\theta_2-\theta_1X_s^{\theta_1,\theta_2})ds+\sigma B_t+\int_0^t\int_{\mathbb{R}_0}z\widetilde{N}(ds,dz),\end{aligned}$$

where $\theta=(\theta_1,\theta_2)$*,* $\theta_1>0$*,* $\sigma>0$*. Assume that the Lévy measure satisfies condition* **(A5)** *which implies* $\int_{|z|>2}\log|z|\nu(dz)<+\infty$*. This integrability, together with* $\theta_1>0$*, ensures that* X^{θ_1,θ_2} *is ergodic in the sense of* **(A4)** *with an invariant probability measure* $\pi_{\theta_1,\theta_2}(dx)$ *which can be computed explicitly (see* [31, Theorem 17.5 and Corollary 17.9] *and* [24, Theorem 2.6]*), and satisfies* $\int_{\mathbb{R}}|x|^p\pi_{\theta_1,\theta_2}(dx)<\infty$ *for any* $p>0$*. In particular,*

$$\begin{aligned}&\lim_{t\to\infty}\widehat{\mathrm{E}}^{\theta_1,\theta_2}[X_t^{\theta_1,\theta_2}]=\int_{\mathbb{R}}x\pi_{\theta_1,\theta_2}(dx)=\frac{\theta_2}{\theta_1},\\&\lim_{t\to\infty}\widehat{\mathrm{E}}^{\theta_1,\theta_2}[(X_t^{\theta_1,\theta_2})^2]=\int_{\mathbb{R}}x^2\pi_{\theta_1,\theta_2}(dx)=\frac{1}{2\theta_1}\left(\sigma^2+\int_{\mathbb{R}_0}z^2\nu(dz)\right)+\left(\frac{\theta_2}{\theta_1}\right)^2.\end{aligned}$$

Then, the matrix $\Gamma(\theta_1,\theta_2)$ *is given by*

$$\Gamma(\theta_1,\theta_2)=\frac{1}{\sigma^2}\begin{pmatrix}\frac{1}{2\theta_1}\left(\sigma^2+\int_{\mathbb{R}_0}z^2\nu(dz)\right)+\left(\frac{\theta_2}{\theta_1}\right)^2 & -\frac{\theta_2}{\theta_1}\\ -\frac{\theta_2}{\theta_1} & 1\end{pmatrix}.$$

Notice that in this case conditions **(A1)**-**(A3)** *hold. As a consequence of Theorem 2.2, under condition* **(A5)***, the LAN property holds with rate of convergence* $\mathrm{diag}(\frac{1}{\sqrt{n\Delta_n}},\frac{1}{\sqrt{n\Delta_n}})$ *and asymptotic Fisher information matrix* $\Gamma(\theta^0)=\Gamma(\theta_1^0,\theta_2^0)$*.*

2) *Consider the one-dimensional Lévy process defined as*

$$\begin{aligned}X_t^{\theta}&=x_0+\theta t+\sigma B_t+Z_t\\&=x_0+\theta t+\sigma B_t+\int_0^t\int_{\mathbb{R}_0}z\widetilde{N}(ds,dz),\end{aligned}$$

where $\theta\in\mathbb{R}$ *and* $\sigma>0$*. Assume that the Lévy measure satisfies condition* **(A5)***. Notice that conditions* **(A1)**-**(A3)** *hold. Then as a consequence of Theorem 2.2, under condition* **(A5)***, the LAN property holds with rate of convergence* $\frac{1}{\sqrt{n\Delta_n}}$ *and asymptotic Fisher information* $\Gamma(\theta^0)=\frac{1}{\sigma^2}$*. In this case, condition* **(A4)** *is not needed since* $\Gamma(\theta^0)$ *can be obtained without using the ergodicity assumption, but thanks to the simple structure of the drift and diffusion coefficients (see* (4.6) *below).*

As usual, constants will be denoted by C which may change of value from one line to the next.

3. Preliminaries

In this section, we are going to present some preliminary results which are needed for the proof of Theorem 2.2. To simplify the exposition, for $i \in \{1, \ldots, m\}$ we use the following notations

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

Conditions **(A1)**-**(A2)** imply that the law of the discrete observation $X^{n,\theta} = (X_{t_0}^\theta, X_{t_1}^\theta, \ldots, X_{t_n}^\theta)$ of the process X^θ has a density in $(\mathbb{R}^d)^{n+1}$ that we denote by $p_n(\cdot;\theta)$. Under conditions **(A1)**, **(A2)** and **(A3)**(a), for any $t > s$ the law of X_t^θ conditioned on $X_s^\theta = x$ possesses a positive transition density $p^\theta(t-s, x, y)$, which is differentiable w.r.t. θ. To analyze the log-likelihood ratio in Theorem 2.2, the Markov property is used to rewrite the global likelihood function in terms of a product of transition densities and then a mean value theorem is applied. More precisely,

$$\begin{aligned}
&\log \frac{d\mathrm{P}_n^{\theta^0+\frac{u}{\sqrt{n\Delta_n}}}}{d\mathrm{P}_n^{\theta^0}}\left(X^{n,\theta^0}\right) = \log \frac{p_n\left(X^{n,\theta^0};\theta^0+\frac{u}{\sqrt{n\Delta_n}}\right)}{p_n\left(X^{n,\theta^0};\theta^0\right)} = \log \frac{p_n\left(X^{n,\theta^0};\theta^{0+}\right)}{p_n\left(X^{n,\theta^0};\theta^0\right)}\\
&= \sum_{k=0}^{n-1} \log \frac{p^{\theta^{0+}}}{p^{\theta^0}}\left(\Delta_n, X_{t_k}^{\theta^0}, X_{t_{k+1}}^{\theta^0}\right) = \sum_{k=0}^{n-1} \log \frac{p^{\theta_1^{0+}}}{p^{\theta^0}}\left(\Delta_n, X_{t_k}^{\theta^0}, X_{t_{k+1}}^{\theta^0}\right)\\
&= \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)\left(\Delta_n, X_{t_k}^{\theta^0}, X_{t_{k+1}}^{\theta^0}\right)\\
&= \sum_{k=0}^{n-1} \log \frac{p^{\theta_1^{0+}}}{p^{\theta_2^{0+}}}\left(\Delta_n, X_{t_k}^{\theta^0}, X_{t_{k+1}}^{\theta^0}\right) + \sum_{k=0}^{n-1} \log \frac{p^{\theta_2^{0+}}}{p^{\theta_3^{0+}}}\left(\Delta_n, X_{t_k}^{\theta^0}, X_{t_{k+1}}^{\theta^0}\right)\\
&\quad + \cdots + \sum_{k=0}^{n-1} \log \frac{p^{\theta_i^{0+}}}{p^{\theta_{i+1}^{0+}}}\left(\Delta_n, X_{t_k}^{\theta^0}, X_{t_{k+1}}^{\theta^0}\right) + \cdots + \sum_{k=0}^{n-1} \log \frac{p^{\theta_m^{0+}}}{p^{\theta^0}}\left(\Delta_n, X_{t_k}^{\theta^0}, X_{t_{k+1}}^{\theta^0}\right)\\
&= \sum_{k=0}^{n-1} \frac{u_1}{\sqrt{n\Delta_n}} \int_0^1 \frac{\partial_{\theta_1} p^{\theta_1^{0+}(\ell)}}{p^{\theta_1^{0+}(\ell)}}\left(\Delta_n, X_{t_k}^{\theta^0}, X_{t_{k+1}}^{\theta^0}\right) d\ell\\
&\quad + \sum_{k=0}^{n-1} \frac{u_2}{\sqrt{n\Delta_n}} \int_0^1 \frac{\partial_{\theta_2} p^{\theta_2^{0+}(\ell)}}{p^{\theta_2^{0+}(\ell)}}\left(\Delta_n, X_{t_k}^{\theta^0}, X_{t_{k+1}}^{\theta^0}\right) d\ell\\
&\quad + \cdots + \sum_{k=0}^{n-1} \frac{u_i}{\sqrt{n\Delta_n}} \int_0^1 \frac{\partial_{\theta_i} p^{\theta_i^{0+}(\ell)}}{p^{\theta_i^{0+}(\ell)}}\left(\Delta_n, X_{t_k}^{\theta^0}, X_{t_{k+1}}^{\theta^0}\right) d\ell
\end{aligned}$$

$$+\cdots+\sum_{k=0}^{n-1}\frac{u_m}{\sqrt{n\Delta_n}}\int_0^1\frac{\partial_{\theta_m}p^{\theta_m^{0+}(\ell)}}{p^{\theta_m^{0+}(\ell)}}\left(\Delta_n,X_{t_k}^{\theta^0},X_{t_{k+1}}^{\theta^0}\right)d\ell. \tag{3.1}$$

Then, using the approach developed in [7, 8], the integration by parts formula of the Malliavin calculus on each interval $[t_k,t_{k+1}]$ will be applied to obtain an explicit expression for the logarithm derivative of the transition density w.r.t. the parameter appearing in the decomposition (3.1). Towards this aim, we introduce canonical filtered probability spaces $(\Omega^i,\mathcal{F}^i,\{\mathcal{F}_t^i\}_{t\geq 0},\mathrm{P}^i)$, $i\in\{1,\ldots,4\}$, associated respectively to each of four processes B,N,W and M. Here $W=(W_t)_{t\geq 0}$ is a d-dimensional Brownian motion, $M(dt,dz)$ is a Poisson random measure with intensity measure $\nu(dz)dt$ associated to a centered Lévy process $\widetilde{Z}=(\widetilde{Z}_t)_{t\geq 0}$ without Gaussian component. The Lévy-Itô decomposition of $\widetilde{Z}$ takes the form $\widetilde{Z}_t=\int_0^t\int_{\mathbb{R}_0^d}z\widetilde{M}(ds,dz)$ for any $t\geq 0$, where $\widetilde{M}(dt,dz):=M(dt,dz)-\nu(dz)dt$ denotes the compensated Poisson random measure, $M(dt,dz):=\sum_{0\leq s\leq t}\mathbf{1}_{\{\Delta\widetilde{Z}_s\neq 0\}}\delta_{(s,\Delta\widetilde{Z}_s)}(ds,dz)$. Here, the jump amplitude of $\widetilde{Z}$ is defined as $\Delta\widetilde{Z}_s:=\widetilde{Z}_s-\widetilde{Z}_{s-}$ for any $s>0$, $\Delta\widetilde{Z}_0:=0$. Four processes B,N,W,M are mutually independent. Let $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geq 0},\mathrm{P})$ be the product filtered probability space of these four canonical spaces. We denote $\widehat{\Omega}=\Omega^1\times\Omega^2$, $\widehat{\mathcal{F}}=\mathcal{F}^1\otimes\mathcal{F}^2$, $\widehat{\mathrm{P}}=\mathrm{P}^1\otimes\mathrm{P}^2$, $\widehat{\mathcal{F}}_t=\mathcal{F}_t^1\otimes\mathcal{F}_t^2$, $\widetilde{\Omega}=\Omega^3\times\Omega^4$, $\widetilde{\mathcal{F}}=\mathcal{F}^3\otimes\mathcal{F}^4$, $\widetilde{\mathrm{P}}=\mathrm{P}^3\otimes\mathrm{P}^4$, and $\widetilde{\mathcal{F}}_t=\mathcal{F}_t^3\otimes\mathcal{F}_t^4$. Thus, $\Omega=\widehat{\Omega}\times\widetilde{\Omega}$, $\mathcal{F}=\widehat{\mathcal{F}}\otimes\widetilde{\mathcal{F}}$, $\mathrm{P}=\widehat{\mathrm{P}}\otimes\widetilde{\mathrm{P}}$, $\mathcal{F}_t=\widehat{\mathcal{F}}_t\otimes\widetilde{\mathcal{F}}_t$, and $\mathrm{E}=\widehat{\mathrm{E}}\otimes\widetilde{\mathrm{E}}$, where E, $\widehat{\mathrm{E}}$, $\widetilde{\mathrm{E}}$ denote the expectation w.r.t. P, $\widehat{\mathrm{P}}$ and $\widetilde{\mathrm{P}}$, respectively.

To avoid confusion with the observed process X^θ, we are going to introduce an independent copy of X^θ, denoted by $Y^\theta=(Y_t^\theta)_{t\geq 0}$, for which the Malliavin calculus will be applied. On the same probability space $(\Omega,\mathcal{F},\mathrm{P})$, we consider the stochastic flow $Y^\theta(s,x)=(Y_t^\theta(s,x),t\geq s)$, $x\in\mathbb{R}^d$ on the time interval $[s,\infty)$ and with initial condition $Y_s^\theta(s,x)=x$ satisfying

$$\begin{aligned}Y_t^\theta(s,x)=x&+\int_s^t b(\theta,Y_u^\theta(s,x))du+\int_s^t\sigma(Y_u^\theta(s,x))dW_u\\&+\int_s^t\int_{\mathbb{R}_0^d}c(Y_{u-}^\theta(s,x),z)\widetilde{M}(du,dz).\end{aligned}\tag{3.2}$$

In particular, we denote $Y_t^\theta\equiv Y_t^\theta(0,x_0)$, for all $t\geq 0$. Thus,

$$Y_t^\theta=x_0+\int_0^t b(\theta,Y_u^\theta)du+\int_0^t\sigma(Y_u^\theta)dW_u+\int_0^t\int_{\mathbb{R}_0^d}c(Y_{u-}^\theta,z)\widetilde{M}(du,dz). \tag{3.3}$$

The Malliavin calculus on the Wiener space induced by the Brownian motion W will be applied. Let D and δ denote respectively the Malliavin derivative and the Skorohod integral w.r.t. W on each interval $[t_k,t_{k+1}]$. We denote by $\mathbb{D}^{1,2}$ the space of random variables which are differentiable in the sense of Malliavin, and by Dom δ the domain of δ. We refer to Nualart [28] for a detailed exposition of the Malliavin calculus on the Wiener space and the Malliavin calculus adapted to our framework is introduced, for instance, in [30]. Recall that for a differentiable random variable $F\in\mathbb{D}^{1,2}$, its Malliavin derivative is denoted by $DF=(D^1F,\ldots,D^dF)$, where D^i is the Malliavin derivative in the ith direction W^i of the Brownian motion $W=(W^1,\ldots,W^d)$, for $i\in\{1,\ldots,d\}$. For a $\mathbb{R}^d$-valued process $U=(U^1,\ldots,U^d)\in$ Dom δ, the Skorohod integral of U is defined as $\delta(U)=\sum_{i=1}^d\delta^i(U^i)$, where δ^i denotes the Skorohod integral w.r.t. W^i.

For any $k \in \{0,\ldots,n-1\}$, under conditions **(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 θ (see Kunita [19]). We denote by $(\nabla_x Y_t^\theta(t_k,x), t \in [t_k,t_{k+1}])$ the Jacobian matrix, and by $(\partial_{\theta_i} Y_t^\theta(t_k,x), t \in [t_k,t_{k+1}])$ the derivative w.r.t. θ_i for $i \in \{1,\ldots,m\}$. These processes solve a system of SDEs

$$\nabla_x Y_t^\theta(t_k,x) = \mathrm{I}_d + \int_{t_k}^t \nabla_x b(\theta, Y_s^\theta(t_k,x))\nabla_x Y_s^\theta(t_k,x)ds \tag{3.4}$$

$$+\sum_{j=1}^d \int_{t_k}^t \nabla_x \sigma_j(Y_s^\theta(t_k,x))\nabla_x Y_s^\theta(t_k,x)dW_s^j + \int_{t_k}^t \int_{\mathbb{R}_0^d} \nabla_x c(Y_{s-}^\theta(t_k,x),z)\nabla_x Y_s^\theta(t_k,x)\widetilde{M}(ds,dz),$$

$$\partial_{\theta_i} Y_t^\theta(t_k,x) = \int_{t_k}^t \Big(\partial_{\theta_i} b(\theta, Y_s^\theta(t_k,x)) + \nabla_x b(\theta, Y_s^\theta(t_k,x))\partial_{\theta_i} Y_s^\theta(t_k,x)\Big) ds \tag{3.5}$$

$$+\sum_{j=1}^d \int_{t_k}^t \nabla_x \sigma_j(Y_s^\theta(t_k,x))\partial_{\theta_i} Y_s^\theta(t_k,x)dW_s^j + \int_{t_k}^t \int_{\mathbb{R}_0^d} \nabla_x c(Y_{s-}^\theta(t_k,x),z)\partial_{\theta_i} Y_s^\theta(t_k,x)\widetilde{M}(ds,dz),$$

for $i \in \{1,\ldots,m\}$, where $\sigma_1,...,\sigma_d : \mathbb{R}^d \to \mathbb{R}^d$ denote the columns of the matrix σ.

Moreover, under conditions **(A1)**, **(A2)** and **(A3)**(a)-(c), the random variables $Y_t^\theta(t_k,x)$, $\nabla_x Y_t^\theta(t_k,x)$, $(\nabla_x Y_t^\theta(t_k,x))^{-1}$ 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 [30, Theorem 3]). Furthermore, the Malliavin derivative $D_s Y_t^\theta(t_k,x)$ satisfies the following equation

$$D_s Y_t^\theta(t_k,x) = \sigma(Y_s^\theta(t_k,x)) + \int_s^t \nabla_x b(\theta, Y_u^\theta(t_k,x))D_s Y_u^\theta(t_k,x)du$$

$$+\sum_{j=1}^d \int_s^t \nabla_x \sigma_j(Y_u^\theta(t_k,x))D_s Y_u^\theta(t_k,x)dW_u^j + \int_s^t \int_{\mathbb{R}_0^d} \nabla_x c(Y_{u-}^\theta(t_k,x),z)D_s Y_u^\theta(t_k,x)\widetilde{M}(du,dz),$$

for $s \leq t$ a.e., and $D_s Y_t^\theta(t_k,x) = 0$ for $s > t$ a.e. By [30, Proposition 7], it holds that

$$D_s Y_t^\theta(t_k,x) = \nabla_x Y_t^\theta(t_k,x)(\nabla_x Y_s^\theta(t_k,x))^{-1}\sigma(Y_s^\theta(t_k,x))\mathbf{1}_{[t_k,t]}(s).$$

Now, for all $k \in \{0,...,n-1\}$ and $x \in \mathbb{R}^d$, we denote by $\widetilde{\mathrm{P}}_{t_k,x}^\theta$ the probability law of Y^θ starting at x at time t_k, 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 by $\widetilde{\mathrm{E}}_{t_k,x}^\theta$ the expectation w.r.t. $\widetilde{\mathrm{P}}_{t_k,x}^\theta$. That is, for all $\widetilde{\mathcal{F}}$-measurable random variables V, we have $\widetilde{\mathrm{E}}_{t_k,x}^\theta[V] = \widetilde{\mathrm{E}}[V|Y_{t_k}^\theta = x]$. Hence, $\widetilde{\mathrm{E}}_{t_k,x}^\theta$ is the expectation under the probability law of Y^θ starting at x at time t_k. Similarly, we denote by $\widehat{\mathrm{P}}_{t_k,x}^\theta$ the probability law of X^θ starting at x at time t_k, 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 by $\widehat{\mathrm{E}}_{t_k,x}^\theta$ the expectation w.r.t. $\widehat{\mathrm{P}}_{t_k,x}^\theta$. That is, for all $\widehat{\mathcal{F}}$-measurable random variables V, we have $\widehat{\mathrm{E}}_{t_k,x}^\theta[V] = \widehat{\mathrm{E}}[V|X_{t_k}^\theta = x]$. Let $\mathrm{P}_{t_k,x}^\theta := \widehat{\mathrm{P}}_{t_k,x}^\theta \otimes \widetilde{\mathrm{P}}_{t_k,x}^\theta$ be the product measure, and $\mathrm{E}_{t_k,x}^\theta = \widehat{\mathrm{E}}_{t_k,x}^\theta \otimes \widetilde{\mathrm{E}}_{t_k,x}^\theta$ denotes the expectation w.r.t. $\mathrm{P}_{t_k,x}^\theta$.

As in [7, Proposition 4.1], we obtain the following explicit expression for the logarithm derivative of the transition density w.r.t. θ in terms of a conditional expectation.

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

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

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

We next derive the following decomposition of the Skorohod integral appearing in the conditional expectation of Lemma 3.1.

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

$$\delta\left(U^\theta(t_k,x)\partial_{\theta_i}Y^\theta_{t_{k+1}}(t_k,x)\right)=\Delta_n(\partial_{\theta_i}b(\theta,x))^*(\sigma\sigma^*)^{-1}(x)\left(Y^\theta_{t_{k+1}}-Y^\theta_{t_k}-b(\theta,Y^\theta_{t_k})\Delta_n\right)$$
$$-R_1^{\theta,k}+R_2^{\theta,k}+R_3^{\theta,k}-R_4^{\theta,k}-R_5^{\theta,k}-R_6^{\theta,k},$$

where

$$R_1^{\theta,k}=\int_{t_k}^{t_{k+1}}\int_s^{t_{k+1}}\mathrm{tr}\left(D_s\left(((\nabla_xY_u^\theta(t_k,x))^{-1}\partial_{\theta_i}b(\theta,Y_u^\theta(t_k,x)))^*\right)\sigma^{-1}(Y_s^\theta(t_k,x))\nabla_xY_s^\theta(t_k,x)\right)duds,$$

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

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

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

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

$$R_6^{\theta,k}=\Delta_n(\partial_{\theta_i}b(\theta,Y^\theta_{t_k}))^*(\sigma\sigma^*)^{-1}(Y^\theta_{t_k})\int_{t_k}^{t_{k+1}}\int_{\mathbb{R}_0^d}c(Y^\theta_{s-},z)\widetilde{M}(ds,dz).$$

As a consequence of Lemma 3.1 and 3.2, we derive the following explicit expression for the logarithm derivative of the transition density.

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

$$\frac{\partial_{\theta_i}p^\theta}{p^\theta}(\Delta_n,x,y)=(\partial_{\theta_i}b(\theta,x))^*(\sigma\sigma^*)^{-1}(x)\left(y-x-b(\theta,x)\Delta_n\right)$$
$$+\frac{1}{\Delta_n}\widetilde{\mathrm{E}}^\theta_{t_k,x}\left[-R_1^{\theta,k}+R_2^{\theta,k}+R_3^{\theta,k}-R_4^{\theta,k}-R_5^{\theta,k}-R_6^{\theta,k}\big|Y^\theta_{t_{k+1}}=y\right].$$

We will use the following estimates for the solution to (3.2).

Lemma 3.4. *Assume conditions* **(A1)**, **(A2)** *and* **(A5)**.

(i) *For any* $p\geq1$ *and* $\theta\in\Theta$, *there exists a constant* $C_p>0$ *such that for all* $k\in\{0,...,n-1\}$ *and* $t\in[t_k,t_{k+1}]$,

$$\widetilde{\mathrm{E}}^\theta_{t_k,x}\left[\left|Y_t^\theta(t_k,x)-Y_{t_k}^\theta(t_k,x)\right|^p\right]\leq C_p\left|t-t_k\right|^{\frac{p}{2}\wedge1}(1+|x|^p).$$

(ii) *For any function g defined on $\Theta\times\mathbb{R}^d$ with polynomial growth in x uniformly in $\theta\in\Theta$, there exist constants $C,q>0$ such that for all $k\in\{0,...,n-1\}$ and $t\in[t_k,t_{k+1}]$,*

$$\widetilde{\mathrm{E}}_{t_k,x}^{\theta}\left[\left|g(\theta,Y_t^{\theta}(t_k,x))\right|\right]\leq C\left(1+|x|^q\right).$$

Moreover, all these statements remain valid for X^{θ}.

Using conditions **(A1)**, **(A2)**, **(A3)**(a)-(c) and **(A5)**, and Gronwall's inequality, it can be checked that for any $\theta\in\Theta$ and $p\geq 2$, there exist constants $C_p,q>0$ such that for all $k\in\{0,...,n-1\}$ and $t\in[t_k,t_{k+1}]$,

$$\begin{aligned}&\widetilde{\mathrm{E}}_{t_k,x}^{\theta}\left[\left|\nabla_x Y_t^{\theta}(t_k,x)\right|^p+\left|(\nabla_x Y_t^{\theta}(t_k,x))^{-1}\right|^p\right]+\sup_{s\in[t_k,t_{k+1}]}\widetilde{\mathrm{E}}_{t_k,x}^{\theta}\left[\left|D_sY_t^{\theta}(t_k,x)\right|^p\right]\leq C_p,\\&\widetilde{\mathrm{E}}_{t_k,x}^{\theta}\left[\left|\partial_{\theta_i}Y_t^{\theta}(t_k,x)\right|^p\right]+\sup_{s\in[t_k,t_{k+1}]}\widetilde{\mathrm{E}}_{t_k,x}^{\theta}\left[\left|D_s\left(\nabla_xY_t^{\theta}(t_k,x)\right)\right|^p\right]\leq C_p\left(1+|x|^q\right),\end{aligned}\tag{3.6}$$

where the constant C_p is uniform in θ. As a consequence, we have the following estimates, which follow easily from (5.3), Lemma 3.4 and properties of the expectation of the Brownian motion and the Skorohod integral.

Lemma 3.5. *Under conditions* **(A1)**, **(A2)**, **(A3)**(a)-(c) *and* **(A5)**, *for any $\theta\in\Theta$ and $p\geq 2$, there exist constants $C_p,q>0$ such that for all $k\in\{0,...,n-1\}$,*

$$\widetilde{\mathrm{E}}_{t_k,x}^{\theta}\left[-R_1^{\theta,k}+R_2^{\theta,k}+R_3^{\theta,k}\right]=0,\tag{3.7}$$

$$\widetilde{\mathrm{E}}_{t_k,x}^{\theta}\left[\left|-R_1^{\theta,k}+R_2^{\theta,k}+R_3^{\theta,k}\right|^p\right]\leq C_p\Delta_n^{\frac{3p+1}{2}}\left(1+|x|^q\right).\tag{3.8}$$

We next recall Girsanov's theorem on each interval $[t_k,t_{k+1}]$. For all $\theta,\theta^1\in\Theta$, $x\in\mathbb{R}^d$ and $k\in\{0,...,n-1\}$, by [13, Theorem III.5.34], 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

$$\begin{aligned}\frac{d\widehat{\mathrm{P}}_{t_k,x}^{\theta}}{d\widehat{\mathrm{P}}_{t_k,x}^{\theta^1}}\left((X_t^{\theta^1})_{t\in[t_k,t_{k+1}]}\right)=\exp\bigg\{&\int_{t_k}^{t_{k+1}}\sigma^{-1}(X_t^{\theta^1})\left(b(\theta,X_t^{\theta^1})-b(\theta^1,X_t^{\theta^1})\right)dB_t\\&-\frac{1}{2}\int_{t_k}^{t_{k+1}}\left|\sigma^{-1}(X_t^{\theta^1})\left(b(\theta,X_t^{\theta^1})-b(\theta^1,X_t^{\theta^1})\right)\right|^2dt\bigg\}.\end{aligned}\tag{3.9}$$

By Girsanov's theorem, the process $B^{\widehat{\mathrm{P}}_{t_k,x}^{\theta}}=(B_t^{\widehat{\mathrm{P}}_{t_k,x}^{\theta}},t\in[t_k,t_{k+1}])$ is a Brownian motion under $\widehat{\mathrm{P}}_{t_k,x}^{\theta}$, where for any $t\in[t_k,t_{k+1}]$,

$$B_t^{\widehat{\mathrm{P}}_{t_k,x}^{\theta}}:=B_t-\int_{t_k}^{t}\sigma^{-1}(X_s^{\theta^1})\left(b(\theta,X_s^{\theta^1})-b(\theta^1,X_s^{\theta^1})\right)ds.$$

Next, we give two following technical lemmas which will be useful in the sequel.

Lemma 3.6. *Assume conditions* **(A1)**, **(A2)** *and* **(A3)**(a). *Then for any $\theta\in\Theta$, $k\in\{0,...,n-1\}$ and $\widetilde{\mathcal{F}}_{t_{k+1}}$-measurable random variable V,*

$$\widehat{\mathrm{E}}^{\theta^0}\left[\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta}\left[V\big|Y_{t_{k+1}}^{\theta}=X_{t_{k+1}}^{\theta^0}\right]\big|\widehat{\mathcal{F}}_{t_k}\right]=\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta}\left[V\right].$$

Now, to simplify the notation, for $j \in \{1, \dots, m\}$ we set

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

Lemma 3.7. *Assume conditions* **(A1)***,* **(A2)***,* **(A3)**(a)-(b) *and* **(A5)***. Let $p, q > 1$ satisfying that $\frac{1}{p} + \frac{1}{q} = 1$. Then for any $k \in \{0, ..., n-1\}$ and $x \in \mathbb{R}^d$, there exist constants $C > 0, q_1 > 0$ such that for any $\widehat{\mathcal{F}}_{t_{k+1}}$-measurable random variable V,*

$$\begin{aligned}
&\left|\widehat{\mathrm{E}}_{t_k,x}^{\theta_i^{0+}(\ell)}\left[V\left(\frac{d\widehat{\mathrm{P}}_{t_k,x}^{\theta^0}}{d\widehat{\mathrm{P}}_{t_k,x}^{\theta_i^{0+}(\ell)}}\left((X_t^{\theta_i^{0+}(\ell)})_{t\in[t_k,t_{k+1}]}\right)-1\right)\right]\right| \\
&\leq C\sqrt{\Delta_n}\,(1+|x|^{q_1})\left(\left|\int_{\theta_i^0+\ell\frac{u_i}{\sqrt{n\Delta_n}}}^{\theta_i^0}\left(\widehat{\mathrm{E}}_{t_k,x}^{\theta_i(0+)}\left[|V|^q\right]\right)^{\frac{1}{q}}d\theta_i\right|\right. \\
&\quad\left.+\left|\int_{\theta_{i+1}^0+\frac{u_{i+1}}{\sqrt{n\Delta_n}}}^{\theta_{i+1}^0}\left(\widehat{\mathrm{E}}_{t_k,x}^{\theta_{i+1}(0+)}\left[|V|^q\right]\right)^{\frac{1}{q}}d\theta_{i+1}\right|+\cdots+\left|\int_{\theta_m^0+\frac{u_m}{\sqrt{n\Delta_n}}}^{\theta_m^0}\left(\widehat{\mathrm{E}}_{t_k,x}^{\theta_m(0+)}\left[|V|^q\right]\right)^{\frac{1}{q}}d\theta_m\right|\right).
\end{aligned}$$

Remark 3.8. *From* (3.7) *in Lemma 3.5, the random variable $-R_1^{\theta,k} + R_2^{\theta,k} + R_3^{\theta,k}$ has zero mean, which turns out to be useful in Lemma 4.4. Furthermore, Lemma 3.7 allows to give the rate $\sqrt{\Delta_n}\frac{1}{\sqrt{n\Delta_n}} = \frac{1}{\sqrt{n}}$ in the estimates, which will be used in Lemma 4.5. Using these technical Lemmas 3.5 and 3.7, we do not require an additional assumption on the decreasing rate of Δ_n, for instance, $n\Delta_n^p \to 0$ for some $p > 1$.*

Next, we prove a discrete ergodic theorem.

Lemma 3.9. *Assume conditions* **(A1)***,* **(A4)** *and* **(A5)***. Let $g : \mathbb{R}^d \to \mathbb{R}$ be a differentiable function satisfying that $|g(x)|$ and $|\nabla g(x)|$ have polynomial growth in x. Then, as $n \to \infty$,*

$$\frac{1}{n}\sum_{k=0}^{n-1} g(X_{t_k}^{\theta^0}) \overset{\widehat{\mathrm{P}}^{\theta^0}}{\longrightarrow} \int_{\mathbb{R}^d} g(x)\pi_{\theta^0}(dx).$$

We finally recall a convergence in probability result and a central limit theorem for triangular arrays of random variables. For each $n \in \mathbb{N}$, let $(\zeta_{k,n})_{k\geq 1}$ be a sequence of random variables defined on the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathrm{P})$, and assume that they are $\mathcal{F}_{t_{k+1}}$-measurable, for all k.

Lemma 3.10. [12, Lemma 3.4] a) *Assume that as $n \to \infty$,*

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

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

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

Lemma 3.11. [12, Lemma 3.6] *Assume that there exist real numbers Q and $V>0$ such that*

$$\sum_{k=0}^{n-1}\mathrm{E}\left[\zeta_{k,n}|\mathcal{F}_{t_k}\right]\xrightarrow{\mathrm{P}}Q,\qquad \sum_{k=0}^{n-1}\left(\mathrm{E}\left[\zeta_{k,n}^2|\mathcal{F}_{t_k}\right]-\left(\mathrm{E}\left[\zeta_{k,n}|\mathcal{F}_{t_k}\right]\right)^2\right)\xrightarrow{\mathrm{P}}V,\ \text{and}$$
$$\sum_{k=0}^{n-1}\mathrm{E}\left[\zeta_{k,n}^4|\mathcal{F}_{t_k}\right]\xrightarrow{\mathrm{P}}0,$$

as $n\to\infty$. Then as $n\to\infty$, $\sum_{k=0}^{n-1}\zeta_{k,n}\xrightarrow{\mathcal{L}(\mathrm{P})}\mathcal{N}(0,V)+Q$, where $\mathcal{N}(0,V)$ is a centered Gaussian random variable with variance V.

4. Proof of Theorem 2.2

In this section, the proof of Theorem 2.2 will be divided into three steps. We begin deriving an appropriate stochastic expansion of the log-likelihood ratio by using Lemma 3.3. The second step will show the main contributions of the stochastic expansion by applying the central limit theorem for triangular arrays of random variables and the ergodicity property. Finally, the last step is devoted to treat the negligible contributions of the stochastic expansion.

4.1. Stochastic expansion of the log-likelihood ratio.

Lemma 4.1. *Assume conditions* **(A1)**, **(A2)** *and* **(A3)**(a)-(c). *Then*

$$\log\frac{d\mathrm{P}_n^{\theta^0+\frac{u}{\sqrt{n\Delta_n}}}}{d\mathrm{P}_n^{\theta^0}}\left(X^{n,\theta^0}\right)=\sum_{k=0}^{n-1}\sum_{i=1}^{m}\xi_{i,k,n}+\sum_{k=0}^{n-1}\sum_{i=1}^{m}\frac{u_i}{\sqrt{n\Delta_n^3}}\int_0^1\Big\{H_4^{\theta^0,i,k}+H_5^{\theta^0,i,k}+H_6^{\theta^0,i,k}$$
$$+\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta_i^{0+}(\ell)}\left[R^{\theta_i^{0+}(\ell),k}-R_4^{\theta_i^{0+}(\ell),k}-R_5^{\theta_i^{0+}(\ell),k}-R_6^{\theta_i^{0+}(\ell),k}\Big|Y_{t_{k+1}}^{\theta_i^{0+}(\ell)}=X_{t_{k+1}}^{\theta^0}\right]\Big\}d\ell,\tag{4.1}$$

where

$$\xi_{i,k,n}=\frac{u_i}{\sqrt{n\Delta_n}}\int_0^1\left(\partial_{\theta_i}b(\theta_i^{0+}(\ell),X_{t_k}^{\theta^0})\right)^*(\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0})$$
$$\cdot\left(\sigma(X_{t_k}^{\theta^0})\left(B_{t_{k+1}}-B_{t_k}\right)+\left(b(\theta^0,X_{t_k}^{\theta^0})-b(\theta_i^{0+}(\ell),X_{t_k}^{\theta^0})\right)\Delta_n\right)d\ell,$$
$$R^{\theta_i^{0+}(\ell),k}=-R_1^{\theta_i^{0+}(\ell),k}+R_2^{\theta_i^{0+}(\ell),k}+R_3^{\theta_i^{0+}(\ell),k},$$
$$H_4^{\theta^0,i,k}=\Delta_n\left(\partial_{\theta_i}b(\theta_i^{0+}(\ell),X_{t_k}^{\theta^0})\right)^*(\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0})\int_{t_k}^{t_{k+1}}\left(b(\theta^0,X_s^{\theta^0})-b(\theta^0,X_{t_k}^{\theta^0})\right)ds,$$
$$H_5^{\theta^0,i,k}=\Delta_n\left(\partial_{\theta_i}b(\theta_i^{0+}(\ell),X_{t_k}^{\theta^0})\right)^*(\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0})\int_{t_k}^{t_{k+1}}\left(\sigma(X_s^{\theta^0})-\sigma(X_{t_k}^{\theta^0})\right)dB_s,$$
$$H_6^{\theta^0,i,k}=\Delta_n\left(\partial_{\theta_i}b(\theta_i^{0+}(\ell),X_{t_k}^{\theta^0})\right)^*(\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0})\int_{t_k}^{t_{k+1}}\int_{\mathbb{R}_0^d}c(X_{s-}^{\theta^0},z)\widetilde{N}(ds,dz).$$

Proof. Using the decomposition (3.1) and Lemma 3.3, we obtain that

$$\log\frac{d\mathrm{P}_n^{\theta^0+\frac{u}{\sqrt{n\Delta_n}}}}{d\mathrm{P}_n^{\theta^0}}\left(X^{n,\theta^0}\right)=\sum_{k=0}^{n-1}\sum_{i=1}^{m}\frac{u_i}{\sqrt{n\Delta_n}}\int_0^1\frac{\partial_{\theta_i}p^{\theta_i^{0+}(\ell)}}{p^{\theta_i^{0+}(\ell)}}\left(\Delta_n,X_{t_k}^{\theta^0},X_{t_{k+1}}^{\theta^0}\right)d\ell$$

$$= \sum_{k=0}^{n-1}\sum_{i=1}^{m} \frac{u_i}{\sqrt{n\Delta_n}} \int_0^1 \bigg((\partial_{\theta_i} b(\theta_i^{0+}(\ell), X_{t_k}^{\theta^0}))^* (\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0}) \Big(X_{t_{k+1}}^{\theta^0} - X_{t_k}^{\theta^0} - b(\theta_i^{0+}(\ell), X_{t_k}^{\theta^0})\Delta_n \Big)$$

$$+ \frac{1}{\Delta_n} \widetilde{\mathrm{E}}_{t_k, X_{t_k}^{\theta^0}}^{\theta_i^{0+}(\ell)} \Big[R^{\theta_i^{0+}(\ell),k} - R_4^{\theta_i^{0+}(\ell),k} - R_5^{\theta_i^{0+}(\ell),k} - R_6^{\theta_i^{0+}(\ell),k} \Big| Y_{t_{k+1}}^{\theta_i^{0+}(\ell)} = X_{t_{k+1}}^{\theta^0} \Big] \bigg) d\ell. \tag{4.2}$$

Next, using equation (1.1), we get that

$$X_{t_{k+1}}^{\theta^0} - X_{t_k}^{\theta^0} = \sigma(X_{t_k}^{\theta^0}) \left(B_{t_{k+1}} - B_{t_k} \right) + b(\theta^0, X_{t_k}^{\theta^0})\Delta_n + \int_{t_k}^{t_{k+1}} \Big(b(\theta^0, X_s^{\theta^0}) - b(\theta^0, X_{t_k}^{\theta^0}) \Big) ds$$

$$+ \int_{t_k}^{t_{k+1}} \Big(\sigma(X_s^{\theta^0}) - \sigma(X_{t_k}^{\theta^0}) \Big) dB_s + \int_{t_k}^{t_{k+1}} \int_{\mathbb{R}_0^d} c(X_{s-}^{\theta^0}, z) \widetilde{N}(ds, dz).$$

This, together with (4.2), gives the desired result. □

In the next two subsections, we will prove that the random variable $\xi_{i,k,n}$ determined by the Gaussian and drift components of equation (1.1) is the only term that contributes to the limit and all the others terms are negligible.

In all what follows, Lemma 3.9 will be used repeatedly without being quoted.

4.2. **Main contributions: LAN property.**

Lemma 4.2. *Assume conditions* **(A1)-(A4)**. *Then as* $n \to \infty$,

$$\sum_{k=0}^{n-1}\sum_{i=1}^{m} \xi_{i,k,n} \overset{\mathcal{L}(\widehat{\mathrm{P}}^{\theta^0})}{\longrightarrow} u^* \mathcal{N}(0, \Gamma(\theta^0)) - \frac{1}{2} u^* \Gamma(\theta^0) u,$$

where $\Gamma(\theta^0)$ *is given by* (2.2).

Proof. Applying Lemma 3.11 to $\sum_{i=1}^m \xi_{i,k,n}$, it suffices to show that as $n \to \infty$,

$$\sum_{k=0}^{n-1} \widehat{\mathrm{E}}^{\theta^0} \left[\xi_{i,k,n} | \widehat{\mathcal{F}}_{t_k} \right] \overset{\widehat{\mathrm{P}}^{\theta^0}}{\longrightarrow} -\frac{1}{2} u_i^2 \Gamma(\theta^0)_{i,i} - u_i u_{i+1} \Gamma(\theta^0)_{i,i+1} - \cdots - u_i u_m \Gamma(\theta^0)_{i,m}, \tag{4.3}$$

$$\sum_{k=0}^{n-1} \Big(\widehat{\mathrm{E}}^{\theta^0} \left[\xi_{i,k,n} \xi_{j,k,n} | \widehat{\mathcal{F}}_{t_k} \right] - \widehat{\mathrm{E}}^{\theta^0} \left[\xi_{i,k,n} | \widehat{\mathcal{F}}_{t_k} \right] \widehat{\mathrm{E}}^{\theta^0} \left[\xi_{j,k,n} | \widehat{\mathcal{F}}_{t_k} \right] \Big) \overset{\widehat{\mathrm{P}}^{\theta^0}}{\longrightarrow} u_i u_j \Gamma(\theta^0)_{i,j}, \tag{4.4}$$

$$\sum_{k=0}^{n-1} \widehat{\mathrm{E}}^{\theta^0} \left[(\xi_{i,k,n})^4 | \widehat{\mathcal{F}}_{t_k} \right] \overset{\widehat{\mathrm{P}}^{\theta^0}}{\longrightarrow} 0. \tag{4.5}$$

Proof of (4.3). Using the fact that $\widehat{\mathrm{E}}^{\theta^0}[B_{t_{k+1}} - B_{t_k} | \widehat{\mathcal{F}}_{t_k}] = 0$, we have

$$\sum_{k=0}^{n-1} \widehat{\mathrm{E}}^{\theta^0} \left[\xi_{i,k,n} | \widehat{\mathcal{F}}_{t_k} \right] = \sum_{k=0}^{n-1} \frac{u_i}{\sqrt{n\Delta_n}} \int_0^1 \Big(\partial_{\theta_i} b(\theta_i^{0+}(\ell), X_{t_k}^{\theta^0}) \Big)^* (\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0})$$

$$\cdot \Big(b(\theta^0, X_{t_k}^{\theta^0}) - b(\theta_i^{0+}(\ell), X_{t_k}^{\theta^0}) \Big) \Delta_n d\ell$$

$$= \sum_{k=0}^{n-1} \frac{u_i}{\sqrt{n\Delta_n}} \int_0^1 \Big(\partial_{\theta_i} b(\theta^0, X_{t_k}^{\theta^0}) \Big)^* (\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0}) \Big(b(\theta^0, X_{t_k}^{\theta^0}) - b(\theta_i^{0+}(\ell), X_{t_k}^{\theta^0}) \Big) \Delta_n d\ell$$

$$+\sum_{k=0}^{n-1}\xi_{i,1,k,n},$$

where

$$\xi_{i,1,k,n}=\frac{u_i}{\sqrt{n\Delta_n}}\int_0^1\Big(\partial_{\theta_i}b(\theta_i^{0+}(\ell),X_{t_k}^{\theta^0})-\partial_{\theta_i}b(\theta^0,X_{t_k}^{\theta^0})\Big)^*(\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0})\\ \cdot\Big(b(\theta^0,X_{t_k}^{\theta^0})-b(\theta_i^{0+}(\ell),X_{t_k}^{\theta^0})\Big)\Delta_n d\ell.$$

Then, using the mean value theorem,

$$b(\theta^0,X_{t_k}^{\theta^0})-b(\theta_i^{0+}(\ell),X_{t_k}^{\theta^0})=-\Big(b(\theta_i^{0+}(\ell),X_{t_k}^{\theta^0})-b(\theta_{i+1}^{0+},X_{t_k}^{\theta^0})+b(\theta_{i+1}^{0+},X_{t_k}^{\theta^0})-b(\theta_{i+2}^{0+},X_{t_k}^{\theta^0})\\ +\cdots+b(\theta_{m-1}^{0+},X_{t_k}^{\theta^0})-b(\theta_m^{0+},X_{t_k}^{\theta^0})+b(\theta_m^{0+},X_{t_k}^{\theta^0})-b(\theta^0,X_{t_k}^{\theta^0})\Big)\\ =-\Big(\ell\frac{u_i}{\sqrt{n\Delta_n}}\int_0^1\partial_{\theta_i}b(\theta_i^{0+}(\alpha\ell),X_{t_k}^{\theta^0})d\alpha+\frac{u_{i+1}}{\sqrt{n\Delta_n}}\int_0^1\partial_{\theta_{i+1}}b(\theta_{i+1}^{0+}(\alpha),X_{t_k}^{\theta^0})d\alpha\\ +\cdots+\frac{u_m}{\sqrt{n\Delta_n}}\int_0^1\partial_{\theta_m}b(\theta_m^{0+}(\alpha),X_{t_k}^{\theta^0})d\alpha\Big)\\ =-\Big(\ell\frac{u_i}{\sqrt{n\Delta_n}}\partial_{\theta_i}b(\theta^0,X_{t_k}^{\theta^0})+\frac{u_{i+1}}{\sqrt{n\Delta_n}}\partial_{\theta_{i+1}}b(\theta^0,X_{t_k}^{\theta^0})+\cdots+\frac{u_m}{\sqrt{n\Delta_n}}\partial_{\theta_m}b(\theta^0,X_{t_k}^{\theta^0})d\alpha\Big)\\ -\Big(\ell\frac{u_i}{\sqrt{n\Delta_n}}\int_0^1\Big(\partial_{\theta_i}b(\theta_i^{0+}(\alpha\ell),X_{t_k}^{\theta^0})-\partial_{\theta_i}b(\theta^0,X_{t_k}^{\theta^0})\Big)d\alpha\\ +\frac{u_{i+1}}{\sqrt{n\Delta_n}}\int_0^1\Big(\partial_{\theta_{i+1}}b(\theta_{i+1}^{0+}(\alpha),X_{t_k}^{\theta^0})-\partial_{\theta_{i+1}}b(\theta^0,X_{t_k}^{\theta^0})\Big)d\alpha\\ +\cdots+\frac{u_m}{\sqrt{n\Delta_n}}\int_0^1\Big(\partial_{\theta_m}b(\theta_m^{0+}(\alpha),X_{t_k}^{\theta^0})-\partial_{\theta_m}b(\theta^0,X_{t_k}^{\theta^0})\Big)d\alpha\Big),$$

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

$$\theta_i^{0+}(\alpha\ell):=(\theta_1^0,\ldots,\theta_{i-1}^0,\theta_i^0+\alpha\ell\frac{u_i}{\sqrt{n\Delta_n}},\theta_{i+1}^0+\frac{u_{i+1}}{\sqrt{n\Delta_n}},\ldots,\theta_m^0+\frac{u_m}{\sqrt{n\Delta_n}}),\\ \theta_j^{0+}(\alpha):=(\theta_1^0,\ldots,\theta_{j-1}^0,\theta_j^0+\alpha\frac{u_j}{\sqrt{n\Delta_n}},\theta_{j+1}^0+\frac{u_{j+1}}{\sqrt{n\Delta_n}},\ldots,\theta_m^0+\frac{u_m}{\sqrt{n\Delta_n}}).$$

Therefore,

$$\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}\Big[\xi_{i,k,n}|\widehat{\mathcal{F}}_{t_k}\Big]=-\sum_{k=0}^{n-1}\frac{u_i}{\sqrt{n\Delta_n}}(\partial_{\theta_i}b(\theta^0,X_{t_k}^{\theta^0}))^*(\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0})\Big(\frac{u_i}{2\sqrt{n\Delta_n}}\partial_{\theta_i}b(\theta^0,X_{t_k}^{\theta^0})\\ +\frac{u_{i+1}}{\sqrt{n\Delta_n}}\partial_{\theta_{i+1}}b(\theta^0,X_{t_k}^{\theta^0})+\cdots+\frac{u_m}{\sqrt{n\Delta_n}}\partial_{\theta_m}b(\theta^0,X_{t_k}^{\theta^0})\Big)\Delta_n\\ +\sum_{k=0}^{n-1}\xi_{i,1,k,n}+\sum_{k=0}^{n-1}(\eta_{i,k,n}+\eta_{i+1,k,n}+\cdots+\eta_{m,k,n})\\ =-\frac{u_i^2}{2n}\sum_{k=0}^{n-1}\Big(\partial_{\theta_i}b(\theta^0,X_{t_k}^{\theta^0})\Big)^*(\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0})\partial_{\theta_i}b(\theta^0,X_{t_k}^{\theta^0})$$

$$
\begin{aligned}
&- \frac{u_i u_{i+1}}{n} \sum_{k=0}^{n-1} \left(\partial_{\theta_i} b(\theta^0, X_{t_k}^{\theta^0}) \right)^* (\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0}) \partial_{\theta_{i+1}} b(\theta^0, X_{t_k}^{\theta^0}) \\
&- \cdots - \frac{u_i u_m}{n} \sum_{k=0}^{n-1} \left(\partial_{\theta_i} b(\theta^0, X_{t_k}^{\theta^0}) \right)^* (\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0}) \partial_{\theta_m} b(\theta^0, X_{t_k}^{\theta^0}) \\
&+ \sum_{k=0}^{n-1} \xi_{i,1,k,n} + \sum_{k=0}^{n-1} \left(\eta_{i,k,n} + \eta_{i+1,k,n} + \cdots + \eta_{m,k,n} \right),
\end{aligned}
$$

where for any $j \in \{i+1, \ldots, m\}$,

$$
\eta_{i,k,n} = -\frac{u_i^2}{n} \int_0^1 \int_0^1 (\partial_{\theta_i} b(\theta^0, X_{t_k}^{\theta^0}))^* (\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0}) \ell \Big(\partial_{\theta_i} b(\theta_i^{0+}(\alpha\ell), X_{t_k}^{\theta^0}) - \partial_{\theta_i} b(\theta^0, X_{t_k}^{\theta^0}) \Big) d\alpha d\ell,
$$

$$
\eta_{j,k,n} = -\frac{u_i u_j}{n} \int_0^1 (\partial_{\theta_i} b(\theta^0, X_{t_k}^{\theta^0}))^* (\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0}) \Big(\partial_{\theta_j} b(\theta_j^{0+}(\alpha), X_{t_k}^{\theta^0}) - \partial_{\theta_j} b(\theta^0, X_{t_k}^{\theta^0}) \Big) d\alpha.
$$

Using Lemma 3.9, as $n \to \infty$,

$$
\begin{aligned}
&- \frac{u_i^2}{2n} \sum_{k=0}^{n-1} \left(\partial_{\theta_i} b(\theta^0, X_{t_k}^{\theta^0}) \right)^* (\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0}) \partial_{\theta_i} b(\theta^0, X_{t_k}^{\theta^0}) \\
&\quad - \frac{u_i u_{i+1}}{n} \sum_{k=0}^{n-1} \left(\partial_{\theta_i} b(\theta^0, X_{t_k}^{\theta^0}) \right)^* (\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0}) \partial_{\theta_{i+1}} b(\theta^0, X_{t_k}^{\theta^0}) \\
&\quad - \cdots - \frac{u_i u_m}{n} \sum_{k=0}^{n-1} \left(\partial_{\theta_i} b(\theta^0, X_{t_k}^{\theta^0}) \right)^* (\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0}) \partial_{\theta_m} b(\theta^0, X_{t_k}^{\theta^0}) \\
&\xrightarrow{\widehat{\mathrm{P}}^{\theta^0}} -\frac{1}{2} u_i^2 \Gamma(\theta^0)_{i,i} - u_i u_{i+1} \Gamma(\theta^0)_{i,i+1} - \cdots - u_i u_m \Gamma(\theta^0)_{i,m}.
\end{aligned}
\tag{4.6}
$$

Next, using conditions **(A2)-(A3)**,

$$
\widehat{\mathrm{E}}^{\theta^0} \left[\left| \sum_{k=0}^{n-1} \xi_{i,1,k,n} \right| \right] \leq \sum_{k=0}^{n-1} \widehat{\mathrm{E}}^{\theta^0} \left[|\xi_{i,1,k,n}| \right] \leq \frac{C}{(\sqrt{n\Delta_n})^{\gamma}},
$$

which tends to zero. Similarly, for any $j \in \{i, \ldots, m\}$,

$$
\widehat{\mathrm{E}}^{\theta^0} \left[\left| \sum_{k=0}^{n-1} \eta_{j,k,n} \right| \right] \leq \sum_{k=0}^{n-1} \widehat{\mathrm{E}}^{\theta^0} \left[|\eta_{j,k,n}| \right] \leq \frac{C}{(\sqrt{n\Delta_n})^{\gamma}}.
$$

Thus, we have shown that as $n \to \infty$,

$$
\sum_{k=0}^{n-1} \xi_{i,1,k,n} + \sum_{k=0}^{n-1} \left(\eta_{i,k,n} + \eta_{i+1,k,n} + \cdots + \eta_{m,k,n} \right) \xrightarrow{\widehat{\mathrm{P}}^{\theta^0}} 0.
$$

Therefore, as $n \to \infty$,

$$
\sum_{k=0}^{n-1} \widehat{\mathrm{E}}^{\theta^0} \left[\xi_{i,k,n} | \widehat{\mathcal{F}}_{t_k} \right] \xrightarrow{\widehat{\mathrm{P}}^{\theta^0}} -\frac{1}{2} u_i^2 \Gamma(\theta^0)_{i,i} - u_i u_{i+1} \Gamma(\theta^0)_{i,i+1} - \cdots - u_i u_m \Gamma(\theta^0)_{i,m},
$$

which gives (4.3).

Proof of (4.4). First, from the previous computations,

$$\begin{aligned}
&\widehat{\mathrm{E}}^{\theta^0}\left[\xi_{i,k,n}|\widehat{\mathcal{F}}_{t_k}\right]\widehat{\mathrm{E}}^{\theta^0}\left[\xi_{j,k,n}|\widehat{\mathcal{F}}_{t_k}\right]=\frac{u_iu_j}{n}\Delta_n\int_0^1\left(\partial_{\theta_i}b(\theta_i^{0+}(\ell),X_{t_k}^{\theta^0})\right)^*(\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0})\\
&\quad\cdot\left(b(\theta^0,X_{t_k}^{\theta^0})-b(\theta_i^{0+}(\ell),X_{t_k}^{\theta^0})\right)d\ell\int_0^1\left(\partial_{\theta_j}b(\theta_j^{0+}(\ell),X_{t_k}^{\theta^0})\right)^*(\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0})\\
&\quad\cdot\left(b(\theta^0,X_{t_k}^{\theta^0})-b(\theta_j^{0+}(\ell),X_{t_k}^{\theta^0})\right)d\ell\\
&=\frac{u_iu_j}{n}\Delta_n\int_0^1\left(\partial_{\theta_i}b(\theta_i^{0+}(\ell),X_{t_k}^{\theta^0})\right)^*(\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0})\Big(\ell\frac{u_i}{\sqrt{n\Delta_n}}\int_0^1\partial_{\theta_i}b(\theta_i^{0+}(\alpha\ell),X_{t_k}^{\theta^0})d\alpha\\
&\quad+\frac{u_{i+1}}{\sqrt{n\Delta_n}}\int_0^1\partial_{\theta_{i+1}}b(\theta_{i+1}^{0+}(\alpha),X_{t_k}^{\theta^0})d\alpha+\cdots+\frac{u_m}{\sqrt{n\Delta_n}}\int_0^1\partial_{\theta_m}b(\theta_m^{0+}(\alpha),X_{t_k}^{\theta^0})d\alpha\Big)d\ell\\
&\quad\cdot\int_0^1\left(\partial_{\theta_j}b(\theta_j^{0+}(\ell),X_{t_k}^{\theta^0})\right)^*(\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0})\Big(\ell\frac{u_j}{\sqrt{n\Delta_n}}\int_0^1\partial_{\theta_j}b(\theta_j^{0+}(\alpha\ell),X_{t_k}^{\theta^0})d\alpha\\
&\quad+\frac{u_{j+1}}{\sqrt{n\Delta_n}}\int_0^1\partial_{\theta_{j+1}}b(\theta_{j+1}^{0+}(\alpha),X_{t_k}^{\theta^0})d\alpha+\cdots+\frac{u_m}{\sqrt{n\Delta_n}}\int_0^1\partial_{\theta_m}b(\theta_m^{0+}(\alpha),X_{t_k}^{\theta^0})d\alpha\Big)d\ell.
\end{aligned}$$

Thus,

$$\begin{aligned}
\left|\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}\left[\xi_{i,k,n}|\widehat{\mathcal{F}}_{t_k}\right]\widehat{\mathrm{E}}^{\theta^0}\left[\xi_{j,k,n}|\widehat{\mathcal{F}}_{t_k}\right]\right|&\leq\sum_{k=0}^{n-1}\left|\widehat{\mathrm{E}}^{\theta^0}\left[\xi_{i,k,n}|\widehat{\mathcal{F}}_{t_k}\right]\widehat{\mathrm{E}}^{\theta^0}\left[\xi_{j,k,n}|\widehat{\mathcal{F}}_{t_k}\right]\right|\\
&\leq\frac{C}{n^2}\sum_{k=0}^{n-1}\left(1+|X_{t_k}|^q\right),
\end{aligned}$$

for some constant $q>0$, which converges to zero in $\widehat{\mathrm{P}}^{\theta^0}$-probability as $n\to\infty$. Thus, as $n\to\infty$,

$$\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}\left[\xi_{i,k,n}|\widehat{\mathcal{F}}_{t_k}\right]\widehat{\mathrm{E}}^{\theta^0}\left[\xi_{j,k,n}|\widehat{\mathcal{F}}_{t_k}\right]\xrightarrow{\widehat{\mathrm{P}}^{\theta^0}}0.\tag{4.7}$$

Next,

$$\begin{aligned}
&\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}\left[\xi_{i,k,n}\xi_{j,k,n}|\widehat{\mathcal{F}}_{t_k}\right]=\frac{u_iu_j}{n\Delta_n}\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}\Big[\int_0^1\int_0^1\left(\partial_{\theta_i}b(\theta_i^{0+}(\ell),X_{t_k}^{\theta^0})\right)^*(\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0})\\
&\cdot\Big(\sigma(X_{t_k}^{\theta^0})\left(B_{t_{k+1}}-B_{t_k}\right)+\left(b(\theta^0,X_{t_k}^{\theta^0})-b(\theta_i^{0+}(\ell),X_{t_k}^{\theta^0})\right)\Delta_n\Big)\Big(\sigma(X_{t_k}^{\theta^0})\left(B_{t_{k+1}}-B_{t_k}\right)\\
&+\left(b(\theta^0,X_{t_k}^{\theta^0})-b(\theta_j^{0+}(\ell'),X_{t_k}^{\theta^0})\right)\Delta_n\Big)^*(\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0})\partial_{\theta_j}b(\theta_j^{0+}(\ell'),X_{t_k}^{\theta^0})d\ell d\ell'|\widehat{\mathcal{F}}_{t_k}\Big]\\
&=\frac{u_iu_j}{n}\sum_{k=0}^{n-1}\int_0^1\int_0^1\left(\partial_{\theta_i}b(\theta_i^{0+}(\ell),X_{t_k}^{\theta^0})\right)^*(\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0})\partial_{\theta_j}b(\theta_j^{0+}(\ell'),X_{t_k}^{\theta^0})d\ell d\ell'\\
&+\frac{u_iu_j}{n}\Delta_n\sum_{k=0}^{n-1}\int_0^1\int_0^1\left(\partial_{\theta_i}b(\theta_i^{0+}(\ell),X_{t_k}^{\theta^0})\right)^*(\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0})\left(b(\theta^0,X_{t_k}^{\theta^0})-b(\theta_i^{0+}(\ell),X_{t_k}^{\theta^0})\right)\\
&\cdot\left(b(\theta^0,X_{t_k}^{\theta^0})-b(\theta_j^{0+}(\ell'),X_{t_k}^{\theta^0})\right)^*(\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0})\partial_{\theta_j}b(\theta_j^{0+}(\ell'),X_{t_k}^{\theta^0})d\ell d\ell'
\end{aligned}$$

$$= \frac{u_i u_j}{n} \sum_{k=0}^{n-1} \Big(\partial_{\theta_i} b(\theta^0, X_{t_k}^{\theta^0})\Big)^* (\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0}) \partial_{\theta_j} b(\theta^0, X_{t_k}^{\theta^0}) + \sum_{k=0}^{n-1} \Big(H_7^{i,k} + H_8^{i,k} + H_9^{i,k}\Big),$$

where $\theta_j^{0+}(\ell') := (\theta_1^0, \ldots, \theta_{j-1}^0, \theta_j^0 + \ell' \frac{u_j}{\sqrt{n\Delta_n}}, \theta_{j+1}^0 + \frac{u_{j+1}}{\sqrt{n\Delta_n}}, \ldots, \theta_m^0 + \frac{u_m}{\sqrt{n\Delta_n}})$ and

$$H_7^{i,k} = \frac{u_i u_j}{n} \int_0^1 \int_0^1 \Big(\partial_{\theta_i} b(\theta_i^{0+}(\ell), X_{t_k}^{\theta^0}) - \partial_{\theta_i} b(\theta^0, X_{t_k}^{\theta^0})\Big)^* (\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0}) \partial_{\theta_j} b(\theta^0, X_{t_k}^{\theta^0}) d\ell d\ell',$$

$$H_8^{i,k} = \frac{u_i u_j}{n} \int_0^1 \int_0^1 (\partial_{\theta_i} b(\theta_i^{0+}(\ell), X_{t_k}^{\theta^0}))^* (\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0}) \big(\partial_{\theta_j} b(\theta_j^{0+}(\ell'), X_{t_k}^{\theta^0}) - \partial_{\theta_j} b(\theta^0, X_{t_k}^{\theta^0})\big) d\ell d\ell',$$

$$H_9^{i,k} = \frac{u_i u_j}{n} \Delta_n \int_0^1 \int_0^1 \Big(\partial_{\theta_i} b(\theta_i^{0+}(\ell), X_{t_k}^{\theta^0})\Big)^* (\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0}) \Big(b(\theta^0, X_{t_k}^{\theta^0}) - b(\theta_i^{0+}(\ell), X_{t_k}^{\theta^0})\Big)$$
$$\cdot \Big(b(\theta^0, X_{t_k}^{\theta^0}) - b(\theta_j^{0+}(\ell'), X_{t_k}^{\theta^0})\Big)^* (\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0}) \partial_{\theta_j} b(\theta_j^{0+}(\ell'), X_{t_k}^{\theta^0}) d\ell d\ell'.$$

Again, by Lemma 3.9, as $n \to \infty$,

$$\frac{1}{n} \sum_{k=0}^{n-1} \Big(\partial_{\theta_i} b(\theta^0, X_{t_k}^{\theta^0})\Big)^* (\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0}) \partial_{\theta_j} b(\theta^0, X_{t_k}^{\theta^0}) \xrightarrow{\widehat{\mathrm{P}}^{\theta^0}} \Gamma(\theta^0)_{i,j}.$$

Next, using conditions **(A2)**-**(A3)**,

$$\widehat{\mathrm{E}}^{\theta^0} \left[\left|\sum_{k=0}^{n-1} \Big(H_7^{i,k} + H_8^{i,k}\Big)\right|\right] \leq \sum_{k=0}^{n-1} \widehat{\mathrm{E}}^{\theta^0} \left[\left|H_7^{i,k} + H_8^{i,k}\right|\right] \leq \frac{C}{(\sqrt{n\Delta_n})^{\gamma}},$$

and

$$\widehat{\mathrm{E}}^{\theta^0} \left[\left|\sum_{k=0}^{n-1} H_9^{i,k}\right|\right] \leq \sum_{k=0}^{n-1} \widehat{\mathrm{E}}^{\theta^0} \left[\left|H_9^{i,k}\right|\right] \leq \frac{C}{n}.$$

Hence, as $n \to \infty$,

$$\sum_{k=0}^{n-1} \Big(H_7^{i,k} + H_8^{i,k} + H_9^{i,k}\Big) \xrightarrow{\widehat{\mathrm{P}}^{\theta^0}} 0.$$

Therefore, as $n \to \infty$,

$$\sum_{k=0}^{n-1} \widehat{\mathrm{E}}^{\theta^0} \left[\xi_{i,k,n} \xi_{j,k,n} | \widehat{\mathcal{F}}_{t_k}\right] \xrightarrow{\widehat{\mathrm{P}}^{\theta^0}} u_i u_j \Gamma(\theta^0)_{i,j},$$

which, together with (4.7), gives (4.4).

Proof of (4.5). Basic computations yield

$$\sum_{k=0}^{n-1} \widehat{\mathrm{E}}^{\theta^0} \left[(\xi_{i,k,n})^4 | \widehat{\mathcal{F}}_{t_k}\right] \leq \frac{C u_i^4}{n^2} \sum_{k=0}^{n-1} \Big(1 + |X_{t_k}^{\theta^0}|^q\Big),$$

for some constants $C, q > 0$. The proof of Lemma 4.2 is completed. □

4.3. **Negligible contributions.**

Lemma 4.3. *Under conditions* **(A1)**-**(A5)**, *as* $n\to\infty$,

$$\sum_{k=0}^{n-1}\sum_{i=1}^{m}\frac{u_i}{\sqrt{n\Delta_n^3}}\int_0^1\Big\{H_4^{\theta^0,i,k}+H_5^{\theta^0,i,k}+H_6^{\theta^0,i,k}$$
$$+\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta_i^{0+}(\ell)}\left[R^{\theta_i^{0+}(\ell),k}-R_4^{\theta_i^{0+}(\ell),k}-R_5^{\theta_i^{0+}(\ell),k}-R_6^{\theta_i^{0+}(\ell),k}\Big|Y_{t_{k+1}}^{\theta_i^{0+}(\ell)}=X_{t_{k+1}}^{\theta^0}\right]\Big\}d\ell\overset{\widehat{\mathrm{P}}^{\theta^0}}{\longrightarrow}0.$$

Proof. The proof is completed by combining the four Lemmas 4.4-4.7 below. □

Consequently, the proof of Theorem 2.2 is now completed from Lemmas 4.1, 4.2 and 4.3.

Lemma 4.4. *Under conditions* **(A1)**, **(A2)**, **(A3)**(a)-(c), **(A4)** *and* **(A5)**, *as* $n\to\infty$,

$$\sum_{k=0}^{n-1}\sum_{i=1}^{m}\frac{u_i}{\sqrt{n\Delta_n^3}}\int_0^1\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta_i^{0+}(\ell)}\left[R^{\theta_i^{0+}(\ell),k}\Big|Y_{t_{k+1}}^{\theta_i^{0+}(\ell)}=X_{t_{k+1}}^{\theta^0}\right]d\ell\overset{\widehat{\mathrm{P}}^{\theta^0}}{\longrightarrow}0.$$

Proof. It suffices to show that conditions (i) and (ii) of Lemma 3.10 a) hold under the measure $\widehat{\mathrm{P}}^{\theta^0}$ applied to the random variable

$$\zeta_{i,k,n}:=\frac{u_i}{\sqrt{n\Delta_n^3}}\int_0^1\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta_i^{0+}(\ell)}\left[R^{\theta_i^{0+}(\ell),k}\Big|Y_{t_{k+1}}^{\theta_i^{0+}(\ell)}=X_{t_{k+1}}^{\theta^0}\right]d\ell,$$

for any $i\in\{1,\ldots,m\}$. We start showing (i) of Lemma 3.10 a). Applying Lemma 3.6 to $\theta=\theta_i^{0+}(\ell)$ and $V=R^{\theta_i^{0+}(\ell),k}$, and using the fact that, by (3.7), $\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta_i^{0+}(\ell)}[R^{\theta_i^{0+}(\ell),k}]=0$, we obtain that

$$\begin{aligned}\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}\left[\zeta_{i,k,n}|\widehat{\mathcal{F}}_{t_k}\right]&=\sum_{k=0}^{n-1}\frac{u_i}{\sqrt{n\Delta_n^3}}\int_0^1\widehat{\mathrm{E}}^{\theta^0}\left[\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta_i^{0+}(\ell)}\left[R^{\theta_i^{0+}(\ell),k}|Y_{t_{k+1}}^{\theta_i^{0+}(\ell)}=X_{t_{k+1}}^{\theta^0}\right]|\widehat{\mathcal{F}}_{t_k}\right]d\ell\\&=\sum_{k=0}^{n-1}\frac{u_i}{\sqrt{n\Delta_n^3}}\int_0^1\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta_i^{0+}(\ell)}[R^{\theta_i^{0+}(\ell),k}]d\ell=0.\end{aligned}$$

Thus, the term appearing in condition (i) of Lemma 3.10 a) actually equals zero.

Next, applying Jensen's inequality and Lemma 3.6 to $\theta=\theta_i^{0+}(\ell)$ and $V=(R^{\theta_i^{0+}(\ell),k})^2$, and (3.8), we obtain that

$$\begin{aligned}\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}\left[\zeta_{i,k,n}^2|\widehat{\mathcal{F}}_{t_k}\right]&=\sum_{k=0}^{n-1}\frac{u_i^2}{n\Delta_n^3}\widehat{\mathrm{E}}^{\theta^0}\left[\left(\int_0^1\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta_i^{0+}(\ell)}\left[R^{\theta_i^{0+}(\ell),k}\Big|Y_{t_{k+1}}^{\theta_i^{0+}(\ell)}=X_{t_{k+1}}^{\theta^0}\right]d\ell\right)^2|\widehat{\mathcal{F}}_{t_k}\right]\\&\leq\sum_{k=0}^{n-1}\frac{u_i^2}{n\Delta_n^3}\int_0^1\widehat{\mathrm{E}}^{\theta^0}\left[\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta_i^{0+}(\ell)}\left[\left(R^{\theta_i^{0+}(\ell),k}\right)^2|Y_{t_{k+1}}^{\theta_i^{0+}(\ell)}=X_{t_{k+1}}^{\theta^0}\right]|\widehat{\mathcal{F}}_{t_k}\right]d\ell\\&=\sum_{k=0}^{n-1}\frac{u_i^2}{n\Delta_n^3}\int_0^1\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta_i^{0+}(\ell)}\left[\left(R^{\theta_i^{0+}(\ell),k}\right)^2\right]d\ell\leq Cu_i^2\frac{\sqrt{\Delta_n}}{n}\sum_{k=0}^{n-1}\left(1+|X_{t_k}^{\theta^0}|^q\right),\end{aligned}$$

for some constant $q>0$, which converges to zero in $\widehat{\mathrm{P}}^{\theta^0}$-probability as $n\to\infty$. Thus, $\sum_{k=0}^{n-1}\zeta_{i,k,n}\overset{\widehat{\mathrm{P}}^{\theta^0}}{\longrightarrow}0$ for any $i\in\{1,\ldots,m\}$. Thus, the result follows. □

Lemma 4.5. *Assume conditions* **(A1)**, **(A2)**, **(A3)**(a)-(b), **(A4)** *and* **(A5)**. *Then as* $n \to \infty$,

$$\sum_{k=0}^{n-1}\sum_{i=1}^{m}\frac{u_i}{\sqrt{n\Delta_n^3}}\int_0^1\left\{H_4^{\theta^0,i,k}-\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta_i^{0+}(\ell)}\left[R_4^{\theta_i^{0+}(\ell),k}\Big|Y_{t_{k+1}}^{\theta_i^{0+}(\ell)}=X_{t_{k+1}}^{\theta^0}\right]\right\}d\ell\xrightarrow{\widehat{\mathrm{P}}^{\theta^0}}0.$$

Proof. We rewrite

$$\begin{aligned}
&H_4^{\theta^0,i,k}-\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta_i^{0+}(\ell)}\left[R_4^{\theta_i^{0+}(\ell),k}\Big|Y_{t_{k+1}}^{\theta_i^{0+}(\ell)}=X_{t_{k+1}}^{\theta^0}\right]\\
&=\Delta_n(\partial_{\theta_i}b(\theta_i^{0+}(\ell),X_{t_k}^{\theta^0}))^*(\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0})\bigg(\int_{t_k}^{t_{k+1}}\Big(b(\theta^0,X_s^{\theta^0})-b(\theta^0,X_{t_k}^{\theta^0})\Big)ds\\
&\quad-\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta_i^{0+}(\ell)}\left[\int_{t_k}^{t_{k+1}}\Big(b(\theta_i^{0+}(\ell),Y_s^{\theta_i^{0+}(\ell)})-b(\theta_i^{0+}(\ell),Y_{t_k}^{\theta_i^{0+}(\ell)})\Big)ds\Big|Y_{t_{k+1}}^{\theta_i^{0+}(\ell)}=X_{t_{k+1}}^{\theta^0}\right]\bigg)\\
&=\Delta_n(\partial_{\theta_i}b(\theta_i^{0+}(\ell),X_{t_k}^{\theta^0}))^*(\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0})\left(M_{i,1,k,n}+M_{i,2,k,n}\right),
\end{aligned}$$

where

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

Thus,

$$\begin{aligned}
\zeta_{i,k,n}:&=\frac{u_i}{\sqrt{n\Delta_n^3}}\int_0^1\left\{H_4^{\theta^0,i,k}-\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta_i^{0+}(\ell)}\left[R_4^{\theta_i^{0+}(\ell),k}\Big|Y_{t_{k+1}}^{\theta_i^{0+}(\ell)}=X_{t_{k+1}}^{\theta^0}\right]\right\}d\ell\\
&=\frac{u_i}{\sqrt{n\Delta_n^3}}\int_0^1\Delta_n(\partial_{\theta_i}b(\theta_i^{0+}(\ell),X_{t_k}^{\theta^0}))^*(\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0})\left(M_{i,1,k,n}+M_{i,2,k,n}\right)d\ell\\
&=\zeta_{i,1,k,n}+\zeta_{i,2,k,n},
\end{aligned}$$

where

$$\begin{aligned}
\zeta_{i,1,k,n}&=\frac{u_i}{\sqrt{n\Delta_n}}\int_0^1(\partial_{\theta_i}b(\theta_i^{0+}(\ell),X_{t_k}^{\theta^0}))^*(\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0})M_{i,1,k,n}d\ell,\\
\zeta_{i,2,k,n}&=\frac{u_i}{\sqrt{n\Delta_n}}\int_0^1(\partial_{\theta_i}b(\theta_i^{0+}(\ell),X_{t_k}^{\theta^0}))^*(\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0})M_{i,2,k,n}d\ell.
\end{aligned}$$

We are going to show that $\sum_{k=0}^{n-1}\zeta_{i,1,k,n}\xrightarrow{\widehat{\mathrm{P}}^{\theta^0}}0$ and $\sum_{k=0}^{n-1}\zeta_{i,2,k,n}\xrightarrow{\widehat{\mathrm{P}}^{\theta^0}}0$.

First, using the mean value theorem,

$$\begin{aligned}
&b(\theta^0,X_s^{\theta^0})-b(\theta_i^{0+}(\ell),X_s^{\theta^0})=b(\theta_{i+1}^{0+},X_s^{\theta^0})-b(\theta_i^{0+}(\ell),X_s^{\theta^0})+b(\theta_{i+2}^{0+},X_s^{\theta^0})-b(\theta_{i+1}^{0+},X_s^{\theta^0})\\
&\qquad+\cdots+b(\theta_m^{0+},X_s^{\theta^0})-b(\theta_{m-1}^{0+},X_s^{\theta^0})+b(\theta^0,X_s^{\theta^0})-b(\theta_m^{0+},X_s^{\theta^0})\\
&=-\ell\frac{u_i}{\sqrt{n\Delta_n}}\int_0^1\partial_{\theta_i}b(\theta_i^{0+}(\alpha\ell),X_s^{\theta^0})d\alpha-\frac{u_{i+1}}{\sqrt{n\Delta_n}}\int_0^1\partial_{\theta_{i+1}}b(\theta_{i+1}^{0+}(\alpha),X_s^{\theta^0})d\alpha
\end{aligned}$$

$$- \cdots - \frac{u_m}{\sqrt{n\Delta_n}} \int_0^1 \partial_{\theta_m} b(\theta_m^{0+}(\alpha), X_s^{\theta^0}) d\alpha.$$

Therefore,

$$\begin{aligned}
&b(\theta^0, X_s^{\theta^0}) - b(\theta^0, X_{t_k}^{\theta^0}) - \left(b(\theta_i^{0+}(\ell), X_s^{\theta^0}) - b(\theta_i^{0+}(\ell), X_{t_k}^{\theta^0})\right) \\
&= b(\theta^0, X_s^{\theta^0}) - b(\theta_i^{0+}(\ell), X_s^{\theta^0}) - \left(b(\theta^0, X_{t_k}^{\theta^0}) - b(\theta_i^{0+}(\ell), X_{t_k}^{\theta^0})\right) \\
&= -\ell \frac{u_i}{\sqrt{n\Delta_n}} \int_0^1 \left(\partial_{\theta_i} b(\theta_i^{0+}(\alpha\ell), X_s^{\theta^0}) - \partial_{\theta_i} b(\theta_i^{0+}(\alpha\ell), X_{t_k}^{\theta^0})\right) d\alpha \\
&\quad - \frac{u_{i+1}}{\sqrt{n\Delta_n}} \int_0^1 \left(\partial_{\theta_{i+1}} b(\theta_{i+1}^{0+}(\alpha), X_s^{\theta^0}) - \partial_{\theta_{i+1}} b(\theta_{i+1}^{0+}(\alpha), X_{t_k}^{\theta^0})\right) d\alpha \\
&\quad - \cdots - \frac{u_m}{\sqrt{n\Delta_n}} \int_0^1 \left(\partial_{\theta_m} b(\theta_m^{0+}(\alpha), X_s^{\theta^0}) - \partial_{\theta_m} b(\theta_m^{0+}(\alpha), X_{t_k}^{\theta^0})\right) d\alpha.
\end{aligned}$$

Next, using the mean value theorem for vector-valued functions,

$$\partial_{\theta_j} b(\theta_j^{0+}(\alpha), X_s^{\theta^0}) - \partial_{\theta_j} b(\theta_j^{0+}(\alpha), X_{t_k}^{\theta^0}) = \left(\int_0^1 J_{\partial_{\theta_j} b}(X_{t_k}^{\theta^0} + v(X_s^{\theta^0} - X_{t_k}^{\theta^0})) dv\right) \cdot (X_s^{\theta^0} - X_{t_k}^{\theta^0}),$$

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

$$J_{\partial_{\theta_j} b}(X_{t_k}^{\theta^0} + v(X_s^{\theta^0} - X_{t_k}^{\theta^0})) = \begin{pmatrix} \partial^2_{\theta_j x_1} b_1 & \ldots & \partial^2_{\theta_j x_d} b_1 \\ \vdots & \ddots & \vdots \\ \partial^2_{\theta_j x_1} b_d & \ldots & \partial^2_{\theta_j x_d} b_d \end{pmatrix} (\theta_j^{0+}(\alpha), X_{t_k}^{\theta^0} + v(X_s^{\theta^0} - X_{t_k}^{\theta^0})).$$

Then, using conditions **(A2)**-**(A3)** and Lemma 3.4 (i), we get that

$$\sum_{k=0}^{n-1} \widehat{\mathrm{E}}^{\theta^0}\left[|\zeta_{i,1,k,n}| \,|\widehat{\mathcal{F}}_{t_k}\right] \leq C \frac{\sqrt{\Delta_n}}{n} \sum_{k=0}^{n-1} \left(1 + |X_{t_k}^{\theta^0}|^q\right),$$

for some constant $q > 0$, which converges to zero in $\widehat{\mathrm{P}}^{\theta^0}$-probability as $n \to \infty$. Thus, by Lemma 3.10 b), $\sum_{k=0}^{n-1} \zeta_{i,1,k,n} \xrightarrow{\widehat{\mathrm{P}}^{\theta^0}} 0$ for any $i \in \{1, \ldots, m\}$.

Next, using Girsanov's theorem and Lemma 3.6, we get that

$$\begin{aligned}
&\widehat{\mathrm{E}}^{\theta^0}\left[M_{i,2,k,n} | \widehat{\mathcal{F}}_{t_k}\right] = \widehat{\mathrm{E}}^{\theta^0}\Bigg[\int_{t_k}^{t_{k+1}} \left(b(\theta_i^{0+}(\ell), X_s^{\theta^0}) - b(\theta_i^{0+}(\ell), X_{t_k}^{\theta^0})\right) ds \\
&\quad - \widetilde{\mathrm{E}}_{t_k, X_{t_k}^{\theta^0}}^{\theta_i^{0+}(\ell)}\left[\int_{t_k}^{t_{k+1}} \left(b(\theta_i^{0+}(\ell), Y_s^{\theta_i^{0+}(\ell)}) - b(\theta_i^{0+}(\ell), Y_{t_k}^{\theta_i^{0+}(\ell)})\right) ds \Big| Y_{t_{k+1}}^{\theta_i^{0+}(\ell)} = X_{t_{k+1}}^{\theta^0}\right] |\widehat{\mathcal{F}}_{t_k}\Bigg] \\
&= \widehat{\mathrm{E}}_{t_k, X_{t_k}^{\theta^0}}^{\theta^0}\left[\int_{t_k}^{t_{k+1}} \left(b(\theta_i^{0+}(\ell), X_s^{\theta^0}) - b(\theta_i^{0+}(\ell), X_{t_k}^{\theta^0})\right) ds\right] \\
&\quad - \widehat{\mathrm{E}}_{t_k, X_{t_k}^{\theta^0}}^{\theta^0}\left[\widetilde{\mathrm{E}}_{t_k, X_{t_k}^{\theta^0}}^{\theta_i^{0+}(\ell)}\left[\int_{t_k}^{t_{k+1}} \left(b(\theta_i^{0+}(\ell), Y_s^{\theta_i^{0+}(\ell)}) - b(\theta_i^{0+}(\ell), Y_{t_k}^{\theta_i^{0+}(\ell)})\right) ds \Big| Y_{t_{k+1}}^{\theta_i^{0+}(\ell)} = X_{t_{k+1}}^{\theta^0}\right]\right] \\
&= \widehat{\mathrm{E}}_{t_k, X_{t_k}^{\theta^0}}^{\theta_i^{0+}(\ell)}\left[\int_{t_k}^{t_{k+1}} \left(b(\theta_i^{0+}(\ell), X_s^{\theta_i^{0+}(\ell)}) - b(\theta_i^{0+}(\ell), X_{t_k}^{\theta_i^{0+}(\ell)})\right) ds \, \frac{d\widehat{\mathrm{P}}_{t_k, X_{t_k}^{\theta^0}}^{\theta^0}}{d\widehat{\mathrm{P}}_{t_k, X_{t_k}^{\theta^0}}^{\theta_i^{0+}(\ell)}}\right]
\end{aligned}$$

$$
\begin{aligned}
&\quad - \widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k, X^{\theta^0}_{t_k}} \left[\int_{t_k}^{t_{k+1}} \Big(b(\theta_i^{0+}(\ell), Y_s^{\theta_i^{0+}(\ell)}) - b(\theta_i^{0+}(\ell), Y_{t_k}^{\theta_i^{0+}(\ell)}) \Big) ds \right] \\
&= \widehat{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k, X^{\theta^0}_{t_k}} \left[\int_{t_k}^{t_{k+1}} \big(b(\theta_i^{0+}(\ell), X_s^{\theta_i^{0+}(\ell)}) - b(\theta_i^{0+}(\ell), X_{t_k}^{\theta_i^{0+}(\ell)}) \big) ds \, \Big(\frac{d\widehat{\mathrm{P}}^{\theta^0}_{t_k, X^{\theta^0}_{t_k}}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k, X^{\theta^0}_{t_k}}} - 1 \Big) \right] \\
&\quad + \widehat{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k, X^{\theta^0}_{t_k}} \left[\int_{t_k}^{t_{k+1}} \big(b(\theta_i^{0+}(\ell), X_s^{\theta_i^{0+}(\ell)}) - b(\theta_i^{0+}(\ell), X_{t_k}^{\theta_i^{0+}(\ell)}) \big) ds \right] \\
&\quad - \widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k, X^{\theta^0}_{t_k}} \left[\int_{t_k}^{t_{k+1}} \Big(b(\theta_i^{0+}(\ell), Y_s^{\theta_i^{0+}(\ell)}) - b(\theta_i^{0+}(\ell), Y_{t_k}^{\theta_i^{0+}(\ell)}) \Big) ds \right] \\
&= \widehat{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k, X^{\theta^0}_{t_k}} \left[\int_{t_k}^{t_{k+1}} \big(b(\theta_i^{0+}(\ell), X_s^{\theta_i^{0+}(\ell)}) - b(\theta_i^{0+}(\ell), X_{t_k}^{\theta_i^{0+}(\ell)}) \big) ds \, \Big(\frac{d\widehat{\mathrm{P}}^{\theta^0}_{t_k, X^{\theta^0}_{t_k}}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k, X^{\theta^0}_{t_k}}} - 1 \Big) \right],
\end{aligned}
$$

where we have used the fact that $X^{\theta_i^{0+}(\ell)}$ is the independent copy of $Y^{\theta_i^{0+}(\ell)}$. Here, to simplify the exposition, we write $\frac{d\widehat{\mathrm{P}}^{\theta^0}_{t_k, X^{\theta^0}_{t_k}}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k, X^{\theta^0}_{t_k}}} = \frac{d\widehat{\mathrm{P}}^{\theta^0}_{t_k, X^{\theta^0}_{t_k}}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k, X^{\theta^0}_{t_k}}} \big((X_t^{\theta_i^{0+}(\ell)})_{t \in [t_k, t_{k+1}]} \big)$.

Then, using Lemma 3.7 with $q = 2$, conditions **(A1)-(A2)** and Lemma 3.4 (i), we get that

$$
\begin{aligned}
&\left| \sum_{k=0}^{n-1} \widehat{\mathrm{E}}^{\theta^0} [\zeta_{i,2,k,n} | \widehat{\mathcal{F}}_{t_k}] \right| = \left| \sum_{k=0}^{n-1} \frac{u_i}{\sqrt{n\Delta_n}} \int_0^1 (\partial_{\theta_i} b(\theta_i^{0+}(\ell), X^{\theta^0}_{t_k}))^* (\sigma\sigma^*)^{-1} (X^{\theta^0}_{t_k}) \widehat{\mathrm{E}}^{\theta^0} [M_{i,2,k,n} | \widehat{\mathcal{F}}_{t_k}] d\ell \right| \\
&= \left| \frac{u_i}{\sqrt{n\Delta_n}} \sum_{k=0}^{n-1} \int_0^1 \widehat{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k, X^{\theta^0}_{t_k}} \Big[(\partial_{\theta_i} b(\theta_i^{0+}(\ell), X^{\theta^0}_{t_k}))^* (\sigma\sigma^*)^{-1} (X^{\theta^0}_{t_k}) \right. \\
&\qquad \left. \cdot \int_{t_k}^{t_{k+1}} \big(b(\theta_i^{0+}(\ell), X_s^{\theta_i^{0+}(\ell)}) - b(\theta_i^{0+}(\ell), X_{t_k}^{\theta_i^{0+}(\ell)}) \big) ds \, \Big(\frac{d\widehat{\mathrm{P}}^{\theta^0}_{t_k, X^{\theta^0}_{t_k}}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k, X^{\theta^0}_{t_k}}} - 1 \Big) \Big] d\ell \right| \\
&\le \frac{|u_i|}{\sqrt{n\Delta_n}} \sum_{k=0}^{n-1} \int_0^1 \left| \widehat{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k, X^{\theta^0}_{t_k}} \Big[(\partial_{\theta_i} b(\theta_i^{0+}(\ell), X^{\theta^0}_{t_k}))^* (\sigma\sigma^*)^{-1} (X^{\theta^0}_{t_k}) \right. \\
&\qquad \left. \cdot \int_{t_k}^{t_{k+1}} \big(b(\theta_i^{0+}(\ell), X_s^{\theta_i^{0+}(\ell)}) - b(\theta_i^{0+}(\ell), X_{t_k}^{\theta_i^{0+}(\ell)}) \big) ds \, \Big(\frac{d\widehat{\mathrm{P}}^{\theta^0}_{t_k, X^{\theta^0}_{t_k}}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k, X^{\theta^0}_{t_k}}} - 1 \Big) \Big] \right| d\ell \\
&\le C \frac{|u_i|}{\sqrt{n\Delta_n}} \sum_{k=0}^{n-1} \int_0^1 \sqrt{\Delta_n} \Big(1 + |X^{\theta^0}_{t_k}|^{q_1} \Big) \left(\left| \int_{\theta_i^0 + \ell \frac{u_i}{\sqrt{n\Delta_n}}}^{\theta_i^0} \Big(\widehat{\mathrm{E}}^{\theta_i(0+)}_{t_k, X^{\theta^0}_{t_k}} \big[|V|^2 \big] \Big)^{\frac{1}{2}} d\theta_i \right| \right. \\
&\quad \left. + \left| \int_{\theta_{i+1}^0 + \frac{u_{i+1}}{\sqrt{n\Delta_n}}}^{\theta_{i+1}^0} \Big(\widehat{\mathrm{E}}^{\theta_{i+1}(0+)}_{t_k, X^{\theta^0}_{t_k}} \big[|V|^2 \big] \Big)^{\frac{1}{2}} d\theta_{i+1} \right| + \cdots + \left| \int_{\theta_m^0 + \frac{u_m}{\sqrt{n\Delta_n}}}^{\theta_m^0} \Big(\widehat{\mathrm{E}}^{\theta_m(0+)}_{t_k, X^{\theta^0}_{t_k}} \big[|V|^2 \big] \Big)^{\frac{1}{2}} d\theta_m \right| \right) d\ell
\end{aligned}
$$

$$\leq C\frac{\Delta_n}{n}\sum_{k=0}^{n-1}\left(1+|X_{t_k}^{\theta^0}|^q\right),$$

for some constants $q_1>0$, $q>0$, which converges to zero in $\widehat{\mathrm{P}}^{\theta^0}$-probability as $n\to\infty$. Here,

$$V:=(\partial_{\theta_i}b(\theta_i^{0+}(\ell),X_{t_k}^{\theta^0}))^*(\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0})\int_{t_k}^{t_{k+1}}\left(b(\theta_i^{0+}(\ell),X_s^{\theta_i^{0+}(\ell)})-b(\theta_i^{0+}(\ell),X_{t_k}^{\theta_i^{0+}(\ell)})\right)ds.$$

and we have used the mean value theorem for vector-valued functions,

$$\begin{aligned}&b(\theta_i^{0+}(\ell),X_s^{\theta_i^{0+}(\ell)})-b(\theta_i^{0+}(\ell),X_{t_k}^{\theta_i^{0+}(\ell)})\\&=\left(\int_0^1 J_b(X_{t_k}^{\theta_i^{0+}(\ell)}+v(X_s^{\theta_i^{0+}(\ell)}-X_{t_k}^{\theta_i^{0+}(\ell)}))dv\right)\cdot(X_s^{\theta_i^{0+}(\ell)}-X_{t_k}^{\theta_i^{0+}(\ell)}),\end{aligned}$$

where the Jacobian matrix is given by

$$\begin{aligned}&J_b(X_{t_k}^{\theta_i^{0+}(\ell)}+v(X_s^{\theta_i^{0+}(\ell)}-X_{t_k}^{\theta_i^{0+}(\ell)}))\\&=\begin{pmatrix}\partial_{x_1}b_1&\ldots&\partial_{x_d}b_1\\\vdots&\ddots&\vdots\\\partial_{x_1}b_d&\ldots&\partial_{x_d}b_d\end{pmatrix}(\theta_i^{0+}(\ell),X_{t_k}^{\theta_i^{0+}(\ell)}+v(X_s^{\theta_i^{0+}(\ell)}-X_{t_k}^{\theta_i^{0+}(\ell)})).\end{aligned}$$

Therefore, $\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}[\zeta_{i,2,k,n}|\widehat{\mathcal{F}}_{t_k}]\overset{\widehat{\mathrm{P}}^{\theta^0}}{\longrightarrow}0$ as $n\to\infty$.

Next, applying Jensen's inequality and Lemma 3.6, conditions **(A1)**-**(A2)**, the mean value theorem for vector-valued functions and Lemma 3.4 (i), we obtain that

$$\begin{aligned}&\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}[\zeta_{i,2,k,n}^2|\widehat{\mathcal{F}}_{t_k}]=\frac{u_i^2}{n\Delta_n}\sum_{k=0}^{n-1}\widehat{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta^0}\left[\left(\int_0^1(\partial_{\theta_i}b(\theta_i^{0+}(\ell),X_{t_k}^{\theta^0}))^*(\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0})M_{i,2,k,n}d\ell\right)^2\right]\\
&\leq\frac{u_i^2}{n\Delta_n}\sum_{k=0}^{n-1}\int_0^1\widehat{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta^0}\left[\left|(\partial_{\theta_i}b(\theta_i^{0+}(\ell),X_{t_k}^{\theta^0}))^*(\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0})M_{i,2,k,n}\right|^2\right]d\ell\\
&\leq 2\frac{u_i^2}{n\Delta_n}\sum_{k=0}^{n-1}\int_0^1\left\{\widehat{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta^0}\left[\left|(\partial_{\theta_i}b(\theta_i^{0+}(\ell),X_{t_k}^{\theta^0}))^*(\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0})M_{i,2,1,k,n}\right|^2\right]\right.\\
&\quad\left.+\widehat{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta^0}\left[\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta_i^{0+}(\ell)}\left[\left|(\partial_{\theta_i}b(\theta_i^{0+}(\ell),X_{t_k}^{\theta^0}))^*(\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0})M_{i,2,2,k,n}\right|^2\Big|\mathrm{Y}_{t_{k+1}}^{\theta_i^{0+}(\ell)}=X_{t_{k+1}}^{\theta^0}\right]\right]\right\}d\ell\\
&=2\frac{u_i^2}{n\Delta_n}\sum_{k=0}^{n-1}\int_0^1\left\{\widehat{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta^0}\left[\left|(\partial_{\theta_i}b(\theta_i^{0+}(\ell),X_{t_k}^{\theta^0}))^*(\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0})M_{i,2,1,k,n}\right|^2\right]\right.\\
&\quad\left.+\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta_i^{0+}(\ell)}\left[\left|(\partial_{\theta_i}b(\theta_i^{0+}(\ell),X_{t_k}^{\theta^0}))^*(\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0})M_{i,2,2,k,n}\right|^2\right]\right\}d\ell\\
&\leq Cu_i^2\frac{\Delta_n^2}{n}\sum_{k=0}^{n-1}\left(1+|X_{t_k}^{\theta^0}|^q\right),\end{aligned}$$

for some constant $q>0$, which converges to zero in $\widehat{\mathrm{P}}^{\theta^0}$-probability as $n\to\infty$. Here

$$M_{i,2,1,k,n}=\int_{t_k}^{t_{k+1}}\Big(b(\theta_i^{0+}(\ell),X_s^{\theta^0})-b(\theta_i^{0+}(\ell),X_{t_k}^{\theta^0})\Big)\,ds,$$
$$M_{i,2,2,k,n}=\int_{t_k}^{t_{k+1}}\Big(b(\theta_i^{0+}(\ell),Y_s^{\theta_i^{0+}(\ell)})-b(\theta_i^{0+}(\ell),Y_{t_k}^{\theta_i^{0+}(\ell)})\Big)ds.$$

Thus, by Lemma 3.10 a), $\sum_{k=0}^{n-1}\zeta_{i,2,k,n}\overset{\widehat{\mathrm{P}}^{\theta^0}}{\longrightarrow}0$ for any $i\in\{1,\ldots,m\}$. Thus, the result follows. □

Lemma 4.6. *Under conditions* **(A1)**, **(A2)**, **(A3)**(b), **(A4)** *and* **(A5)**, *as $n\to\infty$,*

$$\sum_{k=0}^{n-1}\sum_{i=1}^{m}\frac{u_i}{\sqrt{n\Delta_n^3}}\int_0^1\left\{H_5^{\theta^0,i,k}-\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta_i^{0+}(\ell)}\left[R_5^{\theta_i^{0+}(\ell),k}\Big|Y_{t_{k+1}}^{\theta_i^{0+}(\ell)}=X_{t_{k+1}}^{\theta^0}\right]\right\}d\ell\overset{\widehat{\mathrm{P}}^{\theta^0}}{\longrightarrow}0.$$

Proof. For any $i\in\{1,\ldots,m\}$, we set

$$\zeta_{i,k,n}:=\frac{u_i}{\sqrt{n\Delta_n^3}}\int_0^1\left\{H_5^{\theta^0,i,k}-\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta_i^{0+}(\ell)}\left[R_5^{\theta_i^{0+}(\ell),k}\Big|Y_{t_{k+1}}^{\theta_i^{0+}(\ell)}=X_{t_{k+1}}^{\theta^0}\right]\right\}d\ell.$$

Using Lemma 3.6, we get that

$$\begin{aligned}\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}\left[\zeta_{i,k,n}|\widehat{\mathcal{F}}_{t_k}\right]&=\sum_{k=0}^{n-1}\frac{u_i}{\sqrt{n\Delta_n^3}}\int_0^1\widehat{\mathrm{E}}^{\theta^0}\left[H_5^{\theta^0,i,k}-\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta_i^{0+}(\ell)}\left[R_5^{\theta_i^{0+}(\ell),k}|Y_{t_{k+1}}^{\theta_i^{0+}(\ell)}=X_{t_{k+1}}^{\theta^0}\right]|\widehat{\mathcal{F}}_{t_k}\right]d\ell\\&=\sum_{k=0}^{n-1}\frac{u_i}{\sqrt{n\Delta_n^3}}\int_0^1\left(\widehat{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta^0}\left[H_5^{\theta^0,i,k}\right]-\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta_i^{0+}(\ell)}[R_5^{\theta_i^{0+}(\ell),k}]\right)d\ell=0.\end{aligned}$$

Next, proceeding as in the proof of Lemma 4.5 for the term $\zeta_{i,2,k,n}$, we obtain that

$$\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}\left[\zeta_{i,k,n}^2|\widehat{\mathcal{F}}_{t_k}\right]\le Cu_i^2\frac{\Delta_n}{n}\sum_{k=0}^{n-1}\left(1+|X_{t_k}^{\theta^0}|^q\right),$$

for some constant $q>0$, which converges to zero in $\widehat{\mathrm{P}}^{\theta^0}$-probability as $n\to\infty$. Thus, by Lemma 3.10 a), we have shown that $\sum_{k=0}^{n-1}\zeta_{i,k,n}\overset{\widehat{\mathrm{P}}^{\theta^0}}{\longrightarrow}0$ for any $i\in\{1,\ldots,m\}$. Thus, the result follows. □

Lemma 4.7. *Assume conditions* **(A1)**, **(A2)**, **(A3)**(a)-(b), **(A4)** *and* **(A5)**. *Then as $n\to\infty$,*

$$\sum_{k=0}^{n-1}\sum_{i=1}^{m}\frac{u_i}{\sqrt{n\Delta_n^3}}\int_0^1\left\{H_6^{\theta^0,i,k}-\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta_i^{0+}(\ell)}\left[R_6^{\theta_i^{0+}(\ell),k}\Big|Y_{t_{k+1}}^{\theta_i^{0+}(\ell)}=X_{t_{k+1}}^{\theta^0}\right]\right\}d\ell\overset{\widehat{\mathrm{P}}^{\theta^0}}{\longrightarrow}0.$$

Proof. For any $i\in\{1,\ldots,m\}$, we set

$$\zeta_{i,k,n}:=\frac{u_i}{\sqrt{n\Delta_n^3}}\int_0^1\left\{H_6^{\theta^0,i,k}-\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta_i^{0+}(\ell)}\left[R_6^{\theta_i^{0+}(\ell),k}\Big|Y_{t_{k+1}}^{\theta_i^{0+}(\ell)}=X_{t_{k+1}}^{\theta^0}\right]\right\}d\ell.$$

Using Lemma 3.6, we get that

$$\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}[\zeta_{i,k,n}|\widehat{\mathcal{F}}_{t_k}]=\sum_{k=0}^{n-1}\frac{u_i}{\sqrt{n\Delta_n^3}}\int_0^1\widehat{\mathrm{E}}^{\theta^0}\left[H_6^{\theta^0,i,k}-\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta_i^{0+}(\ell)}\left[R_6^{\theta_i^{0+}(\ell),k}|Y_{t_{k+1}}^{\theta_i^{0+}(\ell)}=X_{t_{k+1}}^{\theta^0}\right]|\widehat{\mathcal{F}}_{t_k}\right]d\ell$$
$$=\sum_{k=0}^{n-1}\frac{u_i}{\sqrt{n\Delta_n^3}}\int_0^1\left(\widehat{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta^0}\left[H_6^{\theta^0,i,k}\right]-\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta_i^{0+}(\ell)}[R_6^{\theta_i^{0+}(\ell),k}]\right)d\ell=0.$$

This shows that the term (i) of Lemma 3.10 a) is actually equal to 0 for all $n\geq 1$.

We next show that condition (ii) of Lemma 3.10 a) holds. For this, using Jensen's inequality and Lemma 5.1, we obtain that for any $q>1$,

$$\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}[\zeta_{i,k,n}^2|\widehat{\mathcal{F}}_{t_k}]=\sum_{k=0}^{n-1}\frac{u_i^2}{n\Delta_n^3}\widehat{\mathrm{E}}^{\theta^0}\left[\Big(\int_0^1\Big(H_6^{\theta^0,i,k}-\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta_i^{0+}(\ell)}[R_6^{\theta_i^{0+}(\ell),k}|Y_{t_{k+1}}^{\theta_i^{0+}(\ell)}=X_{t_{k+1}}^{\theta^0}]\Big)d\ell\Big)^2|\widehat{\mathcal{F}}_{t_k}\right]$$
$$\leq\frac{u_i^2}{n\Delta_n^3}\sum_{k=0}^{n-1}\int_0^1\widehat{\mathrm{E}}^{\theta^0}\left[\Big(H_6^{\theta^0,i,k}-\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta_i^{0+}(\ell)}\left[R_6^{\theta_i^{0+}(\ell),k}\Big|Y_{t_{k+1}}^{\theta_i^{0+}(\ell)}=X_{t_{k+1}}^{\theta^0}\right]\Big)^2|\widehat{\mathcal{F}}_{t_k}\right]d\ell$$
$$=\frac{u_i^2}{n\Delta_n}\sum_{k=0}^{n-1}\int_0^1\widehat{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta^0}\Bigg[\Bigg(e_k(\theta_i^{0+}(\ell))\Bigg(\int_{t_k}^{t_{k+1}}\int_{\mathbb{R}_0^d}c(X_{s-}^{\theta^0},z)\widetilde{N}(ds,dz)$$
$$-\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta_i^{0+}(\ell)}\left[\int_{t_k}^{t_{k+1}}\int_{\mathbb{R}_0^d}c(Y_{s-}^{\theta_i^{0+}(\ell)},z)\widetilde{M}(ds,dz)\Big|Y_{t_{k+1}}^{\theta_i^{0+}(\ell)}=X_{t_{k+1}}^{\theta^0}\right]\Bigg)\Bigg)^2\Bigg]d\ell$$
$$\leq C\frac{u_i^2}{n\Delta_n}\sum_{k=0}^{n-1}\int_0^1\left(1+|X_{t_k}^{\theta^0}|^{q_1}\right)\Delta_n\bigg((\lambda_{\upsilon_n}\Delta_n)^{\frac{1}{q}}+\int_{|z|\leq\upsilon_n}\zeta^2(z)\nu(dz)+\Delta_n\Big(\int_{\mathbb{R}_0^d}\zeta(z)\nu(dz)\Big)^2\bigg)d\ell$$
$$=Cu_i^2\bigg((\lambda_{\upsilon_n}\Delta_n)^{\frac{1}{q}}+\int_{|z|\leq\upsilon_n}\zeta^2(z)\nu(dz)+\Delta_n\Big(\int_{\mathbb{R}_0^d}\zeta(z)\nu(dz)\Big)^2\bigg)\frac{1}{n}\sum_{k=0}^{n-1}\left(1+|X_{t_k}^{\theta^0}|^{q_1}\right),$$

for some constant $q_1>0$, where $e_k(\theta_i^{0+}(\ell)):=\left(\partial_{\theta_i}b(\theta_i^{0+}(\ell),X_{t_k}^{\theta^0})\right)^*(\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0})$, and $(\upsilon_n)_{n\geq1}$ defined in Subsection 5.6 is a positive sequence satisfying $\lim_{n\to\infty}\upsilon_n=0$, and $\lambda_{\upsilon_n}:=\int_{|z|>\upsilon_n}\nu(dz)$.

When $\int_{\mathbb{R}^d}\nu(dz)<+\infty$, then $\lambda_{\upsilon_n}\leq\int_{\mathbb{R}^d}\nu(dz)<+\infty$. Therefore, $\lambda_{\upsilon_n}\Delta_n\to0$ as $n\to\infty$.

When $\int_{\mathbb{R}^d}\nu(dz)=+\infty$, then $\lambda_{\upsilon_n}\to\int_{\mathbb{R}^d}\nu(dz)=+\infty$ as $n\to\infty$. Then, there exist $\epsilon\in(0,1)$ and $n_0\in\mathbb{N}$ such that $\lambda_{\upsilon_n}\leq\Delta_n^{\epsilon-1}$ for all $n\geq n_0$. This implies that $\lambda_{\upsilon_n}\Delta_n\leq\Delta_n^{\epsilon}$ for all $n\geq n_0$. Therefore, $\lambda_{\upsilon_n}\Delta_n\to0$ as $n\to\infty$.

Using Lebesgue's dominated convergence theorem, the fact that $\upsilon_n\to0$ and $\zeta(z)\mathbf{1}_{|z|\leq1}\leq C|z|$, and condition **(A5)**, we get that $\int_{|z|\leq\upsilon_n}\zeta^2(z)\nu(dz)\to0$ as $n\to\infty$ and $\int_{\mathbb{R}_0^d}\zeta(z)\nu(dz)<\infty$. Furthermore, using Lemma 3.9, as $n\to\infty$,

$$\frac{1}{n}\sum_{k=0}^{n-1}\left(1+|X_{t_k}^{\theta^0}|^{q_1}\right)\overset{\widehat{\mathrm{P}}^{\theta^0}}{\longrightarrow}\int_{\mathbb{R}^d}(1+|x|^{q_1})\,\pi_{\theta^0}(dx)<+\infty.$$

Hence, we have shown that $\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}[\zeta_{i,k,n}^2|\widehat{\mathcal{F}}_{t_k}]\overset{\widehat{\mathrm{P}}^{\theta^0}}{\longrightarrow}0$ as $n\to\infty$. Thus, by Lemma 3.10 a), the result follows. □

5. Appendix

5.1. **Proof of Lemma 3.1.**

Proof. Let $f : \mathbb{R}^d \to \mathbb{R}$ be a continuously differentiable function with compact support. Fix $t \in [t_k, t_{k+1}]$. The chain rule of the Malliavin calculus gives $(D_t(f(Y^{\theta}_{t_{k+1}}(t_k,x))))^* = (\nabla f(Y^{\theta}_{t_{k+1}}(t_k,x)))^* D_t Y^{\theta}_{t_{k+1}}(t_k,x)$. Since the matrix $D_t Y^{\theta}_{t_{k+1}}(t_k,x)$ is invertible a.s., we have $(\nabla f(Y^{\theta}_{t_{k+1}}(t_k,x)))^* = (D_t(f(Y^{\theta}_{t_{k+1}}(t_k,x))))^* U^{\theta}_t(t_k,x)$, where $U^{\theta}_t(t_k,x) = (D_t Y^{\theta}_{t_{k+1}}(t_k,x))^{-1}$.

Then, using the integration by parts formula of the Malliavin calculus on $[t_k, t_{k+1}]$, we get that for any $i \in \{1,\dots,m\}$,

$$
\begin{aligned}
\partial_{\theta_i}\widetilde{\mathrm{E}}\left[f(Y^{\theta}_{t_{k+1}}(t_k,x))\right] &= \widetilde{\mathrm{E}}\left[(\nabla f(Y^{\theta}_{t_{k+1}}(t_k,x)))^* \,\partial_{\theta_i} Y^{\theta}_{t_{k+1}}(t_k,x)\right]\\
&= \frac{1}{\Delta_n}\widetilde{\mathrm{E}}\left[\int_{t_k}^{t_{k+1}} (\nabla f(Y^{\theta}_{t_{k+1}}(t_k,x)))^* \,\partial_{\theta_i} Y^{\theta}_{t_{k+1}}(t_k,x)dt\right]\\
&= \frac{1}{\Delta_n}\widetilde{\mathrm{E}}\left[\int_{t_k}^{t_{k+1}} (D_t(f(Y^{\theta}_{t_{k+1}}(t_k,x))))^* \,U^{\theta}_t(t_k,x)\,\partial_{\theta_i} Y^{\theta}_{t_{k+1}}(t_k,x)dt\right]\\
&= \frac{1}{\Delta_n}\widetilde{\mathrm{E}}\left[f(Y^{\theta}_{t_{k+1}}(t_k,x))\delta\left(U^{\theta}(t_k,x)\partial_{\theta_i} Y^{\theta}_{t_{k+1}}(t_k,x)\right)\right].
\end{aligned}
\tag{5.1}
$$

Observe that by (3.6), 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 justifies that we can interchange ∂_{θ_i} and $\widetilde{\mathrm{E}}$. Note that here $\delta(V) \equiv \delta(V\mathbf{1}_{[t_k,t_{k+1}]}(\cdot))$ for any $V \in \mathrm{Dom}\,\delta$.

Next, using the stochastic flow property, we have that

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

and

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

which, together with (5.1), finishes the desired proof. □

5.2. **Proof of Lemma 3.2.**

Proof. From (3.4) and Itô's formula,

$$
\begin{aligned}
(\nabla_x Y^{\theta}_t(t_k,x))^{-1} &= \mathrm{I}_d - \int_{t_k}^{t} (\nabla_x Y^{\theta}_s(t_k,x))^{-1}\left(\nabla_x b(\theta, Y^{\theta}_s(t_k,x)) - \sum_{i=1}^{d}(\nabla_x\sigma_i(Y^{\theta}_s(t_k,x)))^2\right)ds\\
&\quad - \sum_{i=1}^{d}\int_{t_k}^{t}(\nabla_x Y^{\theta}_s(t_k,x))^{-1}\nabla_x\sigma_i(Y^{\theta}_s(t_k,x))dW^i_s\\
&\quad + \int_{t_k}^{t}\int_{\mathbb{R}^d_0}(\nabla_x Y^{\theta}_s(t_k,x))^{-1}\left(\mathrm{I}_d + \nabla_x c(Y^{\theta}_{s-}(t_k,x),z)\right)^{-1}(\nabla_x c(Y^{\theta}_{s-}(t_k,x),z))^2\nu(dz)ds
\end{aligned}
$$

$$-\int_{t_k}^t\int_{\mathbb{R}_0^d}(\nabla_x Y_s^\theta(t_k,x))^{-1}\left(\mathrm{I}_d+\nabla_x c(Y_{s-}^\theta(t_k,x),z)\right)^{-1}\nabla_x c(Y_{s-}^\theta(t_k,x),z)\widetilde{M}(ds,dz),$$

which, together with (3.5) and Itô's formula again, implies that

$$(\nabla_x Y_{t_{k+1}}^\theta(t_k,x))^{-1}\partial_{\theta_i}Y_{t_{k+1}}^\theta(t_k,x)=\int_{t_k}^{t_{k+1}}(\nabla_x Y_s^\theta(t_k,x))^{-1}\partial_{\theta_i}b(\theta,Y_s^\theta(t_k,x))ds. \tag{5.2}$$

Then, using the product rule [28, (1.48)], the fact that the Skorohod integral and the Itô integral of an adapted process coincide, and (5.2), we obtain that

$$\begin{aligned}
&\delta\left(U^\theta(t_k,x)\partial_{\theta_i}Y_{t_{k+1}}^\theta(t_k,x)\right)\\
&=\delta\left(\sigma^{-1}(Y_\cdot^\theta(t_k,x))\nabla_x Y_\cdot^\theta(t_k,x)(\nabla_x Y_{t_{k+1}}^\theta(t_k,x))^{-1}\partial_{\theta_i}Y_{t_{k+1}}^\theta(t_k,x)\right)\\
&=(\partial_{\theta_i}Y_{t_{k+1}}^\theta(t_k,x))^*((\nabla_x Y_{t_{k+1}}^\theta(t_k,x))^{-1})^*\int_{t_k}^{t_{k+1}}(\nabla_x Y_s^\theta(t_k,x))^*(\sigma^{-1}(Y_s^\theta(t_k,x)))^*dW_s\\
&\quad-\int_{t_k}^{t_{k+1}}\mathrm{tr}\left(D_s\left((\partial_{\theta_i}Y_{t_{k+1}}^\theta(t_k,x))^*((\nabla_x Y_{t_{k+1}}^\theta(t_k,x))^{-1})^*\right)\sigma^{-1}(Y_s^\theta(t_k,x))\nabla_x Y_s^\theta(t_k,x)\right)ds\\
&=\int_{t_k}^{t_{k+1}}((\nabla_x Y_s^\theta(t_k,x))^{-1}\partial_{\theta_i}b(\theta,Y_s^\theta(t_k,x)))^*ds\int_{t_k}^{t_{k+1}}(\nabla_x Y_s^\theta(t_k,x))^*(\sigma^{-1}(Y_s^\theta(t_k,x)))^*dW_s\\
&-\int_{t_k}^{t_{k+1}}\int_s^{t_{k+1}}\mathrm{tr}\left(D_s\left(((\nabla_x Y_u^\theta(t_k,x))^{-1}\partial_{\theta_i}b(\theta,Y_u^\theta(t_k,x)))^*\right)\sigma^{-1}(Y_s^\theta(t_k,x))\nabla_x Y_s^\theta(t_k,x)\right)duds.
\end{aligned}$$

We next add and subtract the matrix $((\nabla_x Y_{t_k}^\theta(t_k,x))^{-1}\partial_{\theta_i}b(\theta,Y_{t_k}^\theta(t_k,x)))^*$ in the first integral and the matrix $(\nabla_x Y_{t_k}^\theta(t_k,x))^*(\sigma^{-1}(Y_{t_k}^\theta(t_k,x)))^*$ in the second integral. This, together with the fact that $Y_{t_k}^\theta(t_k,x)=Y_{t_k}^\theta=x$, yields

$$\delta\left(U^\theta(t_k,x)\partial_{\theta_i}Y_{t_{k+1}}^\theta(t_k,x)\right)=\Delta_n(\sigma^{-1}(Y_{t_k}^\theta)\partial_{\theta_i}b(\theta,Y_{t_k}^\theta))^*(W_{t_{k+1}}-W_{t_k})-R_1^{\theta,k}+R_2^{\theta,k}+R_3^{\theta,k}. \tag{5.3}$$

On the other hand, by equation (3.3) we have that

$$\begin{aligned}
W_{t_{k+1}}-W_{t_k}=\sigma^{-1}(Y_{t_k}^\theta)\bigg(&Y_{t_{k+1}}^\theta-Y_{t_k}^\theta-b(\theta,Y_{t_k}^\theta)\Delta_n-\int_{t_k}^{t_{k+1}}\left(b(\theta,Y_s^\theta)-b(\theta,Y_{t_k}^\theta)\right)ds\\
&-\int_{t_k}^{t_{k+1}}\left(\sigma(Y_s^\theta)-\sigma(Y_{t_k}^\theta)\right)dW_s-\int_{t_k}^{t_{k+1}}\int_{\mathbb{R}_0^d}c(Y_{s-}^\theta,z)\widetilde{M}(ds,dz)\bigg).
\end{aligned}$$

This, together with (5.3), concludes the desired result. □

5.3. **Proof of Lemma 3.6.**

Proof. For simplicity, we denote $g(y)=g(X_{t_k}^{\theta^0},y):=\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^\theta[V|Y_{t_{k+1}}^\theta=y]$ for all $y\in\mathbb{R}^d$. Then, applying Girsanov's theorem, we obtain that

$$\widehat{\mathrm{E}}^{\theta^0}\left[\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^\theta\left[V\big|Y_{t_{k+1}}^\theta=X_{t_{k+1}}^{\theta^0}\right]|\widehat{\mathcal{F}}_{t_k}\right]=\widehat{\mathrm{E}}^{\theta^0}\left[g(X_{t_{k+1}}^{\theta^0})\big|X_{t_k}^{\theta^0}\right]=\widehat{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta^0}\left[g(X_{t_{k+1}}^{\theta^0})\right]$$

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

where we have used that fact that, by definition of $\mathrm{E}^{\theta}_{t_k,x}$, for any $\widehat{\mathcal{F}}_{t_{k+1}}$-measurable random variable V_1 and $\widetilde{\mathcal{F}}_{t_{k+1}}$-measurable random variable V_2,

$$
\widehat{\mathrm{E}}^{\theta}_{t_k,x} \left[V_1 | X^{\theta}_{t_{k+1}} = y \right] \widetilde{\mathrm{E}}^{\theta}_{t_k,x} \left[V_2 | Y^{\theta}_{t_{k+1}} = y \right] = \mathrm{E}^{\theta}_{t_k,x} \left[V_1 V_2 | X^{\theta}_{t_{k+1}} = y, Y^{\theta}_{t_{k+1}} = y \right],
$$

and the independence between V and $\frac{d\widehat{\mathrm{P}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}}{d\widehat{\mathrm{P}}^{\theta}_{t_k,X^{\theta^0}_{t_k}}}((X^{\theta}_t)_{t\in[t_k,t_{k+1}]})$, together with

$$\widehat{\mathrm{E}}^{\theta}_{t_k,X^{\theta^0}_{t_k}}\Big[\frac{d\widehat{\mathrm{P}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}}{d\widehat{\mathrm{P}}^{\theta}_{t_k,X^{\theta^0}_{t_k}}}((X^{\theta}_t)_{t\in[t_k,t_{k+1}]})\Big]=1.$$

Thus, the result follows. □

5.4. **Proof of Lemma 3.7.**

Proof. Using (3.9), we have that

$$\begin{aligned}
&\frac{d\widehat{\mathrm{P}}^{\theta^0}_{t_k,x}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,x}}-1=\frac{d\widehat{\mathrm{P}}^{\theta^0}_{t_k,x}-d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,x}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,x}}\\
&=\frac{(d\widehat{\mathrm{P}}^{\theta^{0+}_{i+1}}_{t_k,x}-d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,x})+(d\widehat{\mathrm{P}}^{\theta^{0+}_{i+2}}_{t_k,x}-d\widehat{\mathrm{P}}^{\theta^{0+}_{i+1}}_{t_k,x})+\cdots+(d\widehat{\mathrm{P}}^{\theta^{0+}_{m}}_{t_k,x}-d\widehat{\mathrm{P}}^{\theta^{0+}_{m-1}}_{t_k,x})+(d\widehat{\mathrm{P}}^{\theta^{0}}_{t_k,x}-d\widehat{\mathrm{P}}^{\theta^{0+}_{m}}_{t_k,x})}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,x}}\\
&=\int_{\theta_i^0+\ell\frac{u_i}{\sqrt{n\Delta_n}}}^{\theta_i^0}\partial_{\theta_i}\left(\frac{d\widehat{\mathrm{P}}^{\theta_i(0+)}_{t_k,x}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,x}}\right)d\theta_i+\int_{\theta^0_{i+1}+\frac{u_{i+1}}{\sqrt{n\Delta_n}}}^{\theta^0_{i+1}}\partial_{\theta_{i+1}}\left(\frac{d\widehat{\mathrm{P}}^{\theta_{i+1}(0+)}_{t_k,x}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,x}}\right)d\theta_{i+1}\\
&\qquad+\cdots+\int_{\theta^0_{m-1}+\frac{u_{m-1}}{\sqrt{n\Delta_n}}}^{\theta^0_{m-1}}\partial_{\theta_{m-1}}\left(\frac{d\widehat{\mathrm{P}}^{\theta_{m-1}(0+)}_{t_k,x}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,x}}\right)d\theta_{m-1}+\int_{\theta^0_{m}+\frac{u_{m}}{\sqrt{n\Delta_n}}}^{\theta^0_{m}}\partial_{\theta_{m}}\left(\frac{d\widehat{\mathrm{P}}^{\theta_{m}(0+)}_{t_k,x}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,x}}\right)d\theta_{m}\\
&=\int_{\theta_i^0+\ell\frac{u_i}{\sqrt{n\Delta_n}}}^{\theta_i^0}\int_{t_k}^{t_{k+1}}(\partial_{\theta_i}b(\theta_i(0+),X_t^{\theta_i^{0+}(\ell)}))^*(\sigma^*)^{-1}(X_t^{\theta_i^{0+}(\ell)})\\
&\qquad\cdot\left(dB_t-\sigma^{-1}(X_t^{\theta_i^{0+}(\ell)})\big(b(\theta_i(0+),X_t^{\theta_i^{0+}(\ell)})-b(\theta_i^{0+}(\ell),X_t^{\theta_i^{0+}(\ell)})\big)dt\right)\frac{d\widehat{\mathrm{P}}^{\theta_i(0+)}_{t_k,x}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,x}}d\theta_i\\
&+\int_{\theta^0_{i+1}+\frac{u_{i+1}}{\sqrt{n\Delta_n}}}^{\theta^0_{i+1}}\int_{t_k}^{t_{k+1}}(\partial_{\theta_{i+1}}b(\theta_{i+1}(0+),X_t^{\theta_i^{0+}(\ell)}))^*(\sigma^*)^{-1}(X_t^{\theta_i^{0+}(\ell)})\\
&\qquad\cdot\left(dB_t-\sigma^{-1}(X_t^{\theta_i^{0+}(\ell)})\big(b(\theta_{i+1}(0+),X_t^{\theta_i^{0+}(\ell)})-b(\theta_i^{0+}(\ell),X_t^{\theta_i^{0+}(\ell)})\big)dt\right)\frac{d\widehat{\mathrm{P}}^{\theta_{i+1}(0+)}_{t_k,x}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,x}}d\theta_{i+1}\\
&+\cdots+\int_{\theta^0_{m}+\frac{u_{m}}{\sqrt{n\Delta_n}}}^{\theta^0_{m}}\int_{t_k}^{t_{k+1}}(\partial_{\theta_{m}}b(\theta_{m}(0+),X_t^{\theta_i^{0+}(\ell)}))^*(\sigma^*)^{-1}(X_t^{\theta_i^{0+}(\ell)})\\
&\qquad\cdot\left(dB_t-\sigma^{-1}(X_t^{\theta_i^{0+}(\ell)})\big(b(\theta_{m}(0+),X_t^{\theta_i^{0+}(\ell)})-b(\theta_i^{0+}(\ell),X_t^{\theta_i^{0+}(\ell)})\big)dt\right)\frac{d\widehat{\mathrm{P}}^{\theta_{m}(0+)}_{t_k,x}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,x}}d\theta_{m},
\end{aligned}$$

where for $j \in \{i, \dots, m\}$,

$$\frac{d\widehat{\mathrm{P}}_{t_k,x}^{\theta_j(0+)}}{d\widehat{\mathrm{P}}_{t_k,x}^{\theta_i^{0+}(\ell)}} = \frac{d\widehat{\mathrm{P}}_{t_k,x}^{\theta_j(0+)}}{d\widehat{\mathrm{P}}_{t_k,x}^{\theta_i^{0+}(\ell)}}\big((X_t^{\theta_i^{0+}(\ell)})_{t\in[t_k,t_{k+1}]}\big),$$

and

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

Then, using Girsanov's theorem, we get that

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

Here, for $j \in \{i, \dots, m\}$ the process $B^{\widehat{\mathrm{P}}_{t_k,x}^{\theta_j(0+)}} = (B_t^{\widehat{\mathrm{P}}_{t_k,x}^{\theta_j(0+)}}, t \in [t_k, t_{k+1}])$ is a Brownian motion under $\widehat{\mathrm{P}}_{t_k,x}^{\theta_j(0+)}$, where for any $t \in [t_k, t_{k+1}]$,

$$B_t^{\widehat{\mathrm{P}}_{t_k,x}^{\theta_j(0+)}} := B_t - \int_{t_k}^t \sigma^{-1}(X_s^{\theta_i^{0+}(\ell)})\big(b(\theta_j(0+), X_s^{\theta_i^{0+}(\ell)}) - b(\theta_i^{0+}(\ell), X_s^{\theta_i^{0+}(\ell)})\big)ds.$$

Next, using Hölder's and Burkholder-Davis-Gundy's inequalities, conditions **(A2)** and **(A3)**(b), and Lemma 3.4 (ii), we get that

$$\left|\widehat{\mathrm{E}}_{t_k,x}^{\theta_i^{0+}(\ell)}\left[V\left(\frac{d\widehat{\mathrm{P}}_{t_k,x}^{\theta^0}}{d\widehat{\mathrm{P}}_{t_k,x}^{\theta_i^{0+}(\ell)}}\left((X_t^{\theta_i^{0+}(\ell)})_{t\in[t_k,t_{k+1}]}\right)-1\right)\right]\right|$$

$$\leq\left|\int_{\theta_i^0+\ell\frac{u_i}{\sqrt{n\Delta_n}}}^{\theta_i^0}\left|\widehat{\mathrm{E}}_{t_k,x}^{\theta_i(0+)}\left[V\int_{t_k}^{t_{k+1}}(\partial_{\theta_i}b(\theta_i(0+),X_t^{\theta_i(0+)}))^*(\sigma^*)^{-1}(X_t^{\theta_i(0+)})dB_t^{\widehat{\mathrm{P}}_{t_k,x}^{\theta_i(0+)}}\right]\right|d\theta_i\right|$$

$$+\left|\int_{\theta_{i+1}^0+\frac{u_{i+1}}{\sqrt{n\Delta_n}}}^{\theta_{i+1}^0}\left|\widehat{\mathrm{E}}_{t_k,x}^{\theta_{i+1}(0+)}\left[V\int_{t_k}^{t_{k+1}}(\partial_{\theta_{i+1}}b(\theta_{i+1}(0+),X_t^{\theta_{i+1}(0+)}))^*(\sigma^*)^{-1}(X_t^{\theta_{i+1}(0+)})dB_t^{\widehat{\mathrm{P}}_{t_k,x}^{\theta_{i+1}(0+)}}\right]\right|d\theta_{i+1}\right|$$

$$+\cdots+\left|\int_{\theta_m^0+\frac{u_m}{\sqrt{n\Delta_n}}}^{\theta_m^0}\left|\widehat{\mathrm{E}}_{t_k,x}^{\theta_m(0+)}\left[V\int_{t_k}^{t_{k+1}}(\partial_{\theta_m}b(\theta_m(0+),X_t^{\theta_m(0+)}))^*(\sigma^*)^{-1}(X_t^{\theta_m(0+)})dB_t^{\widehat{\mathrm{P}}_{t_k,x}^{\theta_m(0+)}}\right]\right|d\theta_m\right|$$

$$\leq C\Big|\int_{\theta_i^0+\ell\frac{u_i}{\sqrt{n\Delta_n}}}^{\theta_i^0}\left(\widehat{\mathrm{E}}_{t_k,x}^{\theta_i(0+)}\left[|V|^q\right]\right)^{\frac{1}{q}}$$

$$\cdot\left(\Delta_n^{\frac{p}{2}-1}\int_{t_k}^{t_{k+1}}\widehat{\mathrm{E}}_{t_k,x}^{\theta_i(0+)}\left[\left|\partial_{\theta_i}b(\theta_i(0+),X_t^{\theta_i(0+)}))^*(\sigma^*)^{-1}(X_t^{\theta_i(0+)})\right|^p\right]ds\right)^{\frac{1}{p}}d\theta_i\Big|$$

$$+C\Big|\int_{\theta_{i+1}^0+\frac{u_{i+1}}{\sqrt{n\Delta_n}}}^{\theta_{i+1}^0}\left(\widehat{\mathrm{E}}_{t_k,x}^{\theta_{i+1}(0+)}\left[|V|^q\right]\right)^{\frac{1}{q}}$$

$$\cdot\left(\Delta_n^{\frac{p}{2}-1}\int_{t_k}^{t_{k+1}}\widehat{\mathrm{E}}_{t_k,x}^{\theta_{i+1}(0+)}\left[\left|\partial_{\theta_{i+1}}b(\theta_{i+1}(0+),X_t^{\theta_{i+1}(0+)}))^*(\sigma^*)^{-1}(X_t^{\theta_{i+1}(0+)})\right|^p\right]ds\right)^{\frac{1}{p}}d\theta_{i+1}\Big|$$

$$+\cdots+C\Big|\int_{\theta_m^0+\frac{u_m}{\sqrt{n\Delta_n}}}^{\theta_m^0}\left(\widehat{\mathrm{E}}_{t_k,x}^{\theta_m(0+)}\left[|V|^q\right]\right)^{\frac{1}{q}}$$

$$\cdot\left(\Delta_n^{\frac{p}{2}-1}\int_{t_k}^{t_{k+1}}\widehat{\mathrm{E}}_{t_k,x}^{\theta_m(0+)}\left[\left|\partial_{\theta_m}b(\theta_m(0+),X_t^{\theta_m(0+)}))^*(\sigma^*)^{-1}(X_t^{\theta_m(0+)})\right|^p\right]ds\right)^{\frac{1}{p}}d\theta_m\Big|$$

$$\leq C\sqrt{\Delta_n}\,(1+|x|^{q_1})\Bigg(\left|\int_{\theta_i^0+\ell\frac{u_i}{\sqrt{n\Delta_n}}}^{\theta_i^0}\left(\widehat{\mathrm{E}}_{t_k,x}^{\theta_i(0+)}\left[|V|^q\right]\right)^{\frac{1}{q}}d\theta_i\right|$$

$$+\left|\int_{\theta_{i+1}^0+\frac{u_{i+1}}{\sqrt{n\Delta_n}}}^{\theta_{i+1}^0}\left(\widehat{\mathrm{E}}_{t_k,x}^{\theta_{i+1}(0+)}\left[|V|^q\right]\right)^{\frac{1}{q}}d\theta_{i+1}\right|+\cdots+\left|\int_{\theta_m^0+\frac{u_m}{\sqrt{n\Delta_n}}}^{\theta_m^0}\left(\widehat{\mathrm{E}}_{t_k,x}^{\theta_m(0+)}\left[|V|^q\right]\right)^{\frac{1}{q}}d\theta_m\right|\Bigg),$$

for some constants $C>0$, $q_1>0$, where $p,q>1$ and $\frac{1}{p}+\frac{1}{q}=1$. Thus, the result follows. □

5.5. **Proof of Lemma 3.9.**

Proof. Using the mean value theorem, Cauchy-Schwarz inequality, the fact that $|\nabla g(x)|$ has polynomial growth in x, and Lemma 3.4, we get that

$$\widehat{\mathrm{E}}^{\theta^0}\left[\left|\frac{1}{n\Delta_n}\int_0^{n\Delta_n}g(X_s^{\theta^0})ds-\frac{1}{n}\sum_{k=0}^{n-1}g(X_{t_k}^{\theta^0})\right|\right]\leq\frac{1}{n\Delta_n}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}}\widehat{\mathrm{E}}^{\theta^0}\left[\left|g(X_s^{\theta^0})-g(X_{t_k}^{\theta^0})\right|\right]ds$$

$$
\begin{aligned}
&= \frac{1}{n\Delta_n}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \widehat{\mathrm{E}}^{\theta^0}\left[\left|\int_0^1 \nabla g(X_{t_k}^{\theta^0}+u(X_s^{\theta^0}-X_{t_k}^{\theta^0}))du\left(X_s^{\theta^0}-X_{t_k}^{\theta^0}\right)\right|\right]ds\\
&\leq \frac{1}{n\Delta_n}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} \left(\widehat{\mathrm{E}}^{\theta^0}\left[\left|\int_0^1 \nabla g(X_{t_k}^{\theta^0}+u(X_s^{\theta^0}-X_{t_k}^{\theta^0}))du\right|^2\right]\right)^{\frac{1}{2}}\left(\widehat{\mathrm{E}}^{\theta^0}\left[\left|X_s^{\theta^0}-X_{t_k}^{\theta^0}\right|^2\right]\right)^{\frac{1}{2}}ds\\
&\leq \frac{C\sqrt{\Delta_n}}{n\Delta_n}\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}}\left(\int_0^1 \widehat{\mathrm{E}}^{\theta^0}\left[\left|\nabla g(X_{t_k}^{\theta^0}+u(X_s^{\theta^0}-X_{t_k}^{\theta^0}))\right|^2\right]du\right)^{\frac{1}{2}}ds\\
&\leq C\sqrt{\Delta_n},
\end{aligned}\tag{5.4}
$$

which tends to zero as $n\to\infty$. On the other hand, using condition **(A4)** and the fact that $|g(x)|$ has polynomial growth in x, we obtain that as $n\to\infty$,

$$
\frac{1}{n\Delta_n}\int_0^{n\Delta_n} g(X_s^{\theta^0})ds \overset{\widehat{\mathrm{P}}^{\theta^0}}{\longrightarrow} \int_{\mathbb{R}^d} g(x)\pi_{\theta^0}(dx). \tag{5.5}
$$

Thus, the result follows from (5.4) and (5.5). □

5.6. **Large deviation type estimates.** Let $(\upsilon_n)_{n\geq 1}$ be a positive sequence which satisfies $\lim_{n\to\infty}\upsilon_n=0$. The process $Z^{\upsilon_n}=(Z_t^{\upsilon_n})_{t\geq 0}$ defined by $Z_t^{\upsilon_n}=\sum_{0\leq s\leq t}\Delta Z_s\mathbf{1}_{\{|\Delta Z_s|>\upsilon_n\}}$ is a compound Poisson process with intensity of big jumps $\lambda_{\upsilon_n}:=\int_{|z|>\upsilon_n}\nu(dz)$ and distribution of big jumps $\frac{\mathbf{1}_{|z|>\upsilon_n}\nu(dz)}{\lambda_{\upsilon_n}}$. Then, we can split the jumps of the Lévy process Z_t into small jumps and big jumps as follows

$$
\int_0^t\int_{\mathbb{R}_0^d} c(X_{s-}^\theta,z)\widetilde{N}(ds,dz)=\int_0^t\int_{|z|\leq\upsilon_n} c(X_{s-}^\theta,z)\widetilde{N}(ds,dz)+\int_0^t\int_{|z|>\upsilon_n} c(X_{s-}^\theta,z)\widetilde{N}(ds,dz).
$$

Hence, from (1.1), for any $t\geq 0$, we can write

$$
\begin{aligned}
X_t^\theta = x_0 &+ \int_0^t b(\theta,X_s^\theta)ds+\int_0^t\sigma(X_s^\theta)dB_s+\int_0^t\int_{|z|\leq\upsilon_n} c(X_{s-}^\theta,z)\widetilde{N}(ds,dz)\\
&+\int_0^t\int_{|z|>\upsilon_n} c(X_{s-}^\theta,z)\widetilde{N}(ds,dz).
\end{aligned}\tag{5.6}
$$

Let $N^{\upsilon_n}=(N_t^{\upsilon_n})_{t\geq 0}$ denote the Poisson process with intensity λ_{υ_n} counting the big jumps of the compound Poisson process Z^{υ_n}.

Similarly, the process $\widetilde{Z}^{\upsilon_n}=(\widetilde{Z}_t^{\upsilon_n})_{t\geq 0}$ defined by $\widetilde{Z}_t^{\upsilon_n}=\sum_{0\leq s\leq t}\Delta\widetilde{Z}_s\mathbf{1}_{\{|\Delta\widetilde{Z}_s|>\upsilon_n\}}$ is a compound Poisson process with intensity of big jumps λ_{υ_n} and distribution of big jumps $\frac{\mathbf{1}_{|z|>\upsilon_n}\nu(dz)}{\lambda_{\upsilon_n}}$. Then, we can split the jumps of Lévy process $\widetilde{Z}_t$ into small jumps and big jumps as follows

$$
\int_0^t\int_{\mathbb{R}_0^d} c(Y_{s-}^\theta,z)\widetilde{M}(ds,dz)=\int_0^t\int_{|z|\leq\upsilon_n} c(Y_{s-}^\theta,z)\widetilde{M}(ds,dz)+\int_0^t\int_{|z|>\upsilon_n} c(Y_{s-}^\theta,z)\widetilde{M}(ds,dz).
$$

Hence, from (3.3), for any $t \geq 0$, we can write

$$\begin{aligned}Y_t^\theta = x_0 + \int_0^t b(\theta, Y_s^\theta)ds + \int_0^t \sigma(Y_s^\theta)dW_s + \int_0^t \int_{|z|\leq \upsilon_n} c(Y_{s-}^\theta, z)\widetilde{M}(ds,dz) \\ + \int_0^t \int_{|z|>\upsilon_n} c(Y_{s-}^\theta, z)\widetilde{M}(ds,dz).\end{aligned} \tag{5.7}$$

Let $M^{\upsilon_n} = (M_t^{\upsilon_n})_{t\geq 0}$ denote the Poisson process with intensity λ_{υ_n} counting the big jumps of the compound Poisson process $\widetilde{Z}^{\upsilon_n}$.

Now, for $k \in \{0, ..., n-1\}$, we consider the events $\widehat{N}_{0,k}(\upsilon_n) := \{N_{t_{k+1}}^{\upsilon_n} - N_{t_k}^{\upsilon_n} = 0\}$ which have no big jumps of Z^{υ_n} in the interval $[t_k, t_{k+1})$ and $\widehat{N}_{\geq 1,k}(\upsilon_n) := \{N_{t_{k+1}}^{\upsilon_n} - N_{t_k}^{\upsilon_n} \geq 1\}$ which have one or more than one big jump of Z^{υ_n} in the interval $[t_k, t_{k+1})$. Similarly, we consider the events $\widetilde{N}_{0,k}(\upsilon_n) := \{M_{t_{k+1}}^{\upsilon_n} - M_{t_k}^{\upsilon_n} = 0\}$ which have no big jumps of $\widetilde{Z}^{\upsilon_n}$ in the interval $[t_k, t_{k+1})$ and $\widetilde{N}_{\geq 1,k}(\upsilon_n) := \{M_{t_{k+1}}^{\upsilon_n} - M_{t_k}^{\upsilon_n} \geq 1\}$ which have one or more than one big jump of $\widetilde{Z}^{\upsilon_n}$ in the interval $[t_k, t_{k+1})$.

We set $e_k(\theta) := \left(\partial_{\theta_i} b(\theta, X_{t_k}^{\theta^0})\right)^* (\sigma\sigma^*)^{-1}(X_{t_k}^{\theta^0})$. Next, as in [2, Lemma 5.3], we obtain the following large deviation type estimates.

Lemma 5.1. *Assume conditions* **(A1)**, **(A2)**, **(A3)**(b), *and* **(A5)**. *Then, for any $\theta \in \Theta$, there exist constants $C > 0$ and $q_1 > 0$ such that for all $q > 1$, and $k \in \{0, ..., n-1\}$,*

$$\begin{aligned}&\widehat{\mathrm{E}}_{t_k, X_{t_k}^{\theta^0}}^{\theta^0}\Bigg[\Bigg(e_k(\theta)\Bigg(\int_{t_k}^{t_{k+1}}\int_{\mathbb{R}_0^d} c(X_{s-}^{\theta^0}, z)\widetilde{N}(ds,dz) \\ &\qquad - \widetilde{\mathrm{E}}_{t_k, X_{t_k}^{\theta^0}}^{\theta}\Bigg[\int_{t_k}^{t_{k+1}}\int_{\mathbb{R}_0^d} c(Y_{s-}^\theta, z)\widetilde{M}(ds,dz)\Big|Y_{t_{k+1}}^\theta = X_{t_{k+1}}^{\theta^0}\Bigg]\Bigg)\Bigg)^2\Bigg] \\ &\leq C\left(1 + |X_{t_k}^{\theta^0}|^{q_1}\right)\Delta_n\Bigg((\lambda_{\upsilon_n}\Delta_n)^{\frac{1}{q}} + \int_{|z|\leq\upsilon_n}\zeta^2(z)\nu(dz) + \Delta_n\Big(\int_{\mathbb{R}_0^d}\zeta(z)\nu(dz)\Big)^2\Bigg).\end{aligned}$$

Proof. Splitting the Poisson integrals into small jumps and big jumps, we get that

$$\begin{aligned}&\widehat{\mathrm{E}}_{t_k, X_{t_k}^{\theta^0}}^{\theta^0}\Bigg[\Bigg(e_k(\theta)\Bigg(\int_{t_k}^{t_{k+1}}\int_{\mathbb{R}_0^d} c(X_{s-}^{\theta^0}, z)\widetilde{N}(ds,dz) \\ &\quad - \widetilde{\mathrm{E}}_{t_k, X_{t_k}^{\theta^0}}^{\theta}\Bigg[\int_{t_k}^{t_{k+1}}\int_{\mathbb{R}_0^d} c(Y_{s-}^\theta, z)\widetilde{M}(ds,dz)\Big|Y_{t_{k+1}}^\theta = X_{t_{k+1}}^{\theta^0}\Bigg]\Bigg)\Bigg)^2\Bigg] \\ &= \widehat{\mathrm{E}}_{t_k, X_{t_k}^{\theta^0}}^{\theta^0}\Bigg[\Bigg(e_k(\theta)\Bigg(\int_{t_k}^{t_{k+1}}\int_{|z|\leq\upsilon_n} c(X_{s-}^{\theta^0}, z)\widetilde{N}(ds,dz) + \int_{t_k}^{t_{k+1}}\int_{|z|>\upsilon_n} c(X_{s-}^{\theta^0}, z)\widetilde{N}(ds,dz) \\ &\quad - \widetilde{\mathrm{E}}_{t_k, X_{t_k}^{\theta^0}}^{\theta}\Bigg[\int_{t_k}^{t_{k+1}}\int_{|z|\leq\upsilon_n} c(Y_{s-}^\theta, z)\widetilde{M}(ds,dz) \\ &+ \int_{t_k}^{t_{k+1}}\int_{|z|>\upsilon_n} c(Y_{s-}^\theta, z)\widetilde{M}(ds,dz)\Big|Y_{t_{k+1}}^\theta = X_{t_{k+1}}^{\theta^0}\Bigg]\Bigg)\Bigg)^2\Bigg] \leq 3\left(D_{k,n}^1 + D_{k,n}^2 + D_{k,n}^3\right),\end{aligned} \tag{5.8}$$

where

$$D^1_{k,n}=\widehat{\mathrm{E}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}\left[\left(e_k(\theta)\int_{t_k}^{t_{k+1}}\int_{|z|\leq \upsilon_n}c(X^{\theta^0}_{s-},z)\widetilde{N}(ds,dz)\right)^2\right],$$

$$D^2_{k,n}=\widehat{\mathrm{E}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}\left[\left(e_k(\theta)\widetilde{\mathrm{E}}^{\theta}_{t_k,X^{\theta^0}_{t_k}}\left[\int_{t_k}^{t_{k+1}}\int_{|z|\leq \upsilon_n}c(Y^{\theta}_{s-},z)\widetilde{M}(ds,dz)\Big|Y^{\theta}_{t_{k+1}}=X^{\theta^0}_{t_{k+1}}\right]\right)^2\right],$$

$$D^3_{k,n}=\widehat{\mathrm{E}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}\Bigg[\Bigg(e_k(\theta)\Bigg(\int_{t_k}^{t_{k+1}}\int_{|z|>\upsilon_n}c(X^{\theta^0}_{s-},z)\widetilde{N}(ds,dz)$$
$$-\widetilde{\mathrm{E}}^{\theta}_{t_k,X^{\theta^0}_{t_k}}\left[\int_{t_k}^{t_{k+1}}\int_{|z|>\upsilon_n}c(Y^{\theta}_{s-},z)\widetilde{M}(ds,dz)\Big|Y^{\theta}_{t_{k+1}}=X^{\theta^0}_{t_{k+1}}\right]\Bigg)\Bigg)^2\Bigg].$$

First, using Burkholder-Davis-Gundy's inequality and condition **(A1)**,

$$\begin{aligned}D^1_{k,n}&\leq C\int_{t_k}^{t_{k+1}}\int_{|z|\leq \upsilon_n}\widehat{\mathrm{E}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}\left[\left(e_k(\theta)c(X^{\theta^0}_{s-},z)\right)^2\right]\nu(dz)ds\\&\leq C\left(1+|X^{\theta^0}_{t_k}|^{q_1}\right)\int_{t_k}^{t_{k+1}}\int_{|z|\leq \upsilon_n}\zeta^2(z)\nu(dz)ds\\&=C\left(1+|X^{\theta^0}_{t_k}|^{q_1}\right)\Delta_n\int_{|z|\leq \upsilon_n}\zeta^2(z)\nu(dz),\end{aligned}\tag{5.9}$$

for some constant $q_1>0$. Next, using Jensen's inequality, Lemma 3.6, Burkholder-Davis-Gundy's inequality and condition **(A1)**,

$$\begin{aligned}D^2_{k,n}&\leq \widehat{\mathrm{E}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}\left[\widetilde{\mathrm{E}}^{\theta}_{t_k,X^{\theta^0}_{t_k}}\left[\left(e_k(\theta)\int_{t_k}^{t_{k+1}}\int_{|z|\leq \upsilon_n}c(Y^{\theta}_{s-},z)\widetilde{M}(ds,dz)\right)^2\Big|Y^{\theta}_{t_{k+1}}=X^{\theta^0}_{t_{k+1}}\right]\right]\\&=\widetilde{\mathrm{E}}^{\theta}_{t_k,X^{\theta^0}_{t_k}}\left[\left(e_k(\theta)\int_{t_k}^{t_{k+1}}\int_{|z|\leq \upsilon_n}c(Y^{\theta}_{s-},z)\widetilde{M}(ds,dz)\right)^2\right]\\&\leq C\left(1+|X^{\theta^0}_{t_k}|^{q_1}\right)\Delta_n\int_{|z|\leq \upsilon_n}\zeta^2(z)\nu(dz),\end{aligned}\tag{5.10}$$

for some constant $q_1>0$. Next, multiplying the random variable outside the conditional expectation of $D^3_{k,n}$ by $\mathbf{1}_{\widehat{N}_{0,k}(\upsilon_n)}+\mathbf{1}_{\widehat{N}_{\geq 1,k}(\upsilon_n)}$, we get that

$$\begin{aligned}D^3_{k,n}=\widehat{\mathrm{E}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}\Bigg[\left(\mathbf{1}_{\widehat{N}_{0,k}(\upsilon_n)}+\mathbf{1}_{\widehat{N}_{\geq 1,k}(\upsilon_n)}\right)\Bigg(e_k(\theta)\Bigg(\int_{t_k}^{t_{k+1}}\int_{|z|>\upsilon_n}c(X^{\theta^0}_{s-},z)\widetilde{N}(ds,dz)\\-\widetilde{\mathrm{E}}^{\theta}_{t_k,X^{\theta^0}_{t_k}}\left[\int_{t_k}^{t_{k+1}}\int_{|z|>\upsilon_n}c(Y^{\theta}_{s-},z)\widetilde{M}(ds,dz)\Big|Y^{\theta}_{t_{k+1}}=X^{\theta^0}_{t_{k+1}}\right]\Bigg)\Bigg)^2\Bigg]=M^{\theta}_{0,k,n}+M^{\theta}_{\geq 1,k,n},\end{aligned}\tag{5.11}$$

where

$$M^{\theta}_{0,k,n}=\widehat{\mathrm{E}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}\Bigg[\mathbf{1}_{\widehat{N}_{0,k}(\upsilon_n)}\Bigg(e_k(\theta)\Bigg(\int_{t_k}^{t_{k+1}}\int_{|z|>\upsilon_n}c(X^{\theta^0}_{s-},z)\widetilde{N}(ds,dz)$$

$$- \widetilde{\mathrm{E}}^{\theta}_{t_k, X^{\theta^0}_{t_k}} \Big[\int_{t_k}^{t_{k+1}} \int_{|z|>\upsilon_n} c(Y^{\theta}_{s-}, z) \widetilde{M}(ds, dz) \Big| Y^{\theta}_{t_{k+1}} = X^{\theta^0}_{t_{k+1}} \Big] \Big) \Big)^2 \Big],$$

$$M^{\theta}_{\geq 1,k,n} = \widehat{\mathrm{E}}^{\theta^0}_{t_k, X^{\theta^0}_{t_k}} \Big[\mathbf{1}_{\widehat{N}_{\geq 1,k}(\upsilon_n)} \Big(e_k(\theta) \Big(\int_{t_k}^{t_{k+1}} \int_{|z|>\upsilon_n} c(X^{\theta^0}_{s-}, z) \widetilde{N}(ds, dz) - \widetilde{\mathrm{E}}^{\theta}_{t_k, X^{\theta^0}_{t_k}} \Big[\int_{t_k}^{t_{k+1}} \int_{|z|>\upsilon_n} c(Y^{\theta}_{s-}, z) \widetilde{M}(ds, dz) \Big| Y^{\theta}_{t_{k+1}} = X^{\theta^0}_{t_{k+1}} \Big] \Big) \Big)^2 \Big].$$

We start treating $M^{\theta}_{0,k,n}$. Multiplying the random variable inside the conditional expectation of $M^{\theta}_{0,k,n}$ by $\mathbf{1}_{\widetilde{N}_{0,k}(\upsilon_n)} + \mathbf{1}_{\widetilde{N}_{\geq 1,k}(\upsilon_n)}$ and using equation (5.7), we get that

$$\begin{aligned}
M^{\theta}_{0,k,n} &= \widehat{\mathrm{E}}^{\theta^0}_{t_k, X^{\theta^0}_{t_k}} \Big[\mathbf{1}_{\widehat{N}_{0,k}(\upsilon_n)} \Big(e_k(\theta) \Big(- \int_{t_k}^{t_{k+1}} \int_{|z|>\upsilon_n} c(X^{\theta^0}_{s-}, z) \nu(dz) ds \\
&\quad - \widetilde{\mathrm{E}}^{\theta}_{t_k, X^{\theta^0}_{t_k}} \Big[(\mathbf{1}_{\widetilde{N}_{0,k}(\upsilon_n)} + \mathbf{1}_{\widetilde{N}_{\geq 1,k}(\upsilon_n)}) \int_{t_k}^{t_{k+1}} \int_{|z|>\upsilon_n} c(Y^{\theta}_{s-}, z) \widetilde{M}(ds, dz) \big| Y^{\theta}_{t_{k+1}} = X^{\theta^0}_{t_{k+1}} \Big] \Big) \Big)^2 \Big] \\
&= \widehat{\mathrm{E}}^{\theta^0}_{t_k, X^{\theta^0}_{t_k}} \Big[\mathbf{1}_{\widehat{N}_{0,k}(\upsilon_n)} \Big(e_k(\theta) \Big(- \int_{t_k}^{t_{k+1}} \int_{|z|>\upsilon_n} c(X^{\theta^0}_{s-}, z) \nu(dz) ds \\
&\quad + \widetilde{\mathrm{E}}^{\theta}_{t_k, X^{\theta^0}_{t_k}} \Big[\mathbf{1}_{\widetilde{N}_{0,k}(\upsilon_n)} \int_{t_k}^{t_{k+1}} \int_{|z|>\upsilon_n} c(Y^{\theta}_{s-}, z) \nu(dz) ds \big| Y^{\theta}_{t_{k+1}} = X^{\theta^0}_{t_{k+1}} \Big] \\
&\quad - \widetilde{\mathrm{E}}^{\theta}_{t_k, X^{\theta^0}_{t_k}} \Big[\mathbf{1}_{\widetilde{N}_{\geq 1,k}(\upsilon_n)} \int_{t_k}^{t_{k+1}} \int_{|z|>\upsilon_n} c(Y^{\theta}_{s-}, z) \widetilde{M}(ds, dz) \big| Y^{\theta}_{t_{k+1}} = X^{\theta^0}_{t_{k+1}} \Big] \Big) \Big)^2 \Big] \\
&= \widehat{\mathrm{E}}^{\theta^0}_{t_k, X^{\theta^0}_{t_k}} \Big[\mathbf{1}_{\widehat{N}_{0,k}(\upsilon_n)} \Big(e_k(\theta) \Big(- \int_{t_k}^{t_{k+1}} \int_{|z|>\upsilon_n} c(X^{\theta^0}_{s-}, z) \nu(dz) ds \\
&\quad + \widetilde{\mathrm{E}}^{\theta}_{t_k, X^{\theta^0}_{t_k}} \Big[\mathbf{1}_{\widetilde{N}_{0,k}(\upsilon_n)} \int_{t_k}^{t_{k+1}} \int_{|z|>\upsilon_n} c(Y^{\theta}_{s-}, z) \nu(dz) ds \big| Y^{\theta}_{t_{k+1}} = X^{\theta^0}_{t_{k+1}} \Big] \\
&\quad - \widetilde{\mathrm{E}}^{\theta}_{t_k, X^{\theta^0}_{t_k}} \Big[\mathbf{1}_{\widetilde{N}_{\geq 1,k}(\upsilon_n)} \Big(Y^{\theta}_{t_{k+1}} - Y^{\theta}_{t_k} - \int_{t_k}^{t_{k+1}} b(\theta, Y^{\theta}_s) ds - \int_{t_k}^{t_{k+1}} \sigma(Y^{\theta}_s) dW_s \\
&\quad - \int_{t_k}^{t_{k+1}} \int_{|z|\leq\upsilon_n} c(Y^{\theta}_{s-}, z) \widetilde{M}(ds, dz) \Big) \big| Y^{\theta}_{t_{k+1}} = X^{\theta^0}_{t_{k+1}} \Big] \Big) \Big)^2 \Big] \leq 6 \sum_{i=1}^{6} M^{\theta}_{0,i,k,n},
\end{aligned} \tag{5.12}$$

where

$$M^{\theta}_{0,1,k,n} = \widehat{\mathrm{E}}^{\theta^0}_{t_k, X^{\theta^0}_{t_k}} \Big[\mathbf{1}_{\widehat{N}_{0,k}(\upsilon_n)} \Big(e_k(\theta) \int_{t_k}^{t_{k+1}} \int_{|z|>\upsilon_n} c(X^{\theta^0}_{s-}, z) \nu(dz) ds \Big)^2 \Big],$$

$$M^{\theta}_{0,2,k,n} = \widehat{\mathrm{E}}^{\theta^0}_{t_k, X^{\theta^0}_{t_k}} \Big[\mathbf{1}_{\widehat{N}_{0,k}(\upsilon_n)} \Big(e_k(\theta) \widetilde{\mathrm{E}}^{\theta}_{t_k, X^{\theta^0}_{t_k}} \Big[\mathbf{1}_{\widetilde{N}_{0,k}(\upsilon_n)} \int_{t_k}^{t_{k+1}} \int_{|z|>\upsilon_n} c(Y^{\theta}_{s-}, z) \nu(dz) ds \big| Y^{\theta}_{t_{k+1}} = X^{\theta^0}_{t_{k+1}} \Big] \Big)^2 \Big],$$

$$M^{\theta}_{0,3,k,n} = \widehat{\mathrm{E}}^{\theta^0}_{t_k, X^{\theta^0}_{t_k}} \Big[\mathbf{1}_{\widehat{N}_{0,k}(\upsilon_n)} \Big(e_k(\theta) \widetilde{\mathrm{E}}^{\theta}_{t_k, X^{\theta^0}_{t_k}} \Big[\mathbf{1}_{\widetilde{N}_{\geq 1,k}(\upsilon_n)} (Y^{\theta}_{t_{k+1}} - Y^{\theta}_{t_k}) \big| Y^{\theta}_{t_{k+1}} = X^{\theta^0}_{t_{k+1}} \Big] \Big)^2 \Big],$$

$$M^{\theta}_{0,4,k,n} = \widehat{\mathrm{E}}^{\theta^0}_{t_k, X^{\theta^0}_{t_k}} \Big[\mathbf{1}_{\widehat{N}_{0,k}(\upsilon_n)} \Big(e_k(\theta) \widetilde{\mathrm{E}}^{\theta}_{t_k, X^{\theta^0}_{t_k}} \Big[\mathbf{1}_{\widetilde{N}_{\geq 1,k}(\upsilon_n)} \int_{t_k}^{t_{k+1}} b(\theta, Y^{\theta}_s) ds \big| Y^{\theta}_{t_{k+1}} = X^{\theta^0}_{t_{k+1}} \Big] \Big)^2 \Big],$$

$$M^{\theta}_{0,5,k,n}=\widehat{\mathrm{E}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}\left[\mathbf{1}_{\widehat{N}_{0,k}(\upsilon_n)}\left(e_k(\theta)\widetilde{\mathrm{E}}^{\theta}_{t_k,X^{\theta^0}_{t_k}}\left[\mathbf{1}_{\widetilde{N}_{\geq 1,k}(\upsilon_n)}\int_{t_k}^{t_{k+1}}\sigma(Y^{\theta}_s)dW_s\big|Y^{\theta}_{t_{k+1}}=X^{\theta^0}_{t_{k+1}}\right]\right)^2\right],$$

$$M^{\theta}_{0,6,k,n}=\widehat{\mathrm{E}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}\left[\mathbf{1}_{\widehat{N}_{0,k}(\upsilon_n)}\left(e_k(\theta)\widetilde{\mathrm{E}}^{\theta}_{t_k,X^{\theta^0}_{t_k}}\left[\mathbf{1}_{\widetilde{N}_{\geq 1,k}(\upsilon_n)}\int_{t_k}^{t_{k+1}}\int_{|z|\leq\upsilon_n}c(Y^{\theta}_{s-},z)\widetilde{M}(ds,dz)\big|Y^{\theta}_{t_{k+1}}=X^{\theta^0}_{t_{k+1}}\right]\right)^2\right].$$

First,

$$\begin{aligned}M^{\theta}_{0,1,k,n}&\leq\widehat{\mathrm{E}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}\left[\left(e_k(\theta)\int_{t_k}^{t_{k+1}}\int_{|z|>\upsilon_n}c(X^{\theta^0}_{s-},z)\nu(dz)ds\right)^2\right]\\&\leq C\left(1+|X^{\theta^0}_{t_k}|^{q_1}\right)\Delta_n^2\Big(\int_{\mathbb{R}^d_0}\zeta(z)\nu(dz)\Big)^2,\end{aligned}\tag{5.13}$$

for some constant $q_1>0$. Next, using Lemma 3.6,

$$\begin{aligned}M^{\theta}_{0,2,k,n}&\leq\widehat{\mathrm{E}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}\left[\widetilde{\mathrm{E}}^{\theta}_{t_k,X^{\theta^0}_{t_k}}\left[\left(e_k(\theta)\int_{t_k}^{t_{k+1}}\int_{|z|>\upsilon_n}c(Y^{\theta}_{s-},z)\nu(dz)ds\right)^2\big|Y^{\theta}_{t_{k+1}}=X^{\theta^0}_{t_{k+1}}\right]\right]\\&=\widetilde{\mathrm{E}}^{\theta}_{t_k,X^{\theta^0}_{t_k}}\left[\left(e_k(\theta)\int_{t_k}^{t_{k+1}}\int_{|z|>\upsilon_n}c(Y^{\theta}_{s-},z)\nu(dz)ds\right)^2\right]\\&\leq C\left(1+|X^{\theta^0}_{t_k}|^{q_1}\right)\Delta_n^2\Big(\int_{\mathbb{R}^d_0}\zeta(z)\nu(dz)\Big)^2,\end{aligned}\tag{5.14}$$

for some constant $q_1>0$. Using Lemma 3.6, Hölder's inequality with $\frac{1}{p}+\frac{1}{q}=1$, and the fact that $\widetilde{\mathrm{P}}^{\theta}_{t_k,X^{\theta^0}_{t_k}}(\widetilde{N}_{\geq 1,k}(\upsilon_n))\leq C\lambda_{\upsilon_n}\Delta_n e^{-\lambda_{\upsilon_n}\Delta_n}\leq C\lambda_{\upsilon_n}\Delta_n$, we get that

$$\begin{aligned}M^{\theta}_{0,4,k,n}&\leq\widehat{\mathrm{E}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}\left[\widetilde{\mathrm{E}}^{\theta}_{t_k,X^{\theta^0}_{t_k}}\left[\mathbf{1}_{\widetilde{N}_{\geq 1,k}(\upsilon_n)}\left(e_k(\theta)\int_{t_k}^{t_{k+1}}b(\theta,Y^{\theta}_s)ds\right)^2\big|Y^{\theta}_{t_{k+1}}=X^{\theta^0}_{t_{k+1}}\right]\right]\\&=\widetilde{\mathrm{E}}^{\theta}_{t_k,X^{\theta^0}_{t_k}}\left[\mathbf{1}_{\widetilde{N}_{\geq 1,k}(\upsilon_n)}\left(e_k(\theta)\int_{t_k}^{t_{k+1}}b(\theta,Y^{\theta}_s)ds\right)^2\right]\\&\leq\left(\widetilde{\mathrm{E}}^{\theta}_{t_k,X^{\theta^0}_{t_k}}\left[\left(e_k(\theta)\int_{t_k}^{t_{k+1}}b(\theta,Y^{\theta}_s)ds\right)^{2p}\right]\right)^{\frac{1}{p}}\left(\widetilde{\mathrm{P}}^{\theta}_{t_k,X^{\theta^0}_{t_k}}(\widetilde{N}_{\geq 1,k}(\upsilon_n))\right)^{\frac{1}{q}}\\&\leq C\left(1+|X^{\theta^0}_{t_k}|^{q_1}\right)\Delta_n^2\left(\lambda_{\upsilon_n}\Delta_n\right)^{\frac{1}{q}},\end{aligned}\tag{5.15}$$

for some constant $q_1>0$. Similarly,

$$\begin{aligned}M^{\theta}_{0,5,k,n}&\leq\widehat{\mathrm{E}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}\left[\widetilde{\mathrm{E}}^{\theta}_{t_k,X^{\theta^0}_{t_k}}\left[\mathbf{1}_{\widetilde{N}_{\geq 1,k}(\upsilon_n)}\left(e_k(\theta)\int_{t_k}^{t_{k+1}}\sigma(Y^{\theta}_s)dW_s\right)^2\big|Y^{\theta}_{t_{k+1}}=X^{\theta^0}_{t_{k+1}}\right]\right]\\&=\widetilde{\mathrm{E}}^{\theta}_{t_k,X^{\theta^0}_{t_k}}\left[\mathbf{1}_{\widetilde{N}_{\geq 1,k}(\upsilon_n)}\left(e_k(\theta)\int_{t_k}^{t_{k+1}}\sigma(Y^{\theta}_s)dW_s\right)^2\right]\\&\leq\left(\widetilde{\mathrm{E}}^{\theta}_{t_k,X^{\theta^0}_{t_k}}\left[\left(e_k(\theta)\int_{t_k}^{t_{k+1}}\sigma(Y^{\theta}_s)dW_s\right)^{2p}\right]\right)^{\frac{1}{p}}\left(\widetilde{\mathrm{P}}^{\theta}_{t_k,X^{\theta^0}_{t_k}}(\widetilde{N}_{\geq 1,k}(\upsilon_n))\right)^{\frac{1}{q}}\\&\leq C\left(1+|X^{\theta^0}_{t_k}|^{q_1}\right)\Delta_n\left(\lambda_{\upsilon_n}\Delta_n\right)^{\frac{1}{q}}.\end{aligned}\tag{5.16}$$

Next, using Lemma 3.6, Burkholder-Davis-Gundy's inequality and condition **(A1)**,

$$\begin{aligned}M_{0,6,k,n}^{\theta} &\leq \widehat{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta^0}\left[\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta}\left[\left(e_k(\theta)\int_{t_k}^{t_{k+1}}\int_{|z|\leq\upsilon_n}c(Y_{s-}^{\theta},z)\widetilde{M}(ds,dz)\right)^2\big|Y_{t_{k+1}}^{\theta}=X_{t_{k+1}}^{\theta^0}\right]\right]\\ &=\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta}\left[\left(e_k(\theta)\int_{t_k}^{t_{k+1}}\int_{|z|\leq\upsilon_n}c(Y_{s-}^{\theta},z)\widetilde{M}(ds,dz)\right)^2\right]\\ &\leq C\left(1+|X_{t_k}^{\theta^0}|^{q_1}\right)\Delta_n\int_{|z|\leq\upsilon_n}\zeta^2(z)\nu(dz).\end{aligned}\tag{5.17}$$

Now, we treat the term $M_{0,3,k,n}^{\theta}$. For this, using equation (5.6) and the fact that there is no big jump of Z^{υ_n} in the interval $[t_k,t_{k+1})$, we get that

$$\begin{aligned}M_{0,3,k,n}^{\theta} &= \widehat{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta^0}\left[\mathbf{1}_{\widehat{N}_{0,k}(\upsilon_n)}\left(e_k(\theta)\big(X_{t_{k+1}}^{\theta^0}-X_{t_k}^{\theta^0}\big)\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta}\left[\mathbf{1}_{\widetilde{N}_{\geq 1,k}(\upsilon_n)}\big|Y_{t_{k+1}}^{\theta}=X_{t_{k+1}}^{\theta^0}\right]\right)^2\right]\\ &= \widehat{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta^0}\Bigg[\mathbf{1}_{\widehat{N}_{0,k}(\upsilon_n)}\Bigg(e_k(\theta)\Bigg(\int_{t_k}^{t_{k+1}}b(\theta,X_s^{\theta^0})ds+\int_{t_k}^{t_{k+1}}\sigma(X_s^{\theta^0})dB_s\\ &\qquad+\int_{t_k}^{t_{k+1}}\int_{|z|\leq\upsilon_n}c(X_{s-}^{\theta^0},z)\widetilde{N}(ds,dz)-\int_{t_k}^{t_{k+1}}\int_{|z|>\upsilon_n}c(X_{s-}^{\theta^0},z)\nu(dz)ds\Bigg)\\ &\qquad\times\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta}\left[\mathbf{1}_{\widetilde{N}_{\geq 1,k}(\upsilon_n)}\big|Y_{t_{k+1}}^{\theta}=X_{t_{k+1}}^{\theta^0}\right]\Bigg)^2\Bigg]\leq 4\sum_{i=1}^{4}M_{0,3,i,k,n}^{\theta},\end{aligned}\tag{5.18}$$

where

$$\begin{aligned}M_{0,3,1,k,n}^{\theta} &= \widehat{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta^0}\left[\mathbf{1}_{\widehat{N}_{0,k}(\upsilon_n)}\left(e_k(\theta)\int_{t_k}^{t_{k+1}}b(\theta,X_s^{\theta^0})ds\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta}\left[\mathbf{1}_{\widetilde{N}_{\geq 1,k}(\upsilon_n)}\big|Y_{t_{k+1}}^{\theta}=X_{t_{k+1}}^{\theta^0}\right]\right)^2\right],\\ M_{0,3,2,k,n}^{\theta} &= \widehat{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta^0}\left[\mathbf{1}_{\widehat{N}_{0,k}(\upsilon_n)}\left(e_k(\theta)\int_{t_k}^{t_{k+1}}\sigma(X_s^{\theta^0})dB_s\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta}\left[\mathbf{1}_{\widetilde{N}_{\geq 1,k}(\upsilon_n)}\big|Y_{t_{k+1}}^{\theta}=X_{t_{k+1}}^{\theta^0}\right]\right)^2\right],\\ M_{0,3,3,k,n}^{\theta} &= \widehat{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta^0}\left[\mathbf{1}_{\widehat{N}_{0,k}(\upsilon_n)}\left(e_k(\theta)\int_{t_k}^{t_{k+1}}\int_{|z|\leq\upsilon_n}c(X_{s-}^{\theta^0},z)\widetilde{N}(ds,dz)\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta}\left[\mathbf{1}_{\widetilde{N}_{\geq 1,k}(\upsilon_n)}\big|Y_{t_{k+1}}^{\theta}=X_{t_{k+1}}^{\theta^0}\right]\right)^2\right],\\ M_{0,3,4,k,n}^{\theta} &= \widehat{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta^0}\left[\mathbf{1}_{\widehat{N}_{0,k}(\upsilon_n)}\left(e_k(\theta)\int_{t_k}^{t_{k+1}}\int_{|z|>\upsilon_n}c(X_{s-}^{\theta^0},z)\nu(dz)ds\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta}\left[\mathbf{1}_{\widetilde{N}_{\geq 1,k}(\upsilon_n)}\big|Y_{t_{k+1}}^{\theta}=X_{t_{k+1}}^{\theta^0}\right]\right)^2\right].\end{aligned}$$

Using Hölder's inequality with $\frac{1}{p}+\frac{1}{q}=1$ and Jensen's inequality together with Lemma 3.6,

$$\begin{aligned}&M_{0,3,1,k,n}^{\theta} \leq \left(\widehat{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta^0}\left[\left(e_k(\theta)\int_{t_k}^{t_{k+1}}b(\theta,X_s^{\theta^0})ds\right)^{2p}\right]\right)^{\frac{1}{p}}\left(\widehat{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta^0}\left[\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta}\left[\mathbf{1}_{\widetilde{N}_{\geq 1,k}(\upsilon_n)}\big|Y_{t_{k+1}}^{\theta}=X_{t_{k+1}}^{\theta^0}\right]\right]\right)^{\frac{1}{q}}\\ &\leq\left(\Delta_n^{2p-1}\int_{t_k}^{t_{k+1}}\widehat{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta^0}\left[\left(e_k(\theta)b(\theta,X_s^{\theta^0})ds\right)^{2p}\right]ds\right)^{\frac{1}{p}}\left(\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta}\left[\mathbf{1}_{\widetilde{N}_{\geq 1,k}(\upsilon_n)}\right]\right)^{\frac{1}{q}}\\ &=\left(\Delta_n^{2p-1}\int_{t_k}^{t_{k+1}}\widehat{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta^0}\left[\left(e_k(\theta)b(\theta,X_s^{\theta^0})ds\right)^{2p}\right]ds\right)^{\frac{1}{p}}\left(\widetilde{\mathrm{P}}_{t_k,X_{t_k}^{\theta^0}}^{\theta}\big(\widetilde{N}_{\geq 1,k}(\upsilon_n)\big)\right)^{\frac{1}{q}}\\ &\leq C\left(1+|X_{t_k}^{\theta^0}|^{q_1}\right)\Delta_n^2\left(\lambda_{\upsilon_n}\Delta_n\right)^{\frac{1}{q}}.\end{aligned}\tag{5.19}$$

Next, using Hölder's inequality with $\frac{1}{p}+\frac{1}{q}=1$ and Burkholder-Davis-Gundy's inequality together with Lemma 3.6,

$$M^{\theta}_{0,3,2,k,n} \leq C\left(1+|X^{\theta^0}_{t_k}|^{q_1}\right)\Delta_n\left(\lambda_{\upsilon_n}\Delta_n\right)^{\frac{1}{q}}. \tag{5.20}$$

Using Burkholder-Davis-Gundy's inequality and condition **(A1)**,

$$\begin{aligned} M^{\theta}_{0,3,3,k,n} &\leq \widehat{\mathrm{E}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}\left[\left(e_k(\theta)\int_{t_k}^{t_{k+1}}\int_{|z|\leq\upsilon_n} c(X^{\theta^0}_{s-},z)\widetilde{N}(ds,dz)\right)^2\right] \\ &\leq C\left(1+|X^{\theta^0}_{t_k}|^{q_1}\right)\Delta_n\int_{|z|\leq\upsilon_n}\zeta^2(z)\nu(dz). \end{aligned} \tag{5.21}$$

Observe that

$$\begin{aligned} M^{\theta}_{0,3,4,k,n} &\leq \widehat{\mathrm{E}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}\left[\left(e_k(\theta)\int_{t_k}^{t_{k+1}}\int_{|z|>\upsilon_n} c(X^{\theta^0}_{s-},z)\nu(dz)ds\right)^2\right] \\ &\leq C\left(1+|X^{\theta^0}_{t_k}|^{q_1}\right)\Delta_n^2\Big(\int_{\mathbb{R}^d_0}\zeta(z)\nu(dz)\Big)^2. \end{aligned} \tag{5.22}$$

Therefore, from (5.18)-(5.22), we have shown that

$$M^{\theta}_{0,3,k,n} \leq C\left(1+|X^{\theta^0}_{t_k}|^{q_1}\right)\Delta_n\bigg(\left(\lambda_{\upsilon_n}\Delta_n\right)^{\frac{1}{q}}+\int_{|z|\leq\upsilon_n}\zeta^2(z)\nu(dz)+\Delta_n\Big(\int_{\mathbb{R}^d_0}\zeta(z)\nu(dz)\Big)^2\bigg). \tag{5.23}$$

Thus, from (5.12)-(5.17) and (5.23), we have shown that

$$M^{\theta}_{0,k,n} \leq C\left(1+|X^{\theta^0}_{t_k}|^{q_1}\right)\Delta_n\bigg(\left(\lambda_{\upsilon_n}\Delta_n\right)^{\frac{1}{q}}+\int_{|z|\leq\upsilon_n}\zeta^2(z)\nu(dz)+\Delta_n\Big(\int_{\mathbb{R}^d_0}\zeta(z)\nu(dz)\Big)^2\bigg). \tag{5.24}$$

Finally, we treat $M^{\theta}_{\geq 1,k,n}$. Multiplying the random variable inside the conditional expectation of $M^{\theta}_{\geq 1,k,n}$ by $\mathbf{1}_{\widetilde{N}_{0,k}(\upsilon_n)}+\mathbf{1}_{\widetilde{N}_{\geq 1,k}(\upsilon_n)}$ and using equations (5.6) and (5.7), we get that

$$\begin{aligned} M^{\theta}_{\geq 1,k,n} &= \widehat{\mathrm{E}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}\bigg[\mathbf{1}_{\widehat{N}_{\geq 1,k}(\upsilon_n)}\bigg(e_k(\theta)\bigg(\int_{t_k}^{t_{k+1}}\int_{|z|>\upsilon_n} c(X^{\theta^0}_{s-},z)\widetilde{N}(ds,dz) \\ &\quad -\widetilde{\mathrm{E}}^{\theta}_{t_k,X^{\theta^0}_{t_k}}\Big[\left(\mathbf{1}_{\widetilde{N}_{0,k}(\upsilon_n)}+\mathbf{1}_{\widetilde{N}_{\geq 1,k}(\upsilon_n)}\right)\int_{t_k}^{t_{k+1}}\int_{|z|>\upsilon_n} c(Y^{\theta}_{s-},z)\widetilde{M}(ds,dz)\Big|Y^{\theta}_{t_{k+1}}=X^{\theta^0}_{t_{k+1}}\Big]\bigg)\bigg)^2\bigg] \\ &= \widehat{\mathrm{E}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}\bigg[\mathbf{1}_{\widehat{N}_{\geq 1,k}(\upsilon_n)}\bigg(e_k(\theta)\bigg(\int_{t_k}^{t_{k+1}}\int_{|z|>\upsilon_n} c(X^{\theta^0}_{s-},z)\widetilde{N}(ds,dz) \\ &\quad -\widetilde{\mathrm{E}}^{\theta}_{t_k,X^{\theta^0}_{t_k}}\Big[\mathbf{1}_{\widetilde{N}_{\geq 1,k}(\upsilon_n)}\int_{t_k}^{t_{k+1}}\int_{|z|>\upsilon_n} c(Y^{\theta}_{s-},z)\widetilde{M}(ds,dz)\Big|Y^{\theta}_{t_{k+1}}=X^{\theta^0}_{t_{k+1}}\Big] \\ &\quad +\widetilde{\mathrm{E}}^{\theta}_{t_k,X^{\theta^0}_{t_k}}\Big[\mathbf{1}_{\widetilde{N}_{0,k}(\upsilon_n)}\int_{t_k}^{t_{k+1}}\int_{|z|>\upsilon_n} c(Y^{\theta}_{s-},z)\nu(dz)ds\Big|Y^{\theta}_{t_{k+1}}=X^{\theta^0}_{t_{k+1}}\Big]\bigg)\bigg)^2\bigg] \\ &= \widehat{\mathrm{E}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}\bigg[\mathbf{1}_{\widehat{N}_{\geq 1,k}(\upsilon_n)}\bigg(e_k(\theta)\bigg(X^{\theta^0}_{t_{k+1}}-X^{\theta^0}_{t_k}-\int_{t_k}^{t_{k+1}}b(\theta,X^{\theta^0}_s)ds-\int_{t_k}^{t_{k+1}}\sigma(X^{\theta^0}_s)dB_s \end{aligned}$$

$$
\begin{aligned}
&\quad - \int_{t_k}^{t_{k+1}} \int_{|z|\le \upsilon_n} c(X_{s-}^{\theta^0}, z)\widetilde{N}(ds,dz) - \widetilde{\mathrm{E}}_{t_k, X_{t_k}^{\theta^0}}^{\theta}\bigg[\mathbf{1}_{\widetilde{N}_{\ge 1,k}(\upsilon_n)}\bigg(Y_{t_{k+1}}^{\theta} - Y_{t_k}^{\theta}\\
&- \int_{t_k}^{t_{k+1}} b(\theta, Y_s^{\theta})ds - \int_{t_k}^{t_{k+1}} \sigma(Y_s^{\theta})dW_s - \int_{t_k}^{t_{k+1}} \int_{|z|\le \upsilon_n} c(Y_{s-}^{\theta}, z)\widetilde{M}(ds,dz)\bigg)\Big|Y_{t_{k+1}}^{\theta} = X_{t_{k+1}}^{\theta^0}\bigg]\\
&\quad + \widetilde{\mathrm{E}}_{t_k, X_{t_k}^{\theta^0}}^{\theta}\bigg[\mathbf{1}_{\widetilde{N}_{0,k}(\upsilon_n)} \int_{t_k}^{t_{k+1}} \int_{|z|>\upsilon_n} c(Y_{s-}^{\theta}, z)\nu(dz)ds\Big|Y_{t_{k+1}}^{\theta} = X_{t_{k+1}}^{\theta^0}\bigg]\bigg)\bigg)^2\bigg]\\
&\le 8\sum_{i=1}^{8} M_{\ge 1,i,k,n}^{\theta},
\end{aligned}
\tag{5.25}
$$

where

$$
\begin{aligned}
M_{\ge 1,1,k,n}^{\theta} &= \widehat{\mathrm{E}}_{t_k, X_{t_k}^{\theta^0}}^{\theta^0}\bigg[\mathbf{1}_{\widehat{N}_{\ge 1,k}(\upsilon_n)}\bigg(e_k(\theta)\bigg(X_{t_{k+1}}^{\theta^0} - X_{t_k}^{\theta^0}\\
&\qquad - \widetilde{\mathrm{E}}_{t_k, X_{t_k}^{\theta^0}}^{\theta}\Big[\mathbf{1}_{\widetilde{N}_{\ge 1,k}(\upsilon_n)}\big(Y_{t_{k+1}}^{\theta} - Y_{t_k}^{\theta}\big)\Big|Y_{t_{k+1}}^{\theta} = X_{t_{k+1}}^{\theta^0}\Big]\bigg)\bigg)^2\bigg],\\
M_{\ge 1,2,k,n}^{\theta} &= \widehat{\mathrm{E}}_{t_k, X_{t_k}^{\theta^0}}^{\theta^0}\bigg[\mathbf{1}_{\widehat{N}_{\ge 1,k}(\upsilon_n)}\bigg(e_k(\theta)\int_{t_k}^{t_{k+1}} b(\theta, X_s^{\theta^0})ds\bigg)^2\bigg],\\
M_{\ge 1,3,k,n}^{\theta} &= \widehat{\mathrm{E}}_{t_k, X_{t_k}^{\theta^0}}^{\theta^0}\bigg[\mathbf{1}_{\widehat{N}_{\ge 1,k}(\upsilon_n)}\bigg(e_k(\theta)\int_{t_k}^{t_{k+1}} \sigma(X_s^{\theta^0})dB_s\bigg)^2\bigg],\\
M_{\ge 1,4,k,n}^{\theta} &= \widehat{\mathrm{E}}_{t_k, X_{t_k}^{\theta^0}}^{\theta^0}\bigg[\mathbf{1}_{\widehat{N}_{\ge 1,k}(\upsilon_n)}\bigg(e_k(\theta)\int_{t_k}^{t_{k+1}} \int_{|z|\le \upsilon_n} c(X_{s-}^{\theta^0}, z)\widetilde{N}(ds,dz)\bigg)^2\bigg],\\
M_{\ge 1,5,k,n}^{\theta} &= \widehat{\mathrm{E}}_{t_k, X_{t_k}^{\theta^0}}^{\theta^0}\bigg[\mathbf{1}_{\widehat{N}_{\ge 1,k}(\upsilon_n)}\bigg(e_k(\theta)\widetilde{\mathrm{E}}_{t_k, X_{t_k}^{\theta^0}}^{\theta}\Big[\mathbf{1}_{\widetilde{N}_{\ge 1,k}(\upsilon_n)}\int_{t_k}^{t_{k+1}} b(\theta, Y_s^{\theta})ds\Big|Y_{t_{k+1}}^{\theta} = X_{t_{k+1}}^{\theta^0}\Big]\bigg)^2\bigg],\\
M_{\ge 1,6,k,n}^{\theta} &= \widehat{\mathrm{E}}_{t_k, X_{t_k}^{\theta^0}}^{\theta^0}\bigg[\mathbf{1}_{\widehat{N}_{\ge 1,k}(\upsilon_n)}\bigg(e_k(\theta)\widetilde{\mathrm{E}}_{t_k, X_{t_k}^{\theta^0}}^{\theta}\Big[\mathbf{1}_{\widetilde{N}_{\ge 1,k}(\upsilon_n)}\int_{t_k}^{t_{k+1}} \sigma(Y_s^{\theta})dW_s\Big|Y_{t_{k+1}}^{\theta} = X_{t_{k+1}}^{\theta^0}\Big]\bigg)^2\bigg],\\
M_{\ge 1,7,k,n}^{\theta} &= \widehat{\mathrm{E}}_{t_k, X_{t_k}^{\theta^0}}^{\theta^0}\bigg[\mathbf{1}_{\widehat{N}_{\ge 1,k}(\upsilon_n)}\bigg(e_k(\theta)\widetilde{\mathrm{E}}_{t_k, X_{t_k}^{\theta^0}}^{\theta}\Big[\mathbf{1}_{\widetilde{N}_{\ge 1,k}(\upsilon_n)}\int_{t_k}^{t_{k+1}} \int_{|z|\le \upsilon_n} c(Y_{s-}^{\theta}, z)\widetilde{M}(ds,dz)\Big|Y_{t_{k+1}}^{\theta} = X_{t_{k+1}}^{\theta^0}\Big]\bigg)^2\bigg],\\
M_{\ge 1,8,k,n}^{\theta} &= \widehat{\mathrm{E}}_{t_k, X_{t_k}^{\theta^0}}^{\theta^0}\bigg[\mathbf{1}_{\widehat{N}_{\ge 1,k}(\upsilon_n)}\bigg(e_k(\theta)\widetilde{\mathrm{E}}_{t_k, X_{t_k}^{\theta^0}}^{\theta}\Big[\mathbf{1}_{\widetilde{N}_{0,k}(\upsilon_n)}\int_{t_k}^{t_{k+1}} \int_{|z|>\upsilon_n} c(Y_{s-}^{\theta}, z)\nu(dz)ds\Big|Y_{t_{k+1}}^{\theta} = X_{t_{k+1}}^{\theta^0}\Big]\bigg)^2\bigg].
\end{aligned}
$$

First, we treat the term $M_{\ge 1,1,k,n}^{\theta}$. Using equation (5.7) and the fact that there is no big jump of $\widetilde{Z}^{\upsilon_n}$ in the interval $[t_k, t_{k+1})$, we get that

$$
\begin{aligned}
M_{\ge 1,1,k,n}^{\theta} &= \widehat{\mathrm{E}}_{t_k, X_{t_k}^{\theta^0}}^{\theta^0}\bigg[\mathbf{1}_{\widehat{N}_{\ge 1,k}(\upsilon_n)}\bigg(e_k(\theta)\bigg(X_{t_{k+1}}^{\theta^0} - X_{t_k}^{\theta^0}\\
&\quad - \big(X_{t_{k+1}}^{\theta^0} - X_{t_k}^{\theta^0}\big)\widetilde{\mathrm{E}}_{t_k, X_{t_k}^{\theta^0}}^{\theta}\Big[\mathbf{1}_{\widetilde{N}_{\ge 1,k}(\upsilon_n)}\Big|Y_{t_{k+1}}^{\theta} = X_{t_{k+1}}^{\theta^0}\Big]\bigg)\bigg)^2\bigg]\\
&= \widehat{\mathrm{E}}_{t_k, X_{t_k}^{\theta^0}}^{\theta^0}\bigg[\mathbf{1}_{\widehat{N}_{\ge 1,k}(\upsilon_n)}\bigg(e_k(\theta)\big(X_{t_{k+1}}^{\theta^0} - X_{t_k}^{\theta^0}\big)\bigg(1 - \widetilde{\mathrm{E}}_{t_k, X_{t_k}^{\theta^0}}^{\theta}\Big[\mathbf{1}_{\widetilde{N}_{\ge 1,k}(\upsilon_n)}\Big|Y_{t_{k+1}}^{\theta} = X_{t_{k+1}}^{\theta^0}\Big]\bigg)\bigg)^2\bigg]\\
&= \widehat{\mathrm{E}}_{t_k, X_{t_k}^{\theta^0}}^{\theta^0}\bigg[\mathbf{1}_{\widehat{N}_{\ge 1,k}(\upsilon_n)}\bigg(e_k(\theta)\big(X_{t_{k+1}}^{\theta^0} - X_{t_k}^{\theta^0}\big)\widetilde{\mathrm{E}}_{t_k, X_{t_k}^{\theta^0}}^{\theta}\Big[1 - \mathbf{1}_{\widetilde{N}_{\ge 1,k}(\upsilon_n)}\Big|Y_{t_{k+1}}^{\theta} = X_{t_{k+1}}^{\theta^0}\Big]\bigg)^2\bigg]
\end{aligned}
$$

$$
\begin{aligned}
&= \widehat{\mathrm{E}}^{\theta^0}_{t_k, X^{\theta^0}_{t_k}} \Big[\mathbf{1}_{\widehat{N}_{\geq 1,k}(\upsilon_n)} \Big(e_k(\theta) \big(X^{\theta^0}_{t_{k+1}} - X^{\theta^0}_{t_k} \big) \widetilde{\mathrm{E}}^{\theta}_{t_k, X^{\theta^0}_{t_k}} \Big[\mathbf{1}_{\widetilde{N}_{0,k}(\upsilon_n)} \Big| Y^{\theta}_{t_{k+1}} = X^{\theta^0}_{t_{k+1}} \Big] \Big)^2 \Big] \\
&= \widehat{\mathrm{E}}^{\theta^0}_{t_k, X^{\theta^0}_{t_k}} \Big[\mathbf{1}_{\widehat{N}_{\geq 1,k}(\upsilon_n)} \Big(e_k(\theta) \widetilde{\mathrm{E}}^{\theta}_{t_k, X^{\theta^0}_{t_k}} \Big[\mathbf{1}_{\widetilde{N}_{0,k}(\upsilon_n)} \big(Y^{\theta}_{t_{k+1}} - Y^{\theta}_{t_k} \big) \Big| Y^{\theta}_{t_{k+1}} = X^{\theta^0}_{t_{k+1}} \Big] \Big)^2 \Big] \\
&= \widehat{\mathrm{E}}^{\theta^0}_{t_k, X^{\theta^0}_{t_k}} \Big[\mathbf{1}_{\widehat{N}_{\geq 1,k}(\upsilon_n)} \Big(e_k(\theta) \widetilde{\mathrm{E}}^{\theta}_{t_k, X^{\theta^0}_{t_k}} \Big[\mathbf{1}_{\widetilde{N}_{0,k}(\upsilon_n)} \Big(\int_{t_k}^{t_{k+1}} b(\theta, Y^{\theta}_s) ds + \int_{t_k}^{t_{k+1}} \sigma(Y^{\theta}_s) dW_s \\
&\qquad + \int_{t_k}^{t_{k+1}} \int_{|z| \leq \upsilon_n} c(Y^{\theta}_{s-}, z) \widetilde{M}(ds, dz) - \int_{t_k}^{t_{k+1}} \int_{|z| > \upsilon_n} c(Y^{\theta}_{s-}, z) \nu(dz) ds \Big) \Big| Y^{\theta}_{t_{k+1}} = X^{\theta^0}_{t_{k+1}} \Big] \Big)^2 \Big] \\
&\leq 4 \sum_{i=1}^{4} M^{\theta}_{\geq 1,1,i,k,n},
\end{aligned}
\tag{5.26}
$$

where

$$
\begin{aligned}
M^{\theta}_{\geq 1,1,1,k,n} &= \widehat{\mathrm{E}}^{\theta^0}_{t_k, X^{\theta^0}_{t_k}} \Big[\mathbf{1}_{\widehat{N}_{\geq 1,k}(\upsilon_n)} \Big(e_k(\theta) \widetilde{\mathrm{E}}^{\theta}_{t_k, X^{\theta^0}_{t_k}} \Big[\mathbf{1}_{\widetilde{N}_{0,k}(\upsilon_n)} \int_{t_k}^{t_{k+1}} b(\theta, Y^{\theta}_s) ds \Big| Y^{\theta}_{t_{k+1}} = X^{\theta^0}_{t_{k+1}} \Big] \Big)^2 \Big], \\
M^{\theta}_{\geq 1,1,2,k,n} &= \widehat{\mathrm{E}}^{\theta^0}_{t_k, X^{\theta^0}_{t_k}} \Big[\mathbf{1}_{\widehat{N}_{\geq 1,k}(\upsilon_n)} \Big(e_k(\theta) \widetilde{\mathrm{E}}^{\theta}_{t_k, X^{\theta^0}_{t_k}} \Big[\mathbf{1}_{\widetilde{N}_{0,k}(\upsilon_n)} \int_{t_k}^{t_{k+1}} \sigma(Y^{\theta}_s) dW_s \Big| Y^{\theta}_{t_{k+1}} = X^{\theta^0}_{t_{k+1}} \Big] \Big)^2 \Big], \\
M^{\theta}_{\geq 1,1,3,k,n} &= \widehat{\mathrm{E}}^{\theta^0}_{t_k, X^{\theta^0}_{t_k}} \Big[\mathbf{1}_{\widehat{N}_{\geq 1,k}(\upsilon_n)} \Big(e_k(\theta) \widetilde{\mathrm{E}}^{\theta}_{t_k, X^{\theta^0}_{t_k}} \Big[\mathbf{1}_{\widetilde{N}_{0,k}(\upsilon_n)} \int_{t_k}^{t_{k+1}} \int_{|z| \leq \upsilon_n} c(Y^{\theta}_{s-}, z) \widetilde{M}(ds, dz) \Big| Y^{\theta}_{t_{k+1}} = X^{\theta^0}_{t_{k+1}} \Big] \Big)^2 \Big], \\
M^{\theta}_{\geq 1,1,4,k,n} &= \widehat{\mathrm{E}}^{\theta^0}_{t_k, X^{\theta^0}_{t_k}} \Big[\mathbf{1}_{\widehat{N}_{\geq 1,k}(\upsilon_n)} \Big(e_k(\theta) \widetilde{\mathrm{E}}^{\theta}_{t_k, X^{\theta^0}_{t_k}} \Big[\mathbf{1}_{\widetilde{N}_{0,k}(\upsilon_n)} \int_{t_k}^{t_{k+1}} \int_{|z| > \upsilon_n} c(Y^{\theta}_{s-}, z) \nu(dz) ds \Big| Y^{\theta}_{t_{k+1}} = X^{\theta^0}_{t_{k+1}} \Big] \Big)^2 \Big].
\end{aligned}
$$

Proceeding as the terms $M^{\theta}_{0,3,1,k,n}$, $M^{\theta}_{0,3,2,k,n}$, $M^{\theta}_{0,3,3,k,n}$ and $M^{\theta}_{0,3,4,k,n}$, we get that for any $q > 1$,

$$
\begin{aligned}
&M^{\theta}_{\geq 1,1,1,k,n} \leq C \left(1 + |X^{\theta^0}_{t_k}|^{q_1} \right) \Delta_n^2 \left(\lambda_{\upsilon_n} \Delta_n \right)^{\frac{1}{q}}, \quad M^{\theta}_{\geq 1,1,2,k,n} \leq C \left(1 + |X^{\theta^0}_{t_k}|^{q_1} \right) \Delta_n \left(\lambda_{\upsilon_n} \Delta_n \right)^{\frac{1}{q}}, \\
&M^{\theta}_{\geq 1,1,3,k,n} \leq C \left(1 + |X^{\theta^0}_{t_k}|^{q_1} \right) \Delta_n \int_{|z| \leq \upsilon_n} \zeta^2(z) \nu(dz), \\
&M^{\theta}_{\geq 1,1,4,k,n} \leq C \left(1 + |X^{\theta^0}_{t_k}|^{q_1} \right) \Delta_n^2 \Big(\int_{\mathbb{R}^d_0} \zeta(z) \nu(dz) \Big)^2,
\end{aligned}
\tag{5.27}
$$

for some constant $q_1 > 0$. Thus, from (5.26) and (5.27), we have shown that

$$
M^{\theta}_{\geq 1,1,k,n} \leq C \left(1 + |X^{\theta^0}_{t_k}|^{q_1} \right) \Delta_n \Big(\left(\lambda_{\upsilon_n} \Delta_n \right)^{\frac{1}{q}} + \int_{|z| \leq \upsilon_n} \zeta^2(z) \nu(dz) + \Delta_n \Big(\int_{\mathbb{R}^d_0} \zeta(z) \nu(dz) \Big)^2 \Big).
\tag{5.28}
$$

Similarly, we obtain that for any $q > 1$,

$$
\begin{aligned}
M^{\theta}_{\geq 1,2,k,n} + M^{\theta}_{\geq 1,5,k,n} &\leq C \left(1 + |X^{\theta^0}_{t_k}|^{q_1} \right) \Delta_n^2 \left(\lambda_{\upsilon_n} \Delta_n \right)^{\frac{1}{q}}, \\
M^{\theta}_{\geq 1,3,k,n} + M^{\theta}_{\geq 1,6,k,n} &\leq C \left(1 + |X^{\theta^0}_{t_k}|^{q_1} \right) \Delta_n \left(\lambda_{\upsilon_n} \Delta_n \right)^{\frac{1}{q}}, \\
M^{\theta}_{\geq 1,4,k,n} + M^{\theta}_{\geq 1,7,k,n} &\leq C \left(1 + |X^{\theta^0}_{t_k}|^{q_1} \right) \Delta_n \int_{|z| \leq \upsilon_n} \zeta^2(z) \nu(dz),
\end{aligned}
$$

$$M^{\theta}_{\geq 1,8,k,n} \leq C\left(1+|X^{\theta^0}_{t_k}|^{q_1}\right)\Delta_n^2\Big(\int_{\mathbb{R}^d_0}\zeta(z)\nu(dz)\Big)^2.$$

This, together with (5.25) and (5.28), concludes that for any $q>1$,

$$M^{\theta}_{\geq 1,k,n} \leq C\left(1+|X^{\theta^0}_{t_k}|^{q_1}\right)\Delta_n\bigg((\lambda_{\upsilon_n}\Delta_n)^{\frac{1}{q}}+\int_{|z|\leq\upsilon_n}\zeta^2(z)\nu(dz)+\Delta_n\Big(\int_{\mathbb{R}^d_0}\zeta(z)\nu(dz)\Big)^2\bigg), \tag{5.29}$$

for some constant $q_1>0$. Thus, the result follows from (5.8), (5.9), (5.10), (5.11), (5.24) and (5.29). □

References

[1] Applebaum, D. (2009), *Lévy processes and stochastic calculus*, volume 116 of *Cambridge Studies in Advanced Mathematics*, Cambridge University Press, Cambridge, second edition.

[2] Ben Alaya, M., Kebaier, A., Pap, G. and Tran, N.K. (2019), Local asymptotic properties for the growth rate of a jump-type CIR process, Preprint: `https://arxiv.org/pdf/1903.00358.pdf`.

[3] Ben Alaya, M., Kebaier, A. and Tran, N. K. (2017), Local asymptotic properties for Cox-Ingersoll-Ross process with discrete observations, Preprint: `https://arxiv.org/pdf/1708.07070.pdf`.

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

[5] Genon-Catalot, V. and Jacod, J. (1993), On the estimation of the diffusion coefficient for multi-dimensional diffusion processes, *Ann. Inst. H. Poincaré Probab. Statist.*, **29**(1), 119-151.

[6] Gloter, A., Loukianova, D. and Mai, H. (2018), Jump filtering and efficient drift estimation for Lévy-driven SDEs, *Ann. Statist.*, **46**(4), 1445-1480.

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

[8] Gobet, E. (2002), LAN property for ergodic diffusions with discrete observations, *Ann. I. H. Poincaré*, **38**, 711-737.

[9] 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.

[10] Höpfner, R. (2014), *Asymptotics statistics*, De Gruyter.

[11] Jakobsen, N. M and Sørensen, M. (2018), Estimating functions for jump-diffusions, *Stochastic Processes and their Applications*, to appear, DOI: `https://doi.org/10.1016/j.spa.2018.09.006`.

[12] 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.

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

[14] 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.

[15] Kawai, R. (2013), Local Asymptotic Normality Property for Ornstein-Uhlenbeck Processes with Jumps Under Discrete Sampling, *J Theor Probab*, **26**, 932-967.

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

[17] Kohatsu-Higa, A., Nualart, E. and Tran, N.K. (2014), LAN property for a simple Lévy process, *C. R. Acad. Sci. Paris, Ser. I*, **352**(10), 859-864.

[18] Kohatsu-Higa, A., Nualart, E. and Tran, N.K. (2017), LAN property for an ergodic diffusion with jumps, *Statistics: A Journal of Theoretical and Applied Statistics*, **51**(2), 419-454.

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

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

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

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

[23] Mai, H. (2014), Efficient maximum likelihood estimation for Lévy-driven Ornstein-Uhlenbeck processes, *Bernoulli*, **20**(2), 919-957.

[24] Masuda, H. (2007), Ergodicity and exponential β-mixing bounds for multidimensional diffusions with jumps, *Stochastic Processes and their Applications*, **117**, 35-56.

[25] Masuda, H. (2009), Erratum to: "Ergodicity and exponential β-mixing bounds for multidimensional diffusions with jumps", *Stochastic Processes and their Applications*, **119**, 676-678.

[26] Masuda, H. (2013), Convergence of Gaussian quasi-likelihood random fields for ergodic Lévy driven SDE observed at high frequency, *Ann. Statist.*, **41**, 1593-1641.

[27] Meyn, S.P. and Tweedie, R.L. (1993), Stability of Markovian Processes III: Foster-Lyapunov Criteria for Continuous-Time Processes, *Advances in Applied Probability*, **25**, 518-548.

[28] Nualart, D. (2006), *The Malliavin Calculus and Related Topics*, Second Edition, Springer.

[29] Ogihara, T. and Yoshida, N., (2011), Quasi-likelihood analysis for the stochastic differential equation with jumps, *Statistical Inference for Stochastic Processes*, **14**, 189-229.

[30] Petrou, E. (2008), Malliavin Calculus in Lévy spaces and Applications to Finance, *Electron. J. Probab.*, **13**, 852-879.

[31] Sato, K. (1999), *Lévy Processes and Infinitely Divisible Distributions*, Cambridge University Press, Cambridge.

[32] Shimizu, Y. (2006), *M*-Estimation for Discretely Observed Ergodic Diffusion Processes with Infinitely many Jumps, *Statistical Inference for Stochastic Processes*, **9**, 179-225.

[33] Shimizu, Y. and Yoshida, N. (2006), Estimation of Parameters for Diffusion Processes with Jumps from Discrete Observations, *Stat. Inference Stoch. Process.*, **9**(3), 227-277.

[34] Tran, N.K. (2017), LAN property for an ergodic Ornstein-Uhlenbeck process with Poisson jumps, *Communications in Statistics - Theory and Methods*, **46**(16), 7942-7968.

Ngoc Khue Tran, Department of Natural Science Education - Pham Van Dong University, 509 Phan Dinh Phung, Quang Ngai City, Quang Ngai, Vietnam

E-mail address: `tnkhueprob@gmail.com`