Université Paris 13 Nord **Institut Galilée**

Laboratoire Analyse, Géométrie et Applications **UMR 7539**

THÈSE

pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ PARIS 13

Discipline : **Mathématiques**

présentée et soutenue publiquement par

Ngoc Khue Tran

le 18 septembre 2014

Titre :

LAN property for jump-diffusion processes with discrete observations via Malliavin calculus

Directrice de thèse :

Eulalia Nualart

Co-directeur de thèse :

Arturo Kohatsu-Higa

JURY

Président :	Jean-Stéphane Dhersin	Université Paris 13-Nord
Rapporteurs :	Emmanuel Gobet	Ecole Polytechnique
	Jean Jacod	Université Pierre et Marie Curie-Paris 6
Examinateurs :	Emmanuelle Clément	Université Paris-Est Marne-la-Vallée
	Yueyun Hu	Université Paris 13-Nord
	Arturo Kohatsu-Higa	Ritsumeikan University
	Eva Löcherbach	Université de Cergy-Pontoise
	Eulalia Nualart	Universitat Pompeu Fabra and Barcelone GSE
	Anthony Réveillac	INSA de Toulouse

Mis en page avec la classe thloria.

Remerciements

En premier lieu, je tiens à remercier infiniment ma directrice de thèse, Eulalia Nualart, de m'avoir accepté comme étudiant malgré mes connaissances complètement faibles sur le sujet à cet époque-là et de m'avoir fait confiance tout au long de ces quatre années de thèse. Elle a guidé mes premiers pas dans le calcul de Malliavin ainsi que dans le monde de la recherche. Je lui suis extrêmement reconnaissant de sa gentillesse, sa patience, sa persévérance, et toute son énergie qu'elle a investi dans l'accompagnement de cette thèse. Je dois aussi la remercier de m'avoir financé plusieurs longs séjours à Barcelone là où une grande partie de cette thèse s'est réalisée.

C'est grâce à Eulalia que j'ai eu l'occasion de rencontrer et travailler avec mon co-directeur de thèse, Arturo Kohatsu-Higa. C'est avec profonde gratitude que je voudrais le remercier pour sa générosité, sa grande disponibilité et sa rigueur scientifique. J'ai grandement apprécié autant son expertise scientifique que son talent pédagogique. Une partie de cette thèse s'est déroulée dans son établissement au Japon grâce aux deux séjours financés par son projet de recherche.

Je voudrais remercier encore une fois mes deux directeurs d'avoir consacré leur temps à m'aider à la rédaction de ce texte. Leurs conseils avisés, leur soutien précieux, et leurs encouragements dans les moments difficiles et parfois déprimés m'ont permis d'aboutir à ce travail. C'est un très grand honneur pour moi d'être leur étudiant, et un privilège que je ne saurai jamais oublier.

Je tiens à remercier tout particulièrement mes rapporteurs Jean Jacod, Emmanuel Gobet et Nakahiro Yoshida pour leurs nombreuses remarques constructives et suggestions sur la version précédente de ce manuscrit, qui m'ont permis d'améliorer considérablement la qualité de ce texte. Je leur en suis très reconnaissant. Je voudrais remercier vivement Jean-Stéphane Dhersin qui me fait l'honneur de présider ce jury de thèse et Yueyun Hu, Eva Löcherbach, Emmanuelle Clément et Anthony Réveillac qui me font l'honneur de participer au jury de cette thèse.

Je n'aurais pu poursuivre mes études supérieures à Paris et cette thèse n'aurait pu se réaliser s'il n'y avait pas eu le financement du gouvernement vietnamien. Je profite de cet instant pour remercier l'ensemble du personnel du "Projet 322", CROUS de Créteil et Campus France de m'avoir toujours accompagné dans les nombreuses démarches administratives. Je tiens à remercier également Lionel Schwartz pour le financement supplémentaire via le programme LIA CNRS Formath Vietnam et le projet ARCUS MAE/IDF Vietnam, ainsi que pour un voyage au Vietnam. Je suis reconnaissant envers les membres de la Faculté des Sciences Fondamentales de l'Université Pham Van Dong pour les soutiens précieux.

Tout au long de ces quatre années, je me sens redevable au Laboratoire d'Analyse, Géométrie et Applications où j'ai beaucoup apprécié l'ambiance chaleureuse et l'excellent environnement scientifique. Mes remerciements vont à tous les membres du LAGA, plus particulièrement l'équipe de Probabilités et Statistiques au sein de laquelle j'ai été accueilli. Je tiens à dire un énorme merci à Isabelle, Yolande, Jean-Philippe et Philippe pour les démarches administratives ainsi que leur gentillesse. Je tiens à remercier l'équipe de Probabilités et Statistiques et l'ecole doctorale Institut Galilée de m'avoir fourni plusieurs opportunités formidables pour que je puisse participer à l'École d'Été de Probabilités de Saint-Flour, l'École CIMPA en Tunisie, et aux autres conférences.

Je suis également redevable envers Maria Rosa Laorden i Delàs, Yui Makino et Dai Taguchi pour m'avoir beaucoup aidé à organiser mes voyages à Barcelone et au Japon. J'ai une pensée toute particulière à Josefina Dexeus qui m'a accueilli chez elle et m'a préparé les dîners délicieux lors de mes premier et dernier séjours à Barcelone.

Je n'oublie pas de remercier tous mes amis vietnamiens en France dont la liste est très longe : Vu, The Cuong, Tran Cuong, Phong, Quang, Diep, Thi, Tien, Phuong, Nguyen, Nga, Ngà, Chien, Hoan, Duong, Vi, Mai, Dung,..., et avec qui j'ai eu le plaisir de partager de très bons moments dans la vie quotidienne pendant ces années. Je pense tout particulièrement à Quoc Hung, Anh Tuan et Van Tuan pour m'avoir beaucoup aidé dès mes premiers jours à Paris.

Enfin, je ne peux terminer cette page sans remercier du fond du cœur mes parents, mon frère et ma sœur qui m'ont constamment soutenu et encouragé tout au long de mon parcours. Je souhaite leur dédier cette thèse.

Résumé

Dans cette thèse nous appliquons le calcul de Malliavin afin d'obtenir la propriété de normalité asymptotique locale (LAN) à partir d'observations discrètes de certains processus de diffusion uniformément elliptique avec sauts. Dans le Chapitre 2 nous révisons la preuve de la propriété de normalité mixte asymptotique locale (LAMN) pour des processus de diffusion avec sauts à partir d'observations continues, et comme conséquence nous obtenons la propriété LAN en supposant l'ergodicité du processus. Dans le Chapitre 3 nous établissons la propriété LAN pour un processus de Lévy simple dont les paramètres de dérive et de diffusion ainsi que l'intensité sont inconnus. Dans le Chapitre 4, à l'aide du calcul de Malliavin et des estimées de densité de transition, nous démontrons que la propriété LAN est vérifiée pour un processus de diffusion à sauts dont le coefficient de dérive dépends d'un paramètre inconnu. Finalement, dans la même direction nous obtenons dans le Chapitre 5 la propriété LAN pour un processus de diffusion à sauts où les deux paramètres inconnus interviennent dans les coefficients de dérive et de diffusion.

Mots-clés: Calcul de Malliavin ; Estimateur asymptotiquement efficace ; Estimation paramétrique ; Normalité asymptotique locale ; Normalité mixte asymptotique locale ; Processus de Lévy ; Processus de diffusion avec sauts

Abstract

In this thesis we apply the Malliavin calculus in order to obtain the local asymptotic normality (LAN) property from discrete observations for certain uniformly elliptic diffusion processes with jumps. In Chapter 2 we review the proof of the local asymptotic mixed normality (LAMN) property for diffusion processes with jumps from continuous observations, and as a consequence, we derive the LAN property when supposing the ergodicity of the process. In Chapter 3 we establish the LAN property for a simple Lévy process whose drift and diffusion parameters as well as its intensity are unknown. In Chapter 4, using techniques of the Malliavin calculus and the estimates of the transition density, we prove that the LAN property is satisfied for a jump-diffusion process whose drift coefficient depends on an unknown parameter. Finally, in the same direction we obtain in Chapter 5 the LAN property for a jump-diffusion process where two unknown parameters determine the drift and diffusion coefficients of the jump-diffusion process.

Keywords: Malliavin calculus ; Asymptotically efficient estimator ; Parametric estimation ; Local asymptotic normality ; Local asymptotic mixed normality ; Lévy process ; Diffusion process with jumps

Table des matières

Chapitre 1

Introduction

This thesis is concerned with the application of the Malliavin calculus to the study of the local asymptotic normality (LAN) property from discrete observations for a class of uniformly elliptic diffusion processes with jumps. We will also go over the proof of the local asymptotic mixed normality (LAMN) property from continuous observations for jump-diffusion processes. The importance of such property in parametric estimation is characterized by the convolution theorem allowing us to define the asymptotically efficient estimators, and the minimax theorem giving the lower bound for the asymptotic variance of estimators.

The aim of this introduction is to recall several concepts on asymptotic statistical inference, to provide some motivation for the study of such property, to give the main results contained in the thesis, and to explain in detail the main techniques used to obtain such results.

1.1 Some basics on parametric statistical inference

Consider an $\mathbb{R}^n$-valued random vector $X^n = (X_1, \ldots, X_n)$ defined on a probability space $(\Omega, \mathcal{F}, \mathrm{P})$, whose structure will be described later, such that the probability law of X^n depends on a parameter $\theta = (\theta_1, \ldots, \theta_k) \in \Theta$, where the parameter space Θ is an open subset of $\mathbb{R}^k$. We are interested in the case where X^n corresponds to the discrete observations of a stochastic process $X^\theta = (X_t^\theta)_{t \geq 0}$ defined on $(\Omega, \mathcal{F}, \mathrm{P})$ and adapted to a filtration $(\mathcal{F}_t)_{t\geq 0}$, which satisfies a stochastic differential equation with jumps having a Brownian component. We then denote by P^θ the probability law of X^θ on the Skohorod space $(D(\mathbb{R}_+, \mathbb{R}), \mathcal{B}(\mathbb{R}_+, \mathbb{R}))$, where $D(\mathbb{R}_+, \mathbb{R})$ denotes the space of càdlàg functions from $\mathbb{R}$ to $\mathbb{R}$, and $\mathcal{B}(\mathbb{R}_+, \mathbb{R})$ its associated Borel σ-algebra, and by E^θ and E the expectation with respect to P^θ and P, respectively. Let $\overset{\mathrm{P}^\theta}{\longrightarrow}$ and $\overset{\mathcal{L}(\mathrm{P}^\theta)}{\longrightarrow}$ denote the convergence in P^θ-probability and in P^θ-law, respectively. Similarly, $\overset{\mathrm{P}}{\longrightarrow}$ and $\overset{\mathcal{L}(\mathrm{P})}{\longrightarrow}$ denote the convergence in P-probability and in P-law, respectively.

Let $\mathfrak{X}^n$ be the sample space containing all the possible values of X^n, and $\mathcal{B}(\mathfrak{X}^n)$ the Borel σ-algebra of observable events. Let $(\mathrm{P}_n^\theta)_{\theta\in\Theta}$ be the family of probability laws defined on $(\mathfrak{X}^n, \mathcal{B}(\mathfrak{X}^n))$, and induced by $X^n : \Omega \to \mathfrak{X}^n \subset \mathbb{R}^n$. The triplet $(D(\mathbb{R}_+, \mathbb{R}), \mathcal{B}(\mathfrak{X}^n), (\mathrm{P}_n^\theta)_{\theta\in\Theta})$ is called a parametric statistical model, which we denote by $(\mathrm{P}_n^\theta)_{\theta\in\Theta}$. The parametric statistical model $(D(\mathbb{R}_+, \mathbb{R}), \mathcal{B}(\mathbb{R}_+, \mathbb{R}), (\mathrm{P}^\theta)_{\theta\in\Theta})$ is defined similarly, which we denote by $(\mathrm{P}^\theta)_{\theta\in\Theta}$. We denote by E_n^θ the expectation with respect to P_n^θ.

Our objective is to estimate the parameter θ on the basis of the observations X^n. For this, let us introduce the following concepts.

A statistic is any measurable function $T : \mathfrak{X}^n \to \mathbb{R}^m$, which does not depend on θ. Moreover, any statistic $T : \mathfrak{X}^n \to \Theta$ is called an estimator of the parameter $\theta \in \Theta$.

The bias of an estimator $T(X^n)$ is defined as $b_\theta(T(X^n)) = \mathrm{E}^\theta[T(X^n)] - \theta$. An estimator $T(X^n)$ of θ is said to be unbiased if $b_\theta(T(X^n)) = 0$. The asymptotic bias of an estimator $T(X^n)$ is defined as $\lim_{n\to\infty} b_\theta(T(X^n))$. Moreover, if $\lim_{n\to\infty} b_\theta(T(X^n)) = 0$, the estimator $T(X^n)$ is said to be asymptotically unbiased.

The definition of the Fisher information matrix depends on the notion of the score function which plays a central role in parametric statistical inference. In order for this notion to be well-defined, it is necessary to impose certain conditions on the Radon-Nikodym density $p_n(x;\theta)$, $x \in \mathfrak{X}^n$ of P_n^θ with respect to a dominating measure μ_n. We shall utilize the following definition.

Definition 1.1.1. *The parametric statistical model* $(\mathrm{P}_n^\theta)_{\theta\in\Theta}$ *is regular if*

i) *There exists a σ-finite positive measure μ_n on $(\mathfrak{X}^n, \mathcal{B}(\mathfrak{X}^n))$ such that for all $\theta \in \Theta$, the probability measures P_n^θ are absolutely continuous with respect to μ_n, and the Radon-Nikodym density*

$$p_n(x;\theta) = \frac{d\mathrm{P}_n^\theta}{d\mu_n}(x)$$

is continuous on Θ for μ_n-almost all $x \in \mathfrak{X}^n$.

ii) *The function $\sqrt{p_n(x;\theta)}$ is differentiable in θ in $L^2(\mu_n)$ for all $\theta \in \Theta$.*

iii) *The $L^2(\mu_n)$-derivative of $\sqrt{p_n(x;\theta)}$ is continuous in $L^2(\mu_n)$.*

Note that the regularity of the parametric statistical model $(\mathrm{P}^\theta)_{\theta\in\Theta}$ is defined similarly.

Definition 1.1.2. *Let $(\mathrm{P}_n^\theta)_{\theta\in\Theta}$ be a regular parametric statistical model. The likelihood function and log-likelihood function based on X^n are defined as random functions of θ as follows*

$$L_n(\theta) = p_n(X^n;\theta), \qquad \text{and} \quad \ell_n(\theta) = \log p_n(X^n;\theta).$$

The score function is given by the gradient $\nabla_\theta \ell_n(\theta) = \nabla_\theta \log p_n(X^n;\theta)$.

Let us now give some consequences of this regularity condition (see Lemmas 7.1 and 7.2 of [28, Chapter I]).

Lemma 1.1.1. *Let $(\mathrm{P}_n^\theta)_{\theta\in\Theta}$ be a regular parametric statistical model.*

1. *The Fisher information matrix of the model, defined as*

$$I_n(\theta) = \mathrm{E}_n^\theta\left[\nabla_\theta \ell_n(\theta)\nabla_\theta \ell_n(\theta)^\mathsf{T}\right] = \mathrm{E}_n^\theta\left[\nabla_\theta \log p_n(X^n;\theta)\nabla_\theta \log p_n(X^n;\theta)^\mathsf{T}\right]$$

exists and is continuous on Θ.

2. *Let $T : \mathfrak{X}^n \to \mathbb{R}^m$ be a statistic such that $\mathrm{E}_n^\theta[|T(X^n)|^2]$ is bounded in a neighborhood of $\theta \in \Theta$. Then, the function $\mathrm{E}_n^\theta[T(X^n)]$ is continuously differentiable in this neighborhood, and*

$$\nabla_\theta \mathrm{E}_n^\theta\left[T(X^n)\right] = \nabla_\theta \int_{\mathfrak{X}^n} T(x)p_n(x;\theta)\mu_n(dx) = \int_{\mathfrak{X}^n} T(x)\nabla_\theta p_n(x;\theta)\mu_n(dx).$$

Taking $T(X^n) = 1$ in 2., we obtain that $\mathrm{E}_n^\theta[\nabla_\theta \ell_n(\theta)] = 0$. Therefore, $I_n(\theta) = \mathrm{Var}_n^\theta\left(\nabla_\theta \ell_n(\theta)\right)$. Furthermore, if the second order derivative $\nabla_\theta^2 \ell_n(\theta)$ exists, then $I_n(\theta) = -\mathrm{E}_n^\theta[\nabla_\theta^2 \ell_n(\theta)]$.

Let $(\mathrm{P}^\theta)_{\theta\in\Theta}$ be a regular parametric statistical model, and let ν be the measure on the space $(D(\mathbb{R}_+,\mathbb{R}), \mathcal{B}(\mathbb{R}_+,\mathbb{R}))$ from Definition 1.1.1 such that for all $\theta \in \Theta$,

$$p(x;\theta) = \frac{d\mathrm{P}^\theta}{d\nu}(x).$$

The likelihood function and log-likelihood function based on X are defined as random functions of θ as follows

$$L(\theta) = p(X;\theta), \qquad \text{and} \quad \ell(\theta) = \log p(X;\theta).$$

The score function is given by the gradient $\nabla_\theta \ell(\theta) = \nabla_\theta \log p(X;\theta)$. The Fisher information matrix of this model is defined as

$$\begin{aligned} I(\theta) = \mathrm{E}^\theta\left[\nabla_\theta \ell(\theta)\nabla_\theta \ell(\theta)^\mathsf{T}\right] &= \mathrm{E}^\theta\left[\nabla_\theta \log p(X;\theta)\nabla_\theta \log p(X;\theta)^\mathsf{T}\right] \\ &= \int_{D(\mathbb{R}_+,\mathbb{R})} \nabla_\theta \log p(x;\theta)\nabla_\theta \log p(x;\theta)^\mathsf{T} p(x;\theta)\nu(dx). \end{aligned}$$

We are now interested in using the Malliavin calculus in order to write the score function as a conditional expectation involving the Skorohod integral (see [13, Theorem 3.3] for $k = 1$). We refer to Nualart [57] for a detailed exposition of the classical Malliavin calculus on the Wiener space. We now recall briefly the Malliavin calculus for Lévy processes developed by León *et al.* in [50] and Petrou in [61], which will be applied in the thesis.

In all what follows, the observed process is defined by $X^\theta = (X_t^\theta)_{t\geq 0}$, which is driven by a Brownian motion B and a compensated Poisson random measure $\widetilde{N}$. In Chapters 3-5, to avoid confusion with the observed process X^θ, we introduce an independent copy of X^θ, denoted by Y^θ which is driven by a Brownian motion W and a compensated Poisson random measure $\widetilde{M}$, where the Malliavin calculus with respect to W will be applied. Therefore, Y^θ can be considered as the theoretical process.

Definition 1.1.3. *Let us define a Brownian motion $B = (B_t)_{t\geq 0}$ on the canonical probability space $(\Omega^1, \mathcal{F}^1, \mathrm{P}^1)$ with its natural filtration $\{\mathcal{F}_t^1\}_{t\geq 0}$, a Poisson random measure $N(dt, dz)$ on the canonical probability space $(\Omega^2, \mathcal{F}^2, \mathrm{P}^2)$ with intensity measure $\nu(dz)dt$ and its natural filtration $\{\mathcal{F}_t^2\}_{t\geq 0}$, another Brownian motion $W = (W_t)_{t\geq 0}$ on the canonical probability space $(\Omega^3, \mathcal{F}^3, \mathrm{P}^3)$ with its natural filtration $\{\mathcal{F}_t^3\}_{t\geq 0}$, and another Poisson random measure $M(dt, dz)$ on the canonical probability space $(\Omega^4, \mathcal{F}^4, \mathrm{P}^4)$ with intensity measure $\pi(dz)dt$ and its natural filtration $\{\mathcal{F}_t^4\}_{t\geq 0}$. Then, $(\Omega, \mathcal{F}, \mathrm{P})$ is defined as the canonical product probability space, where $\Omega = \Omega^1\times\Omega^2\times\Omega^3\times\Omega^4$, $\mathcal{F} = \mathcal{F}^1\otimes\mathcal{F}^2\otimes\mathcal{F}^3\otimes\mathcal{F}^4$, $\mathrm{P} = \mathrm{P}^1\otimes\mathrm{P}^2\otimes\mathrm{P}^3\otimes\mathrm{P}^4$, and $\mathcal{F}_t = \mathcal{F}_t^1\otimes\mathcal{F}_t^2\otimes\mathcal{F}_t^3\otimes\mathcal{F}_t^4$. Therefore on this space the canonical process represents (B, N, W, M) which are therefore mutually independent.*

We denote by $\widehat{\Omega} = \Omega^1 \times \Omega^2$, $\widehat{\mathcal{F}} = \mathcal{F}^1 \otimes \mathcal{F}^2$, $\widehat{\mathrm{P}} = \mathrm{P}^1 \otimes \mathrm{P}^2$, $\widehat{\mathcal{F}}_t = \mathcal{F}_t^1 \otimes \mathcal{F}_t^2$, and $\widetilde{\Omega} = \Omega^3 \times \Omega^4$, $\widetilde{\mathcal{F}} = \mathcal{F}^3\otimes\mathcal{F}^4$, $\widetilde{\mathrm{P}} = \mathrm{P}^3\otimes\mathrm{P}^4$, $\widetilde{\mathcal{F}}_t = \mathcal{F}_t^3\otimes\mathcal{F}_t^4$. We denote by E, $\widehat{\mathrm{E}}$ and $\widetilde{\mathrm{E}}$ the expectation with respect to P, $\widehat{\mathrm{P}}$ and $\widetilde{\mathrm{P}}$, respectively. Observe that $\Omega = \widehat{\Omega} \times \widetilde{\Omega}$, $\mathcal{F} = \widehat{\mathcal{F}} \otimes \widetilde{\mathcal{F}}$, $\mathrm{P} = \widehat{\mathrm{P}} \otimes \widetilde{\mathrm{P}}$, $\mathcal{F}_t = \widehat{\mathcal{F}}_t \otimes \widetilde{\mathcal{F}}_t$, and $\mathrm{E} = \widehat{\mathrm{E}} \otimes \widetilde{\mathrm{E}}$.

On the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathrm{P})$, consider a two-dimensional centered square integrable Lévy process $Z = (Z^1, Z^2) = (Z_t)_{t\in[0,T]}$ given by

$$
\begin{aligned}
Z_t^1 &= \sigma_1 B_t + \int_0^t \int_{\mathbb{R}_0} z\left(N(dt, dz) - \nu(dz)dt\right), \\
Z_t^2 &= \sigma_2 W_t + \int_0^t \int_{\mathbb{R}_0} z\left(M(dt, dz) - \pi(dz)dt\right),
\end{aligned}
$$

where $\sigma_1, \sigma_2 > 0$ are constant, $\mathbb{R}_0 := \mathbb{R} \setminus \{0\}$. The compensated Poisson random measures are denoted by $\widetilde{N}(dt, dz) := N(dt, dz) - \nu(dz)dt$, and $\widetilde{M}(dt, dz) := M(dt, dz) - \pi(dz)dt$. Here, the intensity measures satisfy that $\int_{\mathbb{R}_0}(1 \wedge |z|^2)\nu(dz) < \infty$ and $\int_{\mathbb{R}_0}(1 \wedge |z|^2)\pi(dz) < \infty$. Remark that the filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$ is the same as the one generated by the Lévy process Z. The main idea of León *et al.* in [50] is to represent random variables on $(\Omega, \mathcal{F}, \mathrm{P})$ via iterated integrals, from which a Mallivin calculus can be defined as in the Gaussian setting in Nualart [57].

To simplify the exposition, we introduce the following unified notation for the Brownian motions and the Poisson random measures

$$
\begin{aligned}
U_1 = U_2 = [0, T], &\quad \text{and} \quad U_3 = U_4 = [0, T] \times \mathbb{R}_0, \\
dQ_1(\cdot) = dB_\cdot, dQ_2(\cdot) = dW_\cdot, &\quad \text{and} \quad Q_3(\cdot, *) = \widetilde{N}(\cdot, *), Q_4(\cdot, *) = \widetilde{M}(\cdot, *), \\
d\langle Q_1\rangle_\cdot = d\langle Q_2\rangle_\cdot = d\cdot, &\quad \text{and} \quad d\langle Q_3\rangle_\cdot = d\cdot \times d\nu(*), d\langle Q_4\rangle_\cdot = d\cdot \times d\pi(*),
\end{aligned}
$$

and for the variables $t_k \in [0, T]$ and $z \in \mathbb{R}_0$,

$$
u_k^\ell := \begin{cases} t_k & , \ell = 1, 2, \\ (t_k, z) & , \ell = 3, 4. \end{cases}
$$

Set $S_n = \{1,2,3,4\}^n$. For $(j_1,\dots,j_n) \in S_n$, define an expanded simplex of the form

$$G_{j_1,\dots,j_n} = \left\{ (u_1^{j_1},\dots,u_n^{j_n}) \in \Pi_{i=1}^n \, U_{j_i} : 0 < t_1 < \cdots < t_n < T \right\}.$$

We next define the iterated integral of the form

$$J_n^{(j_1,\dots,j_n)}(g_{j_1,\dots,j_n}) = \int_{G_{j_1,\dots,j_n}} g_{j_1,\dots,j_n}(u_1^{j_1},\dots,u_n^{j_n}) Q_{j_1}(du_1^{j_1}) \cdots Q_{j_n}(du_n^{j_n}),$$

where $g_{j_1,\dots,j_n}$ is a deterministic function in $L^2(G_{j_1,\dots,j_n}) = L^2(G_{j_1,\dots,j_n}, \otimes_{i=1}^n d\langle Q_{j_i}\rangle)$.

Theorem 1.1.1. *For every random variable $F \in L^2(\Omega, \mathcal{F}_T, \mathrm{P})$, there exists a unique sequence of deterministic functions $\{g_{j_1,\dots,j_n}\}_{n=0}^\infty$, $(j_1,\dots,j_n) \in S_n$, where $g_{j_1,\dots,j_n} \in L^2(G_{j_1,\dots,j_n})$ such that*

$$F = \mathrm{E}[F] + \sum_{n=1}^\infty \sum_{(j_1,\dots,j_n)\in S_n} J_n^{(j_1,\dots,j_n)}(g_{j_1,\dots,j_n}),$$

and we have the isometry

$$\|F\|^2_{L^2(\mathrm{P})} = \mathrm{E}F^2 + \sum_{n=1}^\infty \sum_{(j_1,\dots,j_n)\in S_n} \|J_n^{(j_1,\dots,j_n)}(g_{j_1,\dots,j_n})\|^2_{L^2(G_{j_1,\dots,j_n})}.$$

Using this chaotic representation property, the directional derivatives can be defined with respect to Brownian motion and Poisson random measure. For this, denote

$$G^k_{j_1,\dots,j_n}(t) = \Big\{ (u_1^{j_1},\dots,u_{k-1}^{j_{k-1}}, \hat{u}_k^{j_k}, u_{k+1}^{j_{k+1}},\dots,u_n^{j_n}) \in G_{j_1,\dots,j_{k-1},j_{k+1},\dots,j_n} : \\ 0 < t_1 < \cdots < t_{k-1} < t < t_{k+1} < \cdots < t_n < T \Big\},$$

where $\hat{u}$ means that we omit the u element.

Definition 1.1.4. *Let $g_{j_1,\dots,j_n} \in L^2(G_{j_1,\dots,j_n})$ and $\ell \in \{1,2,3,4\}$. Then*

$$D^{(\ell)}_{u^\ell} J_n^{(j_1,\dots,j_n)}(g_{j_1,\dots,j_n}) = \sum_{i=1}^n \mathbf{1}_{\{j_i=\ell\}} J_{n-1}^{(j_1,\dots,j_{i-1},\hat{j}_i,j_{i+1},\dots,j_n)} \left(g_{j_1,\dots,j_n}(\dots,u^\ell,\dots) \mathbf{1}_{G^i_{j_1,\dots,j_n}(t)} \right)$$

is called the derivative of $J_n^{(j_1,\dots,j_n)}(g_{j_1,\dots,j_n})$ in the ℓ-th direction.

Definition 1.1.5. *Let $\mathbb{D}^{(\ell)}$ be the space of the random variables in $L^2(\Omega)$ that are differentiable in the ℓ-th direction, then*

$$\mathbb{D}^{(\ell)} = \Big\{ F \in L^2(\Omega), F = \mathrm{E}[F] + \sum_{n=1}^\infty \sum_{(j_1,\dots,j_n)\in S_n} J_n^{(j_1,\dots,j_n)}(g_{j_1,\dots,j_n}) : \\ \sum_{n=1}^\infty \sum_{(j_1,\dots,j_n)\in S_n} \sum_{i=1}^n \mathbf{1}_{\{j_i=\ell\}} \int_{U_{j_i}} \left\| g_{j_1,\dots,j_n}(\dots,u^\ell,\dots) \right\|^2_{L^2(G^i_{j_1,\dots,j_n})} d\langle Q_\ell\rangle (u^\ell) < \infty \Big\}.$$

Definition 1.1.6. *Let $F \in \mathbb{D}^{(\ell)}$. Then the derivative in the ℓ-th direction is defined as*

$$D^{(\ell)}_{u^\ell} F = \sum_{n=1}^\infty \sum_{(j_1,\dots,j_n)\in S_n} \sum_{i=1}^n \mathbf{1}_{\{j_i=\ell\}} J_{n-1}^{(j_1,\dots,\hat{j}_i,\dots,j_n)} \left(g_{j_1,\dots,j_n}(\dots,u^\ell,\dots) \mathbf{1}_{G^i_{j_1,\dots,j_n}(t)} \right).$$

On the space $\mathbb{D}^{(1)} \cap \mathbb{D}^{(2)}$, the directional derivatives $(D^{(1)}, D^{(2)})$ with respect to the 2-dimensional Brownian motion (B, W) are equivalent to the classical Malliavin derivative on the Wiener space.

Moreover, the properties like the chain rule, the integration by parts formula, the Skorohod integral and duality relation of the directional derivatives $(D^{(1)}, D^{(2)})$ are preserved as for the classical Malliavin calculus.

Notice that the solution Y^θ to a stochastic differential equation with jumps driven by the Brownian motion W and the compensated Poisson random measure $\widetilde{M}$ can be differentiable with respect to W and $\widetilde{M}$ (see [61]). In Chapters 3-5, we are only concerned with the Malliavin derivative in the 2-th direction, i.e., with respect to W, and let us denote $D \equiv D^{(2)}$, $\mathbb{D}^{1,2} \equiv \mathbb{D}^{(2)}$ and $\mathcal{H} = L^2([0,T], \mathbb{R})$.

Definition 1.1.7. *The divergence operator δ (called the Skorohod integral) is the adjoint of the directional derivative D. That is, δ is an unbounded operator from $L^2([0,T] \times \Omega, \mathbb{R})$ to $L^2(\Omega)$ such that*

(i) *The domain of δ, Dom δ, is the set of random variables $u \in L^2([0,T] \times \Omega, \mathbb{R})$ such that*

$$\left|\mathrm{E}\left[\langle DF, u\rangle_{\mathcal{H}}\right]\right| \leq c_u \left\|F\right\|_{L^2(\Omega)},$$

for all $F \in \mathbb{D}^{1,2}$, where c_u is some positive constant depending on u.

(ii) *If $u \in$ Dom δ then $\delta(u)$ is the element of $L^2(\Omega)$ characterized by the following duality relation*

$$\mathrm{E}\left[F\delta(u)\right] = \mathrm{E}\left[\langle DF, u\rangle_{\mathcal{H}}\right],$$

for any $F \in \mathbb{D}^{1,2}$.

Similarly, X^θ can be differentiable with respect to B and $\widetilde{N}$. Let Dom $\delta^{(1)}$ denote the domain of the Skorohod integral $\delta^{(1)}$, the adjoint operator of the Malliavin derivative $D^{(1)}$ with respect to B, in $L^2([0,T] \times \Omega, \mathbb{R})$.

Proposition 1.1.1. *Let $(\mathrm{P}_n^\theta)_{\theta \in \Theta}$ be a regular parametric statistical model. Assume that the random variables $X_i \in \mathbb{D}^{(1)}$, for all $i \in \{1, \ldots, n\}$, and let $U = (U^1, \ldots, U^k)$ be a k-dimensional stochastic process satisfying that $U^j \in$ Dom $\delta^{(1)}$, for all $j \in \{1, \ldots, k\}$ such that for all $i \in \{1, \ldots, n\}$,*

$$\left\langle D^{(1)} X_i, U \right\rangle_{\mathcal{H}} = \nabla_\theta X_i, \tag{1.1}$$

where $\nabla_\theta X_i = (\partial_{\theta_1} X_i, \ldots, \partial_{\theta_k} X_i)$, and

$$\begin{aligned}\left\langle D^{(1)} X_i, U \right\rangle_{\mathcal{H}} &= \left(\left\langle D^{(1)} X_i, U^1 \right\rangle_{\mathcal{H}}, \ldots, \left\langle D^{(1)} X_i, U^k \right\rangle_{\mathcal{H}}\right) \\ &= \left(\int_0^T D_t^{(1)} X_i U^1(t) dt, \ldots, \int_0^T D_t^{(1)} X_i U^k(t) dt\right).\end{aligned}$$

Moreover, assume that for any $\theta \in \Theta$, there is a neighborhood of θ where $|\nabla_\theta X_i| \leq G$ with $\mathrm{E}[G] < \infty$. Then for all $\theta \in \Theta$,

$$\nabla_\theta \log p_n(x; \theta) = \mathrm{E}\left[\delta^{(1)}(U) | X^n = x\right],$$

for almost all $x \in \mathfrak{X}^n$ and for all $\theta \in \Theta$, where we denote $\delta^{(1)}(U) := (\delta^{(1)}(U^1), \ldots, \delta^{(1)}(U^k))$.

Proof. Let φ be a $\mathcal{C}_b^\infty(\mathbb{R}^n)$ function with compact support. Then, hypothesis (1.1), the chain rule and the duality relation of the Malliavin calculus imply that

$$\begin{aligned}\nabla_\theta \mathrm{E}\left[\varphi(X^n)\right] = \sum_{i=1}^n \mathrm{E}\left[\partial_{x_i}\varphi(X^n)\nabla_\theta X_i\right] &= \sum_{i=1}^n \mathrm{E}\left[\partial_{x_i}\varphi(X^n)\left\langle D^{(1)} X_i, U\right\rangle_{\mathcal{H}}\right] \\ &= \mathrm{E}\left[\left\langle D^{(1)}(\varphi(X^n)), U\right\rangle_{\mathcal{H}}\right] \\ &= \mathrm{E}\left[\varphi(X^n)\delta^{(1)}(U)\right].\end{aligned}$$

On the other hand, by Lemma 1.1.1,

$$\nabla_\theta \mathrm{E}\left[\varphi(X^n)\right] = \int_{\mathfrak{X}^n} \varphi(x)\nabla_\theta p_n(x;\theta)\mu(dx) = \mathrm{E}\left[\varphi(X^n)\nabla_\theta \log p_n(X^n;\theta)\right].$$

Thus, the result follows. □

As a consequence, the Fisher information matrix and the Cramér-Rao lower bound can be obtained without requiring the explicit expression of the density $p_n(\cdot;\theta)$ (see Proposition 1.1.2 below). Observe that by Proposition 1.1.1,

$$I_n(\theta) = \mathrm{Var}\left(\mathrm{E}\left[\delta^{(1)}(U)|X^n\right]\right).$$

Let us now give the classical Cramér-Rao's inequality (see Theorem 7.3 of [28, Chapter I]).

Proposition 1.1.2. *Let $(\mathrm{P}_n^\theta)_{\theta\in\Theta}$ be a regular parametric statistical model, and T be a statistic such that $\mathrm{E}_n^\theta[|T(X^n)|^2]$ is bounded in a neighborhood of $\theta \in \Theta$. Assume that $I_n(\theta)$ is invertible for all $\theta \in \Theta$. Let $g(\theta) = \mathrm{E}_n^\theta\left[T(X^n)\right]$. Then g is continuously differentiable in this neighborhood, and*

$$\mathrm{Var}_n^\theta\left(T(X^n)\right) \geq \nabla_\theta g(\theta) I_n(\theta)^{-1} \nabla_\theta g(\theta)^\mathsf{T}.$$

In this case, $\nabla_\theta g(\theta) I_n(\theta)^{-1} \nabla_\theta g(\theta)^\mathsf{T}$ is called the Cramér-Rao lower bound for estimating $g(\theta)$.

In particular, if $T(X^n)$ is an unbiased estimator of θ, then

$$\mathrm{Var}_n^\theta\left(T(X^n)\right) \geq I_n(\theta)^{-1}.$$

In this case, $I_n(\theta)^{-1}$ is called the Cramér-Rao lower bound for estimating θ.

We remark that if $(\mathrm{P}^\theta)_{\theta\in\Theta}$ is a regular parametric statistical model, then the Cramér-Rao lower bound for an unbiased estimator $T(X^\theta)$ of θ holds with $I(\theta)^{-1}$.

The Cramér-Rao lower bound suggests the following definition.

Definition 1.1.8. *Suppose that $(\mathrm{P}_n^\theta)_{\theta\in\Theta}$ is a regular parametric statistical model. An unbiased estimator $T(X^n)$ of θ is called efficient if its covariance matrix achieves the Cramér-Rao lower bound. That is,*

$$\mathrm{Var}_n^\theta(T(X^n)) = I_n(\theta)^{-1}.$$

Moreover, suppose that the parametric statistical model $(\mathrm{P}^\theta)_{\theta\in\Theta}$ is regular. An estimator $T(X^n)$ of θ is called asymptotically efficient in the Cramér-Rao sense if it is asymptotically normal, and its covariance matrix achieves asymptotically the Cramér-Rao lower bound. That is, there exists a $k \times k$ non-random diagonal matrix $\varphi_n(\theta)$ whose entries are strictly positive and tend to zero as $n \to \infty$ such that as $n \to \infty$,

$$\varphi_n^{-1}(\theta)\left(T(X^n) - \theta\right) \xrightarrow{\mathcal{L}(\mathrm{P}^\theta)} \mathcal{N}\left(0, I(\theta)^{-1}\right), \tag{1.2}$$

where $\mathcal{N}(0, I(\theta)^{-1})$ denotes a centered $\mathbb{R}^k$-valued Gaussian random variable with covariance matrix $I(\theta)^{-1}$, and $\varphi_n^{-1}(\theta)$ is the rate of convergence of $I(\theta)^{-1}$. Here, $I(\theta)$ is the Fisher information matrix of the model $(\mathrm{P}^\theta)_{\theta\in\Theta}$.

Example 1.1.1. *Consider the following diffusion process $X^\theta = (X_t^\theta)_{t\in[0,1]}$*

$$dX_t^\theta = b(\theta, t)dt + dB_t,$$

where $X_0^\theta = x_0$, $\theta \in \Theta \subset \mathbb{R}$, and $b(\theta, \cdot)$ is a continuously differentiable function on $L^2(0,1)$. Let $(\mathrm{P}^\theta)_{\theta\in\Theta}$ be the law of the continuous observation X^θ on the canonical space $(C[0,1], \mathcal{B}[0,1])$, where $C[0,1]$ denotes the space of continuous functions from $[0,1]$ to $\mathbb{R}$, and $\mathcal{B}[0,1]$ its associated

Borel σ-algebra. Let ν be the probability law of $(x_0 + B_t)_{t\in[0,1]}$ on $(C[0,1], \mathcal{B}[0,1])$. Then by Girsanov's Theorem,

$$\frac{d\mathrm{P}^\theta}{d\nu} = \exp\left\{\int_0^1 b(\theta,t)dX_t^\theta - \frac{1}{2}\int_0^1 b^2(\theta,t)dt\right\} = p(X^\theta,\theta).$$

It can be checked that the parametric statistical model $(\mathrm{P}^\theta)_{\theta\in\Theta}$ is regular (see Example 7.3 of [28, Chapter I]). Moreover, the Fisher information matrix is given by

$$I(\theta) = \int_0^1 \left(\partial_\theta b(\theta,t)\right)^2 dt.$$

An estimator $T(X^n)$ of the parameter θ is said to be consistent if it converges in P^θ-probability to θ as $n\to\infty$. When comparing two consistent estimators of the parameter θ, it is natural to compare their rates of convergence and the asymptotic variances of their respective asymptotic distributions, which are in general the normal distribution or mixed normal distribution.

Definition 1.1.9. *A sequence of estimators $(T(X^n))_{n\geq 1}$ of the parameter θ is called asymptotically mixed normal if for any $\theta\in\Theta$, there exists a $k\times k$ non-random diagonal matrix $\varphi_n(\theta)$ whose entries are strictly positive and tend to zero as $n\to\infty$, and a $k\times k$ positive definite random matrix $\Gamma(\theta)$, such that as $n\to\infty$,*

$$\varphi_n^{-1}(\theta)\left(T(X^n)-\theta\right) \overset{\mathcal{L}(\mathrm{P}^\theta)}{\longrightarrow} \Gamma(\theta)^{-1/2}\mathcal{N}(0,I_k),$$

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

When the matrix $\Gamma(\theta)$ is deterministic, the sequence $(T(X^n))_{n\geq 1}$ is asymptotically normal.

When the estimators are asymptotically mixed normal, another important issue is whether these estimators are asymptotically efficient in the sense that they achieve a minimal asymptotic variance. We have given in Definition 1.1.8 the notion of asymptotic efficiency of the estimators in terms of deterministic Cramér-Rao lower bound. Another approach to define the asymptotic efficiency of the estimators is to study the lower bound for asymptotic variances of the estimators via a convolution theorem. This problem is related to a fundamental concept in asymptotic theory of statistics called the local asymptotic normality (LAN) property, which was introduced by Le Cam [48], developed by Hájek [26, 27] and extended by Jeganathan [34], [35] to the local asymptotic mixed normality (LAMN) property. We refer to the monographs by Ibragimov and Has'minskii [28], Kutoyants [47], Le Cam and Lo Yang [49], Van Der Vaart [73] for more detailed expositions of this topic.

Note that solving the issue on the regularity of a parametric statistical model and on the asymptotic mixed normality of the estimators is an interesting topic. However, the purpose of this thesis is to focus on addressing the problem of asymptotic efficiency of the estimators in the latter sense, and more precisely, studying the LAN property for a class of diffusion processes with jumps.

Definition 1.1.10. *The score function is said to be asymptotically mixed normal if for any $\theta\in\Theta$, there exists a $k\times k$ non-random diagonal matrix $\varphi_n(\theta)$ whose entries are strictly positive and tend to zero as $n\to\infty$, and a $k\times k$ positive definite random matrix $\Gamma(\theta)$, such that as $n\to\infty$,*

$$\varphi_n(\theta)\nabla_\theta\ell_n(\theta) \overset{\mathcal{L}(\mathrm{P}^\theta)}{\longrightarrow} \Gamma(\theta)^{1/2}\mathcal{N}\left(0,I_k\right), \tag{1.3}$$

where $\Gamma(\theta)$ and $\mathcal{N}(0,I_k)$ are independent. In this case, the matrix $\Gamma(\theta)$ is called the asymptotic Fisher information matrix of the model. When the matrix $\Gamma(\theta)$ is deterministic, we say that the score function is asymptotically normal.

Definition 1.1.11. *The sequence* $(\mathrm{P}_n^\theta)_{\theta\in\Theta}$ *is said to have the LAMN property if for any* $\theta \in \Theta$ *and* $u \in \mathbb{R}^k$, *as* $n \to \infty$,

$$\log \frac{d\mathrm{P}_n^{\theta+\varphi_n(\theta)u}}{d\mathrm{P}_n^\theta}(X^n) \overset{\mathcal{L}(\mathrm{P}^\theta)}{\longrightarrow} u^\top \Gamma(\theta)^{1/2}\mathcal{N}(0, I_k) - \frac{1}{2}u^\top\Gamma(\theta)u, \tag{1.4}$$

where $\mathcal{N}(0, I_k)$, $\varphi_n^{-1}(\theta)$, *and* $\Gamma(\theta)$ *are as in (1.3). In this case, we say that the LAMN property holds with rate of convergence* $\varphi_n^{-1}(\theta)$ *and asymptotic Fisher information matrix* $\Gamma(\theta)$. *When the matrix* $\Gamma(\theta)$ *is deterministic, the LAN property holds.*

Observe that (1.4) is equivalent to

$$\begin{aligned}\log \frac{d\mathrm{P}_n^{\theta+\varphi_n(\theta)u}}{d\mathrm{P}_n^\theta}(X^n) &= \ell_n(\theta+\varphi_n(\theta)u) - \ell_n(\theta)\\ &= u^\top\varphi_n(\theta)\nabla_\theta\ell_n(\theta) - \frac{1}{2}u^\top\Gamma(\theta)u + o_{\mathrm{P}^\theta}(1),\end{aligned} \tag{1.5}$$

where $\varphi_n(\theta)\nabla_\theta\ell_n(\theta)$ converges in P^θ-law to $\Gamma(\theta)^{1/2}\mathcal{N}(0, I_k)$ as $n \to \infty$.

Two fundamental consequences of the LAMN property are the conditional convolution theorem and the minimax theorem.

Definition 1.1.12. *A sequence of estimators* $(T(X^n))_{n\geq 1}$ *of the parameter* θ *is called regular at* θ *if for any* $u \in \mathbb{R}^k$, *as* $n \to \infty$,

$$\varphi_n^{-1}(\theta)\left(T(X^n) - (\theta + \varphi_n(\theta)u)\right) \overset{\mathcal{L}(\mathrm{P}^{\theta+\varphi_n(\theta)u})}{\longrightarrow} V(\theta),$$

for some $\mathbb{R}^k$*-valued random variable* $V(\theta)$, *independent of* u.

Note that taking $u = 0$, this implies that as $n \to \infty$,

$$\varphi_n^{-1}(\theta)\left(T(X^n) - \theta\right) \overset{\mathcal{L}(\mathrm{P}^\theta)}{\longrightarrow} V(\theta).$$

The conditional convolution theorem says that when the LAMN property holds, then the asymptotic distribution of any regular sequence of estimators of the parameter θ is characterized by a conditional convolution between a Gaussian law and some others laws. More precisely,

Theorem 1.1.2 (Conditional convolution theorem)**.** [34, Corollary 1] *Suppose that the sequence* $(\mathrm{P}_n^\theta)_{\theta\in\Theta}$ *satisfies the LAMN property at a point* θ. *Let* $(T(X^n))_{n\geq 1}$ *be a regular sequence of estimators of the parameter* θ. *Then the law of* $V(\theta)$ *conditionally on* $\Gamma(\theta)$ *is a convolution between the Gaussian law* $\mathcal{N}(0, \Gamma(\theta)^{-1})$ *and some other law* $G_{\Gamma(\theta)}$ *on* $\mathbb{R}^k$, *that is,*

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

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

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

where R is a random variable with distribution $G_{\Gamma(\theta)}$, independent of $\mathcal{N}(0, I_k)$. This implies that, under the conditions of Theorem 1.1.2, as $n \to \infty$,

$$\varphi_n^{-1}(\theta)\left(T(X^n) - \theta\right) \overset{\mathcal{L}(\mathrm{P}^\theta)}{\longrightarrow} \Gamma(\theta)^{-1/2}\mathcal{N}(0, I_k) + R.$$

This theorem suggests the notion of asymptotically efficient estimators in terms of minimal asymptotic variance, when $R = 0$. That is,

Definition 1.1.13. *Assume that the sequence $(\mathrm{P}_n^\theta)_{\theta\in\Theta}$ satisfies the LAMN property at a point θ. A sequence of estimators $(T(X^n))_{n\geq 1}$ of the parameter θ is called asymptotically efficient at θ in the sense of Hájek-Le Cam convolution theorem if as $n\to\infty$,*

$$\varphi_n^{-1}(\theta)\left(T(X^n)-\theta\right) \overset{\mathcal{L}(\mathrm{P}^\theta)}{\longrightarrow} \Gamma(\theta)^{-1/2}\mathcal{N}(0,I_k),$$

where $\Gamma(\theta)$ and $\mathcal{N}(0,I_k)$ are independent.

In particular, when $\Gamma(\theta)$ is deterministic, a sequence of estimators which is asymptotically efficient in the sense of Hájek-Le Cam convolution theorem achieves asymptotically the Cramér-Rao lower bound in (1.2) for the estimation variance, that is, $\Gamma(\theta)=I(\theta)$.

Moreover, as a consequence of the LAMN property, an asymptotic lower bound for risk functions of estimators can be obtained via minimax theorem. More precisely,

Theorem 1.1.3 (Minimax theorem)**.** [34, Proposition 2] *Suppose that the sequence $(\mathrm{P}_n^\theta)_{\theta\in\Theta}$ satisfies the LAMN property at a point $\theta\in\Theta$. Let $(T(X^n))_{n\geq 1}$ be a sequence of estimators of the parameter θ and $l:\mathbb{R}^k\to[0,+\infty)$ be a loss function of the form $l(0)=0, l(x)=l(|x|)$ and $l(|x|)\leq l(|y|)$ if $|x|\leq|y|$. Then*

$$\liminf_{n\to\infty}\mathrm{E}^\theta\left[l\left(\varphi_n^{-1}(\theta)\left(T(X^n)-\theta\right)\right)\right]\geq \mathrm{E}\left[l\left(\Gamma(\theta)^{-1/2}\mathcal{N}(0,I_k)\right)\right].$$

In particular, when we take the quadratic loss function $l(u)=|u|^2$, the above inequality gives an asymptotic lower bound for the covariance matrix of any sequence of unbiased estimators, which is given by $\Gamma(\theta)^{-1}$.

As indicated above, we are concerned with a discrete observation $X^n=(X_0,X_{\Delta_n},\ldots,X_{n\Delta_n})$ at equidistant times $t_k=k\Delta_n$, $k\in\{0,\ldots,n\}$ of a stochastic process $X^\theta=(X_t^\theta)_{t\geq 0}$ defined on $(\Omega,\mathcal{F},\mathrm{P})$. Here n is the observation frequency, and Δ_n is the corresponding time step size.

When the time step size is $\Delta_n=\Delta$, where Δ is a positive constant independent of n, the scheme of observation is called low frequency observation. When $\Delta_n\to 0$ as $n\to\infty$, it is called high frequency observation.

On the other hand, the case of high frequency observation can be divided into two cases depending on $t_n=n\Delta_n$. That is, when $n\Delta_n$ is finite and fixed, we have a discrete observation X^n on a finite fixed interval. When $n\Delta_n\to\infty$ as $n\to\infty$, we have a discrete observation X^n on an increasing interval.

The parametric estimation for discrete observations at high frequency has been developed by, for instance, Florens-Zmirou and Dacunha-Castelle [15], Florens-Zmirou [19], Donhal [17], Yoshida [76], Genon-Catalot and Jacod [21, 22], Kessler [40], Gobet [24, 25] for continuous diffusion processes, and by Aït-Sahalia and Jacod [1, 2], Shimizu and Yoshida [69], Shimizu [67], Ogihara and Yoshida [59], Masuda [55], Kawai [38], Kawai and Masuda [39], Clément, Delattre and Gloter [10, 11] for jump-diffusion processes.

In the other direction, the statistical inference for stochastic processes with continuous-time observations has been widely developed during the last forty years. Several important contribution on this subject can be found in the books and articles of Basawa and Prakasa Rao [6] for stochastic processes, Basawa and Scott [7] for non-ergodic models, Kutoyants [44, 46, 47] for diffusion processes, Sørensen [71] for diffusions with jumps, Barndorff-Nielsen and Sørensen [5] for stochastic processes, and Prakasa Rao [63] for semimartingales.

1.2 Motivation and model setting

In practice, the observations are rather discrete than continuous. The case of discrete-time observations is an interesting subject which has been extensively studied in recent years. However, most of obtained results in the literature are related to the continuous diffusion processes. For the Ornstein-Uhlenbeck process, which possesses an explicit Gaussian law, it has been shown that the LAN and LAMN properties hold true in the ergodic and non-ergodic case, respectively (see [13], [30] and [68]). The LAMN property for one-dimensional diffusion processes was studied by Donhal in [17], where the proof is derived by expanding the transition density with respect to the time and the parameters up to an appropriate order. Later, Genon-Catalot and Jacod [22] showed that the LAMN property can be obtained for a class of diffusion processes by assuming some specific estimates on the transition densities and their derivatives.

Recently, techniques of Malliavin calculus have proved to be a powerful tool for the stochastic analysis of the log-likelihood ratio, which, to our knowledge, was initiated by Gobet in [24]. Concretely, he obtained in this paper the LAMN property from discrete observations at high frequency on the interval $[0,1]$ for multidimensional elliptic diffusion $(X_t^\theta)_{t\in[0,1]}$ defined by

$$X_t^\theta = x + \int_0^t b(\theta, s, X_s^\theta)ds + \int_0^t S(\theta, s, X_s^\theta)dB_s,$$

which generalizes the preceding result obtained by Donhal [17]. For this purpose, the integration by parts formula of the Malliavin calculus on the Gaussian space is applied in order to derive an expansion of the log-likelihood ratio in terms of a sum of conditional expectations involving Skorohod integrals. On the other hand, the upper and lower Gaussian type bounds of the transition density are essentially employed in the analysis of the convergence of the sum of conditional expectations appearing in this expansion. Following the same approach as in [24], the LAN property was next established in [25] from discrete observations at high frequency on an increasing interval for multidimensional ergodic diffusions $(X_t^{\alpha,\beta})_{t\geq 0}$ defined by

$$X_t^{\alpha,\beta} = x_0 + \int_0^t b(\alpha, s, X_s^{\alpha,\beta})ds + \int_0^t S(\beta, s, X_s^{\alpha,\beta})dB_s.$$

Later on, in the same direction Gobet and Gloter [23] showed that the LAMN property is satisfied for integrated diffusions.

In the presence of the jump component, several special cases have been studied. Precisely, the LAN property is established for some Lévy processes whose transition density can be expressed in an explict form, for instance, stable processes and normal inverse Gaussian Lévy processes (see [75, 39]). In addition, Aït-Sahalia and Jacod in [1] established the LAN property for a class of Lévy processes. Recently, Kawai in [38] deals with the particular case of the ergodic Ornstein-Uhlenbeck (O-U) process with jumps whose solution and its respective transition density can be written in semi-explicit form. This implies that a Taylor expansion of the log-density with respect to the parameters can be obtained, therefore reducing the proof to the proof of a classical central limit theorem with independent increments and a residual term. This residual term depends strongly on estimates of the first and second derivatives of the logarithm of the density of the O-U process which are dealt with using the integration by parts formula of Malliavin calculus based on the Brownian motion. However, [38] studies only the case where the unknown parameters determine the drift and diffusion coefficients, but where the jump component does not depend on the parameter.

More recently, using tools of Malliavin calculus as in [24], Clément et *al.* [10] have established the LAMN property for a stochastic process with jumps $(X_t^\lambda)_{t\in[0,1]}$ driven by a compound Poisson process

$$X_t^\lambda = x_0 + \int_0^t b(s, X_s^\lambda)ds + \int_0^t a(s, X_s^\lambda)dW_s + \sum_{k=1}^K c(X_{T_k-}^\lambda, \lambda_k)\mathbf{1}_{t\geq T_k},$$

where the parameter $\lambda = (\lambda_1, \ldots, \lambda_K) \in \mathbb{R}^K$ determines the jump amplitudes, the jump times are given by $0 < T_1 < \cdots < T_K < 1$ and the number of jumps K on $[0,1]$ is deterministic.

Moreover, using the Malliavin calculus for jump processes developed by Bichteler, Gravereaux and Jacod [9], Clément and Gloter [11] prove that the LAMN property holds true for the process solution $(X_t^\theta)_{t\in[0,1]}$ defined by

$$X_t^\theta = x_0 + \int_0^t b(X_s^\theta, \theta)ds + L_t,$$

where $(L_t)_{t\in[0,1]}$ is a pure jump Lévy process whose Lévy measure an α-stable Lévy measure near zero with $\alpha \in (1,2)$.

However, it can be seen that the validity of the LAN property for general stochastic differential equations with jumps having a Brownian component has never been addressed in the literature. One of the reasons could be that the behaviour of the transition density changes strongly due to the presence of jumps in this context. In fact, one expects that the upper bound for the density of such stochastic differential equations with jumps will be controlled by the exponential behaviour of the jump process and that the lower bound will be controlled by the Gaussian behaviour of the Wiener process. For instance, consider a one-dimensional Lévy process $(X_t^x)_{t\geq 0}$ starting from $x \in \mathbb{R}$ defined by

$$X_t^x = x + B_t + \sum_{i=1}^{N_t} Y_i, \tag{1.6}$$

where $B = (B_t)_{t\geq 0}$ is a standard Brownian motion, $N = (N_t)_{t\geq 0}$ is a Poisson process with intensity $\lambda > 0$ independent of B, and $(Y_i)_{i\geq 0}$ are i.i.d. random variables independent of B and N with probability density $\frac{\varphi}{\lambda}$. Here, $\varphi(z)$ is the Lévy density of the Lévy process. When φ is Gaussian, it can be shown that there exist constants $C_1, c_1, C, c > 0$ such that for $0 < t \leq 1$ and $|y-x|$ sufficiently large, the upper and lower bounds of the density $p(t,x,y)$ of X_t^x satisfy

$$C_1 e^{-\lambda t} \exp\left(-c_1|y-x|\sqrt{\left|\ln\frac{|y-x|}{t}\right|}\right) \leq p(t,x,y) \leq \frac{C}{\sqrt{t}} \exp\left(-c|y-x|\sqrt{\left|\ln\frac{|y-x|}{t}\right|}\right), \tag{1.7}$$

and when φ is exponential,

$$C_1 e^{-\lambda t} e^{-c_1|y-x|} \leq p(t,x,y) \leq \frac{C}{\sqrt{t}} e^{-c|y-x|}. \tag{1.8}$$

This shows that the upper and lower bounds of the density are of different characteristic making impossible to implement the argument in Gobet [25].

To resolve the open issue which aims to extend the result obtained by Gobet [25] in the one-dimensional setting, this thesis will deal with independently but connectedly three different cases of jump-diffusion processes by following the Malliavin calculus approach developed by Gobet [24, 25]. In fact in order to be able to determine the strategy and the structure in the study of more general cases, it is essential to first well understand, on the one hand how this Malliavin calculus approach works, and on the other hand how the Gaussian-type estimate for the transition density conditioned on the jump structure is derived and employed for a simple Lévy process defined by

$$X_t^{\theta,\sigma,\lambda} = x_0 + \theta t + \sigma B_t + N_t - \lambda t, \tag{1.9}$$

where $B = (B_t)_{t\geq 0}$ is a standard Brownian motion, $N = (N_t)_{t\geq 0}$ is a Poisson process with intensity $\lambda > 0$ independent of B, and the parameters θ, σ, and λ are unknown.

We next generalise the aforementioned model by considering the following non-linear case

$$dX_t^\theta = b(\theta, X_t^\theta)dt + \sigma(X_t^\theta)dB_t + \int_{\mathbb{R}_0} c(X_{t-}^\theta, z)\left(N(dt,dz) - \nu(dz)dt\right), \tag{1.10}$$

where $N(dt,dz)$ is a Poisson random measure associated with a centered pure-jump Lévy process $\widehat{Z} = (\widehat{Z}_t)_{t\geq 0}$ independent of B, with intensity measure $\nu(dz)dt$, and finite Lévy measure $\lambda = \int_{\mathbb{R}_0} \nu(dz) < \infty$. Here, θ is unknown parameter to be estimated.

Moreover, in the same direction we finally study the following non-linear case

$$dX_t^{\theta,\beta} = b(\theta, X_t^{\theta,\beta})dt + \sigma(\beta, X_t^{\theta,\beta})dB_t + \int_{\mathbb{R}_0} z\left(N(dt, dz) - \nu(dz)dt\right), \tag{1.11}$$

where $N(dt, dz)$ is a Poisson random measure associated with a compensated compound Poisson process $\widehat{Z} = (\widehat{Z}_t)_{t\geq 0}$ independent of B, with intensity measure $\nu(dz)dt$. The random variable that describes the jump sizes of $\widehat{Z}$ takes discrete values. Here, θ and β are unknown parameters to be estimated.

1.3 Goal of the thesis and its main results

The goal of this thesis is to define situations where the jump process will not "deform" the Gaussian nature of the statistical experiment. As commented before, this cannot be achieved by simply obtaining upper and lower bounds of the transition density. Instead, we will condition on the jump structure and use large deviation results that will guarantee that the Gaussian nature of the statistical experiment will remain unchanged. Clearly, this is just a first effort towards a much more general problem where one may have that the Lévy nature of the statistical experiment remains unchanged by the Wiener noise. Even more difficult is to determine the boundary situations. We leave as future research the study of this open problem.

In this thesis we deal with some uniformly elliptic diffusion processes with jumps and study the LAN property from discrete observations of their solution processes. For our objective, we think it is essential to first understand the proof of this property in the continuously observed case. As a result, excluding this introductory chapter, this thesis consists of four self-contained chapters each of which deals with a different jump-diffusion process. Note that the following chapters are independent of each other and the utilized notations are provided inside every chapter. The global bibliography is given at the end of this thesis.

We will next describe the content of each of the chapters in detail.

1.3.1 LAMN property for continuous observations of jump-diffusion processes

In Chapter 2, we consider a d-dimensional process $X^\theta = (X_t)_{t\geq 0}$ solution to the following stochastic differential equation with jumps

$$dX_t = a(\theta, X_t)dt + \sigma(X_t)dB_t + \int_{\mathbb{R}_0^d} c(\theta, X_{t-}, z)(p(dt, dz) - \nu_\theta(dz)dt), \tag{1.12}$$

where $\mathbb{R}_0^d = \mathbb{R}^d \setminus \{0\}$, the unknown parameter θ belongs to an open subset Θ of $\mathbb{R}^k$, $k \geq 1$, $B = (B_t)_{t\geq 0}$ is a d-dimensional standard Brownian motion, and $p(dt, dz)$ is a Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}_0^d$, independent of B with intensity measure $\nu_\theta(dz)dt = f(\theta, z)dzdt$. Here, $\nu_\theta(dz)$ is a Lévy measure on $\mathbb{R}_0^d$ such that $\int_{\mathbb{R}_0^d}(1 \wedge |z|^2)\nu_\theta(dz) < \infty$, for all $\theta \in \Theta$, and $f : \Theta \times \mathbb{R}^d \to \mathbb{R}_+$ is a Borel function strictly positive on $\mathbb{R}_0^d$ with $f(\theta, 0) = 0$.

We give sufficient conditions and follow Luschgy's [52] proof in order to derive the LAMN property (Theorem 2.2.4) when the process is observed continuously in a time interval $[0, T]$ as $T \to +\infty$. We give a Girsanov's theorem and apply the Central Limit theorem for multivariate martingales developed by Crimaldi and Pratelli [14]. Recall that Luschgy's paper shows the LAMN property for general semimartingales using the Girsanov's theorem for semimartingales obtained in Jacod and Shiryaev [33], and the Central Limit theorem for martingales established by Sørensen [71] and Feigin [18]. Here we rewrite the proof of these results without using this abstract semimartingale theory but integral equations with respect to random measures associated with the jumps of the process. Moreover, as a consequence of Theorem 2.2.4, we derive the LAN property in the ergodic case (Theorem 2.4.2).

1.3.2 LAN property for a simple Lévy process

The focus of Chapter 3 is on a simple Lévy process $X^{\theta,\sigma,\lambda} = (X_t^{\theta,\sigma,\lambda})_{t\geq 0}$ in $\mathbb{R}$ defined by

$$X_t^{\theta,\sigma,\lambda} = x_0 + \theta t + \sigma B_t + N_t - \lambda t, \tag{1.13}$$

where $B = (B_t)_{t\geq 0}$ is a standard Brownian motion, $N = (N_t)_{t\geq 0}$ is a Poisson process with intensity $\lambda > 0$ independent of B. The parameters $(\theta, \sigma, \lambda) \in \Theta \times \Sigma \times \Lambda$ are unknown and Θ, Σ and Λ are closed intervals of $\mathbb{R}, \mathbb{R}_+^*$ and $\mathbb{R}_+^*$. In finance this is called the Merton jump-diffusion (MJD) model. The MJD model is one of the first beyond Black-Scholes model in the sense that it tries to capture the negative skewness and excess kurtosis of the log stock price density.

The aim of this chapter is to prove the LAN property under high-frequency observation condition of $X^{\theta,\sigma,\lambda}$ (Theorem 3.1.1). For this, Malliavin calculus and Girsanov's theorem are applied in order to write the log-likelihood ratio in terms of sums of conditional expectations, for which a central limit theorem for triangular arrays can be applied. The techniques used to obtain this result will be discussed in detail in the next subsection, and note that they can be generalized to the case of stochastic differential equations with finite number of jumps and random jump size (Chapters 4-5). With the help of a large deviation principle by conditioning on the number of jumps, our main contribution here is that we find a closed form expression for the corresponding large deviation estimate (Lemma 3.2.6), thereby allowing us to control the jump components in the negligible contribution of the limit (Lemmas 3.3.1, 3.3.2 and 3.3.3).

Since we are dealing with a simple Lévy process with finite jumps, the explicit expression of the density could be used in order to derive the LAN property, as for e.g. in [2]. However, the main purpose of this chapter is to understand and present the methodology for this simple case, which will be next used to prove the LAN property in the non-linear cases where the density function cannot be explicitly written.

1.3.3 LAN property for a jump-diffusion process : drift parameter

In Chapter 4 we address the validity of the LAN property for a jump-diffusion process $X^\theta = (X_t^\theta)_{t\geq 0}$ solution to

$$dX_t^\theta = b(\theta, X_t^\theta)dt + \sigma(X_t^\theta)dB_t + \int_{\mathbb{R}_0} c(X_{t-}^\theta, z)\left(N(dt,dz) - \nu(dz)dt\right), \tag{1.14}$$

where $X_0^\theta = x_0 \in \mathbb{R}$, $\mathbb{R}_0 := \mathbb{R}\setminus\{0\}$, $B = (B_t)_{t\geq 0}$ is a standard Brownian motion, and $N(dt,dz)$ is a Poisson random measure in $(\mathbb{R}_+ \times \mathbb{R}_0, \mathcal{B}(\mathbb{R}_+ \times \mathbb{R}_0))$ associated with a centered pure-jump Lévy process $\widehat{Z} = (\widehat{Z}_t)_{t\geq 0}$ independent of B, with intensity measure $\nu(dz)dt$, and finite Lévy measure $\lambda = \int_{\mathbb{R}_0} \nu(dz) < \infty$. The unknown parameter θ belongs to Θ which is a closed interval of $\mathbb{R}$.

Supposing that the process is observed discretely at high frequency, we then give a set of sufficient conditions **(A1)**-**(A8)** (see page 66) on the regularity of the coefficients, the ergodicity and the behaviour of the Lévy measure in order to obtain the LAN property for X^θ (Theorem 4.1.1). The proof of this result is essentially based on the Malliavin calculus, the Girsanov's theorem and the large deviation principle developed from the aforementioned case.

Notice that the condition on the jump coefficient c in **(A1)** and **(A3)** is needed in order to control the behaviour of the jump amplitudes of X^θ, which can be seen in the discussion in the proof of Lemma 4.2.8.

Several examples of ergodic diffusion processes with jumps are given in [53], [54], and [67]. Moreover, results on ergodicity and exponential ergodicity for diffusion processes with jumps have been established by Masuda in [53, 54]. In addition, Kulik in [43] provides a set of sufficient conditions for the exponential ergodicity of diffusion processes with jumps without Gaussian part and gives some examples. More recently, Qiao in [64] has addressed the exponential ergodicity for stochastic differential equations with jumps and non-Lipschitz coefficients. However, in these papers ergodicity and exponentially ergodicity are understood in the sense of [56], which are both stronger than the ergodicity in the sense **(A6)**.

Note that condition **(A7)** involving the behaviour of the jumps is imposed to ensure that the jump component is "dominated" over by the Gaussian component. Indeed this condition expresses the fact that the small and large jumps do not interfere with the Gaussian behaviour of the transition density. Therefore, the main behaviour in the contribution is given by the Gaussian and drift components of the equation. As a consequence, the asymptotic Fisher information is identical to the one for ergodic diffusion processes without jumps obtained by Gobet [25].

Hypothesis **(A7)** is a sufficient condition that implies that the probability that the jump amplitudes of $\widehat{Z}$ is bounded below by $\rho_1\Delta_n^{\upsilon}$ and above by $\rho_2\Delta_n^{-\gamma}$ on an interval $[0,t_n]$ converges to 0 as $n\to\infty$ (see the computations of page 82). This allows to use this event (see Remark 4.1.3 and page 80), and this is the main trick for the proof of the LAN property. Hypothesis **(A7)** restricts the jump component of the process to have small and large jumps that decay exponentially. For example, a Gaussian amplitude of jumps, with an exponential decay for small jumps. In order to get rid of this hypothesis maybe a convergence argument could be used.

Finally, in order to include the case of unbounded drift coefficient, the squared exponential moment condition **(A8)** is needed. Recall that the problem of the boundedness of the squared exponential moment **(A8)** already appeared in the proof of the LAN property for continuous ergodic diffusion processes (see [25, Proposition 1.1]). In the case of jump-diffusion processes, Masuda gives sufficient conditions on the infinitesimal generator in order to obtain the boundedness of the moments of certain class of unbounded functions (see [53, Theorem 2.2]). Moreover, he establishes in [54, Theorem 1.2] the boundedness of the exponential moments for a class of jump-diffusion processes with finite jump intensity. It is possible that a similar result should be available for hypothesis **(A8)**. We will not discuss this part here in general. For example, **(A8)** is satisfied for the Ornstein-Uhlenbeck process under certain condition on the Lévy measure.

In Chapter 3 we estimate the drift and diffusion parameters and the jump intensity of a simple Lévy process. Therefore, Theorem 4.1.1 is a non-linear extension of the result in Chapter 3 when the unknown parameter is in the drift coefficient.

However, we also remark that condition **(A7)** is not optimal and the condition on the small jumps could be weakened. Indeed, using a convergence argument around the small jumps, a condition $\lambda_n\sqrt{\Delta_n}\to 0$ as $n\to\infty$ may be needed, where

$$\lambda_n=\int_{\{\Delta_n^{\upsilon}\leq|z|\leq\rho\Delta_n^{-\gamma}\}}\nu(dz).$$

This is done in our work in progress where the Lévy measure is assumed to be infinite.

1.3.4 LAN property for a jump-diffusion process : drift and diffusion parameters

In Chapter 5 we consider the process $X^{\theta,\beta}=(X_t^{\theta,\beta})_{t\geq0}$ solution to

$$dX_t^{\theta,\beta}=b(\theta,X_t^{\theta,\beta})dt+\sigma(\beta,X_t^{\theta,\beta})dB_t+\int_{\mathbb{R}_0}z\left(N(dt,dz)-\nu(dz)dt\right), \tag{1.15}$$

where $X_0^{\theta,\beta}=x_0\in\mathbb{R}$, $B=(B_t)_{t\geq0}$ is a standard Brownian motion, and $N(dt,dz)$ is a Poisson random measure in $(\mathbb{R}_+\times\mathbb{R}_0,\mathcal{B}(\mathbb{R}_+\times\mathbb{R}_0))$ associated with a compensated compound Poisson process $\widehat{Z}=(\widehat{Z}_t)_{t\geq0}$ independent of B, with intensity measure $\nu(dz)dt$. Assume that the random variable that describes the jump sizes of $\widehat{Z}$ takes discrete values. The unknown parameters (θ,β) belong to $\Theta\times\Sigma$ which is an open rectangle of $\mathbb{R}^2$.

Supposing that the process is observed discretely at high frequency, we then give a set of sufficient conditions **(A1)**-**(A8)** (see pages 96 and 97) on the regularity of the coefficients, the ergodicity, the behaviour of the Lévy measure and the identification of jumps in order to obtain the LAN property for $X^{\theta,\beta}$ (Theorem 5.1.1). Generally, the proof of this result is essentially based on the Malliavin calculus, the Girsanov's theorem and the large deviation principle developed from the aforementioned case.

The fact that the jump sizes of $\widehat{Z}$ take discrete values and the drift coefficient is bounded is restrictive. Indeed this is just a first step towards treating a much more general parametric model. As explained before, all the problematic of this thesis concerns the argument in the article of Gobet [25] which is based on the fact that the transition densities satisfy the Gaussian type upper and lower bounds. Initially, we had the tendency to believe that "good" upper and lower estimates of the density should essentially solve the problem. Therefore, we started this thesis by studying the type behaviour of the upper and lower bounds of the transition densities for some jump-diffusion processes, for instance (1.6). The obtained estimates (1.7) and (1.8) show that Gobet's argument cannot be implemented. However, the trick of this chapter is to re-employ the "argument of Gaussian-type estimate" which enables to solve the problem as in Gobet [25]. For this, we are restricting ourselves to assuming that the jump sizes of $\widehat{Z}$ take discrete values in order to obtain the Gaussian-type estimate for the conditioned transition density (see Lemma 5.2.5).

We recall that Gobet in [25] deals with the multidimensional ergodic diffusion processes whose diffusion coefficient is assumed to be uniformly strictly elliptic and whose drift coefficient can be bounded or unbounded. The ergodicity result in the case of bounded drift coefficient was addressed in Proposition 5.1 of this article. In this chapter, we are first interested in studying the case of uniformly elliptic diffusion coefficient and bounded drift coefficient expressed by hypothesis **(A2)**. Let us also mention that the result on ergodicity and exponential ergodicity for this class of diffusion processes with jumps was established by Masuda in [54, Theorem 1.2].

Notice that hypotheses **(A5)** and **(A6)** are the same as **(A6)** and **(A7)** in Chapter 4, which were explained in Chapter 4.

As in Chapter 4, in order to deal with the parameter θ in the drift coefficient, we condition on the number of jumps occured in each time interval $[t_k, t_{k+1}]$. As a consequence, the large deviation estimate (Lemma 5.2.14) is obtained. Here, in order to deal with the parameter β in the Brownian component, we will condition on all the possible jump sizes which are assumed to be a countable set $\mathcal{A}$. Therefore, hypotheses **(A7)** and **(A8)** on the behaviour of jumps are added in order to obtain a large deviation principle (Lemma 5.2.15) for the parameter in the Brownian coefficient. In fact, the first condition in **(A7)** is related to the identification of jumps, that is, any two sums of jumps on a small interval for different $\omega \in \Omega$ are either equal or their difference is lower bounded by a value depending on Δ_n. This is used in the computations in pages 117 and 118. Furthermore, the second condition in **(A7)** is used in order to condition on the sum of jumps. Finally, hypothesis **(A8)** on the jump distributions is needed in order for the expression in Remark 5.1.2 to be finite, which ensures the convergence to zero of the negligible contributions. This is because the proof is based on the conditioning on the number of jumps and on the amplitudes of jumps.

Our contribution here is to derive an expression for the derivatives of the log-likelihood function conditioned on the number and the amplitudes of jumps in terms of a conditional expectation by adapting Gobet's Malliavin calculus approach (Lemma 5.2.7).

In Chapter 3 we estimate the drift and diffusion parameters and the jump intensity of a simple Lévy process. Therefore, Theorem 5.1.1 is a non-linear extension of the result in Chapter 3 when the unknown parameters are in the drift and diffusion coefficients.

There are two extensions of the results of this chapter that we should think about in our future research. The first one is to consider an unbounded drift coefficient and to add a non-linear jump coefficient in front of the compound Poisson process (as in equation (1.14)), and the second one is to consider a more general jump size distribution. The main ideas used in this chapter should be enough in order to deal with these two extensions, but of course the computations would be much more difficult, and conditions **(A6)**-**(A8)** need to be adjusted. This is the reason why we have restricted ourselves to these two particular cases, but we do not think that it is a restriction of our methodology.

The remainder of this introductory chapter is devoted to explaining the common techniques used to solve Chapters 3-5.

1.4 Main techniques

Although Chapters 3-5 seem to be different, they are all connected in the sense that the same techniques are used to solve them. We set out here briefly our strategy of the proof whose structure remains the same for each chapter and can be divided into three main steps as described below.

1.4.1 Malliavin calculus approach

The goal of this subsection is to present an adaptation of Gobet's Malliavin calculus approach to our setting. The first step of the proof proceeds with the decomposition of the log-likelihood ratio in terms of sums of transition densities due to the Markov property (see (3.3), (5.33)). As in Gobet [24], with the help of the uniformly elliptic condition on the diffusion coefficient, we can apply the integration by parts formula of the Malliavin calculus on the Wiener space induced by the Brownian motion on each observation interval in order that the derivatives of the log-likelihood function with respect to the parameters are expressed in terms of a conditional expectation involving Skorohod intergrals (Propositions 3.2.1, 4.2.1 and 5.2.1). For this, an independent copy of the observed process needs to be introduced. Using tools of Malliavin calculus, these Skorohod intergrals are decomposed into two parts (Lemmas 4.2.1, 5.2.1 and 5.2.2), where the conditional expectation of the first part can be easily computed, which gives the main contributions in the limit of the convergence of the log-likelihood ratio. More precisely, the main behaviour here will be determined by the Gaussian and drift components of the equation. On the other hand, the second part whose conditional expectation cannot be easily computed will have no contribution in the limit, which causes difficulties in the control of the convergence that we will explain in the next subsection. Consequently, the expansion of the log-likelihood ratio is separated into the main and negligible contributions.

Moreover, adapting Gobet's Malliavin calculus approach can be further expressed by Lemma 5.2.7 where we derive an expression for the derivatives of the log-likelihood function conditioned on the number and the amplitudes of jumps in terms of a conditional expectation involving Skorohod intergrals.

1.4.2 Large deviation principle and Girsanov's theorem

The aim of this subsection is to present how to deal with negligible contributions of the log-likelihood ratio. This is a crucial and technical part of the thesis. For this, we need two general results on convergence in probability for triangular arrays of random variables in order to prove the convergence of a sum of triangular arrays. For each $n \in \mathbb{N}$, consider a sequence of random variables $(Z_{k,n})_{k\geq 1}$ defined on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathrm{P})$, and we assume that they are $\mathcal{F}_{t_{k+1}}$-measurable, for all k.

Lemma 1.4.1. [21, Lemma 9] *Assume that as $n \to \infty$,*

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

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

Lemma 1.4.2. [31, Lemma 4.1] *Assume that as $n \to \infty$,*

$$\sum_{k=0}^{n-1} \mathrm{E}\left[|Z_{k,n}||\mathcal{F}_{t_k}\right] \xrightarrow{\mathrm{P}} 0. \tag{1.16}$$

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

We note that dealing with the negligible contributions in the presence of jumps is different from that in the continuous diffusion case. The point is that in Gobet [24, 25] the upper and lower Gaussian type bounds of the transition density are essentially used in order to control the negligible contribution in the limit. However, in the presence of jumps, we cannot expect to find such good Gaussian type estimates of the transition density due to the mixture of exponential tails coming from the jump process together with the Gaussian tails of the Brownian motion. As a result, a problem arises in dealing with the jump-diffusion processes.

In order to overcome this problem when treating the jump components in the negligible contributions (Lemmas 3.3.1, 3.3.2, 3.3.3, 4.3.5, 5.3.5 and (5.45) of Lemma 5.3.8), one needs to condition on the number of jumps or on the sum of jumps within the conditional expectation which expresses the transition density and outside it. When these two conditionings relate to different jumps or different sum of the jumps one may use a large deviation principle in the estimate. When they are equal one uses the complementary set in order to apply the large deviation principle. The main term can be handled directly. Within all these arguments the Gaussian type upper and lower bounds of the density conditioned on the jumps are again strongly used.

On the other hand, in order to obtain the expected large deviation estimates (see Lemmas 3.2.6, 4.2.8, 5.2.14 and 5.2.15), the condition on the behaviour of the jumps ensuring that the jump component is dominated over by the Gaussian component is needed. In fact this condition expresses the fact that the small and large jumps do not interfere with the Gaussian behaviour of the transition density (see **(A7)** in Chapter 4 and **(A6)** in Chapter 5), which is again employed in Lemmas 4.3.1, 5.3.1 and 5.3.6. Moreover, in Chapter 5 condition **(A7)** on the behaviour of the sum of jumps on a small interval is needed. This condition, on the one hand, allows us to condition on the sum of jumps on each interval. On the other hand, it is related to the identification of jumps, that is, any two sums of jumps on a small interval for different $\omega \in \Omega$ are either equal or their difference is lower bounded by a value depending on Δ_n, from which the large deviation estimate Lemma 5.2.15 can be obtained.

Moreover, in Chapter 5 another difficulty comes from the fact that when applying Lemmas 1.4.1 and 1.4.2 the expectations outside and inside are under two different probability measures. More precisely, the conditional expectation inside needs to be computed under $\mathrm{P}^{\theta_n,\beta(\ell)}$, whereas the convergence is considered with respect to $\mathrm{P}^{\theta_0,\beta_0} \neq \mathrm{P}^{\theta_n,\beta(\ell)}$. To this end, when the two corresponding diffusion parameters are different, we need to condition on the number of jumps and on the amplitudes of jumps in order that the change of measure can be done via the transition density conditioned on the jumps (Lemma 5.2.10). As a consequence, the upper and lower Gaussian type bounds of the transition density conditioned on the jumps in Lemma 5.2.5 are again strongly used. In this case, condition **(A8)** on the jump distributions is needed. On the other hand when two diffusion parameters are the same, which means that only the drift parameters are different, the Girsanov's theorem can be applied (see Lemmas 3.2.2, 4.2.4 and 5.2.11). Note that the technical Lemmas 4.2.5 and 5.2.12 are given in order to measure the deviations of the Girsanov change of measure when the drift parameter changes.

Moreover, in Proposition 3.2.2 of Chapter 3, the derivative of the log-likelihood w.r.t. the intensity parameter λ is expressed in terms of a conditional expectation with the help of the Girsanov's theorem.

1.4.3 Central limit theorem for triangular arrays

To conclude the LAN property, the last step consists in dealing with the main contributions of the log-likelihood ratio. For this, it suffices to apply the central limit theorem for triangular arrays of random variables as indicated just below. For each $n \in \mathbb{N}$, consider a sequence of random variables $(\zeta_{k,n})_{k\geq 1}$ defined on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathrm{P})$, and we assume that they are $\mathcal{F}_{t_{k+1}}$-measurable, for all k.

Lemma 1.4.3. [31, Lemma 4.3] *Assume that there exist real numbers M and $V > 0$ such that as $n \to \infty$,*

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

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

Generally, the mean M and the variance V can be obtained directly in the case of simple Lévy process or by applying a discrete time ergodic theorem (Lemmas 4.2.9, 5.2.16) in the non-linear cases. As a result, the ergodicity condition is needed in Chapters 4-5.

Chapitre 2

LAMN property for continuous observations of diffusion processes with jumps

In this chapter we consider a diffusion process with jumps whose drift and jump coefficient depend on an unknown parameter. We follow Luschgy's [52] proof of the local asymptotic mixed normality (LAMN) property when the process is observed continuously in a time interval $[0, T]$ as $T \to +\infty$, and derive, as a consequence, the local asymptotic normality (LAN) property in the ergodic case. However, we give a Girsanov's theorem and apply the Central Limit theorem for multivariate martingales developed by Crimaldi and Pratelli [14]. Luschgy's paper shows the LAMN property for general semimartingales using the Girsanov's theorem for semimartingales obtained in Jacod and Shiryaev [33], and the Central Limit theorem for martingales established by Sørensen [71] and Feigin [18]. Here we rewrite the proof of these results without using this abstract semimartingale theory but integral equations with respect to random measures associated with the jumps of the process.

2.1 Introduction

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

$$dX_t = a(\theta, X_t)dt + \sigma(X_t)dB_t + \int_{\mathbb{R}_0^d} c(\theta, X_{t-}, z)\left(p(dt, dz) - \nu_\theta(dz)dt\right), \tag{2.1}$$

where the initial condition X_0 is a random variable with finite second moment, $\mathbb{R}_0^d = \mathbb{R}^d \backslash \{0\}$, the unknown parameter θ belongs to an open subset Θ of $\mathbb{R}^k$, for some integer $k \geq 1$, $B = (B_t)_{t\geq 0}$ is a d-dimensional standard Brownian motion, and $p(dt, dz)$ is a Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}_0^d$, independent of B with intensity measure $\nu_\theta(dz)dt = f(\theta, z)dzdt$. Here, $\nu_\theta(dz)$ is a Lévy measure on $\mathbb{R}_0^d$ such that $\int_{\mathbb{R}_0^d}(1 \wedge |z|^2)\nu_\theta(dz) < \infty$, for all $\theta \in \Theta$, and $f : \Theta \times \mathbb{R}^d \to \mathbb{R}_+$ is a Borel function strictly positive on $\mathbb{R}_0^d$ with $f(\theta, 0) = 0$.

The coefficients $a = (a_i)$ and $c = (c_i)$ are $\mathbb{R}^d$-valued Borel functions on $\Theta \times \mathbb{R}^d$ and $\Theta \times \mathbb{R}^d \times \mathbb{R}_0^d$, respectively, and $\sigma = (\sigma_{ij})$ is a $d \times d$ invertible Borel matrix on $\mathbb{R}^d$.

We let $\{\mathcal{F}_t\}_{t\geq 0}$ denote the natural filtration generated by the Brownian motion and the Poisson random measure. By definition, the solution to equation (2.1) is a càdlàg and $\{\mathcal{F}_t\}$-adapted d-dimensional stochastic process $X^\theta = (X_t)_{t\geq 0}$ defined on the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathrm{P})$ such that

$$X_t = X_0 + \int_0^t a(\theta, X_s)ds + \int_0^t \sigma(X_s)dB_s + \int_0^t \int_{\mathbb{R}_0^d} c(\theta, X_{s-}, z)\left(p(ds, dz) - \nu_\theta(dz)ds\right). \tag{2.2}$$

For any $\theta \in \Theta$, we denote by P_θ the probability measure induced by the solution X^θ of (2.1) on the canonical space $(D(\mathbb{R}^d), \mathcal{B}(\mathbb{R}^d))$, where $D(\mathbb{R}^d)$ denotes the space of càdlàg functions from $\mathbb{R}^d$ to $\mathbb{R}^d$, and $\mathcal{B}(\mathbb{R}^d)$ its associated Borel σ-algebra. Moreover, for any $T \geq 0$, we let P_θ^T denote the probability measure generated by the process $X^T = \{X_t, 0 \leq t \leq T\}$ solving equation (2.1) under the parameter θ on the measurable space $(D[0,T], \mathcal{B}[0,T])$. Therefore, P_θ^T is the restriction of P_θ to $\mathcal{F}_T$. For any $\theta \in \Theta$, we denote by E_θ the expectation with respect to the probability law P_θ, and $\overset{\mathrm{P}_\theta}{\longrightarrow}$ and $\overset{\mathcal{L}(\mathrm{P}_\theta)}{\longrightarrow}$ denote the convergence in P_θ-probability and in P_θ-law, respectively.

In this chapter, we are interested in the statistical inference for $\theta \in \Theta$ on the basis of continuous-time observations of the process X^T in the time interval $[0,T]$, as T tends to $+\infty$. Let us start by recalling the concepts on asymptotic statistical inference that we are interested in for our continuously observed parametric model.

We define the log-likelihood function of the family of probability measures $(\mathrm{P}_\theta^T)_{\theta\in\Theta}$ as

$$\ell_T(\theta) = \log \frac{d\mathrm{P}_\theta^T}{d\widetilde{\mathrm{P}}^T},$$

where $\widetilde{\mathrm{P}}^T$ is a probability measure on $(D[0,T], \mathcal{B}[0,T])$, if it exists, satisfying that P_θ^T is absolutely continuous with respect to $\widetilde{\mathrm{P}}^T$, for all $T \geq 0$ and $\theta \in \Theta$.

The score function, when it exists, is given by the gradient $\nabla_\theta \ell_T(\theta)$. We say that the score function is asymptotically mixed normal if, for any $\theta \in \Theta$, there exists a $k \times k$ non-random diagonal matrix $\varphi_T(\theta)$ whose entries are strictly positive and tend to zero as $T \to \infty$, and a $k \times k$ positive definite random matrix $\Gamma(\theta)$, such that as $T \to \infty$,

$$\varphi_T(\theta)\nabla_\theta \ell_T(\theta) \overset{\mathcal{L}(\mathrm{P}_\theta)}{\longrightarrow} \Gamma(\theta)^{1/2}\mathcal{N}(0, I_k), \tag{2.3}$$

where $\mathcal{N}(0, I_k)$ denotes a centered $\mathbb{R}^k$-valued Gaussian random variable independent of $\Gamma(\theta)$ with identity covariance matrix I_k. In this case, the matrix $\Gamma(\theta)$ is called the asymptotic Fisher information matrix of the model. When the matrix $\Gamma(\theta)$ is deterministic, we say that the score function is asymptotically normal.

The family of probability measures $(\mathrm{P}_\theta^T)_{\theta\in\Theta}$ is said to have the LAMN property if for any $\theta \in \Theta$ and $u \in \mathbb{R}^k$, as $T \to \infty$,

$$\log \frac{d\mathrm{P}^T_{\theta+\varphi_T(\theta)u}}{d\mathrm{P}_\theta^T} \overset{\mathcal{L}(\mathrm{P}_\theta)}{\longrightarrow} u^\mathsf{T}\Gamma(\theta)^{1/2}\mathcal{N}(0, I_k) - \frac{1}{2}u^\mathsf{T}\Gamma(\theta)u, \tag{2.4}$$

where $\mathcal{N}(0, I_k)$, $\varphi_T^{-1}(\theta)$, and $\Gamma(\theta)$ are as in (2.3). In this case, we say that the LAMN property holds with rate of convergence $\varphi_T^{-1}(\theta)$ and asymptotic Fisher information matrix $\Gamma(\theta)$. When the matrix $\Gamma(\theta)$ is deterministic, we say that the LAN property holds.

Observe that (2.4) is equivalent to

$$\begin{aligned}\log \frac{d\mathrm{P}^T_{\theta+\varphi_T(\theta)u}}{d\mathrm{P}_\theta^T} &= \ell_T(\theta + \varphi_T(\theta)u) - \ell_T(\theta)\\ &= u^\mathsf{T}\varphi_T(\theta)\nabla_\theta\ell_T(\theta) - \frac{1}{2}u^\mathsf{T}\Gamma(\theta)u + o_{\mathrm{P}_\theta}(1),\end{aligned} \tag{2.5}$$

where $\varphi_T(\theta)\nabla_\theta \ell_T(\theta)$ converges in P_θ-law to $\Gamma(\theta)^{1/2}\mathcal{N}(0, I_k)$ as $T \to \infty$.

The aim of this chapter is to revise sufficient conditions in order to have the asymptotic mixed normality of the score function and the LAMN property for our diffusion model with jumps (2.1). This problem was addressed by Luschgy for semimartingales in [52] by using the Girsanov's theorem for semimartingales established by Jacod and Shiryaev (see [33, Theorem III.3.24 and III.5.19]), and the Central Limit theorem for multivariate martingales proved by Sørensen (see [71, Theorem A.1]), as an extension of the Central Limit theorem for one-dimensional martingales [18, Theorem 2] by Feigin. We remark that the stochastic process with jumps (2.2) is a semimartingale. Therefore, Luschgy's theorem applies and one can derive sufficient conditions on the coefficients in order to have the LAMN property. The aim of this chapter is to present a proof of the LAMN property for the solution X^θ of (2.1) by following the proof of Luschgy but applying the Central Limit theorem for multivariate martingales developed by Crimaldi and Pratelli [14] without using the fact that we have a semimartingale, but using the integral equation (2.2). We then deduce the LAN property with an explicit asymptotic Fisher information matrix in the case where the process X^θ is ergodic. To obtain the desired results, the first step consists in transforming equation (2.1) into a new stochastic differential equation with jumps driven by a random measure associated with the jumps of X^θ. Notice that this approach was also employed by Sørensen in [71]. One of the motivations of writing this chapter is that we are investigating in further chapters the LAMN property for the stochastic differential equations with jumps (2.1) with discrete observations in a time interval $[0, T]$ as $T \to \infty$, which has never been addressed in the literature. For this, we think it is essential to first understand the proof of this property in the continuously observed case but without using the asbtract semimartingale theory, but integral equations with respect to random measures.

This chapter is organized as follows. In Section 2, we provide sufficient conditions and prove the asymptotic mixed normality of the score function as well as the LAMN property for the stochastic differential equation with jumps (2.1). For this purpose, we recall the Central Limit theorem for multivariate martingales developed by Crimaldi and Pratelli [14]. Furthermore, studying the LAMN property from continuous observations is based on the Girsanov's theorem for equivalent probability measures. Therefore, we will give this fundamental result in Section 3. The proof of the LAN property in the ergodic case as a consequence of the LAMN property is given in Section 4. Finally, Section 5 deals with the pure linear birth process and the Ornstein-Uhlenbeck processes with jumps where the LAMN and LAN properties are satisfied and the maximum likelihood estimator is asymptotically efficient in some particular cases.

2.2 LAMN property for jump-diffusion processes

The aim of this section is to give sufficient conditions in order to have the asymptotic mixed normality of the score function and the LAMN property for our stochastic differential equation with jumps (2.1). To this purpose, let us first recall the result on the existence and uniqueness of the solution to our integral equation (2.2), that is,

$$X_t = X_0 + \int_0^t a(\theta, X_s)ds + \int_0^t \sigma(X_s)dB_s + \int_0^t \int_{\mathbb{R}_0^d} c(\theta, X_{s-}, z)\left(p(ds, dz) - \nu_\theta(dz)ds\right).$$

Consider the following Lipschitz continuity and linear growth conditions on the coefficients.

(A1) For any $\theta \in \Theta$, there exist a constant $L > 0$ and a function $\zeta : \mathbb{R}_0^d \to \mathbb{R}_+$ satisfying that $\int_{\mathbb{R}_0^d} \zeta^2(z)\nu_\theta(dz) < \infty$, such that for any $x, y \in \mathbb{R}^d$, $z \in \mathbb{R}_0^d$,

$$|a(\theta, x) - a(\theta, y)| + |\sigma(x) - \sigma(y)| \leq L|x - y|, \quad |a(\theta, x)| \leq L(1 + |x|),$$
$$|c(\theta, x, z) - c(\theta, y, z)| \leq \zeta(z)|x - y|, \quad |c(\theta, x, z)| \leq \zeta(z)(1 + |x|).$$

Theorem 2.2.1. *[33, Theorem III.2.32] Under condition* **(A1)**, *there exists a unique càdlàg and adapted process* $X^\theta = (X_t)_{t\geq 0}$ *solution to equation* (2.1) *on the filtered probability space* $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathrm{P})$. *Moreover, for any fixed* $p > 0$ *and* $T > 0$, *there exists a constant* $C_{p,T} > 0$ *such that for all* $t_0 \in (0, T]$ *and* $t \in [t_0, T]$,

$$\mathrm{E}\left[\sup_{t_0 \leq s \leq t} |X_s - X_{t_0}|^p\right] \leq C_{p,T}(t - t_0)^{\frac{p}{2}\wedge 1}\mathrm{E}\left[(1 + |X_{t_0}|^2)^{p/2}\right].$$

Let us now proceed as in [71] to transform our equation (2.1) into a new stochastic differential equation with jumps driven by a new random measure associated with the jumps of X^θ via a change of variables. To simplify the exposition we assume that $\mathrm{Im}(c) = \mathbb{R}^d$.

For each $(\theta, x) \in \Theta \times \mathbb{R}^d$ fixed, we assume that the mapping $z \in \mathbb{R}_0^d \mapsto y = c(\theta, x, z) \in \mathbb{R}_0^d$ has a continuous differentiable inverse $y \in \mathbb{R}_0^d \mapsto z = c^{-1}(\theta, x, y) \in \mathbb{R}_0^d$ with Jacobian matrix $J(\theta, x, y)$ such that $\det(J(\theta, x, y)) \neq 0$, for all $y \in \mathbb{R}_0^d$.

Set $\Psi(\theta, x, y) = f(\theta, c^{-1}(\theta, x, y))|\det(J(\theta, x, y))|$. Suppose that for any $(\theta, x) \in \Theta \times \mathbb{R}^d$, $\int_{\mathbb{R}_0^d} |c(\theta, x, z)|\nu_\theta(dz) < +\infty$. Then by [33, Proposition II.1.28], equation (2.1) can be rewritten as follows

$$dX_t = b(\theta, X_t)dt + \sigma(X_t)dB_t + \int_{\mathbb{R}_0^d} yN(dt, dy), \tag{2.6}$$

where the function $b : \Theta \times \mathbb{R}^d \to \mathbb{R}^d$ is given by

$$b(\theta, x) = a(\theta, x) - \int_{\mathbb{R}_0^d} c(\theta, x, z)\nu_\theta(dz) = a(\theta, x) - \int_{\mathbb{R}_0^d} y\mu_\theta(x, dy),$$

and $N(dt, dy)$ is the random measure on $\mathbb{R}_+ \times \mathbb{R}_0^d$ associated with the jumps of X with predictable compensator $\mu_\theta(X_{t-}, dy)dt = \Psi(\theta, X_{t-}, y)dydt$, defined by

$$N(\omega; dt, dy) = \sum_{s \geq 0} \mathbf{1}_{\{\Delta X_s(\omega) \neq 0\}} \delta_{(s, \Delta X_s(\omega))}(dt, dy),$$

(see [33, Proposition II.1.16]), where δ_a denotes the Dirac measure at point a. By [33, Theorem II.1.8], the predictable compensator $\mu_\theta(X_{t-}, dy)dt$ is the unique predictable random measure satisfying that

$$\mathrm{E}\left[\int_0^\infty \int_{\mathbb{R}_0^d} \psi(t, y)N(dt, dy)\right] = \mathrm{E}\left[\int_0^\infty \int_{\mathbb{R}_0^d} \psi(t, y)\mu_\theta(X_{t-}, dy)dt\right],$$

for every nonnegative predictable function $\psi(t, y)$ on $\Omega \times \mathbb{R}_+ \times \mathbb{R}_0^d$.

In order to obtain the asymptotic mixed normality of the score function and the LAMN property, we assume that there exists a $k \times k$ non-random diagonal matrix $\varphi_T(\theta)$ whose diagonal entries $\varphi_{i,T}(\theta)$ are strictly positive and tend to zero as $T \to \infty$, and such that the following conditions hold.

(A2) For any $\theta, \theta_0 \in \Theta$, and $T \geq 0$,

$$\mathrm{P}_{\bar{\theta}}\left(\int_0^T |\sigma^{-1}(X_t)\left(b(\theta_0, X_t) - b(\theta, X_t)\right)|^2 dt < \infty\right) = 1, \quad \text{for} \quad \bar{\theta} = \theta, \theta_0.$$

(A3) For any $\theta, \theta_0 \in \Theta$, and $T \geq 0$,

$$\mathrm{P}_{\bar{\theta}}\left(\int_0^T \int_{\mathbb{R}_0^d} \left|\left(\frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} - 1\right) y\right| \Psi(\theta, X_{t-}, y)dydt < \infty\right) = 1,$$

$$\mathrm{P}_{\bar{\theta}}\left(\int_0^T \int_{\mathbb{R}_0^d} \left|\frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} - 1\right| \Psi(\theta, X_{t-}, y)dydt < \infty\right) = 1,$$

and

$$\mathrm{P}_{\bar{\theta}}\left(\int_0^T \int_{\mathbb{R}_0^d} \left(\frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} - 1\right)^2 \Psi(\theta, X_{t-}, y)dydt < \infty\right) = 1, \quad \text{for} \quad \bar{\theta} \in \{\theta, \theta_0\}.$$

(A4) The functions $a(\theta, x)$ and $\Psi(\theta, x, y)$ are differentiable with respect to θ, and the functions $\Psi(\theta, x, y)$ and $\nabla_\theta \Psi(\theta, x, y)$ are continuous in θ.

Under conditions **(A1)-(A3)**, by Girsanov's Theorem 2.3.1, the log-likelihood function is given by

$$\begin{aligned}\ell_T(\theta) &= \log \frac{d\mathrm{P}_\theta^T}{d\mathrm{P}_{\theta_0}^T} \\ &= \int_0^T \sigma^{-1}(X_t)\left(b(\theta,X_t)-b(\theta_0,X_t)\right)\cdot dB_t - \frac{1}{2}\int_0^T \left|\sigma^{-1}(X_t)\left(b(\theta,X_t)-b(\theta_0,X_t)\right)\right|^2 dt \\ &\quad + \int_0^T\int_{\mathbb{R}_0^d} \ln\frac{\Psi(\theta,X_{t-},y)}{\Psi(\theta_0,X_{t-},y)} N(dt,dy) - \int_0^T\int_{\mathbb{R}_0^d}\left(\frac{\Psi(\theta,X_{t-},y)}{\Psi(\theta_0,X_{t-},y)}-1\right)\mu_{\theta_0}(X_{t-},dy)dt,\end{aligned}$$

for any $\theta_0 \in \Theta$.

Therefore, by hypothesis **(A4)**, the score function is given by

$$\begin{aligned}\nabla_\theta \ell_T(\theta) = &\int_0^T \sigma^{-1}(X_t)\nabla_\theta b(\theta,X_t)\cdot\left(dB_t - \sigma^{-1}(X_t)\left(b(\theta,X_t)-b(\theta_0,X_t)\right)\right) \\ &+ \int_0^T\int_{\mathbb{R}_0^d} \nabla_\theta \ln\left(\Psi(\theta,X_{t-},y)\right)\left(N(dt,dy)-\mu_\theta(X_{t-},dy)dt\right).\end{aligned} \tag{2.7}$$

Now, by the Girsanov's Theorem, the process $W=(W_t, 0\le t\le T)$ defined as

$$W_t = B_t - \int_0^t \sigma^{-1}(X_s)\left(b(\theta,X_s)-b(\theta_0,X_s)\right)ds$$

is an $(\mathcal{F}_t, 0\le t\le T)$-Brownian motion under P_θ. Therefore, under P_θ,

$$\begin{aligned}\nabla_\theta \ell_T(\theta) = &\int_0^T \sigma^{-1}(X_t)\nabla_\theta b(\theta,X_t)\cdot dB_t \\ &+ \int_0^T\int_{\mathbb{R}_0^d} \nabla_\theta \ln\left(\Psi(\theta,X_{t-},y)\right)\left(N(dt,dy)-\mu_\theta(X_{t-},dy)dt\right),\end{aligned}$$

which is a $\mathbb{R}^k$-valued P_θ-local martingale whose quadratic variation is given by

$$\begin{aligned}\left[\nabla_\theta \ell(\theta)\right]_T = &\int_0^T (\nabla_\theta b(\theta,X_t))^{\mathsf{T}}(\sigma^{-1}(X_t))^{\mathsf{T}}\sigma^{-1}(X_t)\nabla_\theta b(\theta,X_t)dt \\ &+ \int_0^T\int_{\mathbb{R}_0^d} \left(\nabla_\theta \ln\left(\Psi(\theta,X_{t-},y)\right)\right)^{\mathsf{T}}\nabla_\theta \ln\left(\Psi(\theta,X_{t-},y)\right)N(dt,dy).\end{aligned} \tag{2.8}$$

(A5) As $T\to\infty$,

$$\mathrm{E}_\theta\left[\sup_{0\le t\le T}\left|\varphi_T(\theta)\nabla_\theta\ln(\Psi(\theta,X_{t-},\Delta X_t))\right|\right]\to 0,$$

where $\Delta X_t = X_t - X_{t-}$ denotes the jump size of X_t at time t.

(A6) There exists a $k\times k$ symmetric positive definite random matrix $\Gamma(\theta)$ such that as $T\to\infty$,

$$\varphi_T(\theta)\left[\nabla_\theta\ell(\theta)\right]_T\varphi_T(\theta)\xrightarrow{\mathrm{P}_\theta}\Gamma(\theta),$$

uniformly in $\theta\in\Theta$.

(A7) For all $u\in\mathbb{R}^k$, as $T\to\infty$,

$$\int_0^T\left|\sigma^{-1}(X_t)\left(b(\theta+\varphi_T(\theta)u,X_t)-b(\theta,X_t)-\nabla_\theta b(\theta,X_t)\varphi_T(\theta)u\right)\right|^2dt\xrightarrow{\mathrm{P}_\theta}0,$$

uniformly in $\theta\in\Theta$.

(A8) For all $u \in \mathbb{R}^k$, as $T \to \infty$,

$$\int_0^T \int_{\mathbb{R}_0^d} \left(\frac{\Psi(\theta + \varphi_T(\theta)u, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} - 1 - u^\mathsf{T} \varphi_T(\theta) \nabla_\theta \ln\left(\Psi(\theta, X_{t-}, y)\right) \right)^2 \Psi(\theta, X_{t-}, y) dy dt \xrightarrow{\mathrm{P}_\theta} 0,$$

uniformly in $\theta \in \Theta$.

We next recall a general result on central limit theorem for multivariate martingales developed by Crimaldi and Pratelli [14, Theorem 2.2], which will be important in the sequel. Several versions of the central limit theorem for multivariate martingales were given in the literature. Recall that a central limit theorem for multivariate martingales with a diagonal normalized matrix was established by Sørensen in [71, Theorem A.1], as an extension of the central limit theorem for one-dimensional martingales [18, Theorem 2] by Feigin. Later on, by applying again [18, Theorem 2], Küchler and Sørensen in [42, Theorem 2.1] established a central limit theorem for multivariate martingales with a full normalized matrix, as an extended version of [71, Theorem A.1]. More recently, Crimaldi and Pratelli in [14, Theorem 2.2] have presented a new general version of [42, Theorem 2.1] by eliminating some superfluous hypotheses and replacing the weaker assumptions, whose proof is based on a multidimensional version [14, Prop.3.1] of a convergence result for martingale difference triangular arrays proved in [51].

Theorem 2.2.2. *[14, Theorem 2.2] Let $M = (M_t)_{t\geq 0}$ be a càdlàg and $\{\mathcal{F}_t\}$-adapted k-dimensional martingale defined on the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathrm{P})$, with quadratic variation matrix $[M]$. Let $(\varphi_t)_{t\geq 0}$ be a family of $k \times k$-matrices. Suppose that the following conditions hold as $t \to \infty$,*

(i) $|\varphi_t| \to 0$, *where $|\cdot|$ denotes the sum of the absolute values of the entries of the matrix.*

(ii) $\mathrm{E}[\sup_{0\leq s\leq t} |\varphi_t \Delta M_s|] \to 0$.

(iii) *There exists a $k \times k$ positive definite random matrix U such that $\varphi_t [M]_t \varphi_t \xrightarrow{\mathrm{P}} U$.*

Then $\varphi_t M_t \xrightarrow{\mathcal{L}(\mathrm{P})} U^{1/2}\mathcal{N}(0, I_k)$ as $t \to \infty$, where $\mathcal{N}(0, I_k)$ is a centered $\mathbb{R}^k$-valued Gaussian random variable independent of U.

Remark 2.2.1. *The convergence statement of the previous theorem is established for the stable convergence, which is stronger than the convergence in law.*

We first state the asymptotic mixed normality of the score function.

Theorem 2.2.3. *Assume conditions* **(A1)**-**(A6)**. *Then, the score function is asymptotically mixed normal uniformly for all $\theta \in \Theta$ with asymptotic Fisher information matrix $\Gamma(\theta)$. That is, as $T \to \infty$,*

$$\varphi_T(\theta) \nabla_\theta \ell_T(\theta) \xrightarrow{\mathcal{L}(\mathrm{P}_\theta)} \Gamma(\theta)^{1/2} \mathcal{N}\left(0, I_k\right),$$

uniformly in $\theta \in \Theta$, where $\mathcal{N}(0, I_k)$ is a centered $\mathbb{R}^k$-valued Gaussian random variable independent of $\Gamma(\theta)$.

Proof. Observe that for any $t \in [0, T]$,

$$\Delta \nabla_\theta \ell_t(\theta) = \nabla_\theta \ell_t(\theta) - \nabla_\theta \ell_{t-}(\theta) = \nabla_\theta \ln\left(\Psi(\theta, X_{t-}, \Delta X_t)\right) \mathbf{1}_{\{\Delta X_t \neq 0\}}.$$

Then, from the fact that $|\varphi_T(\theta)| \to 0$ as $T \to \infty$ and hypotheses **(A5)**-**(A6)**, the conditions of Theorem 2.2.2 are satisfied for the local martingale $\nabla_\theta \ell_T(\theta)$. Thus, the result follows. □

We next state the LAMN property for the jump-diffusion process solution to (2.1) on the time interval $[0, T]$.

Theorem 2.2.4. *Assume conditions* **(A1)**-**(A8)**. *Then, the LAMN property holds uniformly for all $\theta \in \Theta$ with rate of convergence $\varphi_T^{-1}(\theta)$ and asymptotic Fisher information matrix $\Gamma(\theta)$. That is, for all $u \in \mathbb{R}^k$, as $T \to \infty$,*

$$\log \frac{d\mathrm{P}^T_{\theta+\varphi_T(\theta)u}}{d\mathrm{P}^T_\theta} \overset{\mathcal{L}(\mathrm{P}_\theta)}{\longrightarrow} u^\top \Gamma(\theta)^{1/2}\mathcal{N}(0, I_k) - \frac{1}{2} u^\top \Gamma(\theta) u,$$

uniformly in $\theta \in \Theta$, where $\mathcal{N}(0, I_k)$ is a centered $\mathbb{R}^k$-valued Gaussian random variable independent of $\Gamma(\theta)$.

Remark 2.2.2. *We observe that conditions* **(A6)**, **(A7)** *and* **(A8)** *are the same as conditions* **(L)**, **(D.1)** *and* **(D.3)** *of Luschgy [52], respectively. Furthermore, notice here that condition* **(A5)** *is weaker than* **(J.1)** *of Luschgy [52], and condition* **(R)** *of Luschgy [52] is not needed. This is because we are applying the central limit theorem for multivariate martingales generalized by Crimaldi and Pratelli [14, Theorem 2.2] as mentioned above.*

Proof. Fix $u \in \mathbb{R}^k$ and $\theta \in \Theta$, and apply Girsanov's Theorem 2.3.1 with $\theta_0 = \theta + \varphi_T(\theta)u$, to get that

$$\log \frac{d\mathrm{P}^T_{\theta_0}}{d\mathrm{P}^T_\theta} = L^c_T + L^d_T,$$

where the continuous part L^c_T and the discontinuous part L^d_T are respectively given by

$$L^c_T = \int_0^T \sigma^{-1}(X_t) \left(b(\theta_0, X_t) - b(\theta, X_t)\right) \cdot dB_t - \frac{1}{2} \int_0^T |\sigma^{-1}(X_t) \left(b(\theta_0, X_t) - b(\theta, X_t)\right)|^2 dt,$$

$$L^d_T = \int_0^T \int_{\mathbb{R}^d_0} \ln \frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} N(dt, dy) - \int_0^T \int_{\mathbb{R}^d_0} \left(\frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} - 1 \right) \mu_\theta(X_{t-}, dy) dt.$$

Adding and substracting the vector $\nabla_\theta b(\theta, X_t)\varphi_T(\theta)u$ in the continuous part L^c_T, and adding and substracting the terms $u^\top \varphi_T(\theta) \nabla_\theta \ln\left(\Psi(\theta, X_{t-}, y)\right)$ and $\frac{1}{2}\left(u^\top \varphi_T(\theta) \nabla_\theta \ln\left(\Psi(\theta, X_{t-}, y)\right)\right)^2$ in the discontinuous part L^d_T, we obtain the following expansion of the log-likelihood ratio

$$\begin{aligned}
\log \frac{d\mathrm{P}^T_{\theta_0}}{d\mathrm{P}^T_\theta} &= u^\top \varphi_T(\theta) \nabla_\theta \ell_T(\theta) - \frac{1}{2} u^\top \varphi_T(\theta) \left[\nabla_\theta \ell(\theta)\right]_T \varphi_T(\theta) u \\
&+ \int_0^T \sigma^{-1}(X_t) \left(b(\theta_0, X_t) - b(\theta, X_t) - \nabla_\theta b(\theta, X_t)\varphi_T(\theta)u\right) \cdot dB_t \\
&- \frac{1}{2} \int_0^T |\sigma^{-1}(X_t) \left(b(\theta_0, X_t) - b(\theta, X_t) - \nabla_\theta b(\theta, X_t)\varphi_T(\theta)u\right)|^2 dt \\
&- \int_0^T u^\top \varphi_T(\theta) \left(\nabla_\theta b(\theta, X_t)\right)^\top \left(\sigma^{-1}(X_t)\right)^\top \sigma^{-1}(X_t) \left(b(\theta_0, X_t) - b(\theta, X_t) - \nabla_\theta b(\theta, X_t)\varphi_T(\theta)u\right) dt \\
&+ \int_0^T \int_{\mathbb{R}^d_0} \left(\frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} - 1 - u^\top \varphi_T(\theta) \nabla_\theta \ln\left(\Psi(\theta, X_{t-}, y)\right) \right) \left(N(dt, dy) - \mu_\theta(X_{t-}, dy)dt\right) \\
&+ \int_0^T \int_{\mathbb{R}^d_0} \left(\ln \frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} - \frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} + 1 + \frac{1}{2} \left(u^\top \varphi_T(\theta) \nabla_\theta \ln\left(\Psi(\theta, X_{t-}, y)\right)\right)^2 \right) N(dt, dy),
\end{aligned}$$

where the P_θ-local martingale $\nabla_\theta \ell_T(\theta)$ is the term that contributes to the limit. In fact, using Theorem 2.2.3 and hypothesis **(A6)**, we get that as $T \to \infty$,

$$u^\top \varphi_T(\theta) \nabla_\theta \ell_T(\theta) - \frac{1}{2} u^\top \varphi_T(\theta) \left[\nabla_\theta \ell(\theta)\right]_T \varphi_T(\theta) u \overset{\mathcal{L}(\mathrm{P}_\theta)}{\longrightarrow} u^\top \Gamma(\theta)^{1/2} \mathcal{N}(0, I_k) - \frac{1}{2} u^\top \Gamma(\theta) u.$$

We next treat the negligible contributions. By hypothesis **(A7)**, the quadratic variation of the local martingale

$$\int_0^T \sigma^{-1}(X_t) \left(b(\theta_0, X_t) - b(\theta, X_t) - \nabla_\theta b(\theta, X_t)\varphi_T(\theta)u\right) \cdot dB_t$$

tends to zero in P_θ-probability as $T \to \infty$ uniformly in $\theta \in \Theta$. Thus, so does the local martingale.

Using the Cauchy-Schwarz inequality and hypotheses **(A6)** and **(A7)**, we get that as $T \to \infty$,

$$\left| \int_0^T u^{\mathsf{T}} \varphi_T(\theta) \left(\nabla_\theta b(\theta, X_t)\right)^{\mathsf{T}} (\sigma^{-1}(X_t))^{\mathsf{T}} \sigma^{-1}(X_t) \left(b(\theta_0, X_t) - b(\theta, X_t) - \partial_\theta b(\theta, X_t) \varphi_T(\theta) u\right) dt \right|$$
$$\leq \left(\int_0^T |\sigma^{-1}(X_t) \nabla_\theta b(\theta, X_t) \varphi_T(\theta) u|^2 dt \right)^{1/2}$$
$$\times \left(\int_0^T |\sigma^{-1}(X_t) \left(b(\theta_0, X_t) - b(\theta, X_t) - \nabla_\theta b(\theta, X_t) \varphi_T(\theta) u\right)|^2 dt \right)^{1/2} \xrightarrow{\mathrm{P}_\theta} 0.$$

By hypothesis **(A8)**, the quadratic characteristic of the local martingale

$$\int_0^T \int_{\mathbb{R}_0^d} \left(\frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} - 1 - u^{\mathsf{T}} \varphi_T(\theta) \nabla_\theta \ln\left(\Psi(\theta, X_{t-}, y)\right) \right) \left(N(dt, dy) - \mu_\theta(X_{t-}, dy) dt\right)$$

tends to zero in P_θ-probability as $T \to \infty$ uniformly in $\theta \in \Theta$. Thus, so does the local martingale.

Finally, appealing to Lemma 2.2.1 below, we conclude the desired proof. □

Lemma 2.2.1. *Assume that the function $\Psi(\theta, x, y)$ is differentiable with respect to θ and that hypotheses* **(A1)**, **(A5)**, **(A6)**, *and* **(A8)** *hold. Then, as $T \to \infty$,*

$$\int_0^T \int_{\mathbb{R}_0^d} \left(\ln \frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} - \frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} + 1 + \frac{1}{2} \left(u^{\mathsf{T}} \varphi_T(\theta) \nabla_\theta \ln\left(\Psi(\theta, X_{t-}, y)\right)\right)^2 \right) N(dt, dy)$$

tends to zero in P_θ-probability uniformly in $\theta \in \Theta$.

Proof. Consider the function $f(y-1) = \ln(y) - (y-1) + \frac{1}{2}(y-1)^2$ defined for all $y > 0$. Then, for all x,

$$\ln(y) - y + 1 + \frac{1}{2} x^2 = f(y-1) - \frac{1}{2}((y-1)^2 - x^2).$$

Therefore,

$$\int_0^T \int_{\mathbb{R}_0^d} \left(\ln \frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} - \frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} + 1 + \frac{1}{2} \left(u^{\mathsf{T}} \varphi_T(\theta) \nabla_\theta \ln\left(\Psi(\theta, X_{t-}, y)\right)\right)^2 \right) N(dt, dy)$$
$$= \int_0^T \int_{\mathbb{R}_0^d} f\left(\frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} - 1 \right) N(dt, dy)$$
$$- \frac{1}{2} \int_0^T \int_{\mathbb{R}_0^d} \left\{ \left(\frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} - 1 \right)^2 - \left(u^{\mathsf{T}} \varphi_T(\theta) \nabla_\theta \ln\left(\Psi(\theta, X_{t-}, y)\right)\right)^2 \right\} N(dt, dy).$$

Now, from [33, Proposition II.1.14], for any $T \geq 0$,

$$\mathrm{E}_\theta \left[\left| \int_0^T \int_{\mathbb{R}_0^d} \left(\frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} - 1 - u^{\mathsf{T}} \varphi_T(\theta) \nabla_\theta \ln\left(\Psi(\theta, X_{t-}, y)\right) \right)^2 N(dt, dy) \right| \right]$$
$$= \mathrm{E}_\theta \left[\int_0^T \int_{\mathbb{R}_0^d} \left(\frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} - 1 - u^{\mathsf{T}} \varphi_T(\theta) \nabla_\theta \ln\left(\Psi(\theta, X_{t-}, y)\right) \right)^2 \mu_\theta(X_{t-}, dy) dt \right].$$

Therefore, the process $\int_0^T \int_{\mathbb{R}_0^d} \left(\frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} - 1 - u^{\mathsf{T}} \varphi_T(\theta) \nabla_\theta \ln(\Psi(\theta, X_{t-}, y)) \right)^2 N(dt, dy)$ is dominated in the sense of Lenglart by its compensator process

$$\int_0^T \int_{\mathbb{R}_0^d} \left(\frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} - 1 - u^{\mathsf{T}} \varphi_T(\theta) \nabla_\theta \ln\left(\Psi(\theta, X_{t-}, y)\right) \right)^2 \mu_\theta(X_{t-}, dy) dt,$$

for any $T \geq 0$. Thus, by Lenglart's inequality [33, Lemma I.3.30 a)], we have that for all $T \geq 0$ and $\epsilon, \eta > 0$,

$$\begin{aligned}
&\mathrm{P}_\theta\left(\left|\int_0^T \int_{\mathbb{R}_0^d} \left(\frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} - 1 - u^\top \varphi_T(\theta) \nabla_\theta \ln\left(\Psi(\theta, X_{t-}, y)\right)\right)^2 N(dt, dy)\right| \geq \epsilon\right) \\
&\leq \frac{\eta}{\epsilon} + \mathrm{P}_\theta\left(\int_0^T \int_{\mathbb{R}_0^d} \left(\frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} - 1 - u^\top \varphi_T(\theta) \nabla_\theta \ln\left(\Psi(\theta, X_{t-}, y)\right)\right)^2 \mu_\theta(X_{t-}, dy) dt \geq \eta\right),
\end{aligned}$$

which, by hypothesis **(A8)**, implies that for all $u \in \mathbb{R}^k$, as $T \to \infty$,

$$\int_0^T \int_{\mathbb{R}_0^d} \left(\frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} - 1 - u^\top \varphi_T(\theta) \nabla_\theta \ln\left(\Psi(\theta, X_{t-}, y)\right)\right)^2 N(dt, dy) \xrightarrow{\mathrm{P}_\theta} 0, \tag{2.9}$$

uniformly in $\theta \in \Theta$. Thus, from hypothesis **(A6)** and the equality $a^2 - b^2 = (a-b)^2 + 2b(a-b)$, we conclude that for all $u \in \mathbb{R}^k$, as $T \to \infty$,

$$\int_0^T \int_{\mathbb{R}_0^d} \left\{\left(\frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} - 1\right)^2 - \left(u^\top \varphi_T(\theta) \nabla_\theta \ln\left(\Psi(\theta, X_{t-}, y)\right)\right)^2\right\} N(dt, dy) \xrightarrow{\mathrm{P}_\theta} 0, \tag{2.10}$$

uniformly in $\theta \in \Theta$.

We next show that for every $\epsilon > 0$, as $T \to \infty$,

$$\int_0^T \int_{\mathbb{R}_0^d} f\left(\frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} - 1\right) \mathbf{1}_{\left\{\left|\frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} - 1\right| > \epsilon\right\}} N(dt, dy) \xrightarrow{\mathrm{P}_\theta} 0, \tag{2.11}$$

uniformly in $\theta \in \Theta$.

For all $a > 0$,

$$\begin{aligned}
&\left\{\left|\int_0^T \int_{\mathbb{R}_0^d} f\left(\frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} - 1\right) \mathbf{1}_{\left\{\left|\frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} - 1\right| > \epsilon\right\}} N(dt, dy)\right| > a\right\} \\
&\subset \left\{\int_0^T \int_{\mathbb{R}_0^d} \mathbf{1}_{\left\{\left|\frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} - 1\right| > \epsilon\right\}} N(dt, dy) \geq 1\right\}.
\end{aligned}$$

Therefore, in order to prove (2.11), it suffices to show that for every $\epsilon > 0$, as $T \to \infty$,

$$\int_0^T \int_{\mathbb{R}_0^d} \mathbf{1}_{\left\{\left|\frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} - 1\right| > \epsilon\right\}} N(dt, dy) \xrightarrow{\mathrm{P}_\theta} 0, \tag{2.12}$$

uniformly in $\theta \in \Theta$. For this, we write

$$\begin{aligned}
&\int_0^T \int_{\mathbb{R}_0^d} \mathbf{1}_{\left\{\left|\frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} - 1\right| > \epsilon\right\}} N(dt, dy) \\
&\leq \int_0^T \int_{\mathbb{R}_0^d} \mathbf{1}_{\left\{\left|u^\mathsf{T} \varphi_T(\theta) \nabla_\theta \ln(\Psi(\theta, X_{t-}, y))\right| > \frac{\epsilon}{2}\right\}} N(dt, dy) \\
&\quad + \int_0^T \int_{\mathbb{R}_0^d} \mathbf{1}_{\left\{\left|\frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} - 1 - u^\mathsf{T} \varphi_T(\theta) \nabla_\theta \ln(\Psi(\theta, X_{t-}, y))\right| > \frac{\epsilon}{2}\right\}} N(dt, dy) \\
&\leq \frac{4}{\epsilon^2} \int_0^T \int_{\mathbb{R}_0^d} \left(u^\mathsf{T} \varphi_T(\theta) \nabla_\theta \ln\left(\Psi(\theta, X_{t-}, y)\right)\right)^2 \mathbf{1}_{\left\{\left|u^\mathsf{T} \varphi_T(\theta) \nabla_\theta \ln(\Psi(\theta, X_{t-}, y))\right| > \frac{\epsilon}{2}\right\}} N(dt, dy) \\
&\quad + \frac{4}{\epsilon^2} \int_0^T \int_{\mathbb{R}_0^d} \left(\frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} - 1 - u^\mathsf{T} \varphi_T(\theta) \nabla_\theta \ln\left(\Psi(\theta, X_{t-}, y)\right)\right)^2 \\
&\qquad \times \mathbf{1}_{\left\{\left|\frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} - 1 - u^\mathsf{T} \varphi_T(\theta) \nabla_\theta \ln(\Psi(\theta, X_{t-}, y))\right| > \frac{\epsilon}{2}\right\}} N(dt, dy).
\end{aligned}$$

Now observe that hypothesis **(A5)** implies that as $T \to \infty$,

$$\sup_{0 \leq t \leq T} \left|u^\mathsf{T} \varphi_T(\theta) \nabla_\theta \ln(\Psi(\theta, X_{t-}, \Delta X_t))\right| \mathbf{1}_{\{\Delta X_t \neq 0\}} \xrightarrow{\mathrm{P}_\theta} 0,$$

which is equivalent to for every $\epsilon > 0$, as $T \to \infty$,

$$\int_0^T \int_{\mathbb{R}_0^d} \left(u^\mathsf{T} \varphi_T(\theta) \nabla_\theta \ln\left(\Psi(\theta, X_{t-}, y)\right)\right)^2 \mathbf{1}_{\left\{\left|u^\mathsf{T} \varphi_T(\theta) \nabla_\theta \ln(\Psi(\theta, X_{t-}, y))\right| > \epsilon\right\}} N(dt, dy) \xrightarrow{\mathrm{P}_\theta} 0.$$

Here we have used the fact that for $0 < a < \epsilon^2$,

$$\begin{aligned}
&\left\{\int_0^T \int_{\mathbb{R}_0^d} \left(u^\mathsf{T} \varphi_T(\theta) \nabla_\theta \ln\left(\Psi(\theta, X_{t-}, y)\right)\right)^2 \mathbf{1}_{\left\{\left|u^\mathsf{T} \varphi_T(\theta) \nabla_\theta \ln(\Psi(\theta, X_{t-}, y))\right| > \epsilon\right\}} N(dt, dy) > a\right\} \\
&= \left\{\sup_{0 \leq t \leq T} \left|u^\mathsf{T} \varphi_T(\theta) \nabla_\theta \ln(\Psi(\theta, X_{t-}, \Delta X_t))\right| \mathbf{1}_{\{\Delta X_t \neq 0\}} > \epsilon\right\}.
\end{aligned}$$

This, together with (2.9), gives (2.12), and hence (2.11).

Now, since $|f(x)| \leq 2|x|^3$ if $|x| \leq \frac{1}{2}$, we have for every $0 < \epsilon \leq \frac{1}{2}$,

$$\begin{aligned}
&\left|\int_0^T \int_{\mathbb{R}_0^d} f\left(\frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} - 1\right) \mathbf{1}_{\left\{\left|\frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} - 1\right| \leq \epsilon\right\}} N(dt, dy)\right| \\
&\leq 2\left|\int_0^T \int_{\mathbb{R}_0^d} \left|\frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} - 1\right|^3 \mathbf{1}_{\left\{\left|\frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} - 1\right| \leq \epsilon\right\}} N(dt, dy)\right| \\
&\leq 2\epsilon\left|\int_0^T \int_{\mathbb{R}_0^d} \left(\frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} - 1\right)^2 N(dt, dy)\right|.
\end{aligned}$$

Thus, from hypothesis **(A6)** and (2.10), we conclude that for every $\epsilon > 0$, as $T \to \infty$,

$$\int_0^T \int_{\mathbb{R}_0^d} f\left(\frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} - 1\right) \mathbf{1}_{\left\{\left|\frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} - 1\right| \leq \epsilon\right\}} N(dt, dy) \xrightarrow{\mathrm{P}_\theta} 0,$$

uniformly in $\theta \in \Theta$, which finishes the desired proof. □

We end this section with an important consequence of the LAMN property, which is the conditional convolution theorem.

First, recall that a family of estimators $(\tilde{\theta}_T)_{T\geq 0}$ of the parameter θ is called regular at θ if for any $u \in \mathbb{R}^k$, as $T \to \infty$,

$$\varphi_T^{-1}(\theta)\left(\tilde{\theta}_T - (\theta + \varphi_T(\theta)u)\right) \overset{\mathcal{L}(\mathrm{P}_{\theta+\varphi_T(\theta)u})}{\longrightarrow} V(\theta),$$

for some $\mathbb{R}^k$-valued random variable $V(\theta)$, independent of u.

Note that taking $u=0$, this implies that as $T \to \infty$,

$$\varphi_T^{-1}(\theta)\left(\tilde{\theta}_T - \theta\right) \overset{\mathcal{L}(\mathrm{P}_\theta)}{\longrightarrow} V(\theta).$$

The conditional convolution theorem says that when the LAMN property holds, then the asymptotic distribution of any regular family of estimators of the parameter θ is characterized by a conditional convolution between a Gaussian law and some others laws. More precisely,

Theorem 2.2.5 (Conditional convolution theorem)**.** [28, Theorem 9.1] *Suppose that the family of probability measures $(\mathrm{P}_\theta^T)_{\theta\in\Theta}$ satisfies the LAMN property at a point θ. Let $(\tilde{\theta}_T)_{T\geq 0}$ be a regular family of estimators of the parameter θ. Then the law of $V(\theta)$ conditionally on $\Gamma(\theta)$ is a convolution between $\mathcal{N}\left(0,\Gamma(\theta)^{-1}\right)$ and some other law $G_{\Gamma(\theta)}$ on $\mathbb{R}^k$, that is,*

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

where $G_{\Gamma(\theta)}$ is the limiting distribution law under P_θ of the difference

$$\varphi_T^{-1}(\theta)\left(\tilde{\theta}_T - \theta\right) - \Gamma(\theta)^{-1}\varphi_T(\theta)\nabla_\theta \ell_T(\theta),$$

as $T \to \infty$, that is,

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

The proof of this theorem uses the change of measure $\theta_0 = \theta + \varphi_T(\theta)u$, which from the LAMN property can be written as

$$\mathrm{E}_{\theta_0}[Z] = \mathrm{E}_\theta\left[Z\frac{d\mathrm{P}_{\theta_0}^T}{d\mathrm{P}_\theta^T}\right] = \mathrm{E}_\theta\left[Ze^{u^\mathsf{T}\varphi_T(\theta)\nabla_\theta\ell_T(\theta) - \frac{1}{2}u^\mathsf{T}\varphi_T(\theta)[\nabla_\theta\ell(\theta)]_T\varphi_T(\theta)u + o_{\mathrm{P}_\theta}(1)}\right],$$

for some random variable Z.

Moreover, the proof shows that the random variable $V(\theta)$ can be written as a sum of two independent random variables

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

where R is a random variable with distribution $G_{\Gamma(\theta)}$. This implies that, under the conditions of Theorem 2.2.5, as $T \to \infty$,

$$\varphi_T^{-1}(\theta)\left(\tilde{\theta}_T - \theta\right) \overset{\mathcal{L}(\mathrm{P}_\theta)}{\longrightarrow} \Gamma(\theta)^{-1/2}\mathcal{N}(0,I_k) + R.$$

This theorem suggests the notion of asymptotically efficient estimators, when $R=0$. That is, a family of estimators $(\tilde{\theta}_T)_{T\geq 0}$ of the parameter θ is called asymptotically efficient at θ if as $T \to \infty$,

$$\varphi_T^{-1}(\theta)\left(\tilde{\theta}_T - \theta\right) \overset{\mathcal{L}(\mathrm{P}_\theta)}{\longrightarrow} \Gamma(\theta)^{-1/2}\mathcal{N}(0,I_k),$$

where $\Gamma(\theta)$ and $\mathcal{N}(0,I_k)$ are independent.

2.3 Girsanov's Theorem

This section is devoted to recall Girsanov's theorem for the diffusion process with jumps (2.1), which is needed in the proof of Theorem 2.2.4. Recall that in [71], Sørensen deals with a more general diffusion process with jumps where the dimension of the space on which the jumps of the Poisson random measure are defined can be different from that of the process. The author gives sufficient conditions for the equivalence of all probability measures and then derives a complicated expression of the Radon-Nikodym derivative (see [71, Theorem 2.1]). The author applies Girsanov's theorem for semimartingales proved by Jacod and Mémin (see [32, Theorem 4.2 and 4.5(b)]). These results are based on the uniqueness of the representation of semimartingales in terms of their local characteristics, and the uniqueness of the solution to the martingale problem associated to this semimartingale. Recall that the diffusion process with jumps X^θ solving (2.1) is a semimartingale, and a weak solution to (2.1) is a solution to the martingale problem associated to X^θ. Furthermore, the set of all weak solutions to (2.1) is the set of all solutions to the martingale problem on the canonical space associated to X^θ. In our context, the proof of Girsanov's theorem below can be based on the uniqueness of the weak solution to equation (2.1).

Finally, Jacod and Shiryaev in [33, Theorem III.3.24 and III.5.19] extend Girsanov's Jacod and Mémin theorem to the multidimensional case. We also refer to [36], [37], [29], [66] for the Girsanov's theorem for semimartingales, multivariate point processes and discontinuous independent increments processes.

Theorem 2.3.1 (Girsanov's theorem). *Assume conditions* **(A1)**-**(A3)**. *Then for all* $\theta, \theta_0 \in \Theta$, *the probability measures* P_θ^T *and* $\mathrm{P}_{\theta_0}^T$ *are equivalent. Furthermore, their Radon-Nikodym derivative is given by*

$$\frac{d\mathrm{P}_{\theta_0}^T}{d\mathrm{P}_\theta^T}$$
$$= \exp\left\{ \int_0^T \sigma^{-1}(X_t)\left(b(\theta_0, X_t) - b(\theta, X_t)\right) \cdot dB_t - \frac{1}{2}\int_0^T \left|\sigma^{-1}(X_t)\left(b(\theta_0, X_t) - b(\theta, X_t)\right)\right|^2 dt \right.$$
$$\left. + \int_0^T \int_{\mathbb{R}_0^d} \ln \frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} N(dt, dy) - \int_0^T \int_{\mathbb{R}_0^d} \left(\frac{\Psi(\theta_0, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} - 1 \right) \mu_\theta(X_{t-}, dy) dt \right\}.$$

2.4 LAN property for ergodic diffusion processes with jumps

In this section, we seek sufficient conditions in order for the LAN property to hold when the diffusion process with jumps X^θ (2.1) is ergodic, as a consequence of Theorem 2.2.4.

Let $X^\theta = (X_t)_{t\geq 0}$ be the solution to equation (2.1), that is,

$$X_t = X_0 + \int_0^t a(\theta, X_s) ds + \int_0^t \sigma(X_s) dB_s + \int_0^t \int_{\mathbb{R}_0^d} c(\theta, X_{s-}, z) \left(p(ds, dz) - \nu_\theta(dz) ds \right).$$

Recall that we have rewritten this equation as

$$X_t = X_0 + \int_0^t b(\theta, X_s) ds + \int_0^t \sigma(X_s) dB_s + \int_0^t \int_{\mathbb{R}_0^d} y N(dt, dy),$$

where $b(\theta, X_t) = a(\theta, X_t) - \int_{\mathbb{R}_0^d} y \mu_\theta(X_{t-}, dy)$, and $N(dt, dy)$ is a jump measure on $\mathbb{R}_+ \times \mathbb{R}_0^d$ with predictable compensator $\mu_\theta(X_{t-}, dy)dt = \Psi(\theta, X_{t-}, y) dy dt$.

As is well-known, X^θ is a homogeneous Markov process (see [3, Theorem 6.4.6]). Let us introduce the ergodic assumption.

(C1) The process X^θ is ergodic in the sense that there exists a unique probability measure $\pi_\theta(dx)$ such that as $T \to \infty$,

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

for any π_θ-integrable function g.

Several examples of ergodic diffusion processes with jumps are given in [53], [54], and [67]. Moreover, results on ergodicity and exponential ergodicity of diffusion processes with jumps have been established by Masuda in [53, 54]. In addition, Kulik in [43] provides a set of sufficient conditions for the exponential ergodicity of diffusion processes with jumps without Gaussian part and gives some examples. More recently, Qiao in [64] has addressed the exponential ergodicity for stochastic differential equations with jumps and non-Lipschitz coefficients. However, in these papers ergodicity and exponentially ergodicity are understood in the sense of [56], which both are stronger than the ergodicity in the sense **(C1)**.

We next show that if a process satisfies the additional Lindeberg condition **(A9)**, then the quadratic characteristic and quadratic variation of the score function are asymptotically equivalent at rate $\varphi_T(\theta)$.

(A9) For all $\epsilon > 0$ and $u \in \mathbb{R}^k$, as $T \to \infty$,

$$\int_0^T \int_{\mathbb{R}_0^d} \Big(u^\mathsf{T}\varphi_T(\theta)\nabla_\theta \ln\left(\Psi(\theta, X_{t-}, y)\right)\Big)^2 \mathbf{1}_{\{|u^\mathsf{T}\varphi_T(\theta)\nabla_\theta \ln(\Psi(\theta,X_{t-},y))|>\epsilon\}}\Psi(\theta, X_{t-}, y)dydt \xrightarrow{\mathrm{P}_\theta} 0,$$

uniformly in $\theta \in \Theta$.

Lemma 2.4.1. *Assume conditions* **(A1)**-**(A4)** *and* **(A9)**. *Then, condition* **(A6)** *is equivalent to the fact that there exists a $k \times k$ symmetric positive definite random matrix $\Gamma(\theta)$ such that as $T \to \infty$,*

$$\varphi_T(\theta)\left\langle \nabla_\theta \ell(\theta)\right\rangle_T \varphi_T(\theta) \xrightarrow{\mathrm{P}_\theta} \Gamma(\theta), \tag{2.13}$$

uniformly in $\theta \in \Theta$, where $\langle \nabla_\theta \ell(\theta)\rangle_T$ is the quadratic characteristic of the score function, that is,

$$\begin{aligned}\langle \nabla_\theta \ell(\theta)\rangle_T &= \int_0^T (\nabla_\theta b(\theta, X_t))^\mathsf{T}(\sigma^{-1}(X_t))^\mathsf{T}\sigma^{-1}(X_t)\nabla_\theta b(\theta, X_t)dt \\ &\quad + \int_0^T\int_{\mathbb{R}_0^d} (\nabla_\theta \ln\left(\Psi(\theta, X_{t-}, y)\right))^\mathsf{T}\nabla_\theta \ln\left(\Psi(\theta, X_{t-}, y)\right)\Psi(\theta, X_{t-}, y)dydt.\end{aligned}$$

Next, observe that the ergodicity assumption implies the convergence of the quadratic characteristic of the score function at rate $\frac{1}{T}$.

Lemma 2.4.2. *Assume conditions* **(A1)**-**(A4)** *and* **(C1)**. *Then, as $T \to \infty$,*

$$\frac{1}{T}\left\langle \nabla_\theta \ell(\theta)\right\rangle_T \xrightarrow{\mathrm{P}_\theta} \Gamma(\theta), \tag{2.14}$$

uniformly in $\theta \in \Theta$, where

$$\begin{aligned}\Gamma(\theta) &= \int_{\mathbb{R}^d} (\nabla_\theta b(\theta, x))^\mathsf{T}(\sigma^{-1}(x))^\mathsf{T}\sigma^{-1}(x)\nabla_\theta b(\theta, x)\pi_\theta(dx) \\ &\quad + \int_{\mathbb{R}^d}\int_{\mathbb{R}_0^d} \nabla_\theta \ln \Psi(\theta, x, y)\left(\nabla_\theta \ln \Psi(\theta, x, y)\right)^\mathsf{T}\Psi(\theta, x, y)dy\pi_\theta(dx).\end{aligned}$$

Therefore, we have the following immediate consequence of Theorem 2.2.3.

Theorem 2.4.1. *Suppose that conditions* **(A1)**-**(A5)**, **(A9)**, *and* **(C1)** *are satisfied with $\varphi_T(\theta)$ the diagonal matrix with entries equal to $\frac{1}{\sqrt{T}}$. Then the score function is asymptotically normal uniformly for all $\theta \in \Theta$ with asymptotic Fisher information matrix $\Gamma(\theta)$. That is, as $T \to \infty$,*

$$\frac{1}{\sqrt{T}} \nabla_\theta \ell_T(\theta) \overset{\mathcal{L}(\mathrm{P}_\theta)}{\longrightarrow} \mathcal{N}\left(0, \Gamma(\theta)\right),$$

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

We next derive the LAN property. For this, we need the following additional assumptions.
(C2) For all $u \in \mathbb{R}^k$, as $T \to \infty$,

$$\int_{\mathbb{R}^d} \left| \sigma^{-1}(x) \left(b\left(\theta + \frac{u}{\sqrt{T}}, x\right) - b(\theta, x) - \nabla_\theta b(\theta, x) \frac{u}{\sqrt{T}} \right) \right|^2 \pi_\theta(dx) = o\left(\frac{1}{T}\right),$$

uniformly in $\theta \in \Theta$.

(C3) For all $u \in \mathbb{R}^k$, as $T \to \infty$,

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}_0^d} \left(\frac{\Psi(\theta + \frac{u}{\sqrt{T}}, x, y)}{\Psi(\theta, x, y)} - 1 - \frac{u^\mathsf{T}}{\sqrt{T}} \nabla_\theta \ln\left(\Psi(\theta, x, y)\right) \right)^2 \Psi(\theta, x, y) dy \pi_\theta(dx) = o\left(\frac{1}{T}\right),$$

uniformly in $\theta \in \Theta$.
We next state the main result of this section.

Theorem 2.4.2. *Suppose that conditions* **(A1)**-**(A5)**, **(A9)**, *and* **(C1)**-**(C3)** *are fulfilled with $\varphi_T(\theta)$ the diagonal matrix with entries equal to $\frac{1}{\sqrt{T}}$. Then the LAN property holds for all $\theta \in \Theta$ with rate of convergence $\sqrt{T}$ and asymptotic Fisher information matrix $\Gamma(\theta)$. That is, for all $u \in \mathbb{R}^k$, as $T \to \infty$,*

$$\log \frac{d\mathrm{P}^T_{\theta + \frac{u}{\sqrt{T}}}}{d\mathrm{P}^T_\theta} \overset{\mathcal{L}(\mathrm{P}_\theta)}{\longrightarrow} u^\mathsf{T} \mathcal{N}\left(0, \Gamma(\theta)\right) - \frac{1}{2} u^\mathsf{T} \Gamma(\theta) u.$$

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

$$\begin{aligned} &\frac{1}{T} \int_0^T \left| \sigma^{-1}(X_t) \left(b(\theta + \frac{u}{\sqrt{T}}, X_t) - b(\theta, X_t) - \nabla_\theta b(\theta, X_t) \frac{u}{\sqrt{T}} \right) \right|^2 dt \\ &\overset{\mathrm{P}_\theta}{\longrightarrow} \int_{\mathbb{R}^d} \left| \sigma^{-1}(x) \left(b(\theta + \frac{u}{\sqrt{T}}, x) - b(\theta, x) - \nabla_\theta b(\theta, x) \frac{u}{\sqrt{T}} \right) \right|^2 \pi_\theta(dx), \end{aligned}$$

which, together with **(C2)** gives **(A7)**.
Again by ergodicity, as $T \to \infty$,

$$\begin{aligned} &\frac{1}{T} \int_0^T \int_{\mathbb{R}_0^d} \left(\frac{\Psi(\theta + \frac{u}{\sqrt{T}}, X_{t-}, y)}{\Psi(\theta, X_{t-}, y)} - 1 - \frac{1}{\sqrt{T}} u^\mathsf{T} \nabla_\theta \ln\left(\Psi(\theta, X_{t-}, y)\right) \right)^2 \Psi(\theta, X_{t-}, y) dy dt \\ &\overset{\mathrm{P}_\theta}{\longrightarrow} \int_{\mathbb{R}^d} \int_{\mathbb{R}_0^d} \left(\frac{\Psi(\theta + \frac{u}{\sqrt{T}}, x, y)}{\Psi(\theta, x, y)} - 1 - \frac{1}{\sqrt{T}} u^\mathsf{T} \nabla_\theta \ln\left(\Psi(\theta, x, y)\right) \right)^2 \Psi(\theta, x, y) dy \pi_\theta(dx), \end{aligned}$$

which, together with **(C3)** gives **(A8)**.
Then, the desired LAN property follows from Lemma 2.4.2 and Theorem 2.2.4. □

As a consequence of the LAN property, an asymptotic lower bound for the variance of any family of unbiased estimators can be obtained. More precisely,

Theorem 2.4.3 (Minimax theorem)**.** [28, Theorem 12.1] *Suppose that the family of probability measures* $(\mathrm{P}_\theta^T)_{\theta\in\Theta}$ *satisfies the LAN property at a point* θ. *Let* $(\tilde{\theta}_T)_{T\geq 0}$ *be a family of estimators of the parameter* θ *and* $l : \mathbb{R}^k \to [0, +\infty)$ *be a loss function of the form* $l(0) = 0, l(x) = l(|x|)$ *and* $l(|x|) \leq l(|y|)$ *if* $|x| \leq |y|$. *Then*

$$\lim_{\delta\to 0} \liminf_{T\to\infty} \sup_{|\theta'-\theta|<\delta} \mathrm{E}_{\theta'} \left[l\left(\varphi_T^{-1}(\theta)\left(\tilde{\theta}_T - \theta'\right)\right)\right] \geq \mathrm{E}_\theta \left[l\left(Z\right)\right],$$

where $\mathcal{L}(Z) = \mathcal{N}(0, \Gamma(\theta)^{-1})$.

In particular, when we take the quadratic loss function $l(u) = |u|^2$, the above inequality gives an asymptotic lower bound for the covariance matrix of any family of unbiased estimators, which is given by $\Gamma(\theta)^{-1}$.

2.5 Examples

Under conditions **(A1)-(A6)**, a family of estimators $(\tilde{\theta}_T)_{T\geq 0}$ of θ satisfying

$$\varphi_T^{-1}(\theta)\left(\tilde{\theta}_T - \theta\right) = \varphi_T^{-1}(\theta)\left[\nabla_\theta \ell(\theta)\right]_T^{-1} \nabla_\theta \ell_T(\theta) + o_{\mathrm{P}_\theta}(1),$$

is asymptotically efficient at θ. Furthermore, assuming the additional conditions **(A7)-(A8)**, this family $(\tilde{\theta}_T)_{T\geq 0}$ is regular at θ.

As a consequence, by Lemma 2.4.1, under conditions **(A1)-(A5)**, **(A9)** and (2.13), a family of estimators $(\tilde{\theta}_T)_{T\geq 0}$ of θ satisfying

$$\varphi_T^{-1}(\theta)\left(\tilde{\theta}_T - \theta\right) = \varphi_T^{-1}(\theta)\left\langle\nabla_\theta \ell(\theta)\right\rangle_T^{-1} \nabla_\theta \ell_T(\theta) + o_{\mathrm{P}_\theta}(1),$$

is asymptotically efficient at θ. Furthermore, assuming the additional conditions **(A7)-(A8)**, this family $(\tilde{\theta}_T)_{T\geq 0}$ is regular at θ.

We next present examples of the LAMN and LAN properties, both taken from [52].

2.5.1 Example 1

Recall that a pure birth process is a counting process $(X_t)_{t\geq 0} = (N_t)_{t\geq 0}$ with predictable intensity $\Psi(\theta, N_{t-}) = \theta N_{t-}$, where $N_0 = 1$ and the birth rate $\theta \in \Theta = (0, \infty)$. The integral $\int_0^t N_{s-}ds$ is the total time lived in the population before time t.

Observe that in this simple example, the process X_t is already written in the form (2.6), there is no dy-dependence, and thus the first condition in **(A3)** is not needed. The other two conditions in **(A3)** and **(A4)** are trivially satisfied.

By (2.7), the score function based on the continuous observation $\{N_t, 0 \leq t \leq T\}$ is given by

$$\nabla_\theta \ell_T(\theta) = \int_0^T \frac{\partial_\theta \Psi}{\Psi}(dN_t - \theta N_{t-}dt) = \frac{N_T - 1}{\theta} - \int_0^T N_{t-}dt.$$

It can be checked that the score function is a P_θ-martingale, and its quadratic variation and quadratic characteristic are respectively given by

$$\left[\nabla_\theta \ell(\theta)\right]_T = \frac{N_T - 1}{\theta^2}, \quad \text{and} \quad \left\langle \nabla_\theta \ell(\theta)\right\rangle_T = \frac{1}{\theta}\int_0^T N_{t-}dt.$$

It can be easily checked that as $T \to \infty$,

$$\theta^2 e^{-\theta T}\left[\nabla_\theta \ell(\theta)\right]_T = e^{-\theta T}(N_T - 1) \xrightarrow{\text{a.s.}} W,$$

where W is a random variable with exponential distribution with parameter 1.

This shows that hypothesis **(A6)** holds with $\varphi_T(\theta) = \theta e^{-\frac{\theta T}{2}}$ and asymptotic Fisher information $\Gamma(\theta) = W$. Finally, hypotheses **(A5)** and **(A8)** hold trivially. Thus, by Theorem 2.2.4 this process satisfies the LAMN property.

On the other hand, the maximum likelihood estimator (MLE) of θ is given by

$$\hat{\theta}_T = \frac{N_T - 1}{\int_0^T N_{s-} ds}.$$

Therefore,

$$\frac{1}{\theta} e^{\frac{\theta T}{2}} \left(\hat{\theta}_T - \theta \right) = \varphi_T^{-1}(\theta) \left\langle \nabla_\theta \ell(\theta) \right\rangle_T^{-1} \nabla_\theta \ell_T(\theta).$$

Observe that as $T \to \infty$,

$$\theta^2 e^{-\theta T} \left\langle \nabla_\theta \ell(\theta) \right\rangle_T = \theta e^{-\theta T} \int_0^T N_{s-} ds \xrightarrow{\text{a.s.}} W,$$

thus, the quadratic characteristic and quadratic variation of the score function are asymptotically equivalent at rate $\theta e^{-\frac{\theta T}{2}}$. Hence, **(A9)** is not needed.

Then, $\hat{\theta}_T$ is a family of regular and asymptotically efficient estimators for all $\theta \in \Theta$.

2.5.2 Example 2

Recall that Jacod [30] and Dietz [16] studied the LAN and LAMN property for Ornstein-Uhlenbeck processes without jumps. In this subsection we consider the Ornstein-Uhlenbeck processes with jumps, a particular case of equation (2.2), defined by

$$X_t = X_0 + \theta_1 \int_0^t X_s ds + \sigma B_t + \int_0^t \int_{\mathbb{R}_0} y N(ds, dy), \tag{2.15}$$

where X_0 is a random variable with finite second moment, $\sigma \geq 0$, $\theta_2 = (\theta_2^1, \ldots, \theta_2^{k-1}) \in \widetilde{\Theta}$, $\theta = (\theta_1, \theta_2) \in \Theta = \mathbb{R} \times \widetilde{\Theta}$, where $\widetilde{\Theta}$ is an open subset of $\mathbb{R}^{k-1}$, for some integer $k > 1$, $N(ds, dy)$ is a Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}_0$ with intensity measure $\mu_{\theta_2}(dy)dt = f(\theta_2, y)dydt$, where $f : \Theta \times \mathbb{R} \to \mathbb{R}_+$ is a Borel function such that $\int_{\mathbb{R}} |y|^2 f(\theta_2, y)dy < \infty$ and $f(\theta_2, 0) = 0$, for all $\theta_2 \in \widetilde{\Theta}$. We also assume that $f \in C^1$ with respect to θ_2.

Assume the following conditions on the Lévy density f.

(H1) For any $\overline{\theta}_2, \theta_2 \in \widetilde{\Theta}$,

$$\int_{\mathbb{R}_0} \left| \left(\frac{f(\overline{\theta}_2, y)}{f(\theta_2, y)} - 1 \right) y \right| f(\theta_2, y)dy + \int_{\mathbb{R}_0} \left(\frac{f(\overline{\theta}_2, y)}{f(\theta_2, y)} - 1 \right) f(\theta_2, y)dy$$
$$+ \int_{\mathbb{R}_0} \left(\frac{f(\overline{\theta}_2, y)}{f(\theta_2, y)} - 1 \right)^2 f(\theta_2, y)dy < \infty.$$

(H2) For any $\theta_2 \in \widetilde{\Theta}$, $p > 2$ and $i \in \{1, \ldots, k-1\}$, $\int_{\mathbb{R}_0} \left| \partial_{\theta_2^i} \ln \left(f(\theta_2, y) \right) \right|^p f(\theta_2, y)dy < \infty$.

(H3) For any $\theta_2 \in \widetilde{\Theta}$, $\int_{\mathbb{R}_0} (\nabla_{\theta_2} \ln \left(f(\theta_2, y) \right))^{\mathsf{T}} \nabla_{\theta_2} \ln \left(f(\theta_2, y) \right) f(\theta_2, y)dy$ is positive definite.

(H4) For all $u \in \mathbb{R}^{k-1}$, as $T \to \infty$,

$$\int_{\mathbb{R}_0} \left| \frac{f(\theta_2 + \frac{u}{\sqrt{T}}, y)}{f(\theta_2, y)} - 1 - \frac{1}{\sqrt{T}} u^{\mathsf{T}} \nabla_{\theta_2} \ln \left(f(\theta_2, y) \right) \right|^2 f(\theta_2, y)dy = o\left(\frac{1}{T} \right),$$

uniformly in $\theta_2 \in \widetilde{\Theta}$.

Let us now consider the following cases :

Case 1 : $\sigma \geq 0$, $\theta_1 = 0$. Assuming **(H1)** (which implies **(A3)**), the score function and its quadratic characteristic are respectively given by

$$\nabla_{\theta_2}\ell_T(\theta_2) = \int_0^T \int_{\mathbb{R}_0} \nabla_{\theta_2} \ln\left(f(\theta_2, y)\right)\left(N(dt,dy) - \mu_{\theta_2}(dy)dt\right),$$

$$\langle \nabla_{\theta_2}\ell(\theta_2)\rangle_T = T\int_{\mathbb{R}_0} \left(\nabla_{\theta_2} \ln\left(f(\theta_2, y)\right)\right)^{\mathsf{T}} \nabla_{\theta_2} \ln\left(f(\theta_2, y)\right)\mu_{\theta_2}(dy).$$

Taking $\varphi_T(\theta_2)$ as the $(k-1)\times(k-1)$ diagonal matrix with entries equal to $\frac{1}{\sqrt{T}}$, then observe that

$$\begin{aligned}
\mathrm{E}_\theta\left[\sup_{0\leq t\leq T}\left|\varphi_T(\theta_2)\nabla_{\theta_2}\ln(f(\theta_2,\Delta X_t))\right|\right] &\leq \frac{1}{\sqrt{T}}\sum_{i=1}^{k-1}\mathrm{E}_\theta\left[\sup_{0\leq t\leq T}\left|\partial_{\theta_2^i}\ln(f(\theta_2,\Delta X_t))\right|\right]\\
&\leq \frac{1}{\sqrt{T}}\sum_{i=1}^{k-1}\left(\mathrm{E}_\theta\left[\sup_{0\leq t\leq T}\left|\partial_{\theta_2^i}\ln(f(\theta_2,\Delta X_t))\right|^p\right]\right)^{\frac{1}{p}}\\
&\leq \frac{1}{\sqrt{T}}\sum_{i=1}^{k-1}\left(\mathrm{E}_\theta\left[\sum_{0\leq t\leq T}\left|\partial_{\theta_2^i}\ln(f(\theta_2,\Delta X_t))\right|^p\right]\right)^{\frac{1}{p}}\\
&= \frac{1}{\sqrt{T}}\sum_{i=1}^{k-1}\left(\mathrm{E}_\theta\left[\int_0^T\int_{\mathbb{R}_0}\left|\partial_{\theta_2^i}\ln(f(\theta_2,y))\right|^p N(dt,dy)\right]\right)^{\frac{1}{p}}\\
&= \frac{1}{T^{\frac{1}{2}-\frac{1}{p}}}\sum_{i=1}^{k-1}\left(\int_{\mathbb{R}_0}\left|\partial_{\theta_2^i}\ln(f(\theta_2,y))\right|^p f(\theta_2,y)dy\right)^{\frac{1}{p}},
\end{aligned}$$

which, by hypothesis **(H2)**, tends to zero as $T\to\infty$. Hence, condition **(A5)** is satisfied.

Condition **(A9)** is satisfied for all $\theta_2 \in \widetilde{\Theta}$ since for all $\epsilon > 0$ and $u \in \mathbb{R}^{k-1}$, as $T\to\infty$,

$$\int_{\mathbb{R}_0}\left(u^{\mathsf{T}}\nabla_{\theta_2}\ln\left(f(\theta_2,y)\right)\right)^2 \mathbf{1}_{\left\{\left|u^{\mathsf{T}}\nabla_{\theta_2}\ln(f(\theta_2,y))\right|>\epsilon\sqrt{T}\right\}} f(\theta_2,y)dy = o(1),$$

uniformly in $\theta_2 \in \widetilde{\Theta}$. On the other hand, for all $T > 0$,

$$\frac{1}{T}\langle\nabla_{\theta_2}\ell(\theta_2)\rangle_T = \int_{\mathbb{R}_0}\left(\nabla_{\theta_2}\ln\left(f(\theta_2,y)\right)\right)^{\mathsf{T}}\nabla_{\theta_2}\ln\left(f(\theta_2,y)\right)\mu_{\theta_2}(dy) =: \widetilde{\Gamma}(\theta_2),$$

which is independent of T. Thus, by Lemma 2.4.1, hypothesis **(A6)** holds. Moreover, it is easy to see that **(A7)** holds true. Finally, hypothesis **(H4)** implies **(A8)**.

As a consequence of Theorem 2.2.4, under conditions **(H1)**-**(H4)**, the LAN property is satisfied with rate of convergence $\sqrt{T}$ and asymptotic Fisher information matrix $\widetilde{\Gamma}(\theta_2)$.

In particular, when X is a one-dimensional Poisson process with intensity $\lambda(\theta)$, the LAN property holds with rate of convergence $\sqrt{T}$ and asymptotic Fisher information $\Gamma(\theta) = \frac{(\lambda'(\theta))^2}{\lambda(\theta)}$.

Case 2 : $\sigma > 0$, $\theta_1 > 0$. Assuming **(H1)**, the score function is given by

$$\nabla_\theta\ell_T(\theta) = \left(\int_0^T\frac{X_t}{\sigma}dB_t, \int_0^T\int_{\mathbb{R}_0}\nabla_{\theta_2}\ln\left(f(\theta_2,y)\right)\left(N(dt,dy)-\mu_{\theta_2}(dy)dt\right)\right),$$

and its quadratic variation and quadratic characteristic are respectively given by

$$[\nabla_\theta\ell(\theta)]_T = \begin{pmatrix}\frac{1}{\sigma^2}\int_0^T X_t^2dt & 0\\ 0 & \int_0^T\int_{\mathbb{R}_0}\left(\nabla_{\theta_2}\ln\left(f(\theta_2,y)\right)\right)^{\mathsf{T}}\nabla_{\theta_2}\ln\left(f(\theta_2,y)\right)N(dt,dy)\end{pmatrix},$$

$$\langle\nabla_\theta\ell(\theta)\rangle_T = \begin{pmatrix}\frac{1}{\sigma^2}\int_0^T X_t^2dt & 0\\ 0 & T\int_{\mathbb{R}_0}\left(\nabla_{\theta_2}\ln\left(f(\theta_2,y)\right)\right)^{\mathsf{T}}\nabla_{\theta_2}\ln\left(f(\theta_2,y)\right)\mu_{\theta_2}(dy)\end{pmatrix}.$$

Taking the diagonal entries of the $k \times k$ diagonal matrix $\varphi_T(\theta)$ as $\varphi_{i,T}(\theta) = \frac{1}{\sqrt{T}}$ for $i = 2, ..., k$. Suppose that $\rho_1 := \int_{\mathbb{R}_0} y\mu_{\theta_2}(dy) < \infty$ and $\rho_2 := \int_{\mathbb{R}_0} y^2\mu_{\theta_2}(dy) < \infty$, for all $\theta_2 \in \widetilde{\Theta}$.

By Itô's formula, the unique solution to equation (2.15) is given by

$$X_t = e^{\theta_1 t}X_0 + \rho_1 \int_0^t e^{\theta_1(t-s)}ds + \sigma \int_0^t e^{\theta_1(t-s)}dB_s + \int_0^t \int_{\mathbb{R}_0} e^{\theta_1(t-s)}y\left(N(ds,dy) - \mu_{\theta_2}(dy)ds\right).$$

Now, consider the martingale

$$\begin{aligned} M_t &= e^{-\theta_1 t}X_t - X_0 + \frac{e^{-\theta_1 t} - 1}{\theta_1}\rho_1 \\ &= \sigma \int_0^t e^{-\theta_1 s}dB_s + \int_0^t \int_{\mathbb{R}_0} e^{-\theta_1 s}y\left(N(ds,dy) - \mu_{\theta_2}(dy)ds\right). \end{aligned}$$

Observe that $\{M_t, \mathcal{F}_t\}_{t\geq 0}$ is a zero-mean square integrable martingale since for all $t \geq 0$,

$$\mathrm{E}_\theta[M_t^2] = -\sigma^2\frac{e^{-2\theta_1 t} - 1}{2\theta_1} - \frac{e^{-2\theta_1 t} - 1}{2\theta_1}\rho_2 \leq \frac{\sigma^2 + \rho_2}{2\theta_1} < \infty.$$

Hence, applying the martingale convergence theorem, M_t converges almost surely as $t \to \infty$ to the random variable

$$M_\infty = \sigma \int_0^\infty e^{-\theta_1 s}dB_s + \int_0^\infty \int_{\mathbb{R}_0} e^{-\theta_1 s}y\left(N(ds,dy) - \mu_{\theta_2}(dy)ds\right).$$

Thus, $e^{-\theta_1 t}X_t$ converges almost surely to $X_0 + \frac{\rho_1}{\theta_1} + M_\infty$ as $t \to \infty$, which implies that $e^{-2\theta_1 t}X_t^2$ converges almost surely to $(X_0 + \frac{\rho_1}{\theta_1} + M_\infty)^2$ as $t \to \infty$. Using the integral version of the Toeplitz lemma, we get that as $t \to \infty$,

$$\frac{\int_0^t X_s^2 ds}{\int_0^t e^{2\theta_1 s}ds} \xrightarrow{\text{a.s.}} \left(X_0 + \frac{\rho_1}{\theta_1} + M_\infty\right)^2,$$

which yields that as $t \to \infty$

$$e^{-2\theta_1 t}\int_0^t X_s^2 ds \xrightarrow{\text{a.s.}} \frac{1}{2\theta_1}\left(X_0 + \frac{\rho_1}{\theta_1} + M_\infty\right)^2.$$

Hence, taking $\varphi_{1,T}(\theta) = e^{-\theta_1 T}$, that is, $\varphi_T(\theta) = \mathrm{diag}(e^{-\theta_1 T}, \frac{1}{\sqrt{T}}, \ldots, \frac{1}{\sqrt{T}})$, together with **(H3)**, condition **(A6)** is satisfied with

$$\Gamma(\theta) = \begin{pmatrix} \frac{1}{2\sigma^2\theta_1}\left(X_0 + \frac{\rho_1}{\theta_1} + M_\infty\right)^2 & 0 \\ 0 & \widetilde{\Gamma}(\theta_2) \end{pmatrix}.$$

Assuming additionally conditions **(H2)** and **(H4)** and applying Theorem 2.2.4, the LAMN property holds with $\varphi_T(\theta) = \mathrm{diag}(e^{-\theta_1 T}, \frac{1}{\sqrt{T}}, \ldots, \frac{1}{\sqrt{T}})$ and asymptotic Fisher information matrix $\Gamma(\theta)$.

Case 3 : $\sigma > 0$, $\theta_1 < 0$. Recall that if

$$\int_{|y|>2} \log|y|\mu_{\theta_2}(dy) < \infty, \tag{2.16}$$

for all $\theta_2 \in \widetilde{\Theta}$, X is ergodic with a unique invariant probability measure $\pi_\theta(dx)$ which can be calculated explicitly (see [65, Theorem 17.5 and Corollary 17.9] and [53, Theorem 2.6]). Therefore, by ergodicity, as $T \to \infty$,

$$\frac{1}{T}\int_0^T X_t^2 dt \xrightarrow{\mathrm{P}_\theta} \int_{\mathbb{R}} x^2\pi_\theta(dx).$$

In this case X_t converges almost surely as $t \to \infty$ to the random variable

$$X_\infty = -\frac{\rho_1}{\theta_1} + \sigma \int_0^\infty e^{\theta_1 s} dB_s + \int_0^\infty \int_{\mathbb{R}_0} e^{\theta_1 s} y \left(N(ds, dy) - \mu_{\theta_2}(dy) ds\right).$$

Again, by ergodicity, as $T \to \infty$,

$$\frac{1}{T}\int_0^T X_t^2 dt \xrightarrow{\mathrm{P}_\theta} \mathrm{E}_\theta \left[X_\infty^2\right].$$

By Itô's formula,

$$X_t^2 = e^{2\theta_1 t} X_0^2 + 2\rho_1 \int_0^t e^{\theta_1 (t-s)} X_s ds + 2\sigma \int_0^t e^{\theta_1 (t-s)} X_s dB_s + \left(\sigma^2 + \rho_2\right) \int_0^t e^{2\theta_1 (t-s)} ds$$
$$+ \int_0^t \int_{\mathbb{R}_0} \left(2 e^{\theta_1 (t-s)} y X_s + e^{2\theta_1 (t-s)} y^2\right) \left(N(ds, dy) - \mu_{\theta_2}(dy) ds\right),$$

together with $\mathrm{E}_\theta[X_t] = e^{\theta_1 t}(\mathrm{E}[X_0] + \frac{\rho_1}{\theta_1}) - \frac{\rho_1}{\theta_1}$, shows that

$$\mathrm{E}_\theta[X_t^2] = e^{2\theta_1 t}\mathrm{E}[X_0^2] + 2\rho_1 \left(\mathrm{E}[X_0] + \frac{\rho_1}{\theta_1}\right) e^{\theta_1 t} t + \frac{2\rho_1^2}{\theta_1^2}\left(1 - e^{\theta_1 t}\right) - \frac{\sigma^2 + \rho_2}{2\theta_1}\left(1 - e^{2\theta_1 t}\right).$$

Therefore,

$$\mathrm{E}_\theta \left[X_\infty^2\right] = \lim_{t \to \infty} \mathrm{E}_\theta \left[X_t^2\right] = \frac{2\rho_1^2}{\theta_1^2} - \frac{\sigma^2 + \rho_2}{2\theta_1},$$

which concludes that

$$\int_{\mathbb{R}} x^2 \pi_\theta(dx) = \frac{2\rho_1^2}{\theta_1^2} - \frac{\sigma^2 + \rho_2}{2\theta_1}.$$

As a consequence, assuming conditions **(H1)-(H4)** and applying Theorem 2.4.2, the LAN property is satisfied with rate of convergence $\sqrt{T}$ and asymptotic Fisher information matrix

$$\Gamma(\theta) = \begin{pmatrix} \frac{1}{\sigma^2}\left(\frac{2\rho_1^2}{\theta_1^2} - \frac{\sigma^2 + \rho_2}{2\theta_1}\right) & 0 \\ 0 & \widetilde{\Gamma}(\theta_2) \end{pmatrix}.$$

Next, we are going to study the asymptotic properties of the MLE of the parameter θ for the parametric model (2.15) in the case that $\sigma > 0$ and $\theta_1 \neq 0$. In particular, we show that the MLE of θ_1 is asymptotically efficient.

Note that the MLE $\hat{\theta}_T = (\hat{\theta}_{1,T}, \hat{\theta}_{2,T})$ of θ satisfies

$$\nabla_\theta \ell_T(\theta) = \left(\int_0^T \frac{X_t}{\sigma} dB_t, \int_0^T \int_{\mathbb{R}_0} \nabla_{\theta_2} \ln\left(f(\theta_2, y)\right)\left(N(dt, dy) - \mu_{\theta_2}(dy) dt\right)\right) = (0, \ldots, 0).$$

First observe that the MLE $\hat{\theta}_{1,T}$ of the drift parameter θ_1 satisfies the following equation

$$\frac{1}{\sigma}\int_0^T X_t dB_t = \frac{1}{\sigma^2}\int_0^T X_t \left(dX_t - \theta_1 X_t dt - \int_{\mathbb{R}_0} y N(dt, dy)\right) = 0,$$

which yields that under P_θ,

$$\hat{\theta}_{1,T} = \frac{\int_0^T X_t dX_t - \int_0^T \int_{\mathbb{R}_0} X_t y N(dt, dy)}{\int_0^T X_t^2 dt} = \theta_1 + \frac{\sigma \int_0^T X_t dB_t}{\int_0^T X_t^2 dt}.$$

When $\theta_1 > 0$, set $\Gamma_{1,1}(\theta) := \frac{1}{2\sigma^2\theta_1}(X_0 + \frac{\rho_1}{\theta_1} + M_\infty)^2$. Then, as $T \to \infty$,

$$e^{\theta_1 T}\left(\hat{\theta}_{1,T} - \theta_1\right) = \frac{\sigma e^{-\theta_1 T}\int_0^T X_t dB_t}{e^{-2\theta_1 T}\int_0^T X_t^2 dt} \xrightarrow{\mathcal{L}(\mathrm{P}_\theta)} \Gamma_{1,1}(\theta)^{-1/2}\mathcal{N}(0,1).$$

When $\theta_1 < 0$, we assume (2.16) and set $\Gamma_{1,1}(\theta) := \frac{1}{\sigma^2}(\frac{2\rho_1^2}{\theta_1^2} - \frac{\sigma^2+\rho_2}{2\theta_1})$. Then, as $T \to \infty$,

$$\sqrt{T}\left(\hat{\theta}_{1,T} - \theta_1\right) = \frac{\sigma \frac{1}{\sqrt{T}} \int_0^T X_t dB_t}{\frac{1}{T}\int_0^T X_t^2 dt} \overset{\mathcal{L}(\mathrm{P}_\theta)}{\longrightarrow} \mathcal{N}\left(0, \Gamma_{1,1}(\theta)^{-1}\right).$$

Consequently, we conclude that the MLE $\hat{\theta}_{1,T}$ of θ_1 is asymptotically efficient for all $\theta_1 \neq 0$. Next, notice that the MLE $\hat{\theta}_{2,T}$ of θ_2 satisfies the following equation

$$\int_0^T \int_{\mathbb{R}_0} \nabla_{\theta_2} \ln\left(f(\theta_2, y)\right)\left(N(dt, dy) - \mu_{\theta_2}(dy)dt\right) = 0, \tag{2.17}$$

which states that the solution $\hat{\theta}_{2,T}$ depends on the Lévy density f. We shall consider here the following two particular cases.

When $(\int_0^t \int_{\mathbb{R}_0} yN(ds, dy))_{t\geq 0}$ is a Poisson process $(N_t)_{t\geq 0}$ with intensity $\theta_2 > 0$. In this case, the solution to equation (2.17) is given by

$$\hat{\theta}_{2,T} = \frac{N_T}{T}.$$

By the Central Limit theorem, as $T \to \infty$,

$$\sqrt{T}\left(\hat{\theta}_{2,T} - \theta_2\right) = \sqrt{T}\left(\frac{N_T}{T} - \theta_2\right) \overset{\mathcal{L}(\mathrm{P}_\theta)}{\longrightarrow} \mathcal{N}\left(0, \theta_2\right).$$

Thus, in this case, as $T \to \infty$, the MLE $\hat{\theta}_T$ satisfies

$$\varphi_T^{-1}(\theta)\left(\hat{\theta}_T - \theta\right) \overset{\mathcal{L}(\mathrm{P}_\theta)}{\longrightarrow} \Gamma(\theta)^{-1/2}\mathcal{N}\left(0, I_2\right),$$

with

$$\Gamma(\theta) = \begin{pmatrix} \Gamma_{1,1}(\theta) & 0 \\ 0 & \frac{1}{\theta_2} \end{pmatrix},$$

which implies that $\hat{\theta}_T$ is asymptotically efficient. Moreover, $\hat{\theta}_T$ is regular since for all $\theta \in \mathbb{R}_0 \times \widetilde{\Theta}$,

$$\varphi_T^{-1}(\theta)\left(\hat{\theta}_T - \theta\right) = \varphi_T^{-1}(\theta)\left\langle \nabla_\theta \ell(\theta)\right\rangle_T^{-1} \nabla_\theta \ell_T(\theta).$$

We next consider the case where the Lévy density $f(\theta_2, y)$ takes the form $\dfrac{\lambda}{\alpha}e^{-y/\alpha}\mathbf{1}_{(0,\infty)}(y)$ with $\lambda, \alpha > 0$ and $\theta_2 = (\lambda, \alpha)$. In this case, solving equation (2.17), we find that the MLE $\hat{\theta}_{2,T} = (\hat{\lambda}_T, \hat{\alpha}_T)$ is given by

$$\hat{\lambda}_T = \frac{N_T}{T}, \quad \text{and} \quad \hat{\alpha}_T = \frac{\int_0^T \int_{\mathbb{R}_0} yN(dt, dy)}{N_T},$$

where N_t is a Poisson process with intensity λ. From the Central Limit theorem, we have that as $T \to \infty$,

$$\sqrt{T}\left(\hat{\lambda}_T - \lambda\right) \overset{\mathcal{L}(\mathrm{P}_\theta)}{\longrightarrow} \mathcal{N}\left(0, \lambda\right).$$

Moreover, applying Theorem 2.2.2, we obtain that as $T \to \infty$,

$$\sqrt{T}\left(\hat{\alpha}_T - \alpha\right) = \frac{T}{N_T}\frac{1}{\sqrt{T}}\int_0^T \int_{\mathbb{R}_0} (y - \alpha)(N(dt, dy) - \mu_{\theta_2}(dy)dt) \overset{\mathcal{L}(\mathrm{P}_\theta)}{\longrightarrow} \mathcal{N}\left(0, \frac{\alpha^2}{\lambda}\right).$$

Hence, we conclude that as $T \to \infty$,

$$\varphi_T^{-1}(\theta)\left(\hat{\theta}_T - \theta\right) \xrightarrow{\mathcal{L}(\mathrm{P}_\theta)} \Gamma(\theta)^{-1/2}\mathcal{N}\left(0, I_3\right),$$

with

$$\Gamma(\theta) = \begin{pmatrix} \Gamma_{1,1}(\theta) & 0 & 0 \\ 0 & \frac{1}{\lambda} & 0 \\ 0 & 0 & \frac{\lambda}{\alpha^2} \end{pmatrix},$$

which implies that $\hat{\theta}_T$ is asymptotically efficient. Moreover, $\hat{\theta}_T$ is regular since for all $\theta \in \mathbb{R}_0 \times \widetilde{\Theta}$,

$$\varphi_T^{-1}(\theta)\left(\hat{\theta}_T - \theta\right) = \varphi_T^{-1}(\theta)\left\langle \nabla_\theta \ell(\theta)\right\rangle_T^{-1} \nabla_\theta \ell_T(\theta) + o_{\mathrm{P}_\theta}(1).$$

Chapitre 3

LAN property for a simple Lévy process

In this chapter, we consider a simple Lévy process given by a Brownian motion and a compensated Poisson process, whose drift and diffusion parameters as well as its intensity are unknown. Supposing that the process is observed discretely at high frequency we derive the local asymptotic normality (LAN) property. In order to obtain this result, Malliavin calculus and Girsanov's theorem are applied in order to write the log-likelihood ratio in terms of sums of conditional expectations, for which a central limit theorem for triangular arrays can be applied.

3.1 Introduction and main result

On a complete probability space $(\Omega^\lambda, \mathcal{F}^\lambda, \mathrm{P}^\lambda)$ defined in Definition 1.1.3, we consider the following stochastic process $X^{\theta,\sigma,\lambda} = (X_t^{\theta,\sigma,\lambda})_{t\geq 0}$ in $\mathbb{R}$ defined by

$$X_t^{\theta,\sigma,\lambda} = x_0 + \theta t + \sigma B_t + N_t - \lambda t, \tag{3.1}$$

where $B = (B_t)_{t\geq 0}$ is a standard Brownian motion, $N = (N_t)_{t\geq 0}$ is a Poisson process with intensity $\lambda > 0$ independent of B, and we denote by $(\widetilde{N}_t^\lambda)_{t\geq 0}$ the compensated Poisson process $\widetilde{N}_t^\lambda := N_t - \lambda t$. The parameters $(\theta, \sigma, \lambda) \in \Theta \times \Sigma \times \Lambda$ are unknown and Θ, Σ and Λ are closed intervals of $\mathbb{R}, \mathbb{R}_+^*$ and $\mathbb{R}_+^*$, where $\mathbb{R}_+^* = \mathbb{R}_+ \setminus \{0\}$. Let $\{\widehat{\mathcal{F}}_t^\lambda\}_{t\geq 0}$ denote the natural filtration generated by B and N. Note that $\{\widehat{\mathcal{F}}_t^\lambda\}_{t\geq 0}$ is also the natural filtration generated by $X^{\theta,\sigma,\lambda}$ for all $(\theta, \sigma, \lambda) \in \Theta \times \Sigma \times \Lambda$. We denote by $\mathrm{P}^{\theta,\sigma,\lambda}$ the probability law induced by the $\widehat{\mathcal{F}}^\lambda$-adapted càdlàg stochastic process $X^{\theta,\sigma,\lambda}$, and by $\mathrm{E}^{\theta,\sigma,\lambda}$ the expectation with respect to $\mathrm{P}^{\theta,\sigma,\lambda}$. Let $\overset{\mathrm{P}^{\theta,\sigma,\lambda}}{\longrightarrow}$ and $\overset{\mathcal{L}(\mathrm{P}^{\theta,\sigma,\lambda})}{\longrightarrow}$ denote the convergence in $\mathrm{P}^{\theta,\sigma,\lambda}$-probability and in $\mathrm{P}^{\theta,\sigma,\lambda}$-law, respectively.

Recall that $(\Omega^2, \mathcal{F}^2, \mathrm{P}^2)$ is the canonical probability space associated with the Poisson process N with intensity λ. Therefore, we denote by $(\Omega^{2,\lambda}, \mathcal{F}^{2,\lambda}, \mathrm{P}^{2,\lambda})$ instead of $(\Omega^2, \mathcal{F}^2, \mathrm{P}^2)$. The structure of the probability space is then given by $\widehat{\Omega}^\lambda = \Omega^1 \times \Omega^{2,\lambda}$, $\widetilde{\Omega}^\lambda = \Omega^3 \times \Omega^{4,\lambda}$, $\widehat{\mathcal{F}}^\lambda = \mathcal{F}^1 \otimes \mathcal{F}^{2,\lambda}$, $\widetilde{\mathcal{F}}^\lambda = \mathcal{F}^3 \otimes \mathcal{F}^{4,\lambda}$, $\widehat{\mathrm{P}}^\lambda = \mathrm{P}^1 \otimes \mathrm{P}^{2,\lambda}$, $\widetilde{\mathrm{P}}^\lambda = \mathrm{P}^3 \otimes \mathrm{P}^{4,\lambda}$, and $\Omega^\lambda = \widehat{\Omega}^\lambda \times \widetilde{\Omega}^\lambda$, $\mathcal{F}^\lambda = \widehat{\mathcal{F}}^\lambda \otimes \widetilde{\mathcal{F}}^\lambda$, $\mathrm{P}^\lambda = \widehat{\mathrm{P}}^\lambda \otimes \widetilde{\mathrm{P}}^\lambda$. We denote by E^λ, $\widehat{\mathrm{E}}^\lambda$, $\widetilde{\mathrm{E}}^\lambda$ the expectation with respect to P^λ, $\widehat{\mathrm{P}}^\lambda$ and $\widetilde{\mathrm{P}}^\lambda$, respectively.

For fixed $(\theta_0, \sigma_0, \lambda_0) \in \Theta \times \Sigma \times \Lambda$, we consider the probability space $(\Omega, \mathcal{F}, \mathrm{P})$, where $\Omega \equiv \Omega^{\lambda_0}$, $\mathcal{F} \equiv \mathcal{F}^{\lambda_0}, \mathrm{P} \equiv \mathrm{P}^{\lambda_0}$. Let us denote $\mathrm{E} \equiv \mathrm{E}^{\lambda_0}$, $\widehat{\mathrm{P}} \equiv \widehat{\mathrm{P}}^{\lambda_0} = \mathrm{P}^1 \otimes \mathrm{P}^{2,\lambda_0}$, and $\widetilde{\mathrm{P}} \equiv \widetilde{\mathrm{P}}^{\lambda_0} = \mathrm{P}^3 \otimes \mathrm{P}^{4,\lambda_0}$. Consider an equidistant discrete observation of the process $X^{\theta_0,\sigma_0,\lambda_0}$ which is denoted by $X^n = (X_{t_0}, X_{t_1}, \ldots, X_{t_n})$, where $t_k = k\Delta_n$ for $k \in \{0, \ldots, n\}$, and $\Delta_n \leq 1$. We assume that the high-frequency observation condition holds. That is,

$$n\Delta_n \to \infty, \quad \text{and} \quad \Delta_n \to 0, \quad \text{as } n \to \infty. \tag{3.2}$$

Given the process $(X_t^{\theta,\sigma,\lambda})_{t\geq 0}$, we denote by $p(\cdot; (\theta, \sigma, \lambda))$ the density of the random vector $(X_{t_0}^{\theta,\sigma,\lambda}, X_{t_1}^{\theta,\sigma,\lambda}, \ldots, X_{t_n}^{\theta,\sigma,\lambda})$. In particular, $p(\cdot; (\theta_0, \sigma_0, \lambda_0))$ denotes the density of the observation X^n. For $(u, v, w) \in \mathbb{R}^3$, set $\theta_n := \theta_0 + \frac{u}{\sqrt{n\Delta_n}}, \sigma_n := \sigma_0 + \frac{v}{\sqrt{n}}, \lambda_n := \lambda_0 + \frac{w}{\sqrt{n\Delta_n}}$.

The aim of this chapter is to prove the following LAN property.

Theorem 3.1.1. *Assume condition* (3.2). *Then, the LAN property holds for the likelihood at* $(\theta_0, \sigma_0, \lambda_0)$ *with rate of convergence* $(\sqrt{n\Delta_n}, \sqrt{n}, \sqrt{n\Delta_n})$ *and asymptotic Fisher information matrix* $\Gamma(\theta_0, \sigma_0, \lambda_0)$. *That is, for all* $z = (u, v, w) \in \mathbb{R}^3$, *as* $n \to \infty$,

$$\log \frac{p\left(X^n; (\theta_n, \sigma_n, \lambda_n)\right)}{p\left(X^n; (\theta_0, \sigma_0, \lambda_0)\right)} \overset{\mathcal{L}(\mathrm{P}^{\theta_0,\sigma_0,\lambda_0})}{\longrightarrow} z^{\mathsf{T}} \mathcal{N}\left(0, \Gamma(\theta_0, \sigma_0, \lambda_0)\right) - \frac{1}{2} z^{\mathsf{T}} \Gamma(\theta_0, \sigma_0, \lambda_0) z,$$

where $\mathcal{N}(0, \Gamma(\theta_0, \sigma_0, \lambda_0))$ *is a centered* $\mathbb{R}^3$*-valued Gaussian vector with covariance matrix*

$$\Gamma(\theta_0, \sigma_0, \lambda_0) = \frac{1}{\sigma_0^2} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & 1 + \frac{\sigma_0^2}{\lambda_0} \end{pmatrix}.$$

Remark 3.1.1. *The LAN property remains valid for the simple Lévy process driven by a Brownian motion and a Poisson process* $N = (N_t)_{t \geq 0}$ *with intensity* $\lambda > 0$, *i.e,* $X_t^{\theta,\sigma,\lambda} = x_0 + \theta t + \sigma B_t + N_t$, *for* $t \geq 0$, *with the same rates of convergence. However, the asymptotic Fisher information matrix changes in this case where there is no correlation between two components involving the parameters* θ *and* λ. *Here, we use the compensated Poisson process because we will try to deal with the infinite Lévy measure for more general cases.*

Recall that Gobet in [25] deals with the multidimensional continuous elliptic diffusions. Some extensions of Gobet's work with the presence of jumps are given for e.g. in [10], [23], and [38]. In the present chapter, we estimate the drift and diffusion parameters and the jump intensity at the same time. One way to proceed in order to prove Theorem 3.1.1 would be to use explicit expression for the density. However, the main motivation for this chapter is to show some of the important properties and arguments in order to prove the LAN property in the non-linear case whose proof is non-trivial.

In fact, the LAN property is a local central limit theorem which is robust to local changes in the values of the parameters to be estimated. The first problem in proving such a result for a combination of drift, Brownian motion and jump process is the fact that the density function cannot be explicitly written. This problem is aggravated in the case of stochastic equations driven by these processes due to the respective non-linear coefficients.

In order to overcome this problem in the general case, one needs such estimates of the derivatives of the log-density. This is quite a difficult problem, which in the Ornstein-Uhlenbeck case [38] is solved due to semi-explicit form and the integration by parts formula with respect to the Brownian motion.

In the present chapter, we present four important Lemmas (Lemmas 3.2.3, 3.2.4, 3.2.5 and 3.2.6) of independent interest which will be the key elements in dealing with the non-linear case. We have preferred to present them in this simpler form as in the general case further arguments need to be added.

We point out that in most cases one cannot expect to find good estimates of the derivatives of the logarithm of the density of X due to the mixture of exponential tails coming from the jump process together with the Gaussian tails of the Brownian motion. In fact, one cannot expect to have upper and lower bounds for the log-density belonging to the same class as in the general arguments of [24] and [25]. Our argument which that will be applicable to the general non-linear case is as follows.

One needs to condition on the number of jumps within the conditional expectation which expresses the transition density and outside it. When these two conditionings relate to different jumps one may use a large deviation principle in the estimate. When they are equal one uses the complementary set in order to apply the large deviation principle. The main term can be handled directly. Within all these arguments the Gaussian type upper and lower bounds of the density conditioned on the jumps are again strongly used.

In fact, the lemmas mentioned previously, deal with this idea. Therefore the semi-explicit Taylor expansion in [38] is replaced by a large deviation analysis within two expectations under contiguous probability measures. This idea seems to have many other uses in the set-up of stochastic differential equations driven by a Brownian motion and a jump process. We remark here that a plain Itô-Taylor expansion would not solve the problem as higher moments of the Poisson process do not become smaller as the expansion order increases.

A related idea to the one presented here appears in [10], where the case of a compound Poisson process is treated. As we will show in forthcoming chapters, the present idea seems to be important in order to obtain many properties for models where one considers a continuous diffusion perturbed by a jump process.

In Section 3.2 we present the integration by parts formulas to be used with respect to the Brownian motion and the Poisson process. We also give our main lemmas to be used for the proof. The proof of the main result is given in Section 3.3. We close with conclusion and further remarks towards the proof in the general non-linear case.

As usual, constants will be denoted by C or c and they will always be independent of time and Δ_n but may depend on bounds for the set $\Theta \times \Sigma \times \Lambda$. They may change of value from one line to the next.

3.2 Preliminaries

In this Section we introduce the preliminary results needed for the proof of Theorem 3.1.1. In order to deal with the likelihood ratio in Theorem 3.1.1, we will use the following decomposition

$$\begin{aligned}\log \frac{p\left(X^n;(\theta_n,\sigma_n,\lambda_n)\right)}{p\left(X^n;(\theta_0,\sigma_0,\lambda_0)\right)} &= \log \frac{p\left(X^n;(\theta_n,\sigma_n,\lambda_n)\right)}{p\left(X^n;(\theta_n,\sigma_0,\lambda_n)\right)} + \log \frac{p\left(X^n;(\theta_n,\sigma_0,\lambda_n)\right)}{p\left(X^n;(\theta_n,\sigma_0,\lambda_0)\right)} \\ &\quad + \log \frac{p\left(X^n;(\theta_n,\sigma_0,\lambda_0)\right)}{p\left(X^n;(\theta_0,\sigma_0,\lambda_0)\right)}.\end{aligned} \tag{3.3}$$

For each of the above terms we will use a mean value theorem on the parameter space and then analyze each term, which will lead to the derivative of the density function. To analyze this derivative, we will use as in Gobet [24] the integration by parts formula of the Malliavin calculus on each interval $[t_k, t_{k+1}]$ in order to obtain the following expressions for the derivatives of the log-likelihood function w.r.t. θ and σ. For this reason, we introduce an extra probabilistic representation of the process $X^{\theta,\sigma,\lambda}$. That is, consider the flow $Y^{\theta,\sigma,\lambda}(s,x) = (Y_t^{\theta,\sigma,\lambda}(s,x), t \geq s)$, $x \in \mathbb{R}$ on the time interval $[s,\infty)$ and with initial condition $Y_s^{\theta,\sigma,\lambda}(s,x) = x$ satisfying

$$Y_t^{\theta,\sigma,\lambda}(s,x) = x + \theta(t-s) + \sigma(W_t - W_s) + \widetilde{M}_t^\lambda - \widetilde{M}_s^\lambda, \tag{3.4}$$

where $W = (W_t)_{t\geq 0}$ is a Brownian motion, $M = (M_t)_{t\geq 0}$ is a Poisson process with intensity λ independent of W, and we denote by $(\widetilde{M}_t^\lambda)_{t\geq 0}$ the compensated Poisson process $\widetilde{M}_t^\lambda := M_t - \lambda t$. In particular, we write $Y_t^{\theta,\sigma,\lambda} \equiv Y_t^{\theta,\sigma,\lambda}(0,x_0)$, for all $t \geq 0$. That is,

$$Y_t^{\theta,\sigma,\lambda} = x_0 + \theta t + \sigma W_t + \widetilde{M}_t^\lambda. \tag{3.5}$$

For any $t > s$, we denote by $p^{\theta,\sigma,\lambda}(t-s,x,y)$ the transition density of $Y_t^{\theta,\sigma,\lambda}$ at y conditioned on $Y_s^{\theta,\sigma,\lambda} = x$. Here, we consider the Malliavin calculus on the Wiener space induced by the Brownian motion W, and we denote by D and δ the Malliavin derivative and the Skorohod integral with respect to W on each interval $[t_k, t_{k+1}]$, respectively (see the Definition 1.1.3 and the discussion following it).

For all $A \in \widetilde{\mathcal{F}}^\lambda$, let us denote $\widetilde{\mathrm{P}}_x^{\theta,\sigma,\lambda}(A) = \widetilde{\mathrm{E}}^\lambda[\mathbf{1}_A | Y_{t_k}^{\theta,\sigma,\lambda} = x]$. We denote by $\widetilde{\mathrm{E}}_x^{\theta,\sigma,\lambda}$ the expectation with respect to $\widetilde{\mathrm{P}}_x^{\theta,\sigma,\lambda}$. That is, for all $\widetilde{\mathcal{F}}^\lambda$-measurable random variable Z, we have that $\widetilde{\mathrm{E}}_x^{\theta,\sigma,\lambda}[Z] = \widetilde{\mathrm{E}}^\lambda[Z | Y_{t_k}^{\theta,\sigma,\lambda} = x]$.

Proposition 3.2.1. *[24, Proposition 4.1] For all $\theta \in \mathbb{R}, \sigma, \lambda \in \mathbb{R}_+^*$, and $k \in \{0, ..., n-1\}$,*

$$\frac{\partial_\theta p^{\theta,\sigma,\lambda}}{p^{\theta,\sigma,\lambda}}(\Delta_n, x, y) = \frac{1}{\sigma}\widetilde{\mathrm{E}}_x^{\theta,\sigma,\lambda}\left[W_{t_{k+1}} - W_{t_k}\middle|Y_{t_{k+1}}^{\theta,\sigma,\lambda} = y\right],$$
$$\frac{\partial_\sigma p^{\theta,\sigma,\lambda}}{p^{\theta,\sigma,\lambda}}(\Delta_n, x, y) = \frac{1}{\sigma\Delta_n}\widetilde{\mathrm{E}}_x^{\theta,\sigma,\lambda}\left[\left(W_{t_{k+1}} - W_{t_k}\right)^2\middle|Y_{t_{k+1}}^{\theta,\sigma,\lambda} = y\right] - \frac{1}{\sigma}.$$

Proof. Let f be a continuously differentiable function with compact support. The chain rule of the Malliavin calculus implies that $f'(Y_{t_{k+1}}^{\theta,\sigma,\lambda}(t_k,x)) = D_t(f(Y_{t_{k+1}}^{\theta,\sigma,\lambda}(t_k,x)))U_t^{\theta,\sigma,\lambda}(t_k,x)$, for all $(\theta,\sigma,\lambda) \in \Theta \times \Sigma \times \Lambda$ and $t \in [t_k, t_{k+1}]$, where

$$U_t^{\theta,\sigma,\lambda}(t_k,x) = \frac{1}{D_t Y_{t_{k+1}}^{\theta,\sigma,\lambda}(t_k,x)}.$$

.

Then, using the integration by parts formula of the Malliavin calculus on the interval $[t_k, t_{k+1}]$, we get that

$$\begin{aligned}
\partial_\theta \widetilde{\mathrm{E}}^\lambda\left[f(Y_{t_{k+1}}^{\theta,\sigma,\lambda}(t_k,x))\right] &= \widetilde{\mathrm{E}}^\lambda\left[f'(Y_{t_{k+1}}^{\theta,\sigma,\lambda}(t_k,x))\partial_\theta Y_{t_{k+1}}^{\theta,\sigma,\lambda}(t_k,x)\right]\\
&= \frac{1}{\Delta_n}\widetilde{\mathrm{E}}^\lambda\left[\int_{t_k}^{t_{k+1}} f'(Y_{t_{k+1}}^{\theta,\sigma,\lambda}(t_k,x))\partial_\theta Y_{t_{k+1}}^{\theta,\sigma,\lambda}(t_k,x)dt\right]\\
&= \frac{1}{\Delta_n}\widetilde{\mathrm{E}}^\lambda\left[\int_{t_k}^{t_{k+1}} D_t(f(Y_{t_{k+1}}^{\theta,\sigma,\lambda}(t_k,x)))U_t^{\theta,\sigma,\lambda}(t_k,x)\partial_\theta Y_{t_{k+1}}^{\theta,\sigma,\lambda}(t_k,x)dt\right]\\
&= \frac{1}{\Delta_n}\widetilde{\mathrm{E}}^\lambda\left[f(Y_{t_{k+1}}^{\theta,\sigma,\lambda}(t_k,x))\delta\left(\partial_\theta Y_{t_{k+1}}^{\theta,\sigma,\lambda}(t_k,x)U^{\theta,\sigma,\lambda}(t_k,x)\right)\right].
\end{aligned}$$

Note that here $\delta(V) \equiv \delta(V\mathbf{1}_{[t_k,t_{k+1}]}(\cdot))$ for any $V \in \mathrm{Dom}\ \delta$. On the other hand,

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

and

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

which concludes that

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

It can be checked that

$$\partial_\theta Y_{t_{k+1}}^{\theta,\sigma,\lambda}(t_k,x) = \Delta_n, \quad \text{and} \quad U_t^{\theta,\sigma,\lambda}(t_k,x) = \frac{1}{D_t Y_{t_{k+1}}^{\theta,\sigma,\lambda}(t_k,x)} = \frac{1}{\sigma}\mathbf{1}_{[t_k,t_{k+1}]}(t).$$

Therefore,

$$\delta\left(\partial_\theta Y_{t_{k+1}}^{\theta,\sigma,\lambda}(t_k,x)U^{\theta,\sigma,\lambda}(t_k,x)\right) = \frac{\Delta_n}{\sigma}\left(W_{t_{k+1}} - W_{t_k}\right),$$

which shows the first equality.

Similarly,

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

where

$$\partial_\sigma Y_{t_{k+1}}^{\theta,\sigma,\lambda}(t_k,x)=W_{t_{k+1}}-W_{t_k}.$$

Then,

$$\begin{aligned}\delta\left(\partial_\sigma Y_{t_{k+1}}^{\theta,\sigma,\lambda}(t_k,x)U^{\theta,\sigma,\lambda}(t_k,x)\right)&=\partial_\sigma Y_{t_{k+1}}^{\theta,\sigma,\lambda}(t_k,x)\delta\left(\frac{1}{\sigma}\mathbf{1}_{[t_k,t_{k+1}]}(\cdot)\right)\\&\quad-\int_{t_k}^{t_{k+1}}D_s\left(\partial_\sigma Y_{t_{k+1}}^{\theta,\sigma,\lambda}(t_k,x)\right)\frac{1}{\sigma}\mathbf{1}_{[t_k,t_{k+1}]}(s)ds\\&=\frac{1}{\sigma}\left(W_{t_{k+1}}-W_{t_k}\right)^2-\frac{\Delta_n}{\sigma},\end{aligned}$$

from where the second expression follows. □

We next recall Girsanov's theorem on each interval $[t_k,t_{k+1}]$.

Lemma 3.2.1. *For all $\theta,\theta_1\in\mathbb{R},\lambda,\lambda_1,\sigma\in\mathbb{R}_+^*$ and $k\in\{0,...,n-1\}$, define a measure $\widehat{Q}_k^{\theta_1,\lambda_1,\theta,\lambda,\sigma}$ by*

$$\widehat{Q}_k^{\theta_1,\lambda_1,\theta,\lambda,\sigma}(A)=\widehat{\mathrm{E}}^\lambda\left[\mathbf{1}_Ae^{-\frac{\theta-\theta_1-(\lambda-\lambda_1)}{\sigma}\left(B_{t_{k+1}}-B_{t_k}\right)+\frac{1}{2}\left(\frac{\theta-\theta_1-(\lambda-\lambda_1)}{\sigma}\right)^2\Delta_n-\left(N_{t_{k+1}}-N_{t_k}\right)\log\frac{\lambda}{\lambda_1}+(\lambda-\lambda_1)\Delta_n}\right],$$

for all $A\in\widehat{\mathcal{F}}^\lambda$. Then $\widehat{Q}_k^{\theta_1,\lambda_1,\theta,\lambda,\sigma}$ is a probability measure and under $\widehat{Q}_k^{\theta_1,\lambda_1,\theta,\lambda,\sigma}$, the process $B_t^{\widehat{Q}_k}=B_t+\frac{\theta-\theta_1-(\lambda-\lambda_1)}{\sigma}(t-t_k)$ is a Brownian motion, and N_t is a Poisson process with intensity λ_1, for all $t\in[t_k,t_{k+1}]$.

Lemma 3.2.2. *For all $\theta,\theta_1\in\mathbb{R},\lambda,\lambda_1,\sigma\in\mathbb{R}_+^*$ and $k\in\{0,...,n-1\}$, define a measure $\widetilde{Q}_k^{\theta_1,\lambda_1,\theta,\lambda,\sigma}$ by*

$$\widetilde{Q}_k^{\theta_1,\lambda_1,\theta,\lambda,\sigma}(A)=\widetilde{\mathrm{E}}^\lambda\left[\mathbf{1}_Ae^{-\frac{\theta-\theta_1-(\lambda-\lambda_1)}{\sigma}\left(W_{t_{k+1}}-W_{t_k}\right)+\frac{1}{2}\left(\frac{\theta-\theta_1-(\lambda-\lambda_1)}{\sigma}\right)^2\Delta_n-\left(M_{t_{k+1}}-M_{t_k}\right)\log\frac{\lambda}{\lambda_1}+(\lambda-\lambda_1)\Delta_n}\right],$$

for all $A\in\widetilde{\mathcal{F}}^\lambda$. Then $\widetilde{Q}_k^{\theta_1,\lambda_1,\theta,\lambda,\sigma}$ is a probability measure and under $\widetilde{Q}_k^{\theta_1,\lambda_1,\theta,\lambda,\sigma}$, the process $W_t^{\widetilde{Q}_k}=W_t+\frac{\theta-\theta_1-(\lambda-\lambda_1)}{\sigma}(t-t_k)$ is a Brownian motion, and M_t is a Poisson process with intensity λ_1, for all $t\in[t_k,t_{k+1}]$.

As a consequence, we have the following expression for the derivative of the log-likelihood w.r.t. λ.

Proposition 3.2.2. *For all $\theta\in\mathbb{R},\sigma,\lambda\in\mathbb{R}_+^*$, and $k\in\{0,...,n-1\}$,*

$$\frac{\partial_\lambda p^{\theta,\sigma,\lambda}}{p^{\theta,\sigma,\lambda}}(\Delta_n,x,y)=\widetilde{\mathrm{E}}_x^{\theta,\sigma,\lambda}\left[-\frac{W_{t_{k+1}}-W_{t_k}}{\sigma}+\frac{\widetilde{M}_{t_{k+1}}^\lambda-\widetilde{M}_{t_k}^\lambda}{\lambda}\Big|Y_{t_{k+1}}^{\theta,\sigma,\lambda}=y\right].$$

Proof. Let f be a continuously differentiable bounded function. Girsanov's theorem yields

$$\widetilde{\mathrm{E}}^\lambda\left[f\left(Y_{t_{k+1}}^{\theta,\sigma,\lambda}(t_k,x)\right)\right]=\widetilde{\mathrm{E}}^{\lambda_1}\left[f\left(Y_{t_{k+1}}^{\theta,\sigma,\lambda_1}(t_k,x)\right)\frac{d\widetilde{\mathrm{P}}^\lambda}{d\widetilde{Q}_k^{\theta,\lambda_1,\theta,\lambda,\sigma}}\right].$$

Taking the derivative with respect to λ in both hand sides of this equality and using Lemma 3.2.2, we get that

$$\begin{aligned}
&\partial_\lambda \widetilde{\mathrm{E}}^\lambda \left[f\left(Y^{\theta,\sigma,\lambda}_{t_{k+1}}(t_k,x)\right)\right] = \widetilde{\mathrm{E}}^{\lambda_1}\left[f\left(Y^{\theta,\sigma,\lambda_1}_{t_{k+1}}(t_k,x)\right)\partial_\lambda\left(\frac{d\widetilde{\mathrm{P}}^\lambda}{d\widetilde{Q}_k^{\theta,\lambda_1,\theta,\lambda,\sigma}}\right)\right]\\
&= \widetilde{\mathrm{E}}^{\lambda_1}\left[f\left(Y^{\theta,\sigma,\lambda_1}_{t_{k+1}}(t_k,x)\right)\left(-\frac{W_{t_{k+1}}-W_{t_k}}{\sigma} - \frac{\lambda-\lambda_1}{\sigma^2}\Delta_n + \frac{\widetilde{M}^\lambda_{t_{k+1}}-\widetilde{M}^\lambda_{t_k}}{\lambda}\right)\frac{d\widetilde{\mathrm{P}}^\lambda}{d\widetilde{Q}_k^{\theta,\lambda_1,\theta,\lambda,\sigma}}\right]\\
&= \widetilde{\mathrm{E}}^{\lambda}\left[f\left(Y^{\theta,\sigma,\lambda}_{t_{k+1}}(t_k,x)\right)\left(-\frac{W_{t_{k+1}}-W_{t_k}}{\sigma} + \frac{\widetilde{M}^\lambda_{t_{k+1}}-\widetilde{M}^\lambda_{t_k}}{\lambda}\right)\right].
\end{aligned}$$

On the other hand,

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

and

$$\begin{aligned}
&\widetilde{\mathrm{E}}^{\lambda}\left[f\left(Y^{\theta,\sigma,\lambda}_{t_{k+1}}(t_k,x)\right)\left(-\frac{W_{t_{k+1}}-W_{t_k}}{\sigma} + \frac{\widetilde{M}^\lambda_{t_{k+1}}-\widetilde{M}^\lambda_{t_k}}{\lambda}\right)\right]\\
&= \widetilde{\mathrm{E}}^{\lambda}\left[f\left(Y^{\theta,\sigma,\lambda}_{t_{k+1}}\right)\left(-\frac{W_{t_{k+1}}-W_{t_k}}{\sigma} + \frac{\widetilde{M}^\lambda_{t_{k+1}}-\widetilde{M}^\lambda_{t_k}}{\lambda}\right)\Big| Y^{\theta,\sigma,\lambda}_{t_k}=x\right]\\
&= \int_{\mathbb{R}} f(y)\widetilde{\mathrm{E}}^{\lambda}\left[-\frac{W_{t_{k+1}}-W_{t_k}}{\sigma} + \frac{\widetilde{M}^\lambda_{t_{k+1}}-\widetilde{M}^\lambda_{t_k}}{\lambda}\Big| Y^{\theta,\sigma,\lambda}_{t_{k+1}}=y, Y^{\theta,\sigma,\lambda}_{t_k}=x\right]p^{\theta,\sigma,\lambda}(\Delta_n,x,y)dy.
\end{aligned}$$

Therefore, the desired result follows. □

The next four lemmas are the main technical core of the chapter. It explains the argument given at the end of the Introduction.

Consider the events $\widehat{J}_{m,k} := \{N_{t_{k+1}} - N_{t_k} = m\}$ and $\widetilde{J}_{m,k} := \{M_{t_{k+1}} - M_{t_k} = m\}$, for all $m \geq 0$ and $k \in \{0, ..., n-1\}$.

Lemma 3.2.3. *For all $\theta \in \mathbb{R}, \sigma, \lambda \in \mathbb{R}^*_+,\ k \in \{0, ..., n-1\},$ and $m \geq 0$,*

$$\widetilde{\mathrm{P}}^{\theta,\sigma,\lambda}_x\left(\widetilde{J}_{m,k}\Big| Y^{\theta,\sigma,\lambda}_{t_{k+1}}=y\right) = \frac{e^{-(y-x-m-(\theta-\lambda)\Delta_n)^2/(2\sigma^2\Delta_n)}\frac{(\lambda\Delta_n)^m}{m!}}{\sum_{i=0}^\infty e^{-(y-x-i-(\theta-\lambda)\Delta_n)^2/(2\sigma^2\Delta_n)}\frac{(\lambda\Delta_n)^i}{i!}}.$$

Proof. For all $i \geq 0$ and $t > s$, we denote by $q^{\theta,\sigma,\lambda}_{(i)}(t-s,x,y)$ the transition density of $Y^{\theta,\sigma,\lambda}_t$ conditioned on $Y^{\theta,\sigma,\lambda}_s = x$ and $M_t - M_s = i$. That is,

$$q^{\theta,\sigma,\lambda}_{(i)}(t-s,x,y) = \frac{1}{\sqrt{2\pi\sigma^2(t-s)}}e^{-\frac{(y-x-i-(\theta-\lambda)(t-s))^2}{2\sigma^2(t-s)}}.$$

By the convolution formula for the sum of independent random variables, we obtain

$$p^{\theta,\sigma,\lambda}(t-s,x,y) = \sum_{i=0}^\infty q^{\theta,\sigma,\lambda}_{(i)}(t-s,x,y)e^{-\lambda(t-s)}\frac{(\lambda(t-s))^i}{i!}.$$

Then, using Bayes' formula, we get that

$$\widetilde{\mathrm{P}}^{\theta,\sigma,\lambda}_x\left(\widetilde{J}_{m,k}\Big| Y^{\theta,\sigma,\lambda}_{t_{k+1}}=y\right) = \frac{q^{\theta,\sigma,\lambda}_{(m)}(\Delta_n,x,y)\widetilde{\mathrm{P}}^{\theta,\sigma,\lambda}_x(\widetilde{J}_{m,k})}{p^{\theta,\sigma,\lambda}(\Delta_n,x,y)}.$$

Hence, the desired result follows. □

For all $j,p \geq 0$ and $k \in \{0,...,n-1\}$, we introduce the random variable

$$S_j^p := \mathbf{1}_{\widehat{J}_{j,k}} \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta,\sigma,\lambda}\left[\left(M_{t_{k+1}}-M_{t_k}\right)^p \mathbf{1}_{\widetilde{J}_{j,k}^c}\middle| Y_{t_{k+1}}^{\theta,\sigma,\lambda}=X_{t_{k+1}}\right].$$

Lemma 3.2.4. *For all $\theta \in \mathbb{R}, \sigma, \lambda \in \mathbb{R}_+^*, j,p \geq 0$ and $k \in \{0,...,n-1\}$,*

$$S_j^p = \mathbf{1}_{\widehat{J}_{j,k}} \frac{\sum_{m=0:m\neq j}^{\infty} m^p e^{-\left(\sigma_0\left(B_{t_{k+1}}-B_{t_k}\right)+j-m+(\theta_0-\theta-\lambda_0+\lambda)\Delta_n\right)^2/(2\sigma^2\Delta_n)} \frac{(\lambda\Delta_n)^m}{m!}}{\sum_{i=0}^{\infty} e^{-\left(\sigma_0\left(B_{t_{k+1}}-B_{t_k}\right)+j-i+(\theta_0-\theta-\lambda_0+\lambda)\Delta_n\right)^2/(2\sigma^2\Delta_n)} \frac{(\lambda\Delta_n)^i}{i!}}. \tag{3.6}$$

Proof. Observe that $\widetilde{J}_{j,k}^c := \cup_{m=0:m\neq j}^{\infty}\{M_{t_{k+1}}-M_{t_k}=m\}$. Thus, appealing to Lemma 3.2.3 and equation (3.1), we get

$$\begin{aligned}
S_j^p &= \mathbf{1}_{\widehat{J}_{j,k}} \sum_{m=0:m\neq j}^{\infty} m^p \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta,\sigma,\lambda}\left[\mathbf{1}_{\widetilde{J}_{m,k}}\middle| Y_{t_{k+1}}^{\theta,\sigma,\lambda}=X_{t_{k+1}}\right]\\
&= \mathbf{1}_{\widehat{J}_{j,k}} \frac{\sum_{m=0:m\neq j}^{\infty} m^p e^{-\left(X_{t_{k+1}}-X_{t_k}-m-(\theta-\lambda)\Delta_n\right)^2/(2\sigma^2\Delta_n)} \frac{(\lambda\Delta_n)^m}{m!}}{\sum_{i=0}^{\infty} e^{-\left(X_{t_{k+1}}-X_{t_k}-i-(\theta-\lambda)\Delta_n\right)^2/(2\sigma^2\Delta_n)} \frac{(\lambda\Delta_n)^i}{i!}}\\
&= \mathbf{1}_{\widehat{J}_{j,k}} \frac{\sum_{m=0:m\neq j}^{\infty} m^p e^{-\left(\sigma_0\left(B_{t_{k+1}}-B_{t_k}\right)+j-m+(\theta_0-\theta-\lambda_0+\lambda)\Delta_n\right)^2/(2\sigma^2\Delta_n)} \frac{(\lambda\Delta_n)^m}{m!}}{\sum_{i=0}^{\infty} e^{-\left(\sigma_0\left(B_{t_{k+1}}-B_{t_k}\right)+j-i+(\theta_0-\theta-\lambda_0+\lambda)\Delta_n\right)^2/(2\sigma^2\Delta_n)} \frac{(\lambda\Delta_n)^i}{i!}},
\end{aligned}$$

which concludes the desired result. □

We next fix $\alpha \in (0,\frac{1}{2})$, and split S_j^p in two separate terms as follows

$$S_j^p = S_j^p \mathbf{1}_{\{|B_{t_{k+1}}-B_{t_k}|\leq \Delta_n^{\alpha}\}} + S_j^p \mathbf{1}_{\{|B_{t_{k+1}}-B_{t_k}|> \Delta_n^{\alpha}\}} =: S_{1,j}^p + S_{2,j}^p.$$

Furthermore, we write $S_{1,j}^p = S_{1,1,j}^p + S_{1,2,j}^p$, and $S_{2,j}^p = S_{2,1,j}^p + S_{2,2,j}^p$, where $S_{1,1,j}^p$ and $S_{2,1,j}^p$ contain the terms $\sum_{m<j}$, and $S_{1,2,j}^p$ and $S_{2,2,j}^p$ contain the terms $\sum_{m>j}$ in (3.6).

We have the following estimates.

Lemma 3.2.5. *Assume that $|\theta_0-\theta| \leq \frac{C}{\sqrt{n\Delta_n}}$ and $|\lambda_0-\lambda| \leq \frac{C}{\sqrt{n\Delta_n}}$, for some constant $C>0$. Then for all $\sigma \in \mathbb{R}_+^*, j,p \geq 0,\ k \in \{0,...,n-1\}$, and for n large enough,*

$$S_{1,1,j}^p \leq \mathbf{1}_{\widehat{J}_{j,k}} \frac{j!}{(\lambda\Delta_n)^j} \sum_{m<j} m^p e^{-\frac{(j-m)^2}{4\sigma^2\Delta_n}} \frac{(\lambda\Delta_n)^m}{m!}, \tag{3.7}$$

$$S_{1,2,j}^p \leq \mathbf{1}_{\widehat{J}_{j,k}} e^{-\frac{1}{4\sigma^2\Delta_n}} \sum_{\ell>0} (\ell+j)^p \frac{(\lambda\Delta_n)^\ell}{\ell!}, \tag{3.8}$$

$$S_{2,1,j}^p \leq j^p \mathbf{1}_{\widehat{J}_{j,k}} \mathbf{1}_{\{|B_{t_{k+1}}-B_{t_k}|> \Delta_n^{\alpha}\}}, \tag{3.9}$$

$$S_{2,2,j}^p \leq \mathbf{1}_{\widehat{J}_{j,k}} \mathbf{1}_{\{|B_{t_{k+1}}-B_{t_k}|> \Delta_n^{\alpha}\}} \sum_{\ell=0}^{\infty} (\ell+j+1)^p \frac{(\lambda\Delta_n)^\ell}{\ell!}, \tag{3.10}$$

where (3.9) and (3.10) hold for all $n \geq 1$.

Proof. By lower bounding the denominator by the term $i=j$, we obtain that

$$\begin{aligned}
S_{1,1,j}^p &\leq \mathbf{1}_{\{|B_{t_{k+1}}-B_{t_k}|\leq \Delta_n^{\alpha}\}} \mathbf{1}_{\widehat{J}_{j,k}} \frac{\sum_{m<j} m^p e^{-\left(\sigma_0\left(B_{t_{k+1}}-B_{t_k}\right)+j-m+(\theta_0-\theta-\lambda_0+\lambda)\Delta_n\right)^2/(2\sigma^2\Delta_n)} \frac{(\lambda\Delta_n)^m}{m!}}{e^{-\left(\sigma_0\left(B_{t_{k+1}}-B_{t_k}\right)+(\theta_0-\theta-\lambda_0+\lambda)\Delta_n\right)^2/(2\sigma^2\Delta_n)} \frac{(\lambda\Delta_n)^j}{j!}}\\
&= \mathbf{1}_{\{|B_{t_{k+1}}-B_{t_k}|\leq \Delta_n^{\alpha}\}} \mathbf{1}_{\widehat{J}_{j,k}} \frac{j!}{(\lambda\Delta_n)^j} \sum_{m<j} m^p e^{-\frac{(j-m)^2}{2\sigma^2\Delta_n}} e^{-\frac{2(j-m)}{2\sigma^2\Delta_n}(\sigma_0(B_{t_{k+1}}-B_{t_k})+(\theta_0-\theta-\lambda_0+\lambda)\Delta_n)} \frac{(\lambda\Delta_n)^m}{m!}.
\end{aligned}$$

Conditionally on the event $\{|B_{t_{k+1}}-B_{t_k}|\leq \Delta_n^\alpha\}$ and choosing n large enough, we have that

$$\left|\sigma_0(B_{t_{k+1}}-B_{t_k})+(\theta_0-\theta-\lambda_0+\lambda)\Delta_n\right|\leq \sigma\Delta_n^\alpha+C\frac{\sqrt{\Delta_n}}{\sqrt{n}}\leq \frac{j-m}{4},$$

for some positive constant C. Thus, (3.7) holds true.

Now, observing that the function $e^{-(\sigma_0(B_{t_{k+1}}-B_{t_k})+j-m+(\theta_0-\theta-\lambda_0+\lambda)\Delta_n)^2/(2\sigma^2\Delta_n)}$ is decreasing w.r.t. m, we can upper bound each term by the term $m=j+1$ in the sum $\sum_{m>j}$. Moreover, lower bounding again the denominator by the term $i=j$, it follows that

$$\begin{aligned}S_{1,2,j}^p &\leq \mathbf{1}_{\{|B_{t_{k+1}}-B_{t_k}|\leq \Delta_n^\alpha\}}\mathbf{1}_{\widehat{J}_{j,k}}\frac{e^{-\left(\sigma_0\left(B_{t_{k+1}}-B_{t_k}\right)-1+(\theta_0-\theta-\lambda_0+\lambda)\Delta_n\right)^2/(2\sigma^2\Delta_n)}}{e^{-\left(\sigma_0\left(B_{t_{k+1}}-B_{t_k}\right)+(\theta_0-\theta-\lambda_0+\lambda)\Delta_n\right)^2/(2\sigma^2\Delta_n)}}\\
&\quad\times\sum_{m>j}m^p\frac{(\lambda\Delta_n)^{m-j}}{(m-j)!}\frac{j!(m-j)!}{m!}\\
&\leq \mathbf{1}_{\{|B_{t_{k+1}}-B_{t_k}|\leq \Delta_n^\alpha\}}\mathbf{1}_{\widehat{J}_{j,k}}e^{-\frac{1}{2\sigma^2\Delta_n}\left(1-2\sigma_0\left(B_{t_{k+1}}-B_{t_k}\right)-2(\theta_0-\theta-\lambda_0+\lambda)\Delta_n\right)}\sum_{\ell=1}^\infty(\ell+j)^p\frac{(\lambda\Delta_n)^\ell}{\ell!},\end{aligned}$$

where we have used the fact that $\frac{j!(m-j)!}{m!}\leq 1$.

Conditionally on the event $\{|B_{t_{k+1}}-B_{t_k}|\leq \Delta_n^\alpha\}$ and choosing n large enough, we get that

$$\left|\sigma_0\left(B_{t_{k+1}}-B_{t_k}\right)+(\theta_0-\theta-\lambda_0+\lambda)\Delta_n\right|\leq \sigma_0\Delta_n^\alpha+C\frac{\sqrt{\Delta_n}}{\sqrt{n}}\leq \frac{1}{4},$$

for some positive constant C, from where we deduce (3.8).

Inequality (3.9) follows easily bounding m^p by j^p. Thus, it remains to show (3.10), which follows similarly as in (3.8) but choosing the term $i=j+1$ to lower bound the denominator, which yields

$$\begin{aligned}S_{2,2,j}^p &\leq \mathbf{1}_{\{|B_{t_{k+1}}-B_{t_k}|>\Delta_n^\alpha\}}\mathbf{1}_{\widehat{J}_{j,k}}\frac{e^{-\left(\sigma_0\left(B_{t_{k+1}}-B_{t_k}\right)-1+(\theta_0-\theta-\lambda_0+\lambda)\Delta_n\right)^2/(2\sigma^2\Delta_n)}\sum_{m>j}m^p\frac{(\lambda\Delta_n)^m}{m!}}{e^{-\left(\sigma_0\left(B_{t_{k+1}}-B_{t_k}\right)-1+(\theta_0-\theta-\lambda_0+\lambda)\Delta_n\right)^2/(2\sigma^2\Delta_n)}\frac{(\lambda\Delta_n)^{j+1}}{(j+1)!}}\\
&= \mathbf{1}_{\{|B_{t_{k+1}}-B_{t_k}|>\Delta_n^\alpha\}}\mathbf{1}_{\widehat{J}_{j,k}}\sum_{\ell=0}^\infty(\ell+j+1)^p\frac{(\lambda\Delta_n)^\ell}{\ell!}\frac{\ell!(j+1)!}{(\ell+j+1)!}\\
&\leq \mathbf{1}_{\{|B_{t_{k+1}}-B_{t_k}|>\Delta_n^\alpha\}}\mathbf{1}_{\widehat{J}_{j,k}}\sum_{\ell=0}^\infty(\ell+j+1)^p\frac{(\lambda\Delta_n)^\ell}{\ell!},\end{aligned}$$

where we have used the fact that $\frac{\ell!(j+1)!}{(\ell+j+1)!}\leq 1$. □

For all $p\geq 0$ and $k\in\{0,...,n-1\}$, set

$$\begin{aligned}M_{1,p}^{\theta,\sigma,\lambda}&:=\sum_{j=0}^\infty j^p\mathrm{E}\left[\mathbf{1}_{\widehat{J}_{j,k}}\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta,\sigma,\lambda}\left[\mathbf{1}_{\widetilde{J}_{j,k}^c}\middle|Y_{t_{k+1}}^{\theta,\sigma,\lambda}=X_{t_{k+1}}\right]\middle|X_{t_k}\right],\\
M_{2,p}^{\theta,\sigma,\lambda}&:=\sum_{j=0}^\infty \mathrm{E}\left[\mathbf{1}_{\widehat{J}_{j,k}}\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta,\sigma,\lambda}\left[\left(M_{t_{k+1}}-M_{t_k}\right)^p\mathbf{1}_{\widetilde{J}_{j,k}^c}\middle|Y_{t_{k+1}}^{\theta,\sigma,\lambda}=X_{t_{k+1}}\right]\middle|X_{t_k}\right].\end{aligned}$$

Lemma 3.2.6. *Assume that $|\theta_0-\theta|\leq \frac{C}{\sqrt{n\Delta_n}}$ and $|\lambda_0-\lambda|\leq\frac{C}{\sqrt{n\Delta_n}}$, for some constant $C>0$. Then, for all $\sigma\in\Sigma$, $p\geq 0$, and n large enough, there exist constants $C_1,C_2>0$ such that for any $\alpha\in(0,\frac{1}{2})$, $k\in\{0,...,n-1\}$,*

$$M_{1,p}^{\theta,\sigma,\lambda}+M_{2,p}^{\theta,\sigma,\lambda}\leq C_1e^{-\frac{1}{C_2\Delta_n^{1-2\alpha}}}.$$

Proof. We can write

$$M_{1,p}^{\theta,\sigma,\lambda} = \sum_{j=0}^{\infty} j^p \mathrm{E}\left[S_j^0 | X_{t_k}\right] = \sum_{j=0}^{\infty} j^p \mathrm{E}\left[S_{1,1,j}^0 + S_{1,2,j}^0 + S_{2,1,j}^0 + S_{2,2,j}^0 \Big| X_{t_k}\right].$$

Set $a_{\lambda_0} := 2\sigma^2 \Delta_n \log(\lambda_0 \lambda^{-1})$ and appeal to (3.7) to get that

$$\begin{aligned}
\sum_{j=0}^{\infty} j^p \mathrm{E}\left[S_{1,1,j}^0 | X_{t_k}\right] &\leq \sum_{j=0}^{\infty} j^p \mathrm{E}\left[\mathbf{1}_{\widehat{J}_{j,k}} \frac{j!}{(\lambda\Delta_n)^j} \sum_{m<j} e^{-\frac{(j-m)^2}{4\sigma^2\Delta_n}} \frac{(\lambda\Delta_n)^m}{m!} \Big| X_{t_k}\right] \\
&= \sum_{j=0}^{\infty} j^p e^{-\lambda_0\Delta_n} \frac{(\lambda_0\Delta_n)^j}{j!} \frac{j!}{(\lambda\Delta_n)^j} \sum_{m<j} e^{-\frac{(j-m)^2}{4\sigma^2\Delta_n}} \frac{(\lambda\Delta_n)^m}{m!} \\
&= e^{(\lambda-\lambda_0)\Delta_n} \sum_{m=0}^{\infty} \sum_{j=m+1}^{\infty} j^p e^{-\frac{(j-m)^2}{4\sigma^2\Delta_n}} \left(\frac{\lambda_0}{\lambda}\right)^j e^{-\lambda\Delta_n} \frac{(\lambda\Delta_n)^m}{m!} \\
&= e^{(\lambda-\lambda_0)\Delta_n} \sum_{m=0}^{\infty} \sum_{j=m+1}^{\infty} j^p e^{-\frac{(j-m-a_{\lambda_0})^2}{4\sigma^2\Delta_n}} e^{\frac{2a_{\lambda_0}m + a_{\lambda_0}^2}{4\sigma^2\Delta_n}} e^{-\lambda\Delta_n} \frac{(\lambda\Delta_n)^m}{m!} \\
&\leq e^{(\lambda-\lambda_0)\Delta_n} \sum_{m=0}^{\infty} \int_{m+1}^{\infty} x^p e^{-\frac{(x-m-a_{\lambda_0})^2}{4\sigma^2\Delta_n}} dx \, e^{\frac{2a_{\lambda_0}m + a_{\lambda_0}^2}{4\sigma^2\Delta_n}} e^{-\lambda\Delta_n} \frac{(\lambda\Delta_n)^m}{m!} \\
&\overset{u\sqrt{2\sigma^2\Delta_n} := x-m-a_{\lambda_0}}{=} e^{(\lambda-\lambda_0)\Delta_n} \sqrt{2\sigma^2\Delta_n} \sum_{m=0}^{\infty} \int_{\frac{1-a_{\lambda_0}}{\sqrt{2\sigma^2\Delta_n}}}^{\infty} (u\sqrt{2\sigma^2\Delta_n} + m + a_{\lambda_0})^p e^{-\frac{u^2}{2}} du \\
&\qquad \times e^{\frac{2a_{\lambda_0}m + a_{\lambda_0}^2}{4\sigma^2\Delta_n}} e^{-\lambda\Delta_n} \frac{(\lambda\Delta_n)^m}{m!} \\
&\leq C\Phi\left(-\frac{1-a_{\lambda_0}}{\sqrt{2\sigma^2\Delta_n}}\right) \leq Ce^{-\frac{1}{c\Delta_n}},
\end{aligned}$$

for some constants $c, C > 0$, where $\Phi(\cdot)$ denotes the distribution function of the standard normal distribution and we have used the fact that $\frac{a_{\lambda_0}}{\Delta_n} = 2\sigma^2 \log(\lambda_0\lambda^{-1})$.

On the other hand, using (3.8),

$$\sum_{j=0}^{\infty} j^p \mathrm{E}\left[S_{1,2,j}^0 | X_{t_k}\right] \leq e^{-\frac{1}{4\sigma^2\Delta_n}} \sum_{\ell>0} \frac{(\lambda\Delta_n)^\ell}{\ell!} \sum_{j=0}^{\infty} j^p \mathrm{E}\left[\mathbf{1}_{\widehat{J}_{j,k}} \Big| X_{t_k}\right] \leq Ce^{-\frac{1}{c\Delta_n}},$$

for some constants $c, C > 0$.

Moreover, using the independence between N and B and (3.9), we get

$$\begin{aligned}
\sum_{j=0}^{\infty} j^p \mathrm{E}\left[S_{2,1,j}^0 | X_{t_k}\right] &\leq \sum_{j=0}^{\infty} j^p \mathrm{E}\left[\mathbf{1}_{\widehat{J}_{j,k}} \mathbf{1}_{\{|B_{t_{k+1}} - B_{t_k}| > \Delta_n^\alpha\}} \Big| X_{t_k}\right] \leq C\Phi\left(-\frac{\Delta_n^\alpha}{\sqrt{\Delta_n}}\right) \\
&\leq Ce^{-\frac{1}{c\Delta_n^{1-2\alpha}}},
\end{aligned}$$

for some constants $c, C > 0$.

Finally, (3.10) yields

$$\sum_{j=0}^{\infty} j^p \mathrm{E}\left[S_{2,2,j}^0 | X_{t_k}\right] \leq \sum_{\ell=0}^{\infty} \frac{(\lambda\Delta_n)^\ell}{\ell!} \sum_{j=0}^{\infty} j^p \mathrm{E}\left[\mathbf{1}_{\widehat{J}_{j,k}} \mathbf{1}_{\{|B_{t_{k+1}} - B_{t_k}| > \Delta_n^\alpha\}} \Big| X_{t_k}\right] \leq Ce^{-\frac{1}{c\Delta_n^{1-2\alpha}}},$$

for some constants $c, C > 0$. This shows the estimate for $M_{1,p}^{\theta,\sigma,\lambda}$.

We next treat $M_{2,p}^{\theta,\sigma,\lambda}$. Observe that

$$M_{2,p}^{\theta,\sigma,\lambda} = \sum_{j=0}^{\infty} \mathrm{E}\left[S_j^p | X_{t_k}\right] = \sum_{j=0}^{\infty} \mathrm{E}\left[S_{1,1,j}^p + S_{1,2,j}^p + S_{2,1,j}^p + S_{2,2,j}^p \Big| X_{t_k}\right].$$

Proceeding as for the term $M_{1,p}^{\theta,\sigma,\lambda}$, we conclude the desired result. □

The next technical lemma will be used several times in the sequel.

Lemma 3.2.7. *For all $w \in \mathbb{R}$, and $k \in \{0, ..., n-1\}$,*

$$\mathrm{E}\left[\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\sigma_0,\lambda(\ell)}\left[\widetilde{M}_{t_{k+1}}^{\lambda(\ell)} - \widetilde{M}_{t_k}^{\lambda(\ell)}\Big| Y_{t_{k+1}}^{\theta_n,\sigma_0,\lambda(\ell)} = X_{t_{k+1}}\right]\Big| X_{t_k}\right] = -\frac{\ell w}{\sqrt{n\Delta_n}}\Delta_n,$$

where $\lambda(\ell) := \lambda_0 + \dfrac{\ell w}{\sqrt{n\Delta_n}}$ *and* $\ell \in [0,1]$.

Proof. Using Lemma 3.2.1, we have

$$\begin{aligned}
&\mathrm{E}\left[\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\sigma_0,\lambda(\ell)}\left[\widetilde{M}_{t_{k+1}}^{\lambda(\ell)} - \widetilde{M}_{t_k}^{\lambda(\ell)}\Big| Y_{t_{k+1}}^{\theta_n,\sigma_0,\lambda(\ell)} = X_{t_{k+1}}\right]\Big| X_{t_k}\right]\\
&= \mathrm{E}_{\widehat{Q}_k}^{\lambda(\ell)}\left[\left(\widetilde{M}_{t_{k+1}}^{\lambda(\ell)} - \widetilde{M}_{t_k}^{\lambda(\ell)}\right)\frac{d\widehat{\mathrm{P}}^{\lambda_0}}{d\widehat{Q}_k^{\theta_n,\lambda(\ell),\theta_0,\lambda_0,\sigma_0}}\Big| X_{t_k}\right]\\
&= \mathrm{E}_{\widehat{Q}_k}^{\lambda(\ell)}\left[\left(\widetilde{M}_{t_{k+1}}^{\lambda(\ell)} - \widetilde{M}_{t_k}^{\lambda(\ell)}\right) e^{\frac{-u+\ell w}{\sigma\sqrt{n\Delta_n}}(B_{t_{k+1}}-B_{t_k}) - \frac{(-u+\ell w)^2}{2\sigma^2 n} + (N_{t_{k+1}}-N_{t_k})\log\frac{\lambda_0}{\lambda(\ell)} + \frac{\ell w\Delta_n}{\sqrt{n\Delta_n}}}\Big| X_{t_k}\right]\\
&= \mathrm{E}\left[\widetilde{M}_{t_{k+1}}^{\lambda(\ell)} - \widetilde{M}_{t_k}^{\lambda(\ell)}\right]\mathrm{E}_{\widehat{Q}_k}^{\lambda(\ell)}\left[e^{\frac{-u+\ell w}{\sigma\sqrt{n\Delta_n}}(B_{t_{k+1}}-B_{t_k}) - \frac{(-u+\ell w)^2}{2\sigma^2 n} + (N_{t_{k+1}}-N_{t_k})\log\frac{\lambda_0}{\lambda(\ell)} + \frac{\ell w\Delta_n}{\sqrt{n\Delta_n}}}\right]\\
&= -\frac{\ell w}{\sqrt{n\Delta_n}}\Delta_n,
\end{aligned}$$

where the second expectation equals 1 and $\widehat{Q}_k \equiv \widehat{Q}_k^{\theta_n,\lambda(\ell),\theta_0,\lambda_0,\sigma_0}$. Here we have used the independence between $M_{t_{k+1}} - M_{t_k}$, $N_{t_{k+1}} - N_{t_k}$, $B_{t_{k+1}} - B_{t_k}$, and X_{t_k}. Thus, the result follows. □

3.3 Proof of Theorem 3.1.1

In this section, the proof of Theorem 3.1.1 will be divided into several steps. Recall the decomposition in (3.3). Then we begin deriving a stochastic expansion of the log-likelihood ratio using Propositions 3.2.1 and 3.2.2. The second step is devoted to treat the negligible contributions of this expansion. Finally, the last step concludes the LAN property by applying the central limit theorem for triangular arrays.

3.3.1 Expansion of the log-likelihood ratio

For $\ell \in [0,1]$, set $\theta(\ell) := \theta_n(\ell, u) := \theta_0 + \dfrac{\ell u}{\sqrt{n\Delta_n}}$, $\sigma(\ell) := \sigma_n(\ell, v) := \sigma_0 + \dfrac{\ell v}{\sqrt{n}}$, $\lambda(\ell) := \lambda_n(\ell, w) := \lambda_0 + \dfrac{\ell w}{\sqrt{n\Delta_n}}$. Then, from the Markov property and Proposition 3.2.1,

$$\begin{aligned}
\log\frac{p\left(X^n; (\theta_n, \sigma_0, \lambda_0)\right)}{p\left(X^n; (\theta_0, \sigma_0, \lambda_0)\right)} &= \sum_{k=0}^{n-1}\log\frac{p^{\theta_n,\sigma_0,\lambda_0}}{p^{\theta_0,\sigma_0,\lambda_0}}\left(\Delta_n, X_{t_k}, X_{t_{k+1}}\right)\\
&= \sum_{k=0}^{n-1}\frac{u}{\sqrt{n\Delta_n}}\int_0^1\frac{\partial_\theta p^{\theta(\ell),\sigma_0,\lambda_0}}{p^{\theta(\ell),\sigma_0,\lambda_0}}\left(\Delta_n, X_{t_k}, X_{t_{k+1}}\right)d\ell\\
&= \sum_{k=0}^{n-1}\frac{u}{\sqrt{n\Delta_n}}\frac{1}{\sigma_0}\int_0^1\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell),\sigma_0,\lambda_0}\left[W_{t_{k+1}} - W_{t_k}\Big| Y_{t_{k+1}}^{\theta(\ell),\sigma_0,\lambda_0} = X_{t_{k+1}}\right]d\ell.
\end{aligned}$$

Equation (3.5) yields that

$$\sigma_0(W_{t_{k+1}} - W_{t_k}) = Y^{\theta(\ell),\sigma_0,\lambda_0}_{t_{k+1}} - Y^{\theta(\ell),\sigma_0,\lambda_0}_{t_k} - \theta(\ell)\Delta_n - \left(\widetilde{M}^{\lambda_0}_{t_{k+1}} - \widetilde{M}^{\lambda_0}_{t_k}\right), \tag{3.11}$$

which gives

$$\log \frac{p\left(X^n; (\theta_n, \sigma_0, \lambda_0)\right)}{p\left(X^n; (\theta_0, \sigma_0, \lambda_0)\right)} = \sum_{k=0}^{n-1} (\xi_{k,n} + H_{k,n}),$$

where

$$\begin{aligned}
\xi_{k,n} &:= \frac{u}{\sqrt{n\Delta_n}} \frac{1}{\sigma_0^2} \left(\sigma_0 \left(B_{t_{k+1}} - B_{t_k}\right) - \frac{u\Delta_n}{2\sqrt{n\Delta_n}} \right), \\
H_{k,n} &:= \frac{u}{\sqrt{n\Delta_n}} \frac{1}{\sigma_0^2} \left(\widetilde{N}^{\lambda_0}_{t_{k+1}} - \widetilde{N}^{\lambda_0}_{t_k} - \int_0^1 \widetilde{\mathrm{E}}^{\theta(\ell),\sigma_0,\lambda_0}_{X_{t_k}} \left[\widetilde{M}^{\lambda_0}_{t_{k+1}} - \widetilde{M}^{\lambda_0}_{t_k} \middle| Y^{\theta(\ell),\sigma_0,\lambda_0}_{t_{k+1}} = X_{t_{k+1}} \right] d\ell \right).
\end{aligned}$$

Again the Markov property and Proposition 3.2.1 give

$$\begin{aligned}
&\log \frac{p\left(X^n; (\theta_n, \sigma_n, \lambda_n)\right)}{p\left(X^n; (\theta_n, \sigma_0, \lambda_n)\right)} = \sum_{k=0}^{n-1} \frac{v}{\sqrt{n}} \int_0^1 \frac{\partial_\sigma p^{\theta_n,\sigma(\ell),\lambda_n}}{p^{\theta_n,\sigma(\ell),\lambda_n}} \left(\Delta_n, X_{t_k}, X_{t_{k+1}}\right) d\ell \\
&= \sum_{k=0}^{n-1} \frac{v}{\sqrt{n}} \int_0^1 \left(\frac{1}{\sigma(\ell)\Delta_n} \widetilde{\mathrm{E}}^{\theta_n,\sigma(\ell),\lambda_n}_{X_{t_k}} \left[\left(W_{t_{k+1}} - W_{t_k}\right)^2 \middle| Y^{\theta_n,\sigma(\ell),\lambda_n}_{t_{k+1}} = X_{t_{k+1}} \right] - \frac{1}{\sigma(\ell)} \right) d\ell.
\end{aligned}$$

Then, using (3.11) with $(\theta_n, \sigma(\ell), \lambda_n)$ instead of $(\theta(\ell), \sigma_0, \lambda_0)$, we get that

$$\log \frac{p\left(X^n; (\theta_n, \sigma_n, \lambda_n)\right)}{p\left(X^n; (\theta_n, \sigma_0, \lambda_n)\right)} = \sum_{k=0}^{n-1} (\eta_{k,n} + M_{k,n}),$$

where

$$\begin{aligned}
\eta_{k,n} &:= \frac{v}{\sqrt{n}} \int_0^1 \frac{1}{\Delta_n} \left(\frac{\sigma_0^2}{\sigma(\ell)^3} \left(B_{t_{k+1}} - B_{t_k}\right)^2 - \frac{\Delta_n}{\sigma(\ell)} \right) d\ell, \\
M_{k,n} &:= \frac{v}{\sqrt{n}} \int_0^1 \frac{1}{\Delta_n} \frac{1}{\sigma(\ell)^3} \Big\{ \left(\theta_0\Delta_n + \widetilde{N}^{\lambda_0}_{t_{k+1}} - \widetilde{N}^{\lambda_0}_{t_k}\right)^2 + 2\sigma_0 \left(B_{t_{k+1}} - B_{t_k}\right) \left(\theta_0\Delta_n + \widetilde{N}^{\lambda_0}_{t_{k+1}} - \widetilde{N}^{\lambda_0}_{t_k}\right) \\
&\quad - \widetilde{\mathrm{E}}^{\theta_n,\sigma(\ell),\lambda_n}_{X_{t_k}} \Big[\left(\theta_n\Delta_n + \widetilde{M}^{\lambda_n}_{t_{k+1}} - \widetilde{M}^{\lambda_n}_{t_k}\right)^2 \\
&\quad + 2\sigma(\ell) \left(W_{t_{k+1}} - W_{t_k}\right) \left(\theta_n\Delta_n + \widetilde{M}^{\lambda_n}_{t_{k+1}} - \widetilde{M}^{\lambda_n}_{t_k}\right) \Big| Y^{\theta_n,\sigma(\ell),\lambda_n}_{t_{k+1}} = X_{t_{k+1}} \Big] \Big\} d\ell.
\end{aligned}$$

Finally, using the Markov property and Proposition 3.2.2,

$$\begin{aligned}
&\log \frac{p\left(X^n; (\theta_n, \sigma_0, \lambda_n)\right)}{p\left(X^n; (\theta_n, \sigma_0, \lambda_0)\right)} = \sum_{k=0}^{n-1} \frac{w}{\sqrt{n\Delta_n}} \int_0^1 \frac{\partial_\lambda p^{\theta_n,\sigma_0,\lambda(\ell)}}{p^{\theta_n,\sigma_0,\lambda(\ell)}} \left(\Delta_n, X_{t_k}, X_{t_{k+1}}\right) d\ell \\
&= \sum_{k=0}^{n-1} \frac{w}{\sqrt{n\Delta_n}} \int_0^1 \widetilde{\mathrm{E}}^{\theta_n,\sigma_0,\lambda(\ell)}_{X_{t_k}} \left[-\frac{W_{t_{k+1}} - W_{t_k}}{\sigma_0} + \frac{\widetilde{M}^{\lambda(\ell)}_{t_{k+1}} - \widetilde{M}^{\lambda(\ell)}_{t_k}}{\lambda(\ell)} \middle| Y^{\theta_n,\sigma_0,\lambda(\ell)}_{t_{k+1}} = X_{t_{k+1}} \right] d\ell.
\end{aligned}$$

Now, from (3.11) with $(\theta_n, \sigma_0, \lambda(\ell))$ instead of $(\theta(\ell), \sigma_0, \lambda_0)$, we deduce that

$$\log \frac{p\left(X^n; (\theta_n, \sigma, \lambda_n)\right)}{p\left(X^n; (\theta_n, \sigma, \lambda)\right)} = \sum_{k=0}^{n-1} (\beta_{k,n} - R_{k,n}),$$

where

$$\beta_{k,n} := -\frac{w}{\sqrt{n\Delta_n}}\frac{1}{\sigma_0^2}\left(\sigma_0\left(B_{t_{k+1}}-B_{t_k}\right)+\frac{w\Delta_n}{2\sqrt{n\Delta_n}}-\frac{u\Delta_n}{\sqrt{n\Delta_n}}\right)$$
$$+\frac{w}{\sqrt{n\Delta_n}}\int_0^1 \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\sigma_0,\lambda(\ell)}\left[\frac{\widetilde{M}_{t_{k+1}}^{\lambda(\ell)}-\widetilde{M}_{t_k}^{\lambda(\ell)}}{\lambda(\ell)}\Big|Y_{t_{k+1}}^{\theta_n,\sigma_0,\lambda(\ell)}=X_{t_{k+1}}\right]d\ell, \tag{3.12}$$
$$R_{k,n} := \frac{w}{\sqrt{n\Delta_n}}\frac{1}{\sigma_0^2}\int_0^1\left(N_{t_{k+1}}-N_{t_k}-\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\sigma_0,\lambda(\ell)}\left[M_{t_{k+1}}-M_{t_k}\Big|Y_{t_{k+1}}^{\theta_n,\sigma_0,\lambda(\ell)}=X_{t_{k+1}}\right]\right)d\ell.$$

Therefore, we have obtained the following expansion of the log-likelihood ratio

$$\log\frac{p\left(X^n;(\theta_n,\sigma_n,\lambda_n)\right)}{p\left(X^n;(\theta_0,\sigma_0,\lambda_0)\right)}=\sum_{k=0}^{n-1}\left(\xi_{k,n}+\eta_{k,n}+\beta_{k,n}+H_{k,n}+M_{k,n}-R_{k,n}\right),$$

where, as we will see in the next subsections, the random variables $\xi_{k,n},\eta_{k,n},\beta_{k,n}$ are the terms that contribute to the limit in Theorem 3.1.1, and $H_{k,n}, M_{k,n}$ and $R_{k,n}$ are the negligible contributions.

3.3.2 Negligible contributions

Lemma 3.3.1. *Assume condition* (3.2). *Then, as* $n\to\infty$, $\sum_{k=0}^{n-1}H_{k,n}\overset{\mathrm{P}^{\theta_0,\sigma_0,\lambda_0}}{\longrightarrow}0$.

Proof. Since the $H_{k,n}$ are $\widehat{\mathcal{F}}_{t_{k+1}}$-measurable it suffices to show that conditions (i) and (ii) of Lemma 1.4.1 hold for the sequence $(H_{k,n})_{k\geq 1}$ under the measure $\mathrm{P}^{\theta_0,\sigma_0,\lambda_0}$.

First, using the fact that $\mathrm{E}[\widetilde{N}_{t_{k+1}}^{\lambda_0}-\widetilde{N}_{t_k}^{\lambda_0}|X_{t_k}]=0$, and Lemma 3.2.1, we get that

$$\sum_{k=0}^{n-1}\mathrm{E}^{\theta_0,\sigma_0,\lambda_0}\left[H_{k,n}|\widehat{\mathcal{F}}_{t_k}\right]=\sum_{k=0}^{n-1}\mathrm{E}\left[H_{k,n}|X_{t_k}\right]$$
$$=-\frac{u}{\sqrt{n\Delta_n}}\frac{1}{\sigma_0^2}\sum_{k=0}^{n-1}\int_0^1\mathrm{E}\left[\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell),\sigma_0,\lambda_0}\left[\widetilde{M}_{t_{k+1}}^{\lambda_0}-\widetilde{M}_{t_k}^{\lambda_0}\Big|Y_{t_{k+1}}^{\theta(\ell),\sigma_0,\lambda_0}=X_{t_{k+1}}\right]\Big|X_{t_k}\right]d\ell$$
$$=-\frac{u}{\sqrt{n\Delta_n}}\frac{1}{\sigma_0^2}\sum_{k=0}^{n-1}\int_0^1\mathrm{E}_{\widehat{Q}_k}\left[\left(\widetilde{M}_{t_{k+1}}^{\lambda_0}-\widetilde{M}_{t_k}^{\lambda_0}\right)\frac{d\widehat{\mathrm{P}}^{\lambda_0}}{d\widehat{Q}_k^{\theta(\ell),\lambda_0,\theta_0,\lambda_0,\sigma_0}}\Big|X_{t_k}\right]d\ell$$
$$=-\frac{u}{\sqrt{n\Delta_n}}\frac{1}{\sigma_0^2}\sum_{k=0}^{n-1}\int_0^1\mathrm{E}_{\widehat{Q}_k}\left[\left(\widetilde{M}_{t_{k+1}}^{\lambda_0}-\widetilde{M}_{t_k}^{\lambda_0}\right)e^{-\frac{\ell u}{\sigma\sqrt{n\Delta_n}}\left(B_{t_{k+1}}-B_{t_k}\right)-\frac{\ell^2u^2}{2\sigma^2n}}\Big|X_{t_k}\right]d\ell,$$

where $\widehat{Q}_k\equiv\widehat{Q}_k^{\theta(\ell),\lambda_0,\theta_0,\lambda_0,\sigma_0}$. Thus, using the independence between $M_{t_{k+1}}-M_{t_k}$, $B_{t_{k+1}}-B_{t_k}$, and X_{t_k}, together with $\mathrm{E}_{\widehat{Q}_k}[\widetilde{M}_{t_{k+1}}^{\lambda_0}-\widetilde{M}_{t_k}^{\lambda_0}]=\mathrm{E}[\widetilde{M}_{t_{k+1}}^{\lambda_0}-\widetilde{M}_{t_k}^{\lambda_0}]=0$, we conclude that the term (i) of Lemma 1.4.1 is actually equal to 0 for all $n\geq 1$.

We next show that (ii) of Lemma 1.4.1 holds. By Cauchy-Schwarz's and Jensen's inequalities, together with Lemma 3.2.6, for all $\alpha \in (0, \frac{1}{2})$ and $j\mathbf{1}_{\widetilde{J}_{j,k}} = j - j\mathbf{1}_{\widetilde{J}^c_{j,k}}$, we get that

$$\begin{aligned}
&\sum_{k=0}^{n-1} \mathrm{E}^{\theta_0,\sigma_0,\lambda_0}\left[H^2_{k,n}|\widehat{\mathcal{F}}_{t_k}\right] \\
&\leq \sum_{k=0}^{n-1} \frac{u^2}{n\Delta_n}\frac{1}{\sigma_0^4}\int_0^1 \mathrm{E}\left[\left(N_{t_{k+1}} - N_{t_k} - \widetilde{\mathrm{E}}^{\theta(\ell),\sigma_0,\lambda_0}_{X_{t_k}}\left[M_{t_{k+1}} - M_{t_k}\middle| Y^{\theta(\ell),\sigma_0,\lambda_0}_{t_{k+1}} = X_{t_{k+1}}\right]\right)^2 \middle| X_{t_k}\right] d\ell \\
&= \sum_{k=0}^{n-1} \frac{u^2}{n\Delta_n}\frac{1}{\sigma_0^4}\int_0^1 \\
&\times \sum_{j=0}^{\infty} \mathrm{E}\left[\mathbf{1}_{\widehat{J}_{j,k}}\left(j - \widetilde{\mathrm{E}}^{\theta(\ell),\sigma_0,\lambda_0}_{X_{t_k}}\left[\left(M_{t_{k+1}} - M_{t_k}\right)\left(\mathbf{1}_{\widetilde{J}_{j,k}} + \mathbf{1}_{\widetilde{J}^c_{j,k}}\right)\middle| Y^{\theta(\ell),\sigma_0,\lambda_0}_{t_{k+1}} = X_{t_{k+1}}\right]\right)^2 \middle| X_{t_k}\right] d\ell \\
&\leq \sum_{k=0}^{n-1} \frac{u^2}{n\Delta_n}\frac{2}{\sigma_0^4}\int_0^1 \left(M^{\theta(\ell),\sigma_0,\lambda_0}_{1,2} + M^{\theta(\ell),\sigma_0,\lambda_0}_{2,2}\right) d\ell \\
&\leq C_1\frac{u^2}{\Delta_n}e^{-\frac{1}{C_2\Delta_n^{1-2\alpha}}},
\end{aligned}$$

for some constants $C_1, C_2 > 0$. This concludes the desired result. □

Lemma 3.3.2. *Assume condition* (3.2)*. Then, as $n \to \infty$, $\sum_{k=0}^{n-1} R_{k,n} \overset{\mathrm{P}^{\theta_0,\sigma_0,\lambda_0}}{\longrightarrow} 0$.*

Proof. Since the $R_{k,n}$ are $\widehat{\mathcal{F}}_{t_{k+1}}$-measurable, it suffices to show that conditions (i) and (ii) of Lemma 1.4.1 under the measure $\mathrm{P}^{\theta_0,\sigma_0,\lambda_0}$ hold true for the sequence $(R_{k,n})_{k\geq 1}$. We start showing (i). Using the fact that $\mathrm{E}[N_{t_{k+1}} - N_{t_k}|X_{t_k}] = \lambda_0\Delta_n$, and $M_{t_{k+1}} - M_{t_k} = \widetilde{M}^{\lambda(\ell)}_{t_{k+1}} - \widetilde{M}^{\lambda(\ell)}_{t_k} + \lambda(\ell)\Delta_n$, together with Lemma 3.2.7, we get that

$$\begin{aligned}
\sum_{k=0}^{n-1} \mathrm{E}^{\theta_0,\sigma_0,\lambda_0}\left[R_{k,n}|\widehat{\mathcal{F}}_{t_k}\right] &= \sum_{k=0}^{n-1}\mathrm{E}\left[R_{k,n}|X_{t_k}\right] \\
&= \frac{w}{\sigma_0^2\sqrt{n\Delta_n}}\sum_{k=0}^{n-1}\int_0^1\left(\lambda_0\Delta_n - \mathrm{E}\left[\widetilde{\mathrm{E}}^{\theta_n,\sigma_0,\lambda(\ell)}_{X_{t_k}}\left[M_{t_{k+1}} - M_{t_k}\middle| Y^{\theta_n,\sigma_0,\lambda(\ell)}_{t_{k+1}} = X_{t_{k+1}}\right]\middle| X_{t_k}\right]\right) d\ell \\
&= \frac{w}{\sigma_0^2\sqrt{n\Delta_n}}\sum_{k=0}^{n-1}\int_0^1 (\lambda_0\Delta_n - \lambda_0\Delta_n)\, d\ell \\
&= 0.
\end{aligned}$$

Hence, we conclude that the term (i) of Lemma 1.4.1 is actually equal to 0 for all $n \geq 1$.

Next, we show condition Lemma 1.4.1(ii). Proceeding as in Lemma 3.3.1, for all $\alpha \in (0, \frac{1}{2})$, we have that

$$\begin{aligned}
&\sum_{k=0}^{n-1} \mathrm{E}^{\theta_0,\sigma_0,\lambda_0}\left[R^2_{k,n}|\widehat{\mathcal{F}}_{t_k}\right] \\
&\leq \sum_{k=0}^{n-1} \frac{w^2}{n\Delta_n}\frac{1}{\sigma_0^4}\int_0^1 \mathrm{E}\left[\left(N_{t_{k+1}} - N_{t_k} - \widetilde{\mathrm{E}}^{\theta_n,\sigma_0,\lambda(\ell)}_{X_{t_k}}\left[M_{t_{k+1}} - M_{t_k}\middle| Y^{\theta_n,\sigma_0,\lambda(\ell)}_{t_{k+1}} = X_{t_{k+1}}\right]\right)^2 \middle| X_{t_k}\right] d\ell \\
&\leq \sum_{k=0}^{n-1} \frac{w^2}{n\Delta_n}\frac{2}{\sigma_0^4}\int_0^1 \left(M^{\theta_n,\sigma_0,\lambda(\ell)}_{1,2} + M^{\theta_n,\sigma_0,\lambda(\ell)}_{2,2}\right) d\ell \\
&\leq C_1\frac{w^2}{\Delta_n}e^{-\frac{1}{C_2\Delta_n^{1-2\alpha}}},
\end{aligned}$$

for some constants $C_1, C_2 > 0$. Thus, the result follows. □

Lemma 3.3.3. *Assume condition* (3.2). *Then, as* $n \to \infty$, $\sum_{k=0}^{n-1} M_{k,n} \overset{\mathrm{P}^{\theta_0,\sigma_0,\lambda_0}}{\longrightarrow} 0$.

Proof. Since the $M_{k,n}$ are $\widehat{\mathcal{F}}_{t_{k+1}}$-measurable, it suffices to show that conditions (i) and (ii) of Lemma 1.4.1 under the measure $\mathrm{P}^{\theta_0,\sigma_0,\lambda_0}$ hold for the sequence $(M_{k,n})_{k\geq 1}$. We start proving (i). We have

$$\begin{aligned}&\left(\theta_0\Delta_n + \widetilde{N}^{\lambda_0}_{t_{k+1}} - \widetilde{N}^{\lambda_0}_{t_k}\right)^2 + 2\sigma_0\left(B_{t_{k+1}} - B_{t_k}\right)\left(\theta_0\Delta_n + \widetilde{N}^{\lambda_0}_{t_{k+1}} - \widetilde{N}^{\lambda_0}_{t_k}\right)\\&= 2(\theta_0-\lambda_0)\Delta_n(X_{t_{k+1}} - X_{t_k}) - (\theta_0-\lambda_0)^2\Delta_n^2 - (N_{t_{k+1}} - N_{t_k})^2\\&\quad + 2(X_{t_{k+1}} - X_{t_k})(N_{t_{k+1}} - N_{t_k}) - 2(\theta_0-\lambda_0)\Delta_n(N_{t_{k+1}} - N_{t_k}),\end{aligned}$$

and

$$\begin{aligned}&\left(\theta_n\Delta_n + \widetilde{M}^{\lambda_n}_{t_{k+1}} - \widetilde{M}^{\lambda_n}_{t_k}\right)^2 + 2\sigma(\ell)\left(W_{t_{k+1}} - W_{t_k}\right)\left(\theta_n\Delta_n + \widetilde{M}^{\lambda_n}_{t_{k+1}} - \widetilde{M}^{\lambda_n}_{t_k}\right)\\&= 2\left(Y^{\theta_n,\sigma(\ell),\lambda_n}_{t_{k+1}} - Y^{\theta_n,\sigma(\ell),\lambda_n}_{t_k}\right)\left(\theta_n\Delta_n + M_{t_{k+1}} - M_{t_k} - \lambda_n\Delta_n\right)\\&\quad - \left(\theta_n\Delta_n + M_{t_{k+1}} - M_{t_k} - \lambda_n\Delta_n\right)^2.\end{aligned}$$

This implies that

$$\sum_{k=0}^{n-1} \mathrm{E}^{\theta_0,\sigma_0,\lambda_0}\left[M_{k,n}|\widehat{\mathcal{F}}_{t_k}\right] = \sum_{k=0}^{n-1} \frac{v}{\sqrt{n}}\int_0^1 \frac{1}{\Delta_n}\frac{1}{\sigma(\ell)^3}\left(T_1 + T_2 - T_3 + T_4 - T_5\right)d\ell,$$

where

$$\begin{aligned}T_1 &= 2\frac{w-u}{\sqrt{n\Delta_n}}\Delta_n \mathrm{E}\left[X_{t_{k+1}} - X_{t_k}|X_{t_k}\right] = 2\theta_0\frac{w-u}{\sqrt{n\Delta_n}}\Delta_n^2,\\T_2 &= -(\theta_0-\lambda_0)^2\Delta_n^2 + (\theta_n-\lambda_n)^2\Delta_n^2 + 2\lambda_0\frac{w-u}{\sqrt{n\Delta_n}}\Delta_n^2,\\T_3 &= \mathrm{E}\left[(N_{t_{k+1}} - N_{t_k})^2 - \widetilde{\mathrm{E}}^{\theta_n,\sigma(\ell),\lambda_n}_{X_{t_k}}\left[(M_{t_{k+1}} - M_{t_k})^2\Big|Y^{\theta_n,\sigma(\ell),\lambda_n}_{t_{k+1}} = X_{t_{k+1}}\right]\Big|X_{t_k}\right],\\T_4 &= 2\mathrm{E}\left[(X_{t_{k+1}} - X_{t_k})\left(N_{t_{k+1}} - N_{t_k} - \widetilde{\mathrm{E}}^{\theta_n,\sigma(\ell),\lambda_n}_{X_{t_k}}\left[M_{t_{k+1}} - M_{t_k}\Big|Y^{\theta_n,\sigma(\ell),\lambda_n}_{t_{k+1}} = X_{t_{k+1}}\right]\right)\Big|X_{t_k}\right],\\T_5 &= 2\Delta_n(\theta_n-\lambda_n)\mathrm{E}\left[N_{t_{k+1}} - N_{t_k} - \widetilde{\mathrm{E}}^{\theta_n,\sigma(\ell),\lambda_n}_{X_{t_k}}\left[M_{t_{k+1}} - M_{t_k}\Big|Y^{\theta_n,\sigma(\ell),\lambda_n}_{t_{k+1}} = X_{t_{k+1}}\right]\Big|X_{t_k}\right].\end{aligned}$$

Clearly,

$$\left|\sum_{k=0}^{n-1} \frac{v}{\sqrt{n}}\int_0^1 \frac{1}{\Delta_n}\frac{1}{\sigma(\ell)^3}\left(T_1 + T_2\right)d\ell\right| \leq C_1\sqrt{\Delta_n} + \frac{C_2}{\sqrt{n}},$$

for some constants $C_1, C_2 > 0$.

Moreover,

$$\begin{aligned}T_3 &= \sum_{j=0}^{\infty}\mathrm{E}\left[\mathbf{1}_{\widehat{J}_{j,k}}\left(j^2 - \widetilde{\mathrm{E}}^{\theta_n,\sigma(\ell),\lambda_n}_{X_{t_k}}\left[(M_{t_{k+1}} - M_{t_k})^2\left(\mathbf{1}_{\widetilde{J}_{j,k}} + \mathbf{1}_{\widetilde{J}^c_{j,k}}\right)\Big|Y^{\theta_n,\sigma(\ell),\lambda_n}_{t_{k+1}} = X_{t_{k+1}}\right]\right)\Big|X_{t_k}\right]\\&= M^{\theta_n,\sigma(\ell),\lambda_n}_{1,2} - M^{\theta_n,\sigma(\ell),\lambda_n}_{2,2},\end{aligned}$$

which, together with Lemma 3.2.6, implies that, for all $\alpha \in (0, \frac{1}{2})$,

$$\begin{aligned}\left|\sum_{k=0}^{n-1} \frac{v}{\sqrt{n}}\int_0^1 \frac{1}{\Delta_n}\frac{1}{\sigma(\ell)^3}T_3 d\ell\right| &= \left|\sum_{k=0}^{n-1} \frac{v}{\sqrt{n}}\int_0^1 \frac{1}{\Delta_n}\frac{1}{\sigma(\ell)^3}\left(M^{\theta_n,\sigma(\ell),\lambda_n}_{1,2} - M^{\theta_n,\sigma(\ell),\lambda_n}_{2,2}\right)d\ell\right|\\&\leq C_1\frac{\sqrt{n}}{\Delta_n}e^{-\frac{1}{C_2\Delta_n^{1-2\alpha}}},\end{aligned}$$

for some constants $C_1, C_2 > 0$.

In order to treat T_5, we proceed as for the term T_3 to get that, for all $\alpha \in (0, \frac{1}{2})$,

$$\left|\sum_{k=0}^{n-1} \frac{v}{\sqrt{n}} \int_0^1 \frac{1}{\Delta_n} \frac{1}{\sigma(\ell)^3} T_5 d\ell\right| \le C_1 \frac{\sqrt{n}}{\Delta_n} e^{-\frac{1}{C_2 \Delta_n^{1-2\alpha}}},$$

for some constants $C_1, C_2 > 0$.

Using Cauchy-Schwarz's and Jensen's inequalities, and proceeding as in Lemma 3.3.1, together with Lemma 3.2.6, we get that, for all $\alpha \in (0, \frac{1}{2})$,

$$\begin{aligned}
\left|\sum_{k=0}^{n-1} \frac{v}{\sqrt{n}} \int_0^1 \frac{1}{\Delta_n} \frac{1}{\sigma(\ell)^3} T_4 d\ell\right| &\le \sum_{k=0}^{n-1} \frac{2|v|}{\sqrt{n\Delta_n}} \int_0^1 \frac{1}{|\sigma(\ell)|^3} \\
&\times \left(\mathrm{E}\left[\left(N_{t_{k+1}} - N_{t_k} - \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\sigma(\ell),\lambda_n}\left[M_{t_{k+1}} - M_{t_k} \middle| Y_{t_{k+1}}^{\theta_n,\sigma(\ell),\lambda_n} = X_{t_{k+1}}\right]\right)^2 \middle| X_{t_k}\right]\right)^{1/2} d\ell \\
&\le \sum_{k=0}^{n-1} \frac{2\sqrt{2}|v|}{\sqrt{n\Delta_n}} \int_0^1 \frac{1}{|\sigma(\ell)|^3} \left(M_{1,2}^{\theta_n,\sigma(\ell),\lambda_n} + M_{2,2}^{\theta_n,\sigma(\ell),\lambda_n}\right)^{1/2} d\ell \\
&\le C_1 \frac{\sqrt{n}}{\sqrt{\Delta_n}} e^{-\frac{1}{C_2 \Delta_n^{1-2\alpha}}},
\end{aligned}$$

for some constants $C_1, C_2 > 0$. Consequently, condition (i) of Lemma 1.4.1 holds. We next show (ii). Applying Hölder's inequality, and the same decomposition as for T_3, we get that

$$\begin{aligned}
&\sum_{k=0}^{n-1} \mathrm{E}^{\theta_0,\sigma_0,\lambda_0}\left[M_{k,n}^2 | \widehat{\mathcal{F}}_{t_k}\right] \\
&\le \frac{v^2}{n\Delta_n^2} \sum_{k=0}^{n-1} \int_0^1 \frac{1}{\sigma(\ell)^6} \mathrm{E}\left[\left\{H_{\theta_0,\lambda_0} - \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\sigma(\ell),\lambda_n}\left[\bar{H}_{\theta_n,\lambda_n} \middle| Y_{t_{k+1}}^{\theta_n,\sigma(\ell),\lambda_n} = X_{t_{k+1}}\right]\right\}^2 \middle| X_{t_k}\right] d\ell \\
&\le \frac{v^2}{n\Delta_n^2} \sum_{k=0}^{n-1} \int_0^1 \frac{2}{\sigma(\ell)^6} (V_1 + V_2)\, d\ell,
\end{aligned}$$

where

$$\begin{aligned}
V_1 &:= \sum_{j=0}^{\infty} \mathrm{E}\left[\mathbf{1}_{\widehat{J}_{j,k}} \left\{H_{\theta_0,\lambda_0} - \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\sigma(\ell),\lambda_n}\left[\mathbf{1}_{\widetilde{J}_{j,k}} \bar{H}_{\theta_n,\lambda_n} \middle| Y_{t_{k+1}}^{\theta_n,\sigma(\ell),\lambda_n} = X_{t_{k+1}}\right]\right\}^2 \middle| X_{t_k}\right], \\
V_2 &:= \sum_{j=0}^{\infty} \mathrm{E}\left[\mathbf{1}_{\widehat{J}_{j,k}} \left\{\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\sigma(\ell),\lambda_n}\left[\mathbf{1}_{\widetilde{J}_{j,k}^c} \bar{H}_{\theta_n,\lambda_n} \middle| Y_{t_{k+1}}^{\theta_n,\sigma(\ell),\lambda_n} = X_{t_{k+1}}\right]\right\}^2 \middle| X_{t_k}\right],
\end{aligned}$$

and

$$\begin{aligned}
H_{\theta_0,\lambda_0} &:= \left(\theta_0 \Delta_n + \widetilde{N}_{t_{k+1}}^{\lambda_0} - \widetilde{N}_{t_k}^{\lambda_0}\right)^2 + 2\sigma_0 \left(B_{t_{k+1}} - B_{t_k}\right) \left(\theta_0 \Delta_n + \widetilde{N}_{t_{k+1}}^{\lambda_0} - \widetilde{N}_{t_k}^{\lambda_0}\right). \\
\bar{H}_{\theta_n,\lambda_n} &:= 2\left(Y_{t_{k+1}}^{\theta_n,\sigma(\ell),\lambda_n} - Y_{t_k}^{\theta_n,\sigma(\ell),\lambda_n}\right)\left(\theta_n \Delta_n + M_{t_{k+1}} - M_{t_k} - \lambda_n \Delta_n\right) \\
&\quad - \left(\theta_n \Delta_n + M_{t_{k+1}} - M_{t_k} - \lambda_n \Delta_n\right)^2.
\end{aligned}$$

Using equation (3.1) and Jensen's inequality, adding and substracting one term, we get

$$\begin{aligned}
V_1 &= \sum_{j=0}^{\infty} \mathrm{E}\Bigg[\mathbf{1}_{\widehat{J}_{j,k}} \Bigg\{ (\theta_0 \Delta_n + j - \lambda_0 \Delta_n)^2 + 2\sigma_0 \left(B_{t_{k+1}} - B_{t_k}\right)(\theta_0 \Delta_n + j - \lambda_0 \Delta_n) \\
&\quad - \left((\theta_n \Delta_n + j - \lambda_n \Delta_n)^2 + 2\left(\sigma_0 (B_{t_{k+1}} - B_{t_k}) + \frac{(w-u)\Delta_n}{\sqrt{n\Delta_n}}\right)(\theta_n \Delta_n + j - \lambda_n \Delta_n)\right) \\
&\quad\quad \times \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\sigma(\ell),\lambda_n}\left[\mathbf{1}_{\widetilde{J}_{j,k}} \middle| Y_{t_{k+1}}^{\theta_n,\sigma(\ell),\lambda_n} = X_{t_{k+1}}\right] \Bigg\}^2 \Bigg| X_{t_k}\Bigg] \\
&\le 2\left(V_{1,1} + V_{1,2}\right),
\end{aligned}$$

where

$$V_{1,1} = \sum_{j=0}^{\infty} \mathrm{E}\Big[\mathbf{1}_{\widehat{J}_{j,k}} \Big\{ (\theta_0\Delta_n + j - \lambda_0\Delta_n)^2 + 2\sigma_0 \left(B_{t_{k+1}} - B_{t_k}\right) (\theta_0\Delta_n + j - \lambda_0\Delta_n)$$
$$- \Big((\theta_n\Delta_n + j - \lambda_n\Delta_n)^2 + 2 \Big(\sigma_0(B_{t_{k+1}} - B_{t_k}) + \frac{(w-u)\Delta_n}{\sqrt{n\Delta_n}} \Big) (\theta_n\Delta_n + j - \lambda_n\Delta_n) \Big) \Big\}^2 \Big| X_{t_k} \Big],$$
$$V_{1,2} = \sum_{j=0}^{\infty} \mathrm{E}\Big[\mathbf{1}_{\widehat{J}_{j,k}} \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\sigma(\ell),\lambda_n} \Big[\mathbf{1}_{\widetilde{J}_{j,k}^c} \Big| Y_{t_{k+1}}^{\theta_n,\sigma(\ell),\lambda_n} = X_{t_{k+1}} \Big]$$
$$\times \Big((\theta_n\Delta_n + j - \lambda_n\Delta_n)^2 + 2 \Big(\sigma_0(B_{t_{k+1}} - B_{t_k}) + \frac{(w-u)\Delta_n}{\sqrt{n\Delta_n}} \Big) (\theta_n\Delta_n + j - \lambda_n\Delta_n) \Big)^2 \Big| X_{t_k} \Big].$$

Basic computations yield that

$$\left| \frac{v^2}{n\Delta_n^2} \sum_{k=0}^{n-1} \int_0^1 \frac{1}{\sigma(\ell)^6} V_{1,1} d\ell \right| \le \frac{C}{n}.$$

Now, we treat $V_{1,2}$. In order to deal with the terms that contain the increments of the Brownian motion, we multiply those increments by $\mathbf{1}_{\{|B_{t_{k+1}} - B_{t_k}| \le \Delta_n^\alpha\}} + \mathbf{1}_{\{|B_{t_{k+1}} - B_{t_k}| > \Delta_n^\alpha\}}$, for $\alpha \in (0, \frac{1}{2})$. Then, for the terms involving $\mathbf{1}_{\{|B_{t_{k+1}} - B_{t_k}| > \Delta_n^\alpha\}}$, we bound the conditional expectation by one, use the independence between B and N, and Cauchy-Schwarz inequality, to ultimately conclude that these terms can be bounded by

$$C_1\Delta_n \left(\mathrm{E}\left[(B_{t_{k+1}} - B_{t_k})^4 | X_{t_k}\right]\right)^{1/4} \left(\mathrm{P}\left(|B_{t_{k+1}} - B_{t_k}| > \Delta_n^\alpha | X_{t_k}\right)\right)^{1/2} \le C_1\Delta_n^2 e^{-\frac{1}{C_2\Delta_n^{1-2\alpha}}}.$$

On the other hand, the term involving $\mathbf{1}_{\{|B_{t_{k+1}} - B_{t_k}| \le \Delta_n^\alpha\}}$ can be bounded by $M_{1,0}^{\theta_n,\sigma(\ell),\lambda_n}$. The other terms that do not involve the increment of the Brownian motion can be bounded by $M_{1,p}^{\theta_n,\sigma(\ell),\lambda_n}$ for $p \in \{0, \dots, 4\}$. Thus, using Lemma 3.2.6, we obtain that, for all $\alpha \in (0, \frac{1}{2})$,

$$\frac{v^2}{n\Delta_n^2} \sum_{k=0}^{n-1} \int_0^1 \frac{1}{\sigma(\ell)^6} V_{1,2} d\ell \le C_1 \frac{1}{\Delta_n^2} e^{-\frac{1}{C_2\Delta_n^{1-2\alpha}}},$$

for some constants $C_1, C_2 > 0$. Thus, $V_1 \overset{\mathrm{P}^{\theta_0,\sigma_0,\lambda_0}}{\longrightarrow} 0$ as $n \to \infty$.

Applying Jensen's inequality and equation (3.1), we obtain that

$$V_2 \le \sum_{j=0}^{\infty} \mathrm{E}\left[\mathbf{1}_{\widehat{J}_{j,k}} \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\sigma(\ell),\lambda_n} \left[\mathbf{1}_{\widetilde{J}_{j,k}^c} \bar{H}_{\theta_n,\lambda_n}^2 \Big| Y_{t_{k+1}}^{\theta_n,\sigma(\ell),\lambda_n} = X_{t_{k+1}} \right] \Big| X_{t_k} \right]$$
$$= \sum_{j=0}^{\infty} \mathrm{E}\left[\mathbf{1}_{\widehat{J}_{j,k}} \sum_{m=0: m\neq j}^{\infty} \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\sigma(\ell),\lambda_n} \left[\mathbf{1}_{\widetilde{J}_{m,k}} \bar{H}_{\theta_n,\lambda_n}^2 \Big| Y_{t_{k+1}}^{\theta_n,\sigma(\ell),\lambda_n} = X_{t_{k+1}} \right] \Big| X_{t_k} \right]$$
$$= \sum_{j=0}^{\infty} \mathrm{E}\Big[\mathbf{1}_{\widehat{J}_{j,k}} \sum_{m=0: m\neq j}^{\infty} \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\sigma(\ell),\lambda_n} \left[\mathbf{1}_{\widetilde{J}_{m,k}} \Big| Y_{t_{k+1}}^{\theta_n,\sigma(\ell),\lambda_n} = X_{t_{k+1}} \right] \Big((\theta_n\Delta_n + m - \lambda_n\Delta_n)^2$$
$$+ 2 \Big(\sigma_0(B_{t_{k+1}} - B_{t_k}) + j - m + \frac{(w-u)\Delta_n}{\sqrt{n\Delta_n}} \Big) (\theta_n\Delta_n + m - \lambda_n\Delta_n) \Big)^2 \Big| X_{t_k} \Big].$$

Observe that V_2 can be upper bounded by a sum of the terms $M_{1,p}^{\theta_n,\sigma(\ell),\lambda_n}$, for $p \in \{0, 1, 2\}$ and $M_{2,p}^{\theta_n,\sigma(\ell),\lambda_n}$, for $p \in \{0, \dots, 4\}$. Then, from Lemma 3.2.6, we get that, for all $\alpha \in (0, \frac{1}{2})$,

$$\frac{v^2}{n\Delta_n^2} \sum_{k=0}^{n-1} \int_0^1 \frac{1}{\sigma(\ell)^6} V_2 d\ell \le C_1 \frac{1}{\Delta_n^2} e^{-\frac{1}{C_2\Delta_n^{1-2\alpha}}},$$

for some constants $C_1, C_2 > 0$. This concludes the desired result. □

3.3.3 Main contribution : LAN property

Proof. Applying Lemma 1.4.3 to $\zeta_{k,n} = \xi_{k,n}+\eta_{k,n}+\beta_{k,n}$, we need to consider $\mathrm{E}^{\theta_0,\sigma_0,\lambda_0}[(\xi_{k,n}+\eta_{k,n}+\beta_{k,n})^r|\widehat{\mathcal{F}}_{t_k}]$ for $r=1,2$ and 4 but this conditional expectation equals $\mathrm{E}[(\xi_{k,n}+\eta_{k,n}+\beta_{k,n})^r|\widehat{\mathcal{F}}_{t_k}]$. Therefore, it suffices to show that as $n\to\infty$:

$$\sum_{k=0}^{n-1}\mathrm{E}\left[\xi_{k,n}|\widehat{\mathcal{F}}_{t_k}\right] \overset{\mathrm{P}^{\theta_0,\sigma_0,\lambda_0}}{\longrightarrow} -\frac{u^2}{2\sigma_0^2}, \tag{3.13}$$

$$\sum_{k=0}^{n-1}\mathrm{E}\left[\eta_{k,n}|\widehat{\mathcal{F}}_{t_k}\right] \overset{\mathrm{P}^{\theta_0,\sigma_0,\lambda_0}}{\longrightarrow} -\frac{v^2}{2}\frac{2}{\sigma_0^2}, \tag{3.14}$$

$$\sum_{k=0}^{n-1}\mathrm{E}\left[\beta_{k,n}|\widehat{\mathcal{F}}_{t_k}\right] \overset{\mathrm{P}^{\theta_0,\sigma_0,\lambda_0}}{\longrightarrow} -\frac{w^2}{2\sigma_0^2}\left(1+\frac{\sigma_0^2}{\lambda_0}\right)+\frac{uw}{\sigma_0^2}, \tag{3.15}$$

$$\sum_{k=0}^{n-1}\left(\mathrm{E}\left[\xi_{k,n}^2|\widehat{\mathcal{F}}_{t_k}\right]-\left(\mathrm{E}\left[\xi_{k,n}|\widehat{\mathcal{F}}_{t_k}\right]\right)^2\right) \overset{\mathrm{P}^{\theta_0,\sigma_0,\lambda_0}}{\longrightarrow} \frac{u^2}{\sigma_0^2}, \tag{3.16}$$

$$\sum_{k=0}^{n-1}\left(\mathrm{E}\left[\eta_{k,n}^2|\widehat{\mathcal{F}}_{t_k}\right]-\left(\mathrm{E}\left[\eta_{k,n}|\widehat{\mathcal{F}}_{t_k}\right]\right)^2\right) \overset{\mathrm{P}^{\theta_0,\sigma_0,\lambda_0}}{\longrightarrow} v^2\frac{2}{\sigma_0^2}, \tag{3.17}$$

$$\sum_{k=0}^{n-1}\left(\mathrm{E}\left[\beta_{k,n}^2|\widehat{\mathcal{F}}_{t_k}\right]-\left(\mathrm{E}\left[\beta_{k,n}|\widehat{\mathcal{F}}_{t_k}\right]\right)^2\right) \overset{\mathrm{P}^{\theta_0,\sigma_0,\lambda_0}}{\longrightarrow} \frac{w^2}{\sigma_0^2}\left(1+\frac{\sigma_0^2}{\lambda_0}\right), \tag{3.18}$$

$$\sum_{k=0}^{n-1}\left(\mathrm{E}\left[\xi_{k,n}\eta_{k,n}|\widehat{\mathcal{F}}_{t_k}\right]-\mathrm{E}\left[\xi_{k,n}|\widehat{\mathcal{F}}_{t_k}\right]\mathrm{E}\left[\eta_{k,n}|\widehat{\mathcal{F}}_{t_k}\right]\right) \overset{\mathrm{P}^{\theta_0,\sigma_0,\lambda_0}}{\longrightarrow} 0, \tag{3.19}$$

$$\sum_{k=0}^{n-1}\left(\mathrm{E}\left[\xi_{k,n}\beta_{k,n}|\widehat{\mathcal{F}}_{t_k}\right]-\mathrm{E}\left[\xi_{k,n}|\widehat{\mathcal{F}}_{t_k}\right]\mathrm{E}\left[\beta_{k,n}|\widehat{\mathcal{F}}_{t_k}\right]\right) \overset{\mathrm{P}^{\theta_0,\sigma_0,\lambda_0}}{\longrightarrow} -\frac{uw}{\sigma_0^2}, \tag{3.20}$$

$$\sum_{k=0}^{n-1}\left(\mathrm{E}\left[\eta_{k,n}\beta_{k,n}|\widehat{\mathcal{F}}_{t_k}\right]-\mathrm{E}\left[\eta_{k,n}|\widehat{\mathcal{F}}_{t_k}\right]\mathrm{E}\left[\beta_{k,n}|\widehat{\mathcal{F}}_{t_k}\right]\right) \overset{\mathrm{P}^{\theta_0,\sigma_0,\lambda_0}}{\longrightarrow} 0, \tag{3.21}$$

$$\sum_{k=0}^{n-1}\mathrm{E}\left[\xi_{k,n}^4|\widehat{\mathcal{F}}_{t_k}\right] \overset{\mathrm{P}^{\theta_0,\sigma_0,\lambda_0}}{\longrightarrow} 0, \tag{3.22}$$

$$\sum_{k=0}^{n-1}\mathrm{E}\left[\eta_{k,n}^4|\widehat{\mathcal{F}}_{t_k}\right] \overset{\mathrm{P}^{\theta_0,\sigma_0,\lambda_0}}{\longrightarrow} 0, \tag{3.23}$$

$$\sum_{k=0}^{n-1}\mathrm{E}\left[\beta_{k,n}^4|\widehat{\mathcal{F}}_{t_k}\right] \overset{\mathrm{P}^{\theta_0,\sigma_0,\lambda_0}}{\longrightarrow} 0. \tag{3.24}$$

The validity of (3.13), (3.16), (3.22), (3.14), and (3.19) is easily checked by using moment properties of the Brownian motion and the definitions of $\xi_{k,n}$ and $\eta_{k,n}$.

Proof of (3.17). First, we observe that

$$\sum_{k=0}^{n-1}\left(\mathrm{E}\left[\eta_{k,n}|\widehat{\mathcal{F}}_{t_k}\right]\right)^2=\frac{v^2}{n}\sum_{k=0}^{n-1}\left(\int_0^1\frac{\sigma_0^2-\sigma(\ell)^2}{\sigma(\ell)^3}d\ell\right)^2\leq\frac{C}{n},$$

for some constant $C>0$.

On the other hand, since $\mathrm{E}[(B_{t_{k+1}}-B_{t_k})^2|\widehat{\mathcal{F}}_{t_k}]=\Delta_n$ and $\mathrm{E}[(B_{t_{k+1}}-B_{t_k})^4|\widehat{\mathcal{F}}_{t_k}]=3\Delta_n^2$, we deduce that as $n \to \infty$,

$$\sum_{k=0}^{n-1} \mathrm{E}\left[\eta_{k,n}^2 | \widehat{\mathcal{F}}_{t_k}\right] = \frac{v^2}{n\Delta_n^2} \sum_{k=0}^{n-1} \left\{ \left(\int_0^1 \frac{\sigma_0^2}{\sigma(\ell)^3} d\ell \right)^2 \mathrm{E}\left[\left(B_{t_{k+1}} - B_{t_k} \right)^4 | \widehat{\mathcal{F}}_{t_k} \right] + \left(\int_0^1 \frac{\Delta_n}{\sigma(\ell)} d\ell \right)^2 \right.$$
$$\left. - 2 \int_0^1 \frac{\Delta_n}{\sigma(\ell)} d\ell \int_0^1 \frac{\sigma_0^2}{\sigma(\ell)^3} d\ell \mathrm{E}\left[\left(B_{t_{k+1}} - B_{t_k} \right)^2 | \widehat{\mathcal{F}}_{t_k} \right] \right\}$$
$$\to v^2 \frac{2}{\sigma_0^2},$$

which completes the proof of (3.17).

Proof of (3.23). It is easy to see that

$$\sum_{k=0}^{n-1} \mathrm{E}\left[\eta_{k,n}^4 | \widehat{\mathcal{F}}_{t_k}\right] \leq \frac{C}{n},$$

for some constant $C > 0$.

Proof of (3.15). Using (3.12) in page 50 and Lemma 3.2.7, we get that as $n \to \infty$,

$$\sum_{k=0}^{n-1} \mathrm{E}\left[\beta_{k,n} | \widehat{\mathcal{F}}_{t_k}\right] = -\frac{w^2}{2\sigma_0^2} + \frac{uw}{\sigma_0^2}$$
$$+ \sum_{k=0}^{n-1} \frac{w}{\sqrt{n\Delta_n}} \int_0^1 \frac{1}{\lambda(\ell)} \mathrm{E}\left[\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\sigma_0,\lambda(\ell)} \left[\widetilde{M}_{t_{k+1}}^{\lambda(\ell)} - \widetilde{M}_{t_k}^{\lambda(\ell)} \middle| Y_{t_{k+1}}^{\theta_n,\sigma_0,\lambda(\ell)} = X_{t_{k+1}} \right] \middle| X_{t_k} \right] d\ell$$
$$= -\frac{w^2}{2\sigma_0^2} + \frac{uw}{\sigma_0^2} - \sum_{k=0}^{n-1} \frac{w}{\sqrt{n\Delta_n}} \int_0^1 \frac{1}{\lambda(\ell)} \frac{\ell w}{\sqrt{n\Delta_n}} \Delta_n d\ell$$
$$\overset{\mathrm{P}^{\theta_0,\sigma_0,\lambda_0}}{\longrightarrow} -\frac{w^2}{2\sigma_0^2} + \frac{uw}{\sigma_0^2} - \frac{w^2}{2\lambda_0},$$

which concludes (3.15).

Proof of (3.18). First, Lemma 3.2.7 yields that as $n \to \infty$,

$$\sum_{k=0}^{n-1} \left(\mathrm{E}\left[\beta_{k,n} | \widehat{\mathcal{F}}_{t_k}\right] \right)^2 = \sum_{k=0}^{n-1} \left(-\frac{w^2}{2\sigma_0^2 n} + \frac{uw}{\sigma_0^2 n} \right.$$
$$\left. + \frac{w}{\sqrt{n\Delta_n}} \int_0^1 \frac{1}{\lambda(\ell)} \mathrm{E}\left[\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\sigma_0,\lambda(\ell)} \left[\widetilde{M}_{t_{k+1}}^{\lambda(\ell)} - \widetilde{M}_{t_k}^{\lambda(\ell)} \middle| Y_{t_{k+1}}^{\theta_n,\sigma_0,\lambda(\ell)} = X_{t_{k+1}} \right] \middle| X_{t_k} \right] d\ell \right)^2$$
$$\overset{\mathrm{P}^{\theta_0,\sigma_0,\lambda_0}}{\longrightarrow} 0.$$

Next, we write

$$\sum_{k=0}^{n-1} \mathrm{E}\left[\beta_{k,n}^2 | \widehat{\mathcal{F}}_{t_k}\right] = S_{n,1} + S_{n,2} - 2S_{n,3},$$

where

$$S_{n,1} := \sum_{k=0}^{n-1} \frac{w^2}{n\Delta_n} \frac{1}{\sigma_0^4} \mathrm{E}\left[\left(\sigma_0 \left(B_{t_{k+1}} - B_{t_k} \right) + \frac{w\Delta_n}{2\sqrt{n\Delta_n}} - \frac{u\Delta_n}{\sqrt{n\Delta_n}} \right)^2 \middle| X_{t_k} \right],$$

$$S_{n,2} := \sum_{k=0}^{n-1} \frac{w^2}{n\Delta_n} \mathrm{E}\left[\left(\int_0^1 \frac{1}{\lambda(\ell)} \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\sigma_0,\lambda(\ell)} \left[\widetilde{M}_{t_{k+1}}^{\lambda(\ell)} - \widetilde{M}_{t_k}^{\lambda(\ell)} \middle| Y_{t_{k+1}}^{\theta_n,\sigma_0,\lambda(\ell)} = X_{t_{k+1}} \right] d\ell \right)^2 \middle| X_{t_k} \right],$$

$$S_{n,3} := \sum_{k=0}^{n-1} \frac{w^2}{n\Delta_n} \frac{1}{\sigma_0^2} \mathrm{E}\Bigg[\left(\sigma_0 \left(B_{t_{k+1}} - B_{t_k} \right) + \frac{w\Delta_n}{2\sqrt{n\Delta_n}} - \frac{u\Delta_n}{\sqrt{n\Delta_n}} \right)$$
$$\times \int_0^1 \frac{1}{\lambda(\ell)} \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\sigma_0,\lambda(\ell)} \left[\widetilde{M}_{t_{k+1}}^{\lambda(\ell)} - \widetilde{M}_{t_k}^{\lambda(\ell)} \middle| Y_{t_{k+1}}^{\theta_n,\sigma_0,\lambda(\ell)} = X_{t_{k+1}} \right] d\ell \middle| X_{t_k} \Bigg].$$

Using moment properties of Brownian motion, we get that as $n \to \infty$,

$$S_{n,1} \overset{\mathrm{P}^{\theta_0,\sigma_0,\lambda_0}}{\longrightarrow} \frac{w^2}{\sigma_0^2}.$$

Since $\widetilde{M}^{\lambda(\ell)}_{t_{k+1}} - \widetilde{M}^{\lambda(\ell)}_{t_k} = M_{t_{k+1}} - M_{t_k} - \lambda(\ell)\Delta_n$, we write $S_{n,2} = S_{n,2,1} - 2S_{n,2,2} + w^2\Delta_n$, where

$$S_{n,2,1} := \sum_{k=0}^{n-1} \frac{w^2}{n\Delta_n} \mathrm{E}\left[\left(\int_0^1 \frac{1}{\lambda(\ell)} \widetilde{\mathrm{E}}^{\theta_n,\sigma_0,\lambda(\ell)}_{X_{t_k}}\left[M_{t_{k+1}} - M_{t_k}\Big|Y^{\theta_n,\sigma_0,\lambda(\ell)}_{t_{k+1}} = X_{t_{k+1}}\right] d\ell\right)^2 \Big|X_{t_k}\right],$$

$$S_{n,2,2} := \sum_{k=0}^{n-1} \frac{w^2}{n} \int_0^1 \frac{1}{\lambda(\ell)} \mathrm{E}\left[\widetilde{\mathrm{E}}^{\theta_n,\sigma_0,\lambda(\ell)}_{X_{t_k}}\left[M_{t_{k+1}} - M_{t_k}\Big|Y^{\theta_n,\sigma_0,\lambda(\ell)}_{t_{k+1}} = X_{t_{k+1}}\right]\Big|X_{t_k}\right] d\ell.$$

Observe that Lemma 3.2.7 yields $S_{n,2,2} \overset{\mathrm{P}^{\theta_0,\sigma_0,\lambda_0}}{\longrightarrow} 0$ as $n \to \infty$. Moreover, adding and substracting the term $N_{t_{k+1}} - N_{t_k}$, we write $S_{n,2,1} = S_{n,2,1,1} + S_{n,2,1,2} - 2S_{n,2,1,3}$, where

$$S_{n,2,1,1} := \sum_{k=0}^{n-1} \frac{w^2}{n\Delta_n} \mathrm{E}\left[\left(\int_0^1 \frac{1}{\lambda(\ell)} \left(N_{t_{k+1}} - N_{t_k}\right) d\ell\right)^2 \Big|X_{t_k}\right],$$

$$S_{n,2,1,2} := \sum_{k=0}^{n-1} \frac{w^2}{n\Delta_n}$$

$$\times \mathrm{E}\left[\left(\int_0^1 \frac{1}{\lambda(\ell)} \left(N_{t_{k+1}} - N_{t_k} - \widetilde{\mathrm{E}}^{\theta_n,\sigma_0,\lambda(\ell)}_{X_{t_k}}\left[M_{t_{k+1}} - M_{t_k}\Big|Y^{\theta_n,\sigma_0,\lambda(\ell)}_{t_{k+1}} = X_{t_{k+1}}\right]\right) d\ell\right)^2 \Big|X_{t_k}\right],$$

$$S_{n,2,1,3} := \sum_{k=0}^{n-1} \frac{w^2}{n\Delta_n} \mathrm{E}\Bigg[\int_0^1 \frac{1}{\lambda(\ell)} \left(N_{t_{k+1}} - N_{t_k}\right) d\ell$$

$$\times \int_0^1 \frac{1}{\lambda(\ell)} \left(N_{t_{k+1}} - N_{t_k} - \widetilde{\mathrm{E}}^{\theta_n,\sigma_0,\lambda(\ell)}_{X_{t_k}}\left[M_{t_{k+1}} - M_{t_k}\Big|Y^{\theta_n,\sigma_0,\lambda(\ell)}_{t_{k+1}} = X_{t_{k+1}}\right]\right) d\ell\Big|X_{t_k}\Bigg].$$

Proceeding as in the proof of Lemma 3.3.1, one can easily show that $S_{n,2,1,2}$ and $S_{n,2,1,3}$ converge to zero in $\mathrm{P}^{\theta_0,\sigma_0,\lambda_0}$-probability as $n \to \infty$. Moreover, since $\mathrm{E}[(N_{t_{k+1}} - N_{t_k})^2|X_{t_k}] = \lambda_0\Delta_n + (\lambda_0\Delta_n)^2$, we deduce that as $n \to \infty$,

$$S_{n,2,1,1} \overset{\mathrm{P}^{\theta_0,\sigma_0,\lambda_0}}{\longrightarrow} \frac{w^2}{\lambda_0},$$

which implies that as $n \to \infty$,

$$S_{n,2} \overset{\mathrm{P}^{\theta_0,\sigma_0,\lambda_0}}{\longrightarrow} \frac{w^2}{\lambda_0}.$$

Next, we show that $n \to \infty$, $S_{n,3} \overset{\mathrm{P}^{\theta_0,\sigma_0,\lambda_0}}{\longrightarrow} 0$. Using Lemma 3.2.7, it suffices to show that as $n \to \infty$,

$$S_{n,3,1} = \sum_{k=0}^{n-1} \frac{w^2}{n\Delta_n} \frac{1}{\sigma_0} \int_0^1 \frac{1}{\lambda(\ell)}$$

$$\times \mathrm{E}\left[\left(B_{t_{k+1}} - B_{t_k}\right) \widetilde{\mathrm{E}}^{\theta_n,\sigma_0,\lambda(\ell)}_{X_{t_k}}\left[\widetilde{M}^{\lambda(\ell)}_{t_{k+1}} - \widetilde{M}^{\lambda(\ell)}_{t_k}\Big|Y^{\theta_n,\sigma_0,\lambda(\ell)}_{t_{k+1}} = X_{t_{k+1}}\right]\Big|X_{t_k}\right] d\ell \overset{\mathrm{P}^{\theta_0,\sigma_0,\lambda_0}}{\longrightarrow} 0.$$

Using the independence between B and N, we have that

$$S_{n,3,1} := -\sum_{k=0}^{n-1} \frac{w^2}{n\Delta_n} \frac{1}{\sigma_0} \int_0^1 \frac{1}{\lambda(\ell)} \mathrm{E}\Bigg[\left(B_{t_{k+1}} - B_{t_k}\right)$$

$$\times \left(N_{t_{k+1}} - N_{t_k} - \widetilde{\mathrm{E}}^{\theta_n,\sigma_0,\lambda(\ell)}_{X_{t_k}}\left[M_{t_{k+1}} - M_{t_k}\Big|Y^{\theta_n,\sigma_0,\lambda(\ell)}_{t_{k+1}} = X_{t_{k+1}}\right]\right)\Big|X_{t_k}\Bigg] d\ell.$$

Now, using Cauchy-Schwarz inequality, we get that

$$|S_{n,3,1}| \leq \sum_{k=0}^{n-1} \frac{w^2}{n\sqrt{\Delta_n}} \frac{1}{\sigma_0} \int_0^1 \frac{1}{|\lambda(\ell)|} \times \left(\mathrm{E}\left[\left(N_{t_{k+1}} - N_{t_k} - \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\sigma_0,\lambda(\ell)} \left[M_{t_{k+1}} - M_{t_k} \Big| Y_{t_{k+1}}^{\theta_n,\sigma_0,\lambda(\ell)} = X_{t_{k+1}} \right] \right)^2 \Big| X_{t_k} \right] \right)^{1/2} d\ell,$$

which converges to zero in $\mathrm{P}^{\theta_0,\sigma_0,\lambda_0}$-probability as $n \to \infty$ by proceeding as in Lemma 3.3.1.

Consequently, the proof of (3.18) is now completed.

Proof of (3.24). Applying Jensen's inequality, we get that

$$\sum_{k=0}^{n-1} \mathrm{E}\left[\beta_{k,n}^4 | \widehat{\mathcal{F}}_{t_k} \right] \leq 8 \sum_{k=0}^{n-1} \frac{w^4}{n^2 \Delta_n^2 \sigma_0^8} \mathrm{E}\left[\left(\sigma_0 \left(B_{t_{k+1}} - B_{t_k} \right) + \frac{w\Delta_n}{2\sqrt{n\Delta_n}} - \frac{u\Delta_n}{\sqrt{n\Delta_n}} \right)^4 \Big| X_{t_k} \right] + 8 \sum_{k=0}^{n-1} \frac{w^4}{n^2 \Delta_n^2} \int_0^1 \frac{1}{(\lambda(\ell))^4} \mathrm{E}\left[\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\sigma_0,\lambda(\ell)} \left[\left(\widetilde{M}_{t_{k+1}}^{\lambda(\ell)} - \widetilde{M}_{t_k}^{\lambda(\ell)} \right)^4 \Big| Y_{t_{k+1}}^{\theta_n,\sigma_0,\lambda(\ell)} = X_{t_{k+1}} \right] \Big| X_{t_k} \right] d\ell,$$

which converges to zero in $\mathrm{P}^{\theta_0,\sigma_0,\lambda_0}$-probability as $n \to \infty$, since $\mathrm{E}[(B_{t_{k+1}} - B_{t_k})^4 | X_{t_k}] = 3\Delta_n^2$ and for n large enough,

$$\left| \mathrm{E}\left[\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\sigma_0,\lambda(\ell)} \left[\left(\widetilde{M}_{t_{k+1}}^{\lambda(\ell)} - \widetilde{M}_{t_k}^{\lambda(\ell)} \right)^4 \Big| Y_{t_{k+1}}^{\theta_n,\sigma_0,\lambda(\ell)} = X_{t_{k+1}} \right] \Big| X_{t_k} \right] \right| \leq C\Delta_n,$$

for some constant $C > 0$, by using the same arguments as in the proof of Lemma 3.2.7.

Proof of (3.20). Using again Lemma 3.2.7, we get that as $n \to \infty$,

$$\sum_{k=0}^{n-1} \mathrm{E}\left[\xi_{k,n} | \widehat{\mathcal{F}}_{t_k} \right] \mathrm{E}\left[\beta_{k,n} | \widehat{\mathcal{F}}_{t_k} \right] = -\frac{u^2}{2\sigma_0^2 n} \sum_{k=0}^{n-1} \Bigg(-\frac{w^2}{2\sigma_0^2 n} + \frac{uw}{\sigma_0^2 n} + \frac{w}{\sqrt{n\Delta_n}} \int_0^1 \frac{1}{\lambda(\ell)} \mathrm{E}\left[\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\sigma_0,\lambda(\ell)} \left[\widetilde{M}_{t_{k+1}}^{\lambda(\ell)} - \widetilde{M}_{t_k}^{\lambda(\ell)} \Big| Y_{t_{k+1}}^{\theta_n,\sigma_0,\lambda(\ell)} = X_{t_{k+1}} \right] \Big| X_{t_k} \right] d\ell \Bigg) \overset{\mathrm{P}^{\theta_0,\sigma_0,\lambda_0}}{\longrightarrow} 0.$$

Moreover, basic computations yield that as $n \to \infty$,

$$\sum_{k=0}^{n-1} \mathrm{E}\left[\xi_{k,n} \beta_{k,n} | \widehat{\mathcal{F}}_{t_k} \right] = -\frac{uw}{\sigma_0^2} + \frac{u^2 w}{4n\sigma_0^2} (w - 2u) - \sum_{k=0}^{n-1} \frac{u^2 w \Delta_n}{2\sigma_0^2 n \Delta_n \sqrt{n\Delta_n}} \int_0^1 \frac{1}{\lambda(\ell)} \mathrm{E}\left[\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\sigma_0,\lambda(\ell)} \left[\widetilde{M}_{t_{k+1}}^{\lambda(\ell)} - \widetilde{M}_{t_k}^{\lambda(\ell)} \Big| Y_{t_{k+1}}^{\theta_n,\sigma_0,\lambda(\ell)} = X_{t_{k+1}} \right] \Big| X_{t_k} \right] d\ell + \sum_{k=0}^{n-1} \frac{uw}{n\Delta_n} \frac{1}{\sigma_0} \int_0^1 \frac{1}{\lambda(\ell)} \mathrm{E}\left[\left(B_{t_{k+1}} - B_{t_k} \right) \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\sigma_0,\lambda(\ell)} \left[\widetilde{M}_{t_{k+1}}^{\lambda(\ell)} - \widetilde{M}_{t_k}^{\lambda(\ell)} \Big| Y_{t_{k+1}}^{\theta_n,\sigma_0,\lambda(\ell)} = X_{t_{k+1}} \right] \Big| X_{t_k} \right] d\ell, \overset{\mathrm{P}^{\theta_0,\sigma_0,\lambda_0}}{\longrightarrow} -\frac{uw}{\sigma_0^2},$$

where we have used again Lemma 3.2.7 and proceeded as for the term $S_{n,3,1}$.

Similarly, we can show (3.21) and the proof of Theorem 3.1.1 is now completed. □

3.4 Conclusion and Final Comments

As we explained in the Introduction, the argument given here can be extended to more general cases with further arguments. We try to explain here in few words the strategy in the general case. In fact, Lemma 3.2.3 can be generalized to the case of stochastic differential equations with finite number of jumps and random jump size. Looking at the structure of the definition of S_j^p just before Lemma 3.2.4 one can see that the structure belongs to a large deviation principle for the process X. In fact, S_j^p describes a conditional expectation under $\widetilde{J}_{j,k}^c$ while the observation process satisfies $\widehat{J}_{j,k}$. The corresponding large deviation estimates are obtained in Lemmas 3.2.5 and 3.2.6. Finally one has to take limits in the above argument to obtain the aforementioned result for general jump driving processes.

3.5 Maximum likelihood estimator

By the Markov property, the log-likelihood function based on X^n can be written as follows

$$\begin{aligned}\ell_n(\theta_0,\sigma_0,\lambda_0) &= \log p(X^n;(\theta_0,\sigma_0,\lambda_0))\\ &= \sum_{k=0}^{n-1} \log p^{\theta_0,\sigma_0,\lambda_0}(\Delta_n, X_{t_k}, X_{t_{k+1}}).\end{aligned} \tag{3.25}$$

The maximum likelihood estimator $(\widehat{\theta}_n,\widehat{\sigma}_n,\widehat{\lambda}_n)$ of $(\theta_0,\sigma_0,\lambda_0)$ is defined as the solution to the system of equations

$$\begin{cases}\partial_\theta \ell_n(\theta_0,\sigma_0,\lambda_0) = 0\\ \partial_\sigma \ell_n(\theta_0,\sigma_0,\lambda_0) = 0\\ \partial_\lambda \ell_n(\theta_0,\sigma_0,\lambda_0) = 0.\end{cases} \tag{3.26}$$

Theorem 3.5.1. *Assume condition* (3.2). *Then, the maximum likelihood estimators* $(\widehat{\theta}_n,\widehat{\sigma}_n,\widehat{\lambda}_n)$ *of* $(\theta_0,\sigma_0,\lambda_0)$ *are consistent and asymptotically efficient. That is, as* $n\to\infty$,

$$(\widehat{\theta}_n,\widehat{\sigma}_n,\widehat{\lambda}_n) \overset{\mathrm{P}^{\theta_0,\sigma_0,\lambda_0}}{\longrightarrow} (\theta_0,\sigma_0,\lambda_0),$$

and

$$\left(\sqrt{n\Delta_n}(\widehat{\theta}_n-\theta_0), \sqrt{n}(\widehat{\sigma}_n-\sigma_0), \sqrt{n\Delta_n}(\widehat{\lambda}_n-\lambda_0)\right) \overset{\mathcal{L}(\mathrm{P}^{\theta_0,\sigma_0,\lambda_0})}{\longrightarrow} \mathcal{N}\left(0,\Gamma(\theta_0,\sigma_0,\lambda_0)^{-1}\right),$$

where $\mathcal{N}(0,\Gamma(\theta_0,\sigma_0,\lambda_0)^{-1})$ *is a centered* $\mathbb{R}^3$*-valued Gaussian vector with covariance matrix*

$$\Gamma(\theta_0,\sigma_0,\lambda_0)^{-1} = \begin{pmatrix}\lambda_0+\sigma_0^2 & 0 & \lambda_0\\ 0 & \frac{\sigma_0^2}{2} & 0\\ \lambda_0 & 0 & \lambda_0\end{pmatrix}.$$

Proof. Using (3.25) and Propositions 3.2.1 and 3.2.2, (3.26) is equivalent to

$$\begin{cases}\sum_{k=0}^{n-1} \dfrac{1}{\sigma_0}\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_0,\sigma_0,\lambda_0}\left[W_{t_{k+1}}-W_{t_k}\Big|Y_{t_{k+1}}^{\theta_0,\sigma_0,\lambda_0}=X_{t_{k+1}}\right]=0\\ \sum_{k=0}^{n-1}\left(\dfrac{1}{\sigma_0\Delta_n}\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_0,\sigma_0,\lambda_0}\left[\left(W_{t_{k+1}}-W_{t_k}\right)^2\Big|Y_{t_{k+1}}^{\theta_0,\sigma_0,\lambda_0}=X_{t_{k+1}}\right]-\frac{1}{\sigma_0}\right)=0\\ \sum_{k=0}^{n-1}\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_0,\sigma_0,\lambda_0}\left[-\dfrac{W_{t_{k+1}}-W_{t_k}}{\sigma_0}+\dfrac{\widetilde{M}_{t_{k+1}}^{\lambda_0}-\widetilde{M}_{t_k}^{\lambda_0}}{\lambda_0}\Big|Y_{t_{k+1}}^{\theta_0,\sigma_0,\lambda_0}=X_{t_{k+1}}\right]=0.\end{cases}$$

Using (3.11) with $(\theta_0,\sigma_0,\lambda_0)$ instead of $(\theta(\ell),\sigma_0,\lambda_0)$ and taking the conditional expectation, we obtain that

$$\widehat{\theta}_n = \theta_0 + \frac{1}{n\Delta_n}\sum_{k=0}^{n-1}\Big(\sigma_0\left(B_{t_{k+1}} - B_{t_k}\right) + \widetilde{N}^{\lambda_0}_{t_{k+1}} - \widetilde{N}^{\lambda_0}_{t_k}\Big)$$
$$+ \widehat{\lambda}_n - \frac{1}{n\Delta_n}\sum_{k=0}^{n-1}\widetilde{\mathrm{E}}^{\theta_0,\sigma_0,\lambda_0}_{X_{t_k}}\left[M_{t_{k+1}} - M_{t_k}\Big| Y^{\theta_0,\sigma_0,\lambda_0}_{t_{k+1}} = X_{t_{k+1}}\right],$$
$$\widehat{\sigma}^2_n = \frac{\sigma_0^2}{n\Delta_n}\sum_{k=0}^{n-1}\left(B_{t_{k+1}} - B_{t_k}\right)^2 + \frac{1}{n\Delta_n}\sum_{k=0}^{n-1}\bigg\{\Big(\theta_0\Delta_n + \widetilde{N}^{\lambda_0}_{t_{k+1}} - \widetilde{N}^{\lambda_0}_{t_k}\Big)^2 + 2\sigma_0\left(B_{t_{k+1}} - B_{t_k}\right)$$
$$\times\Big(\theta_0\Delta_n + \widetilde{N}^{\lambda_0}_{t_{k+1}} - \widetilde{N}^{\lambda_0}_{t_k}\Big) - \widetilde{\mathrm{E}}^{\theta_0,\sigma_0,\lambda_0}_{X_{t_k}}\bigg[\Big(\theta_0\Delta_n + \widetilde{M}^{\lambda_0}_{t_{k+1}} - \widetilde{M}^{\lambda_0}_{t_k}\Big)^2$$
$$+ 2\sigma_0\left(W_{t_{k+1}} - W_{t_k}\right)\Big(\theta_0\Delta_n + \widetilde{M}^{\lambda_0}_{t_{k+1}} - \widetilde{M}^{\lambda_0}_{t_k}\Big)\Big| Y^{\theta_0,\sigma_0,\lambda_0}_{t_{k+1}} = X_{t_{k+1}}\bigg]\bigg\},$$
$$\widehat{\lambda}_n = \frac{\frac{1}{n\Delta_n}\sum_{k=0}^{n-1}\widetilde{\mathrm{E}}^{\theta_0,\sigma_0,\lambda_0}_{X_{t_k}}\left[M_{t_{k+1}} - M_{t_k}\Big| Y^{\theta_0,\sigma_0,\lambda_0}_{t_{k+1}} = X_{t_{k+1}}\right]}{\frac{1}{\sigma_0 n\Delta_n}\sum_{k=0}^{n-1}\widetilde{\mathrm{E}}^{\theta_0,\sigma_0,\lambda_0}_{X_{t_k}}\left[W_{t_{k+1}} - W_{t_k}\Big| Y^{\theta_0,\sigma_0,\lambda_0}_{t_{k+1}} = X_{t_{k+1}}\right] + 1}.$$

Next, using the fact that as $n \to \infty$,

$$\frac{1}{n\Delta_n}\sum_{k=0}^{n-1}\Big(\sigma_0\left(B_{t_{k+1}} - B_{t_k}\right) + \widetilde{N}^{\lambda_0}_{t_{k+1}} - \widetilde{N}^{\lambda_0}_{t_k}\Big) \overset{\mathrm{P}^{\theta_0,\sigma_0,\lambda_0}}{\longrightarrow} 0,$$
$$\frac{1}{n\Delta_n}\sum_{k=0}^{n-1}\widetilde{\mathrm{E}}^{\theta_0,\sigma_0,\lambda_0}_{X_{t_k}}\left[M_{t_{k+1}} - M_{t_k}\Big| Y^{\theta_0,\sigma_0,\lambda_0}_{t_{k+1}} = X_{t_{k+1}}\right] \overset{\mathrm{P}^{\theta_0,\sigma_0,\lambda_0}}{\longrightarrow} \lambda_0,$$
$$\frac{1}{n\Delta_n}\sum_{k=0}^{n-1}\left(B_{t_{k+1}} - B_{t_k}\right)^2 \overset{\mathrm{P}^{\theta_0,\sigma_0,\lambda_0}}{\longrightarrow} 1,$$
$$\frac{1}{n\Delta_n}\sum_{k=0}^{n-1}\widetilde{\mathrm{E}}^{\theta_0,\sigma_0,\lambda_0}_{X_{t_k}}\left[W_{t_{k+1}} - W_{t_k}\Big| Y^{\theta_0,\sigma_0,\lambda_0}_{t_{k+1}} = X_{t_{k+1}}\right] \overset{\mathrm{P}^{\theta_0,\sigma_0,\lambda_0}}{\longrightarrow} 0,$$

and, by proceeding as for the term $M_{k,n}$,

$$\frac{1}{n\Delta_n}\sum_{k=0}^{n-1}\bigg\{\Big(\theta_0\Delta_n + \widetilde{N}^{\lambda_0}_{t_{k+1}} - \widetilde{N}^{\lambda_0}_{t_k}\Big)^2 + 2\sigma_0\left(B_{t_{k+1}} - B_{t_k}\right)\Big(\theta_0\Delta_n + \widetilde{N}^{\lambda_0}_{t_{k+1}} - \widetilde{N}^{\lambda_0}_{t_k}\Big)$$
$$- \widetilde{\mathrm{E}}^{\theta_0,\sigma_0,\lambda_0}_{X_{t_k}}\bigg[\Big(\theta_0\Delta_n + \widetilde{M}^{\lambda_0}_{t_{k+1}} - \widetilde{M}^{\lambda_0}_{t_k}\Big)^2$$
$$+ 2\sigma_0\left(W_{t_{k+1}} - W_{t_k}\right)\Big(\theta_0\Delta_n + \widetilde{M}^{\lambda_0}_{t_{k+1}} - \widetilde{M}^{\lambda_0}_{t_k}\Big)\Big| Y^{\theta_0,\sigma_0,\lambda_0}_{t_{k+1}} = X_{t_{k+1}}\bigg]\bigg\} \overset{\mathrm{P}^{\theta_0,\sigma_0,\lambda_0}}{\longrightarrow} 0,$$

we conclude that as $n \to \infty$,

$$(\widehat{\theta}_n, \widehat{\sigma}_n, \widehat{\lambda}_n) \overset{\mathrm{P}^{\theta_0,\sigma_0,\lambda_0}}{\longrightarrow} (\theta_0, \sigma_0, \lambda_0).$$

On the other hand, we can write

$$\sqrt{n\Delta_n}(\widehat{\theta}_n - \theta_0) = \sum_{k=0}^{n-1}\zeta_{k,n,1} + \sum_{k=0}^{n-1}R_{k,n,1},$$
$$\sqrt{n}(\widehat{\sigma}^2_n - \sigma^2_0) = \sum_{k=0}^{n-1}\zeta_{k,n,2} + \sum_{k=0}^{n-1}R_{k,n,2},$$
$$\sqrt{n\Delta_n}(\widehat{\lambda}_n - \lambda_0) = \sum_{k=0}^{n-1}\zeta_{k,n,3} - \sum_{k=0}^{n-1}R_{k,n,3},$$

where

$$\zeta_{k,n,1} = \frac{1}{\sqrt{n\Delta_n}}\left(\sigma_0\left(B_{t_{k+1}} - B_{t_k}\right) + \widetilde{N}^{\lambda_0}_{t_{k+1}} - \widetilde{N}^{\lambda_0}_{t_k}\right),$$
$$\zeta_{k,n,2} = \frac{\sigma_0^2}{\sqrt{n\Delta_n}}\left(\left(B_{t_{k+1}} - B_{t_k}\right)^2 - \Delta_n\right),$$
$$\zeta_{k,n,3} = \frac{1}{\sqrt{n\Delta_n}}\left(\widetilde{N}^{\lambda_0}_{t_{k+1}} - \widetilde{N}^{\lambda_0}_{t_k}\right),$$

and

$$R_{k,n,1} = \sqrt{n\Delta_n}\left(\widehat{\lambda}_n - \frac{1}{n\Delta_n}\sum_{k=0}^{n-1}\widetilde{\mathrm{E}}^{\theta_0,\sigma_0,\lambda_0}_{X_{t_k}}\left[M_{t_{k+1}} - M_{t_k}\middle| Y^{\theta_0,\sigma_0,\lambda_0}_{t_{k+1}} = X_{t_{k+1}}\right]\right),$$

$$\begin{aligned}R_{k,n,2} = \frac{1}{\sqrt{n\Delta_n}}\sum_{k=0}^{n-1}\Bigg\{&\left(\theta_0\Delta_n + \widetilde{N}^{\lambda_0}_{t_{k+1}} - \widetilde{N}^{\lambda_0}_{t_k}\right)^2 + 2\sigma_0\left(B_{t_{k+1}} - B_{t_k}\right)\left(\theta_0\Delta_n + \widetilde{N}^{\lambda_0}_{t_{k+1}} - \widetilde{N}^{\lambda_0}_{t_k}\right)\\
&- \widetilde{\mathrm{E}}^{\theta_0,\sigma_0,\lambda_0}_{X_{t_k}}\Bigg[\left(\theta_0\Delta_n + \widetilde{M}^{\lambda_0}_{t_{k+1}} - \widetilde{M}^{\lambda_0}_{t_k}\right)^2\\
&+ 2\sigma_0\left(W_{t_{k+1}} - W_{t_k}\right)\left(\theta_0\Delta_n + \widetilde{M}^{\lambda_0}_{t_{k+1}} - \widetilde{M}^{\lambda_0}_{t_k}\right)\Bigg| Y^{\theta_0,\sigma_0,\lambda_0}_{t_{k+1}} = X_{t_{k+1}}\Bigg]\Bigg\},\end{aligned}$$

$$\begin{aligned}R_{k,n,3} = &\frac{1}{\sqrt{n\Delta_n}}\sum_{k=0}^{n-1}\left(N_{t_{k+1}} - N_{t_k} - \widetilde{\mathrm{E}}^{\theta_0,\sigma_0,\lambda_0}_{X_{t_k}}\left[M_{t_{k+1}} - M_{t_k}\middle| Y^{\theta_0,\sigma_0,\lambda_0}_{t_{k+1}} = X_{t_{k+1}}\right]\right)\\
&+ \frac{1}{\sqrt{n\Delta_n}}\sum_{k=0}^{n-1}\widetilde{\mathrm{E}}^{\theta_0,\sigma_0,\lambda_0}_{X_{t_k}}\left[M_{t_{k+1}} - M_{t_k}\middle| Y^{\theta_0,\sigma_0,\lambda_0}_{t_{k+1}} = X_{t_{k+1}}\right]\\
&\times\frac{\frac{1}{\sigma_0 n\Delta_n}\sum_{k=0}^{n-1}\widetilde{\mathrm{E}}^{\theta_0,\sigma_0,\lambda_0}_{X_{t_k}}\left[W_{t_{k+1}} - W_{t_k}\middle| Y^{\theta_0,\sigma_0,\lambda_0}_{t_{k+1}} = X_{t_{k+1}}\right]}{\frac{1}{\sigma_0 n\Delta_n}\sum_{k=0}^{n-1}\widetilde{\mathrm{E}}^{\theta_0,\sigma_0,\lambda_0}_{X_{t_k}}\left[W_{t_{k+1}} - W_{t_k}\middle| Y^{\theta_0,\sigma_0,\lambda_0}_{t_{k+1}} = X_{t_{k+1}}\right] + 1}.\end{aligned}$$

Notice that the random variables $\zeta_{k,n,1}, \zeta_{k,n,2}, \zeta_{k,n,3}$ are the terms that contribute to the limit in Theorem 3.5.1. On the other hand, it can be checked that the random variables $R_{k,n,1}, R_{k,n,2}$ and $R_{k,n,3}$ are the negligible terms. Then, applying the central limit theorem for triangular arrays, we obtain that as $n \to \infty$,

$$\left(\sqrt{n\Delta_n}(\widehat{\theta}_n - \theta_0), \sqrt{n}(\widehat{\sigma}_n^2 - \sigma_0^2), \sqrt{n\Delta_n}(\widehat{\lambda}_n - \lambda_0)\right) \overset{\mathcal{L}(\mathrm{P}^{\theta_0,\sigma_0,\lambda_0})}{\longrightarrow} \mathcal{N}\left(0, I(\theta_0,\sigma_0,\lambda_0)\right),$$

where

$$I(\theta_0,\sigma_0,\lambda_0) = \begin{pmatrix}\lambda_0 + \sigma_0^2 & 0 & \lambda_0\\ 0 & 2\sigma_0^4 & 0\\ \lambda_0 & 0 & \lambda_0\end{pmatrix}.$$

This, together with the fact that $\widehat{\sigma}_n \overset{\mathrm{P}^{\theta_0,\sigma_0,\lambda_0}}{\longrightarrow} \sigma_0$ as $n \to \infty$, yields

$$\left(\sqrt{n\Delta_n}(\widehat{\theta}_n - \theta_0), \sqrt{n}(\widehat{\sigma}_n - \sigma_0), \sqrt{n\Delta_n}(\widehat{\lambda}_n - \lambda_0)\right) \overset{\mathcal{L}(\mathrm{P}^{\theta_0,\sigma_0,\lambda_0})}{\longrightarrow} \mathcal{N}\left(0, \Gamma(\theta_0,\sigma_0,\lambda_0)^{-1}\right),$$

which finishes the proof of Theorem 3.5.1. □

Chapitre 4

LAN property for a jump-diffusion process : drift parameter

In this chapter, we consider an ergodic diffusion process with jumps driven by a Brownian motion and a Poisson random measure associated with a centered pure-jump Lévy process with finite Lévy measure, whose drift coefficient depends on an unknown parameter. Supposing that the process is observed discretely at high frequency, we derive the local asymptotic normality (LAN) property. In order to obtain this result, Malliavin calculus and Girsanov's theorem are applied in order to write the log-likelihood ratio in terms of sums of conditional expectations, for which a central limit theorem for triangular arrays can be applied.

4.1 Introduction and main result

On a complete probability space $(\Omega, \mathcal{F}, \mathrm{P})$ defined in Definition 1.1.3, we consider the process $X^\theta = (X_t^\theta)_{t\geq 0}$ solution to the following stochastic differential equation with jumps

$$dX_t^\theta = b(\theta, X_t^\theta)dt + \sigma(X_t^\theta)dB_t + \int_{\mathbb{R}_0} c(X_{t-}^\theta, z)\left(N(dt,dz) - \nu(dz)dt\right), \tag{4.1}$$

where $X_0^\theta = x_0 \in \mathbb{R}$, $\mathbb{R}_0 := \mathbb{R}\setminus\{0\}$, $B = (B_t)_{t\geq 0}$ is a standard Brownian motion, and $N(dt,dz)$ is a Poisson random measure in $(\mathbb{R}_+ \times \mathbb{R}_0, \mathcal{B}(\mathbb{R}_+ \times \mathbb{R}_0))$ independent of B, with intensity measure $\nu(dz)dt$, and finite Lévy measure $\lambda = \int_{\mathbb{R}_0}\nu(dz) < \infty$. The compensated Poisson random measure is denoted by $\widetilde{N}(dt,dz) := N(dt,dz) - \nu(dz)dt$. Let $\widehat{Z} = (\widehat{Z}_t)_{t\geq 0}$ be a centered pure-jump Lévy process associated with $N(dt,dz)$, i.e., $\widehat{Z}_t = \int_0^t\int_{\mathbb{R}_0} z(N(ds,dz) - \nu(dz)ds)$, for $t \geq 0$. Let $\{\widehat{\mathcal{F}}_t\}_{t\geq 0}$ denote the natural filtration generated by B and N. The unknown parameter θ belongs to Θ which is a closed interval of $\mathbb{R}$. The coefficients $b : \Theta\times\mathbb{R} \to \mathbb{R}, \sigma : \mathbb{R} \to \mathbb{R}$ and $c : \mathbb{R}\times\mathbb{R}_0 \to \mathbb{R}$ are measurable functions satisfying condition **(A1)** below under which equation (4.1) has a unique $\widehat{\mathcal{F}}_t$-adapted càdlàg solution X^θ. We denote by P^θ the probability law induced by X^θ, and by E^θ the expectation with respect to P^θ. Let $\overset{\mathrm{P}^\theta}{\longrightarrow}$ and $\overset{\mathcal{L}(\mathrm{P}^\theta)}{\longrightarrow}$ denote the convergence in P^θ-probability and in P^θ-law, respectively.

Recall that the structure of the probability space is given by $\widehat{\Omega} = \Omega^1 \times \Omega^2$, $\widetilde{\Omega} = \Omega^3 \times \Omega^4$, $\widehat{\mathcal{F}} = \mathcal{F}^1 \otimes \mathcal{F}^2$, $\widetilde{\mathcal{F}} = \mathcal{F}^3 \otimes \mathcal{F}^4$, $\widehat{\mathrm{P}} = \mathrm{P}^1 \otimes \mathrm{P}^2$, $\widetilde{\mathrm{P}} = \mathrm{P}^3 \otimes \mathrm{P}^4$, and $\Omega = \widehat{\Omega}\times\widetilde{\Omega}$, $\mathcal{F} = \widehat{\mathcal{F}}\otimes\widetilde{\mathcal{F}}$, $\mathrm{P} = \widehat{\mathrm{P}}\otimes\widetilde{\mathrm{P}}$. We denote by E, $\widehat{\mathrm{E}}$, $\widetilde{\mathrm{E}}$ the expectation with respect to P, $\widehat{\mathrm{P}}$ and $\widetilde{\mathrm{P}}$, respectively.

For fixed $\theta_0 \in \Theta$ and $n \geq 1$, we consider a discrete observation scheme at equidistant times $t_k = k\Delta_n$, $k \in \{0, ..., n\}$ of the diffusion process X^{θ_0}, which is denoted by $X^n = (X_{t_0}, X_{t_1}, ..., X_{t_n})$, where $\Delta_n \leq 1$. We assume that the sequence of time-step sizes Δ_n satisfies the high-frequency observation condition

$$n\Delta_n \to \infty, \quad \text{and} \quad \Delta_n \to 0, \quad \text{as } n \to \infty.$$

We consider the following hypotheses on equation (4.1).

(A1) For any $\theta \in \Theta$, there exists a constant $C > 0$ such that for all $x, y \in \mathbb{R}, z, z_1, z_2 \in \mathbb{R}_0$,

$$\begin{aligned}
&|b(\theta,x) - b(\theta,y)| + |\sigma(x) - \sigma(y)| \leq C|x-y|, \quad |b(\theta,x)| \leq C\left(1+|x|\right),\\
&|c(x,z) - c(y,z)| \leq C|x-y||z|, \quad |c(x,z)| \leq C(1+|x|)|z|,\\
&|c(x,z_1) - c(x,z_2)| \leq C\left(1+|x|\right)|z_1 - z_2|.
\end{aligned}$$

(A2) There exists a constant $c \geq 1$ such that for all $x \in \mathbb{R}$,

$$\frac{1}{c} \leq |\sigma(x)| \leq c.$$

(A3) For all $(x,z) \in \mathbb{R} \times \mathbb{R}_0, c(x,z) \neq 0$, and $c(x,0) = 0$. Moreover, there exists a constant $C > 0$ such that for all $z \in \mathbb{R}_0$,

$$\inf_{x \in \mathbb{R}} |c(x,z)| \geq C|z|.$$

(A4) The functions b, σ and c are of class C^1 w.r.t. θ and x. Each partial derivative $\partial_\theta b$, $\partial_x b$, $\partial_x \sigma$ and $\partial_x c$ is of class C^1 w.r.t. x. Moreover, there exist positive constants C, q, ϵ, η, independent of $(\theta, \theta_1, \theta_2, x, y, z) \in \Theta^3 \times \mathbb{R}^2 \times \mathbb{R}_0$ such that

(a) $|\partial_x b(\theta,x)| + |\partial_x \sigma(x)| + |\partial_x c(x,z)| \leq C$;

(b) $|h(\cdot,x)| \leq C\left(1+|x|^q\right)$ for $h(\cdot,x) = \partial_\theta b(\theta,x), \partial_x^2 b(\theta,x), \partial_{x,\theta}^2 b(\theta,x)$ or $\partial_x^2 \sigma(x)$;

(c) $|\partial_\theta b(\theta_1,x) - \partial_\theta b(\theta_2,x)| \leq C|\theta_1 - \theta_2|^\epsilon \left(1+|x|^q\right)$;

(d) $|\partial_\theta b(\theta,x) - \partial_\theta b(\theta,y)| \leq C|x-y|$;

(e) $|\partial_x^2 c(x,z)| \leq C|z|\left(1+|x|\right)$ and $|1+\partial_x c(x,z)| \geq \eta$.

(A5) For any $p \geq 2$, $\int_{\mathbb{R}_0} |z|^p \nu(dz) < \infty$.

(A6) The process X^{θ_0} is ergodic in the sense that there exists a unique probability measure $\pi_{\theta_0}(dx)$ such that as $T \to \infty$,

$$\frac{1}{T}\int_0^T g(X_t^{\theta_0})dt \xrightarrow{\mathrm{P}^{\theta_0}} \int_{\mathbb{R}} g(x)\pi_{\theta_0}(dx),$$

for any π_{θ_0}-integrable function $g : \mathbb{R} \to \mathbb{R}$.

(A7) There exist constants $\epsilon > 0$, $q > 1$, $\rho_1, \rho_2 > 0$ and $0 < \upsilon, \gamma < \frac{1}{2}$ such that as $n \to \infty$,

$$\frac{\sqrt{n}}{\Delta_n^\epsilon}\left(n\Delta_n\left(\int_{\{|z|\geq \rho_2\Delta_n^{-\gamma}\}} \nu(dz) + \int_{\{|z|\leq \rho_1\Delta_n^{\upsilon}\}} \nu(dz)\right)\right)^{\frac{1}{q}} \to 0.$$

(A8) There exist constants $n_0 \geq 1$ and $C > 0$ such that

$$\sup_{n \geq n_0} \max_{k \in \{0,\ldots,n\}} \mathrm{E}\left[e^{C\Delta_n^{1-2\gamma}X_{t_k}^2}\right] < \infty,$$

where γ is as in **(A7)**.

Remark 4.1.1. *In the case where the jump coefficient c is lower bounded, then* **(A8)** *implies that $\int_{\mathbb{R}} e^{Cz^2}\nu(dz) < \infty$ for some $C > 0$, which in particular implies* **(A5)**.

A detailed explanation on the hypotheses is given in the subsection 1.3.3 of the introductory chapter.

Conditions **(A1)-(A2)** imply that the law of the discrete observation $(X_{t_0}^\theta, X_{t_1}^\theta, \ldots, X_{t_n}^\theta)$ of the process $(X_t^\theta)_{t\geq 0}$ has a density in $\mathbb{R}^{n+1}$ that we denote by $p(\cdot;\theta)$. In particular, $p(\cdot;\theta_0)$ denotes the density of the random vector X^n. The main result of this chapter is the following LAN property.

Theorem 4.1.1. *Assume conditions* **(A1)-(A8)**. *Then, the LAN property holds for the likelihood at θ_0 with rate of convergence $\sqrt{n\Delta_n}$ and asymptotic Fisher information $\Gamma(\theta_0)$. That is, for all $u \in \mathbb{R}$, as $n \to \infty$,*

$$\log \frac{p(X^n;\theta_n)}{p(X^n;\theta_0)} \overset{\mathcal{L}(\mathrm{P}^{\theta_0})}{\longrightarrow} u\mathcal{N}\left(0, \Gamma(\theta_0)\right) - \frac{u^2}{2}\Gamma\left(\theta_0\right),$$

where $\theta_n = \theta_0 + \frac{u}{\sqrt{n\Delta_n}}$, and $\mathcal{N}(0,\Gamma(\theta_0))$ is a centered Gaussian random variable with variance

$$\Gamma\left(\theta_0\right) = \int_{\mathbb{R}} \left(\frac{\partial_\theta b(\theta_0, x)}{\sigma(x)}\right)^2 \pi_{\theta_0}(dx).$$

Remark 4.1.2. *In the case where the drift coefficient is bounded, condition* **(A8)** *is not needed.*

Remark 4.1.3. *We remark that $\Gamma(\theta_0)$ is identical to the asymptotic Fisher information for ergodic diffusion processes without jumps (see [25, Theorem 4.1]). This is due to the fact that the jump component is dominated over by the Gaussian component, which will be seen in the discussion of the subsection 4.3.1.*

Example 4.1.1. 1) *Consider the Ornstein-Uhlenbeck process with jumps defined as*

$$X_t^\theta = x_0 - \theta \int_0^t X_s^\theta ds + \sigma B_t + \int_0^t \int_{\mathbb{R}_0} z\widetilde{N}(ds, dz),$$

where $\theta > 0$, $\sigma \in \mathbb{R}_0$ and the Lévy measure satisfies **(A5)**, **(A7)**, *and is finite. Then X^θ is ergodic in the sense of* **(A6)**, *and the invariant probability measure $\pi_\theta(dx)$ can be calculated explicitly (see [65, Theorem 17.5 and Corollary 17.9] and [53, Theorem 2.6]). In particular,*

$$\Gamma(\theta) := \int_{\mathbb{R}} \frac{x^2}{\sigma^2} \pi_\theta(dx) = \frac{1}{2\theta}\left(1 + \frac{1}{\sigma^2}\int_{\mathbb{R}_0} z^2 \nu(dz)\right).$$

In addition, assume that there exists a constant $C > 0$ such that $\int_{\mathbb{R}_0} e^{Cz^2}\nu(dz) < \infty$. Then, the infinitesimal generator of X^θ satisfies that $\mathcal{A}e^{Cx^2} \leq -c_1 e^{Cx^2} + c_2$, for some constants $c_1, c_2 > 0$. Then, by [53, Theorem 2.2], condition **(A8)** *is satisfied.*

As a consequence of Theorem 4.1.1, the LAN property holds with rate of convergence $\sqrt{n\Delta_n}$ and asymptotic Fisher information $\Gamma(\theta_0)$.

2) *Consider the process*

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

where $\theta \in \mathbb{R}$ and $\sigma \in \mathbb{R}_0$, and the Lévy measure is finite. Under conditions **(A5)** *and* **(A7)-(A8)**, *the LAN property holds with rate of convergence $\sqrt{n\Delta_n}$ and asymptotic Fisher information $\Gamma(\theta_0) = \frac{1}{\sigma^2}$. In this case condition* **(A6)** *fails.*

3) *Assume that $\nu(dz)$ has compact support on $\{c \leq |z| \leq C\}$, for some constants $c, C > 0$. In this case, condition* **(A7)** *holds.*

4) *Assume that* $\nu(dz)=\varphi(z)\mathbf{1}_{\{|z|\geq 1\}}dz$, *where* φ *is the standard Gaussian density. Then, for* n *sufficiently large,*

$$\frac{\sqrt{n}}{\Delta_n^{\epsilon}}\left(n\Delta_n\left(\int_{\{|z|\geq\rho_2\Delta_n^{-\gamma}\}}\nu(dz)+\int_{\{|z|\leq\rho_1\Delta_n^{\upsilon}\}}\nu(dz)\right)\right)^{\frac{1}{q}}$$
$$=\frac{\sqrt{n}}{\Delta_n^{\epsilon}}\left(n\Delta_n\int_{\{|z|\geq\rho_2\Delta_n^{-\gamma}\}}\nu(dz)\right)^{\frac{1}{q}}\leq c_q\frac{\sqrt{n}}{\Delta_n^{\epsilon}}\left(n\Delta_n\right)^{\frac{1}{q}}e^{-c_q\Delta_n^{-2\gamma}},$$

which tends to zero as $n\to\infty$ *for all* $\epsilon>0$, $q>1$, $\rho_1,\rho_2>0$ *and* $0<\upsilon,\gamma<\frac{1}{2}$, *and thus* **(A7)** *holds.*

As usual, constants will be denoted by C or c and they will always be independent of time and Δ_n but may depend on bounds for the set Θ. They may change of value from one line to the next.

4.2 Preliminaries

In this section we introduce some preliminary results needed for the proof of Theorem 4.1.1.

We start as in Gobet [24] applying the integration by parts formula of the Malliavin calculus on the Wiener space to analyze the log-likelihood function. In order to avoid confusion with the observed process X^{θ}, we introduce an extra probabilistic representation of X^{θ} where the Malliavin calculus will be applied. That is, consider the flow $Y^{\theta}(s,x)=(Y_t^{\theta}(s,x),t\geq s)$, $x\in\mathbb{R}$ on the time interval $[s,\infty)$ and with initial condition $Y_s^{\theta}(s,x)=x$ satisfying

$$\begin{aligned}Y_t^{\theta}(s,x)=x&+\int_s^t b(\theta,Y_u^{\theta}(s,x))du+\int_s^t\sigma(Y_u^{\theta}(s,x))dW_u\\&+\int_s^t\int_{\mathbb{R}_0}c(Y_{u-}^{\theta}(s,x),z)\left(M(du,dz)-\nu(dz)du\right),\end{aligned}\tag{4.2}$$

where $W=(W_t)_{t\geq 0}$ is a Brownian motion, $M(dt,dz)$ is a Poisson random measure with intensity measure $\nu(dz)dt$ associated with a centered pure-jump Lévy process $\widetilde{Z}=(\widetilde{Z}_t)_{t\geq 0}$ independent of W, and we denote by $\widetilde{M}(dt,dz):=M(dt,dz)-\nu(dz)dt$ the compensated Poisson random measure. In particular, we write $Y_t^{\theta}\equiv Y_t^{\theta}(0,x_0)$, for all $t\geq 0$. That is,

$$Y_t^{\theta}=x_0+\int_0^t b(\theta,Y_u^{\theta})du+\int_0^t\sigma(Y_u^{\theta})dW_u+\int_0^t\int_{\mathbb{R}_0}c(Y_{u-}^{\theta},z)\left(M(du,dz)-\nu(dz)du\right).\tag{4.3}$$

Here, we consider the Malliavin calculus on the Wiener space induced by the Brownian motion W, and we denote by D and δ the Malliavin derivative and the Skorohod integral with respect to W on each interval $[t_k,t_{k+1}]$, respectively (see the Definition 1.1.3 and the discussion following it). For all $A\in\widetilde{\mathcal{F}}$, let us denote $\widetilde{\mathrm{P}}_x^{\theta}(A)=\widetilde{\mathrm{E}}[\mathbf{1}_A|Y_{t_k}^{\theta}=x]$. We denote by $\widetilde{\mathrm{E}}_x^{\theta}$ the expectation with respect to $\widetilde{\mathrm{P}}_x^{\theta}$. That is, for all $\widetilde{\mathcal{F}}$-measurable random variable V, we have that $\widetilde{\mathrm{E}}_x^{\theta}[V]=\widetilde{\mathrm{E}}[V|Y_{t_k}^{\theta}=x]$.

Under conditions **(A1)-(A4)**, for any $t>s$ the law of Y_t^{θ} conditioned on $Y_s^{\theta}=x$ admits a positive transition density $p^{\theta}(t-s,x,y)$, which is differentiable w.r.t. θ. As a consequence of [24, Proposition 4.1], we have the following expression for the derivative of the log-likelihood function w.r.t. θ in terms of a conditional expectation, although one can also follow the same steps as in the proof of Proposition 3.2.1.

Proposition 4.2.1. *Assume conditions* **(A1)**-**(A4)**. *Then for all* $k\in\{0,...,n-1\}$ *and* $\theta\in\Theta$,

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

where $U_t^\theta(t_k,x) = (D_t Y_{t_{k+1}}^\theta(t_k,x))^{-1} = (\partial_x Y_{t_{k+1}}^\theta(t_k,x))^{-1}\partial_x Y_t^\theta(t_k,x)\sigma^{-1}(Y_t^\theta(t_k,x))$ *for all* $t \in [t_k,t_{k+1}]$*, and the processes* $(\partial_\theta Y_t^\theta(t_k,x), t\in[t_k,t_{k+1}])$ *and* $(\partial_x Y_t^\theta(t_k,x), t\in[t_k,t_{k+1}])$ *denote the solutions to linear equations*

$$\partial_\theta Y_t^\theta(t_k,x) = \int_{t_k}^t \Big(\partial_\theta b(\theta, Y_s^\theta(t_k,x)) + \partial_x b(\theta, Y_s^\theta(t_k,x))\partial_\theta Y_s^\theta(t_k,x)\Big)ds$$
$$+ \int_{t_k}^t \partial_x\sigma(Y_s^\theta(t_k,x))\partial_\theta Y_s^\theta(t_k,x)dW_s + \int_{t_k}^t\int_{\mathbb{R}_0}\partial_x c(Y_{s-}^\theta(t_k,x),z)\partial_\theta Y_s^\theta(t_k,x)\widetilde{M}(ds,dz),$$
$$\partial_x Y_t^\theta(t_k,x) = 1 + \int_{t_k}^t \partial_x b(\theta, Y_s^\theta(t_k,x))\partial_x Y_s^\theta(t_k,x)ds + \int_{t_k}^t \partial_x\sigma(Y_s^\theta(t_k,x))\partial_x Y_s^\theta(t_k,x)dW_s$$
$$+ \int_{t_k}^t\int_{\mathbb{R}_0}\partial_x c(Y_{s-}^\theta(t_k,x),z)\partial_x Y_s^\theta(t_k,x)\widetilde{M}(ds,dz).$$

We have the following decomposition of the Skorohod integral appearing in the conditional expectation of Proposition 4.2.1.

Lemma 4.2.1. *Under conditions* **(A1)**-**(A4)***, for all* $\theta\in\Theta$ *and* $k\in\{0,...,n-1\}$*,*

$$\delta\Big(\partial_\theta Y_{t_{k+1}}^\theta(t_k,x)U^\theta(t_k,x)\Big) = \Delta_n\partial_\theta b(\theta,Y_{t_k}^\theta)\sigma^{-2}(Y_{t_k}^\theta)\Big(Y_{t_{k+1}}^\theta - Y_{t_k}^\theta - b(\theta,Y_{t_k}^\theta)\Delta_n\Big)$$
$$+ R_1^{\theta,k} + R_2^{\theta,k} + R_3^{\theta,k} - R_4^{\theta,k} - R_5^{\theta,k} - R_6^{\theta,k},$$

where

$$R_1^{\theta,k} := -\int_{t_k}^{t_{k+1}} D_s\left(\frac{\partial_\theta Y_{t_{k+1}}^\theta(t_k,x)}{\partial_x Y_{t_{k+1}}^\theta(t_k,x)}\right)\frac{\partial_x Y_s^\theta(t_k,x)}{\sigma(Y_s^\theta(t_k,x))}ds,$$
$$R_2^{\theta,k} := \int_{t_k}^{t_{k+1}}\frac{\partial_\theta b(\theta,Y_s^\theta(t_k,x))}{\partial_x Y_s^\theta(t_k,x)}ds\int_{t_k}^{t_{k+1}}\left(\frac{\partial_x Y_s^\theta(t_k,x)}{\sigma(Y_s^\theta(t_k,x))} - \frac{\partial_x Y_{t_k}^\theta(t_k,x)}{\sigma(Y_{t_k}^\theta(t_k,x))}\right)dW_s,$$
$$R_3^{\theta,k} := \int_{t_k}^{t_{k+1}}\left(\frac{\partial_\theta b(\theta,Y_s^\theta(t_k,x))}{\partial_x Y_s^\theta(t_k,x)} - \frac{\partial_\theta b(\theta,Y_{t_k}^\theta(t_k,x))}{\partial_x Y_{t_k}^\theta(t_k,x)}\right)ds\int_{t_k}^{t_{k+1}}\frac{\partial_x Y_{t_k}^\theta(t_k,x)}{\sigma(Y_{t_k}^\theta(t_k,x))}dW_s,$$
$$R_4^{\theta,k} := \Delta_n\partial_\theta b(\theta,Y_{t_k}^\theta)\sigma^{-2}(Y_{t_k}^\theta)\int_{t_k}^{t_{k+1}}\Big(b(\theta,Y_s^\theta) - b(\theta,Y_{t_k}^\theta)\Big)ds,$$
$$R_5^{\theta,k} := \Delta_n\partial_\theta b(\theta,Y_{t_k}^\theta)\sigma^{-2}(Y_{t_k}^\theta)\int_{t_k}^{t_{k+1}}\Big(\sigma(Y_s^\theta) - \sigma(Y_{t_k}^\theta)\Big)dW_s,$$
$$R_6^{\theta,k} := \Delta_n\partial_\theta b(\theta,Y_{t_k}^\theta)\sigma^{-2}(Y_{t_k}^\theta)\int_{t_k}^{t_{k+1}}\int_{\mathbb{R}_0} c(Y_{s-}^\theta,z)\widetilde{M}(ds,dz),$$

and

$$D_s\left(\frac{\partial_\theta Y_{t_{k+1}}^\theta(t_k,x)}{\partial_x Y_{t_{k+1}}^\theta(t_k,x)}\right) = \int_s^{t_{k+1}}\Bigg(-\frac{\partial_\theta b(\theta,Y_u^\theta(t_k,x))}{(\partial_x Y_u^\theta(t_k,x))^2}D_s(\partial_x Y_u^\theta(t_k,x))$$
$$+ \partial_{x,\theta}^2 b(\theta,Y_u^\theta(t_k,x))\frac{D_s Y_u^\theta(t_k,x)}{\partial_x Y_u^\theta(t_k,x)}\Bigg)du.$$

Proof. By Itô's formula,

$$\frac{1}{\partial_x Y_t^\theta(t_k,x)} = 1 - \int_{t_k}^t\frac{\partial_x b(\theta,Y_s^\theta(t_k,x)) - \big(\partial_x\sigma(Y_s^\theta(t_k,x))\big)^2}{\partial_x Y_s^\theta(t_k,x)}ds - \int_{t_k}^t\frac{\partial_x\sigma(Y_s^\theta(t_k,x))}{\partial_x Y_s^\theta(t_k,x)}dW_s$$
$$+ \int_{t_k}^t\int_{\mathbb{R}_0}\frac{\big(\partial_x c(Y_{s-}^\theta(t_k,x),z)\big)^2}{\big(1+\partial_x c(Y_{s-}^\theta(t_k,x),z)\big)\partial_x Y_s^\theta(t_k,x)}\nu(dz)ds$$
$$- \int_{t_k}^t\int_{\mathbb{R}_0}\frac{\partial_x c(Y_{s-}^\theta(t_k,x),z)}{\big(1+\partial_x c(Y_{s-}^\theta(t_k,x),z)\big)\partial_x Y_s^\theta(t_k,x)}\widetilde{M}(ds,dz),$$

which implies that

$$\frac{\partial_\theta Y^\theta_{t_{k+1}}(t_k,x)}{\partial_x Y^\theta_{t_{k+1}}(t_k,x)} = \int_{t_k}^{t_{k+1}} \frac{\partial_\theta b(\theta, Y^\theta_s(t_k,x))}{\partial_x Y^\theta_s(t_k,x)} ds.$$

Then, using the product rule [57, (1.48)], we obtain that

$$\begin{aligned}\delta\left(\partial_\theta Y^\theta_{t_{k+1}}(t_k,x) U^\theta(t_k,x)\right) &= \int_{t_k}^{t_{k+1}} \frac{\partial_\theta b(\theta, Y^\theta_s(t_k,x))}{\partial_x Y^\theta_s(t_k,x)} ds \int_{t_k}^{t_{k+1}} \frac{\partial_x Y^\theta_s(t_k,x)}{\sigma(Y^\theta_s(t_k,x))} dW_s \\ &\quad - \int_{t_k}^{t_{k+1}} D_s\left(\frac{\partial_\theta Y^\theta_{t_{k+1}}(t_k,x)}{\partial_x Y^\theta_{t_{k+1}}(t_k,x)}\right) \frac{\partial_x Y^\theta_s(t_k,x)}{\sigma(Y^\theta_s(t_k,x))} ds.\end{aligned}$$

We next add and substract the term $\frac{\partial_x Y^\theta_{t_k}(t_k,x)}{\sigma(Y^\theta_{t_k}(t_k,x))}$ in the second integral above, and next we add and substract the term $\frac{\partial_\theta b(\theta, Y^\theta_{t_k}(t_k,x))}{\partial_x Y^\theta_{t_k}(t_k,x)}$ in the first one. This, together the fact that $Y^\theta_{t_k}(t_k,x) = Y^\theta_{t_k} = x$, yields

$$\delta\left(\partial_\theta Y^\theta_{t_{k+1}}(t_k,x) U^\theta(t_k,x)\right) = \Delta_n \partial_\theta b(\theta, Y^\theta_{t_k}) \sigma^{-1}(Y^\theta_{t_k}) \left(W_{t_{k+1}} - W_{t_k}\right) + R_1^{\theta,k} + R_2^{\theta,k} + R_3^{\theta,k}. \quad (4.4)$$

On the other hand, by equation (4.3) we have that

$$\begin{aligned}W_{t_{k+1}} - W_{t_k} &= \sigma^{-1}(Y^\theta_{t_k}) \Bigg(Y^\theta_{t_{k+1}} - Y^\theta_{t_k} - b(\theta, Y^\theta_{t_k})\Delta_n - \int_{t_k}^{t_{k+1}} \left(b(\theta, Y^\theta_s) - b(\theta, Y^\theta_{t_k})\right) ds \\ &\quad - \int_{t_k}^{t_{k+1}} \left(\sigma(Y^\theta_s) - \sigma(Y^\theta_{t_k})\right) dW_s - \int_{t_k}^{t_{k+1}} \int_{\mathbb{R}_0} c(Y^\theta_{s-}, z) \widetilde{M}(ds, dz) \Bigg),\end{aligned}$$

which concludes the desired result. □

We will use the following estimates for the solution to (4.2).

Lemma 4.2.2. *Assume conditions* **(A1)** *and* **(A5)**.

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

$$\mathrm{E}\left[\left|Y^\theta_t(t_k,x) - Y^\theta_{t_k}(t_k,x)\right|^p \big| Y^\theta_{t_k}(t_k,x) = x\right] \leq C_p \left|t - t_k\right|^{\frac{p}{2}\wedge 1} \left(1 + |x|^p\right).$$

(ii) *For any function* $g : \Theta \times \mathbb{R} \to \mathbb{R}$ *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* $t \in [t_k, t_{k+1}]$,

$$\mathrm{E}\left[\left|g(\theta, Y^\theta_t(t_k,x))\right| \big| Y^\theta_{t_k}(t_k,x) = x\right] \leq C\left(1 + |x|^q\right).$$

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

Under conditions **(A1)**, **(A2)** and **(A4)**, for any $k \in \{0, ..., n-1\}$ and $t \geq t_k$, the random variables $Y^\theta_t(t_k,x)$, $\partial_x Y^\theta_t(t_k,x)$, $(\partial_x Y^\theta_t(t_k,x))^{-1}$ and $\partial_\theta Y^\theta_t(t_k,x)$ belong to $\mathbb{D}^{1,2}$ (see [61, Theorem 3]).

Assuming conditions **(A1)-(A5)** and using Gronwall's inequality, one can easily check 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 $t \in [t_k, t_{k+1}]$,

$$\begin{aligned}&\mathrm{E}\left[\left|\partial_x Y^\theta_t(t_k,x))\right|^p + \frac{1}{\left|\partial_x Y^\theta_t(t_k,x))\right|^p} \Big| Y^\theta_{t_k}(t_k,x) = x\right] \\ &\quad + \sup_{s\in[t_k,t_{k+1}]} \mathrm{E}\left[\left|D_s Y^\theta_t(t_k,x)\right|^p \big| Y^\theta_{t_k}(t_k,x) = x\right] \leq C_p, \quad \text{and} \\ &\sup_{s\in[t_k,t_{k+1}]} \mathrm{E}\left[\left|D_s\left(\partial_x Y^\theta_t(t_k,x)\right)\right|^p \big| Y^\theta_{t_k}(t_k,x) = x\right] \leq C_p\left(1 + |x|^q\right).\end{aligned}$$

As a consequence, we have the following estimates, which follow easily from (4.4), Lemma 4.2.2 and properties of the moments of the Brownian motion.

Lemma 4.2.3. *Under conditions* **(A1)**-**(A5)**, *for any* $\theta \in \Theta$ *and* $p \geq 2$, *there exist constants* $C_p, q > 0$ *such that for all* $k \in \{0, ..., n-1\}$,

$$\mathrm{E}\left[R_1^{\theta,k} + R_2^{\theta,k} + R_3^{\theta,k} - R_5^{\theta,k} \big| Y_{t_k}^{\theta}(t_k, x) = x\right] = 0, \tag{4.5}$$

$$\mathrm{E}\left[\left|R_1^{\theta,k} + R_2^{\theta,k} + R_3^{\theta,k} - R_5^{\theta,k}\right|^p \big| Y_{t_k}^{\theta}(t_k, x) = x\right] \leq C_p \Delta_n^{\frac{3p+1}{2}} (1 + |x|^q), \tag{4.6}$$

$$\mathrm{E}\left[\left|\delta\left(\partial_\theta Y_{t_{k+1}}^{\theta}(t_k, x) U^{\theta}(t_k, x)\right)\right|^p \big| Y_{t_k}^{\theta}(t_k, x) = x\right] \leq C_p \Delta_n^{\frac{3p}{2}} (1 + |x|^q). \tag{4.7}$$

We next recall Girsanov's theorem on each interval $[t_k, t_{k+1}]$.

Lemma 4.2.4. *Under conditions* **(A1)** *and* **(A2)**, *for all* $\theta, \theta_1 \in \Theta$, *and* $k \in \{0, ..., n-1\}$, *define a measure*

$$\widehat{Q}_k^{\theta_1,\theta} = \widehat{\mathrm{E}}\left[\mathbf{1}_A e^{-\int_{t_k}^{t_{k+1}} \frac{b(\theta,X_t)-b(\theta_1,X_t)}{\sigma(X_t)} dB_t + \frac{1}{2}\int_{t_k}^{t_{k+1}} \left(\frac{b(\theta,X_t)-b(\theta_1,X_t)}{\sigma(X_t)}\right)^2 dt}\right],$$

for all $A \in \widehat{\mathcal{F}}$. *Then* $\widehat{Q}_k^{\theta_1,\theta}$ *is a probability measure and under* $\widehat{Q}_k^{\theta_1,\theta}$, *the process* $B_t^{\widehat{Q}_k^{\theta_1,\theta}} = B_t + \int_{t_k}^{t_{k+1}} \frac{b(\theta,X_t)-b(\theta_1,X_t)}{\sigma(X_t)} dt$ *is a Brownian motion, for all* $t \in [t_k, t_{k+1}]$.

Lemma 4.2.5. *Assume conditions* **(A1)**, **(A2)**, *and* **(A4)**(b). *Let* $\theta, \theta_1 \in \Theta$ *such that* $|\theta - \theta_1| \leq \frac{C}{\sqrt{n\Delta_n}}$, *for some constant* $C > 0$. *Then there exist constants* $C, q > 0$ *such that for any random variable* V, *and* $k \in \{0, ..., n-1\}$,

$$\left|\mathrm{E}_{\widehat{Q}_k^{\theta_1,\theta}}\left[V\left(\frac{d\widehat{\mathrm{P}}}{d\widehat{Q}_k^{\theta_1,\theta}} - 1\right)\Big|X_{t_k}^{\theta}\right]\right| \leq \frac{C}{\sqrt{n}}\left(1 + |X_{t_k}^{\theta}|^q\right)\int_0^1 \left(\mathrm{E}_{\widehat{\mathrm{P}}^{\alpha}}\left[V^2\Big|X_{t_k}^{\theta}\right]\right)^{1/2} d\alpha,$$

where $\mathrm{E}_{\widehat{\mathrm{P}}^{\alpha}}$ *denotes the expectation under the probability measure* $\widehat{\mathrm{P}}^{\alpha}$ *defined as*

$$\frac{d\widehat{\mathrm{P}}^{\alpha}}{d\widehat{Q}_k^{\theta_1,\theta}} := e^{\alpha\int_{t_k}^{t_{k+1}} \frac{b(\theta,X_t)-b(\theta_1,X_t)}{\sigma(X_t)} dB_t - \frac{\alpha^2}{2}\int_{t_k}^{t_{k+1}} \left(\frac{b(\theta,X_t)-b(\theta_1,X_t)}{\sigma(X_t)}\right)^2 dt},$$

for all $\alpha \in [0, 1]$.

Proof. Observe that

$$\frac{d\widehat{\mathrm{P}}}{d\widehat{Q}_k^{\theta_1,\theta}} - 1 = \int_0^1 \int_{t_k}^{t_{k+1}} \frac{b(\theta, X_t) - b(\theta_1, X_t)}{\sigma(X_t)}\left(dB_t - \alpha\frac{b(\theta, X_t) - b(\theta_1, X_t)}{\sigma(X_t)} dt\right)\frac{d\widehat{\mathrm{P}}^{\alpha}}{d\widehat{Q}_k^{\theta_1,\theta}} d\alpha.$$

Consider the process $W = (W_t)_{t\in[t_k,t_{k+1}]}$ defined by

$$W_t := B_t - \alpha\int_{t_k}^{t} \frac{b(\theta, X_s) - b(\theta_1, X_s)}{\sigma(X_s)} ds.$$

By Girsanov's theorem, W is a Brownian motion under $\widehat{\mathrm{P}}^{\alpha}$.

Then, using Girsanov's theorem, Cauchy-Schwarz inequality, and hypotheses **(A2)**, **(A4)**(b), together with Lemma 4.2.2 (ii), we get that

$$\left|\mathrm{E}_{\widehat{Q}_k^{\theta_1,\theta}}\left[V\left(\frac{d\widehat{\mathrm{P}}}{d\widehat{Q}_k^{\theta_1,\theta}} - 1\right)\Big|X_{t_k}^{\theta}\right]\right| = \left|\int_0^1 \mathrm{E}_{\widehat{\mathrm{P}}^{\alpha}}\left[V\int_{t_k}^{t_{k+1}} \frac{b(\theta, X_t) - b(\theta_0, X_t)}{\sigma(X_t)} dW_t\Big|X_{t_k}^{\theta}\right] d\alpha\right|$$

$$\leq \int_0^1 \left(\mathrm{E}_{\widehat{\mathrm{P}}^{\alpha}}\left[V^2\Big|X_{t_k}^{\theta}\right]\right)^{1/2}\left(\mathrm{E}_{\widehat{\mathrm{P}}^{\alpha}}\left[\left|\int_{t_k}^{t_{k+1}} \frac{b(\theta, X_t) - b(\theta_0, X_t)}{\sigma(X_t)} dW_t\right|^2 \Big|X_{t_k}^{\theta}\right]\right)^{1/2} d\alpha$$

$$\leq \frac{C}{\sqrt{n}}\left(1 + |X_{t_k}^{\theta}|^q\right)\int_0^1 \left(\mathrm{E}_{\widehat{\mathrm{P}}^{\alpha}}\left[V^2\Big|X_{t_k}^{\theta}\right]\right)^{1/2} d\alpha,$$

for some constants $C, q > 0$. Thus, the result follows. □

For any $t > s$ and $i \geq 0$, we denote by $q^{\theta}_{(i)}(t-s, x, y)$ the transition density of Y^{θ}_t conditioned on $Y^{\theta}_s = x$ and $M_t - M_s = i$, where $M_t = M([0,t] \times \mathbb{R})$. That is,

$$p^{\theta}(t-s, x, y) = \sum_{i=0}^{\infty} q^{\theta}_{(i)}(t-s, x, y) e^{-\lambda(t-s)} \frac{(\lambda(t-s))^i}{i!}. \tag{4.8}$$

From [25, Proposition 1.2], for any $\theta \in \Theta$ there exist constants $c, C > 1$ such that for all $0 < t \leq 1$, and $x, y \in \mathbb{R}$,

$$\frac{1}{C\sqrt{t}} e^{-c\frac{(y-x)^2}{t}} e^{-ctx^2} \leq q^{\theta}_{(0)}(t, x, y) \leq \frac{C}{\sqrt{t}} e^{-\frac{(y-x)^2}{ct}} e^{ctx^2}. \tag{4.9}$$

For any $t > s$ and $i \geq 1$, we denote by $q^{\theta}_{(i)}(t-s, x, y; a_1, \ldots, a_i)$ the transition density of Y^{θ}_t conditioned on $Y^{\theta}_s = x, M_t - M_s = i$ and $\widetilde{\Lambda}_{[s,t]} = \{a_1, \ldots, a_i\}$, where $\widetilde{\Lambda}_{[s,t]}$ are the jump amplitudes of $\widetilde{Z}$ on the interval $[s,t]$, i.e, $\widetilde{\Lambda}_{[s,t]} := \{\Delta \widetilde{Z}_u; s \leq u \leq t\}$.

Lemma 4.2.6. *Under conditions* **(A1)**-**(A4)**, *for all $\theta \in \Theta$ and n large enough, there exist constants $C_1, C_2, C_3 > 0$ such that for all $a, x, y \in \mathbb{R}$,*

$$q^{\theta}_{(1)}(\Delta_n, x, y; a) \leq C_1 e^{C_2 \Delta_n (x^2 + (1+x^2)a^2)} \frac{1}{\sqrt{\Delta_n}} e^{-\frac{(y-x-c(x,a))^2}{C_3 \Delta_n}}.$$

Proof. Using the Chapman-Kolmogorov equation in terms of transition density and the fact that the distribution of the jump time conditioned on $M_{t_{k+1}} - M_{t_k} = 1$ is a uniform distribution on $[t_k, t_{k+1}]$, together with (4.9), we get that

$$\begin{aligned} q^{\theta}_{(1)}(\Delta_n, x, y; a) &= \frac{1}{\Delta_n} \int_{t_k}^{t_{k+1}} \int_{\mathbb{R}} q^{\theta}_{(0)}(t-t_k, x, z) q^{\theta}_{(0)}(t_{k+1}-t, z+c(z,a), y) dz dt \\ &\leq \frac{C}{\Delta_n} \int_{t_k}^{t_{k+1}} \int_{\mathbb{R}} \frac{1}{\sqrt{t-t_k}} e^{-\frac{(z-x)^2}{c(t-t_k)}} e^{c(t-t_k)x^2} \frac{1}{\sqrt{t_{k+1}-t}} e^{-\frac{(y-z-c(z,a))^2}{c(t_{k+1}-t)}} e^{c(t_{k+1}-t)(z+c(z,a))^2} dz dt, \end{aligned}$$

for some constants $c, C > 1$.

We next use the change of variables $u := \varphi(z) := z + c(z,a) - x - c(x,a)$. Observe that $\varphi(x) = 0$. Moreover, from hypotheses **(A4)**(a) and (e),

$$\eta \leq |\varphi'(z)| = |1 + \partial_z c(z,a)| \leq \beta,$$

for some constants $\beta, \eta > 0$. Therefore, the mapping $z \to \varphi(z)$ admits an inverse function φ^{-1}. Thus, for any $u \in \mathbb{R}$, there exists $\xi \in (0,u)$ or $\xi \in (u,0)$ such that

$$\frac{|u|}{\beta} \leq |\varphi^{-1}(u) - \varphi^{-1}(0)| = \frac{|u|}{|\varphi'(\varphi^{-1}(\xi))|} \leq \frac{|u|}{\eta},$$

which yields

$$\begin{aligned} q^{\theta}_{(1)}(\Delta_n, x, y; a) &\leq \frac{C}{\Delta_n} \int_{t_k}^{t_{k+1}} \int_{\mathbb{R}} \frac{1}{\sqrt{t-t_k}} e^{-\frac{(\varphi^{-1}(u)-\varphi^{-1}(0))^2}{c(t-t_k)}} \frac{1}{\sqrt{t_{k+1}-t}} e^{-\frac{(y-u-x-c(x,a))^2}{c(t_{k+1}-t)}} \\ &\qquad \times e^{c(t-t_k)x^2} e^{c(t_{k+1}-t)(u+x+c(x,a))^2} du dt \\ &\leq \frac{C}{\Delta_n} \int_{t_k}^{t_{k+1}} \int_{\mathbb{R}} \frac{1}{\sqrt{t-t_k}} e^{-\frac{u^2}{c\beta^2(t-t_k)}} \frac{1}{\sqrt{t_{k+1}-t}} e^{-\frac{(y-u-x-c(x,a))^2}{c(t_{k+1}-t)}} \\ &\qquad \times e^{2c(t_{k+1}-t)u^2} e^{c\Delta_n x^2} e^{2c\Delta_n(x+c(x,a))^2} du dt. \end{aligned}$$

We next use the fact that

$$e^{-\frac{u^2}{c\beta^2(t-t_k)} + 2c(t_{k+1}-t)u^2} = e^{-\frac{1-2c^2\beta^2(t-t_k)(t_{k+1}-t)}{c\beta^2(t-t_k)}u^2} \leq e^{-\frac{(1-c^2\beta^2\Delta_n^2)u^2}{c\beta^2(t-t_k)}},$$

to get that for n large enough, there exists $C>0$ such that

$$q^{\theta}_{(1)}(\Delta_n,x,y;a)\leq e^{5c\Delta_n x^2}e^{4c\Delta_n c^2(x,a)}\frac{C}{\Delta_n}\int_{t_k}^{t_{k+1}}\int_{\mathbb{R}}\frac{1}{\sqrt{\frac{\beta^2}{1-c^2\beta^2\Delta_n^2}(t-t_k)}}e^{-\frac{(1-c^2\beta^2\Delta_n^2)u^2}{c\beta^2(t-t_k)}}$$
$$\times\frac{1}{\sqrt{t_{k+1}-t}}e^{-\frac{(y-u-x-c(x,a))^2}{c(t_{k+1}-t)}}dudt$$
$$=e^{5c\Delta_n x^2}e^{4c\Delta_n c^2(x,a)}\frac{C}{\Delta_n}\int_{t_k}^{t_{k+1}}\frac{1}{\sqrt{\frac{\beta^2}{1-c^2\beta^2\Delta_n^2}(t-t_k)+t_{k+1}-t}}e^{-\frac{(y-x-c(x,a))^2}{c\left(\frac{\beta^2}{1-c^2\beta^2\Delta_n^2}(t-t_k)+t_{k+1}-t\right)}}dt.$$

Next, observe that for n large enough there exist constants $\beta_1,\beta_2>0$ such that

$$\beta_1\Delta_n\leq\frac{\beta^2}{1-\frac{1}{2}c^2\beta^2\Delta_n^2}(t-t_k)+t_{k+1}-t\leq\beta_2\Delta_n,$$

from where we deduce that

$$q^{\theta}_{(1)}(\Delta_n,x,y;a)\leq Ce^{5c\Delta_n x^2}e^{4c\Delta_n c^2(x,a)}\frac{1}{\sqrt{\Delta_n}}e^{-\frac{(y-x-c(x,a))^2}{c\Delta_n}}.$$

Finally, hypothesis **(A1)** implies the desired result. □

Consider the events $\widehat{J}_{i,k}=\{N_{t_{k+1}}-N_{t_k}=i\}$ and $\widetilde{J}_{i,k}=\{M_{t_{k+1}}-M_{t_k}=i\}$, for $i=0,1$ and $k\in\{0,...,n-1\}$, where $N_t=N([0,t]\times\mathbb{R})$. We denote by $\widehat{\Lambda}_{[s,t]}$ the jump amplitudes of $\widehat{Z}$ on the interval $[s,t]$, i.e, $\widehat{\Lambda}_{[s,t]}:=\{\Delta\widehat{Z}_u;s\leq u\leq t\}$, and by $\mu(dz)=\frac{\nu(dz)}{\lambda}$ the jump size distribution of $\widehat{Z}$. As in Lemma 3.2.4, we have the following expressions for the conditional expectations in terms of the transition densities.

Lemma 4.2.7. *Under conditions* **(A1)**-**(A4)**, *for all* $k\in\{0,...,n-1\}$ *and* $\theta\in\Theta$,

$$\mathrm{E}_{\widehat{Q}_k^{\theta,\theta_0}}\left[\mathbf{1}_{\widehat{J}_{0,k}}\left(\widetilde{\mathrm{E}}^{\theta}_{X_{t_k}}\left[\mathbf{1}_{\widetilde{J}_{1,k}}\int_{t_k}^{t_{k+1}}\int_{\mathbb{R}_0}c(Y^{\theta}_{t_k},z)M(ds,dz)\middle|Y^{\theta}_{t_{k+1}}=X_{t_{k+1}}\right]\right)^2\middle|X_{t_k}\right]$$
$$=\int_{\mathbb{R}}\left(\frac{\int_{\mathbb{R}_0}q^{\theta}_{(1)}(\Delta_n,X_{t_k},y;a)c\left(X_{t_k},a\right)\mu(da)e^{-\lambda\Delta_n}\lambda\Delta_n}{p^{\theta}(\Delta_n,X_{t_k},y)}\right)^2 q^{\theta}_{(0)}(\Delta_n,X_{t_k},y)e^{-\lambda\Delta_n}dy,\tag{4.10}$$

$$\mathrm{E}_{\widehat{Q}_k^{\theta,\theta_0}}\left[\mathbf{1}_{\widehat{J}_{1,k}}\left(\int_{t_k}^{t_{k+1}}\int_{\mathbb{R}_0}c(X_{t_k},z)N(ds,dz)\widetilde{\mathrm{E}}^{\theta}_{X_{t_k}}\left[\mathbf{1}_{\widetilde{J}_{0,k}}\middle|Y^{\theta}_{t_{k+1}}=X_{t_{k+1}}\right]\right)^2\middle|X_{t_k}\right]$$
$$=\int_{\mathbb{R}_0}\int_{\mathbb{R}}\left(\frac{q^{\theta}_{(0)}(\Delta_n,X_{t_k},y)e^{-\lambda\Delta_n}}{p^{\theta}(\Delta_n,X_{t_k},y)}\right)^2 q^{\theta}_{(1)}(\Delta_n,X_{t_k},y;a)e^{-\lambda\Delta_n}\lambda\Delta_n c^2(X_{t_k},a)dy\mu(da),\tag{4.11}$$

and

$$\mathrm{E}_{\widehat{Q}_k^{\theta,\theta_0}}\left[\mathbf{1}_{\widehat{J}_{1,k}}\left(\int_{t_k}^{t_{k+1}}\int_{\mathbb{R}_0}c(X_{t_k},z)N(ds,dz)\widetilde{\mathrm{E}}^{\theta}_{X_{t_k}}\left[\mathbf{1}_{\widetilde{J}_{1,k}}\middle|Y^{\theta}_{t_{k+1}}=X_{t_{k+1}}\right]\right.\right.$$
$$\left.\left.-\widetilde{\mathrm{E}}^{\theta}_{X_{t_k}}\left[\mathbf{1}_{\widetilde{J}_{1,k}}\int_{t_k}^{t_{k+1}}\int_{\mathbb{R}_0}c(Y^{\theta}_{t_k},z)M(ds,dz)\middle|Y^{\theta}_{t_{k+1}}=X_{t_{k+1}}\right]\right)^2\middle|X_{t_k}\right]$$
$$=\int_{\mathbb{R}_0}\int_{\mathbb{R}}\left(\frac{\int_{\mathbb{R}_0}\left(c(X_{t_k},z)-c(X_{t_k},a)\right)q^{\theta}_{(1)}(\Delta_n,X_{t_k},y;a)\mu(da)e^{-\lambda\Delta_n}\lambda\Delta_n}{p^{\theta}(\Delta_n,X_{t_k},y)}\right)^2$$
$$\times q^{\theta}_{(1)}(\Delta_n,X_{t_k},y;z)e^{-\lambda\Delta_n}\lambda\Delta_n dy\mu(dz).\tag{4.12}$$

Proof. Using Bayes' formula, we get that

$$\begin{aligned}
&\widetilde{\mathrm{E}}^{\theta}_{X_{t_k}}\left[\mathbf{1}_{\widetilde{J}_{1,k}}\int_{t_k}^{t_{k+1}}\int_{\mathbb{R}_0} c(Y^{\theta}_{t_k},z)M(ds,dz)\Big|Y^{\theta}_{t_{k+1}}=X_{t_{k+1}}\right]\\
&=\frac{\widetilde{\mathrm{E}}^{\theta}_{X_{t_k}}\left[c(Y^{\theta}_{t_k},\widetilde{\Lambda}_{[t_k,t_{k+1}]})\mathbf{1}_{\{Y^{\theta}_{t_{k+1}}=X_{t_{k+1}}\}}\Big|\widetilde{J}_{1,k}\right]\widetilde{\mathrm{P}}^{\theta}_{X_{t_k}}\left(\widetilde{J}_{1,k}\right)}{p^{\theta}(\Delta_n,X_{t_k},X_{t_{k+1}})}\\
&=\frac{\int_{\mathbb{R}_0} q^{\theta}_{(1)}(\Delta_n,X_{t_k},X_{t_{k+1}};a)c\left(X_{t_k},a\right)\mu(da)e^{-\lambda\Delta_n}\lambda\Delta_n}{p^{\theta}(\Delta_n,X_{t_k},X_{t_{k+1}})}.
\end{aligned}$$

This, together with Bayes' formula again, implies that

$$\begin{aligned}
&\mathrm{E}_{\widehat{Q}^{\theta,\theta_0}_k}\left[\mathbf{1}_{\widehat{J}_{0,k}}\left(\widetilde{\mathrm{E}}^{\theta}_{X_{t_k}}\left[\mathbf{1}_{\widetilde{J}_{1,k}}\int_{t_k}^{t_{k+1}}\int_{\mathbb{R}_0} c(Y^{\theta}_{t_k},z)M(ds,dz)\Big|Y^{\theta}_{t_{k+1}}=X_{t_{k+1}}\right]\right)^2\Big|X_{t_k}\right]\\
&=\widehat{Q}^{\theta,\theta_0}_k\left(\widehat{J}_{0,k}\big|X_{t_k}\right)\mathrm{E}_{\widehat{Q}^{\theta,\theta_0}_k}\left[\left(\frac{\int_{\mathbb{R}_0} q^{\theta}_{(1)}(\Delta_n,X_{t_k},X_{t_{k+1}};a)c\left(X_{t_k},a\right)\mu(da)e^{-\lambda\Delta_n}\lambda\Delta_n}{p^{\theta}(\Delta_n,X_{t_k},X_{t_{k+1}})}\right)^2\Big|\widehat{J}_{0,k},X_{t_k}\right],
\end{aligned}$$

which implies (4.10). Similarly,

$$\begin{aligned}
&\mathrm{E}_{\widehat{Q}^{\theta,\theta_0}_k}\left[\mathbf{1}_{\widehat{J}_{1,k}}\left(\int_{t_k}^{t_{k+1}}\int_{\mathbb{R}_0} c(X_{t_k},z)N(ds,dz)\widetilde{\mathrm{E}}^{\theta}_{X_{t_k}}\left[\mathbf{1}_{\widetilde{J}_{0,k}}\Big|Y^{\theta}_{t_{k+1}}=X_{t_{k+1}}\right]\right)^2\Big|X_{t_k}\right]\\
&=\mathrm{E}_{\widehat{Q}^{\theta,\theta_0}_k}\left[\mathbf{1}_{\widehat{J}_{1,k}}\left(c\left(X_{t_k},\widehat{\Lambda}_{[t_k,t_{k+1}]}\right)\right)^2\left(\frac{q^{\theta}_{(0)}(\Delta_n,X_{t_k},X_{t_{k+1}})e^{-\lambda\Delta_n}}{p^{\theta}(\Delta_n,X_{t_k},X_{t_{k+1}})}\right)^2\Big|X_{t_k}\right]\\
&=\int_{\mathbb{R}_0}\mathrm{E}_{\widehat{Q}^{\theta,\theta_0}_k}\left[\left(\frac{q^{\theta}_{(0)}(\Delta_n,X_{t_k},X_{t_{k+1}})e^{-\lambda\Delta_n}}{p^{\theta}(\Delta_n,X_{t_k},X_{t_{k+1}})}\right)^2\Big|\widehat{J}_{1,k},\widehat{\Lambda}_{[t_k,t_{k+1}]}=\{a\},X_{t_k}\right]c^2(X_{t_k},a)\\
&\qquad\times\widehat{Q}^{\theta,\theta_0}_k\left(\widehat{\Lambda}_{[t_k,t_{k+1}]}\in da,\widehat{J}_{1,k}\Big|X_{t_k}\right)\\
&=\int_{\mathbb{R}_0}\int_{\mathbb{R}}\left(\frac{q^{\theta}_{(0)}(\Delta_n,X_{t_k},y)e^{-\lambda\Delta_n}}{p^{\theta}(\Delta_n,X_{t_k},y)}\right)^2 q^{\theta}_{(1)}(\Delta_n,X_{t_k},y;a)e^{-\lambda\Delta_n}\lambda\Delta_n c^2(X_{t_k},a)dy\mu(da),
\end{aligned}$$

which shows (4.11). The proof of (4.12) follows along the same lines and is therefore omitted. □

By abuse of notation, consider the events $\widehat{J}_{2,k}=\{N_{t_{k+1}}-N_{t_k}\geq 2\}$ and $\widetilde{J}_{2,k}=\{M_{t_{k+1}}-M_{t_k}\geq 2\}$. Set $I=\{\rho_1\Delta_n^{\upsilon}\leq|a|\leq\rho_2\Delta_n^{-\gamma}\}$ and $\lambda_n=\int_I\nu(da)$, where $\rho_1,\rho_2>0$ and $0<\upsilon,\gamma<\frac{1}{2}$ are from hypothesis **(A7)**. For $i=0,1,2$, set

$$\begin{aligned}
M^{\theta}_i=\mathrm{E}_{\widehat{Q}^{\theta,\theta_0}_k}\Bigg[\mathbf{1}_{\widehat{J}_{i,k}}\Bigg(&\int_{t_k}^{t_{k+1}}\int_I c(X_{t_k},z)N(ds,dz)\\
&-\widetilde{\mathrm{E}}^{\theta}_{X_{t_k}}\left[\int_{t_k}^{t_{k+1}}\int_I c(Y^{\theta}_{t_k},z)M(ds,dz)\Big|Y^{\theta}_{t_{k+1}}=X_{t_{k+1}}\right]\Bigg)^2\Big|X_{t_k}\Bigg].
\end{aligned}$$

Recall that for the simple Lévy process (3.1), we used a large deviation principle by conditioning on the number of jumps within the conditional expectation in order to obtain the large deviation estimates (see Lemma 3.2.6). For the non-linear model (4.1), now we will obtain the parallel of Lemma 3.2.6 in our case.

Lemma 4.2.8. *Under conditions* **(A1)-(A5)**, *for any $\theta \in \Theta$ and n large enough, there exist constants $C, C_0, C_1 > 0$, such that for all $\alpha \in (\upsilon, \frac{1}{2})$, $\alpha_0 \in (\frac{1}{4}, \frac{1}{2})$, and $k \in \{0, ..., n-1\}$,*

$$M_0^\theta \leq C e^{C_1 \Delta_n^{1-2\gamma} X_{t_k}^2} \left(1 + |X_{t_k}|^2\right) \left(\lambda_n \Delta_n^{3/2} + \Delta_n^{-2\gamma} e^{-C_0 \Delta_n^{2\alpha - 1}}\right), \tag{4.13}$$

$$M_1^\theta \leq C e^{C_1 \Delta_n^{1-2\gamma} X_{t_k}^2} \left(1 + |X_{t_k}|^3\right) \left(\lambda_n \Delta_n^{3/2} + \Delta_n^{-\frac{1}{2} - 3\gamma} e^{-C_0 \Delta_n^{2(\alpha \vee \alpha_0) - 1}}\right), \tag{4.14}$$

$$M_2^\theta \leq C \lambda_n \Delta_n^{3/2} (1 + |X_{t_k}|^2). \tag{4.15}$$

In particular, (4.15) *holds for all $n \geq 1$.*

Proof. We start showing (4.13). Multiplying the random variable inside the conditional expectation of M_0^θ by $(\mathbf{1}_{\widetilde{J}_{0,k}} + \mathbf{1}_{\widetilde{J}_{1,k}} + \mathbf{1}_{\widetilde{J}_{2,k}})$, we get that $M_0^\theta \leq 2(M_{0,1}^\theta + M_{0,2}^\theta)$, where for $i = 1, 2$,

$$M_{0,i}^\theta = \mathrm{E}_{\widehat{Q}_k^{\theta,\theta_0}} \left[\mathbf{1}_{\widehat{J}_{0,k}} \left(\widetilde{\mathrm{E}}_{X_{t_k}}^\theta \left[\mathbf{1}_{\widetilde{J}_{i,k}} \int_{t_k}^{t_{k+1}} \int_I c(Y_{t_k}^\theta, z) M(ds, dz) \middle| Y_{t_{k+1}}^\theta = X_{t_{k+1}} \right] \right)^2 \middle| X_{t_k} \right].$$

By (4.10), we have that

$$M_{0,1}^\theta = \int_{\mathbb{R}} \left(\frac{\int_I q_{(1)}^\theta(\Delta_n, X_{t_k}, y; a) c\left(X_{t_k}, a\right) \mu(da) e^{-\lambda_n \Delta_n} \lambda_n \Delta_n}{p^\theta(\Delta_n, X_{t_k}, y)} \right)^2 q_{(0)}^\theta(\Delta_n, X_{t_k}, y) e^{-\lambda_n \Delta_n} dy.$$

We next divide the integral in $M_{0,1}^\theta$ into the subdomains $\{y : |y - X_{t_k}| > \Delta_n^\alpha\}$ and $\{y : |y - X_{t_k}| \leq \Delta_n^\alpha\}$, where $\alpha \in (\upsilon, \frac{1}{2})$, and call each integral $M_{0,1,1}^\theta$ and $M_{0,1,2}^\theta$. Therefore, the estimation of $M_{0,1}^\theta$ will be divided into two parts. One will use large deviation for the continuous process in the first part. The other will use the fact that the jump parts are significantly bigger than the continuous parts. This fact will be obtained under condition **(A3)**.

We start bounding $M_{0,1,1}^\theta$. By (4.8),

$$p^\theta(\Delta_n, X_{t_k}, y) \geq \int_I q_{(1)}^\theta(\Delta_n, X_{t_k}, y; a) \mu(da) e^{-\lambda_n \Delta_n} \lambda_n \Delta_n. \tag{4.16}$$

Then, using the fact that by **(A1)**, on I, $|c(X_{t_k}, a)| \leq C \Delta_n^{-\gamma}(1 + |X_{t_k}|)$ for some constant $C > 0$, and (4.9), we get that

$$\begin{aligned} M_{0,1,1}^\theta &\leq C \Delta_n^{-2\gamma} \left(1 + |X_{t_k}|^2\right) \int_{\{|y - X_{t_k}| > \Delta_n^\alpha\}} q_{(0)}^\theta(\Delta_n, X_{t_k}, y) dy \\ &\leq C \Delta_n^{-2\gamma} \left(1 + |X_{t_k}|^2\right) \int_{\{|y - X_{t_k}| > \Delta_n^\alpha\}} \frac{1}{\sqrt{\Delta_n}} e^{-\frac{(y - X_{t_k})^2}{c \Delta_n}} e^{c \Delta_n X_{t_k}^2} dy \\ &\leq C \Delta_n^{-2\gamma} \left(1 + |X_{t_k}|^2\right) e^{c \Delta_n X_{t_k}^2} e^{-C_1 \Delta_n^{2\alpha - 1}}, \end{aligned}$$

for some constants $C, C_1 > 0$ and $c > 1$. We next treat $M_{0,1,2}^\theta$. Observe that (4.8) yields

$$\left(p^\theta(\Delta_n, X_{t_k}, y)\right)^2 \geq q_{(0)}^\theta(\Delta_n, X_{t_k}, y) e^{-\lambda_n \Delta_n} \int_I q_{(1)}^\theta(\Delta_n, X_{t_k}, y; a) \mu(da) e^{-\lambda_n \Delta_n} \lambda_n \Delta_n. \tag{4.17}$$

Therefore, using hypothesis **(A1)** and Lemma 4.2.6, we get that for n large enough

$$
\begin{aligned}
M_{0,1,2}^{\theta} &\leq C\Delta_n^{-2\gamma}\left(1+|X_{t_k}|^2\right)e^{-\lambda_n\Delta_n}\lambda_n\Delta_n\int_{\left\{|y-X_{t_k}|\leq\Delta_n^{\alpha}\right\}}\int_I q_{(1)}^{\theta}(\Delta_n,X_{t_k},y;a)\mu(da)dy\\
&\leq C\Delta_n^{-2\gamma}\left(1+|X_{t_k}|^2\right)e^{C_2\Delta_n^{1-2\gamma}X_{t_k}^2}\int_I\int_{\left\{|y-X_{t_k}|\leq\Delta_n^{\alpha}\right\}}\frac{1}{\sqrt{\Delta_n}}e^{-\frac{(y-X_{t_k}-c(X_{t_k},a))^2}{C_3\Delta_n}}dy\mu(da)\\
&\leq C\Delta_n^{-2\gamma}\left(1+|X_{t_k}|^2\right)e^{C_2\Delta_n^{1-2\gamma}X_{t_k}^2}\int_I\Bigg\{\int_{-\infty}^{\frac{\Delta_n^{\alpha}-C\rho_1\Delta_n^{\upsilon}}{\sqrt{C_3\Delta_n}}}e^{-w^2}dw\mathbf{1}_{\left\{c(X_{t_k},a)\geq C\rho_1\Delta_n^{\upsilon}\right\}}\\
&\qquad+\int_{\frac{-\Delta_n^{\alpha}+C\rho_1\Delta_n^{\upsilon}}{\sqrt{C_3\Delta_n}}}^{+\infty}e^{-w^2}dw\mathbf{1}_{\left\{c(X_{t_k},a)\leq -C\rho_1\Delta_n^{\upsilon}\right\}}\Bigg\}\mu(da)\\
&\leq C\Delta_n^{-2\gamma}\left(1+|X_{t_k}|^2\right)e^{C_2\Delta_n^{1-2\gamma}X_{t_k}^2}e^{-C_0\Delta_n^{2\upsilon-1}},
\end{aligned}
$$

for some constants $C, C_0, C_2, C_3 > 0$, where we have applied Fubini's theorem, the change of variables $w=\frac{y-X_{t_k}-c(X_{t_k},a)}{\sqrt{C_3\Delta_n}}$, and the fact that by **(A3)**, on I, $|c(X_{t_k},a)|\geq C|a|\geq C\rho_1\Delta_n^{\upsilon}$ for some constant $C>0$, together with $e^{-\lambda_n\Delta_n}\lambda_n\Delta_n\leq\lambda$. This shows that for n large enough and $\alpha\in(\upsilon,\frac{1}{2})$,

$$
M_{0,1}^{\theta}\leq C\Delta_n^{-2\gamma}\left(1+|X_{t_k}|^2\right)e^{C_1\Delta_n^{1-2\gamma}X_{t_k}^2}e^{-C_0\Delta_n^{2\alpha-1}}, \tag{4.18}
$$

for some constants $C, C_0, C_1 > 0$.

In order to treat $M_{0,2}^{\theta}$, observe that by Jensen and Cauchy-Schwarz inequalities, and hypotheses **(A1)** and **(A5)**, it holds that

$$
M_{0,2}^{\theta}\leq \mathrm{E}\left[\mathbf{1}_{\widetilde{J}_{2,k}}\left(\int_{t_k}^{t_{k+1}}\int_I c(Y_{t_k}^{\theta},z)M(ds,dz)\right)^2\Big|Y_{t_k}^{\theta}=X_{t_k}\right]\leq C\lambda_n\Delta_n^{3/2}\left(1+|X_{t_k}|^2\right).
$$

This shows (4.13).

We next show (4.14). As for the term M_0^{θ}, we have that $M_1^{\theta}\leq 2(M_{1,1}^{\theta}+M_{1,2}^{\theta})$, where

$$
\begin{aligned}
M_{1,1}^{\theta} &= \mathrm{E}_{\widehat{Q}_k^{\theta,\theta_0}}\Bigg[\mathbf{1}_{\widehat{J}_{1,k}}\Bigg(\int_{t_k}^{t_{k+1}}\int_I c(X_{t_k},z)N(ds,dz)\\
&\qquad-\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta}\left[\mathbf{1}_{\widetilde{J}_{1,k}}\int_{t_k}^{t_{k+1}}\int_I c(Y_{t_k}^{\theta},z)M(ds,dz)\Big|Y_{t_{k+1}}^{\theta}=X_{t_{k+1}}\right]\Bigg)^2\Big|X_{t_k}\Bigg],\\
M_{1,2}^{\theta} &= \mathrm{E}_{\widehat{Q}_k^{\theta,\theta_0}}\left[\mathbf{1}_{\widehat{J}_{1,k}}\left(\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta}\left[\mathbf{1}_{\widetilde{J}_{2,k}}\int_{t_k}^{t_{k+1}}\int_I c(Y_{t_k}^{\theta},z)M(ds,dz)\Big|Y_{t_{k+1}}^{\theta}=X_{t_{k+1}}\right]\right)^2\Big|X_{t_k}\right].
\end{aligned}
$$

We start bounding $M_{1,1}^{\theta}$. Adding and substracting the term

$$
\int_{t_k}^{t_{k+1}}\int_I c(X_{t_k},z)N(ds,dz)\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta}\left[\mathbf{1}_{\widetilde{J}_{1,k}}|Y_{t_{k+1}}^{\theta}=X_{t_{k+1}}\right]
$$

inside the square, we get that $M_{1,1}^{\theta}\leq 2(M_{1,1,1}^{\theta}+M_{1,1,2}^{\theta})$, where

$$
\begin{aligned}
M_{1,1,1}^{\theta} &= \mathrm{E}_{\widehat{Q}_k^{\theta,\theta_0}}\Bigg[\mathbf{1}_{\widehat{J}_{1,k}}\Bigg(\int_{t_k}^{t_{k+1}}\int_I c(X_{t_k},z)N(ds,dz)\\
&\qquad-\int_{t_k}^{t_{k+1}}\int_I c(X_{t_k},z)N(ds,dz)\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta}\left[\mathbf{1}_{\widetilde{J}_{1,k}}\Big|Y_{t_{k+1}}^{\theta}=X_{t_{k+1}}\right]\Bigg)^2\Big|X_{t_k}\Bigg],\\
M_{1,1,2}^{\theta} &= \mathrm{E}_{\widehat{Q}_k^{\theta,\theta_0}}\Bigg[\mathbf{1}_{\widehat{J}_{1,k}}\Bigg(\int_{t_k}^{t_{k+1}}\int_I c(X_{t_k},z)N(ds,dz)\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta}\left[\mathbf{1}_{\widetilde{J}_{1,k}}\Big|Y_{t_{k+1}}^{\theta}=X_{t_{k+1}}\right]\\
&\qquad-\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta}\left[\mathbf{1}_{\widetilde{J}_{1,k}}\int_{t_k}^{t_{k+1}}\int_I c(Y_{t_k}^{\theta},z)M(ds,dz)\Big|Y_{t_{k+1}}^{\theta}=X_{t_{k+1}}\right]\Bigg)^2\Big|X_{t_k}\Bigg].
\end{aligned}
$$

Observe that $M^{\theta}_{1,1,1} \leq 2(M^{\theta}_{1,1,1,0} + M^{\theta}_{1,1,1,2})$, where for $i = 0, 2$,

$$M^{\theta}_{1,1,1,i} = \mathrm{E}_{\widehat{Q}^{\theta,\theta_0}_k}\left[\mathbf{1}_{\widehat{J}_{1,k}}\left(\int_{t_k}^{t_{k+1}}\int_I c(X_{t_k},z)N(ds,dz)\widetilde{\mathrm{E}}^{\theta}_{X_{t_k}}\left[\mathbf{1}_{\widetilde{J}_{i,k}}\middle|Y^{\theta}_{t_{k+1}} = X_{t_{k+1}}\right]\right)^2\middle|X_{t_k}\right].$$

By (4.11),

$$M^{\theta}_{1,1,1,0} = \int_I\int_{\mathbb{R}}\left(\frac{q^{\theta}_{(0)}(\Delta_n,X_{t_k},y)e^{-\lambda_n\Delta_n}}{p^{\theta}(\Delta_n,X_{t_k},y)}\right)^2 q^{\theta}_{(1)}(\Delta_n,X_{t_k},y;a)e^{-\lambda_n\Delta_n}\lambda_n\Delta_n c^2(X_{t_k},a)dy\mu(da).$$

Again we divide the dy integral into the subdomains $\{y : |y-X_{t_k}| > \Delta_n^{\alpha}\}$ and $\{y : |y-X_{t_k}| \leq \Delta_n^{\alpha}\}$, where $\alpha \in (\upsilon, \frac{1}{2})$, and call the terms $M^{\theta}_{1,1,1,0,1}$ and $M^{\theta}_{1,1,1,0,2}$. In the same way the term $M^{\theta}_{0,1,1}$ was treated, we use (4.17) and hypothesis **(A1)**, we obtain that

$$\begin{aligned} M^{\theta}_{1,1,1,0,1} &\leq C\Delta_n^{-2\gamma}\left(1+|X_{t_k}|^2\right)\int_{\{|y-X_{t_k}|>\Delta_n^{\alpha}\}} q^{\theta}_{(0)}(\Delta_n,X_{t_k},y)dy \\ &\leq C\Delta_n^{-2\gamma}\left(1+|X_{t_k}|^2\right)e^{c\Delta_n X^2_{t_k}}e^{-C_1\Delta_n^{2\alpha-1}}, \end{aligned}$$

for some constants $C, C_1 > 0$ and $c > 1$. Next, (4.8) yields

$$p^{\theta}(\Delta_n,X_{t_k},y) \geq q^{\theta}_{(0)}(\Delta_n,X_{t_k},y)e^{-\lambda_n\Delta_n}. \tag{4.19}$$

Then, as for the term $M^{\theta}_{0,1,2}$, using hypothesis **(A1)** and Lemma 4.2.6, we get that for n large enough

$$\begin{aligned} M^{\theta}_{1,1,1,0,2} &\leq C\Delta_n^{-2\gamma}\left(1+|X_{t_k}|^2\right)e^{-\lambda_n\Delta_n}\lambda_n\Delta_n\int_I\int_{\{|y-X_{t_k}|\leq\Delta_n^{\alpha}\}} q^{\theta}_{(1)}(\Delta_n,X_{t_k},y;a)dy\mu(da) \\ &\leq C\Delta_n^{-2\gamma}\left(1+|X_{t_k}|^2\right)e^{C_2\Delta_n^{1-2\gamma}X^2_{t_k}}e^{-C_0\Delta_n^{2\upsilon-1}}, \end{aligned}$$

for some constants $C, C_0, C_2 > 0$. Therefore, the term $M^{\theta}_{1,1,1,0}$ satisfies (4.18).

As for the term $M^{\theta}_{0,2}$, we have that $M^{\theta}_{1,1,1,2} \leq C\lambda_n\Delta_n^{3/2}(1+|X_{t_k}|^2)$ for some constant $C > 0$. Therefore, the term $M^{\theta}_{1,1,1}$ satisfies (4.13).

We next treat $M^{\theta}_{1,1,2}$. Using (4.12), we have that

$$\begin{aligned} M^{\theta}_{1,1,2} = \int_I\int_{\mathbb{R}}&\left(\frac{\int_I\left(c(X_{t_k},z)-c(X_{t_k},a)\right)q^{\theta}_{(1)}(\Delta_n,X_{t_k},y;a)\mu(da)e^{-\lambda_n\Delta_n}\lambda_n\Delta_n}{p^{\theta}(\Delta_n,X_{t_k},y)}\right)^2 \\ &\times q^{\theta}_{(1)}(\Delta_n,X_{t_k},y;z)e^{-\lambda_n\Delta_n}\lambda_n\Delta_n dy\mu(dz). \end{aligned}$$

We next fix α_0 and ε such that $\frac{1}{4} < \varepsilon < \alpha_0 < \frac{1}{2}$, and consider the set

$$E^k_z = \left\{a \in I : |c(X_{t_k},z)-c(X_{t_k},a)| \leq \Delta_n^{\varepsilon},\ \text{for all } z \in I\right\}.$$

We next split the integral inside the square of $M^{\theta}_{1,1,2}$ over the sets $\mathbf{1}_{E^k_z}$ and $\mathbf{1}_{(E^k_z)^c}$ and call both terms $M^{\theta}_{1,1,2,1}$ and $M^{\theta}_{1,1,2,2}$. First, (4.16) and Lemma 4.2.6 yield that

$$M^{\theta}_{1,1,2,1} \leq Ce^{-\lambda_n\Delta_n}\lambda_n\Delta_n^{1+2\varepsilon}\int_I\int_{\mathbb{R}} q^{\theta}_{(1)}(\Delta_n,X_{t_k},y;z)dy\mu(dz) \leq C\lambda_n\Delta_n^{1+2\varepsilon}e^{C_1\Delta_n^{1-2\gamma}X^2_{t_k}}, \tag{4.20}$$

for some constants $C, C_1 > 0$.

Next, we treat $M^{\theta}_{1,1,2,2}$ by dividing the domain of the dy integral into the subdomains $I_1 := \{y : |y-X_{t_k}-c(X_{t_k},z)| > \Delta_n^{\alpha_0}\}$ and $I_2 := \{y : |y-X_{t_k}-c(X_{t_k},z)| \leq \Delta_n^{\alpha_0}\}$, and call both terms

$M^{\theta}_{1,1,2,2,1}$ and $M^{\theta}_{1,1,2,2,2}$. Then, using hypothesis **(A1)**, together with (4.16) and Lemma 4.2.6, we get that

$$\begin{aligned} M^{\theta}_{1,1,2,2,1} &\leq C\Delta_n^{-2\gamma}\left(1+|X_{t_k}|^2\right)e^{-\lambda_n\Delta_n}\lambda_n\Delta_n\int_I\int_{I_1} q^{\theta}_{(1)}(\Delta_n, X_{t_k}, y; z)dy\mu(dz)\\ &\leq C\Delta_n^{-2\gamma}\left(1+|X_{t_k}|^2\right)e^{C_2\Delta_n^{1-2\gamma}X_{t_k}^2}\int_I\int_{I_1}\frac{1}{\sqrt{\Delta_n}}e^{-\frac{(y-X_{t_k}-c(X_{t_k},z))^2}{C_3\Delta_n}}dy\mu(dz)\\ &\leq C\Delta_n^{-2\gamma}\left(1+|X_{t_k}|^2\right)e^{C_2\Delta_n^{1-2\gamma}X_{t_k}^2}e^{-C_0\Delta_n^{2\alpha_0-1}}, \end{aligned}$$

for some constants $C, C_0, C_2, C_3 > 0$.

Next, (4.8) yields

$$\left(p^{\theta}(\Delta_n, X_{t_k}, y)\right)^2 \geq p^{\theta}(\Delta_n, X_{t_k}, y)\int_I q^{\theta}_{(1)}(\Delta_n, X_{t_k}, y; a)\mu(da)e^{-\lambda_n\Delta_n}\lambda_n\Delta_n.$$

Then, using hypothesis **(A1)** and Lemma 4.2.6, we obtain that

$$\begin{aligned} M^{\theta}_{1,1,2,2,2} &\leq C\Delta_n^{-2\gamma}\left(1+|X_{t_k}|^2\right)e^{-\lambda_n\Delta_n}\lambda_n\Delta_n\\ &\quad\times\int_I\int_{I_2}\int_I \mathbf{1}_{(E^k_z)^c}q^{\theta}_{(1)}(\Delta_n, X_{t_k}, y; a)\mu(da)\frac{q^{\theta}_{(1)}(\Delta_n, X_{t_k}, y; z)e^{-\lambda_n\Delta_n}\lambda_n\Delta_n}{p^{\theta}(\Delta_n, X_{t_k}, y)}dy\mu(dz)\\ &\leq C\Delta_n^{-2\gamma}\left(1+|X_{t_k}|^2\right)e^{C_2\Delta_n^{1-2\gamma}X_{t_k}^2}\int_I\int_{I_2}\int_I \mathbf{1}_{(E^k_z)^c}\frac{1}{\sqrt{\Delta_n}}e^{-\frac{\left(y-X_{t_k}-c(X_{t_k},a)\right)^2}{C_3\Delta_n}}\mu(da)\\ &\quad\times\frac{q^{\theta}_{(1)}(\Delta_n, X_{t_k}, y; z)e^{-\lambda_n\Delta_n}\lambda_n\Delta_n}{p^{\theta}(\Delta_n, X_{t_k}, y)}dy\mu(dz)\\ &\leq C\Delta_n^{-2\gamma}\left(1+|X_{t_k}|^2\right)e^{C_2\Delta_n^{1-2\gamma}X_{t_k}^2}\int_I\int_I\int_{\{|h|\leq\Delta_n^{\alpha_0}\}}\mathbf{1}_{(E^k_z)^c}\frac{1}{\sqrt{\Delta_n}}e^{-\frac{\left(h+c(X_{t_k},z)-c(X_{t_k},a)\right)^2}{C_3\Delta_n}}\\ &\quad\times\frac{q^{\theta}_{(1)}(\Delta_n, X_{t_k}, h+X_{t_k}+c(X_{t_k},z); z)e^{-\lambda_n\Delta_n}\lambda_n\Delta_n}{p^{\theta}(\Delta_n, X_{t_k}, h+X_{t_k}+c(X_{t_k},z))}dh\mu(da)\mu(dz), \end{aligned}$$

for some constants $C, C_2, C_3 > 0$, where we have used the change of variable $h := y - X_{t_k} - c(X_{t_k}, z)$.

Since $|h| \leq \Delta_n^{\alpha_0}$ and $|c(X_{t_k}, z) - c(X_{t_k}, a)| > \Delta_n^{\varepsilon}$ on $(E^k_z)^c$, for n large enough there exists a constant $C_4 \in (0,1)$ such that $|h + c(X_{t_k}, z) - c(X_{t_k}, a)| \geq C_4 \Delta_n^{\varepsilon}$. Then, we deduce that

$$\begin{aligned}
M_{1,1,2,2,2}^{\theta} &\leq C\Delta_n^{-\frac{1}{2}-2\gamma} e^{-\frac{C_4^2 \Delta_n^{2\varepsilon-1}}{C_3}} \left(1+|X_{t_k}|^2\right) e^{C_2 \Delta_n^{1-2\gamma} X_{t_k}^2} \\
&\quad \times \int_I \int_{\{|h| \leq \Delta_n^{\alpha_0}\}} \frac{q_{(1)}^{\theta}(\Delta_n, X_{t_k}, h + X_{t_k} + c(X_{t_k}, z); z) e^{-\lambda_n \Delta_n} \lambda_n \Delta_n}{p^{\theta}(\Delta_n, X_{t_k}, h + X_{t_k} + c(X_{t_k}, z))} dh \mu(dz) \\
&= C\Delta_n^{-\frac{1}{2}-2\gamma} e^{-\frac{C_4^2 \Delta_n^{2\varepsilon-1}}{C_3}} \left(1+|X_{t_k}|^2\right) e^{C_2 \Delta_n^{1-2\gamma} X_{t_k}^2} \\
&\quad \times \int_I \int_{\{|y - X_{t_k} - c(X_{t_k}, z)| \leq \Delta_n^{\alpha_0}\}} \frac{q_{(1)}^{\theta}(\Delta_n, X_{t_k}, y; z) e^{-\lambda_n \Delta_n} \lambda_n \Delta_n}{p^{\theta}(\Delta_n, X_{t_k}, y)} dy \mu(dz) \\
&\leq C\Delta_n^{-\frac{1}{2}-2\gamma} e^{-\frac{C_4^2 \Delta_n^{2\varepsilon-1}}{C_3}} \left(1+|X_{t_k}|^2\right) e^{C_2 \Delta_n^{1-2\gamma} X_{t_k}^2} \\
&\quad \times \int_{\{|y - X_{t_k}| \leq \Delta_n^{\alpha_0} + C\Delta_n^{-\gamma}(1+|X_{t_k}|)\}} \frac{\int_I q_{(1)}^{\theta}(\Delta_n, X_{t_k}, y; z) \mu(dz) e^{-\lambda_n \Delta_n} \lambda_n \Delta_n}{p^{\theta}(\Delta_n, X_{t_k}, y)} dy \\
&= C\Delta_n^{-\frac{1}{2}-2\gamma+\alpha_0} e^{-\frac{C_4^2 \Delta_n^{2\varepsilon-1}}{C_3}} \left(1+|X_{t_k}|^2\right) e^{C_2 \Delta_n^{1-2\gamma} X_{t_k}^2} \\
&\qquad + C\Delta_n^{-\frac{1}{2}-3\gamma} e^{-\frac{C_4^2 \Delta_n^{2\varepsilon-1}}{C_3}} \left(1+|X_{t_k}|^3\right) e^{C_2 \Delta_n^{1-2\gamma} X_{t_k}^2} \\
&\leq C\Delta_n^{-\frac{1}{2}-3\gamma} e^{-\frac{C_4^2 \Delta_n^{2\varepsilon-1}}{C_3}} \left(1+|X_{t_k}|^3\right) e^{C_2 \Delta_n^{1-2\gamma} X_{t_k}^2},
\end{aligned}$$

where we have used the change of variable $y := h + X_{t_k} + c(X_{t_k}, z)$, the linear growth condition on c, together with (4.16).

Therefore, we have shown that for n large enough and $\alpha_0 \in (\epsilon, \frac{1}{2})$,

$$M_{1,1,2,2}^{\theta} \leq C\Delta_n^{-\frac{1}{2}-3\gamma} \left(1+|X_{t_k}|^3\right) e^{C_1 \Delta_n^{1-2\gamma} X_{t_k}^2} e^{-C_0 \Delta_n^{2\alpha_0 - 1}},$$

for some constants $C, C_0, C_1 > 0$, which together with (4.20) gives

$$M_{1,1,2}^{\theta} \leq C e^{C_1 \Delta_n^{1-2\gamma} X_{t_k}^2} \left(1+|X_{t_k}|^3\right) \left(\lambda_n \Delta_n^{1+2\epsilon} + \Delta_n^{-\frac{1}{2}-3\gamma} e^{-C_0 \Delta_n^{2\alpha_0-1}}\right).$$

Finally, as for the term $M_{0,2}^{\theta}$, we obtain that $M_{1,2}^{\theta} + M_2^{\theta} \leq C\lambda_n \Delta_n^{3/2}(1 + |X_{t_k}|^2)$, which concludes the proof of (4.14) and (4.15). □

Finally, we recall a discrete ergodic theorem.

Lemma 4.2.9. [40, Lemma 8] *Assume conditions* **(A1)** *and* **(A6)**. *Consider a differentiable function $g : \mathbb{R} \to \mathbb{R}$, whose derivatives have polynomial growth in x. Then, as $n \to \infty$,*

$$\frac{1}{n} \sum_{k=0}^{n-1} g(X_{t_k}) \xrightarrow{\mathrm{P}^{\theta_0}} \int_{\mathbb{R}} g(x) \pi_{\theta_0}(dx).$$

4.3 Proof of Theorem 4.1.1

In this section, the proof of Theorem 4.1.1 will be divided into several steps. We begin deriving a stochastic expansion of the log-likelihood ratio using Proposition 4.2.1 and Lemma 4.2.1. The second step is devoted to treat the negligible contributions of this expansion. Finally, the last step concludes the LAN property by applying the central limit theorem for triangular arrays.

4.3.1 Expansion of the log-likelihood ratio

By the Markov property and Proposition 4.2.1,

$$\begin{aligned}\log\frac{p(X^n;\theta_n)}{p(X^n;\theta_0)} &= \sum_{k=0}^{n-1}\log\frac{p^{\theta_n}}{p^{\theta_0}}(\Delta_n,X_{t_k},X_{t_{k+1}})\\ &= \sum_{k=0}^{n-1}\frac{u}{\sqrt{n\Delta_n}}\int_0^1\frac{\partial_\theta p^{\theta(\ell)}}{p^{\theta(\ell)}}(\Delta_n,X_{t_k},X_{t_{k+1}})d\ell\\ &= \sum_{k=0}^{n-1}\frac{u}{\sqrt{n\Delta_n}}\int_0^1\frac{1}{\Delta_n}\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)}\left[\delta\left(\partial_\theta Y_{t_{k+1}}^{\theta(\ell)}(t_k,X_{t_k})U^{\theta(\ell)}(t_k,X_{t_k})\right)\Big|Y_{t_{k+1}}^{\theta(\ell)}=X_{t_{k+1}}\right]d\ell,\end{aligned}$$

where $\theta(\ell):=\theta_n(\ell,u):=\theta_0+\frac{\ell u}{\sqrt{n\Delta_n}}$.

We next consider the stopping time

$$\widehat{\tau}:=\inf\left\{s\geq 0:|\Delta\widehat{Z}_s|<\rho_1\Delta_n^{\upsilon}\text{ or }\ |\Delta\widehat{Z}_s|>\rho_2\Delta_n^{-\gamma}\right\},\tag{4.21}$$

and

$$\widetilde{\tau}:=\inf\left\{s\geq 0:|\Delta\widetilde{Z}_s|<\rho_1\Delta_n^{\upsilon}\text{ or }\ |\Delta\widetilde{Z}_s|>\rho_2\Delta_n^{-\gamma}\right\},\tag{4.22}$$

where $\rho_1,\rho_2>0$ and $0<\upsilon,\gamma<\frac{1}{2}$ are from hypothesis **(A7)**.

Observe that on the event $\{\widehat{\tau}>n\Delta_n\}$, all the jumps of $\widehat{Z}$ in the interval $[0,n\Delta_n]$ are in the interval $[\rho_1\Delta_n^{\upsilon},\rho_2\Delta_n^{-\gamma}]$. Hence, for all $\omega\in\{\widehat{\tau}>n\Delta_n\}$, X^θ satisfies

$$X_t^\theta=x_0+\int_0^t b(\theta,X_s^\theta)ds+\int_0^t\sigma(X_s^\theta)dB_s+\int_0^t\int_I c(X_{s-}^\theta,z)\left(N(ds,dz)-\nu(dz)ds\right),\tag{4.23}$$

for all $t\in[0,n\Delta_n]$, where recall that $I=\{z\in\mathbb{R}_0:\rho_1\Delta_n^{\upsilon}\leq|z|\leq\rho_2\Delta_n^{-\gamma}\}$. A similar statement is true for Y^θ.

Then, multiplying by $\mathbf{1}_{\{\widetilde{\tau}>n\Delta_n\}}+\mathbf{1}_{\{\widetilde{\tau}\leq n\Delta_n\}}$ inside and $\mathbf{1}_{\{\widehat{\tau}>n\Delta_n\}}+\mathbf{1}_{\{\widehat{\tau}\leq n\Delta_n\}}$ outside the conditional expectation above, we get that

$$\log\frac{p(X^n;\theta_n)}{p(X^n;\theta_0)}=\frac{u}{\sqrt{n\Delta_n^3}}\sum_{k=0}^{n-1}\int_0^1\left(Z_{k,n}^{1,\ell}+Z_{k,n}^{2,\ell}+Z_{k,n}^{3,\ell}\right)d\ell,$$

where

$$\begin{aligned}Z_{k,n}^{1,\ell}&=\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)}\left[\delta\left(\partial_\theta Y_{t_{k+1}}^{\theta(\ell)}(t_k,X_{t_k})U^{\theta(\ell)}(t_k,X_{t_k})\right)\Big|Y_{t_{k+1}}^{\theta(\ell)}=X_{t_{k+1}}\right]\mathbf{1}_{\{\widehat{\tau}\leq n\Delta_n\}},\\ Z_{k,n}^{2,\ell}&=\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)}\left[\delta\left(\partial_\theta Y_{t_{k+1}}^{\theta(\ell)}(t_k,X_{t_k})U^{\theta(\ell)}(t_k,X_{t_k})\right)\mathbf{1}_{\{\widetilde{\tau}\leq n\Delta_n\}}\Big|Y_{t_{k+1}}^{\theta(\ell)}=X_{t_{k+1}}\right]\mathbf{1}_{\{\widehat{\tau}>n\Delta_n\}},\\ Z_{k,n}^{3,\ell}&=\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)}\left[\delta\left(\partial_\theta Y_{t_{k+1}}^{\theta(\ell)}(t_k,X_{t_k})U^{\theta(\ell)}(t_k,X_{t_k})\right)\mathbf{1}_{\{\widetilde{\tau}>n\Delta_n\}}\Big|Y_{t_{k+1}}^{\theta(\ell)}=X_{t_{k+1}}\right]\mathbf{1}_{\{\widehat{\tau}>n\Delta_n\}}.\end{aligned}$$

We will later see that the terms concerning $Z_{k,n}^{1,\ell}$ and $Z_{k,n}^{2,\ell}$ are negligible (Lemma 4.3.1). The main contribution in the asymptotics will be given by $Z_{k,n}^{3,\ell}$, which expresses the fact that the small and large jumps do not interfere with the Gaussian behaviour of the transition density. In fact to see this, applying Lemma 4.2.1 to $Z_{k,n}^{3,\ell}$, and using equation (4.1) for the term $X_{t_{k+1}}-X_{t_k}$ coming from the term $Y_{t_{k+1}}^{\theta(\ell)}-Y_{t_k}^{\theta(\ell)}$ in Lemma 4.2.1, we obtain the following expansion of the log-likehood ratio

$$\begin{aligned}\log\frac{p(X^n;\theta_n)}{p(X^n;\theta_0)}=\sum_{k=0}^{n-1}\xi_{k,n}+\frac{u}{\sqrt{n\Delta_n^3}}\sum_{k=0}^{n-1}\int_0^1\bigg\{&Z_{k,n}^{1,\ell}+Z_{k,n}^{2,\ell}\\ &+\left(Z_{k,n}^{4,\ell}+Z_{k,n}^{5,\ell}+Z_{k,n}^{6,\ell}\right)\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)}\left[\mathbf{1}_{\{\widetilde{\tau}>n\Delta_n\}}\Big|Y_{t_{k+1}}^{\theta(\ell)}=X_{t_{k+1}}\right]\mathbf{1}_{\{\widehat{\tau}>n\Delta_n\}}\\ &+\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)}\left[\left(R^{\theta(\ell),k}-R_4^{\theta(\ell),k}-R_6^{\theta(\ell),k}\right)\mathbf{1}_{\{\widetilde{\tau}>n\Delta_n\}}\Big|Y_{t_{k+1}}^{\theta(\ell)}=X_{t_{k+1}}\right]\mathbf{1}_{\{\widehat{\tau}>n\Delta_n\}}\bigg\}d\ell,\end{aligned}$$

where

$$
\begin{aligned}
\xi_{k,n} &= \frac{u}{\sqrt{n\Delta_n}} \int_0^1 \frac{\partial_\theta b(\theta(\ell), X_{t_k})}{\sigma^2(X_{t_k})} \left(\sigma(X_{t_k}) \left(B_{t_{k+1}} - B_{t_k}\right) + \left(b(\theta_0, X_{t_k}) - b(\theta(\ell), X_{t_k})\right) \Delta_n\right) \\
&\quad \times \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)} \left[\mathbf{1}_{\{\widetilde{\tau} > n\Delta_n\}} \Big| Y_{t_{k+1}}^{\theta(\ell)} = X_{t_{k+1}} \right] \mathbf{1}_{\{\widehat{\tau} > n\Delta_n\}} d\ell, \\
Z_{k,n}^{4,\ell} &= \Delta_n \partial_\theta b(\theta(\ell), X_{t_k}) \sigma^{-2}(X_{t_k}) \int_{t_k}^{t_{k+1}} \left(b(\theta_0, X_s^{\theta_0}) - b(\theta_0, X_{t_k}) \right) ds, \\
Z_{k,n}^{5,\ell} &= \Delta_n \partial_\theta b(\theta(\ell), X_{t_k}) \sigma^{-2}(X_{t_k}) \int_{t_k}^{t_{k+1}} \left(\sigma(X_s^{\theta_0}) - \sigma(X_{t_k}) \right) dB_s, \\
Z_{k,n}^{6,\ell} &= \Delta_n \partial_\theta b(\theta(\ell), X_{t_k}) \sigma^{-2}(X_{t_k}) \int_{t_k}^{t_{k+1}} \int_{\mathbb{R}_0} c(X_{s-}^{\theta_0}, z) \widetilde{N}(ds, dz), \\
R^{\theta(\ell),k} &= R_1^{\theta(\ell),k} + R_2^{\theta(\ell),k} + R_3^{\theta(\ell),k} - R_5^{\theta(\ell),k}.
\end{aligned}
$$

In the next subsections we will show that $\xi_{k,n}$ is the only term that contributes to the limit in Theorem 4.1.1, and all the others are negligible contributions. Therefore again, the main behaviour is given by the Gaussian and drift components of the equation (4.1).

4.3.2 Negligible contributions

Lemma 4.3.1. *Under conditions* **(A1)**-**(A5)** *and* **(A7)**, *as $n \to \infty$,*

$$
\frac{u}{\sqrt{n\Delta_n^3}} \sum_{k=0}^{n-1} \int_0^1 \left(Z_{k,n}^{1,\ell} + Z_{k,n}^{2,\ell} \right) d\ell \xrightarrow{\mathrm{P}^{\theta_0}} 0.
$$

Proof. It suffices to show that condition (1.16) of Lemma 1.4.2 holds for each sequence $(Z_{k,n}^{i,\ell})_{k\geq 1}$ under the measure P^{θ_0}.

First, applying Hölder's and Jensen's inequalities, Girsanov's theorem, Lemma 4.2.5, and (4.7), we obtain that for some constants $C, q_0 > 0$,

$$
\begin{aligned}
&\frac{|u|}{\sqrt{n\Delta_n^3}} \sum_{k=0}^{n-1} \mathrm{E}\left[\left| \int_0^1 Z_{k,n}^{1,\ell} d\ell \right| \Big| \widehat{\mathcal{F}}_{t_k} \right] \\
&\leq \frac{|u|}{\sqrt{n\Delta_n^3}} \sum_{k=0}^{n-1} \int_0^1 \left(\mathrm{E}\left[\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)} \left[\left| \delta\left(\partial_\theta Y_{t_{k+1}}^{\theta(\ell)}(t_k, X_{t_k}) U^{\theta(\ell)}(t_k, X_{t_k}) \right) \right|^p \Big| Y_{t_{k+1}}^{\theta(\ell)} = X_{t_{k+1}} \right] \Big| X_{t_k} \right] \right)^{\frac{1}{p}} \\
&\quad \times \left(\mathrm{P}\left(\widehat{\tau} \leq n\Delta_n | X_{t_k} \right) \right)^{\frac{1}{q}} d\ell \\
&\leq \frac{|u|}{\sqrt{n\Delta_n^3}} \sum_{k=0}^{n-1} \int_0^1 \left(\mathrm{E}_{\widehat{Q}_k^{\theta(\ell),\theta_0}} \left[\left| \delta\left(\partial_\theta Y_{t_{k+1}}^{\theta(\ell)}(t_k, X_{t_k}) U^{\theta(\ell)}(t_k, X_{t_k}) \right) \right|^p \left(\frac{d\widehat{\mathrm{P}}}{d\widehat{Q}_k^{\theta(\ell),\theta_0}} - 1 \right) \Big| X_{t_k} \right] \right. \\
&\quad \left. + \mathrm{E}_{\widehat{Q}_k^{\theta(\ell),\theta_0}} \left[\left| \delta\left(\partial_\theta Y_{t_{k+1}}^{\theta(\ell)}(t_k, X_{t_k}) U^{\theta(\ell)}(t_k, X_{t_k}) \right) \right|^p \Big| X_{t_k} \right] \right)^{\frac{1}{p}} \left(\mathrm{P}\left(\widehat{\tau} \leq n\Delta_n | X_{t_k} \right) \right)^{\frac{1}{q}} d\ell \\
&\leq \frac{C|u|}{\sqrt{n}} \sum_{k=0}^{n-1} \left(1 + |X_{t_k}|^{q_0} \right) \left(\mathrm{P}\left(\widehat{\tau} \leq n\Delta_n | X_{t_k} \right) \right)^{\frac{1}{q}},
\end{aligned}
$$

where $p, q > 1$ and $\frac{1}{p} + \frac{1}{q} = 1$. On the other hand,

$$
\begin{aligned}
&\mathrm{P}(\widehat{\tau} > n\Delta_n | X_{t_k}) = \mathrm{P}(\forall s \in [0, n\Delta_n], \;\; \rho_1 \Delta_n^{\upsilon} \leq |\Delta \widehat{Z}_s| \leq \rho_2 \Delta_n^{-\gamma} | X_{t_k}) \\
&= \mathrm{P}(\forall s \in [0, k\Delta_n], \;\; \rho_1 \Delta_n^{\upsilon} \leq |\Delta \widehat{Z}_s| \leq \rho_2 \Delta_n^{-\gamma} | X_{t_k}) \\
&\qquad \times \sum_{j=0}^{\infty} \mathrm{P}(\{\forall s \in [k\Delta_n, n\Delta_n], \;\; \rho_1 \Delta_n^{\upsilon} \leq |\Delta \widehat{Z}_s| \leq \rho_2 \Delta_n^{-\gamma}\} \cap \{N_{n\Delta_n} - N_{k\Delta_n} = j\}) \\
&\leq \sum_{j=0}^{\infty} e^{-\lambda_n (n-k)\Delta_n} \frac{(\lambda_n (n-k)\Delta_n)^j}{j!} \mathrm{P}(\rho_1 \Delta_n^{\upsilon} \leq |\widehat{\Lambda}| \leq \rho_2 \Delta_n^{-\gamma})^j \\
&= e^{-\lambda_n (n-k)\Delta_n \left(1 - \mathrm{P}\left(\rho_1 \Delta_n^{\upsilon} \leq |\widehat{\Lambda}| \leq \rho_2 \Delta_n^{-\gamma}\right)\right)},
\end{aligned}
$$

where $\widehat{\Lambda}$ is a random variable with distribution $\frac{\nu}{\lambda}$. Therefore, we obtain that

$$
\begin{aligned}
&\frac{|u|}{\sqrt{n\Delta_n^3}} \sum_{k=0}^{n-1} \mathrm{E}\left[\left| \int_0^1 Z_{k,n}^{1,\ell} d\ell \right| \Big| \widehat{\mathcal{F}}_{t_k} \right] \\
&\leq \frac{C|u|}{\sqrt{n}} \sum_{k=0}^{n-1} (1 + |X_{t_k}|^{q_0}) \left(1 - e^{-\lambda_n (n-k)\Delta_n \left(1 - \mathrm{P}\left(\rho_1 \Delta_n^{\upsilon} \leq |\widehat{\Lambda}| \leq \rho_2 \Delta_n^{-\gamma}\right)\right)}\right)^{\frac{1}{q}} \\
&\leq \left(1 - e^{-\lambda_n n\Delta_n \left(1 - \mathrm{P}\left(\rho_1 \Delta_n^{\upsilon} \leq |\widehat{\Lambda}| \leq \rho_2 \Delta_n^{-\gamma}\right)\right)}\right)^{\frac{1}{q}} \frac{C|u|}{\sqrt{n}} \sum_{k=0}^{n-1} (1 + |X_{t_k}|^{q_0}).
\end{aligned}
$$

Then, using the fact that $1 - e^{-x} \leq x$, for all $x \geq 0$, and that $\lambda_n \leq \lambda$, we get that

$$
\begin{aligned}
&\left(1 - e^{-\lambda_n n\Delta_n \left(1 - \mathrm{P}\left(\rho_1 \Delta_n^{\upsilon} \leq |\widehat{\Lambda}| \leq \rho_2 \Delta_n^{-\gamma}\right)\right)}\right)^{\frac{1}{q}} \\
&\leq \left(\lambda n\Delta_n \left(1 - \mathrm{P}\left(\rho_1 \Delta_n^{\upsilon} \leq |\widehat{\Lambda}| \leq \rho_2 \Delta_n^{-\gamma}\right)\right)\right)^{\frac{1}{q}} \\
&\leq c_q \left\{ \left(\lambda n\Delta_n \mathrm{P}\left(|\widehat{\Lambda}| \geq \rho_2 \Delta_n^{-\gamma}\right)\right)^{\frac{1}{q}} + \left(\lambda n\Delta_n \mathrm{P}\left(\rho_1 \Delta_n^{\upsilon} \geq |\widehat{\Lambda}|\right)\right)^{\frac{1}{q}} \right\} \\
&= c_q \left\{ \left(n\Delta_n \int_{\{|z| \geq \rho_2 \Delta_n^{-\gamma}\}} \nu(dz)\right)^{\frac{1}{q}} + \left(n\Delta_n \int_{\{|z| \leq \rho_1 \Delta_n^{\upsilon}\}} \nu(dz)\right)^{\frac{1}{q}} \right\}.
\end{aligned}
$$

Therefore, by **(A7)** we conclude that (1.16) holds true, and by Lemma 1.4.2, as $n \to \infty$,

$$
\frac{u}{\sqrt{n\Delta_n^3}} \sum_{k=0}^{n-1} \int_0^1 Z_{k,n}^{1,\ell} d\ell \xrightarrow{\mathrm{P}^{\theta_0}} 0.
$$

Next, as for the term $Z_{k,n}^{1,\ell}$, applying Girsanov's theorem, Lemma 4.2.5, and (4.7), we obtain that for some constants $C, q_0 > 0$,

$$\begin{aligned}
&\frac{|u|}{\sqrt{n\Delta_n^3}}\sum_{k=0}^{n-1}\mathrm{E}\left[\left|\int_0^1 Z_{k,n}^{2,\ell}d\ell\right|\Big|\widehat{\mathcal{F}}_{t_k}\right]\\
&\leq \frac{|u|}{\sqrt{n\Delta_n^3}}\sum_{k=0}^{n-1}\int_0^1 \mathrm{E}\left[\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)}\left[\left|\delta\left(\partial_\theta Y_{t_{k+1}}^{\theta(\ell)}(t_k,X_{t_k})U^{\theta(\ell)}(t_k,X_{t_k})\right)\right|\mathbf{1}_{\{\widetilde{\tau}\leq n\Delta_n\}}\Big|Y_{t_{k+1}}^{\theta(\ell)}=X_{t_{k+1}}\right]\Big|X_{t_k}\right]d\ell\\
&\leq \frac{|u|}{\sqrt{n\Delta_n^3}}\sum_{k=0}^{n-1}\int_0^1\left(\left|\mathrm{E}_{\widehat{Q}_k^{\theta(\ell),\theta_0}}\left[\left|\delta\left(\partial_\theta Y_{t_{k+1}}^{\theta(\ell)}(t_k,X_{t_k})U^{\theta(\ell)}(t_k,X_{t_k})\right)\right|\mathbf{1}_{\{\widetilde{\tau}\leq n\Delta_n\}}\left(\frac{d\widehat{\mathrm{P}}}{d\widehat{Q}_k^{\theta(\ell),\theta_0}}-1\right)\Big|X_{t_k}\right]\right|\right.\\
&\quad\left.+\,\mathrm{E}_{\widehat{Q}_k^{\theta(\ell),\theta_0}}\left[\left|\delta\left(\partial_\theta Y_{t_{k+1}}^{\theta(\ell)}(t_k,X_{t_k})U^{\theta(\ell)}(t_k,X_{t_k})\right)\right|\mathbf{1}_{\{\widetilde{\tau}\leq n\Delta_n\}}\Big|X_{t_k}\right]\right)d\ell\\
&\leq \frac{C|u|}{\sqrt{n}}\sum_{k=0}^{n-1}\left(1+|X_{t_k}|^{q_0}\right)\left(\mathrm{P}\left(\widetilde{\tau}\leq n\Delta_n|X_{t_k}\right)\right)^{\frac{1}{q}},
\end{aligned}$$

where we have used Hölder's inequality with $p > 1$ and $q > 1$ conjugate. On the other hand,

$$\begin{aligned}
\mathrm{P}(\widetilde{\tau} > n\Delta_n|X_{t_k}) &= \mathrm{P}(\widetilde{\tau} > n\Delta_n) = \mathrm{P}(\forall s\in[0,n\Delta_n],\quad \rho_1\Delta_n^{v}\leq|\Delta\widetilde{Z}_s|\leq\rho_2\Delta_n^{-\gamma})\\
&= \sum_{j=0}^{\infty}\mathrm{P}(\{\forall s\in[0,n\Delta_n],\quad \rho_1\Delta_n^{v}\leq|\Delta\widetilde{Z}_s|\leq\rho_2\Delta_n^{-\gamma}\}\cap\{M_{n\Delta_n}-M_0=j\})\\
&= \sum_{j=0}^{\infty}e^{-\lambda_n n\Delta_n}\frac{(\lambda_n n\Delta_n)^j}{j!}\mathrm{P}(\rho_1\Delta_n^{v}\leq|\widetilde{\Lambda}|\leq\rho_2\Delta_n^{-\gamma})^j\\
&= e^{-\lambda_n n\Delta_n\left(1-\mathrm{P}\left(\rho_1\Delta_n^{v}\leq|\widetilde{\Lambda}|\leq\rho_2\Delta_n^{-\gamma}\right)\right)},
\end{aligned}$$

where $\widetilde{\Lambda}$ is a random variable with distribution $\frac{\nu}{\lambda}$. Therefore, we obtain that

$$\begin{aligned}
&\frac{|u|}{\sqrt{n\Delta_n^3}}\sum_{k=0}^{n-1}\mathrm{E}\left[\left|\int_0^1 Z_{k,n}^{2,\ell}d\ell\right|\Big|\widehat{\mathcal{F}}_{t_k}\right]\\
&\leq \frac{|u|}{\sqrt{n}}\sum_{k=0}^{n-1}\left(1+|X_{t_k}|^{q_0}\right)\left(1-e^{-\lambda_n n\Delta_n\left(1-\mathrm{P}\left(\rho_1\Delta_n^{v}\leq|\widetilde{\Lambda}|\leq\rho_2\Delta_n^{-\gamma}\right)\right)}\right)^{\frac{1}{q}}\\
&\leq c_q\left\{\left(n\Delta_n\int_{\{|z|\geq\rho_2\Delta_n^{-\gamma}\}}\nu(dz)\right)^{\frac{1}{q}}+\left(n\Delta_n\int_{\{|z|\leq\rho_1\Delta_n^{v}\}}\nu(dz)\right)^{\frac{1}{q}}\right\}\frac{|u|}{\sqrt{n}}\sum_{k=0}^{n-1}\left(1+|X_{t_k}|^{q_0}\right).
\end{aligned}$$

Therefore, by **(A7)** we conclude that (1.16) holds true, and by Lemma 1.4.2, as $n \to \infty$,

$$\frac{u}{\sqrt{n\Delta_n^3}}\sum_{k=0}^{n-1}\int_0^1 Z_{k,n}^{2,\ell}d\ell \xrightarrow{\mathrm{P}^{\theta_0}} 0.$$

Thus, the result follows. □

Lemma 4.3.2. *Under conditions* **(A1)**-**(A5)** *and* **(A7)**, *as* $n \to \infty$,

$$\sum_{k=0}^{n-1}\frac{u}{\sqrt{n\Delta_n^3}}\int_0^1\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)}\left[R^{\theta(\ell),k}\mathbf{1}_{\{\widetilde{\tau}>n\Delta_n\}}\Big|Y_{t_{k+1}}^{\theta(\ell)}=X_{t_{k+1}}\right]\mathbf{1}_{\{\widehat{\tau}>n\Delta_n\}}d\ell \xrightarrow{\mathrm{P}^{\theta_0}} 0.$$

Proof. Using the fact that $\mathbf{1}_{\{\widehat{\tau}>n\Delta_n\}} = 1-\mathbf{1}_{\{\widehat{\tau}\leq n\Delta_n\}}$ and $\mathbf{1}_{\{\widetilde{\tau}>n\Delta_n\}} = 1-\mathbf{1}_{\{\widetilde{\tau}\leq n\Delta_n\}}$, it suffices to show that as $n\to\infty$,

$$\sum_{k=0}^{n-1}\frac{u}{\sqrt{n\Delta_n^3}}\int_0^1 \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)}\left[R^{\theta(\ell),k}\mathbf{1}_{\{\widetilde{\tau}>n\Delta_n\}}\Big|Y_{t_{k+1}}^{\theta(\ell)}=X_{t_{k+1}}\right]\mathbf{1}_{\{\widehat{\tau}\leq n\Delta_n\}}d\ell \overset{\mathrm{P}^{\theta_0}}{\longrightarrow} 0, \tag{4.24}$$

$$\sum_{k=0}^{n-1}\frac{u}{\sqrt{n\Delta_n^3}}\int_0^1 \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)}\left[R^{\theta(\ell),k}\mathbf{1}_{\{\widetilde{\tau}\leq n\Delta_n\}}\Big|Y_{t_{k+1}}^{\theta(\ell)}=X_{t_{k+1}}\right]d\ell \overset{\mathrm{P}^{\theta_0}}{\longrightarrow} 0, \tag{4.25}$$

$$\sum_{k=0}^{n-1}\frac{u}{\sqrt{n\Delta_n^3}}\int_0^1 \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)}\left[R^{\theta(\ell),k}\Big|Y_{t_{k+1}}^{\theta(\ell)}=X_{t_{k+1}}\right]d\ell \overset{\mathrm{P}^{\theta_0}}{\longrightarrow} 0. \tag{4.26}$$

The convergences (4.24) and (4.25) are treated similarly as for the terms $Z_{k,n}^{1,\ell}$ and $Z_{k,n}^{2,\ell}$. To treat (4.26), it suffices to show that conditions (i) and (ii) of Lemma 1.4.1 hold under the measure P^{θ_0}. We start showing (i). Applying Girsanov's theorem, Lemma 4.2.5, (4.5), and (4.6) with $p=2$, we get that

$$\begin{aligned}
&\left|\sum_{k=0}^{n-1}\frac{u}{\sqrt{n\Delta_n^3}}\int_0^1 \mathrm{E}\left[\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)}\left[R^{\theta(\ell),k}\Big|Y_{t_{k+1}}^{\theta(\ell)}=X_{t_{k+1}}\right]\Big|\widehat{\mathcal{F}}_{t_k}\right]d\ell\right|\\
&\leq \sum_{k=0}^{n-1}\frac{|u|}{\sqrt{n\Delta_n^3}}\int_0^1\left|\mathrm{E}_{\widehat{Q}_k^{\theta(\ell),\theta_0}}\left[R^{\theta(\ell),k}\left(\frac{d\widehat{\mathrm{P}}}{d\widehat{Q}_k^{\theta(\ell),\theta_0}}-1\right)\Big|X_{t_k}\right]\right|d\ell\\
&\leq \frac{C|u|\Delta_n^{1/4}}{n}\sum_{k=0}^{n-1}\left(1+|X_{t_k}|^q\right),
\end{aligned}$$

for some constants $C,q>0$. Observe that (4.6) and (4.7) remain valid under the measure $\widehat{\mathrm{P}}^{\alpha}$ defined in Lemma 4.2.5. This shows Lemma 1.4.1(i). Similarly, applying Jensen's inequality, Girsanov's theorem, Lemma 4.2.5, and (4.6) with $p\in\{2,4\}$, we obtain that

$$\begin{aligned}
&\sum_{k=0}^{n-1}\frac{u^2}{n\Delta_n^3}\mathrm{E}\left[\left(\int_0^1 \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)}\left[R^{\theta(\ell),k}\Big|Y_{t_{k+1}}^{\theta(\ell)}=X_{t_{k+1}}\right]d\ell\right)^2\Big|\widehat{\mathcal{F}}_{t_k}\right]\\
&\leq \sum_{k=0}^{n-1}\frac{u^2}{n\Delta_n^3}\int_0^1\left\{\mathrm{E}_{\widehat{Q}_k^{\theta(\ell),\theta_0}}\left[\left(R^{\theta(\ell),k}\right)^2\Big|X_{t_k}\right]+\left|\mathrm{E}_{\widehat{Q}_k^{\theta(\ell),\theta_0}}\left[\left(R^{\theta(\ell),k}\right)^2\left(\frac{d\widehat{\mathrm{P}}}{d\widehat{Q}_k^{\theta(\ell),\theta_0}}-1\right)\Big|X_{t_k}\right]\right|\right\}d\ell\\
&\leq \frac{Cu^2\Delta_n^{1/4}}{n}\sum_{k=0}^{n-1}\left(1+|X_{t_k}|^q\right),
\end{aligned}$$

which concludes the desired result. □

Lemma 4.3.3. *Under conditions* **(A1)**-**(A2)**, **(A4)**-**(A5)** *and* **(A7)**, *as* $n\to\infty$,

$$\sum_{k=0}^{n-1}\frac{u}{\sqrt{n\Delta_n^3}}\int_0^1 Z_{k,n}^{5,\ell}\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)}\left[\mathbf{1}_{\{\widetilde{\tau}>n\Delta_n\}}\Big|Y_{t_{k+1}}^{\theta(\ell)}=X_{t_{k+1}}\right]\mathbf{1}_{\{\widehat{\tau}>n\Delta_n\}}d\ell \overset{\mathrm{P}^{\theta_0}}{\longrightarrow} 0.$$

Proof. Using the fact that $\mathbf{1}_{\{\widehat{\tau}>n\Delta_n\}} = 1-\mathbf{1}_{\{\widehat{\tau}\leq n\Delta_n\}}$ and $\mathbf{1}_{\{\widetilde{\tau}>n\Delta_n\}} = 1-\mathbf{1}_{\{\widetilde{\tau}\leq n\Delta_n\}}$, it suffices to show that as $n\to\infty$,

$$\sum_{k=0}^{n-1}\frac{u}{\sqrt{n\Delta_n^3}}\int_0^1 Z_{k,n}^{5,\ell}\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)}\left[\mathbf{1}_{\{\widetilde{\tau}>n\Delta_n\}}\Big|Y_{t_{k+1}}^{\theta(\ell)}=X_{t_{k+1}}\right]\mathbf{1}_{\{\widehat{\tau}\leq n\Delta_n\}}d\ell \overset{\mathrm{P}^{\theta_0}}{\longrightarrow} 0, \tag{4.27}$$

$$\sum_{k=0}^{n-1}\frac{u}{\sqrt{n\Delta_n^3}}\int_0^1 Z_{k,n}^{5,\ell}\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)}\left[\mathbf{1}_{\{\widetilde{\tau}\leq n\Delta_n\}}\Big|Y_{t_{k+1}}^{\theta(\ell)}=X_{t_{k+1}}\right]d\ell \overset{\mathrm{P}^{\theta_0}}{\longrightarrow} 0, \tag{4.28}$$

$$\sum_{k=0}^{n-1}\frac{u}{\sqrt{n\Delta_n^3}}\int_0^1 Z_{k,n}^{5,\ell}d\ell \overset{\mathrm{P}^{\theta_0}}{\longrightarrow} 0. \tag{4.29}$$

The convergences (4.27) and (4.28) are treated similarly as for the terms $Z_{k,n}^{1,\ell}$ and $Z_{k,n}^{2,\ell}$. We next treat (4.29). Clearly, for all $n \geq 1$,

$$\sum_{k=0}^{n-1} \frac{u}{\sqrt{n\Delta_n^3}} \int_0^1 \mathrm{E}\left[Z_{k,n}^{5,\ell} | \widehat{\mathcal{F}}_{t_k} \right] d\ell = 0,$$

and by Lemma 4.2.2(i),

$$\sum_{k=0}^{n-1} \frac{u^2}{n\Delta_n^3} \mathrm{E}\left[\left(\int_0^1 Z_{k,n}^{5,\ell} d\ell \right)^2 \Big| \widehat{\mathcal{F}}_{t_k} \right] \leq \frac{Cu^2\Delta_n}{n} \sum_{k=0}^{n-1} (1 + |X_{t_k}|^q),$$

for some constants $C, q > 0$. Thus, Lemma 1.4.1 concludes the desired result. □

Lemma 4.3.4. *Assume conditions* **(A1)**-**(A2)**, **(A4)**-**(A5)** *and* **(A7)**. *Then as* $n \to \infty$,

$$\sum_{k=0}^{n-1} \frac{u}{\sqrt{n\Delta_n^3}} \int_0^1 \bigg(Z_{k,n}^{4,\ell} \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)} \left[\mathbf{1}_{\{\widetilde{\tau} > n\Delta_n\}} \Big| Y_{t_{k+1}}^{\theta(\ell)} = X_{t_{k+1}} \right] \mathbf{1}_{\{\widehat{\tau} > n\Delta_n\}}$$
$$- \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)} \left[R_4^{\theta(\ell),k} \mathbf{1}_{\{\widetilde{\tau} > n\Delta_n\}} \Big| Y_{t_{k+1}}^{\theta(\ell)} = X_{t_{k+1}} \right] \mathbf{1}_{\{\widehat{\tau} > n\Delta_n\}} \bigg) d\ell \xrightarrow{\mathrm{P}^{\theta_0}} 0.$$

Proof. Using the fact that $\mathbf{1}_{\{\widehat{\tau} > n\Delta_n\}} = 1 - \mathbf{1}_{\{\widehat{\tau} \leq n\Delta_n\}}$ and $\mathbf{1}_{\{\widetilde{\tau} > n\Delta_n\}} = 1 - \mathbf{1}_{\{\widetilde{\tau} \leq n\Delta_n\}}$, it suffices to show that as $n \to \infty$,

$$\sum_{k=0}^{n-1} \frac{u}{\sqrt{n\Delta_n^3}} \int_0^1 Z_{k,n}^{4,\ell} \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)} \left[\mathbf{1}_{\{\widetilde{\tau} > n\Delta_n\}} \Big| Y_{t_{k+1}}^{\theta(\ell)} = X_{t_{k+1}} \right] \mathbf{1}_{\{\widehat{\tau} \leq n\Delta_n\}} d\ell \xrightarrow{\mathrm{P}^{\theta_0}} 0, \tag{4.30}$$

$$\sum_{k=0}^{n-1} \frac{u}{\sqrt{n\Delta_n^3}} \int_0^1 Z_{k,n}^{4,\ell} \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)} \left[\mathbf{1}_{\{\widetilde{\tau} \leq n\Delta_n\}} \Big| Y_{t_{k+1}}^{\theta(\ell)} = X_{t_{k+1}} \right] d\ell \xrightarrow{\mathrm{P}^{\theta_0}} 0, \tag{4.31}$$

$$\sum_{k=0}^{n-1} \frac{u}{\sqrt{n\Delta_n^3}} \int_0^1 \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)} \left[R_4^{\theta(\ell),k} \mathbf{1}_{\{\widetilde{\tau} > n\Delta_n\}} \Big| Y_{t_{k+1}}^{\theta(\ell)} = X_{t_{k+1}} \right] \mathbf{1}_{\{\widehat{\tau} \leq n\Delta_n\}} d\ell \xrightarrow{\mathrm{P}^{\theta_0}} 0, \tag{4.32}$$

$$\sum_{k=0}^{n-1} \frac{u}{\sqrt{n\Delta_n^3}} \int_0^1 \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)} \left[R_4^{\theta(\ell),k} \mathbf{1}_{\{\widetilde{\tau} \leq n\Delta_n\}} \Big| Y_{t_{k+1}}^{\theta(\ell)} = X_{t_{k+1}} \right] d\ell \xrightarrow{\mathrm{P}^{\theta_0}} 0, \tag{4.33}$$

$$\sum_{k=0}^{n-1} \frac{u}{\sqrt{n\Delta_n^3}} \int_0^1 \left(Z_{k,n}^{4,\ell} - \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)} \left[R_4^{\theta(\ell),k} \Big| Y_{t_{k+1}}^{\theta(\ell)} = X_{t_{k+1}} \right] \right) d\ell \xrightarrow{\mathrm{P}^{\theta_0}} 0. \tag{4.34}$$

The convergences (4.30), (4.31), (4.32) and (4.33) are treated similarly as for the terms $Z_{k,n}^{1,\ell}$ and $Z_{k,n}^{2,\ell}$. We next treat (4.34). By the mean value theorem,

$$Z_{k,n}^{4,\ell} - \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)} \left[R_4^{\theta(\ell),k} \Big| Y_{t_{k+1}}^{\theta(\ell)} = X_{t_{k+1}} \right] = M_{k,n,1} + M_{k,n,2},$$

where

$$M_{k,n,1} := -\frac{\ell u \Delta_n}{\sqrt{n\Delta_n}} \frac{\partial_\theta b(\theta(\ell), X_{t_k})}{\sigma^2(X_{t_k})} \int_{t_k}^{t_{k+1}} \left(\partial_\theta b(\theta_0 + \frac{\ell u w}{\sqrt{n\Delta_n}}, X_s^{\theta_0}) - \partial_\theta b(\theta_0 + \frac{\ell u w}{\sqrt{n\Delta_n}}, X_{t_k}) \right) ds,$$

$$M_{k,n,2} := \Delta_n \frac{\partial_\theta b(\theta(\ell), X_{t_k})}{\sigma^2(X_{t_k})} \bigg(\int_{t_k}^{t_{k+1}} \left(b(\theta(\ell), X_s^{\theta_0}) - b(\theta(\ell), X_{t_k}) \right) ds$$
$$- \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)} \left[\int_{t_k}^{t_{k+1}} \left(b(\theta(\ell), Y_s^{\theta(\ell)}) - b(\theta(\ell), Y_{t_k}^{\theta(\ell)}) \right) ds \Big| Y_{t_{k+1}}^{\theta(\ell)} = X_{t_{k+1}} \right] \bigg),$$

for some $w \in (0, 1)$.

Using Lemma 4.2.2(i), we get that

$$\sum_{k=0}^{n-1} \frac{|u|}{\sqrt{n\Delta_n^3}} \mathrm{E}\left[\left|\int_0^1 M_{k,n,1} d\ell\right| \Big| \widehat{\mathcal{F}}_{t_k}\right] \leq \frac{Cu^2\sqrt{\Delta_n}}{n} \sum_{k=0}^{n-1} \left(1+|X_{t_k}|^q\right),$$

for some constants $C, q > 0$. Therefore, by Lemma 1.4.2, we conclude that as $n \to \infty$,

$$\sum_{k=0}^{n-1} \frac{u}{\sqrt{n\Delta_n^3}} \int_0^1 M_{k,n,1} d\ell \xrightarrow{\mathrm{P}^{\theta_0}} 0.$$

We next show that as $n \to \infty$,

$$\sum_{k=0}^{n-1} \frac{u}{\sqrt{n\Delta_n^3}} \int_0^1 M_{k,n,2} d\ell \xrightarrow{\mathrm{P}^{\theta_0}} 0.$$

Using Girsanov's theorem, and Lemmas 4.2.5 and 4.2.2(i), we obtain that

$$\begin{aligned}
&\left|\sum_{k=0}^{n-1} \frac{u}{\sqrt{n\Delta_n^3}} \int_0^1 \mathrm{E}\left[M_{k,n,2}|\widehat{\mathcal{F}}_{t_k}\right] d\ell\right| = \left|\sum_{k=0}^{n-1} \frac{u}{\sqrt{n\Delta_n}} \int_0^1 \frac{\partial_\theta b(\theta(\ell), X_{t_k})}{\sigma^2(X_{t_k})}\right.\\
&\times \left\{\int_{t_k}^{t_{k+1}} \mathrm{E}_{\widehat{Q}_k^{\theta(\ell),\theta_0}}\left[\left(b(\theta(\ell), X_s^{\theta(\ell)}) - b(\theta(\ell), X_{t_k}^{\theta(\ell)})\right)\left(\frac{d\widehat{\mathrm{P}}}{d\widehat{Q}_k^{\theta(\ell),\theta_0}} - 1\right) \Big| X_{t_k}\right] ds\right.\\
&\left.\left. - \mathrm{E}_{\widehat{Q}_k^{\theta(\ell),\theta_0}}\left[\int_{t_k}^{t_{k+1}} \left(b(\theta(\ell), Y_s^{\theta(\ell)}) - b(\theta(\ell), Y_{t_k}^{\theta(\ell)})\right) ds \left(\frac{d\widehat{\mathrm{P}}}{d\widehat{Q}_k^{\theta(\ell),\theta_0}} - 1\right) \Big| X_{t_k}\right]\right\} d\ell\right|\\
&\leq \frac{C|u|\Delta_n}{n} \sum_{k=0}^{n-1} \left(1+|X_{t_k}|^q\right),
\end{aligned}$$

for some constants $C, q > 0$, which shows Lemma 1.4.1(i).

Finally, proceeding as in the proof of Lemma 4.3.2, we get that condition (ii) of Lemma 1.4.1 holds. Thus, the result follows. □

Lemma 4.3.5. *Assume conditions* **(A1)**-**(A2)**, **(A4)**-**(A5)** *and* **(A7)**. *Then as* $n \to \infty$,

$$\begin{aligned}
\sum_{k=0}^{n-1} \frac{u}{\sqrt{n\Delta_n^3}} \int_0^1 &\left(Z_{k,n}^{6,\ell} \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)}\left[\mathbf{1}_{\{\widetilde{\tau}>n\Delta_n\}} \Big| Y_{t_{k+1}}^{\theta(\ell)} = X_{t_{k+1}}\right] \mathbf{1}_{\{\widehat{\tau}>n\Delta_n\}}\right.\\
&\left. - \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)}\left[R_6^{\theta(\ell),k} \mathbf{1}_{\{\widetilde{\tau}>n\Delta_n\}} \Big| Y_{t_{k+1}}^{\theta(\ell)} = X_{t_{k+1}}\right] \mathbf{1}_{\{\widehat{\tau}>n\Delta_n\}}\right) d\ell \xrightarrow{\mathrm{P}^{\theta_0}} 0.
\end{aligned}$$

Proof. Using the fact that $\mathbf{1}_{\{\widehat{\tau}>n\Delta_n\}} = 1 - \mathbf{1}_{\{\widehat{\tau}\leq n\Delta_n\}}$ and $\mathbf{1}_{\{\widetilde{\tau}>n\Delta_n\}} = 1 - \mathbf{1}_{\{\widetilde{\tau}\leq n\Delta_n\}}$, it suffices to show that as $n \to \infty$,

$$\sum_{k=0}^{n-1} \frac{u}{\sqrt{n\Delta_n^3}} \int_0^1 Z_{k,n}^{6,\ell} \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)} \left[\mathbf{1}_{\{\widehat{\tau} > n\Delta_n\}} \Big| Y_{t_{k+1}}^{\theta(\ell)} = X_{t_{k+1}} \right] \mathbf{1}_{\{\widehat{\tau} \leq n\Delta_n\}} d\ell \xrightarrow{\mathrm{P}^{\theta_0}} 0, \tag{4.35}$$

$$\sum_{k=0}^{n-1} \frac{u}{\sqrt{n\Delta_n^3}} \int_0^1 Z_{k,n}^{6,\ell} \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)} \left[\mathbf{1}_{\{\widehat{\tau} \leq n\Delta_n\}} \Big| Y_{t_{k+1}}^{\theta(\ell)} = X_{t_{k+1}} \right] d\ell \xrightarrow{\mathrm{P}^{\theta_0}} 0, \tag{4.36}$$

$$\sum_{k=0}^{n-1} \frac{u}{\sqrt{n\Delta_n^3}} \int_0^1 \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)} \left[R_6^{\theta(\ell),k} \mathbf{1}_{\{\widehat{\tau} > n\Delta_n\}} \Big| Y_{t_{k+1}}^{\theta(\ell)} = X_{t_{k+1}} \right] \mathbf{1}_{\{\widehat{\tau} \leq n\Delta_n\}} d\ell \xrightarrow{\mathrm{P}^{\theta_0}} 0, \tag{4.37}$$

$$\sum_{k=0}^{n-1} \frac{u}{\sqrt{n\Delta_n^3}} \int_0^1 \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)} \left[R_6^{\theta(\ell),k} \mathbf{1}_{\{\widehat{\tau} \leq n\Delta_n\}} \Big| Y_{t_{k+1}}^{\theta(\ell)} = X_{t_{k+1}} \right] d\ell \xrightarrow{\mathrm{P}^{\theta_0}} 0, \tag{4.38}$$

$$\sum_{k=0}^{n-1} \frac{u}{\sqrt{n\Delta_n}} \int_0^1 \frac{\partial_\theta b(\theta(\ell), X_{t_k})}{\sigma^2(X_{t_k})} \bigg(\int_{t_k}^{t_{k+1}} \int_I c(X_{s-}^{\theta_0}, z) \widetilde{N}(ds, dz) \tag{4.39}$$

$$- \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)} \left[\int_{t_k}^{t_{k+1}} \int_I c(Y_{s-}^{\theta(\ell)}, z) \widetilde{M}(ds, dz) \Big| Y_{t_{k+1}}^{\theta(\ell)} = X_{t_{k+1}} \right] \bigg) d\ell \xrightarrow{\mathrm{P}^{\theta_0}} 0. \tag{4.40}$$

The convergences (4.35)-(4.38) are treated similarly as for the terms $Z_{k,n}^{1,\ell}$ and $Z_{k,n}^{2,\ell}$. We next treat (4.40). By Girsanov's theorem,

$$\begin{aligned}
&\mathrm{E}\left[\int_{t_k}^{t_{k+1}} \int_I c(X_{s-}^{\theta_0}, z) \widetilde{N}(ds, dz) - \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)} \left[\int_{t_k}^{t_{k+1}} \int_I c(Y_{s-}^{\theta(\ell)}, z) \widetilde{M}(ds, dz) \Big| Y_{t_{k+1}}^{\theta(\ell)} = X_{t_{k+1}} \right] \Big| \widehat{\mathcal{F}}_{t_k} \right] \\
&= -\mathrm{E}_{\widehat{Q}_k^{\theta(\ell),\theta_0}} \left[\int_{t_k}^{t_{k+1}} \int_I c(Y_{s-}^{\theta(\ell)}, z) \widetilde{M}(ds, dz) \frac{d\widehat{\mathrm{P}}}{d\widehat{Q}_k^{\theta(\ell),\theta_0}} \Big| X_{t_k} \right] \\
&= 0,
\end{aligned}$$

where we have used the independence between $\int_{t_k}^{t_{k+1}} \int_I c(Y_{s-}^{\theta(\ell)}, z) \widetilde{M}(ds, dz)$ and $\frac{d\widehat{\mathrm{P}}}{d\widehat{Q}_k^{\theta(\ell),\theta_0}}$ together with the fact that $\mathrm{E}_{\widehat{Q}_k^{\theta(\ell),\theta_0}}[\int_{t_k}^{t_{k+1}} \int_I c(Y_{s-}^{\theta(\ell)}, z) \widetilde{M}(ds, dz)] = 0$. This shows that the term (i) of Lemma 1.4.1 is actually equal to 0 for all $n \geq 1$.

We next show that condition (ii) of Lemma 1.4.1 holds. Cauchy-Schwarz inequality gives

$$\begin{aligned}
&\sum_{k=0}^{n-1} \frac{u^2}{n\Delta_n} \mathrm{E}\bigg[\bigg(\int_0^1 \frac{\partial_\theta b(\theta(\ell), X_{t_k})}{\sigma^2(X_{t_k})} \bigg(\int_{t_k}^{t_{k+1}} \int_I c(X_{s-}^{\theta_0}, z) \widetilde{N}(ds, dz) \\
&- \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)} \left[\int_{t_k}^{t_{k+1}} \int_I c(Y_{s-}^{\theta(\ell)}, z) \widetilde{M}(ds, dz) \Big| Y_{t_{k+1}}^{\theta(\ell)} = X_{t_{k+1}} \right] \bigg) d\ell \bigg)^2 \Big| \widehat{\mathcal{F}}_{t_k} \bigg] \leq 3\left(D_1 + D_2 + D_3 \right),
\end{aligned}$$

where

$$
\begin{aligned}
D_1 &= \frac{u^2}{n\Delta_n}\sum_{k=0}^{n-1}\int_0^1 \left(\frac{\partial_\theta b(\theta(\ell),X_{t_k})}{\sigma^2(X_{t_k})}\right)^2 \\
&\quad\times \mathrm{E}\left[\left(\int_{t_k}^{t_{k+1}}\int_I \left(c(X_{s-}^{\theta_0},z)-c(X_{t_k},z)\right)\widetilde{N}(ds,dz)\right)^2\Big|X_{t_k}\right]d\ell,\\
D_2 &= \frac{u^2}{n\Delta_n}\sum_{k=0}^{n-1}\int_0^1 \left(\frac{\partial_\theta b(\theta(\ell),X_{t_k})}{\sigma^2(X_{t_k})}\right)^2 \\
&\quad\times \mathrm{E}\left[\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)}\left[\left(\int_{t_k}^{t_{k+1}}\int_I \left(c(Y_{s-}^{\theta(\ell)},z)-c(Y_{t_k}^{\theta(\ell)},z)\right)\widetilde{M}(ds,dz)\right)^2\Big|Y_{t_{k+1}}^{\theta(\ell)}=X_{t_{k+1}}\right]\Big|X_{t_k}\right]d\ell,\\
D_3 &= \frac{u^2}{n\Delta_n}\sum_{k=0}^{n-1}\int_0^1 \left(\frac{\partial_\theta b(\theta(\ell),X_{t_k})}{\sigma^2(X_{t_k})}\right)^2 \mathrm{E}\Bigg[\Bigg(\int_{t_k}^{t_{k+1}}\int_I c(X_{t_k},z)N(ds,dz)\\
&\quad-\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)}\left[\int_{t_k}^{t_{k+1}}\int_I c(Y_{t_k}^{\theta(\ell)},z)M(ds,dz)\Big|Y_{t_{k+1}}^{\theta(\ell)}=X_{t_{k+1}}\right]\Bigg)^2\Big|X_{t_k}\Bigg]d\ell.
\end{aligned}
$$

Using Burkhölder's inequality, the Lipschitz property of c and Lemma 4.2.2(i), together with hypotheses **(A1)-(A2)** and **(A4)-(A5)**, we get that for some constants $C,q>0$,

$$
D_1 \leq \frac{Cu^2\Delta_n}{n}\sum_{k=0}^{n-1}\left(1+|X_{t_k}|^q\right).
$$

Moreover, using Girsanov's theorem, Burkhölder's inequality, Lemmas 4.2.5 and 4.2.2(i), together with hypotheses **(A1)-(A2)** and **(A4)-(A5)**, we obtain that for some constants $C,q>0$,

$$
\begin{aligned}
D_2 &\leq \frac{u^2}{n\Delta_n}\sum_{k=0}^{n-1}\int_0^1 \left(\frac{\partial_\theta b(\theta(\ell),X_{t_k})}{\sigma^2(X_{t_k})}\right)^2 \Bigg\{\mathrm{E}_{\widehat{Q}_k^{\theta(\ell),\theta_0}}\left[\left(\int_{t_k}^{t_{k+1}}\int_I \left(c(Y_{s-}^{\theta(\ell)},z)-c(Y_{t_k}^{\theta(\ell)},z)\right)\widetilde{M}(ds,dz)\right)^2\Big|X_{t_k}\right]\\
&\quad+\left|\mathrm{E}_{\widehat{Q}_k^{\theta(\ell),\theta_0}}\left[\left(\int_{t_k}^{t_{k+1}}\int_I \left(c(Y_{s-}^{\theta(\ell)},z)-c(Y_{t_k}^{\theta(\ell)},z)\right)\widetilde{M}(ds,dz)\right)^2\left(\frac{d\widehat{\mathrm{P}}}{d\widehat{Q}_k^{\theta(\ell),\theta_0}}-1\right)\Big|X_{t_k}\right]\right|\Bigg\}d\ell\\
&\leq C\left(\Delta_n+\frac{1}{\sqrt{n}}\right)\frac{u^2}{n}\sum_{k=0}^{n-1}\left(1+|X_{t_k}|^q\right).
\end{aligned}
$$

Again, Girsanov's theorem yields $D_3=D_{3,1}+D_{3,2}$, where

$$
\begin{aligned}
D_{3,1} &:= \frac{u^2}{n\Delta_n}\sum_{k=0}^{n-1}\int_0^1 \left(\frac{\partial_\theta b(\theta(\ell),X_{t_k})}{\sigma^2(X_{t_k})}\right)^2 \mathrm{E}_{\widehat{Q}_k^{\theta(\ell),\theta_0}}\Bigg[\Bigg(\int_{t_k}^{t_{k+1}}\int_I c(X_{t_k},z)N(ds,dz)\\
&\quad-\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)}\left[\int_{t_k}^{t_{k+1}}\int_I c(Y_{t_k}^{\theta(\ell)},z)M(ds,dz)\Big|Y_{t_{k+1}}^{\theta(\ell)}=X_{t_{k+1}}\right]\Bigg)^2\left(\frac{d\widehat{\mathrm{P}}}{d\widehat{Q}_k^{\theta(\ell),\theta_0}}-1\right)\Big|X_{t_k}\Bigg]d\ell,\\
D_{3,2} &:= \frac{u^2}{n\Delta_n}\sum_{k=0}^{n-1}\int_0^1 \left(\frac{\partial_\theta b(\theta(\ell),X_{t_k})}{\sigma^2(X_{t_k})}\right)^2 \mathrm{E}_{\widehat{Q}_k^{\theta(\ell),\theta_0}}\Bigg[\Bigg(\int_{t_k}^{t_{k+1}}\int_I c(X_{t_k},z)N(ds,dz)\\
&\quad-\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)}\left[\int_{t_k}^{t_{k+1}}\int_I c(Y_{t_k}^{\theta(\ell)},z)M(ds,dz)\Big|Y_{t_{k+1}}^{\theta(\ell)}=X_{t_{k+1}}\right]\Bigg)^2\Big|X_{t_k}\Bigg]d\ell.
\end{aligned}
$$

Observe that $|D_{3,1}| \leq 2(D_{3,1,1}+D_{3,1,2})$, where

$$D_{3,1,1} := \frac{u^2}{n\Delta_n}\sum_{k=0}^{n-1}\int_0^1\left(\frac{\partial_\theta b(\theta(\ell),X_{t_k})}{\sigma^2(X_{t_k})}\right)^2$$
$$\times\left|\mathrm{E}_{\widehat{Q}_k^{\theta(\ell),\theta_0}}\left[\left(\int_{t_k}^{t_{k+1}}\int_I c(X_{t_k},z)N(ds,dz)\right)^2\left(\frac{d\widehat{\mathrm{P}}}{d\widehat{Q}_k^{\theta(\ell),\theta_0}}-1\right)\Big|X_{t_k}\right]\right|d\ell,$$
$$D_{3,1,2} := \frac{u^2}{n\Delta_n}\sum_{k=0}^{n-1}\int_0^1\left(\frac{\partial_\theta b(\theta(\ell),X_{t_k})}{\sigma^2(X_{t_k})}\right)^2$$
$$\times\left|\mathrm{E}_{\widehat{Q}_k^{\theta(\ell),\theta_0}}\left[\left(\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)}\left[\int_{t_k}^{t_{k+1}}\int_I c(Y_{t_k}^{\theta(\ell)},z)M(ds,dz)\Big|Y_{t_{k+1}}^{\theta(\ell)}=X_{t_{k+1}}\right]\right)^2\left(\frac{d\widehat{\mathrm{P}}}{d\widehat{Q}_k^{\theta(\ell),\theta_0}}-1\right)\Big|X_{t_k}\right]\right|d\ell.$$

Using the same arguments as for the term D_2, we get that for some constants $C,q>0$,

$$D_{3,1,1} \leq \frac{C}{\sqrt{n\Delta_n}}\frac{u^2}{n}\sum_{k=0}^{n-1}\left(1+|X_{t_k}|^q\right).$$

Applying Lemma 4.2.5, Jensen's inequality and **(A1)**, **(A5)**, we obtain that

$$D_{3,1,2} \leq \frac{Cu^2}{n\Delta_n\sqrt{n}}\sum_{k=0}^{n-1}\left(1+|X_{t_k}|^q\right)$$
$$\times\int_0^1\left(\widehat{\mathrm{E}}^{\alpha}\left[\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)}\left[\left(\int_{t_k}^{t_{k+1}}\int_I c(Y_{t_k}^{\theta(\ell)},z)M(ds,dz)\right)^4\Big|Y_{t_{k+1}}^{\theta(\ell)}=X_{t_{k+1}}\right]\Big|X_{t_k}\right]\right)^{1/2}d\ell$$
$$\leq \frac{C}{\sqrt{n\Delta_n}}\frac{u^2}{n}\sum_{k=0}^{n-1}\left(1+|X_{t_k}|^q\right),$$

for some constants $C,q>0$.

Next, hypotheses **(A2)** and **(A4)**(b) yield that

$$D_{3,2} \leq \frac{Cu^2}{n\Delta_n}\sum_{k=0}^{n-1}\left(1+|X_{t_k}|^q\right)\int_0^1\mathrm{E}_{\widehat{Q}_k^{\theta(\ell),\theta_0}}\left[\left(\int_{t_k}^{t_{k+1}}\int_I c(X_{t_k},z)N(ds,dz)\right.\right.$$
$$\left.\left.-\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)}\left[\int_{t_k}^{t_{k+1}}\int_I c(Y_{t_k}^{\theta(\ell)},z)M(ds,dz)\Big|Y_{t_{k+1}}^{\theta(\ell)}=X_{t_{k+1}}\right]\right)^2\Big|X_{t_k}\right]d\ell,$$

for some constants $C,q>0$.

Multiplying the random variable inside the expectation by $\left(\mathbf{1}_{\widehat{J}_{0,k}}+\mathbf{1}_{\widehat{J}_{1,k}}+\mathbf{1}_{\widehat{J}_{2,k}}\right)$ and applying Lemma 4.2.8, we get that for any $\alpha\in(\upsilon,\frac{1}{2})$ and $\alpha_0\in(\frac{1}{4},\frac{1}{2})$,

$$D_{3,2} \leq \frac{Cu^2}{n\Delta_n}\sum_{k=0}^{n-1}\left(1+|X_{t_k}|^q\right)\int_0^1\left(M_0^{\theta(\ell)}+M_1^{\theta(\ell)}+M_2^{\theta(\ell)}\right)d\ell$$
$$\leq C\left(\lambda_n\sqrt{\Delta_n}+\Delta_n^{-\frac{3}{2}-3\gamma}e^{-C_0\Delta_n^{2(\alpha\vee\alpha_0)-1}}\right)\frac{u^2}{n}\sum_{k=0}^{n-1}\left(1+|X_{t_k}|^q\right)e^{c_0\Delta_n^{1-2\gamma}X_{t_k}^2},$$

for some constants $c_0,C_0,C,q>0$. By hypothesis **(A8)**, $D_{3,2}$ converges to zero in P^{θ_0}-probability as $n\to\infty$. The desired proof is now finished. □

4.3.3 Main contributions : LAN property

Proof. We write $\xi_{k,n} = \xi_{k,n,1} - \xi_{k,n,2} - \xi_{k,n,3}$, where

$$\xi_{k,n,1} = \frac{u}{\sqrt{n\Delta_n}} \int_0^1 \frac{\partial_\theta b(\theta(\ell), X_{t_k})}{\sigma^2(X_{t_k})} \left(\sigma(X_{t_k}) \left(B_{t_{k+1}} - B_{t_k}\right) + \left(b(\theta_0, X_{t_k}) - b(\theta(\ell), X_{t_k})\right) \Delta_n\right) d\ell,$$

$$\xi_{k,n,2} = \frac{u}{\sqrt{n\Delta_n}} \int_0^1 \frac{\partial_\theta b(\theta(\ell), X_{t_k})}{\sigma^2(X_{t_k})} \left(\sigma(X_{t_k}) \left(B_{t_{k+1}} - B_{t_k}\right) + \left(b(\theta_0, X_{t_k}) - b(\theta(\ell), X_{t_k})\right) \Delta_n\right) \times \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)} \left[\mathbf{1}_{\{\widetilde{\tau} > n\Delta_n\}} \Big| Y_{t_{k+1}}^{\theta(\ell)} = X_{t_{k+1}}\right] \mathbf{1}_{\{\widehat{\tau} \leq n\Delta_n\}} d\ell,$$

$$\xi_{k,n,3} = \frac{u}{\sqrt{n\Delta_n}} \int_0^1 \frac{\partial_\theta b(\theta(\ell), X_{t_k})}{\sigma^2(X_{t_k})} \left(\sigma(X_{t_k}) \left(B_{t_{k+1}} - B_{t_k}\right) + \left(b(\theta_0, X_{t_k}) - b(\theta(\ell), X_{t_k})\right) \Delta_n\right) \times \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell)} \left[\mathbf{1}_{\{\widetilde{\tau} \leq n\Delta_n\}} \Big| Y_{t_{k+1}}^{\theta(\ell)} = X_{t_{k+1}}\right] d\ell.$$

First, proceeding as for the terms $Z_{k,n}^{1,\ell}$ and $Z_{k,n}^{2,\ell}$, we get that as $n \to \infty$,

$$\sum_{k=0}^{n-1} \left(\xi_{k,n,2} + \xi_{k,n,3}\right) \overset{\mathrm{P}^{\theta_0}}{\longrightarrow} 0.$$

Next, applying Lemma 1.4.3 to $\xi_{k,n,1}$, we need to consider $\mathrm{E}^{\theta_0}[\xi_{k,n,1}^r | \widehat{\mathcal{F}}_{t_k}]$ for $r = 1, 2$ and 4 but this conditional expectation equals $\mathrm{E}[\xi_{k,n,1}^r | \widehat{\mathcal{F}}_{t_k}]$. Therefore, it suffices to show that as $n \to \infty$:

$$\sum_{k=0}^{n-1} \mathrm{E}\left[\xi_{k,n,1} | \widehat{\mathcal{F}}_{t_k}\right] \overset{\mathrm{P}^{\theta_0}}{\longrightarrow} -\frac{u^2}{2}\Gamma(\theta_0), \tag{4.41}$$

$$\sum_{k=0}^{n-1} \left(\mathrm{E}\left[\xi_{k,n,1}^2 | \widehat{\mathcal{F}}_{t_k}\right] - \left(\mathrm{E}\left[\xi_{k,n,1} | \widehat{\mathcal{F}}_{t_k}\right]\right)^2\right) \overset{\mathrm{P}^{\theta_0}}{\longrightarrow} u^2\Gamma(\theta_0), \tag{4.42}$$

$$\sum_{k=0}^{n-1} \mathrm{E}\left[\xi_{k,n,1}^4 | \widehat{\mathcal{F}}_{t_k}\right] \overset{\mathrm{P}^{\theta_0}}{\longrightarrow} 0, \tag{4.43}$$

where

$$\Gamma(\theta_0) = \int_{\mathbb{R}} \left(\frac{\partial_\theta b(\theta_0, x)}{\sigma(x)}\right)^2 \pi_{\theta_0}(dx).$$

Proof of (4.41). Since $\mathrm{E}[B_{t_{k+1}} - B_{t_k} | \widehat{\mathcal{F}}_{t_k}] = 0$, we get that

$$\begin{aligned}\sum_{k=0}^{n-1} \mathrm{E}\left[\xi_{k,n,1} | \widehat{\mathcal{F}}_{t_k}\right] &= -\frac{u^2}{n} \sum_{k=0}^{n-1} \int_0^1 \ell \frac{\partial_\theta b(\theta(\ell), X_{t_k})}{\sigma^2(X_{t_k})} \partial_\theta b(\theta_0 + \frac{\ell u v}{\sqrt{n\Delta_n}}, X_{t_k}) d\ell \\ &= -\frac{u^2}{2n} \sum_{k=0}^{n-1} \left(\frac{\partial_\theta b(\theta_0, X_{t_k})}{\sigma(X_{t_k})}\right)^2 - H_1 - H_2,\end{aligned}$$

where $v \in (0,1)$, $H_1 = \sum_{k=0}^{n-1} H_{k,n}$, and

$$H_{k,n} := \frac{u^2}{n} \int_0^1 \ell \frac{\partial_\theta b(\theta(\ell), X_{t_k})}{\sigma^2(X_{t_k})} \left(\partial_\theta b(\theta_0 + \frac{\ell u v}{\sqrt{n\Delta_n}}, X_{t_k}) - \partial_\theta b(\theta_0, X_{t_k})\right) d\ell,$$

$$H_2 := \frac{u^2}{n} \sum_{k=0}^{n-1} \int_0^1 \ell \frac{\partial_\theta b(\theta_0, X_{t_k})}{\sigma^2(X_{t_k})} \left(\partial_\theta b(\theta(\ell), X_{t_k}) - \partial_\theta b(\theta_0, X_{t_k})\right) d\ell.$$

Using hypotheses **(A2)** and **(A4)**(b), (c), we have that for some constants $C, \epsilon, q > 0$,

$$\sum_{k=0}^{n-1} \mathrm{E}\left[|H_{k,n}| | \widehat{\mathcal{F}}_{t_k}\right] \leq \frac{C|u|^{\epsilon+2}|v|^\epsilon}{(\sqrt{n\Delta_n})^\epsilon} \frac{1}{n} \sum_{k=0}^{n-1} \left(1 + |X_{t_k}|^q\right),$$

which, by Lemma 1.4.2, implies that $H_1 \xrightarrow{\mathrm{P}^{\theta_0}} 0$ as $n \to \infty$. Thus, so does H_2 by using the same argument. On the other hand, applying Lemma 4.2.9, we obtain that as $n \to \infty$,

$$\frac{1}{n}\sum_{k=0}^{n-1}\left(\frac{\partial_\theta b(\theta_0, X_{t_k})}{\sigma(X_{t_k})}\right)^2 \xrightarrow{\mathrm{P}^{\theta_0}} \Gamma(\theta_0), \tag{4.44}$$

which gives (4.41).

Proof of (4.42). First, from the previous computations, we have that

$$\begin{aligned}\sum_{k=0}^{n-1}\left(\mathrm{E}\left[\xi_{k,n,1}|\widehat{\mathcal{F}}_{t_k}\right]\right)^2 &= \frac{u^4}{n^2}\sum_{k=0}^{n-1}\left(\int_0^1 \ell\frac{\partial_\theta b(\theta(\ell), X_{t_k})}{\sigma^2(X_{t_k})}\partial_\theta b(\theta_0 + \frac{\ell uv}{\sqrt{n\Delta_n}}, X_{t_k})d\ell\right)^2\\ &\leq \frac{Cu^4}{n^2}\sum_{k=0}^{n-1}\left(1+|X_{t_k}|^q\right),\end{aligned}$$

for some constants $C, q > 0$, which converges to zero in P^{θ_0}-probability as $n \to \infty$.

Next, using properties of the moments of the Brownian motion, we can write

$$\sum_{k=0}^{n-1}\mathrm{E}\left[\xi_{k,n,1}^2|\widehat{\mathcal{F}}_{t_k}\right] = \frac{u^2}{n}\sum_{k=0}^{n-1}\left(\frac{\partial_\theta b(\theta_0, X_{t_k})}{\sigma(X_{t_k})}\right)^2 + H_3 + H_4 + H_5,$$

where

$$\begin{aligned}H_3 &:= \frac{2u^2}{n}\sum_{k=0}^{n-1}\frac{\partial_\theta b(\theta_0, X_{t_k})}{\sigma(X_{t_k})}\int_0^1 \frac{\partial_\theta b(\theta(\ell), X_{t_k}) - \partial_\theta b(\theta_0, X_{t_k})}{\sigma(X_{t_k})}d\ell,\\ H_4 &:= \frac{u^2}{n}\sum_{k=0}^{n-1}\left(\int_0^1 \frac{\partial_\theta b(\theta(\ell), X_{t_k}) - \partial_\theta b(\theta_0, X_{t_k})}{\sigma(X_{t_k})}d\ell\right)^2,\\ H_5 &:= \frac{u^2\Delta_n}{n}\sum_{k=0}^{n-1}\left(\int_0^1 \frac{\partial_\theta b(\theta(\ell), X_{t_k})}{\sigma^2(X_{t_k})}\left(b(\theta_0, X_{t_k}) - b(\theta(\ell), X_{t_k})\right)d\ell\right)^2.\end{aligned}$$

As for the term H_1, using hypotheses **(A2)** and **(A4)**(b), (c), we get that H_3, H_4, H_5 converge to zero in P^{θ_0}-probability as $n \to \infty$. Moreover, using again (4.44), we conclude (4.42).

Proof of (4.43). Basic computation yields

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

for some constants $C, q > 0$. The proof of Theorem 4.1.1 is now completed. □

4.4 Maximum likelihood estimator for Ornstein-Uhlenbeck process with jumps

Consider the Ornstein-Uhlenbeck process with jumps defined in Example 4.1.1 1)

$$X_t^\theta = x_0 - \theta\int_0^t X_s^\theta ds + \sigma B_t + \int_0^t\int_{\mathbb{R}_0} z\widetilde{N}(ds, dz), \tag{4.45}$$

where $\theta > 0$, $\sigma \in \mathbb{R}_0$ and the Lévy measure satisfies **(A5)**, **(A7)**, and is finite. Assume that there exists a constant $C > 0$ such that $\int_{\mathbb{R}_0} e^{Cz^2}\nu(dz) < \infty$.

By the Markov property, the log-likelihood function based on X^n can be written as follows

$$\ell_n(\theta_0) = \log p(X^n; \theta_0) = \sum_{k=0}^{n-1} \log p^{\theta_0}(\Delta_n, X_{t_k}, X_{t_{k+1}}). \tag{4.46}$$

The maximum likelihood estimator $\widehat{\theta}_n$ of θ_0 is defined as the solution to the likelihood equation $\partial_\theta \ell_n(\theta_0) = 0$.

Theorem 4.4.1. *Assume conditions* **(A5)** *and* **(A7)**. *Then, the maximum likelihood estimators* $\widehat{\theta}_n$ *of* θ_0 *are consistent and asymptotically efficient. That is, as* $n \to \infty$,

$$\widehat{\theta}_n \overset{\mathrm{P}^{\theta_0}}{\longrightarrow} \theta_0,$$

and

$$\sqrt{n\Delta_n}(\widehat{\theta}_n - \theta_0) \overset{\mathcal{L}(\mathrm{P}^{\theta_0})}{\longrightarrow} \mathcal{N}\left(0, \Gamma(\theta_0)^{-1}\right),$$

where

$$\Gamma(\theta_0) = \frac{1}{2\theta_0}\left(1 + \frac{1}{\sigma^2}\int_{\mathbb{R}_0} z^2 \nu(dz)\right).$$

Proof. Using (4.46) and Proposition 4.2.1, the likelihood equation is equivalent to

$$\sum_{k=0}^{n-1} \frac{1}{\Delta_n} \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_0}\left[\delta\left(\partial_\theta Y_{t_{k+1}}^{\theta_0}(t_k, X_{t_k}) U^{\theta_0}(t_k, X_{t_k})\right) \Big| Y_{t_{k+1}}^{\theta_0} = X_{t_{k+1}}\right] = 0, \tag{4.47}$$

where

$$Y_t^\theta(s,x) = x - \theta \int_s^t Y_u^\theta(s,x) du + \sigma \int_s^t dW_u + \int_s^t \int_{\mathbb{R}_0} z\widetilde{M}(ds,dz),$$
$$Y_t^\theta = x_0 - \theta \int_0^t Y_u^\theta du + \sigma W_t + \int_0^t \int_{\mathbb{R}_0} z\widetilde{M}(ds,dz).$$

From Lemma 4.2.1,

$$\begin{aligned}\delta\left(\partial_\theta Y_{t_{k+1}}^{\theta_0}(t_k, X_{t_k}) U^{\theta_0}(t_k, X_{t_k})\right) &= -\Delta_n Y_{t_k}^{\theta_0} \sigma^{-2}\left(Y_{t_{k+1}}^{\theta_0} - Y_{t_k}^{\theta_0} + \theta_0 Y_{t_k}^{\theta_0}\Delta_n\right) \\ &\quad + R_1^{\theta_0,k} + R_2^{\theta_0,k} + R_3^{\theta_0,k} - R_4^{\theta_0,k} - R_6^{\theta_0,k},\end{aligned} \tag{4.48}$$

where

$$\begin{aligned}
R_1^{\theta_0,k} &:= -\sigma^{-1} \int_{t_k}^{t_{k+1}} D_s \left(\frac{\partial_\theta Y_{t_{k+1}}^{\theta_0}(t_k, X_{t_k})}{\partial_x Y_{t_{k+1}}^{\theta_0}(t_k, X_{t_k})}\right) \partial_x Y_s^{\theta_0}(t_k, X_{t_k}) ds, \\
R_2^{\theta_0,k} &:= -\sigma^{-1} \int_{t_k}^{t_{k+1}} \frac{Y_s^{\theta_0}(t_k, X_{t_k})}{\partial_x Y_s^{\theta_0}(t_k, X_{t_k})} ds \int_{t_k}^{t_{k+1}} \left(\partial_x Y_s^{\theta_0}(t_k, X_{t_k}) - \partial_x Y_{t_k}^{\theta_0}(t_k, X_{t_k})\right) dW_s, \\
R_3^{\theta_0,k} &:= -\sigma^{-1} \int_{t_k}^{t_{k+1}} \left(\frac{Y_s^{\theta_0}(t_k, X_{t_k})}{\partial_x Y_s^{\theta_0}(t_k, X_{t_k})} - \frac{Y_{t_k}^{\theta_0}(t_k, X_{t_k})}{\partial_x Y_{t_k}^{\theta_0}(t_k, X_{t_k})}\right) ds \int_{t_k}^{t_{k+1}} \partial_x Y_{t_k}^{\theta_0}(t_k, X_{t_k}) dW_s, \\
R_4^{\theta_0,k} &:= \Delta_n Y_{t_k}^{\theta_0} \sigma^{-2} \theta_0 \int_{t_k}^{t_{k+1}} \left(Y_s^{\theta_0} - Y_{t_k}^{\theta_0}\right) ds, \\
R_6^{\theta_0,k} &:= -\Delta_n Y_{t_k}^{\theta_0} \sigma^{-2} \int_{t_k}^{t_{k+1}} \int_{\mathbb{R}_0} z\widetilde{M}(ds,dz).
\end{aligned}$$

Plugging (4.48) into (4.47), taking the conditional expectation, and using equation (4.45), we obtain that

$$\widehat{\theta}_n = \frac{\sum_{k=0}^{n-1}\left(-X_{t_k}\left(X_{t_{k+1}}-X_{t_k}\right)+\frac{\sigma^2}{\Delta_n}\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_0}\left[R_1^{\theta_0,k}+R_2^{\theta_0,k}+R_3^{\theta_0,k}-R_4^{\theta_0,k}-R_6^{\theta_0,k}\Big|Y_{t_{k+1}}^{\theta_0}=X_{t_{k+1}}\right]\right)}{\Delta_n\sum_{k=0}^{n-1}X_{t_k}^2}$$

$$= \theta_0 + \frac{-\frac{1}{\sigma n\Delta_n}\sum_{k=0}^{n-1}X_{t_k}\left(B_{t_{k+1}}-B_{t_k}\right)+S_1+S_2-S_3}{\frac{1}{\sigma^2 n}\sum_{k=0}^{n-1}X_{t_k}^2},$$

where

$$S_1 = \frac{1}{n\Delta_n^2}\sum_{k=0}^{n-1}\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_0}\left[R_1^{\theta_0,k}+R_2^{\theta_0,k}+R_3^{\theta_0,k}\Big|Y_{t_{k+1}}^{\theta_0}=X_{t_{k+1}}\right],$$

$$S_2 = \frac{\theta_0}{\sigma^2 n\Delta_n}\sum_{k=0}^{n-1}X_{t_k}\left(\int_{t_k}^{t_{k+1}}\left(X_s^{\theta_0}-X_{t_k}^{\theta_0}\right)ds-\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_0}\left[\int_{t_k}^{t_{k+1}}\left(Y_s^{\theta_0}-Y_{t_k}^{\theta_0}\right)ds\Big|Y_{t_{k+1}}^{\theta_0}=X_{t_{k+1}}\right]\right),$$

$$S_3 = \frac{1}{\sigma^2 n\Delta_n}\sum_{k=0}^{n-1}X_{t_k}\left(\int_{t_k}^{t_{k+1}}\int_{\mathbb{R}_0}z\widetilde{N}(ds,dz)-\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_0}\left[\int_{t_k}^{t_{k+1}}\int_{\mathbb{R}_0}z\widetilde{M}(ds,dz)\Big|Y_{t_{k+1}}^{\theta_0}=X_{t_{k+1}}\right]\right).$$

Using the ergodicity property and applying Lemma 4.2.9, we obtain that as $n\to\infty$,

$$\frac{1}{\sigma^2 n}\sum_{k=0}^{n-1}X_{t_k}^2 \xrightarrow{\mathrm{P}^{\theta_0}} \Gamma(\theta_0), \tag{4.49}$$

On the other hand, it can be checked that under conditions **(A5)** and **(A7)**, $S_1, S_2, S_3 \xrightarrow{\mathrm{P}^{\theta_0}} 0$ as $n\to\infty$. Moreover, applying Lemma 1.4.1, we get that as $n\to\infty$,

$$\frac{1}{\sigma n\Delta_n}\sum_{k=0}^{n-1}X_{t_k}\left(B_{t_{k+1}}-B_{t_k}\right)\xrightarrow{\mathrm{P}^{\theta_0}} 0.$$

Therefore, we have shown that $\widehat{\theta}_n \xrightarrow{\mathrm{P}^{\theta_0}} \theta_0$ as $n\to\infty$.

Next, we can write

$$\sqrt{n\Delta_n}(\widehat{\theta}_n-\theta_0) = \frac{-\frac{1}{\sigma\sqrt{n\Delta_n}}\sum_{k=0}^{n-1}X_{t_k}\left(B_{t_{k+1}}-B_{t_k}\right)+\sqrt{n\Delta_n}\left(S_1+S_2-S_3\right)}{\frac{1}{\sigma^2 n}\sum_{k=0}^{n-1}X_{t_k}^2}.$$

Then, using Lemma 1.4.3 and (4.49), we obtain that as $n\to\infty$,

$$\frac{1}{\sigma\sqrt{n\Delta_n}}\sum_{k=0}^{n-1}X_{t_k}\left(B_{t_{k+1}}-B_{t_k}\right)\xrightarrow{\mathcal{L}(\mathrm{P}^{\theta_0})}\mathcal{N}\left(0,\Gamma(\theta_0)\right).$$

This, together with (4.49) and the fact that under conditions **(A5)** and **(A7)** as $n\to\infty$,

$$\sqrt{n\Delta_n}\left(S_1+S_2-S_3\right)\xrightarrow{\mathrm{P}^{\theta_0}} 0,$$

concludes that $n\to\infty$,

$$\sqrt{n\Delta_n}(\widehat{\theta}_n-\theta_0)\xrightarrow{\mathcal{L}(\mathrm{P}^{\theta_0})}\mathcal{N}\left(0,\Gamma(\theta_0)^{-1}\right).$$

The proof of Theorem 4.4.1 is now completed. □

Chapitre 5

LAN property for a jump-diffusion process : drift and diffusion parameters

In this chapter, we consider an ergodic diffusion process with jumps driven by a Brownian motion and a Poisson random measure associated with a compensated compound Poisson process, whose drift and diffusion coefficients depend on unknown parameters. Supposing that the process is observed discretely at high frequency, we derive the local asymptotic normality (LAN) property. In order to obtain this result, Malliavin calculus and Girsanov's theorem are applied in order to write the log-likelihood ratio in terms of sums of conditional expectations, for which a central limit theorem for triangular arrays can be applied.

5.1 Introduction and main result

On a complete probability space $(\Omega, \mathcal{F}, \mathrm{P})$ defined in Definition 1.1.3, we consider the process $X^{\theta,\beta} = (X_t^{\theta,\beta})_{t\geq 0}$ solution to the following stochastic differential equation with jumps

$$dX_t^{\theta,\beta} = b(\theta, X_t^{\theta,\beta})dt + \sigma(\beta, X_t^{\theta,\beta})dB_t + \int_{\mathbb{R}_0} z\left(N(dt,dz) - \nu(dz)dt\right), \tag{5.1}$$

where $X_0^{\theta,\beta} = x_0 \in \mathbb{R}$, $\mathbb{R}_0 := \mathbb{R}\setminus\{0\}$, $B = (B_t)_{t\geq 0}$ is a standard Brownian motion, and $N(dt,dz)$ is a Poisson random measure in $(\mathbb{R}_+\times\mathbb{R}_0, \mathcal{B}(\mathbb{R}_+\times\mathbb{R}_0))$ independent of B, with intensity measure $\nu(dz)dt$, and finite Lévy measure $\lambda = \int_{\mathbb{R}_0}\nu(dz) < \infty$. The compensated Poisson random measure is denoted by $\widetilde{N}(dt,dz) := N(dt,dz) - \nu(dz)dt$. Let $\widehat{Z} = (\widehat{Z}_t)_{t\geq 0}$ be a compensated compound Poisson process associated with $N(dt,dz)$, i.e., $\widehat{Z}_t = \int_0^t\int_{\mathbb{R}_0} z(N(ds,dz)-\nu(dz)ds)$, for $t\geq 0$. The random variable $\widehat{\Lambda}$ that describes the jump sizes of $\widehat{Z}$ takes values in $A = \{a_i, i\in\mathbb{N}\}$, $a_i\in\mathbb{R}_0$, and has distribution $\mu(dz) = \frac{\nu(dz)}{\lambda} = \sum_{i=1}^{\infty} p_{a_i}\delta_{a_i}(dz)$, where $0\leq p_{a_i}\leq 1$, and $\sum_{i=1}^{\infty} p_{a_i} = 1$. Let $\{\widehat{\mathcal{F}}_t\}_{t\geq 0}$ denote the natural filtration generated by B and N. The unknown parameters (θ,β) belong to $\Theta\times\Sigma$ which is an open rectangle of $\mathbb{R}^2$. The coefficients $b:\Theta\times\mathbb{R}\to\mathbb{R}$ and $\sigma:\Sigma\times\mathbb{R}\to\mathbb{R}$ are measurable functions satisfying condition **(A1)** below under which equation (5.1) has a unique $\widehat{\mathcal{F}}_t$-adapted càdlàg solution $X^{\theta,\beta}$. We denote by $\mathrm{P}^{\theta,\beta}$ the probability law induced by $X^{\theta,\beta}$, and by $\mathrm{E}^{\theta,\beta}$ the expectation with respect to $\mathrm{P}^{\theta,\beta}$. Let $\stackrel{\mathrm{P}^{\theta,\beta}}{\longrightarrow}$ and $\stackrel{\mathcal{L}(\mathrm{P}^{\theta,\beta})}{\longrightarrow}$ denote the convergence in $\mathrm{P}^{\theta,\beta}$-probability and in $\mathrm{P}^{\theta,\beta}$-law, respectively.

Recall that the structure of the probability space is given by $\widehat{\Omega} = \Omega^1\times\Omega^2$, $\widetilde{\Omega} = \Omega^3\times\Omega^4$, $\widehat{\mathcal{F}} = \mathcal{F}^1\otimes\mathcal{F}^2$, $\widetilde{\mathcal{F}} = \mathcal{F}^3\otimes\mathcal{F}^4$, $\widehat{\mathrm{P}} = \mathrm{P}^1\otimes\mathrm{P}^2$, $\widetilde{\mathrm{P}} = \mathrm{P}^3\otimes\mathrm{P}^4$, and $\Omega = \widehat{\Omega}\times\widetilde{\Omega}$, $\mathcal{F} = \widehat{\mathcal{F}}\otimes\widetilde{\mathcal{F}}$, $\mathrm{P} = \widehat{\mathrm{P}}\otimes\widetilde{\mathrm{P}}$. We denote by E, $\widehat{\mathrm{E}}$, $\widetilde{\mathrm{E}}$ the expectation with respect to P, $\widehat{\mathrm{P}}$ and $\widetilde{\mathrm{P}}$, respectively.

For fixed $(\theta_0,\beta_0)\in\Theta\times\Sigma$ and $n\geq 1$, we consider a discrete observation scheme at equidistant times $t_k = k\Delta_n$, $k\in\{0,...,n\}$ of the diffusion process X^{θ_0,β_0}, which is denoted by $X^n = (X_{t_0}, X_{t_1}, ..., X_{t_n})$, where $\Delta_n\leq 1$. We assume that the sequence of time-step sizes Δ_n satisfies the high-frequency observation condition

$$n\Delta_n \to \infty, \quad \text{and} \quad \Delta_n \to 0, \quad \text{as } n \to \infty.$$

We consider the following hypotheses on equation (5.1).

(A1) For any $(\theta, \beta) \in \Theta \times \Sigma$, there exists a constant $C > 0$ such that for all $x, y \in \mathbb{R}$,

$$|b(\theta, x) - b(\theta, y)| + |\sigma(\beta, x) - \sigma(\beta, y)| \leq C|x - y|.$$

(A2) For any $(\theta, \beta) \in \Theta \times \Sigma$, there exist constants $C > 0$ and $c \geq 1$ such that for all $x \in \mathbb{R}$,

$$|b(\theta, x)| \leq C, \quad \text{and} \quad \frac{1}{c} \leq |\sigma(\beta, x)| \leq c.$$

(A3) The functions b and σ are of class C^1 w.r.t. θ, β, and x. Each partial derivative $\partial_\theta b$, $\partial_x b$, $\partial_\beta \sigma$ and $\partial_x \sigma$ is of class C^1 w.r.t. x. Moreover, there exist positive constants C, q, ϵ, η, independent of $(\theta, \theta_1, \theta_2, \beta, \beta_1, \beta_2, x, y) \in \Theta^3 \times \Sigma^3 \times \mathbb{R}^2$ such that
(a) $|\partial_x b(\theta, x)| + |\partial_x \sigma(\beta, x)| \leq C$;
(b) $|h(\cdot, x)| \leq C(1 + |x|^q)$ for $h = \partial_\theta b, \partial_x^2 b, \partial_{x,\theta}^2 b$, $\partial_\beta \sigma$, $\partial_x^2 \sigma$ or $\partial_{x,\beta}^2 \sigma$;
(c) $|\partial_\theta b(\theta_1, x) - \partial_\theta b(\theta_2, x)| \leq C|\theta_1 - \theta_2|^\epsilon (1 + |x|^q)$;
(d) $|\partial_\beta \sigma(\beta_1, x) - \partial_\beta \sigma(\beta_2, x)| \leq C|\beta_1 - \beta_2|^\epsilon (1 + |x|^q)$;
(e) $|\partial_\theta b(\theta, x) - \partial_\theta b(\theta, y)| + |\partial_\beta \sigma(\beta, x) - \partial_\beta \sigma(\beta, y)| \leq C|x - y|$.

(A4) For any $p \geq 2$, $\int_{\mathbb{R}_0} |z|^p \nu(dz) < \infty$.

(A5) The process X^{θ_0, β_0} is ergodic in the sense that there exists a unique probability measure $\pi_{\theta_0, \beta_0}(dx)$ such that as $T \to \infty$,

$$\frac{1}{T} \int_0^T g(X_t^{\theta_0, \beta_0}) dt \overset{\mathrm{P}^{\theta_0, \beta_0}}{\longrightarrow} \int_{\mathbb{R}} g(x) \pi_{\theta_0, \beta_0}(dx),$$

for any π_{θ_0, β_0}-integrable function $g : \mathbb{R} \to \mathbb{R}$.

(A6) There exist constants $\epsilon > 0$, $q > 1$, $\rho_1, \rho_2 > 0$ and $0 < \upsilon, \gamma < \frac{1}{2}$ such that as $n \to \infty$,

$$\frac{\sqrt{n}}{\Delta_n^\epsilon} \left(n\Delta_n \left(\int_{\{|z| \geq \rho_2 \Delta_n^{-\gamma}\}} \nu(dz) + \int_{\{|z| \leq \rho_1 \Delta_n^{\upsilon}\}} \nu(dz) \right) \right)^{\frac{1}{q}} \to 0.$$

(A7) For any $\omega, \omega' \in \Omega$, there exist constants $C > 0$ and $n_0 \geq 1$ such that for all $n \geq n_0$ and $k \in \{0, \ldots, n-1\}$,

$$\left| \left(\widehat{L}_{t_{k+1}} - \widehat{L}_{t_k} \right)(\omega) - \left(\widehat{L}_{t_{k+1}} - \widehat{L}_{t_k} \right)(\omega') \right| \begin{cases} = 0, & \text{or} \\ \geq C\Delta_n^{\upsilon}, \end{cases}$$

where υ is as in **(A6)** and $\widehat{L}_t = \int_0^t \int_{\mathbb{R}_0} zN(ds, dz)$ is defined to be the sum of the jumps of $\widehat{Z}$ on the interval $[0, t]$.
Furthermore, for all $(\theta, \beta) \in \Theta \times \Sigma$, $q > 1$ and $p \in \{2, 4\}$,

$$\sum_{r \in \mathcal{A}} r^p \left(\mathrm{P}\left(\widehat{L}_{t_{k+1}} - \widehat{L}_{t_k} = r \middle| X_{t_k} \right) \right)^{\frac{1}{q}} < \infty,$$

$$\sum_{r \in \mathcal{A}} \left(\mathrm{P}\left(\widehat{L}_{t_{k+1}} - \widehat{L}_{t_k} = r \middle| X_{t_k} \right) \right)^{\frac{1}{q}} < \infty,$$

where we denote $\mathcal{A} := \{\sum_{i=1}^j a_i, a_i \in A, j \in \mathbb{N}\}$.

(A8) For any $q > 1$,

$$\sum_{a_i \in A} p_{a_i}^{\frac{1}{q}} < \infty.$$

A detailed explanation on the hypotheses is given in the subsection 1.3.4 of the introductory chapter.

Conditions **(A1)-(A2)** imply that the law of the discrete observation $(X_{t_0}^{\theta}, X_{t_1}^{\theta}, \ldots, X_{t_n}^{\theta})$ of the process $(X_t^{\theta,\beta})_{t\geq 0}$ has a density in $\mathbb{R}^{n+1}$ that we denote by $p(\cdot;(\theta,\beta))$. In particular, $p(\cdot;(\theta_0,\beta_0))$ denotes the density of the random vector X^n. The main result of this chapter is the following LAN property.

Theorem 5.1.1. *Assume conditions* **(A1)**-**(A8)**. *Then, the LAN property holds for the likelihood at* $(\theta_0, \beta_0) \in \Theta \times \Sigma$ *with rate of convergence* $(\sqrt{n\Delta_n}, \sqrt{n})$ *and asymptotic Fisher information matrix* $\Gamma(\theta_0, \beta_0)$. *That is, for all* $w = (u, v) \in \mathbb{R}^2$, *as* $n \to \infty$,

$$\log \frac{p(X^n;(\theta_n,\beta_n))}{p(X^n;(\theta_0,\beta_0))} \overset{\mathcal{L}(\mathrm{P}^{\theta_0,\beta_0})}{\longrightarrow} w^{\mathsf{T}}\mathcal{N}(0,\Gamma(\theta_0,\beta_0)) - \frac{1}{2}w^{\mathsf{T}}\Gamma(\theta_0,\beta_0)w,$$

where $\theta_n = \theta_0 + \frac{u}{\sqrt{n\Delta_n}}$, $\beta_n = \beta_0 + \frac{v}{\sqrt{n}}$, *and* $\mathcal{N}(0,\Gamma(\theta_0,\beta_0))$ *is a centered* $\mathbb{R}^2$*-valued Gaussian random variable with covariance matrix*

$$\Gamma(\theta_0,\beta_0) = \begin{pmatrix} \int_{\mathbb{R}} \left(\frac{\partial_\theta b(\theta_0,x)}{\sigma(\beta_0,x)}\right)^2 \pi_{\theta_0,\beta_0}(dx) & 0 \\ 0 & 2\int_{\mathbb{R}} \left(\frac{\partial_\beta \sigma(\beta_0,x)}{\sigma(\beta_0,x)}\right)^2 \pi_{\theta_0,\beta_0}(dx) \end{pmatrix}.$$

Remark 5.1.1. *Observe that as seen in Remark 4.1.3, we obtain the same asymptotic Fisher information as in the continuous case (see [25, Theorem 4.1]).*

Remark 5.1.2. *Assume condition* **(A8)**. *Then for all* $q \geq 1$, $p > 0$ *and* $n \geq 1$,

$$\sum_{m=1}^{\infty} \sum_{(a_1,\ldots,a_m)\in A} \left(p_{a_1}\cdots p_{a_m}\frac{(C^p\lambda\Delta_n)^m}{m!}\right)^{\frac{1}{q}} < \infty,$$

where $C > 1$ *is the constant in (5.13), since*

$$\begin{aligned}\sum_{m=1}^{\infty} \sum_{(a_1,\ldots,a_m)\in A} \left(p_{a_1}\cdots p_{a_m}\frac{(C^p\lambda\Delta_n)^m}{m!}\right)^{\frac{1}{q}} &= \sum_{m=1}^{\infty} \frac{\left\{(C^p\lambda\Delta_n)^{\frac{1}{q}}\right\}^m}{(m!)^{\frac{1}{q}}} \sum_{(a_1,\ldots,a_m)\in A} p_{a_1}^{\frac{1}{q}}\cdots p_{a_m}^{\frac{1}{q}} \\ &= \sum_{m=1}^{\infty} \frac{\left\{(C^p\lambda\Delta_n)^{\frac{1}{q}}\right\}^m}{(m!)^{\frac{1}{q}}} \left(\sum_{a_i\in A} p_{a_i}^{\frac{1}{q}}\right)^m \\ &= \sum_{m=1}^{\infty} \frac{\left\{(C^p\lambda\Delta_n)^{\frac{1}{q}} \sum_{a_i\in A} p_{a_i}^{\frac{1}{q}}\right\}^m}{(m!)^{\frac{1}{q}}} \\ &< \infty.\end{aligned}$$

Example 5.1.1. 1) *Consider the process*

$$X_t^{\theta} = x_0 + \theta t + \beta B_t + \int_0^t \int_{\mathbb{R}_0} z\widetilde{N}(ds,dz),$$

where $\theta \in \mathbb{R}$, $\beta \in \mathbb{R}_0$, and the Lévy measure is finite and satifies **(A4)**, **(A6)** *and* **(A8)**. *Assume further condition* **(A7)**. *Then, the LAN property holds with rate of convergence $(\sqrt{n\Delta_n}, \sqrt{n})$ and asymptotic Fisher information matrix*

$$\Gamma(\theta_0, \beta_0) = \frac{1}{\beta_0^2} \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}.$$

In this case condition **(A5)** *fails.*

2) *Assume that there exist constants $c, C > 0$ such that $c \leq |a_i| \leq C$, for all $i \in \{1, \dots, \infty\}$. In this case, condition* **(A6)** *holds.*

3) *Assume that $\widehat{\Lambda}$ has distribution $\sum_{i=1}^{\infty} \frac{1}{2^i}\delta_i(dz)$. Then, for n sufficiently large*

$$\frac{\sqrt{n}}{\Delta_n^{\epsilon}} \left(n\Delta_n \left(\int_{\{|z| \geq \rho_2 \Delta_n^{-\gamma}\}} \nu(dz) + \int_{\{|z| \leq \rho_1 \Delta_n^{\upsilon}\}} \nu(dz) \right) \right)^{\frac{1}{q}} = \frac{\sqrt{n}}{\Delta_n^{\epsilon}} \left(\lambda^2 n\Delta_n \sum_{\{i : i \geq \rho_2 \Delta_n^{-\gamma}\}} \frac{1}{2^i} \right)^{\frac{1}{q}}$$
$$\leq \frac{\sqrt{n}}{\Delta_n^{\epsilon}} \left(3\lambda^2 n\Delta_n 2^{-\frac{\rho_2}{2}\Delta_n^{-\gamma}} \right)^{\frac{1}{q}} \to 0,$$

for all $\epsilon > 0$, $q > 1$, $\rho_1, \rho_2 > 0$ and $0 < \upsilon, \gamma < \frac{1}{2}$, and thus, condition **(A6)** *holds.*

In this case, condition **(A8)** *holds since for all $q > 1$,*

$$\sum_{i=1}^{\infty} \frac{1}{2^{\frac{i}{q}}} < \infty.$$

4) *Suppose that $\widehat{L}_t$ has the form $\widehat{L}_t = \sum_{i=1}^{N_t} Y_i$, where $N = (N_t)_{t \geq 0}$ is a Poisson process with intensity $\lambda > 0$, and $(Y_i)_{i \in \mathbb{N}}$ is a sequence of independent and identically distributed positive random variables, independent of N, with distribution $\mu(dz)$ satisfying condition* **(A8)**. *For any $k \in \{0, \dots, n-1\}$, let $q_j = \mathrm{P}(N_{t_{k+1}} - N_{t_k} = j)$, for $j \in \{0, \dots, \infty\}$ and for all $m \geq 0$ set $b_m = \mathrm{P}(N_{t_{k+1}} - N_{t_k} > m) = \sum_{j=m+1}^{\infty} q_j$.*

Observe that for all $m \geq 0$ and n sufficiently large,

$$\frac{b_{m+1}}{b_m} = \frac{\sum_{j=m+2}^{\infty} e^{-\lambda\Delta_n} \frac{(\lambda\Delta_n)^j}{j!}}{\sum_{j=m+1}^{\infty} e^{-\lambda\Delta_n} \frac{(\lambda\Delta_n)^j}{j!}} = \frac{(\lambda\Delta_n)^{m+2} \sum_{i=0}^{\infty} \frac{(\lambda\Delta_n)^i}{(i+m+2)!}}{(\lambda\Delta_n)^{m+1} \sum_{i=0}^{\infty} \frac{(\lambda\Delta_n)^i}{(i+m+1)!}} < \lambda\Delta_n < \frac{1}{2}.$$

Assume that there exists a constant $c > 0$ such that

$$\int_0^{\infty} e^{cz} \mu(dz) \leq 2.$$

Then by [74, Theorem 1], for any $k \in \{0, \dots, n-1\}$, $r \geq 0$ and n sufficiently large,

$$\mathrm{P}\left(\widehat{L}_{t_{k+1}} - \widehat{L}_{t_k} \geq r \middle| X_{t_k} \right) \leq 2\left(1 - e^{-\lambda\Delta_n}\right) e^{-cr} \leq 2e^{-cr},$$

which implies that the second statement of condition **(A7)** *holds.*

As usual, constants will be denoted by C or c and they will always be independent of time and Δ_n but may depend on bounds for the set Θ. They may change of value from one line to the next.

5.2 Preliminaries

In this section we introduce some preliminary results needed for the proof of Theorem 5.1.1.

We start as in Gobet [24] applying the integration by parts formula of the Malliavin calculus on the Wiener space to analyze the log-likelihood function. In order to avoid confusion with the observed process $X^{\theta,\beta}$, we introduce an extra probabilistic representation of $X^{\theta,\beta}$ where the Malliavin calculus will be applied. That is, consider the flow $Y^{\theta,\beta}(s,x)=(Y_t^{\theta,\beta}(s,x), t\geq s)$, $x\in\mathbb{R}$ on the time interval $[s,\infty)$ and with initial condition $Y_s^{\theta,\beta}(s,x)=x$ satisfying

$$\begin{aligned} Y_t^{\theta,\beta}(s,x) &= x+\int_s^t b(\theta,Y_u^{\theta,\beta}(s,x))du+\int_s^t \sigma(\beta,Y_u^{\theta,\beta}(s,x))dW_u \\ &\quad+\int_s^t\int_{\mathbb{R}_0} z\left(M(du,dz)-\nu(dz)du\right), \end{aligned} \tag{5.2}$$

where $W=(W_t)_{t\geq 0}$ is a Brownian motion, $M(dt,dz)$ is a Poisson random measure with intensity measure $\nu(dz)dt$ associated with a centered pure-jump Lévy process $\widetilde{Z}=(\widetilde{Z}_t)_{t\geq 0}$ independent of W, and we denote by $\widetilde{M}(dt,dz):=M(dt,dz)-\nu(dz)dt$ the compensated Poisson random measure. In particular, we write $Y_t^{\theta,\beta}\equiv Y_t^{\theta,\beta}(0,x_0)$, for all $t\geq 0$. That is,

$$Y_t^{\theta,\beta}=x_0+\int_0^t b(\theta,Y_u^{\theta,\beta})du+\int_0^t\sigma(\beta,Y_u^{\theta,\beta})dW_u+\int_0^t\int_{\mathbb{R}_0} z\left(M(du,dz)-\nu(dz)du\right). \tag{5.3}$$

Here, we consider the Malliavin calculus on the Wiener space induced by the Brownian motion W, and we denote by D and δ the Malliavin derivative and the Skorohod integral with respect to W on each interval $[t_k,t_{k+1}]$, respectively (see the Definition 1.1.3 and the discussion following it). For all $A\in\widetilde{\mathcal{F}}$, let us denote $\widetilde{\mathrm{P}}_x^{\theta,\beta}(A)=\widetilde{\mathrm{E}}[\mathbf{1}_A|Y_{t_k}^{\theta,\beta}=x]$. We denote by $\widetilde{\mathrm{E}}_x^{\theta,\beta}$ the expectation with respect to $\widetilde{\mathrm{P}}_x^{\theta,\beta}$. That is, for all $\widetilde{\mathcal{F}}$-measurable random variable V, we have that $\widetilde{\mathrm{E}}_x^{\theta,\beta}[V]=\widetilde{\mathrm{E}}[V|Y_{t_k}^{\theta,\beta}=x]$.

Under conditions **(A1)**-**(A3)**, for any $t>s$ the law of $Y_t^{\theta,\beta}$ conditioned on $Y_s^{\theta,\beta}=x$ admits a positive transition density $p^{\theta,\beta}(t-s,x,y)$, which is differentiable w.r.t. θ and β. As a consequence of [24, Proposition 4.1], we have the following expression for the derivatives of the log-likelihood function w.r.t. θ and β in terms of a conditional expectation.

Proposition 5.2.1. *Assume conditions* **(A1)**-**(A3)**. *Then for all $k\in\{0,...,n-1\}$ and $(\theta,\beta)\in\Theta\times\Sigma$,*

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

where $U_t^{\theta,\beta}(t_k,x)=(D_tY_{t_{k+1}}^{\theta,\beta}(t_k,x))^{-1}=(\partial_xY_{t_{k+1}}^{\theta,\beta}(t_k,x))^{-1}\partial_xY_t^{\theta,\beta}(t_k,x)\sigma^{-1}(Y_t^{\theta,\beta}(t_k,x))$ for all $t\in[t_k,t_{k+1}]$, and the processes $(\partial_\theta Y_t^{\theta,\beta}(t_k,x),t\in[t_k,t_{k+1}])$, $(\partial_\beta Y_t^{\theta,\beta}(t_k,x),t\in[t_k,t_{k+1}])$, and $(\partial_x Y_t^{\theta,\beta}(t_k,x),t\in[t_k,t_{k+1}])$ denote the solutions to linear equations

$$\begin{aligned} \partial_\theta Y_t^{\theta,\beta}(t_k,x) &= \int_{t_k}^t\left(\partial_\theta b(\theta,Y_s^{\theta,\beta}(t_k,x))+\partial_x b(\theta,Y_s^{\theta,\beta}(t_k,x))\partial_\theta Y_s^{\theta,\beta}(t_k,x)\right)ds \\ &\quad+\int_{t_k}^t\partial_x\sigma(\beta,Y_s^{\theta,\beta}(t_k,x))\partial_\theta Y_s^{\theta,\beta}(t_k,x)dW_s, \\ \partial_\beta Y_t^{\theta,\beta}(t_k,x) &= \int_{t_k}^t\partial_x b(\theta,Y_s^{\theta,\beta}(t_k,x))\partial_\beta Y_s^{\theta,\beta}(t_k,x)ds \\ &\quad+\int_{t_k}^t\left(\partial_\beta\sigma(\beta,Y_s^{\theta,\beta}(t_k,x))+\partial_x\sigma(\beta,Y_s^{\theta,\beta}(t_k,x))\partial_\beta Y_s^{\theta,\beta}(t_k,x)\right)dW_s, \\ \partial_x Y_t^{\theta,\beta}(t_k,x) &= 1+\int_{t_k}^t\partial_x b(\theta,Y_s^{\theta,\beta}(t_k,x))\partial_x Y_s^{\theta,\beta}(t_k,x)ds \\ &\quad+\int_{t_k}^t\partial_x\sigma(\beta,Y_s^{\theta,\beta}(t_k,x))\partial_x Y_s^{\theta,\beta}(t_k,x)dW_s. \end{aligned}$$

We have the following decompositions of the Skorohod integral appearing in the conditional expectations of Proposition 5.2.1.

Lemma 5.2.1. *Under conditions* **(A1)**-**(A3)**, *for all* $(\theta,\beta)\in\Theta\times\Sigma$ *and* $k\in\{0,...,n-1\}$,

$$\delta\left(\partial_\theta Y_{t_{k+1}}^{\theta,\beta}(t_k,x)U^{\theta,\beta}(t_k,x)\right)=\Delta_n\partial_\theta b(\theta,Y_{t_k}^{\theta,\beta})\sigma^{-2}(\beta,Y_{t_k}^{\theta,\beta})\left(Y_{t_{k+1}}^{\theta,\beta}-Y_{t_k}^{\theta,\beta}-b(\theta,Y_{t_k}^{\theta,\beta})\Delta_n\right)$$
$$+R_1^{\theta,\beta}+R_2^{\theta,\beta}+R_3^{\theta,\beta}-R_4^{\theta,\beta}-R_5^{\theta,\beta}-R_6^{\theta,\beta},$$

where

$$R_1^{\theta,\beta}:=-\int_{t_k}^{t_{k+1}}D_s\left(\frac{\partial_\theta Y_{t_{k+1}}^{\theta,\beta}(t_k,x)}{\partial_x Y_{t_{k+1}}^{\theta,\beta}(t_k,x)}\right)\frac{\partial_x Y_s^{\theta,\beta}(t_k,x)}{\sigma(\beta,Y_s^{\theta,\beta}(t_k,x))}ds,$$
$$R_2^{\theta,\beta}:=\int_{t_k}^{t_{k+1}}\frac{\partial_\theta b(\theta,Y_s^{\theta,\beta}(t_k,x))}{\partial_x Y_s^{\theta,\beta}(t_k,x)}ds\int_{t_k}^{t_{k+1}}\left(\frac{\partial_x Y_s^{\theta,\beta}(t_k,x)}{\sigma(\beta,Y_s^{\theta,\beta}(t_k,x))}-\frac{\partial_x Y_{t_k}^{\theta,\beta}(t_k,x)}{\sigma(\beta,Y_{t_k}^{\theta,\beta}(t_k,x))}\right)dW_s,$$
$$R_3^{\theta,\beta}:=\int_{t_k}^{t_{k+1}}\left(\frac{\partial_\theta b(\theta,Y_s^{\theta,\beta}(t_k,x))}{\partial_x Y_s^{\theta,\beta}(t_k,x)}-\frac{\partial_\theta b(\theta,Y_{t_k}^{\theta,\beta}(t_k,x))}{\partial_x Y_{t_k}^{\theta,\beta}(t_k,x)}\right)ds\int_{t_k}^{t_{k+1}}\frac{\partial_x Y_{t_k}^{\theta,\beta}(t_k,x)}{\sigma(\beta,Y_{t_k}^{\theta,\beta}(t_k,x))}dW_s,$$
$$R_4^{\theta,\beta}:=\Delta_n\partial_\theta b(\theta,Y_{t_k}^{\theta,\beta})\sigma^{-2}(\beta,Y_{t_k}^{\theta,\beta})\int_{t_k}^{t_{k+1}}\left(b(\theta,Y_s^{\theta,\beta})-b(\theta,Y_{t_k}^{\theta,\beta})\right)ds,$$
$$R_5^{\theta,\beta}:=\Delta_n\partial_\theta b(\theta,Y_{t_k}^{\theta,\beta})\sigma^{-2}(\beta,Y_{t_k}^{\theta,\beta})\int_{t_k}^{t_{k+1}}\left(\sigma(\beta,Y_s^{\theta,\beta})-\sigma(\beta,Y_{t_k}^{\theta,\beta})\right)dW_s,$$
$$R_6^{\theta,\beta}:=\Delta_n\partial_\theta b(\theta,Y_{t_k}^{\theta,\beta})\sigma^{-2}(\beta,Y_{t_k}^{\theta,\beta})\int_{t_k}^{t_{k+1}}\int_{\mathbb{R}_0}z\widetilde{M}(ds,dz),$$

and

$$D_s\left(\frac{\partial_\theta Y_{t_{k+1}}^{\theta,\beta}(t_k,x)}{\partial_x Y_{t_{k+1}}^{\theta,\beta}(t_k,x)}\right)=\int_s^{t_{k+1}}\bigg(-\frac{\partial_\theta b(\theta,Y_u^{\theta,\beta}(t_k,x))}{(\partial_x Y_u^{\theta,\beta}(t_k,x))^2}D_s(\partial_x Y_u^{\theta,\beta}(t_k,x))$$
$$+\partial_{x,\theta}^2 b(\theta,Y_u^{\theta,\beta}(t_k,x))\frac{D_sY_u^{\theta,\beta}(t_k,x)}{\partial_x Y_u^{\theta,\beta}(t_k,x)}\bigg)du.$$

Proof. By Itô's formula,

$$\frac{1}{\partial_x Y_t^{\theta,\beta}(t_k,x)}=1-\int_{t_k}^{t}\frac{\partial_x b(\theta,Y_s^{\theta,\beta}(t_k,x))-\left(\partial_x\sigma(\beta,Y_s^{\theta,\beta}(t_k,x))\right)^2}{\partial_x Y_s^{\theta,\beta}(t_k,x)}ds$$
$$-\int_{t_k}^{t}\frac{\partial_x\sigma(\beta,Y_s^{\theta,\beta}(t_k,x))}{\partial_x Y_s^{\theta,\beta}(t_k,x)}dW_s,$$

which implies that

$$\frac{\partial_\theta Y_{t_{k+1}}^{\theta,\beta}(t_k,x)}{\partial_x Y_{t_{k+1}}^{\theta,\beta}(t_k,x)}=\int_{t_k}^{t_{k+1}}\frac{\partial_\theta b(\theta,Y_s^{\theta,\beta}(t_k,x))}{\partial_x Y_s^{\theta,\beta}(t_k,x)}ds.$$

Then, using the product rule [57, (1.48)], we obtain that

$$\delta\left(\partial_\theta Y_{t_{k+1}}^{\theta,\beta}(t_k,x)U^{\theta,\beta}(t_k,x)\right)=\int_{t_k}^{t_{k+1}}\frac{\partial_\theta b(\theta,Y_s^{\theta,\beta}(t_k,x))}{\partial_x Y_s^{\theta,\beta}(t_k,x)}ds\int_{t_k}^{t_{k+1}}\frac{\partial_x Y_s^{\theta,\beta}(t_k,x)}{\sigma(\beta,Y_s^{\theta,\beta}(t_k,x))}dW_s$$
$$-\int_{t_k}^{t_{k+1}}D_s\left(\frac{\partial_\theta Y_{t_{k+1}}^{\theta,\beta}(t_k,x)}{\partial_x Y_{t_{k+1}}^{\theta,\beta}(t_k,x)}\right)\frac{\partial_x Y_s^{\theta,\beta}(t_k,x)}{\sigma(\beta,Y_s^{\theta,\beta}(t_k,x))}ds.$$

We next add and substract the term $\frac{\partial_x Y^{\theta,\beta}_{t_k}(t_k,x)}{\sigma(\beta,Y^{\theta,\beta}_{t_k}(t_k,x))}$ in the second integral above, and next we add and substract the term $\frac{\partial_\theta b(\theta,Y^{\theta,\beta}_{t_k}(t_k,x))}{\partial_x Y^{\theta,\beta}_{t_k}(t_k,x)}$ in the first one. This, together the fact that $Y^{\theta,\beta}_{t_k}(t_k,x) = Y^{\theta,\beta}_{t_k} = x$, yields

$$\delta\left(\partial_\theta Y^{\theta,\beta}_{t_{k+1}}(t_k,x)U^{\theta,\beta}(t_k,x)\right) = \Delta_n \partial_\theta b(\theta,Y^{\theta,\beta}_{t_k})\sigma^{-1}(\beta,Y^{\theta,\beta}_{t_k})\left(W_{t_{k+1}}-W_{t_k}\right) + R^{\theta,\beta}_1 + R^{\theta,\beta}_2 + R^{\theta,\beta}_3. \tag{5.4}$$

On the other hand, by equation (5.3) we have that

$$\begin{aligned} W_{t_{k+1}} - W_{t_k} = \sigma^{-1}(\beta,Y^{\theta,\beta}_{t_k})\bigg(&Y^{\theta,\beta}_{t_{k+1}} - Y^{\theta,\beta}_{t_k} - b(\theta,Y^{\theta,\beta}_{t_k})\Delta_n - \int_{t_k}^{t_{k+1}}\left(b(\theta,Y^{\theta,\beta}_s) - b(\theta,Y^{\theta,\beta}_{t_k})\right)ds \\ &- \int_{t_k}^{t_{k+1}}\left(\sigma(\beta,Y^{\theta,\beta}_s) - \sigma(\beta,Y^{\theta,\beta}_{t_k})\right)dW_s - \int_{t_k}^{t_{k+1}}\int_{\mathbb{R}_0} z\widetilde{M}(ds,dz)\bigg), \end{aligned}$$

which concludes the desired result. □

Lemma 5.2.2. *Under conditions* **(A1)**-**(A3)**, *for all* $(\theta,\beta)\in\Theta\times\Sigma$ *and* $k\in\{0,...,n-1\}$,

$$\begin{aligned} \delta\left(\partial_\beta Y^{\theta,\beta}_{t_{k+1}}(t_k,x)U^{\theta,\beta}(t_k,x)\right) &= \frac{\partial_\beta\sigma}{\sigma^3}(\beta,Y^{\theta,\beta}_{t_k})\left(Y^{\theta,\beta}_{t_{k+1}} - Y^{\theta,\beta}_{t_k}\right)^2 - \frac{\partial_\beta\sigma}{\sigma}(\beta,Y^{\theta,\beta}_{t_k})\Delta_n \\ &+ H^{\theta,\beta}_3 + H^{\theta,\beta}_4 + H^{\theta,\beta}_5 + H^{\theta,\beta}_6 + H^{\theta,\beta}_7 - \frac{\partial_\beta\sigma}{\sigma^3}(\beta,Y^{\theta,\beta}_{t_k})\bigg\{\left(H^{\theta,\beta}_8 + H^{\theta,\beta}_9 + H^{\theta,\beta}_{10}\right)^2 \\ &+ 2\sigma(\beta,Y^{\theta,\beta}_{t_k})(W_{t_{k+1}} - W_{t_k})\left(H^{\theta,\beta}_8 + H^{\theta,\beta}_9 + H^{\theta,\beta}_{10}\right)\bigg\}, \end{aligned}$$

where

$$\begin{aligned} H^{\theta,\beta}_3 &:= -\int_{t_k}^{t_{k+1}} \frac{\partial_x\sigma(\beta,Y^{\theta,\beta}_s(t_k,x))\partial_\beta\sigma(\beta,Y^{\theta,\beta}_s(t_k,x))}{\partial_x Y^{\theta,\beta}_s(t_k,x)}ds\int_{t_k}^{t_{k+1}}\frac{\partial_x Y^{\theta,\beta}_s(t_k,x)}{\sigma(\beta,Y^{\theta,\beta}_s(t_k,x))}dW_s, \\ H^{\theta,\beta}_4 &:= -\int_{t_k}^{t_{k+1}}\left(H^{\theta,\beta}_1 + H^{\theta,\beta}_2\right)\frac{\partial_x Y^{\theta,\beta}_s(t_k,x)}{\sigma(\beta,Y^{\theta,\beta}_s(t_k,x))}ds, \\ H^{\theta,\beta}_5 &:= \int_{t_k}^{t_{k+1}}\left(\frac{\partial_\beta\sigma(\beta,Y^{\theta,\beta}_s(t_k,x))}{\partial_x Y^{\theta,\beta}_s(t_k,x)} - \frac{\partial_\beta\sigma(\beta,Y^{\theta,\beta}_{t_k}(t_k,x))}{\partial_x Y^{\theta,\beta}_{t_k}(t_k,x)}\right)dW_s\int_{t_k}^{t_{k+1}}\frac{\partial_x Y^{\theta,\beta}_s(t_k,x)}{\sigma(\beta,Y^{\theta,\beta}_s(t_k,x))}dW_s, \\ H^{\theta,\beta}_6 &:= \int_{t_k}^{t_{k+1}}\frac{\partial_\beta\sigma(\beta,Y^{\theta,\beta}_{t_k}(t_k,x))}{\partial_x Y^{\theta,\beta}_{t_k}(t_k,x)}dW_s\int_{t_k}^{t_{k+1}}\left(\frac{\partial_x Y^{\theta,\beta}_s(t_k,x)}{\sigma(\beta,Y^{\theta,\beta}_s(t_k,x))} - \frac{\partial_x Y^{\theta,\beta}_{t_k}(t_k,x)}{\sigma(\beta,Y^{\theta,\beta}_{t_k}(t_k,x))}\right)dW_s, \\ H^{\theta,\beta}_7 &:= -\int_{t_k}^{t_{k+1}}\left(\frac{\partial_\beta\sigma(\beta,Y^{\theta,\beta}_s(t_k,x))}{\sigma(\beta,Y^{\theta,\beta}_s(t_k,x))} - \frac{\partial_\beta\sigma(\beta,Y^{\theta,\beta}_{t_k}(t_k,x))}{\sigma(\beta,Y^{\theta,\beta}_{t_k}(t_k,x))}\right)ds, \quad H^{\theta,\beta}_8 := \int_{t_k}^{t_{k+1}} b(\theta,Y^{\theta,\beta}_s)ds, \\ H^{\theta,\beta}_9 &:= \int_{t_k}^{t_{k+1}}\left(\sigma(\beta,Y^{\theta,\beta}_s) - \sigma(\beta,Y^{\theta,\beta}_{t_k})\right)dW_s, \quad H^{\theta,\beta}_{10} := \int_{t_k}^{t_{k+1}}\int_{\mathbb{R}_0} z\widetilde{M}(ds,dz), \end{aligned}$$

and

$$\begin{aligned} H^{\theta,\beta}_1 &= -\int_s^{t_{k+1}} D_s\left(\frac{\partial_x\sigma(\beta,Y^{\theta,\beta}_u(t_k,x))\partial_\beta\sigma(\beta,Y^{\theta,\beta}_u(t_k,x))}{\partial_x Y^{\theta,\beta}_u(t_k,x)}\right)du, \\ H^{\theta,\beta}_2 &= \int_s^{t_{k+1}} D_s\left(\frac{\partial_\beta\sigma(\beta,Y^{\theta,\beta}_u(t_k,x))}{\partial_x Y^{\theta,\beta}_u(t_k,x)}\right)dW_u, \end{aligned}$$

where

$$D_s\left(\frac{\partial_x\sigma\partial_\beta\sigma(\beta,Y_u^{\theta,\beta}(t_k,x))}{\partial_xY_u^{\theta,\beta}(t_k,x)}\right)=-\frac{\partial_x\sigma\partial_\beta\sigma(\beta,Y_u^{\theta,\beta}(t_k,x))}{(\partial_xY_u^{\theta,\beta}(t_k,x))^2}D_s(\partial_xY_u^{\theta,\beta}(t_k,x))$$
$$+\left(\partial_x^2\sigma\partial_\beta\sigma(\beta,Y_u^{\theta,\beta}(t_k,x))+\partial_x\sigma\partial_{x,\beta}^2\sigma(\beta,Y_u^{\theta,\beta}(t_k,x))\right)\frac{D_sY_u^{\theta,\beta}(t_k,x)}{\partial_xY_u^{\theta,\beta}(t_k,x)},$$
$$D_s\left(\frac{\partial_\beta\sigma(\beta,Y_u^{\theta,\beta})}{\partial_xY_u^{\theta,\beta}(t_k,x)}\right)=-\frac{\partial_\beta\sigma(\beta,Y_u^{\theta,\beta}(t_k,x))}{(\partial_xY_u^{\theta,\beta}(t_k,x))^2}D_s(\partial_xY_u^{\theta,\beta}(t_k,x))+\partial_{x,\beta}^2\sigma(\beta,Y_u^{\theta,\beta}(t_k,x))\frac{D_sY_u^{\theta,\beta}(t_k,x)}{\partial_xY_u^{\theta,\beta}(t_k,x)}.$$

Proof. By Itô's formula,

$$\frac{\partial_\beta Y_{t_{k+1}}^{\theta,\beta}(t_k,x)}{\partial_xY_{t_{k+1}}^{\theta,\beta}(t_k,x)}=-\int_{t_k}^{t_{k+1}}\frac{\partial_x\sigma(\beta,Y_s^{\theta,\beta}(t_k,x))\partial_\beta\sigma(\beta,Y_s^{\theta,\beta}(t_k,x))}{\partial_xY_s^{\theta,\beta}(t_k,x)}ds+\int_{t_k}^{t_{k+1}}\frac{\partial_\beta\sigma(\beta,Y_s^{\theta,\beta}(t_k,x))}{\partial_xY_s^{\theta,\beta}(t_k,x)}dW_s.$$

Then, using again the product rule [57, (1.48)], we obtain that

$$\delta\left(\partial_\beta Y_{t_{k+1}}^{\theta,\beta}(t_k,x)U^{\theta,\beta}(t_k,x)\right)=\left(-\int_{t_k}^{t_{k+1}}\frac{\partial_x\sigma\partial_\beta\sigma(\beta,Y_s^{\theta,\beta}(t_k,x))}{\partial_xY_s^{\theta,\beta}(t_k,x)}ds+\int_{t_k}^{t_{k+1}}\frac{\partial_\beta\sigma(\beta,Y_s^{\theta,\beta}(t_k,x))}{\partial_xY_s^{\theta,\beta}(t_k,x)}dW_s\right)$$
$$\times\int_{t_k}^{t_{k+1}}\frac{\partial_xY_s^{\theta,\beta}(t_k,x)}{\sigma(\beta,Y_s^{\theta,\beta}(t_k,x))}dW_s-\int_{t_k}^{t_{k+1}}\left\{\frac{\partial_\beta\sigma(\beta,Y_s^{\theta,\beta}(t_k,x))}{\partial_xY_s^{\theta,\beta}(t_k,x)}+\int_s^{t_{k+1}}D_s\left(\frac{\partial_\beta\sigma(\beta,Y_u^{\theta,\beta}(t_k,x))}{\partial_xY_u^{\theta,\beta}(t_k,x)}\right)dW_u\right.$$
$$\left.-\int_s^{t_{k+1}}D_s\left(\frac{\partial_x\sigma(\beta,Y_u^{\theta,\beta}(t_k,x))\partial_\beta\sigma(\beta,Y_u^{\theta,\beta}(t_k,x))}{\partial_xY_u^{\theta,\beta}(t_k,x)}\right)du\right\}\frac{\partial_xY_s^{\theta,\beta}(t_k,x)}{\sigma(\beta,Y_s^{\theta,\beta}(t_k,x))}ds.$$

We next add and substract the term $\frac{\partial_\beta\sigma(\beta,Y_{t_k}^{\theta,\beta}(t_k,x))}{\partial_xY_{t_k}^{\theta,\beta}(t_k,x)}$ in the second integral above, the term $\frac{\partial_xY_{t_k}^{\theta,\beta}(t_k,x)}{\sigma(\beta,Y_{t_k}^{\theta,\beta}(t_k,x))}$ in the third one, and the term $\frac{\partial_\beta\sigma(\beta,Y_{t_k}^{\theta,\beta}(t_k,x))}{\sigma(\beta,Y_{t_k}^{\theta,\beta}(t_k,x))}$ in the last one. This, together the fact that $Y_{t_k}^{\theta,\beta}(t_k,x)=Y_{t_k}^{\theta,\beta}=x$, yields

$$\delta\left(\partial_\beta Y_{t_{k+1}}^{\theta,\beta}(t_k,x)U^{\theta,\beta}(t_k,x)\right)=\frac{\partial_\beta\sigma}{\sigma}(\beta,Y_{t_k}^{\theta,\beta})\left(W_{t_{k+1}}-W_{t_k}\right)^2-\frac{\partial_\beta\sigma}{\sigma}(\beta,Y_{t_k}^{\theta,\beta})\Delta_n$$
$$+H_3^{\theta,\beta}+H_4^{\theta,\beta}+H_5^{\theta,\beta}+H_6^{\theta,\beta}+H_7^{\theta,\beta}. \tag{5.5}$$

On the other hand, by equation (5.3) we have that

$$W_{t_{k+1}}-W_{t_k}=\sigma^{-1}(\beta,Y_{t_k}^{\theta,\beta})\Bigg(Y_{t_{k+1}}^{\theta,\beta}-Y_{t_k}^{\theta,\beta}-\int_{t_k}^{t_{k+1}}b(\theta,Y_s^{\theta,\beta})ds$$
$$-\int_{t_k}^{t_{k+1}}\left(\sigma(\beta,Y_s^{\theta,\beta})-\sigma(\beta,Y_{t_k}^{\theta,\beta})\right)dW_s-\int_{t_k}^{t_{k+1}}\int_{\mathbb{R}_0}z\widetilde{M}(ds,dz)\Bigg),$$

which concludes the desired result. □

We will use the following estimates for the solution to (5.2).

Lemma 5.2.3. *Assume conditions* **(A1)** *and* **(A4)**.

(i) *For any $p\geq 2$ and $(\theta,\beta)\in\Theta\times\Sigma$, there exists a constant $C_p>0$ such that for all $k\in\{0,...,n-1\}$ and $t\in[t_k,t_{k+1}]$,*

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

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

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

Moreover, all these statements remain valid for $X^{\theta,\beta}$.

Under conditions **(A1)-(A3)**, for any $k \in \{0, ..., n-1\}$ and $t \geq t_k$, the random variables $Y_t^{\theta,\beta}(t_k, x)$, $\partial_x Y_t^{\theta,\beta}(t_k, x)$, $(\partial_x Y_t^{\theta,\beta}(t_k, x))^{-1}$, $\partial_\theta Y_t^{\theta,\beta}(t_k, x)$ and $\partial_\beta Y_t^{\theta,\beta}(t_k, x)$ belong to $\mathbb{D}^{1,2}$ (see [61, Theorem 3]).

Assuming conditions **(A1)-(A4)** and using Gronwall's inequality, one can easily check that for any $(\theta, \beta) \in \Theta \times \Sigma$ and $p \geq 2$, there exist constants $C_p, q > 0$ such that for all $k \in \{0, ..., n-1\}$ and $t \in [t_k, t_{k+1}]$,

$$\mathrm{E}\left[\left|\partial_x Y_t^{\theta,\beta}(t_k, x))\right|^p + \frac{1}{\left|\partial_x Y_t^{\theta,\beta}(t_k, x))\right|^p} \Big| Y_{t_k}^{\theta,\beta}(t_k, x) = x\right]$$
$$+ \sup_{s \in [t_k, t_{k+1}]} \mathrm{E}\left[\left|D_s Y_t^{\theta,\beta}(t_k, x)\right|^p \Big| Y_{t_k}^{\theta,\beta}(t_k, x) = x\right] \leq C_p, \quad \text{and}$$
$$\sup_{s \in [t_k, t_{k+1}]} \mathrm{E}\left[\left|D_s\left(\partial_x Y_t^{\theta,\beta}(t_k, x)\right)\right|^p \Big| Y_{t_k}^{\theta,\beta}(t_k, x) = x\right] \leq C_p\left(1 + |x|^q\right).$$

As a consequence, we have the following estimates, which follow easily from (5.4), (5.5), Lemma 5.2.3 and properties of the moments of the Brownian motion.

Lemma 5.2.4. *Under conditions* **(A1)-(A4)**, *for any $(\theta, \beta) \in \Theta \times \Sigma$ and $p \geq 2$, there exist constants $C_p, q > 0$ such that for all $k \in \{0, ..., n-1\}$,*

$$\mathrm{E}\left[R_1^{\theta,\beta} + R_2^{\theta,\beta} + R_3^{\theta,\beta} - R_5^{\theta,\beta} \big| Y_{t_k}^{\theta,\beta}(t_k, x) = x\right] = 0, \tag{5.6}$$

$$\mathrm{E}\left[\left|R_1^{\theta,\beta} + R_2^{\theta,\beta} + R_3^{\theta,\beta} - R_5^{\theta,\beta}\right|^p \big| Y_{t_k}^{\theta,\beta}(t_k, x) = x\right] \leq C_p \Delta_n^{\frac{3p+1}{2}}\left(1 + |x|^q\right), \tag{5.7}$$

$$\mathrm{E}\left[\left|\delta\left(\partial_\theta Y_{t_{k+1}}^{\theta,\beta}(t_k, x) U^{\theta,\beta}(t_k, x)\right)\right|^p \big| Y_{t_k}^{\theta,\beta}(t_k, x) = x\right] \leq C_p \Delta_n^{\frac{3p}{2}}\left(1 + |x|^q\right), \tag{5.8}$$

$$\mathrm{E}\left[H_3^{\theta,\beta} + H_4^{\theta,\beta} + H_5^{\theta,\beta} + H_6^{\theta,\beta} + H_7^{\theta,\beta} \big| Y_{t_k}^{\theta,\beta}(t_k, x) = x\right] = 0, \tag{5.9}$$

$$\mathrm{E}\left[\left|H_3^{\theta,\beta} + H_4^{\theta,\beta} + H_5^{\theta,\beta} + H_6^{\theta,\beta} + H_7^{\theta,\beta}\right|^p \big| Y_{t_k}^{\theta,\beta}(t_k, x) = x\right] \leq C_p \Delta_n^{p+\frac{1}{2}}\left(1 + |x|^q\right), \tag{5.10}$$

$$\mathrm{E}\left[\left|\delta\left(\partial_\beta Y_{t_{k+1}}^{\theta,\beta}(t_k, x) U^{\theta,\beta}(t_k, x)\right)\right|^p \big| Y_{t_k}^{\theta,\beta}(t_k, x) = x\right] \leq C_p \Delta_n^{p}\left(1 + |x|^q\right). \tag{5.11}$$

For any $t > s$ and $j \geq 0$, we denote by $q_{(j)}^{\theta,\beta}(t-s, x, y)$ the transition density of $Y_t^{\theta,\beta}$ conditioned on $Y_s^{\theta,\beta} = x$ and $M_t - M_s = j$, where $M_t = M([0, t] \times \mathbb{R})$. That is,

$$p^{\theta,\beta}(t-s, x, y) = \sum_{j=0}^{\infty} q_{(j)}^{\theta,\beta}(t-s, x, y) e^{-\lambda(t-s)} \frac{(\lambda(t-s))^j}{j!}. \tag{5.12}$$

From [24, Proposition 5.1], for any $(\theta, \beta) \in \Theta \times \Sigma$ there exist constants $c, C > 1$ such that for all $0 < t \leq 1$, and $x, y \in \mathbb{R}$,

$$\frac{1}{C\sqrt{t}} e^{-c\frac{(y-x)^2}{t}} \leq q_{(0)}^{\theta,\beta}(t, x, y) \leq \frac{C}{\sqrt{t}} e^{-\frac{(y-x)^2}{ct}}. \tag{5.13}$$

For any $t > s$ and $j \geq 1$, we denote by $q_{(j)}^{\theta,\beta}(t-s, x, y; a_1, \ldots, a_j)$ the transition density of $Y_t^{\theta,\beta}$ conditioned on $Y_s^{\theta,\beta} = x, M_t - M_s = j$ and $\widetilde{\Lambda}_{[s,t]} = \{a_1, \ldots, a_i\}$, where $\widetilde{\Lambda}_{[s,t]}$ are the jump amplitudes of $\widetilde{Z}$ on the interval $[s,t]$, i.e, $\widetilde{\Lambda}_{[s,t]} := \{\Delta\widetilde{Z}_u; s \leq u \leq t\}$. Consider the events $\widehat{J}_{j,k} = \{N_{t_{k+1}} - N_{t_k} = j\}$ and $\widetilde{J}_{j,k} = \{M_{t_{k+1}} - M_{t_k} = j\}$, for $j \geq 0$ and $k \in \{0,...,n-1\}$, where $N_t = N([0,t]\times\mathbb{R})$. We denote by $\widehat{\Lambda}_{[s,t]}$ the jump amplitudes of $\widehat{Z}$ on the interval $[s,t]$, i.e, $\widehat{\Lambda}_{[s,t]} := \{\Delta\widehat{Z}_u; s \leq u \leq t\}$, and by $\{\widehat{J}_{j,k}, a_1,\ldots,a_j\} := \{N_{t_{k+1}} - N_{t_k} = j\} \cap \{\widehat{\Lambda}_{[t_k,t_{k+1}]} = \{a_1,\ldots,a_j\}\}$, $\{\widetilde{J}_{j,k}, a_1,\ldots,a_j\} := \{M_{t_{k+1}} - M_{t_k} = j\} \cap \{\widetilde{\Lambda}_{[t_k,t_{k+1}]} = \{a_1,\ldots,a_j\}\}$, for any $j \geq 1$ and $a_1,\ldots,a_j \in \mathbb{R}_0$.

In what follows, by abuse of notation we will let $\widetilde{\Lambda}_{[s,t]}(\omega) = a_1$ in the case that $M_t(\omega) - M_s(\omega) = 1$, similarly for $\widehat{\Lambda}$.

Lemma 5.2.5. *Under conditions* **(A1)**-**(A3)**, *for all* $(\theta,\beta) \in \Theta\times\Sigma$, $j \geq 1$, $x,y \in \mathbb{R}$ *and* $a_1,\ldots,a_j \in \mathbb{R}_0$,

$$\frac{1}{C^{j+1}\sqrt{\Delta_n}} e^{-c\frac{(y-x-a)^2}{\Delta_n}} \leq q_{(j)}^{\theta,\beta}(\Delta_n,x,y;a_1,\ldots,a_j) \leq \frac{C^{j+1}}{\sqrt{\Delta_n}} e^{-\frac{(y-x-a)^2}{c\Delta_n}}, \tag{5.14}$$

where $a = a_1 + \cdots + a_j$ *and* C, c *are the constants in* (5.13).

Proof. Using (5.13) and the Chapman-Kolmogorov equation in terms of transition density repeatedly, we get that

$$\begin{aligned} q_{(j)}^{\theta,\beta}(\Delta_n,x,y;a_1,\ldots,a_j) &= \frac{j!}{\Delta_n^j}\int_{t_k}^{t_{k+1}}\int_{t_k}^{s_j}\int_{t_k}^{s_{j-1}}\cdots\int_{t_k}^{s_2}\int_{\mathbb{R}^j} q_{(0)}^{\theta,\beta}(s_1-t_k,x,z_1) \\ &\quad \times q_{(0)}^{\theta,\beta}(s_2-s_1,z_1+a_1,z_2)\cdots q_{(0)}^{\theta,\beta}(t_{k+1}-s_j,z_j+a_j,y)dz_1\cdots dz_j ds_1\cdots ds_j \\ &\leq \frac{j!}{\Delta_n^j}\int_{t_k}^{t_{k+1}}\int_{t_k}^{s_j}\int_{t_k}^{s_{j-1}}\cdots\int_{t_k}^{s_2}\int_{\mathbb{R}^j} \frac{C}{\sqrt{s_1-t_k}}e^{-\frac{(z_1-x)^2}{c(s_1-t_k)}}\frac{C}{\sqrt{s_2-s_1}}e^{-\frac{(z_2-z_1-a_1)^2}{c(s_2-s_1)}} \\ &\quad \times\cdots\times\frac{C}{\sqrt{t_{k+1}-s_j}}e^{-\frac{\left(y-z_j-a_j\right)^2}{c(t_{k+1}-s_j)}}dz_1\cdots dz_j ds_1\cdots ds_j \\ &= \frac{j!}{\Delta_n^j}\int_{t_k}^{t_{k+1}}\int_{t_k}^{s_j}\int_{t_k}^{s_{j-1}}\cdots\int_{t_k}^{s_2}\frac{C^{j+1}}{\sqrt{\Delta_n}}e^{-\frac{\left(y-x-a_1-\cdots-a_j\right)^2}{c\Delta_n}}ds_1\cdots ds_j, \end{aligned}$$

which concludes the upper bound of (5.14) using the fact that

$$\frac{j!}{\Delta_n^j}\int_{t_k}^{t_{k+1}}\int_{t_k}^{s_j}\int_{t_k}^{s_{j-1}}\cdots\int_{t_k}^{s_2}ds_1\cdots ds_j = 1.$$

Similarly,

$$\begin{aligned} q_{(j)}^{\theta,\beta}(\Delta_n,x,y;a_1,\ldots,a_j) &\geq \frac{j!}{\Delta_n^j}\int_{t_k}^{t_{k+1}}\int_{t_k}^{s_j}\int_{t_k}^{s_{j-1}}\cdots\int_{t_k}^{s_2}\int_{\mathbb{R}^j}\frac{1}{C\sqrt{s_1-t_k}}e^{-c\frac{(z_1-x)^2}{s_1-t_k}} \\ &\quad \times\frac{1}{C\sqrt{s_2-s_1}}e^{-c\frac{(z_2-z_1-a_1)^2}{s_2-s_1}}\times\cdots\times\frac{1}{C\sqrt{t_{k+1}-s_j}}e^{-c\frac{\left(y-z_j-a_j\right)^2}{t_{k+1}-s_j}}dz_1\cdots dz_j ds_1\cdots ds_j \\ &= \frac{j!}{\Delta_n^j}\int_{t_k}^{t_{k+1}}\int_{t_k}^{s_j}\int_{t_k}^{s_{j-1}}\cdots\int_{t_k}^{s_2}\frac{1}{C^{j+1}\sqrt{\Delta_n}}e^{-c\frac{\left(y-x-a_1-\cdots-a_j\right)^2}{\Delta_n}}ds_1\cdots ds_j, \end{aligned}$$

which concludes the lower bound of (5.14), and finishes the desired proof. □

Lemma 5.2.6. *Assume conditions* **(A1)**-**(A3)**. *Then for any* $(\theta,\beta),(\bar{\theta},\bar{\beta}) \in \Theta\times\Sigma$, *and* $p>1$ *close to* 1, *there exists a constant* $C_0 > 0$ *such that for all* $a_1,\ldots,a_j \in \mathbb{R}_0$, $k \in \{0,...,n-1\}$, *and* $j \geq 1$,

$$\begin{aligned} &\mathrm{E}\left[\mathbf{1}_{\{\widehat{J}_{j,k},a_1,\ldots,a_j\}}\left(\frac{q_{(j)}^{\bar{\theta},\bar{\beta}}}{q_{(j)}^{\theta,\beta}}(\Delta_n,X_{t_k}^{\theta,\beta},X_{t_{k+1}}^{\theta,\beta};a_1,\ldots,a_j)\right)^p\Big|X_{t_k}^{\theta,\beta}=x\right] \\ &\leq C_0C^{(2p-1)(j+1)}p_{a_1}\cdots p_{a_j}e^{-\lambda\Delta_n}\frac{(\lambda\Delta_n)^j}{j!}, \end{aligned} \tag{5.15}$$

where C is as in (5.14), and

$$\mathrm{E}\left[\mathbf{1}_{\widehat{J}_{0,k}}\left(\frac{q_{(0)}^{\bar{\theta},\bar{\beta}}}{q_{(0)}^{\theta,\beta}}(\Delta_n, X_{t_k}^{\theta,\beta}, X_{t_{k+1}}^{\theta,\beta})\right)^p \Big| X_{t_k}^{\theta,\beta}=x\right] \leq C_0. \tag{5.16}$$

Moreover, all these statements remain valid for $Y^{\theta,\beta}$.

Proof. Applying the upper and lower bound of (5.14) to $q_{(j)}^{\bar{\theta},\bar{\beta}}$ and $q_{(j)}^{\theta,\beta}$, respectively, together with the independence between N and $\widehat{\Lambda}$, we get that for all $p>1$,

$$\begin{aligned}
&\mathrm{E}\left[\mathbf{1}_{\{\widehat{J}_{j,k},a_1,\dots,a_j\}}\left(\frac{q_{(j)}^{\bar{\theta},\bar{\beta}}}{q_{(j)}^{\theta,\beta}}(\Delta_n, X_{t_k}^{\theta,\beta}, X_{t_{k+1}}^{\theta,\beta}; a_1,\dots,a_j)\right)^p \Big| X_{t_k}^{\theta,\beta}=x\right]\\
&= p_{a_1}\cdots p_{a_j} e^{-\lambda\Delta_n}\frac{(\lambda\Delta_n)^j}{j!}\int_{\mathbb{R}} q_{(j)}^{\bar{\theta},\bar{\beta}}(\Delta_n,x,y;a_1,\dots,a_j)^p q_{(j)}^{\theta,\beta}(\Delta_n,x,y;a_1,\dots,a_j)^{1-p}dy\\
&\leq C^{(j+1)(2p-1)} p_{a_1}\cdots p_{a_j} e^{-\lambda\Delta_n}\frac{(\lambda\Delta_n)^j}{j!}\int_{\mathbb{R}}\frac{1}{\sqrt{\Delta_n}} e^{(-\frac{p}{c}+(p-1)c)\frac{\left(y-x-a_1-\cdots-a_j\right)^2}{\Delta_n}}dy.
\end{aligned}$$

If we choose $p\in(1,\frac{c^2}{c^2-1})$, then $-\frac{p}{c}+(p-1)c<0$, and the integral above is finite and equal to C_0. This concludes (5.15). The proof of (5.16) can be done similarly by using (5.13). □

As in [24, Proposition 4.1], we have the following expression for the derivatives of the log-likelihood function conditioned on the number and the amplitudes of jumps w.r.t. θ and β in terms of a conditional expectation.

Lemma 5.2.7. *Assume conditions* **(A1)**-**(A3)**. *Then for all $(\theta,\beta)\in\Theta\times\Sigma$, $k\in\{0,...,n-1\}$, $j\geq 1$, $x,y\in\mathbb{R}$, and $a_1,\dots,a_j\in\mathbb{R}_0$,*

$$\begin{aligned}
\frac{\partial_\theta q_{(j)}^{\theta,\beta}}{q_{(j)}^{\theta,\beta}}(\Delta_n,x,y;a_1,\dots,a_j) &= \frac{1}{\Delta_n}\widetilde{\mathrm{E}}_x^{\theta,\beta}\left[\delta\left(\partial_\theta Y_{t_{k+1}}^{\theta,\beta}(t_k,x)U^{\theta,\beta}(t_k,x)\right)\Big| Y_{t_{k+1}}^{\theta,\beta}=y,\widetilde{J}_{j,k},a_1,\dots,a_j\right],\\
\frac{\partial_\beta q_{(j)}^{\theta,\beta}}{q_{(j)}^{\theta,\beta}}(\Delta_n,x,y;a_1,\dots,a_j) &= \frac{1}{\Delta_n}\widetilde{\mathrm{E}}_x^{\theta,\beta}\left[\delta\left(\partial_\beta Y_{t_{k+1}}^{\theta,\beta}(t_k,x)U^{\theta,\beta}(t_k,x)\right)\Big| Y_{t_{k+1}}^{\theta,\beta}=y,\widetilde{J}_{j,k},a_1,\dots,a_j\right],\\
\frac{\partial_\theta q_{(0)}^{\theta,\beta}}{q_{(0)}^{\theta,\beta}}(\Delta_n,x,y) &= \frac{1}{\Delta_n}\widetilde{\mathrm{E}}_x^{\theta,\beta}\left[\delta\left(\partial_\theta Y_{t_{k+1}}^{\theta,\beta}(t_k,x)U^{\theta,\beta}(t_k,x)\right)\Big| Y_{t_{k+1}}^{\theta,\beta}=y,\widetilde{J}_{0,k}\right],\\
\frac{\partial_\beta q_{(0)}^{\theta,\beta}}{q_{(0)}^{\theta,\beta}}(\Delta_n,x,y) &= \frac{1}{\Delta_n}\widetilde{\mathrm{E}}_x^{\theta,\beta}\left[\delta\left(\partial_\beta Y_{t_{k+1}}^{\theta,\beta}(t_k,x)U^{\theta,\beta}(t_k,x)\right)\Big| Y_{t_{k+1}}^{\theta,\beta}=y,\widetilde{J}_{0,k}\right],
\end{aligned}$$

where the process $U^{\theta,\beta}=(U_t^{\theta,\beta},t\in[t_k,t_{k+1}])$ is defined in Proposition 5.2.1.

Proof. Let f be a continuously differentiable function with compact support. The chain rule of the Malliavin calculus implies that $f'(Y_{t_{k+1}}^{\theta,\beta}(t_k,x)) = D_t(f(Y_{t_{k+1}}^{\theta,\beta}(t_k,x)))U_t^{\theta,\beta}(t_k,x)$, for all $(\theta,\beta)\in\Theta\times\Sigma$ and $t\in[t_k,t_{k+1}]$, where

$$U_t^{\theta,\beta}(t_k,x)=\frac{1}{D_tY_{t_{k+1}}^{\theta,\beta}(t_k,x)}.$$

Then, using the integration by parts formula of the Malliavin calculus on the interval $[t_k,t_{k+1}]$

and the independence between W, $\widetilde{\Lambda}$ and M, we get that

$$\begin{aligned}
&\partial_\theta \widetilde{\mathrm{E}}\left[\mathbf{1}_{\{\widetilde{J}_{j,k},a_1,\ldots,a_j\}} f(Y^{\theta,\beta}_{t_{k+1}}(t_k,x))\right] = \widetilde{\mathrm{E}}\left[\mathbf{1}_{\{\widetilde{J}_{j,k},a_1,\ldots,a_j\}} f'(Y^{\theta,\beta}_{t_{k+1}}(t_k,x))\partial_\theta Y^{\theta,\beta}_{t_{k+1}}(t_k,x)\right]\\
&= \frac{1}{\Delta_n}\widetilde{\mathrm{E}}\left[\mathbf{1}_{\{\widetilde{J}_{j,k},a_1,\ldots,a_j\}}\int_{t_k}^{t_{k+1}} f'(Y^{\theta,\beta}_{t_{k+1}}(t_k,x))\partial_\theta Y^{\theta,\beta}_{t_{k+1}}(t_k,x)dt\right]\\
&= \frac{1}{\Delta_n}\widetilde{\mathrm{E}}\left[\int_{t_k}^{t_{k+1}} D_t(f(Y^{\theta,\beta}_{t_{k+1}}(t_k,x)))U^{\theta,\beta}_t(t_k,x)\partial_\theta Y^{\theta,\beta}_{t_{k+1}}(t_k,x)\mathbf{1}_{\{\widetilde{J}_{j,k},a_1,\ldots,a_j\}}dt\right]\\
&= \frac{1}{\Delta_n}\widetilde{\mathrm{E}}\left[\mathbf{1}_{\{\widetilde{J}_{j,k},a_1,\ldots,a_j\}} f(Y^{\theta,\beta}_{t_{k+1}}(t_k,x))\delta\left(\partial_\theta Y^{\theta,\beta}_{t_{k+1}}(t_k,x)U^{\theta,\beta}(t_k,x)\right)\right].
\end{aligned}$$

On the other hand,

$$\begin{aligned}
&\partial_\theta \widetilde{\mathrm{E}}\left[\mathbf{1}_{\{\widetilde{J}_{j,k},a_1,\ldots,a_j\}} f(Y^{\theta,\beta}_{t_{k+1}}(t_k,x))\right]\\
&= \int_{\mathbb{R}} f(y)\partial_\theta q^{\theta,\beta}_{(j)}(\Delta_n,x,y;a_1,\ldots,a_j)p_{a_1}\cdots p_{a_j}e^{-\lambda\Delta_n}\frac{(\lambda\Delta_n)^j}{j!}dy,
\end{aligned}$$

and

$$\begin{aligned}
&\widetilde{\mathrm{E}}\left[\mathbf{1}_{\{\widetilde{J}_{j,k},a_1,\ldots,a_j\}} f(Y^{\theta,\beta}_{t_{k+1}}(t_k,x))\delta\left(\partial_\theta Y^{\theta,\beta}_{t_{k+1}}(t_k,x)U^{\theta,\beta}(t_k,x)\right)\right]\\
&= \widetilde{\mathrm{E}}\left[\mathbf{1}_{\{\widetilde{J}_{j,k},a_1,\ldots,a_j\}} f(Y^{\theta,\beta}_{t_{k+1}})\delta\left(\partial_\theta Y^{\theta,\beta}_{t_{k+1}}(t_k,x)U^{\theta,\beta}(t_k,x)\right)\Big| Y^{\theta,\beta}_{t_k}=x\right]\\
&= \int_{\mathbb{R}} f(y)\widetilde{\mathrm{E}}\left[\delta\left(\partial_\theta Y^{\theta,\beta}_{t_{k+1}}(t_k,x)U^{\theta,\beta}(t_k,x)\right)\Big| Y^{\theta,\beta}_{t_{k+1}}=y, Y^{\theta,\beta}_{t_k}=x,\widetilde{J}_{j,k},a_1,\ldots,a_j\right]\\
&\quad\times q^{\theta,\beta}_{(j)}(\Delta_n,x,y;a_1,\ldots,a_j)p_{a_1}\cdots p_{a_j}e^{-\lambda\Delta_n}\frac{(\lambda\Delta_n)^j}{j!}dy.
\end{aligned}$$

This shows the first equality. The proof of the other equalities follow along the same lines and are omitted. □

As in [25, Proposition 1.2], we have the following estimates.

Lemma 5.2.8. *Assume conditions* **(A1)-(A4)**. *Then for any* $(\theta,\beta),(\bar{\theta},\bar{\beta})\in\Theta\times\Sigma$, $p>1$, *and* $p_1>1$ *close to* 1, *there exist constants* $C_0,q>0$ *such that for all* $k\in\{0,...,n-1\}$, $a_1,\ldots,a_j\in\mathbb{R}_0$, *and* $j\geq 1$,

$$\begin{aligned}
&\mathrm{E}\left[\mathbf{1}_{\{\widehat{J}_{j,k},a_1,\ldots,a_j\}}\left(\frac{\partial_\theta q^{\theta,\beta}_{(j)}}{q^{\theta,\beta}_{(j)}}(\Delta_n,X^{\bar{\theta},\bar{\beta}}_{t_k},X^{\bar{\theta},\bar{\beta}}_{t_{k+1}};a_1,\ldots,a_j)\right)^p\Big| X^{\bar{\theta},\bar{\beta}}_{t_k}=x\right]\\
&\leq C_0\Delta_n^{\frac{p}{2}}\left(C^{(2p_1-1)(j+1)}p_{a_1}\cdots p_{a_j}e^{-\lambda\Delta_n}\frac{(\lambda\Delta_n)^j}{j!}\right)^{\frac{1}{p_1}}(1+|x|^q),\\
&\mathrm{E}\left[\mathbf{1}_{\{\widehat{J}_{j,k},a_1,\ldots,a_j\}}\left(\frac{\partial_\beta q^{\theta,\beta}_{(j)}}{q^{\theta,\beta}_{(j)}}(\Delta_n,X^{\bar{\theta},\bar{\beta}}_{t_k},X^{\bar{\theta},\bar{\beta}}_{t_{k+1}};a_1,\ldots,a_j)\right)^p\Big| X^{\bar{\theta},\bar{\beta}}_{t_k}=x\right]\\
&\leq C_0\left(C^{(2p_1-1)(j+1)}p_{a_1}\cdots p_{a_j}e^{-\lambda\Delta_n}\frac{(\lambda\Delta_n)^j}{j!}\right)^{\frac{1}{p_1}}(1+|x|^q),
\end{aligned}$$

where C *is as in* (5.14), *and*

$$\begin{aligned}
&\mathrm{E}\left[\mathbf{1}_{\widehat{J}_{0,k}}\left(\frac{\partial_\theta q^{\theta,\beta}_{(0)}}{q^{\theta,\beta}_{(0)}}(\Delta_n,X^{\bar{\theta},\bar{\beta}}_{t_k},X^{\bar{\theta},\bar{\beta}}_{t_{k+1}})\right)^p\Big| X^{\bar{\theta},\bar{\beta}}_{t_k}=x\right]\leq C_0\Delta_n^{\frac{p}{2}}(1+|x|^q),\\
&\mathrm{E}\left[\mathbf{1}_{\widehat{J}_{0,k}}\left(\frac{\partial_\beta q^{\theta,\beta}_{(0)}}{q^{\theta,\beta}_{(0)}}(\Delta_n,X^{\bar{\theta},\bar{\beta}}_{t_k},X^{\bar{\theta},\bar{\beta}}_{t_{k+1}})\right)^p\Big| X^{\bar{\theta},\bar{\beta}}_{t_k}=x\right]\leq C_0(1+|x|^q).
\end{aligned}$$

Proof. Applying Lemma 5.2.7 and Jensen's inequality, for any $j \geq 1$ and $p > 1$, we obtain that

$$\mathrm{E}\left[\mathbf{1}_{\{\widehat{J}_{j,k},a_1,\ldots,a_j\}}\left(\frac{\partial_\beta q_{(j)}^{\theta,\beta}}{q_{(j)}^{\theta,\beta}}(\Delta_n, X_{t_k}^{\bar{\theta},\bar{\beta}}, X_{t_{k+1}}^{\bar{\theta},\bar{\beta}}; a_1,\ldots,a_j)\right)^p \Big| X_{t_k}^{\bar{\theta},\bar{\beta}} = x\right] = \mathrm{E}\Bigg[\mathbf{1}_{\{\widehat{J}_{j,k},a_1,\ldots,a_j\}}$$
$$\times\left(\frac{1}{\Delta_n}\widetilde{\mathrm{E}}_x^{\theta,\beta}\left[\delta\left(\partial_\beta Y_{t_{k+1}}^{\theta,\beta}(t_k,x)U^{\theta,\beta}(t_k,x)\right)\Big| Y_{t_{k+1}}^{\theta,\beta} = X_{t_{k+1}}^{\bar{\theta},\bar{\beta}}, \widetilde{J}_{j,k}, a_1,\ldots,a_j\right]\right)^p \Big| X_{t_k}^{\bar{\theta},\bar{\beta}} = x\Bigg]$$
$$\leq \frac{1}{\Delta_n^p}\int_{\mathbb{R}}\widetilde{\mathrm{E}}_x^{\theta,\beta}\left[\left|\delta\left(\partial_\beta Y_{t_{k+1}}^{\theta,\beta}(t_k,x)U^{\theta,\beta}(t_k,x)\right)\right|^p \Big| Y_{t_{k+1}}^{\theta,\beta} = y, \widetilde{J}_{j,k}, a_1,\ldots,a_j\right]$$
$$\times q_{(j)}^{\bar{\theta},\bar{\beta}}(\Delta_n,x,y;a_1,\ldots,a_j)p_{a_1}\cdots p_{a_j}e^{-\lambda\Delta_n}\frac{(\lambda\Delta_n)^j}{j!}dy$$
$$= \frac{1}{\Delta_n^p}\widetilde{\mathrm{E}}_x^{\theta,\beta}\left[\left|\delta\left(\partial_\beta Y_{t_{k+1}}^{\theta,\beta}(t_k,x)U^{\theta,\beta}(t_k,x)\right)\right|^p \mathbf{1}_{\{\widetilde{J}_{j,k},a_1,\ldots,a_j\}}\frac{q_{(j)}^{\bar{\theta},\bar{\beta}}}{q_{(j)}^{\theta,\beta}}(\Delta_n, Y_{t_k}^{\theta,\beta}, Y_{t_{k+1}}^{\theta,\beta}; a_1,\ldots,a_j)\right].$$

Then, applying Hölder's inequality with $\frac{1}{p_1}+\frac{1}{p_2}=1$, together with Lemma 5.2.6 and (5.11), we get that if p_1 is close to 1,

$$\mathrm{E}\left[\mathbf{1}_{\{\widehat{J}_{j,k},a_1,\ldots,a_j\}}\left(\frac{\partial_\beta q_{(j)}^{\theta,\beta}}{q_{(j)}^{\theta,\beta}}(\Delta_n, X_{t_k}^{\bar{\theta},\bar{\beta}}, X_{t_{k+1}}^{\bar{\theta},\bar{\beta}}; a_1,\ldots,a_j)\right)^p \Big| X_{t_k}^{\bar{\theta},\bar{\beta}} = x\right]$$
$$\leq \frac{1}{\Delta_n^p}\left(\widetilde{\mathrm{E}}_x^{\theta,\beta}\left[\left|\delta\left(\partial_\beta Y_{t_{k+1}}^{\theta,\beta}(t_k,x)U^{\theta,\beta}(t_k,x)\right)\right|^{pp_2}\right]\right)^{\frac{1}{p_2}}$$
$$\times\left(\widetilde{\mathrm{E}}_x^{\theta,\beta}\left[\mathbf{1}_{\{\widetilde{J}_{j,k},a_1,\ldots,a_j\}}\left(\frac{q_{(j)}^{\bar{\theta},\bar{\beta}}}{q_{(j)}^{\theta,\beta}}(\Delta_n, Y_{t_k}^{\theta,\beta}, Y_{t_{k+1}}^{\theta,\beta}; a_1,\ldots,a_j)\right)^{p_1}\right]\right)^{\frac{1}{p_1}}$$
$$\leq C_0\left(C^{(2p_1-1)(j+1)}p_{a_1}\cdots p_{a_j}e^{-\lambda\Delta_n}\frac{(\lambda\Delta_n)^j}{j!}\right)^{\frac{1}{p_1}}(1+|x|^q),$$

for some constants $C_0, q > 0$. This concludes the second inequality. The proof of the other inequalities follow along the same lines and are omitted. □

Lemma 5.2.9. *Assume conditions* **(A1)-(A4)**. *Then for any $q_1, q_2, q_3 > 1$ conjugate, q_3 close to 1, and $p_1 > 1$ close to 1, there exist constants $C_0, q > 0$ such that for any random variable Y, $k \in \{0, ..., n-1\}$, $a_1, \ldots, a_j \in \mathbb{R}_0$, and $j \geq 1$,*

$$\left|\mathrm{E}\left[Y\mathbf{1}_{\{\widehat{J}_{j,k},a_1,\ldots,a_j\}}\left(\frac{q_{(j)}^{\theta_0,\beta_0}}{q_{(j)}^{\theta_n,\beta(\ell)}}(\Delta_n, X_{t_k}^{\theta_n,\beta(\ell)}, X_{t_{k+1}}^{\theta_n,\beta(\ell)}; a_1,\ldots,a_j)-1\right)\Big| X_{t_k}^{\theta_n,\beta(\ell)} = x\right]\right|$$
$$\leq \frac{C_0}{\sqrt{n}}\left(\mathrm{E}\left[|Y|^{q_1}\big|X_{t_k}^{\theta_n,\beta(\ell)} = x\right]\right)^{\frac{1}{q_1}}\left(C^{(p_1\vee q_3)(j+1)}p_{a_1}\cdots p_{a_j}e^{-\lambda\Delta_n}\frac{(\lambda\Delta_n)^j}{j!}\right)^{\frac{1}{p_1q_2}+\frac{1}{q_3}}(1+|x|^q),$$

where C is as in (5.14), *and*

$$\left|\mathrm{E}\left[Y\mathbf{1}_{\widehat{J}_{0,k}}\left(\frac{q_{(0)}^{\theta_0,\beta_0}}{q_{(0)}^{\theta_n,\beta(\ell)}}(\Delta_n, X_{t_k}^{\theta_n,\beta(\ell)}, X_{t_{k+1}}^{\theta_n,\beta(\ell)})-1\right)\Big| X_{t_k}^{\theta_n,\beta(\ell)} = x\right]\right|$$
$$\leq \frac{C_0}{\sqrt{n}}\left(\mathrm{E}\left[|Y|^{q_1}\big|X_{t_k}^{\theta_n,\beta(\ell)} = x\right]\right)^{\frac{1}{q_1}}(1+|x|^q),$$

where $\beta(\ell) := \beta_0 + \dfrac{\ell v}{\sqrt{n}}$, and $\ell \in [0,1]$.

Proof. Observe that we can write

$$\frac{q_{(j)}^{\theta_0,\beta_0}}{q_{(j)}^{\theta_n,\beta(\ell)}}(\Delta_n, X_{t_k}^{\theta_n,\beta(\ell)}, X_{t_{k+1}}^{\theta_n,\beta(\ell)}; a_1,\ldots,a_j) - 1 = \frac{q_{(j)}^{\theta_0,\beta_0} - q_{(j)}^{\theta_0,\beta(\ell)} + q_{(j)}^{\theta_0,\beta(\ell)} - q_{(j)}^{\theta_n,\beta(\ell)}}{q_{(j)}^{\theta_n,\beta(\ell)}}$$

$$= -\frac{\ell v}{\sqrt{n}}\int_0^1 \frac{\partial_\beta q_{(j)}^{\theta_0,\beta_0+\frac{\ell vh}{\sqrt{n}}}}{q_{(j)}^{\theta_0,\beta_0+\frac{\ell vh}{\sqrt{n}}}}\frac{q_{(j)}^{\theta_0,\beta_0+\frac{\ell vh}{\sqrt{n}}}}{q_{(j)}^{\theta_n,\beta(\ell)}}dh - \frac{u}{\sqrt{n\Delta_n}}\int_0^1 \frac{\partial_\theta q_{(j)}^{\theta_0+\frac{uh}{\sqrt{n\Delta_n}},\beta(\ell)}}{q_{(j)}^{\theta_0+\frac{uh}{\sqrt{n\Delta_n}},\beta(\ell)}}\frac{q_{(j)}^{\theta_0+\frac{uh}{\sqrt{n\Delta_n}},\beta(\ell)}}{q_{(j)}^{\theta_n,\beta(\ell)}}dh.$$

Here, to simplify the exposition, let us write $\frac{q_{(j)}^{\theta_0,\beta_0}}{q_{(j)}^{\theta_n,\beta(\ell)}} \equiv \frac{q_{(j)}^{\theta_0,\beta_0}}{q_{(j)}^{\theta_n,\beta(\ell)}}(\Delta_n, X_{t_k}^{\theta_n,\beta(\ell)}, X_{t_{k+1}}^{\theta_n,\beta(\ell)}; a_1,\ldots,a_j)$. This implies that

$$\left|\mathrm{E}\left[Y\mathbf{1}_{\{\widehat{J}_{j,k},a_1,\ldots,a_j\}}\left(\frac{q_{(j)}^{\theta_0,\beta_0}}{q_{(j)}^{\theta_n,\beta(\ell)}}-1\right)\Big|X_{t_k}^{\theta_n,\beta(\ell)}=x\right]\right| \leq S_1+S_2,$$

where

$$S_1 = \frac{|v|}{\sqrt{n}}\int_0^1\left|\mathrm{E}\left[Y\mathbf{1}_{\{\widehat{J}_{j,k},a_1,\ldots,a_j\}}\frac{\partial_\beta q_{(j)}^{\theta_0,\beta_0+\frac{\ell vh}{\sqrt{n}}}}{q_{(j)}^{\theta_0,\beta_0+\frac{\ell vh}{\sqrt{n}}}}\frac{q_{(j)}^{\theta_0,\beta_0+\frac{\ell vh}{\sqrt{n}}}}{q_{(j)}^{\theta_n,\beta(\ell)}}\Big|X_{t_k}^{\theta_n,\beta(\ell)}=x\right]\right|dh,$$

$$S_2 = \frac{|u|}{\sqrt{n\Delta_n}}\int_0^1\left|\mathrm{E}\left[Y\mathbf{1}_{\{\widehat{J}_{j,k},a_1,\ldots,a_j\}}\frac{\partial_\theta q_{(j)}^{\theta_0+\frac{uh}{\sqrt{n\Delta_n}},\beta(\ell)}}{q_{(j)}^{\theta_0+\frac{uh}{\sqrt{n\Delta_n}},\beta(\ell)}}\frac{q_{(j)}^{\theta_0+\frac{uh}{\sqrt{n\Delta_n}},\beta(\ell)}}{q_{(j)}^{\theta_n,\beta(\ell)}}\Big|X_{t_k}^{\theta_n,\beta(\ell)}=x\right]\right|dh.$$

First, we treat S_1. Using Hölder's inequality with $q_1,q_2,q_3>1$ conjugate, together with Lemmas 5.2.6 and 5.2.8, we get that if q_3 is close to 1, for any $p_1>1$ close to 1,

$$S_1 \leq \frac{|v|}{\sqrt{n}}\int_0^1\left(\mathrm{E}\left[|Y|^{q_1}|X_{t_k}^{\theta_n,\beta(\ell)}=x\right]\right)^{\frac{1}{q_1}}\left(\mathrm{E}\left[\mathbf{1}_{\{\widehat{J}_{j,k},a_1,\ldots,a_j\}}\left(\frac{\partial_\beta q_{(j)}^{\theta_0,\beta_0+\frac{\ell vh}{\sqrt{n}}}}{q_{(j)}^{\theta_0,\beta_0+\frac{\ell vh}{\sqrt{n}}}}\right)^{q_2}\Big|X_{t_k}^{\theta_n,\beta(\ell)}=x\right]\right)^{\frac{1}{q_2}}$$

$$\times\left(\mathrm{E}\left[\mathbf{1}_{\{\widehat{J}_{j,k},a_1,\ldots,a_j\}}\left(\frac{q_{(j)}^{\theta_0,\beta_0+\frac{\ell vh}{\sqrt{n}}}}{q_{(j)}^{\theta_n,\beta(\ell)}}\right)^{q_3}\Big|X_{t_k}^{\theta_n,\beta(\ell)}=x\right]\right)^{\frac{1}{q_3}}dh$$

$$\leq C_0\frac{|v|}{\sqrt{n}}\left(\mathrm{E}\left[|Y|^{q_1}|X_{t_k}^{\theta_n,\beta(\ell)}=x\right]\right)^{\frac{1}{q_1}}\left(C^{(2p_1-1)(j+1)}p_{a_1}\cdots p_{a_j}e^{-\lambda\Delta_n}\frac{(\lambda\Delta_n)^j}{j!}\right)^{\frac{1}{p_1q_2}}$$

$$\times\left(C^{(2q_3-1)(j+1)}p_{a_1}\cdots p_{a_j}e^{-\lambda\Delta_n}\frac{(\lambda\Delta_n)^j}{j!}\right)^{\frac{1}{q_3}}(1+|x|^q),$$

for some constants $C_0,q>0$.

Similarly,

$$S_2 \leq C_0\frac{|u|}{\sqrt{n}}\left(\mathrm{E}\left[|Y|^{q_1}|X_{t_k}^{\theta_n,\beta(\ell)}=x\right]\right)^{\frac{1}{q_1}}\left(C^{(2p_1-1)(j+1)}p_{a_1}\cdots p_{a_j}e^{-\lambda\Delta_n}\frac{(\lambda\Delta_n)^j}{j!}\right)^{\frac{1}{p_1q_2}}$$

$$\times\left(C^{(2q_3-1)(j+1)}p_{a_1}\cdots p_{a_j}e^{-\lambda\Delta_n}\frac{(\lambda\Delta_n)^j}{j!}\right)^{\frac{1}{q_3}}(1+|x|^q).$$

This concludes the first inequality. The second one can be done similarly. Thus, the result follows. □

Lemma 5.2.10. *Assume conditions* **(A1)**-**(A3)**, *and let f be any bounded function. Then for any $k \in \{0, ..., n-1\}$, and $(\theta, \beta) \in \Theta \times \Sigma$,*

$$\mathrm{E}\left[f(X_{t_{k+1}})|X_{t_k}\right] = \mathrm{E}\left[f(X_{t_{k+1}}^{\theta_n,\beta(\ell)})\mathbf{1}_{\widehat{J}_{0,k}}\frac{q_{(0)}^{\theta_0,\beta_0}}{q_{(0)}^{\theta_n,\beta(\ell)}}(\Delta_n, X_{t_k}^{\theta_n,\beta(\ell)}, X_{t_{k+1}}^{\theta_n,\beta(\ell)})\Big|X_{t_k}^{\theta_n,\beta(\ell)} = X_{t_k}\right]$$
$$+\sum_{j=1}^{\infty}\sum_{(a_1,\ldots,a_j)\in A}\mathrm{E}\left[f(X_{t_{k+1}}^{\theta_n,\beta(\ell)})\mathbf{1}_{\{\widehat{J}_{j,k},a_1,\ldots,a_j\}}\frac{q_{(j)}^{\theta_0,\beta_0}}{q_{(j)}^{\theta_n,\beta(\ell)}}(\Delta_n, X_{t_k}^{\theta_n,\beta(\ell)}, X_{t_{k+1}}^{\theta_n,\beta(\ell)}; a_1, \ldots, a_j)\Big|X_{t_k}^{\theta_n,\beta(\ell)} = X_{t_k}\right].$$

Proof. Observe that

$$\mathrm{E}\left[f(X_{t_{k+1}})|X_{t_k}\right] = \mathrm{E}\left[\mathbf{1}_{\widehat{J}_{0,k}}f(X_{t_{k+1}})|X_{t_k}\right] + \sum_{j=1}^{\infty}\sum_{(a_1,\ldots,a_j)\in A}\mathrm{E}\left[\mathbf{1}_{\{\widehat{J}_{j,k},a_1,\ldots,a_j\}}f(X_{t_{k+1}})|X_{t_k}\right]$$
$$= \int_{\mathbb{R}} f(y)q_{(0)}^{\theta_0,\beta_0}(\Delta_n, X_{t_k}, y)e^{-\lambda\Delta_n}dy$$
$$+\sum_{j=1}^{\infty}\sum_{(a_1,\ldots,a_j)\in A}\int_{\mathbb{R}} f(y)q_{(j)}^{\theta_0,\beta_0}(\Delta_n, X_{t_k}, y; a_1, \ldots, a_j)p_{a_1}\cdots p_{a_j}e^{-\lambda\Delta_n}\frac{(\lambda\Delta_n)^j}{j!}dy$$
$$= \int_{\mathbb{R}} f(y)\frac{q_{(0)}^{\theta_0,\beta_0}}{q_{(0)}^{\theta_n,\beta(\ell)}}q_{(0)}^{\theta_n,\beta(\ell)}(\Delta_n, X_{t_k}, y)e^{-\lambda\Delta_n}dy$$
$$+\sum_{j=1}^{\infty}\sum_{(a_1,\ldots,a_j)\in A}\int_{\mathbb{R}} f(y)\frac{q_{(j)}^{\theta_0,\beta_0}}{q_{(j)}^{\theta_n,\beta(\ell)}}q_{(j)}^{\theta_n,\beta(\ell)}(\Delta_n, X_{t_k}, y; a_1, \ldots, a_j)p_{a_1}\cdots p_{a_j}e^{-\lambda\Delta_n}\frac{(\lambda\Delta_n)^j}{j!}dy,$$

which concludes the desired result. □

We next recall Girsanov's theorem on each interval $[t_k, t_{k+1}]$.

Lemma 5.2.11. *Under conditions* **(A1)** *and* **(A2)**, *for all $\theta, \theta_1 \in \Theta$, $\beta \in \Sigma$, and $k \in \{0, ..., n-1\}$, define a measure*

$$\widehat{Q}_k^{\theta_1,\theta,\beta} = \widehat{\mathrm{E}}\left[\mathbf{1}_A e^{-\int_{t_k}^{t_{k+1}}\frac{b(\theta,X_t)-b(\theta_1,X_t)}{\sigma(\beta,X_t)}dB_t + \frac{1}{2}\int_{t_k}^{t_{k+1}}\left(\frac{b(\theta,X_t)-b(\theta_1,X_t)}{\sigma(\beta,X_t)}\right)^2dt}\right],$$

for all $A \in \widehat{\mathcal{F}}$. Then $\widehat{Q}_k^{\theta_1,\theta,\beta}$ is a probability measure and under $\widehat{Q}_k^{\theta_1,\theta,\beta}$, the process $B_t^{\widehat{Q}_k^{\theta_1,\theta,\beta}} = B_t + \int_{t_k}^{t_{k+1}}\frac{b(\theta,X_t)-b(\theta_1,X_t)}{\sigma(\beta,X_t)}dt$ is a Brownian motion, for all $t \in [t_k, t_{k+1}]$.

Lemma 5.2.12. *Assume conditions* **(A1)**, **(A2)**, *and* **(A3)**(b). *Let $\theta, \theta_1 \in \Theta$ and $\beta \in \Sigma$ such that $|\theta - \theta_1| \leq \frac{C}{\sqrt{n\Delta_n}}$, for some constant $C > 0$. Then there exist constants $C, q > 0$ such that for any random variable V, and $k \in \{0, ..., n-1\}$,*

$$\left|\mathrm{E}_{\widehat{Q}_k^{\theta_1,\theta,\beta}}\left[V\left(\frac{d\widehat{\mathrm{P}}}{d\widehat{Q}_k^{\theta_1,\theta,\beta}} - 1\right)\Big|X_{t_k}^{\theta,\beta}\right]\right| \leq \frac{C}{\sqrt{n}}\left(1 + |X_{t_k}^{\theta,\beta}|^q\right)\int_0^1\left(\mathrm{E}_{\widehat{\mathrm{P}}^\alpha}\left[V^2\Big|X_{t_k}^{\theta,\beta}\right]\right)^{1/2}d\alpha,$$

where $\mathrm{E}_{\widehat{\mathrm{P}}^\alpha}$ denotes the expectation under the probability measure $\widehat{\mathrm{P}}^\alpha$ defined as

$$\frac{d\widehat{\mathrm{P}}^\alpha}{d\widehat{Q}_k^{\theta_1,\theta,\beta}} := e^{\alpha\int_{t_k}^{t_{k+1}}\frac{b(\theta,X_t)-b(\theta_1,X_t)}{\sigma(\beta,X_t)}dB_t - \frac{\alpha^2}{2}\int_{t_k}^{t_{k+1}}\left(\frac{b(\theta,X_t)-b(\theta_1,X_t)}{\sigma(\beta,X_t)}\right)^2dt},$$

for all $\alpha \in [0, 1]$.

Proof. Observe that

$$\frac{d\widehat{\mathrm{P}}}{d\widehat{Q}_k^{\theta_1,\theta,\beta}}-1=\int_0^1\int_{t_k}^{t_{k+1}}\frac{b(\theta,X_t)-b(\theta_1,X_t)}{\sigma(\beta,X_t)}\left(dB_t-\alpha\frac{b(\theta,X_t)-b(\theta_1,X_t)}{\sigma(\beta,X_t)}dt\right)\frac{d\widehat{\mathrm{P}}^{\alpha}}{d\widehat{Q}_k^{\theta_1,\theta,\beta}}d\alpha.$$

Consider the process $W=(W_t)_{t\in[t_k,t_{k+1}]}$ defined by

$$W_t:=B_t-\alpha\int_{t_k}^{t}\frac{b(\theta,X_s)-b(\theta_1,X_s)}{\sigma(\beta,X_s)}ds.$$

By Girsanov's theorem, W is a Brownian motion under $\widehat{\mathrm{P}}^{\alpha}$.

Then, using Girsanov's theorem, Cauchy-Schwarz inequality, and hypotheses **(A2)**, **(A3)**(b), together with Lemma 4.2.2 (ii), we get that

$$\begin{aligned}&\left|\mathrm{E}_{\widehat{Q}_k^{\theta_1,\theta,\beta}}\left[V\left(\frac{d\widehat{\mathrm{P}}}{d\widehat{Q}_k^{\theta_1,\theta,\beta}}-1\right)\Big|X_{t_k}^{\theta,\beta}\right]\right|=\left|\int_0^1\mathrm{E}_{\widehat{\mathrm{P}}^{\alpha}}\left[V\int_{t_k}^{t_{k+1}}\frac{b(\theta,X_t)-b(\theta_0,X_t)}{\sigma(\beta,X_t)}dW_t\Big|X_{t_k}^{\theta,\beta}\right]d\alpha\right|\\&\leq\int_0^1\left(\mathrm{E}_{\widehat{\mathrm{P}}^{\alpha}}\left[V^2\Big|X_{t_k}^{\theta,\beta}\right]\right)^{1/2}\left(\mathrm{E}_{\widehat{\mathrm{P}}^{\alpha}}\left[\left|\int_{t_k}^{t_{k+1}}\frac{b(\theta,X_t)-b(\theta_0,X_t)}{\sigma(\beta,X_t)}dW_t\right|^2\Big|X_{t_k}^{\theta,\beta}\right]\right)^{1/2}d\alpha\\&\leq\frac{C}{\sqrt{n}}\left(1+|X_{t_k}^{\theta,\beta}|^q\right)\int_0^1\left(\mathrm{E}_{\widehat{\mathrm{P}}^{\alpha}}\left[V^2\Big|X_{t_k}^{\theta,\beta}\right]\right)^{1/2}d\alpha,\end{aligned}$$

for some constants $C,q>0$. Thus, the result follows. ☐

Lemma 5.2.13. *Under conditions* **(A1)**-**(A3)**, *for all* $k\in\{0,...,n-1\}$ *and* $\theta\in\Theta$,

$$\begin{aligned}&\mathrm{E}_{\widehat{Q}_k^{\theta,\theta_0,\beta_0}}\left[\mathbf{1}_{\widehat{J}_{0,k}}\left(\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta,\beta_0}\left[\mathbf{1}_{\widetilde{J}_{1,k}}\int_{t_k}^{t_{k+1}}\int_{\mathbb{R}_0}zM(ds,dz)\Big|Y_{t_{k+1}}^{\theta,\beta_0}=X_{t_{k+1}}\right]\right)^2\Big|X_{t_k}\right]\\&=\int_{\mathbb{R}}\left(\frac{\int_{\mathbb{R}_0}q_{(1)}^{\theta,\beta_0}(\Delta_n,X_{t_k},y;a)a\mu(da)e^{-\lambda\Delta_n}\lambda\Delta_n}{p^{\theta,\beta_0}(\Delta_n,X_{t_k},y)}\right)^2q_{(0)}^{\theta,\beta_0}(\Delta_n,X_{t_k},y)e^{-\lambda\Delta_n}dy,\end{aligned}\tag{5.17}$$

$$\begin{aligned}&\mathrm{E}_{\widehat{Q}_k^{\theta,\theta_0,\beta_0}}\left[\mathbf{1}_{\widehat{J}_{1,k}}\left(\int_{t_k}^{t_{k+1}}\int_{\mathbb{R}_0}zN(ds,dz)\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta,\beta_0}\left[\mathbf{1}_{\widetilde{J}_{0,k}}\Big|Y_{t_{k+1}}^{\theta,\beta_0}=X_{t_{k+1}}\right]\right)^2\Big|X_{t_k}\right]\\&=\int_{\mathbb{R}_0}\int_{\mathbb{R}}\left(\frac{q_{(0)}^{\theta,\beta_0}(\Delta_n,X_{t_k},y)e^{-\lambda\Delta_n}}{p^{\theta,\beta_0}(\Delta_n,X_{t_k},y)}\right)^2q_{(1)}^{\theta,\beta_0}(\Delta_n,X_{t_k},y;a)e^{-\lambda\Delta_n}\lambda\Delta_na^2dy\mu(da),\end{aligned}\tag{5.18}$$

and

$$\begin{aligned}&\mathrm{E}_{\widehat{Q}_k^{\theta,\theta_0,\beta_0}}\Bigg[\mathbf{1}_{\widehat{J}_{1,k}}\Bigg(\int_{t_k}^{t_{k+1}}\int_{\mathbb{R}_0}zN(ds,dz)\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta,\beta_0}\left[\mathbf{1}_{\widetilde{J}_{1,k}}\Big|Y_{t_{k+1}}^{\theta,\beta_0}=X_{t_{k+1}}\right]\\&\qquad-\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta,\beta_0}\left[\mathbf{1}_{\widetilde{J}_{1,k}}\int_{t_k}^{t_{k+1}}\int_{\mathbb{R}_0}zM(ds,dz)\Big|Y_{t_{k+1}}^{\theta,\beta_0}=X_{t_{k+1}}\right]\Bigg)^2\Big|X_{t_k}\Bigg]\\&=\int_{\mathbb{R}_0}\int_{\mathbb{R}}\left(\frac{\int_{\mathbb{R}_0}(z-a)\,q_{(1)}^{\theta,\beta_0}(\Delta_n,X_{t_k},y;a)\mu(da)e^{-\lambda\Delta_n}\lambda\Delta_n}{p^{\theta,\beta_0}(\Delta_n,X_{t_k},y)}\right)^2\\&\qquad\times q_{(1)}^{\theta,\beta_0}(\Delta_n,X_{t_k},y;z)e^{-\lambda\Delta_n}\lambda\Delta_ndy\mu(dz).\end{aligned}\tag{5.19}$$

Proof. Using Bayes' formula, we get that

$$\begin{aligned}
&\widetilde{\mathrm{E}}^{\theta,\beta_0}_{X_{t_k}}\left[\mathbf{1}_{\widetilde{J}_{1,k}}\int_{t_k}^{t_{k+1}}\int_{\mathbb{R}_0} zM(ds,dz)\middle| Y^{\theta,\beta_0}_{t_{k+1}}=X_{t_{k+1}}\right]\\
&=\frac{\widetilde{\mathrm{E}}^{\theta,\beta_0}_{X_{t_k}}\left[\widetilde{\Lambda}_{[t_k,t_{k+1}]}\mathbf{1}_{\{Y^{\theta,\beta_0}_{t_{k+1}}=X_{t_{k+1}}\}}\middle|\widetilde{J}_{1,k}\right]\widetilde{\mathrm{P}}^{\theta,\beta_0}_{X_{t_k}}\left(\widetilde{J}_{1,k}\right)}{p^{\theta,\beta_0}(\Delta_n,X_{t_k},X_{t_{k+1}})}\\
&=\frac{\int_{\mathbb{R}_0} q^{\theta,\beta_0}_{(1)}(\Delta_n,X_{t_k},X_{t_{k+1}};a)a\mu(da)e^{-\lambda\Delta_n}\lambda\Delta_n}{p^{\theta,\beta_0}(\Delta_n,X_{t_k},X_{t_{k+1}})}.
\end{aligned}$$

This, together with Bayes' formula again, implies that

$$\begin{aligned}
&\mathrm{E}_{\widehat{Q}^{\theta,\theta_0,\beta_0}_k}\left[\mathbf{1}_{\widehat{J}_{0,k}}\left(\widetilde{\mathrm{E}}^{\theta,\beta_0}_{X_{t_k}}\left[\mathbf{1}_{\widetilde{J}_{1,k}}\int_{t_k}^{t_{k+1}}\int_{\mathbb{R}_0} zM(ds,dz)\middle| Y^{\theta,\beta_0}_{t_{k+1}}=X_{t_{k+1}}\right]\right)^2\middle| X_{t_k}\right]\\
&=\widehat{Q}^{\theta,\theta_0,\beta_0}_k\left(\widehat{J}_{0,k}\middle|X_{t_k}\right)\mathrm{E}_{\widehat{Q}^{\theta,\theta_0,\beta_0}_k}\left[\left(\frac{\int_{\mathbb{R}_0} q^{\theta,\beta_0}_{(1)}(\Delta_n,X_{t_k},X_{t_{k+1}};a)a\mu(da)e^{-\lambda\Delta_n}\lambda\Delta_n}{p^{\theta,\beta_0}(\Delta_n,X_{t_k},X_{t_{k+1}})}\right)^2\middle|\widehat{J}_{0,k},X_{t_k}\right],
\end{aligned}$$

which implies (5.17). Similarly,

$$\begin{aligned}
&\mathrm{E}_{\widehat{Q}^{\theta,\theta_0,\beta_0}_k}\left[\mathbf{1}_{\widehat{J}_{1,k}}\left(\int_{t_k}^{t_{k+1}}\int_{\mathbb{R}_0} zN(ds,dz)\widetilde{\mathrm{E}}^{\theta,\beta_0}_{X_{t_k}}\left[\mathbf{1}_{\widetilde{J}_{0,k}}\middle|Y^{\theta,\beta_0}_{t_{k+1}}=X_{t_{k+1}}\right]\right)^2\middle|X_{t_k}\right]\\
&=\mathrm{E}_{\widehat{Q}^{\theta,\theta_0,\beta_0}_k}\left[\mathbf{1}_{\widehat{J}_{1,k}}\widehat{\Lambda}^2_{[t_k,t_{k+1}]}\left(\frac{q^{\theta,\beta_0}_{(0)}(\Delta_n,X_{t_k},X_{t_{k+1}})e^{-\lambda\Delta_n}}{p^{\theta,\beta_0}(\Delta_n,X_{t_k},X_{t_{k+1}})}\right)^2\middle|X_{t_k}\right]\\
&=\int_{\mathbb{R}_0}\mathrm{E}_{\widehat{Q}^{\theta,\theta_0,\beta_0}_k}\left[\left(\frac{q^{\theta,\beta_0}_{(0)}(\Delta_n,X_{t_k},X_{t_{k+1}})e^{-\lambda\Delta_n}}{p^{\theta,\beta_0}(\Delta_n,X_{t_k},X_{t_{k+1}})}\right)^2\middle|\widehat{J}_{1,k},\widehat{\Lambda}_{[t_k,t_{k+1}]}=\{a\},X_{t_k}\right]a^2\\
&\qquad\times\widehat{Q}^{\theta,\theta_0,\beta_0}_k\left(\widehat{\Lambda}_{[t_k,t_{k+1}]}\in da,\widehat{J}_{1,k}\middle|X_{t_k}\right)\\
&=\int_{\mathbb{R}_0}\int_{\mathbb{R}}\left(\frac{q^{\theta,\beta_0}_{(0)}(\Delta_n,X_{t_k},y)e^{-\lambda\Delta_n}}{p^{\theta,\beta_0}(\Delta_n,X_{t_k},y)}\right)^2 q^{\theta,\beta_0}_{(1)}(\Delta_n,X_{t_k},y;a)e^{-\lambda\Delta_n}\lambda\Delta_n a^2dy\mu(da),
\end{aligned}$$

which shows (5.18). The proof of (5.19) follows along the same lines and is therefore omitted. □

By abuse of notation in this subsection relating to the term M_i^{θ,β_0} below, consider the events $\widehat{J}_{2,k}=\{N_{t_{k+1}}-N_{t_k}\geq 2\}$ and $\widetilde{J}_{2,k}=\{M_{t_{k+1}}-M_{t_k}\geq 2\}$. Set $I=\{a\in A:\rho_1\Delta_n^{\upsilon}\leq|a|\leq\rho_2\Delta_n^{-\gamma}\}$ and $\lambda_n=\int_I\nu(da)$, where $\rho_1,\rho_2>0$ and $0<\upsilon,\gamma<\frac{1}{2}$ are from hypothesis **(A6)**. For $i=0,1,2$, set

$$\begin{aligned}
M_i^{\theta,\beta_0}=\mathrm{E}_{\widehat{Q}^{\theta,\theta_0,\beta_0}_k}&\left[\mathbf{1}_{\widehat{J}_{i,k}}\left(\int_{t_k}^{t_{k+1}}\int_I zN(ds,dz)\right.\right.\\
&\left.\left.-\widetilde{\mathrm{E}}^{\theta,\beta_0}_{X_{t_k}}\left[\int_{t_k}^{t_{k+1}}\int_I zM(ds,dz)\middle|Y^{\theta,\beta_0}_{t_{k+1}}=X_{t_{k+1}}\right]\right)^2\middle|X_{t_k}\right].
\end{aligned}$$

Recall that for the simple Lévy process (3.1), we used a large deviation principle by conditioning on the number of jumps. For the non-linear model (5.1), we will obtain the parallel of Lemma 3.2.6 in our case.

Lemma 5.2.14. *Under conditions* **(A1)-(A4)**, *for any* $\theta \in \Theta$ *and* n *large enough, there exist constants* $C, C_0 > 0$, *such that for all* $\alpha \in (\upsilon, \frac{1}{2})$, $\alpha_0 \in (\frac{1}{4}, \frac{1}{2})$, *and* $k \in \{0, ..., n-1\}$,

$$M_0^{\theta,\beta_0} \leq C\left(\lambda_n \Delta_n^{3/2} + \Delta_n^{-2\gamma} e^{-C_0 \Delta_n^{2\alpha-1}}\right), \tag{5.20}$$

$$M_1^{\theta,\beta_0} \leq C\left(\lambda_n \Delta_n^{3/2} + \Delta_n^{-\frac{1}{2}-3\gamma} e^{-C_0 \Delta_n^{2(\alpha\vee\alpha_0)-1}}\right), \tag{5.21}$$

$$M_2^{\theta,\beta_0} \leq C\lambda_n \Delta_n^{3/2}. \tag{5.22}$$

In particular, (5.22) *holds for all* $n \geq 1$.

Proof. We start showing (5.20). Multiplying the random variable inside the conditional expectation of M_0^{θ,β_0} by $(\mathbf{1}_{\widetilde{J}_{0,k}} + \mathbf{1}_{\widetilde{J}_{1,k}} + \mathbf{1}_{\widetilde{J}_{2,k}})$, we get that $M_0^{\theta,\beta_0} \leq 2(M_{0,1}^{\theta,\beta_0} + M_{0,2}^{\theta,\beta_0})$, where for $i = 1, 2$,

$$M_{0,i}^{\theta,\beta_0} = \mathrm{E}_{\widehat{Q}_k^{\theta,\theta_0,\beta_0}}\left[\mathbf{1}_{\widehat{J}_{0,k}}\left(\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta,\beta_0}\left[\mathbf{1}_{\widetilde{J}_{i,k}}\int_{t_k}^{t_{k+1}}\int_I zM(ds,dz)\Big|Y_{t_{k+1}}^{\theta,\beta_0} = X_{t_{k+1}}\right]\right)^2\Big|X_{t_k}\right].$$

By (5.17), we have that

$$M_{0,1}^{\theta,\beta_0} = \int_{\mathbb{R}}\left(\frac{\int_I q_{(1)}^{\theta,\beta_0}(\Delta_n, X_{t_k}, y; a)a\mu(da)e^{-\lambda_n\Delta_n}\lambda_n\Delta_n}{p^{\theta,\beta_0}(\Delta_n, X_{t_k}, y)}\right)^2 q_{(0)}^{\theta,\beta_0}(\Delta_n, X_{t_k}, y)e^{-\lambda_n\Delta_n}dy,$$

We next divide the integral in $M_{0,1}^{\theta,\beta_0}$ into the subdomains $\{y : |y-X_{t_k}| > \Delta_n^\alpha\}$ and $\{y : |y-X_{t_k}| \leq \Delta_n^\alpha\}$, where $\alpha \in (\upsilon, \frac{1}{2})$, and call each integral $M_{0,1,1}^{\theta,\beta_0}$ and $M_{0,1,2}^{\theta,\beta_0}$. We start bounding $M_{0,1,1}^{\theta,\beta_0}$. By (5.12),

$$p^{\theta,\beta_0}(\Delta_n, X_{t_k}, y) \geq \int_I q_{(1)}^{\theta,\beta_0}(\Delta_n, X_{t_k}, y; a)\mu(da)e^{-\lambda_n\Delta_n}\lambda_n\Delta_n. \tag{5.23}$$

Then, using (5.13), we get that

$$\begin{aligned} M_{0,1,1}^{\theta,\beta_0} &\leq C\Delta_n^{-2\gamma}\int_{\{|y-X_{t_k}|>\Delta_n^\alpha\}} q_{(0)}^{\theta,\beta_0}(\Delta_n, X_{t_k}, y)dy \\ &\leq C\Delta_n^{-2\gamma}\int_{\{|y-X_{t_k}|>\Delta_n^\alpha\}} \frac{1}{\sqrt{\Delta_n}} e^{-\frac{(y-X_{t_k})^2}{c\Delta_n}} dy \leq C\Delta_n^{-2\gamma}e^{-C_0\Delta_n^{2\alpha-1}}, \end{aligned}$$

for some constants $C, C_0 > 0$ and $c \geq 1$. We next treat $M_{0,1,2}^{\theta,\beta_0}$. Observe that (5.12) yields

$$\left(p^{\theta,\beta_0}(\Delta_n, X_{t_k}, y)\right)^2 \geq q_{(0)}^{\theta,\beta_0}(\Delta_n, X_{t_k}, y)e^{-\lambda_n\Delta_n}\int_I q_{(1)}^{\theta,\beta_0}(\Delta_n, X_{t_k}, y; a)\mu(da)e^{-\lambda_n\Delta_n}\lambda_n\Delta_n. \tag{5.24}$$

Therefore, using Lemma 5.2.5, we get that for n large enough

$$\begin{aligned} M_{0,1,2}^{\theta,\beta_0} &\leq C\Delta_n^{-2\gamma}e^{-\lambda_n\Delta_n}\lambda_n\Delta_n\int_{\{|y-X_{t_k}|\leq\Delta_n^\alpha\}}\int_I q_{(1)}^{\theta,\beta_0}(\Delta_n, X_{t_k}, y; a)\mu(da)dy \\ &\leq C\Delta_n^{-2\gamma}\int_I\int_{\{|y-X_{t_k}|\leq\Delta_n^\alpha\}}\frac{1}{\sqrt{\Delta_n}}e^{-\frac{(y-X_{t_k}-a)^2}{c\Delta_n}}dy\mu(da) \\ &\leq C\Delta_n^{-2\gamma}\int_I\left\{\int_{-\infty}^{\frac{\Delta_n^\alpha-\rho_1\Delta_n^\upsilon}{\sqrt{c\Delta_n}}} e^{-w^2}dw\mathbf{1}_{\{a\geq\rho_1\Delta_n^\upsilon\}} + \int_{\frac{-\Delta_n^\alpha+\rho_1\Delta_n^\upsilon}{\sqrt{c\Delta_n}}}^{+\infty} e^{-w^2}dw\mathbf{1}_{\{a\leq-\rho_1\Delta_n^\upsilon\}}\right\}\mu(da) \\ &\leq C\Delta_n^{-2\gamma}e^{-C_0\Delta_n^{2\upsilon-1}}, \end{aligned}$$

for some constants $C, C_0 > 0$ and $c \geq 1$, where we have applied Fubini's theorem, the change of variables $w = \frac{y - X_{t_k} - a}{\sqrt{c\Delta_n}}$, and the fact that on I, $|a| \geq \rho_1 \Delta_n^{\upsilon}$, together with $e^{-\lambda_n \Delta_n} \lambda_n \Delta_n \leq \lambda$. This shows that for n large enough and for $\alpha \in (\upsilon, \frac{1}{2})$,

$$M_{0,1}^{\theta,\beta_0} \leq C\Delta_n^{-2\gamma} e^{-C_0 \Delta_n^{2\alpha - 1}}, \tag{5.25}$$

for some constants $C, C_0 > 0$.

In order to treat $M_{0,2}^{\theta,\beta_0}$, observe that by Jensen and Cauchy-Schwarz inequalities, and hypothesis **(A4)**, it holds that

$$M_{0,2}^{\theta,\beta_0} \leq \mathrm{E}\left[\mathbf{1}_{\widetilde{J}_{2,k}} \left(\int_{t_k}^{t_{k+1}} \int_I z M(ds, dz)\right)^2 \middle| Y_{t_k}^{\theta,\beta_0} = X_{t_k}\right] \leq C\lambda_n \Delta_n^{3/2}.$$

This shows (5.20).

We next show (5.21). As for the term M_0^{θ,β_0}, we have that $M_1^{\theta,\beta_0} \leq 2(M_{1,1}^{\theta,\beta_0} + M_{1,2}^{\theta,\beta_0})$, where

$$\begin{aligned} M_{1,1}^{\theta,\beta_0} &= \mathrm{E}_{\widehat{Q}_k^{\theta,\theta_0,\beta_0}}\Bigg[\mathbf{1}_{\widehat{J}_{1,k}}\Bigg(\int_{t_k}^{t_{k+1}} \int_I z N(ds, dz) \\ &\quad - \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta,\beta_0}\left[\mathbf{1}_{\widetilde{J}_{1,k}} \int_{t_k}^{t_{k+1}} \int_I z M(ds, dz) \middle| Y_{t_{k+1}}^{\theta,\beta_0} = X_{t_{k+1}}\right]\Bigg)^2 \Bigg| X_{t_k}\Bigg], \\ M_{1,2}^{\theta,\beta_0} &= \mathrm{E}_{\widehat{Q}_k^{\theta,\theta_0,\beta_0}}\left[\mathbf{1}_{\widehat{J}_{1,k}}\left(\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta,\beta_0}\left[\mathbf{1}_{\widetilde{J}_{2,k}} \int_{t_k}^{t_{k+1}} \int_I z M(ds, dz) \middle| Y_{t_{k+1}}^{\theta,\beta_0} = X_{t_{k+1}}\right]\right)^2 \middle| X_{t_k}\right]. \end{aligned}$$

Adding and substracting the term $\int_{t_k}^{t_{k+1}} \int_I z N(ds, dz) \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta,\beta_0}\left[\mathbf{1}_{\widetilde{J}_{1,k}} | Y_{t_{k+1}}^{\theta,\beta_0} = X_{t_{k+1}}\right]$ inside the square to start bounding $M_{1,1}^{\theta,\beta_0}$, we get that $M_{1,1}^{\theta,\beta_0} \leq 2(M_{1,1,1}^{\theta,\beta_0} + M_{1,1,2}^{\theta,\beta_0})$, where

$$\begin{aligned} M_{1,1,1}^{\theta,\beta_0} &= \mathrm{E}_{\widehat{Q}_k^{\theta,\theta_0,\beta_0}}\Bigg[\mathbf{1}_{\widehat{J}_{1,k}}\Bigg(\int_{t_k}^{t_{k+1}} \int_I z N(ds, dz) \\ &\quad - \int_{t_k}^{t_{k+1}} \int_I z N(ds, dz) \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta,\beta_0}\left[\mathbf{1}_{\widetilde{J}_{1,k}} | Y_{t_{k+1}}^{\theta,\beta_0} = X_{t_{k+1}}\right]\Bigg)^2 \Bigg| X_{t_k}\Bigg], \\ M_{1,1,2}^{\theta,\beta_0} &= \mathrm{E}_{\widehat{Q}_k^{\theta,\theta_0,\beta_0}}\Bigg[\mathbf{1}_{\widehat{J}_{1,k}}\Bigg(\int_{t_k}^{t_{k+1}} \int_I z N(ds, dz) \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta,\beta_0}\left[\mathbf{1}_{\widetilde{J}_{1,k}} | Y_{t_{k+1}}^{\theta,\beta_0} = X_{t_{k+1}}\right] \\ &\quad - \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta,\beta_0}\left[\mathbf{1}_{\widetilde{J}_{1,k}} \int_{t_k}^{t_{k+1}} \int_I z M(ds, dz) \middle| Y_{t_{k+1}}^{\theta,\beta_0} = X_{t_{k+1}}\right]\Bigg)^2 \Bigg| X_{t_k}\Bigg]. \end{aligned}$$

Observe that $M_{1,1,1}^{\theta,\beta_0} \leq 2(M_{1,1,1,0}^{\theta,\beta_0} + M_{1,1,1,2}^{\theta,\beta_0})$, where for $i = 0, 2$,

$$M_{1,1,1,i}^{\theta,\beta_0} = \mathrm{E}_{\widehat{Q}_k^{\theta,\theta_0,\beta_0}}\left[\mathbf{1}_{\widehat{J}_{1,k}}\left(\int_{t_k}^{t_{k+1}} \int_I z N(ds, dz) \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta,\beta_0}\left[\mathbf{1}_{\widetilde{J}_{i,k}} \middle| Y_{t_{k+1}}^{\theta,\beta_0} = X_{t_{k+1}}\right]\right)^2 \middle| X_{t_k}\right].$$

By (5.18),

$$M_{1,1,1,0}^{\theta,\beta_0} = \int_I \int_{\mathbb{R}} \left(\frac{q_{(0)}^{\theta,\beta_0}(\Delta_n, X_{t_k}, y) e^{-\lambda_n \Delta_n}}{p^{\theta}(\Delta_n, X_{t_k}, y)}\right)^2 q_{(1)}^{\theta,\beta_0}(\Delta_n, X_{t_k}, y; a) e^{-\lambda_n \Delta_n} \lambda_n \Delta_n a^2 dy \mu(da).$$

Again we divide the dy integral into the subdomains $\{y : |y - X_{t_k}| > \Delta_n^{\alpha}\}$ and $\{y : |y - X_{t_k}| \leq \Delta_n^{\alpha}\}$, where $\alpha \in (\upsilon, \frac{1}{2})$, and call the terms $M_{1,1,1,0,1}^{\theta,\beta_0}$ and $M_{1,1,1,0,2}^{\theta,\beta_0}$. As for the term $M_{0,1,1}^{\theta,\beta_0}$, using (5.24), we obtain that

$$M_{1,1,1,0,1}^{\theta,\beta_0} \leq C\Delta_n^{-2\gamma} \int_{\left\{|y - X_{t_k}| > \Delta_n^{\alpha}\right\}} q_{(0)}^{\theta,\beta_0}(\Delta_n, X_{t_k}, y) dy \leq C\Delta_n^{-2\gamma} e^{-C_0 \Delta_n^{2\alpha - 1}},$$

for some constants $C, C_0 > 0$. Next, (5.12) yields

$$p^{\theta,\beta_0}(\Delta_n, X_{t_k}, y) \geq q_{(0)}^{\theta,\beta_0}(\Delta_n, X_{t_k}, y)e^{-\lambda_n\Delta_n}. \tag{5.26}$$

Then, as for the term $M_{0,1,2}^{\theta,\beta_0}$, using Lemma 5.2.5, we get that for n large enough

$$\begin{aligned} M_{1,1,1,0,2}^{\theta,\beta_0} &\leq C\Delta_n^{-2\gamma}e^{-\lambda_n\Delta_n}\lambda_n\Delta_n \int_I \int_{\{|y-X_{t_k}|\leq \Delta_n^{\alpha}\}} q_{(1)}^{\theta,\beta_0}(\Delta_n, X_{t_k}, y; a)dy\mu(da) \\ &\leq C\Delta_n^{-2\gamma}e^{-C_0\Delta_n^{2\upsilon-1}}, \end{aligned}$$

for some constants $C, C_0 > 0$. Therefore, the term $M_{1,1,1,0}^{\theta,\beta_0}$ satisfies (5.25).

As for the term $M_{0,2}^{\theta,\beta_0}$, we have that $M_{1,1,1,2}^{\theta,\beta_0} \leq C\lambda_n\Delta_n^{3/2}$ for some constant $C > 0$. Therefore, the term $M_{1,1,1}^{\theta,\beta_0}$ satisfies (5.20).

We next treat $M_{1,1,2}^{\theta,\beta_0}$. Using (5.19), we have that

$$\begin{aligned} M_{1,1,2}^{\theta,\beta_0} = \int_I \int_{\mathbb{R}} &\left(\frac{\int_I (z-a)\, q_{(1)}^{\theta,\beta_0}(\Delta_n, X_{t_k}, y; a)\mu(da)e^{-\lambda_n\Delta_n}\lambda_n\Delta_n}{p^{\theta}(\Delta_n, X_{t_k}, y)} \right)^2 \\ &\times q_{(1)}^{\theta,\beta_0}(\Delta_n, X_{t_k}, y; z)e^{-\lambda_n\Delta_n}\lambda_n\Delta_n dy\mu(dz). \end{aligned}$$

We next fix α_0 and ε such that $\frac{1}{4} < \varepsilon < \alpha_0 < \frac{1}{2}$, and consider the set

$$E_z^k = \{a \in I : |z-a| \leq \Delta_n^{\varepsilon}, \text{ for all } z \in I\}.$$

We next split the integral inside the square of $M_{1,1,2}^{\theta,\beta_0}$ over the sets $\mathbf{1}_{E_z^k}$ and $\mathbf{1}_{(E_z^k)^c}$ and call both terms $M_{1,1,2,1}^{\theta,\beta_0}$ and $M_{1,1,2,2}^{\theta,\beta_0}$. First, (5.23) and Lemma 5.2.5 yield that

$$M_{1,1,2,1}^{\theta,\beta_0} \leq Ce^{-\lambda_n\Delta_n}\lambda_n\Delta_n^{1+2\varepsilon} \int_I \int_{\mathbb{R}} q_{(1)}^{\theta,\beta_0}(\Delta_n, X_{t_k}, y; z)dy\mu(dz) \leq C\lambda_n\Delta_n^{1+2\varepsilon}, \tag{5.27}$$

for some constant $C > 0$.

Next, we treat $M_{1,1,2,2}^{\theta,\beta_0}$ by dividing the domain of the dy integral into the subdomains $I_1 := \{y : |y - X_{t_k} - z| > \Delta_n^{\alpha_0}\}$ and $I_2 := \{y : |y - X_{t_k} - z| \leq \Delta_n^{\alpha_0}\}$, and call both terms $M_{1,1,2,2,1}^{\theta,\beta_0}$ and $M_{1,1,2,2,2}^{\theta,\beta_0}$. Then, using (5.23) and Lemma 5.2.5, we get that

$$\begin{aligned} M_{1,1,2,2,1}^{\theta,\beta_0} &\leq C\Delta_n^{-2\gamma}e^{-\lambda_n\Delta_n}\lambda_n\Delta_n \int_I \int_{I_1} q_{(1)}^{\theta,\beta_0}(\Delta_n, X_{t_k}, y; z)dy\mu(dz) \\ &\leq C\Delta_n^{-2\gamma} \int_I \int_{I_1} \frac{1}{\sqrt{\Delta_n}} e^{-\frac{(y-X_{t_k}-z)^2}{c\Delta_n}} dy\mu(dz) \leq C\Delta_n^{-2\gamma}e^{-C_0\Delta_n^{2\alpha_0-1}}, \end{aligned}$$

for some constants $C, C_0 > 0$ and $c \geq 1$.

Next, (5.12) yields

$$\left(p^{\theta,\beta_0}(\Delta_n, X_{t_k}, y)\right)^2 \geq p^{\theta,\beta_0}(\Delta_n, X_{t_k}, y) \int_I q_{(1)}^{\theta,\beta_0}(\Delta_n, X_{t_k}, y; a)\mu(da)e^{-\lambda_n\Delta_n}\lambda_n\Delta_n.$$

Then, using Lemma 5.2.5, we obtain that

$$
\begin{aligned}
&M_{1,1,2,2,2}^{\theta,\beta_0} \leq C\Delta_n^{-2\gamma} e^{-\lambda_n\Delta_n}\lambda_n\Delta_n \\
&\quad\times \int_I\int_{I_2}\int_I \mathbf{1}_{(E_z^k)^c} q_{(1)}^{\theta,\beta_0}(\Delta_n, X_{t_k}, y; a)\mu(da) \frac{q_{(1)}^{\theta,\beta_0}(\Delta_n, X_{t_k}, y; z)e^{-\lambda_n\Delta_n}\lambda_n\Delta_n}{p^{\theta,\beta_0}(\Delta_n, X_{t_k}, y)} dy\mu(dz) \\
&\leq C\Delta_n^{-2\gamma}\int_I\int_{I_2}\int_I \mathbf{1}_{(E_z^k)^c}\frac{1}{\sqrt{\Delta_n}} e^{-\frac{\left(y-X_{t_k}-a\right)^2}{c\Delta_n}}\mu(da)\frac{q_{(1)}^{\theta,\beta_0}(\Delta_n, X_{t_k}, y; z)e^{-\lambda_n\Delta_n}\lambda_n\Delta_n}{p^{\theta,\beta_0}(\Delta_n, X_{t_k}, y)} dy\mu(dz) \\
&\leq C\Delta_n^{-2\gamma}\int_I\int_I\int_{\{|h|\leq\Delta_n^{\alpha_0}\}} \mathbf{1}_{(E_z^k)^c}\frac{1}{\sqrt{\Delta_n}} e^{-\frac{(h+z-a)^2}{c\Delta_n}} \\
&\quad\times\frac{q_{(1)}^{\theta,\beta_0}(\Delta_n, X_{t_k}, h+X_{t_k}+z; z)e^{-\lambda_n\Delta_n}\lambda_n\Delta_n}{p^{\theta,\beta_0}(\Delta_n, X_{t_k}, h+X_{t_k}+z, z))} dh\mu(da)\mu(dz),
\end{aligned}
$$

for some constants $C > 0$ and $c \geq 1$, where we have used the change of variable $h := y - X_{t_k} - z$.

Since $|h| \leq \Delta_n^{\alpha_0}$ and $|z-a| > \Delta_n^{\varepsilon}$ on $(E_z^k)^c$, for n large enough there exists a constant $C_1 \in (0,1)$ such that $|h+z-a| \geq C_1\Delta_n^{\varepsilon}$. Then, we deduce that

$$
\begin{aligned}
&M_{1,1,2,2,2}^{\theta,\beta_0} \leq C\Delta_n^{-\frac{1}{2}-2\gamma}e^{-\frac{C_1^2\Delta_n^{2\varepsilon-1}}{c}}\int_I\int_{\{|h|\leq\Delta_n^{\alpha_0}\}}\frac{q_{(1)}^{\theta,\beta_0}(\Delta_n, X_{t_k}, h+X_{t_k}+z; z)e^{-\lambda_n\Delta_n}\lambda_n\Delta_n}{p^{\theta,\beta_0}(\Delta_n, X_{t_k}, h+X_{t_k}+z)} dh\mu(dz) \\
&= C\Delta_n^{-\frac{1}{2}-2\gamma}e^{-\frac{C_1^2\Delta_n^{2\varepsilon-1}}{c}}\int_I\int_{\{|y-X_{t_k}-z|\leq\Delta_n^{\alpha_0}\}}\frac{q_{(1)}^{\theta,\beta_0}(\Delta_n, X_{t_k}, y; z)e^{-\lambda_n\Delta_n}\lambda_n\Delta_n}{p^{\theta,\beta_0}(\Delta_n, X_{t_k}, y)} dy\mu(dz) \\
&\leq C\Delta_n^{-\frac{1}{2}-2\gamma}e^{-\frac{C_1^2\Delta_n^{2\varepsilon-1}}{c}}\int_{\{|y-X_{t_k}|\leq\Delta_n^{\alpha_0}+\rho_2\Delta_n^{-\gamma}\}}\frac{\int_I q_{(1)}^{\theta,\beta_0}(\Delta_n, X_{t_k}, y; z)\mu(dz)e^{-\lambda_n\Delta_n}\lambda_n\Delta_n}{p^{\theta,\beta_0}(\Delta_n, X_{t_k}, y)} dy \\
&= C\Delta_n^{-\frac{1}{2}-2\gamma+\alpha_0}e^{-\frac{C_1^2\Delta_n^{2\varepsilon-1}}{c}} + C\Delta_n^{-\frac{1}{2}-3\gamma}e^{-\frac{C_1^2\Delta_n^{2\varepsilon-1}}{c}} \\
&\leq C\Delta_n^{-\frac{1}{2}-3\gamma}e^{-\frac{C_1^2\Delta_n^{2\varepsilon-1}}{c}},
\end{aligned}
$$

where we have used the change of variable $y := h + X_{t_k} + z$, and (5.23).

Therefore, we have shown that for n large enough and $\alpha_0 \in (\epsilon, \frac{1}{2})$,

$$
M_{1,1,2,2}^{\theta,\beta_0} \leq C\Delta_n^{-\frac{1}{2}-3\gamma}e^{-C_0\Delta_n^{2\alpha_0-1}},
$$

for some constants $C, C_0 > 0$, which together with (5.27) gives

$$
M_{1,1,2}^{\theta,\beta_0} \leq C\left(\lambda_n\Delta_n^{1+2\epsilon} + \Delta_n^{-\frac{1}{2}-3\gamma}e^{-C_0\Delta_n^{2\alpha_0-1}}\right).
$$

Finally, as for the term $M_{0,2}^{\theta,\beta_0}$, we obtain that $M_{1,2}^{\theta,\beta_0} + M_2^{\theta,\beta_0} \leq C\lambda_n\Delta_n^{3/2}$, which concludes the proof of (5.21) and (5.22). □

For all $k \in \{0,\ldots,n-1\}$ and $p \in \{2,4\}$, set $\widehat{A}_{k,r} = \{\widehat{L}_{t_{k+1}} - \widehat{L}_{t_k} = r\}$, $\widehat{A}_{k,r}^c = \{\widehat{L}_{t_{k+1}} - \widehat{L}_{t_k} \neq r\}$, $\widetilde{A}_{k,r} = \{\widetilde{L}_{t_{k+1}} - \widetilde{L}_{t_k} = r\}$, $\widetilde{A}_{k,r}^c = \{\widetilde{L}_{t_{k+1}} - \widetilde{L}_{t_k} \neq r\}$, where $\widetilde{L}_t = \int_0^t\int_{\mathbb{R}_0} zM(ds,dz)$ and

$$
\begin{aligned}
M_{1,p}^{\theta_n,\beta(\ell)} &= \sum_{r\in\overline{\mathcal{A}}} r^p \mathrm{E}\left[\mathbf{1}_{\widehat{A}_{k,r}}\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[\mathbf{1}_{\widetilde{A}_{k,r}^c}\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)} = X_{t_{k+1}}\right]\Big|X_{t_k}\right], \\
M_{2,p}^{\theta_n,\beta(\ell)} &= \sum_{r\in\overline{\mathcal{A}}} \mathrm{E}\left[\mathbf{1}_{\widehat{A}_{k,r}}\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[\mathbf{1}_{\widetilde{A}_{k,r}^c}\left(\int_{t_k}^{t_{k+1}}\int_I zM(ds,dz)\right)^p\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)} = X_{t_{k+1}}\right]\Big|X_{t_k}\right],
\end{aligned}
$$

where $\overline{\mathcal{A}} := \{\sum_{i=1}^j a_i, a_i \in I, j \in \mathbb{N}\}$

As in Lemma 3.2.6, we obtain the following large deviation estimates for the non-linear model (5.1).

Lemma 5.2.15. *Under conditions* **(A1)**-**(A4)** *and* **(A7)**-**(A8)**, *for n large enough, there exist constants $C_0, C_1 > 0$, such that for all $\alpha \in (\upsilon, \frac{1}{2})$, and $k \in \{0, ..., n-1\}$,*

$$M_{1,p}^{\theta_n,\beta(\ell)} + M_{2,p}^{\theta_n,\beta(\ell)} \leq C_1 e^{-C_0 \Delta_n^{2\alpha-1}}.$$

Proof. We start bounding $M_{1,p}^{\theta_n,\beta(\ell)}$. For this, we fix $\alpha \in (\upsilon, \frac{1}{2})$, and write $M_{1,p}^{\theta_n,\beta(\ell)} = M_{1,p,1}^{\theta_n,\beta(\ell)} + M_{1,p,2}^{\theta_n,\beta(\ell)}$, where

$$M_{1,p,1}^{\theta_n,\beta(\ell)} = \sum_{r \in \overline{\mathcal{A}}} r^p \mathrm{E}\left[\mathbf{1}_{\left\{\left|\int_{t_k}^{t_{k+1}} \sigma(\beta, X_s) dB_s\right| > \Delta_n^{\alpha}\right\}} \mathbf{1}_{\widehat{A}_{k,r}} \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[\mathbf{1}_{\widetilde{A}_{k,r}^c} \middle| Y_{t_{k+1}}^{\theta_n,\beta(\ell)} = X_{t_{k+1}}\right] \middle| X_{t_k}\right],$$
$$M_{1,p,2}^{\theta_n,\beta(\ell)} = \sum_{r \in \overline{\mathcal{A}}} r^p \mathrm{E}\left[\mathbf{1}_{\left\{\left|\int_{t_k}^{t_{k+1}} \sigma(\beta, X_s) dB_s\right| \leq \Delta_n^{\alpha}\right\}} \mathbf{1}_{\widehat{A}_{k,r}} \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[\mathbf{1}_{\widetilde{A}_{k,r}^c} \middle| Y_{t_{k+1}}^{\theta_n,\beta(\ell)} = X_{t_{k+1}}\right] \middle| X_{t_k}\right].$$

Applying Cauchy-Schwarz and the exponential martingale inequalities, together with **(A2)**, **(A4)**, we get that

$$\begin{aligned}
M_{1,p,1}^{\theta_n,\beta(\ell)} &\leq \mathrm{E}\left[\mathbf{1}_{\left\{\left|\int_{t_k}^{t_{k+1}} \sigma(\beta, X_s) dB_s\right| > \Delta_n^{\alpha}\right\}} \left(\int_{t_k}^{t_{k+1}} \int_I z N(ds, dz)\right)^p \middle| X_{t_k}\right] \\
&\leq \left(\mathrm{P}\left(\left|\int_{t_k}^{t_{k+1}} \sigma(\beta, X_s) dB_s\right| > \Delta_n^{\alpha} \middle| X_{t_k}\right)\right)^{\frac{1}{2}} \left(\mathrm{E}\left[\left(\int_{t_k}^{t_{k+1}} \int_I z N(ds, dz)\right)^{2p} \middle| X_{t_k}\right]\right)^{\frac{1}{2}} \\
&\leq \sqrt{2\Delta_n} e^{-\frac{\Delta_n^{2\alpha-1}}{4c^2}},
\end{aligned}$$

since the quadratic variation of the continuous martingale $\int_{t_k}^{t_{k+1}} \sigma(\beta, X_s) dB_s$ is upper bounded by $c^2 \Delta_n$, where the constant c is as in **(A2)**.

Next, applying Hölder's and Jensen's inequalities with q_1, q_2 conjugate, we get that

$$M_{1,p,2}^{\theta_n,\beta(\ell)} \leq \sum_{r \in \overline{\mathcal{A}}} r^p \left(\mathrm{P}\left(\widehat{A}_{k,r} | X_{t_k}\right)\right)^{\frac{1}{q_1}} \left(H_{k,r}^{\theta_n,\beta(\ell)}\right)^{\frac{1}{q_2}}, \tag{5.28}$$

where

$$H_{k,r}^{\theta_n,\beta(\ell)} = \mathrm{E}\left[\mathbf{1}_{\left\{\left|\int_{t_k}^{t_{k+1}} \sigma(\beta, X_s) dB_s\right| \leq \Delta_n^{\alpha}\right\}} \mathbf{1}_{\widehat{A}_{k,r}} \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[\mathbf{1}_{\widetilde{A}_{k,r}^c} \middle| Y_{t_{k+1}}^{\theta_n,\beta(\ell)} = X_{t_{k+1}}\right] \middle| X_{t_k}\right].$$

Since $1 = \mathbf{1}_{\widehat{J}_{0,k}} + \sum_{j=1}^{\infty} \sum_{(a_1,...,a_j) \in I} \mathbf{1}_{\{\widehat{J}_{j,k}, a_1, ..., a_j\}}$, and set $a := a_1 + \cdots + a_j$, we can write $H_{k,r}^{\theta_n,\beta(\ell)} = H_{k,r,1}^{\theta_n,\beta(\ell)} + H_{k,r,2}^{\theta_n,\beta(\ell)}$, where

$$\begin{aligned}
H_{k,r,1}^{\theta_n,\beta(\ell)} &= \mathrm{E}\left[\mathbf{1}_{\left\{\left|\int_{t_k}^{t_{k+1}} \sigma(\beta, X_s) dB_s\right| \leq \Delta_n^{\alpha}\right\}} \mathbf{1}_{\widehat{J}_{0,k}} \mathbf{1}_{\{r=0\}} \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[\mathbf{1}_{\widetilde{A}_{k,0}^c} \middle| Y_{t_{k+1}}^{\theta_n,\beta(\ell)} = X_{t_{k+1}}\right] \middle| X_{t_k}\right], \\
H_{k,r,2}^{\theta_n,\beta(\ell)} &= \sum_{j=1}^{\infty} \sum_{(a_1,...,a_j) \in I} \mathrm{E}\left[\mathbf{1}_{\left\{\left|\int_{t_k}^{t_{k+1}} \sigma(\beta, X_s) dB_s\right| \leq \Delta_n^{\alpha}\right\}} \mathbf{1}_{\{\widehat{J}_{j,k}, a_1, ..., a_j\}} \mathbf{1}_{\{r=a\}} \right. \\
&\qquad \left. \times \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[\mathbf{1}_{\widetilde{A}_{k,a}^c} \middle| Y_{t_{k+1}}^{\theta_n,\beta(\ell)} = X_{t_{k+1}}\right] \middle| X_{t_k}\right].
\end{aligned}$$

We first treat $H_{k,r,1}^{\theta_n,\beta(\ell)}$. Using Bayes's formula, (5.12), (5.13) and Lemma 5.2.5, we get that

$$
\begin{aligned}
&\mathbf{1}_{\widehat{J}_{0,k}} \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[\mathbf{1}_{\widetilde{A}_{k,0}^c}\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}\right]\\
&=\mathbf{1}_{\widehat{J}_{0,k}}\sum_{m=1}^{\infty}\sum_{(z_1,\ldots,z_m)\in I}\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[\mathbf{1}_{\{\widetilde{J}_{m,k,z_1,\ldots,z_m}\}}\mathbf{1}_{\{z\neq 0\}}\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}\right]\\
&=\mathbf{1}_{\widehat{J}_{0,k}}\frac{\sum_{m=1}^{\infty}\sum_{(z_1,\ldots,z_m)\in I}q_{(m)}^{\theta_n,\beta(\ell)}(\Delta_n,X_{t_k},X_{t_{k+1}};z_1,\ldots,z_m)\mathbf{1}_{\{z\neq 0\}}p_{z_1}\ldots p_{z_m}e^{-\lambda_n\Delta_n}\frac{(\lambda_n\Delta_n)^m}{m!}}{p^{\theta_n,\beta(\ell)}(\Delta_n,X_{t_k},X_{t_{k+1}})}\\
&=\mathbf{1}_{\widehat{J}_{0,k}}\frac{\sum_{m=1}^{\infty}\sum_{(z_1,\ldots,z_m)\in I}q_{(m)}^{\theta_n,\beta(\ell)}(\Delta_n,X_{t_k},X_{t_{k+1}};z_1,\ldots,z_m)\mathbf{1}_{\{z\neq 0\}}p_{z_1}\ldots p_{z_m}e^{-\lambda_n\Delta_n}\frac{(\lambda_n\Delta_n)^m}{m!}}{\sum_{i=0}^{\infty}q_{(i)}^{\theta_n,\beta(\ell)}(\Delta_n,X_{t_k},X_{t_{k+1}})e^{-\lambda_n\Delta_n}\frac{(\lambda_n\Delta_n)^i}{i!}}\\
&\leq\mathbf{1}_{\widehat{J}_{0,k}}\frac{\sum_{m=1}^{\infty}\sum_{(z_1,\ldots,z_m)\in I}q_{(m)}^{\theta_n,\beta(\ell)}(\Delta_n,X_{t_k},X_{t_{k+1}};z_1,\ldots,z_m)\mathbf{1}_{\{z\neq 0\}}p_{z_1}\ldots p_{z_m}e^{-\lambda_n\Delta_n}\frac{(\lambda_n\Delta_n)^m}{m!}}{q_{(0)}^{\theta_n,\beta(\ell)}(\Delta_n,X_{t_k},X_{t_{k+1}})e^{-\lambda_n\Delta_n}}\\
&\leq\mathbf{1}_{\widehat{J}_{0,k}}\frac{\sum_{m=1}^{\infty}\sum_{(z_1,\ldots,z_m)\in I}\frac{C^{m+1}}{\sqrt{\Delta_n}}e^{-\frac{(X_{t_{k+1}}-X_{t_k}-z)^2}{c\Delta_n}}\mathbf{1}_{\{z\neq 0\}}p_{z_1}\ldots p_{z_m}e^{-\lambda_n\Delta_n}\frac{(\lambda_n\Delta_n)^m}{m!}}{\frac{1}{C\sqrt{\Delta_n}}e^{-c\frac{(X_{t_{k+1}}-X_{t_k})^2}{\Delta_n}}e^{-\lambda_n\Delta_n}},
\end{aligned}
$$

for some constants $C>0$ and $c\geq 1$, where $z:=z_1+\cdots+z_m$, and we have lower bounded the denominator by the term $i=0$.

Conditioning on $|\int_{t_k}^{t_{k+1}}\sigma(\beta,X_s)dB_s|\leq\Delta_n^{\alpha}$ and $\widehat{J}_{0,k}$, using equation (5.1), the boundedness of b and the fact that $|z|\geq C\Delta_n^{\upsilon}$ for some constant $C>0$, together with $c>1$, we have that for n sufficiently large,

$$
\begin{aligned}
&e^{-\frac{(X_{t_{k+1}}-X_{t_k}-z)^2}{c\Delta_n}+c\frac{(X_{t_{k+1}}-X_{t_k})^2}{\Delta_n}}\leq e^{\left(c-\frac{1}{c}\right)\frac{\left(C_0\Delta_n+\Delta_n^{\alpha}+\Delta_n^{1-\gamma}\lambda\right)^2}{\Delta_n}+2\frac{|z|\left(C_0\Delta_n+\Delta_n^{\alpha}+\Delta_n^{1-\gamma}\lambda\right)}{c\Delta_n}-\frac{|z|^2}{c\Delta_n}}\\
&\leq e^{C_1\Delta_n^{2\alpha-1}-\frac{|z|}{c\Delta_n}(|z|-C_2\Delta_n^{\alpha})}\left(\mathbf{1}_{\{|z|\leq C_2\Delta_n^{\alpha}\}}+\mathbf{1}_{\{|z|>C_2\Delta_n^{\alpha}\}}\right)\\
&\leq e^{C_1\Delta_n^{2\alpha-1}}\left(e^{-\frac{C_2}{c}\Delta_n^{\alpha-1}(C\Delta_n^{\upsilon}-C_2\Delta_n^{\alpha})}\mathbf{1}_{\{|z|\leq C_2\Delta_n^{\alpha}\}}+e^{-\frac{C}{c}\Delta_n^{\upsilon-1}(C\Delta_n^{\upsilon}-C_2\Delta_n^{\alpha})}\mathbf{1}_{\{|z|>C_2\Delta_n^{\alpha}\}}\right) \qquad (5.29)\\
&\leq e^{C_1\Delta_n^{2\alpha-1}}\left(e^{-\frac{C_2}{c}C\Delta_n^{\upsilon+\alpha-1}+\frac{C_2^2}{c}\Delta_n^{2\alpha-1}}+e^{-\frac{C^2}{c}\Delta_n^{2\upsilon-1}+\frac{C}{c}C_2\Delta_n^{\alpha+\upsilon-1}}\right)\\
&\leq e^{-C_3\Delta_n^{\upsilon+\alpha-1}}+e^{-C_4\Delta_n^{2\upsilon-1}}\leq C_5e^{-C_6\Delta_n^{\alpha+\upsilon-1}},
\end{aligned}
$$

for some constants $C_0,\ldots,C_6>0$.

On the other hand,

$$
\sum_{m=1}^{\infty}\sum_{(z_1,\ldots,z_m)\in I}C^m p_{z_1}\ldots p_{z_m}e^{-\lambda_n\Delta_n}\frac{(\lambda_n\Delta_n)^m}{m!}<\infty, \tag{5.30}
$$

which concludes that $H_{k,r,1}^{\theta_n,\beta(\ell)}\leq C_1e^{-C_0\Delta_n^{\alpha+\upsilon-1}}$, for some constants $C_0,C_1>0$.

Next, applying Hölder's and Jensen's inequalities with p_1,p_2 conjugate, we get that

$$
\begin{aligned}
H_{k,r,2}^{\theta_n,\beta(\ell)}&\leq\sum_{j=1}^{\infty}\sum_{(a_1,\ldots,a_j)\in I}\left(\mathrm{P}\left(\widehat{J}_{j,k},a_1,\ldots,a_j\right)|X_{t_k}\right)^{\frac{1}{p_1}}\left(H_{k,r,3}^{\theta_n,\beta(\ell)}\right)^{\frac{1}{p_2}}\\
&=\sum_{j=1}^{\infty}\sum_{(a_1,\ldots,a_j)\in I}\left(p_{a_1}\ldots p_{a_j}e^{-\lambda_n\Delta_n}\frac{(\lambda_n\Delta_n)^j}{j!}\right)^{\frac{1}{p_1}}\left(H_{k,r,3}^{\theta_n,\beta(\ell)}\right)^{\frac{1}{p_2}},
\end{aligned}
$$

where

$$
H_{k,r,3}^{\theta_n,\beta(\ell)}=\mathrm{E}\left[\mathbf{1}_{\left\{\left|\int_{t_k}^{t_{k+1}}\sigma(\beta,X_s)dB_s\right|\leq\Delta_n^{\alpha}\right\}}\mathbf{1}_{\{\widehat{J}_{j,k},a_1,\ldots,a_j\}}\mathbf{1}_{\{r=a\}}\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[\mathbf{1}_{\widetilde{A}_{k,a}^c}\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}\right]\Big|X_{t_k}\right].
$$

Using Bayes's formula, (5.12), (5.13) and Lemma 5.2.5, we get that

$$\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[\mathbf{1}_{\widetilde{A}_{k,a}^c}\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}\right]=\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[\mathbf{1}_{\widetilde{J}_{0,k}}\mathbf{1}_{\{a\neq 0\}}\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}\right]$$
$$+\sum_{m=1}^{\infty}\sum_{(z_1,\ldots,z_m)\in I}\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[\mathbf{1}_{\{\widetilde{J}_{m,k},z_1,\ldots,z_m\}}\mathbf{1}_{\{a\neq z\}}\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}\right]$$
$$=\frac{q_{(0)}^{\theta_n,\beta(\ell)}(\Delta_n,X_{t_k},X_{t_{k+1}})\mathbf{1}_{\{a\neq 0\}}e^{-\lambda_n\Delta_n}}{\sum_{i=0}^{\infty}q_{(i)}^{\theta_n,\beta(\ell)}(\Delta_n,X_{t_k},X_{t_{k+1}})e^{-\lambda_n\Delta_n}\frac{(\lambda_n\Delta_n)^i}{i!}}$$
$$+\frac{\sum_{m=1}^{\infty}\sum_{(z_1,\ldots,z_m)\in I}q_{(m)}^{\theta_n,\beta(\ell)}(\Delta_n,X_{t_k},X_{t_{k+1}};z_1,\ldots,z_m)\mathbf{1}_{\{a\neq z\}}p_{z_1}\ldots p_{z_m}e^{-\lambda_n\Delta_n}\frac{(\lambda_n\Delta_n)^m}{m!}}{\sum_{i=0}^{\infty}q_{(i)}^{\theta_n,\beta(\ell)}(\Delta_n,X_{t_k},X_{t_{k+1}})e^{-\lambda_n\Delta_n}\frac{(\lambda_n\Delta_n)^i}{i!}}$$
$$\leq\frac{q_{(0)}^{\theta_n,\beta(\ell)}(\Delta_n,X_{t_k},X_{t_{k+1}})\mathbf{1}_{\{a\neq 0\}}e^{-\lambda_n\Delta_n}}{q_{(j)}^{\theta_n,\beta(\ell)}(\Delta_n,X_{t_k},X_{t_{k+1}};a_1,\ldots,a_j)p_{a_1}\ldots p_{a_j}e^{-\lambda_n\Delta_n}\frac{(\lambda_n\Delta_n)^j}{j!}}$$
$$+\frac{\sum_{m=1}^{\infty}\sum_{(z_1,\ldots,z_m)\in I}q_{(m)}^{\theta_n,\beta(\ell)}(\Delta_n,X_{t_k},X_{t_{k+1}};z_1,\ldots,z_m)\mathbf{1}_{\{a\neq z\}}p_{z_1}\ldots p_{z_m}e^{-\lambda_n\Delta_n}\frac{(\lambda_n\Delta_n)^m}{m!}}{q_{(j)}^{\theta_n,\beta(\ell)}(\Delta_n,X_{t_k},X_{t_{k+1}};a_1,\ldots,a_j)p_{a_1}\ldots p_{a_j}e^{-\lambda_n\Delta_n}\frac{(\lambda_n\Delta_n)^j}{j!}}$$
$$\leq\frac{\frac{C}{\sqrt{\Delta_n}}e^{-\frac{(X_{t_{k+1}}-X_{t_k})^2}{c\Delta_n}}\mathbf{1}_{\{a\neq 0\}}e^{-\lambda_n\Delta_n}}{\frac{1}{C^{j+1}\sqrt{\Delta_n}}e^{-\frac{c(X_{t_{k+1}}-X_{t_k}-a)^2}{\Delta_n}}p_{a_1}\ldots p_{a_j}e^{-\lambda_n\Delta_n}\frac{(\lambda_n\Delta_n)^j}{j!}}$$
$$+\frac{\sum_{m=1}^{\infty}\sum_{(z_1,\ldots,z_m)\in I}\frac{C^{m+1}}{\sqrt{\Delta_n}}e^{-\frac{(X_{t_{k+1}}-X_{t_k}-z)^2}{c\Delta_n}}\mathbf{1}_{\{a\neq z\}}p_{z_1}\ldots p_{z_m}e^{-\lambda_n\Delta_n}\frac{(\lambda_n\Delta_n)^m}{m!}}{\frac{1}{C^{j+1}\sqrt{\Delta_n}}e^{-\frac{c(X_{t_{k+1}}-X_{t_k}-a)^2}{\Delta_n}}p_{a_1}\ldots p_{a_j}e^{-\lambda_n\Delta_n}\frac{(\lambda_n\Delta_n)^j}{j!}},$$

for some constants $C>0$ and $c\geq 1$, where $z:=z_1+\cdots+z_m$ and we have lower bounded the denominator by the term $q_{(j)}^{\theta_n,\beta(\ell)}(\Delta_n,X_{t_k},X_{t_{k+1}};a_1,\ldots,a_j)p_{a_1}\ldots p_{a_j}e^{-\lambda_n\Delta_n}\frac{(\lambda_n\Delta_n)^j}{j!}$.

Conditioning on $|\int_{t_k}^{t_{k+1}}\sigma(\beta,X_s)dB_s|\leq\Delta_n^{\alpha}$ and $\{\widehat{J}_{j,k},a_1,\ldots,a_j\}$, using equation (5.1), the boundedness of b and the fact that $|a|\geq C\Delta_n^{v}$ for some constant $C>0$, we get that, by proceeding as in (5.29), for n sufficiently large,

$$e^{-\frac{(X_{t_{k+1}}-X_{t_k})^2}{c\Delta_n}+c\frac{(X_{t_{k+1}}-X_{t_k}-a)^2}{\Delta_n}}\leq e^{\left(c-\frac{1}{c}\right)\frac{\left(C_0\Delta_n+\Delta_n^{\alpha}+\Delta_n^{1-\gamma}\lambda\right)^2}{\Delta_n}+2\frac{|a|\left(C_0\Delta_n+\Delta_n^{\alpha}+\Delta_n^{1-\gamma}\lambda\right)}{c\Delta_n}-\frac{|a|^2}{c\Delta_n}}$$
$$\leq C_5e^{-C_6\Delta_n^{\alpha+v-1}},$$

for some constants $C_5,C_6>0$.

Similarly, using the same above arguments, together with the fact that by hypothesis **(A7)**, $|a-z|\geq C\Delta_n^{v}$ for some constant $C>0$, we get that, by proceeding as in (5.29), for n sufficiently large,

$$e^{-\frac{(X_{t_{k+1}}-X_{t_k}-z)^2}{c\Delta_n}+c\frac{(X_{t_{k+1}}-X_{t_k}-a)^2}{\Delta_n}}\leq e^{\left(c-\frac{1}{c}\right)\frac{\left(C_0\Delta_n+\Delta_n^{\alpha}+\Delta_n^{1-\gamma}\lambda\right)^2}{\Delta_n}+2\frac{|a-z|\left(C_0\Delta_n+\Delta_n^{\alpha}+\Delta_n^{1-\gamma}\lambda\right)}{c\Delta_n}-\frac{|a-z|^2}{c\Delta_n}}$$
$$\leq C_5e^{-C_6\Delta_n^{\alpha+v-1}},$$

for some constants $C_5,C_6>0$.

Then using again (5.30), we conclude that $H_{k,r,3}^{\theta_n,\beta(\ell)}\leq C_1C^{j+1}e^{-C_0\Delta_n^{\alpha+v-1}}$, which, together with hypothesis **(A8)**, implies that

$$H_{k,r,2}^{\theta_n,\beta(\ell)}\leq\sum_{j=1}^{\infty}\sum_{(a_1,\ldots,a_j)\in I}\left(p_{a_1}\ldots p_{a_j}e^{-\lambda_n\Delta_n}\frac{(\lambda_n\Delta_n)^j}{j!}\right)^{\frac{1}{p_1}}\left(C_1C^{j+1}e^{-C_0\Delta_n^{\alpha+v-1}}\right)^{\frac{1}{p_2}}$$
$$\leq c_1e^{-c_2\Delta_n^{\alpha+v-1}},$$

for some constants $c_1, c_2 > 0$. Therefore, we have shown that

$$H_{k,r}^{\theta_n,\beta(\ell)} \leq C_1 e^{-C_0 \Delta_n^{\alpha+v-1}}, \tag{5.31}$$

for some constants $C_0, C_1 > 0$, which, together with hypothesis **(A7)** and (5.28), yields that $M_{1,p,2}^{\theta_n,\beta(\ell)} \leq Ce^{-\frac{C_0}{q_2}\Delta_n^{\alpha+v-1}}$. Thus, we have obtained that

$$M_{1,p}^{\theta_n,\beta(\ell)} \leq C_1 e^{-C_0 \Delta_n^{2\alpha-1}}, \tag{5.32}$$

for some constants $C_0, C_1 > 0$.

We next bound $M_{2,p}^{\theta_n,\beta(\ell)}$. As for the term $M_{1,p}^{\theta_n,\beta(\ell)}$, we write $M_{2,p}^{\theta_n,\beta(\ell)} = M_{2,p,1}^{\theta_n,\beta(\ell)} + M_{2,p,2}^{\theta_n,\beta(\ell)}$, where

$$\begin{aligned} M_{2,p,1}^{\theta_n,\beta(\ell)} &= \sum_{r\in\overline{\mathcal{A}}} \mathrm{E}\Bigg[\mathbf{1}_{\left\{\left|\int_{t_k}^{t_{k+1}}\sigma(\beta,X_s)dB_s\right|>\Delta_n^{\alpha}\right\}} \\ &\quad \times \mathbf{1}_{\widehat{A}_{k,r}} \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[\mathbf{1}_{\widetilde{A}_{k,r}^c}\left(\int_{t_k}^{t_{k+1}}\int_I zM(ds,dz)\right)^p \Big| Y_{t_{k+1}}^{\theta_n,\beta(\ell)} = X_{t_{k+1}}\right] \Big| X_{t_k}\Bigg], \\ M_{2,p,2}^{\theta_n,\beta(\ell)} &= \sum_{r\in\overline{\mathcal{A}}} \mathrm{E}_{X_{t_k}}^{\theta,\beta}\Bigg[\mathbf{1}_{\left\{\left|\int_{t_k}^{t_{k+1}}\sigma(\beta,X_s)dB_s\right|\leq\Delta_n^{\alpha}\right\}} \\ &\quad \times \mathbf{1}_{\widehat{A}_{k,r}} \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[\mathbf{1}_{\widetilde{A}_{k,r}^c}\left(\int_{t_k}^{t_{k+1}}\int_I zM(ds,dz)\right)^p \Big| Y_{t_{k+1}}^{\theta_n,\beta(\ell)} = X_{t_{k+1}}\right] \Big| X_{t_k}\Bigg]. \end{aligned}$$

First, applying the exponential martingale, Cauchy-Schwarz and Jensen's inequalities, we get that

$$\begin{aligned} M_{2,p,1}^{\theta_n,\beta(\ell)} &\leq \mathrm{E}\Bigg[\mathbf{1}_{\left\{\left|\int_{t_k}^{t_{k+1}}\sigma(\beta,X_s)dB_s\right|>\Delta_n^{\alpha}\right\}} \\ &\quad \times \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[\left(\int_{t_k}^{t_{k+1}}\int_I zM(ds,dz)\right)^p \Big| Y_{t_{k+1}}^{\theta_n,\beta(\ell)} = X_{t_{k+1}}\right] \Big| X_{t_k}\Bigg] \\ &\leq \left(\mathrm{P}\left(\left|\int_{t_k}^{t_{k+1}}\sigma(\beta,X_s)dB_s\right| > \Delta_n^{\alpha}\Big| X_{t_k}\right)\right)^{\frac{1}{2}} (T_{k,p})^{\frac{1}{2}} \\ &\leq \sqrt{2}e^{-\frac{\Delta_n^{2\alpha-1}}{4c^2}} (T_{k,p})^{\frac{1}{2}}, \end{aligned}$$

where the constant c is as in **(A2)** and

$$T_{k,p} = \mathrm{E}\left[\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[\left(\int_{t_k}^{t_{k+1}}\int_I zM(ds,dz)\right)^{2p} \Big| Y_{t_{k+1}}^{\theta_n,\beta(\ell)} = X_{t_{k+1}}\right] \Big| X_{t_k}\right].$$

Using Lemma 5.2.10 and Hölder's inequality with q_1, q_2 conjugate, together with **(A4)**, **(A8)**, we get that

$$
\begin{aligned}
T_{k,p} &= \mathrm{E}\left[\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[\left(\int_{t_k}^{t_{k+1}}\int_I zM(ds,dz)\right)^{2p}\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}^{\theta_n,\beta(\ell)}\right]\mathbf{1}_{\widehat{J}_{0,k}}\frac{q_{(0)}^{\theta,\beta}}{q_{(0)}^{\theta_n,\beta(\ell)}}\Big|X_{t_k}^{\theta_n,\beta(\ell)}=X_{t_k}\right]\\
&\quad+\sum_{j=1}^{\infty}\sum_{(a_1,\ldots,a_j)\in I}\mathrm{E}\Bigg[\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[\left(\int_{t_k}^{t_{k+1}}\int_I zM(ds,dz)\right)^{2p}\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}^{\theta_n,\beta(\ell)}\right]\\
&\quad\times\mathbf{1}_{\{\widehat{J}_{j,k},a_1,\ldots,a_j\}}\frac{q_{(j)}^{\theta,\beta}}{q_{(j)}^{\theta_n,\beta(\ell)}}\Big|X_{t_k}^{\theta_n,\beta(\ell)}=X_{t_k}\Bigg]\\
&\leq\left(\mathrm{E}\left[\left(\int_{t_k}^{t_{k+1}}\int_I zM(ds,dz)\right)^{2pq_1}\Big|Y_{t_k}^{\theta_n,\beta(\ell)}=X_{t_k}\right]\right)^{\frac{1}{q_1}}\left(\mathrm{E}\left[\mathbf{1}_{\widehat{J}_{0,k}}\left(\frac{q_{(0)}^{\theta,\beta}}{q_{(0)}^{\theta_n,\beta(\ell)}}\right)^{q_2}\Big|X_{t_k}^{\theta_n,\beta(\ell)}=X_{t_k}\right]\right)^{\frac{1}{q_2}}\\
&\quad+\sum_{j=1}^{\infty}\sum_{(a_1,\ldots,a_j)\in I}\left(\mathrm{E}\left[\left(\int_{t_k}^{t_{k+1}}\int_I zM(ds,dz)\right)^{2pq_1}\Big|Y_{t_k}^{\theta_n,\beta(\ell)}=X_{t_k}\right]\right)^{\frac{1}{q_1}}\\
&\quad\times\left(\mathrm{E}\left[\mathbf{1}_{\{\widehat{J}_{j,k},a_1,\ldots,a_j\}}\left(\frac{q_{(j)}^{\theta,\beta}}{q_{(j)}^{\theta_n,\beta(\ell)}}\right)^{q_2}\Big|X_{t_k}^{\theta_n,\beta(\ell)}=X_{t_k}\right]\right)^{\frac{1}{q_2}}\\
&\leq C\Delta_n^{\frac{1}{q_1}}\left(1+\sum_{j=1}^{\infty}\sum_{(a_1,\ldots,a_j)\in A}\left(C^{(2q_2-1)(j+1)}p_{a_1}\cdots p_{a_j}e^{-\lambda\Delta_n}\frac{(\lambda\Delta_n)^j}{j!}\right)^{\frac{1}{q_2}}\right)\\
&<\infty.
\end{aligned}
$$

We then deduce that $M_{2,p,1}^{\theta_n,\beta(\ell)}\leq Ce^{-\frac{\Delta_n^{2\alpha-1}}{4c^2}}$, for some constants $C>0$ and $c\geq 1$.

Next, applying Hölder's and Jensen's inequalities with q_1,q_2 conjugate, together with (5.31) and **(A7)**, we get that

$$
\begin{aligned}
M_{2,p,2}^{\theta_n,\beta(\ell)} &\leq \sum_{r\in\overline{\mathcal{A}}}\left(\mathrm{P}\left(\widehat{A}_{k,r}\big|X_{t_k}\right)\right)^{\frac{1}{q_1}}\Bigg(\mathrm{E}\Bigg[\mathbf{1}_{\left\{\left|\int_{t_k}^{t_{k+1}}\sigma(\beta,X_s)dB_s\right|\leq\Delta_n^{\alpha}\right\}}\mathbf{1}_{\widehat{A}_{k,r}}\\
&\quad\times\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[\mathbf{1}_{\widetilde{A}_{k,r}^c}\left(\int_{t_k}^{t_{k+1}}\int_I zM(ds,dz)\right)^{pq_2}\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}\right]\Big|X_{t_k}\Bigg]\Bigg)^{\frac{1}{q_2}}\\
&\leq\sum_{r\in\overline{\mathcal{A}}}\left(\mathrm{P}\left(\widehat{A}_{k,r}\big|X_{t_k}\right)\right)^{\frac{1}{q_1}}\left(H_{k,n}\right)^{\frac{1}{2q_2}}\left(T_{k,pq_2}\right)^{\frac{1}{2q_2}}\\
&\leq C_1e^{-C_0\Delta_n^{\alpha+v-1}},
\end{aligned}
$$

for some constants $C_0,C_1>0$. This shows that $M_{2,p}^{\theta_n,\beta(\ell)}$ satisfies (5.32), thus the result follows. □

Finally, we recall a discrete time ergodic theorem.

Lemma 5.2.16. [40, Lemma 8] *Assume conditions* **(A1)** *and* **(A5)**. *Consider a differentiable function $g:\mathbb{R}\to\mathbb{R}$, whose derivatives have polynomial growth in x. Then, as $n\to\infty$,*

$$\frac{1}{n}\sum_{k=0}^{n-1}g(X_{t_k})\overset{\mathrm{P}^{\theta_0,\beta_0}}{\longrightarrow}\int_{\mathbb{R}}g(x)\pi_{\theta_0,\beta_0}(dx).$$

5.3 Proof of Theorem 5.1.1

In this section, the proof of Theorem 5.1.1 will be divided into several steps. We begin deriving a stochastic expansion of the log-likelihood ratio using Proposition 5.2.1 and Lemmas 5.2.1, 5.2.2. The second step is devoted to treat the negligible contributions of this expansion. Finally, the last step concludes the LAN property by applying the central limit theorem for triangular arrays.

5.3.1 Expansion of the log-likelihood ratio

In order to deal with the log-likelihood ratio in Theorem 5.1.1, we will use the following decomposition

$$\log \frac{p(X^n;(\theta_n,\beta_n))}{p(X^n;(\theta_0,\beta_0))} = \log \frac{p(X^n;(\theta_n,\beta_0))}{p(X^n;(\theta_0,\beta_0))} + \log \frac{p(X^n;(\theta_n,\beta_n))}{p(X^n;(\theta_n,\beta_0))}. \tag{5.33}$$

For $\ell \in [0,1]$, set $\theta(\ell) := \theta_n(\ell,u) := \theta_0 + \frac{\ell u}{\sqrt{n\Delta_n}}, \beta(\ell) := \beta_n(\ell,v) := \beta_0 + \frac{\ell v}{\sqrt{n}}$. Then, from the Markov property and Proposition 5.2.1,

$$\begin{aligned}
\log \frac{p(X^n;(\theta_n,\beta_0))}{p(X^n;(\theta_0,\beta_0))} &= \sum_{k=0}^{n-1} \log \frac{p^{\theta_n,\beta_0}}{p^{\theta_0,\beta_0}}(\Delta_n, X_{t_k}, X_{t_{k+1}}) \\
&= \sum_{k=0}^{n-1} \frac{u}{\sqrt{n\Delta_n}} \int_0^1 \frac{\partial_\theta p^{\theta(\ell),\beta_0}}{p^{\theta(\ell),\beta_0}}(\Delta_n, X_{t_k}, X_{t_{k+1}}) d\ell \\
&= \sum_{k=0}^{n-1} \frac{u}{\sqrt{n\Delta_n}} \int_0^1 \frac{1}{\Delta_n} \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell),\beta_0} \left[\delta\left(\partial_\theta Y_{t_{k+1}}^{\theta(\ell),\beta_0}(t_k, X_{t_k}) U^{\theta(\ell),\beta_0}(t_k, X_{t_k}) \right) \Big| Y_{t_{k+1}}^{\theta(\ell),\beta_0} = X_{t_{k+1}} \right] d\ell.
\end{aligned}$$

We next consider the stopping time

$$\widehat{\tau} := \inf\left\{ s \geq 0 : |\Delta \widehat{Z}_s| < \rho_1 \Delta_n^{\upsilon} \text{ or } |\Delta \widehat{Z}_s| > \rho_2 \Delta_n^{-\gamma} \right\}, \tag{5.34}$$

and

$$\widetilde{\tau} := \inf\left\{ s \geq 0 : |\Delta \widetilde{Z}_s| < \rho_1 \Delta_n^{\upsilon} \text{ or } |\Delta \widetilde{Z}_s| > \rho_2 \Delta_n^{-\gamma} \right\}, \tag{5.35}$$

where $\rho_1, \rho_2 > 0$ and $0 < \upsilon, \gamma < \frac{1}{2}$ are from hypothesis **(A6)**.

Observe that on the event $\{\widehat{\tau} > n\Delta_n\}$, all the jumps of $\widehat{Z}$ in the interval $[0, n\Delta_n]$ are in the interval $[\rho_1\Delta_n^{\upsilon}, \rho_2\Delta_n^{-\gamma}]$. Hence, for all $\omega \in \{\widehat{\tau} > n\Delta_n\}$, $X^{\theta,\beta}$ satisfies

$$X_t^{\theta,\beta} = x_0 + \int_0^t b(\theta, X_s^{\theta,\beta})ds + \int_0^t \sigma(\beta, X_s^{\theta,\beta})dB_s + \int_0^t \int_I z\left(N(ds,dz) - \nu(dz)ds\right), \tag{5.36}$$

for all $t \in [0, n\Delta_n]$, where recall that $I = \{z \in A : \rho_1\Delta_n^{\upsilon} \leq |z| \leq \rho_2\Delta_n^{-\gamma}\}$. A similar statement is true for $Y^{\theta,\beta}$.

Then, multiplying by $\mathbf{1}_{\{\widetilde{\tau} > n\Delta_n\}} + \mathbf{1}_{\{\widetilde{\tau} \leq n\Delta_n\}}$ inside and $\mathbf{1}_{\{\widehat{\tau} > n\Delta_n\}} + \mathbf{1}_{\{\widehat{\tau} \leq n\Delta_n\}}$ outside the conditional expectation above, we get that

$$\log \frac{p(X^n;(\theta_n,\beta_0))}{p(X^n;(\theta_0,\beta_0))} = \frac{u}{\sqrt{n\Delta_n^3}} \sum_{k=0}^{n-1} \int_0^1 \left(Z_{k,n}^{1,\ell} + Z_{k,n}^{2,\ell} + Z_{k,n}^{3,\ell} \right) d\ell,$$

where

$$\begin{aligned}
Z_{k,n}^{1,\ell} &= \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell),\beta_0} \left[\delta\left(\partial_\theta Y_{t_{k+1}}^{\theta(\ell),\beta_0}(t_k, X_{t_k}) U^{\theta(\ell),\beta_0}(t_k, X_{t_k}) \right) \Big| Y_{t_{k+1}}^{\theta(\ell),\beta_0} = X_{t_{k+1}} \right] \mathbf{1}_{\{\widehat{\tau} \leq n\Delta_n\}}, \\
Z_{k,n}^{2,\ell} &= \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell),\beta_0} \left[\delta\left(\partial_\theta Y_{t_{k+1}}^{\theta(\ell),\beta_0}(t_k, X_{t_k}) U^{\theta(\ell),\beta_0}(t_k, X_{t_k}) \right) \mathbf{1}_{\{\widetilde{\tau} \leq n\Delta_n\}} \Big| Y_{t_{k+1}}^{\theta(\ell),\beta_0} = X_{t_{k+1}} \right] \mathbf{1}_{\{\widehat{\tau} > n\Delta_n\}}, \\
Z_{k,n}^{3,\ell} &= \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell),\beta_0} \left[\delta\left(\partial_\theta Y_{t_{k+1}}^{\theta(\ell),\beta_0}(t_k, X_{t_k}) U^{\theta(\ell),\beta_0}(t_k, X_{t_k}) \right) \mathbf{1}_{\{\widetilde{\tau} > n\Delta_n\}} \Big| Y_{t_{k+1}}^{\theta(\ell),\beta_0} = X_{t_{k+1}} \right] \mathbf{1}_{\{\widehat{\tau} > n\Delta_n\}}.
\end{aligned}$$

We will later see that the terms concerning $Z_{k,n}^{1,\ell}$ and $Z_{k,n}^{2,\ell}$ are negligible (Lemma 5.3.1). The main contribution in the asymptotics will be given by $Z_{k,n}^{3,\ell}$, which expresses the fact that the small and large jumps do not interfere with the Gaussian behaviour of the transition density. In fact to see this, applying Lemma 5.2.1 to $Z_{k,n}^{3,\ell}$, and using equation (5.1) for the term $X_{t_{k+1}}-X_{t_k}$ coming from the term $Y_{t_{k+1}}^{\theta(\ell),\beta_0}-Y_{t_k}^{\theta(\ell),\beta_0}$ in Lemma 5.2.1, we obtain the following expansion of the log-likelihood ratio

$$\log\frac{p(X^n;(\theta_n,\beta_0))}{p(X^n;(\theta_0,\beta_0))}=\sum_{k=0}^{n-1}\xi_{k,n}+\frac{u}{\sqrt{n\Delta_n^3}}\sum_{k=0}^{n-1}\int_0^1\Big\{Z_{k,n}^{1,\ell}+Z_{k,n}^{2,\ell}$$
$$+\left(Z_{k,n}^{4,\ell}+Z_{k,n}^{5,\ell}+Z_{k,n}^{6,\ell}\right)\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell),\beta_0}\left[\mathbf{1}_{\{\widetilde{\tau}>n\Delta_n\}}\Big|Y_{t_{k+1}}^{\theta(\ell),\beta_0}=X_{t_{k+1}}\right]\mathbf{1}_{\{\widehat{\tau}>n\Delta_n\}}$$
$$+\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell),\beta_0}\left[\left(R^{\theta(\ell),\beta_0}-R_4^{\theta(\ell),\beta_0}-R_6^{\theta(\ell),\beta_0}\right)\mathbf{1}_{\{\widetilde{\tau}>n\Delta_n\}}\Big|Y_{t_{k+1}}^{\theta(\ell),\beta_0}=X_{t_{k+1}}\right]\mathbf{1}_{\{\widehat{\tau}>n\Delta_n\}}\Big\}d\ell,$$

where

$$\xi_{k,n}=\frac{u}{\sqrt{n\Delta_n}}\int_0^1\frac{\partial_\theta b(\theta(\ell),X_{t_k})}{\sigma^2(\beta_0,X_{t_k})}\left(\sigma(\beta_0,X_{t_k})\left(B_{t_{k+1}}-B_{t_k}\right)+\left(b(\theta_0,X_{t_k})-b(\theta(\ell),X_{t_k})\right)\Delta_n\right)$$
$$\times\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell),\beta_0}\left[\mathbf{1}_{\{\widetilde{\tau}>n\Delta_n\}}\Big|Y_{t_{k+1}}^{\theta(\ell),\beta_0}=X_{t_{k+1}}\right]\mathbf{1}_{\{\widehat{\tau}>n\Delta_n\}}d\ell,$$
$$Z_{k,n}^{4,\ell}=\Delta_n\partial_\theta b(\theta(\ell),X_{t_k})\sigma^{-2}(\beta_0,X_{t_k})\int_{t_k}^{t_{k+1}}\left(b(\theta_0,X_s^{\theta_0,\beta_0})-b(\theta_0,X_{t_k})\right)ds,$$
$$Z_{k,n}^{5,\ell}=\Delta_n\partial_\theta b(\theta(\ell),X_{t_k})\sigma^{-2}(\beta_0,X_{t_k})\int_{t_k}^{t_{k+1}}\left(\sigma(\beta_0,X_s^{\theta_0,\beta_0})-\sigma(\beta_0,X_{t_k})\right)dB_s,$$
$$Z_{k,n}^{6,\ell}=\Delta_n\partial_\theta b(\theta(\ell),X_{t_k})\sigma^{-2}(\beta_0,X_{t_k})\int_{t_k}^{t_{k+1}}\int_{\mathbb{R}_0}z\widetilde{N}(ds,dz),$$
$$R^{\theta(\ell),\beta_0}=R_1^{\theta(\ell),\beta_0}+R_2^{\theta(\ell),\beta_0}+R_3^{\theta(\ell),\beta_0}-R_5^{\theta(\ell),\beta_0}.$$

Again the Markov property and Proposition 5.2.1 give

$$\log\frac{p\left(X^n;(\theta_n,\beta_n)\right)}{p\left(X^n;(\theta_n,\beta_0)\right)}=\sum_{k=0}^{n-1}\log\frac{p^{\theta_n,\beta_n}}{p^{\theta_n,\beta_0}}\left(\Delta_n,X_{t_k},X_{t_{k+1}}\right)$$
$$=\sum_{k=0}^{n-1}\frac{v}{\sqrt{n}}\int_0^1\frac{\partial_\beta p^{\theta_n,\beta(\ell)}}{p^{\theta_n,\beta(\ell)}}\left(\Delta_n,X_{t_k},X_{t_{k+1}}\right)d\ell$$
$$=\sum_{k=0}^{n-1}\frac{v}{\sqrt{n}}\int_0^1\frac{1}{\Delta_n}\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[\delta\left(\partial_\beta Y_{t_{k+1}}^{\theta_n,\beta(\ell)}(t_k,X_{t_k})U^{\theta_n,\beta(\ell)}(t_k,X_{t_k})\right)\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}\right]d\ell.$$

Then, multiplying by $\mathbf{1}_{\{\widetilde{\tau}>n\Delta_n\}}+\mathbf{1}_{\{\widetilde{\tau}\leq n\Delta_n\}}$ inside and $\mathbf{1}_{\{\widehat{\tau}>n\Delta_n\}}+\mathbf{1}_{\{\widehat{\tau}\leq n\Delta_n\}}$ outside the conditional expectation above, we get that

$$\log\frac{p\left(X^n;(\theta_n,\beta_n)\right)}{p\left(X^n;(\theta_n,\beta_0)\right)}=\frac{v}{\sqrt{n\Delta_n^2}}\sum_{k=0}^{n-1}\int_0^1\left(Q_{k,n}^{1,\ell}+Q_{k,n}^{2,\ell}+Q_{k,n}^{3,\ell}\right)d\ell,$$

where

$$Q_{k,n}^{1,\ell}=\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[\delta\left(\partial_\beta Y_{t_{k+1}}^{\theta_n,\beta(\ell)}(t_k,X_{t_k})U^{\theta_n,\beta(\ell)}(t_k,X_{t_k})\right)\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}\right]\mathbf{1}_{\{\widehat{\tau}\leq n\Delta_n\}},$$
$$Q_{k,n}^{2,\ell}=\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[\delta\left(\partial_\beta Y_{t_{k+1}}^{\theta_n,\beta(\ell)}(t_k,X_{t_k})U^{\theta_n,\beta(\ell)}(t_k,X_{t_k})\right)\mathbf{1}_{\{\widetilde{\tau}\leq n\Delta_n\}}\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}\right]\mathbf{1}_{\{\widehat{\tau}>n\Delta_n\}},$$
$$Q_{k,n}^{3,\ell}=\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[\delta\left(\partial_\beta Y_{t_{k+1}}^{\theta_n,\beta(\ell)}(t_k,X_{t_k})U^{\theta_n,\beta(\ell)}(t_k,X_{t_k})\right)\mathbf{1}_{\{\widetilde{\tau}>n\Delta_n\}}\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}\right]\mathbf{1}_{\{\widehat{\tau}>n\Delta_n\}}.$$

Similarly, the terms concerning $Q_{k,n}^{1,\ell}$ and $Q_{k,n}^{2,\ell}$ are negligible (Lemma 5.3.6), whereas the main contribution in the asymptotics will be determined by $Q_{k,n}^{3,\ell}$. In fact, applying Lemma 5.2.2 to $Q_{k,n}^{3,\ell}$, and using equation (5.1) for the term $X_{t_{k+1}}-X_{t_k}$ coming from the term $Y_{t_{k+1}}^{\theta_n,\beta(\ell)}-Y_{t_k}^{\theta_n,\beta(\ell)}$ in Lemma 5.2.2, we obtain the following expansion of the log-likelihood ratio

$$
\begin{aligned}
\log\frac{p\left(X^n;(\theta_n,\beta_n)\right)}{p\left(X^n;(\theta_n,\beta_0)\right)} &= \sum_{k=0}^{n-1}\eta_{k,n}+\frac{v}{\sqrt{n\Delta_n^2}}\sum_{k=0}^{n-1}\int_0^1\left(Q_{k,n}^{1,\ell}+Q_{k,n}^{2,\ell}\right)d\ell\\
&+\sum_{k=0}^{n-1}\frac{v}{\sqrt{n\Delta_n^2}}\int_0^1\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[H^{\theta_n,\beta(\ell)}\mathbf{1}_{\{\widetilde{\tau}>n\Delta_n\}}\middle|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}\right]\mathbf{1}_{\{\widehat{\tau}>n\Delta_n\}}d\ell\\
&+\sum_{k=0}^{n-1}\frac{v}{\sqrt{n\Delta_n^2}}\int_0^1\frac{\partial_\beta\sigma}{\sigma^3}\left(\beta(\ell),X_{t_k}\right)\Bigg\{\bigg(\left(H_{11}+H_{12}+H_{13}\right)^2+2\sigma(\beta_0,X_{t_k})\left(B_{t_{k+1}}-B_{t_k}\right)\\
&\quad\times\left(H_{11}+H_{12}+H_{13}\right)\bigg)\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[\mathbf{1}_{\{\widetilde{\tau}>n\Delta_n\}}\middle|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}\right]\mathbf{1}_{\{\widehat{\tau}>n\Delta_n\}}\\
&-\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\bigg[\bigg(\left(H_8^{\theta_n,\beta(\ell)}+H_9^{\theta_n,\beta(\ell)}+H_{10}^{\theta_n,\beta(\ell)}\right)^2+2\sigma(\beta(\ell),Y_{t_k}^{\theta_n,\beta(\ell)})\left(W_{t_{k+1}}-W_{t_k}\right)\\
&\quad\times\left(H_8^{\theta_n,\beta(\ell)}+H_9^{\theta_n,\beta(\ell)}+H_{10}^{\theta_n,\beta(\ell)}\right)\bigg)\mathbf{1}_{\{\widetilde{\tau}>n\Delta_n\}}\bigg|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}\bigg]\mathbf{1}_{\{\widehat{\tau}>n\Delta_n\}}\Bigg\}d\ell,
\end{aligned}
$$

where

$$
\begin{aligned}
\eta_{k,n}&=\frac{v}{\sqrt{n\Delta_n^2}}\int_0^1\left(\frac{\partial_\beta\sigma}{\sigma^3}\left(\beta(\ell),X_{t_k}\right)\sigma^2(\beta_0,X_{t_k})\left(B_{t_{k+1}}-B_{t_k}\right)^2-\frac{\partial_\beta\sigma}{\sigma}\left(\beta(\ell),X_{t_k}\right)\Delta_n\right)\\
&\quad\times\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[\mathbf{1}_{\{\widetilde{\tau}>n\Delta_n\}}\middle|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}\right]\mathbf{1}_{\{\widehat{\tau}>n\Delta_n\}}d\ell,\\
H_{11}&=\int_{t_k}^{t_{k+1}}b(\theta_0,X_s^{\theta_0,\beta_0})ds,\quad H_{12}=\int_{t_k}^{t_{k+1}}\left(\sigma(\beta_0,X_s^{\theta_0,\beta_0})-\sigma(\beta_0,X_{t_k})\right)dB_s,\\
H_{13}&=\int_{t_k}^{t_{k+1}}\int_I z\widetilde{N}(ds,dz),\\
H^{\theta_n,\beta(\ell)}&=H_3^{\theta_n,\beta(\ell)}+H_4^{\theta_n,\beta(\ell)}+H_5^{\theta_n,\beta(\ell)}+H_6^{\theta_n,\beta(\ell)}+H_7^{\theta_n,\beta(\ell)}.
\end{aligned}
$$

Therefore, we have obtained the following expansion of the log-likelihood ratio

$$
\begin{aligned}
\log\frac{p(X^n;(\theta_n,\beta_n))}{p(X^n;(\theta_0,\beta_0))}&=\sum_{k=0}^{n-1}\left(\xi_{k,n}+\eta_{k,n}\right)+\frac{u}{\sqrt{n\Delta_n^3}}\sum_{k=0}^{n-1}\int_0^1\Bigg\{Z_{k,n}^{1,\ell}+Z_{k,n}^{2,\ell}+\left(Z_{k,n}^{4,\ell}+Z_{k,n}^{5,\ell}+Z_{k,n}^{6,\ell}\right)\\
&\quad\times\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell),\beta_0}\left[\mathbf{1}_{\{\widetilde{\tau}>n\Delta_n\}}\middle|Y_{t_{k+1}}^{\theta(\ell),\beta_0}=X_{t_{k+1}}\right]\mathbf{1}_{\{\widehat{\tau}>n\Delta_n\}}\\
&+\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell),\beta_0}\left[\left(R^{\theta(\ell),\beta_0}-R_4^{\theta(\ell),\beta_0}-R_6^{\theta(\ell),\beta_0}\right)\mathbf{1}_{\{\widetilde{\tau}>n\Delta_n\}}\middle|Y_{t_{k+1}}^{\theta(\ell),\beta_0}=X_{t_{k+1}}\right]\mathbf{1}_{\{\widehat{\tau}>n\Delta_n\}}\Bigg\}d\ell\\
&+\frac{v}{\sqrt{n\Delta_n^2}}\sum_{k=0}^{n-1}\int_0^1\left(Q_{k,n}^{1,\ell}+Q_{k,n}^{2,\ell}+\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[H^{\theta_n,\beta(\ell)}\mathbf{1}_{\{\widetilde{\tau}>n\Delta_n\}}\middle|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}\right]\mathbf{1}_{\{\widehat{\tau}>n\Delta_n\}}\right)d\ell\\
&+\sum_{k=0}^{n-1}\frac{v}{\sqrt{n\Delta_n^2}}\int_0^1\frac{\partial_\beta\sigma}{\sigma^3}\left(\beta(\ell),X_{t_k}\right)\Bigg\{\bigg(\left(H_{11}+H_{12}+H_{13}\right)^2+2\sigma(\beta_0,X_{t_k})\left(B_{t_{k+1}}-B_{t_k}\right)\\
&\quad\times\left(H_{11}+H_{12}+H_{13}\right)\bigg)\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[\mathbf{1}_{\{\widetilde{\tau}>n\Delta_n\}}\middle|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}\right]\mathbf{1}_{\{\widehat{\tau}>n\Delta_n\}}\\
&-\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\bigg[\bigg(\left(H_8^{\theta_n,\beta(\ell)}+H_9^{\theta_n,\beta(\ell)}+H_{10}^{\theta_n,\beta(\ell)}\right)^2+2\sigma(\beta(\ell),Y_{t_k}^{\theta_n,\beta(\ell)})\left(W_{t_{k+1}}-W_{t_k}\right)\\
&\quad\times\left(H_8^{\theta_n,\beta(\ell)}+H_9^{\theta_n,\beta(\ell)}+H_{10}^{\theta_n,\beta(\ell)}\right)\bigg)\mathbf{1}_{\{\widetilde{\tau}>n\Delta_n\}}\bigg|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}\bigg]\mathbf{1}_{\{\widehat{\tau}>n\Delta_n\}}\Bigg\}d\ell.
\end{aligned}
$$

In the next subsections we will show that $\xi_{k,n}$ and $\eta_{k,n}$ are the terms that contribute to the limit in Theorem 5.1.1, and all the others are negligible contributions. Therefore again, the main behaviour is given by the Gaussian and drift components of the equation (5.1).

5.3.2 Negligible contributions

To simplify the exposition, let us denote $U_1^{\theta,\beta}(t_k,x)=\partial_\theta Y_{t_{k+1}}^{\theta,\beta}(t_k,x)U^{\theta,\beta}(t_k,x)$ and $U_2^{\theta,\beta}(t_k,x)=\partial_\beta Y_{t_{k+1}}^{\theta,\beta}(t_k,x)U^{\theta,\beta}(t_k,x)$.

Lemma 5.3.1. *Under conditions* **(A1)-(A4)** *and* **(A6)**, *as* $n\to\infty$,

$$\frac{u}{\sqrt{n\Delta_n^3}}\sum_{k=0}^{n-1}\int_0^1\left(Z_{k,n}^{1,\ell}+Z_{k,n}^{2,\ell}\right)d\ell\overset{\mathrm{P}^{\theta_0,\beta_0}}{\longrightarrow}0.$$

Proof. It suffices to show that condition (1.16) of Lemma 1.4.2 holds for each sequence $(Z_{k,n}^{i,\ell})_{k\geq 1}$ under the measure $\mathrm{P}^{\theta_0,\beta_0}$.

First, applying Hölder's and Jensen's inequalities, Girsanov's theorem, Lemma 5.2.12, and (5.8), we obtain that for some constants $C,q_0>0$,

$$\begin{aligned}
&\frac{|u|}{\sqrt{n\Delta_n^3}}\sum_{k=0}^{n-1}\mathrm{E}\left[\left|\int_0^1 Z_{k,n}^{1,\ell}d\ell\right|\Big|\widehat{\mathcal{F}}_{t_k}\right]\leq\frac{|u|}{\sqrt{n\Delta_n^3}}\sum_{k=0}^{n-1}\int_0^1\left(\mathrm{P}\left(\widehat{\tau}\leq n\Delta_n|X_{t_k}\right)\right)^{\frac{1}{q}}\\
&\quad\times\left(\mathrm{E}\left[\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell),\beta_0}\left[\left|\delta\left(U_1^{\theta(\ell),\beta_0}(t_k,X_{t_k})\right)\right|^p\Big|Y_{t_{k+1}}^{\theta(\ell),\beta_0}=X_{t_{k+1}}\right]\Big|X_{t_k}\right]\right)^{\frac{1}{p}}d\ell\\
&\leq\frac{|u|}{\sqrt{n\Delta_n^3}}\sum_{k=0}^{n-1}\int_0^1\left(\mathrm{E}_{\widehat{Q}_k^{\theta(\ell),\theta_0,\beta_0}}\left[\left|\delta\left(U_1^{\theta(\ell),\beta_0}(t_k,X_{t_k})\right)\right|^p\left(\frac{d\widehat{\mathrm{P}}}{d\widehat{Q}_k^{\theta(\ell),\theta_0,\beta_0}}-1\right)\Big|X_{t_k}\right]\right.\\
&\quad\left.+\mathrm{E}_{\widehat{Q}_k^{\theta(\ell),\theta_0,\beta_0}}\left[\left|\delta\left(U_1^{\theta(\ell),\beta_0}(t_k,X_{t_k})\right)\right|^p\Big|X_{t_k}\right]\right)^{\frac{1}{p}}\left(\mathrm{P}\left(\widehat{\tau}\leq n\Delta_n|X_{t_k}\right)\right)^{\frac{1}{q}}d\ell\\
&\leq C\frac{|u|}{\sqrt{n}}\sum_{k=0}^{n-1}\left(1+|X_{t_k}|^{q_0}\right)\left(\mathrm{P}\left(\widehat{\tau}\leq n\Delta_n|X_{t_k}\right)\right)^{\frac{1}{q}},
\end{aligned}$$

where $p,q>1$ and $\frac{1}{p}+\frac{1}{q}=1$. On the other hand,

$$\begin{aligned}
&\mathrm{P}(\widehat{\tau}>n\Delta_n|X_{t_k})=\mathrm{P}(\forall s\in[0,n\Delta_n],\ \ \rho_1\Delta_n^{\upsilon}\leq|\Delta\widehat{Z}_s|\leq\rho_2\Delta_n^{-\gamma}|X_{t_k})\\
&=\mathrm{P}(\forall s\in[0,k\Delta_n],\ \ \rho_1\Delta_n^{\upsilon}\leq|\Delta\widehat{Z}_s|\leq\rho_2\Delta_n^{-\gamma}|X_{t_k})\\
&\quad\times\sum_{j=0}^{\infty}\mathrm{P}(\{\forall s\in[k\Delta_n,n\Delta_n],\ \ \rho_1\Delta_n^{\upsilon}\leq|\Delta\widehat{Z}_s|\leq\rho_2\Delta_n^{-\gamma}\}\cap\{N_{n\Delta_n}-N_{k\Delta_n}=j\})\\
&\leq\sum_{j=0}^{\infty}e^{-\lambda_n(n-k)\Delta_n}\frac{(\lambda_n(n-k)\Delta_n)^j}{j!}\mathrm{P}(\rho_1\Delta_n^{\upsilon}\leq|\widehat{\Lambda}|\leq\rho_2\Delta_n^{-\gamma})^j\\
&=e^{-\lambda_n(n-k)\Delta_n\left(1-\mathrm{P}\left(\rho_1\Delta_n^{\upsilon}\leq|\widehat{\Lambda}|\leq\rho_2\Delta_n^{-\gamma}\right)\right)},
\end{aligned}\tag{5.37}$$

where $\widehat{\Lambda}$ is a random variable with distribution $\frac{\nu}{\lambda}$. Therefore, we obtain that

$$\begin{aligned}
&\frac{|u|}{\sqrt{n\Delta_n^3}}\sum_{k=0}^{n-1}\mathrm{E}\left[\left|\int_0^1 Z_{k,n}^{1,\ell}d\ell\right|\Big|\widehat{\mathcal{F}}_{t_k}\right]\\
&\leq\frac{C|u|}{\sqrt{n}}\sum_{k=0}^{n-1}\left(1+|X_{t_k}|^{q_0}\right)\left(1-e^{-\lambda_n(n-k)\Delta_n\left(1-\mathrm{P}\left(\rho_1\Delta_n^{\upsilon}\leq|\widehat{\Lambda}|\leq\rho_2\Delta_n^{-\gamma}\right)\right)}\right)^{\frac{1}{q}}\\
&\leq\left(1-e^{-\lambda_n n\Delta_n\left(1-\mathrm{P}\left(\rho_1\Delta_n^{\upsilon}\leq|\widehat{\Lambda}|\leq\rho_2\Delta_n^{-\gamma}\right)\right)}\right)^{\frac{1}{q}}\frac{C|u|}{\sqrt{n}}\sum_{k=0}^{n-1}\left(1+|X_{t_k}|^{q_0}\right).
\end{aligned}$$

Then, using the fact that $1-e^{-x}\leq x$, for all $x\geq 0$, and that $\lambda_n\leq\lambda$, we get that

$$\begin{aligned}
&\left(1-e^{-\lambda_n n\Delta_n\left(1-\mathrm{P}\left(\rho_1\Delta_n^{\upsilon}\leq|\widehat{\Lambda}|\leq\rho_2\Delta_n^{-\gamma}\right)\right)}\right)^{\frac{1}{q}}\\
&\leq\left(\lambda n\Delta_n\left(1-\mathrm{P}\left(\rho_1\Delta_n^{\upsilon}\leq|\widehat{\Lambda}|\leq\rho_2\Delta_n^{-\gamma}\right)\right)\right)^{\frac{1}{q}}\\
&\leq c_q\left\{\left(\lambda n\Delta_n\mathrm{P}\left(|\widehat{\Lambda}|\geq\rho_2\Delta_n^{-\gamma}\right)\right)^{\frac{1}{q}}+\left(\lambda n\Delta_n\mathrm{P}\left(\rho_1\Delta_n^{\upsilon}\geq|\widehat{\Lambda}|\right)\right)^{\frac{1}{q}}\right\}\\
&=c_q\left\{\left(n\Delta_n\int_{\{|z|\geq\rho_2\Delta_n^{-\gamma}\}}\nu(dz)\right)^{\frac{1}{q}}+\left(n\Delta_n\int_{\{|z|\leq\rho_1\Delta_n^{\upsilon}\}}\nu(dz)\right)^{\frac{1}{q}}\right\}.
\end{aligned}$$

Therefore, by **(A6)** we conclude that (1.16) holds true, and by Lemma 1.4.2, as $n\to\infty$,

$$\frac{u}{\sqrt{n\Delta_n^3}}\sum_{k=0}^{n-1}\int_0^1 Z_{k,n}^{1,\ell}d\ell\overset{\mathrm{P}^{\theta_0,\beta_0}}{\longrightarrow}0.$$

Next, as for the term $Z_{k,n}^{1,\ell}$, applying Girsanov's theorem, Lemma 5.2.12, and (5.8), we obtain that for some constants $C,q_0>0$,

$$\begin{aligned}
&\frac{|u|}{\sqrt{n\Delta_n^3}}\sum_{k=0}^{n-1}\mathrm{E}\left[\left|\int_0^1 Z_{k,n}^{2,\ell}d\ell\right|\Big|\widehat{\mathcal{F}}_{t_k}\right]\\
&\leq\frac{|u|}{\sqrt{n\Delta_n^3}}\sum_{k=0}^{n-1}\int_0^1\mathrm{E}\left[\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell),\beta_0}\left[\left|\delta\left(U_1^{\theta(\ell),\beta_0}(t_k,X_{t_k})\right)\right|\mathbf{1}_{\{\widetilde{\tau}\leq n\Delta_n\}}\Big|Y_{t_{k+1}}^{\theta(\ell),\beta_0}=X_{t_{k+1}}\right]\Big|X_{t_k}\right]d\ell\\
&\leq\frac{|u|}{\sqrt{n\Delta_n^3}}\sum_{k=0}^{n-1}\int_0^1\left(\left|\mathrm{E}_{\widehat{Q}_k^{\theta(\ell),\theta_0,\beta_0}}\left[\left|\delta\left(U_1^{\theta(\ell),\beta_0}(t_k,X_{t_k})\right)\right|\mathbf{1}_{\{\widetilde{\tau}\leq n\Delta_n\}}\left(\frac{d\widehat{\mathrm{P}}}{d\widehat{Q}_k^{\theta(\ell),\theta_0,\beta_0}}-1\right)\Big|X_{t_k}\right]\right|\right.\\
&\quad\left.+\mathrm{E}_{\widehat{Q}_k^{\theta(\ell),\theta_0,\beta_0}}\left[\left|\delta\left(U_1^{\theta(\ell),\beta_0}(t_k,X_{t_k})\right)\right|\mathbf{1}_{\{\widetilde{\tau}\leq n\Delta_n\}}\Big|X_{t_k}\right]\right)d\ell\\
&\leq\frac{C|u|}{\sqrt{n}}\sum_{k=0}^{n-1}\left(1+|X_{t_k}|^{q_0}\right)\left(\mathrm{P}\left(\widetilde{\tau}\leq n\Delta_n|X_{t_k}\right)\right)^{\frac{1}{q}},
\end{aligned}$$

where we have used Hölder's inequality with $p>1$ and $q>1$ conjugate. On the other hand,

$$\begin{aligned}
\mathrm{P}(\widetilde{\tau}>n\Delta_n|X_{t_k})&=\mathrm{P}(\widetilde{\tau}>n\Delta_n)=\mathrm{P}(\forall s\in[0,n\Delta_n],\ \ \rho_1\Delta_n^{\upsilon}\leq|\Delta\widetilde{Z}_s|\leq\rho_2\Delta_n^{-\gamma})\\
&=\sum_{j=0}^{\infty}\mathrm{P}(\{\forall s\in[0,n\Delta_n],\ \ \rho_1\Delta_n^{\upsilon}\leq|\Delta\widetilde{Z}_s|\leq\rho_2\Delta_n^{-\gamma}\}\cap\{M_{n\Delta_n}-M_0=j\})\\
&=\sum_{j=0}^{\infty}e^{-\lambda_n n\Delta_n}\frac{(\lambda_n n\Delta_n)^j}{j!}\mathrm{P}(\rho_1\Delta_n^{\upsilon}\leq|\widetilde{\Lambda}|\leq\rho_2\Delta_n^{-\gamma})^j\\
&=e^{-\lambda_n n\Delta_n\left(1-\mathrm{P}\left(\rho_1\Delta_n^{\upsilon}\leq|\widetilde{\Lambda}|\leq\rho_2\Delta_n^{-\gamma}\right)\right)},
\end{aligned}$$

where $\widetilde{\Lambda}$ is a random variable with distribution $\frac{\nu}{\lambda}$. Therefore, we obtain that

$$\begin{aligned}
&\frac{|u|}{\sqrt{n\Delta_n^3}}\sum_{k=0}^{n-1}\mathrm{E}\left[\left|\int_0^1 Z_{k,n}^{2,\ell}d\ell\right|\Big|\widehat{\mathcal{F}}_{t_k}\right]\\
&\leq\frac{|u|}{\sqrt{n}}\sum_{k=0}^{n-1}\left(1+|X_{t_k}|^{q_0}\right)\left(1-e^{-\lambda_n n\Delta_n\left(1-\mathrm{P}\left(\rho_1\Delta_n^{\upsilon}\leq|\widetilde{\Lambda}|\leq\rho_2\Delta_n^{-\gamma}\right)\right)}\right)^{\frac{1}{q}}\\
&\leq c_q\left\{\left(n\Delta_n\int_{\{|z|\geq\rho_2\Delta_n^{-\gamma}\}}\nu(dz)\right)^{\frac{1}{q}}+\left(n\Delta_n\int_{\{|z|\leq\rho_1\Delta_n^{\upsilon}\}}\nu(dz)\right)^{\frac{1}{q}}\right\}\frac{|u|}{\sqrt{n}}\sum_{k=0}^{n-1}\left(1+|X_{t_k}|^{q_0}\right).
\end{aligned}$$

Therefore, by **(A6)** we conclude that (1.16) holds true, and by Lemma 1.4.2, as $n \to \infty$,

$$\frac{u}{\sqrt{n\Delta_n^3}} \sum_{k=0}^{n-1} \int_0^1 Z_{k,n}^{2,\ell} d\ell \overset{\mathrm{P}^{\theta_0}}{\longrightarrow} 0.$$

Thus, the result follows. □

Lemma 5.3.2. *Under conditions* **(A1)**-**(A4)** *and* **(A6)***, as* $n \to \infty$,

$$\sum_{k=0}^{n-1} \frac{u}{\sqrt{n\Delta_n^3}} \int_0^1 \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell),\beta_0} \left[R^{\theta(\ell),\beta_0} \mathbf{1}_{\{\widetilde{\tau} > n\Delta_n\}} \Big| Y_{t_{k+1}}^{\theta(\ell),\beta_0} = X_{t_{k+1}} \right] \mathbf{1}_{\{\widehat{\tau} > n\Delta_n\}} d\ell \overset{\mathrm{P}^{\theta_0,\beta_0}}{\longrightarrow} 0.$$

Proof. Using the fact that $\mathbf{1}_{\{\widehat{\tau} > n\Delta_n\}} = 1 - \mathbf{1}_{\{\widehat{\tau} \leq n\Delta_n\}}$ and $\mathbf{1}_{\{\widetilde{\tau} > n\Delta_n\}} = 1 - \mathbf{1}_{\{\widetilde{\tau} \leq n\Delta_n\}}$, it suffices to show that as $n \to \infty$,

$$\sum_{k=0}^{n-1} \frac{u}{\sqrt{n\Delta_n^3}} \int_0^1 \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell),\beta_0} \left[R^{\theta(\ell),\beta_0} \mathbf{1}_{\{\widetilde{\tau} > n\Delta_n\}} \Big| Y_{t_{k+1}}^{\theta(\ell),\beta_0} = X_{t_{k+1}} \right] \mathbf{1}_{\{\widehat{\tau} \leq n\Delta_n\}} d\ell \overset{\mathrm{P}^{\theta_0,\beta_0}}{\longrightarrow} 0, \tag{5.38}$$

$$\sum_{k=0}^{n-1} \frac{u}{\sqrt{n\Delta_n^3}} \int_0^1 \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell),\beta_0} \left[R^{\theta(\ell),\beta_0} \mathbf{1}_{\{\widetilde{\tau} \leq n\Delta_n\}} \Big| Y_{t_{k+1}}^{\theta(\ell),\beta_0} = X_{t_{k+1}} \right] d\ell \overset{\mathrm{P}^{\theta_0,\beta_0}}{\longrightarrow} 0, \tag{5.39}$$

$$\sum_{k=0}^{n-1} \frac{u}{\sqrt{n\Delta_n^3}} \int_0^1 \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell),\beta_0} \left[R^{\theta(\ell),\beta_0} \Big| Y_{t_{k+1}}^{\theta(\ell),\beta_0} = X_{t_{k+1}} \right] d\ell \overset{\mathrm{P}^{\theta_0,\beta_0}}{\longrightarrow} 0. \tag{5.40}$$

The convergences (5.38) and (5.39) are treated similarly as for the terms $Z_{k,n}^{1,\ell}$ and $Z_{k,n}^{2,\ell}$. To treat (5.40), it suffices to show that conditions (i) and (ii) of Lemma 1.4.1 hold under the measure $\mathrm{P}^{\theta_0,\beta_0}$. We start showing (i). Applying Girsanov's theorem, Lemma 5.2.12, (5.6) and (5.7) with $p = 2$, we get that

$$\begin{aligned}
&\left| \sum_{k=0}^{n-1} \frac{u}{\sqrt{n\Delta_n^3}} \int_0^1 \mathrm{E} \left[\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell),\beta_0} \left[R^{\theta(\ell),\beta_0} \Big| Y_{t_{k+1}}^{\theta(\ell),\beta_0} = X_{t_{k+1}} \right] \Big| \widehat{\mathcal{F}}_{t_k} \right] d\ell \right| \\
&\leq \sum_{k=0}^{n-1} \frac{|u|}{\sqrt{n\Delta_n^3}} \int_0^1 \left| \mathrm{E}_{\widehat{Q}_k^{\theta(\ell),\theta_0,\beta_0}} \left[R^{\theta(\ell),\beta_0} \left(\frac{d\widehat{\mathrm{P}}}{d\widehat{Q}_k^{\theta(\ell),\theta_0,\beta_0}} - 1 \right) \Big| X_{t_k} \right] \right| d\ell \\
&\leq \frac{C|u|\Delta_n^{1/4}}{n} \sum_{k=0}^{n-1} (1 + |X_{t_k}|^q).
\end{aligned}$$

for some constants $C, q > 0$. Observe that (5.7) and (5.8) remain valid under $\widehat{\mathrm{P}}^{\alpha}$, the measure defined in Lemma 5.2.12. This shows Lemma 1.4.1 (i). Similarly, applying Jensen's inequality, Girsanov's theorem, Lemma 5.2.12 and (5.7) with $p \in \{2, 4\}$, we obtain that

$$\begin{aligned}
&\sum_{k=0}^{n-1} \frac{u^2}{n\Delta_n^3} \mathrm{E} \left[\left(\int_0^1 \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell),\beta_0} \left[R^{\theta(\ell),\beta_0} \Big| Y_{t_{k+1}}^{\theta(\ell),\beta_0} = X_{t_{k+1}} \right] d\ell \right)^2 \Big| \widehat{\mathcal{F}}_{t_k} \right] \\
&\leq \sum_{k=0}^{n-1} \frac{u^2}{n\Delta_n^3} \int_0^1 \left\{ \mathrm{E}_{\widehat{Q}_k^{\theta(\ell),\theta_0,\beta_0}} \left[\left(R^{\theta(\ell),\beta_0} \right)^2 \Big| X_{t_k} \right] \right. \\
&\qquad \left. + \left| \mathrm{E}_{\widehat{Q}_k^{\theta(\ell),\theta_0,\beta_0}} \left[\left(R^{\theta(\ell),\beta_0} \right)^2 \left(\frac{d\widehat{\mathrm{P}}}{d\widehat{Q}_k^{\theta(\ell),\theta_0,\beta_0}} - 1 \right) \Big| X_{t_k} \right] \right| \right\} d\ell \\
&\leq \frac{Cu^2\Delta_n^{1/4}}{n} \sum_{k=0}^{n-1} (1 + |X_{t_k}|^q),
\end{aligned}$$

which concludes the desired result. □

Lemma 5.3.3. *Under conditions* **(A1)**-**(A4)** *and* **(A6)**, *as* $n \to \infty$,

$$\sum_{k=0}^{n-1} \frac{u}{\sqrt{n\Delta_n^3}} \int_0^1 Z_{k,n}^{5,\ell} \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell),\beta_0} \left[\mathbf{1}_{\{\widetilde{\tau}>n\Delta_n\}} \middle| Y_{t_{k+1}}^{\theta(\ell),\beta_0} = X_{t_{k+1}} \right] \mathbf{1}_{\{\widehat{\tau}>n\Delta_n\}} d\ell \overset{\mathrm{P}^{\theta_0,\beta_0}}{\longrightarrow} 0.$$

Proof. Using the fact that $\mathbf{1}_{\{\widehat{\tau}>n\Delta_n\}} = 1 - \mathbf{1}_{\{\widehat{\tau}\leq n\Delta_n\}}$ and $\mathbf{1}_{\{\widetilde{\tau}>n\Delta_n\}} = 1 - \mathbf{1}_{\{\widetilde{\tau}\leq n\Delta_n\}}$, it suffices to show that as $n \to \infty$,

$$\sum_{k=0}^{n-1} \frac{u}{\sqrt{n\Delta_n^3}} \int_0^1 Z_{k,n}^{5,\ell} d\ell \overset{\mathrm{P}^{\theta_0,\beta_0}}{\longrightarrow} 0.$$

Clearly, for all $n \geq 1$,

$$\sum_{k=0}^{n-1} \frac{u}{\sqrt{n\Delta_n^3}} \int_0^1 \mathrm{E}\left[Z_{k,n}^{5,\ell} | \widehat{\mathcal{F}}_{t_k} \right] d\ell = 0,$$

and by Lemma 5.2.3(i),

$$\sum_{k=0}^{n-1} \frac{u^2}{n\Delta_n^3} \mathrm{E}\left[\left(\int_0^1 Z_{k,n}^{5,\ell} d\ell \right)^2 \middle| \widehat{\mathcal{F}}_{t_k} \right] \leq \frac{Cu^2\Delta_n}{n} \sum_{k=0}^{n-1} \left(1 + |X_{t_k}|^q\right),$$

for some constant $C, q > 0$. Thus, Lemma 1.4.1 concludes the desired result. □

Lemma 5.3.4. *Assume conditions* **(A1)**-**(A4)** *and* **(A6)**. *Then as* $n \to \infty$,

$$\sum_{k=0}^{n-1} \frac{u}{\sqrt{n\Delta_n^3}} \int_0^1 \Bigg(Z_{k,n}^{4,\ell} \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell),\beta_0} \left[\mathbf{1}_{\{\widetilde{\tau}>n\Delta_n\}} \middle| Y_{t_{k+1}}^{\theta(\ell),\beta_0} = X_{t_{k+1}} \right] \mathbf{1}_{\{\widehat{\tau}>n\Delta_n\}}$$
$$- \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell),\beta_0} \left[R_4^{\theta(\ell),\beta_0} \mathbf{1}_{\{\widetilde{\tau}>n\Delta_n\}} \middle| Y_{t_{k+1}}^{\theta(\ell),\beta_0} = X_{t_{k+1}} \right] \mathbf{1}_{\{\widehat{\tau}>n\Delta_n\}} \Bigg) d\ell \overset{\mathrm{P}^{\theta_0,\beta_0}}{\longrightarrow} 0.$$

Proof. Using the fact that $\mathbf{1}_{\{\widehat{\tau}>n\Delta_n\}} = 1 - \mathbf{1}_{\{\widehat{\tau}\leq n\Delta_n\}}$ and $\mathbf{1}_{\{\widetilde{\tau}>n\Delta_n\}} = 1 - \mathbf{1}_{\{\widetilde{\tau}\leq n\Delta_n\}}$, it suffices to show that as $n \to \infty$,

$$\sum_{k=0}^{n-1} \frac{u}{\sqrt{n\Delta_n^3}} \int_0^1 \left(Z_{k,n}^{4,\ell} - \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell),\beta_0} \left[R_4^{\theta(\ell),\beta_0} \middle| Y_{t_{k+1}}^{\theta(\ell),\beta_0} = X_{t_{k+1}} \right] \right) d\ell \overset{\mathrm{P}^{\theta_0,\beta_0}}{\longrightarrow} 0.$$

By the mean value theorem,

$$Z_{k,n}^{4,\ell} - \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell),\beta_0} \left[R_4^{\theta(\ell),\beta_0} \middle| Y_{t_{k+1}}^{\theta(\ell),\beta_0} = X_{t_{k+1}} \right] = M_{k,n,1} + M_{k,n,2},$$

where

$$M_{k,n,1} := -\frac{\ell u \Delta_n}{\sqrt{n\Delta_n}} \frac{\partial_\theta b(\theta(\ell), X_{t_k})}{\sigma^2(\beta_0, X_{t_k})} \int_{t_k}^{t_{k+1}} \left(\partial_\theta b(\theta_0 + \frac{\ell u w}{\sqrt{n\Delta_n}}, X_s^{\theta_0,\beta_0}) - \partial_\theta b(\theta_0 + \frac{\ell u w}{\sqrt{n\Delta_n}}, X_{t_k}) \right) ds,$$

$$M_{k,n,2} := \Delta_n \frac{\partial_\theta b(\theta(\ell), X_{t_k})}{\sigma^2(\beta_0, X_{t_k})} \Bigg(\int_{t_k}^{t_{k+1}} \left(b(\theta(\ell), X_s^{\theta_0,\beta_0}) - b(\theta(\ell), X_{t_k}) \right) ds$$
$$- \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell),\beta_0} \left[\int_{t_k}^{t_{k+1}} \left(b(\theta(\ell), Y_s^{\theta(\ell),\beta_0}) - b(\theta(\ell), Y_{t_k}^{\theta(\ell),\beta_0}) \right) ds \middle| Y_{t_{k+1}}^{\theta(\ell),\beta_0} = X_{t_{k+1}} \right] \Bigg),$$

for some $w \in (0, 1)$.

Using Lemma 5.2.3(i), we get that

$$\sum_{k=0}^{n-1} \frac{|u|}{\sqrt{n\Delta_n^3}} \mathrm{E}\left[\left| \int_0^1 M_{k,n,1} d\ell \right| \middle| \widehat{\mathcal{F}}_{t_k} \right] \leq \frac{Cu^2\sqrt{\Delta_n}}{n} \sum_{k=0}^{n-1} \left(1 + |X_{t_k}|^q\right),$$

for some constants $C, q > 0$. Therefore, by Lemma 1.4.2, we conclude that as $n \to \infty$,

$$\sum_{k=0}^{n-1} \frac{u}{\sqrt{n\Delta_n^3}} \int_0^1 M_{k,n,1} d\ell \overset{\mathrm{P}^{\theta_0,\beta_0}}{\longrightarrow} 0.$$

We next show that as $n \to \infty$,

$$\sum_{k=0}^{n-1} \frac{u}{\sqrt{n\Delta_n^3}} \int_0^1 M_{k,n,2} d\ell \overset{\mathrm{P}^{\theta_0,\beta_0}}{\longrightarrow} 0.$$

Using Girsanov's theorem, and Lemmas 5.2.12 and 5.2.3(i), we obtain that

$$\begin{aligned}
&\left|\sum_{k=0}^{n-1} \frac{u}{\sqrt{n\Delta_n^3}} \int_0^1 \mathrm{E}\left[M_{k,n,2} | \widehat{\mathcal{F}}_{t_k}\right] d\ell\right| = \left|\sum_{k=0}^{n-1} \frac{u}{\sqrt{n\Delta_n}} \int_0^1 \frac{\partial_\theta b(\theta(\ell), X_{t_k})}{\sigma^2(\beta_0, X_{t_k})}\right.\\
&\times \left\{ \int_{t_k}^{t_{k+1}} \mathrm{E}_{\widehat{Q}_k^{\theta(\ell),\theta_0,\beta_0}} \left[\left(b(\theta(\ell), X_s^{\theta(\ell),\beta_0}) - b(\theta(\ell), X_{t_k}^{\theta(\ell),\beta_0}) \right) \left(\frac{d\widehat{\mathrm{P}}}{d\widehat{Q}_k^{\theta(\ell),\theta_0,\beta_0}} - 1 \right) \Big| X_{t_k} \right] ds \right.\\
&\left.\left. - \mathrm{E}_{\widehat{Q}_k^{\theta(\ell),\theta_0,\beta_0}} \left[\int_{t_k}^{t_{k+1}} \left(b(\theta(\ell), Y_s^{\theta(\ell),\beta_0}) - b(\theta(\ell), Y_{t_k}^{\theta(\ell),\beta_0}) \right) ds \left(\frac{d\widehat{\mathrm{P}}}{d\widehat{Q}_k^{\theta(\ell),\theta_0,\beta_0}} - 1 \right) \Big| X_{t_k} \right] \right\} d\ell \right|\\
&\leq \frac{C|u|\Delta_n}{n} \sum_{k=0}^{n-1} (1 + |X_{t_k}|^q),
\end{aligned}$$

for some constants $C, q > 0$, which shows Lemma 1.4.1(i).

Next, proceeding as in the proof of Lemma 5.3.2 to show that condition (ii) of Lemma 1.4.1 holds. Thus, the result follows. □

Lemma 5.3.5. *Assume conditions* **(A1)**-**(A4)** *and* **(A6)**. *Then as* $n \to \infty$,

$$\begin{aligned}
&\sum_{k=0}^{n-1} \frac{u}{\sqrt{n\Delta_n^3}} \int_0^1 \left(Z_{k,n}^{6,\ell} \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell),\beta_0} \left[\mathbf{1}_{\{\widetilde{\tau} > n\Delta_n\}} \Big| Y_{t_{k+1}}^{\theta(\ell),\beta_0} = X_{t_{k+1}} \right] \mathbf{1}_{\{\widehat{\tau} > n\Delta_n\}} \right.\\
&\qquad \left. - \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell),\beta_0} \left[R_6^{\theta(\ell),\beta_0} \mathbf{1}_{\{\widetilde{\tau} > n\Delta_n\}} \Big| Y_{t_{k+1}}^{\theta(\ell),\beta_0} = X_{t_{k+1}} \right] \mathbf{1}_{\{\widehat{\tau} > n\Delta_n\}} \right) d\ell \overset{\mathrm{P}^{\theta_0,\beta_0}}{\longrightarrow} 0.
\end{aligned}$$

Proof. Using the fact that $\mathbf{1}_{\{\widehat{\tau} > n\Delta_n\}} = 1 - \mathbf{1}_{\{\widehat{\tau} \leq n\Delta_n\}}$ and $\mathbf{1}_{\{\widetilde{\tau} > n\Delta_n\}} = 1 - \mathbf{1}_{\{\widetilde{\tau} \leq n\Delta_n\}}$, it suffices to show that as $n \to \infty$,

$$\begin{aligned}
&\sum_{k=0}^{n-1} \frac{u}{\sqrt{n\Delta_n}} \int_0^1 \frac{\partial_\theta b(\theta(\ell), X_{t_k})}{\sigma^2(\beta_0, X_{t_k})} \left(\int_{t_k}^{t_{k+1}} \int_I z\widetilde{N}(ds, dz) \right.\\
&\qquad \left. - \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell),\beta_0} \left[\int_{t_k}^{t_{k+1}} \int_I z\widetilde{M}(ds, dz) \Big| Y_{t_{k+1}}^{\theta(\ell),\beta_0} = X_{t_{k+1}} \right] \right) d\ell \overset{\mathrm{P}^{\theta_0,\beta_0}}{\longrightarrow} 0.
\end{aligned}$$

First, by Girsanov's theorem,

$$\begin{aligned}
&\mathrm{E}\left[\int_{t_k}^{t_{k+1}} \int_I z\widetilde{N}(ds, dz) - \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell),\beta_0} \left[\int_{t_k}^{t_{k+1}} \int_I z\widetilde{M}(ds, dz) \Big| Y_{t_{k+1}}^{\theta(\ell),\beta_0} = X_{t_{k+1}} \right] \Big| \widehat{\mathcal{F}}_{t_k} \right]\\
&= -\mathrm{E}_{\widehat{Q}_k^{\theta(\ell),\theta_0,\beta_0}} \left[\int_{t_k}^{t_{k+1}} \int_I z\widetilde{M}(ds, dz) \frac{d\widehat{\mathrm{P}}}{d\widehat{Q}_k^{\theta(\ell),\theta_0,\beta_0}} \Big| X_{t_k} \right]\\
&= 0,
\end{aligned}$$

where we have used the independence between $\int_{t_k}^{t_{k+1}} \int_I z\widetilde{M}(ds, dz)$ and $\frac{d\widehat{\mathrm{P}}}{d\widehat{Q}_k^{\theta(\ell),\theta_0,\beta_0}}$ together with the fact that $\mathrm{E}_{\widehat{Q}_k^{\theta(\ell),\theta_0,\beta_0}}[\int_{t_k}^{t_{k+1}} \int_I z\widetilde{M}(ds, dz)] = 0$. This shows that the term (i) of Lemma 1.4.1 is actually equal to 0 for all $n \geq 1$.

We next show that condition (ii) of Lemma 1.4.1 holds. Cauchy-Schwarz inequality and Girsanov's theorem give

$$\sum_{k=0}^{n-1}\frac{u^2}{n\Delta_n}\mathrm{E}\Big[\Big(\int_0^1\frac{\partial_\theta b(\theta(\ell),X_{t_k})}{\sigma^2(\beta_0,X_{t_k})}\Big(\int_{t_k}^{t_{k+1}}\int_I z\widetilde{N}(ds,dz)$$
$$-\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell),\beta_0}\left[\int_{t_k}^{t_{k+1}}\int_I z\widetilde{M}(ds,dz)\Big|Y_{t_{k+1}}^{\theta(\ell),\beta_0}=X_{t_{k+1}}\right]\Big)d\ell\Big)^2\Big|\widehat{\mathcal{F}}_{t_k}\Big]\leq D_1+D_2,$$

where

$$D_1:=\frac{u^2}{n\Delta_n}\sum_{k=0}^{n-1}\int_0^1\left(\frac{\partial_\theta b(\theta(\ell),X_{t_k})}{\sigma^2(\beta_0,X_{t_k})}\right)^2\mathrm{E}_{\widehat{Q}_k^{\theta(\ell),\theta_0,\beta_0}}\Big[\Big(\int_{t_k}^{t_{k+1}}\int_I zN(ds,dz)$$
$$-\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell),\beta_0}\left[\int_{t_k}^{t_{k+1}}\int_I zM(ds,dz)\Big|Y_{t_{k+1}}^{\theta(\ell),\beta_0}=X_{t_{k+1}}\right]\Big)^2\left(\frac{d\widehat{\mathrm{P}}}{d\widehat{Q}_k^{\theta(\ell),\theta_0,\beta_0}}-1\right)\Big|X_{t_k}\Big]d\ell,$$

$$D_2:=\frac{u^2}{n\Delta_n}\sum_{k=0}^{n-1}\int_0^1\left(\frac{\partial_\theta b(\theta(\ell),X_{t_k})}{\sigma^2(\beta_0,X_{t_k})}\right)^2\mathrm{E}_{\widehat{Q}_k^{\theta(\ell),\theta_0,\beta_0}}\Big[\Big(\int_{t_k}^{t_{k+1}}\int_I zN(ds,dz)$$
$$-\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell),\beta_0}\left[\int_{t_k}^{t_{k+1}}\int_I zM(ds,dz)\Big|Y_{t_{k+1}}^{\theta(\ell),\beta_0}=X_{t_{k+1}}\right]\Big)^2\Big|X_{t_k}\Big]d\ell.$$

Observe that $|D_1|\leq 2(D_{1,1}+D_{1,2})$, where

$$D_{1,1}:=\frac{u^2}{n\Delta_n}\sum_{k=0}^{n-1}\int_0^1\left(\frac{\partial_\theta b(\theta(\ell),X_{t_k})}{\sigma^2(\beta_0,X_{t_k})}\right)^2$$
$$\times\left|\mathrm{E}_{\widehat{Q}_k^{\theta(\ell),\theta_0,\beta_0}}\left[\left(\int_{t_k}^{t_{k+1}}\int_I zN(ds,dz)\right)^2\left(\frac{d\widehat{\mathrm{P}}}{d\widehat{Q}_k^{\theta(\ell),\theta_0,\beta_0}}-1\right)\Big|X_{t_k}\right]\right|d\ell,$$

$$D_{1,2}:=\frac{u^2}{n\Delta_n}\sum_{k=0}^{n-1}\int_0^1\left(\frac{\partial_\theta b(\theta(\ell),X_{t_k})}{\sigma^2(\beta_0,X_{t_k})}\right)^2$$
$$\times\left|\mathrm{E}_{\widehat{Q}_k^{\theta(\ell),\theta_0,\beta_0}}\left[\left(\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell),\beta_0}\left[\int_{t_k}^{t_{k+1}}\int_I zM(ds,dz)\Big|Y_{t_{k+1}}^{\theta(\ell),\beta_0}=X_{t_{k+1}}\right]\right)^2\left(\frac{d\widehat{\mathrm{P}}}{d\widehat{Q}_k^{\theta(\ell),\theta_0,\beta_0}}-1\right)\Big|X_{t_k}\right]\right|d\ell.$$

Using Lemma 5.2.12 and hypotheses **(A2)-(A4)**, we get that for some constants $C,q>0$,

$$D_{1,1}\leq\frac{C}{\sqrt{n\Delta_n}}\frac{u^2}{n}\sum_{k=0}^{n-1}\left(1+|X_{t_k}|^q\right).$$

Applying Lemma 4.2.5, Jensen's inequality and **(A2)**, **(A4)**, we obtain that

$$D_{3,1,2}\leq\frac{Cu^2}{n\Delta_n\sqrt{n}}\sum_{k=0}^{n-1}\left(1+|X_{t_k}|^q\right)$$
$$\times\int_0^1\left(\widehat{\mathrm{E}}^\alpha\left[\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell),\beta_0}\left[\left(\int_{t_k}^{t_{k+1}}\int_I zM(ds,dz)\right)^4\Big|Y_{t_{k+1}}^{\theta(\ell),\beta_0}=X_{t_{k+1}}\right]\Big|X_{t_k}\right]\right)^{1/2}d\ell$$
$$\leq\frac{C}{\sqrt{n\Delta_n}}\frac{u^2}{n}\sum_{k=0}^{n-1}\left(1+|X_{t_k}|^q\right),$$

for some constants $C,q>0$.

Next, hypotheses **(A2)** and **(A3)**(b) yield that

$$D_{3,2} \leq \frac{Cu^2}{n\Delta_n}\sum_{k=0}^{n-1}(1+|X_{t_k}|^q)\int_0^1 \mathrm{E}_{\widehat{Q}_k^{\theta(\ell),\theta_0,\beta_0}}\Bigg[\Bigg(\int_{t_k}^{t_{k+1}}\int_I zN(ds,dz) - \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell),\beta_0}\Big[\int_{t_k}^{t_{k+1}}\int_I zM(ds,dz)\Big|Y_{t_{k+1}}^{\theta(\ell),\beta_0}=X_{t_{k+1}}\Big]\Bigg)^2\Big|X_{t_k}\Bigg]d\ell,$$

for some constants $C,q>0$.

Multiplying the random variable inside the expectation by $\left(\mathbf{1}_{\widehat{J}_{0,k}}+\mathbf{1}_{\widehat{J}_{1,k}}+\mathbf{1}_{\widehat{J}_{2,k}}\right)$ and applying Lemma 5.2.14, we get that for any $\alpha\in(\upsilon,\frac{1}{2})$ and $\alpha_0\in(\frac{1}{4},\frac{1}{2})$,

$$\begin{aligned} D_{3,2} &\leq \frac{Cu^2}{n\Delta_n}\sum_{k=0}^{n-1}(1+|X_{t_k}|^q)\int_0^1\left(M_0^{\theta(\ell),\beta_0}+M_1^{\theta(\ell),\beta_0}+M_2^{\theta(\ell),\beta_0}\right)d\ell \\ &\leq C\left(\lambda_n\sqrt{\Delta_n}+\Delta_n^{-\frac{3}{2}-3\gamma}e^{-C_0\Delta_n^{2(\alpha\vee\alpha_0)-1}}\right)\frac{u^2}{n}\sum_{k=0}^{n-1}(1+|X_{t_k}|^q), \end{aligned}$$

for some constants $C_0,C>0$. Thus, D_2 converges to zero in $\mathrm{P}^{\theta_0,\beta_0}$-probability as $n\to\infty$. The desired proof is now finished. □

Lemma 5.3.6. *Under conditions* **(A1)**-**(A4)**, *and* **(A6)**,**(A8)**, *as* $n\to\infty$,

$$\frac{\upsilon}{\sqrt{n\Delta_n^2}}\sum_{k=0}^{n-1}\int_0^1\left(Q_{k,n}^{1,\ell}+Q_{k,n}^{2,\ell}\right)d\ell \overset{\mathrm{P}^{\theta_0,\beta_0}}{\longrightarrow} 0.$$

Proof. It suffices to show that (1.16) holds for each sequence $(Q_{k,n}^{i,\ell})_{k\geq 1}$ under the measure $\mathrm{P}^{\theta_0,\beta_0}$. First, as for the term $Z_{k,n}^{1,\ell}$, applying Hölder's and Jensen's inequalities repeatedly, Lemmas 5.2.10, 5.2.6 and (5.11), together with **(A8)**, we obtain that for any $p,q>1$ conjugate, q_1,q_2 conjugate, with q_2 close to 1, there exist constants $C_0, q_0 > 0$ such that

$$
\begin{aligned}
&\frac{|v|}{\sqrt{n\Delta_n^2}} \sum_{k=0}^{n-1} \mathrm{E}\left[\left|\int_0^1 Q_{k,n}^{1,\ell} d\ell\right| \Big| \widehat{\mathcal{F}}_{t_k}\right] \leq \frac{|v|}{\sqrt{n\Delta_n^2}} \sum_{k=0}^{n-1} \int_0^1 \left(\mathrm{P}\left(\widehat{\tau} \leq n\Delta_n | X_{t_k}\right)\right)^{\frac{1}{q}} \\
&\quad \times \left(\mathrm{E}\left[\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[\left|\delta\left(U_2^{\theta_n,\beta(\ell)}(t_k, X_{t_k})\right)\right|^p \Big| Y_{t_{k+1}}^{\theta_n,\beta(\ell)} = X_{t_{k+1}}\right] \big| X_{t_k}\right]\right)^{\frac{1}{p}} d\ell \\
&= \frac{|v|}{\sqrt{n\Delta_n^2}} \sum_{k=0}^{n-1} \int_0^1 \left(\mathrm{P}\left(\widehat{\tau} \leq n\Delta_n | X_{t_k}\right)\right)^{\frac{1}{q}} \\
&\quad \times \Bigg\{\mathrm{E}\left[\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[\left|\delta\left(U_2^{\theta_n,\beta(\ell)}(t_k, X_{t_k})\right)\right|^p \Big| Y_{t_{k+1}}^{\theta_n,\beta(\ell)} = X_{t_{k+1}}^{\theta_n,\beta(\ell)}\right] \mathbf{1}_{\widehat{J}_{0,k}} \frac{q_{(0)}^{\theta,\beta}}{q_{(0)}^{\theta_n,\beta(\ell)}} \Big| X_{t_k}^{\theta_n,\beta(\ell)} = X_{t_k}\right] \\
&\qquad + \sum_{j=1}^{\infty} \sum_{(a_1,\ldots,a_j)\in A} \mathrm{E}\Bigg[\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[\left|\delta\left(U_2^{\theta_n,\beta(\ell)}(t_k, X_{t_k})\right)\right|^p \Big| Y_{t_{k+1}}^{\theta_n,\beta(\ell)} = X_{t_{k+1}}^{\theta_n,\beta(\ell)}\right] \\
&\qquad \times \mathbf{1}_{\{\widehat{J}_{j,k},a_1,\ldots,a_j\}} \frac{q_{(j)}^{\theta,\beta}}{q_{(j)}^{\theta_n,\beta(\ell)}} \Big| X_{t_k}^{\theta_n,\beta(\ell)} = X_{t_k}\Bigg]\Bigg\}^{\frac{1}{p}} d\ell \\
&\leq \frac{|v|}{\sqrt{n\Delta_n^2}} \sum_{k=0}^{n-1} \int_0^1 \left(\mathrm{P}\left(\widehat{\tau} \leq n\Delta_n | X_{t_k}\right)\right)^{\frac{1}{q}} \Bigg\{\left(\mathrm{E}\left[\left|\delta\left(U_2^{\theta_n,\beta(\ell)}(t_k, X_{t_k})\right)\right|^{pq_1} \Big| Y_{t_k}^{\theta_n,\beta(\ell)} = X_{t_k}\right]\right)^{\frac{1}{q_1}} \\
&\quad \times \left(\mathrm{E}\left[\mathbf{1}_{\widehat{J}_{0,k}} \left(\frac{q_{(0)}^{\theta,\beta}}{q_{(0)}^{\theta_n,\beta(\ell)}}\right)^{q_2} \Big| X_{t_k}^{\theta_n,\beta(\ell)} = X_{t_k}\right]\right)^{\frac{1}{q_2}} \\
&\quad + \sum_{j=1}^{\infty} \sum_{(a_1,..,a_j)\in A} \left(\mathrm{E}\left[\mathbf{1}_{\{\widehat{J}_{j,k},a_1,\ldots,a_j\}} \left(\frac{q_{(j)}^{\theta,\beta}}{q_{(j)}^{\theta_n,\beta(\ell)}}\right)^{q_2} \Big| X_{t_k}^{\theta_n,\beta(\ell)} = X_{t_k}\right]\right)^{\frac{1}{q_2}} \\
&\quad \times \left(\mathrm{E}\left[\left|\delta\left(U_2^{\theta_n,\beta(\ell)}(t_k, X_{t_k})\right)\right|^{pq_1} \Big| Y_{t_k}^{\theta_n,\beta(\ell)} = X_{t_k}\right]\right)^{\frac{1}{q_1}} \Bigg\}^{\frac{1}{p}} d\ell \\
&\leq C_0 \frac{|v|}{\sqrt{n}} \sum_{k=0}^{n-1} \left(1 + |X_{t_k}|^{q_0}\right) \left(\mathrm{P}\left(\widehat{\tau} \leq n\Delta_n | X_{t_k}\right)\right)^{\frac{1}{q}} \\
&\quad \times \left(1 + \sum_{j=1}^{\infty} \sum_{(a_1,\ldots,a_j)\in A} \left(C^{(2q_2-1)(j+1)} p_{a_1}\cdots p_{a_j} e^{-\lambda\Delta_n} \frac{(\lambda\Delta_n)^j}{j!}\right)^{\frac{1}{q_2}}\right)^{\frac{1}{p}} \\
&\leq C_0 \left(n\Delta_n \left(\int_{\{|z|\geq \rho_2 \Delta_n^{-\gamma}\}} \nu(dz) + \int_{\{|z|\leq \rho_1 \Delta_n^{v}\}} \nu(dz)\right)\right)^{\frac{1}{q}} \frac{|v|}{\sqrt{n}} \sum_{k=0}^{n-1} \left(1 + |X_{t_k}|^{q_0}\right),
\end{aligned}
$$

where we have used (5.37), and the inequality $1-e^{-x} \leq x$ valid for all $x \geq 0$. Then by hypothesis **(A6)**, we conclude that (1.16) holds true, and by Lemma 1.4.2, as $n \to \infty$,

$$
\frac{v}{\sqrt{n\Delta_n^2}} \sum_{k=0}^{n-1} \int_0^1 Q_{k,n}^{1,\ell} d\ell \overset{\mathrm{P}^{\theta_0,\beta_0}}{\longrightarrow} 0.
$$

The term $Q_{k,n}^{2,\ell}$ is treated similarly. Thus, the result follows. □

Lemma 5.3.7. *Under conditions* **(A1)**-**(A4)** *and* **(A6)**, *as* $n \to \infty$,

$$
\sum_{k=0}^{n-1} \frac{v}{\sqrt{n\Delta_n^2}} \int_0^1 \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[H^{\theta_n,\beta(\ell)} \mathbf{1}_{\{\widehat{\tau}>n\Delta_n\}} \Big| Y_{t_{k+1}}^{\theta_n,\beta(\ell)} = X_{t_{k+1}}\right] \mathbf{1}_{\{\widehat{\tau}>n\Delta_n\}} d\ell \overset{\mathrm{P}^{\theta_0,\beta_0}}{\longrightarrow} 0.
$$

Proof. Using the fact that $\mathbf{1}_{\{\widehat{\tau}>n\Delta_n\}}=1-\mathbf{1}_{\{\widehat{\tau}\leq n\Delta_n\}}$ and $\mathbf{1}_{\{\widetilde{\tau}>n\Delta_n\}}=1-\mathbf{1}_{\{\widetilde{\tau}\leq n\Delta_n\}}$, it suffices to show that as $n\to\infty$,

$$\sum_{k=0}^{n-1}\frac{v}{\sqrt{n\Delta_n^2}}\int_0^1\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[H^{\theta_n,\beta(\ell)}\mathbf{1}_{\{\widetilde{\tau}>n\Delta_n\}}\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}\right]\mathbf{1}_{\{\widehat{\tau}\leq n\Delta_n\}}d\ell\overset{\mathrm{P}^{\theta_0,\beta_0}}{\longrightarrow}0,\tag{5.41}$$

$$\sum_{k=0}^{n-1}\frac{v}{\sqrt{n\Delta_n^2}}\int_0^1\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[H^{\theta_n,\beta(\ell)}\mathbf{1}_{\{\widetilde{\tau}\leq n\Delta_n\}}\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}\right]d\ell\overset{\mathrm{P}^{\theta_0,\beta_0}}{\longrightarrow}0,\tag{5.42}$$

$$\sum_{k=0}^{n-1}\frac{v}{\sqrt{n\Delta_n^2}}\int_0^1\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[H^{\theta_n,\beta(\ell)}\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}\right]d\ell\overset{\mathrm{P}^{\theta_0,\beta_0}}{\longrightarrow}0.\tag{5.43}$$

The covergences (5.41) and (5.42) are treated similarly as for the terms $Q_{k,n}^{1,\ell}$ and $Q_{k,n}^{2,\ell}$. To treat (5.43), it suffices to show that conditions (i) and (ii) of Lemma 1.4.1 hold under the measure $\mathrm{P}^{\theta_0,\beta_0}$. We start showing (i). Using Lemma 5.2.10, (5.9), Lemma 5.2.9, Jensen's inequality, and (5.10), we get that

$$\begin{aligned}&\left|\sum_{k=0}^{n-1}\frac{v}{\sqrt{n\Delta_n^2}}\int_0^1\mathrm{E}\left[\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[H^{\theta_n,\beta(\ell)}\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}\right]\Big|\widehat{\mathcal{F}}_{t_k}\right]d\ell\right|\\
&=\Bigg|\sum_{k=0}^{n-1}\frac{v}{\sqrt{n\Delta_n^2}}\int_0^1\Bigg(\mathrm{E}\Bigg[\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[H^{\theta_n,\beta(\ell)}\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}^{\theta_n,\beta(\ell)}\right]\mathbf{1}_{\widehat{J}_{0,k}}\left(\frac{q_{(0)}^{\theta_0,\beta_0}}{q_{(0)}^{\theta_n,\beta(\ell)}}-1\right)\Big|X_{t_k}^{\theta_n,\beta(\ell)}=X_{t_k}\Bigg]\\
&\quad+\sum_{j=1}^{\infty}\sum_{(a_1,\ldots,a_j)\in I}\mathrm{E}\Bigg[\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[H^{\theta_n,\beta(\ell)}\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}^{\theta_n,\beta(\ell)}\right]\\
&\quad\times\mathbf{1}_{\{\widehat{J}_{j,k,a_1,\ldots,a_j}\}}\left(\frac{q_{(j)}^{\theta_0,\beta_0}}{q_{(j)}^{\theta_n,\beta(\ell)}}-1\right)\Big|X_{t_k}^{\theta_n,\beta(\ell)}=X_{t_k}\Bigg]\Bigg)d\ell\Bigg|\\
&\leq\Delta_n^{\frac{1}{2q_1}}\left(1+\sum_{j=1}^{\infty}\sum_{(a_1,\ldots,a_j)\in A}\left(C^{(p_1\vee q_3)(j+1)}p_{a_1}\cdots p_{a_j}e^{-\lambda_n\Delta_n}\frac{(\lambda_n\Delta_n)^j}{j!}\right)^{\frac{1}{p_1q_2}+\frac{1}{q_3}}\right)\\
&\quad\times\frac{C_0}{n}\sum_{k=0}^{n-1}\left(1+|X_{t_k}|^q\right).\end{aligned}$$

By **(A8)**, this converges to zero in $\mathrm{P}^{\theta_0,\beta_0}$-probability as $n\to\infty$, which shows Lemma 1.4.1 (i).

Next, applying Jensen's and Hölder's inequalities with q_1,q_2 conjugate, q_2 close to 1, together with Lemmas 5.2.10, 5.2.6, and (5.10), we get that

$$\sum_{k=0}^{n-1} \frac{v^2}{n\Delta_n^2} \mathrm{E}\left[\left(\int_0^1 \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[H^{\theta_n,\beta(\ell)}\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}\right]d\ell\right)^2\Big|\widehat{\mathcal{F}}_{t_k}\right]$$
$$\leq \sum_{k=0}^{n-1} \frac{v^2}{n\Delta_n^2}\int_0^1\Bigg(\mathrm{E}\left[\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[H^2\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}^{\theta_n,\beta(\ell)}\right]\mathbf{1}_{\widehat{J}_{0,k}}\frac{q_{(0)}^{\theta_0,\beta_0}}{q_{(0)}^{\theta_n,\beta(\ell)}}\Big|X_{t_k}^{\theta_n,\beta(\ell)}=X_{t_k}\right]$$
$$+\sum_{j=1}^{\infty}\sum_{(a_1,\dots,a_j)\in I}\mathrm{E}\left[\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[H^2\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}^{\theta_n,\beta(\ell)}\right]\mathbf{1}_{\{\widehat{J}_{j,k},a_1,\dots,a_j\}}\frac{q_{(j)}^{\theta_0,\beta_0}}{q_{(j)}^{\theta_n,\beta(\ell)}}\Big|X_{t_k}^{\theta_n,\beta(\ell)}=X_{t_k}\right]\Bigg)d\ell$$
$$\leq \sum_{k=0}^{n-1}\frac{v^2}{n\Delta_n^2}\int_0^1\Bigg\{\left(\mathrm{E}\left[|H|^{2q_1}\Big|Y_{t_k}^{\theta_n,\beta(\ell)}=X_{t_k}\right]\right)^{\frac{1}{q_1}}\left(\mathrm{E}\left[\mathbf{1}_{\widehat{J}_{0,k}}\left(\frac{q_{(0)}^{\theta_0,\beta_0}}{q_{(0)}^{\theta_n,\beta(\ell)}}\right)^{q_2}\Big|X_{t_k}^{\theta_n,\beta(\ell)}=X_{t_k}\right]\right)^{\frac{1}{q_2}}$$
$$+\sum_{j=1}^{\infty}\sum_{(a_1,\dots,a_j)\in I}\left(\mathrm{E}\left[|H|^{2q_1}\Big|Y_{t_k}^{\theta_n,\beta(\ell)}=X_{t_k}\right]\right)^{\frac{1}{q_1}}\left(\mathrm{E}\left[\mathbf{1}_{\{\widehat{J}_{j,k},a_1,\dots,a_j\}}\left(\frac{q_{(j)}^{\theta_0,\beta_0}}{q_{(j)}^{\theta_n,\beta(\ell)}}\right)^{q_2}\Big|X_{t_k}^{\theta_n,\beta(\ell)}=X_{t_k}\right]\right)^{\frac{1}{q_2}}\Bigg\}d\ell$$
$$\leq \Delta_n^{\frac{1}{2q_1}}\frac{C_0}{n}\sum_{k=0}^{n-1}(1+|X_{t_k}|^q)$$
$$\times\left(1+\sum_{j=1}^{\infty}\sum_{(a_1,\dots,a_j)\in I}\left(C^{(2q_2-1)(j+1)}p_{a_1}\cdots p_{a_j}e^{-\lambda_n\Delta_n}\frac{(\lambda_n\Delta_n)^j}{j!}\right)^{\frac{1}{q_2}}\right).$$

By hypothesis **(A8)**, this converges to zero in $\mathrm{P}^{\theta_0,\beta_0}$-probability as $n\to\infty$. Thus, the result follows. □

Lemma 5.3.8. *Under conditions* **(A1)**-**(A4)** *and* **(A6)-(A8)**, *as* $n\to\infty$,

$$\sum_{k=0}^{n-1}\frac{v}{\sqrt{n\Delta_n^2}}\int_0^1\frac{\partial_\beta\sigma}{\sigma^3}(\beta(\ell),X_{t_k})\Bigg\{\Bigg((H_{11}+H_{12}+H_{13})^2+2\sigma(\beta_0,X_{t_k})\left(B_{t_{k+1}}-B_{t_k}\right)$$
$$\times(H_{11}+H_{12}+H_{13})\Bigg)\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[\mathbf{1}_{\{\widetilde{\tau}>n\Delta_n\}}\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}\right]\mathbf{1}_{\{\widehat{\tau}>n\Delta_n\}}$$
$$-\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\Bigg[\Bigg(\left(H_8^{\theta_n,\beta(\ell)}+H_9^{\theta_n,\beta(\ell)}+H_{10}^{\theta_n,\beta(\ell)}\right)^2+2\sigma(\beta(\ell),Y_{t_k}^{\theta_n,\beta(\ell)})\left(W_{t_{k+1}}-W_{t_k}\right)$$
$$\times\left(H_8^{\theta_n,\beta(\ell)}+H_9^{\theta_n,\beta(\ell)}+H_{10}^{\theta_n,\beta(\ell)}\right)\Bigg)\mathbf{1}_{\{\widetilde{\tau}>n\Delta_n\}}\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}\Bigg]\mathbf{1}_{\{\widehat{\tau}>n\Delta_n\}}\Bigg\}d\ell\xrightarrow{\mathrm{P}^{\theta_0,\beta_0}}0.$$

Proof. Using the fact that $\mathbf{1}_{\{\widetilde{\tau}>n\Delta_n\}}=1-\mathbf{1}_{\{\widetilde{\tau}\leq n\Delta_n\}}$ and $\mathbf{1}_{\{\widehat{\tau}>n\Delta_n\}}=1-\mathbf{1}_{\{\widehat{\tau}\leq n\Delta_n\}}$, it suffices to show that as $n\to\infty$,

$$\sum_{k=0}^{n-1}\frac{v}{\sqrt{n\Delta_n^2}}\int_0^1\frac{\partial_\beta\sigma}{\sigma^3}(\beta(\ell),X_{t_k})\Bigg\{(H_{11}+H_{12}+H_{13})^2+2\sigma(\beta_0,X_{t_k})\left(B_{t_{k+1}}-B_{t_k}\right)$$
$$\times(H_{11}+H_{12}+H_{13})-\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\Bigg[\left(H_8^{\theta_n,\beta(\ell)}+H_9^{\theta_n,\beta(\ell)}+H_{10}^{\theta_n,\beta(\ell)}\right)^2+2\sigma(\beta(\ell),Y_{t_k}^{\theta_n,\beta(\ell)})$$
$$\times\left(W_{t_{k+1}}-W_{t_k}\right)\left(H_8^{\theta_n,\beta(\ell)}+H_9^{\theta_n,\beta(\ell)}+H_{10}^{\theta_n,\beta(\ell)}\right)\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}\Bigg]\Bigg\}d\ell\xrightarrow{\mathrm{P}^{\theta_0,\beta_0}}0.$$

For this, we split the term inside the conditional expectation as

$$\left(H_8^{\theta_n,\beta(\ell)}+H_9^{\theta_n,\beta(\ell)}+H_{10}^{\theta_n,\beta(\ell)}\right)^2+2\sigma(\beta(\ell),Y_{t_k}^{\theta_n,\beta(\ell)})\left(W_{t_{k+1}}-W_{t_k}\right)\\ \times\left(H_8^{\theta_n,\beta(\ell)}+H_9^{\theta_n,\beta(\ell)}+H_{10}^{\theta_n,\beta(\ell)}\right)=K^{\theta_n,\beta(\ell)}+F^{\theta_n,\beta(\ell)},$$

where

$$K^{\theta_n,\beta(\ell)}:=\left(\int_{t_k}^{t_{k+1}}b(\theta_n,Y_s^{\theta_n,\beta(\ell)})ds\right)^2+\left(\int_{t_k}^{t_{k+1}}\sigma(\beta(\ell),Y_s^{\theta_n,\beta(\ell)})dW_s\right)^2\\ -\left(\sigma(\beta(\ell),Y_{t_k}^{\theta_n,\beta(\ell)})(W_{t_{k+1}}-W_{t_k})\right)^2+2\int_{t_k}^{t_{k+1}}b(\theta_n,Y_s^{\theta_n,\beta(\ell)})ds\\ \times\left(\int_{t_k}^{t_{k+1}}\sigma(\beta(\ell),Y_s^{\theta_n,\beta(\ell)})dB_s+\int_{t_k}^{t_{k+1}}\int_I z\widetilde{M}(ds,dz)\right),$$

and

$$F^{\theta_n,\beta(\ell)}:=\left(\int_{t_k}^{t_{k+1}}\int_I z\widetilde{M}(ds,dz)\right)^2+2\int_{t_k}^{t_{k+1}}\int_I z\widetilde{M}(ds,dz)\int_{t_k}^{t_{k+1}}\sigma(\beta(\ell),Y_s^{\theta_n,\beta(\ell)})dW_s.$$

Similarly,

$$(H_{11}+H_{12}+H_{13})^2+2\sigma(\beta_0,X_{t_k})\left(B_{t_{k+1}}-B_{t_k}\right)(H_{11}+H_{12}+H_{13})=K^{\theta_0,\beta_0}+F^{\theta_0,\beta_0},$$

where

$$K^{\theta_0,\beta_0}:=\left(\int_{t_k}^{t_{k+1}}b(\theta_0,X_s^{\theta_0,\beta_0})ds\right)^2+\left(\int_{t_k}^{t_{k+1}}\sigma(\beta_0,X_s^{\theta_0,\beta_0})dB_s\right)^2-\left(\sigma(\beta_0,X_{t_k})(B_{t_{k+1}}-B_{t_k})\right)^2\\ +2\int_{t_k}^{t_{k+1}}b(\theta_0,X_s^{\theta_0,\beta_0})ds\left(\int_{t_k}^{t_{k+1}}\sigma(\beta_0,X_s^{\theta_0,\beta_0})dB_s+\int_{t_k}^{t_{k+1}}\int_I z\widetilde{N}(ds,dz)\right),$$

and

$$F^{\theta_0,\beta_0}:=\left(\int_{t_k}^{t_{k+1}}\int_I z\widetilde{N}(ds,dz)\right)^2+2\int_{t_k}^{t_{k+1}}\int_I z\widetilde{N}(ds,dz)\int_{t_k}^{t_{k+1}}\sigma(\beta_0,X_s^{\theta_0,\beta_0})dB_s.$$

Therefore, it suffices to show that as $n\to\infty$,

$$\sum_{k=0}^{n-1}\frac{v}{\sqrt{n\Delta_n^2}}\int_0^1\frac{\partial_\beta\sigma}{\sigma^3}(\beta(\ell),X_{t_k})\left(K^{\theta_0,\beta_0}-\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[K^{\theta_n,\beta(\ell)}\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}\right]\right)d\ell\overset{\mathrm{P}^{\theta_0,\beta_0}}{\longrightarrow}0,\tag{5.44}$$

$$\sum_{k=0}^{n-1}\frac{v}{\sqrt{n\Delta_n^2}}\int_0^1\frac{\partial_\beta\sigma}{\sigma^3}(\beta(\ell),X_{t_k})\left(F^{\theta_0,\beta_0}-\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[F^{\theta_n,\beta(\ell)}\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}\right]\right)d\ell\overset{\mathrm{P}^{\theta_0,\beta_0}}{\longrightarrow}0.\tag{5.45}$$

We first show (5.44), by proving that conditions (i) and (ii) of Lemma 1.4.1 hold under the measure $\mathrm{P}^{\theta_0,\beta_0}$. We start showing (i). Applying Lemma 5.2.10 to the conditional expectation, we get that

$$\mathrm{E}\left[K^{\theta_0,\beta_0}-\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[K^{\theta_n,\beta(\ell)}\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}\right]\big|\widehat{\mathcal{F}}_{t_k}\right]=S_1+S_2,$$

where

$$S_1 = -\mathrm{E}\left[\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[K^{\theta_n,\beta(\ell)}\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)} = X_{t_{k+1}}^{\theta_n,\beta(\ell)}\right]\mathbf{1}_{\widehat{J}_{0,k}}\left(\frac{q_{(0)}^{\theta_0,\beta_0}}{q_{(0)}^{\theta_n,\beta(\ell)}}-1\right)\Big|X_{t_k}^{\theta_n,\beta(\ell)} = X_{t_k}\right]$$
$$-\sum_{j=1}^{\infty}\sum_{(a_1,\ldots,a_j)\in I}\mathrm{E}\Bigg[\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[K^{\theta_n,\beta(\ell)}\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)} = X_{t_{k+1}}^{\theta_n,\beta(\ell)}\right]$$
$$\times \mathbf{1}_{\{\widehat{J}_{j,k},a_1,\ldots,a_j\}}\left(\frac{q_{(j)}^{\theta_0,\beta_0}}{q_{(j)}^{\theta_n,\beta(\ell)}}-1\right)\Big|X_{t_k}^{\theta_n,\beta(\ell)} = X_{t_k}\Bigg],$$
$$S_2 = \mathrm{E}\left[K^{\theta_0,\beta_0}\big|X_{t_k}\right] - \mathrm{E}\left[K^{\theta_n,\beta(\ell)}\big|\big|Y_{t_k}^{\theta_n,\beta(\ell)} = X_{t_k}\right].$$

Applying Lemma 5.2.9 and Jensen's inequality, we conclude that

$$\left|\sum_{k=0}^{n-1}\frac{v}{\sqrt{n\Delta_n^2}}\int_0^1\frac{\partial_\beta\sigma}{\sigma^3}\left(\beta(\ell),X_{t_k}\right)S_1 d\ell\right| \leq \Delta_n^{\frac{1}{2q_1}}\frac{C_0}{n}\sum_{k=0}^{n-1}(1+|X_{t_k}|^q)$$
$$\times\left(1+\sum_{j=1}^{\infty}\sum_{(a_1,\ldots,a_j)\in I}\left(C^{(p_1\vee q_3)(j+1)}p_{a_1}\cdots p_{a_j}e^{-\lambda_n\Delta_n}\frac{(\lambda_n\Delta_n)^j}{j!}\right)^{\frac{1}{p_1q_2}+\frac{1}{q_3}}\right).$$

By hypothesis **(A8)**, this converges to zero in $\mathrm{P}^{\theta_0,\beta_0}$-probability as $n\to\infty$.

On the other hand, observe that

$$\mathrm{E}_{X_{t_k}}\left[K^{\theta_0,\beta_0}\big|X_{t_k}\right] = 2\int_{t_k}^{t_{k+1}}\int_{t_k}^{s}\mathrm{E}\left[b(\theta_0,X_s^{\theta_0,\beta_0})b(\theta_0,X_u^{\theta_0,\beta_0})\big|X_{t_k}\right]duds$$
$$+\int_{t_k}^{t_{k+1}}\mathrm{E}\left[\sigma^2(\beta_0,X_s^{\theta_0,\beta_0})-\sigma^2(\beta_0,X_{t_k})\big|X_{t_k}\right]ds$$
$$+2\int_{t_k}^{t_{k+1}}\mathrm{E}\left[b(\theta_0,X_s^{\theta_0,\beta_0})\left(\int_{t_k}^{s}\sigma(\beta_0,X_u^{\theta_0,\beta_0})dB_u+\int_{t_k}^{s}\int_I z\widetilde{N}(du,dz)\right)\big|X_{t_k}\right]ds$$
$$=2\int_{t_k}^{t_{k+1}}\int_{t_k}^{s}\mathrm{E}\left[b(\theta_0,X_s^{\theta_0,\beta_0})b(\theta_0,X_u^{\theta_0,\beta_0})\big|X_{t_k}\right]duds$$
$$+\int_{t_k}^{t_{k+1}}\mathrm{E}\left[\sigma^2(\beta_0,X_s^{\theta_0,\beta_0})-\sigma^2(\beta_0,X_{t_k})\big|X_{t_k}\right]ds$$
$$+2\int_{t_k}^{t_{k+1}}\mathrm{E}\left[b(\theta_0,X_s^{\theta_0,\beta_0})\left(X_s^{\theta_0,\beta_0}-X_{t_k}-\int_{t_k}^{s}b(\theta_0,X_u^{\theta_0,\beta_0})du\right)\big|X_{t_k}\right]ds$$
$$=\int_{t_k}^{t_{k+1}}\mathrm{E}\left[\sigma^2(\beta_0,X_s^{\theta_0,\beta_0})-\sigma^2(\beta_0,X_{t_k})\big|X_{t_k}\right]ds+2\int_{t_k}^{t_{k+1}}\mathrm{E}\left[b(\theta_0,X_s^{\theta_0,\beta_0})\left(X_s-X_{t_k}\right)\big|X_{t_k}\right]ds,$$

which implies that $S_2 = S_{2,1}+2S_{2,2}$, where

$$S_{2,1} = \int_{t_k}^{t_{k+1}}\Bigg(\mathrm{E}\left[\sigma^2(\beta_0,X_s^{\theta_0,\beta_0})-\sigma^2(\beta_0,X_{t_k})\big|X_{t_k}\right]$$
$$-\mathrm{E}\left[\sigma^2(\beta(\ell),Y_s^{\theta_n,\beta(\ell)})-\sigma^2(\beta(\ell),Y_{t_k}^{\theta_n,\beta(\ell)})\big|Y_{t_k}^{\theta_n,\beta(\ell)} = X_{t_k}\right]\Bigg)ds,$$
$$S_{2,2} = \int_{t_k}^{t_{k+1}}\Bigg(\mathrm{E}\left[b(\theta_0,X_s^{\theta_0,\beta_0})\left(X_s^{\theta_0,\beta_0}-X_{t_k}\right)\big|X_{t_k}\right]$$
$$-\mathrm{E}\left[b(\theta_n,Y_s^{\theta_n,\beta(\ell)})\left(Y_s^{\theta_n,\beta(\ell)}-Y_{t_k}^{\theta_n,\beta(\ell)}\right)\big|Y_{t_k}^{\theta_n,\beta(\ell)} = X_{t_k}\right]\Bigg)ds.$$

By the mean value theorem, there exists $r \in (0,1)$ such that $S_{2,1} = S_{2,1,1} - S_{2,1,2}$, where

$$S_{2,1,1} = \int_{t_k}^{t_{k+1}} \Big(\mathrm{E}\Big[\sigma^2(\beta_0, X_s^{\theta_0,\beta_0}) - \sigma^2(\beta_0, X_{t_k}) \big| X_{t_k}\Big] - \mathrm{E}\Big[\sigma^2(\beta_0, Y_s^{\theta_n,\beta(\ell)}) - \sigma^2(\beta_0, Y_{t_k}^{\theta_n,\beta(\ell)}) \big| Y_{t_k}^{\theta_n,\beta(\ell)} = X_{t_k}\Big] \Big) ds,$$

$$S_{2,1,2} = 2\frac{\ell v}{\sqrt{n}} \int_{t_k}^{t_{k+1}} \mathrm{E}\Big[\sigma\partial_\beta\sigma(\beta_0 + \frac{\ell v r}{\sqrt{n}}, Y_s^{\theta_n,\beta(\ell)}) - \sigma\partial_\beta\sigma(\beta_0 + \frac{\ell v r}{\sqrt{n}}, Y_{t_k}^{\theta_n,\beta(\ell)}) \big| Y_{t_k}^{\theta_n,\beta(\ell)} = X_{t_k}\Big] ds.$$

Using hypotheses **(A1)**, **(A3)**(e) and Lemma 5.2.3, we get that for some constants $C, q > 0$,

$$\left| \sum_{k=0}^{n-1} \frac{v}{\sqrt{n\Delta_n^2}} \int_0^1 \frac{\partial_\beta \sigma}{\sigma^3}\left(\beta(\ell), X_{t_k}\right) S_{2,1,2} d\ell \right| \leq C\frac{\sqrt{\Delta_n}}{n} \sum_{k=0}^{n-1} \left(1 + |X_{t_k}|^q\right).$$

Next, using Lemma 5.2.10, we have that

$$\begin{aligned} S_{2,1,1} = \int_{t_k}^{t_{k+1}} \Bigg\{ &\mathrm{E}\Bigg[\Big(\sigma^2(\beta_0, X_s^{\theta_n,\beta(\ell)}) - \sigma^2(\beta_0, X_{t_k}^{\theta_n,\beta(\ell)})\Big) \mathbf{1}_{\widehat{J}_{0,k}} \\ &\times \left(\frac{q_{(0)}^{\theta_0,\beta_0}}{q_{(0)}^{\theta_n,\beta(\ell)}}(s - t_k, X_{t_k}^{\theta_n,\beta(\ell)}, X_s^{\theta_n,\beta(\ell)}) - 1 \right) \big| X_{t_k}^{\theta_n,\beta(\ell)} = X_{t_k} \Bigg] \\ &+ \sum_{j=1}^{\infty} \sum_{(a_1,\ldots,a_j)\in I} \mathrm{E}\Bigg[\Big(\sigma^2(\beta_0, X_s^{\theta_n,\beta(\ell)}) - \sigma^2(\beta_0, X_{t_k}^{\theta_n,\beta(\ell)})\Big) \\ &\times \mathbf{1}_{\{\widehat{J}_{j,k}, a_1,\ldots,a_j\}} \left(\frac{q_{(j)}^{\theta_0,\beta_0}}{q_{(j)}^{\theta_n,\beta(\ell)}}(s - t_k, X_{t_k}^{\theta_n,\beta(\ell)}, X_s^{\theta_n,\beta(\ell)}; a_1, \ldots, a_j) - 1 \right) \big| X_{t_k}^{\theta_n,\beta(\ell)} = X_{t_k} \Bigg] \Bigg\} ds, \end{aligned}$$

where in this case we denote $\{\widehat{J}_{j,k}, a_1, \ldots, a_j\} = \{N_s - N_{t_k} = j\} \cap \{\widehat{\Lambda}_{[t_k,s]} = \{a_1, \ldots, a_j\}\}$.

Applying Lemmas 5.2.9, 5.2.3 and **(A1)** to conclude that for some constants $C_0, q > 0$,

$$\begin{aligned} &\left| \sum_{k=0}^{n-1} \frac{v}{\sqrt{n\Delta_n^2}} \int_0^1 \frac{\partial_\beta \sigma}{\sigma^3}\left(\beta(\ell), X_{t_k}\right) W_{2,1,1} d\ell \right| \leq \Delta_n^{\frac{1}{q_1}} \frac{C_0}{n} \sum_{k=0}^{n-1} \left(1 + |X_{t_k}|^q\right) \\ &\times \left(1 + \sum_{j=1}^{\infty} \sum_{(a_1,\ldots,a_j)\in I} \left(C^{(p_1 \vee q_3)(j+1)} p_{a_1} \cdots p_{a_j} e^{-\lambda_n \Delta_n} \frac{(\lambda_n \Delta_n)^j}{j!} \right)^{\frac{1}{p_1 q_2} + \frac{1}{q_3}} \right), \end{aligned}$$

which, by hypothesis **(A8)**, converges to zero in $\mathrm{P}^{\theta_0,\beta_0}$-probability as $n \to \infty$.

We next treat $S_{2,2}$. By the mean value theorem, there exists $r \in (0,1)$ such that $S_{2,2} = S_{2,2,1} - S_{2,2,2}$, where

$$S_{2,2,1} = \int_{t_k}^{t_{k+1}} \Big(\mathrm{E}\Big[b(\theta_0, X_s^{\theta_0,\beta_0}) \left(X_s^{\theta_0,\beta_0} - X_{t_k}\right) \big| X_{t_k}\Big] - \mathrm{E}\Big[b(\theta_0, Y_s^{\theta_n,\beta(\ell)}) \left(Y_s^{\theta_n,\beta(\ell)} - Y_{t_k}^{\theta_n,\beta(\ell)}\right) \big| Y_{t_k}^{\theta_n,\beta(\ell)} = X_{t_k}\Big] \Big) ds,$$

$$S_{2,2,2} = \frac{u}{\sqrt{n\Delta_n}} \int_{t_k}^{t_{k+1}} \mathrm{E}\Big[\partial_\theta b(\theta(r), Y_s^{\theta_n,\beta(\ell)}) \left(Y_s^{\theta_n,\beta(\ell)} - Y_{t_k}^{\theta_n,\beta(\ell)}\right) \big| Y_{t_k}^{\theta_n,\beta(\ell)} = X_{t_k}\Big] ds,$$

and $\theta(r) = \theta_0 + \frac{ur}{\sqrt{n\Delta_n}}$. Proceeding as for the term $S_{2,1,1}$, we conclude that as $n \to \infty$,

$$\sum_{k=0}^{n-1} \frac{v}{\sqrt{n\Delta_n^2}} \int_0^1 \frac{\partial_\beta \sigma}{\sigma^3}\left(\beta(\ell), X_{t_k}\right) S_{2,2,1} d\ell \overset{\mathrm{P}^{\theta_0,\beta_0}}{\longrightarrow} 0.$$

We next add and substract the term $\partial_\theta b(\theta(r), Y_{t_k}^{\theta_n,\beta(\ell)})$ and use equation (5.3) to get that $S_{2,2,2} = S_{2,2,2,1} + S_{2,2,2,2}$, where

$$\begin{aligned} S_{2,2,2,1} &= \frac{u}{\sqrt{n\Delta_n}} \int_{t_k}^{t_{k+1}} \mathrm{E}\Big[\Big(\partial_\theta b(\theta(r), Y_s^{\theta_n,\beta(\ell)}) - \partial_\theta b(\theta(r), Y_{t_k}^{\theta_n,\beta(\ell)})\Big) \\ &\quad \times \Big(Y_s^{\theta_n,\beta(\ell)} - Y_{t_k}^{\theta_n,\beta(\ell)}\Big) \big| Y_{t_k}^{\theta_n,\beta(\ell)} = X_{t_k}\Big] ds, \\ S_{2,2,2,2} &= \frac{u}{\sqrt{n\Delta_n}} \partial_\theta b(\theta(r), X_{t_k}) \int_{t_k}^{t_{k+1}} \int_{t_k}^{s} \mathrm{E}\left[b(\theta_n, Y_u^{\theta_n,\beta(\ell)}) \big| Y_{t_k}^{\theta_n,\beta(\ell)} = X_{t_k}\right] duds. \end{aligned}$$

Therefore, using **(A3)**(e) and Lemma 5.2.3, we get that for some constants $C, q > 0$,

$$\left| \sum_{k=0}^{n-1} \frac{v}{\sqrt{n\Delta_n^2}} \int_0^1 \frac{\partial_\beta \sigma}{\sigma^3} \left(\beta(\ell), X_{t_k}\right) S_{2,2,2} d\ell \right| \leq C \frac{\sqrt{\Delta_n}}{n} \sum_{k=0}^{n-1} \left(1 + |X_{t_k}|^q\right),$$

which concludes that as $n \to \infty$,

$$\sum_{k=0}^{n-1} \frac{v}{\sqrt{n\Delta_n^2}} \int_0^1 \frac{\partial_\beta \sigma}{\sigma^3} \left(\beta(\ell), X_{t_k}\right) S_2 d\ell \overset{\mathrm{P}^{\theta_0,\beta_0}}{\longrightarrow} 0.$$

This finishes the proof of Lemma 1.4.1 (i).

Next, applying Jensen's and Hölder's inequalities with q_1, q_2 conjugate, with q_2 close to 1, together with Lemmas 5.2.10, 5.2.6, we get that

$$\sum_{k=0}^{n-1}\frac{v^2}{n\Delta_n^2}\mathrm{E}\left[\left(\int_0^1\frac{\partial_\beta\sigma}{\sigma^3}(\beta(\ell),X_{t_k})\left(K^{\theta_0,\beta_0}-\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[K^{\theta_n,\beta(\ell)}\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}\right]\right)d\ell\right)^2\Big|\widehat{\mathcal{F}}_{t_k}\right]$$
$$\leq\sum_{k=0}^{n-1}\frac{2v^2}{n\Delta_n^2}\int_0^1\left(\frac{\partial_\beta\sigma}{\sigma^3}\right)^2(\beta(\ell),X_{t_k})$$
$$\times\left(\mathrm{E}\left[\left(K^{\theta_0,\beta_0}\right)^2\Big|X_{t_k}\right]+\mathrm{E}\left[\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[\left(K^{\theta_n,\beta(\ell)}\right)^2\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}\right]\Big|X_{t_k}\right]\right)d\ell$$
$$=\sum_{k=0}^{n-1}\frac{2v^2}{n\Delta_n^2}\int_0^1\left(\frac{\partial_\beta\sigma}{\sigma^3}\right)^2(\beta(\ell),X_{t_k})\Bigg\{\mathrm{E}\left[\left(K^{\theta_0,\beta_0}\right)^2\Big|X_{t_k}\right]$$
$$+\mathrm{E}\left[\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[\left(K^{\theta_n,\beta(\ell)}\right)^2\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}^{\theta_n,\beta(\ell)}\right]\mathbf{1}_{\widehat{J}_{0,k}}\frac{q_{(0)}^{\theta_0,\beta_0}}{q_{(0)}^{\theta_n,\beta(\ell)}}\Big|X_{t_k}^{\theta_n,\beta(\ell)}=X_{t_k}\right]$$
$$+\sum_{j=1}^{\infty}\sum_{(a_1,\ldots,a_j)\in I}\mathrm{E}\Bigg[\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[\left(K^{\theta_n,\beta(\ell)}\right)^2\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}^{\theta_n,\beta(\ell)}\right]$$
$$\times\mathbf{1}_{\{\widehat{J}_{j,k},a_1,\ldots,a_j\}}\frac{q_{(j)}^{\theta_0,\beta_0}}{q_{(j)}^{\theta_n,\beta(\ell)}}\Big|X_{t_k}^{\theta_n,\beta(\ell)}=X_{t_k}\Bigg]d\ell\Bigg\}$$
$$\leq\sum_{k=0}^{n-1}\frac{2v^2}{n\Delta_n^2}\int_0^1\left(\frac{\partial_\beta\sigma}{\sigma^3}\right)^2(\beta(\ell),X_{t_k})\Bigg\{C\Delta_n^{\frac{5}{2}}\left(1+|X_{t_k}|^q\right)$$
$$+\left(\mathrm{E}\left[\left(K^{\theta_n,\beta(\ell)}\right)^{2q_1}\Big|Y_{t_k}^{\theta_n,\beta(\ell)}=X_{t_k}\right]\right)^{\frac{1}{q_1}}\left(\mathrm{E}\left[\mathbf{1}_{\widehat{J}_{0,k}}\left(\frac{q_{(0)}^{\theta_0,\beta_0}}{q_{(0)}^{\theta_n,\beta(\ell)}}\right)^{q_2}\Big|X_{t_k}^{\theta_n,\beta(\ell)}=X_{t_k}\right]\right)^{\frac{1}{q_2}}$$
$$+\sum_{j=1}^{\infty}\sum_{(a_1,\ldots,a_j)\in I}\left(\mathrm{E}\left[\left(K^{\theta_n,\beta(\ell)}\right)^{2q_1}\Big|Y_{t_k}^{\theta_n,\beta(\ell)}=X_{t_k}\right]\right)^{\frac{1}{q_1}}$$
$$\times\left(\mathrm{E}\left[\mathbf{1}_{\{\widehat{J}_{j,k},a_1,\ldots,a_j\}}\left(\frac{q_{(j)}^{\theta_0,\beta_0}}{q_{(j)}^{\theta_n,\beta(\ell)}}\right)^{q_2}\Big|X_{t_k}^{\theta_n,\beta(\ell)}=X_{t_k}\right]\right)^{\frac{1}{q_2}}\Bigg\}d\ell$$
$$\leq\Delta_n^{\frac{1}{2q_1}}\frac{C_0}{n}\sum_{k=0}^{n-1}\left(1+|X_{t_k}|^q\right)$$
$$\times\left(1+\sum_{j=1}^{\infty}\sum_{(a_1,\ldots,a_j)\in I}\left(C^{(2q_2-1)(j+1)}p_{a_1}\cdots p_{a_j}e^{-\lambda_n\Delta_n}\frac{(\lambda_n\Delta_n)^j}{j!}\right)^{\frac{1}{q_2}}\right),$$

By hypothesis **(A8)**, this converges to zero in $\mathrm{P}^{\theta_0,\beta_0}$-probability as $n\to\infty$. This finishes the proof of (5.44).

Now it remains to treat (5.45). Using equation (5.1), it suffices to show that as $n\to\infty$,

$$\sum_{k=0}^{n-1}\frac{v}{\sqrt{n\Delta_n^2}}\int_0^1\frac{\partial_\beta\sigma}{\sigma^3}(\beta(\ell),X_{t_k})\Bigg(\left(\int_{t_k}^{t_{k+1}}\int_I zN(ds,dz)\right)^2$$
$$-\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[\left(\int_{t_k}^{t_{k+1}}\int_I zM(ds,dz)\right)^2\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}\right]\Bigg)d\ell\overset{\mathrm{P}^{\theta_0,\beta_0}}{\longrightarrow}0,\tag{5.46}$$

and

$$\sum_{k=0}^{n-1}\frac{v}{\sqrt{n\Delta_n^2}}\int_0^1\frac{\partial_\beta\sigma}{\sigma^3}(\beta(\ell),X_{t_k})\left(X_{t_{k+1}}-X_{t_k}\right)\Bigg(\int_{t_k}^{t_{k+1}}\int_I zN(ds,dz) - \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[\int_{t_k}^{t_{k+1}}\int_I zM(ds,dz)\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}\right]\Bigg)d\ell\overset{\mathrm{P}^{\theta_0,\beta_0}}{\longrightarrow}0. \tag{5.47}$$

First, we treat (5.46) by showing that conditions (i) and (ii) of Lemma 1.4.1 hold under the measure $\mathrm{P}^{\theta_0,\beta_0}$. We start showing (i). Recall that the events $\widehat{A}_{k,r}$, $\widetilde{A}_{k,r}$ and $\widetilde{A}_{k,r}^c$ are introduced just before Lemma 5.2.15. Using $\mathbf{1}_{\widetilde{A}_{k,r}^c}=1-\mathbf{1}_{\widetilde{A}_{k,r}}$, we have that

$$\begin{aligned}
&\mathrm{E}\left[\left(\int_{t_k}^{t_{k+1}}\int_I zN(ds,dz)\right)^2-\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[\left(\int_{t_k}^{t_{k+1}}\int_I zM(ds,dz)\right)^2\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}\right]\Big|\widehat{\mathcal{F}}_{t_k}\right]\\
&=\sum_{r\in\overline{\mathcal{A}}}\mathrm{E}\Bigg[\mathbf{1}_{\widehat{A}_{k,r}}\Bigg(\left(\int_{t_k}^{t_{k+1}}\int_I zN(ds,dz)\right)^2\\
&\qquad-\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[(\mathbf{1}_{\widetilde{A}_{k,r}}+\mathbf{1}_{\widetilde{A}_{k,r}^c})\left(\int_{t_k}^{t_{k+1}}\int_I zM(ds,dz)\right)^2\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}\right]\Bigg)\Big|\widehat{\mathcal{F}}_{t_k}\Bigg]\\
&=M_{1,2}^{\theta_n,\beta(\ell)}-M_{2,2}^{\theta_n,\beta(\ell)},
\end{aligned}$$

which, together with Lemma 5.2.15 and hypotheses **(A2)**, **(A3)**(b), implies that for any $\alpha\in(\upsilon,\frac{1}{2})$,

$$\begin{aligned}
&\Bigg|\sum_{k=0}^{n-1}\frac{v}{\sqrt{n\Delta_n^2}}\int_0^1\frac{\partial_\beta\sigma}{\sigma^3}(\beta(\ell),X_{t_k})\,\mathrm{E}\Bigg[\left(\int_{t_k}^{t_{k+1}}\int_I zN(ds,dz)\right)^2\\
&\qquad-\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[\left(\int_{t_k}^{t_{k+1}}\int_I zM(ds,dz)\right)^2\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}\right]\Big|\widehat{\mathcal{F}}_{t_k}\Bigg]d\ell\Bigg|\\
&\le\sum_{k=0}^{n-1}\frac{|v|}{\sqrt{n\Delta_n^2}}\int_0^1\left|\frac{\partial_\beta\sigma}{\sigma^3}\right|(\beta(\ell),X_{t_k})\left(M_{1,2}^{\theta_n,\beta(\ell)}+M_{2,2}^{\theta_n,\beta(\ell)}\right)d\ell\\
&\le C\frac{\sqrt{n}}{\Delta_n}e^{-C_0\Delta_n^{2\alpha-1}}\frac{1}{n}\sum_{k=0}^{n-1}\left(1+|X_{t_k}|^q\right),
\end{aligned}$$

for some constants $C,C_0,q>0$. This shows Lemma 1.4.1 (i).

Next, Jensen's inequality gives

$$\begin{aligned}
&\mathrm{E}\left[\left(\left(\int_{t_k}^{t_{k+1}}\int_I zN(ds,dz)\right)^2-\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[\left(\int_{t_k}^{t_{k+1}}\int_I zM(ds,dz)\right)^2\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}\right]\right)^2\Big|\widehat{\mathcal{F}}_{t_k}\right]\\
&=\sum_{r\in\overline{\mathcal{A}}}\mathrm{E}\Bigg[\mathbf{1}_{\widehat{A}_{k,r}}\Bigg(\left(\int_{t_k}^{t_{k+1}}\int_I zN(ds,dz)\right)^2\\
&\qquad-\widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\left[(\mathbf{1}_{\widetilde{A}_{k,r}}+\mathbf{1}_{\widetilde{A}_{k,r}^c})\left(\int_{t_k}^{t_{k+1}}\int_I zM(ds,dz)\right)^2\Big|Y_{t_{k+1}}^{\theta_n,\beta(\ell)}=X_{t_{k+1}}\right]\Bigg)^2\Big|\widehat{\mathcal{F}}_{t_k}\Bigg]\\
&\le 2\left(M_{1,4}^{\theta_n,\beta(\ell)}+M_{2,4}^{\theta_n,\beta(\ell)}\right),
\end{aligned}$$

which, together with Lemma 5.2.15 and hypotheses **(A2)**, **(A3)**(b), implies that for any $\alpha\in$

$(v, \frac{1}{2})$,

$$\sum_{k=0}^{n-1} \frac{v^2}{n\Delta_n^2} \mathrm{E}\Big[\Big(\int_0^1 \frac{\partial_\beta \sigma}{\sigma^3}(\beta(\ell), X_{t_k}) \Big(\Big(\int_{t_k}^{t_{k+1}} \int_I z N(ds,dz)\Big)^2 - \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\Big[\Big(\int_{t_k}^{t_{k+1}} \int_I z M(ds,dz)\Big)^2 \Big| Y_{t_{k+1}}^{\theta_n,\beta(\ell)} = X_{t_{k+1}}\Big]\Big) d\ell\Big)^2 \Big| \widehat{\mathcal{F}}_{t_k}\Big]$$
$$\leq \sum_{k=0}^{n-1} \frac{2v^2}{n\Delta_n^2} \int_0^1 \Big(\frac{\partial_\beta \sigma}{\sigma^3}\Big)^2 (\beta(\ell), X_{t_k}) \Big(M_{1,4}^{\theta_n,\beta(\ell)} + M_{2,4}^{\theta_n,\beta(\ell)}\Big) d\ell$$
$$\leq \frac{C}{\Delta_n^2} e^{-C_0 \Delta_n^{2\alpha-1}} \frac{1}{n} \sum_{k=0}^{n-1} (1 + |X_{t_k}|^q),$$

for some constants $C, C_0, q > 0$. This finishes the proof of (5.46).

Finally, using Cauchy-Schwarz inequality and Lemma 5.2.3 (i), and proceeding as for (5.46), we conclude (5.47). Thus, the desired result follows. □

5.3.3 Main contributions : LAN property

Proof. Using the fact that $\mathbf{1}_{\{\widehat{\tau} > n\Delta_n\}} = 1 - \mathbf{1}_{\{\widehat{\tau} \leq n\Delta_n\}}$ and $\mathbf{1}_{\{\widetilde{\tau} > n\Delta_n\}} = 1 - \mathbf{1}_{\{\widetilde{\tau} \leq n\Delta_n\}}$, we write $\xi_{k,n} = \xi_{k,n,1} - \xi_{k,n,2} - \xi_{k,n,3}$, where

$$\xi_{k,n,1} = \frac{u}{\sqrt{n\Delta_n}} \int_0^1 \frac{\partial_\theta b(\theta(\ell), X_{t_k})}{\sigma^2(\beta_0, X_{t_k})} \big(\sigma(\beta_0, X_{t_k}) \big(B_{t_{k+1}} - B_{t_k}\big) + (b(\theta_0, X_{t_k}) - b(\theta(\ell), X_{t_k}))\Delta_n\big) d\ell,$$
$$\xi_{k,n,2} = \frac{u}{\sqrt{n\Delta_n}} \int_0^1 \frac{\partial_\theta b(\theta(\ell), X_{t_k})}{\sigma^2(\beta_0, X_{t_k})} \big(\sigma(\beta_0, X_{t_k}) \big(B_{t_{k+1}} - B_{t_k}\big) + (b(\theta_0, X_{t_k}) - b(\theta(\ell), X_{t_k}))\Delta_n\big)$$
$$\times \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell),\beta_0}\Big[\mathbf{1}_{\{\widetilde{\tau} > n\Delta_n\}} \Big| Y_{t_{k+1}}^{\theta(\ell),\beta_0} = X_{t_{k+1}}\Big] \mathbf{1}_{\{\widehat{\tau} \leq n\Delta_n\}} d\ell,$$
$$\xi_{k,n,3} = \frac{u}{\sqrt{n\Delta_n}} \int_0^1 \frac{\partial_\theta b(\theta(\ell), X_{t_k})}{\sigma^2(\beta_0, X_{t_k})} \big(\sigma(\beta_0, X_{t_k}) \big(B_{t_{k+1}} - B_{t_k}\big) + (b(\theta_0, X_{t_k}) - b(\theta(\ell), X_{t_k}))\Delta_n\big)$$
$$\times \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta(\ell),\beta_0}\Big[\mathbf{1}_{\{\widetilde{\tau} \leq n\Delta_n\}} \Big| Y_{t_{k+1}}^{\theta(\ell),\beta_0} = X_{t_{k+1}}\Big] d\ell.$$

Similarly, we write $\eta_{k,n} = \eta_{k,n,1} - \eta_{k,n,2} - \eta_{k,n,3}$, where

$$\eta_{k,n,1} = \frac{v}{\sqrt{n\Delta_n^2}} \int_0^1 \Big(\frac{\partial_\beta \sigma}{\sigma^3}(\beta(\ell), X_{t_k}) \sigma^2(\beta_0, X_{t_k}) \big(B_{t_{k+1}} - B_{t_k}\big)^2 - \frac{\partial_\beta \sigma}{\sigma}(\beta(\ell), X_{t_k}) \Delta_n\Big) d\ell,$$
$$\eta_{k,n,2} = \frac{v}{\sqrt{n\Delta_n^2}} \int_0^1 \Big(\frac{\partial_\beta \sigma}{\sigma^3}(\beta(\ell), X_{t_k}) \sigma^2(\beta_0, X_{t_k}) \big(B_{t_{k+1}} - B_{t_k}\big)^2 - \frac{\partial_\beta \sigma}{\sigma}(\beta(\ell), X_{t_k}) \Delta_n\Big)$$
$$\times \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\Big[\mathbf{1}_{\{\widetilde{\tau} > n\Delta_n\}} \Big| Y_{t_{k+1}}^{\theta_n,\beta(\ell)} = X_{t_{k+1}}\Big] \mathbf{1}_{\{\widehat{\tau} \leq n\Delta_n\}} d\ell,$$
$$\eta_{k,n,3} = \frac{v}{\sqrt{n\Delta_n^2}} \int_0^1 \Big(\frac{\partial_\beta \sigma}{\sigma^3}(\beta(\ell), X_{t_k}) \sigma^2(\beta_0, X_{t_k}) \big(B_{t_{k+1}} - B_{t_k}\big)^2 - \frac{\partial_\beta \sigma}{\sigma}(\beta(\ell), X_{t_k}) \Delta_n\Big)$$
$$\times \widetilde{\mathrm{E}}_{X_{t_k}}^{\theta_n,\beta(\ell)}\Big[\mathbf{1}_{\{\widetilde{\tau} \leq n\Delta_n\}} \Big| Y_{t_{k+1}}^{\theta_n,\beta(\ell)} = X_{t_{k+1}}\Big] d\ell.$$

Proceeding as for the terms $Z_{k,n}^{1,\ell}$ and $Z_{k,n}^{2,\ell}$, we get that as $n \to \infty$,

$$\sum_{k=0}^{n-1} (\xi_{k,n,2} + \xi_{k,n,3}) \overset{\mathrm{P}^{\theta_0,\beta_0}}{\longrightarrow} 0.$$

Proceeding as for the terms $Q_{k,n}^{1,\ell}$ and $Q_{k,n}^{2,\ell}$, we get that as $n \to \infty$,

$$\sum_{k=0}^{n-1} (\eta_{k,n,2} + \eta_{k,n,3}) \overset{\mathrm{P}^{\theta_0,\beta_0}}{\longrightarrow} 0.$$

Next, applying Lemma 1.4.3 to $\zeta_{k,n} = \xi_{k,n,1} + \eta_{k,n,1}$, it suffices to show that as $n \to \infty$,

$$\sum_{k=0}^{n-1} \mathrm{E}\left[\xi_{k,n,1} | \widehat{\mathcal{F}}_{t_k}\right] \overset{\mathrm{P}^{\theta_0,\beta_0}}{\longrightarrow} -\frac{u^2}{2}\Gamma_b(\theta_0,\beta_0), \tag{5.48}$$

$$\sum_{k=0}^{n-1} \mathrm{E}\left[\eta_{k,n,1} | \widehat{\mathcal{F}}_{t_k}\right] \overset{\mathrm{P}^{\theta_0,\beta_0}}{\longrightarrow} -\frac{v^2}{2}\Gamma_\sigma(\theta_0,\beta_0), \tag{5.49}$$

$$\sum_{k=0}^{n-1} \left(\mathrm{E}\left[\xi_{k,n,1}^2 | \widehat{\mathcal{F}}_{t_k}\right] - \left(\mathrm{E}\left[\xi_{k,n,1} | \widehat{\mathcal{F}}_{t_k}\right]\right)^2\right) \overset{\mathrm{P}^{\theta_0,\beta_0}}{\longrightarrow} u^2\Gamma_b(\theta_0,\beta_0), \tag{5.50}$$

$$\sum_{k=0}^{n-1} \left(\mathrm{E}\left[\eta_{k,n,1}^2 | \widehat{\mathcal{F}}_{t_k}\right] - \left(\mathrm{E}\left[\eta_{k,n,1} | \widehat{\mathcal{F}}_{t_k}\right]\right)^2\right) \overset{\mathrm{P}^{\theta_0,\beta_0}}{\longrightarrow} v^2\Gamma_\sigma(\theta_0,\beta_0), \tag{5.51}$$

$$\sum_{k=0}^{n-1} \left(\mathrm{E}\left[\xi_{k,n,1}\eta_{k,n,1} | \widehat{\mathcal{F}}_{t_k}\right] - \mathrm{E}\left[\xi_{k,n,1} | \widehat{\mathcal{F}}_{t_k}\right] \mathrm{E}\left[\eta_{k,n,1} | \widehat{\mathcal{F}}_{t_k}\right]\right) \overset{\mathrm{P}^{\theta_0,\beta_0}}{\longrightarrow} 0, \tag{5.52}$$

$$\sum_{k=0}^{n-1} \mathrm{E}\left[\xi_{k,n,1}^4 | \widehat{\mathcal{F}}_{t_k}\right] \overset{\mathrm{P}^{\theta_0,\beta_0}}{\longrightarrow} 0, \tag{5.53}$$

$$\sum_{k=0}^{n-1} \mathrm{E}\left[\eta_{k,n,1}^4 | \widehat{\mathcal{F}}_{t_k}\right] \overset{\mathrm{P}^{\theta_0,\beta_0}}{\longrightarrow} 0, \tag{5.54}$$

where

$$\Gamma_b(\theta_0,\beta_0) = \int_{\mathbb{R}} \left(\frac{\partial_\theta b(\theta_0,x)}{\sigma(\beta_0,x)}\right)^2 \pi_{\theta_0,\beta_0}(dx), \quad \text{and} \quad \Gamma_\sigma(\theta_0,\beta_0) = 2\int_{\mathbb{R}} \left(\frac{\partial_\beta \sigma(\beta_0,x)}{\sigma(\beta_0,x)}\right)^2 \pi_{\theta_0,\beta_0}(dx).$$

Proof of (5.48). Using $\mathrm{E}[B_{t_{k+1}} - B_{t_k} | \widehat{\mathcal{F}}_{t_k}] = 0$, and the mean value theorem, we get

$$\begin{aligned}
\sum_{k=0}^{n-1} \mathrm{E}\left[\xi_{k,n,1} | \widehat{\mathcal{F}}_{t_k}\right] &= -\frac{u^2}{n}\sum_{k=0}^{n-1}\int_0^1 \ell \frac{\partial_\theta b(\theta(\ell), X_{t_k})}{\sigma^2(\beta_0, X_{t_k})} \partial_\theta b(\theta_0 + \frac{\ell u r}{\sqrt{n\Delta_n}}, X_{t_k}) d\ell \\
&= -\frac{u^2}{2n}\sum_{k=0}^{n-1}\left(\frac{\partial_\theta b(\theta_0, X_{t_k})}{\sigma(\beta_0, X_{t_k})}\right)^2 - T_1 - T_2,
\end{aligned}$$

for some $r \in (0,1)$, $T_1 = \sum_{k=0}^{n-1} T_{k,n}$, and

$$\begin{aligned}
T_{k,n} &:= \frac{u^2}{n}\int_0^1 \ell \frac{\partial_\theta b(\theta(\ell), X_{t_k})}{\sigma^2(\beta_0, X_{t_k})} \left(\partial_\theta b(\theta_0 + \frac{\ell u r}{\sqrt{n\Delta_n}}, X_{t_k}) - \partial_\theta b(\theta_0, X_{t_k})\right) d\ell, \\
T_2 &:= \frac{u^2}{n}\sum_{k=0}^{n-1}\int_0^1 \ell \frac{\partial_\theta b(\theta_0, X_{t_k})}{\sigma^2(\beta_0, X_{t_k})} \left(\partial_\theta b(\theta(\ell), X_{t_k}) - \partial_\theta b(\theta_0, X_{t_k})\right) d\ell.
\end{aligned}$$

Using hypotheses **(A2)** and **(A3)**(b), (c), we have that for some constants $C, \epsilon, q > 0$,

$$\sum_{k=0}^{n-1} \mathrm{E}\left[|T_{k,n}| | \widehat{\mathcal{F}}_{t_k}\right] \leq \frac{C|u|^{\epsilon+2}|r|^\epsilon}{(\sqrt{n\Delta_n})^\epsilon} \frac{1}{n}\sum_{k=0}^{n-1}\left(1 + |X_{t_k}|^q\right),$$

which, by Lemma 1.4.2, implies that $T_1 \overset{\mathrm{P}^{\theta_0,\beta_0}}{\longrightarrow} 0$ as $n \to \infty$. Thus, so does T_2 by using the same argument. On the other hand, applying Lemma 5.2.16, we obtain that as $n \to \infty$,

$$\frac{1}{n}\sum_{k=0}^{n-1}\left(\frac{\partial_\theta b(\theta_0, X_{t_k})}{\sigma(\beta_0, X_{t_k})}\right)^2 \overset{\mathrm{P}^{\theta_0,\beta_0}}{\longrightarrow} \Gamma_b(\theta_0, \beta_0), \tag{5.55}$$

which gives (5.48).

Proof of (5.50). First, from the previous computations, we have that

$$\begin{aligned}\sum_{k=0}^{n-1}\left(\mathrm{E}\left[\xi_{k,n,1}|\widehat{\mathcal{F}}_{t_k}\right]\right)^2 &= \frac{u^4}{n^2}\sum_{k=0}^{n-1}\left(\int_0^1 \ell\frac{\partial_\theta b(\theta(\ell), X_{t_k})}{\sigma^2(\beta_0, X_{t_k})}\partial_\theta b(\theta_0 + \frac{\ell u r}{\sqrt{n\Delta_n}}, X_{t_k})d\ell\right)^2\\ &\le \frac{Cu^4}{n^2}\sum_{k=0}^{n-1}\left(1+|X_{t_k}|^q\right),\end{aligned}$$

for some constants $C, q > 0$.

Next, using properties of the moments of the Brownian motion, we can write

$$\sum_{k=0}^{n-1}\mathrm{E}\left[\xi_{k,n,1}^2|\widehat{\mathcal{F}}_{t_k}\right] = \frac{u^2}{n}\sum_{k=0}^{n-1}\left(\frac{\partial_\theta b(\theta_0, X_{t_k})}{\sigma(\beta_0, X_{t_k})}\right)^2 + T_3 + T_4 + T_5,$$

where

$$\begin{aligned}T_3 &:= \frac{2u^2}{n}\sum_{k=0}^{n-1}\frac{\partial_\theta b(\theta_0, X_{t_k})}{\sigma(\beta_0, X_{t_k})}\int_0^1\frac{\partial_\theta b(\theta(\ell), X_{t_k}) - \partial_\theta b(\theta_0, X_{t_k})}{\sigma(\beta_0, X_{t_k})}d\ell,\\ T_4 &:= \frac{u^2}{n}\sum_{k=0}^{n-1}\left(\int_0^1\frac{\partial_\theta b(\theta(\ell), X_{t_k}) - \partial_\theta b(\theta_0, X_{t_k})}{\sigma(\beta_0, X_{t_k})}d\ell\right)^2,\\ T_5 &:= \frac{u^2\Delta_n}{n}\sum_{k=0}^{n-1}\left(\int_0^1\frac{\partial_\theta b(\theta(\ell), X_{t_k})}{\sigma^2(\beta_0, X_{t_k})}\left(b(\theta_0, X_{t_k}) - b(\theta(\ell), X_{t_k})\right)d\ell\right)^2.\end{aligned}$$

As for the term T_1, using hypotheses **(A2)** and **(A3)**(b), (c), we get that T_3, T_4, T_5 converge to zero in $\mathrm{P}^{\theta_0,\beta_0}$-probability as $n \to \infty$. Moreover, using again (5.55), we conclude (5.50).

Proof of (5.53). Basic computation yields that

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

for some constants $C, q > 0$.

Proof of (5.49). Again, using properties of the moments of the Brownian motion and the mean value theorem, we have

$$\sum_{k=0}^{n-1}\mathrm{E}\left[\eta_{k,n,1}|\widehat{\mathcal{F}}_{t_k}\right] = -\frac{v^2}{2}\frac{2}{n}\sum_{k=0}^{n-1}\left(\frac{\partial_\beta\sigma(\beta_0, X_{t_k})}{\sigma(\beta_0, X_{t_k})}\right)^2 - T_6 - T_7 - T_8,$$

where, for some $r \in (0,1)$,

$$T_6 = \frac{2v^2}{n}\sum_{k=0}^{n-1}\int_0^1 \ell\frac{\partial_\beta\sigma}{\sigma^3}(\beta(\ell),X_{t_k})\,\sigma(\beta_0+\frac{\ell vr}{\sqrt{n}},X_{t_k})\left(\partial_\beta\sigma(\beta_0+\frac{\ell vr}{\sqrt{n}},X_{t_k})-\partial_\beta\sigma(\beta_0,X_{t_k})\right)d\ell,$$

$$T_7 = \frac{2v^2}{n}\sum_{k=0}^{n-1}\int_0^1 \ell\frac{\partial_\beta\sigma}{\sigma^3}(\beta(\ell),X_{t_k})\left(\sigma(\beta_0+\frac{\ell vr}{\sqrt{n}},X_{t_k})-\sigma(\beta_0,X_{t_k})\right)\partial_\beta\sigma(\beta_0,X_{t_k})d\ell,$$

$$T_8 = \frac{2v^2}{n}\sum_{k=0}^{n-1}\int_0^1 \ell\left(\frac{\partial_\beta\sigma}{\sigma^3}(\beta(\ell),X_{t_k})-\frac{\partial_\beta\sigma}{\sigma^3}(\beta_0,X_{t_k})\right)\sigma(\beta_0,X_{t_k})\partial_\beta\sigma(\beta_0,X_{t_k})d\ell.$$

As for the term T_1, using **(A2)** and **(A3)**(b), (d), together with Lemma 1.4.2, we conclude that T_6, T_7, T_8 converge to zero in $\mathrm{P}^{\theta_0,\beta_0}$-probability as $n\to\infty$. Again, applying Lemma 5.2.16, we obtain that as $n\to\infty$,

$$\frac{2}{n}\sum_{k=0}^{n-1}\left(\frac{\partial_\beta\sigma(\beta_0,X_{t_k})}{\sigma(\beta_0,X_{t_k})}\right)^2 \overset{\mathrm{P}^{\theta_0,\beta_0}}{\longrightarrow} \Gamma_\sigma(\theta_0,\beta_0), \tag{5.56}$$

which gives (5.49).

Proof of (5.51). First, from the previous computations, we have that

$$\begin{aligned}\sum_{k=0}^{n-1}\left(\mathrm{E}\left[\eta_{k,n,1}|\widehat{\mathcal{F}}_{t_k}\right]\right)^2 &= \frac{4v^4}{n^2}\sum_{k=0}^{n-1}\left(\int_0^1 \ell\frac{\partial_\beta\sigma}{\sigma^3}(\beta(\ell),X_{t_k})\,\sigma\partial_\beta\sigma(\beta_0+\frac{\ell vr}{\sqrt{n}},X_{t_k})d\ell\right)^2\\ &\leq \frac{Cv^4}{n^2}\sum_{k=0}^{n-1}\left(1+|X_{t_k}|^q\right),\end{aligned}$$

for some constants $C,q>0$.

Next, using the fact that $\mathrm{E}[(B_{t_{k+1}}-B_{t_k})^2|\widehat{\mathcal{F}}_{t_k}]=\Delta_n$ and $\mathrm{E}[(B_{t_{k+1}}-B_{t_k})^4|\widehat{\mathcal{F}}_{t_k}]=3\Delta_n^2$, we can write

$$\sum_{k=0}^{n-1}\mathrm{E}\left[\eta_{k,n,1}^2|\widehat{\mathcal{F}}_{t_k}\right] = \frac{2v^2}{n}\sum_{k=0}^{n-1}\left(\frac{\partial_\beta\sigma(\beta_0,X_{t_k})}{\sigma(\beta_0,X_{t_k})}\right)^2+\frac{v^2}{n}\sum_{k=0}^{n-1}S_{k,n},$$

where for some constants $C,q>0$,

$$\frac{v^2}{n}\sum_{k=0}^{n-1}\mathrm{E}\left[|S_{k,n}|\,|\widehat{\mathcal{F}}_{t_k}\right]\leq\frac{Cv^2}{n\sqrt{n}}\sum_{k=0}^{n-1}\left(1+|X_{t_k}|^q\right),$$

which, together with Lemma 1.4.2 and (5.56), concludes (5.51).

Proof of (5.52). Using properties of the moments of the Brownian motion, we get that

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

for some constants $C,q>0$.

Proof of (5.54). Basic computation yields that

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

for some constants $C,q>0$. The proof of Theorem 5.1.1 is now completed. □

Bibliographie

[1] Aït-Sahalia, Y. and Jacod, J., Volatility estimators for discretely sampled Lévy processes. *Ann. Statist.*, 35(1) :355–392, 2007.

[2] Aït-Sahalia, Y. and Jacod, J., Fisher's information for discretely sampled Lévy processes. *Econometrica*, 76(4) :727–761, 2008.

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

[4] Azencott, R., Densité des diffusions en temps petit : développements asymptotiques. In *Seminar on probability, XVIII*, volume 1059 of *Lecture Notes in Math.*, pages 402–498. Springer, Berlin, 1984.

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

[6] Basawa, I.V. and Prakasa Rao, B. L. S., *Statistical inference for stochastic processes.* Probability and Mathematical Statistics. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], London-New York, 1980.

[7] Basawa, I.V. and Scott, D. J., *Asymptotic Optimal Inference for Non-ergodic Models.* Lecture Notes in Statistics. Springer-Verlag New York, 1983.

[8] Bertoin, J., *Lévy processes.* Cambridge University Press, 1998.

[9] Bichteler, K., Gravereaux, J.B. and Jacod, J., *Malliavin calculus for processes with jumps*, volume 2 of *Stochastics Monographs.* Gordon and Breach Science Publishers, New York, 1987.

[10] Clément, E., Delattre, S. and Gloter, A., Asymptotic lower bounds in estimating jumps. *Bernoulli*, 20(3) :1059-1096, 2014.

[11] Clément, E. and Gloter, A., Local asymptotic mixed normality property for discretely observed stochastic differential equations driven by stable lévy processes. 2013. Preprint.

[12] Cont, R. and Tankov, P., *Financial modelling with jump processes.* Chapman & Hall/CRC, Boca Raton, FL, 2004.

[13] Corcuera, J.M. and Kohatsu-Higa, A., Statistical inference and Malliavin calculus. In *Seminar on Stochastic Analysis, Random Fields and Applications VI*, volume 63 of *Progr. Probab.*, pages 59–82. Birkhäuser/Springer Basel AG, Basel, 2011.

[14] Crimaldi, I. and Pratelli, L., Convergence results for multivariate martingales. *Stochastic Process. Appl.*, 115(4) :571–577, 2005.

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

[16] Dietz, H.M., A non-Markovian relative of the Ornstein-Uhlenbeck process and some of its local statistical properties. *Scand. J. Statist.*, 19(4) :363–379, 1992.

[17] Dohnal, G., On estimating the diffusion coefficient. *J. Appl. Probab.*, 24(1) :105–114, 1987.

[18] Feigin, P.D., Stable convergence of semimartingales. *Stochastic Process. Appl.*, 19(1) :125–134, 1985.

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

[20] Fournier, N. and Printems, J., Absolute continuity for some one-dimensional processes. *Bernoulli*, 16(2) :343–360, 2010.

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

[22] Genon-Catalot, V. and Jacod, J., Estimation of the diffusion coefficient for diffusion processes : random sampling. *Scand. J. Statist.*, 21(3) :193–221, 1994.

[23] Gloter, A. and Gobet, E., LAMN property for hidden processes : the case of integrated diffusions. *Ann. Inst. Henri Poincaré Probab. Stat.*, 44(1) :104–128, 2008.

[24] Gobet, E., Local asymptotic mixed normality property for elliptic diffusion : a Malliavin calculus approach. *Bernoulli*, 7(6) :899–912, 2001.

[25] Gobet, E., LAN property for ergodic diffusions with discrete observations. *Ann. Inst. H. Poincaré Probab. Statist.*, 38(5) :711–737, 2002.

[26] Hájek, J., A characterization of limiting distributions of regular estimates. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 14 :323–330, 1969/1970.

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

[28] Ibragimov, I.A, and Has′minskiĭ, R.Z., *Statistical estimation*, volume 16 of *Applications of Mathematics*. Springer-Verlag, New York-Berlin, 1981. Asymptotic theory, Translated from the Russian by Samuel Kotz.

[29] Jacod, J., Multivariate point processes : predictable projection, Radon-Nikodým derivatives, representation of martingales. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 31 :235–253, 1974/75.

[30] Jacod, J., *Inference for Stochastic Processes - Handbook of Financial Econometrics*, volume 2 : Applications, Edited by Yacine Ait-Sahalia. North-Holland, 2009.

[31] Jacod, J., Statistics and high-frequency data. In *Statistical methods for stochastic differential equations*, volume 124 of *Monogr. Statist. Appl. Probab.*, pages 191–310. CRC Press, Boca Raton, FL, 2012.

[32] Jacod, J. and Mémin, J., Caractéristiques locales et conditions de continuité absolue pour les semi-martingales. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 35(1) :1–37, 1976.

[33] Jacod, J. and Shiryaev, A.N., *Limit theorems for stochastic processes*, volume 288 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 2003.

[34] Jeganathan, P., On the asymptotic theory of estimation when the limit of the log-likelihood ratios is mixed normal. *Sankhyā Ser. A*, 44(2) :173–212, 1982.

[35] Jeganathan, P., Some asymptotic properties of risk functions when the limit of the experiment is mixed normal. *Sankhyā Ser. A*, 45(1) :66–87, 1983.

[36] Kabanov, Ju. M., Lipcer, R. Š. and Širjaev, A. N., Absolute continuity and singularity of locally absolutely continuous probability distributions. I. *Mat. Sb. (N.S.)*, 107(149) :364–415, 1978.

[37] Kabanov, Ju. M., Lipcer, R. Š. and Širjaev, A. N., Absolute continuity and singularity of locally absolutely continuous probability distributions. II. *Mat. Sb. (N.S.)*, 108(150) :32–61, 1979.

[38] Kawai, R., Local asymptotic normality property for Ornstein-Uhlenbeck processes with jumps under discrete sampling. *J. Theoret. Probab.*, 26(4) :932–967, 2013.

[39] Kawai, R. and Masuda, H., Local asymptotic normality for normal inverse Gaussian Lévy processes with high-frequency sampling. *ESAIM Probab. Stat.*, 17 :13–32, 2013.

[40] Kessler, M., Estimation of an ergodic diffusion from discrete observations. *Scand. J. Statist.*, 24(2) :211–229, 1997.

[41] Kohatsu-Higa, A., Nualart, E. and Tran, N.K., LAN property for a simple Lévy process. Accepted for publication in the *C. R. Math. Acad. Sci. Paris, Ser. I.* 2014.

[42] Küchler, U. and Sørensen, M., A note on limit theorems for multivariate martingales. *Bernoulli*, 5(3) :483–493, 1999.

[43] Kulik, A.M., Exponential ergodicity of the solutions to SDE's with a jump noise. *Stochastic Process. Appl.*, 119(2) :602–632, 2009.

[44] Kutoyants, Y. A., *Parameter estimation for stochastic processes*, volume 6 of *Research and Exposition in Mathematics.* Heldermann Verlag, Berlin, 1984. Translated from the Russian and edited by B. L. S. Prakasa Rao.

[45] Kutoyants, Y. A., *Identification of dynamical systems with small noise.* Mathematics and its Applications. Kluwer Academic Publishers Group, Dordrecht, 1994.

[46] Kutoyants, Y. A., *Statistical inference for spatial Poisson processes.* Lecture Notes in Statistics. Springer-Verlag, New York, 1998.

[47] Kutoyants, Y. A., *Statistical inference for ergodic diffusion processes.* Springer Series in Statistics. Springer-Verlag London, Ltd., London, 2004.

[48] Le Cam, L., Locally asymptotically normal families of distributions. Certain approximations to families of distributions and their use in the theory of estimation and testing hypotheses. *Univ. california Publ. Statist.*, 3 :37–98, 1960.

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

[50] León, J.A., Solé, J.L., Utzet, F. and Vives, J., On Lévy processes, Malliavin calculus and market models with jumps. *Finance Stoch.*, 6(2) :197–225, 2002.

[51] Letta, G. and Pratelli, L., Convergence stable vers un noyau gaussien. *Rend. Accad. Naz. Sci. XL Mem. Mat. Appl. (5)*, 20 :205–213, 1996.

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

[53] Masuda, H., Ergodicity and exponential β-mixing bounds for multidimensional diffusions with jumps. *Stochastic Process. Appl.*, 117(1) :35–56, 2007.

[54] Masuda, H., On stability of diffusions with compound-Poisson jumps. *Bull. Inform. Cybernet.*, 40 :61–74, 2008.

[55] Masuda, H., Convergence of Gaussian quasi-likelihood random fields for ergodic Lévy driven SDE observed at high frequency. *Ann. Statist.*, 41(3) :1593–1641, 2013.

[56] Meyn, S.P. and Tweedie, R.L., Stability of Markovian processes. III. Foster-Lyapunov criteria for continuous-time processes. *Adv. in Appl. Probab.*, 25(3) :518–548, 1993.

[57] Nualart, D., *The Malliavin calculus and related topics.* Probability and its Applications (New York). Springer-Verlag, Berlin, second edition, 2006.

[58] Ogihara, T., Local Asymptotic Mixed Normality Property for Nonsynchronously Observed Diffusion Processes. 2014. Preprint at arxiv.org/abs/1310.5304

[59] Ogihara, T. and Yoshida, N., Quasi-likelihood analysis for the stochastic differential equation with jumps. *Stat. Inference Stoch. Process.*, 14(3) :189–229, 2011.

[60] Ogihara, T. and Yoshida, N., Quasi-Likelihood Analysis for Stochastic Regression Models with Nonsynchronous Observations. 2012. Preprint at arxiv.org/abs/1212.4911

[61] Petrou, E., Malliavin Calculus in Lévy spaces and Applications to Finance. *Electron. J. Probab.*, 13 :852–879, 2008.

[62] Picard, J., On the existence of smooth densities for jump processes. *Probab. Theory Related Fields*, 105(4) :481–511, 1996.

[63] Prakasa Rao, B. L. S., *Semimartingales and their statistical inference*, volume 83 of *Monographs on Statistics and Applied Probability.* Chapman & Hall/CRC, Boca Raton, FL, 1999.

[64] Qiao, H., Exponential Ergodicity for SDEs with Jumps and Non-Lipschitz Coefficients. *J. Theoret. Probab.*, 27(1) :137–152, 2014.

[65] Sato, K., *Lévy processes and infinitely divisible distributions*, volume 68 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, Cambridge, 1999. Translated from the 1990 Japanese original, Revised by the author.

[66] Segall, A. and Kailath, T., Radon-Nikodým derivatives with respect to measures induced by discontinuous independent-increment processes. *Ann. Probability*, 3(3) :449–464, 1975.

[67] Shimizu, Y., M-estimation for discretely observed ergodic diffusion processes with infinitely many jumps. *Stat. Inference Stoch. Process.*, 9(2) :179–225, 2006.

[68] Shimizu, Y., Local asymptotic mixed normality for discretely observed non-recurrent Ornstein-Uhlenbeck processes. *Ann. Inst. Statist. Math.*, 64(1) :193–211, 2012.

[69] Shimizu, Y. and Yoshida, N., Estimation of parameters for diffusion processes with jumps from discrete observations. *Stat. Inference Stoch. Process.*, 9(3) :227–277, 2006.

[70] Solé, J.L., Utzet, F. and Vives, J., Canonical Lévy process and Malliavin calculus. In *Stochastic Process. Appl.*, 117(2) :165–187, 2007.

[71] Sørensen, M., Likelihood methods for diffusions with jumps. In *Statistical inference in stochastic processes*, volume 6 of *Probab. Pure Appl.*, pages 67–105. Dekker, New York, 1991.

[72] Uchida, M., Contrast-based information criterion for ergodic diffusion processes from discrete observations. *Ann. Inst. Statist. Math.*, 62(1) :161–187, 2010.

[73] van der Vaart, A.W., *Asymptotic statistics*, volume 3 of *Cambridge Series in Statistical and Probabilistic Mathematics.* Cambridge University Press, Cambridge, 1998.

[74] Willmot, G.E. and Lin, X.D., Lundberg bounds on the tails of compound distributions. *J. Appl. Probab.*, 31(3) :743–756, 1994.

[75] Woerner, J.H.C., Local asymptotic normality for the scale parameter of stable processes. *Statist. Probab. Lett.*, 63(1) :61–65, 2003.

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