Contrast function estimation for the drift parameter of ergodic jump diffusion process.

Chiara Amorino[1] and Arnaud Gloter[1]

[1]Laboratoire de Mathématiques et Modélisation d'Evry, Université Paris-Saclay

November 24, 2020

Abstract

In this paper we consider an ergodic diffusion process with jumps whose drift coefficient depends on an unknown parameter. We suppose that the process is discretely observed. We introduce an estimator based on a contrast function, which is efficient without requiring any conditions on the rate at which the step discretization goes to zero, and where we allow the observed process to have non summable jumps. This extends earlier results where the condition on the step discretization was needed and where the process was supposed to have summable jumps. In general situations, our contrast function is not explicit and one has to resort to some approximation. In the case of a finite jump activity, we propose explicit approximations of the contrast function, such that the efficient estimation of the drift parameter is feasible. This extends the results obtained by Kessler in the case of continuous processes.

Efficient drift estimation, ergodic properties, high frequency data, Lévy-driven SDE, thresholding methods.

1 Introduction

Diffusion processes with jumps have been widely used to describe the evolution of phenomenon arising in various fields. In finance, jump-processes were introduced to model the dynamic of asset prices (Merton, 1976),(Kou,2002), exchange rates (Bates, 1996), or volatility processes (Barndorff-Nielsen & Shephard, 2001),(Eraker, Johannes, & N, 2003). Utilization of jump-processes in neuroscience can be found for instance in (Ditlevsen & Greenwood, 2013).

Practical applications of these models has lead to the recent development of many statistical methods. In this work, our aim is to estimate the drift parameter θ from a discrete sampling of the process X^θ solution to

$$X_t^\theta = X_0^\theta + \int_0^t b(\theta, X_s^\theta)ds + \int_0^t \sigma(X_s^\theta)dW_s + \int_0^t \int_{\mathbb{R}\setminus\{0\}} \gamma(X_{s-}^\theta)z\tilde{\mu}(ds, dz),$$

where W is a one dimensional Brownian motion and $\tilde{\mu}$ a compensated Poisson random measure, with a possible infinite jump activity. We assume that the process is sampled at the times $(t_i^n)_{i=0,\dots,n}$ where the sampling step $\Delta_n := \sup_{i=0,\dots,n-1} t_{i+1}^n - t_i^n$ goes to zero. Due to the presence of a Gaussian component, we know that it is impossible to estimate the drift parameter on a finite horizon of time. Thus, we assume that $t_n \to \infty$ and the ergodicity of the process X^θ.

Generally, the main difficulty while considering statistical inference of discretely observed stochastic processes comes from the lack of explicit expression for the likelihood. Indeed, the transition density of a jump-diffusion process is usually unknown explicitly. Several methods have been developed to circumvent this difficulty. For instance, closed form expansions of the transition density of jump-diffusions is studied in (Aït-Sahalia & Yu, 2006), (Li & Chen, 2016). In the context of high frequency observation, the asymptotic behaviour of estimating functions are studied in (Jakobsen & Sørensen, 2017), and conditions are given to ensure rate optimality and efficiency. Another approach, fruitful in the case of high frequency observation, is to consider pseudo-likelihood method, for instance based on the high frequency approximation of the dynamic of the process by the one of the Euler scheme. This leads to explicit contrast functions with Gaussian structures (see e.g. (Shimizu & Yoshida, 2006),(Shimizu, 2006),(Masuda, 2013)).

The validity of the approximation by the Euler pseudo-likelihood is justified by the high frequency assumption of the observations, and actually proving that the estimators are asymptotic normal usually necessitates some conditions on the rate at which Δ_n should tend to zero. For applications, it is important that the condition on $\Delta_n \to 0$ is less stringent as possible.

In the case of continuous processes, Florens-Zmirou (Florens-Zmirou, 1989) proposes estimation of drift and diffusion parameters under the fast sampling assumption $n\Delta_n^2 \to 0$. Yoshida (Yoshida, 1992) suggests a correction of the contrast function that yields to the condition $n\Delta_n^3 \to 0$. In Kessler (Kessler, 1997), the author introduces an explicit modification of the Euler scheme contrast such that the associated estimators are asymptotically normal, under the condition $n\Delta_n^k \to 0$ where $k \geq 2$ is arbitrarily large. Hence, the result by Kessler allows for any arbitrarily slow polynomial decay to zero of the sampling step.

In the case of jump-diffusions, Shimizu (Shimizu, 2006) proposes parametric estimation of drift, diffusion and jump coefficients. The asymptotic normality of the estimators are obtained under some explicit conditions relating the sampling step and jump intensity of the process. These conditions on Δ_n are more restrictive as the intensity of jumps near zero is high. In the situation where this jump intensity is finite, the conditions of (Shimizu,2006) reduces to $n\Delta_n^2 \to 0$. In (Gloter, Loukianova, & Mai, 2018), the condition on the sampling step is relaxed to $n\Delta_n^3 \to 0$, when one estimates the drift parameter only.

In this paper, we focus on the estimation of the drift parameter, and our aim is to weaken the conditions on the decay of the sampling step in way comparable to Kessler's work (Kessler, 1997), but in the framework of jump-diffusion processes.

One of the idea in Kessler's paper is to replace, in the Euler scheme contrast function, the contribution of the drift by the exact value of the first conditional moment $m^{(1)}_{\theta,t_i,t_{i+1}}(x) = E[X^\theta_{t_{i+1}} \mid X^\theta_{t_i} = x]$ or some explicit approximation with arbitrarily high order when $\Delta_n \to 0$. In presence of jumps, the contrasts functions in (Shimizu & Yoshida, 2006) (see also (Shimizu, 2006), (Gloter, Loukianova, & Mai, 2018)) resort to a filtering procedure in order to suppress the contribution of jumps and recover the continuous part of the process. Based on those ideas, we introduce a contrast function (see Definition 1), whose expression relies on the quantity $m_{\theta,t_i,t_{i+1}}(x) = \frac{E[X^\theta_{t_{i+1}}\varphi((X^\theta_{t_{i+1}}-X^\theta_{t_i})/(t_{i+1}-t_i)^\beta)|X^\theta_{t_i}=x]}{E[\varphi((X^\theta_{t_{i+1}}-X^\theta_{t_i})/(t_{i+1}-t_i)^\beta)|X^\theta_{t_i}=x]}$, where φ is some compactly supported function and $\beta < 1/2$. The function φ is such that $\varphi((X^\theta_{t_{i+1}} - X^\theta_{t_i})/(t_{i+1} - t_i)^\beta)$ vanishes when the increments of the data are too large compared to the typical increments of a continuous diffusion process, and thus can be used to filter the contribution of the jumps.

The main result of our paper is that the associated estimator converges at rate $\sqrt{t_n}$, with some explicit asymptotic variance and is efficient. Comparing to earlier results ((Shimizu & Yoshida, 2006), (Shimizu, 2006), (Gloter, Loukianova, & Mai, 2018)), the sampling step $(t_i^n)_{i=0,\dots,n}$ can be irregular, no condition is needed on the rate at which $\Delta_n \to 0$ and we have suppressed the assumption that the jumps of the process are summable. Let us stress that when the jumps activity is so high that the jumps are not summable, we have to choose $\beta < 1/3$ (see Assumption A_β).

Moreover, in the case where the intensity is finite and with the specific choice of φ being an oscillating function, we prove that we can approximate our contrast function by a completely explicit one, exactly as in the paper by Kessler (Kessler, 1997). This yields to an efficient estimator under the condition $n\Delta_n^k \to 0$, where k is related to the oscillating properties of the function φ. As k can be chosen arbitrarily high, up to a proper choice of φ, our method allows to estimate efficiently the drift parameter, under the assumption that the sampling step tends to zero at some polynomial rate.
We also show numerically that, when the jump activity is finite, the estimator we deduce from the explicit approximation of the contrast function performs well, making the bias visibly reduced.
On the other side, considering the case of infinite jumps activity (taking in particular a tempered α-stable jump process with $\alpha < 1$), we implement our main results building an approximation of m (see Theorem 2 below) from which we deduce an approximation of the contrast that we minimize in order to get the estimator of the drift coefficient. The estimator we found is a corrected version of the estimator that would result from the choice of an Euler scheme approximation. We see numerically that our estimator is well-performed and that the correction term we give drastically reduces the bias, especially as α gets bigger.

The outline of the paper is the following. In Section 2 we present the assumptions on the process X. The Section 3 contains the main results of the paper: in Section 3.1 we define the contrast function while the consistency and asymptotic normality of the estimator are stated in Section 3.2. In Section 4 we explain how to use in practice the contrast function and so we deal with its approximations in Section 4.1 while its explicit modification is presented in the case of finite jump activity in Section 4.2. The Section 5 is devoted to numerical results and perspectives for practical applications. In Section 6 we state limit theorems useful to study the asymptotic behavior of the contrast function. The proofs of the main statistical results are given in Section 7, while the proofs of the limit theorems and some technical results are presented in the Appendix.

2 Model, assumptions

Let Θ be a compact subset of $\mathbb{R}$ and X^θ a solution to

$$X_t^\theta = X_0^\theta + \int_0^t b(\theta, X_s^\theta)ds + \int_0^t a(X_s^\theta)dW_s + \int_0^t \int_{\mathbb{R}\setminus\{0\}} \gamma(X_{s^-}^\theta)z\tilde{\mu}(ds, dz), \quad t \in \mathbb{R}_+, \tag{1}$$

where $W = (W_t)_{t\geq 0}$ is a one dimensional Brownian motion, μ is a Poisson random measure associated to the Lévy process $L = (L_t)_{t\geq 0}$, with $L_t := \int_0^t \int_{\mathbb{R}} z\tilde{\mu}(ds, dz)$ and $\tilde{\mu} = \mu - \bar{\mu}$ is the compensated one, on $[0, \infty) \times \mathbb{R}$. We denote $(\Omega, \mathcal{F}, \mathbb{P})$ the probability space on which W and μ are defined.
We suppose that the compensator has the following form: $\bar{\mu}(dt, dz) := F(dz)dt$, where conditions on the Levy measure F will be given later.
The initial condition X_0^θ, W and L are independent.

2.1 Assumptions

We suppose that the functions $b : \Theta \times \mathbb{R} \to \mathbb{R}$, $a : \mathbb{R} \to \mathbb{R}$ and $\gamma : \mathbb{R} \to \mathbb{R}$ satisfy the following assumptions:

ASSUMPTION 1: *The functions $a(x)$, $\gamma(x)$ and, for all $\theta \in \Theta$, $b(x, \theta)$ are globally Lipschitz. Moreover, the Lipschitz constant of b is uniformly bounded on Θ.*

Under Assumption 1 the equation (1) admits a unique non-explosive càdlàg adapted solution possessing the strong Markov property, cf (Applebaum, 2009) (Theorems 6.2.9. and 6.4.6.).

ASSUMPTION 2: *For all $\theta \in \Theta$ there exists a constant $t > 0$ such that X_t^θ admits a density $p_t^\theta(x, y)$ with respect to the Lebesgue measure on $\mathbb{R}$; bounded in $y \in \mathbb{R}$ and in $x \in K$ for every compact $K \subset \mathbb{R}$. Moreover, for every $x \in \mathbb{R}$ and every open ball $U \in \mathbb{R}$, there exists a point $z = z(x, U) \in supp(F)$ such that $\gamma(x)z \in U$.*

The last assumption was used in (Masuda, 2007) to prove the irreducibility of the process X^θ. Other sets of conditions, sufficient for irreducibility, are in (Masuda, 2007).

ASSUMPTION 3 (Ergodicity):

1. *For all $q > 0$, $\int_{|z|>1} |z|^q F(z)dz < \infty$.*
2. *For all $\theta \in \Theta$ there exists $C > 0$ such that $xb(x, \theta) \leq -C|x|^2$, if $|x| \to \infty$.*
3. *$|\gamma(x)|/|x| \to 0$ as $|x| \to \infty$.*
4. *$|a(x)|/|x| \to 0$ as $|x| \to \infty$.*
5. *$\forall \theta \in \Theta$, $\forall q > 0$ we have $\mathbb{E}|X_0^\theta|^q < \infty$.*

Assumption 2 ensures, together with the Assumption 3, the existence of unique invariant distribution π^θ, as well as the ergodicity of the process X^θ, as stated in the Lemma 2 below.

ASSUMPTION 4 (Jumps):

1. *The jump coefficient γ is bounded from below, that is $\inf_{x\in\mathbb{R}} |\gamma(x)| := \gamma_{min} > 0$.*
2. *The Lévy measure F is absolutely continuous with respect to the Lebesgue measure and we denote $F(z) = \frac{F(dz)}{dz}$.*
3. *We suppose that $\exists\, c > 0$ s.t., for all $z \in \mathbb{R}$, $F(z) \leq \frac{c}{|z|^{1+\alpha}}$, with $\alpha \in (0, 2)$.*

Assumptions 4.1 is useful to compare size of jumps of X and L.

ASSUMPTION 5 (Non-degeneracy): *There exists some $\alpha > 0$, such that $a^2(x) \geq \alpha$ for all $x \in \mathbb{R}$*

The Assumption 5 ensures the existence of the contrast function defined in Section 3.1.

ASSUMPTION 6 (Identifiability): *For all $\theta \neq \theta'$, $(\theta, \theta') \in \Theta^2$,*

$$\int_{\mathbb{R}} \frac{(b(\theta, x) - b(\theta', x))^2}{a^2(x)} d\pi^\theta(x) > 0$$

We can see that this last assumption is equivalent to

$$\forall \theta \neq \theta', \qquad (\theta, \theta') \in \Theta^2, \qquad b(\theta, .) \neq b(\theta', .). \tag{2}$$

We also need the following technical assumption:

ASSUMPTION 7:

1. The derivatives $\frac{\partial^{k_1+k_2} b}{\partial x^{k_1} \partial \theta^{k_2}}$, with $k_1 + k_2 \leq 4$ and $k_2 \leq 3$, exist and they are bounded if $k_1 \geq 1$. If $k_1 = 0$, for each $k_2 \leq 3$ they have polynomial growth.
2. The derivatives $a^{(k)}(x)$ exist and they are bounded for each $1 \leq k \leq 4$.
3. The derivatives $\gamma^{(k)}(x)$ exist and they are bounded for each $1 \leq k \leq 4$.

Define the asymptotic Fisher information by

$$I(\theta) = \int_{\mathbb{R}} \frac{(\dot{b}(\theta, x))^2}{a^2(x)} \pi^\theta(dx). \tag{3}$$

ASSUMPTION 8: *For all* $\theta \in \Theta$, $I(\theta) > 0$.

Remark 1. *If* $\alpha < 1$, *using Assumption 4.3 the stochastic differential equation (1) can be rewritten as follows:*

$$X_t^\theta = X_0^\theta + \int_0^t \bar{b}(\theta, X_s^\theta) ds + \int_0^t a(X_s^\theta) dW_s + \int_0^t \int_{\mathbb{R}\setminus\{0\}} \gamma(X_{s^-}^\theta) z \mu(ds, dz), \quad t \in \mathbb{R}_+, \tag{4}$$

where $\bar{b}(\theta, X_s^\theta) = b(\theta, X_s^\theta) - \int_{\mathbb{R}\setminus\{0\}} \gamma(X_{s^-}^\theta) z F(z) dz$.
This expression implies that X *follows diffusion equation* $X_t^\theta = X_0^\theta + \int_0^t \bar{b}(\theta, X_s^\theta) ds + \int_0^t a(X_s^\theta) dW_s$ *in the interval in which no jump occurred.*

From now on we denote the true parameter value by θ_0, an interior point of the parameter space Θ that we want to estimate. We shorten X for X^{θ_0}.
We will use some moment inequalities for jump diffusions, gathered in the following lemma:
Lemma 1. *Let* X *satisfies Assumptions 1-4. Let* $L_t := \int_0^t \int_{\mathbb{R}} z \tilde{\mu}(ds, dz)$ *and let* $\mathcal{F}_s := \sigma\left\{(W_u)_{0<u\leq s}, (L_u)_{0<u\leq s}, X_0\right\}$. *Then, for all* $t > s$,
1) for all $p \geq 2$, $\mathbb{E}[|X_t - X_s|^p]^{\frac{1}{p}} \leq c|t-s|^{\frac{1}{p}}$,
2) for all $p \geq 2$, $p \in \mathbb{N}$, $\mathbb{E}[|X_t - X_s|^p | \mathcal{F}_s] \leq c|t-s|(1 + |X_s|^p)$.
3) for all $p \geq 2$, $p \in \mathbb{N}$, $\sup_{h \in [0,1]} \mathbb{E}[|X_{s+h}|^p | \mathcal{F}_s] \leq c(1 + |X_s|^p)$.

The first two points follow from Theorem 66 of (Protter, 2005) and Proposition 3.1 in (Shimizu & Yoshida, 2006). The last point is a consequence of the second one: $\forall h \in [0,1]$,

$$\mathbb{E}[|X_{s+h}|^p | \mathcal{F}_s] = \mathbb{E}[|X_{s+h} - X_s + X_s|^p | \mathcal{F}_s] \leq c(\mathbb{E}[|X_{s+h} - X_s|^p | \mathcal{F}_s] + \mathbb{E}[|X_s|^p | \mathcal{F}_s]),$$

where c may change value line to line. Using the second point of Lemma 1 and the measurability of X_s with respect to $\mathcal{F}_s$, it is upper bounded by $c|h|(1 + |X_s|^p) + c|X_s|^p$. Therefore

$$\sup_{h \in [0,1]} \mathbb{E}[|X_{s+h}|^p | \mathcal{F}_s] \leq \sup_{h \in [0,1]} c|h|(1 + |X_s|^p) + c|X_s|^p \leq c(1 + |X_s|^p).$$

2.2 Ergodic properties of solutions

An important role is playing by ergodic properties of solution of equation (1)
The following Lemma states that Assumptions $1-4$ are sufficient for the existence of an invariant measure π^θ such that an ergodic theorem holds and moments of all order exist.
Lemma 2. *Under assumptions 1 to 4, for all* $\theta \in \Theta$, X^θ *admits a unique invariant distribution* π^θ *and the ergodic theorem holds:*

1. *For every measurable function* $g : \mathbb{R} \to \mathbb{R}$ *satisfying* $\pi^\theta(g) < \infty$, *we have a.s.*

$$\lim_{t \to \infty} \frac{1}{t} \int_0^t g(X_s^\theta) ds = \pi^\theta(g).$$

2. *For all $q > 0$, $\pi^\theta(|x|^q) < \infty$.*
3. *For all $q > 0$, $\sup_{t\geq 0} \mathbb{E}[|X_t^\theta|^q] < \infty$.*

A proof is in (Gloter, Loukianova, & Mai, 2018) (Section 8 of Supplement) in the case $\alpha \in (0,1)$, the proof relies on (Masuda, 2007). In order to use it also in the case $\alpha \geq 1$ we have to show that, taken $q > 2$ q even and $f^\star(x) = |x|^q$, $f^\star$ satisfies the drift condition $Af^\star = A_d f^\star + A_c f^\star \leq -c_1 f^\star + c_2$, where $c_1 > 0$ and $c_2 > 0$.
Using Taylor's formula up to second order we have

$$|A_d f^\star(x)| \leq c \int_{\mathbb{R}} \int_0^1 |z|^2 \, \|\gamma\|_\infty \, |f''^\star(x + sz\gamma(y))| F(z) ds dz =$$

$$= c \int_{\mathbb{R}} \int_0^1 |z|^2 \, \|\gamma\|_\infty \, q(q-1)|x + sz\gamma(y)|^{q-2} F(z) ds dz = o(|x|^q). \tag{5}$$

Concerning the generator's continuous part, we use the second point of Assumption 3 to get

$$A_c f^\star(x) = \frac{1}{2}\sigma^2(x) q(q-1) x^{q-2} + b(\theta, x) q \, x \, x^{q-2} \leq o(|x|^q) - cq|x|^2 x^{q-2} \leq o(|x|^q) - cf^\star(x). \tag{6}$$

By (5) and (6), the drift condition holds.

3 Construction of the estimator and main results

We exhibit a contrast function for the estimation of a parameter in the drift coefficient. We prove that the derived estimator is consistent and asymptotically normal.

3.1 Construction of the estimator

Let X^θ be the solution to (1). Suppose that we observe a finite sample

$$X_{t_0}, ..., X_{t_n}; \qquad 0 = t_0 \leq t_1 \leq ... \leq t_n,$$

where X is the solution to (1) with $\theta = \theta_0$. Every observation time point depends also on n, but to simplify the notation we suppress this index. We will be working in a high-frequency setting, i.e.

$$\Delta_n := \sup_{i=0,\dots,n-1} \Delta_{n,i} \longrightarrow 0, \quad n \to \infty,$$

with $\Delta_{n,i} := (t_{i+1} - t_i)$.
We assume $\lim_{n\to\infty} t_n = \infty$ and $n\Delta_n = O(t_n)$ as $n \to \infty$.
We introduce a jump filtered version of the gaussian quasi-likelihood. This leads to the following contrast function:
Definition 1. *For $\beta \in (0, \frac{1}{2})$ and $k > 0$, we define the contrast function $U_n(\theta)$ as follows:*

$$U_n(\theta) := \sum_{i=0}^{n-1} \frac{(X_{t_{i+1}} - m_{\theta,t_i,t_{i+1}}(X_{t_i}))^2}{a^2(X_{t_i})(t_{i+1} - t_i)} \varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}} - X_{t_i}) 1_{\{|X_{t_i}| \leq \Delta_{n,i}^{-k}\}} \tag{7}$$

where

$$m_{\theta,t_i,t_{i+1}}(x) := \frac{\mathbb{E}[X_{t_{i+1}}^\theta \varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}^\theta - X_{t_i}^\theta) | X_{t_i}^\theta = x]}{\mathbb{E}[\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}^\theta - X_{t_i}^\theta) | X_{t_i}^\theta = x]} \tag{8}$$

and

$$\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}} - X_{t_i}) = \varphi\Big(\frac{X_{t_{i+1}} - X_{t_i}}{\Delta_{n,i}^\beta}\Big),$$

with φ a smooth version of the indicator function, such that $\varphi(\zeta) = 0$ for each ζ, with $|\zeta| \geq 2$ and $\varphi(\zeta) = 1$ for each ζ, with $|\zeta| \leq 1$.
The last indicator aims to avoid the possibility that $|X_{t_i}|$ is big. The constant k is positive and it will be choosen later, related to the development of $m_{\theta,t_i,t_{i+1}}(x)$ (cf. Remark 2 below).
Moreover we define

$$m_{\theta,h}(x) := \frac{\mathbb{E}[X_h^\theta \varphi_{h^\beta}(X_h^\theta - X_0^\theta) | X_0^\theta = x]}{\mathbb{E}[\varphi_{h^\beta}(X_h^\theta - X_0^\theta) | X_0^\theta = x]}.$$

By the homogeneity of the equation we get that $m_{\theta,t_i,t_{i+1}}(x)$ depends only on the difference $t_{i+1} - t_i$ and so $m_{\theta,t_i,t_{i+1}}(x) = m_{\theta,t_{i+1}-t_i}(x)$ that we may denote simply as $m_\theta(x)$, in order to make the notation easier.

We define an estimator $\hat{\theta}_n$ of θ_0 as

$$\hat{\theta}_n \in \arg\min_{\theta\in\Theta} U_n(\theta). \tag{9}$$

The idea, with a finite intensity, is to use the size of $X_{t_{i+1}} - X_{t_i}$ in order to judge the existence of a jump in an interval $[t_i, t_{i+1})$. The increment of X with continuous transition could hardly exceed the threshold $\Delta_{n,i}^{\beta}$ with $\beta \in (0, \frac{1}{2})$. Therefore we can judge a jump occurred if $|X_{t_{i+1}} - X_{t_i}| > \Delta_{n,i}^{\beta}$. We keep the idea even when the intensity is no longer finite.
With a such defined $m_\theta(X_{t_i})$, using the true parameter value θ_0, we have that

$$\mathbb{E}[(X_{t_{i+1}} - m_{\theta_0,t_i,t_{i+1}}(X_{t_i}))\varphi_{\Delta_{n,i}^{\beta}}(X_{t_{i+1}} - X_{t_i})|X_{t_i} = x] = \mathbb{E}[X_{t_{i+1}}\varphi_{\Delta_{n,i}^{\beta}}(X_{t_{i+1}} - x)|X_{t_i} = x] +$$

$$- \frac{\mathbb{E}[X_{t_{i+1}}\varphi_{\Delta_{n,i}^{\beta}}(X_{t_{i+1}} - X_{t_i})|X_{t_i} = x]}{\mathbb{E}[\varphi_{\Delta_{n,i}^{\beta}}(X_{t_{i+1}} - X_{t_i})|X_{t_i} = x]}\mathbb{E}[\varphi_{\Delta_{n,i}^{\beta}}(X_{t_{i+1}} - X_{t_i})|X_{t_i} = x] = 0,$$

where we have just used the definition and the measurability of $m_{\theta_0,t_i,t_{i+1}}(X_{t_i})$.
But, as the transition density is unknown, in general there is no closed expression for $m_{\theta,h}(x)$, hence the contrast is not explicit. However, in the proof of our results we will need an explicit development of (7).

In the sequel, for $\delta \geq 0$, we will denote $R(\theta, \Delta_{n,i}^{\delta}, x)$ for any function $R(\theta, \Delta_{n,i}^{\delta}, x) = R_{i,n}(\theta, x)$, where $R_{i,n} : \Theta \times \mathbb{R} \longrightarrow \mathbb{R}$, $(\theta, x) \mapsto R_{i,n}(\theta, x)$ is such that

$$\exists c > 0 \qquad |R_{i,n}(\theta, x)| \leq c(1 + |x|^c)\Delta_{n,i}^{\delta} \tag{10}$$

uniformly in θ and with c independent of i, n.
The functions R represent the term of rest and have the following useful property, consequence of the just given definition:

$$R(\theta, \Delta_{n,i}^{\delta}, x) = \Delta_{n,i}^{\delta} R(\theta, \Delta_{n,i}^{0}, x). \tag{11}$$

We point out that it does not involve the linearity of R, since the functions R on the left and on the right side are not necessarily the same but only two functions on which the control (10) holds with $\Delta_{n,i}^{\delta}$ and $\Delta_{n,i}^{0}$, respectively.

We state asymptotic expansions for $m_{\theta,\Delta_{n,i}}$. The cases $\alpha < 1$ and $\alpha \geq 1$ yield to different magnitude for the rest term.

Case $\alpha \in (0, 1)$:
Theorem 1. *Suppose that Assumptions 1 to 4 hold and that $\beta \in (0, \frac{1}{2})$ and $\alpha \in (0, 1)$ are given in definition* 1 *and the third point of Assumption* 4*, respectively. Then*

$$\mathbb{E}[\varphi_{\Delta_{n,i}^{\beta}}(X_{t_{i+1}}^{\theta} - X_{t_i}^{\theta})|X_{t_i}^{\theta} = x] = 1 + R(\theta, \Delta_{n,i}^{(1-\alpha\beta)\wedge(2-3\beta)}, x). \tag{12}$$

Theorem 2. *Suppose that Assumptions 1 to 4 hold and that $\beta \in (0, \frac{1}{2})$ and $\alpha \in (0, 1)$ are given in definition* 1 *and the third point of Assumption* 4*, respectively. Then*

$$\mathbb{E}[(X_{t_{i+1}}^{\theta} - x)\varphi_{\Delta_{n,i}^{\beta}}(X_{t_{i+1}}^{\theta} - X_{t_i}^{\theta})|X_{t_i}^{\theta} = x] = \Delta_{n,i}\, b(x, \theta) + \tag{13}$$

$$-\Delta_{n,i} \int_{\mathbb{R}\setminus\{0\}} z\,\gamma(x)\,[1 - \varphi_{\Delta_{n,i}^{\beta}}(\gamma(x)z)]\,F(z)dz \,+\, R(\theta, \Delta_{n,i}^{2-2\beta}, x).$$

There exists $k_0 > 0$ such that, for $|x| \leq \Delta_{n,i}^{-k_0}$,

$$m_{\theta,\Delta_{n,i}}(x) = x + \Delta_{n,i}\, b(x, \theta) + \tag{14}$$

$$-\Delta_{n,i} \int_{\mathbb{R}\setminus\{0\}} z\,\gamma(x)\,[1 - \varphi_{\Delta_{n,i}^{\beta}}(\gamma(x)z)]\,F(z)dz \,+\, R(\theta, \Delta_{n,i}^{2-2\beta}, x).$$

.

Case $\alpha \in [1, 2)$:
Theorem 3. *Suppose that Assumptions 1 to 4 hold and that $\beta \in (0, \frac{1}{2})$ and $\alpha \in [1, 2)$ are given in definition* 1 *and the third point of Assumption* 4*, respectively. Then*

$$\mathbb{E}[\varphi_{\Delta_{n,i}^{\beta}}(X_{t_{i+1}}^{\theta} - X_{t_i}^{\theta})|X_{t_i}^{\theta} = x] = 1 + R(\theta, \Delta_{n,i}^{(1-\alpha\beta)\wedge(2-4\beta)}, x). \tag{15}$$

Theorem 4. *Suppose that Assumptions 1 to 4 hold and that* $\beta \in (0, \frac{1}{3})$ *and* $\alpha \in [1, 2)$ *are given in definition* 1 *and the third point of Assumption* 4*, respectively. Then*

$$\mathbb{E}[(X^\theta_{t_{i+1}} - x)\varphi_{\Delta^\beta_{n,i}}(X^\theta_{t_{i+1}} - X^\theta_{t_i})|X^\theta_{t_i} = x] = \Delta_{n,i}\, b(x, \theta)+ \tag{16}$$

$$-\Delta_{n,i} \int_{\mathbb{R}\setminus\{0\}} z\,\gamma(x)\,[1 - \varphi_{\Delta^\beta_{n,i}}(\gamma(x)z)]\, F(z)dz \,+\, R(\theta, \Delta^{2-3\beta}_{n,i}, x).$$

There exists $k_0 > 0$ *such that, for* $|x| \le \Delta^{-k_0}_{n,i}$,

$$m_{\theta,\Delta_{n,i}}(x) = x + \Delta_{n,i}\, b(x, \theta)+ \tag{17}$$

$$-\Delta_{n,i} \int_{\mathbb{R}\setminus\{0\}} z\,\gamma(x)\,[1 - \varphi_{\Delta^\beta_{n,i}}(\gamma(x)z)]\, F(z)dz \,+\, R(\theta, \Delta^{2-3\beta}_{n,i}, x).$$

Remark 2. *The constant* k *in the definition* (7) *of contrast function can be taken in the interval* $(0, k_0]$*. In this way* $\Delta^{-k}_{n,i} \le \Delta^{-k_0}_{n,i}$ *and so* (14) *or* (17) *holds for* $|x| = |X_{t_i}|$ *smaller than* $\Delta^{-k}_{n,i}$*.*
If it is not the case the contribution of the observation X_{t_i} *in the contrast function is just* 0*. However we will see that suppressing the contribution of too big* $|X_{t_i}|$ *does not effect the efficiency property of our estimator.*
Remark 3. *In the development* (13) *or* (16) *the term* $\Delta_{n,i} \int_{\mathbb{R}\setminus\{0\}} z\,\gamma(x)\,[1 - \varphi_{\Delta^\beta_{n,i}}(\gamma(x)z)]\, F(z)dz$ *is independent of* θ*, hence it will disappear in the difference* $m_\theta(x) - m_{\theta_0}(x)$*, but it is not negligible compared to* $\Delta_{n,i}\, b(x, \theta)$ *since its order is* $\Delta_{n,i}$ *if* $\alpha \in (0, 1)$ *and at most* $\Delta^{\frac{1}{2}}_{n,i}$ *if* $\alpha \in [1, 2)$*. Indeed, by the definition of the function* φ*, we know that we can consider as support of* $\varphi_{\Delta^\beta_{n,i}}(0) - \varphi_{\Delta^\beta_{n,i}}(\gamma(x)z)$ *the interval* $c \times [-\frac{\Delta^\beta_{n,i}}{\|\gamma\|_\infty}, \frac{\Delta^\beta_{n,i}}{\|\gamma\|_\infty}]^c$*. If* $\alpha < 1$*, using moreover the third point of Assumption 4 we get the following estimation:*

$$|\Delta_{n,i} \int_{\mathbb{R}\setminus\{0\}} z\,\gamma(x)\,[1 - \varphi_{\Delta^\beta_{n,i}}(\gamma(x)z)]\, F(z)dz| \le R(\theta_0, \Delta^1_{n,i}, X_{t_i}). \tag{18}$$

Otherwise, if $\alpha \ge 1$*, we have*

$$|\Delta_{n,i} \int_{\mathbb{R}\setminus\{0\}} z\,\gamma(x)\,[1 - \varphi_{\Delta^\beta_{n,i}}(\gamma(x)z)]\, F(z)dz| \le c|\Delta_{n,i}| \int_{c\times[-\frac{\Delta^\beta_{n,i}}{\|\gamma\|_\infty}, \frac{\Delta^\beta_{n,i}}{\|\gamma\|_\infty}]^c} |z|^{-\alpha} = R(\theta, \Delta^{1+\beta(1-\alpha)}_{n,i}, x),$$

with $\beta \in (0, \frac{1}{2})$ *and* $\alpha \in [1, 2)$*, hence the exponent on* $\Delta_{n,i}$ *is always more than* $\frac{1}{2}$*.*
We can therefore write in the first case

$$m_{\theta,\Delta_{n,i}}(x) = x + R(\theta, \Delta_{n,i}, x) = R(\theta, \Delta^0_{n,i}, x) \tag{19}$$

and in the second

$$m_{\theta,\Delta_{n,i}}(x) = x + R(\theta, \Delta^{1+\beta(1-\alpha)}_{n,i}, x) = R(\theta, \Delta^0_{n,i}, x). \tag{20}$$

Remark 4. *In Theorems 1 - 3 we do not need conditions on* β *because, for each* $\beta \in (0, \frac{1}{2})$ *and for each* $\alpha \in (0, 2)$ *the exponent on* $\Delta_{n,i}$ *is positive and therefore the last term of* (15) *is negligible compared to* 1*. In Theorem 4, instead,* R *is a negligible function if and only if* $2 - 3\beta \ge 1$*, it means that it must be* $\beta \le \frac{1}{3}$*. We have taken* $\beta \in (0, \frac{1}{3})$ *and so such a condition is always respected.*

3.2 Main results

Let us introduce the Assumption A_β that turns out starting from Theorems 1, 2, 3 and 4:

ASSUMPTION A_β: We choose $\beta \in (0, \frac{1}{2})$ if $\alpha \in (0, 1)$. If on the contrary $\alpha \in [1, 2)$, then we take β in $(0, \frac{1}{3})$.

The following theorems give a general consistency result and the asymptotic normality of the estimator $\hat{\theta}_n$, that hold without further assumptions on n and Δ_n.
Theorem 5. *(Consistency)*
Suppose that Assumptions 1 to 7 and A_β *hold and let* k *of the definition of the contrast function* (7) *be in* $(0, k_0)$*. Then the estimator* $\hat{\theta}_n$ *is consistent in probability:*

$$\hat{\theta}_n \xrightarrow{\mathbb{P}} \theta_0, \qquad n \to \infty.$$

Recalling that the Fisher information I is given by (3), we give the following theorem.

Theorem 6. *(Asymptotic normality)*
Suppose that Assumptions 1 to 8 and A_β hold, and $0 < k < k_0$.
Then the estimator $\hat{\theta}_n$ is asymptotically normal:

$$\sqrt{t_n}(\hat{\theta}_n - \theta_0) \xrightarrow{\mathcal{L}} N(0, I^{-1}(\theta_0)), \qquad n \to \infty.$$

Remark 5. *Furthermore, the estimator $\hat{\theta}_n$ is asymptotically efficient in the sense of the Hájek-Le Cam convolution theorem.*
The Hájek–LeCam convolution theorem states that any regular estimator in a parametric model which satisfies LAN property is asymptotically equivalent to a sum of two independent random variables, one of which is normal with asymptotic variance equal to the inverse of Fisher information, and the other having arbitrary distribution. The efficient estimators are those with the second component identically equal to zero.
The model (1) is LAN with Fisher information $I(\theta) = \int_{\mathbb{R}} \frac{(\dot{b}(\theta,x))^2}{a^2(x)} \pi^\theta(dx)$ (see (Gloter, Loukianova, & Mai, 2018)) and thus $\hat{\theta}_n$ is efficient.
Remark 6. *We point out that, contrary to the papers (Gloter, Loukianova, & Mai, 2018) and (Shimizu & Yoshida, 2006), in this case there is not any condition on the sampling, that can be irregular and with Δ_n that goes slowly to zero. On the other hand, our contrast function relies on the quantity $m_{\theta,h}(x)$ which is not explicit in general.*

4 Practical implementation of the contrast method

In order to use in practice the contrast function (7), one need to know the values of the quantities $m_{\theta,t_i,t_{i+1}}(X_{t_i})$. In most cases, it seems impossible to find an explicit expression for the function $m_{\theta,h}$ appearing in Definition 1. However, explicit or numerical approximations of this function seem available in many situations.

4.1 Approximate contrast function

Let us assume that one has at disposal an approximation of the function $m_{\theta,h}(x)$, denoted by $\widetilde{m}_{\theta,h}(x)$ which satisfies, for $|x| \le h^{-k_0}$,

$$|\widetilde{m}_{\theta,h}(x) - m_{\theta,h}(x)| \le R(\theta, h^\rho, x)$$

where the constant $\rho > 1$ assesses the quality of the approximation. We assume that the first three derivatives of $\tilde{m}_{h,\theta}$ with respect to the parameter provide approximation of the derivatives of $m_{h,\theta}$, in the following way

$$|\frac{\partial^i \widetilde{m}_{\theta,h}(x)}{\partial \theta^i} - \frac{\partial^i m_{\theta,h}(x)}{\partial \theta^i}| \le R(\theta, h^{1+\epsilon}, x), \quad \text{for } i = 1, 2, \tag{21}$$

$$|\frac{\partial^3 \widetilde{m}_{\theta,h}(x)}{\partial \theta^3} - \frac{\partial^3 m_{\theta,h}(x)}{\partial \theta^3}| \le R(\theta, h, x), \tag{22}$$

for all $|x| \le h^{-k_0}$ and where $\epsilon > 0$. Let us stress that from Proposition 8 below, we know the derivatives with respect to θ of the quantity $m_{h,\theta}$.

Now, we consider $\widetilde{\theta}_n$ the estimator obtained from minimization of the contrast function (7) where one has replaced $m_{\theta,t_i,t_{i+1}}(X_{t_i})$ by its approximation $\widetilde{m}_{\theta,\Delta_{n,i}}(X_{t_i})$. Then, the result of Theorem 6 can be extended as follows.

Proposition 1. *Suppose that Assumptions 1 to 8 and A_β hold, with $0 < k < k_0$, and that $\sqrt{n}\Delta_n^{\rho-1/2} \to 0$ as $n \to \infty$.*
Then, the estimator $\widetilde{\theta}_n$ is asymptotically normal:

$$\sqrt{t_n}(\widetilde{\theta}_n - \theta_0) \xrightarrow{\mathcal{L}} N(0, I^{-1}(\theta_0)), \qquad n \to \infty.$$

We give below several examples of approximations of $m_{\theta,h}$. Let us stress that, in general, Theorem 2 (resp. Theorem 4) provides an explicit approximation of $m_{\theta,\Delta_{n,i}}(x)$ with an error of order $\Delta_{n,i}^{2-2\beta}$ (resp. of order $\Delta_{n,i}^{2-3\beta}$). They can be used to construct an explicit contrast function. In the next section we show that when the intensity is finite, it is possible to construct an explicit approximation of $m_{\theta,h}$ with arbitrarily high order.

4.2 Explicit contrast in the finite intensity case.

In the case with finite intensity it is possible to make the contrast explicit, using the development of $m_{\theta,\Delta_{n,i}}$ proved in the next proposition. We need the following assumption:

ASSUMPTION A_f:

1. We have $F(z) = \lambda F_0(z)$, $\int_{\mathbb{R}} F_0(z)dz = 1$ and F is a $\mathcal{C}^\infty$ function.
2. We assume that $x \mapsto a(x)$, $x \mapsto b(x, \theta)$ and $x \mapsto \gamma(x)$ are $\mathcal{C}^\infty$ functions, they have at most uniform in θ polynomial growth as well as their derivatives.

Let us define $A_K^{(k)}(x) = \bar{A}_c^k(g)(x)$, with $g(y) = (y - x)$ and $\bar{A}_c(f) = \bar{b}f' + \frac{1}{2}a^2 f''$; $\bar{b}(\theta, y) = b(\theta, y) - \int_{\mathbb{R}} \gamma(y)zF(z)dz$ as in the Remark 1.

Proposition 2. *Assume that A_f holds and let φ be a $\mathcal{C}^\infty$ function that has compact support and such that $\varphi \equiv 1$ on $[-1, 1]$ and $\forall k \in \{0, ..., M\}$, $\int_{\mathbb{R}} x^k \varphi(x)dx = 0$ for $M \geq 0$. Then, for $|x| \leq \Delta_{n,i}^{-k_0}$ with some $k_0 > 0$,*

$$m_{\theta,\Delta_{n,i}}(x) = x + \sum_{k=1}^{\lfloor \beta(M+2) \rfloor} A_K^{(k)}(x) \frac{\Delta_{n,i}^k}{k!} + R(\theta, \Delta_{n,i}^{\beta(M+2)}, x). \tag{23}$$

In order to say that (23) holds, we have to prove the existence of a function φ with a compact support such that $\varphi \equiv 1$ on $[-1, 1]$ and, $\forall k \in \{0, ..., M\}$, $\int_{\mathbb{R}} x^k \varphi(x)dx$. We build it through ψ, a function with compact support, $\mathcal{C}^\infty$, such that $\psi|_{[-1,1]}(x) = \frac{x^M}{M!}$. We then define $\varphi(x) := \frac{\partial^M}{\partial x^M} \psi(x)$.
In this way we have $\varphi \equiv 1$ on $[-1, 1]$, φ is $\mathcal{C}^\infty$, with compact support and such that for each $l \in \{0, ...M\}$, using the integration by parts, $\int_{\mathbb{R}} x^l \varphi(x)dx = 0$, as we wanted.

Remark 7. *The development (23) is the same found in Kessler (Kessler, 1997) in the case without jumps and it is obtained by the iteration of the continuous generator $\bar{A}_c$. Hence, it is completely explicit. Let us stress that in Kessler (Kessler, 1997) the right hand side of (23) stands for an approximation of $E[\bar{X}^\theta_{\Delta_{n,i}} \mid \bar{X}^\theta_0 = x]$ where $\bar{X}^\theta$ is the continuous diffusion solution of $d\bar{X}^\theta_t = \bar{b}(\theta, \bar{X}^\theta_s)ds + \sigma(\bar{X}^\theta_s)dW_s$. From Proposition 2, the right hand side of (23) is also an approximation of $m_{\theta,\Delta_{n,i}}(x) = \dfrac{E[X^\theta_{\Delta_{n,i}} \varphi_{\Delta^\beta_{n,i}}(X^\theta_{\Delta_{n,i}} - x) \mid X^\theta_0 = x]}{E[\varphi_{\Delta^\beta_{n,i}}(X^\theta_{\Delta_{n,i}} - x) \mid X^\theta_0 = x]}$ in the case of finite activity jumps, and for a truncation kernel φ satisfying $\forall k \in \{0, ..., M\}$, $\int_{\mathbb{R}} x^k \varphi(x)dx = 0$. We emphasize that in the expansion of $m_{\theta,\Delta_{i,n}}$ given in Proposition 2, the contribution of the discontinuous part of the generator disappears only thanks to the choice of an oscillating function φ.*

Remark 8. *In the definition of the contrast function (7) we can replace $m_{\theta,\Delta_{i,n}}(x)$ with the explicit approximation $\tilde{m}^k_{\theta,\Delta_{n,i}}(x) := x + \sum_{h=1}^k \frac{\Delta^h_{n,i}}{h!} A_K^{(h)}(x)$, with an error $R(\theta, \Delta^k_{n,i}, x)$, for $k \leq \lfloor 2(M+1)\beta \rfloor$. Using $A_K^{(1)}(x) = \Delta_{n,i}[b(\theta, x) - \int_{\mathbb{R}} \gamma(y)zF(z)dz]$ and the expansions (120)–(122) we deduce that the conditions (21) – (22) are valid. Then, by application of Proposition 1, we can see that the associated estimator is efficient under the assumption $\sqrt{n}\Delta_n^{k-\frac{1}{2}} \to 0$ for $n \to \infty$. As M, and thus k, can be chosen arbitrarily large, we see that the sampling step Δ_n is allowed to converge to zero in a arbitrarily slow polynomial rate as a function of n. It turns out that a slow sampling step necessitates to choose a truncation function with more vanishing moments.*

5 Numerical experiments

5.1 Finite jump activity

Let us consider the model

$$X_t = X_0 + \int_0^t (\theta_1 X_s + \theta_2)ds + \sigma W_t + \gamma \int_0^t \int_{\mathbb{R}\setminus\{0\}} z\tilde{\mu}(ds, dz), \tag{24}$$

where the compensator of the jump measure is $\overline{\mu}(ds, dz) = \lambda F_0(z)dsdz$ for F_0 the probability density of the law $\mathcal{N}(\mu_J, \sigma_J^2)$ with $\mu_J \in \mathbb{R}$, $\sigma_J > 0$, $\sigma > 0$, $\theta_1 < 0$, $\theta_2 \in \mathbb{R}$, $\gamma \geq 0$, $\lambda \geq 0$. Since the jump activity is finite, we know from Section 4.2 that the function $m_{(\theta_1,\theta_2),\Delta_{n,i}}(x)$ can be approximated at any order using (23). As the latter is also the asymptotic expansion of the first conditional moment for the continuous S.D.E. $\bar{X}_t = \bar{X}_0 + \int_0^t (\theta_1 \bar{X}_s + \theta_2 - \gamma\lambda\mu_J)ds + \sigma W_t$, which is explicit due to the linearity of the model, we decide to directly use the expression of the conditional moment and set

$$\widetilde{m}_{(\theta_1,\theta_2),\Delta_{n,i}}(x) = (x + \frac{\theta_2}{\theta_1} - \frac{\gamma\lambda\mu_J}{\theta_1})e^{\theta_1\Delta_{n,i}} + \frac{\gamma\lambda\mu_J - \theta_2}{\theta_1}. \tag{25}$$

Following Nikolskii (Nikolskii, 1977), we construct oscillating truncation functions in the following way. First, we choose $\varphi^{(0)} : \mathbb{R} \to [0,1]$ a $\mathcal{C}^\infty$ symmetric function with support on $[-2,2]$ such that $\varphi^{(0)}(x) = 1$ for $|x| \leq 1$. We let, for $d > 1$, $\varphi_d^{(1)}(x) = (d\varphi^{(0)}(x) - \varphi^{(0)}(x/d))/(d-1)$, which is a function equal to 1 on $[-1,1]$, vanishing on $[-d,d]^c$ and such that $\int_\mathbb{R} \varphi_d^{(1)}(x)dx = 0$. For $l \in \mathbb{N}$, $l \geq 1$, and $d > 1$, we set $\varphi_d^{(l)}(x) = c_d^{-1} \sum_{k=1}^l C_l^k (-1)^{k+1} \frac{1}{k} \varphi_d^{(1)}(x/k)$, where $c_d = \sum_{k=1}^l C_l^k (-1)^{k+1} \frac{1}{k}$. One can check that $\varphi_d^{(l)}$ is compactly supported, equal to 1 on $[-1,1]$, and that for all $k \in \{0,\dots,l\}$, $\int_\mathbb{R} x^k \varphi_d^{(l)}(x)dx = 0$, for $l \geq 1$. With these notations, we estimate the parameter $\theta = (\theta_1,\theta_2)$ by minimization of the contrast function

$$U_n(\theta) = \sum_{i=0}^{n-1} (X_{t_{i+1}} - \widetilde{m}_{(\theta_1,\theta_2),\Delta_{n,i}}(X_{t_i}))^2 \varphi^{(l)}_{c\Delta^\beta_{n,i}}(X_{t_{i+1}} - X_{t_i}), \tag{26}$$

where $l \in \mathbb{N}$ and $c > 0$ will be specified latter.

For numerical simulations, we choose $T = 2000$, $n = 10^4$, $\Delta_{i,n} = \Delta_n = 1/5$, $\theta_1 = -0.5$, $\theta_2 = 2$ and $X_0 = x_0 = 4$. We estimate the bias and standard deviation of our estimators using a Monte Carlo method based on 5000 replications. As a start, we consider a situation without jumps $\lambda = 0$, in which we remove the truncation function φ in the contrast, as it is useless in absence of jumps. In Table 1, we compare the estimator $\widetilde{\theta}_n$ which uses the Kessler exact bias correction given by (25), with an estimator based on the Euler scheme approximation where one uses the approximation $\widetilde{m}^{\text{Euler}}_{(\theta_1,\theta_2),\Delta_{n,i}}(x) = x + \Delta_{n,i}(\theta_1 x + \theta_2)$. From Table 1 we see that the estimator $\tilde{\theta}_n^{\text{Euler}}$ based on Euler contrast exhibits some bias which is completely removed using Kessler's correction. Next, we set a jump intensity $\lambda = 0.1$, with jumps size whose common law is $\mathcal{N}(0,2)$ and set $\gamma = 1$. We use the contrast function relying on (25). Results are given for three choices of truncation function, $\varphi^{(0)}$, $\varphi_d^{(2)}$ and $\varphi_d^{(3)}$ where $d = 3$. Plots of these functions are given in Figure 1. We choose $\beta = 0.49$ and $c = 1$. As the true value of the volatility is $\sigma = 0.3$, this choice enables most of the increments without jumps of X on $[t_i, t_{i+1}]$ to be such that $\varphi^{(l)}_{c\Delta^\beta_{n,i}}(X_{t_{i+1}} - X_{t_i}) = 1$. Let us stress that, if σ is unknown, it is possible to estimate, even roughly, the local value of the volatility in order to choose c accordingly (see (Gloter, Loukianova, & Mai, 2018) for analogous discussion). Results in Table 2 show that the estimator works well, with a reduced bias for all choices of truncation function. Especially the bias is much smaller than the one of the Euler scheme contrast in absence of jumps. It shows the benefit of using (25) in the contrast function, even if the truncation function is not oscillating as is it when we consider $\varphi^{(0)}$. We remark that by the choice of a symmetric truncation function one has $\int_\mathbb{R} u\varphi^{(0)}(u)du = 0$ and inspecting the proof of Proposition 8 it can be seen that this conditions is sufficient, in the expansion of $m_{\theta,\Delta_{n,i}}$, to suppress the largest contribution of the discrete part of the generator.

If the number of jumps is greater, e.g. for $\lambda = 1$, we see in Table 3 that using the oscillating kernels $\varphi_d^{(2)}$, $\varphi_d^{(3)}$ yields to a smaller bias than using $\varphi^{(0)}$, whereas it tends to increase the standard deviation of the estimator. The estimator we get using $\varphi_d^{(3)}$ performs well in this situation, it has a negligible bias and a standard deviation comparable to the one in the case where the process has no jump.

	Mean (std) for $\theta_1 = -0.5$	Mean (std) for $\theta_2 = 2$
$\tilde{\theta}_n^{\text{Euler}}$	-0.4783 (0.0213)	1.9133 (0.0856)
$\widetilde{\theta}_n$	-0.5021 (0.0236)	2.0084 (0.0947)

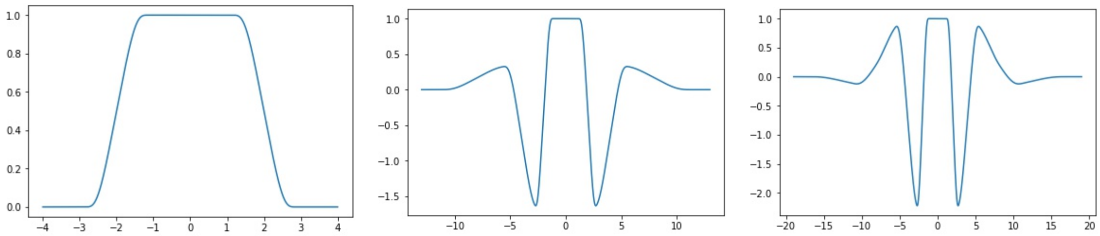
Table 1: Process without jump

	Mean (std) for $\theta_1 = -0.5$	Mean (std) for $\theta_2 = 2$
$\widetilde{\theta_n}$ using $\varphi^{(0)}$	-0.4967 (0.0106)	1.9869 (0.0430)
$\widetilde{\theta_n}$ using $\varphi_d^{(2)}$	-0.4990 (0.0153)	1.9959 (0.0622)
$\widetilde{\theta_n}$ using $\varphi_d^{(3)}$	-0.5006 (0.0196)	2.0023 (0.0798)

Table 2: Gaussian jumps with $\lambda = 0.1$

	Mean (std) for $\theta_1 = -0.5$	Mean (std) for $\theta_2 = 2$
$\widetilde{\theta}_n$ using $\varphi^{(0)}$	-0.4623 (0.0059)	1.8495 (0.0256)
$\widetilde{\theta}_n$ using $\varphi_d^{(2)}$	-0.4886 (0.0161)	1.9549 (0.0710)
$\widetilde{\theta}_n$ using $\varphi_d^{(3)}$	-0.5033 (0.0243)	2.0136 (0.1059)

Table 3: Gaussian jumps with $\lambda = 1$



(a) $\varphi^{(0)}$ (b) $\varphi_d^{(2)}$ with $d = 3$ (c) $\varphi_d^{(3)}$ with $d = 3$

Figure 1: Plot of the truncation functions

5.2 Infinite jumps activity

Let us consider X solution to the stochastic differential equation (24), where the compensator of the jump measure is $\overline{\mu}(ds, dz) = \frac{e^{-z}}{z^{1+\alpha}} 1_{(0,\infty)}(z) ds dz$ with $\alpha \in (0, 1)$. This situation corresponds to the choice of the Levy process $(\int_0^t \int_{\mathbb{R}\setminus\{0\}} z\tilde{\mu}(ds, dz))_t$ being a tempered α-stable jump process. In the case of infinite jump activity, we have no result providing approximations at any arbitrary order of $m_{\theta,\Delta_{n,i}}$. However, we can use Theorem 2 to find some useful explicit approximation.

According to (14) and taking into account that the threshold level is $c\Delta_{n,i}^{\beta}$ for some $c > 0$, we have

$$\begin{aligned} m_{\theta,\Delta_{n,i}}(x) &= x + \Delta_{n,i}(\theta_1 x + \theta_2) - \Delta_{n,i}\gamma \int_0^\infty \frac{e^{-z}}{z^\alpha} dz + \Delta_{n,i}\gamma \int_0^\infty \varphi_{c\Delta_{n,i}^\beta}(\gamma z) \frac{e^{-z}}{z^\alpha} dz + R(\theta, \Delta_{n,i}^{2-2\beta}, x) \\ &= x + \Delta_{n,i}\overline{b}(x, \theta_1, \theta_2) + \Delta_{n,i}^{1+\beta(1-\alpha)} c^{1-\alpha}\gamma^\alpha \int_0^\infty \varphi(v) \frac{e^{-\frac{cv\Delta_{n,i}^\beta}{\gamma}}}{v^\alpha} dv + R(\theta, \Delta_{n,i}^{2-2\beta}, x), \end{aligned}$$

where in the last line, following the notation of Remark 1, we have set $\overline{b}(x, \theta_1, \theta_2) = (\theta_1 x + \theta_2) - \gamma \int_0^\infty \frac{e^{-z}}{z^\alpha} dz$, and we make the change of variable $v = \frac{\gamma z}{c\Delta_{n,i}^\beta}$. This leads us to consider the approximation

$$\widetilde{m}_{\theta,\Delta_{n,i}}(x) = x + \Delta_{n,i}\overline{b}(x, \theta_1, \theta_2) + \Delta_{n,i}^{1+\beta(1-\alpha)} c^{1-\alpha}\gamma^\alpha \int_0^\infty \varphi(v) \frac{1}{v^\alpha} dv, \tag{27}$$

which is such that $|\widetilde{m}_{\theta,\Delta_{n,i}}(x) - m_{\theta,\Delta_{n,i}}(x)| \leq R(\theta, \Delta_{n,i}^{(2-2\beta)\wedge(1+\beta(2-\alpha))}, x)$.

For numerical simulations, we choose $T = 100$, $n = 10^4$, $\Delta_{i,n} = \Delta_n = 1/100$, $\theta_1 = -0.5$, $\theta_2 = 2$, $X_0 = x_0 = 4$, $\gamma = 1$, $\sigma = 0.3$ and $\alpha \in \{0.1, \ 0.3, \ 0.5\}$. To illustrate the estimation method, we focus on the estimation of the parameter θ_2 only, as the minimisation of the contrast defined by (26)–(27) yields to the explicit estimator,

$$\begin{aligned} \widetilde{\theta}_{2,n} = &\frac{\sum_{i=0}^{n-1}(X_{t_{i+1}} - X_{t_i} - \Delta_n\theta_1 X_{t_i})\varphi_{c\Delta_n^\beta}(X_{t_{i+1}} - X_{t_i})}{\Delta_n \sum_{i=0}^{n-1} \varphi_{c\Delta_n^\beta}(X_{t_{i+1}} - X_{t_i})} - \gamma \int_0^\infty \frac{e^{-z}}{z^\alpha} dz \\ &- \Delta_n^{\beta(1-\alpha)} c^{1-\alpha}\gamma^\alpha \int_0^\infty \varphi(v) \frac{1}{v^\alpha} dv \\ =: &\ \widetilde{\theta}_{2,n}^{\text{Euler}} - \Delta_n^{\beta(1-\alpha)} c^{1-\alpha}\gamma^\alpha \int_0^\infty \varphi(v) \frac{1}{v^\alpha} dv. \end{aligned} \tag{28}$$

We can see that the estimator $\widetilde{\theta}_{2,n}$ is a corrected version of the estimator $\widetilde{\theta}_{2,n}^{\text{Euler}}$, that would result from the choice of the approximation $m_{\theta,\Delta_n}(x) \approx x + \Delta_n\overline{b}(x, \theta_1, \theta_2)$ in the definition of the contrast function. Comparing with estimators of earlier works (e.g. (Gloter, Loukianova, & Mai, 2018), (Shimizu, 2006)), the presence of this correction term appears new. In lines 2–3 of Table 4, we compare the mean and standard deviation of $\widetilde{\theta}_{2,n}$ and $\widetilde{\theta}_{2,n}^{\text{Euler}}$ for $\alpha \in \{0.1, \ 0.3, \ 0.5\}$ and with the choice $c = 1$, $\beta = 0.49$

and $\varphi = \varphi^{(0)}$ (see Figure 1). We see that the estimator $\widetilde{\theta}_{2,n}$ performs well and the correction term in (28) drastically reduces the bias present in $\widetilde{\theta}_{2,n}^{\text{Euler}}$, especially when the jump activity near 0 is high, corresponding to larger values of α. If we take a threshold level $c = 1.5$ higher, we see in line 5 of Table 4 that the bias of the estimator $\widetilde{\theta}_{2,n}^{\text{Euler}}$ increases, since the estimator $\widetilde{\theta}_{2,n}^{\text{Euler}}$ keeps more jumps that induce a stronger bias. On the other hand, the bias of the estimator $\widetilde{\theta}_{2,n}$ remains small (see line 4 of Table 4), as the correction term in (28) increases with c.

		$\alpha = 0.1$	$\alpha = 0.3$	$\alpha = 0.5$
c=1	$\widetilde{\theta}_{2,n}$	1.99 (0.0315)	1.98 (0.0340)	1.97 (0.0367)
	$\widetilde{\theta}_{2,n}^{\text{Euler}}$	2.20 (0.0315)	2.37 (0.0340)	2.76 (0.0367)
c=1.5	$\widetilde{\theta}_{2,n}$	1.97 (0.0340)	1.96 (0.0363)	1.94 (0.0397)
	$\widetilde{\theta}_{2,n}^{\text{Euler}}$	2.28 (0.0340)	2.48 (0.0363)	2.90 (0.0397)

Table 4: Mean (std) for the estimation of $\theta_2 = 2$

5.3 Conclusion and perspectives for practical applications

In this paper, we have shown that it is theoretically possible to estimate the drift parameter efficiently under the sole condition of a sampling step converging to zero. However, the contrast function relies on the quantity $m_{h,\theta}(x)$ which is usually not explicit. For practical implementation, the question of approximation of $m_{h,\theta}(x)$ is crucial, and one also has face the question of choosing the threshold level, characterized here by c, β and φ. On contrary to more conventional threshold methods, it appears here that the estimation quality seems less sensitive to choice of the threshold level, as the quantity $m_{h,\theta}(x)$ depends by construction on this threshold level and may compensate for too large threshold. On the other hand, the quantity $m_{h,\theta}(x)$ can be numerically very far from the approximation derived form the Euler scheme approximation. This can be seen in the example of Section 5.2, where the correction term of the estimator is, on this finite sample example, essentially of the same magnitude as the estimated quantity. A perspective, in the situation of infinite jump activity, would be to numerically approximate the function $x \mapsto m_{h,\theta}(x)$, using for instance a Monte Carlo approach, and provide more precise corrections than the explicit correction used in Section 5.2.

In the specific situation of finite activity, we proposed an explicit approximation of $m_{h,\theta}(x)$ with arbitrary order. This approximation is the same one as Kessler's approximation in absence of jumps, and it relies on the choice of oscillating truncation functions. A crucial point in the proof of the expansion of $m_{h,\theta}(x)$ given in Proposition 2 is that the support of the truncation function $\varphi_{c\Delta_{n,i}^{\beta}}$ is small compared to the typical scale where the density of the jumps law varies. However, our construction of oscillating function is such that the support of $\varphi = \varphi_d^{(l)}$ tends to be larger as the number of oscillations l is larger, which yields to restrictions for the choice of l on finite sample. Moreover, the truncation function takes large negative values as well, which makes the minimization of the contrast function unstable if the parameter set is too large. Perspective for further works would be to extend Proposition 2 for a non oscillating function φ. We expect that the resulting asymptotic expansion would involve additional terms related to the quantities $\int u^k \varphi(u) du$.

6 Limit theorems

The asymptotic properties of estimators are deduced from the asymptotic behavior of the contrast function. We therefore prepare some limit theorems for triangular arrays of the data, that we will prove in the Appendix.

Proposition 3. *Suppose that Assumptions 1 to 4 hold, $\Delta_n \to 0$ and $t_n \to \infty$.*
Moreover suppose that f is a differentiable function $\mathbb{R} \times \Theta \to \mathbb{R}$ such that $|f(x,\theta)| \le c(1+|x|)^c$, $|\partial_x f(x,\theta)| \le c(1+|x|)^c$ and $|\partial_\theta f(x,\theta)| \le c(1+|x|)^c$.
Then, $x \mapsto f(x,\theta)$ is a π-integrable function for any $\theta \in \Theta$ and the following convergence result holds as $n \to \infty$:
(i) $\sup_{\theta \in \Theta} |\frac{1}{t_n} \sum_{i=0}^{n-1} \Delta_{n,i} f(X_{t_i}, \theta) 1_{\{|X_{t_i}| \le \Delta_{n,i}^{-k}\}} - \int_{\mathbb{R}} f(x,\theta) \pi(dx)| \xrightarrow{\mathbb{P}} 0$,
(ii) $\sup_{\theta \in \Theta} |\frac{1}{t_n} \sum_{i=0}^{n-1} \Delta_{n,i} f(X_{t_i}, \theta) \varphi_{\Delta_{n,i}^{\beta}}(X_{t_{i+1}} - X_{t_i}) 1_{\{|X_{t_i}| \le \Delta_{n,i}^{-k}\}} - \int_{\mathbb{R}} f(x,\theta) \pi(dx)| \xrightarrow{\mathbb{P}} 0$.

The next proposition will be used in order to prove the consistency.
First, we prepare some notations. We define

$$\zeta_i := \int_{t_i}^{t_{i+1}} a(X_s)dW_s + \int_{t_i}^{t_{i+1}} \int_{\mathbb{R}\setminus\{0\}} \gamma(X_{s^-}) z\tilde{\mu}(ds,dz) + \Delta_{n,i} \int_{\mathbb{R}\setminus\{0\}} z\,\gamma(X_{t_i})\,[1-\varphi_{\Delta_{n,i}^\beta}(\gamma(X_{t_i})z)]\,F(z)dz. \tag{29}$$

We now observe that using the dynamic of the process X and the development (14) of m we get

$$X_{t_{i+1}} - m_\theta(X_{t_i}) + R(\theta, \Delta_{n,i}^{2-2\beta}, X_{t_i}) = (\int_{t_i}^{t_{i+1}} b(X_s,\theta_0)ds - \Delta_{n,i} b(X_{t_i},\theta)) + \zeta_i, \tag{30}$$

if $\alpha < 1$ and the same but with the different rest term $R(\theta, \Delta_{n,i}^{2-3\beta}, X_{t_i})$ if $\alpha \geq 1$. From the choice that we have made on α and β in Theorems 2 and 4, the exponent on $\Delta_{n,i}$ in the rest function is always more than 1. Hence, from now on, we will call it simply $R(\theta, \Delta_{n,i}^{1+\delta}, X_{t_i})$, with $\delta > 0$. That is the reason why we choose such a definition for ζ_i.

Proposition 4. *Suppose that Assumptions 1 to 4 and A_β hold, $\Delta_n \to 0$ and $t_n \to \infty$ and, $\forall i \in \{0,...,n-1\}$, $f_{i,n}$: $\mathbb{R}\times\Theta \to \mathbb{R}$. Moreover we suppose that $\exists c$: $|f_{i,n}(x,\theta)| \leq c(1+|x|^c)$ $\forall i,n$.*
Then, $\forall \theta \in \Theta$,

$$\frac{1}{t_n}\sum_{i=0}^{n-1} f_{i,n}(X_{t_i},\theta)\,\zeta_i\,\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}} - X_{t_i})1_{\{|X_{t_i}|\leq \Delta_{n,i}^{-k}\}} \xrightarrow{\mathbb{P}} 0.$$

The proof relies on the following lemma:

Lemma 3. *Suppose that Assumptions 1 to 4 and A_β hold. Then*

$$1. \quad \mathbb{E}[\zeta_i \varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}} - X_{t_i})1_{\{|X_{t_i}|\leq \Delta_{n,i}^{-k}\}}|\mathcal{F}_{t_i}] = R(\theta_0, \Delta_{n,i}^{(1+\delta)\wedge \frac{3}{2}}, X_{t_i}), \tag{31}$$

$$2. \quad \mathbb{E}[\zeta_i^2\,\varphi^2_{\Delta_{n,i}^\beta}(X_{t_{i+1}} - X_{t_i})1_{\{|X_{t_i}|\leq \Delta_{n,i}^{-k}\}}|\mathcal{F}_{t_i}] = R(\theta_0, \Delta_{n,i}, X_{t_i}), \tag{32}$$

and

$$3. \quad \mathbb{E}[(X_{t_{i+1}} - m_{\theta_0}(X_{t_i}))^2\,\varphi^2_{\Delta_{n,i}^\beta}(X_{t_{i+1}} - X_{t_i})1_{\{|X_{t_i}|\leq \Delta_{n,i}^{-k}\}}|\mathcal{F}_{t_i}] = R(\theta_0, \Delta_{n,i}, X_{t_i}), \tag{33}$$

where $(\mathcal{F}_s)_s$ is the filtration defined in Lemma 1 and δ is positive as defined above.

We now give an asymptotic normality result:

Proposition 5. *Suppose that Assumptions 1 to 4 and A_β hold, $\Delta_n \to 0$, $t_n \to \infty$.*
Moreover suppose that f is a continuous function $\Theta\times\mathbb{R}\to\mathbb{R}$ that satisfies conditions in Proposition 3.
Then for all θ

$$\frac{1}{\sqrt{t_n}}\sum_{i=0}^{n-1}(X_{t_{i+1}} - m_{\theta_0}(X_{t_i}))f(X_{t_i},\theta)\,\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}} - X_{t_i})1_{\{|X_{t_i}|\leq \Delta_{n,i}^{-k}\}} \xrightarrow{\mathcal{L}} N(0, \int_{\mathbb{R}} f^2(x,\theta)\,a^2(x)\,\pi(dx)).$$

7 Proof of main results

We state a proposition that will be used repeatedly in the proof of Theorems 1,2,3 and 4. This proposition is an estimation of some expectations related to the event that increments of the process X lies where $\varphi_{\Delta_{n,i}}$, that is the smooth version of the indicator function, becomes singular for $\Delta_n \to 0$. The proof is postponed to Section A.3.

Proposition 6. *Suppose that Assumptions 1 to 4 and A_β hold. Moreover suppose that $h: \mathbb{R}\times\Theta \longrightarrow \mathbb{R}$ is a function for which $\exists c > 0: \sup_{\theta\in\Theta}|h(x,\theta)| \leq c(1+|x|)^c$. Then $\forall k \geq 1$ $\forall \epsilon > 0$, we have*

$$\sup_{u\in[t_i,t_{i+1}]} \mathbb{E}[|h(X_u^\theta,\theta)|\,|\varphi^{(k)}_{\Delta_{n,i}^\beta}(X_u^\theta - X_{t_i}^\theta)|\,\big|X_{t_i}^\theta = x] = R(\theta, \Delta_{n,i}^{1-\alpha\beta-\epsilon}, x).$$

with α and β given in the third point of Assumption 4 and Definition 1. We have used $\varphi^{(k)}_{\Delta_{n,i}^\beta}(y)$ in order to denote $\varphi^{(k)}(\frac{y}{\Delta_{n,i}^\beta})\Delta_{n,i}^{-\beta}$.

Proposition 6 is a consequence of the following more general proposition:

Proposition 7. *Suppose that Assumption 1 to 4 and A_β hold. For $c>0$, we define*

$$\mathcal{Z}_{h,c,p} := \left\{ Z=(Z_\theta)_{\theta\in\Theta} \textit{family of random variables } \mathcal{F}_h \textit{ measurable such that } \sup_{\theta\in\Theta}\mathbb{E}[|Z_\theta|^p|X_0^\theta=x]\le c(1+|x|^c)\right\}.$$

Then $\forall k\ge 1$ we have, $\forall \epsilon \ge \frac{1}{p}$,

$$\sup_{Z\in\mathcal{Z}_{h,c,p}}\mathbb{E}[\,|Z_\theta|\,|\varphi_{h^\beta}^{(k)}(X_h^\theta - X_0^\theta)|\,|X_0^\theta = x]\le R(\theta, h^{(1-\alpha\beta)(1-\epsilon)}, x),$$

where $R(\theta,h^\delta,x)$ denotes any function such that $\exists c>0$: $|R(\theta,h^\delta,x)|\le c(1+|x|^c)h^\delta$ uniformly in θ, with c independent of h.

7.1 Development of $m_{\theta,\Delta_{n,i}}(x)$

In order to study the asymptotic behavior of the contrast function we need some explicit approximation of $m_{\theta,\Delta_{n,i}}$. We study the asymptotic expansion of $m_{\theta,\Delta_{n,i}}(x)$ as $\Delta_{n,i}\to 0$. The main tools is the iteration of the Dynkin's formula that provides us the following expansion for every function f: $\mathbb{R}\to\mathbb{R}$ such that f is in $C^{2(k+1)}$:

$$\mathbb{E}[f(X_{t_{i+1}}^\theta)|X_{t_i}^\theta=x] = \sum_{j=0}^k \frac{\Delta_{n,i}^j}{j!}A^j f(x) + \int_{t_i}^{t_{i+1}}\int_{t_i}^{u_1}\dots\int_{t_i}^{u_k}\mathbb{E}[A^{k+1}f(X_{u_{k+1}}^\theta)|X_{t_i}^\theta = x]\,du_{k+1}\dots du_2\,du_1 \tag{34}$$

where A denotes the generator of the diffusion. A is the sum of the continuous and discrete part: $A:=A_c+A_d$, with

$$A_c f(x) = \frac{1}{2}a^2(x)f''(x)+b(x,\theta)f'(x)$$

and

$$A_d f(x) = \int_{\mathbb{R}}(f(x+\gamma(x)z)-f(x)-z\gamma(x)f'(x))\,F(z)dz.$$

We set $A^0=Id$.

7.1.1 Proof of Theorem 1:

Proof. We have to show (12). Using the formula (34) in the case $k=1$, we get

$$\mathbb{E}[\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}^\theta - X_{t_i}^\theta)|X_{t_i}^\theta = x] =$$

$$= A^0\varphi_{\Delta_{n,i}^\beta}(0)+(t_{i+1}-t_i)A\varphi_{\Delta_{n,i}^\beta}(0)+\int_{t_i}^{t_{i+1}}\int_{t_i}^{u_1}\mathbb{E}[A^2\varphi_{\Delta_{n,i}^\beta}(X_{u_2}^\theta)|X_{t_i}^\theta = x]du_2du_1. \tag{35}$$

We have defined φ as a smooth version of the indicator function, it means that in a neighborhood of 0 its value is 1 and so that $\varphi^{(k)}(0)=0$ for each $k\ge1$.
We denote $f_{i,n}(y):=\varphi_{\Delta_{n,i}^\beta}(y-x)=\varphi(\frac{y-x}{\Delta_{n,i}^\beta})$, with $\beta\in(0,\frac{1}{2})$. By the building, $f_{i,n}(x)=1$ and $f_{i,n}^{(k)}(x)=0$ for each $k\ge 1$, so we get $A_c f_{i,n}(x)=0$ and $A_d f_{i,n}(x)=\int_{\mathbb{R}\setminus\{0\}}[f_{i,n}(x+\gamma(x)z)-1]\,F(z)dz$.
In the sequel the constant $c>0$ may change from line to line.
From the definition of $f_{i,n}$ and the fact that $\varphi=1$ on $[-1,1]$ we have that $f_{i,n}(y)=1$ for $|y-x|\le\Delta_{n,i}^\beta$. Thus

$$|A_d f_{i,n}(x)|\le 2\left\|\varphi_{\Delta_{n,i}^\beta}\right\|_\infty\int_{\{z:|z\gamma(x)|\ge\Delta_{n,i}^\beta\}}F(z)dz\le$$

$$\le 2\left\|\varphi_{\Delta_{n,i}^\beta}\right\|_\infty\int_{\left\{z:|z|\ge\frac{\Delta_{n,i}^\beta}{|\gamma(x)|}\right\}}|z|^{-1-\alpha}dz\le c\left\|\varphi_{\Delta_{n,i}^\beta}\right\|_\infty|\gamma(x)|^\alpha\Delta_{n,i}^{-\beta\alpha}=R(\theta,\Delta_{n,i}^{-\alpha\beta},x),$$

where the second inequality follows from point 3 of Assumption 4. Substituting in (35) we get

$$\mathbb{E}[\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}^\theta - X_{t_i}^\theta)|X_{t_i}^\theta = x]=1+\Delta_{n,i}R(\theta,\Delta_{n,i}^{-\alpha\beta},x)+\int_{t_i}^{t_{i+1}}\int_{t_i}^{u_1}\mathbb{E}[A^2\varphi_{\Delta_{n,i}^\beta}(X_{u_2}^\theta)|X_{t_i}^\theta=x]du_2du_1. \tag{36}$$

In order to prove (12), we want to show that the last term is negligible.
We consider the generator's decomposition in discrete and continuous part $A=A_c+A_d$ that yields:

$A^2 f_{i,n}(y) = (A_c^2 f_{i,n})(y) + A_c(A_d f_{i,n})(y) + A_d(A_c f_{i,n})(y) + (A_d^2 f_{i,n})(y)$.
We observe that we can write $(A_c^2 f_{i,n})(y)$ as

$$\sum_{j=1}^{4} \Delta_{n,i}^{-\beta j} h_j(y,\theta) \varphi^{(j)}_{\Delta_{n,i}^{\beta}}(y-x),$$

where $\varphi^{(j)}_{\Delta_{n,i}^{\beta}}(y-x) = \varphi^{(j)}(\frac{(y-x)}{\Delta_{n,i}^{\beta}})$. For each $j \in \{1,2,3,4\}$, h_j is a function of a, b and their derivatives up to second order: $h_1 = \frac{1}{2}a^2 b'' + bb'$, $h_2 = \frac{1}{2}a^2(a')^2 + \frac{1}{2}a^3 a'' + a^2 b' + aa'b + b^2$, $h_3 = a^3 a' + a^2 b$ and $h_4 = \frac{1}{4}a^4$.
Using the Proposition 6 we get that $\sup_{u_2 \in [t_i, t_{i+1}]} |\mathbb{E}[(A_c^2 f_{i,n})(X_{u_2}^\theta)|X_{t_i}^\theta = x]|$ is upper bounded by

$$\sup_{u_2 \in [t_i, t_{i+1}]} |\sum_{j=1}^{4} \Delta_{n,i}^{-\beta j}\, \mathbb{E}[h_j(X_{u_2}^\theta, \theta) \varphi^{(j)}_{\Delta_{n,i}^{\beta}}(X_{u_2}^\theta - X_{t_i}^\theta)|X_{t_i}^\theta = x]| =$$

$$= |\sum_{j=1}^{4} \Delta_{n,i}^{-\beta j} R(\theta, \Delta_{n,i}^{1-\alpha\beta-\epsilon}, x)| = R(\theta, \Delta_{n,i}^{1-\alpha\beta-\epsilon-4\beta}, x).$$

Let us now consider $A_c(A_d f_{i,n})(y)$. Substituting the definition of $A_d f_{i,n}$ we get

$$A_c(A_d f_{i,n})(y) = A_c(\int_{\mathbb{R}} g_n(\cdot, z) F(z) dz)(y), \tag{37}$$

where

$$g_n(y,z) := \varphi_{\Delta_{n,i}^{\beta}}(y - x + z\gamma(y)) - \varphi_{\Delta_{n,i}^{\beta}}(y-x) - \Delta_{n,i}^{-\beta} \varphi'_{\Delta_{n,i}^{\beta}}(y-x)\gamma(y) z \tag{38}$$

and where the notation used means that we are applying the differential operator A_c with respect to the variable represented with a dot. In order to estimate it we observe that

$$|g_n(y,z)| \le \Delta_{n,i}^{-\beta} \left\|\varphi'\right\|_\infty |z||\gamma(y)|, \tag{39}$$

$$|\frac{\partial}{\partial y} g_n(y,z)| \le \Delta_{n,i}^{-2\beta} P(y)|z| \qquad \text{and} \tag{40}$$

$$|\frac{\partial^2}{\partial y^2} g_n(y,z)| \le \Delta_{n,i}^{-3\beta} P(y)(|z| + |z|^2); \tag{41}$$

where $P(y)$ is a polynomial function in y, that may change from line to line.
Since the functions a^2 and b have polynomial growth, we obtain

$$|A_c g_n(\cdot, z)(y)| \le \Delta_{n,i}^{-3\beta} P(y)(|z| + |z|^2). \tag{42}$$

Using the dominated convergence theorem we get

$$A_c(\int_{\mathbb{R}} g_n(\cdot, z) F(z) dz)_{(y)} = \int_{\mathbb{R}} (A_c g_n)(\cdot, z)_{(y)} F(z) dz,$$

Therefore, using (42),

$$|A_c(\int_{\mathbb{R}} g_n(\cdot, z) F(z) dz)_{(y)}| \le \Delta_{n,i}^{-3\beta} P(y) \int_{\mathbb{R}} (|z| + |z|^2) F(z) dz,$$

that is upper bounded by $c\Delta_{n,i}^{-3\beta} P(y)$ since α is less than 1. It turns

$$\sup_{u_2 \in [t_i, t_{i+1}]} |\mathbb{E}[(A_c(A_d f_{i,n}))(X_{u_2}^\theta)|X_{t_i}^\theta = x]| \le \sup_{u_2 \in [t_i, t_{i+1}]} |\mathbb{E}[c\Delta_{n,i}^{-3\beta} P(X_{u_2}^\theta)|X_{t_i}^\theta = x]| = R(\theta, \Delta_{n,i}^{-3\beta}, x)$$

where, in the last equality, we have used the third point of Lemma 1.
We reason in the same way on $A_d(A_c f_{i,n})(y)$, which is equal to

$$\int_{\mathbb{R}} [A_c f_{i,n}(y + z\gamma(y)) - A_c f_{i,n}(y) - z\gamma(y)(A_c f_{i,n})'(y)] F(z) dz. \tag{43}$$

It is, in module, upper bounded by

$$c \int_0^1 \int_{\mathbb{R}} [|(A_c f_{i,n})'(y + z\gamma(y)s)| + |(A_c f_{i,n})'(y)|]|z||\gamma(y)| F(z) ds\, dz. \tag{44}$$

We observe that, $\forall y'$, $(A_c f_{i,n})'(y') = (b' f'_{i,n} + b f''_{i,n} + a a' f''_{i,n} + \frac{1}{2} a^2 f'''_{i,n})(y')$.
By the fact that $|\frac{\partial^j}{\partial y^j} \varphi_{\Delta^\beta_{n,i}}(y)| \le c \Delta_{n,i}^{-\beta j}$ for $j = 1, 2, 3$ and recalling $f_{i,n}(y) = \varphi_{\Delta^\beta_{n,i}}(y - x)$, we get that

$$|(A_c f_{i,n})'(y')| \le c\, P(y') \Delta_{n,i}^{-3\beta}, \tag{45}$$

where we have used that b and a^2 have polynomial growth. We obtain that (44) is upper bounded by

$$\Delta_{n,i}^{-3\beta} \int_0^1 \int_{\mathbb{R}} (P(y + z\gamma(y)s) + P(y)) |z| |\gamma(y)| F(z) ds\, dz \le \Delta_{n,i}^{-3\beta} \int_{\mathbb{R}} P(y) P(z) |z| F(z) dz \le c \Delta_{n,i}^{-3\beta} P(y),$$

where we have used the first point of Assumptions 3 and the third of Assumption 4, with $\alpha \in (0,1)$, in order to get $\int_{\mathbb{R}} P(z)|z| F(z) dz \le \infty$.
Considering the controls (44) and (45) on (43) it yields, using again the third point of Lemma 1,

$$\sup_{u_2 \in [t_i, t_{i+1}]} |\mathbb{E}[(A_d(A_c f_{i,n}))(X^\theta_{u_2}) | X^\theta_{t_i} = x]| = R(\theta, \Delta_{n,i}^{-3\beta}, x).$$

To conclude, we consider $A_d(A_d f_{i,n})(y)$:

$$\int_{\mathbb{R}} [A_d f_{i,n}(y + z\gamma(y)) - A_d f_{i,n}(y) - z\gamma(y)(A_d f_{i,n})'(y)] F(z) dz. \tag{46}$$

Again, (46) is, in module, upper bounded by

$$c \int_0^1 \int_{\mathbb{R}} [|(A_d f_{i,n})'(y - x + z\gamma(y)s)| + |(A_d f_{i,n})'(y)|] |z| |\gamma(y)| F(z) ds\, dz \tag{47}$$

But

$$A_d f_{i,n}(y') = \int_{\mathbb{R}} g_n(y', z) F(z) dz, \tag{48}$$

with $g_n(y', z)$ given in (38) Using control equation (40) and dominated convergence theorem, we get that its derivative is upper bounded by $c \Delta_{n,i}^{-2\beta} P(y')$.
Using also (46) and (47),

$$|A_d^2 f_{i,n}(y)| \le \Delta_{n,i}^{-2\beta} P(y) \int_{\mathbb{R}} |z| F(z) dz$$

and it turns, using third point of Lemma 1,

$$\sup_{u_2 \in [t_i, t_{i+1}]} |\mathbb{E}[(A_d^2 f_{i,n})(X^\theta_{u_2}) | X^\theta_{t_i} = x]| = R(\theta, \Delta_n^{-2\beta}, x).$$

By the decomposition of the generator in A_c and A_d we get

$$\sup_{u_2 \in [t_i, t_{i+1}]} |\mathbb{E}[A^2 f_{i,n}(X^\theta_{u_2}) | X^\theta_{t_i} = x]| = R(\theta, \Delta_{n,i}^{1 - \alpha\beta - 4\beta - \epsilon}, x) + R(\theta, \Delta_{n,i}^{-3\beta}, x) + R(\theta, \Delta_{n,i}^{-2\beta}, x),$$

with $\alpha \in (0,1)$ and $\beta \in (0, \frac{1}{2})$, so it is $R(\theta, \Delta_{n,i}^{-3\beta}, x)$, since the other R functions are always negligible compared to it.
Using (36) we get

$$\mathbb{E}[\varphi_{\Delta^\beta_{n,i}}(X^\theta_{t_{i+1}} - X^\theta_{t_i}) | X^\theta_{t_i} = x] = 1 + \Delta_{n,i} R(\theta, \Delta_{n,i}^{-\alpha\beta}, x) + \frac{\Delta_{n,i}^2}{2} R(\theta, \Delta_{n,i}^{-3\beta}, x).$$

We deduce, using the definition of $\Delta_{n,i}$ and (11), that it is

$$1 + R(\theta, \Delta_{n,i}^{1-\alpha\beta}, x) + R(\theta, \Delta_{n,i}^{2-3\beta}, x) = 1 + R(\theta, \Delta_{n,i}^{(1-\alpha\beta)\wedge(2-3\beta)}, x),$$

as we wanted. □

7.2 Proof of Theorem 3

Proof. Let α now be in $[1, 2)$. In the sequel we skip the study of the case $\alpha = 1$ for simplicity, in order to avoid the appearance of logarithmic functions. However, such a specific case is embedded in the case $\alpha > 1$ by taking $\alpha = 1 + \epsilon$ with a choice of $\epsilon > 0$ arbitrarily small.
Using again Dynkin formula, we have that (36) is still true. Considering the generator's decomposition, we act like in the case where α is less than 1 to get that

$$\sup_{u_2 \in [t_i, t_{i+1}]} |\mathbb{E}[(A_c^2 f_{i,n})(X^\theta_{u_2}) | X^\theta_{t_i} = x]| = R(\theta, \Delta_{n,i}^{1 - \alpha\beta - \epsilon - 4\beta}, x). \tag{49}$$

Concerning $A_c(A_d f_{i,n})(y)$, we use (37) with g_n defined in (38). Using Taylor development to the second order we get

$$|g_n(y,z)| \leq \left\| \varphi''_{\Delta^\beta_{n,i}} \right\|_\infty |\Delta_{n,i}|^{-2\beta} \frac{|z|^2\gamma(y)^2}{2}. \tag{50}$$

In the same way we get the following two estimations:

$$|\frac{\partial}{\partial y} g_n(y,z)| \leq |\Delta_{n,i}|^{-2\beta} \left\| \varphi''_{\Delta^\beta_{n,i}} \right\|_\infty |\gamma(y)||\gamma'(y)||z|^2 + |\frac{\Delta_{n,i}|^{-3\beta}}{2} \left\| \varphi'''_{\Delta^\beta_{n,i}} \right\|_\infty |z|^2\gamma^2(y)|1+\gamma'(y)z|,$$

$$|\frac{\partial^2}{\partial y^2} g_n(y,z)| \leq |\Delta_{n,i}|^{-2\beta}|z|^2 P(y) + |\Delta_{n,i}|^{-3\beta}|P(y)(|z|^2+|z|^3) + |\Delta_{n,i}|^{-4\beta} P(y)(|z|^2+|z|^3). \tag{51}$$

Since a^2 and b have polynomial growth, (51) provides us an estimation on $|A_c g_n(\cdot,z)(y)|$. Using dominated convergence theorem, (37), the estimation of $|A_c g_n(\cdot,z)(y)|$ obtained from (51) and the fact that $\int_{\mathbb{R}}(|z|^2+|z|^3)F(z)dz < \infty$, we get

$$\sup_{u_2\in[t_i,t_{i+1}]} |\mathbb{E}[(A_c A_d f_{i,n})(X^\theta_{u_2})|X^\theta_{t_i}=x]| = R(\theta,\Delta^{-2\beta}_{n,i},x)+R(\theta,\Delta^{-3\beta}_{n,i},x)+R(\theta,\Delta^{-4\beta}_{n,i},x) = R(\theta,\Delta^{-4\beta}_{n,i},x). \tag{52}$$

We now consider $A_d(A_c f_{i,n})(y)$. Using (43) and the development to the second order of the function $A_c f_{i,n}(y+z\gamma(y))$ we obtain

$$|A_d(A_c f_{i,n})(y)| \leq c\int_{\mathbb{R}}\int_0^1 |(A_c f_{i,n})''(y+s\,z\gamma(y))||z|^2|\gamma^2(y)|F(z)dsdz. \tag{53}$$

We observe that $(A_c f_{i,n})''(y') = [b''f'_{i,n} + 2b'f''_{i,n} + bf'''_{i,n} + (a')^2 f''_{i,n} + a(a''f''_{i,n} + a'f'''_{i,n}) + 2aa'f'''_{i,n} + \frac{1}{2}a^2 f^{(4)}_{i,n}](y')$. By the fact that $|\frac{\partial^j}{\partial y^j}\varphi_{\Delta^\beta_{n,i}}(y)| \leq c\Delta^{-\beta j}_{n,i}$ for $j=1,2,3$ and recalling $f_{i,n}(y)=\varphi_{\Delta^\beta_{n,i}}(y-x)$, we get that

$$|(A_c f_{i,n})''(y')| \leq c\,P(y')\Delta^{-4\beta}_{n,i}. \tag{54}$$

Using (53) and (54) it yields

$$\sup_{u_2\in[t_i,t_{i+1}]} |\mathbb{E}[(A_d A_c f_{i,n})(X^\theta_{u_2})|X^\theta_{t_i}=x]| = R(\theta,\Delta^{-4\beta}_{n,i},x). \tag{55}$$

To conclude, we consider $A_d A_d f_{i,n}$. Using (46) and the development up to the second order we get

$$|A_d(A_d f_{i,n})(y)| \leq c\int_{\mathbb{R}}\int_0^1 |(A_d f)''(y+s\,z\gamma(y))||z|^2|\gamma^2(y)|F(z)dsdz.$$

We recall that (48) still holds, with g_n defined in (38). In order to estimate $(A_d f)''(y)$ in the case where $\alpha\in[1,2)$ we use therefore (51) joint with dominated convergence theorem. It provides us

$$\sup_{u_2\in[t_i,t_{i+1}]} |\mathbb{E}[(A_d A_d f_{i,n})(X^\theta_{u_2})|X^\theta_{t_i}=x]| = R(\theta,\Delta^{-4\beta}_{n,i},x). \tag{56}$$

Using (49), (52), (55) and (56) we put the pieces together and so we obtain

$$\sup_{u_2\in[t_i,t_{i+1}]} |\mathbb{E}[A^2 f_{i,n}(X^\theta_{u_2})|X^\theta_{t_i}=x]| = R(\theta,\Delta^{1-\alpha\beta-4\beta-\epsilon}_{n,i},x)+R(\theta,\Delta^{-4\beta}_{n,i},x).$$

We replace it in the Dynkin formula (36) getting

$$\mathbb{E}[\varphi_{\Delta^\beta_{n,i}}(X^\theta_{t_{i+1}}-X^\theta_{t_i})|X^\theta_{t_i}=x] = 1+\Delta_{n,i}R(\theta,\Delta^{-\alpha\beta}_{n,i},x)+\frac{\Delta^2_{n,i}}{2}R(\theta,\Delta^{(1-\alpha\beta-4\beta-\epsilon)\wedge(-4\beta)}_{n,i},x).$$

Using the definition of $\Delta_{n,i}$ and (11) it is

$$1+R(\theta,\Delta^{(1-\alpha\beta)\wedge(3-\alpha\beta-4\beta-\epsilon)\wedge(2-4\beta)}_{n,i},x). \tag{57}$$

Since ϵ is arbitrarily small, for each choice of α and β there exists ϵ such that $3-\alpha\beta-4\beta-\epsilon$ is greater than $2-4\beta$ and (15) follows. □

7.3 Proof of Theorem 2

Proof. We observe that

$$m_{\theta,\Delta_{n,i}}(x) := \frac{\mathbb{E}[X^\theta_{t_{i+1}}\varphi_{\Delta^\beta_{n,i}}(X^\theta_{t_{i+1}} - X^\theta_{t_i})|X^\theta_{t_i} = x]}{\mathbb{E}[\varphi_{\Delta^\beta_{n,i}}(X^\theta_{t_{i+1}} - X^\theta_{t_i})|X^\theta_{t_i} = x]} = x + \frac{\mathbb{E}[g_{i,n}(X^\theta_{t_{i+1}})|X^\theta_{t_i} = x]}{\mathbb{E}[\varphi_{\Delta^\beta_{n,i}}(X^\theta_{t_{i+1}} - X^\theta_{t_i})|X^\theta_{t_i} = x]}, \tag{58}$$

with $g_{i,n}(y) = (y-x)\varphi_{\Delta^\beta_{n,i}}(y-x)$.

We have already found a development for the denominator of (58) given by (12), we use again the Dynkin's formula (34) for $k = 1$ in order to find a development for the numerator. By the building, $g_{i,n}(x) = 0$, $g'_{i,n}(x) = 1$ and $g''_{i,n}(x) = 0$, so we get

$$A_c g_{i,n}(x) = b(x,\theta)$$

and

$$A_d g_{i,n}(x) = \int_{\mathbb{R}\setminus\{0\}}[g_{i,n}(x + z\gamma(x)) - z\gamma(x)]F(z)dz = \int_{\mathbb{R}\setminus\{0\}} z\gamma(x)[\varphi_{\Delta^\beta_{n,i}}(z\gamma(x)) - 1]F(z)dz$$

where we have used, in the last equality, simply the definition of $g_{i,n}$.
Substituting in the Dynkin's formula we get

$$\mathbb{E}[g_{i,n}(X^\theta_{t_{i+1}})|X^\theta_{t_i} = x] = \Delta_{n,i}(b(x,\theta) + \int_{\mathbb{R}\setminus\{0\}} z\gamma(x)[\varphi_{\Delta^\beta_{n,i}}(z\gamma(x)) - 1]F(z)dz) +$$

$$+\int_{t_i}^{t_{i+1}}\int_{t_i}^{u_1} \mathbb{E}[A^2 g_{i,n}(X_{u_2})|X_{t_i} = x]du_2 du_1. \tag{59}$$

In order to show that the last term is negligible, we have to estimate $(A^2 g_{i,n})(y)$ using the decomposition in continuous and discrete part of the generator, as we have already done.
Since $g_{i,n}(y) = (y-x)\varphi_{\Delta^\beta_{n,i}}(y-x)$, we have

$$g^{(h)}_{i,n}(y) = \sum_{k=0}^h \binom{h}{k}\frac{\partial^k}{\partial y^k}(y-x)\frac{\partial^{h-k}}{\partial y^{h-k}}(\varphi_{\Delta^\beta_{n,i}}(y-x)),$$

with $\binom{h}{k}$ binomial coefficients. So we get, observing that the derivatives of $(y-x)$ after the second order are zero, the following useful control for $h \ge 1$:

$$|g^{(h)}_{i,n}(y)| \le |\varphi^{(h)}_{\Delta^\beta_{n,i}}(y-x)|\Delta_{n,i}^{-\beta h}|y-x| + |\varphi^{(h-1)}_{\Delta^\beta_{n,i}}(y-x)|\Delta_{n,i}^{-\beta(h-1)}|h|. \tag{60}$$

By the definition of φ as a smooth version of the indicator function, we know that it exists $c > 0$ such that if $\frac{|y-x|}{\Delta^\beta_{n,i}} > c$, then φ and its derivatives are zero when evaluated at the point $\frac{(y-x)}{\Delta^\beta_{n,i}}$.
So we can say that $|\varphi^{(h)}_{\Delta^\beta_{n,i}}(y-x)||y-x| \le c|\varphi^{(h)}_{\Delta^\beta_{n,i}}(y-x)|\Delta^\beta_{n,i}$ and consequently

$$|g^{(h)}_{i,n}(y)| \le c|\varphi^{(h)}_{\Delta^\beta_{n,i}}(y-x)|\Delta_{n,i}^{-\beta(h-1)} + c|\varphi^{(h-1)}_{\Delta^\beta_{n,i}}(y-x)|\Delta_{n,i}^{-\beta(h-1)}. \tag{61}$$

Reasoning as in the proof of Theorem 1, we start with $(A_c^2 g_{i,n})(y)$ and we get that it is $\sum_{j=1}^4 h_j(y,\theta)g^{(j)}_{i,n}(y)$ where again, for each $j \in \{1,2,3,4\}$, h_j is a function of a, b and their derivatives up to second order.
We substitute in $\mathbb{E}[(A_c^2 g_{i,n})(X^\theta_{u_2})|X^\theta_{t_i} = x]$, getting $\sum_{j=1}^4 \mathbb{E}[h_j(X^\theta_{u_2},\theta)g^{(j)}_{i,n}(X^\theta_{u_2})|X^\theta_{t_i} = x]$. Using the estimation (60) we obtain

$$\sup_{u_2\in[t_i,t_{i+1}]}|\mathbb{E}[|A_c^2 g_{i,n}(X^\theta_{u_2})||X^\theta_{t_i} = x]| \le$$

$$\le \sup_{u_2\in[t_i,t_{i+1}]}|\sum_{j=1}^4 c\Delta_{n,i}^{-\beta(j-1)}\mathbb{E}[|h_j(X^\theta_{u_2},\theta)|(|\varphi^{(j)}_{\Delta^\beta_{n,i}}(X^\theta_{u_2} - X^\theta_{t_i})| + |\varphi^{(j-1)}_{\Delta^\beta_{n,i}}(X^\theta_{u_2} - X^\theta_{t_i})|)|X^\theta_{t_i} = x]|.$$

We observe that we can see $\sup_{u_2\in[t_i,t_{i+1}]}|\mathbb{E}[|h_1(X^\theta_{u_2},\theta)||\varphi_{\Delta^\beta_{n,i}}(X^\theta_{u_2} - X^\theta_{t_i})||X^\theta_{t_i} = x]|$ as $R(\theta,\Delta^0_{n,i},x) = R(\theta,1,x)$ and we use the Proposition 6 on the other terms, getting

$$\sup_{u_2\in[t_i,t_{i+1}]}|\mathbb{E}[|A_c^2 g_{i,n}(X^\theta_{u_2})||X^\theta_{t_i} = x]| \le R(\theta,\Delta_{n,i}^{1-\alpha\beta-\epsilon},x) + R(\theta,1,x) + \sum_{j=2}^4 c\Delta_{n,i}^{-\beta(j-1)}R(\theta,\Delta_{n,i}^{1-\alpha\beta-\epsilon},x) =$$

$$= R(\theta,\Delta_{n,i}^{1-\alpha\beta-\epsilon-3\beta},x) + R(\theta,1,x). \tag{62}$$

Let us now consider $A_c(A_d g_{i,n})(y)$

$$A_c(A_d g_{i,n})(y) = A_c(\int_{\mathbb{R}} [g_{i,n}(\cdot + z\gamma(\cdot)) - g_{i,n}(\cdot) - z\gamma(\cdot)g'_{i,n}(\cdot)]F(z)dz)_{(y)}. \tag{63}$$

Let us denote

$$h_{i,n}(y,z) := g_{i,n}(y + z\gamma(y)) - g_{i,n}(y) - z\gamma(y)g'_{i,n}(y). \tag{64}$$

We observe that

$$\frac{\partial h_{i,n}}{\partial y}(y,z) = g'_{i,n}(y + z\gamma(y)) - g'_{i,n}(y) + z\gamma'(y)(g'_{i,n}(y + z\gamma(y)) - g'_{i,n}(y)) - z\gamma(y)g''_{i,n}(y), \tag{65}$$

$$\frac{\partial^2 h_{i,n}}{\partial y^2}(y,z) = g''_{i,n}(y + z\gamma(y))(1 + z\gamma'(y))^2 + g'_{i,n}(y + z\gamma(y))z\gamma''(y) + \\ - g''_{i,n}(y) - g'''_{i,n}(y)\gamma(y)z - 2g''_{i,n}(y)z\gamma'(y) - g'_{i,n}(y)z\gamma''(y). \tag{66}$$

Using the estimation (61), we have

$$|g'_{i,n}(y)| \le c \qquad |g''_{i,n}(y)| \le c\Delta_{n,i}^{-\beta} \qquad |g'''_{i,n}(y)| \le c\Delta_{n,i}^{-2\beta} \tag{67}$$

Hence

$$|\frac{\partial h_{i,n}}{\partial y}(y,z)| \le \left\| g''_{i,n} \right\|_\infty P(y)(|z| + |z|^2) \le (|z| + |z|^2)P(y)\Delta_{n,i}^{-\beta}, \tag{68}$$

and similarly

$$|\frac{\partial^2 h_{i,n}}{\partial y^2}(y,z)| \le \Delta_{n,i}^{-2\beta} P(y)(|z| + |z|^2 + |z|^3).$$

Since functions a^2 and b have polynomial growth, we obtain

$$|A_c h_{i,n}(y,z)| \le \Delta_{n,i}^{-2\beta} P(y)(|z| + |z|^2 + |z|^3).$$

Using dominated convergence theorem we get

$$|A_c(\int_{\mathbb{R}} h_{i,n}(\cdot, z)F(z)dz)_{(y)}| \le \Delta_{n,i}^{-2\beta} P(y) \int_{\mathbb{R}} (|z| + |z|^2 + |z|^3)F(z)dz$$

and so, using also the third point of Lemma 1 and (63), we get

$$\sup_{u_2 \in [t_i, t_{i+1}]} |\mathbb{E}[A_c(A_d g_{n,i})(X^\theta_{u_2})|X^\theta_{t_i} = x]| = R(\theta, \Delta_{n,i}^{-2\beta}, x). \tag{69}$$

We reason on the same way on $A_d(A_c g_{n,i})(y)$:

$$A_d(A_c g_{n,i})(y) = \int_{\mathbb{R}} [A_c g_{n,i}(y + z\gamma(y)) - A_c g_{n,i}(y) - z\gamma(y)(A_c g_{n,i})'(y)]F(z)dz. \tag{70}$$

It is, in module, upper bounded by

$$c\int_0^1 \int_{\mathbb{R}} [|(A_c g_{n,i})'(y + z\gamma(y)s)| + |(A_c g_{n,i})'(y)|]|z||\gamma(y)|F(z)dsdz.$$

In order to estimate it we observe that, $\forall y'$,

$$(A_c g_{n,i})'(y') = (aa'g''_{n,i} + \frac{1}{2}a^2 g'''_{n,i} + b'g'_{n,i} + bg''_{n,i})(y').$$

Using (67) and the polynomial growth of a, b and their derivatives, we get

$$|(A_c g_{n,i})'(y')| \le c + P(y')\Delta_{n,i}^{-\beta} + P(y')\Delta_{n,i}^{-2\beta} \le P(y')\Delta_{n,i}^{-2\beta}.$$

It yields

$$c\int_0^1 \int_{\mathbb{R}} [|(A_c g_{n,i})'(y + z\gamma(y)s)| + |(A_c g_{n,i})'(y)|]|z||\gamma(y)|F(z)dsdz \le$$

$$\le \Delta_{n,i}^{-2\beta} \int_0^1 \int_{\mathbb{R}} (P(y + z\gamma(y)s) + P(y)]|z||\gamma(y)|F(z)dsdz \le \Delta_{n,i}^{-2\beta} \int_{\mathbb{R}} P(y)P(z)|z|F(z)dz \le c\Delta_{n,i}^{-2\beta} P(y),$$

where we have used the first point of Assumptions 3 and the second of Assumption 4. Hence $|A_d(A_c g)(y)| \leq \Delta_{n,i}^{-2\beta} P(y)$.
Taking the expected value and using the third point of Lemma 1, we obtain

$$\sup_{u_2 \in [t_i, t_{i+1}]} |\mathbb{E}[A_d(A_c g_{n,i})(X_{u_2}^\theta)|X_{t_i}^\theta = x]| = R(\theta, \Delta_{n,i}^{-2\beta}, x).$$

In conclusion, we consider $A_d^2(g_{n,i})(y)$

$$A_d^2(g_{n,i})(y) = \int_{\mathbb{R}} [A_d g_{n,i}(y + z\gamma(y)) - A_d g_{n,i}(y) - z\gamma(y)(A_d g_{n,i})'(y)] F(z) dz. \tag{71}$$

Again it is, in module, upper bounded by

$$c \int_0^1 \int_{\mathbb{R}} [|(A_d g_{n,i})'(y + z\gamma(y)s)| + |(A_d g_{n,i})'(y)|] |z| |\gamma(y)| F(z) ds\, dz \tag{72}$$

But

$$A_d g_{n,i}(y') = \int_{\mathbb{R}} [g_{n,i}(y' + z\gamma(y')) - g_{n,i}(y') - z\gamma(y') g'_{n,i}(y')] F(z) dz = \int_{\mathbb{R}} h_{i,n}(y', z) F(z) dz, \tag{73}$$

with $h_{i,n}$ defined in (64). Using control equation (68) and dominated convergence theorem, we get that (73) is upper bounded by $P(y')\Delta_{n,i}^{-\beta}$.
It follows from (71) and (72) that

$$|A_d^2 g_{n,i}(y)| \leq c \Delta_{n,i}^{-\beta} P(y) \int_{\mathbb{R}} P(z) F(z) dz$$

and it turns, using again the third point of Lemma 1,

$$\sup_{u_2 \in [t_i, t_{i+1}]} |\mathbb{E}[(A_d^2 g_{n,i})(X_{u_2}^\theta)|X_{t_i}^\theta = x]| = R(\theta, \Delta_{n,i}^{-\beta}, x).$$

Pieces things together we get

$$\sup_{u_2 \in [t_i, t_{i+1}]} |\mathbb{E}[A^2 g_{n,i}(X_{u_2}^\theta)|X_{t_i}^\theta = x]| = R(\theta, \Delta_{n,i}^{1-\alpha\beta-\epsilon-3\beta}, x) + R(\theta, \Delta_{n,i}^{-2\beta}, x) + R(\theta, \Delta_{n,i}^{-\beta}, x) =$$

$$= R(\theta, \Delta_{n,i}^{-2\beta}, x),$$

where $R(\theta, \Delta_{n,i}^{1-\alpha\beta-\epsilon-3\beta}, x)$ is negligible compared to $R(\theta, \Delta_{n,i}^{-2\beta}, x)$ because, for each choice of α and β, we can find an ϵ arbitrarily small such that $1 - \alpha\beta - \epsilon - \beta$ is more than 0. We substitute it in Dynkin's formula and we obtain

$$\mathbb{E}[g_{n,i}(X_{t_{i+1}}^\theta)|X_{t_i}^\theta = x] =$$

$$= \Delta_{n,i}(b(x,\theta) + \int_{\mathbb{R}\setminus\{0\}} z\gamma(x)[\varphi_{\Delta_{n,i}^\beta}(z\gamma(x)) - 1] F(z) dz) + \frac{\Delta_{n,i}^2}{2} R(\theta, \Delta_{n,i}^{-2\beta}, x). \tag{74}$$

We use the definition of $\Delta_{n,i}$ and the property (11) on R, then we substitute in (74) getting (13).
We now want to prove (14). From the expansion (13) and the property (10) of R, there exists $k_0 > 0$ such that for $|x| \leq \Delta_{n,i}^{k_0}$, $\mathbb{E}[\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}^\theta - X_{t_i}^\theta)|X_{t_i}^\theta = x] \geq \frac{1}{2}$ $\forall n, i \leq n$: we are avoiding the possibility that the denominator is in the neighborhood of 0. Using (58), (74) and (12) we have that

$$m_\theta(x) = x + \frac{\Delta_{n,i}(b(x,\theta) + \int_{\mathbb{R}\setminus\{0\}} z\gamma(x)[\varphi_{\Delta_{n,i}^\beta}(z\gamma(x)) - 1] F(z) dz) + R(\theta, \Delta_{n,i}^{2-2\beta}, x)}{1 + R(\theta, \Delta_{n,i}^{(1-\alpha\beta)\wedge(2-3\beta)}, x)}. \tag{75}$$

Now we can use that R in the denominator is a rest function and so we obtain

$$\frac{1}{1 + R(\theta, \Delta_{n,i}^{(1-\alpha\beta)\wedge(2-3\beta)}, x)} \sim 1 - R(\theta, \Delta_{n,i}^{(1-\alpha\beta)\wedge(2-3\beta)}, x). \tag{76}$$

Replacing (76) in (75) we get

$$m_\theta(x) = x + [\Delta_{n,i}(b(x,\theta) + \int_{\mathbb{R}\setminus\{0\}} z\gamma(x)[\varphi_{\Delta_{n,i}^\beta}(z\gamma(x)) - 1] F(z) dz) + R(\theta, \Delta_{n,i}^{2-2\beta}, x)](1 - R(\theta, \Delta_{n,i}^{(1-\alpha\beta)\wedge(2-3\beta)}, x)).$$

The expansion (14) follows. □

7.4 Proof of Theorem 4

Proof. Let us now consider an expansion of (58) in the case where α is in $[1, 2)$. Again, we skip the study of $\alpha = 1$ to avoid the emergence of logarithmic functions; as it is embedded in the study of $\alpha > 1$ with the choice of α arbitrarily close to 1.
We start observing that (59) and (61) still hold; we want to show that even in this case the last term of (59) is negligible compared to the others. Again, we consider its decomposition in continuous and discrete part.
Concerning $A_c^2 g_{i,n}$, (62) is still true. Let us now consider $A_c(A_d g_{i,n})(y)$ as written in (63). We act as in the proof of Theorem 3, using Taylor development up to second order, on the function $h_{i,n}$ defined in (64). Hence we obtain the following estimation:

$$|h_{i,n}(y,z)| \leq \left\|g''_{i,n}\right\|_\infty \frac{|z|^2\gamma(y)^2}{2}$$

and in the same way, using also (67),

$$|\frac{\partial h_{i,n}}{\partial y}(y,z)| \leq \left\|g''_{i,n}\right\|_\infty |z|^2|\gamma(y)\gamma'(y)| + \left\|g'''_{i,n}\right\|_\infty |z|^2\frac{\gamma^2(y)}{2}|1+\gamma'(y)z| \leq$$

$$\leq |z|^2 P(y)|\Delta_{n,i}|^{-\beta} + |\Delta_{n,i}|^{-2\beta}P(y)(|z|^2+|z|^3), \tag{77}$$

$$|\frac{\partial^2 h_{i,n}}{\partial y^2}(y,z)| \leq |\Delta_{n,i}^{-\beta}||z|^2 P(y) + |\Delta_{n,i}|^{-2\beta}P(y)(|z|^2+|z|^3) + |\Delta_{n,i}|^{-3\beta}P(y)(|z|^2+|z|^3). \tag{78}$$

Since a^2 and b have polynomial growth, (78) provides us an estimation on $|A_c h_{i,n}(\cdot,z)(y)|$. Using dominated convergence theorem, (63), the estimation of $|A_c h_{i,n}(\cdot,z)(y)|$ obtained from (78) and the fact that both $\int_{\mathbb{R}}(|z|^2+|z|^3)F(z)dz$ and $\int_{\mathbb{R}}(|z|^2+|z|^3)F(z)dz$ are finite, we get

$$\sup_{u_2\in[t_i,t_{i+1}]}|\mathbb{E}[A_c(A_d g_{n,i})(X^\theta_{u_2})|X^\theta_{t_i}=x]| = R(\theta,\Delta_{n,i}^{-\beta},x) + R(\theta,\Delta_{n,i}^{-2\beta},x) + R(\theta,\Delta_{n,i}^{-3\beta},x) = R(\theta,\Delta_{n,i}^{-3\beta},x). \tag{79}$$

We now consider $A_d(A_c g_{i,n})(y)$. Using (70) and the development to the second order of the function $A_c g_{i,n}(y+z\gamma(y))$ we obtain

$$|A_d(A_c g_{i,n})(y)| \leq c\int_{\mathbb{R}}\int_0^1 |(A_c g_{i,n})''(y+s\,z\gamma(y))||z|^2|\gamma^2(y)|F(z)dsdz. \tag{80}$$

We observe that $(A_c g_{i,n})''(y) = [b'' g'_{i,n} + 2b' g''_{i,n} + b g'''_{i,n} + (a')^2 g''_{i,n} + a(a'' g''_{i,n} + a' g'''_{i,n}) + 2aa' g'''_{i,n} + \frac{1}{2}a^2 g^{(4)}_{i,n}](y)$. Using (67), to which we add $|g^{(4)}_{i,n}(y)| \leq c\Delta_{n,i}^{-3\beta}$, we get

$$|(A_c g_{i,n})''(y)| \leq c\,P(y)\Delta_{n,i}^{-3\beta}. \tag{81}$$

Using (80) and (81) it yields

$$\sup_{u_2\in[t_i,t_{i+1}]}|\mathbb{E}[(A_d A_c g_{i,n})(X^\theta_{u_2})|X^\theta_{t_i}=x]| = R(\theta,\Delta_{n,i}^{-3\beta},x). \tag{82}$$

To conclude, we consider $A_d A_d g_{i,n}$. Using (71) and the development up to the second order we get

$$|A_d(A_d g_{i,n})(y)| \leq c\int_{\mathbb{R}}\int_0^1 |(A_d g_{i,n})''(y+s\,z\gamma(y))||z|^2|\gamma^2(y)|F(z)dsdz.$$

We recall that (73) still holds, with $h_{i,n}$ defined in (64). In order to estimate $(A_d g_{i,n})''(y)$ in the case where $\alpha \in [1,2)$ we use therefore (78) joint with dominated convergence theorem. It provides us

$$\sup_{u_2\in[t_i,t_{i+1}]}|\mathbb{E}[(A_d A_d g_{i,n})(X^\theta_{u_2})|X^\theta_{t_i}=x]| = R(\theta,\Delta_{n,i}^{-3\beta},x). \tag{83}$$

Using (62), (79), (82) and (83) we put the pieces together and so we obtain

$$\sup_{u_2\in[t_i,t_{i+1}]}|\mathbb{E}[A^2 f_{i,n}(X^\theta_{u_2})|X^\theta_{t_i}=x]| = R(\theta,\Delta_{n,i}^{1-\alpha\beta-3\beta-\epsilon},x) + R(\theta,1,x) + R(\theta,\Delta_{n,i}^{-3\beta},x) = R(\theta,\Delta_{n,i}^{-3\beta},x).$$

Indeed, since ϵ is arbitrarily small, for each choice of α and β we can find ϵ such that $1-\alpha\beta-3\beta-\epsilon > -3\beta$. We substitute in the Dynkin formula (59) and so we get

$$\mathbb{E}[g_{n,i}(X^\theta_{t_{i+1}})|X^\theta_{t_i}=x] =$$

$$= \Delta_{n,i}(b(x,\theta) + \int_{\mathbb{R}\setminus\{0\}} z\gamma(x)[\varphi_{\Delta_{n,i}^{\beta}}(z\gamma(x)) - 1]F(z)dz + \frac{\Delta_{n,i}^2}{2} R(\theta, \Delta_{n,i}^{-3\beta}, x). \quad (84)$$

We use the definition of $\Delta_{n,i}$ and the property (11) on R, then we substitute in (84) getting (16).
In order to prove (17), we observe again that from the expansion (16) and the property (10) of R, there exists $k_0 > 0$ such that for $|x| \le \Delta_{n,i}^{k_0}$, $\mathbb{E}[\varphi_{\Delta_{n,i}^{\beta}}(X_{t_{i+1}}^{\theta} - X_{t_i}^{\theta})|X_{t_i}^{\theta} = x] \ge \frac{1}{2}$ $\forall n, i \le n$. Using (58), (84) and (15) we have that

$$m_\theta(x) = x + \frac{\Delta_{n,i}(b(x,\theta) + \int_{\mathbb{R}\setminus\{0\}} z\gamma(x)[\varphi_{\Delta_{n,i}^{\beta}}(z\gamma(x)) - 1]F(z)dz + R(\theta, \Delta_{n,i}^{2-3\beta}, x)}{1 + R(\theta, \Delta_{n,i}^{(1-\alpha\beta)\wedge(2-4\beta)}, x)}. \quad (85)$$

Now R in the denominator is a rest function and so

$$\frac{1}{1 + R(\theta, \Delta_{n,i}^{(1-\alpha\beta)\wedge(2-4\beta)}, x)} \sim 1 - R(\theta, \Delta_{n,i}^{(1-\alpha\beta)\wedge(2-4\beta)}, x). \quad (86)$$

We now replace (86) in (85) and we observe that multiplying by R we obtain negligible functions, hence we get (17). □

Let us now prove the development of $m_{\theta,\Delta_{n,i}}$ in the particular case with finite intensity that makes possible to approximate explicitly the contrast function.

7.5 Proof of Proposition 2

Proof. We want to use again Dynkin's formula (34). We consider the decomposition of the generator: $A = A_c + A_d$ and, by the Remark 1 and the fact that we are in the finite intensity case, we can take $A_d f(x) = \int_{\mathbb{R}} \lambda[f(x + \gamma(x)z) - f(x)]F_0(z)dz$, where $F(z) = \lambda F_0(z)$ and $\int_{\mathbb{R}} F_0(z)dz = 1$.
Concerning the denominator, we denote again $f_{i,n}(y) := \varphi_{\Delta_{n,i}^{\beta}}(y - x)$ and, in order to calculate $A^k f_{i,n}(y)$ we introduce the following set of functions:

$$\mathcal{F}^p := \left\{ g(y)\, s.t.\, g(y) = \sum_{k=0}^{p} \varphi^{(k)}((y-x)\Delta_{n,i}^{-\beta})\Delta_{n,i}^{-k\beta}(\sum_{j=0}^{k} h_{k,j}(y)\Delta_{n,i}^{\beta j}) \right\}$$

where, $\forall k, j$, $\forall l \ge 0$ $\exists c$ such that $|\frac{\partial^l}{\partial y^l} h_{k,j}(y)| \le c(1 + |y|^c)$ and $\forall k, j$ $h_{k,j}$ is $\mathcal{C}^\infty$. We observe that, if $g \in \mathcal{F}^p$, then $g' \in \mathcal{F}^{p+1}$, bg and $a^2 g$ are in $\mathcal{F}^p$ and therefore if $g \in \mathcal{F}^p$, then $Ag \in \mathcal{F}^{p+2}$.
We now want to show that, for $g \in \mathcal{F}^p$, A_d acts like $-\lambda I_d$ up to an error term. Indeed,

$$A_d g(y) = \int_{\mathbb{R}} \lambda[g(y + \gamma(y)z) - g(y)]F_0(z)dz = \lambda \int_{\mathbb{R}} g(y + \gamma(y)z)F_0(z)dz - \lambda g(y). \quad (87)$$

Let us start considering $g(y) = \varphi^{(k)}((y-x)\Delta_{n,i}^{-\beta})h(y)$, where $k \le p$ and $h \in \mathcal{C}^\infty$ is such that $\forall l \ge 0$ $\exists c$: $|\frac{\partial^l}{\partial y^l} h(y)| \le c(1 + |y|^c)$. Then,

$$\int_{\mathbb{R}} g(y + \gamma(y)z)F_0(z)dz = \int_{\mathbb{R}} \varphi^{(k)}((y + \gamma(y)z - x)\Delta_{n,i}^{-\beta})\, h(y + \gamma(y)z)F_0(z)dz.$$

With the change of variable $u := (y + \gamma(y)z - x)\Delta_{n,i}^{-\beta}$ it becomes equal to

$$\frac{\Delta_{n,i}^{\beta}}{\gamma(y)} \int_{\mathbb{R}} \varphi^{(k)}(u)h(x + u\Delta_{n,i}^{\beta})F_0(\frac{x-y}{\gamma(y)} + \frac{\Delta_{n,i}^{\beta}u}{\gamma(y)})du. \quad (88)$$

We define $\tilde{F}(x,y,s) := \frac{h(x+s)}{\gamma(y)} F_0(\frac{x-y}{\gamma(y)} + \frac{s}{\gamma(y)})$ and we develop it up to the M-order, getting

$$\tilde{F}(x,y,\Delta_{n,i}^{\beta}u) = \sum_{j=0}^{M} \frac{\partial^j \tilde{F}}{\partial s^j}(x,y,0)(\Delta_{n,i}^{\beta}u)^j + \int_0^1 \frac{\partial^{M+1}}{\partial s^{M+1}} \tilde{F}(x,y,t\Delta_{n,i}^{\beta}u)\frac{(1-t)^M}{M!}(\Delta_{n,i}^{\beta}u)^{M+1}dt.$$

Replacing the development in (88) and recalling that by the definition of φ we have $\int_{\mathbb{R}} u^j \varphi^{(k)}(u)du = 0$, we get

$$\int_{\mathbb{R}} \varphi^{(k)}(u)\tilde{F}(x,y,\Delta_{n,i}^{\beta}u)du = \sum_{j=0}^{M} 0 + \int_{\mathbb{R}}\int_0^1 \varphi^{(k)}(u)\frac{\partial^{M+1}}{\partial s^{M+1}}\tilde{F}(x,y,t\Delta_{n,i}^{\beta}u)\frac{(1-t)^M}{M!}(\Delta_{n,i}^{\beta}u)^{M+1}dt\, du. \quad (89)$$

We observe that it is $|\frac{\partial^{l_1+l_2+l_3}}{\partial s^{l_1}\partial x^{l_2}\partial y^{l_3}}\tilde{F}(x,y,s)| \le c(1+|x|^c+|y|^c+|s|^c)$. Therefore, since the support of $\varphi^{(k)}$ is compact, we get

$$\int_{\mathbb{R}}\int_0^1 \varphi^{(k)}(u)\frac{\partial^{M+1}}{\partial s^{M+1}}\tilde{F}(x,y,t\Delta_{n,i}^{\beta}u)\frac{(1-t)^M}{M!}(\Delta_{n,i}^{\beta}u)^{M+1}dt\,du \le c(\Delta_{n,i}^{\beta})^{M+1}(1+|x|^c+|y|^c). \tag{90}$$

Hence using (88) and (90) on $|\int_{\mathbb{R}} g(y+\gamma(y)z)F_0(z)dz|$ and the differentiation of (89) on $|\frac{\partial^l}{\partial y^l}\int_{\mathbb{R}} g(y+\gamma(y)z)F_0(z)dz|$ we get that both of them are upper bounded by $c(1+|x|^c+|y|^c)\Delta_{n,i}^{\beta(M+2)}$, where in the second case the constant c depends on l.
Turning to a general function $g\in\mathcal{F}^p$, the estimations above become

$$|\int_{\mathbb{R}} g(y+\gamma(y)z)F_0(z)dz| \le c(1+|x|^c+|y|^c)\Delta_{n,i}^{\beta(M+2)}\Delta_{n,i}^{-\beta p} \tag{91}$$

and, $\forall l\ge 1$,

$$|\frac{\partial^l}{\partial y^l}\int_{\mathbb{R}} g(y+\gamma(y)z)F_0(z)dz| \le c_l(1+|x|^{c_l}+|y|^{c_l})\Delta_{n,i}^{\beta(M+2)}\Delta_{n,i}^{-\beta p}. \tag{92}$$

We introduce the set of functions

$$\mathcal{R}^p := \left\{ r(x,y,\Delta_{n,i}^p) \text{ such that } \forall l\ge 0\ \exists c_l\, |\frac{\partial^l}{\partial y^l}r(x,y,\Delta_{n,i}^p)| \le c_l(1+|x|^{c_l}+|y|^{c_l})\Delta_{n,i}^p \right\}.$$

Hence, using (87), (91) and (92) we have proved that, $\forall g\in\mathcal{F}^p$,

$$A_d g(y) = -\lambda g(y) + r(x,y,\Delta_{n,i}^{\beta(M+2-p)}). \tag{93}$$

We observe that if a function r is in $\mathcal{R}^p$, then both $A_d r$ and $A_c r$ are in $\mathcal{R}^p$. We can therefore now calculate for $f_{i,n}(y)=\varphi((y-x)\Delta_{n,i}^{-\beta})$, $f_{i,n}\in\mathcal{F}^0$,

$$A_{i_1}f_{i,n}(y) = \begin{cases} A_c f_{i,n}(y) & \text{if } i_1 = c \\ A_d f_{i,n}(y) = -\lambda f_{i,n}(y) + r(x,y,\Delta_{n,i}^{\beta(M+2)}) & \text{if } i_1 = d, \end{cases} \tag{94}$$

We want to show, by recurrence, that

$$A_{i_N}\circ\ldots\circ A_{i_1}(f_{i,n})(y) = A_c^{l(i_1,\ldots,i_N)}f_{i,n}(y)(-\lambda)^{N-l(i_1,\ldots,i_N)} + r(x,y,\Delta_{n,i}^{\beta(M+2)-2\beta l(i_1,\ldots,i_N)}), \tag{95}$$

with $l(i_1,\ldots,i_N)$ the number of c in $\{i_1,\ldots,i_N\}$. Let us consider the base case

$$A_{i_2}\circ A_{i_1}f_{i,n}(y) = \begin{cases} A_c^2 f_{i,n}(y) & \text{if } i_2=i_1=c \\ A_c(-\lambda f_{i,n}(y)+r(x,y,\Delta_{n,i}^{\beta(M+2)})) = -\lambda A_c f_{i,n}(y) + r(x,y,\Delta_{n,i}^{\beta(M+2)}) & \text{if } i_2=c,\ i_1=d \\ -\lambda A_c f_{i,n}(y) + r(x,y,\Delta_{n,i}^{\beta(M+2)-2\beta}) & \text{if } i_2=d,\ i_1=c \\ A_d(-\lambda f_{i,n}(y)+r(x,y,\Delta_{n,i}^{\beta(M+2)})) = \lambda^2 f_{i,n}(y) + r(x,y,\Delta_{n,i}^{\beta(M+2)}) & \text{if } i_2=i_1=d, \end{cases} \tag{96}$$

where in the third case we have used $A_c f_{i,n}\in\mathcal{F}^2$. So we have

$$A_{i_2}\circ A_{i_1}f_{i,n}(y) = A_c^{l(i_1,i_2)}f_{i,n}(y)(-\lambda)^{2-l(i_1,i_2)} + r(x,y,\Delta_{n,i}^{\beta(M+2)-2\beta l(i_1,i_2)}),$$

as we wanted. For the inductive step, we assume that (95) holds, now

$$A_{i_{N+1}}\circ A_{i_N}\circ\ldots\circ A_{i_1}(f_{i,n})(y) =$$

$$= \begin{cases} A_c\circ A_c^{l(i_1,\ldots,i_N)}f_{i,n}(y)(-\lambda)^{N-l(i_1,\ldots,i_N)} + r(x,y,\Delta_{n,i}^{\beta(M+2)-2\beta l(i_1,\ldots,i_N)}) & \text{if } i_{N+1}=c, \\ (-\lambda)A_c^{l(i_1,\ldots,i_N)}f_{i,n}(y)(-\lambda)^{N-l(i_1,\ldots,i_N)} + r(x,y,\Delta_{n,i}^{\beta(M+2)-2\beta l(i_1,\ldots,i_N)}) & \text{if } i_{N+1}=d, \end{cases} \tag{97}$$

where in the first case we have used that $A_c r(x,y,\Delta_{n,i}^h)\in\mathcal{R}^h$, $\forall h$, and in the second case that $A_d r(x,y,\Delta_{n,i}^h)\in\mathcal{R}^h$ and that $A_c^{l(i_1,\ldots,i_N)}f_{i,n}\in\mathcal{F}^{2l(i_1,\ldots,i_N)}$ while using (93).
It is equal to $A_c^{l(i_1,\ldots,i_N,i_{N+1})}f_{i,n}(y)(-\lambda)^{N+1-l(i_1,\ldots,i_N,i_{N+1})} + r(x,y,\Delta_{n,i}^{\beta(M+2)-2\beta l(i_1,\ldots,i_N,i_{N+1})})$ and therefore the recurrence is proved. We can now calculate $A^k f_{i,n}(x)$ in the Dynkin's formula (34) using (95):

$$A^k f_{i,n}(x) = \sum_{(i_1,\ldots,i_k)\in\{c,d\}^k}(A_{i_k}\circ\ldots\circ A_{i_1})f_{i,n}(x) =$$

$$= \sum_{(i_1,...,i_k)\in\{c,d\}^k} A_c^{l(i_1,...,i_k)} f_{i,n}(x)(-\lambda)^{k-l(i_1,...,i_k)} + r(x,x,\Delta_{n,i}^{\beta(M+2)-2\beta l(i_1,...,i_k)}). \tag{98}$$

Recalling that $A_c^l f_{i,n}(x) = 0 \ \forall l \geq 1$, (98) becomes $(-\lambda)^k f_{i,n}(x) + r(x,x,\Delta_{n,i}^{\beta(M+2)-2\beta k})$.
Therefore, the principal term in the development of the denominator of $m_{\theta,\Delta_{n,i}}(x)$ from Dynkin's formula up to order N is

$$\sum_{k=0}^{N} \frac{\Delta_{n,i}^k}{k!} A^k f_{i,n}(x) = \sum_{k=0}^{N} \frac{\Delta_{n,i}^k}{k!} (-\lambda)^k f_{i,n}(x) + r(x,x,\Delta_{n,i}^{\beta(M+2)-2\beta k+k}).$$

Let us now consider the term of rest in the Dynkin's formula (34). Observing that

$$|A_c^{N+1} f_{i,n}(y)| \leq \Delta_{n,i}^{-2\beta(N+1)}(1+|y|^c)$$

using (95) and the definition of the function r, we get that

$$|A^{N+1} f_{i,n}(y)| \leq c(\Delta_{n,i}^{-2\beta(N+1)} + \Delta_{n,i}^{\beta(M+2)-2\beta(N+1)})(1+|y|^c). \tag{99}$$

Therefore

$$\mathbb{E}[|A^{N+1} f_{i,n}(X_{u_{n+1}})||X_{t_i}=x] \leq c(\Delta_{n,i}^{-2\beta(N+1)} + \Delta_{n,i}^{\beta(M+2)-2\beta(N+1)})(1+|x|^c). \tag{100}$$

Replacing in (34) it yields

$$|\int_{t_i}^{t_{i+1}} \int_{t_i}^{u_1} ... \int_{t_i}^{u_N} \mathbb{E}[A^{N+1} f_{i,n}(X_{u_{n+1}})|X_{t_i}=x] du_{N+1}...du_2 du_1| \leq$$

$$\leq c\Delta_{n,i}^{N+1}(\Delta_{n,i}^{-2\beta(N+1)} + \Delta_{n,i}^{\beta(M+2)-2\beta(N+1)})(1+|x|^c).$$

Since $\Delta_{n,i}^{\beta(M+2)-2\beta(N+1)}$ is negligible compared to $\Delta_{n,i}^{-2\beta(N+1)}$, it is enough to have $(N+1)(1-2\beta) \geq \lfloor \beta(M+2) \rfloor$ in order to get the following development of the denominator $d_{\Delta_{n,i}}(x)$ of $m_{\theta,\Delta_{n,i}}(x)$:

$$d_{\Delta_{n,i}}(x) = \sum_{k=0}^{N} \frac{\Delta_{n,i}^k}{k!}(-\lambda)^k f_{i,n}(x) + r(x,x,\Delta_{n,i}^{\beta(M+2)+(1-2\beta)k}) + r(x,x,\Delta_{n,i}^{(1-2\beta)(N+1)}) =$$

$$= \sum_{k=0}^{\lfloor \beta(M+2) \rfloor} \frac{\Delta_{n,i}^k}{k!}(-\lambda)^k + r(x,x,\Delta_{n,i}^{\beta(M+2)}),$$

where we have also used that, by the definition of $f_{i,n}$, $f_{i,n}(x) = 1$ and in the sum we have considered only the terms up to $k = \lfloor \beta(M+2) \rfloor$ because the others are rest terms.
Let us now study the numerator $n_{\Delta_{n,i}}(x)$ of $m_{\theta,\Delta_{n,i}}(x)$: acting like in the proof of Theorem 2 we consider $\bar{g}(y) := (y-x)\varphi((y-x)\Delta_{n,i}^{-\beta})$. Let us introduce, in place of $\mathcal{F}^p$, the set $\tilde{\mathcal{F}}^p$.

$$\tilde{\mathcal{F}}^p := \left\{ \tilde{g}(y) \ s.t. \ \tilde{g}(y) = \sum_{k=0}^{p} \varphi^{(k)}((y-x)\Delta_{n,i}^{-\beta})\Delta_{n,i}^{-k\beta}(\sum_{j=0}^{k} h_{k,j}(x,y)\Delta_{n,i}^{\beta j}) \right\}$$

where, $\forall k, j$, $\forall l \geq 0$, $\exists c_l$ such that $|\frac{\partial^l}{\partial y^l} h_{k,j}(x,y)| \leq c_l(1+|x|^{c_l}+|y|^{c_l})$. We observe that, as it was for $\mathcal{F}^p$, if $\tilde{g} \in \tilde{\mathcal{F}}^p$ then $A\tilde{g} \in \tilde{\mathcal{F}}^{p+2}$ and, for all $\tilde{g} \in \tilde{\mathcal{F}}^p$,

$$A_d \tilde{g}(y) = -\lambda \tilde{g}(y) + r(x,y,\Delta_{n,i}^{\beta(M+2-p)}). \tag{101}$$

It turns that the same relation as (95) holds with $\bar{g}$ in place of $f_{i,n}$. Hence we get

$$A^k \bar{g}(y) = (A_c + A_d)^k \bar{g}(y) = \sum_{(i_1,...,i_k)\in\{c,d\}^k} A_c^{l(i_1,...,l_k)} \bar{g}(y)(-\lambda)^{k-l(i_1,...,i_k)} + r(x,y,\Delta_{n,i}^{\beta(M+2)-2\beta l(i_1,...,i_k)}) =$$

$$= \sum_{l=0}^{k} \binom{k}{l} (-\lambda)^{k-l} A_c^l \bar{g}(y) + r(x,y,\Delta_{n,i}^{\beta(M+2)-2\beta k}), \tag{102}$$

where $l(i_1,...,i_k)$ is the number of c in $\{i_1,...,i_k\}$ and $\binom{k}{l}$ are the binomial coefficients. Now, concerning the continuous part of the generator, since it is local and $\bar{g}(y) = (y-x)$ in the neighborhood of x, we find $A_c^l \bar{g}(x) = A_K^{(l)}(x)$, which are exactly the coefficients found in the case without jump studied by Kessler

in (Kessler, 1997).
By (102), the principal term in the development of the numerator is therefore

$$\sum_{k=0}^{N}\frac{\Delta_{n,i}^k}{k!}A^k g(x)=\sum_{k=0}^{N}\frac{\Delta_{n,i}^k}{k!}(\sum_{l=0}^{k}\binom{k}{l}(-\lambda)^{k-l}A_K^{(l)}(x)+r(x,x,\Delta_{n,i}^{\beta(M+2)-2\beta k}))=$$

$$=\sum_{k=0}^{N}\frac{\Delta_{n,i}^k}{k!}(\sum_{l=0}^{k}\binom{k}{l}(-\lambda)^{k-l}A_K^{(l)}(x))+r(x,x,\Delta_{n,i}^{\beta(M+2)}). \tag{103}$$

Changing the order of summation and introducing $k' := k-l$ we get that the first term of the previous equation is equal to

$$\sum_{l=0}^{N}\sum_{k=l}^{N}\frac{\Delta_{n,i}^k}{k!}\binom{k}{l}(-\lambda)^{k-l}A_K^{(l)}(x)=\sum_{l=0}^{N}\frac{\Delta_{n,i}^l}{l!}A_K^{(l)}(x)\sum_{k'=0}^{N-l}\Delta_{n,i}^{k'}(-\lambda)^{k'}\frac{l!}{(k'+l)!}\binom{l+k'}{l}=$$

$$=\sum_{l=0}^{N}\frac{\Delta_{n,i}^l}{l!}A_K^{(l)}(x)\sum_{k'=0}^{N-l}\frac{\Delta_{n,i}^{k'}(-\lambda)^{k'}}{k'!}, \tag{104}$$

where in the last equality we have used the definition of binomial coefficients. Concerning the rest term in the Dynkin's formula, we use again (99) and (100) with $\bar{g}$ in place of $f_{i,n}$ and it turns again

$$|\int_{t_i}^{t_{i+1}}\int_{t_i}^{u_1}...\int_{t_i}^{u_N}\mathbb{E}[A^{N+1}\bar{g}(X_{u_{N+1}})|X_{t_i}=x]du_{N+1}...du_2du_1|\le r(x,x,\Delta_{n,i}^{(1-2\beta)(N+1)}). \tag{105}$$

Hence, using (103), (104) and (105) we have the following development:

$$n_{\Delta_{n,i}}(x)=\sum_{l=0}^{N}\frac{\Delta_{n,i}^l}{l!}A_K^{(l)}(x)\sum_{k'=0}^{N-l}\frac{\Delta_{n,i}^{k'}(-\lambda)^{k'}}{k'!}+r(x,x,\Delta_{n,i}^{\beta(M+2)})+r(x,x,\Delta_{n,i}^{(1-2\beta)(N+1)}). \tag{106}$$

If $(N+1)(1-2\beta)\ge\beta(M+2)$, it entails

$$n_{\Delta_{n,i}}(x)=\sum_{l=0}^{\lfloor\beta(M+2)\rfloor}\frac{\Delta_{n,i}^l}{l!}A_K^{(l)}(x)\sum_{k'=0}^{\lfloor\beta(M+2)\rfloor}\frac{\Delta_{n,i}^{k'}(-\lambda)^{k'}}{k'!}+r(x,x,\Delta_{n,i}^{\beta(M+2)}).$$

Acting as in the proof of the development of m_θ given in Theorem 2 we can say that it exists $k_0>0$ such that, for $|x|\le\Delta_{n,i}^{-k_0}$, the development of $m_{\theta,\Delta_{n,i}}(x)$ is

$$x+\frac{n_{\Delta_{n,i}}(x)}{d_{\Delta_{n,i}}(x)}=x+\sum_{l=0}^{\lfloor\beta(M+2)\rfloor}\frac{\Delta_{n,i}^l}{l!}A_K^{(l)}(x)+r(x,x,\Delta_{n,i}^{\beta(M+2)}). \tag{107}$$

The expansion (23) follows after remarking that $A_K^{(0)}(x)=0$. □

7.6 Contrast convergence

Before proving the contrast convergence, let us define $r(\theta,x)$ as the particular rest function that turns out from the development of $m_{\theta,\Delta_{n,i}}$:

$$r(\theta,x):=m_{\theta,\Delta_{n,i}}(x)-x-\Delta_{n,i}\,b(x,\theta)-\Delta_{n,i}\int_{\mathbb{R}\setminus\{0\}}z\,\gamma(x)\,[1-\varphi_{\Delta_{n,i}^{\beta}}(\gamma(x)z)]\,F(z)dz. \tag{108}$$

We recall that $r(\theta,x)$ is $R(\theta,\Delta_{n,i}^{1+\delta},x)$ with $\delta>0$ as defined below equation (30).
In order to prove the consistency and asymptotic normality of the estimator, the first step is the following Lemma:

Lemma 4. *Suppose that Assumptions 1-5 and A_β are satisfied. Then*

$$\frac{U_n(\theta)-U_n(\theta_0)}{t_n}\xrightarrow{\mathbb{P}}\int_{\mathbb{R}}\frac{(b(x,\theta)-b(x,\theta_0))^2}{a^2(x)}\pi(dx) \tag{109}$$

Proof. By the definition,

$$U_n(\theta)=\sum_{i=0}^{n-1}\frac{(X_{t_{i+1}}-m_\theta(X_{t_i}))^2}{a^2(X_{t_i})\Delta_{n,i}}\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}.$$

We want to reformulate the contrast function, in order to compensate for the terms not depending on θ in the difference $U_n(\theta)-U_n(\theta_0)$.
The dynamic of the process X is known and so we can write

$$X_{t_{i+1}}=X_{t_i}+\int_{t_i}^{t_{i+1}}b(X_s,\theta_0)ds+\int_{t_i}^{t_{i+1}}a(X_s)dW_s+\int_{t_i}^{t_{i+1}}\int_{\mathbb{R}}\gamma(X_{s^-})z\tilde{\mu}(ds,dz).$$

We have proved the development (14) of m_θ, too. We can substitute both of them in $U_n(\theta)$, getting

$$U_n(\theta)=\sum_{i=0}^{n-1}\frac{1}{a^2(X_{t_i})\Delta_{n,i}}[X_{t_i}+\int_{t_i}^{t_{i+1}}b(X_s,\theta_0)ds+\int_{t_i}^{t_{i+1}}a(X_s)dW_s+\int_{t_i}^{t_{i+1}}\int_{\mathbb{R}}\gamma(X_{s^-})z\tilde{\mu}(ds,dz)-X_{t_i}+$$

$$+\Delta_{n,i}(-b(X_{t_i},\theta)+\int_{\mathbb{R}\setminus\{0\}}z\,\gamma(X_{t_i})\,(1-\varphi_{\Delta_{n,i}^\beta}(\gamma(X_{t_i})z))\,F(z)dz)+r(\theta,X_{t_i})]^2\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}=$$

$$=\sum_{i=0}^{n-1}\frac{1}{a^2(X_{t_i})\Delta_{n,i}}(\int_{t_i}^{t_{i+1}}b(X_s,\theta_0)ds-\Delta_{n,i}\,b(X_{t_i},\theta)+\zeta_i+r(\theta,X_{t_i}))^2\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}},$$

we recall the definition of $\zeta_i:=\int_{t_i}^{t_{i+1}}a(X_s)dW_s+\int_{t_i}^{t_{i+1}}\int_{\mathbb{R}\setminus\{0\}}\gamma(X_{s^-})z\tilde{\mu}(ds,dz)+$
$+\Delta_{n,i}\int_{\mathbb{R}\setminus\{0\}}z\,\gamma(X_{t_i})\,[1-\varphi_{\Delta_{n,i}^\beta}(\gamma(X_{t_i})z)]\,F(z)dz$, as in (29); we point out that ζ_i does not depend on θ.
In the same way

$$U_n(\theta_0)=\sum_{i=0}^{n-1}\frac{1}{a^2(X_{t_i})\Delta_{n,i}}(\int_{t_i}^{t_{i+1}}b(X_s,\theta_0)ds-\Delta_{n,i}\,b(X_{t_i},\theta_0)+\zeta_i+r(\theta_0,X_{t_i}))^2\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}$$

and so

$$\frac{U_n(\theta)-U_n(\theta_0)}{t_n}=\frac{1}{t_n}\sum_{i=0}^{n-1}\frac{\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}}{a^2(X_{t_i})\Delta_{n,i}}[\Delta_{n,i}^2(b(X_{t_i},\theta)^2-b(X_{t_i},\theta_0)^2)+$$

$$+2\Delta_{n,i}\int_{t_i}^{t_{i+1}}b(X_s,\theta_0)ds(b(X_{t_i},\theta_0)-b(X_{t_i},\theta))+A_i+B_i+C_i+D_i+E_i], \qquad (110)$$

with

$$A_i=2\zeta_i\Delta_{n,i}(b(X_{t_i},\theta_0)-b(X_{t_i},\theta)),\qquad B_i=2\zeta_i(r(\theta,X_{t_i})-r(\theta_0,X_{t_i})),$$
$$C_i=2\Delta_{n,i}(r(\theta_0,X_{t_i})b(X_{t_i},\theta_0)-r(\theta,X_{t_i})b(X_{t_i},\theta)),\qquad D_i=r(\theta,X_{t_i})^2-r(\theta_0,X_{t_i})^2,$$
$$E_i=2\int_{t_i}^{t_{i+1}}b(X_s,\theta_0)ds(r(\theta,X_{t_i})-r(\theta_0,X_{t_i})).$$

Our goal is to show that the contribution of A_i, B_i, C_i, D_i and E_i go to zero in probability as $n\to\infty$ and to prove that the other terms converge to $\int_{\mathbb{R}}\frac{(b(x,\theta)-b(x,\theta_0))^2}{a^2(x)}\pi(dx)$.
We observe that the rest function $r(\theta,x)$ is present in all the terms that have to converge to 0 but A_i, on which we use a different motivation to obtain the convergence:

$$\frac{1}{t_n}\sum_{i=0}^{n-1}\frac{\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}}{a^2(X_{t_i})\Delta_{n,i}}A_i=\frac{1}{t_n}\sum_{i=0}^{n-1}\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})f_{i,n}(X_{t_i},\theta)1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}\zeta_i,$$

with $f_{i,n}(X_{t_i},\theta):=\frac{2}{a^2(X_{t_i})}(b(X_{t_i},\theta_0)-b(X_{t_i},\theta))$.
In order to apply Proposition 4 we observe that, by the assumptions done on the coefficients, $f_{i,n}$ has polynomial growth. We therefore get the convergence to zero in probability, using Proposition 4.
We want to show that $\frac{1}{t_n}\sum_{i=0}^{n-1}\frac{\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}}{a^2(X_{t_i})\Delta_{n,i}}B_i\xrightarrow{\mathbb{P}}0$ and so we observe that, by the definition of the function r and by (11) we have that

$$r(\theta,X_{t_i})1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}=R(\theta,\Delta_{n,i}^{1+\delta},X_{t_i})=\Delta_{n,i}^{1+\delta}R(\theta,1,X_{t_i}). \qquad (111)$$

Hence

$$\frac{1}{t_n}\sum_{i=0}^{n-1}\frac{\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}}{a^2(X_{t_i})\Delta_{n,i}}B_i=$$

$$=\frac{1}{t_n}\sum_{i=0}^{n-1}\Delta_{n,i}^\delta\zeta_i\frac{\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}(R(\theta,1,X_{t_i})-R(\theta_0,1,X_{t_i}))}{a^2(X_{t_i})}$$

To prove the convergence, we have to show that

$$\frac{1}{t_n}\sum_{i=0}^{n-1}|\mathbb{E}[\Delta_{n,i}^\delta f_{i,n}(X_{t_i},\theta)1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}\zeta_i\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})|\mathcal{F}_{t_i}]|\xrightarrow{\mathbb{P}}0, \tag{112}$$

and

$$\frac{1}{(t_n)^2}\sum_{i=0}^{n-1}\mathbb{E}[\Delta_{n,i}^{2\delta}f_{i,n}^2(X_{t_i},\theta)1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}\zeta_i^2\varphi_{\Delta_{n,i}^\beta}^2(X_{t_{i+1}}-X_{t_i})|\mathcal{F}_{t_i}]\xrightarrow{\mathbb{P}}0,$$

with

$$f_{i,n}(X_{t_i},\theta)=\frac{R(\theta,1,X_{t_i})-R(\theta_0,1,X_{t_i})}{a^2(X_{t_i})}.$$

By the measurability of X_{t_i} with respect to $\mathcal{F}_{t_i}$, by the fact that $|\Delta_{n,i}|\le\Delta_n$ and that $t_n=0(n\Delta_n)$ we get

$$\frac{1}{t_n}\sum_{i=0}^{n-1}|\mathbb{E}[\Delta_{n,i}^\delta f_{i,n}(X_{t_i},\theta)1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}\zeta_i\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})|\mathcal{F}_{t_i}]|=$$

$$=\frac{1}{t_n}\sum_{i=0}^{n-1}\Delta_{n,i}^\delta|f_{i,n}(X_{t_i},\theta)||\mathbb{E}[1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}\zeta_i\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})|\mathcal{F}_{t_i}]|\le$$

$$\le\Delta_n^\delta\frac{c}{n\Delta_n}\sum_{i=0}^{n-1}|f_{i,n}(X_{t_i},\theta)||\mathbb{E}[1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}\zeta_i\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})|\mathcal{F}_{t_i}]|.$$

We recall that δ is positive. Using (31), we get the convergence (112) in L^1 and thus in probability.
In the same way,

$$\frac{1}{(t_n)^2}\sum_{i=0}^{n-1}\mathbb{E}[\Delta_{n,i}^{2\delta}f_{i,n}^2(X_{t_i},\theta)1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}\zeta_i^2\varphi_{\Delta_{n,i}^\beta}^2(X_{t_{i+1}}-X_{t_i})|\mathcal{F}_{t_i}]\le$$

$$\le\Delta_n^{2\delta}\frac{c}{n^2\Delta_n^2}\sum_{i=0}^{n-1}f_{i,n}^2(X_{t_i},\theta)\mathbb{E}[1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}\zeta_i^2\varphi_{\Delta_{n,i}^\beta}^2(X_{t_{i+1}}-X_{t_i})|\mathcal{F}_{t_i}],$$

that goes to zero in probability using (32).

$$\frac{1}{t_n}\sum_{i=0}^{n-1}\frac{\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}}{a^2(X_{t_i})\Delta_{n,i}}C_i=\frac{2}{t_n}\sum_{i=0}^{n-1}\Delta_{n,i}^{1+\delta}f_{i,n}(X_{t_i},\theta)\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i}),$$

with $f_{i,n}(X_{t_i},\theta):=\frac{R(\theta_0,1,X_{t_i})b(X_{t_i},\theta_0)-R(\theta,1,X_{t_i})b(X_{t_i},\theta)}{a^2(X_{t_i})}1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}$, where we have used (111).
In module, it is upper bounded by $\Delta_n^\delta\frac{c}{n}\sum_{i=0}^{n-1}|f_{i,n}(X_{t_i},\theta)\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})|$.
We observe that the exponent on Δ_n is positive so it goes to zero as $n\to\infty$ and that $|\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})|\le c$. By the polynomial growth of $f_{i,n}$ and the third point of Lemma 2, we get that $\frac{1}{n}\sum_{i=0}^{n-1}|f_{i,n}(X_{t_i},\theta)|$ is bounded in L^1. It yields the convergence in probability that we were looking for.
Let us consider D_i. Using triangle inequality, we can just prove the convergence of the following:

$$|\frac{1}{t_n}\sum_{i=0}^{n-1}\frac{\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}}{a^2(X_{t_i})\Delta_{n,i}}r(\theta,X_{t_i})^2|=$$

$$=|\frac{1}{t_n}\sum_{i=0}^{n-1}\Delta_{n,i}^{1+2\delta}\frac{\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}}{a^2(X_{t_i})}R(\theta,1,X_{t_i})^2|\le$$

$$\le\Delta_n^{2\delta}\frac{1}{n}\sum_{i=0}^{n-1}|f_{i,n}(X_{t_i},\theta)\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})|,$$

$f_{i,n}(X_{t_i},\theta) = \frac{R(\theta,1,X_{t_i})^2}{a^2(X_{t_i})}$, using also the indicator is always upper bounded by 1.
Also this time the exponent on Δ_n is positive. We can use the boundedness of $|\varphi_{\Delta_{n,i}^\beta}|$, the polynomial growth of $f_{i,n}$ and third point of Lemma 2 in order to get that $\frac{1}{n}\sum_{i=0}^{n-1}|f_{i,n}(X_{t_i},\theta)\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})|$ is bounded in L^1 . It turns

$$|\frac{1}{t_n}\sum_{i=0}^{n-1}\frac{\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}}{a^2(X_{t_i})\Delta_{n,i}}r(\theta,X_{t_i})^2| \xrightarrow{\mathbb{P}} 0.$$

Considering E_i, we use again the triangle inequality in order to prove only the convergence to zero of the following:

$$|\frac{1}{t_n}\sum_{i=0}^{n-1}2\frac{\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}}{a^2(X_{t_i})\Delta_{n,i}}\int_{t_i}^{t_{i+1}}b(X_s,\theta_0)ds\,r(\theta,X_{t_i})|. \tag{113}$$

In the sequel it will be useful to substitute $\int_{t_i}^{t_{i+1}}b(X_s,\theta_0)ds$ with $\Delta_{n,i}\,b(X_{t_i},\theta_0)$.

$$\int_{t_i}^{t_{i+1}}b(X_s,\theta_0)ds = \int_{t_i}^{t_{i+1}}[b(X_s,\theta_0)-b(X_{t_i},\theta_0)]ds + \Delta_{n,i}b(X_{t_i},\theta_0). \tag{114}$$

In order to show that the first term is negligible compared to $\Delta_{n,i}$, we consider the following expected value:

$$\sup_{u\in[0,\Delta_{n,i}]}\mathbb{E}[|b(X_{t_i+u},\theta_0)-b(X_{t_i},\theta_0)||\mathcal{F}_{t_i}] \le \sup_{u\in[0,\Delta_{n,i}]}\mathbb{E}[\left\|\frac{\partial b}{\partial x}\right\|_\infty|X_{t_i+u}-X_{t_i}||\mathcal{F}_{t_i}] \le$$
$$\le c\sup_{u\in[0,\Delta_{n,i}]}\mathbb{E}[|X_{t_i+u}-X_{t_i}||\mathcal{F}_{t_i}].$$

In the last inequality we have used that the derivative of b is supposed bounded.
Using Holder inequality we get that it is, for each $p\ge 2$, upper bounded by

$$c\sup_{u\in[0,\Delta_{n,i}]}(\mathbb{E}[|X_{t_i+u}-X_{t_i}|^p|\mathcal{F}_{t_i}])^{\frac{1}{p}} \le$$
$$\le c\sup_{u\in[0,\Delta_{n,i}]}(|t_i+u-t_i|(1+|X_{t_i}|^p))^{\frac{1}{p}} = R(\theta,\Delta_{n,i}^{\frac{1}{p}},X_{t_i}). \tag{115}$$

Where, in the last inequality, we have used the second point of Lemma 1.
For $p=2$, $\mathbb{E}[|b(X_{t_i+u},\theta_0)-b(X_{t_i},\theta_0)||\mathcal{F}_{t_i}] \le R(\theta,\Delta_{n,i}^{\frac{1}{2}},X_{t_i})$ and therefore

$$\int_{t_i}^{t_{i+1}}\mathbb{E}[|b(X_s,\theta_0)-b(X_{t_i},\theta_0)||\mathcal{F}_{t_i}]ds \le \int_{t_i}^{t_{i+1}}R(\theta,\Delta_{n,i}^{\frac{1}{2}},X_{t_i})ds = R(\theta,\Delta_{n,i}^{\frac{3}{2}},X_{t_i}), \tag{116}$$

negligible compared to $\Delta_{n,i}$, that is the order of the second term of (114).
Using (111) and (114), (113) can be reformulated as

$$|\frac{1}{t_n}\sum_{i=0}^{n-1}2\frac{\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}\Delta_{n,i}^\delta R(\theta,1,X_{t_i})}{a^2(X_{t_i})}[\Delta_{n,i}\,b(X_{t_i},\theta_0)+$$
$$+\int_{t_i}^{t_{i+1}}[b(X_s,\theta_0)-b(X_{t_i},\theta_0)]ds]|. \tag{117}$$

The first term is upper bounded by

$$\Delta_n^\delta\frac{1}{n}\sum_{i=0}^{n-1}|f_{i,n}(X_{t_i},\theta)\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})|,$$

where $f_{i,n}(X_{t_i},\theta) = \frac{2b(X_{t_i},\theta_0)R(\theta,1,X_{t_i})}{a^2(X_{t_i})}$.
Again, the exponent on Δ_n is positive and $\frac{1}{n}\sum_{i=0}^{n-1}|f_{i,n}(X_{t_i},\theta)\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})|$ is bounded in L^1 using the boundedness of $\varphi_{\Delta_{n,i}^\beta}$, the polynomial growth of $f_{i,n}$ and the third point of Lemma 2.
Concerning the second term of (117), we observe it is upper bounded by

$$\Delta_n^\delta\frac{1}{n\Delta_n}\sum_{i=0}^{n-1}|f_{i,n}(X_{t_i},\theta)\int_{t_i}^{t_{i+1}}[b(X_s,\theta_0)-b(X_{t_i},\theta_0)]ds\,\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})|,$$

where $f_{i,n}(X_{t_i},\theta) = \frac{2R(\theta,1,X_{t_i})}{a^2(X_{t_i})}$. The exponent on Δ_n is still positive and $\frac{1}{n\Delta_n}\sum_{i=0}^{n-1}|f_{i,n}(X_{t_i},\theta)\int_{t_i}^{t_{i+1}}[b(X_s,\theta_0)-b(X_{t_i},\theta_0)]ds\,\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})|$ is bounded in L^1. Indeed,

$$\frac{1}{n\Delta_n}\sum_{i=0}^{n-1}\mathbb{E}[|f_{i,n}(X_{t_i},\theta)\int_{t_i}^{t_{i+1}}[b(X_s,\theta_0)-b(X_{t_i},\theta_0)]ds\,\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})|]\leq$$

$$\leq \frac{c}{n\Delta_n}\sum_{i=0}^{n-1}\mathbb{E}[|f_{i,n}(X_{t_i},\theta)\int_{t_i}^{t_{i+1}}[b(X_s,\theta_0)-b(X_{t_i},\theta_0)]ds|]=$$

$$=\frac{c}{n\Delta_n}\sum_{i=0}^{n-1}\mathbb{E}[|f_{i,n}(X_{t_i},\theta)\mathbb{E}[\int_{t_i}^{t_{i+1}}[b(X_s,\theta_0)-b(X_{t_i},\theta_0)]ds\,|\mathcal{F}_{t_i}]|]=\frac{c}{n\Delta_n}\sum_{i=0}^{n-1}\mathbb{E}[|f_{i,n}(X_{t_i},\theta)R(\theta,\Delta_{n,i}^{\frac{3}{2}},X_{t_i})|], \tag{118}$$

where we have used the definition of conditional expectation and (116).
From (11), we can upper bound (118) by $\Delta_n^{\frac{1}{2}}\frac{1}{n}\sum_{i=0}^{n-1}\mathbb{E}[|f_{i,n}(X_{t_i},\theta)R(\theta,1,X_{t_i})|]$.
The exponent of Δ_n is clearly positive and $\frac{1}{n}\sum_{i=0}^{n-1}\mathbb{E}[|f_{i,n}(X_{t_i},\theta)R(\theta,1,X_{t_i})|]$ is bounded using again the polynomial growth of both $f_{n,i}$ and R and the third point of Lemma 2.
We have obtained the wanted convergence.

Let us now consider the main terms of (110): we will show that they converge to $\int_{\mathbb{R}}\frac{(b(x,\theta)-b(x,\theta_0))^2}{a^2(x)}\pi(dx)$.
In order to do it, we want to replace $\int_{t_i}^{t_{i+1}}b(X_s,\theta_0)ds$ with $\Delta_{n,i}\,b(X_{t_i},\theta_0)$ in (110), getting:

$$\Delta_{n,i}^2[b(X_{t_i},\theta)^2-b(X_{t_i},\theta_0)^2]+2\Delta_{n,i}^2 b(X_{t_i},\theta_0)[b(X_{t_i},\theta_0)-b(X_{t_i},\theta)]=\Delta_{n,i}^2[b(X_{t_i},\theta)-b(X_{t_i},\theta_0)]^2.$$

Hence, we can reformulate (110) adding and subtracting $\Delta_{n,i}b(X_{t_i},\theta_0)$. We obtain

$$\frac{U_n(\theta)-U_n(\theta_0)}{t_n}=\frac{1}{t_n}\sum_{i=0}^{n-1}\frac{\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})1_{\{|X_{t_i}|\leq\Delta_{n,i}^{-k}\}}}{a^2(X_{t_i})}\Delta_{n,i}[b(X_{t_i},\theta)-b(X_{t_i},\theta_0)]^2+$$

$$+\frac{1}{t_n}\sum_{i=0}^{n-1}\frac{2\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})1_{\{|X_{t_i}|\leq\Delta_{n,i}^{-k}\}}}{a^2(X_{t_i})}(\int_{t_i}^{t_{i+1}}[b(X_s,\theta_0)-b(X_{t_i},\theta_0)]ds)\,[b(X_{t_i},\theta_0)-b(X_{t_i},\theta)]+R_i, \tag{119}$$

where R_i represents the rest terms, for which we have already shown the convergence to 0 in probability.
The second term of (119) goes to 0 in L^1, in fact

$$\mathbb{E}[|\frac{1}{t_n}\sum_{i=0}^{n-1}f(X_{t_i},\theta)\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})(\int_{t_i}^{t_{i+1}}(b(X_s,\theta_0)ds-b(X_{t_i},\theta_0))ds|]=$$

$$=\mathbb{E}[|\frac{1}{t_n}\sum_{i=0}^{n-1}f(X_{t_i},\theta)\mathbb{E}[\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})(\int_{t_i}^{t_{i+1}}(b(X_s,\theta_0)-b(X_{t_i},\theta_0))ds|\mathcal{F}_{t_i}]|],$$

With $f(X_{t_i},\theta) := \frac{2(b(X_{t_i},\theta_0)-b(X_{t_i},\theta))}{a^2(X_{t_i})}1_{\{|X_{t_i}|\leq\Delta_{n,i}^{-k}\}}$.
Using that $\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})$ is bounded by a constant and the estimation (116), we get that it is upper bounded by

$$\mathbb{E}[|\frac{1}{n\Delta_n}\sum_{i=0}^{n-1}f(X_{t_i},\theta)R(\theta,\Delta_{n,i}^{\frac{3}{2}},X_{t_i})|]\leq\Delta_n^{\frac{1}{2}}\frac{1}{n}\sum_{i=0}^{n-1}\mathbb{E}[|f(X_{t_i},\theta)R(\theta,1,X_{t_i})|],$$

where in the last inequality we have used (11), the triangle inequality and that $|\Delta_{n,i}|\leq\Delta_n$. Using the third point of Lemma 2, we obtain that $\frac{1}{n}\sum_{i=0}^{n-1}|f(X_{t_i},\theta)R(\theta,1,X_{t_i})|$ is bounded in L^1 and so the convergence wanted.
To conclude, we use the second point of Proposition 3 on the first term of (119). It yields

$$\frac{1}{t_n}\sum_{i=0}^{n-1}\frac{\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})1_{\{|X_{t_i}|\leq\Delta_{n,i}^{-k}\}}}{a^2(X_{t_i})}\Delta_{n,i}(b(X_{t_i},\theta_0)-b(X_{t_i},\theta))^2\xrightarrow{\mathbb{P}}\int_{\mathbb{R}}\frac{(b(x,\theta)-b(x,\theta_0))^2}{a^2(x)}\pi(dx).$$

Therefore,

$$\frac{U_n(\theta)-U_n(\theta_0)}{t_n}\xrightarrow{\mathbb{P}}\int_{\mathbb{R}}\frac{(b(x,\theta)-b(x,\theta_0))^2}{a^2(x)}\pi(dx).$$

□

Remark 9. *We observe that the contrast function does not converge:* $\forall \theta \in \Theta$

$$\lim_{n\to\infty} \frac{U_n(\theta)}{t_n} = \infty.$$

It happens because, in the expansion

$$X_{t_{i+1}} - m_\theta(X_{t_i}) = \zeta_i + \int_{t_i}^{t_{i+1}} b(X_s, \theta_0)ds - \Delta_{n,i}\, b(X_{t_i}, \theta) + R(\theta, \Delta_{n,i}^{1+\delta}, X_{t_i}),$$

ζ_i *is of the order* $\Delta_n^{\frac{1}{2}}$ *while the order of the part dependent on* θ *is* Δ_n*.*
That is the reason why we consider the difference between $U_n(\theta)$ *and* $U_n(\theta_0)$*: stressing that* ζ_i *does not depend on* θ*, we get that in the difference it does not contribute anymore.*
The asymptotic behavior of $(U_n(\theta) - U_n(\theta_0))$ *is therefore governed by the part depending on* θ*.*

7.7 Consistency of the estimator

In order to prove the consistency of $\hat{\theta}_n$, we need that the convergence (109) takes place in probability uniformly in the parameter θ, we want therefore to show the uniformity of the convergence in θ.
Let $S_n(\theta) := \frac{U_n(\theta) - U_n(\theta_0)}{t_n}$; we regard this as a random element taking values in $(C(\Theta), \|.\|_\infty)$. It suffices to prove the tightness of this sequence, to do it we need an explicit approximation of $\dot{m}_{\theta,h}$. Such an approximation, together with the approximation of $\ddot{m}_{\theta,h}$, will be also useful to study the asymptotic behavior of the derivatives of the contrast function. In the following proposition we study their asymptotic expansions as $\Delta_{n,i} \to 0$:

Proposition 8. *Suppose that Assumptions 1 to 4 and 7 hold, with* $\alpha \in (0, 2)$*,* $\alpha \neq 1$ *and* $\beta \in (0, \frac{1}{1+\alpha} - \epsilon)$*. Then, for* $|y| \le h^{-k_0}$ *(where* k_0 *is the same as in Theorem 2 or 4, according to* $\alpha < 1$ *or* $\alpha > 1$*),*

$$\dot{m}_{\theta,h}(y) = h\dot{b}(y, \theta) + R(\theta, h^{\frac{3}{2} \wedge (2 - \alpha\beta - \epsilon - \beta)}, y) \tag{120}$$

and

$$\ddot{m}_{\theta,h}(y) = h\ddot{b}(y, \theta) + R(\theta, h^{\frac{3}{2} \wedge (2 - \alpha\beta - \epsilon - \beta)}, y). \tag{121}$$

Remark 10. *It is also possible to show that*

$$|\dddot{m}_{\theta,h}(y)| = R(\theta, h, y). \tag{122}$$

The proposition above will be proved in the Appendix A.1, where we will also justify (122). We can now show the tightness of $S_n(\theta)$:

Lemma 5. *Suppose that Assumptions 1 - 8 and* A_β *are satisfied. Then*

$$S_n(\theta) := \frac{U_n(\theta) - U_n(\theta_0)}{t_n}$$

is a tight sequence in $(C(\Theta), \|.\|_\infty)$*.*

Proof. In the proof we use the notation of Section 5.3 and especially of the proof of Lemma 4. Since the sum of tight sequences is also tight, we can see $S_n(\theta)$ as $S_{n1}(\theta) + S_{n2}(\theta)$, where

$$S_{n1}(\theta) := \frac{1}{t_n} \sum_{i=0}^{n-1} \frac{\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}} - X_{t_i}) 1_{\{|X_{t_i}| \le \Delta_{n,i}^{-k}\}}}{a^2(X_{t_i})\Delta_{n,i}} [\Delta_{n,i}^2 (b(X_{t_i}, \theta)^2 - b(X_{t_i}, \theta_0)^2) +$$

$$+ 2\Delta_{n,i} \int_{t_i}^{t_{i+1}} b(X_s, \theta_0) ds (b(X_{t_i}, \theta_0) - b(X_{t_i}, \theta)) + C_i + D_i + E_i] +$$

$$+ \frac{2}{t_n} \sum_{i=0}^{n-1} \frac{1_{\{|X_{t_i}| \le \Delta_{n,i}^{-k}\}} \mathbb{E}[\zeta_i \varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}} - X_{t_i}) | \mathcal{F}_{t_i}]}{a^2(X_{t_i})\Delta_{n,i}} (\Delta_{n,i}(b(X_{t_i}, \theta) - b(X_{t_i}, \theta_0)) + (r(\theta, X_{t_i}) - r(\theta_0, X_{t_i}))),$$

$$S_{n2}(\theta) := \frac{2}{t_n} \sum_{i=0}^{n-1} \frac{1_{\{|X_{t_i}| \le \Delta_{n,i}^{-k}\}}}{a^2(X_{t_i})\Delta_{n,i}} [\zeta_i \varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}} - X_{t_i}) +$$

$$-\mathbb{E}[\zeta_i\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})|\mathcal{F}_{t_i}]](\Delta_{n,i}(b(X_{t_i},\theta)-b(X_{t_i},\theta_0))+(r(\theta,X_{t_i})-r(\theta_0,X_{t_i}))),$$

and show the tightness of the two sequences individually, using two different criteria.
In order to prove that S_{n1} is tight, we want to show that $\sup_n \mathbb{E}[\sup_{\theta\in\Theta}|\frac{\partial}{\partial\theta}S_{n1}(\theta)|]<\infty$. As concerns $S_{n2}(\theta)$, according to Theorem 20 in Appendix 1 from Ibragimov and Has' Minskii (Ibragimov & Has' Minskii, 2013), we should verify the following: for some positive constant H independent of n, ,

$$\mathbb{E}[(S_{n2}(\theta))^2]\le H \qquad \forall\theta\in\Theta, \tag{123}$$

$$\mathbb{E}[(S_{n2}(\theta_1)-S_{n2}(\theta_2))^2]\le H(\theta_1-\theta_2)^2 \qquad \forall\theta_1,\theta_2\in\Theta. \tag{124}$$

The derivative that we want to estimate is, using the expressions of C_i, D_i and E_i,

$$\frac{\partial S_{n1}(\theta)}{\partial\theta}=\frac{1}{t_n}\sum_{i=0}^{n-1}\frac{\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}}{a^2(X_{t_i})\Delta_{n,i}}[2\Delta_{n,i}^2 b(X_{t_i},\theta)\dot{b}(X_{t_i},\theta)+ \tag{125}$$

$$+2\Delta_{n,i}\int_{t_i}^{t_{i+1}}b(X_s,\theta_0)ds(-\dot{b}(X_{t_i},\theta))-2\Delta_{n,i}(\dot{b}r)(X_{t_i},\theta)-2\Delta_{n,i}(b\dot{r})(\theta,X_{t_i})+2(\dot{r}r)(\theta,X_{t_i})+$$

$$+2\dot{r}(\theta,X_{t_i})\int_{t_i}^{t_{i+1}}b(X_s,\theta_0)ds]+\frac{2}{t_n}\sum_{i=0}^{n-1}\frac{\mathbb{E}[\zeta_i\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})|\mathcal{F}_{t_i}]1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}}{a^2(X_{t_i})\Delta_{n,i}}(\dot{r}(\theta,X_{t_i})+\Delta_{n,i}\dot{b}(X_{t_i},\theta)).$$

Using triangle inequality, we can just estimate each term in L^1 norm.
Using the polynomial growth of both b and $\dot{b}$, the fact that φ and the indicator function are bounded, that a^2 is bigger than a constant from Assumption 5 and that $|\Delta_{n,i}|\le\Delta_n$, we get the first term of (125) is upper bounded by

$$\mathbb{E}[\sup_{\theta\in\Theta}|\frac{1}{n}\sum_{i=0}^{n-1}(1+|X_{t_i}|^c)|],$$

that is bounded by the third point of Lemma 2.
On the second term of (125) we can use that φ and the indicator function are bounded, that a^2 is bigger than a constant from Assumption 5, that both b and $\dot{b}$ have polynomial growth, from the integral we get a $|\Delta_{n,i}|$ (using (114) and (116)) that is smaller than Δ_n and so we have just to use the third point of Lemma 2 in order to say that the moments of X are bounded. Hence

$$\mathbb{E}[\sup_{\theta\in\Theta}|\frac{1}{t_n}\sum_{i=0}^{n-1}\frac{\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}}{a^2(X_{t_i})\Delta_{n,i}}2\Delta_{n,i}\int_{t_i}^{t_{i+1}}b(X_s,\theta_0)ds(-\dot{b}(X_{t_i},\theta))|]\le c.$$

Concerning the third and the fourth terms of (125), we use again that φ and the indicator function are bounded, that a^2 is bigger than a constant from Assumption 5 and that $\dot{b}$ has polynomial growth. We recall that

$$r(\theta,X_{t_i})1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}=R(\theta,\Delta_{n,i}^{1+\delta},X_{t_i})=\Delta_{n,i}^{1+\delta}R(\theta,1,X_{t_i}), \tag{126}$$

using (111). By the definition (108) and the development (120) of $\dot{m}_\theta$ we get also the following estimation:

$$\sup_{\theta\in\Theta}|\dot{r}(\theta,x)|\le\Delta_{n,i}(1+|x|^c). \tag{127}$$

We obtain in this way a $|\Delta_{n,i}|$ that is always smaller than Δ_n and so we can simplify the Δ_n in the denominator. Now we use the third point of Lemma 2 and we get also this time that the expectation is bounded.
Also on the fifth we use that φ and the indicator function are bounded, a^2 is bigger than a constant from Assumption 5, (126) and (127) on $\dot{r}$. Therefore the fifth term of (125) is upper bounded by $\Delta_n^\delta\mathbb{E}[|\frac{1}{n}\sum_{i=0}^{n-1}(1+|X_{t_i}|^c)|]$.
Since the exponent on Δ_n is positive and by the third point of Lemma 2, it is upper bounded by a constant.
As concerns the expected value of the sixth term of (125), we use again that φ and the indicator function are both bounded, a^2 is bigger than a constant from Assumption 5 and (127) on $\dot{r}$. Moreover, we get a $|\Delta_{n,i}|$ from the integral (using (114) and (116)). The third point of Lemma 2 is sufficient to assure the boundedness of the considered expectation.
Let us now consider

$$\mathbb{E}[\sup_{\theta\in\Theta}|\frac{2}{t_n}\sum_{i=0}^{n-1}\frac{\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}}{a^2(X_{t_i})}\dot{b}(X_{t_i},\theta)\mathbb{E}[\zeta_i\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})|\mathcal{F}_{t_i}]|].$$

By the boundedness of φ, the Assumption 5 on a and the polynomial growth of $\dot{b}$, it is upper bounded by

$$\mathbb{E}[\sup_{\theta \in \Theta} |\frac{1}{n\Delta_n} \sum_{i=0}^{n-1} \mathbb{E}[1_{\{|X_{t_i}| \le \Delta_{n,i}^{-k}\}} \zeta_i \varphi_{\Delta_{n,i}^{\beta}}(X_{t_{i+1}} - X_{t_i}) | \mathcal{F}_{t_i}](1 + |X_{t_i}|^c)|] \le$$

$$\le \mathbb{E}[\sup_{\theta \in \Theta} |\frac{1}{n\Delta_n} \sum_{i=0}^{n-1} R(\theta, \Delta_{n,i}^{(1+\delta)\wedge \frac{3}{2}})(1 + |X_{t_i}|^c)|] \le c\Delta_n^{\delta \wedge \frac{1}{2}},$$

where we have used (31), $|\Delta_{n,i}| \le \Delta_n$ and the third point of Lemma 2. Since the exponent on Δ_n is positive, it is bounded by a constant.
In order to conclude the proof of the S_{n1}'s tightness, we observe that by the boundedness of both φ and the indicator function, the Assumption 5 on a and (127) on $\dot{r}$ we get

$$\mathbb{E}[\sup_{\theta \in \Theta} |\frac{2}{t_n} \sum_{i=0}^{n-1} \frac{\varphi_{\Delta_{n,i}^{\beta}}(X_{t_{i+1}} - X_{t_i}) 1_{\{|X_{t_i}| \le \Delta_{n,i}^{-k}\}}}{a^2(X_{t_i})\Delta_{n,i}} \dot{r}(\theta, X_{t_i}) \mathbb{E}[\zeta_i \varphi_{\Delta_{n,i}^{\beta}}(X_{t_{i+1}} - X_{t_i}) | \mathcal{F}_{t_i}]|] \le$$

$$\le \mathbb{E}[|\frac{c}{n\Delta_n} \sum_{i=0}^{n-1} \mathbb{E}[1_{\{|X_{t_i}| \le \Delta_{n,i}^{-k}\}} \zeta_i \varphi_{\Delta_{n,i}^{\beta}}(X_{t_{i+1}} - X_{t_i}) | \mathcal{F}_{t_i}](1 + |X_{t_i}|^c)],$$

on which we can act exactly like above, getting the wanted boundedness.

Let us now consider S_{n2}. In order to prove (124), we observe that

$$\mathbb{E}[(S_{n2}(\theta_1) - S_{n2}(\theta_2))^2] \le \frac{c}{n^2\Delta_n^2} \mathbb{E}[(\sum_{i=0}^{n-1} \frac{1_{\{|X_{t_i}| \le \Delta_{n,i}^{-k}\}}}{a^2(X_{t_i})\Delta_{n,i}} [\zeta_i \varphi_{\Delta_{n,i}^{\beta}}(X_{t_{i+1}} - X_{t_i}) + \tag{128}$$

$$- \mathbb{E}[\zeta_i \varphi_{\Delta_{n,i}^{\beta}}(X_{t_{i+1}} - X_{t_i}) | \mathcal{F}_{t_i}]](\Delta_{n,i}(b(X_{t_i}, \theta_2) - b(X_{t_i}, \theta_1)) + r(\theta_1, X_{t_i}) - r(\theta_2, X_{t_i})))^2]$$

By the building the sum is a square integrable martingale. The Pythagoras' theorem on a square integrable martingale yields that (128) is equal to

$$\frac{c}{n^2\Delta_n^2} \sum_{i=0}^{n-1} \mathbb{E}[\frac{1_{\{|X_{t_i}| \le \Delta_{n,i}^{-k}\}}}{a^4(X_{t_i})\Delta_{n,i}^2} [\zeta_i \varphi_{\Delta_{n,i}^{\beta}}(X_{t_{i+1}} - X_{t_i}) + \tag{129}$$

$$- \mathbb{E}[\zeta_i \varphi_{\Delta_{n,i}^{\beta}}(X_{t_{i+1}} - X_{t_i}) | \mathcal{F}_{t_i}]]^2 (\Delta_{n,i}(b(X_{t_i}, \theta_2) - b(X_{t_i}, \theta_1)) + r(\theta_1, X_{t_i}) - r(\theta_2, X_{t_i}))^2].$$

We now observe that

$$(\Delta_{n,i}(b(X_{t_i}, \theta_2) - b(X_{t_i}, \theta_1)) + r(\theta_1, X_{t_i}) - r(\theta_2, X_{t_i}))^2 \le c\Delta_{n,i}^2(b(X_{t_i}, \theta_2) - b(X_{t_i}, \theta_1))^2 +$$

$$+ c(r(\theta_1, X_{t_i}) - r(\theta_2, X_{t_i}))^2 \le c\Delta_{n,i}^2 \dot{b}(X_{t_i}, \theta_u)^2(\theta_1 - \theta_2)^2 + c\dot{r}(\theta_u, X_{t_i})^2(\theta_1 - \theta_2)^2,$$

where $\theta_u \in [\theta_1, \theta_2]$. Using (127), it is upper bounded by

$$c\Delta_{n,i}^2[\dot{b}(X_{t_i}, \theta_u)^2 + (1 + |X_{t_i}|^c)^2](\theta_1 - \theta_2)^2. \tag{130}$$

Replacing (130) in (129), using that the indicator function is bounded by a constant, the Assumption 5 on a and that $\dot{b}$ has polynomial growth, we get that (129) is upper bounded by

$$\frac{c}{n^2\Delta_n^2} \sum_{i=0}^{n-1} \mathbb{E}[(\zeta_i \varphi_{\Delta_{n,i}^{\beta}}(X_{t_{i+1}} - X_{t_i}) - \mathbb{E}[\zeta_i \varphi_{\Delta_{n,i}^{\beta}}(X_{t_{i+1}} - X_{t_i}) | \mathcal{F}_{t_i}])^2 (1 + |X_{t_i}|^c)^2](\theta_1 - \theta_2)^2 =$$

$$= \frac{c}{n^2\Delta_n^2} \sum_{i=0}^{n-1} \mathbb{E}[\mathbb{E}[(\zeta_i \varphi_{\Delta_{n,i}^{\beta}}(X_{t_{i+1}} - X_{t_i}) - \mathbb{E}[\zeta_i \varphi_{\Delta_{n,i}^{\beta}}(X_{t_{i+1}} - X_{t_i}) | \mathcal{F}_{t_i}])^2 | \mathcal{F}_{t_i}](1 + |X_{t_i}|^c)^2](\theta_1 - \theta_2)^2, \tag{131}$$

by the definition of conditional expected value and the measurability of X_{t_i}.
We observe that $\mathbb{E}[(\zeta_i \varphi_{\Delta_{n,i}^{\beta}}(X_{t_{i+1}} - X_{t_i}) - \mathbb{E}[\zeta_i \varphi_{\Delta_{n,i}^{\beta}}(X_{t_{i+1}} - X_{t_i}) | \mathcal{F}_{t_i}])^2 | \mathcal{F}_{t_i}]$ is the conditional variance of $\zeta_i \varphi$ and so it is always smaller then $\mathbb{E}[\zeta_i^2 \varphi_{\Delta_{n,i}^{\beta}}^2(X_{t_{i+1}} - X_{t_i}) | \mathcal{F}_{t_i}]$ that is, using (32), $R(\theta, \Delta_{n,i}, X_{t_i})$. We get that (131) is upper bounded by

$$\frac{1}{n^2\Delta_n^2} \sum_{i=0}^{n-1} \mathbb{E}[R(\theta, \Delta_{n,i}, X_{t_i})(1 + |X_{t_i}|^c)^2](\theta_1 - \theta_2)^2 \le \frac{1}{n\Delta_n} c(\theta_1 - \theta_2)^2,$$

where in the last inequality we have used (11) in order to say that $R(\theta, \Delta_{n,i}, X_{t_i}) = \Delta_{n,i} R(\theta, 1, X_{t_i})$, the fact that $|\Delta_{n,i}| \le \Delta_n$, the natural polynomial growth of the function derived from its definition (10) and the third point of Lemma 2 in order to assure the boundedness of the expected value.
Hence, recalling that $n\Delta_n \to \infty$, we get (124) since $\frac{1}{n\Delta_n} c(\theta_1 - \theta_2)^2 \le c(\theta_1 - \theta_2)^2$.

Concerning (123), we act exactly like we have already done in order to prove (124), getting $\mathbb{E}[(S_{n2}(\theta))^2] \le c(\theta - \theta_0)^2$. Θ is a compact set and so Θ's diameter $d := \sup_{\theta_1, \theta_2 \in \Theta} |\theta_1 - \theta_2|$ is $< \infty$. We therefore deduce (123): $c(\theta - \theta_0)^2 \le cd^2 \le c$.
The tightness of $S_n(\theta) = \frac{U_n(\theta) - U_n(\theta_0)}{t_n}$ follows. □

We are now ready to show the consistence of the estimator $\hat{\theta}_n := arg \min_{\theta \in \Theta} U_n(\theta)$.
We want to prove that $\hat{\theta}_n \xrightarrow{\mathbb{P}} \theta_0$ when $n \to \infty$, that is equivalent to show that $\forall \left\{ \hat{\theta}_{n_k} \right\} \subset \hat{\theta}_n$, $\exists \left\{ \hat{\theta}_{n_{k_j}} \right\} \subset \left\{ \hat{\theta}_{n_k} \right\}$ such that $\hat{\theta}_{n_{k_j}} \to \theta_0$ a.s.
Let $\left\{ \hat{\theta}_{n_k} \right\}$ be a subsequence of $\left\{ \hat{\theta}_n \right\}$. By the uniform convergence in probability of the contrast function given by Lemma 4 and Lemma 5, we get the a.s. convergence along some subsequence of n_k, denoted n_{k_j}:

$$\sup_{\theta \in \Theta} |\frac{U_{n_{k_j}}(\theta) - U_{n_{k_j}}(\theta_0)}{t_{n_{k_j}}} - l(\theta, \theta_0)| \xrightarrow{a.s.} 0, \quad n_{k_j} \to \infty,$$

where $l(\theta, \theta_0) = \int_{\mathbb{R}} \frac{(b(x,\theta) - b(x, \theta_0))^2}{a^2(x)} \pi(dx) \ge 0$.
Now, for fixed $\omega \in \Omega$, thanks to the compactness of Θ, there exists a subsequence of n_{k_j}, that we still denote n_{k_j}, and a θ_∞ such that $\hat{\theta}_{n_{k_j}} \to \theta_\infty$.
Since the mapping $\theta \mapsto l(\theta, \theta_0)$ is continuous, we have $l(\hat{\theta}_{n_{k_j}}, \theta_0) \to l(\theta_\infty, \theta_0)$.
Then, by the definition of $\hat{\theta}_n$ as the argmin of $U_n(\theta)$, we have

$$0 \ge \frac{U_{n_{k_j}}(\hat{\theta}_{n_{k_j}}) - U_{n_{k_j}}(\theta_0)}{t_{n_{k_j}}} \to l(\theta_\infty, \theta_0) \ge 0$$

and so $l(\theta_\infty, \theta_0) = 0$. The Assumption 6 of identifiability leads that $\theta_\infty = \theta_0$.
This implies that any convergent subsequence of $\hat{\theta}_n$ tends to θ_0; this means the consistency of $\hat{\theta}_n$.

7.8 Contrast's derivatives convergence

We are now ready to show the convergence of the derivative of the contrast function through the following lemma:

Lemma 6. *Suppose that Assumptions 1 - 8 and A_β are satisfied. Then*

$$\frac{\dot{U}_n(\theta_0)}{\sqrt{t_n}} \xrightarrow{\mathcal{L}} N(0, 4 \int_{\mathbb{R}} (\frac{\dot{b}(x, \theta_0)}{a(x)})^2 \pi(dx)).$$

Proof. We recall that

$$U_n(\theta_0) = \sum_{i=0}^{n-1} \frac{(X_{t_{i+1}} - m_{\theta_0}(X_{t_i}))^2}{a^2(X_{t_i}) \Delta_{n,i}} \varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}} - X_{t_i}) 1_{\{|X_{t_i}| \le \Delta_{n,i}^{-k}\}},$$

hence

$$\dot{U}_n(\theta_0) = 2 \sum_{i=0}^{n-1} \frac{(X_{t_{i+1}} - m_{\theta_0}(X_{t_i})) \dot{m}_{\theta_0}(X_{t_i})}{a^2(X_{t_i}) \Delta_{n,i}} \varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}} - X_{t_i}) 1_{\{|X_{t_i}| \le \Delta_{n,i}^{-k}\}}. \tag{132}$$

It means that

$$\frac{\dot{U}_n(\theta_0)}{\sqrt{t_n}} = \frac{2}{\sqrt{t_n}} \sum_{i=0}^{n-1} (X_{t_{i+1}} - m_{\theta_0}(X_{t_i})) \frac{\dot{b}(X_{t_i}, \theta_0)}{a^2(X_{t_i})} \varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}} - X_{t_i}) 1_{\{|X_{t_i}| \le \Delta_{n,i}^{-k}\}} +$$
$$+ \frac{2}{\sqrt{t_n}} \sum_{i=0}^{n-1} (X_{t_{i+1}} - m_{\theta_0}(X_{t_i})) \frac{R(\theta_0, \Delta_{n,i}^{\frac{1}{2} \wedge (1 - \alpha\beta - \epsilon - \beta)}, X_{t_i})}{a^2(X_{t_i})} \varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}} - X_{t_i}) 1_{\{|X_{t_i}| \le \Delta_{n,i}^{-k}\}}, \tag{133}$$

where we have used the development (120) of $\dot{m}_\theta(X_{t_i})$.
We now use Proposition 5 on the first term of (133), getting that it converges in distribution to a Gaussian random variable with mean 0 and variance $\int_{\mathbb{R}} \frac{4\dot{b}^2(x,\theta_0)}{a^4(x)} a^2(x)\pi(dx) = \int_{\mathbb{R}} 4(\frac{\dot{b}(x,\theta_0)}{a(x)})^2\pi(dx)$, as we wanted. In order to get the thesis we want to show that the second term of (133) goes to zero in probability as $t_n \to \infty$. In order to do this, we we want to use Lemma 9 of (Genon Catalot & Jacod 1993)) and so we have to prove the following:

$$\frac{2}{\sqrt{t_n}} \sum_{i=0}^{n-1} \mathbb{E}[(X_{t_{i+1}} - m_{\theta_0}(X_{t_i})) \frac{R(\theta_0, \Delta_{n,i}^{\frac{1}{2}\wedge(1-\alpha\beta-\epsilon-\beta)}, X_{t_i})}{a^2(X_{t_i})} \varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}} - X_{t_i}) 1_{\{|X_{t_i}| \le \Delta_{n,i}^{-k}\}} | \mathcal{F}_{t_i}] \to 0 \quad (134)$$

$$\frac{4}{t_n} \sum_{i=0}^{n-1} \mathbb{E}[((X_{t_{i+1}} - m_{\theta_0}(X_{t_i})) \frac{R(\theta_0, \Delta_{n,i}^{\frac{1}{2}\wedge(1-\alpha\beta-\epsilon-\beta)}, X_{t_i})}{a^2(X_{t_i})} \varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}} - X_{t_i}) 1_{\{|X_{t_i}| \le \Delta_{n,i}^{-k}\}})^2 | \mathcal{F}_{t_i}] \to 0 \quad (135)$$

Using the measurability and the fact that

$$\mathbb{E}[(X_{t_{i+1}} - m_{\theta_0}(X_{t_i})) \varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}} - X_{t_i}) | \mathcal{F}_{t_i}] = 0 \quad (136)$$

we get (134). Let us consider (135). Using the Assumption 5 on a, the measurability of R and the expression (33) we can upper bound it with

$$\frac{c}{n\Delta_n} \sum_{i=0}^{n-1} R(\theta_0, \Delta_{n,i}^{1\wedge 2(1-\alpha\beta-\epsilon-\beta)}, X_{t_i}) R(\theta_0, \Delta_{n,i}, X_{t_i}) \le \Delta_n^{1\wedge 2(1-\alpha\beta-\epsilon-\beta)} \frac{1}{n} \sum_{i=0}^{n-1} R(\theta_0, 1, X_{t_i}),$$

that goes to zero in norm 1 by the polynomial growth of R, the third point of Lemma 2 and A_β. Therefore it converges to zero also in probability.
It follows that

$$\frac{2}{\sqrt{t_n}} \sum_{i=0}^{n-1} (X_{t_{i+1}} - m_{\theta_0}(X_{t_i})) \frac{R(\theta_0, \Delta_{n,i}^{\frac{1}{2}\wedge(1-\alpha\beta-\epsilon-\beta)}, X_{t_i})}{a^2(X_{t_i})} \varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}} - X_{t_i}) 1_{\{|X_{t_i}| \le \Delta_{n,i}^{-k}\}} \xrightarrow{\mathbb{P}} 0,$$

as we wanted. ☐

Concerning the second derivative of the contrast function, we have the following convergence:

Lemma 7. *Suppose that Assumptions 1 - 8 and A_β hold. Then*

$$\frac{\ddot{U}_n(\theta_0)}{t_n} \xrightarrow{\mathbb{P}} -2 \int_{\mathbb{R}} (\frac{\dot{b}(x,\theta_0)}{a(x)})^2 \, \pi(dx).$$

Proof. Derivating twice the expression of U_n we get

$$\ddot{U}_n(\theta_0) = -2 \sum_{i=0}^{n-1} \frac{\dot{m}_{\theta_0}^2(X_{t_i})}{a^2(X_{t_i})\Delta_{n,i}} \varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}} - X_{t_i}) 1_{\{|X_{t_i}| \le \Delta_{n,i}^{-k}\}} +$$

$$+2 \sum_{i=0}^{n-1} \frac{(X_{t_{i+1}} - m_{\theta_0}(X_{t_i}))\ddot{m}_{\theta_0}(X_{t_i})}{a^2(X_{t_i})\Delta_{n,i}} \varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}} - X_{t_i}) 1_{\{|X_{t_i}| \le \Delta_{n,i}^{-k}\}} \quad (137)$$

First of all we show that the second term of (137), divided by $n\Delta_n$, goes to zero in probability. We use again Lemma 9 of (Genon Catalot & Jacod 1993). Hence, our goal is to prove the following:

$$\frac{2}{t_n} \sum_{i=0}^{n-1} \mathbb{E}[\frac{(X_{t_{i+1}} - m_{\theta_0}(X_{t_i}))\ddot{m}_{\theta_0}(X_{t_i})}{a^2(X_{t_i})\Delta_{n,i}} \varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}} - X_{t_i}) 1_{\{|X_{t_i}| \le \Delta_{n,i}^{-k}\}} | \mathcal{F}_{t_i}] \to 0 \quad (138)$$

$$\frac{4}{(t_n)^2} \sum_{i=0}^{n-1} \mathbb{E}[(\frac{(X_{t_{i+1}} - m_{\theta_0}(X_{t_i}))\ddot{m}_{\theta_0}(X_{t_i})}{a^2(X_{t_i})\Delta_{n,i}} \varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}} - X_{t_i}) 1_{\{|X_{t_i}| \le \Delta_{n,i}^{-k}\}})^2 | \mathcal{F}_{t_i}] \to 0 \quad (139)$$

As we acted in the last proof, we use (136) in order to get (138).
Concerning (139), using Assumption 5 on a, the measurability of R, the development (121) of $\ddot{m}_{\theta_0}(X_{t_i})$ and the expression (33) we can upper bound it with

$$\frac{c}{n^2\Delta_n^2} \sum_{i=0}^{n-1} [R(\theta_0, \Delta_{n,i}, X_{t_i}) \frac{\Delta_{n,i}^2 \ddot{b}^2(X_{t_i}, \theta_0) + R(\theta_0, \Delta_{n,i}^{3\wedge 2(2-\alpha\beta-\epsilon-\beta)}, X_{t_i})}{\Delta_{n,i}^2}] \le$$

$$\leq \frac{c}{n^2\Delta_n}\sum_{i=0}^{n-1} R(\theta_0, 1, X_{t_i}),$$

where in the last inequality we have used the polynomial growth of $\ddot{b}$, the property (11) on R and that $|\Delta_{n,i}| \leq \Delta_n$. Since $n\Delta_n \to \infty$ and $\frac{1}{n}\sum_{i=0}^{n-1} R(\theta_0, 1, X_{t_i})$ is bounded in L^1, we get the convergence en probability wanted.
Let us now consider the first term of (137). Using the development (120) we get

$$\frac{-2}{t_n}\sum_{i=0}^{n-1}\frac{(\Delta_{n,i}\dot{b}(X_{t_i},\theta_0) + R(\theta_0, \Delta_{n,i}^{\frac{3}{2}\wedge(2-\beta-\beta\alpha-\epsilon)}, X_{t_i}))^2}{a^2(X_{t_i})\Delta_{n,i}}\varphi_{\Delta_{n,i}^{\beta}}(X_{t_{i+1}} - X_{t_i})1_{\{|X_{t_i}|\leq \Delta_{n,i}^{-k}\}}. \tag{140}$$

Hence, we obtain three terms by expanding the square. Using on the first Proposition 3, we get the convergence

$$\frac{-2}{t_n}\sum_{i=0}^{n-1}\frac{\Delta_{n,i}\dot{b}^2(X_{t_i},\theta_0)}{a^2(X_{t_i})}\varphi_{\Delta_{n,i}^{\beta}}(X_{t_{i+1}} - X_{t_i})1_{\{|X_{t_i}|\leq \Delta_{n,i}^{-k}\}} \xrightarrow{\mathbb{P}} -2\int_{\mathbb{R}}\frac{\dot{b}^2(x,\theta_0)}{a^2(x)}\pi(dx). \tag{141}$$

The second term of (140) is

$$\frac{-4}{t_n}\sum_{i=0}^{n-1}\frac{2\dot{b}(X_{t_i},\theta_0)R(\theta_0, \Delta_{n,i}^{\frac{3}{2}\wedge(2-\beta-\beta\alpha-\epsilon)}, X_{t_i})}{a^2(X_{t_i})}\varphi_{\Delta_{n,i}^{\beta}}(X_{t_{i+1}} - X_{t_i})1_{\{|X_{t_i}|\leq \Delta_{n,i}^{-k}\}}.$$

Using Assumption 5 on a, the fact that both φ and the indicator function are bounded, the polynomial growth of both $\dot{b}$ and R and the third point of Lemma 2 we get that its L^1 norm is upper bounded by $c\Delta_n^{\frac{1}{2}\wedge(1-\beta-\beta\alpha-\epsilon)}$. Since the exponent on Δ_n is positive, the convergence in norm L^1 and therefore in probability follows.
Concerning the last term of (140), using again Assumption 5 on a, the fact that both φ and the indicator function are bounded, the polynomial growth of R and the third point of Lemma 2 we get that its L^1 norm is upper bounded by $c\Delta_n^{1\wedge(2-2\beta-2\beta\alpha-2\epsilon)}$. Once again, since the exponent on Δ_n is positive, the convergence in norm L^1 and therefore in probability follows.
It yields

$$\frac{\ddot{U}_n(\theta_0)}{t_n} \xrightarrow{\mathbb{P}} -2\int_{\mathbb{R}}\frac{\dot{b}^2(x,\theta_0)}{a^2(x)}\pi(dx).$$

□

7.9 Asymptotic normality of the estimator

In order to show the asymptotic normality of the estimator we need the following lemma:

Lemma 8. *Suppose that Assumptions 1 - 8 and A_β hold. Then*

$$\frac{1}{t_n}\sup_{t\in[0,1]}|\ddot{U}_n(\theta_0 + t(\hat{\theta}_n - \theta_0)) - \ddot{U}_n(\theta_0)| \xrightarrow{\mathbb{P}} 0, \tag{142}$$

where $\hat{\theta}_n$ is the estimator defined in (9).

Proof. Let us define

$$\tilde{\theta}_n := \theta_0 + t(\hat{\theta}_n - \theta_0). \tag{143}$$

Using (137),

$$\begin{aligned}\frac{\ddot{U}_n(\tilde{\theta}_n) - \ddot{U}_n(\theta_0)}{t_n} &= -\frac{2}{t_n}\sum_{i=0}^{n-1}\frac{(\dot{m}^2_{\tilde{\theta}_n}(X_{t_i}) - \dot{m}^2_{\theta_0}(X_{t_i}))}{a^2(X_{t_i})\Delta_{n,i}}\varphi_{\Delta_{n,i}^{\beta}}(X_{t_{i+1}} - X_{t_i})1_{\{|X_{t_i}|\leq \Delta_{n,i}^{-k}\}}+\\ &+\frac{2}{t_n}\sum_{i=0}^{n-1}\frac{(X_{t_{i+1}} - m_{\theta_0}(X_{t_i}))(\ddot{m}_{\tilde{\theta}_n}(X_{t_i}) - \ddot{m}_{\theta_0}(X_{t_i}))}{a^2(X_{t_i})\Delta_{n,i}}\varphi_{\Delta_{n,i}^{\beta}}(X_{t_{i+1}} - X_{t_i})1_{\{|X_{t_i}|\leq \Delta_{n,i}^{-k}\}}+\\ &+\frac{2}{t_n}\sum_{i=0}^{n-1}\frac{(m_{\theta_0}(X_{t_i}) - m_{\tilde{\theta}_n}(X_{t_i}))\ddot{m}_{\tilde{\theta}_n}(X_{t_i})}{a^2(X_{t_i})\Delta_{n,i}}\varphi_{\Delta_{n,i}^{\beta}}(X_{t_{i+1}} - X_{t_i})1_{\{|X_{t_i}|\leq \Delta_{n,i}^{-k}\}}.\end{aligned} \tag{144}$$

Concerning the first term of (144), we use the following estimation:

$$|\dot{m}^2_{\tilde{\theta}_n}(X_{t_i}) - \dot{m}^2_{\theta_0}(X_{t_i})| \leq 2|\ddot{m}_{\theta_u}(X_{t_i})\dot{m}_{\theta_u}(X_{t_i})(\tilde{\theta}_n - \theta_0)|, \tag{145}$$

where $\theta_u \in [\theta_0, \tilde{\theta}_n]$. We replace the development (120) and (121) of $\dot{m}$ and $\ddot{m}$. Hence the first term of (144) is, in module, upper bounded by

$$\frac{2}{n}\sum_{i=0}^{n-1}|2(\dot{b}(X_{t_i},\theta_u)+R(\theta_u,\Delta_{n,i}^{\frac{1}{2}\wedge(1-\beta-\beta\alpha-\epsilon)},X_{t_i}))(\ddot{b}(X_{t_i},\theta_u)+R(\theta_u,\Delta_{n,i}^{\frac{1}{2}\wedge(1-\beta-\beta\alpha-\epsilon)},X_{t_i}))||\tilde{\theta}_n-\theta_0| = \tag{146}$$

$$= \frac{1}{n}\sum_{i=0}^{n-1}|R(\theta_u,1,X_{t_i})||\tilde{\theta}_n-\theta_0| \le \frac{1}{n}\sum_{i=0}^{n-1}c(1+|X_{t_i}|^c)|\hat{\theta}_n-\theta_0|,$$

where we have used Assumption 5 on a, the boundedness of both φ and the indicator function, the property (11) on R that $|\Delta_{n,i}| \le \Delta_n$ and the definition (143) of $\tilde{\theta}_n$ joint with the fact that $|t| \le 1$. By the consistency of $\hat{\theta}_n$ that we have already proved, we get that the first term of (144) converges to zero in probability uniformly in t, since the right hand side of (146) is bounded in L^1 by the third point of Lemma 2 and it does not depend on t.
On the third term of (144) we use again the Assumption 5 on a, the fact that both φ and the indicator function are bounded, the development (121) of $\ddot{m}_\theta$ and the following estimation: $|m_{\theta_0}(X_{t_i})-m_{\tilde{\theta}_n}(X_{t_i})| \le |\dot{m}_{\theta_u}(X_{t_i})||\theta_0-\tilde{\theta}_n|$, on which we can use the development (120) of $\dot{m}_\theta$. We can hence upper bound the third term with

$$\frac{2}{n}\sum_{i=0}^{n-1}2|(\dot{b}(X_{t_i},\theta_u)+R(\theta_u,\Delta_{n,i}^{\frac{1}{2}\wedge(1-\beta-\beta\alpha-\epsilon)},X_{t_i}))(\ddot{b}(X_{t_i},\tilde{\theta}_n)+R(\tilde{\theta}_n,\Delta_{n,i}^{\frac{1}{2}\wedge(1-\beta-\beta\alpha-\epsilon)},X_{t_i}))||\theta_0-\tilde{\theta}_n| = \tag{147}$$

$$= \frac{1}{n}\sum_{i=0}^{n-1}|R(\theta,1,X_{t_i})||\tilde{\theta}_n-\theta_0| \le \frac{1}{n}\sum_{i=0}^{n-1}c(1+|X_{t_i}|^c)|\hat{\theta}_n-\theta_0|.$$

The consistency of $\hat{\theta}_n$ yields the convergence in probability uniformly in t wanted, by the boundedness in L^1 of the sum, that does not depend on t.
It remains to prove the convergence to zero, uniformly in t, for the second term of (144); it is sufficient to prove that the following sequence $S_n(\theta)$ converges to zero uniformly with respect to θ:

$$S_n(\theta) := \frac{2}{t_n}\sum_{i=0}^{n-1}\frac{(X_{t_{i+1}}-m_{\theta_0}(X_{t_i}))(\ddot{m}_\theta(X_{t_i})-\ddot{m}_{\theta_0}(X_{t_i}))}{a^2(X_{t_i})\Delta_{n,i}}\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})1_{\left\{|X_{t_i}|\le\Delta_{n,i}^{-k}\right\}}.$$

The pointwise convergence is already proved (it is enough to repeat the proof of (138) and (139) with $\ddot{m}_\theta(X_{t_i})-\ddot{m}_{\theta_0}(X_{t_i})$ in place of $\ddot{m}_{\theta_0}(X_{t_i})$). In order to show that the convergence takes place uniformly in θ, we prove the tightness of $S_n(\theta)$, using the criterion analogues to (123) and (124).
Let us consider (124) first. We observe that

$$\mathbb{E}[(S_n(\theta_1)-S_n(\theta_2))^2] \le$$

$$\le \frac{c}{n^2\Delta_n^2}\mathbb{E}[(\sum_{i=0}^{n-1}\frac{(X_{t_{i+1}}-m_{\theta_0}(X_{t_i}))(\ddot{m}_{\theta_1}(X_{t_i})-\ddot{m}_{\theta_2}(X_{t_i}))}{a^2(X_{t_i})\Delta_{n,i}}\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})1_{\left\{|X_{t_i}|\le\Delta_{n,i}^{-k}\right\}})^2]. \tag{148}$$

By the building the sum is a square integrable martingale. The Pythagoras' theorem on a square integrable martingale yields that (148) is equal to

$$\frac{c}{n^2\Delta_n^2}\mathbb{E}[\sum_{i=0}^{n-1}\frac{(X_{t_{i+1}}-m_{\theta_0}(X_{t_i}))^2(\ddot{m}_{\theta_1}(X_{t_i})-\ddot{m}_{\theta_2}(X_{t_i}))^2}{a^4(X_{t_i})\Delta_{n,i}^2}\varphi^2_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})1_{\left\{|X_{t_i}|\le\Delta_{n,i}^{-k}\right\}}]. \tag{149}$$

We now use the following estimation:

$$|\ddot{m}_{\theta_1}(X_{t_i})-\ddot{m}_{\theta_2}(X_{t_i})| \le |\dddot{m}_{\theta_u}(X_{t_i})||\theta_1-\theta_2|. \tag{150}$$

Replacing (150) in (148) and using (122) on $\dddot{m}_{\theta_u}(X_{t_i})$, we can upper bound (148) with

$$\frac{c}{n^2\Delta_n^2}\mathbb{E}[\sum_{i=0}^{n-1}\frac{(X_{t_{i+1}}-m_{\theta_0}(X_{t_i}))^2R(\theta_u,\Delta_{n,i}^2,X_{t_i})}{a^4(X_{t_i})\Delta_{n,i}^2}\varphi^2_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})1_{\left\{|X_{t_i}|\le\Delta_{n,i}^{-k}\right\}}](\theta_1-\theta_2)^2 \le$$

$$\le \frac{c}{n^2\Delta_n^2}\mathbb{E}[\sum_{i=0}^{n-1}f(X_{t_i},\theta_u)\mathbb{E}[(X_{t_{i+1}}-m_{\theta_0}(X_{t_i}))^2\varphi^2_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})1_{\left\{|X_{t_i}|\le\Delta_{n,i}^{-k}\right\}}|\mathcal{F}_{t_i}]](\theta_1-\theta_2)^2, \tag{151}$$

with $f(X_{t_i},\theta_u) = \frac{R(\theta_u,1,X_{t_i})}{a^4(X_{t_i})}$ and where we have used the property (11) of the functions R and the definition of conditional expected value.
Using (33), the property (11) and that $|\Delta_{n,i}| \le \Delta_n$, we can upper bound (151) with $\frac{4}{n^2\Delta_n}\sum_{i=0}^{n-1}\mathbb{E}[f(X_{t_i},\theta_u)R(\theta_0,1,X_{t_i})](\theta_1-\theta_2)^2$.
By the Assumption 5 on a and the polynomial growth of R derived by its definition, f has polynomial growth. Using the third point of Lemma 2 we get that the expected value is bounded. Hence, since $n\Delta_n \to \infty$, it yields

$$\frac{4}{n^2\Delta_n}\sum_{i=0}^{n-1}\mathbb{E}[f(X_{t_i},\theta_u)R(\theta_0,1,X_{t_i})] \le c, \tag{152}$$

therefore we obtain (124) on S_n.
Concerning (123), we can act exactly in the same way, using (152) and the compactness of Θ. The tightness of $S_n(\theta)$ follows. □

We are now ready to prove the asymptotic normality of the estimator. Using (142) we have that

$$\frac{1}{t_n}\int_0^1 [\ddot{U}_n(\theta_0+t(\hat{\theta}_n-\theta_0))-\ddot{U}_n(\theta_0)]dt \xrightarrow{\mathbb{P}} 0. \tag{153}$$

We observe that

$$\frac{1}{t_n}\int_0^1 [\ddot{U}_n(\theta_0+t(\hat{\theta}_n-\theta_0))]dt\sqrt{t_n}(\hat{\theta}_n-\theta_0) =$$

$$= \frac{1}{\sqrt{t_n}}\int_0^1 [\ddot{U}_n(\theta_0+t(\hat{\theta}_n-\theta_0))]dt\,(\hat{\theta}_n-\theta_0) = \frac{1}{\sqrt{t_n}}(\dot{U}_n(\hat{\theta}_n)-\dot{U}_n(\theta_0)) = -\frac{\dot{U}_n(\theta_0)}{\sqrt{t_n}}, \tag{154}$$

where in the last equality we have used that, on the set $\left\{\hat{\theta}_n \in \mathring{\Theta}\right\}$, $\dot{U}_n(\hat{\theta}_n) = 0$ since $\hat{\theta}_n$ is a minimum. Hence

$$\sqrt{t_n}(\hat{\theta}_n-\theta_0) = \frac{-\frac{\dot{U}_n(\theta_0)}{\sqrt{t_n}}}{\frac{1}{t_n}\int_0^1[\ddot{U}_n(\theta_0+t(\hat{\theta}_n-\theta_0))]dt}. \tag{155}$$

Using Lemma 6 we have the convergence in distribution of the numerator of (155) to $N(0, 4\int_{\mathbb{R}}(\frac{\dot{b}(x,\theta_0)}{a(x)})^2\pi(dx))$ and, by the equation (153), the denominator converges in probability to $-2\int_{\mathbb{R}}(\frac{\dot{b}(x,\theta_0)}{a(x)})^2\pi(dx)$.
Therefore$\sqrt{t_n}(\hat{\theta}_n-\theta_0)$ converges in distribution to $N(0, \frac{4\int_{\mathbb{R}}(\frac{\dot{b}(x,\theta_0)}{a(x)})^2\pi(dx)}{4(\int_{\mathbb{R}}(\frac{\dot{b}(x,\theta_0)}{a(x)})^2\pi(dx))^2})$, i. e. it is $N(0, (\int_{\mathbb{R}}(\frac{\dot{b}(x,\theta_0)}{a(x)})^2\pi(dx))^{-1})$, as we wanted.

7.10 Proof of Proposition 1

The proof of the proposition is essentially similar to the the proof of the asymptotic normality of $\hat{\theta}_n$ given in Sections 7.6–7.9 and we skip it. The main difference comes from the fact that Proposition 5 holds true with $\widetilde{m}_{\theta_0}(X_{t_i})$ replacing $m_{\theta_0}(X_{t_i})$ under the condition that $\sqrt{n}\Delta_n^{\rho-1/2} \to 0$.

References

Aït-Sahalia, Y., & Yu, J. (2006). Saddlepoint Approximations for Continuous-Time Markov Processes, Journal of Econometrics, 134, 507–551.

Applebaum, D. L. processes, & stochastic calculus. Cambridge university press. (2009).

Barndorff-Nielsen, O. E., & Shephard, N. (2001). Non-Gaussian Ornstein-Uhlenbeck-Based Models and Some of Their Uses in Financial Economics. J. R. Stat. Soc., Ser. B, Stat. Methodol., 63, 167-241.

Bates, D. S. (1996). Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark. The Review of Financial Studies, 9(1), 69-107.

Bichteler, K. (1987). Malliavin calculus for processes with jumps. Stochastics Monographs.

Ditlevsen, S., & Greenwood, P. (2013). The Morris–lecar Neuron Model Embeds a Leaky Integrate-And-Fire Model. Journal of Mathematical Biology 67 239-259.

Eraker, B., Johannes, M., & N, P. (2003). The Impact of Jumps in Volatility and Returns. J. Finance, 58(3), 1269.

Florens Zmirou, Danielle. "Approximate discrete-time schemes for statistics of diffusion processes." Statistics: A Journal of Theoretical and Applied Statistics 20.4 (1989): 547-557.

Genon Catalot, V. and Jacod, J. (1993). On the estimation of the diffusion coefficient for multi- dimensional diffusion processes. Annales de l'institut Henri Poincaré (B) Probabilités et Statistiques, 29, 119-151.

Gloter, A., Loukianova, D., & Mai, H. (2018). Jump Filtering and Efficient Drift Estimation for Lévy-Driven Sdes. The Annals of Statistics, 46(4), 1445.

Ibragimov, I. A., & Has' Minskii, R. Z. (2013). Statistical Estimation: Asymptotic Theory (vol. 16). Springer Science and Business Media.

Jacod, J., & Protter, P. (2011). Discretization of Processes (vol. 67). Springer Science and Business Media.

Jacod, J., and Shiryaev, A. (2013). Limit theorems for stochastic processes (Vol. 288). Springer Science and Business Media.

Jakobsen, N., & Sørensen, M. (2017). Estimating Functions for Jump-Diffusions, Preprint.

Kessler, M. (1997). Estimation of an Ergodic Diffusion from Discrete Observations. Scandinavian Journal of Statistics, 24(2), 211-229.

Kou, S.G. (2002). A Jump Diffusion Model for Option Pricing. Management Science, 48, 1086-1101.

Li, C., & Chen, D. (2016). Estimating Jump-Diffusions Using Closed Form Likelihood Expansions. Journal of Econometrics, 195, 51–71.

Masuda, H. (2007). Ergodicity and exponential beta mixing bounds for multidimensional diffusions with jumps. Stochastic processes and their applications, 117(1), 35-56.

Masuda, H. (2009). Erratum to: Ergodicity and exponential beta mixing bound for multidimensional diffusions with jumps, Stochastic Process. Appl. 117 (2007) 35–56. Stochastic Processes and their Applications, 119(2), 676-678.

Masuda, H. (2013). Convergence of Gaussian Quasi-Likelihood Random Fields for Ergodic Lévy Driven Sde Observed At High Frequency. Annals of Stat., 41(3), 1593.

Merton, R. C. (1976). Option Pricing When Underlying Stock Returns Are Discontinuous. Journal of Financial Economics, 3, 125-144.

Nikolskii, S. M. (1977). Approximation of Functions of Several Variables and Imbedding Theorems (russian). Sec. Ed., Moskva, Nauka 1977 English Translation of the First Ed., Berlin 1975.

Protter, P. E. (2005). Stochastic differential equations. In Stochastic integration and differential equations (pp. 249-361). Springer, Berlin, Heidelberg.

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

Shimizu, Y., & Yoshida, N. (2006). Estimation of Parameters for Diffusion Processes with Jumps from Discrete Observations. Statistical Inference for Stochastic Processes, 9(3), 227-277.

Yoshida, N. (1992). Estimation for Diffusion Processes from Discrete Observation. Journal of Multivariate Analysis, 41, 220-242.

LaMME, UMR CNRS 8071
Université d'Evry Val d'Essonne
91037 Évry Cedex
France
E-mail: chiara.amorino@univ-evry.fr
E-mail: arnaud.gloter@univ-evry.fr

A Appendix

In this section we will prove the technical lemmas that we have used in order to show the main theorems.

A.1 Proof of expansions of the derivatives of the function $m_{\theta,h}$

In order to prove the explicit approximation of $\dot{m}_{\theta,h}$ and $\ddot{m}_{\theta,h}$ provided in Proposition 8, the following lemma will be useful. We point out that X_t^θ is $X_t^{\theta,x}$ and so the process starts in 0: $X_0^{\theta,x} = x$.

Lemma 9. *Suppose that Assumptions 1 to 4 and 7 hold. Let us define* $\dot{X}_t^{\theta,x} := \frac{\partial X_t^{\theta,x}}{\partial \theta}$ *and* $\ddot{X}_t^{\theta,x} := \frac{\partial^2 X_t^{\theta,x}}{\partial \theta^2}$. *Then, for all* $p \ge 2$ $\exists c > 0$: $\forall h \le \Delta_n$ $\forall x$,

$$\mathbb{E}[|\frac{\dot{X}_h^{\theta,x}}{h}|^p] \le c(1 + |x|^c), \tag{156}$$

$$\mathbb{E}[|\frac{\ddot{X}_h^{\theta,x}}{h}|^p] \le c(1 + |x|^c). \tag{157}$$

Proof. The dynamic of the process X is known. The same applies to the processes $\dot{X}_t^{\theta,x}$ and $\ddot{X}_t^{\theta,x}$ (cf. (missing citation), section 5).

$$\dot{X}_h^{\theta,x} = \int_0^h (b'(X_s^{\theta,x},\theta)\dot{X}_s^{\theta,x} + \dot{b}(X_s^{\theta,x},\theta))ds + \int_0^h a'(X_s^{\theta,x})\dot{X}_s^{\theta,x} dW_s + \int_0^h \int_{\mathbb{R}} \gamma'(X_{s^-}^{\theta,x})\dot{X}_s^{\theta,x} z\tilde{\mu}(dz,ds) \tag{158}$$

and

$$\ddot{X}_h^{\theta,x} = \int_0^h (b''(X_s^{\theta,x},\theta)(\dot{X}_s^{\theta,x})^2 + 2\dot{b}'(X_s^{\theta,x},\theta)\dot{X}_s^{\theta,x} + b'(X_s^{\theta,x},\theta)\ddot{X}_s^{\theta,x} + \ddot{b}(X_s^{\theta,x},\theta))ds + \tag{159}$$

$$+ \int_0^h (a''(X_s^{\theta,x})(\dot{X}_s^{\theta,x})^2 + a'(X_s^{\theta,x})\ddot{X}_s^{\theta,x})dW_s + \int_0^h \int_{\mathbb{R}} (\gamma''(X_{s^-}^{\theta,x})(\dot{X}_s^{\theta,x})^2 + \gamma'(X_{s^-}^{\theta,x})\ddot{X}_s^{\theta,x}) z\tilde{\mu}(dz,ds).$$

From now on, we will drop the dependence of the starting point in order to make the notation easier. Let us start with the proof of (156). We observe that, taking the L^p norm of (158), we have the following estimation:

$$\mathbb{E}[|\dot{X}_h^\theta|^p] \le c\mathbb{E}[|\int_0^h (b'(X_s^\theta,\theta)\dot{X}_s^\theta + \dot{b}(X_s^\theta,\theta))ds|^p] + c\mathbb{E}[|\int_0^h a'(X_s^\theta)\dot{X}_s^\theta dW_s|^p] + c\mathbb{E}[|\int_0^h \int_{\mathbb{R}} \gamma'(X_{s^-}^\theta)\dot{X}_s^\theta z\tilde{\mu}(dz,ds)|^p]. \tag{160}$$

Concerning the first term of (160),

$$\mathbb{E}[|\int_0^h (b'(X_s^\theta,\theta)\dot{X}_s^\theta + \dot{b}(X_s^\theta,\theta))ds|^p] \le c\mathbb{E}[|\int_0^h b'(X_s^\theta,\theta)\dot{X}_s^\theta ds|^p] + c\mathbb{E}[|\int_0^h \dot{b}(X_s^\theta,\theta)ds|^p].$$

Then, using Jensen inequality on the first, we obtain

$$\mathbb{E}[|\int_0^h b'(X_s^\theta,\theta)\dot{X}_s^\theta ds|^p] = \mathbb{E}[h^p|\frac{1}{h}\int_0^h b'(X_s^\theta,\theta)\dot{X}_s^\theta ds|^p] \le$$

$$\le \mathbb{E}[h^{p-1}\int_0^h |b'(X_s^\theta,\theta)|^p|\dot{X}_s^\theta|^p ds] = h^{p-1}\int_0^h \mathbb{E}[|b'(X_s^\theta,\theta)|^p|\dot{X}_s^\theta|^p]ds.$$

The derivatives of b with respect to x are supposed bounded, it yields

$$\mathbb{E}[|\int_0^h b'(X_s^\theta,\theta)\dot{X}_s^\theta ds|^p] \le ch^{p-1}\int_0^h \mathbb{E}[|\dot{X}_s^\theta|^p]ds. \tag{161}$$

Let us now consider the second term of (160). Using Burkholder-Davis-Gundy and Jensen inequalities we get

$$\mathbb{E}[|\int_0^h a'(X_s^\theta)\dot{X}_s^\theta dW_s|^p] \le c\mathbb{E}[|\int_0^h (a'(X_s^\theta)\dot{X}_s^\theta)^2 ds|^{\frac{p}{2}}] =$$

$$= c\mathbb{E}[h^{\frac{p}{2}}|\frac{1}{h}\int_0^h (a'(X_s^\theta)\dot{X}_s^\theta)^2 ds|^{\frac{p}{2}}] \le ch^{\frac{p}{2}-1}\mathbb{E}[\int_0^h |a'(X_s^\theta)\dot{X}_s^\theta|^p].$$

Therefore

$$\mathbb{E}[|\int_0^h a'(X_s^\theta)\dot{X}_s^\theta dW_s|^p] \le ch^{\frac{p}{2}-1}\int_0^h \mathbb{E}[|\dot{X}_s^\theta|^p]ds, \tag{162}$$

where we have used that the derivatives of a are bounded.
The third term of (160) can be estimed using Kunita inequality (cf. the Appendix of (Jacod & Protter, 2011)):

$$\mathbb{E}[|\int_0^h \int_{\mathbb{R}} \gamma'(X_{s-}^\theta)\dot{X}_s^\theta z\tilde{\mu}(dz,ds)|^p] \leq$$

$$\leq \mathbb{E}[\int_0^h \int_{\mathbb{R}} |\gamma'(X_s^\theta)\dot{X}_s^\theta|^p |z|^p \bar{\mu}(dz,ds)] + \mathbb{E}[|\int_0^h \int_{\mathbb{R}} (\gamma'(X_s^\theta)\dot{X}_s^\theta)^2 z^2 \bar{\mu}(dz,ds)|^{\frac{p}{2}}] \leq$$

$$\leq \int_0^h \mathbb{E}[|\gamma'(X_s^\theta)|^p |\dot{X}_s^\theta|^p](\int_{\mathbb{R}} |z|^p F(z)dz)ds + \mathbb{E}[|\int_0^h (\gamma'(X_s^\theta)\dot{X}_s^\theta)^2 (\int_{\mathbb{R}} z^2 F(z)dz)ds|^{\frac{p}{2}}] \leq$$

$$\leq c\int_0^h \mathbb{E}[|\gamma'(X_s^\theta)|^p |\dot{X}_s^\theta|^p]ds + c\mathbb{E}[|\int_0^h (\gamma'(X_s^\theta)\dot{X}_s^\theta)^2 ds|^{\frac{p}{2}}],$$

where in the last two inequalities we have just used the definition of the compensated measure $\bar{\mu}$ and the third point of Assumption 4.
Since the derivatives of γ are supposed bounded and by the Jensen inequality we get it is upper bounded by

$$c\int_0^h \mathbb{E}[|\dot{X}_s^\theta|^p]ds + c\mathbb{E}[h^{\frac{p}{2}-1}\int_0^h |\gamma'(X_s^\theta)|^p |\dot{X}_s^\theta|^p ds] \leq c\int_0^h \mathbb{E}[|\dot{X}_s^\theta|^p]ds + ch^{\frac{p}{2}-1}\int_0^h \mathbb{E}[|\dot{X}_s^\theta|^p]ds.$$

Hence

$$\mathbb{E}[|\int_0^h \int_{\mathbb{R}} \gamma'(X_{s-}^\theta)\dot{X}_s^\theta z\tilde{\mu}(dz,ds)|^p] \leq c(1+h^{\frac{p}{2}-1})\int_0^h \mathbb{E}[|\dot{X}_s^\theta|^p]ds. \tag{163}$$

From (161), (162) and (163), we obtain

$$\mathbb{E}[|\dot{X}_h^\theta|^p] \leq c\mathbb{E}[|\int_0^h \dot{b}(X_s^\theta,\theta)ds|^p] + c(1+h^{\frac{p}{2}-1}+h^{p-1})\int_0^h \mathbb{E}[|\dot{X}_s^\theta|^p]ds.$$

Let M_h be $\mathbb{E}[|\dot{X}_h^\theta|^p]$, then the equation above can be seen as

$$M_h \leq c\mathbb{E}[|\int_0^h \dot{b}(X_s^\theta,\theta)ds|^p] + c(1+h^{\frac{p}{2}-1}+h^{p-1})\int_0^h M_s ds.$$

Using Gronwall lemma, it yields $M_h \leq c\mathbb{E}[|\int_0^h \dot{b}(X_s^\theta,\theta)ds|^p]e^{ch(1+h^{\frac{p}{2}-1}+h^{p-1})}$.
By the polynomial growth of $\dot{b}$ and the third point of Lemma 1,

$$\mathbb{E}[|\int_0^h \dot{b}(X_s^\theta,\theta)ds|^p] \leq ch^p(1+|X_0^{0,x}|^c) = ch^p(1+|x|^c).$$

Hence $\mathbb{E}[|\dot{X}_h^\theta|^p] \leq ch^p(1+|x|^c)$.
Our goal is now to prove (157). In order to do it, we take the L^p norm of (159), getting the following estimation:

$$\mathbb{E}[|\ddot{X}_h^\theta|^p] \leq \mathbb{E}[|\int_0^h (b''(X_s^\theta,\theta)(\dot{X}_s^\theta)^2 + 2\dot{b}'(X_s^\theta,\theta)\dot{X}_s^\theta + b'(X_s^\theta,\theta)\ddot{X}_s^\theta + \ddot{b}(X_s^\theta,\theta))ds|^p] +$$

$$+\mathbb{E}[|\int_0^h (a''(X_s^\theta)(\dot{X}_s^\theta)^2 + a'(X_s^\theta)\ddot{X}_s^\theta)dW_s|^p] + \mathbb{E}[|\int_0^h \int_{\mathbb{R}} (\gamma''(X_{s-}^\theta)(\dot{X}_s^\theta)^2 + \gamma'(X_{s-}^\theta)\ddot{X}_s^\theta)z\tilde{\mu}(dz,ds)|^p] \tag{164}$$

The first term of (164) is upper bounded by

$$\mathbb{E}[|\int_0^h (b''(X_s^\theta,\theta)(\dot{X}_s^\theta)^2 ds|^p] + \mathbb{E}[|\int_0^h 2\dot{b}'(X_s^\theta,\theta)\dot{X}_s^\theta ds|^p] + \mathbb{E}[|\int_0^h b'(X_s^\theta,\theta)\ddot{X}_s^\theta ds|^p] + \mathbb{E}[|\int_0^h \ddot{b}(X_s^\theta,\theta)ds|^p] \leq$$

$$\leq ch^{p-1}\int_0^h \mathbb{E}[|\dot{X}_s^\theta|^{2p}]ds + ch^{p-1}\int_0^h \mathbb{E}[|\dot{b}'(X_s^\theta,\theta)|^p |\dot{X}_s^\theta|^p]ds + ch^{p-1}\int_0^h \mathbb{E}[|\ddot{X}_s^\theta|^p]ds + \mathbb{E}[|\int_0^h \ddot{b}(X_s^\theta,\theta)ds|^p], \tag{165}$$

where we have used Jensen inequality and that the derivatives of b with respect to x are supposed bounded.
By Holder inequality

$$\mathbb{E}[|\dot{b}'(X_s^\theta,\theta)|^p |\dot{X}_s^\theta|^p] \leq (\mathbb{E}[|\dot{b}'(X_s^\theta,\theta)|^{pp_1}])^{\frac{1}{p_1}}(\mathbb{E}[|\dot{X}_s^\theta|^{pp_2}]^{\frac{1}{p_2}} \leq c(h^{pp_2})^{\frac{1}{p_2}}(1+|x|^c) = ch^p(1+|x|^c),$$

where in the last inequality we have used the boundedness of $\dot{b}'$ and (156). Since $\ddot{b}$ has polynomial growth and by the third point of Lemma 1, $\mathbb{E}[|\int_0^h \ddot{b}(X_s^\theta,\theta))ds|^p] \le ch^p(1+|x|^c)$. Replacing in (165) and using also on its first term (156) we obtain it is upper bounded by

$$\mathbb{E}[|\int_0^h (b''(X_s^\theta,\theta)(\dot{X}_s^\theta)^2 + 2\dot{b}'(X_s^\theta,\theta)\dot{X}_s^\theta + b'(X_s^\theta,\theta)\ddot{X}_s^\theta + \ddot{b}(X_s^\theta,\theta))ds|^p] \le \tag{166}$$

$$\le c(1+|x|^c)(h^{3p}+h^{2p}+h^p) + ch^{p-1}\int_0^h \mathbb{E}[|\ddot{X}_s^\theta|^p]ds.$$

Let us now consider the second term of (164). By Burkholder-Davis-Gundy and Jensen inequalities we get

$$\mathbb{E}[|\int_0^h (a''(X_s^\theta)(\dot{X}_s^\theta)^2 + a'(X_s^\theta)\ddot{X}_s^\theta)dW_s|^p] \le \mathbb{E}[|\int_0^h (a''(X_s^\theta)(\dot{X}_s^\theta)^2 + a'(X_s^\theta)\ddot{X}_s^\theta)^2 ds|^{\frac{p}{2}}] \le$$

$$\le h^{\frac{p}{2}-1}\mathbb{E}[\int_0^h |a''(X_s^\theta)(\dot{X}_s^\theta)^2|^p + |a'(X_s^\theta)\ddot{X}_s^\theta|^p ds] \le ch^{\frac{p}{2}+2p}(1+|x|^c) + ch^{\frac{p}{2}-1}\int_0^h \mathbb{E}[|\ddot{X}_s^\theta|^p]ds, \tag{167}$$

where in the last inequality we have used that the derivatives of a are supposed bounded and (156). Concerning the last term of (164), by Kunita inequality it is upper bounded by

$$\mathbb{E}[\int_0^h \int_{\mathbb{R}} |\gamma''(X_s^\theta)(\dot{X}_s^\theta)^2 + \gamma'(X_s^\theta)\ddot{X}_s^\theta|^p |z|^p \bar{\mu}(dz,ds)] + \mathbb{E}[|\int_0^h \int_{\mathbb{R}} (\gamma''(X_s^\theta)(\dot{X}_s^\theta)^2 + \gamma'(X_s^\theta)\ddot{X}_s^\theta)^2 z^2 \bar{\mu}(dz,ds)|^{\frac{p}{2}}] \le$$

$$\le c\int_0^h \mathbb{E}[|\gamma''(X_s^\theta)(\dot{X}_s^\theta)^2 + \gamma'(X_s^\theta)\ddot{X}_s^\theta|^p ds] + h^{\frac{p}{2}-1}\int_0^h \mathbb{E}[|\gamma''(X_s^\theta)(\dot{X}_s^\theta)^2 + \gamma'(X_s^\theta)\ddot{X}_s^\theta|^p ds],$$

having used Jensen inequality and the third point of Assumption 4 in order to say that $\int_{\mathbb{R}} |z|^p F(z)dz < c$. Using (156) and the boundedness of the derivatives of γ, it is upper bounded by

$$c(1+|x|^c)h^{2p+1} + c\int_0^h \mathbb{E}[|\ddot{X}_s^\theta|^p]ds + ch^{\frac{p}{2}+2p}(1+|x|^c) + ch^{\frac{p}{2}-1}\int_0^h \mathbb{E}[|\ddot{X}_s^\theta|^p]ds. \tag{168}$$

From (166), (167) and (168) we get

$$\mathbb{E}[|\ddot{X}_h^\theta|^p] \le c(1+|x|^c)h^p(1+h^p+h^{2p}+h^{p+\frac{p}{2}}+h^{p+1}) + c(1+h^{p-1}+h^{\frac{p}{2}-1})\int_0^h \mathbb{E}[|\ddot{X}_s^\theta|^p]ds.$$

Using Gronwall Lemma we obtain $\mathbb{E}[|\ddot{X}_h^\theta|^p] \le c(1+|x|^c)h^p(1+h^p+h^{2p}+h^{p+\frac{p}{2}}+h^{p+1})$ and so $\mathbb{E}[|\ddot{X}_h^\theta|^p] \le c(1+|x|^c)h^p$, as we wanted. □

Remark 11. *Supposing that the same assumptions as in Lemma 5 hold and acting as we have done in order to get the estimations (156) and (157) it is possible to prove that, for all $p \ge 2$ $\exists c > 0$: $\forall h \le \Delta_n$, $\forall x$,*

$$\mathbb{E}[|\frac{\partial^3}{\partial\theta^3}X_h^{\theta,x}|^p \frac{1}{h^p}] \le c(1+|x|^c). \tag{169}$$

A.1.1 Proof of Proposition 8

Proof. As in the proof of Lemma 9, we drop the dependence on the starting point in order to make the notation easier.
We recall the definition of $m_{\theta,h}(x)$:

$$m_{\theta,h}(x) := \frac{\mathbb{E}[X_h^\theta \varphi_{h^\beta}(X_h^\theta - X_0^\theta)|X_0^\theta = x]}{\mathbb{E}[\varphi_{h^\beta}(X_h^\theta - X_0^\theta)|X_0^\theta = x]} = \frac{\mathbb{E}[X_h^\theta \varphi_{h^\beta}(X_h^\theta - x)]}{\mathbb{E}[\varphi_{h^\beta}(X_h^\theta - x)]}.$$

Its derivative with respect to θ is

$$\frac{\mathbb{E}[\dot{X}_h^\theta \varphi_{h^\beta}(X_h^\theta - x)] + \mathbb{E}[X_h^\theta h^{-\beta}\dot{X}_h^\theta \varphi'_{h^\beta}(X_h^\theta - x)]}{\mathbb{E}[\varphi_{h^\beta}(X_h^\theta - x)]} - m_{\theta,h}(x)\frac{\mathbb{E}[h^{-\beta}\dot{X}_h^\theta \varphi'_{h^\beta}(X_h^\theta - x)]}{\mathbb{E}[\varphi_{h^\beta}(X_h^\theta - x)]}. \tag{170}$$

On the second and on the third term of (170) we divide and we multiply by h and then we use Proposition 7, taking $Z_1 = \frac{\dot{X}_h^\theta}{h}X_h^\theta$ and $Z_2 = \frac{\dot{X}_h^\theta}{h}$, respectively. We are allowed to do that because they are both bounded in L^p, with p arbitrary high, since we can use (156) on Z_2 and Holder inequality, (156) and the third point of Lemma 1 on Z_1. For $|x| \le h^{-k_0}$ we have

$$m_{\theta,h}(x) = x + \frac{\mathbb{E}[(X_h^\theta - x)\varphi_{h^\beta}(X_h^\theta - x)]}{\mathbb{E}[\varphi_{h^\beta}(X_h^\theta - x)]} = R(\theta,1,x), \tag{171}$$

where we have used that k_0 turns out in the proof of theorems 2 and 4, hence it has been chosen such that, for $|x| \le h^{-k_0}$ we have that $\mathbb{E}[\varphi_{h^\beta}(X_h^\theta - x)] \ge \frac{1}{2}$. Moreover the expected value is bounded as a result of the boundedness of φ and the third point of Lemma 1. It yields, for $\epsilon > 0$ arbitrary small,

$$\dot{m}_{\theta,h} = \frac{\mathbb{E}[\dot{X}_h^\theta \varphi_{h^\beta}(X_h^\theta - x)] + R(\theta, h^{2-\alpha\beta-\epsilon-\beta}, x)}{\mathbb{E}[\varphi_{h^\beta}(X_h^\theta - x)]}. \tag{172}$$

Let us now consider the first term. Replacing the dynamic of the process $\dot{X}_h^\theta$, we get

$$\mathbb{E}[\int_0^h (b'(X_s^\theta,\theta)\dot{X}_s^\theta + \dot{b}(X_s^\theta,\theta))ds\, \varphi_{h^\beta}(X_h^\theta - x)] + \mathbb{E}[(\int_0^h a'(X_s^\theta)\dot{X}_s^\theta dW_s +$$

$$+ \int_0^h \int_{\mathbb{R}} \gamma'(X_{s^-}^\theta)\dot{X}_s^\theta z\tilde{\mu}(dz,ds))\, \varphi_{h^\beta}(X_h^\theta - x)] = \mathbb{E}[\int_0^h \dot{b}(X_s^\theta,\theta)ds\, \varphi_{h^\beta}(X_h^\theta - x)] + R(\theta, h^{\frac{3}{2}}, x). \tag{173}$$

In fact, using Holder inequality,

$$|\mathbb{E}[\int_0^h b'(X_s^\theta,\theta)\dot{X}_s^\theta ds\, \varphi_{h^\beta}(X_h^\theta - x)]| \le$$

$$\le (\mathbb{E}[|\int_0^h b'(X_s^\theta,\theta)\dot{X}_s^\theta ds|^p])^{\frac{1}{p}} (\mathbb{E}[\varphi_{h^\beta}^q(X_h^\theta - x)])^{\frac{1}{q}} \le (ch^{p-1}\int_0^h \mathbb{E}[|\dot{X}_s^\theta|^p]ds)^{\frac{1}{p}},$$

where in the last inequality we have used that φ is bounded and (161). By (156), it is upper bounded by $(ch^{2p}(1+|x|^c))^{\frac{1}{p}}$. It turns

$$\mathbb{E}[\int_0^h b'(X_s^\theta,\theta)\dot{X}_s^\theta ds\, \varphi_{h^\beta}(X_h^\theta - x)] = R(\theta, h^2, x). \tag{174}$$

In the same way, from Holder inequality, (162) and the fact that φ is bounded, we get $|\mathbb{E}[\int_0^h a'(X_s^\theta)\dot{X}_s^\theta dW_s\, \varphi_{h^\beta}(X_h^\theta - x)]| \le (ch^{\frac{p}{2}-1}\int_0^h \mathbb{E}[|\dot{X}_s^\theta|^p]ds)^{\frac{1}{p}}$. Using (156), it yields

$$\mathbb{E}[\int_0^h a'(X_s^\theta)\dot{X}_s^\theta dW_s\, \varphi_{h^\beta}(X_h^\theta - x)] = R(\theta, h^{\frac{3}{2}}, x). \tag{175}$$

Using again Holder inequality, the fact that φ is bounded and (163) we obtain

$$|\mathbb{E}[\int_0^h \int_{\mathbb{R}} \gamma'(X_{s^-}^\theta)\dot{X}_s^\theta z\tilde{\mu}(dz,ds)\, \varphi_{h^\beta}(X_h^\theta - x)]| \le (c(1+h^{\frac{p}{2}-1})\int_0^h \mathbb{E}[|\dot{X}_s^\theta|^p]ds)^{\frac{1}{p}}.$$

Using (156), we obtain $\mathbb{E}[\int_0^h \int_{\mathbb{R}} \gamma'(X_{s^-}^\theta)\dot{X}_s^\theta z\tilde{\mu}(dz,ds)\, \varphi_{h^\beta}(X_h^\theta - x)] = R(\theta, h^{1+\frac{1}{p}}, x)$, where p turns out from Holder inequality. We can choose $p = 2$, getting

$$\mathbb{E}[\int_0^h \int_{\mathbb{R}} \gamma'(X_{s^-}^\theta)\dot{X}_s^\theta z\tilde{\mu}(dz,ds)\, \varphi_{h^\beta}(X_h^\theta - x)] = R(\theta, h^{\frac{3}{2}}, x). \tag{176}$$

Using (174), (175) and (176) we have (173), as we wanted.
The first term of (173) can be seen as

$$\mathbb{E}[\int_0^h (\dot{b}(X_s^\theta,\theta) - \dot{b}(x,\theta))ds\, \varphi_{h^\beta}(X_h^\theta - x)] + \mathbb{E}[\int_0^h \dot{b}(x,\theta)ds\, \varphi_{h^\beta}(X_h^\theta - x)].$$

Using Holder inequality and the fact that φ is bounded we get

$$\mathbb{E}[\int_0^h (\dot{b}(X_s^\theta,\theta) - \dot{b}(x,\theta))ds\, \varphi_{h^\beta}(X_h^\theta - x)] \le$$

$$\le c(\mathbb{E}[(\int_0^h (\dot{b}(X_s^\theta,\theta) - \dot{b}(x,\theta))ds)^p])^{\frac{1}{p}} \le c(\mathbb{E}[(\int_0^h \left\|\frac{\partial \dot{b}}{\partial x}\right\|_\infty |X_s^\theta - x|ds)^p])^{\frac{1}{p}}.$$

From Jensen inequality we get it is upper bounded by $c(h^{p-1}\int_0^h \mathbb{E}[|X_s^\theta - x|^p]ds)^{\frac{1}{p}} \le c(h^{p+1}(1+|x|^p))^{\frac{1}{p}}$, where we have used the second point of Lemma 1. It yields

$$\mathbb{E}[\int_0^h (\dot{b}(X_s^\theta,\theta) - \dot{b}(x,\theta))ds\, \varphi_{h^\beta}(X_h^\theta - x)] = R(\theta, h^{1+\frac{1}{p}}, x).$$

Taking $p=2$, the equation (173) becomes

$$\mathbb{E}[\int_0^h \dot{b}(x,\theta)ds\,\varphi_{h^\beta}(X_h^\theta - x)] + R(\theta, h^{\frac{3}{2}}, x) + R(\theta, h^{\frac{3}{2}}, x) = \mathbb{E}[h\dot{b}(x,\theta)\varphi_{h^\beta}(X_h^\theta - x)] + R(\theta, h^{\frac{3}{2}}, x). \quad (177)$$

Replacing in (172), we get

$$\dot{m}_{\theta,h}(x) = \frac{\mathbb{E}[h\dot{b}(x,\theta)\varphi_{h^\beta}(X_h^\theta - x)] + R(\theta, h^{\frac{3}{2}}, x) + R(\theta, h^{2-\alpha\beta-\epsilon-\beta}, x)}{\mathbb{E}[\varphi_{h^\beta}(X_h^\theta - x)]} = h\dot{b}(x,\theta) + \frac{R(\theta, h^{\frac{3}{2}\wedge(2-\alpha\beta-\epsilon-\beta)}, x)}{\mathbb{E}[\varphi_{h^\beta}(X_h^\theta - x)]}.$$

We use the developments (12) and (15) on the denominator; in both of them the function R is negligible compared to 1 without any condition on α and β. Hence for $|x| \le h^{-k_0}$ we get the expression (120).

In order to prove (121), we have to compute the second derivative of $m_{\theta,h}(x)$. From now on we will write only $\varphi^{(k)}$ for $\varphi^{(k)}_{h^\beta}(X_h^\theta - x)$, $k \ge 0$.

$$\ddot{m}_{\theta,h}(x) = \frac{\mathbb{E}[\ddot{X}_h^\theta\varphi] + h^{-\beta}\mathbb{E}[(\dot{X}_h^\theta)^2\varphi']}{\mathbb{E}[\varphi]} - \frac{h^{-\beta}\mathbb{E}[\dot{X}_h^\theta\varphi]\mathbb{E}[\dot{X}_h^\theta\varphi']}{(\mathbb{E}[\varphi])^2} + h^{-\beta}\frac{\mathbb{E}[(\dot{X}_h^\theta)^2\varphi'] + \mathbb{E}[(\dot{X}_h^\theta)^2\varphi'' X_h^\theta h^{-\beta}] + \mathbb{E}[X_h^\theta\varphi'\ddot{X}_h^\theta]}{\mathbb{E}\varphi} +$$

$$+ h^{-2\beta}\mathbb{E}[\dot{X}_h^\theta\varphi']\frac{m_{\theta,h}(x)\mathbb{E}[\dot{X}_h^\theta\varphi'] - \mathbb{E}[X_h^\theta\dot{X}_h^\theta\varphi']}{(\mathbb{E}[\varphi])^2} - h^{-\beta}\frac{\dot{m}_{\theta,h}(x)\mathbb{E}[\dot{X}_h^\theta\varphi'] + h^{-\beta}m_{\theta,h}(x)\mathbb{E}[(\dot{X}_h^\theta)^2\varphi''] + m_{\theta,h}(x)\mathbb{E}[\ddot{X}_h^\theta\varphi']}{\mathbb{E}[\varphi]}.$$

As for the study of (170), we want to rely on Proposition 6 to treat each term of the form $\mathbb{E}[Z\varphi^{(k)}]$, with $k \ge 1$, where Z is bounded in L^p and use $|\mathbb{E}[Z\varphi^{(k)}]| \le \mathbb{E}[|Z||\varphi^{(k)}|] = R(\theta, h^{1-\alpha\beta-\epsilon}, x)$.
We take successively the following variables as choice for Z: $(\frac{\dot{X}_h^\theta}{h})^2$, $\frac{\dot{X}_h^\theta}{h}$, $(\frac{\dot{X}_h^\theta}{h})^2$, $(\frac{\dot{X}_h^\theta}{h})^2 X_h^\theta$, $\frac{\ddot{X}_h^\theta}{h}X_h^\theta$, $\frac{\dot{X}_h^\theta}{h}$, $\frac{\dot{X}_h^\theta}{h}X_h^\theta$, $\frac{\dot{X}_h^\theta}{h}$, $(\frac{\dot{X}_h^\theta}{h})^2$, $\frac{\ddot{X}_h^\theta}{h}$.
All those variable Z are bounded in L^p for $p \ge 2$ by (156) - (157), the third point of Lemma 1 and Holder inequality. We deduce

$$\ddot{m}_{\theta,h}(x) = \frac{\mathbb{E}[\ddot{X}_h^\theta\varphi] + R(\theta, h^{3-\alpha\beta-\epsilon-\beta}, x)}{\mathbb{E}[\varphi]} - \frac{\mathbb{E}[\dot{X}_h^\theta\varphi]R(\theta, h^{2-\alpha\beta-\epsilon-\beta}, x)}{(\mathbb{E}[\varphi])^2} + \quad (178)$$

$$+ \frac{R(\theta, h^{3-\alpha\beta-\epsilon-\beta}, x) + R(\theta, h^{3-\alpha\beta-\epsilon-2\beta}, x) + R(\theta, h^{2-\alpha\beta-\epsilon-\beta}, x)}{\mathbb{E}[\varphi]} + \frac{R(\theta, h^{4-2\alpha\beta-\epsilon-2\beta}, x)}{(\mathbb{E}[\varphi])^2} +$$

$$- \frac{R(\theta, h^{2-\alpha\beta-\epsilon-\beta}, x)\dot{m}_{\theta,h}(x) + R(\theta, h^{3-\alpha\beta-\epsilon-2\beta}, x) + R(\theta, h^{2-\alpha\beta-\epsilon-\beta}, x)}{\mathbb{E}[\varphi]}.$$

We are no longer considering $m_{\theta,h}(x)$ because, by the expression (171), we can include it in the function R.
Using (173) and (177), $\mathbb{E}[\dot{X}_h^\theta\varphi] = h\dot{b}(x,\theta)\mathbb{E}[\varphi] + R(\theta, h^{\frac{3}{2}}, x)$.
Hence $\mathbb{E}[\dot{X}_h^\theta\varphi]R(\theta, h^{2-\alpha\beta-\epsilon-\beta}, x) = R(\theta, h^{3-\alpha\beta-\epsilon-\beta}, x)$, by the definition of rest function R.
We have already proved (120), so

$$R(\theta, h^{2-\alpha\beta-\epsilon-\beta}, x)\dot{m}_{\theta,h}(x) = R(\theta, h^{3-\alpha\beta-\epsilon-\beta}, x).$$

Let us now consider $\mathbb{E}[\ddot{X}_h^\theta\varphi]$. Replacing the dynamic of $\ddot{X}_h^\theta$ by (159), it is

$$\mathbb{E}[\varphi\int_0^h \ddot{b}(X_s^\theta,\theta)ds] + \mathbb{E}[\varphi\int_0^h (b''(X_s^\theta,\theta)(\dot{X}_s^\theta)^2 + 2\dot{b}'(X_s^\theta,\theta)\dot{X}_s^\theta + b'(X_s^\theta,\theta)\ddot{X}_s^\theta)ds] +$$

$$+ \mathbb{E}[\varphi\int_0^h (a''(X_s^\theta)(\dot{X}_s^\theta)^2 + a'(X_s^\theta)\ddot{X}_s^\theta)dW_s] + \mathbb{E}[\varphi\int_0^h\int_{\mathbb{R}} (\gamma''(X_{s^-}^\theta)(\dot{X}_s^\theta)^2 + \gamma'(X_{s^-}^\theta)\ddot{X}_s^\theta)z\tilde{\mu}(dz,ds)] =$$

$$= \mathbb{E}[\varphi\int_0^h \ddot{b}(X_s^\theta,\theta)ds] + R(\theta, h^{\frac{3}{2}}, x). \quad (179)$$

Indeed, using Holder inequality,

$$|\mathbb{E}[\varphi\int_0^h (b''(X_s^\theta,\theta)(\dot{X}_s^\theta)^2 + 2\dot{b}'(X_s^\theta,\theta)\dot{X}_s^\theta + b'(X_s^\theta,\theta)\ddot{X}_s^\theta)ds]| \le$$

$$\le (\mathbb{E}[\varphi^q])^{\frac{1}{q}}(\mathbb{E}[(\int_0^h (b''(X_s^\theta,\theta)(\dot{X}_s^\theta)^2 + 2\dot{b}'(X_s^\theta,\theta)\dot{X}_s^\theta + b'(X_s^\theta,\theta)\ddot{X}_s^\theta)ds)^p])^{\frac{1}{p}} \le$$

$$\leq (c(1+|x|^c)h^{3p} + c(1+|x|^c)h^{2p} + ch^{p-1}\int_0^h \mathbb{E}[|\dot{X}_s|^p]ds)^{\frac{1}{p}},$$

where in the last inequality we have used that φ is bounded and we acted as in (166). By (157), it is upper bounded by $(ch^{3p} + ch^{2p})^{\frac{1}{p}}(1+|x|^c)$. It turns

$$\mathbb{E}[\varphi\int_0^h (b''(X_s^\theta,\theta)(\dot{X}_s^\theta)^2 + 2\dot{b}'(X_s^\theta,\theta)\dot{X}_s^\theta + b'(X_s^\theta,\theta)\ddot{X}_s^\theta)ds] = R(\theta,h^2,x). \tag{180}$$

In the same way, from Holder inequality, (167) and the fact that φ is bounded we get

$$|\mathbb{E}[\varphi\int_0^h (a''(X_s^\theta)(\dot{X}_s^\theta)^2 + a'(X_s^\theta)\ddot{X}_s^\theta)dW_s]| \leq (c(1+|x|^c)h^{2p+\frac{p}{2}} + ch^{\frac{p}{2}-1}\int_0^h \mathbb{E}[|\dot{X}_s|^p]ds)^{\frac{1}{p}}.$$

Using (157), it is upper bounded by $\leq (ch^{2p+\frac{p}{2}} + ch^{p+\frac{p}{2}})^{\frac{1}{p}}(1+|x|^c)$ and so we obtain

$$|\mathbb{E}[\varphi\int_0^h (a''(X_s^\theta)(\dot{X}_s^\theta)^2 + a'(X_s^\theta)\ddot{X}_s^\theta)dW_s]| = R(\theta,h^{\frac{3}{2}},x). \tag{181}$$

Using again Holder inequality, (168), the fact that φ is bounded and (157), we have

$$|\mathbb{E}[\varphi\int_0^h\int_{\mathbb{R}} (\gamma''(X_{s^-}^\theta)(\dot{X}_s^\theta)^2 + \gamma'(X_{s^-}^\theta)\ddot{X}_s^\theta)z\tilde{\mu}(dz,ds)]| \leq c(h^{2p+1} + h^{p+1} + h^{2p+\frac{p}{2}} + h^{p+\frac{p}{2}})^{\frac{1}{p}}(1+|x|^c).$$

Hence, since $p \geq 2$,

$$|\mathbb{E}[\varphi\int_0^h\int_{\mathbb{R}} (\gamma''(X_{s^-}^\theta)(\dot{X}_s^\theta)^2 + \gamma'(X_{s^-}^\theta)\ddot{X}_s^\theta)z\tilde{\mu}(dz,ds)]| = R(\theta,h^{1+\frac{1}{p}},x).$$

Since p turns out from Holder inequality and on which we have only the constraint $p \geq 2$, we can choose $p = 2$, getting

$$|\mathbb{E}[\varphi\int_0^h\int_{\mathbb{R}} (\gamma''(X_{s^-}^\theta)(\dot{X}_s^\theta)^2 + \gamma'(X_{s^-}^\theta)\ddot{X}_s^\theta)z\tilde{\mu}(dz,ds)]| = R(\theta,h^{\frac{3}{2}},x). \tag{182}$$

From (180), (181) and (182) we have (179) as we wanted.
The first term of (179) can be seen as $\mathbb{E}[\varphi\int_0^h(\ddot{b}(X_s^\theta,\theta) - \ddot{b}(x,\theta))ds] + \mathbb{E}[\varphi\int_0^h \ddot{b}(x,\theta)ds]$.
Using Holder inequality and the fact that φ is bounded we get

$$\mathbb{E}[\int_0^h (\ddot{b}(X_s^\theta,\theta) - \ddot{b}(x,\theta))ds\,\varphi_{h^\beta}(X_h^\theta - x)] \leq$$

$$\leq c(\mathbb{E}[(\int_0^h (\ddot{b}(X_s^\theta,\theta) - \ddot{b}(x,\theta))ds)^p])^{\frac{1}{p}} \leq c(\mathbb{E}[(\int_0^h \left\|\frac{\partial\ddot{b}}{\partial x}\right\|_\infty |X_s^\theta - x|ds)^p])^{\frac{1}{p}}.$$

From Jensen inequality we get it is upper bounded by $c(h^{p-1}\int_0^h \mathbb{E}[|X_s^\theta - x|^p]ds)^{\frac{1}{p}} \leq c(h^{p+1}(1+|x|^p))^{\frac{1}{p}}$, where we have used the second point of Lemma 1. It yields

$$\mathbb{E}[\varphi\int_0^h (\ddot{b}(X_s^\theta,\theta) - \ddot{b}(x,\theta))ds] = R(\theta,h^{1+\frac{1}{p}},x).$$

Therefore, considering $p = 2$, (179) becomes $\mathbb{E}[\varphi\ddot{X}_h^\theta] = \mathbb{E}[\varphi\ddot{b}(x,\theta)h] + R(\theta,h^{\frac{3}{2}},x)$.
Replacing in (178) and using the development (12) or (15) of the denominator we obtain, for $|x| \leq h^{-k_0}$,

$$\ddot{m}_{\theta,h}(x) = h\ddot{b}(x,\theta) + R(\theta,h^{\frac{3}{2}},x) + R(\theta,h^{3-\alpha\beta-\epsilon-\beta},x) + R(\theta,h^{3-\alpha\beta-\epsilon-2\beta},x) +$$

$$+R(\theta,h^{2-\alpha\beta-\epsilon-\beta},x) + R(\theta,h^{4-2\alpha\beta-\epsilon-2\beta},x) = h\ddot{b}(x,\theta) + R(\theta,h^{\frac{3}{2}\wedge(2-\alpha\beta-\epsilon-\beta)},x).$$

□

We want now to justify (122).
In the expression of $\dddot{m}_{\theta,h}(y)$, the numerator is the sum of product of terms with the following form: $\mathbb{E}[\varphi^{(k)}X_h^{h_0}\dot{X}_h^{h_1}\ddot{X}_h^{h_2}\dddot{X}_h^{h_3}]h^{-\beta k}$, where $k \geq 1$ and $h_1 + h_2 + h_3 \geq k$.
The only term with a different form is $\mathbb{E}[\varphi\dddot{X}]$, that is $R(\theta,h,y)$ by the boundedness of φ and the equation (169).
We observe that, using Proposition 7 defining $Z = \frac{X_h^{h_0}\dot{X}_h^{h_1}\ddot{X}_h^{h_2}\dddot{X}_h^{h_3}}{h^{h_1+h_2+h_3}}$, we get

$$|\mathbb{E}[\varphi^{(k)}X_h^{h_0}\dot{X}_h^{h_1}\ddot{X}_h^{h_2}\dddot{X}_h^{h_3}]|h^{-\beta k} \leq h^{-\beta k + h_1 + h_2 + h_3 + 1 - \alpha\beta - \epsilon} \leq h^{(1-\beta)k + 1 - \alpha\beta - \epsilon}.$$

We observe that the exponent on h is more then 1 if and only if $\beta < \frac{k}{k+\alpha} - \frac{\epsilon}{k+\alpha}$, with $k \geq 1$. Since $\frac{1}{1+\alpha} - \frac{\epsilon}{1+\alpha}$ is the smallest, the Assumption $\beta < \frac{1}{1+\alpha} - \frac{\epsilon}{1+\alpha}$ that we added in Proposition 8 assures that $|\dddot{m}_{\theta,h}(y)| = R(\theta,h,y)$, as we wanted.

A.2 Proof of limit theorems

In this subsection we prove the theorems stated in Section 6.

A.2.1 Proof of Proposition 3

Proof. (i) follows from Lemma 4.4 in (Gloter, Loukianova, & Mai, 2018), ergodic theorem and the L^1 convergence to zero of $\frac{1}{n}\sum_{i=0}^{n-1}(1+|X_{t_i}|)^c 1_{\{|X_{t_i}|>\Delta_{n,i}^{-k}\}}$, which is a consequence of the third point of Lemma 2. Remark that in (Gloter, Loukianova, & Mai, 2018) the Lemma 4.4 is stated the for $\alpha \in (0,1)$ only. However an inspection of the proof shows that it is valid for $\alpha \in (0,2)$.
Concerning (ii), we can see $\frac{1}{t_n}\sum_{i=0}^{n-1}\Delta_{n,i} f(X_{t_i},\theta)\varphi_{\Delta_{n,i}^{\beta}}(X_{t_{i+1}}-X_{t_i})1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}$ as

$$\frac{1}{t_n}\sum_{i=0}^{n-1}\Delta_{n,i} f(X_{t_i},\theta)1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}} + \frac{1}{t_n}\sum_{i=0}^{n-1}\Delta_{n,i} f(X_{t_i},\theta)(\varphi_{\Delta_{n,i}^{\beta}}(X_{t_{i+1}}-X_{t_i})-1)1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}. \tag{183}$$

We have already showed in (i) that on the first term of (183) we have the convergence wanted and so, in order to get the thesis, it is enough to prove the following:

$$\sup_{\theta\in\Theta}|\frac{1}{t_n}\sum_{i=0}^{n-1}\Delta_{n,i} f(X_{t_i},\theta)(\varphi_{\Delta_{n,i}^{\beta}}(X_{t_{i+1}}-X_{t_i})-1)1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}| \xrightarrow{\mathbb{P}} 0 \tag{184}$$

We observe that

$$|\frac{1}{t_n}\sum_{i=0}^{n-1}\Delta_{n,i} f(X_{t_i},\theta)(\varphi_{\Delta_{n,i}^{\beta}}(X_{t_{i+1}}-X_{t_i})-1)1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}| \le |\frac{1}{n\Delta_n}\sum_{i=0}^{n-1}\Delta_{n,i} f(X_{t_i},\theta)(\varphi_{\Delta_{n,i}^{\beta}}(X_{t_{i+1}}-X_{t_i})-1)|.$$

By the definition of φ, it is different from zero only if $|\Delta X_i| > \Delta_{n,i}^{\beta}$. Using Markov inequality and Lemma 1,

$$\mathbb{P}(|X_{t_{i+1}}-X_{t_i}| > \Delta_{n,i}^{\beta}) \le \mathbb{E}[|X_{t_{i+1}}-X_{t_i}|^2]\Delta_{n,i}^{-2\beta} \le c\Delta_{n,i}^{1-2\beta}. \tag{185}$$

It means that the left hand side of (184) converges to zero in L^1 and so in probability, indeed

$$\mathbb{E}[\sup_{\theta\in\Theta}|\frac{1}{t_n}\sum_{i=0}^{n-1}\Delta_{n,i} f(X_{t_i},\theta)(\varphi_{\Delta_{n,i}^{\beta}}(X_{t_{i+1}}-X_{t_i})-1)1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}|] \le$$

$$\le \mathbb{E}[\sup_{\theta\in\Theta}|\frac{1}{n\Delta_n}\sum_{i=0}^{n-1}\Delta_{n,i} f(X_{t_i},\theta)1_{\{|X_{t_{i+1}}-X_{t_i}|>\Delta_{n,i}^{\beta}\}}|] \le \frac{1}{n\Delta_n}\sum_{i=0}^{n-1}\Delta_{n,i}\mathbb{E}[\sup_{\theta\in\Theta}|f(X_{t_i},\theta)|^2]^{\frac{1}{2}}\mathbb{E}[|1_{\{|X_{t_{i+1}}-X_{t_i}|>\Delta_{n,i}^{\beta}\}}|^2]^{\frac{1}{2}} \le$$

$$\le \frac{c}{n\Delta_n}\sum_{i=0}^{n-1}\Delta_{n,i}\mathbb{P}(|X_{t_{i+1}}-X_{t_i}| > \Delta_{n,i}^{\beta})^{\frac{1}{2}} \le c\Delta_{n,i}^{\frac{1}{2}-\beta},$$

where we have first used Cauchy-Schwarz inequality and then the polynomial growth of $|\sup_{\theta\in\Theta} f|$and the third point of Lemma 2 and (185). Since the exponent on $\Delta_{n,i}$ is positive we get the thesis. □

A.2.2 Proof of Proposition 4 and Lemma 3

Proof of Proposition 4.
In order to show that $\frac{1}{t_n}\sum_{i=0}^{n-1} f_{i,n}(X_{t_i},\theta)\zeta_i\varphi_{\Delta_{n,i}^{\beta}}(X_{t_{i+1}}-X_{t_i})1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}$ converges to zero in probability, we want to use the Lemma 9 of (missing citation) and so we have to show the following:

$$\frac{1}{t_n}\sum_{i=0}^{n-1}\mathbb{E}[f_{i,n}(X_{t_i},\theta)\zeta_i\varphi_{\Delta_{n,i}^{\beta}}(X_{t_{i+1}}-X_{t_i})1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}|\mathcal{F}_{t_i}] \to 0, \tag{186}$$

$$\frac{1}{(t_n)^2}\sum_{i=0}^{n-1}\mathbb{E}[f_{i,n}^2(X_{t_i},\theta)\zeta_i^2\varphi_{\Delta_{n,i}^{\beta}}^2(X_{t_{i+1}}-X_{t_i})1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}|\mathcal{F}_{t_i}] \to 0. \tag{187}$$

If Lemma 3 holds we have that, using (31), the left hand side of (186) results upper bounded by $\Delta_n^{\delta\wedge\frac{1}{2}}\frac{1}{n}\sum_{i=0}^{n-1}|f_{i,n}(X_{t_i},\theta)|R(\theta,1,X_{t_i})$, where we have used the property (11) on R and the fact that $|\Delta_{n,i}| \le \Delta_n$. Since the exponent on Δ_n is positive and $\frac{1}{n}\sum_{i=0}^{n-1} f_{i,n}(X_{t_i},\theta)R(\theta,1,X_{t_i})$ is bounded in L^1 using the polynomial growth of both $f_{i,n}$ and R and the third point of Lemma 2, we get the convergence in probability (186).
Concerning (187), if Lemma 3 holds we can use (32) getting that (187) is

$\frac{1}{n^2\Delta_n}\sum_{i=0}^{n-1} f_{i,n}^2(X_{t_i},\theta)R(\theta,1,X_{t_i})$, where we have used also the property (11) on R and the fact that $|\Delta_{n,i}| \le \Delta_n$. Since $n\Delta_n \to \infty$ and $\frac{1}{n}\sum_{i=0}^{n-1} f_{i,n}^2(X_{t_i},\theta)R(\theta,1,X_{t_i})$ is bounded in L^1 by the polynomial growth of both $f_{i,n}$ and R and the third point of Lemma 2, we get the convergence (187) as we wanted.
Hence, if Lemma 3 holds, then Proposition 4 is proved. □

Proof of Lemma 3.
By the definition (29) of ζ_i and the dynamic of the process X, we get

$$\zeta_i = X_{t_{i+1}} - X_{t_i} - \int_{t_i}^{t_{i+1}} b(\theta_0, X_s)ds + \Delta_{n,i}\int_{\mathbb{R}\setminus\{0\}} z\,\gamma(X_{t_i})\,[1-\varphi_{\Delta_{n,i}^\beta}(\gamma(X_{t_i})z)]\,F(z)dz. \tag{188}$$

We write the left hand side of (31) by using the last equation and adding and subtracting $m_{\theta_0,\Delta_{n,i}}(X_{t_i})$:

$$\mathbb{E}[(X_{t_{i+1}}-m_{\theta_0,\Delta_{n,i}}(X_{t_i}))\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}|\mathcal{F}_{t_i}]+\mathbb{E}[(m_{\theta_0,\Delta_{n,i}}(X_{t_i})-X_{t_i}-\int_{t_i}^{t_{i+1}} b(\theta_0,X_s)ds+ \tag{189}$$
$$+\Delta_{n,i}\int_{\mathbb{R}\setminus\{0\}} z\,\gamma(X_{t_i})\,[1-\varphi_{\Delta_{n,i}^\beta}(\gamma(X_{t_i})z)]\,F(z)dz)\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}|\mathcal{F}_{t_i}].$$

By the $\mathcal{F}_{t_i}$-measurability of X_{t_i}, the first term of (189) is equal to

$$\mathbb{E}[(X_{t_{i+1}}-m_{\theta_0,\Delta_{n,i}}(X_{t_i}))\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})|\mathcal{F}_{t_i}]1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}},$$

that is zero by the definition of $m_{\theta_0,\Delta_{n,i}}$.
On the second term of (189) we use the development (14) or (17), respectively for $\alpha<1$ and $\alpha>1$. Hence, we obtain

$$\mathbb{E}[(\int_{t_i}^{t_{i+1}}(b(\theta_0,X_{t_i})-b(\theta_0,X_s))ds+R(\theta_0,\Delta_{n,i}^{1+\delta},X_{t_i}))\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}|\mathcal{F}_{t_i}], \tag{190}$$

where $\delta>0$ is defined below equation (30). Using the boundedness of both φ and the indicator function and (116) on the first term of (190), we get that (190) is upper bounded by

$$R(\theta_0,\Delta_{n,i}^{\frac{3}{2}},X_{t_i})+R(\theta_0,\Delta_{n,i}^{1+\delta},X_{t_i})=R(\theta_0,\Delta_{n,i}^{(1+\delta)\wedge\frac{3}{2}},X_{t_i}),$$

as we wanted.
Concerning the second point of Lemma 3, we use (29) in order to say that

$$\zeta_i^2\le c(\int_{t_i}^{t_{i+1}}a(X_s)dW_s)^2+c(\int_{t_i}^{t_{i+1}}\int_{\mathbb{R}\setminus\{0\}}z\,\gamma(X_{s^-})\tilde{\mu}(ds,dz))^2+c\Delta_{n,i}^2(\int_{\mathbb{R}\setminus\{0\}}z\,\gamma(X_{t_i})\,[1-\varphi_{\Delta_{n,i}^\beta}(\gamma(X_{t_i})z)]\,F(z)dz)^2. \tag{191}$$

Using this estimation in the left hand side of (32) we obtain three terms, the first is

$$\mathbb{E}[c(\int_{t_i}^{t_{i+1}}a(X_s)dW_s)^2\varphi_{\Delta_{n,i}^\beta}^2(X_{t_{i+1}}-X_{t_i})1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}|\mathcal{F}_{t_i}]\le\mathbb{E}[c(\int_{t_i}^{t_{i+1}}a(X_s)dW_s)^2|\mathcal{F}_{t_i}],$$

by the boundedness of both φ and the indicator function. Using the conditional form of Ito's isometry it is

$$c\mathbb{E}[\int_{t_i}^{t_{i+1}}a^2(X_s)ds|\mathcal{F}_{t_i}]=R(\theta_0,\Delta_{n,i},X_{t_i}), \tag{192}$$

by the polynomial growth of a, the third point of Lemma 1 and the definition of the function R.
We can upper bound the second term of (191) using first of all the boundedness of both φ and the indicator function, and then Kunita's inequality in the conditional form (Appendix of (Jacod & Protter, 2011)). We get the following estimation:

$$\mathbb{E}[c(\int_{t_i}^{t_{i+1}}\int_{\mathbb{R}\setminus\{0\}}z\,\gamma(X_{s^-})\tilde{\mu}(ds,dz))^2\varphi_{\Delta_{n,i}^\beta}^2(X_{t_{i+1}}-X_{t_i})1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}|\mathcal{F}_{t_i}]\le$$
$$\le\mathbb{E}[c(\int_{t_i}^{t_{i+1}}\int_{\mathbb{R}\setminus\{0\}}z\,\gamma(X_{s^-})\tilde{\mu}(ds,dz))^2|\mathcal{F}_{t_i}]\le c\mathbb{E}[\int_{t_i}^{t_{i+1}}\int_{\mathbb{R}\setminus\{0\}}|z|^2\,\gamma^2(X_{s^-})\bar{\mu}(ds,dz)|\mathcal{F}_{t_i}]\le$$
$$\le c\mathbb{E}[\int_{t_i}^{t_{i+1}}\gamma^2(X_{s^-})ds|\mathcal{F}_{t_i}]=R(\theta_0,\Delta_{n,i},X_{t_i}), \tag{193}$$

where in the last inequality and equality we have used, respectively, the definition of the compensator measure $\bar{\mu}$ and the polynomial growth of γ and the third point of Lemma 1.
Concerning the third term of (191), we have already showed in Remark 3 an estimation, depending on α, that is at most $\Delta_{n,i}^{\frac{1}{2}}$. Its square is therefore at least a $R(\theta, \Delta_{n,i}, X_{t_i})$ function, it follows that (32) holds.
We now want to prove (33). Using (30),

$$(X_{t_{i+1}} - m_{\theta_0}(X_{t_i}))^2 \le c\zeta_i^2 + c(\int_{t_i}^{t_{i+1}} b(X_s, \theta_0)ds - \Delta_{n,i}b(X_{t_i}, \theta_0))^2 + R(\theta_0, \Delta_{n,i}^{2+2\delta}, X_{t_i}). \tag{194}$$

We can replace it in (33), getting three terms that are of magnitude at most $\Delta_{n,i}$.
Indeed, on the first we can use (32).
On the second term we can use the boundedness of both φ and the indicator function and Jensen inequality, getting

$$c\mathbb{E}[(\int_{t_i}^{t_{i+1}} b(X_s, \theta_0)ds - \Delta_{n,i}b(X_{t_i}, \theta_0))^2\varphi^2_{\Delta_{n,i}^\beta}(X_{t_{i+1}} - X_{t_i})1_{\left\{|X_{t_i}| \le \Delta_{n,i}^{-k}\right\}}|\mathcal{F}_{t_i}] \le$$

$$c\Delta_{n,i}\mathbb{E}[\int_{t_i}^{t_{i+1}} (b(X_s, \theta_0)ds - b(X_{t_i}, \theta_0))^2|\mathcal{F}_{t_i}] \le$$

$$\le c\Delta_{n,i}\mathbb{E}[\int_{t_i}^{t_{i+1}} b^2(X_s, \theta_0)ds|\mathcal{F}_{t_i}] + c\Delta_{n,i}^2\mathbb{E}[b^2(X_{t_i}, \theta_0)|\mathcal{F}_{t_i}] = R(\theta_0, \Delta_{n,i}^2, X_{t_i}), \tag{195}$$

where in the last equality we have used the polynomial growth of b on both of the two terms and moreover the third point of Lemma 1 on the first term.
In conclusion, we obtain

$$\mathbb{E}[(X_{t_{i+1}} - m_{\theta_0}(X_{t_i}))^2\varphi^2_{\Delta_{n,i}^\beta}(X_{t_{i+1}} - X_{t_i})1_{\left\{|X_{t_i}| \le \Delta_{n,i}^{-k}\right\}}|\mathcal{F}_{t_i}] =$$

$$= R(\theta_0, \Delta_{n,i}, X_{t_i}) + R(\theta_0, \Delta_{n,i}^2, X_{t_i}) + R(\theta_0, \Delta_{n,i}^{2+2\delta}, X_{t_i}) = R(\theta_0, \Delta_{n,i}, X_{t_i}).$$

Hence, we have the thesis. □

A.2.3 Proof of Proposition 5.

In order to prove Proposition 5, the following lemma will be useful:

Lemma 10. *Let us denote by $\tilde{X}^J$ the jump part of X given by*

$$\tilde{X}_t^J := \int_0^t \int_{\mathbb{R}\setminus\{0\}} z\,\gamma(X_{s^-})\tilde{\mu}(ds, dz), \qquad t \ge 0 \tag{196}$$

and $\Delta_i\tilde{X}^J := \tilde{X}^J_{t_{i+1}} - \tilde{X}^J_{t_i}$.
Then, for each $q \ge 2$, $\exists \epsilon > 0$ such that

$$\mathbb{E}[|\Delta_i\tilde{X}^J\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}} - X_{t_i})|^q|\mathcal{F}_{t_i}] = R(\theta_0, \Delta_{n,i}^{1+\beta(q-\alpha)}, X_{t_i}) = R(\theta_0, \Delta_{n,i}^{1+\epsilon}, X_{t_i}). \tag{197}$$

Proof of Lemma 10.
For all $n \in \mathbb{N}$ and $i \in \mathbb{N}$ we define the set on which all the jumps of L on the interval $(t_i, t_{i+1}]$ are small:

$$N_n^i := \left\{|\Delta L_s| \le \frac{4\Delta_{n,i}^\beta}{\gamma_{min}}; \quad \forall s \in (t_i, t_{i+1}]\right\}, \tag{198}$$

where $\Delta L_s := L_s - L_{s^-}$. We hence split the left hand side of (197) as

$$\mathbb{E}[|\Delta_i\tilde{X}^J\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}} - X_{t_i})|^q 1_{N_n^i}|\mathcal{F}_{t_i}] + \mathbb{E}[|\Delta_i\tilde{X}^J\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}} - X_{t_i})|^q 1_{(N_n^i)^c}|\mathcal{F}_{t_i}]. \tag{199}$$

We now observe that, by the definition of N_n^i,

$$|\mathbb{E}[|\Delta_i\tilde{X}^J\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}} - X_{t_i})|^q 1_{N_n^i}|\mathcal{F}_{t_i}]| \le$$

$$\leq c\mathbb{E}[|\int_{t_i}^{t_{i+1}}\int_{|z|\leq\frac{4\Delta_{n,i}^\beta}{\gamma_{min}}} z\,\gamma(X_{s^-})\tilde{\mu}(ds,dz)|^q + |\int_{t_i}^{t_{i+1}}\int_{|z|\geq\frac{4\Delta_{n,i}^\beta}{\gamma_{min}}} |z|\,|\gamma(X_{s^-})|\bar{\mu}(ds,dz)|^q|\mathcal{F}_{t_i}]. \tag{200}$$

We observe that the order of the second term depend on α. Acting as in Remark 3, we get that its order is $\Delta_{n,i}^q$ if $\alpha\in(0,1)$ while it is $\Delta_{n,i}^{q+q\beta(1-\alpha)}$ if $\alpha\in(1,2)$. Since q is more than $q+q\beta(1-\alpha)$ if and only if $\alpha>1$, we can say that the second term of (200) is upper bounded by $c\Delta_{n,i}^{q\wedge(q+q\beta(1-\alpha))}$. The first term of (200) is instead upper bounded by

$$c\mathbb{E}[|\int_{t_i}^{t_{i+1}}\int_{|z|\leq\frac{4\Delta_{n,i}^\beta}{\gamma_{min}}} |z|^2\,|\gamma(X_{s^-})|^2\bar{\mu}(ds,dz)|^{\frac{q}{2}}|\mathcal{F}_{t_i}] + c\mathbb{E}[\int_{t_i}^{t_{i+1}}\int_{|z|\leq\frac{4\Delta_{n,i}^\beta}{\gamma_{min}}} |z|^q\,|\gamma(X_{s^-})|^q\bar{\mu}(ds,dz)|\mathcal{F}_{t_i}]\leq$$

$$\leq c\,\|\gamma\|_\infty^q\,(\mathbb{E}[|\int_{t_i}^{t_{i+1}}\int_{|z|\leq\frac{4\Delta_{n,i}^\beta}{\gamma_{min}}} |z|^{1-\alpha}dzds|^{\frac{q}{2}}|\mathcal{F}_{t_i}] + \mathbb{E}[\int_{t_i}^{t_{i+1}}\int_{|z|\leq\frac{4\Delta_{n,i}^\beta}{\gamma_{min}}} |z|^{q-1-\alpha}dzds|\mathcal{F}_{t_i}])\leq$$

$$\leq c(\int_{t_i}^{t_{i+1}}\Delta_{n,i}^{(2-\alpha)\beta}ds)^{\frac{q}{2}}+c\int_{t_i}^{t_{i+1}}\Delta_{n,i}^{(q-\alpha)\beta}ds+\Delta_{n,i}^q\leq c(\Delta_{n,i}^{(1+(2-\alpha)\beta)\frac{q}{2}}+\Delta_{n,i}^{(q-\alpha)\beta+1}) = R(\theta_0,\Delta_{n,i}^{(q-\alpha)\beta+1},X_{t_i}), \tag{201}$$

where we have used Kunita inequality, the definition of $\bar{\mu}$ and the second point of Assumption 4. Using the consideration below equation (200) and (201) we get

$$|\mathbb{E}[|\Delta_i\tilde{X}^J\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})|^q 1_{N_n^i}|\mathcal{F}_{t_i}]|\leq R(\theta_0,\Delta_{n,i}^{(q-\alpha)\beta+1},X_{t_i})+R(\theta_0,\Delta_{n,i}^{q\wedge(q+q\beta(1-\alpha))},X_{t_i}) = R(\theta_0,\Delta_{n,i}^{(q-\alpha)\beta+1},X_{t_i}). \tag{202}$$

For $\alpha\in(0,2)$, $\alpha\neq1$ and $\beta\in(0,\frac{1}{2})$ the exponent on $\Delta_{n,i}$ can be seen as $1+\epsilon$, with $\epsilon>0$.
Concerning the second term of (199), we have

$$\mathbb{E}[|\Delta_i\tilde{X}^J\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})|^q 1_{(N_n^i)^c}|\mathcal{F}_{t_i}]\leq c\mathbb{E}[(|\Delta_iX|^q+|\Delta X_i^c|^q)|\varphi_{\Delta_{n,i}^\beta}^q(X_{t_{i+1}}-X_{t_i})|1_{(N_n^i)^c}|\mathcal{F}_{t_i}], \tag{203}$$

where $|\Delta_iX| := |X_{t_{i+1}}-X_{t_i}|$ and ΔX_i^c is the increment of the continuous part of X in the interval $(t_i,t_{i+1}]$. We observe that, by the definition of $\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})$, the first term in the right hand side is different from zero only if $|\Delta_iX|^q\leq\Delta_{n,i}^{\beta q}$. Therefore

$$\mathbb{E}[|\Delta_iX|^q\varphi_{\Delta_{n,i}^\beta}^q(X_{t_{i+1}}-X_{t_i})1_{(N_n^i)^c}|\mathcal{F}_{t_i}]\leq\Delta_{n,i}^{\beta q}\mathbb{P}_i((N_n^i)^c)\leq c\Delta_{n,i}^{\beta q+1-\alpha\beta}. \tag{204}$$

Indeed

$$\mathbb{P}_i((N_n^i)^c) = \mathbb{P}_i(\exists s\in(t_i,t_{i+1}]\,:\,|\Delta L_s|>\frac{4\Delta_{n,i}^\beta}{\gamma_{min}})\leq c\int_{t_i}^{t_{i+1}}\int_{\frac{4\Delta_{n,i}^\beta}{\gamma_{min}}}^\infty F(z)dzds\leq c\Delta_{n,i}^{1-\alpha\beta}, \tag{205}$$

where we have used the third point of Assumption 4. Since $q\geq2$, $\beta q+1-\alpha\beta$ is always more than 1.
In the same way

$$\mathbb{E}[|\Delta X_i^c|^q\varphi_{\Delta_{n,i}^\beta}^q(X_{t_{i+1}}-X_{t_i})1_{(N_n^i)^c}|\mathcal{F}_{t_i}]\leq c\Delta_{n,i}^{\frac{1}{2}q}\Delta_{n,i}^{1-\alpha\beta}(1+|X_{t_i}|^c), \tag{206}$$

that is again more than 1. Using (199), (202), (204) and (206) we get the thesis. □

We can now prove Proposition 5.

Proof of Proposition 5.
We denote

$$s_i^n := \frac{1}{\sqrt{t_n}}(X_{t_{i+1}}-m_{\theta_0}(X_{t_i}))f(X_{t_i},\theta)\varphi_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})1_{\{|X_{t_i}|\leq\Delta_{n,i}^{-k}\}}. \tag{207}$$

In order to show the asymptotic normality we have to prove that s^n is a martingale difference array such that

$$\sum_{i=0}^{n-1}\mathbb{E}[|s_i^n|^{2+r}|\mathcal{F}_{t_i}]\xrightarrow{\mathbb{P}}0, \tag{208}$$

for a constant $\delta>0$, and

$$\sum_{i=0}^{n-1}\mathbb{E}[|s_i^n|^2|\mathcal{F}_{t_i}]\xrightarrow{\mathbb{P}}\int_{\mathbb{R}}f^2(x,\theta)a^2(x)\pi(dx), \tag{209}$$

c.f. Theorem A2 in the Appendix of (Shimizu & Yoshida, 2006).
We observe that s_i^n is a martingale difference array since, $\forall i \ge 0$,

$$\mathbb{E}[s_i^n|\mathcal{F}_{t_i}] = \frac{f(X_{t_i}, \theta)1_{\{|X_{t_i}| \le \Delta_{n,i}^{-k}\}}}{\sqrt{t_n}} \mathbb{E}[(X_{t_{i+1}} - m_{\theta_0}(X_{t_i}))\varphi_{\Delta_{n,i}^{\beta}}(X_{t_{i+1}} - X_{t_i})|\mathcal{F}_{t_i}] = 0,$$

by the measurability of f and the indicator function and the definition of $m_{\theta_0}(X_{t_i})$.
We now want to prove (209). Using (30) and the definition of ζ_i we have that

$$(X_{t_{i+1}} - m_{\theta_0}(X_{t_i}))^2 = (\int_{t_i}^{t_{i+1}} a(X_s)dW_s)^2 + 2B_{i,n}\int_{t_i}^{t_{i+1}} a(X_s)dW_s + B_{i,n}^2, \tag{210}$$

where

$$B_{i,n} := \int_{t_i}^{t_{i+1}} (b(X_s, \theta_0) - b(X_{t_i}, \theta_0))ds + R(\theta_0, \Delta_{n,i}^{1+\delta}, X_{t_i}) +$$

$$+ \int_{t_i}^{t_{i+1}} \int_{\mathbb{R}\setminus\{0\}} z\,\gamma(X_{s^-})\tilde{\mu}(ds, dz) + \Delta_{n,i} \int_{\mathbb{R}\setminus\{0\}} z\,\gamma(X_{s^-})[1 - \varphi_{\Delta_{n,i}^{\beta}}(\gamma(X_{t_i})z)]F(z)dz.$$

Replacing (210) in the definition (207) of s_i^n we get three terms. We start proving that

$$\frac{1}{t_n}\sum_{i=0}^{n-1} \mathbb{E}[B_{i,n}^2 f^2(X_{t_i}, \theta)\varphi_{\Delta_{n,i}^{\beta}}^2(X_{t_{i+1}} - X_{t_i})1_{\{|X_{t_i}| \le \Delta_{n,i}^{-k}\}}|\mathcal{F}_{t_i}] \xrightarrow{\mathbb{P}} 0. \tag{211}$$

Indeed,

$$\mathbb{E}[B_{i,n}^2\varphi_{\Delta_{n,i}^{\beta}}^2(X_{t_{i+1}} - X_{t_i})|\mathcal{F}_{t_i}] \le c\mathbb{E}[[(\int_{t_i}^{t_{i+1}} (b(X_s, \theta_0) - b(X_{t_i}, \theta_0))ds)^2 +$$

$$+ R(\theta_0, \Delta_{n,i}^{2+2\delta}, X_{t_i}) + (\int_{t_i}^{t_{i+1}} \int_{\mathbb{R}\setminus\{0\}} z\,\gamma(X_{s^-})\tilde{\mu}(ds, dz))^2 + \tag{212}$$

$$+(\Delta_{n,i}\int_{\mathbb{R}\setminus\{0\}} z\,\gamma(X_{s^-})[1 - \varphi_{\Delta_{n,i}^{\beta}}(\gamma(X_{t_i})z)]F(z)dz)^2]\varphi_{\Delta_{n,i}^{\beta}}^2(X_{t_{i+1}} - X_{t_i})|\mathcal{F}_{t_i}] \le$$

$$\le R(\theta_0, \Delta_{n,i}^2, X_{t_i}) + R(\theta_0, \Delta_{n,i}^{2+2\delta}, X_{t_i}) + R(\theta_0, \Delta_{n,i}^{(1+\beta(q-\alpha))\wedge(1+\epsilon)}, X_{t_i}), \tag{213}$$

where we have used (195) on the first term of (212), (197) of the previous lemma on the third and Remark 3 on the fourth. Indeed, in Remark 3, we found that the last term in less than $R(\theta_0, \Delta_{n,i}^2, X_{t_i})$ if $\alpha \le 1$ and less than $R(\theta_0, \Delta_{n,i}^{2+2\beta(1-\alpha)}, X_{t_i})$ if $\alpha > 1$; in both cases the exponent on $\Delta_{n,i}$ is always more than 1, hence we can write it as $1 + \epsilon$.
We can upper bound with (213) the left hand side of (211) getting $\frac{1}{t_n}\sum_{i=0}^{n-1} f^2(X_{t_i}, \theta)R(\theta_0, \Delta_{n,i}^{1+\epsilon}, X_{t_i})1_{\{|X_{t_i}| \le \Delta_{n,i}^{-k}\}}$, that converges to 0 in norm L^1 by the polynomial growth of both f and R and the third point of Lemma 2 and using that $|\Delta_{n,i}| \le \Delta_n$. We obtain therefore the convergence in probability (211) wanted.
Let us now consider the contribution of the first term of (210) for the proof of (209). We can see it as

$$\frac{1}{t_n}\sum_{i=0}^{n-1} f^2(X_{t_i}, \theta)\mathbb{E}[(\int_{t_i}^{t_{i+1}} a(X_s)dW_s)^2|\mathcal{F}_{t_i}]1_{\{|X_{t_i}| \le \Delta_{n,i}^{-k}\}} + \tag{214}$$

$$+\frac{1}{t_n}\sum_{i=0}^{n-1} f^2(X_{t_i}, \theta)\mathbb{E}[(\int_{t_i}^{t_{i+1}} a(X_s)dW_s)^2(\varphi_{\Delta_{n,i}^{\beta}}^2(X_{t_{i+1}} - X_{t_i}) - 1)|\mathcal{F}_{t_i}]1_{\{|X_{t_i}| \le \Delta_{n,i}^{-k}\}}.$$

On the first term of (214) we use Ito's isometry, getting

$$\frac{1}{t_n}\sum_{i=0}^{n-1} f^2(X_{t_i}, \theta)\mathbb{E}[\int_{t_i}^{t_{i+1}} a(X_s)^2 ds|\mathcal{F}_{t_i}]1_{\{|X_{t_i}| \le \Delta_{n,i}^{-k}\}} =$$

$$= \frac{1}{t_n}\sum_{i=0}^{n-1} f^2(X_{t_i}, \theta)(\Delta_{n,i}a^2(X_{t_i}) + R(\theta_0, \Delta_{n,i}^{\frac{3}{2}}, X_{t_i}))1_{\{|X_{t_i}| \le \Delta_{n,i}^{-k}\}}, \tag{215}$$

where we have used (116) with a^2 in place of b. Using the first point of Proposition 1 we get that

$$\frac{1}{t_n}\sum_{i=0}^{n-1} f^2(X_{t_i}, \theta)\Delta_{n,i}a^2(X_{t_i})1_{\{|X_{t_i}| \le \Delta_{n,i}^{-k}\}} \xrightarrow{\mathbb{P}} \int_{\mathbb{R}} f^2(x, \theta)a^2(x)\pi(dx), \tag{216}$$

while $\frac{1}{t_n}\sum_{i=0}^{n-1} f^2(X_{t_i},\theta)R(\theta_0,\Delta_{n,i}^{\frac{3}{2}},X_{t_i})1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}$ goes to zero in norm L^1 and therefore in probability. Let us now consider the second term of (214). Using Cauchy- Schwarz inequality we get it is upper bounded by

$$\frac{1}{n\Delta_n}\sum_{i=0}^{n-1} f^2(X_{t_i},\theta)\mathbb{E}[|\int_{t_i}^{t_{i+1}} a(X_s)dW_s|^4|\mathcal{F}_{t_i}]^{\frac{1}{2}}\mathbb{E}[|\varphi^2_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})-1|^2|\mathcal{F}_{t_i}]^{\frac{1}{2}}\le$$

$$\le\frac{1}{n\Delta_n}\sum_{i=0}^{n-1} f^2(X_{t_i},\theta)\mathbb{E}[|\int_{t_i}^{t_{i+1}} a^2(X_s)ds|^2|\mathcal{F}_{t_i}]^{\frac{1}{2}}\mathbb{E}[1_{\{|X_{t_{i+1}}-X_{t_i}|>\Delta_{n,i}^\beta\}}|\mathcal{F}_{t_i}]^{\frac{1}{2}},$$

where we have used Burkholder Davis Gundy inequality and the fact that, by the definition of φ, it is different from 0 only if $|X_{t_{i+1}}-X_{t_i}|>\Delta_{n,i}^\beta$. Using Jensen inequality and (185) in the conditional form we can upper bound it with

$$\frac{1}{n\Delta_n}\sum_{i=0}^{n-1} f^2(X_{t_i},\theta)\mathbb{E}[\Delta_{n,i}^2|\frac{1}{\Delta_{n,i}}\int_{t_i}^{t_{i+1}} a^2(X_s)ds|^2|\mathcal{F}_{t_i}]^{\frac{1}{2}}\mathbb{P}(|X_{t_{i+1}}-X_{t_i}|>\Delta_{n,i}^\beta|\mathcal{F}_{t_i})^{\frac{1}{2}}\le$$

$$\frac{c}{n\Delta_n}\sum_{i=0}^{n-1} f^2(X_{t_i},\theta)\mathbb{E}[\Delta_{n,i}\int_{t_i}^{t_{i+1}} a^4(X_s)ds|\mathcal{F}_{t_i}]^{\frac{1}{2}}R(\theta_0,\Delta_{n,i}^{\frac{1}{2}-\beta},X_{t_i})\le$$

$$\le\frac{c}{n\Delta_n}\sum_{i=0}^{n-1} f^2(X_{t_i},\theta)\Delta_{n,i}^{\frac{1}{2}}[\Delta_{n,i}a^4(X_{t_i})+R(\theta_0,\Delta_{n,i}^{\frac{3}{2}},X_{t_i})]^{\frac{1}{2}}R(\theta_0,\Delta_{n,i}^{\frac{1}{2}-\beta},X_{t_i}),\tag{217}$$

where we have also used (116) with a^4 in place of b. We observe that (217) goes to 0 in L^1 and therefore in probability, indeed its L^1 norm is upper bounded by

$$\le\Delta_n^{\frac{1}{2}-\beta}\frac{c}{n}\sum_{i=0}^{n-1}\mathbb{E}[f^2(X_{t_i},\theta)R(\theta_0,1,X_{t_i})(a^2(X_{t_i})+R(\theta_0,\Delta_{n,i}^{\frac{3}{4}},X_{t_i}))],$$

that goes to 0 by the polynomial growth of f, R and a and the third point of Lemma 1 and since $\beta<\frac{1}{2}$. Let us now consider the second term of (210) for the proof of (209). Using Cauchy-Schwarz inequality, (213) and Ito's isometry we get

$$\frac{2}{t_n}\sum_{i=0}^{n-1} f^2(X_{t_i},\theta)\mathbb{E}[B_{i,n}\int_{t_i}^{t_{i+1}} a(X_s)dW_s\,\varphi^2_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})1_{\{|X_{t_i}|\le\Delta_{n,i}^{-k}\}}|\mathcal{F}_{t_i}]\le$$

$$\le\frac{c}{n\Delta_n}\sum_{i=0}^{n-1} f^2(X_{t_i},\theta)R(\theta_0,\Delta_{n,i}^{1+\epsilon},X_{t_i})^{\frac{1}{2}}\mathbb{E}[\int_{t_i}^{t_{i+1}} a(X_s)^2ds|\mathcal{F}_{t_i}]^{\frac{1}{2}}\le$$

$$\le\Delta_n^{\frac{\epsilon}{2}}\frac{c}{n}\sum_{i=0}^{n-1} f^2(X_{t_i},\theta)R(\theta_0,1,X_{t_i})(a^2(X_{t_i})+R(\theta_0,\Delta_{n,i}^{\frac{1}{4}},X_{t_i})),\tag{218}$$

where in the last inequality we have used the property (11) of R and (215) with the trivial estimation $|\Delta_{n,i}|\le\Delta_n$. By the polynomial growth of both a, f and R and the fact that the exponent on Δ_n is positive we have that (218) converges to 0 en norm L^1. Hence it converges to 0 in probability, (209) follows.

Our goal is now to prove (208). Using (210) we have that

$$\sum_{i=0}^{n-1}\mathbb{E}[|s_i^n|^{2+r}|\mathcal{F}_{t_i}]\le$$

$$\le c\frac{1}{(n\Delta_n)^{1+\frac{r}{2}}}\sum_{i=0}^{n-1} f^{2+r}(X_{t_i},\theta)(\mathbb{E}[B_{i,n}^{2+r}\varphi^{2+r}_{\Delta_{n,i}^\beta}(X_{t_{i+1}}-X_{t_i})|\mathcal{F}_{t_i}]+\mathbb{E}[(\int_{t_i}^{t_{i+1}} a(X_s)dW_s)^{2+r}|\mathcal{F}_{t_i}]).\tag{219}$$

We act as we have already done in the proof of (209) on the first term of (219): using (197) we get it is upper bounded by

$$\frac{c}{(n\Delta_n)^{1+\frac{r}{2}}}\sum_{i=0}^{n-1} f^{2+r}(X_{t_i},\theta)R(\theta_0,\Delta_{n,i}^{1+\epsilon},X_{t_i})\le\Delta_n^\epsilon\frac{c}{(n\Delta_n)^{\frac{r}{2}}}\frac{1}{n}\sum_{i=0}^{n-1} f^{2+r}(X_{t_i},\theta)R(\theta_0,1,X_{t_i}),$$

that converges to 0 in norm L^1 (and therefore in probability) since $\epsilon > 0$ and $n\Delta_n \to \infty$ for $n \to \infty$. Concerning the second term of (219), using Burkholder-Davis-Gundy inequality and (215) we have

$$\mathbb{E}[(\int_{t_i}^{t_{i+1}} a(X_s)dW_s)^{2+r}|\mathcal{F}_{t_i}]) \le R(\theta_0, \Delta_{n,i}^{1+\frac{r}{2}}, X_{t_i}). \tag{220}$$

Using (220) we get that the second term of (219) is upper bounded by $\frac{c}{n^{\frac{r}{2}}}\frac{1}{n}\sum_{i=0}^{n-1} f^{2+r}(X_{t_i},\theta)R(\theta_0,1,X_{t_i})$, that converges to 0 in norm L^1 and hence in probability since $n^{\frac{r}{2}} \to \infty$. We deduce (208) and therefore the wanted asymptotic normality. □

A.3 Proof of Propositions 6 and 7.

Since Proposition 6 is a consequence of Proposition 7, let us start with the proof of Proposition 7. To lighten the notation we forget the dependence on θ of X^θ and Z_θ.
Proof of Proposition 7.
Using $\tilde{X}_t^J$ defined in (196), we introduce the event

$$E_h := \left\{ \tilde{X}_h^J := \int_0^t \int_{\mathbb{R}\setminus\{0\}} z\,\gamma(X_{s^-})\tilde{\mu}(ds,dz) \in [\frac{1}{2}h^\beta, 4h^\beta] \right\}. \tag{221}$$

We have that

$$\mathbb{E}[|Z\varphi_{h^\beta}^{(k)}(X_h - x)|] = \mathbb{E}[|Z\varphi_{h^\beta}^{(k)}(X_h - x)|1_{E_h}] + \mathbb{E}[|Z\varphi_{h^\beta}^{(k)}(X_h - x)|1_{E_h^c}]. \tag{222}$$

We observe that, by its definition, $\varphi_{h^\beta}^{(k)}(X_h - x)$ is different from 0 only if $|X_h - x| \in [h^\beta, 2h^\beta]$. But $\Delta_h X := |X_h - x| = |X_h^c - x + \tilde{X}_h^J|$ hence on E_h^c, where $\tilde{X}_h^J \notin [\frac{1}{2}h^\beta, 4h^\beta]$, from $|X_h - x| \in [h^\beta, 2h^\beta]$ we deduce that it must be $|X_h^c - x| \ge \frac{1}{2}h^\beta$. Using this observation and Holder inequality we have that the second term on the right hand side of (222) is upper bounded by

$$(\mathbb{E}[|Z|^p])^{\frac{1}{p}}(\mathbb{E}[|\varphi_{h^\beta}^{(k)}(X_h - x)|^q 1_{E_h^c}])^{\frac{1}{q}} \le c(\mathbb{P}(|X_h^c - x| \ge \frac{1}{2}h^\beta))^{\frac{1}{q}} \le ch^{\frac{r}{q}(\frac{1}{2}-\beta)}$$

$\forall r > 1$, where we have also used that Z is bounded in L^p and Remark 2 in (Gloter, Loukianova, & Mai, 2018).
In order to estimate the first term on the right hand side of (222) we need the following lemma that we will prove at the end of the section:
Lemma 11. *Let us consider E_h, the set defined in (221). We have*

$$\mathbb{P}(E_h) \le R(\theta, h^{1-\beta\alpha}, x). \tag{223}$$

If $Z \in \mathcal{Z}_{h,c,p}$, then using Holder inequality, the estimation (223) and the boundedness of Z in L^p we get

$$\mathbb{E}[|Z\varphi_{h^\beta}^{(k)}(X_h - x)|1_{E_h}] \le (\mathbb{E}[|Z|^p]^{\frac{1}{p}})(\mathbb{E}[|\varphi_{h^\beta}^{(k)}(X_h - x)|^q 1_{E_h}])^{\frac{1}{q}} \le$$

$$\le cR(\theta, h^{1-\beta\alpha}, x)^{\frac{1}{q}} = cR(\theta, h^{\frac{1-\beta\alpha}{q}}, x),$$

with $\frac{1}{p} + \frac{1}{q} = 1$. Hence, we get the Proposition 7. □

Proposition 6 is a consequence of Proposition 7, observing that $(h(X_u,\theta))_{\theta\in\Theta} \in \mathcal{Z}_{t_{i+1}-t_i,c,p}$, for $u \in [t_i, t_{i+1}]$, and the Markov property.

In conclusion, we prove Lemma 11.

Proof of Lemma 11.
We use again the set N_n^i defined in (198). We have

$$\mathbb{P}(E_h) = \mathbb{P}(E_h \cap N_n^i) + \mathbb{P}(E_h \cap (N_n^i)^c). \tag{224}$$

On the second term of (224) we use (205), getting

$$\mathbb{P}(E_h \cap (N_n^i)^c) \le \mathbb{P}((N_n^i)^c) \le ch^{1-\alpha\beta}. \tag{225}$$

Concerning the set $E_h \cap N_n^i$, we use Markov inequality and we obtain, $\forall r > 1$,

$$\mathbb{P}(E_h \cap N_n^i) \le c\mathbb{E}[|\tilde{X}_h^J|^r 1_{N_n^i}]h^{-\beta r} \le ch^{-\beta r}h^{1+\beta(r-\alpha)} = ch^{1-\beta\alpha}, \tag{226}$$

where in the last inequality we used (202).
Using (224), (225) and (226) we get the Lemma 11. □