LAMN property for Euler type scheme of stochastic differential equations driven by a stable Lévy process and applications

Alexandre Brouste, Laurent Denis, Trâm Thi-Bao Ngô

Efficient estimation for stochastic differential equations driven by a stable Lévy process

ALEXANDRE BROUSTE[1,a], LAURENT DENIS[1,b] and THI-BAO-TRÂM NGÔ[1,c]

[1] *Le Mans Université, Laboratoire Manceau de Mathématiques, Avenue Olivier Messiaen, 72085 Le Mans Cedex 09, France* [a]*Alexandre.Brouste@univ-lemans.fr,* [b]*Laurent.Denis@univ-lemans.fr,* [c]*Thi_Bao_Tram.Ngo@univ-lemans.fr*

The joint parametric estimation of the drift coefficient, the scale coefficient and the jump activity in stochastic differential equations driven by a symmetric stable Lévy process is considered, based on high-frequency observations. Firstly, the LAMN property for the corresponding Euler-type scheme is proved and lower bounds for the estimation risk in this setting are deduced. When the approximation scheme experiment is asymptotically equivalent to the original one, these bounds can be transferred. Secondly, a one-step procedure is proposed which is shown to be fast and asymptotically efficient. The performances in terms of asymptotical variance and computation time on samples of finite size are illustrated with simulations.

Keywords: Lévy process, stable process; stochastic differential equation; LAMN property; parametric estimation, one-step procedure

1. Introduction

Local Asymptotic Mixed Normality (LAMN) provides a powerful framework under which the asymptotical optimality of estimators can be studied. More precisely, for a statistical experiment satisfying the LAMN property, minimax theorems can be applied and a lower bound for the variance of the estimators can be derived (see e.g. Jeganathan (1982)).

The LAMN property of the likelihoods has been of great interest by many authors. It was established for the estimation of the parameters of the drift and the diffusion coefficient for $\mathbb{R}^d$-valued solutions of diffusion processes observed at high-frequency (infill asymptotics) in Gobet (2001). This result confirmed that the estimators elicited in Genon-Catelot and Jacod (1993) were asymptotically efficient. Recently, the LAMN property has been extended to the solution of stochastic differential equations driven by stable Lévy processes in the high-frequency setting for the parameter of a general drift but with constant scale coefficient in Clément and Gloter (2015) and for the parameters in drift and scale in Clément, Gloter and Nguyen (2019). In these works, the stability index was supposed to be known. For an unknown stability index, the LAN property with a non-singular Fisher information matrix was set in Brouste and Masuda (2018) using a nondiagonal rate matrix for a stable Lévy process with drift. Later on, quasi-likelihood estimation procedures were proposed to estimate jointly drift parameter, scale parameter and stability index in Clément and Gloter (2020) but the asymptotic efficiency remains an open question in the general setting. The present paper gives an answer for the asymptotic efficiency, in some particular cases, in the joint parametric estimation (drift, scale and stability index) for stochastic differential equations driven by stable Lévy processes in the high-frequency setting.

Precisely, in this work, the LAMN property of the likelihoods is established for the statistical experiment of observing a generalized Euler scheme of the stochastic differential equations driven by stable Lévy processes. This asymptotic property is obtained for both multiplicative and non-multiplicative scale coefficients. For some specific scale coefficients, it can be shown theoretically with the results in Clément (2023) that this statistical experiment used here is asymptotically equivalent (see Shiryaev

and Spokoiny (2000) and the reference therein) to the high-frequency observation of the solution of the considered stochastic differential equation. In this case, an alternative Le Cam's one-step procedure is proposed (see Le Cam (1956)) since the maximum likelihood estimator (MLE), which is asymptotically efficient, can be time-consuming for large samples. It is based on an initial guess estimator which is a combination of generalized variations of the trajectory for the scale and index parameters and a maximum likelihood type estimator for the drift parameter. Since the Fisher information matrix can be expressed explicitly, the initial guess estimation is corrected by a single step of the Fisher scoring method on the log-likelihood function. This new estimator is fast to be computed and is shown to be asymptotically efficient.

The rest of this paper is organized as follows. Section 2 is dedicated to the notations and the assumptions made. The main results on the LAMN property of the likelihoods for the aforementioned statistical experiment and the asymptotical efficiency of the one-step procedure are shown in Section 3. Numerical simulations in Section 4 illustrate the performance of the procedure on samples of finite size in terms of asymptotic variance and computation time. The proofs for LAMN properties are postponed in Section 5 and the accompanying paper.

2. Assumptions and settings

We consider the stochastic differential equation driven by a stable Lévy process as follows

$$X_t = x_0 + \int_0^t b(X_s, \mu)ds + \int_0^t a(X_{s-}, \sigma)dJ_s, \quad t \in [0,T], \tag{2.1}$$

where (J_t) denotes the standard symmetric β-stable Lévy process whose characteristic function is

$$\mathbb{E}_\theta(e^{iuJ_1}) = e^{-|u|^\beta}, \quad u \in \mathbb{R}.$$

The distribution of (X_t) associated with parameter $\theta = (\mu, \sigma, \beta) \in \mathbb{R} \times (0,\infty) \times (0,2)$ is denoted $\mathbb{P}_\theta$ and the expectation under $\mathbb{P}_\theta$ is denoted $\mathbb{E}_\theta$.

Without loss of generality, we assume that $T = 1$. We observe the process $(X_{t_i^n})_{0\le i\le n}$ on the time grid $t_i^n = i/n$ for $i \in \{0, 1, \dots, n\}$ that solves (2.1) for the parameter value $\theta_0 = (\mu_0, \sigma_0, \beta_0) \in \Theta$ where Θ is an open subset of $\mathbb{R} \times (0,\infty) \times (0,2)$. The unknown parameter θ_0 is to be estimated.

In addition, we make the following assumption on the coefficient functions. Note that from now on, we denote by C and p some generic positive constants whose value may change from line to line.

(A) We assume that $x \mapsto a(x, \sigma_0)$ is C^2 on $\mathbb{R}$ and that there exists a neighborhood $V_{\mu_0} \times V_{\sigma_0}$ of (μ_0, σ_0), such that b is C^3 on $\mathbb{R} \times V_{\mu_0}$ and

$$\sup_x \Big(\sup_{\mu \in V_{\mu_0}} |\partial_x b(x,\mu)| + |\partial_x a(x,\sigma_0)| \Big) \le C,$$

$$\max_{2\le \ell \le 3} \sup_{\mu \in V_{\mu_0}} |\partial_x^\ell b(x,\mu)| + |\partial_x^2 a(x,\sigma_0)| \le C(1+|x|^p),$$

$$\forall x \in \mathbb{R}, \forall \sigma \in V_{\sigma_0}, \quad a(x,\sigma) > 0 \quad \text{and} \quad \sup_{\sigma \in V_{\sigma_0}} \frac{1}{a(x,\sigma)} \le C(1+|x|^p).$$

We also assume that for any $x \in \mathbb{R}$, $\mu \mapsto b(x,\mu)$ and $\sigma \mapsto a(x,\sigma)$ are C^4 and

$$\sup_{(\mu,\sigma) \in V_{\mu_0} \times V_{\sigma_0}} \max_{1\le \ell \le 4} (|\partial_\mu^\ell b(x,\mu)| + |\partial_\sigma^\ell a(x,\sigma)|) \le C(1+|x|^p),$$

$$\sup_{\mu\in V_{\mu_0}} |\partial_x^k \partial_\mu^\ell b(x,\mu)| \le C(1+|x|^p), \quad \text{for } k\in\{2,3\},\ (k,\ell)\neq(2,2)\ .$$

Under the boundedness assumption on the derivative with respect to x, the coefficients a and b are globally Lipschitz and equation (2.1) admits a unique solution. However, it is difficult to deal with transition density ratio of X due to the lack of its explicit form. Now, to solve this problem, on the same probability space for (X_t), we define the alternative scheme for (2.1) (proposed in Clément and Gloter (2019)) on the time grid $t_i^n = i/n$ for $i\in\{0,1,\dots,n\}$,

$$\overline{X}_{t_{i+1}^n} = \xi_{t_{i+1}^n - t_i^n}(\overline{X}_{t_i^n},\mu) + a(\overline{X}_{t_i^n},\sigma)(J_{t_{i+1}^n} - J_{t_i^n}) \tag{2.2}$$

where $(\xi_t(x,\mu))$ solves the ODE

$$\xi_t(x,\mu) = x + \int_0^t b(\xi_s(x,\mu),\mu)ds, \quad t\ge 0.$$

To help analyse the asymptotic properties for the statistical experiment using the observations $(X_{t_i^n})$, we work with the observation $(\overline{X}_{t_i^n})$ that solves (2.2) with θ_0. In some particular cases, it can be deduced (see Section 3) that the estimation based on $(\overline{X}_{t_i^n})$ has the same asymptotic properties as estimation based on $(X_{t_i^n})$. Let

$$z_n(x,y,\theta) = \frac{y-\xi_{1/n}(x,\mu)}{n^{-1/\beta}a(x,\sigma)}.$$

Note that $(z_n(\overline{X}_{t_i^n},\overline{X}_{t_{i+1}^n},\theta_0))_i$ are i.i.d. β_0-stable random variables. Consequently, the log-likelihood function based on the observations $(\overline{X}_{t_i^n})_{0\le i\le n}$ has an explicit form and is given by

$$\ell_n(\theta) = \sum_{i=0}^{n-1} \log(n^{1/\beta}a(\overline{X}_{t_i^n},\sigma)^{-1}\phi_\beta(z_n(\overline{X}_{t_i^n},\overline{X}_{t_{i+1}^n},\theta))), \tag{2.3}$$

where ϕ_β denotes the density of J_1.

In order to prove the LAMN property for $(\overline{X}_{t_i^n})_{0\le i\le n}$, we use the Taylor's formula with the help of our law of large number (see Theorem 5.2) and the explicit expression of the probability density functions which allows us to avoid the need of Malliavin calculus. Let $\varphi_n(\cdot)$ a scale such that $\varphi_n(\theta)$ is a sequence of 3×3 non-singular matrices satisfying $\|\varphi_n(\theta)\|\to 0$ as $n\to\infty$. In what follows, by Taylor's expansion, for arbitrary bounded $u\in\mathbb{R}^3$, there exists $0<\varepsilon_n<1$ such that

$$\begin{aligned}&\ell_n(\theta+\varphi_n(\theta)u) - \ell_n(\theta)\\ &\quad = u^\top\Delta_n(\theta) - \frac{1}{2}u^\top\mathcal{J}_n(\theta)u + \frac{1}{3!}((\varphi_n(\theta)u)^\top\cdot\partial_\theta\mathcal{I}_n(\theta+\varepsilon_n\varphi_n(\theta)u)\cdot(\varphi_n(\theta)u))^\top(\varphi_n(\theta)u)\end{aligned} \tag{2.4}$$

where
$$\Delta_n(\theta) := \varphi_n(\theta)^\top\partial_\theta\ell_n(\theta),$$
$$\mathcal{I}_n(\theta) := -\partial_\theta^2\ell_n(\theta) \quad \text{and} \quad \mathcal{J}_n(\theta) := \varphi_n(\theta)^\top\mathcal{I}_n(\theta)\varphi_n(\theta).$$

For any $h \in \mathbb{R}^3$ and $\widetilde{\theta} \in \Theta$, we use the notation $h^\top \cdot \partial_\theta \mathcal{I}_n(\widetilde{\theta}) \cdot h = \begin{pmatrix} h^\top \partial_\mu \mathcal{I}_n(\widetilde{\theta}) h \\ h^\top \partial_\sigma \mathcal{I}_n(\widetilde{\theta}) h \\ h^\top \partial_\beta \mathcal{I}_n(\widetilde{\theta}) h \end{pmatrix} \in \mathbb{R}^3$.

Also, we note that

$$\partial_\mu z_n = -n^{1/\beta} \frac{\partial_\mu \xi_{1/n}(x,\mu)}{a(x,\sigma)}, \quad , \partial_\sigma z_n = -\frac{\partial_\sigma a}{a} z_n, \quad \partial_\beta z_n = -\frac{\log n}{\beta^2} z_n. \tag{2.5}$$

To simplify the notations, as in Clément and Gloter (2020), we introduce the functions

$$\begin{aligned} h_\beta(z) =& (\partial_z \phi_\beta / \phi_\beta)(z) \quad k_\beta(z) = 1 + z h_\beta(z), \\ q_\beta(z) =& z \partial_z k_\beta(z), \quad r_\beta(z) = z \partial_\beta h_\beta(z), \quad f_\beta(z) = (\partial_\beta \phi_\beta / \phi_\beta)(z). \end{aligned}$$

From these notations, we easily see that $\partial_z k_\beta(z) = h_\beta(z) + z \partial_z h_\beta(z)$ and $\partial_z f_\beta(z) = (\partial_\beta h_\beta)(z)$.

3. Main results

First, we state in Section 3.1 the main results on the LAMN property which differs depending on the form of the scale coefficient, namely the non-multiplicative (NM) and multiplicative case (M). Then, we recall in Section 3.2 the asymptotic equivalence between the discrete time observations of the solution of the SDE (2.1) and the corresponding approximation scheme (2.2) and the possible transfer of the lower bounds for estimation risk. Third, a one-step procedure is described and shown to be asymptotically efficient in Section 3.3.

3.1. LAMN properties

Here, we assume that (A) holds and now present our main results, namely LAMN property for $(\overline{X}_{t_i^n})$ in two cases the non-multiplicative (NM) and multiplicative (M) cases. The proofs of these results are postponed in Section 5. We define the sequence $(\overline{P}_n^{\theta_0})$ as distribution of $(\overline{X}_{t_i^n})$.

3.1.1. Non-multiplicative case

Here, we assume in addition of (A):

- $s \mapsto \frac{\partial_\sigma a}{a}(X_s, \sigma_0)$ is almost surely non constant.
- Almost surely, $\exists t \in (0,1)$, such that $\partial_\mu b(X_t, \mu_0) \neq 0$, where $(X_t)_{t\in[0,1]}$ solves (2.1) for the parameter value θ_0.

Theorem 3.1. *Let X be the solution of (2.1) with the parameter value θ_0. For our non-multiplicative case, we take $\varphi_n(\theta_0) = \begin{pmatrix} n^{1/2-1/\beta_0} & 0 & 0 \\ 0 & n^{-1/2} & 0 \\ 0 & 0 & \frac{1}{\log n \sqrt{n}} \end{pmatrix}$. Then, under assumption (A), the family $(\overline{P}_n^{\theta_0})$ satisfies the LAMN property. More precisely, we have the following convergences under $\mathbb{P}_{\theta_0}$:*

$$\left| \ell_n(\theta_0 + \varphi_n(\theta_0)u) - \ell_n(\theta_0) - \left(u^\top \Delta_n(\theta_0) - \frac{1}{2} u^\top \mathcal{J}_n(\theta_0) u \right) \right| \xrightarrow{\mathbb{P}_{\theta_0}} 0$$

for any $u \in \mathbb{R}^3$, *and* $(\Delta_n(\theta_0), \mathcal{J}_n(\theta_0)) \overset{stably}{\Longrightarrow} (\Delta(\theta_0), \mathcal{I}(\theta_0))$ *where* $\Delta(\theta_0) = \mathcal{I}(\theta_0)^{1/2}\mathcal{N}$ *with* $\mathcal{N}$ *a standard Gaussian variable independent of* $\mathcal{I}(\theta_0)$ *and* $\mathcal{I}(\theta_0)$ *is the random asymptotic information of the statistical model*

$$\mathcal{I}(\theta_0) = \begin{pmatrix} \int_0^1 \frac{\partial_\mu b(X_s,\mu_0)^2}{a(X_s,\sigma_0)^2} ds \mathbb{E}_{\theta_0}(h^2_{\beta_0}(J_1)) & 0 \\ 0 & \overline{\mathcal{I}}(\theta_0) \end{pmatrix} \tag{3.1}$$

with $$\overline{\mathcal{I}}(\theta_0) = \begin{pmatrix} \int_0^1 \frac{\partial_\sigma a(X_s,\sigma_0)^2}{a(X_s,\sigma_0)^2} ds \mathbb{E}_{\theta_0}(k^2_{\beta_0}(J_1)) & \frac{1}{\beta_0^2}\int_0^1 \frac{\partial_\sigma a(X_s,\sigma_0)}{a(X_s,\sigma_0)} ds \mathbb{E}_{\theta_0}(k^2_{\beta_0}(J_1)) \\ \frac{1}{\beta_0^2}\int_0^1 \frac{\partial_\sigma a(X_s,\sigma_0)}{a(X_s,\sigma_0)} ds \mathbb{E}_{\theta_0}(k^2_{\beta_0}(J_1)) & \frac{1}{\beta_0^4}\mathbb{E}_{\theta_0}(k^2_{\beta_0}(J_1)) \end{pmatrix}.$$

Moreover, there exists a local maximum $\widehat{\theta}_n$ *of* ℓ_n *with probability tending to* 1, *for which*

$$\varphi_n(\theta_0)^{-1}(\widehat{\theta}_n - \theta_0) \overset{stably}{\Longrightarrow} \mathcal{I}(\theta_0)^{-1/2}\mathcal{N}.$$

Remark 3.2. Note that the matrix $\mathcal{I}(\theta_0)$ is invertible a.s. since

$$\frac{1}{\beta_0^4}\mathbb{E}_{\theta_0}(k^2_{\beta_0}(J_1))\left(\int_0^1 \frac{\partial_\sigma a(X_s,\sigma_0)^2}{a(X_s,\sigma_0)^2} ds - \left(\int_0^1 \frac{\partial_\sigma a(X_s,\sigma_0)}{a(X_s,\sigma_0)} ds\right)^2\right) > 0, \quad a.s.$$

3.1.2. Multiplicative case

Here, we assume in addition of (A):

- $a(x,\sigma) = \sigma\bar{a}(x)$ for all x.
- Almost surely, $\exists t \in (0,1)$, such that $\partial_\mu b(X_t,\mu_0) \neq 0$, where $(X_t)_{t\in[0,1]}$solves (2.1) for the parameter value θ_0.

Theorem 3.3. *Let* X *be the solution of* (2.1) *with the parameter value* θ_0. *For our multiplicative case, we take* $\varphi_n(\theta_0) = \frac{1}{\sqrt{n}}\begin{pmatrix} n^{1-1/\beta_0} & 0 & 0 \\ 0 & \varphi_{11,n}(\theta_0) & \varphi_{12,n}(\theta_0) \\ 0 & \varphi_{21,n}(\theta_0) & \varphi_{22,n}(\theta_0) \end{pmatrix}$, *where*

$$\begin{cases} \varphi_{11,n}(\theta_0)\frac{1}{\sigma_0} + \varphi_{21,n}(\theta_0)\frac{\log n}{\beta_0^2} \to \overline{\varphi}_{11}, \quad \varphi_{12,n}(\theta_0)\frac{1}{\sigma_0} + \varphi_{22,n}(\theta_0)\frac{\log n}{\beta_0^2} \to \overline{\varphi}_{12}, \\ \varphi_{21,n}(\theta_0) \to \overline{\varphi}_{21}, \quad \varphi_{22,n}(\theta_0) \to \overline{\varphi}_{22}, \\ \overline{\varphi}_{11}\overline{\varphi}_{22} - \overline{\varphi}_{12}\overline{\varphi}_{21} > 0 \end{cases}.$$

Then, under assumption (A), the family $(\overline{P}^n_{\theta_0})$ *satisfies the LAMN property with asymptotic score function* $\Delta(\theta_0)$ *and random asymptotic information matrix* $\mathcal{I}(\theta_0)$ *where* $\Delta(\theta_0) = \mathcal{I}(\theta_0)^{1/2}\mathcal{N}$ *with* $\mathcal{N}$ *a standard Gaussian variable independent of* $\mathcal{I}(\theta_0)$ *and*

$$\mathcal{I}(\theta_0) = \begin{pmatrix} \int_0^1 \frac{\partial_\mu b(X_s,\mu_0)^2}{a(X_s,\sigma_0)^2} ds \mathbb{E}_{\theta_0}(h^2_{\beta_0}(J_1)) & 0 \\ 0 & \overline{\varphi}^\top \overline{\mathcal{I}}(\theta_0)\overline{\varphi} \end{pmatrix} \tag{3.2}$$

with $\overline{\varphi} = \begin{pmatrix} \overline{\varphi}_{11} & \overline{\varphi}_{12} \\ \overline{\varphi}_{21} & \overline{\varphi}_{22} \end{pmatrix}$ *and* $\overline{\mathcal{I}}(\theta_0) = \begin{pmatrix} \mathbb{E}_{\theta_0}(k^2_{\beta_0}(J_1)) & -\mathbb{E}_{\theta_0}((k_{\beta_0} f_{\beta_0})(J_1)) \\ -\mathbb{E}_{\theta_0}((k_{\beta_0} f_{\beta_0})(J_1)) & \mathbb{E}_{\theta_0}(f^2_{\beta_0}(J_1)) \end{pmatrix}$.
Moreover, there exists a local maximum $\widehat{\theta}_n$ *of* ℓ_n *with probability tending to* 1, *for which*

$$\varphi_n(\theta_0)^{-1}(\widehat{\theta}_n - \theta_0) \overset{stably}{\Longrightarrow} \mathcal{I}(\theta_0)^{-1/2}\mathcal{N}, \quad under\ \mathbb{P}_{\theta_0}$$

3.2. Asymptotic equivalence

Let us first recall the definition of total variation (see e.g. Strasser (1985))

Definition 3.4. *The total variation between two probabilities measures* P *and* Q *on* $(\Omega, \mathcal{F})$ *dominated by* ν *is defined by*

$$d_{TV}(P,Q) = \sup_{A\in\mathcal{F}} |P(A) - Q(A)| = \frac{1}{2}\int \left|\frac{dP}{d\nu} - \frac{dQ}{d\nu}\right| d\nu.$$

Now, let us consider the two experiments as follows.

- Experiment $\mathcal{E}^n$ with $P^n_{\theta_0}$ distribution of $(X_{t_i^n})_{i\in\{0,\dots,n\}}$.
- Experiment $\overline{\mathcal{E}}^n$ with $\overline{P}^n_{\theta_0}$ distribution of $(\overline{X}_{t_i^n})_{i\in\{0,\dots,n\}}$.

For any subset $K \subseteq \Theta$ containing θ_0, the Le Cam distance, Δ, (see Le Cam (1964)) is bounded by

$$\Delta(\mathcal{E}^n, \overline{\mathcal{E}}^n; K) \leq \sup_{\theta\in K} d_{TV}(P^n_\theta, \overline{P}^n_\theta).$$

Let $w : \mathbb{R}^3 \to \mathbb{R}_+$ be any non-constant continuous bounded subconvex loss function. The minimax estimation risk for the statistical experiment $\mathcal{G}^n$ is defined by

$$\mathcal{R}(\mathcal{G}^n, w; K) = \inf_{T_n} \sup_{\vartheta\in K} \mathbb{E}_\vartheta \left(w\left(\varphi_n^{-1}(\theta_0)(T_n - \vartheta)\right)\right),$$

where the infimum is taken over all the estimator T_n of θ_0. From Theorem 3.1 and Theorem 3.3, combined with (Höpfner, 2014, Sections 1.11 and 7.12), we have

$$\lim_{n\to\infty} \mathcal{R}(\overline{\mathcal{E}}^n, w; K) \geq \mathbb{E}(w(\mathcal{I}(\theta_0)^{-1/2}\mathcal{N})).$$

From (Shiryaev and Spokoiny, 2000, Theorem 2.1), it can be shown that if $\lim_{n\to\infty} \Delta(\mathcal{E}^n, \overline{\mathcal{E}}^n; K) = 0$, then

$$\lim_{n\to\infty} \mathcal{R}(\mathcal{E}^n, w; K) = \lim_{n\to\infty} \mathcal{R}(\overline{\mathcal{E}}^n, w; K). \tag{3.3}$$

In addition, by (Shiryaev and Spokoiny, 2000, Section 2.1.3), it turns out that for the unbounded loss functions $w(\varphi_n^{-1}(\theta_0)(T_n - \vartheta)) = \|\varphi_n^{-1}(\theta_0)(T_n - \vartheta)\|^p$, $p > 0$ typically used in estimation problems, the assertion (3.3) remains valid.
Consequently, the statistical inference in experiment $\mathcal{E}^n$ inherits the same asymptotic properties as in $\overline{\mathcal{E}}^n$ (see e.g. (Clément, 2023, Remark 4.3)). Hence, for instance, we deduce the following result.

Theorem 3.5. *Let $K_1 \subset \mathbb{R}$ be compact and K_2 be a compact subset of $(0,2)$ such that $K = K_1 \times (0,\infty) \times K_2 \subseteq \Theta$, we assume that the function $a(x,\sigma) = \sigma$ constant for any $x \in \mathbb{R}$, (A) holds with further*

$$\sup_{\substack{\mu \in K_1 \\ x \in \mathbb{R}}} (|\partial_x b(x,\mu)| + |\partial_x^2 b(x,\mu)|) \leq C$$

then, the experiments are asymptotically equivalent in Le Cam sense since

$$\lim_{n\to\infty} \Delta(\mathcal{E}^n, \overline{\mathcal{E}}^n; K) = 0.$$

Proof. This is direct from (Clément, 2023, Theorem 4.1 (i)) that

$$d_{TV}(P_\theta^n, \overline{P}_\theta^n) \leq C(\sigma, b, \beta) \max\left\{\frac{1}{\sqrt{n}}, \frac{1}{n^{4\beta/(\beta+2)}}\right\}$$

where $C(\sigma,b,\beta)$ has exponential growth in $\|\partial_x b\|_\infty$ and polynomial growth in $\|\partial_x^2 b\|_\infty$, $1/\sigma$, σ, $1/\beta$ and $1/(\beta-2)$. □

Remark 3.6. From Theorem 3.5, when a is constant, we have the asymptotic equivalence between the two experiments $\overline{\mathcal{E}}^n$ and $\mathcal{E}^n$. Then, thanks to the LAMN property proven above, the quasi-likelihood estimator proposed in Clément and Gloter (2020) for the original experiment $\mathcal{E}^n$ is consequently efficient in the sense of the Convolution Theorem and attains the local asymptotic minimax bound (see e.g (Höpfner, 2014, Theorems 7.10 & 7.12)). For the general form of the function a, the equivalence between these two experiments remains an open question.

3.3. One-step statistical procedure

We present in this section the construction of the one-step procedure when the asymptotic equivalence between the two experiments $\overline{\mathcal{E}}^n$ and $\mathcal{E}^n$ and the result on asymptotic efficiency hold. It is based on an initial guess estimation which is corrected by a single step of the Fisher scoring method on the log-likelihood function.

3.3.1. Moment estimator

For the one-step procedure, our initial estimators is built with the ratio of generalized p-variations (see Todorov (2013)). Namely, denoting $\Delta_i^n X = X_{t_i^n} - X_{t_{i-1}^n}$,

$$V_n^1(p,X) = \sum_{i=2}^n |\Delta_i^n X - \Delta_{i-1}^n X|^p \text{ and } V_n^2(p,X) = \sum_{i=4}^n |\Delta_i^n X - \Delta_{i-1}^n X + \Delta_{i-2}^n X - \Delta_{i-3}^n X|^p,$$

we define

$$\widehat{\beta}_n^0 = \frac{p \log 2}{\log(V_n^2(p,X)/V_n^1(p,X))} \mathbb{1}_{\{V_n^2(p,X) \neq V_n^1(p,X)\}}$$

and $\widehat{\sigma}_n^0$ which satisfies

$$n^{\frac{p}{\widehat{\beta}_n^0} - 1} V_n^1(p,X) = \mu_p(\widehat{\beta}_n^0) \int_0^1 |a(X_s, \widehat{\sigma}_n^0)|^p ds. \tag{3.4}$$

For the multiplicative case where $a(X_s, \widehat{\sigma}_n^0) = \widehat{\sigma}_n^0\, \overline{a}(X_s)$, we get the form explicit of

$$\widehat{\sigma}_n^0 = \left[n^{\frac{p}{\widehat{\beta}_n^0}-1} V_n^1(p,X) \left(\mu_p(\widehat{\beta}_n^0) \int_0^1 |\overline{a}(X_s)|^p ds \right)^{-1} \right]^{1/p}$$

where $\mu_p(\widehat{\beta}_n^0) = 2^{p/\widehat{\beta}_n^0} \frac{2^p \Gamma(\frac{p+1}{2}) \Gamma(1-p/\widehat{\beta}_n^0)}{\sqrt{\pi}\Gamma(1-p/2)}$. The parameter μ_0 is estimated by maximizing the log-likelihood function $\ell_n\left(\mu, \widehat{\sigma}_n^0, \widehat{\beta}_n^0\right)$ with respect to μ. Therefore,

$$\widehat{\theta}_n^0 = (\widehat{\mu}_n^0, \widehat{\sigma}_n^0, \widehat{\beta}_n^0). \tag{3.5}$$

Remark 3.7. We have from (Todorov, 2013, Corollary 1, Theorem 3 and (21)) that if p takes its values in $\left(\frac{|\beta_0-1|}{2(\beta_0\wedge 1)}, \frac{\beta_0}{2}\right)$ and provided $\beta_0 > 2/3$, we have that $\sqrt{n}(\widehat{\beta}_n^0 - \beta_0)$ is tight for both multiplicative (M) and non-multiplicative (NM) cases and that $(\sqrt{n}/\log n)(\widehat{\sigma}_n^0 - \sigma_0)$ is tight for the multiplicative (M) case. For the non-multiplicative case (NM), the tightness of $\frac{n^{1/2}}{\log n}(\widehat{\sigma}_n^0 - \sigma_0)$ is proved in the next Lemma 3.8.

The next Lemma 3.8 and Theorem 3.9 are proven in the supplementary document.

Lemma 3.8. *Let us assume that $\widehat{\beta}_n^0 \xrightarrow{\mathbb{P}} \beta_0$, (A) holds and that a and $\partial_\sigma a$ are non-singular and positive. Then $\widehat{\sigma}_n^0$ defined in (3.4) satisfies that $\frac{n^{1/2}}{\log n}(\widehat{\sigma}_n^0 - \sigma_0)$ is tight.*

For this choice of $\widehat{\theta}_n^0$, we get in the following theorem the global uniqueness of $\widehat{\mu}_n^0$ under some further assumptions on β_0 and the coefficient function b.

Theorem 3.9. *Assume that (A) holds and further $\mu \in int(A)$ (the interior of A) for A a compact subset of $\mathbb{R}$, $\beta_0 > 1$ and*

$$\sup_{\substack{\mu \in int(A) \\ x \in \mathbb{R}}} (|\partial_x b(x,\mu)| + |\partial_\mu b(x,\mu)| + |\partial_\mu^2 b(x,\mu)|) \le C.$$

Let us denote

$$G_n(\mu) = n^{1-2/\widehat{\beta}_n^0} \partial_\mu \ell_n(\mu, \widehat{\sigma}_n^0, \widehat{\beta}_n^0)$$

where $(\sqrt{n}/\log n)(\widehat{\sigma}_n^0 - \sigma_0, \widehat{\beta}_n^0 - \beta_0)$ is tight. Then, any sequence $(\widehat{\mu}_n^0, n \ge 1)$ that solves $G_n(\mu) = 0$ is consistent and eventually unique. Moreover, we have $\frac{n^{1/\beta_0 - 1/2}}{(\log n)^2}(\widehat{\mu}_n^0 - \mu_0)$ is tight.

3.3.2. Asymptotic efficiency

The one-step estimator $\widehat{\theta}_n^1$ is defined by

$$\widehat{\theta}_n^1 = \widehat{\theta}_n^0 + (\varphi_n(\widehat{\theta}_n^0)^{-1^\top} \mathcal{I}(\widehat{\theta}_n^0) \varphi_n(\widehat{\theta}_n^0)^{-1})^{-1} \partial_\theta \ell_n(\widehat{\theta}_n^0). \tag{3.6}$$

Theorem 3.10. *Suppose that (A) holds and $(\frac{n^{1/\beta_0-1/2}}{(\log n)^2}(\widehat{\mu}_n^0 - \mu_0), (\sqrt{n}/\log n)(\widehat{\sigma}_n^0 - \sigma_0), \sqrt{n}(\widehat{\beta}_n^0 - \beta_0))$ is tight and β_0 in a compact subset of $(0,2)$, under the LAMN property and sufficiently regular Fisher information matrix, the sequence $(\widehat{\theta}_n^1, n \ge 1)$ is asymptotically efficient.*

Proof. With a choice of initial estimator satisfying the tightness assumption, the asymptotic efficiency of the one-step estimation depends on the form of $\varphi_n(\theta_0)$ defined in the LAMN property. More precisely, if the initial estimator $\widehat{\theta}_n^0$ is such that

$$\varphi_n(\theta_0)^{-1}(\widehat{\theta}_n^0 - \theta_0) = O_P(1), \tag{3.7}$$

then we say that it has a good rate of convergence and the proof of efficiency for the corresponding one-step estimation is straight-forward by (Höpfner, 2014, Theorem 7.19 (a)&(c)). Otherwise, we need to prove its efficiency using some classical techniques and the help of Theorem 5.2. Here and in the sequel, $O_P(n^{-c})$ and $o_P(n^{-c})$ mean that $n^c O_P(n^{-c})$ is bounded in probability and $n^c o_P(n^{-c})$ converges in probability, i.e., for any $\varepsilon > 0$, there exists $C > 0$ such that

$$\mathbb{P}_{\theta_0}(n^c|O_P(n^{-c})| > C) \le \varepsilon \text{ and } \lim_{n\to\infty} \mathbb{P}_{\theta_0}(n^c|o_P(n^{-c})| \ge \varepsilon) = 0.$$

Now, let us have a closer look into our two cases:

• For the non-multiplicative (NM) case, its initial estimator does not possess a good rate of convergence for the estimation of θ_0. Precisely, we only have that $\nu_n^{-1}(\theta_0)(\widehat{\theta}_n^0 - \theta_0)$ is tight for $\nu_n(\theta_0) = diag(\frac{(\log n)^2}{n^{1/\beta_0 - 1/2}}, \frac{\log n}{\sqrt{n}}, \frac{1}{\sqrt{n}})$. Here, we need to prove that for $\varphi_n(\theta_0) = diag(\frac{1}{n^{1/\beta_0 - 1/2}}, \frac{1}{\sqrt{n}}, \frac{1}{\log n\sqrt{n}})$, the one-step estimator defined in (3.6) satisfies

$$\varphi_n^{-1}(\theta_0)(\widehat{\theta}_n^1 - \theta_0) \xrightarrow{\mathcal{L}} \mathcal{I}^{-1/2}(\theta_0)\mathcal{N},$$

with $\mathcal{N}$ a standard Gaussian variable independent of $\mathcal{I}(\theta_0)$. First, from (3.6) and Taylor's formula,

$$\begin{aligned}
\varphi_n^{-1}(\theta_0)(\widehat{\theta}_n^1 - \theta_0) &= \varphi_n^{-1}(\theta_0)(\widehat{\theta}_n^0 - \theta_0) + \mathcal{I}^{-1}(\widehat{\theta}_n^0)\varphi_n(\theta_0)\partial\ell_n(\widehat{\theta}_n^0) \\
&= \varphi_n^{-1}(\theta_0)(\widehat{\theta}_n^0 - \theta_0) + \mathcal{I}^{-1}(\widehat{\theta}_n^0)\varphi_n(\theta_0)[\partial\ell_n(\theta_0) + \partial^2\ell_n(\tilde{\theta}_n)(\widehat{\theta}_n^0 - \theta_0)] \\
&= \mathcal{I}^{-1}(\widehat{\theta}_n^0)\varphi_n(\theta_0)\partial\ell_n(\theta_0) + \mathcal{I}^{-1}(\widehat{\theta}_n^0)[\mathcal{I}(\widehat{\theta}_n^0) + \varphi_n(\theta_0)\partial^2\ell_n(\tilde{\theta}_n)\varphi_n(\theta_0)]\varphi_n^{-1}(\theta_0)(\widehat{\theta}_n^0 - \theta_0) \\
&= \mathcal{I}^{-1}(\widehat{\theta}_n^0)\varphi_n(\theta_0)\partial\ell_n(\theta_0) + \mathcal{I}^{-1}(\widehat{\theta}_n^0)[\mathcal{I}(\widehat{\theta}_n^0) - \mathcal{I}(\theta_0)]\varphi_n^{-1}(\theta_0)(\widehat{\theta}_n^0 - \theta_0) \\
&\quad + \mathcal{I}^{-1}(\widehat{\theta}_n^0)[\mathcal{I}(\theta_0) + \varphi_n(\theta_0)\partial^2\ell_n(\tilde{\theta}_n)\varphi_n(\theta_0)]\varphi_n^{-1}(\theta_0)(\widehat{\theta}_n^0 - \theta_0)
\end{aligned}$$

where $\tilde{\theta}_n$ is some value between $\widehat{\theta}_n^0$ and θ_0. From this, the asymptotic normality of the first term $\mathcal{I}^{-1}(\widehat{\theta}_n^0)\varphi_n(\theta_0)\partial\ell_n(\theta_0)$ is obtained from the LAMN property. Next, by Asumption (A) and approximation (5.5) with β_0 in a compact set, we have

$$|\mathcal{I}(\widehat{\theta}_n^0) - \mathcal{I}(\theta_0)| \le C(1 + \sup_{s\in[0,1]} |X_s|^p)\,|\widehat{\theta}_n^0 - \theta_0| = O_P(\nu_n(\theta_0)), \quad \text{for some } p > 0,$$

and since $\nu_n(\theta_0)\varphi_n^{-1}(\theta_0)\nu_n(\theta_0) \to 0$, the second term

$$\mathcal{I}^{-1}(\widehat{\theta}_n^0)[\mathcal{I}(\widehat{\theta}_n^0) - \overline{\mathcal{I}}(\theta_0)]\varphi_n^{-1}(\theta_0)(\widehat{\theta}_n^0 - \theta_0) \xrightarrow{\mathbb{P}_{\theta_0}} 0.$$

Considering the last term, since $\varphi_n^{-1}(\theta_0)\nu_n(\theta_0) = O_P((\log n)^2)$, we have

$$\begin{aligned}
&\mathcal{I}^{-1}(\widehat{\theta}_n^0)[\mathcal{I}(\theta_0) + \varphi_n(\theta_0)\partial^2\ell_n(\tilde{\theta}_n)\varphi_n(\theta_0)]\varphi_n^{-1}(\theta_0)(\widehat{\theta}_n^0 - \theta_0) \\
=&\mathcal{I}^{-1}(\widehat{\theta}_n^0)[\mathcal{I}(\theta_0) + \varphi_n(\theta_0)\partial^2\ell_n(\tilde{\theta}_n)\varphi_n(\theta_0)]O_P((\log n)^2)\nu_n^{-1}(\theta_0)(\widehat{\theta}_n^0 - \theta_0)
\end{aligned}$$

Since $\|\nu_n^{-1}(\theta_0)(\tilde{\theta}_n - \theta_0)\| \leq C$, one can follow the arguments in (Clément and Gloter, 2020, The proof of Theorem 3.1), combined with our Theorem 5.2 to get that

$$\mathcal{I}(\theta_0) + \varphi_n(\theta_0)\partial^2\ell_n(\tilde{\theta}_n)\varphi_n(\theta_0) = o_P(n^{-\varepsilon}).$$

Finally, we obtain that

$$\varphi_n^{-1}(\theta_0)(\widehat{\theta}_n^1 - \theta_0) = \mathcal{I}^{-1}(\widehat{\theta}_n^0)\varphi_n(\theta_0)\partial\ell_n(\theta_0) + o_P(1).$$

Then by (Höpfner, 2014, Theorem 7.11), $\widehat{\theta}_n^1$ is regular and efficient at θ_0.

• For the multiplicative (M) case, the initial estimator possesses a good rate for σ_0 and β_0 but not for μ_0. Thanks to the diagonal and the block-diagonal forms of $\varphi_n(\theta_0)$ and the Fisher information matrix respectively, we can consider the estimation of μ_0 separately from the one of pair (σ_0, β_0). Then, the asymptotic efficiency of the estimation is immediate for σ_0 and β_0 by (Höpfner, 2014, Theorem 7.19) and is obtainable for μ_0 by similar arguments as in NM case above. □

Remark 3.11. The explicit observed information matrices (3.1) and (3.2) can not be computed since the trajectory of X which is only observed discretely. In practice, the integrals are replaced by Riemann sums based on the discrete observations. Under regularity conditions, the proof of Theorem 3.10 can be rewritten.

4. Numerical simulations

In this section, the performances in terms of asymptotic variance and computational time on samples of finite size are illustrated for the moment estimator (ME), the one-step estimator (OS) from Section 3.3 and the maximum likelihood estimator (MLE) from Theorems 3.1 and 3.3 considering two cases discussed above: multiplicative (M) and non-multiplicative (NM). It is important to note that here our numerical test results are obtained by using the statistical experiment $\mathcal{E}^n$ corresponding to the observations on the true trajectory (X_t), instead of the experiment $\overline{\mathcal{E}^n}$ as considered in the theoretical Section 3.1. The OS performs quite similarly to the MLE in terms of variance but it reduces significantly the computational time.

All the models have linear drift $b(x, \mu) = \mu x$. For these choices, we can easily obtain the explicit solution of the ordinary differential equation $\xi_{1/n}(x, \mu) = xe^{\mu/n}$. Moreover, the conditions of Theorem 3.9 are satisfied which affirm the uniqueness of the initial estimator for the drift. Note that for the more complicated form of drift where the exact solution can not be found, the quantity $\xi_{1/n}(x, \mu)$ can be replaced by its Euler approximation $\xi_{1/n}(x, \mu) \simeq x + b(x, \mu)/n$. Indeed, as shown in (Clément and Gloter, 2020, Remark 3.2), they state that when $\beta > 2/3$, the quality of estimation is the same when using the approximation of $\xi_{1/n}(x, \mu)$ as when using its true value.

Here, we generate the vector of observations $(X_{t_i^n})_{0\leq i\leq n}$ by sub-sampling a refined simulation of the process $(X_t)_{t\geq 0}$ (by an Euler scheme with time-step $(1000n)^{-1}$).

4.1. Multiplicative case

Here, we take

$$\varphi_n(\theta_0) = \frac{1}{\sqrt{n}}\begin{pmatrix} n^{1-1/\beta_0} & 0 & 0 \\ 0 & 1 & -\beta_0^{-2}\sigma_0 \log n \\ 0 & 0 & 1 \end{pmatrix}$$

which yield $\overline{\varphi}_{11} = \sigma_0^{-1}$, $\overline{\varphi}_{12} = \overline{\varphi}_{21} = 0$ and $\overline{\varphi}_{22} = 1$.
We emphasize that, as the asymptotic law of $\widehat{\mu}_n$ is mixed normal, the estimation error $\widehat{\mu}_n - \mu_0$ is re-scaled by a factor involving the random quantity

$$U_{\mu_0} = \int_0^1 \frac{\partial_\mu b(X_s, \mu_0)^2}{a(X_s, \sigma_0)^2} ds$$

that we approximate, in practice, by the Riemann sum based on the observations $(X_{t_i^n})_{0 \le i \le n}$.
From this, following the theoretical results obtained by using the statistical experiment $\overline{\mathcal{E}}^n$ in Section 3, we can easily deduce the following asymptotic properties. In particular, using Theorem 3.3, the re-normalized error by maximum likelihood estimation converges to a Gaussian limit and the mixed normal form at the limit is eliminated, namely,

$$\widetilde{\varphi_n}(\theta_0)^{-1}(\widehat{\theta}_n - \theta_0) \overset{stably}{\Longrightarrow} \mathcal{N}(0, \widetilde{\mathcal{I}}(\theta_0)^{-1}),$$

where

$$\widetilde{\varphi_n}(\theta_0) = \frac{1}{\sqrt{n}} \begin{pmatrix} U_{\mu_0}^{-\frac{1}{2}} n^{1-1/\beta_0} & 0 & 0 \\ 0 & 1 & -\beta_0^{-2}\sigma_0 \log n \\ 0 & 0 & 1 \end{pmatrix},$$

$$\widetilde{\mathcal{I}}(\theta_0) = \begin{pmatrix} \mathbb{E}_{\theta_0}(h_{\beta_0}^2(J_1)) & 0 & 0 \\ 0 & \frac{1}{\sigma_0^2}\mathbb{E}_{\theta_0}(k_{\beta_0}^2(J_1)) & -\frac{1}{\sigma_0}\mathbb{E}_{\theta_0}(k_{\beta_0}(J_1) f_{\beta_0}(J_1)) \\ 0 & -\frac{1}{\sigma_0}\mathbb{E}_{\theta_0}(k_{\beta_0}(J_1) f_{\beta_0}(J_1)) & \mathbb{E}_{\theta_0}(f_{\beta_0}^2(J_1)) \end{pmatrix} \text{ non random.}$$

Considering our one-step procedure, we take the moment estimators described in (3.5) as the initial estimators. In this case, from Theorem 3.3 and Remark 3.7 we have that $\widetilde{\varphi_n}(\theta_0)^{-1}(\widehat{\theta}_n^0 - \theta_0)$ is tight and from Theorem 3.10

$$\widetilde{\varphi_n}(\theta_0)^{-1}(\widehat{\theta}_n^1 - \theta_0) \overset{stably}{\Longrightarrow} \mathcal{N}(0, \widetilde{\mathcal{I}}(\theta_0)^{-1}).$$

In what follows, we show the numerical simulations based on these re-normalized statistical errors but using instead the statistical experiment $\mathcal{E}^n$. For this multiplicative case (M), we consider two kinds of models: the Ornsteins-Uhlenbeck (O-U) model

$$X_t = \int_0^t \mu X_s ds + \sigma J_t, \quad t \in [0,1]. \tag{4.1}$$

and the square root model (SR)

$$X_t = \int_0^t \mu X_s ds + \sigma \int_0^t \sqrt{1 + X_{s-}^2} dJ_s, \quad t \in [0,1]. \tag{4.2}$$

Figures 1-3, we plot the histograms of the re-scaled errors of estimation together with the density of their Gaussian limits in red lines for the O-U model. In each panel, this solid line represents the asymptotic normal distribution with an efficient variance. The implementations of the likelihood, the score and Fisher information for computing the sequences of the MLE and the OS are based on the techniques of Matsui and Takemura (2004). From here, we can observe that the histograms of the ME are far from the efficient asymptotic normal distributions, whereas both of the MLE and the OS

sequences show much better performances. For the second model, the comparison between the three methods is shown in Table 1 by analyzing their renormalized mean squared errors. Almost the same as the first model, it is shown that MLE and OS have similar performance in terms of asymptotic variance and but OS consumes much less time for the computations than MLE.
In the figures 1-3 and Table 1, the asymptotically efficient variance is $1/\mathbb{E}_{\theta_0}(h^2_{\beta_0}(J_1))$ for the estimation of μ_0 and is

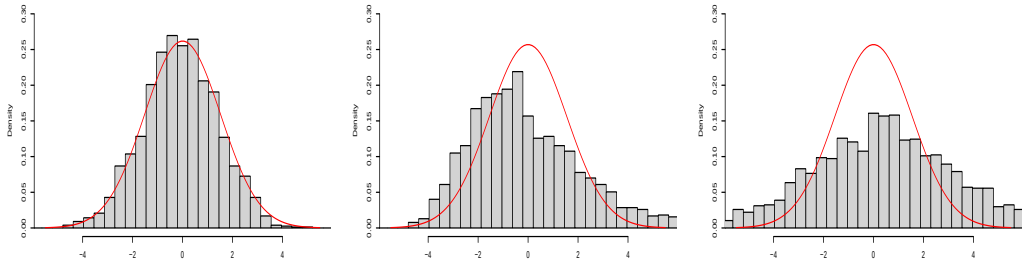
$$\frac{\mathbb{E}_{\theta_0}(k^2_{\beta_0}(J_1))}{\mathbb{E}_{\theta_0}(f^2_{\beta_0}(J_1))\mathbb{E}_{\theta_0}(k^2_{\beta_0}(J_1)) - (\mathbb{E}_{\theta_0}(k_{\beta_0}(J_1)f_{\beta_0}(J_1)))^2} \tag{4.3}$$

for the estimations of σ_0 and β_0 (see e.g. the calculations in (Brouste and Masuda, 2018, (9) and (10))).
Now, we present in details our simulation results.

For the O-U model (4.1), we simulated 2000 Monte-Carlo samples of $n = 2^{10}$ random variables $\Delta_j X = X_{t^n_{j+1}} - X_{t^n_j}$ with $\mu_0 = -0.7$, $\sigma_0 = 1$ and $\beta_0 = 1.3$, $p = 0.55$.

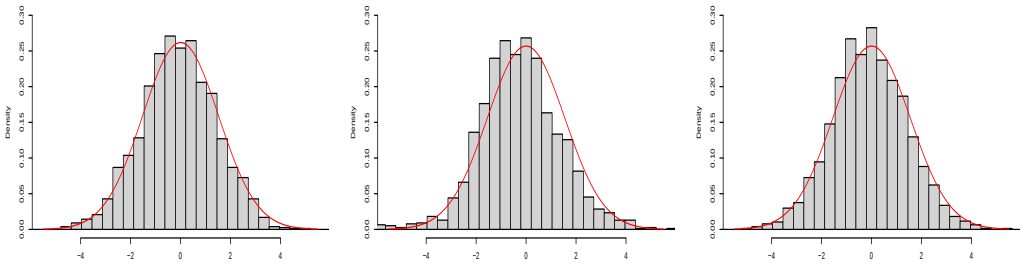
Remark 4.1. The O-U model is the simplest model that satisfies the conditions in the (M) case. Since the coefficient function a is constant, theoretically from Theorem 3.5, we do have the equivalence between the estimation based on $(X_{t^n_i})$ and the estimation based on $(\overline{X}_{t^n_i})$.

Remark 4.2. Here, for this O-U model, the computational time of MLE is 594459 secs while the one by one-step estimation is 55590 secs. This means that the estimation by one-step procedure is about 10 times faster than the maximum likelihood estimation, but gives similar approximation as shown in Figure 3 and 2.



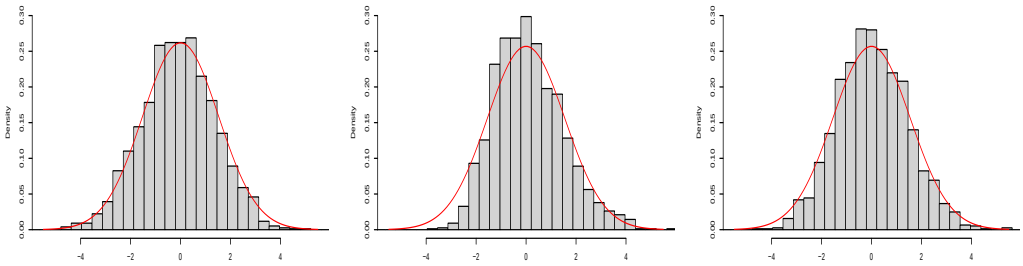
(a) $U^{\frac{1}{2}}_{\mu_0} n^{\frac{1}{\beta_0}-\frac{1}{2}}(\widehat{\mu}^0_n - \mu_0)$ (b) $\frac{\sqrt{n}}{\log n}(\widehat{\sigma}^0_n - \sigma_0)$ (c) $\sqrt{n}(\widehat{\beta}^0_n - \beta_0)$

Figure 1: Distributions of the re-scaled errors of moment estimation (ME) and comparison with the asymptotic normal distribution with efficient variance given by (4.3)



(a) $U_{\widehat{\mu}_0}^{\frac{1}{2}} n^{\frac{1}{\widehat{\beta}_0}-\frac{1}{2}} (\widehat{\mu}_n^1 - \mu_0)$ (b) $\frac{\sqrt{n}}{\log n}(\widehat{\sigma}_n^1 - \sigma_0)$ (c) $\sqrt{n}(\widehat{\beta}_n^1 - \beta_0)$

Figure 2: Distributions of the re-scaled errors of one-step estimation (OS) and comparison with their theoretical Gaussian limits



(a) $U_{\widehat{\mu}_0}^{\frac{1}{2}} n^{\frac{1}{\widehat{\beta}_0}-\frac{1}{2}} (\widehat{\mu}_n - \mu_0)$ (b) $\frac{\sqrt{n}}{\log n}(\widehat{\sigma}_n - \sigma_0)$ (c) $\sqrt{n}(\widehat{\beta}_n - \beta_0)$

Figure 3: Distributions of the re-scaled errors ofthe maximum likelihood estimation (MLE) and comparison with their theoretical Gaussian limits

For the SR model (4.2), we simulated 2000 Monte-Carlo samples of $n = 2^{10}$ random variables $\Delta_j X = X_{t_{j+1}^n} - X_{t_j^n}$ with $\mu_0 = -0.5$, $\sigma_0 = 0.5$ and $\beta_0 = 1.5$, $p = 0.7$. Our numerical results are summarized in the following table.

	MSE of $\hat{\mu}$	MSE of $\hat{\sigma}$	MSE of $\hat{\beta}$	Computational time
Initial estimation (ME)	2.36	6.80	6.34	20749 secs
One-step	2.36	2.97	2.53	87060 secs
MLE	2.34	2.71	2.39	487063 secs
Efficient variance	2.33	2.38	2.38	

Table 1. Comparison on the truncated re-normalized mean squared errors and on the computational times between the three methods of estimation ME, OS and MLE for the model (4.2).

Remark 4.3. The scale coefficient a of the SR model is not constant, then, it is clear that the Theorem 3.5 can not be applied in this case. The equivalence between the estimation based on $(X_{t_i^n})$ and the estimation based on $(\overline{X}_{t_i^n})$ remains as an open theoretical question. However, from the numerical point of view, it is shown in Table 1 that the asymptotic efficiency of the OS and MLE estimations using the experiment $\mathcal{E}^n$ is ensured. Once again, the one-step estimation appears to be fast and efficient.

4.2. Non-multiplicative case

Comparing to multiplicative case (M), the rate of estimation in the non-multiplicative case (NM) is faster for both σ_0 and β_0 by a factor of $\log n$. Here, the rate is $n^{1/\beta_0-1/2}$ for μ_0, $\sqrt{n}$ for σ_0 and $\sqrt{n}\log n$ for β_0. Similarly as the second model of case (M) above, first, we recall the theoretical results from the alternative Euler scheme, then, we do some numerical tests but using the observations on the true trajectory. To begin with, the asymptotic law of the estimation error is mixed Gaussian by Theorem 3.1. Therefore, we define re-scaled errors of estimation that have Gaussian laws. First of all, we define

$$\begin{aligned}
U_{\mu_0} &= \int_0^1 \frac{\partial_\mu b(X_s,\mu_0)^2}{a(X_s,\sigma_0)^2}ds \\
U_{\sigma_0} &= \int_0^1 \frac{\partial_\sigma a(X_s,\sigma_0)^2}{a(X_s,\sigma_0)^2}ds - \left(\int_0^1 \frac{\partial_\sigma a(X_s,\sigma_0)}{a(X_s,\sigma_0)}ds\right)^2 \\
U_{\beta_0} &= U_{\sigma_0}\left(\beta_0^4\int_0^1 \frac{\partial_\sigma a(X_s,\sigma_0)^2}{a(X_s,\sigma_0)^2}ds\right)^{-1}.
\end{aligned}$$

Then, from the stable convergence result of Theorem 3.1, we have

$$\begin{aligned}
U_{\mu_0}^{1/2} n^{1/\beta_0-1/2}(\widehat{\mu}_n-\mu_0) &\overset{stably}{\Longrightarrow} \mathcal{N}(0,(\mathbb{E}_{\theta_0} h_{\beta_0}^2(J_1))^{-1}) \\
U_{\sigma_0}^{1/2}\sqrt{n}(\widehat{\sigma}_n-\sigma_0) &\overset{stably}{\Longrightarrow} \mathcal{N}(0,(\mathbb{E}_{\theta_0} k_{\beta_0}^2(J_1))^{-1}) \\
U_{\beta_0}^{1/2}\sqrt{n}\log n(\widehat{\beta}_n-\beta_0) &\overset{stably}{\Longrightarrow} \mathcal{N}(0,(\mathbb{E}_{\theta_0} k_{\beta_0}^2(J_1))^{-1}).
\end{aligned}$$

For our one-step procedure, we take the moment estimators described in (3.5) as the initial estimators. In this case, the rates of convergence for $\widehat{\sigma}_n^0-\sigma_0$ and $\widehat{\beta}_n^0-\beta_0$ are worse than the ones in the formula of $\varphi_n(\theta_0)$ chosen. Despite of this obstacle, from Theorem 3.10, we still have that

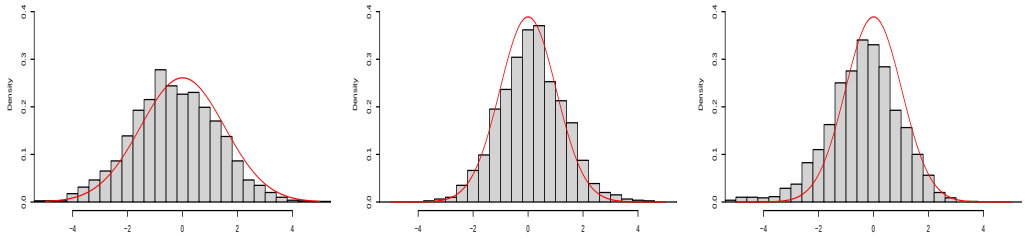
$$\begin{aligned}
U_{\mu_0}^{1/2} n^{1/\beta_0-1/2}(\widehat{\mu}_n^1-\mu_0) &\overset{stably}{\Longrightarrow} \mathcal{N}(0,(\mathbb{E}_{\theta_0} h_{\beta_0}^2(J_1))^{-1}) \\
U_{\sigma_0}^{1/2}\sqrt{n}(\widehat{\sigma}_n^1-\sigma_0) &\overset{stably}{\Longrightarrow} \mathcal{N}(0,(\mathbb{E}_{\theta_0} k_{\beta_0}^2(J_1))^{-1}) \\
U_{\beta_0}^{1/2}\sqrt{n}\log n(\widehat{\beta}_n^1-\beta_0) &\overset{stably}{\Longrightarrow} \mathcal{N}(0,(\mathbb{E}_{\theta_0} k_{\beta_0}^2(J_1))^{-1}).
\end{aligned}$$

We consider the following model

$$X_t = 1 + \int_0^t \mu X_s ds + \int_0^t \exp\left(\sigma\sin^2(X_s)\right) dJ_s, \quad t\in[0,1]. \tag{4.4}$$

For this model, we simulated 2000 Monte-Carlo samples of $n = 2^{10}$ random variables $\Delta_j X = X_{t_{j+1}^n} - X_{t_j^n}$ with $\mu_0 = -0.5$, $\sigma_0 = 1$ and $\beta_0 = 1.5$, $p = 0.7$. Our numerical results for one-step estimations are shown in the following histograms.

Remark 4.4. Here, for this model, the computational time of MLE is 207328 secs while the one by one-step estimation is 37063 secs. This means that the estimation by one-step procedure is about 5 times faster than the maximum likelihood estimation, but gives similar approximation as shown in Figure 3 and 2.



(a) $U_{\mu_0}^{\frac{1}{2}} n^{\frac{1}{\beta_0}-\frac{1}{2}} (\widehat{\mu}_n^1 - \mu_0)$ (b) $\frac{\sqrt{n}}{\beta_0^2} U_{\sigma_0}^{1/2} (\widehat{\sigma}_n^1 - \sigma_0)$ (c) $U_{\beta_0}^{1/2} \sqrt{n} \log n (\widehat{\beta}_n^1 - \beta_0)$

Figure 4: Histograms of the re-scaled errors of one-step estimation and comparison with their theoretical Gaussian limits

5. Proofs of Theorems 3.1 and 3.3

We start from the Taylor's expansion (2.4), we rewrite

$$\Delta_n(\theta_0) = -\varphi_n(\theta_0)^\top \sum_{i=0}^{n-1} \begin{pmatrix} g^1(\overline{X}_{t_i^n}, \overline{X}_{t_{i+1}^n}, \theta_0) \\ g^2(\overline{X}_{t_i^n}, \overline{X}_{t_{i+1}^n}, \theta_0) \\ g^3(\overline{X}_{t_i^n}, \overline{X}_{t_{i+1}^n}, \theta_0) \end{pmatrix} \text{ and } \mathcal{I}_n(\theta_0) = \sum_{i=0}^{n-1} \begin{pmatrix} \partial_\theta g^1(\overline{X}_{t_i^n}, \overline{X}_{t_{i+1}^n}, \theta_0)^\top \\ \partial_\theta g^2(\overline{X}_{t_i^n}, \overline{X}_{t_{i+1}^n}, \theta_0)^\top \\ \partial_\theta g^3(\overline{X}_{t_i^n}, \overline{X}_{t_{i+1}^n}, \theta_0)^\top \end{pmatrix}.$$

For any $\theta \in \Theta$, it can be addressed to (Clément and Gloter, 2020, formulas (2.6)-(2.8)) for the explicit expressions of $g^k(x, y, \theta)$ and (Clément and Gloter, 2020, formulas (3.6)-(3.9)) for all elements of $\partial_\theta g^k(x, y, \theta)$. Here, we recall that

$$g^1(x, y, \theta) = n^{1/\beta} \frac{\partial_\mu \xi_{1/n}^x(\mu)}{a(x, \sigma)} \frac{\partial_z \phi_\beta}{\phi_\beta}(z_n(x, y, \theta))$$

$$g^2(x, y, \theta) = \frac{\partial_\sigma a(x, \sigma)}{a(x, \sigma)} (1 + z_n(x, y, \theta) \frac{\partial_z \phi_\beta}{\phi_\beta}(z_n(x, y, \theta)))$$

$$g^3(x, y, \theta) = \frac{\log n}{\beta^2} (1 + z_n(x, y, \theta) \frac{\partial_z \phi_\beta}{\phi_\beta}(z_n(x, y, \theta)) - \frac{\partial_\beta \phi_\beta}{\phi_\beta}(z_n(x, y, \theta)).$$

By denoting

$$z_n^i(\theta) = z_n(\overline{X}_{t_i^n}, \overline{X}_{t_{i+1}^n}, \theta) \tag{5.1}$$

and the elements of $\mathcal{I}_n(\theta)$ can be expressed as follows

$$\mathcal{I}_n^{1,1}(\theta) = n^{1/\beta}\sum_{i=0}^{n-1}\frac{\partial_\mu^2\xi_{1/n}(\overline{X}_{t_i^n},\mu)}{a(\overline{X}_{t_i^n},\sigma)}h_\beta(z_n^i(\theta)) - n^{2/\beta}\sum_{i=0}^{n-1}\frac{(\partial_\mu\xi_{1/n}(\overline{X}_{t_i^n},\mu))^2}{a(\overline{X}_{t_i^n},\sigma)^2}\partial_z h_\beta(z_n^i(\theta))$$

$$\mathcal{I}_n^{2,2}(\theta) = \sum_{i=0}^{n-1}\left[\partial_\sigma\left(\frac{\partial_\sigma a(\overline{X}_{t_i^n},\sigma)}{a(\overline{X}_{t_i^n},\sigma)}\right)k_\beta(z_n^i(\theta)) - \left(\frac{\partial_\sigma a(\overline{X}_{t_i^n},\sigma)}{a(\overline{X}_{t_i^n},\sigma)}\right)^2 z_n^i(\theta)\partial_z k_\beta(z_n^i(\theta))\right]$$

$$\mathcal{I}_n^{3,3}(\theta) = -\sum_{i=0}^{n-1}\left[\partial_\beta f_\beta(z_n^i(\theta)) - 2\frac{\log n}{\beta^2}z_n^i(\theta)(\partial_\beta h_\beta)(z_n^i(\theta)) + 2\frac{\log n}{\beta^3}k_\beta(z_n^i(\theta)) + \frac{(\log n)^2}{\beta^4}z_n^i(\theta)\partial_z k_\beta(z_n^i(\theta))\right]$$

$$\mathcal{I}_n^{1,2}(\theta) = \mathcal{I}_n^{2,1}(\theta) = -n^{1/\beta}\sum_{i=0}^{n-1}\frac{\partial_\sigma a(\overline{X}_{t_i^n},\sigma)}{a(\overline{X}_{t_i^n},\sigma)^2}\partial_\mu\xi_{1/n}(\overline{X}_{t_i^n},\mu)\partial_z k_\beta(z_n^i(\theta))$$

$$\mathcal{I}_n^{1,3}(\theta) = \mathcal{I}_n^{3,1}(\theta) = n^{1/\beta}\sum_{i=0}^{n-1}\frac{\partial_\mu\xi_{1/n}(\overline{X}_{t_i^n},\mu)}{a(\overline{X}_{t_i^n},\sigma)}\left[-\frac{\log n}{\beta^2}\partial_z k_\beta(z_n^i(\theta)) + \partial_z f_\beta(z_n^i(\theta))\right]$$

$$\mathcal{I}_n^{2,3}(\theta) = \mathcal{I}_n^{3,2}(\theta) = \sum_{i=0}^{n-1}\frac{\partial_\sigma a(\overline{X}_{t_i^n},\sigma)}{a(\overline{X}_{t_i^n},\sigma)}\left[-\frac{\log n}{\beta^2}z_n^i(\theta)\partial_z k_\beta(z_n^i(\theta)) + z_n^i(\theta)(\partial_\beta h_\beta)(z_n^i(\theta))\right].$$

Let us first remark that in both cases, the multiplicative (M) and non-multiplicative (NM) ones, the convergence of the score function and of the information matrix has been established in (Clément and Gloter, 2020, Theorems 3.1 and 3.2) using observations $(X_{t_i^n})$. For us, since we use instead the observations $(\overline{X}_{t_i^n})$, similar results can be obtained by similar analyses. It just remains to prove from equality (2.4) that the term

$$((\varphi_n(\theta_0)u)^\top\cdot\partial_\theta\mathcal{I}_n(\theta_0+\varepsilon_n\varphi_n(\theta_0)u)\cdot(\varphi_n(\theta_0)u))^\top(\varphi_n(\theta_0)u)$$

tends to zero in probability. To do so, we denote $\mathcal{M}_n(\theta) = \partial_\theta\mathcal{I}_n(\theta)$, for any $\theta\in\Theta$, we have the explicit elements of the matrix $\mathcal{M}_n(\theta)$ as follows

$$\mathcal{M}_n^{1,1,1}(\theta) = n^{1/\beta}\sum_{i=0}^{n-1}\left(\frac{\partial_\mu^3\xi_{1/n}(\overline{X}_{t_i^n},\mu)}{a(\overline{X}_{t_i^n},\sigma)}h_\beta(z_n^i(\theta)) + \frac{\partial_\mu^2\xi_{1/n}(\overline{X}_{t_i^n},\mu)}{a(\overline{X}_{t_i^n},\sigma)}\partial_z h_\beta(z_n^i(\theta))\partial_\mu z_n^i(\theta)\right)$$

$$- n^{2/\beta}\sum_{i=0}^{n-1}\left(\frac{2(\partial_\mu\xi_{1/n}(\overline{X}_{t_i^n},\mu))\partial_\mu^2\xi_{1/n}(\overline{X}_{t_i^n},\mu)}{a(\overline{X}_{t_i^n},\sigma)^2}\partial_z h_\beta(z_n^i(\theta)) + \frac{(\partial_\mu\xi_{1/n}(\overline{X}_{t_i^n},\mu))^2}{a(\overline{X}_{t_i^n},\sigma)^2}\partial_z^2 h_\beta(z_n^i(\theta))\partial_\mu z_n^i(\theta)\right)$$

$$\mathcal{M}_n^{2,2,2}(\theta) = \sum_{i=0}^{n-1}\left[\partial_\sigma^2\left(\frac{\partial_\sigma a(\overline{X}_{t_i^n},\sigma)}{a(\overline{X}_{t_i^n},\sigma)}\right)k_\beta(z_n^i(\theta)) + \partial_\sigma\left(\frac{\partial_\sigma a(\overline{X}_{t_i^n},\sigma)}{a(\overline{X}_{t_i^n},\sigma)}\right)\partial_z k_\beta(z_n^i(\theta))\partial_\sigma z_n^i(\theta)\right.$$

$$\left.-2\left(\frac{\partial_\sigma a(\overline{X}_{t_i^n},\sigma)}{a(\overline{X}_{t_i^n},\sigma)}\right)\partial_\sigma\left(\frac{\partial_\sigma a(\overline{X}_{t_i^n},\sigma)}{a(\overline{X}_{t_i^n},\sigma)}\right)q_\beta(z_n^i(\theta)) - \left(\frac{\partial_\sigma a(\overline{X}_{t_i^n},\sigma)}{a(\overline{X}_{t_i^n},\sigma)}\right)^2\partial_z q_\beta(z_n^i(\theta))\partial_\sigma z_n^i(\theta)\right]$$

$$\mathcal{M}_n^{3,3,3}(\theta) = -\sum_{i=0}^{n-1}\left[(\partial_\beta^2 f_\beta)(z_n^i(\theta)) + 4\frac{\log n}{\beta^3}r_\beta(z_n^i(\theta)) - 2\frac{\log n}{\beta^2}(\partial_z r_\beta(z_n^i(\theta))\partial_\beta z_n^i(\theta) + (\partial_\beta r_\beta)(z_n^i(\theta)))\right.$$

$$-6\frac{\log n}{\beta^4}k_\beta(z_n^i(\theta)) + 2\frac{\log n}{\beta^3}(\partial_z k_\beta(z_n^i(\theta))\partial_\beta z_n^i(\theta) + (\partial_\beta k_\beta)(z_n^i(\theta))) - 4\frac{(\log n)^2}{\beta^5}q_\beta(z_n^i(\theta))$$

$$\left.+\frac{(\log n)^2}{\beta^4}(\partial_z q_\beta(z_n^i(\theta))\partial_\beta z_n^i(\theta) + (\partial_\beta q_\beta)(z_n^i(\theta)))\right]$$

$$
\begin{aligned}
\mathcal{M}_n^{1,2,1}(\theta) =& \mathcal{M}_n^{1,1,2}(\theta) = \mathcal{M}_n^{2,1,1}(\theta) \\
=& -n^{1/\beta}\sum_{i=0}^{n-1}\frac{\partial_\sigma a(\overline{X}_{t_i^n},\sigma)}{a(\overline{X}_{t_i^n},\sigma)^2}(\partial_\mu^2\xi_{1/n}(\overline{X}_{t_i^n},\mu)\partial_z k_\beta(z_n^i(\theta)) + \partial_\mu\xi_{1/n}(\overline{X}_{t_i^n},\mu)\partial_z^2 k_\beta(z_n^i(\theta))\partial_\mu z_n^i(\theta)) \\
\mathcal{M}_n^{2,1,2}(\theta) =& \mathcal{M}_n^{2,2,1}(\theta) = \mathcal{M}_n^{1,2,2}(\theta) \\
=& -n^{1/\beta}\sum_{i=0}^{n-1}\partial_\mu\xi_{1/n}(\overline{X}_{t_i^n},\mu)\left[\partial_\sigma\left(\frac{\partial_\sigma a(\overline{X}_{t_i^n},\sigma)}{a(\overline{X}_{t_i^n},\sigma)^2}\right)\partial_z k_\beta(z_n^i(\theta)) + \frac{\partial_\sigma a(\overline{X}_{t_i^n},\sigma)}{a(\overline{X}_{t_i^n},\sigma)^2}\partial_z^2 k_\beta(z_n^i(\theta))\partial_\sigma z_n^i(\theta)\right] \\
\mathcal{M}_n^{3,1,2}(\theta) =& \mathcal{M}_n^{1,2,3}(\theta) = \mathcal{M}_n^{2,3,1}(\theta) = \mathcal{M}_n^{3,2,1}(\theta) = \mathcal{M}_n^{2,1,3}(\theta) = \mathcal{M}_n^{1,3,2}(\theta) \\
=& -n^{1/\beta}\sum_{i=0}^{n-1}\frac{\partial_\sigma a(\overline{X}_{t_i^n},\sigma)}{a(\overline{X}_{t_i^n},\sigma)^2}\partial_\mu\xi_{1/n}(\overline{X}_{t_i^n},\mu)\left[-\frac{\log n}{\beta^2}\partial_z k_\beta(z_n^i(\theta)) + \partial_z^2 k_\beta(z_n^i(\theta))\partial_\beta z_n^i(\theta) + (\partial_\beta\partial_z k_\beta)(z_n^i(\theta))\right] \\
\mathcal{M}_n^{1,3,1}(\theta) =& \mathcal{M}_n^{1,1,3}(\theta) = \mathcal{M}_n^{3,1,1}(\theta) = n^{1/\beta}\sum_{i=0}^{n-1}\frac{\partial_\mu^2\xi_{1/n}(\overline{X}_{t_i^n},\mu)}{a(\overline{X}_{t_i^n},\sigma)}\left[-\frac{\log n}{\beta^2}\partial_z k_\beta(z_n^i(\theta)) + \partial_z f_\beta(z_n^i(\theta))\right] \\
&+ n^{1/\beta}\sum_{i=0}^{n-1}\frac{\partial_\mu\xi_{1/n}(\overline{X}_{t_i^n},\mu)}{a(\overline{X}_{t_i^n},\sigma)}\left[-\frac{\log n}{\beta^2}\partial_z^2 k_\beta(z_n^i(\theta))\partial_\mu z_n^i(\theta) + \partial_z^2 f_\beta(z_n^i(\theta))\partial_\mu z_n^i(\theta)\right] \\
\mathcal{M}_n^{3,1,3}(\theta) =& \mathcal{M}_n^{3,3,1}(\theta) = \mathcal{M}_n^{1,3,3}(\theta) = -\frac{\log n}{\beta^2}n^{1/\beta}\sum_{i=0}^{n-1}\frac{\partial_\mu\xi_{1/n}(\overline{X}_{t_i^n},\mu)}{a(\overline{X}_{t_i^n},\sigma)}\left[-\frac{\log n}{\beta^2}\partial_z k_\beta(z_n^i(\theta)) + \partial_z f_\beta(z_n^i(\theta))\right] \\
&+ n^{1/\beta}\sum_{i=0}^{n-1}\frac{\partial_\mu\xi_{1/n}(\overline{X}_{t_i^n},\mu)}{a(\overline{X}_{t_i^n},\sigma)}\left[2\frac{\log n}{\beta^3}\partial_z k_\beta(z_n^i(\theta)) - \frac{\log n}{\beta^2}(\partial_z^2 k_\beta(z_n^i(\theta))\partial_\beta z_n^i(\theta) + (\partial_\beta\partial_z k_\beta)(z_n^i(\theta)))\right. \\
&\left.+\partial_z^2 f_\beta(z_n^i(\theta))\partial_\beta z_n^i(\theta) + (\partial_\beta\partial_z f_\beta)(z_n^i(\theta))\right] \\
\mathcal{M}_n^{2,3,2}(\theta) =& \mathcal{M}_n^{2,2,3}(\theta) = \mathcal{M}_n^{3,2,2}(\theta) = \sum_{i=0}^{n-1}\frac{\partial_\sigma a(\overline{X}_{t_i^n},\sigma)}{a(\overline{X}_{t_i^n},\sigma)}\left[-\frac{\log n}{\beta^2}\partial_z q_\beta(z_n^i(\theta))\partial_\sigma z_n^i(\theta) + \partial_z r_\beta(z_n^i(\theta))\partial_\sigma z_n^i(\theta)\right] \\
&+ \sum_{i=0}^{n-1}\partial_\sigma\left(\frac{\partial_\sigma a(\overline{X}_{t_i^n},\sigma)}{a(\overline{X}_{t_i^n},\sigma)}\right)\left[-\frac{\log n}{\beta^2}q_\beta(z_n^i(\theta)) + r_\beta(z_n^i(\theta))\right] \\
\mathcal{M}_n^{3,2,3}(\theta) =& \mathcal{M}_n^{3,3,2}(\theta) = \mathcal{M}_n^{2,3,3}(\theta) \\
=& \sum_{i=0}^{n-1}\frac{\partial_\sigma a(\overline{X}_{t_i^n},\sigma)}{a(\overline{X}_{t_i^n},\sigma)}\left[2\frac{\log n}{\beta^3}q_\beta(z_n^i(\theta)) - \frac{\log n}{\beta^2}(\partial_z q_\beta(z_n^i(\theta))\partial_\beta z_n^i(\theta) + (\partial_\beta q_\beta)(z_n^i(\theta)))\right. \\
&\left.+\partial_z r_\beta(z_n^i(\theta))\partial_\beta z_n^i(\theta) + (\partial_\beta r_\beta)(z_n^i(\theta))\right].
\end{aligned}
$$

The next Proposition, whose proof is given in the supplementary document, is widely used.

Proposition 5.1. *Under the Assumption (A), we have*

$$
\begin{cases}
\sup_{\mu\in V_{\mu_0}}|\partial_\mu\xi_{1/n}(x,\mu) - \frac{1}{n}\partial_\mu b(x,\mu)| \le C(1+|x|^p)/n^2, \\
\sup_{\mu\in V_{\mu_0}}|\partial_\mu^2\xi_{1/n}(x,\mu) - \frac{1}{n}\partial_\mu^2 b(x,\mu)| \le C(1+|x|^p)/n^2, \quad \textit{for some } p>0. \\
\sup_{\mu\in V_{\mu_0}}|\partial_\mu^3\xi_{1/n}(x,\mu) - \frac{1}{n}\partial_\mu^3 b(x,\mu)| \le C(1+|x|^p)/n^2,
\end{cases} \tag{5.2}
$$

The next Theorem, which is a modified version of (Clément and Gloter, 2020, Theorem 4.2), plays an important role in the sequel. It is also proved in the accompanying document of this paper.

Theorem 5.2. *Under the Assumption (A), let f be a continuous function such that*

$$|f(x,\mu,\sigma)-f(y,\mu_0,\sigma_0)|\le C(1+|x|^p+|y|^p)(|\mu-\mu_0|+|\sigma-\sigma_0|+|x-y|),$$

and let $(z,\beta)\mapsto g_\beta(z)$ be a C^1 function (with respect to (z,β)) such that $\partial_z g_\beta$ is bounded (uniformly in β on a compact subset of $(0,2)$) and

$$|g_\beta(z)|+|\partial_\beta g_\beta(z)|\le C(1+(\log(1+|z|))^p),\quad p>0.$$

Then, for $\varepsilon<\left(\frac{1}{\beta_0}-\frac{1}{2}\right)\wedge\frac{1}{2}$ and $\eta>0$, we have the convergence in probability under $\mathbb{P}_{\theta_0}$:

$$\sup_{\theta\in V_n^{(\eta)}(\theta_0)} n^{\varepsilon}\left|\frac{1}{n}\sum_{i=0}^{n-1} f(\overline{X}_{t_i^n},\mu,\sigma)g_\beta(z_n^i(\theta))-\int_0^1 f(X_s,\mu_0,\sigma_0)ds\,\mathbb{E}_{\theta_0}(g_{\beta_0}(J_1))\right|\underset{n\to\infty}{\longrightarrow}0, \tag{5.3}$$

where $V_n^{(\eta)}(\theta_0)=\{\theta:\|diag(\frac{n^{1/\beta_0-1/2}}{(\log n)^2},\frac{n^{1/2}}{\log n},\frac{n^{1/2}}{\log n})(\theta-\theta_0)\|\le\eta\}$ and $z_n^i(\theta)$ is defined by (5.1). Moreover, if $\mathbb{E}_{\theta_0}(g_{\beta_0}(J_1))=0$, the following convergence in probability under $\mathbb{P}_{\theta_0}$ holds:

$$\sup_{\theta\in V_n^{(\eta)}(\theta_0)}\left|n^{\varepsilon-1/\beta_0}\sum_{i=0}^{n-1} f(\overline{X}_{t_i^n},\mu,\sigma)g_\beta(z_n^i(\theta))\right|\underset{n\to\infty}{\longrightarrow}0 \tag{5.4}$$

Proof of Theorem 3.1. First of all, similarly as in (Clément and Gloter, 2020, Remark 3.3), one can easily follow the proof of Theorem 2.10 in Masuda (2015) and Theorems 1 and 2 in Sweeting (1980) to prove the last assertion.
Now, to obtain LAMN property, we only need to prove the convergence to zero in probability of the last term in the Taylor's expansion (2.4). To do so, for $\boxed{\hat\theta_n=\theta_0+\varepsilon_n\varphi_n(\theta_0)u}$, we use (Clément and Gloter, 2020, Proof of Theorem 4.2) for the convergence to zero of the elements of the following matrices

$$\mathcal{H}_n^1(\hat\theta_n)=n^{1/2-1/\beta_0}\varphi_n^\top(\theta_0)\partial_\mu\mathcal{I}_n(\hat\theta_n)\varphi_n(\theta_0)=\begin{pmatrix}\frac{\mathcal{M}_n^{1,1,1}(\hat\theta_n)}{n^{3/\beta_0-3/2}} & \frac{\mathcal{M}_n^{1,1,2}(\hat\theta_n)}{n^{2/\beta_0-1/2}} & \frac{\mathcal{M}_n^{1,1,3}(\hat\theta_n)}{n^{2/\beta_0-1/2}\log n}\\ \frac{\mathcal{M}_n^{1,1,2}(\hat\theta_n)}{n^{2/\beta_0-1/2}} & \frac{\mathcal{M}_n^{1,2,2}(\hat\theta_n)}{n^{1/\beta_0+1/2}} & \frac{\mathcal{M}_n^{1,2,3}(\hat\theta_n)}{n^{1/\beta_0+1/2}\log n}\\ \frac{\mathcal{M}_n^{1,1,3}(\hat\theta_n)}{n^{2/\beta_0-1/2}\log n} & \frac{\mathcal{M}_n^{1,2,3}(\hat\theta_n)}{n^{1/\beta_0+1/2}\log n} & \frac{\mathcal{M}_n^{1,3,3}(\hat\theta_n)}{n^{1/\beta_0+1/2}(\log n)^2}\end{pmatrix},$$

$$\mathcal{H}_n^2(\hat\theta_n)=n^{-1/2}\varphi_n^\top(\theta_0)\partial_\sigma\mathcal{I}_n(\hat\theta_n)\varphi_n(\theta_0)=\begin{pmatrix}\frac{\mathcal{M}_n^{2,1,1}(\hat\theta_n)}{n^{2/\beta_0-1/2}} & \frac{\mathcal{M}_n^{2,1,2}(\hat\theta_n)}{n^{1/\beta_0+1/2}} & \frac{\mathcal{M}_n^{2,1,3}(\hat\theta_n)}{n^{1/\beta_0+1/2}\log n}\\ \frac{\mathcal{M}_n^{2,1,2}(\hat\theta_n)}{n^{1/\beta_0+1/2}} & \frac{\mathcal{M}_n^{2,2,2}(\hat\theta_n)}{n^{3/2}} & \frac{\mathcal{M}_n^{2,2,3}(\hat\theta_n)}{n^{3/2}\log n}\\ \frac{\mathcal{M}_n^{2,1,3}(\hat\theta_n)}{n^{1/\beta_0+1/2}\log n} & \frac{\mathcal{M}_n^{2,2,3}(\hat\theta_n)}{n^{3/2}\log n} & \frac{\mathcal{M}_n^{2,3,3}(\hat\theta_n)}{n^{3/2}(\log n)^2}\end{pmatrix},$$

$$\mathcal{H}_n^3(\hat\theta_n)=\frac{1}{\sqrt{n}\log n}\varphi_n^\top(\theta_0)\partial_\beta\mathcal{I}_n(\hat\theta_n)\varphi_n(\theta_0)=\begin{pmatrix}\frac{\mathcal{M}_n^{3,1,1}(\hat\theta_n)}{n^{2/\beta_0-1/2}\log n} & \frac{\mathcal{M}_n^{3,1,2}(\hat\theta_n)}{n^{1/\beta_0+1/2}\log n} & \frac{\mathcal{M}_n^{3,1,3}(\hat\theta_n)}{n^{1/\beta_0+1/2}(\log n)^2}\\ \frac{\mathcal{M}_n^{3,1,2}(\hat\theta_n)}{n^{1/\beta_0+1/2}\log n} & \frac{\mathcal{M}_n^{3,2,2}(\hat\theta_n)}{n^{3/2}\log n} & \frac{\mathcal{M}_n^{3,2,3}(\hat\theta_n)}{n^{3/2}(\log n)^2}\\ \frac{\mathcal{M}_n^{3,1,3}(\hat\theta_n)}{n^{1/\beta_0+1/2}(\log n)^2} & \frac{\mathcal{M}_n^{3,2,3}(\hat\theta_n)}{n^{3/2}(\log n)^2} & \frac{\mathcal{M}_n^{3,3,3}(\hat\theta_n)}{n^{3/2}(\log n)^3}\end{pmatrix}.$$

First, it is followed from the series expansion of the density (see e.g. (Sato, 2000, Remark 14.18) or Masuda (2015)) that for any non-negative integers k and k',

$$|\partial_z^k \partial_\beta^{k'} \phi_\beta(z)| \sim C_{k,k',\beta} (\log|z|)^{k'} |z|^{-\beta-1-k}, \quad |z| \to \infty, \tag{5.5}$$

for some positive constant $C_{k,k',\beta}$. From this, since $\partial_z \phi_\beta(z) = h_\beta(z)\phi_\beta(z)$, it is easy to get thanks to the Leibniz formula:

$$|\partial_z^k \partial_\beta^{k'} h_\beta(z)| \sim C'_{k,k',\beta} (\log|z|)^{k'} |z|^{-k-1}, \quad |z| \to \infty, \tag{5.6}$$

for some positive constant $C'_{k,k',\beta}$.
From (Clément and Gloter, 2020, Proof of Theorem 3.1), this equivalence (5.5) permits to deduce that h_β, $\partial_z h_\beta$, k_β, $\partial_z k_\beta$, $\partial_z f_\beta (= \partial_\beta h_\beta)$ and q_β are bounded, and $|r_\beta(z)| = |(\partial_\beta k_\beta)(z)| = |z(\partial_\beta h_\beta)(z)| \le C\log|z|$, $|f_\beta(z)| \le C\log|z|$ and that $|(\partial_\beta f_\beta)(z)| \le C(\log|z|)^2$. We also have from the symmetry of ϕ_β and the integration by part formula, $\mathbb{E}_\theta(h_\beta(J_1)) = 0$.

Similarly, it is easy to see from (5.6) that $\partial_\beta \partial_z h_\beta$, $z\partial_\beta \partial_z h_\beta$, $z^2 \partial_\beta \partial_z h_\beta$, $\partial_z^2 h_\beta$, $\partial_\beta \partial_z^2 h_\beta$ $z\partial_\beta \partial_z^2 h_\beta$, $z^2 \partial_\beta \partial_z^2 h_\beta$, $\partial_z^3 h_\beta$, $z\partial_z h_\beta$, $z\partial_z^2 h_\beta$, $z^2 \partial_z^2 h_\beta$, $z\partial_z^3 h_\beta$ and $z^2 \partial_z^3 h_\beta$ are also bounded. This leads to

$$\begin{aligned}
\partial_z^2 k_\beta(z) &= 2\partial_z h_\beta(z) + z\partial_z^2 h_\beta(z), \quad z\partial_z^2 k_\beta, \quad z^2 \partial_z^2 k_\beta, \\
\partial_z r_\beta(z) &= (\partial_\beta \partial_z k_\beta)(z) = (\partial_\beta h_\beta)(z) + z(\partial_\beta \partial_z h_\beta)(z), \quad z\partial_z r_\beta, \\
\partial_z^2 r_\beta(z) &= 2(\partial_z \partial_\beta h_\beta)(z) + z(\partial_\beta \partial_z^2 h_\beta)(z), \quad \partial_z(z\partial_z r_\beta)(z) = z(\partial_z^2 r_\beta)(z) + (\partial_z r_\beta)(z), \\
(\partial_\beta q_\beta)(z) &= z(\partial_\beta \partial_z k_\beta)(z) = z(\partial_z r_\beta)(z), \quad (\partial_z q_\beta)(z) = z(\partial_z^2 k_\beta)(z) + (\partial_z k_\beta)(z), \quad z\partial_z q_\beta, \\
(\partial_z^2 q_\beta)(z) &= z(\partial_z^3 k_\beta)(z) + 2(\partial_z^2 k_\beta)(z), \quad z\partial_z^2 q_\beta, \\
\partial_z(z\partial_z q_\beta)(z) &= z(\partial_z^2 q_\beta)(z) + (\partial_z q_\beta)(z), \quad (\partial_z \partial_\beta q_\beta)(z) = z(\partial_\beta \partial_z^2 k_\beta)(z) + (\partial_\beta \partial_z k_\beta)(z), \\
(\partial_\beta \partial_z^2 k_\beta)(z) &= 2(\partial_\beta \partial_z h_\beta)(z) + z(\partial_\beta \partial_z^2 h_\beta)(z), \quad z\partial_\beta \partial_z^2 k_\beta, \\
\partial_z^3 k_\beta(z) &= 3\partial_z h_\beta(z) + z\partial_z^3 h_\beta(z), \quad z\partial_z^3 k_\beta(z), \quad \partial_z(z\partial_z^2 k_\beta)(z) = z\partial_z^3 k_\beta(z) + \partial_z^2 k_\beta(z)
\end{aligned}$$

are bounded. Moreover, on the one hand, we also have $\partial_\beta^2 \partial_z k_\beta$, $\partial_z \partial_\beta r_\beta$, $\partial_z^2 f_\beta$, $z\partial_z^2 f_\beta$, $\partial_\beta(z\partial_z^2 f_\beta)$, $\partial_z^3 f_\beta$, $\partial_\beta \partial_z f_\beta$, $\partial_\beta^2 \partial_z f_\beta$, $\partial_\beta \partial_z^2 f_\beta$ and $\partial_z(z\partial_z^2 f_\beta)$ are bounded. On the other hand, we have $\partial_\beta(z\partial_z q_\beta)$, $\partial_\beta(z\partial_z r_\beta)$, $\partial_\beta^2 q_\beta$, $\partial_\beta r_\beta$, $\partial_\beta^2 r_\beta$, $\partial_\beta^2 f_\beta$ and $\partial_\beta^3 f_\beta$ bounded by logarithm. All these analyses make sure that we are in good conditions for applying the Theorem 5.2 in the following.

Concerning the elements of the matrix $\mathcal{H}_n^1(\hat\theta_n)$: Applying the formulas in (2.5) for the $\mathcal{M}_n^{1,1,1}(\hat\theta_n)$, by some basic calculations, we easily have the following inequality

$$\begin{aligned}
&\frac{|\mathcal{M}_n^{1,1,1}(\hat\theta_n)|}{n^{3/\beta_0-3/2}} \\
&\le n^{3/2-3/\beta_0+1/\hat\beta_n} \left(\left| \sum_{i=0}^{n-1} \frac{\frac{1}{n}\partial_\mu^3 b(\overline{X}_{t_i^n}, \hat\mu_n)}{a(\overline{X}_{t_i^n}, \hat\sigma_n)} h_{\hat\beta_n}(z_n^i(\hat\theta_n)) \right| + \sum_{i=0}^{n-1} \left| \frac{\partial_\mu^3 \xi_{1/n}(\overline{X}_{t_i^n}, \hat\mu_n) - \frac{1}{n}\partial_\mu^3 b(\overline{X}_{t_i^n}, \hat\mu_n)}{a(\overline{X}_{t_i^n}, \hat\sigma_n)} h_{\hat\beta_n}(z_n^i(\hat\theta_n)) \right| \right) \\
&+ n^{3/2-1/\beta_0+2(1/\hat\beta_n-1/\beta_0)} \sum_{i=0}^{n-1} \left| \frac{(\partial_\mu^2 \xi_{1/n}(\overline{X}_{t_i^n}, \hat\mu_n) - \frac{1}{n}\partial_\mu^2 b(\overline{X}_{t_i^n}, \hat\mu_n))\partial_\mu \xi_{1/n}(\overline{X}_{t_i^n}, \hat\mu_n)}{a(\overline{X}_{t_i^n}, \hat\sigma_n)^2} \partial_z h_{\hat\beta_n}(z_n^i(\hat\theta_n)) \right. \\
&\left. + \left(\frac{\frac{1}{n}\partial_\mu^2 b(\overline{X}_{t_i^n}, \hat\mu_n)(\partial_\mu \xi_{1/n}(\overline{X}_{t_i^n}, \hat\mu_n) - \frac{1}{n}\partial_\mu b(\overline{X}_{t_i^n}, \hat\mu_n))}{a(\overline{X}_{t_i^n}, \hat\sigma_n)^2} + \frac{\frac{1}{n^2}\partial_\mu^2 b(\overline{X}_{t_i^n}, \hat\mu_n)\partial_\mu b(\overline{X}_{t_i^n}, \hat\mu_n)}{a(\overline{X}_{t_i^n}, \hat\sigma_n)^2} \right) \partial_z h_{\hat\beta_n}(z_n^i(\hat\theta_n)) \right|
\end{aligned}$$

$$+n^{3/2-1/\beta_0+2(1/\hat\beta_n-1/\beta_0)}\sum_{i=0}^{n-1}\left|\frac{2(\partial_\mu\xi_{1/n}(\overline{X}_{t_i^n},\hat\mu_n)-\frac{1}{n}\partial_\mu b(\overline{X}_{t_i^n},\hat\mu_n))\partial_\mu^2\xi_{1/n}(\overline{X}_{t_i^n},\hat\mu_n)}{a(\overline{X}_{t_i^n},\hat\sigma_n)^2}\partial_z h_{\hat\beta_n}(z_n^i(\hat\theta_n))\right.$$
$$\left.+\left(\frac{\frac{2}{n}\partial_\mu b(\overline{X}_{t_i^n},\hat\mu_n)(\partial_\mu^2\xi_{1/n}(\overline{X}_{t_i^n},\hat\mu_n)-\frac{1}{n}\partial_\mu^2 b(\overline{X}_{t_i^n},\hat\mu_n))}{a(\overline{X}_{t_i^n},\hat\sigma_n)^2}+\frac{\frac{2}{n^2}\partial_\mu b(\overline{X}_{t_i^n},\hat\mu_n)\partial_\mu^2 b(\overline{X}_{t_i^n},\hat\mu_n)}{a(\overline{X}_{t_i^n},\hat\sigma_n)^2}\right)\partial_z h_{\hat\beta_n}(z_n^i(\hat\theta_n))\right|$$
$$+n^{3/2+3(1/\hat\beta_n-1/\beta_0)}\sum_{i=0}^{n-1}\left|\frac{(\partial_\mu\xi_{1/n}(\overline{X}_{t_i^n},\hat\mu_n)-\frac{1}{n}\partial_\mu b(\overline{X}_{t_i^n},\hat\mu_n)+\frac{1}{n}\partial_\mu b(\overline{X}_{t_i^n},\hat\mu_n))^3}{a(\overline{X}_{t_i^n},\hat\sigma_n)^3}\partial_z^2 h_{\hat\beta_n}(z_n^i(\hat\theta_n))\right|.$$

Then, thanks to the inequalities in (5.2) and Assumption (A), the first sum in the upper bound of $\frac{|\mathcal{M}_n^{1,1,1}(\hat\theta_n)|}{n^{3/\beta_0-3/2}}$ above verifies the hypotheses of Theorem 5.2 and the other sums are bounded by the convenient terms that enable us to use the result from Theorem 5.2. More precisely, we have

$$\begin{aligned}\frac{|\mathcal{M}_n^{1,1,1}(\hat\theta_n)|}{n^{3/\beta_0-3/2}}&\le n^{1/2-1/\beta_0+1/\hat\beta_n-1/\beta_0}\frac{1}{n^{1/\beta_0}}\left\{\left|\sum_{i=0}^{n-1}\frac{\partial_\mu^3 b(\overline{X}_{t_i^n},\hat\mu)}{a(\overline{X}_{t_i^n},\hat\sigma_n)}h_{\hat\beta_n}(z_n^i(\hat\theta_n))\right|+\sum_{i=0}^{n-1}\frac{C}{n}(1+|\overline{X}_{t_i^n}|^p)|h_{\hat\beta_n}(z_n^i(\hat\theta_n))|\right\}\\&+n^{1/2-1/\beta_0+2(1/\hat\beta_n-1/\beta_0)}\frac{1}{n}\sum_{i=0}^{n-1}C(1+|\overline{X}_{t_i^n}|^p)(1+1/n)|\partial_z h_{\hat\beta_n}(z_n^i(\hat\theta_n))|\\&+n^{-1/2+3(1/\hat\beta_n-1/\beta_0)}\frac{1}{n}\sum_{i=0}^{n-1}C(1+|\overline{X}_{t_i^n}|^p)(1+1/n)|\partial_z^2 h_{\hat\beta_n}(z_n^i(\hat\theta_n))|.\end{aligned}\tag{5.7}$$

Now, using the fact that $(\hat\theta_n)_n$ converges to θ_0 and applying Theorem 5.2, the convergence to zero in probability of the first sum and the remaining sums in the r.h.s. of the above inequality (5.7) are obtained respectively by (5.4) since $\mathbb{E}_{\theta_0}(h_{\beta_0}(J_1))=0$ and by (5.3) as n tends to infinity. Similarly, the convergences to zero of other elements of $\mathcal{H}_n^1(\hat\theta_n)$ since we have that

$$\begin{aligned}&\frac{|\mathcal{M}_n^{1,1,2}(\hat\theta_n)|}{n^{2/\beta_0-1/2}}\\&\le n^{1/2-2/\beta_0+1/\hat\beta_n}\sum_{i=0}^{n-1}\left|\frac{\partial_\sigma a(\overline{X}_{t_i^n},\hat\sigma_n)}{a(\overline{X}_{t_i^n},\hat\sigma_n)^2}\right|\left|(\partial_\mu^2\xi_{1/n}(\overline{X}_{t_i^n},\hat\mu_n)-\frac{1}{n}\partial_\mu^2 b(\overline{X}_{t_i^n},\hat\mu_n))+\frac{1}{n}\partial_\mu^2 b(\overline{X}_{t_i^n},\hat\mu_n)\right||\partial_z k_{\hat\beta_n}(z_n^i(\hat\theta_n))|\\&+n^{1/2-2/\beta_0+2/\hat\beta_n}\sum_{i=0}^{n-1}\left|\frac{\partial_\sigma a(\overline{X}_{t_i^n},\hat\sigma_n)}{a(\overline{X}_{t_i^n},\hat\sigma_n)^3}\right|(\partial_\mu\xi_{1/n}(\overline{X}_{t_i^n},\hat\mu_n)-\frac{1}{n}\partial_\mu b(\overline{X}_{t_i^n},\hat\mu_n)+\frac{1}{n}\partial_\mu b(\overline{X}_{t_i^n},\hat\mu_n))^2|\partial_z^2 k_{\hat\beta_n}(z_n^i(\hat\theta_n))|\\&\le n^{-1/2-1/\beta_0+1/\hat\beta_n}\frac{1}{n^{1/\beta_0}}\sum_{i=0}^{n-1}C(1+|\overline{X}_{t_i^n}|^p)|\partial_z k_{\hat\beta_n}(z_n^i(\hat\theta_n))|\\&+n^{-1/2-2(1/\beta_0-1/\hat\beta_n)}\frac{1}{n}\sum_{i=0}^{n-1}C(1+|\overline{X}_{t_i^n}|^p)|\partial_z^2 k_{\hat\beta_n}(z_n^i(\hat\theta_n))|,\end{aligned}\tag{5.8}$$

$$\begin{aligned}\frac{|\mathcal{M}_n^{1,1,3}(\hat\theta_n)|}{n^{2/\beta_0-1/2}\log n}&\le\frac{n^{1/2-2/\beta_0+1/\hat\beta_n}}{\log n}\sum_{i=0}^{n-1}\frac{|\partial_\mu^2\xi_{1/n}(\overline{X}_{t_i^n},\hat\mu_n)-\frac{1}{n}\partial_\mu^2 b(\overline{X}_{t_i^n},\hat\mu_n))+\frac{1}{n}\partial_\mu^2 b(\overline{X}_{t_i^n},\hat\mu_n)|}{|a(\overline{X}_{t_i^n},\hat\sigma_n)|}\\&\times\left|-\frac{\log n}{\hat\beta_n^2}\partial_z k_{\hat\beta_n}(z_n^i(\hat\theta_n))+\partial_z f_{\hat\beta_n}(z_n^i(\hat\theta_n))\right|\\&+\frac{n^{1/2-2/\beta_0+2/\hat\beta_n}}{\log n}\sum_{i=0}^{n-1}\frac{(\partial_\mu\xi_{1/n}(\overline{X}_{t_i^n},\hat\mu_n)-\frac{1}{n}\partial_\mu b(\overline{X}_{t_i^n},\hat\mu_n)+\frac{1}{n}\partial_\mu b(\overline{X}_{t_i^n},\hat\mu_n))^2}{a(\overline{X}_{t_i^n},\hat\sigma_n)^2}\end{aligned}$$

$$\times\left|-\frac{\log n}{\hat\beta_n^2}\partial_z^2 k_{\hat\beta_n}(z_n^i(\hat\theta_n))+\partial_z^2 f_{\hat\beta_n}(z_n^i(\hat\theta_n))\right|$$

$$\le n^{-1/2-1/\beta_0+1/\hat\beta_n}\frac{1}{n^{1/\beta_0}}\sum_{i=0}^{n-1}C(1+|\overline{X}_{t_i^n}|^p)\left(\frac{1}{\hat\beta_n}|\partial_z k_{\hat\beta_n}(z_n^i(\hat\theta_n))|+\frac{1}{\log n}|\partial_z f_{\hat\beta_n}(z_n^i(\hat\theta_n))|\right)$$

$$+n^{-1/2-2(1/\beta_0-1/\hat\beta_n)}\frac{1}{n}\sum_{i=0}^{n-1}C(1+|\overline{X}_{t_i^n}|^p)\left(\frac{1}{\hat\beta_n^2}|\partial_z^2 k_{\hat\beta_n}(z_n^i(\hat\theta_n))|+\frac{1}{\log n}|\partial_z^2 f_{\hat\beta_n}(z_n^i(\hat\theta_n))|\right), \tag{5.9}$$

$$\frac{|\mathcal{M}_n^{1,2,2}(\hat\theta_n)|}{n^{1/\beta_0+1/2}}\le n^{1/\hat\beta_n-1/\beta_0-1/2}\sum_{i=0}^{n-1}|\partial_\mu\xi_{1/n}(\overline{X}_{t_i^n},\hat\mu_n)-\frac{1}{n}\partial_\mu b(\overline{X}_{t_i^n},\hat\mu_n)+\frac{1}{n}\partial_\mu b(\overline{X}_{t_i^n},\hat\mu_n)|$$

$$\times\left|\partial_\sigma\left(\frac{\partial_\sigma a(\overline{X}_{t_i^n},\hat\sigma_n)}{a(\overline{X}_{t_i^n},\hat\sigma_n)^2}\right)\partial_z k_{\hat\beta_n}(z_n^i(\hat\theta_n))+\frac{(\partial_\sigma a(\overline{X}_{t_i^n},\hat\sigma_n))^2}{a(\overline{X}_{t_i^n},\hat\sigma_n)^3}z_n^i(\hat\theta_n)\partial_z^2 k_{\hat\beta_n}(z_n^i(\hat\theta_n))\right|$$

$$\le n^{-1/2-1/\beta_0+1/\hat\beta_n}\frac{1}{n}\sum_{i=0}^{n-1}C(1+|\overline{X}_{t_i^n}|^p)\left(|\partial_z k_{\hat\beta_n}(z_n^i(\hat\theta_n))|+|\partial_z^2 k_{\hat\beta_n}(z_n^i(\hat\theta_n))|\right), \tag{5.10}$$

$$\frac{|\mathcal{M}_n^{1,2,3}(\hat\theta_n)|}{n^{1/\beta_0+1/2}\log n}\le\frac{n^{1/\hat\beta_n-1/\beta_0-1/2}}{\log n}\sum_{i=0}^{n-1}\left|\frac{\partial_\sigma a(\overline{X}_{t_i^n},\hat\sigma_n)}{a(\overline{X}_{t_i^n},\hat\sigma_n)^2}\right||\partial_\mu\xi_{1/n}(\overline{X}_{t_i^n},\hat\mu_n)-\frac{1}{n}\partial_\mu b(\overline{X}_{t_i^n},\hat\mu_n)+\frac{1}{n}\partial_\mu b(\overline{X}_{t_i^n},\hat\mu_n)|$$

$$\times\left|\frac{\log n}{\hat\beta_n^2}\partial_z k_{\hat\beta_n}(z_n^i(\hat\theta_n))+\frac{\log n}{\hat\beta_n^2}z_n^i(\hat\theta_n)\partial_z^2 k_{\hat\beta_n}(z_n^i(\hat\theta_n))-(\partial_\beta\partial_z k_{\hat\beta_n})(z_n^i(\hat\theta_n))\right|$$

$$\le\frac{n^{-1/2-1/\beta_0+1/\hat\beta_n}}{\log n}\frac{1}{n}\sum_{i=0}^{n-1}C(1+|\overline{X}_{t_i^n}|^p)\left(\frac{\log n}{\hat\beta_n^2}(|\partial_z k_{\hat\beta_n}(z_n^i(\hat\theta_n))|\right.$$

$$\left.+|z_n^i(\hat\theta_n)\partial_z^2 k_{\hat\beta_n}(z_n^i(\hat\theta_n))|)+|(\partial_\beta\partial_z k_{\hat\beta_n})(z_n^i(\hat\theta_n))|\right), \tag{5.11}$$

$$\frac{|\mathcal{M}_n^{1,3,3}(\hat\theta_n)|}{n^{1/\beta_0+1/2}(\log n)^2}\le\frac{n^{1/\hat\beta_n-1/\beta_0-1/2}}{\hat\beta_n^2\log n}\sum_{i=0}^{n-1}\frac{|\partial_\mu\xi_{1/n}(\overline{X}_{t_i^n},\hat\mu_n)-\frac{1}{n}\partial_\mu b(\overline{X}_{t_i^n},\hat\mu_n)+\frac{1}{n}\partial_\mu b(\overline{X}_{t_i^n},\hat\mu_n)|}{|a(\overline{X}_{t_i^n},\hat\sigma_n)|}$$

$$\times\left|-\frac{\log n}{\hat\beta_n^2}\partial_z k_{\hat\beta_n}(z_n^i(\hat\theta_n))+\partial_z f_{\hat\beta_n}(z_n^i(\hat\theta_n))\right|$$

$$+\frac{n^{1/\hat\beta_n-1/\beta_0-1/2}}{(\log n)^2}\sum_{i=0}^{n-1}\frac{|\partial_\mu\xi_{1/n}(\overline{X}_{t_i^n},\hat\mu_n)-\frac{1}{n}\partial_\mu b(\overline{X}_{t_i^n},\hat\mu_n)+\frac{1}{n}\partial_\mu b(\overline{X}_{t_i^n},\hat\mu_n|)}{|a(\overline{X}_{t_i^n},\hat\sigma_n)|}$$

$$\times\left|2\frac{\log n}{\hat\beta_n^3}\partial_z k_{\hat\beta_n}(z_n^i(\hat\theta_n))-\frac{\log n}{\hat\beta_n^2}((\partial_\beta\partial_z k_{\hat\beta_n})(z_n^i(\hat\theta_n))-\frac{\log n}{\hat\beta_n^2}z_n^i(\hat\theta_n)\partial_z^2 k_{\hat\beta_n}(z_n^i(\hat\theta_n)))\right.$$

$$\left.-\frac{\log n}{\hat\beta_n^2}z_n^i(\hat\theta_n)\partial_z^2 f_{\hat\beta_n}(z_n^i(\hat\theta_n))+(\partial_\beta\partial_z f_{\hat\beta_n})(z_n^i(\hat\theta_n))\right|$$

$$\le\frac{n^{1/\hat\beta_n-1/\beta_0-1/2}}{\hat\beta_n^2\log n}\frac{1}{n}\sum_{i=0}^{n-1}C(1+|\overline{X}_{t_i^n}|^p)\left[\frac{\log n}{\hat\beta_n^2}|\partial_z k_{\hat\beta_n}(z_n^i(\hat\theta_n))|+|\partial_z f_{\hat\beta_n}(z_n^i(\hat\theta_n))|\right]$$

$$+\frac{n^{1/\hat\beta_n-1/\beta_0-1/2}}{(\log n)^2}\frac{1}{n}\sum_{i=0}^{n-1}C(1+|\overline{X}_{t_i^n}|^p)\left[2\frac{\log n}{\hat\beta_n^3}|\partial_z k_{\hat\beta_n}(z_n^i(\hat\theta_n))|+\frac{\log n}{\hat\beta_n^2}(|(\partial_\beta\partial_z k_{\hat\beta_n})(z_n^i(\hat\theta_n))|\right.$$

$$\left.+\frac{\log n}{\hat\beta_n^2}|z_n^i(\hat\theta_n)\partial_z^2 k_{\hat\beta_n}(z_n^i(\hat\theta_n))|)+\frac{\log n}{\hat\beta_n^2}|z_n^i(\hat\theta_n)\partial_z^2 f_{\hat\beta_n}(z_n^i(\hat\theta_n))|+|(\partial_\beta\partial_z f_{\hat\beta_n})(z_n^i(\hat\theta_n))|\right]. \tag{5.12}$$

Concerning the elements of the matrix $\mathcal{H}_n^2(\theta)$: By symmetricity of this matrix, we have

$$\frac{\mathcal{M}_n^{2,1,1}(\hat{\theta}_n)}{n^{2/\beta_0-1/2}} = \frac{\mathcal{M}_n^{1,1,2}(\hat{\theta}_n)}{n^{2/\beta-1/2}}, \quad \frac{\mathcal{M}_n^{2,1,2}(\hat{\theta}_n)}{n^{1/\beta_0+1/2}} = \frac{\mathcal{M}_n^{1,2,2}(\hat{\theta}_n)}{n^{1/\beta_0+1/2}}, \quad \frac{\mathcal{M}_n^{2,1,3}(\hat{\theta}_n)}{n^{1/\beta_0+1/2}\log n} = \frac{\mathcal{M}_n^{1,2,3}(\hat{\theta}_n)}{n^{1/\beta_0+1/2}\log n},$$

then, the proof of their convergences are treated as above. For the rest elements, we rewrite as follows

$$\begin{aligned}
\frac{|\mathcal{M}_n^{2,2,3}(\hat{\theta}_n)|}{n^{3/2}\log n} &\le \frac{1}{n^{3/2}\log n}\sum_{i=0}^{n-1}\left(\frac{\partial_\sigma a(\overline{X}_{t_i^n},\hat{\sigma}_n)}{a(\overline{X}_{t_i^n},\hat{\sigma}_n)}\right)^2\left|\frac{\log n}{\hat{\beta}_n^2}z_n^i(\hat{\theta}_n)\partial_z q_{\hat{\beta}_n}(z_n^i(\hat{\theta}_n)) - z_n^i(\hat{\theta}_n)\partial_z r_{\hat{\beta}_n}(z_n^i(\hat{\theta}_n))\right| \\
&\quad + \frac{1}{n^{3/2}\log n}\sum_{i=0}^{n-1}\partial_\sigma\left(\frac{\partial_\sigma a(\overline{X}_{t_i^n},\hat{\sigma}_n)}{a(\overline{X}_{t_i^n},\hat{\sigma}_n)}\right)\left|-\frac{\log n}{\hat{\beta}_n^2}q_{\hat{\beta}_n}(z_n^i(\hat{\theta}_n)) + r_{\hat{\beta}_n}(z_n^i(\hat{\theta}_n))\right| \\
&\le \frac{1}{n^{1/2}\log n}\frac{1}{n}\sum_{i=0}^{n-1}C(1+|\overline{X}_{t_i^n}|^p)\left[\frac{\log n}{\hat{\beta}_n^2}(|z_n^i(\hat{\theta}_n)\partial_z q_{\hat{\beta}_n}(z_n^i(\hat{\theta}_n))| + |q_{\hat{\beta}_n}(z_n^i(\hat{\theta}_n))|)\right. \\
&\quad \left. + |z_n^i(\hat{\theta}_n)\partial_z r_{\hat{\beta}_n}(z_n^i(\hat{\theta}_n))| + |r_{\hat{\beta}_n}(z_n^i(\hat{\theta}_n))|\right], \qquad (5.13)
\end{aligned}$$

$$\begin{aligned}
&\frac{|\mathcal{M}_n^{2,2,2}(\hat{\theta}_n)|}{n^{3/2}} \\
&\le \frac{1}{n^{3/2}}\sum_{i=0}^{n-1}\left|\partial_\sigma^2\left(\frac{\partial_\sigma a(\overline{X}_{t_i^n},\hat{\sigma}_n)}{a(\overline{X}_{t_i^n},\hat{\sigma}_n)}\right)k_{\hat{\beta}_n}(z_n^i(\hat{\theta}_n)) - \frac{\partial_\sigma a(\overline{X}_{t_i^n},\hat{\sigma}_n)}{a(\overline{X}_{t_i^n},\hat{\sigma}_n)}\partial_\sigma\left(\frac{\partial_\sigma a(\overline{X}_{t_i^n},\hat{\sigma}_n)}{a(\overline{X}_{t_i^n},\hat{\sigma}_n)}\right)z_n^i(\hat{\theta}_n)\partial_z k_{\hat{\beta}_n}(z_n^i(\hat{\theta}_n))\right. \\
&\left. -2\left(\frac{\partial_\sigma a(\overline{X}_{t_i^n},\hat{\sigma}_n)}{a(\overline{X}_{t_i^n},\hat{\sigma}_n)}\right)\partial_\sigma\left(\frac{\partial_\sigma a(\overline{X}_{t_i^n},\hat{\sigma}_n)}{a(\overline{X}_{t_i^n},\hat{\sigma}_n)}\right)q_{\hat{\beta}_n}(z_n^i(\hat{\theta}_n)) + \left(\frac{\partial_\sigma a(\overline{X}_{t_i^n},\hat{\sigma}_n)}{a(\overline{X}_{t_i^n},\hat{\sigma}_n)}\right)^3 z_n^i(\hat{\theta}_n)\partial_z q_{\hat{\beta}_n}(z_n^i(\hat{\theta}_n))\right| \\
&\le \frac{1}{n^{1/2}}\frac{1}{n}\sum_{i=0}^{n-1}C(1+|\overline{X}_{t_i^n}|^p)(|k_{\hat{\beta}_n}(z_n^i(\hat{\theta}_n))| + |z_n^i(\hat{\theta}_n)\partial_z k_{\hat{\beta}_n}(z_n^i(\hat{\theta}_n))| + |q_{\hat{\beta}_n}(z_n^i(\hat{\theta}_n))| + |z_n^i(\hat{\theta}_n)\partial_z q_{\hat{\beta}_n}(z_n^i(\hat{\theta}_n))|), \qquad (5.14)
\end{aligned}$$

$$\begin{aligned}
\frac{|\mathcal{M}_n^{2,3,3}(\hat{\theta}_n)|}{n^{3/2}(\log n)^2} &\le \frac{1}{n^{3/2}(\log n)^2}\sum_{i=0}^{n-1}\left|\frac{\partial_\sigma a(\overline{X}_{t_i^n},\hat{\sigma}_n)}{a(\overline{X}_{t_i^n},\hat{\sigma}_n)}\right|\left|2\frac{\log n}{\hat{\beta}_n^3}q_\beta(z_n^i(\hat{\theta}_n)) - \frac{\log n}{\hat{\beta}_n^2}(-\frac{\log n}{\hat{\beta}_n^2}z_n^i(\hat{\theta}_n)\partial_z q_{\hat{\beta}_n}(z_n^i(\hat{\theta}_n))\right. \\
&\quad \left. + (\partial_\beta q_{\hat{\beta}_n})(z_n^i(\hat{\theta}_n))) - \frac{\log n}{\hat{\beta}_n^2}z_n^i(\hat{\theta}_n)\partial_z r_{\hat{\beta}_n}(z_n^i(\hat{\theta}_n)) + (\partial_\beta r_{\hat{\beta}_n})(z_n^i(\hat{\theta}_n))\right| \\
&\le \frac{1}{n^{1/2}(\log n)^2}\frac{1}{n}\sum_{i=0}^{n-1}C(1+|\overline{X}_{t_i^n}|^p)\left[\frac{\log n}{\hat{\beta}_n^3}|q_\beta(z_n^i(\hat{\theta}_n))| + \frac{\log n}{\hat{\beta}_n^2}(\frac{\log n}{\hat{\beta}_n^2}|z_n^i(\hat{\theta}_n)\partial_z q_{\hat{\beta}_n}(z_n^i(\hat{\theta}_n))|\right. \\
&\quad \left. + (|\partial_\beta q_{\hat{\beta}_n})(z_n^i(\hat{\theta}_n)))| + \frac{\log n}{\hat{\beta}_n^2}|z_n^i(\hat{\theta}_n)\partial_z r_{\hat{\beta}_n}(z_n^i(\hat{\theta}_n))| + |(\partial_\beta r_{\hat{\beta}_n})(z_n^i(\hat{\theta}_n))|\right]. \qquad (5.15)
\end{aligned}$$

Then, by similar arguments as above, we also obtain their convergences to zero using Theorem 5.2.

Concerning the elements of the matrix $\mathcal{H}_n^3(\theta)$: Since from symmetricity of this matrix

$$\frac{\mathcal{M}_n^{3,1,1}(\hat{\theta}_n)}{n^{2/\beta_0-1/2}\log n} = \frac{\mathcal{M}_n^{1,1,3}(\hat{\theta}_n)}{n^{2/\beta_0-1/2}\log n}, \quad \frac{\mathcal{M}_n^{3,1,2}(\hat{\theta}_n)}{n^{1/\beta_0+1/2}\log n} = \frac{\mathcal{M}_n^{1,2,3}(\hat{\theta}_n)}{n^{1/\beta_0+1/2}\log n},$$

$$\frac{\mathcal{M}_n^{3,1,3}(\hat\theta_n)}{n^{1/\beta_0+1/2}(\log n)^2}=\frac{\mathcal{M}_n^{1,3,3}(\hat\theta_n)}{n^{1/\beta_0+1/2}(\log n)^2},\quad \frac{\mathcal{M}_n^{3,2,2}(\hat\theta_n)}{n^{3/2}\log n}=\frac{\mathcal{M}_n^{2,2,3}(\hat\theta_n)}{n^{3/2}\log n},\quad \frac{\mathcal{M}_n^{3,2,3}(\hat\theta_n)}{n^{3/2}(\log n)^2}=\frac{\mathcal{M}_n^{2,3,3}(\hat\theta_n)}{n^{3/2}(\log n)^2},$$

the proof of their convergences are treated as above. For the element resting, we rewrite as follows

$$\begin{aligned}\frac{|\mathcal{M}_n^{3,3,3}(\hat\theta_n)|}{n^{3/2}(\log n)^3}\le{}&\frac{1}{n^{3/2}(\log n)^3}\sum_{i=0}^{n-1}\Big|(\partial_\beta^2 f_{\hat\beta_n})(z_n^i(\hat\theta_n))+4\frac{\log n}{\hat\beta_n^3}r_{\hat\beta_n}(z_n^i(\hat\theta_n))-2\frac{\log n}{\hat\beta_n^2}(-\frac{\log n}{\hat\beta_n^2}z_n^i(\hat\theta_n)\partial_z r_{\hat\beta_n}(z_n^i(\hat\theta_n))\\&+(\partial_\beta r_{\hat\beta_n})(z_n^i(\hat\theta_n)))-6\frac{\log n}{\hat\beta_n^4}k_{\hat\beta_n}(z_n^i(\hat\theta_n))+2\frac{\log n}{\hat\beta_n^3}(-\frac{\log n}{\hat\beta_n^2}z_n^i(\hat\theta_n)\partial_z k_{\hat\beta_n}(z_n^i(\hat\theta_n))+(\partial_\beta k_{\hat\beta_n})(z_n^i(\hat\theta_n)))\\&-4\frac{(\log n)^2}{\hat\beta_n^5}q_{\hat\beta_n}(z_n^i(\hat\theta_n))+\frac{(\log n)^2}{\hat\beta_n^4}(-\frac{\log n}{\hat\beta_n^2}z_n^i(\hat\theta_n)\partial_z q_{\hat\beta_n}(z_n^i(\hat\theta_n))+(\partial_\beta q_{\hat\beta_n})(z_n^i(\hat\theta_n)))\Big|\\\le{}&\frac{1}{n^{1/2}(\log n)^3}\frac{1}{n}\sum_{i=0}^{n-1}C\Big[|(\partial_\beta^2 f_{\hat\beta_n})(z_n^i(\hat\theta_n))|+\frac{\log n}{\hat\beta_n^3}|r_{\hat\beta_n}(z_n^i(\hat\theta_n))|+\frac{\log n}{\hat\beta_n^2}(\frac{\log n}{\hat\beta_n^2}|z_n^i(\hat\theta_n)\partial_z r_{\hat\beta_n}(z_n^i(\hat\theta_n))|\\&+|(\partial_\beta r_{\hat\beta_n})(z_n^i(\hat\theta_n)))|+\frac{\log n}{\hat\beta_n^4}|k_{\hat\beta_n}(z_n^i(\hat\theta_n))|+\frac{\log n}{\hat\beta_n^3}(\frac{\log n}{\hat\beta_n^2}|z_n^i(\hat\theta_n)\partial_z k_{\hat\beta_n}(z_n^i(\hat\theta_n))|+|(\partial_\beta k_{\hat\beta_n})(z_n^i(\hat\theta_n))|)\\&+\frac{(\log n)^2}{\hat\beta_n^5}|q_{\hat\beta_n}(z_n^i(\hat\theta_n))|+\frac{(\log n)^2}{\hat\beta_n^4}(\frac{\log n}{\hat\beta_n^2}|z_n^i(\hat\theta_n)\partial_z q_{\hat\beta_n}(z_n^i(\hat\theta_n))|+|(\partial_\beta q_{\hat\beta_n})(z_n^i(\hat\theta_n))|)\Big].\end{aligned}\tag{5.16}$$

Thus, by similar arguments as above, we obtain their convergences to zero. □

Proof of Theorem 3.3. Now, here again we can follow (Masuda, 2015, Proof of Theorem 2.10) and (Sweeting, 1980, Theorems 1 and 2) to prove the last assertion.
To obtain LAMN property, we only need to prove the convergence to zeros in probability of the last term in the Taylor's expansion above. To do so, for $\boxed{\hat\theta_n=\theta_0+\varepsilon_n\varphi_n(\theta_0)u}$, we use some inequalities in the proof of Theorem 3.1 for the convergence to zero of the elements of the following matrices

$$\begin{aligned}\mathcal{H}_n^1(\hat\theta_n)={}&n^{1/2-1/\beta_0}\varphi_n^\top(\theta_0)\partial_\mu\mathcal{I}_n(\hat\theta_n)\varphi_n(\theta_0)\\={}&\begin{pmatrix}\frac{\mathcal{M}_n^{1,1,1}(\hat\theta_n)}{n^{3/\beta_0-3/2}} & \frac{\varphi_{11,n}\mathcal{M}_n^{1,1,2}(\hat\theta_n)+\varphi_{21,n}\mathcal{M}_n^{1,1,3}(\hat\theta_n)}{n^{2/\beta_0-1/2}} & \frac{\varphi_{12,n}\mathcal{M}_n^{1,1,2}(\hat\theta_n)+\varphi_{22,n}\mathcal{M}_n^{1,1,3}(\hat\theta_n)}{n^{2/\beta_0-1/2}}\\ \frac{\varphi_{11,n}\mathcal{M}_n^{1,1,2}(\hat\theta_n)+\varphi_{21,n}\mathcal{M}_n^{1,1,3}(\hat\theta_n)}{n^{2/\beta_0-1/2}} & \frac{\mathcal{Q}_{n,1}(\hat\theta_n)}{n^{1/\beta_0+1/2}} & \frac{\mathcal{K}_{n,1}(\hat\theta_n)}{n^{1/\beta_0+1/2}}\\ \frac{\varphi_{12,n}\mathcal{M}_n^{1,1,2}(\hat\theta_n)+\varphi_{22,n}\mathcal{M}_n^{1,1,3}(\hat\theta_n)}{n^{2/\beta_0-1/2}} & \frac{\mathcal{K}_{n,1}(\hat\theta_n)}{n^{1/\beta_0+1/2}} & \frac{\mathcal{P}_{n,1}(\hat\theta_n)}{n^{1/\beta_0+1/2}}\end{pmatrix},\\\mathcal{H}_n^2(\hat\theta_n)={}&n^{-1/2}(\varphi_{11,n}+\varphi_{12,n})\varphi_n^\top(\theta_0)\partial_\sigma\mathcal{I}_n(\hat\theta_n)\varphi_n(\theta_0)\\={}&\begin{pmatrix}\frac{(\varphi_{11,n}+\varphi_{12,n})\mathcal{M}_n^{2,1,1}(\hat\theta_n)}{n^{2/\beta_0-1/2}} & \frac{\varphi_{11,n}\mathcal{M}_n^{2,1,2}(\hat\theta_n)+\varphi_{21,n}\mathcal{M}_n^{2,1,3}(\hat\theta_n)}{(\varphi_{11,n}+\varphi_{12,n})^{-1}n^{1/\beta_0+1/2}} & \frac{\varphi_{12,n}\mathcal{M}_n^{2,1,2}(\hat\theta_n)+\varphi_{22,n}\mathcal{M}_n^{2,1,3}(\hat\theta_n)}{(\varphi_{11,n}+\varphi_{12,n})^{-1}n^{1/\beta_0+1/2}}\\ \frac{\varphi_{11,n}\mathcal{M}_n^{2,1,2}(\hat\theta_n)+\varphi_{21,n}\mathcal{M}_n^{2,1,3}(\hat\theta_n)}{(\varphi_{11,n}+\varphi_{12,n})^{-1}n^{1/\beta_0+1/2}} & \frac{\mathcal{Q}_{n,2}(\hat\theta_n)}{(\varphi_{11,n}+\varphi_{12,n})^{-1}n^{3/2}} & \frac{\mathcal{K}_{n,2}(\hat\theta_n)}{(\varphi_{11,n}+\varphi_{12,n})^{-1}n^{3/2}}\\ \frac{\varphi_{12,n}\mathcal{M}_n^{2,1,2}(\hat\theta_n)+\varphi_{22,n}\mathcal{M}_n^{2,1,3}(\hat\theta_n)}{(\varphi_{11,n}+\varphi_{12,n})^{-1}n^{1/\beta_0+1/2}} & \frac{\mathcal{K}_{n,2}(\hat\theta_n)}{(\varphi_{11,n}+\varphi_{12,n})^{-1}n^{3/2}} & \frac{\mathcal{P}_{n,2}(\hat\theta_n)}{(\varphi_{11,n}+\varphi_{12,n})^{-1}n^{3/2}}\end{pmatrix},\\\mathcal{H}_n^3(\hat\theta_n)={}&n^{-1/2}(\varphi_{21,n}+\varphi_{22,n})\varphi_n^\top(\theta)\partial_\beta\mathcal{I}_n(\hat\theta_n)\varphi_n(\theta)\\={}&\begin{pmatrix}\frac{(\varphi_{21,n}+\varphi_{22,n})\mathcal{M}_n^{3,1,1}(\hat\theta_n)}{n^{2/\beta_0-1/2}} & \frac{\varphi_{11,n}\mathcal{M}_n^{3,1,2}(\hat\theta_n)+\varphi_{21,n}\mathcal{M}_n^{3,1,3}(\hat\theta_n)}{(\varphi_{21,n}+\varphi_{22,n})^{-1}n^{1/\beta_0+1/2}} & \frac{\varphi_{12,n}\mathcal{M}_n^{3,1,2}(\hat\theta_n)+\varphi_{22,n}\mathcal{M}_n^{3,1,3}(\hat\theta_n)}{(\varphi_{21,n}+\varphi_{22,n})^{-1}n^{1/\beta_0+1/2}}\\ \frac{\varphi_{11,n}\mathcal{M}_n^{3,1,2}(\hat\theta_n)+\varphi_{21,n}\mathcal{M}_n^{3,1,3}(\hat\theta_n)}{(\varphi_{21,n}+\varphi_{22,n})^{-1}n^{1/\beta_0+1/2}} & \frac{\mathcal{Q}_{n,3}(\hat\theta_n)}{(\varphi_{21,n}+\varphi_{22,n})^{-1}n^{3/2}} & \frac{\mathcal{K}_{n,3}(\hat\theta_n)}{(\varphi_{21,n}+\varphi_{22,n})^{-1}n^{3/2}}\\ \frac{\varphi_{12,n}\mathcal{M}_n^{3,1,2}(\hat\theta_n)+\varphi_{22,n}\mathcal{M}_n^{3,1,3}(\hat\theta_n)}{(\varphi_{21,n}+\varphi_{22,n})^{-1}n^{1/\beta_0+1/2}} & \frac{\mathcal{K}_{n,3}(\hat\theta_n)}{(\varphi_{21,n}+\varphi_{22,n})^{-1}n^{3/2}} & \frac{\mathcal{P}_{n,3}(\hat\theta_n)}{(\varphi_{21,n}+\varphi_{22,n})^{-1}n^{3/2}}\end{pmatrix}.\end{aligned}$$

where for convenience $\varphi_{ij,n}$ signifies $\varphi_{ij,n}(\theta_0)$, $i,j \in \{1,2\}$ and for $i \in \{1,2,3\}$,

$$\mathcal{K}_{n,i}(\hat{\theta}_n) = \varphi_{11,n}\varphi_{12,n}\mathcal{M}_n^{i,2,2}(\hat{\theta}_n) + (\varphi_{11,n}\varphi_{22,n} + \varphi_{21,n}\varphi_{12,n})\mathcal{M}_n^{i,2,3}(\hat{\theta}_n) + \varphi_{21,n}\varphi_{22,n}\mathcal{M}_n^{i,3,3}(\hat{\theta}_n),$$
$$\mathcal{Q}_{n,i}(\hat{\theta}_n) = \varphi_{11,n}^2\mathcal{M}_n^{i,2,2}(\hat{\theta}_n) + 2\varphi_{11,n}\varphi_{21,n}\mathcal{M}_n^{i,2,3}(\hat{\theta}_n) + \varphi_{21,n}^2\mathcal{M}_n^{i,3,3}(\hat{\theta}_n),$$
$$\mathcal{P}_{n,i}(\hat{\theta}_n) = \varphi_{12,n}^2\mathcal{M}_n^{i,2,2}(\hat{\theta}_n) + 2\varphi_{12,n}\varphi_{22,n}\mathcal{M}_n^{i,2,3}(\hat{\theta}_n) + \varphi_{22,n}^2\mathcal{M}_n^{i,3,3}(\hat{\theta}_n).$$

First, from our choice of $\varphi_n(\theta_0)$ above, $|\varphi_{ij,n}(\theta_0)| \leq C(1+\log n)$, $i,j \in \{1,2\}$. This deduces that

$$\frac{v_n(\theta_0)}{\sqrt{n}} \underset{n\to\infty}{\longrightarrow} 0_{4\times 4} \tag{5.17}$$

where $v_n(\theta_0) = (v_{ij,n}(\theta_0))_{i,j\in\{1,\dots,4\}} = (\,\varphi_{11,n}\ \varphi_{12,n}\ \varphi_{21,n}\ \varphi_{21,n}\,)^\top(\,\varphi_{11,n}\ \varphi_{12,n}\ \varphi_{21,n}\ \varphi_{21,n}\,)$.
Now, since $\hat{\beta}_n = \beta_0 + \frac{\varepsilon_n}{\sqrt{n}}(u_2\varphi_{21,n}(\theta_0) + u_3\varphi_{22,n}(\theta_0))$, we have

$$\begin{aligned} n^{1/\hat{\beta}_n - 1/\beta_0} &= \exp\left(\frac{\log n}{\beta_0}\left(\frac{1}{1+\frac{\varepsilon_n}{\sqrt{n}}(u_2\varphi_{21,n}(\theta_0)+u_3\varphi_{22,n}(\theta_0))} - 1\right)\right) \\ &\sim \exp\left(-\frac{\log n}{\beta_0}\frac{\varepsilon_n}{\sqrt{n}}(u_2\varphi_{21,n}(\theta_0)+u_3\varphi_{22,n}(\theta_0))\right) \underset{n\to\infty}{\longrightarrow} 1. \end{aligned}$$

Thus, combined with the previous analyses for the NM case, we can obtain easily the convergence in probability of the considering terms. In particular, from (5.7), $\frac{\mathcal{M}_n^{1,1,1}(\hat{\theta}_n)}{n^{3/\beta_0-3/2}} \underset{n\to\infty}{\longrightarrow} 0$. And from (5.8), (5.9), (5.10), (5.11), (5.12), (5.13), (5.15), (5.14) and (5.16), respectively we get that

$$\frac{v_{ij,n}(\theta_0)}{\sqrt{n}}\sqrt{n}\frac{\mathcal{M}_n^{1,1,2}(\hat{\theta}_n)}{n^{2/\beta_0-1/2}},\ \frac{v_{ij,n}(\theta_0)\log n}{\sqrt{n}}\sqrt{n}\frac{\mathcal{M}_n^{1,1,3}(\hat{\theta}_n)}{n^{2/\beta_0-1/2}\log n},\ \frac{v_{ij,n}(\theta_0)}{\sqrt{n}}\sqrt{n}\frac{\mathcal{M}_n^{1,2,2}(\hat{\theta}_n)}{n^{1/\beta_0+1/2}},$$
$$\frac{v_{ij,n}(\theta_0)\log n}{\sqrt{n}}\sqrt{n}\frac{\mathcal{M}_n^{1,2,3}(\hat{\theta}_n)}{n^{1/\beta_0+1/2}\log n},\ \frac{v_{ij,n}(\theta_0)(\log n)^2}{\sqrt{n}}\sqrt{n}\frac{\mathcal{M}_n^{1,3,3}(\hat{\theta}_n)}{n^{1/\beta_0+1/2}(\log n)^2},\ \frac{v_{ij,n}(\theta_0)\log n}{\sqrt{n}}\sqrt{n}\frac{\mathcal{M}_n^{2,2,3}(\hat{\theta}_n)}{n^{3/2}\log n},$$
$$\frac{v_{ij,n}(\theta_0)(\log n)^2}{\sqrt{n}}\sqrt{n}\frac{\mathcal{M}_n^{2,3,3}(\hat{\theta}_n)}{n^{3/2}(\log n)^2},\ \frac{v_{ij,n}(\theta_0)}{\sqrt{n}}\sqrt{n}\frac{\mathcal{M}_n^{2,2,2}(\hat{\theta}_n)}{n^{3/2}},\ \frac{v_{ij,n}(\theta_0)(\log n)^3}{\sqrt{n}}\sqrt{n}\frac{\mathcal{M}_n^{3,3,3}(\hat{\theta}_n)}{n^{3/2}(\log n)^3}$$

are convergent to zero as n tends to infinity for any $i,j \in \{1,\dots,4\}$ thanks to (5.37). □

Acknowledgements

We would like to thank Emmanuelle Clément for valuable comments and suggestions.
This research benefited from the support of the ANR project "Efficient inference for large and high-frequency data" (ANR-21-CE40-0021).

Supplementary Material

The proofs of the technical results (Proposition 5.1, Lemma 3.8, Theorem 3.9 and Theorem 3.10) are shown in the supplementary material.

References

BAYRAKTAR, E. and CLÉMENT, E. (2023). Estimation of a pure-jump stable Cox-Ingersoll-Ross process. working paper or preprint.

BROUSTE, A. and MASUDA, H. (2018). Efficient estimation of stable Lévy process with symmetric jumps. *Statistical Inference for Stochastic Processes* **21** 289-307. https://doi.org/10.1007/s11203-018-9181-0

CLÉMENT, E. (2023). Hellinger and total variation distance in approximating Lévy driven SDEs. *Annals of Applied Probability* **33** 2176-2209. https://doi.org/10.1214/22-AAP1863

CLÉMENT, E. and GLOTER, A. (2015). Local Asymptotic Mixed Normality property for discretely observed stochastic differential equations driven by stable Lévy processes. *Stochastic Processes and their Applications* **125** 2316-2352.

CLÉMENT, E. and GLOTER, A. (2019). Estimating functions for SDE driven by stable Lévy processes. *Annales de l'Institut Henri Poincaré* **55** 1316–1348.

CLÉMENT, E., GLOTER, A. and NGUYEN, H. (2019). LAMN property for the drift and volatility parameters of a SDE driven by a stable Lévy Process. *ESAIM:P & S* **23** 136–175.

CLÉMENT, E. and GLOTER, A. (2020). Joint estimation for SDE driven by locally stable Lévy processes. *Electronic Journal of Statistics* **14** 2922 – 2956. https://doi.org/10.1214/20-EJS1737

GENON-CATELOT, V. and JACOD, J. (1993). On the estimation of the diffusion coefficient for multidimensional diffusion processes. *Ann. Inst. H. Poincaré Probab.Statist.* **29** 119-151.

GOBET, E. (2001). Local asymptotic mixed normality property for elliptic diffusion: a Malliavin calculus approach. *Bernoulli* **7** 899–912.

HÖPFNER, R. (2014). *Asymptotic Statistics*. De Gruyter, Berlin, Boston. https://doi.org/doi:10.1515/9783110250282

JACOD, J. and PROTTER, P. (2012). *Discretization of Processes*, 1 ed. *Stochastic Modelling and Applied Probability*. Springer Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-24127-7

JACOD, J. and SORENSEN, M. (2017). A review of asymptotic theory of estimating functions. *Statistical Inference for Stochastic Processes* **21** 415-434.

JEGANATHAN, P. (1982). On the Asymptotic Theory of Estimation When the Limit of the Log-Likelihood Ratios Is Mixed Normal. *Sankhyā: The Indian Journal of Statistics, Series A (1961-2002)* **44** 173–212.

KUNITA, H. (2004). *Stochastic Differential Equations Based on Lévy Processes and Stochastic Flows of Diffeomorphisms* In *Real and Stochastic Analysis: New Perspectives* 305–373. Birkhäuser Boston, Boston, MA. https://doi.org/10.1007/978-1-4612-2054-1_6

LE CAM, L. (1956). On the Asymptotic Theory of Estimation and Testing Hypotheses. In *Proc. Third Berkeley Symp. Math. Statist. Prob.* **1** 355–368.

LE CAM, L. (1964). Sufficiency and Approximate Sufficiency. *The Annals of Mathematical Statistics* **35** 1419 – 1455. https://doi.org/10.1214/aoms/1177700372

MASUDA, H. (2015). *Parametric Estimation of Lévy Processes* In *Lévy Matters IV: Estimation for Discretely Observed Lévy Processes* 179–286. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-319-12373-8_3

MATSUI, M. and TAKEMURA, A. (2004). Some Improvements in Numerical Evaluation of Symmetric Stable Density and Its Derivatives. *Communications in Statistics - Theory and Methods* **35** 149 - 172.

PROTTER, P. (1992). *Stochastic Integration and Differential Equation*, Second ed. Springer-Verlag, Berlin, Heidelberg.

SATO, K. I. (2000). Density transformation in Lévy processes. *MaPhySto Lecture Notes 7*.

SHIRYAEV, A. and SPOKOINY, V. (2000). *Statistical Experiments and Decisions - Asymptotic Theory*. World Scientific.

STRASSER, H. (1985). *Mathematical Theory of Statistics*. De Gruyter, Berlin, New York. https://doi.org/doi:10.1515/9783110850826

SWEETING, T. J. (1980). Uniform Asymptotic Normality of the Maximum Likelihood Estimator. *The Annals of Statistics* **8** 1375 – 1381. https://doi.org/10.1214/aos/1176345208

TODOROV, V. (2013). Power variation from second order differences for pure jump semimartingales. *Stochastic Processes and their Applications* **123** 2829-2850. A Special Issue on the Occasion of the 2013 International Year of Statistics. https://doi.org/10.1016/j.spa.2013.04.005

Supplement to "Efficient estimation for stochastic differential equations driven by a stable Lévy process"

ALEXANDRE BROUSTE[1,a], LAURENT DENIS[1,b] and THI-BAO-TRÂM NGÔ[1,c]

[1] *Le Mans Université, Laboratoire Manceau de Mathématiques, Avenue Olivier Messiaen, 72085 Le Mans Cedex 09, France,* [a]*Alexandre.Brouste@univ-lemans.fr,* [b]*Laurent.Denis@univ-lemans.fr,* [c]*Thi_Bao_Tram.Ngo@univ-lemans.fr*

Keywords: Lévy process, stable process; stochastic differential equation; LAMN property; parametric estimation, one-step procedure

Proofs of some technical results

Proof of Proposition 5.1. The first two inequalities are recalled from Clément and Gloter (2020), here, we only need to prove the last one. First, we note that

$$\partial_\mu \xi_{1/n}(x,\mu) = \int_0^{1/n} \partial_x b(\xi_s(x,\mu),\mu)\partial_\mu \xi_s(x,\mu)ds + \int_0^{1/n} (\partial_\mu b)(\xi_s(x,\mu),\mu)ds,$$

$$\begin{aligned}\partial_\mu^2 \xi_{1/n}(x,\mu) &= \int_0^{1/n} \partial_x b(\xi_s(x,\mu),\mu)\partial_\mu^2 \xi_s(x,\mu)ds + \int_0^{1/n} \partial_x^2 b(\xi_s(x,\mu),\mu)(\partial_\mu \xi_s(x,\mu))^2 ds \\ &\quad + 2\int_0^{1/n} (\partial_\mu \partial_x b)(\xi_s(x,\mu),\mu)\partial_\mu \xi_s(x,\mu)ds + \int_0^{1/n} (\partial_\mu^2 b)(\xi_s(x,\mu),\mu)ds,\end{aligned}$$

$$\begin{aligned}\partial_\mu^3 \xi_{1/n}(x,\mu) &= \int_0^{1/n} \partial_x b(\xi_s(x,\mu),\mu)\partial_\mu^3 \xi_s(x,\mu)ds + \int_0^{1/n} \partial_x^2 b(\xi_s(x,\mu),\mu)\partial_\mu^2 \xi_s(x,\mu)ds \\ &\quad + \int_0^{1/n} (\partial_\mu \partial_x b)(\xi_s(x,\mu),\mu)\partial_\mu^2 \xi_s(x,\mu)ds + \int_0^{1/n} \partial_x^3 b(\xi_s(x,\mu),\mu)(\partial_\mu \xi_s(x,\mu))^3 ds \\ &\quad + 3\int_0^{1/n} (\partial_\mu \partial_x^2 b)(\xi_s(x,\mu),\mu)(\partial_\mu \xi_s(x,\mu))^2 ds \\ &\quad + 3\int_0^{1/n} (\partial_\mu^2 \partial_x b)(\xi_s(x,\mu),\mu)(\partial_\mu \xi_s(x,\mu))ds + \int_0^{1/n} (\partial_\mu^3 b)(\xi_s(x,\mu),\mu)ds.\end{aligned}$$

Therefore, we get

$$|\partial_\mu^3 \xi_{1/n}(x,\mu) - \frac{1}{n}\partial_\mu^3 b(x,\mu)| \le \int_0^{1/n} |\partial_x b(\xi_s(x,\mu),\mu)||\partial_\mu^3 \xi_s(x,\mu) - \frac{1}{n}\partial_\mu^3 b(x,\mu)|ds + |\mathcal{R}_n(\mu)|,$$

where

$$\mathcal{R}_n(\mu)$$

$$
\begin{aligned}
=&\frac{1}{n}\partial_\mu^3 b(x,\mu)\int_0^{1/n}\partial_x b(\xi_s(x,\mu),\mu)ds+\int_0^{1/n}\partial_x^2 b(\xi_s(x,\mu),\mu)\partial_\mu^2\xi_s(x,\mu)ds\\
&+\int_0^{1/n}(\partial_\mu\partial_x b)(\xi_s(x,\mu),\mu)\partial_\mu^2\xi_s(x,\mu)ds+\int_0^{1/n}\partial_x^3 b(\xi_s(x,\mu),\mu)(\partial_\mu\xi_s(x,\mu))^3 ds\\
&+3\int_0^{1/n}(\partial_\mu\partial_x^2 b)(\xi_s(x,\mu),\mu)(\partial_\mu\xi_s(x,\mu))^2 ds+3\int_0^{1/n}(\partial_\mu^2\partial_x b)(\xi_s(x,\mu),\mu)(\partial_\mu\xi_s(x,\mu))ds\\
&+\int_0^{1/n}[(\partial_\mu^3 b)(\xi_s(x,\mu),\mu)-\partial_\mu^3 b(x,\mu)]ds.
\end{aligned}
$$

Now, for some $s\in(0,1/n)$, since $b(.)$ is Lipschitz, we have

$$
\begin{aligned}
|\xi_s(x,\mu)| =&|x+\int_0^s b(\xi_v(x,\mu),\mu)dv|\le|x|+\int_0^s C(1+|\xi_v(x,\mu)|)dv,\\
|\xi_s(x,\mu)-x| =&|\int_0^s b(\xi_v(x,\mu),\mu)dv|\le C(|x|+1)s+\int_0^s C|\xi_v(x,\mu)-x|dv.
\end{aligned}
$$

By Gronwall's lemma

$$
|\xi_s(x,\mu)|\le(|x|+Cs)e^{Cs},\quad\text{and }|\xi_s(x,\mu)-x|\le C(|x|+1)se^{Cs},\quad s\in(0,1/n). \tag{5.18}
$$

Then, from Assumption (A) and Gronwall's lemma again, we have

$$
\sup_{\mu\in V_{\mu_0}}\sup_{k\in\{1,2,3\}}\partial_\mu^k\xi_{1/n}(x,\mu)\le C(1+|x|^p)/n \tag{5.19}
$$

for some $p>0$. Thus, from Assumption (A) and Jensen's inequality for convex function $x\in[1,\infty)\mapsto x^m$, $m>1$, we easily deduce that $\sup_{\mu\in V_{\mu_0}}|\mathcal{R}_n(\mu)|\le C(1+|x|^p)/n^2$ for some $p>0$. Finally, combined with Gronwall's lemma, we then obtain (5.2). □

Proof of Theorem 5.2. First of all, we introduce for any $q>0$

$$
\begin{aligned}
J_t^q &= J_t-\sum_{s\le t}\Delta J_s\mathbb{1}_{\{|\Delta J_s|>q\}},\\
T_q &= \inf\{t>0:|\Delta J_t|\ge q\}.
\end{aligned}
$$

and X^q the solution of the following SDE:

$$
X_t^q=x_0+\int_0^t b(X_s^q,\mu_0)ds+\int_0^t a(X_{s-}^q,\sigma_0)dJ_s^q,\quad t\ge0, \tag{5.20}
$$

In a natural way, we denote by $\overline{X}^q$ the associate scheme, for any $i\in\{0,1,\dots,n-1\}$, we denote

$$
\overline{X}^q_{t_{i+1}^n}=\xi_{t_{i+1}^n-t_i^n}(\overline{X}^q_{t_i^n},\mu)+a(\overline{X}^q_{t_i^n},\sigma)(J^q_{t_{i+1}^n}-J^q_{t_i^n}). \tag{5.21}
$$

Let $q>1$ and $p\ge2$ be fixed. It is standard that

$$
\mathbb{E}(\sup_{t\in[0,1]}|X_t^q|^p)<+\infty. \tag{5.22}
$$

Concerning the scheme, we also have for any $q > 1$ and n

$$\mathbb{E}(\sup_{i\in\{0,1,\cdots,n\}} |\overline{X}^q_{t_i^n}|^p) \leq C_{q,p}. \tag{5.23}$$

Indeed, let $n \geq 1$ and $i \in \{1, 2, \cdots, n-1\}$ then we have

$$\begin{aligned}\overline{X}^q_{t_{i+1}^n} &= \overline{X}^q_{t_i^n} + \int_0^{\frac{1}{n}} b(\xi_s(\overline{X}^q_{t_i^n}, \mu_0), \mu_0)ds + a(\overline{X}^q_{t_i^n}, \sigma_0)(J^q_{t_{i+1}^n} - J^q_{t_i^n}) \\ &= x_0 + \int_{t_i^n}^{t_{i+1}^n} b(\xi_{u-\eta_n(u)}(\overline{X}^q_{\eta_n(u)}, \mu_0), \mu_0)\, du + \int_{t_i^n}^{t_{i+1}^n} a(\overline{X}^q_{\eta_n(u)}, \sigma_0)dJ^q_u\end{aligned}$$

Let us now introduce the process Z defined by

$$\forall t \in [0,1], Z_t = x_0 + \int_0^t b(\xi_{u-\eta_n(u)}(\overline{X}^q_{\eta_n(u)}, \mu_0), \mu_0)\, du + \int_0^t a(\overline{X}^q_{\eta_n(u)}, \sigma_0)dJ^q_u$$

so that for any i, $Z_{t_i^n} = \overline{X}^q_{t_i^n}$ and so

$$\forall t \in [0,1], Z^*_t = \sup_{s\in[0,t]} |Z_s| \geq \sup_i |\overline{X}^q_{t_i^n}|. \tag{5.24}$$

Following for example Protter (1992) Section I.4, we know that J^q admits the following decomposition:

$$J^q_t = M^q_t + A^q_t, \tag{5.25}$$

with $M^q_t = \int_0^t \int_{|z|\leq q} z\tilde{N}(ds, dz)$, where $\tilde{N}$ is the compensated Poisson measure with Lévy measure $\frac{1}{|z|^{1+\beta}}\mathbf{1}_{[-q,q]}(z)dz$. M^q is a martingale belonging to all the L^p spaces and A^q is a finite variation process such that its total variation satisfies

$$d|A|_t \leq \left(2\int_1^q \frac{z}{z^{1+\beta}}dz\right)dt.$$

Then since coefficients a and b are Lipschitz w.r.t. variable x and thanks to estimates (5.18) and (5.24), we have:

$$\begin{aligned}|Z^*_t|^p \leq & C\left(|x_0|^p + \left(\int_0^t (Z^*_u + C/n)e^{C/n}\, du\right)^p + \left(\int_0^t (1 + Z^*_u)\, du\right)^p\right. \\ & \left. + \sup_{s\in[0,t]}\left|\int_0^s a(\overline{X}^q_{\eta_n(u)}, \sigma_0)dM^q_u\right|^p\right),\end{aligned} \tag{5.26}$$

where constant C above but also below denotes a constant which may vary from line to line but does not depend on n. Using (Jacod and Protter, 2012, Inequality (2.1.36) in Lemma 2.1.5) (see also (Kunita,

2004, Theorem A.3)) we have:

$$\begin{aligned}
&\mathbb{E}_{\theta_0}\left(\sup_{s\in[0,t]}\left|\int_0^s a(\overline{X}^q_{\eta_n(u)},\sigma_0)dM^q_u\right|^p\right)\\
&\le C\left(\mathbb{E}_{\theta_0}\left[\left(\int_0^t\int_{|z|\le q}|a(\overline{X}^q_{\eta_n(u)},\sigma_0)|^2\frac{z^2}{|z|^{1+\beta}}dzds\right)^{p/2}\right]+\mathbb{E}_{\theta_0}\left[\int_0^t\int_{|z|\le q}|a(\overline{X}^q_{\eta_n(u)},\sigma_0)|^p\frac{z^p}{|z|^{1+\beta}}dzds\right]\right)\\
&\le C\left(\mathbb{E}_{\theta_0}\left[\left(\int_0^t|a(\overline{X}^q_{\eta_n(u)},\sigma_0)|^2ds\right)^{p/2}\right]+\mathbb{E}_{\theta_0}\left[\int_0^t|a(\overline{X}^q_{\eta_n(u)},\sigma_0)|^pds\right]\right)\\
&\le C\mathbb{E}_{\theta_0}\left[\int_0^t|a(\overline{X}^q_{\eta_n(u)},\sigma_0)|^pdu\right]\le C\left(1+\mathbb{E}_{\theta_0}\left[\int_0^t(Z^*_u)^pdu\right]\right)
\end{aligned}\tag{5.27}$$

From this we easily get that

$$\forall t\in[0,1],\ \mathbb{E}_{\theta_0}[(Z^*_t)^p]\le C(1+\int_0^t\mathbb{E}_{\theta_0}[(Z^*_s)^p]\,ds)$$

which ensures by Gronwall's Lemma that $\mathbb{E}_{\theta_0}[(Z^*_1)^p]\le C_{q,p}$.
Clearly, on $N_q=\{T_q>1\}$, $X_t=X^q_t$ a.s. for all $t\in[0,1]$ and $\overline{X}=\overline{X}^q$. As $\lim_{q\to+\infty}\mathbb{P}_{\theta_0}(T_q\le 1)=0$, we only need to prove that for any $q>0$,

$$\sup_{\theta\in V_n^{(\eta)}(\theta_0)}n^{\varepsilon}\left|\frac{1}{n}\sum_{i=0}^{n-1}f(\overline{X}^q_{t_i^n},\mu,\sigma)g_\beta(z^i_n(\theta))-\int_0^1 f(X^q_s,\mu_0,\sigma_0)ds\mathbb{E}_{\theta_0}(g_{\beta_0}(J_1))\right|\mathbb{1}_{N_q}\underset{n\to\infty}{\longrightarrow}0,\tag{5.28}$$

in probability under $\mathbb{P}_{\theta_0}$. This is implied by the following convergences in probability,

$$\sup_{\theta\in V_n^{(\eta)}(\theta_0)}n^{\varepsilon}\left|\frac{1}{n}\sum_{i=0}^{n-1}[f(\overline{X}^q_{t_i^n},\mu,\sigma)-f(\overline{X}^q_{t_i^n},\mu_0,\sigma_0)]g_\beta(z^i_n(\theta))\right|\mathbb{1}_{N_q}\overset{\mathbb{P}_{\theta_0}}{\underset{n\to\infty}{\longrightarrow}}0,\tag{5.29}$$

$$\sup_{\theta\in V_n^{(\eta)}(\theta_0)}n^{\varepsilon}\left|\frac{1}{n}\sum_{i=0}^{n-1}f(\overline{X}^q_{t_i^n},\mu_0,\sigma_0)[g_\beta(z^i_n(\theta))-g_{\beta_0}(z^i_n(\theta_0))]\right|\mathbb{1}_{N_q}\overset{\mathbb{P}_{\theta_0}}{\underset{n\to\infty}{\longrightarrow}}0,\tag{5.30}$$

$$n^{\varepsilon}\left|\frac{1}{n}\sum_{i=0}^{n-1}f(\overline{X}^q_{t_i^n},\mu_0,\sigma_0)[g_{\beta_0}(z^i_n(\theta_0))-\mathbb{E}_{\theta_0}(g_{\beta_0}(J_1))]\right|\overset{\mathbb{P}_{\theta_0}}{\underset{n\to\infty}{\longrightarrow}}0,\tag{5.31}$$

$$n^{\varepsilon}\left|\frac{1}{n}\sum_{i=0}^{n-1}f(\overline{X}^q_{t_i^n},\mu_0,\sigma_0)-\frac{1}{n}\sum_{i=0}^{n-1}f(X^q_{t_i^n},\mu_0,\sigma_0)\right|\overset{\mathbb{P}_{\theta_0}}{\underset{n\to\infty}{\longrightarrow}}0.\tag{5.32}$$

$$n^{\varepsilon}\left|\frac{1}{n}\sum_{i=0}^{n-1}f(X^q_{t_i^n},\mu_0,\sigma_0)-\int_0^1 f(X^q_s,\mu_0,\sigma_0)ds\right|\overset{\mathbb{P}_{\theta_0}}{\underset{n\to\infty}{\longrightarrow}}0.\tag{5.33}$$

We remark then that

$$\begin{aligned}
z^i_n(\theta_0)=&n^{1/\beta_0}(J_{t^n_{i+1}}-J_{t^n_i}),\\
z^i_n(\theta)=&\frac{n^{1/\beta}}{a(\overline{X}_{t^n_i},\sigma)}[\xi_{1/n}(\overline{X}_{t^n_i},\mu_0)+a(\overline{X}_{t^n_i},\sigma_0)(J_{t^n_{i+1}}-J_{t^n_i})-\xi_{1/n}(\overline{X}_{t^n_i},\mu)]
\end{aligned}$$

and that if $\theta \in V_n^{(\eta)}(\theta_0)$, then

$$|\mu-\mu_0| \le \eta\, n^{1/2-1/\beta_0}(\log n)^2,\ |\sigma-\sigma_0| \le C\,\eta\, n^{-1/2}\log n \text{ and } |\beta-\beta_0| \le C\,\eta\, n^{-1/2}\log n. \quad (5.34)$$

For (5.29) we remark that

$$\begin{aligned}
&\mathbb{E}_{\theta_0}\left(\sup_{\theta\in V_n^{(\eta)}(\theta_0)} n^{\varepsilon}\left|\frac{1}{n}\sum_{i=0}^{n-1}[f(\overline{X}^q_{t_i^n},\mu,\sigma)-f(\overline{X}^q_{t_i^n},\mu_0,\sigma_0)]g_\beta(z_n^i(\theta))\right|\mathbb{1}_{N_q}\right)\\
&\le \mathbb{E}_{\theta_0}\left(C(1+\sup_i|\overline{X}^q_{t_i^n}|^p)(\log n)^2(n^{1/2-1/\beta_0+\varepsilon}+n^{-1/2+\varepsilon})|g_\beta(z_n^i(\theta)|\mathbb{1}_{N_q}\right)\\
&\le C(1+\mathbb{E}_{\theta_0}(\sup_i|\overline{X}^q_{t_i^n}|^{2p}))^{1/2}(1+\mathbb{E}_{\theta_0}(\sup_{\theta\in V_n^{(\eta)}(\theta_0)}|z_n^i(\theta)|^{2\kappa}\mathbb{1}_{N_q}))^{1/2}(\log n)^2(n^{1/2-1/\beta_0+\varepsilon}+n^{-1/2+\varepsilon})\\
&\le C(\log n)^2(n^{1/2-1/\beta_0+\varepsilon}+n^{-1/2+\varepsilon}) \underset{n\to+\infty}{\longrightarrow} 0,
\end{aligned}$$

where κ may be chosen in $]0,\beta/2[$ thanks to the assumptions made on g_β. The last line is followed from (Clément and Gloter, 2020, the proof of (4.11)) that $\mathbb{E}_{\theta_0}(|z_n^i(\theta_0)|^k|\mathcal{F}_{t_i^n}) \le C$, for any $k<\beta_0$ and $\sup_{\theta\in V_n^{(\eta)}(\theta_0)}|z_n^i(\theta)-z_n^i(\theta_0)|\mathbb{1}_{N_q} \le C\ (1+|\overline{X}^q_{t_i^n}|^j)(1+|z_n^i(\theta_0)|)(\log n)^2/\sqrt{n}$, for some $j>0$.

To prove (5.30), we proceed similarly by splitting the difference $g_\beta(z_n^i(\theta))-g_{\beta_0}(z_n^i(\theta_0))$ into two parts: $g_\beta(z_n^i(\theta))-g_{\beta_0}(z_n^i(\theta))$ and $g_{\beta_0}(z_n^i(\theta))-g_{\beta_0}(z_n^i(\theta_0))$ using the fact that f is Lipschitz w.r.t. variable x.

For (5.31), let $\xi_i^n = n^{\varepsilon-1}f(\overline{X}^q_{t_i^n},\mu_0,\sigma_0)[g_{\beta_0}(z_n^i(\theta_0))-\mathbb{E}_{\theta_0}(g_{\beta_0}(J_1))]$, we simply use the fact that $z_n^i(\theta_0)=n^{1/\beta_0}(J_{t_{i+1}^n}-J_{t_i^n}) \overset{\mathcal{L}}{=} J_1$, using one more time the fact that f is Lipschitz we have:

$$n\mathbb{E}_{\theta_0}(\xi_i^n|\mathcal{F}_{t_i^n})=0,\quad n\mathbb{E}_{\theta_0}((\xi_i^n)^2|\mathcal{F}_{t_i^n})=n^{2\varepsilon-1}f(\overline{X}^q_{t_i^n},\mu_0,\sigma_0)^2 Var_{\theta_0}(g_{\beta_0}(J_1)) \to 0.$$

we then conclude that $\sum_{i=1}^n \xi_i^n$ tends to 0 in probability using the results in Jacod and Protter (2012), Section (2.2.4) for triangular array. For (5.32), we note that

$$\begin{aligned}
\overline{X}^q_{t_i^n}-X^q_{t_i^n} = Z_{t_i^n}-X^q_{t_i^n} &= \int_0^{t_i^n}(b(\xi_{u-\eta_n(u)}(\overline{X}^q_{\eta_n(u)},\mu_0),\mu_0)-b(X^q_u,\mu_0))\,du\\
&+\int_0^{t_i^n}(a(\overline{X}^q_{\eta_n(u)},\sigma_0)-a(X^q_{u-},\sigma_0))dJ^q_u.
\end{aligned}$$

Then, we separate the first integral in the r.h.s. into three parts concerning the differences

$$\begin{aligned}
&b(\xi_{u-\eta_n(u)}(\overline{X}^q_{\eta_n(u)},\mu_0),\mu_0)-b(\overline{X}^q_{\eta_n(u)},\mu_0)\\
&b(\overline{X}^q_{\eta_n(u)},\mu_0)-b(X^q_{\eta_n(u)},\mu_0)\\
&b(X^q_{\eta_n(u)},\mu_0)-b(X^q_u,\mu_0)
\end{aligned}$$

and we shall do similarly for the second integral. Now, thanks to the Lipschitz property of the coefficient functions $x\mapsto a(x,\sigma_0)$ and $x\mapsto b(x,\mu_0)$, we can easily recycle the arguments from (5.26) to obtain

$$\mathbb{E}_{\theta_0}(\sup_{s\in[0,1]}|\overline{X}^q_s-X^q_s|^2)=O(n^{-1}).$$

Finally, using Lipschitz assumption on f and Cauchy-Schwarz inequality, we have

$$\mathbb{E}\left(n^{\varepsilon}\left|\frac{1}{n}\sum_{i=0}^{n-1} f(\overline{X}^{q}_{t_i^n},\mu_0,\sigma_0)-\frac{1}{n}\sum_{i=0}^{n-1} f(X^{q}_{t_i^n},\mu_0,\sigma_0)\right|\right)\underset{n\to+\infty}{\longrightarrow} 0.$$

For (5.33), we rewrite the considering error as follows

$$n^{\varepsilon}\left|\int_0^1 (f(X^{q}_{\eta_n(s)},\mu_0,\sigma_0)-f(X^{q}_{s},\mu_0,\sigma_0))ds\right|$$

and its convergence to zero in probability is directly using a part of the arguments for (5.32).

□

Proof of Lemma 3.8. To begin with, we denote

$$F_n(\sigma)=n^{\widehat{\beta}^0_n-1}V^1_n(p,X)-\mu_p(\widehat{\beta}^0_n)\int_0^1|a(X_s,\sigma)|^p ds.$$

We prove that if $(\sqrt{n}/\log n)F_n(\sigma_0)$ is tight which is true from (Todorov, 2013, (21)), then, any sequence $(\widehat{\sigma}^0_n, n\geq 1)$ belonging to V_{σ_0} (a neighborhood of σ_0 defined in assumption (A)) that solves $F_n(\sigma)=0$ is consistent and eventually unique. To do so, using (Jacod and Sorensen, 2017, Theorem 2.7.a), since $F_n(\sigma_0)$ converges to zero in probability, we need to verify the following two conditions:

(i) There exists F defined on A compact subset of V_{σ_0}, continuously differentiable, such that $F(\sigma_0)=0$ and σ_0 is the unique root of $F(\sigma)=0$.

(ii) The following convergence in probability holds

$$\sup_{\sigma\in A}|\partial_\sigma F_n(\sigma)-\partial_\sigma F(\sigma)|\to 0,$$

and $\partial_\sigma F(\sigma)$ is non-singular with probability one.

For this proof, we choose

$$F(\sigma)=-p\mu_p(\beta_0)\int_{\sigma_0}^{\sigma}\int_0^1\frac{\partial_\sigma a}{a^{1-p}}(X_s,u)dsdu.$$

For this choice, the criteria (ii) is straightforward from the fact that $\widehat{\beta}^0_n\overset{\mathbb{P}}{\to}\beta_0$. Considering (i), it is obvious that $F(\sigma_0)=0$ and σ_0 is the unique solution of $F(\sigma)=0$ since from those assumptions on the function a,

$$\int_{\sigma_0}^{\sigma}\int_0^1\frac{\partial_\sigma a}{a^{1-p}}(X_s,u)dsdu=0\ \Leftrightarrow\ \sigma=\sigma_0.$$

Now, from this, we prove that $\frac{n^{1/2}}{\log n}(\widehat{\sigma}^0_n-\sigma_0)$ is tight. From Taylor's formula, we have

$$F_n(\sigma_0)=-p\mu_p(\widehat{\beta}^0_n)\int_0^1\frac{\partial_\sigma a}{a^{1-p}}(X_s,\tilde{\sigma}_n)ds\ (\widehat{\sigma}^0_n-\sigma_0)$$

where $\tilde{\sigma}_n$ lies between $\widehat{\sigma}^0_n$ and σ_0. Then, using the consistency of $\widehat{\sigma}^0_n$ proven above and the tightness of $(\sqrt{n}/\log n)F_n(\sigma_0)$, we complete the proof. □

Proof of Theorem 3.9. Inspired by the work in Bayraktar and Clément (2023), using (Jacod and Sorensen, 2017, Theorem 2.7.a)), for the consistency and uniqueness, we need to verify the following two conditions:

(i) There exists G defined on A, continuously differentiable, such that $G_n(\mu_0)$ converges to zero in probability, $G(\mu_0) = 0$ and μ_0 is the unique root of $G(\mu) = 0$.

(ii) The following convergence in probability holds

$$\sup_{\mu \in A} |\partial_\mu G_n(\mu) - \partial_\mu G(\mu)| \to 0,$$

and $\partial_\mu G(\mu)$ is non-singular with probability one.

For some $\eta > 0$, we set

$$W_n^{(\eta)} = \left\{ (\sigma, \beta) : \left\| \frac{\sqrt{n}}{\log n} \begin{pmatrix} \sigma - \sigma_0 \\ \beta - \beta_0 \end{pmatrix} \right\| \leq \eta \right\}.$$

First, we prove the second assertion (ii) by setting

$$G(\mu) = \int_{\mu_0}^{\mu} \int_0^1 \frac{\partial_\mu b(X_s, z)^2}{a(X_s, \sigma_0)^2} ds dz \, \mathbb{E}(h^2_{\beta_0}(J_1)).$$

Since $(\widehat{\sigma}_n^0, \widehat{\beta}_n^0) \in W_n^{(\eta)}$, we know that $\widehat{\beta}_n^0 > 1$ a.s. for any n large enough. We have $\partial_\mu G_n(\mu) = -n^{1-2/\widehat{\beta}_n^0} \mathcal{I}_n^{1,1}(\mu, \widehat{\sigma}_n^0, \widehat{\beta}_n^0)$ where

$$\begin{aligned}
&\mathcal{I}_n^{1,1}(\mu, \widehat{\sigma}_n^0, \widehat{\beta}_n^0) \\
=&n^{1/\widehat{\beta}_n^0} \sum_{i=0}^{n-1} \frac{\partial^2_\mu \xi_{1/n}(X_{t_i^n}, \mu)}{a(X_{t_i^n}, \widehat{\sigma}_n^0)} h_\beta(z_n^i(\mu, \widehat{\sigma}_n^0, \widehat{\beta}_n^0)) - n^{2/\widehat{\beta}_n^0} \sum_{i=0}^{n-1} \frac{(\partial_\mu \xi_{1/n}(X_{t_i^n}, \mu))^2}{a(X_{t_i^n}, \widehat{\sigma}_n^0)^2} \partial_z h_\beta(z_n^i(\mu, \widehat{\sigma}_n^0, \widehat{\beta}_n^0)).
\end{aligned}$$

We rewrite $\partial_\mu G_n(\mu) - \partial_\mu G(\mu) = A_n(\mu) + B_n(\mu)$, where

$$\begin{aligned}
A_n(\mu) =& n \sum_{i=0}^{n-1} \frac{(\partial_\mu \xi_{1/n}(X_{t_i^n}, \mu))^2}{a(X_{t_i^n}, \widehat{\sigma}_n^0)^2} \partial_z h_\beta(z_n^i(\mu, \widehat{\sigma}_n^0, \widehat{\beta}_n^0)) - \int_0^1 \frac{\partial_\mu b(X_s, \mu)^2}{a(X_s, \sigma_0)^2} ds \, \mathbb{E}(h^2_{\beta_0}(J_1)), \\
B_n(\mu) =& - n^{1-1/\widehat{\beta}_n^0} \sum_{i=0}^{n-1} \frac{\partial^2_\mu \xi_{1/n}(X_{t_i^n}, \mu)}{a(X_{t_i^n}, \widehat{\sigma}_n^0)} h_\beta(z_n^i(\mu, \widehat{\sigma}_n^0, \widehat{\beta}_n^0)).
\end{aligned}$$

Our aim is to prove that $\sup_{\mu \in A} |A_n(\mu)| \xrightarrow{\mathbb{P}} 0$ and $\sup_{\mu \in A} |B_n(\mu)| \xrightarrow{\mathbb{P}} 0$. To do so, we first prove that

$$\sup_{\mu \in A} |\partial_\mu \xi_{1/n}(x, \mu) - \frac{1}{n} \partial_\mu b(x, \mu)| \leq C(1 + |x|^p)/n^2,$$

$$\sup_{\mu \in A} |\partial^2_\mu \xi_{1/n}(x, \mu) - \frac{1}{n} \partial^2_\mu b(x, \mu)| \leq C(1 + |x|^p)/n^2.$$

Indeed, under our assumptions on the regularity of the coefficient function $b(.)$, the proof is classic and can be obtained by Gronwall's lemma, similarly to the proof of Proposition 5.1. Thanks to these results,

we can replace $\partial_\mu^k \xi_{1/n}(x,\mu)$ by $\frac{1}{n}\partial_\mu^k b(x,\mu)$ for $k \in \{1,2\}$ in the expressions of $A_n(\mu)$ and $B_n(\mu)$, the error for this replacement is negligible.
For $A_n(\mu)$: We rewrite this term as the sum of the two following terms

$$A_{n,1}(\mu) = \frac{1}{n}\sum_{i=0}^{n-1} \frac{(\partial_\mu b(X_{t_i^n},\mu))^2}{a(X_{t_i^n},\widehat{\sigma}_n^0)^2}(\partial_z h_\beta(z_n^i(\mu,\widehat{\sigma}_n^0,\widehat{\beta}_n^0)) - \partial_z h_{\beta_0}(z_n^i(\theta_0))),$$

$$A_{n,2}(\mu) = \frac{1}{n}\sum_{i=0}^{n-1} \frac{(\partial_\mu b(X_{t_i^n},\mu))^2}{a(X_{t_i^n},\widehat{\sigma}_n^0)^2}\partial_z h_{\beta_0}(z_n^i(\theta_0)) - \int_0^1 \frac{\partial_\mu b(X_s,\mu)^2}{a(X_s,\sigma_0)^2} ds\, \mathbb{E}(h_{\beta_0}^2(J_1)).$$

The convergence to zero of $\sup_{\mu\in A}|A_{n,2}(\mu)|$ can be deduced using the assumptions on the regularity of functions a and b and similar arguments as for (5.29) and (5.31). Here, we only take into account the term $A_{n,1}(\mu)$. For this, we separate the difference into two parts: $\partial_z h_\beta(z_n^i(\mu,\widehat{\sigma}_n^0,\widehat{\beta}_n^0)) - \partial_z h_{\beta_0}(z_n^i(\mu,\widehat{\sigma}_n^0,\widehat{\beta}_n^0))$ and $\partial_z h_{\beta_0}(z_n^i(\mu,\widehat{\sigma}_n^0,\widehat{\beta}_n^0)) - \partial_z h_{\beta_0}(z_n^i(\theta_0))$. First, since $\partial_\beta \partial_z h_\beta(z)$ is bounded for any values of β and z large enough, by intermediate value theorem, we have

$$|\partial_z h_\beta(z_n^i(\mu,\widehat{\sigma}_n^0,\widehat{\beta}_n^0)) - \partial_z h_{\beta_0}(z_n^i(\mu,\widehat{\sigma}_n^0,\widehat{\beta}_n^0))| \le C\frac{\log n}{\sqrt{n}}. \tag{5.35}$$

Second, since $\partial_z^2 h_\beta(z)$ is bounded for any values of β and z, again by intermediate value theorem,

$$z_n^i(\mu,\widehat{\sigma}_n^0,\widehat{\beta}_n^0) - z_n^i(\theta_0) = \left(n^{1/\widehat{\beta}_n^0 - 1/\beta_0}\frac{a(X_{t_i^n},\sigma_0)}{a(X_{t_i^n},\widehat{\sigma}_n^0)} - 1\right) z_n^i(\theta_0) + n^{1/\widehat{\beta}_n^0}\frac{\xi_{1/n}(X_{t_i^n},\mu_0) - \xi_{1/n}(X_{t_i^n},\mu)}{a(X_{t_i^n},\widehat{\sigma}_n^0)}$$

and assumption (A), we have

$$\begin{aligned}
&|\partial_z h_{\beta_0}(z_n^i(\mu,\widehat{\sigma}_n^0,\widehat{\beta}_n^0)) - \partial_z h_{\beta_0}(z_n^i(\theta_0))| \le C|z_n^i(\mu,\widehat{\sigma}_n^0,\widehat{\beta}_n^0) - z_n^i(\theta_0)| \\
&\le C\left(\left|n^{1/\widehat{\beta}_n^0 - 1/\beta_0}\frac{a(X_{t_i^n},\sigma_0)}{a(X_{t_i^n},\widehat{\sigma}_n^0)} - 1\right||z_n^i(\theta_0)| + n^{1/\widehat{\beta}_n^0}\left|\frac{\xi_{1/n}(X_{t_i^n},\mu_0) - \xi_{1/n}(X_{t_i^n},\mu)}{a(X_{t_i^n},\widehat{\sigma}_n^0)}\right|\right) \\
&\le \frac{C}{a(X_{t_i^n},\widehat{\sigma}_n^0)}\left(\left|(n^{1/\widehat{\beta}_n^0 - 1/\beta_0} - 1)a(X_{t_i^n},\sigma_0) + a(X_{t_i^n},\sigma_0) - a(X_{t_i^n},\widehat{\sigma}_n^0)\right||z_n^i(\theta_0)|\right. \\
&\qquad \left. + n^{1/\widehat{\beta}_n^0}\left|\frac{\xi_{1/n}(X_{t_i^n},\mu_0) - \xi_{1/n}(X_{t_i^n},\mu)}{a(X_{t_i^n},\widehat{\sigma}_n^0)}\right|\right) \\
&\le C(1 + |X_{t_i^n}|^p)\left(\frac{(\log n)^2}{\sqrt{n}}|z_n^i(\theta_0)| + n^{1/\widehat{\beta}_n^0 - 1}\right),
\end{aligned} \tag{5.36}$$

for some $p > 0$. The last inequality is obtained by $n^{1/\widehat{\beta}_n^0 - 1/\beta_0} - 1 \sim \frac{|\widehat{\beta}_n^0 - \beta_0|}{\beta_0\widehat{\beta}_n^0}$, the assumption (A), the intermediate value theorem using the Lipchitz property of a and (5.19). Therefore, under our assumptions on the regularity of a and b, using the techniques in Jacod and Protter (2012) for triangular array combined with the fact that $\mathbb{E}_{\theta_0}(|z_n^i(\theta_0)|^\delta|\mathcal{F}_{t_i^n}) \le C$ for $\delta < \beta_0$ (similar arguments as for (5.29)), the convergence to zero of $\sup_{\mu\in A}|A_{n,1}(\mu)|$ is guaranteed from (5.35) and (5.36).
For $B_n(\mu)$: Since $h_\beta(z)$ is bounded for any values of β and z big enough, from our assumptions on the

regularity of the coefficient functions, we have

$$\sup_{\mu\in A}|B_n(\mu)| \le Cn^{-1-1/\widehat{\beta}_n^0}\sum_{i=0}^{n-1}(1+|X_{t_i^n}|^{p'}) \xrightarrow{\mathbb{P}} 0.$$

Thus, we obtain (ii). Considering the assertion (i), by the definition of G, we see immediately that $G(\mu_0)=0$. In addition, μ_0 is the unique solution of $G(\mu)=0$ thanks to the assumption non degeneracy that there exists $s\in[0,1]$ such that $\partial_\mu b(X_s,\mu)\neq 0$. Now, it rests to prove that $G_n(\mu_0)\xrightarrow{\mathbb{P}}0$. To do so, by Taylor's expansion, we have

$$\begin{aligned} G_n(\mu_0) &= n^{1-2/\beta_0}\partial_\mu \ell_n(\mu_0,\sigma_0,\beta_0) \\ &+ n^{1-2/\beta_0}\int_0^1 \begin{pmatrix} \mathcal{I}_n^{1,2}(\mu_0,\sigma_0+t(\widehat{\sigma}_n^0-\sigma_0),\beta_0+t(\widehat{\beta}_n^0-\beta_0)) \\ \mathcal{I}_n^{1,3}(\mu_0,\sigma_0+t(\widehat{\sigma}_n^0-\sigma_0),\beta_0+t(\widehat{\beta}_n^0-\beta_0)) \end{pmatrix}^\top \begin{pmatrix} \widehat{\sigma}_n^0-\sigma_0 \\ \widehat{\beta}_n^0-\beta_0 \end{pmatrix} dt. \end{aligned} \tag{5.37}$$

From this, on the one hand, from (Clément and Gloter, 2020, section 3.2.1.), $n^{1-2/\beta_0}\partial_\mu\ell_n(\mu_0,\sigma_0,\beta_0)$ is tight and converges to zero in probability. On the other hand, from (Clément and Gloter, 2020, section 3.2.2.), we have

$$\sup_{(\sigma,\beta)\in W_n^{(\eta)}}\left\|\frac{1}{n^{1/\beta}\log n}\begin{pmatrix}\mathcal{I}_n^{1,2}(\mu_0,\sigma,\beta)\\ \mathcal{I}_n^{1,3}(\mu_0,\sigma,\beta)\end{pmatrix}\right\|$$

is tight and since

$$\begin{aligned}\sup_{(\sigma,\beta)\in W_n^{(\eta)}}\left\|\frac{n^{1/\beta}\log n}{n^{2/\beta-1}}\begin{pmatrix}\sigma-\sigma_0\\ \beta-\beta_0\end{pmatrix}\right\| &\le C\,\frac{(\log n)^2}{n^{1/\beta_0-1/2}}\sup_{(\sigma,\beta)\in W_n^{(\eta)}} n^{\frac{1}{\beta_0}-\frac{1}{\beta}} \\ &\le C\,\frac{(\log n)^2}{n^{1/\beta_0-1/2}}\exp\left((\log n)^2/\sqrt{n}\right)\underset{n\to\infty}{\longrightarrow}0,\end{aligned}$$

the second term in the Taylor's expansion above converges to zero.

Now, for the tightness of $\frac{n^{1/\beta_0-1/2}}{(\log n)^2}(\widehat{\mu}_n^0-\mu_0)$, we proceed as follows. Since $G_n(\widehat{\mu}_n^0)=0$, by Taylor's expansion, we get

$$G_n(\mu_0) = -n^{1-2/\widehat{\beta}_n^0}\int_0^1 \partial_\mu^2\ell_n(\mu_0+t(\widehat{\mu}_n^0-\mu_0),\widehat{\sigma}_n^0,\widehat{\beta}_n^0)dt\ (\widehat{\mu}_n^0-\mu_0).$$

Then, reusing the decomposition of $G_n(\mu_0)$ in (5.37) and its arguments for its convergence above, combined with μ_0-consistency of $\widehat{\mu}_n^0$ proven above and the fact that $n^{1-2/\beta_0}\partial_\mu^2\ell_n(\mu_0+t(\widehat{\mu}_n^0-\mu_0),\widehat{\sigma}_n^0,\widehat{\beta}_n^0)$ converges to non-singular $\mathcal{I}^{1,1}(\theta_0)$ in probability uniformly from (Clément and Gloter, 2020, section 3.2.2.), the proof is completed. □