Electronic Journal of Statistics
Vol. 14 (2020) 2922–2956
ISSN: 1935-7524
https://doi.org/10.1214/20-EJS1737

Joint estimation for SDE driven by locally stable Lévy processes

Emmanuelle Clément[*,†1] **and Arnaud Gloter**[2]

[1] *CY Cergy Paris Université*
Laboratoire AGM, UMR 8088
F-95000 Cergy, France
e-mail: emmanuelle.clement@univ-eiffel.fr

[2] *LaMME, Université d'Evry, CNRS*
Université Paris-Saclay
91025 Evry, France
e-mail: arnaud.gloter@univ-evry.fr

Abstract: Considering a class of stochastic differential equations driven by a locally stable process, we address the joint parametric estimation, based on high frequency observations of the process on a fixed time interval, of the drift coefficient, the scale coefficient and the jump activity of the process. Extending the methodology proposed in [6], where the jump activity was assumed to be known, we obtain two different rates of convergence in estimating simultaneously the scale parameter and the jump activity, depending on the scale coefficient. If the scale coefficient is multiplicative: $a(x, \sigma) = \sigma \overline{a}(x)$, the joint estimation of the scale coefficient and the jump activity behaves as for the translated stable process studied in [5] and the rate of convergence of our estimators is non diagonal. In the non multiplicative case, the results are different and we obtain a diagonal and faster rate of convergence which coincides with the one obtained in estimating marginally each parameter. In both cases, the estimation method is illustrated by numerical simulations showing that our estimators are rather easy to implement.

MSC2020 subject classifications: Primary 60G51, 60G52, 60J75, 62F12; secondary 60H07, 60F05.
Keywords and phrases: Lévy process, stable process, stochastic differential equation, parametric inference, estimating functions.

Received September 2019.

1. Introduction

In this paper, we consider a class of stochastic differential equations driven by a symmetric locally α-stable process

$$X_t = x_0 + \int_0^t b(X_s, \theta) ds + \int_0^t a(X_{s-}, \sigma) dL_s^\alpha,$$

and we study the joint estimation of (θ, σ, α) based on high-frequency observations of the process on the time interval $[0, T]$ with T fixed (without restriction

*Corresponding author.
†This research is supported by the Paris Seine Initiative.

we will next assume that $T = 1$). In recent years, there has been growing interest in modeling with pure-jump Lévy processes (see for example Jing et al. [13] and [17]) and estimation of such processes is of particular interest.

A large literature is devoted to parametric estimation of jump-diffusions from high-frequency observations and we know that, due to the Brownian component, the estimation of the drift coefficient is not possible without assuming that T goes to infinity. For pure-jump processes, assuming that the jump activity $\alpha \in (0, 2)$, the situation is completely different and we can estimate all the parameters on a fixed time interval. When X is a Lévy process, the first results in that direction have been established among others by Aït-Sahalia and Jacod [1] [2], Kawai and Masuda [14] [16], Masuda [18], Ivanenko, Kulik and Masuda [10] and more recently by Brouste and Masuda [5]. Concerning the parametric estimation of pure-jump driven stochastic equations the literature is less abundant and only partial results are available. The estimation of (θ, σ) is performed by Masuda in [19], assuming that α is known and with the restriction $\alpha \in [1, 2)$. The estimation method proposed in [19] is based on an approximation (for small h) of the distribution of the normalized increment $h^{-1/\alpha}(X_{t+h} - X_t - hb(X_t, \theta))/a(X_t, \sigma)$ by the α-stable distribution. However this approximation is not relevant if $\alpha < 1$. To solve this problem, Clément and Gloter [6] consider the following modified increment $h^{-1/\alpha}(X_{t+h} - \xi_h^{X_t}(\theta))/a(X_t, \sigma)$, where $(\xi_t^x(\theta))_{t \geq 0}$ solves the ordinary equation

$$\xi_t^x(\theta) = x + \int_0^t b(\xi_s^{x_0}(\theta), \theta) ds, \quad t \geq 0.$$

This permits to estimate (θ, σ), for $\alpha \in (0, 2)$ known. Turning to the efficiency of these estimation methods, the LAMN property is established in Clément and al. [7] for the estimation of (θ, σ) assuming that the scale coefficient a is constant and that $(L_t^\alpha)_t$ is a truncated stable process.

In this paper, we perform the joint estimation of the three parameters (θ, σ, α) assuming that $\alpha \in (0, 2)$. Our methodology follows the ideas of [6] and is based on estimating functions (we refer to Sørensen [22] and to the recent survey by Jacod and Sørensen [12] for asymptotics in estimating function methods). Let us recall brieflty the methodology developed in [6]. Observing that the conditional distribution of $h^{-1/\alpha}(X_{t+h} - \xi_h^{X_t}(\theta))/a(X_t, \sigma)$ is close to the α-stable distribution (this is estimated in total variation distance in [6]) the idea is to approximate the transition density $p_h(x, y)$ of the process $(X_t)_t$ by

$$\frac{h^{-1/\alpha}}{a(x, \sigma)} \varphi_\alpha \left(h^{-1/\alpha} \frac{(y - \xi_h^x(\theta))}{a(x, \sigma)} \right),$$

where φ_α is the density of a symmetric α-stable variable S_1^α. This approximation permits to construct a quasi-likelihood function and then a natural choice of estimating function is to consider the associated score function. In the present paper, the additional estimation of the jump activity α requires extensions to non bounded functions of total variation distance estimates and limit theorems established in [6], to prove the asymptotic properties of our estimators. We stress on the fact that these asymptotic properties are established without restriction on the jump activity α.

The estimation of θ achieves the optimal rate and the information established in [7] for a simplified stochastic equation but the rate of convergence and the asymptotic variance-covariance matrix in estimating (σ, α) depend on the function a. To take into account this new phenomenon, we distinguish between two cases.

If the function a is multiplicative (multiplicative case), $a(x, \sigma) = \sigma\overline{a}(x)$, then we show that the rate of convergence is non diagonal and we compute the asymptotic variance of the estimator. This case extends the previous results established respectively in [18] and [5] for a translated α-stable process, where it is shown that the Fisher information matrix is singular in estimating (σ, α) with a diagonal norming rate, but that the LAN property holds with a non singular information matrix using a non diagonal norming rate. Furthermore, we can conjecture that in the multiplicative case our estimator is efficient since the asymptotic variance in estimating (σ, α) is the inverse of the information matrix appearing in the LAN property established in [5] for the translated α-stable process. A consequence of the non diagonal rate is that the asymptotic errors in estimating σ and α jointly are proportional, which is supported also by our numerical simulations.

On the other hand, if the scale coefficient a does not separate σ and x (non multiplicative case), $s \to \frac{\partial_\sigma a}{a}(X_s, \sigma_0)$ is almost surely non constant, the result is new and surprising. Indeed our estimator is asymptotically mixed normal with a diagonal norming rate, faster than in the multiplicative case. Moreover, this rate achieves the optimal rate of convergence in estimating marginally σ and α. Especially this shows that, contrarily to the multiplicative case, the rate in estimating jointly (θ, σ) and α coincides with the one obtained assuming that α is known. Remark that the efficiency in the non multiplicative case is still an open problem since the LAMN property is not yet established for a non constant scale coefficient a.

The paper is organized as follows. Section 2 introduces the notation and assumptions. In Section 3 we state our main results: estimation method and asymptotic properties of the estimators. The main limit theorems to prove consistency and asymptotic mixed normality of our estimators are established in Section 4. Section 5 contains some simulation results that illustrate the asymptotic properties of the estimators.

2. Notation and assumptions

We consider the class of stochastic one-dimensional equations:

$$X_t = x_0 + \int_0^t b(X_s, \theta)ds + \int_0^t a(X_{s-}, \sigma)dL_s^\alpha \tag{2.1}$$

where (L_t^α) is a pure-jump locally α-stable process defined on a filtered space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,1]}, \mathbb{P})$. To simplify the notation we assume that θ, σ are real parameters. We observe the discrete time process $(X_{t_i})_{0\leq i\leq n}$ with $t_i = i/n$, for

$i = 0, \dots, n$ that solves (2.1) for the parameter value $\beta_0 = (\theta_0, \sigma_0, \alpha_0)$ and our aim is to estimate the parameter β_0.

We make some regularity assumptions on the coefficients a and b that ensure in particular that (2.1) admits a unique strong solution. We also specify the behavior of the Lévy measure near zero of the process $(L_t^\alpha)_{t\in[0,1]}$.

H1(Regularity): (a) Let $V_{\theta_0} \times V_{\sigma_0}$ be a neighborhood of (θ_0, σ_0). We assume that $x \mapsto a(x, \sigma_0)$ is $\mathcal{C}^2$ on $\mathbb{R}$, b is $\mathcal{C}^2$ on $\mathbb{R} \times V_{\theta_0}$ and

$$\sup_x \big(\sup_{\theta \in V_{\theta_0}} |\partial_x b(x,\theta)| + |\partial_x a(x,\sigma_0)| \big) \leq C,$$
$$\exists p > 0 \text{ s.t. } |\partial_x^2 b(x,\theta_0)| + |\partial_x^2 a(x,\sigma_0)| \leq C(1+|x|^p),$$
$$a \text{ is non negative and } \exists p \geq 0 \text{ s.t. } \sup_{\sigma \in V_{\sigma_0}} \frac{1}{a(x,\sigma)} \leq C(1+|x|^p),$$

(b) $\forall x \in \mathbb{R}$, $\theta \mapsto b(x,\theta)$ and $\sigma \mapsto a(x,\sigma)$ are $\mathcal{C}^3$

$$\exists p > 0 \text{ s.t. } \sup_{(\theta,\sigma)\in V_{\theta_0}\times V_{\sigma_0}} \max_{1\leq l \leq 3} (|\partial_\theta^l b(x,\theta)| + |\partial_\sigma^l a(x,\sigma)|) \leq C(1+|x|^p),$$
$$\exists p > 0 \text{ s.t. } \sup_{\theta \in V_{\theta_0}} |\partial_x \partial_\theta b(x,\theta)| \leq C(1+|x|^p).$$

H2 (Lévy measure): (a) The Lévy measure of (L_t^α) satisfies

$$\nu(dz) = \frac{g(z)}{|z|^{\alpha+1}} 1_{\mathbb{R}\setminus\{0\}}(z)dz,$$

where $\alpha \in (0,2)$ and $g : \mathbb{R} \mapsto \mathbb{R}$ is a continuous symmetric non negative bounded function with $g(0) = 1$.

(b) g is differentiable on $\{0 < |z| \leq \eta\}$ for some $\eta > 0$ with continuous derivative such that $\sup_{0<|z|\leq\eta} \left|\frac{\partial_z g(z)}{g(z)}\right| < \infty$.

This assumption is satisfied by a large class of processes: α-stable process ($g = 1$), truncated α-stable process ($g = \tau$ a truncation function), tempered stable process ($g(z) = e^{-\lambda|z|}$, $\lambda > 0$).

Remark 2.1. *Our results rely on Theorem 4.1 and Theorem 4.2 in [6], obtained under* **H2***, that give a rate of convergence in total variation distance between respectively the rescaled distributions of $X_{1/n}$ and $L_{1/n}^\alpha$, and the locally α-stable distribution and the stable distribution. The key point is that the rate of convergence ε_n satisfies $\sqrt{n}\varepsilon_n \to 0$. However, as in [3], [10] and [24], we could consider, with some proof modifications (in this paper and in [6]), a more general class of locally stable processes and weaken* **H2***. In particular, our methodology permits to consider ν symmetric admitting the decomposition*

$$\nu(dz) = \frac{g_0(z)}{|z|^{\alpha+1}} 1_{\{0<|z|\leq\eta\}} dz + \nu_1(dz).$$

If ν_1, possibly singular, is supported on $\{|z| > \eta\}$, then due to the localization introduced in Section 4.1 of [6], Theorem 4.1 and Theorem 4.2 remain true. Moreover the result of Proposition 4.1 (in this paper) can be obtained (with a different proof) assuming that $\int_{\{|z|>\eta\}} |z|^\delta \nu_1(dz) < \infty$, for $0 < \delta < \min(1, \alpha)$.

If ν_1 is supported on $\mathbb{R} \setminus 0$, we assume additionally that ν_1 is absolutely continuous for $|z| \leq \eta$ with

$$1_{\{0<|z|\leq\eta\}}\nu_1(dz)/dz = 1_{\{0<|z|\leq\eta\}}g_1(z)/|z|^{\beta+1}, \quad 0 \leq \beta < \alpha,$$

where g_0 and g_1 are continuously differentiable on $\{|z| \leq \eta\}$ and $g_0(0) = 1$. Then setting $g(z) = g_0(z) + g_1(z)|z|^{\alpha-\beta}$, we have

$$1_{\{0<|z|\leq\eta\}}\nu(dz) = 1_{\{0<|z|\leq\eta\}}g(z)/|z|^{\alpha+1}.$$

One can check that **H2(b)** *is not satisfied for this function g since $\partial_z g$ is not bounded on $\{|z| \leq \eta\}$. But it can be proven that the result of Theorem 4.1 in [6] remains true under the weaker assumption $z \mapsto z\partial_z g(z)$ bounded, which is satisfied by g defined above. Turning to the result of Theorem 4.2 in [6] (established under the condition $g(z) = 1 + O(|z|)$), we can obtain (with a different proof) the slower rate of convergence $\varepsilon_n = \min(n^{-1/\alpha}, n^{-(\alpha-\beta)/\alpha})$ if $g(z) = 1 + O(|z|) + O(|z|^{\alpha-\beta})$ and $0 < \beta < \alpha$. Consequently to ensure the convergence $\sqrt{n}\varepsilon_n \to 0$, we need the additional restriction $\beta < \alpha/2$.*

The rate of convergence and the information in the joint estimation of $(\theta_0, \sigma_0, \alpha_0)$ depend crucially on the function a and we will prove that if a separates the parameter σ (multiplicative case), the rate of convergence is not diagonal.

NDNM (non degeneracy in the non multiplicative case): $s \to \frac{\partial_\sigma a}{a}(X_s, \sigma_0)$ is almost surely non constant. Almost surely, $\exists t_1 \in (0, 1)$, such that $\partial_\theta b(X_{t_1}, \theta_0) \neq 0$, where $(X_t)_{t\in[0,1]}$ solves (2.1) for the parameter value β_0.

NDM (non degeneracy in the multiplicative case): $a(x, \sigma) = \sigma\overline{a}(x)$. Almost surely, $\exists t_1 \in (0, 1)$, such that $\partial_\theta b(X_{t_1}, \theta_0) \neq 0$, where $(X_t)_{t\in[0,1]}$ solves (2.1) for the parameter value β_0.

We observe that in the multiplicative case the assumptions H1 can be written simply in terms of the function $\overline{a}$ as soon as $\sigma_0 > 0$.

To estimate the parameter $\beta_0 = (\theta_0, \sigma_0, \alpha_0)$, we extend the methodology proposed in [6] based on estimating equations (see also [22]). Considering $X_{1/n}$ solution of (2.1) (with $\beta = (\theta, \sigma, \alpha)$) and introducing the ordinary differential equation

$$\xi_t^{x_0}(\theta) = x_0 + \int_0^t b(\xi_s^{x_0}(\theta), \theta)ds, \quad t \in [0, 1], \tag{2.2}$$

it is proved in [6] (combining Theorem 4.1 and Theorem 4.2) that $n^{1/\alpha}(X_{1/n} - \xi_{1/n}^{x_0}(\theta))/a(x_0, \sigma)$ converges in total variation distance to S_1^α, a stable random variable with characteristic function $e^{-C(\alpha)|u|^\alpha}$. Thus if $X_{1/n}$ admits a density, denoted by $p_{1/n}(x_0, y, \beta)$, then $p_{1/n}$ converges in L^1-norm to

$$\frac{n^{1/\alpha}}{a(x_0,\sigma)}\varphi_\alpha\left(n^{1/\alpha}\frac{(y-\xi^{x_0}_{1/n}(\theta))}{a(x_0,\sigma)}\right)$$

where φ_α is the density of S^α_1. We mention that the existence of the density $p_{1/n}$ is established under stronger assumptions on the Lévy measure (essentially integrability conditions for the large jumps part), see for example [4] or [9], but is not required in our method. So to estimate β, the previous convergence suggests to consider the following approximation of the likelihood function

$$\log L_n(\theta,\sigma,\alpha) = \sum_{i=1}^n \log\left(\frac{n^{1/\alpha}}{a(X_{\frac{i-1}{n}},\sigma)}\varphi_\alpha(z_n(X_{\frac{i-1}{n}}, X_{\frac{i}{n}}, \theta, \sigma, \alpha))\right) \tag{2.3}$$

where

$$z_n(x,y,\theta,\sigma,\alpha) = z_n(x,y,\beta) = n^{1/\alpha}\frac{(y-\xi^x_{1/n}(\theta))}{a(x,\sigma)}. \tag{2.4}$$

Note that φ_α can be computed numerically (see for example [21]). A natural choice of estimating functions is therefore the score function. This leads to the following functions

$$G_n(\beta) = \begin{pmatrix} G^1_n(\beta) \\ G^2_n(\beta) \\ G^3_n(\beta) \end{pmatrix} = -\nabla_\beta \log L_n(\theta,\sigma,\alpha), \tag{2.5}$$

with for $k = 1, 2, 3$

$$G^k_n(\beta) = \sum_{i=1}^n g^k\left(X_{\frac{i-1}{n}}, X_{\frac{i}{n}}, \beta\right),$$

$$g^1(x,y,\beta) = n^{1/\alpha}\frac{\partial_\theta \xi^x_{1/n}(\theta)}{a(x,\sigma)}\frac{\partial_z\varphi_\alpha}{\varphi_\alpha}(z_n(x,y,\beta)), \tag{2.6}$$

$$g^2(x,y,\beta) = \frac{\partial_\sigma a(x,\sigma)}{a(x,\sigma)}(1+z_n(x,y,\beta)\frac{\partial_z\varphi_\alpha}{\varphi_\alpha}(z_n(x,y,\beta))), \tag{2.7}$$

$$\begin{aligned} g^3(x,y,\beta) = \frac{\log n}{\alpha^2}&(1+z_n(x,y,\beta)\frac{\partial_z\varphi_\alpha}{\varphi_\alpha}(z_n(x,y,\beta))) \\ &-\frac{\partial_\alpha\varphi_\alpha}{\varphi_\alpha}(z_n(x,y,\beta)). \end{aligned} \tag{2.8}$$

Note that to compute the above functions, we used

$$\partial_\theta z_n = -n^{1/\alpha}\frac{\partial_\theta\xi^x_{1/n}(\theta)}{a(x,\sigma)},\ \partial_\sigma z_n = -\frac{\partial_\sigma a}{a}z_n,\ \partial_\alpha z_n = -\frac{\log n}{\alpha^2}z_n.$$

To simplify the notation, we introduce the functions

$$\begin{aligned} h_\alpha(z) &= \partial_z \varphi_\alpha(z)/\varphi_\alpha(z) \\ k_\alpha(z) &= 1 + z h_\alpha(z), \qquad \partial_z k_\alpha(z) = h_\alpha(z) + z\partial_z h_\alpha(z) \\ f_\alpha(z) &= \partial_\alpha \varphi_\alpha(z)/\varphi_\alpha(z). \end{aligned}$$

Note that we have the relation $\partial_\alpha h_\alpha = \partial_z f_\alpha$.

From Dumouchel [8], we know that

$$|\partial_z^{k_1}\partial_\alpha^{k_2}\varphi_\alpha(z)| \leq C\frac{(\log(|z|))^{k_2}}{|z|^{k_1+\alpha+1}},$$

as $|z|$ goes to infinity. This permits to deduce that h_α, $\partial_z h_\alpha$, k_α, $\partial_z k_\alpha$ are bounded on $\mathbb{R}\times(0,2)$ and that for $|z|$ large enough

$$|f_\alpha(z)| \leq C\log|z|, \quad |\partial_\alpha f_\alpha(z)| \leq C(\log|z|)^2.$$

We also observe that $\partial_z f_\alpha$ and $z \mapsto z\partial_z k_\alpha(z)$ are bounded and that $z \mapsto z\partial_\alpha h_\alpha(z)$ is bounded, for $|z|$ large, by $C\log|z|$.

Throughout the paper, we denote by C a generic constant whose value may change from line to line.

3. Joint estimation

3.1. Main results

We estimate β by solving the equation $G_n(\beta) = 0$, where G_n is defined by (2.5) with g^1, g^2 and g^3 given by (2.6), (2.7), (2.8). We prove that the resulting estimator is consistent and asymptotically mixed normal. However the rate of convergence and the asymptotic information matrix depend on the function a. Let us define the matrix rate u_n by

$$u_n = \begin{pmatrix} \frac{1}{n^{1/\alpha_0-1/2}} & 0 \\ 0 & \frac{1}{\sqrt{n}}v_n \end{pmatrix}, \quad v_n = \begin{pmatrix} v_n^{1,1} & v_n^{1,2} \\ v_n^{2,1} & v_n^{2,2} \end{pmatrix},$$

where v_n is specified below, depending on the coefficient a.

Under the assumption NDNM, we obtain a diagonal rate of convergence as stated in the following theorem.

Theorem 3.1. *We assume that assumptions H1, H2 and NDNM hold and that v_n is given by (diagonal rate)*

$$v_n = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{\log n} \end{pmatrix}.$$

Then there exists an estimator $(\hat\theta_n, \hat\sigma_n, \hat\alpha_n)$ solving the equation $G_n(\beta) = 0$ with probability tending to 1, that converges in probability to $(\theta_0, \sigma_0, \alpha_0)$. Moreover *we have the stable convergence in law with respect to* $\sigma(L_s^{\alpha_0}, s \le 1)$

$$u_n^{-1}\begin{pmatrix} \hat{\theta}_n - \theta_0 \\ \hat{\sigma}_n - \sigma_0 \\ \hat{\alpha}_n - \alpha_0 \end{pmatrix} \xrightarrow{\mathcal{L}_s} I(\beta_0)^{-1/2}\mathcal{N},$$

where $\mathcal{N}$ *is a standard Gaussian variable independent of* $I(\beta_0)$ *and*

$$I(\beta_0) = \begin{pmatrix} \int_0^1 \frac{\partial_\theta b(X_s,\theta_0)^2}{a(X_s,\sigma_0)^2} ds \mathbb{E} h_{\alpha_0}^2(S_1^{\alpha_0}) & 0 \\ 0 & I_{\sigma\alpha}(\beta_0) \end{pmatrix} \tag{3.1}$$

with

$$I_{\sigma\alpha}(\beta_0) = \begin{pmatrix} \int_0^1 \frac{\partial_\sigma a(X_s,\sigma_0)^2}{a(X_s,\sigma_0)^2} ds \mathbb{E} k_{\alpha_0}^2(S_1^{\alpha_0}) & \frac{1}{\alpha_0^2}\int_0^1 \frac{\partial_\sigma a(X_s,\sigma_0)}{a(X_s,\sigma_0)} ds \mathbb{E} k_{\alpha_0}^2(S_1^{\alpha_0}) \\ \frac{1}{\alpha_0^2}\int_0^1 \frac{\partial_\sigma a(X_s,\sigma_0)}{a(X_s,\sigma_0)} ds \mathbb{E} k_{\alpha_0}^2(S_1^{\alpha_0}) & \frac{1}{\alpha_0^4}\mathbb{E} k_{\alpha_0}^2(S_1^{\alpha_0}) \end{pmatrix}.$$

Note that the matrix $I(\beta_0)$ is invertible a.s. since from NDNM

$$\frac{1}{\alpha_0^4}\mathbb{E} k_{\alpha_0}^2(S_1^{\alpha_0}) \left(\int_0^1 \frac{\partial_\sigma a(X_s,\sigma_0)^2}{a(X_s,\sigma_0)^2} ds - \left(\int_0^1 \frac{\partial_\sigma a(X_s,\sigma_0)}{a(X_s,\sigma_0)} ds \right)^2 \right) > 0, \quad a.s.$$

Turning to the multiplicative case (assumption NDM), we have the following result.

Theorem 3.2. *We assume that H1, H2 and NDM hold. We assume moreover that*

$$\begin{aligned} v_n^{1,1}\frac{1}{\sigma_0} + v_n^{2,1}\frac{\log n}{\alpha_0^2} &\to \overline{v}^{1,1} \qquad v_n^{1,2}\frac{1}{\sigma_0} + v_n^{2,2}\frac{\log n}{\alpha_0^2} \to \overline{v}^{1,2} \\ v_n^{2,1} &\to \overline{v}^{2,1} \qquad v_n^{2,2} \to \overline{v}^{2,2} \end{aligned} \tag{3.2}$$

and that $\overline{v}^{1,1}\overline{v}^{2,2} - \overline{v}^{1,2}\overline{v}^{2,1} > 0$*. Then there exists an estimator* $(\hat{\theta}_n, \hat{\sigma}_n, \hat{\alpha}_n)$ *solving the equation* $G_n(\beta) = 0$ *with probability tending to 1, that converges in probability to* $(\theta_0, \sigma_0, \alpha_0)$*. Moreover we have the stable convergence in law with respect to* $\sigma(L_s^{\alpha_0}, s \le 1)$

$$u_n^{-1}\begin{pmatrix} \hat{\theta}_n - \theta_0 \\ \hat{\sigma}_n - \sigma_0 \\ \hat{\alpha}_n - \alpha_0 \end{pmatrix} \xrightarrow{\mathcal{L}_s} \overline{I}(\beta_0)^{-1/2}\mathcal{N},$$

where $\mathcal{N}$ *is a standard Gaussian variable independent of* $\overline{I}(\beta_0)$ *and*

$$\overline{I}(\beta_0) = \begin{pmatrix} \int_0^1 \frac{\partial_\theta b(X_s,\theta_0)^2}{a(X_s,\sigma_0)^2} ds \mathbb{E} h_{\alpha_0}^2(S_1^{\alpha_0}) & 0 \\ 0 & \overline{v}^T \overline{I}_{\sigma\alpha}(\beta_0)\overline{v} \end{pmatrix} \tag{3.3}$$

with

$$\overline{v} = \begin{pmatrix} \overline{v}^{1,1} & \overline{v}^{1,2} \\ \overline{v}^{2,1} & \overline{v}^{2,2} \end{pmatrix},$$

$$\overline{I}_{\sigma\alpha}(\beta_0) = \begin{pmatrix} \mathbb{E} k_{\alpha_0}^2(S_1^{\alpha_0}) & -\mathbb{E}(k_{\alpha_0} f_{\alpha_0})(S_1^{\alpha_0}) \\ -\mathbb{E}(k_{\alpha_0} f_{\alpha_0})(S_1^{\alpha_0}) & \mathbb{E} f_{\alpha_0}^2(S_1^{\alpha_0}) \end{pmatrix}.$$

Remark 3.1. *In the particular case of constant coefficients a and b (where assumption NDM holds), our estimator is efficient. Indeed the rate of convergence and the asymptotic Fisher information* $\overline{I}$ *are the one obtained recently by Brouste and Masuda [5], where the LAN property is established from high frequency observations, for the translated* α*-stable process*

$$X_t = \theta t + \sigma S_t^{\alpha}.$$

Remark 3.2. *If we have some additional information on the parameter* α_0*, we can replace the solution to the ordinary equation* (2.2) *by an approximation (see also Proposition 3.1 in [6]). In particular, if* $\alpha_0 \in (2/3, 2)$*, we can check from H1 that* $\sup_{\theta \in V_{\theta_0}} |\xi^x_{1/n}(\theta) - x - b(x,\theta)/n| \leq C(1+|x|)/n^2$ *and consequently setting* $\overline{z}(x,y,\beta) = n^{1/\alpha}(y - x - b(x,\theta)/n)/a(x,\sigma)$*, we deduce that (with* $V_n^{(\eta)}(\beta_0)$ *defined by* (3.4)*)*

$$\sup_{\beta \in V_n^{(\eta)}(\beta_0)} |z_n(x,y,\beta) - \overline{z}_n(x,y,\beta)| \leq C(1+|x|^p)\varepsilon_n,$$

where $n^{1/2}\varepsilon_n$ *goes to zero. This control is sufficient to show that the results of Theorem 3.1 and Theorem 3.2 hold with the estimating functions* $\overline{G}_n(\beta) = -\nabla_\beta \log \overline{L}_n(\beta)$ *where* $\overline{L}_n$ *is the quasi-likelihood function obtained by replacing* z_n *by* $\overline{z}_n$ *in the expression* (2.3).

Remark 3.3. *Since* $I(\beta_0)$ *and* $\overline{I}(\beta_0)$ *are positive definite a.s., we can check that the estimator* $(\hat{\theta}_n, \hat{\sigma}_n, \hat{\alpha}_n)$ *proposed in Theorem 3.1 and Theorem 3.2 is also a local maximum of the quasi-likelihood function* L_n *defined by* (2.3)*, on a set with probability tending to one (see Sweeting [23]).*

For the reader convenience we recall the sufficient conditions established in Sørensen [22] to prove the existence, consistency and asymptotic normality of estimating functions based estimators. To this end, we define the matrix $J_n(\beta_1, \beta_2, \beta_3)$ by

$$J_n(\beta_1,\beta_2,\beta_3) = \sum_{i=1}^{n} \begin{pmatrix} \nabla_\beta g^1(X_{\frac{i-1}{n}}, X_{\frac{i}{n}}, \beta_1)^T \\ \nabla_\beta g^2(X_{\frac{i-1}{n}}, X_{\frac{i}{n}}, \beta_2)^T \\ \nabla_\beta g^3(X_{\frac{i-1}{n}}, X_{\frac{i}{n}}, \beta_3)^T \end{pmatrix}.$$

For $\eta > 0$, we also define

$$V_n^{(\eta)}(\beta_0) = \{(\theta,\sigma,\alpha); ||(u_n)^{-1}(\beta - \beta_0)^T|| \leq \eta\}, \tag{3.4}$$

where $||.||$ is a vector or a matrix norm and A^T is the transpose of the matrix A.

With these notations, Theorem 3.1 and Theorem 3.2 are consequence of the two following conditions:

C1: $\forall \eta > 0$, we have the convergence in probability

$$\sup_{\beta_1,\beta_2,\beta_3 \in V_n^{(\eta)}(\beta_0)} ||u_n^T J_n(\beta_1,\beta_2,\beta_3) u_n - W(\beta_0)|| \to 0,$$

where $W(\beta_0) = I(\beta_0)$ (assumption NDNM) or $W(\beta_0) = \overline{I}(\beta_0)$ (assumption NDM).

C2: $(u_n^T G_n(\beta_0))_n$ stably converges in law to $W(\beta_0)^{1/2}\mathcal{N}$ where $\mathcal{N}$ is a standard Gaussian variable independent of $W(\beta_0)$ and the convergence is stable with respect to the σ-field $\sigma(L_s^{\alpha_0}, s \leq 1)$.

Before starting the proof, we compute explicitly $u_n^T G_n(\beta_0)$ and J_n. This permits to understand how appear the conditions on the matrix v_n depending on the assumptions on a. We have

$$u_n^T G_n(\beta_0) = \begin{pmatrix} \sqrt{n}\sum_{i=1}^n \frac{\partial_\theta \xi^i_{1/n}(\theta_0)}{a(X_{\frac{i-1}{n}},\sigma_0)} h_{\alpha_0}(z_n^i(\beta_0)) \\ \frac{1}{\sqrt{n}}\sum_{i=1}^n \left((v_n^{1,1}\frac{\partial_\sigma a(X_{\frac{i-1}{n}},\sigma_0)}{a(X_{\frac{i-1}{n}},\sigma_0)} + v_n^{2,1}\frac{\log n}{\alpha_0^2}) k_{\alpha_0}(z_n^i(\beta_0)) - v_n^{2,1} f_{\alpha_0}(z_n^i(\beta_0))\right) \\ \frac{1}{\sqrt{n}}\sum_{i=1}^n \left((v_n^{1,2}\frac{\partial_\sigma a(X_{\frac{i-1}{n}},\sigma_0)}{a(X_{\frac{i-1}{n}},\sigma_0)} + v_n^{2,2}\frac{\log n}{\alpha_0^2}) k_{\alpha_0}(z_n^i(\beta_0)) - v_n^{2,2} f_{\alpha_0}(z_n^i(\beta_0))\right) \end{pmatrix}$$

where we have used the short notation

$$z_n^i(\beta_0) = z_n(X_{\frac{i-1}{n}}, X_{\frac{i}{n}}, \beta_0), \tag{3.5}$$

with z_n defined by (2.4) and

$$\xi^i_{1/n}(\theta_0) = \xi^{X_{(i-1)/n}}_{1/n}(\theta_0),$$

with ξ solving (2.2). Using the relation $\partial_\alpha h_\alpha = \partial_z f_\alpha$, we now express each term of the matrix J_n. We have

$$\begin{aligned} J_n^{1,1}(\beta_0) &= n^{1/\alpha_0}\sum_{i=1}^n \frac{\partial^2_\theta \xi^i_{1/n}(\theta_0)}{a(X_{\frac{i-1}{n}},\sigma_0)} h_{\alpha_0}(z_n^i(\beta_0)) \\ &\quad - n^{2/\alpha_0}\sum_{i=1}^n \frac{(\partial_\theta \xi^i_{1/n}(\theta_0))^2}{a(X_{\frac{i-1}{n}},\sigma_0)^2}\partial_z h_{\alpha_0}(z_n^i(\beta_0)) \end{aligned} \tag{3.6}$$

$$J_n^{1,2}(\beta_0) = J_n^{2,1}(\beta_0) = -n^{1/\alpha_0}\sum_{i=1}^n \frac{\partial_\sigma a(X_{\frac{i-1}{n}},\sigma_0)}{a(X_{\frac{i-1}{n}},\sigma_0)^2}\partial_\theta \xi^i_{1/n}(\theta_0)\partial_z k_{\alpha_0}(z_n^i(\beta_0))$$

$$\begin{aligned} J_n^{1,3}(\beta_0) &= J_n^{3,1}(\beta_0) = \\ & n^{1/\alpha_0}\sum_{i=1}^n \frac{\partial_\theta \xi^i_{1/n}(\theta_0)}{a(X_{\frac{i-1}{n}},\sigma_0)}\left[-\frac{\log n}{\alpha_0^2}\partial_z k_{\alpha_0}(z_n^i(\beta_0)) + \partial_z f_{\alpha_0}(z_n^i(\beta_0))\right] \end{aligned}$$

$$\begin{aligned} J_n^{2,2}(\beta_0) &= \sum_{i=1}^n \left[\partial_\sigma\left(\frac{\partial_\sigma a}{a}\right)(X_{\frac{i-1}{n}},\sigma_0) k_{\alpha_0}(z_n^i(\beta_0)) \right. \\ &\quad \left. -(\frac{\partial_\sigma a}{a})^2(X_{\frac{i-1}{n}},\sigma_0) z_n^i(\beta_0)\partial_z k_{\alpha_0}(z_n^i(\beta_0))\right] \end{aligned} \tag{3.7}$$

$$J_n^{3,3}(\beta_0) = -\sum_{i=1}^n \Big[\partial_\alpha f_{\alpha_0}(z_n^i(\beta_0)) - 2\frac{\log n}{\alpha_0^2} z_n^i(\beta_0)\partial_\alpha h_{\alpha_0}(z_n^i(\beta_0)) + 2\frac{\log n}{\alpha_0^3} k_{\alpha_0}(z_n^i(\beta_0)) + \frac{(\log n)^2}{\alpha_0^4} z_n^i(\beta_0)\partial_z k_{\alpha_0}(z_n^i(\beta_0))\Big] \tag{3.8}$$

$$J_n^{2,3}(\beta_0) = J_n^{3,2}(\beta_0) = \sum_{i=1}^n \frac{\partial_\sigma a}{a}(X_{\frac{i-1}{n}}, \sigma_0)\Big[-\frac{\log n}{\alpha_0^2} z_n^i(\beta_0)\partial_z k_{\alpha_0}(z_n^i(\beta_0)) + z_n^i(\beta_0)\partial_\alpha h_{\alpha_0}(z_n^i(\beta_0))\Big]. \tag{3.9}$$

From these computations and using the limit theorems established in Section 4, we can check conditions **C1** and **C2** and proceed to the proof of Theorem 3.1 and Theorem 3.2. We first remark that in the above expressions we can replace $\partial_\theta \xi_{1/n}^x(\theta)$ by $\partial_\theta b(x,\theta)/n$. Indeed from H1 and Gronwall's Lemma we have

$$\sup_{\theta\in V_{\theta_0}} |\partial_\theta \xi_{1/n}^x(\theta) - \frac{1}{n}\partial_\theta b(x,\theta)| \leq C(1+|x|^p)/n^2, \tag{3.10}$$

$$\sup_{\theta\in V_{\theta_0}} |\partial_\theta^2 \xi_{1/n}^x(\theta) - \frac{1}{n}\partial_\theta^2 b(x,\theta)| \leq C(1+|x|^p)/n^2. \tag{3.11}$$

Furthermore, by a standard localization procedure we can assume that a is bounded. Indeed setting $a^K(x,\sigma) = a(x,\sigma)\mathcal{I}_K(a(x,\sigma))$ where $\mathcal{I}_K$ is a smooth real function, equal to 1 on $[-K,K]$ and vanishing outside $[-2K,2K]$, and considering the process X^K solution of (2.1) with coefficients b and a^K, then $X = X^K$ on $\Omega^K = \{\omega\in\Omega; \sup_{0\leq t\leq 1}|a(X_{t^-}(\omega),\sigma_0)| \leq K\}$ and $\mathbb{P}(\Omega^K)\to 1$ as K goes to infinity. Consequently, in the next proof sections, we assume that a is bounded.

3.2. Proof of Theorem 3.1

3.2.1. Condition C2

We recall that $h_{\alpha_0}, k_{\alpha_0}$ are bounded and that f_{α_0} is asymptotically equivalent to the logarithm. Moreover some straightforward computations permit to show that $\mathbb{E}h_{\alpha_0}(S_1^{\alpha_0}) = \mathbb{E}k_{\alpha_0}(S_1^{\alpha_0}) = \mathbb{E}f_{\alpha_0}(S_1^{\alpha_0}) = 0$ and $\mathbb{E}(h_{\alpha_0}k_{\alpha_0})(S_1^{\alpha_0}) = 0$. Therefore from Corollary 4.1, we deduce the convergence in probability

$$\frac{1}{\log n\sqrt{n}}\sum_{i=1}^n f_{\alpha_0}(z_n^i(\beta_0)) \to 0$$

and from Theorem 4.1 we obtain the stable convergence in law

$$\begin{pmatrix} \frac{1}{\sqrt{n}}\sum_{i=1}^n \frac{\partial_\theta b(X_{\frac{i-1}{n}},\theta_0)}{a(X_{\frac{i-1}{n}},\sigma_0)} h_{\alpha_0}(z_n^i(\beta_0)) \\ \frac{1}{\sqrt{n}}\sum_{i=1}^n \frac{\partial_\sigma a(X_{\frac{i-1}{n}},\sigma_0)}{a(X_{\frac{i-1}{n}},\sigma_0)} k_{\alpha_0}(z_n^i(\beta_0)) \\ \frac{1}{\sqrt{n}}\sum_{i=1}^n \left(\frac{1}{\alpha_0^2}k_{\alpha_0}(z_n^i(\beta_0)) - \frac{1}{\log n} f_{\alpha_0}(z_n^i(\beta_0))\right) \end{pmatrix} \xrightarrow{\mathcal{L}_s} I(\beta_0)^{1/2}\mathcal{N},$$

where $I(\beta_0)$ is given by (3.1) and $\mathcal{N}$ is a standard Gaussian variable independent of $I(\beta_0)$.

Now with u_n given by

$$u_n = \begin{pmatrix} \frac{1}{n^{1/\alpha_0 - 1/2}} & 0 & 0 \\ 0 & \frac{1}{n^{1/2}} & 0 \\ 0 & 0 & \frac{1}{n^{1/2}\log n} \end{pmatrix}$$

and using the approximation (3.10) it yields

$$u_n^T G_n(\beta_0) = \begin{pmatrix} \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\partial_\theta b(X_{\frac{i-1}{n}}, \theta_0)}{a(X_{\frac{i-1}{n}}, \sigma_0)} h_{\alpha_0}(z_n^i(\beta_0)) \\ \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\partial_\sigma a(X_{\frac{i-1}{n}}, \sigma_0)}{a(X_{\frac{i-1}{n}}, \sigma_0)} k_{\alpha_0}(z_n^i(\beta_0)) \\ \frac{1}{\sqrt{n}} \sum_{i=1}^n \left(\frac{1}{\alpha_0^2} k_{\alpha_0}(z_n^i(\beta_0)) - \frac{1}{\log n} f_{\alpha_0}(z_n^i(\beta_0)) \right) \end{pmatrix} + o_P(1),$$

and the stable convergence in law of $u_n^T G_n(\beta_0)$ is proved.

3.2.2. Condition C1

We have to check the uniform convergence in probability

$$\sup_{\beta_1, \beta_2, \beta_3 \in V_n^{(\eta)}(\beta_0)} ||u_n^T J_n(\beta_1, \beta_2, \beta_3) u_n - I(\beta_0)|| \to 0,$$

with $V_n^{(\eta)}(\beta_0)$ defined by (3.4) and

$$u_n^T J_n(\beta_1, \beta_2, \beta_3) u_n = \begin{pmatrix} \frac{J_n^{1,1}(\beta_1)}{n^{2/\alpha_0 - 1}} & \frac{J_n^{1,2}(\beta_1)}{n^{1/\alpha_0}} & \frac{J_n^{1,3}(\beta_1)}{n^{1/\alpha_0}\log n} \\ \frac{J_n^{2,1}(\beta_2)}{n^{1/\alpha_0}} & \frac{J_n^{2,2}(\beta_2)}{n} & \frac{J_n^{2,3}(\beta_2)}{n\log n} \\ \frac{J_n^{3,1}(\beta_3)}{n^{1/\alpha_0}\log n} & \frac{J_n^{3,2}(\beta_3)}{n\log n} & \frac{J_n^{3,3}(\beta_3)}{n(\log n)^2} \end{pmatrix}$$

where the coefficients of the matrix J_n are given by (3.6)–(3.9).

After a meticulous study of each term appearing in the matrix $u_n^T J_n(\beta_1, \beta_2, \beta_3)u_n$ and using the approximations (3.10) and (3.11), condition C_1 reduces to prove the following uniform convergence in probability

$$\sup_{\beta \in V_n^{(\eta)}(\beta_0)} |\frac{1}{n}\sum_{i=1}^n f(X_{\frac{i-1}{n}}, \theta, \sigma) g_\alpha(z_n^i(\beta)) - \int_0^1 f(X_s, \theta_0, \sigma_0) ds \mathbb{E} g_{\alpha_0}(S_1^{\alpha_0})| \to 0,$$

$$\sup_{\beta \in V_n^{(\eta)}(\beta_0)} |\frac{1}{n^{1/\alpha_0}}\sum_{i=1}^n f(X_{\frac{i-1}{n}}, \theta, \sigma) g_\alpha(z_n^i(\beta))| \to 0, \quad \text{if} \quad \mathbb{E} g_{\alpha_0}(S_1^{\alpha_0}) = 0,$$

for functions f depending on a, b and their partial derivatives with respect to the parameters θ, σ and g_α belonging to the set of functions $h_\alpha, k_\alpha, \partial_z k_\alpha$,

$\partial_z f_\alpha, \partial_z h_\alpha, z\partial_z k_\alpha, \partial_\alpha h_\alpha, \partial_\alpha f_\alpha, z\partial_\alpha h_\alpha$. These functions satisfy the assumptions of Theorem 4.2. Moreover, using the symmetry of φ_α (φ_α and f_α are even) and the integration by part formula, we can prove

$$\begin{aligned} \mathbb{E}h_\alpha(S_1^\alpha) = \mathbb{E}k_\alpha(S_1^\alpha) = \mathbb{E}\partial_z k_\alpha(S_1^\alpha) = \mathbb{E}\partial_\alpha h_\alpha(S_1^\alpha) = \mathbb{E}\partial_z f_\alpha(S_1^\alpha) = 0 \\ \mathbb{E}\partial_z h_\alpha(S_1^\alpha) = -\mathbb{E}h_\alpha^2(S_1^\alpha) \\ \mathbb{E}S_1^\alpha \partial_z k_\alpha(S_1^\alpha) = -\mathbb{E}k_\alpha^2(S_1^\alpha) \\ \mathbb{E}\partial_\alpha f_\alpha(S_1^\alpha) = -\mathbb{E}f_\alpha^2(S_1^\alpha) \\ \mathbb{E}S_1^\alpha \partial_\alpha h_\alpha(S_1^\alpha) = -\mathbb{E}S_1^\alpha f_\alpha(S_1^\alpha) h_\alpha(S_1^\alpha) = -\mathbb{E}(k_\alpha f_\alpha)(S_1^\alpha). \end{aligned} \tag{3.12}$$

The result follows then from Theorem 4.2 (convergence (4.3) and (4.4)).

3.3. *Proof of Theorem 3.2*

We first observe that from NDM $\partial_\sigma a/a = 1/\sigma$.

3.3.1. *Condition C2*

Since $\mathbb{E}h_{\alpha_0}(S_1^{\alpha_0}) = \mathbb{E}k_{\alpha_0}(S_1^{\alpha_0}) = \mathbb{E}f_{\alpha_0}(S_1^{\alpha_0}) = 0$, we deduce from Theorem 4.1 the stable convergence in law

$$\frac{1}{\sqrt{n}} \begin{pmatrix} 1 & 0 \\ 0 & \overline{v}^T \end{pmatrix} \sum_{i=1}^n \begin{pmatrix} \frac{\partial_\theta b(X_{\frac{i-1}{n}},\theta_0)}{a(X_{\frac{i-1}{n}},\sigma_0)} h_{\alpha_0}(z_n^i(\beta_0)) \\ k_{\alpha_0}(z_n^i(\beta_0)) \\ -f_{\alpha_0}(z_n^i(\beta_0)) \end{pmatrix} \xrightarrow{\mathcal{L}_s} \overline{I}(\beta_0)^{1/2}\mathcal{N},$$

where $\overline{I}(\beta_0)$ is given by (3.3) and $\mathcal{N}$ is a standard Gaussian variable independent of $\overline{I}(\beta_0)$.

Using the approximation (3.10) and the property of v_n (3.2), we deduce

$$u_n^T G_n(\beta_0) = \frac{1}{\sqrt{n}} \begin{pmatrix} 1 & 0 \\ 0 & \overline{v}^T \end{pmatrix} \sum_{i=1}^n \begin{pmatrix} \frac{\partial_\theta b(X_{\frac{i-1}{n}},\theta_0)}{a(X_{\frac{i-1}{n}},\sigma_0)} h_{\alpha_0}(z_n^i(\beta_0)) \\ k_{\alpha_0}(z_n^i(\beta_0)) \\ -f_{\alpha_0}(z_n^i(\beta_0)) \end{pmatrix} + o_P(1),$$

and C2 is proved.

3.3.2. *Condition C1*

We will prove

$$\sup_{\beta_1,\beta_2,\beta_3 \in V_n^{(\eta)}(\beta_0)} ||u_n^T J_n(\beta_1,\beta_2,\beta_3) u_n - \overline{I}(\beta_0)|| \to 0.$$

We have:

$$u_n^T J_n(\beta_1,\beta_2,\beta_3)u_n =$$
$$\begin{pmatrix} \frac{J_n^{1,1}(\beta_1)}{n^{2/\alpha_0-1}} & \frac{1}{n^{1/\alpha_0}}(J_n^{1,2}(\beta_1),J_n^{1,3}(\beta_1))v_n \\ \frac{1}{n^{1/\alpha_0}}v_n^T(J_n^{2,1}(\beta_2),J_n^{3,1}(\beta_3))^T & \frac{1}{n}v_n^T\begin{pmatrix} J_n^{2,2}(\beta_2) & J_n^{2,3}(\beta_2) \\ J_n^{3,2}(\beta_3) & J_n^{3,3}(\beta_3)\end{pmatrix}v_n \end{pmatrix},$$

and using the symmetry of J_n, the proof reduces to the following convergence in probability

$$\sup_{\beta\in V_n^{(\eta)}(\beta_0)} \left|\frac{J_n^{1,1}(\beta)}{n^{2/\alpha_0-1}} - \int_0^1 \frac{\partial_\theta b(X_s,\theta_0)^2}{a(X_s,\sigma_0)^2}ds\mathbb{E}h_{\alpha_0}^2(S_1^{\alpha_0})\right| \to 0, \tag{3.13}$$

$$\sup_{\beta_2,\beta_3\in V_n^{(\eta)}(\beta_0)} \left|\frac{1}{n^{1/\alpha_0}}(J_n^{1,2}(\beta_2),J_n^{1,3}(\beta_3))v_n\right| \to 0, \tag{3.14}$$

$$\sup_{\beta_2,\beta_3\in V_n^{(\eta)}(\beta_0)} \left|\left|\frac{1}{n}v_n^T\begin{pmatrix} J_n^{2,2}(\beta_2) & J_n^{2,3}(\beta_2) \\ J_n^{3,2}(\beta_3) & J_n^{3,3}(\beta_3)\end{pmatrix}v_n - \overline{v}^T\overline{I}_{\sigma\alpha}(\beta_0)\overline{v}\right|\right| \to 0. \tag{3.15}$$

From the expression of J_n given in (3.6)–(3.9) and using the approximations (3.10) and (3.11), convergence (3.13) follows from (4.3) and (4.4) in Theorem 4.2 and (3.14) is a consequence of (4.5) in Theorem 4.2, since the terms of the matrix v_n are bounded by $\log n$. To study the convergence (3.15) we observe that

$$\overline{v} = \begin{pmatrix} \frac{1}{\sigma_0} & \frac{\log n}{\alpha_0^2} \\ 0 & 1 \end{pmatrix} \times v_n + o(1)$$

and consequently we just have to prove

$$\sup_{\beta_2,\beta_3\in V_n^{(\eta)}(\beta_0)} \left|\left|v_n^T\left(\frac{1}{n}\begin{pmatrix} J_n^{2,2}(\beta_2) & J_n^{2,3}(\beta_2) \\ J_n^{3,2}(\beta_3) & J_n^{3,3}(\beta_3)\end{pmatrix} - \overline{J}_n(\beta_0)\right)v_n\right|\right| \to 0 \tag{3.16}$$

where

$$\overline{J}_n(\beta_0) = r_n^T\begin{pmatrix} \mathbb{E}k_{\alpha_0}^2(S_1^{\alpha_0}) & -\mathbb{E}(k_{\alpha_0}f_{\alpha_0})(S_1^{\alpha_0}) \\ -\mathbb{E}(k_{\alpha_0}f_{\alpha_0})(S_1^{\alpha_0}) & \mathbb{E}f_{\alpha_0}^2(S_1^{\alpha_0})\end{pmatrix} r_n,$$

with

$$r_n = \begin{pmatrix} \frac{1}{\sigma_0} & \frac{\log n}{\alpha_0^2} \\ 0 & 1 \end{pmatrix}.$$

To simplify the notation we introduce the following normalized sums:

$$S_n^{(k)}(\beta) = \frac{1}{n}\sum_{i=1}^n k_\alpha(z_n^i(\beta))$$
$$S_n^{(zk')}(\beta) = \frac{1}{n}\sum_{i=1}^n z_n^i(\beta)\partial_z k_\alpha(z_n^i(\beta))$$

$$S_n^{(z\partial h)}(\beta) = \frac{1}{n}\sum_{i=1}^{n} z_n^i(\beta)\partial_\alpha h_\alpha(z_n^i(\beta))$$
$$S_n^{(\partial f)}(\beta) = \frac{1}{n}\sum_{i=1}^{n} \partial_\alpha f_\alpha(z_n^i(\beta)),$$

and from (3.7) (3.8), (3.9) we obtain

$$\frac{1}{n}J_n^{2,2}(\beta) = -\frac{1}{\sigma^2}(S_n^{(k)}(\beta) + S_n^{(zk')}(\beta))$$
$$\frac{1}{n}J_n^{3,2}(\beta) = -\frac{\log n}{\sigma\alpha^2}S_n^{(zk')}(\beta) + \frac{1}{\sigma}S_n^{(z\partial h)}(\beta)$$
$$\frac{1}{n}J_n^{3,3}(\beta) = -\frac{(\log n)^2}{\alpha^4}S_n^{(zk')}(\beta) + 2\frac{\log n}{\alpha^2}S_n^{(z\partial h)}(\beta) - 2\frac{\log n}{\alpha^3}S_n^{(k)}(\beta) - S_n^{(\partial f)}(\beta).$$

A simple computation gives moreover

$$\overline{J}_n(\beta_0) = \begin{pmatrix} \frac{1}{\sigma_0^2}\mathbb{E}k_{\alpha_0}^2(S_1^{\alpha_0}) & \text{sym} \\ \frac{\log n}{\sigma_0\alpha_0^2}\mathbb{E}k_{\alpha_0}^2(S_1^{\alpha_0}) - \frac{1}{\sigma_0}\mathbb{E}(k_{\alpha_0}f_{\alpha_0})(S_1^{\alpha_0}) & \overline{J}_n^{2,2}(\beta_0) \end{pmatrix}.$$

with $\overline{J}_n^{2,2}(\beta_0) = \frac{(\log n)^2}{\alpha_0^4}\mathbb{E}k_{\alpha_0}^2(S_1^{\alpha_0}) - 2\frac{\log n}{\alpha_0^2}\mathbb{E}(k_{\alpha_0}f_{\alpha_0})(S_1^{\alpha_0}) + \mathbb{E}f_{\alpha_0}^2(S_1^{\alpha_0})$. Then using once again that v_n is bounded by $\log n$, (3.16) is proved as soon as we have the following convergence in probability (with $q \le 4$)

$$\sup_{\beta\in V_n^{(\eta)}(\beta_0)} |(\log n)^q S_n^{(k)}(\beta)| \to 0,$$
$$\sup_{\beta\in V_n^{(\eta)}(\beta_0)} (\log n)^q |S_n^{(zk')}(\beta) + \mathbb{E}k_{\alpha_0}^2(S_1^{\alpha_0})| \to 0,$$
$$\sup_{\beta\in V_n^{(\eta)}(\beta_0)} (\log n)^q |S_n^{(z\partial h)}(\beta) + \mathbb{E}(k_{\alpha_0}f_{\alpha_0})(S_1^{\alpha_0})| \to 0,$$
$$\sup_{\beta\in V_n^{(\eta)}(\beta_0)} (\log n)^q |S_n^{(\partial f)}(\beta) + \mathbb{E}f_{\alpha_0}^2(S_1^{\alpha_0})| \to 0.$$

Recalling the equalities (3.12), the above convergence results from (4.5) in Theorem 4.2.

4. Limit theorems

We state in this section some limit theorems (Central Limit Theorem and uniform Law of Large Numbers) that are crucial to obtain the asymptotic properties of our estimators. We follow the approach proposed in [6], extending the results to non bounded functions, with some uniformity with respect to the parameter α.

The next key proposition extends to non bounded functions the control in total variation distance established in [6] (Theorem 4.1 and Theorem 4.2).

Proposition 4.1. *Let f be a real function such that*

$$\forall x \in \mathbb{R} \quad |f(x)| \leq C(1 + (\log(1 + |x|))^q),$$

for some constants $C > 0$ and $q > 0$. Then assuming H1, H2 and a bounded, we have

i)

$$|\mathbb{E}f\left(n^{1/\alpha_0}\frac{(X_{1/n} - \xi^{x_0}_{1/n}(\theta_0))}{a(x_0,\sigma_0)}\right) - \mathbb{E}f(n^{1/\alpha_0}L^{\alpha_0}_{1/n})| \leq C(1 + |x_0|)\varepsilon_n,$$

ii)

$$|\mathbb{E}f(n^{1/\alpha_0}L^{\alpha_0}_{1/n}) - \mathbb{E}f(S^{\alpha_0}_1)| \leq C\varepsilon_n,$$

where $n^{1/2}\varepsilon_n \to 0$ as n goes to infinity.

Proof. We set $z_n = n^{1/\alpha_0}\frac{(X_{1/n} - \xi^{x_0}_{1/n}(\theta_0))}{a(x_0,\sigma_0)}$ and we consider the truncation $f_{K_n} = f1_{\{|x| \leq K_n\}}$. By assumption, $||f_{K_n}||_\infty \leq C(\log K_n)^q$ and from Theorem 4.1 and Theorem 4.2 in [6] we have

$$|\mathbb{E}f_{K_n}(z_n) - \mathbb{E}f_{K_n}(n^{1/\alpha_0}L^{\alpha_0}_{1/n})| \leq C(1 + |x_0|)(\log K_n)^q\tilde{\varepsilon}_n,$$
$$|\mathbb{E}f_{K_n}(n^{1/\alpha_0}L^{\alpha_0}_{1/n}) - \mathbb{E}f_{K_n}(S^{\alpha_0}_1)| \leq C(\log K_n)^q\tilde{\varepsilon}_n,$$

where $\tilde{\varepsilon}_n = 1/n^{1-\varepsilon}$ if $\alpha_0 \leq 1$, for any $\varepsilon \in (0,1)$, and $\tilde{\varepsilon}_n = 1/n^{1/\alpha_0 - \varepsilon}$ if $\alpha_0 > 1$, for any $\varepsilon \in (0, 1/\alpha_0)$. Then, if $K_n = n^p$ for any $p > 0$, we deduce

$$\sqrt{n}(\log K_n)^q\tilde{\varepsilon}_n \to 0.$$

It remains to bound $|\mathbb{E}f(S^{\alpha_0}_1) - \mathbb{E}f_{K_n}(S^{\alpha_0}_1)|$, $|\mathbb{E}f(z_n) - \mathbb{E}f_{K_n}(z_n)|$ and $|\mathbb{E}f(n^{1/\alpha_0}L^{\alpha_0}_{1/n}) - \mathbb{E}f_{K_n}(n^{1/\alpha_0}L^{\alpha_0}_{1/n})|$.

For $\delta < \alpha_0$, we have

$$\begin{aligned}|\mathbb{E}f(S^{\alpha_0}_1) - \mathbb{E}f_{K_n}(S^{\alpha_0}_1)| &= \mathbb{E}|f(S^{\alpha_0}_1)|1_{\{|S^{\alpha_0}_1| > K_n\}} \\ &\leq \frac{C}{K_n^\delta}(1 + \mathbb{E}(\log^q(1 + |S^{\alpha_0}_1|)|S^{\alpha_0}_1|^\delta)) \leq \frac{C}{K_n^\delta},\end{aligned}$$

and we conclude choosing $K_n = n^{1/\delta}$. Turning to the second term and proceeding similarly we just have to check that for $\delta < \alpha_0$

$$\mathbb{E}|n^{1/\alpha_0}(X_{1/n} - \xi^{x_0}_{1/n}(\theta_0))|^\delta \leq C, \tag{4.1}$$

since from H1 a is lower bounded. Using Gronwall's Lemma, we have

$$\sup_{s \leq 1/n}|n^{1/\alpha_0}(X_s - \xi^{x_0}_s(\theta_0))| \leq Cn^{1/\alpha_0}\sup_{s \leq 1/n}|\int_0^s a(X_{u-},\sigma_0)dL^{\alpha_0}_u|,$$

and (4.1) holds if $\mathbb{E} n^{\delta/\alpha_0} \sup_{s \leq 1/n} |\int_0^s a(X_{u-}, \sigma_0) dL_u^{\alpha_0}|^\delta \leq C$. This is obtained by rescaling. Setting $L_t^n = n^{1/\alpha_0} L_{t/n}^{\alpha_0}$ for $t \in [0,1]$ then $(L_t^n)_{t \in [0,1]}$ is a Lévy process with Lévy measure ν^n given by

$$\nu^n(dz) = \frac{1}{|z|^{\alpha_0+1}} g(z/n^{1/\alpha_0}) dz.$$

Considering now $(X_t^n)_{t \in [0,1]}$ that solves the equation

$$X_t^n = x_0 + \frac{1}{n} \int_0^t b(X_s^n, \theta_0) ds + \frac{1}{n^{1/\alpha_0}} \int_0^t a(X_{s-}^n, \sigma_0) dL_s^n,$$

we can check that the processes $(X_{t/n}, n^{1/\alpha_0} L_{t/n}^{\alpha_0})_{t \in [0,1]}$ and $(X_t^n, L_t^n)_{t \in [0,1]}$ have the same law, and (4.1) reduces to prove

$$\mathbb{E} \sup_{s \leq 1} |\int_0^s a(X_{u-}^n, \sigma_0) dL_u^n|^\delta \leq C \tag{4.2}$$

We can split (L_t^n) in two parts (small jumps and large jumps): $L_t^n = L_t^{n,1} + L_t^{n,2}$ with

$$L_t^{n,1} = \int_0^t \int_{\{0<|z|\leq 1\}} z \tilde{\mu^n}(ds, dz),$$
$$L_t^{n,2} = \int_0^t \int_{\{|z|>1\}} z \mu^n(ds, dz),$$

where μ^n and $\tilde{\mu^n}$ are respectively the Poisson measure and the compensated Poisson random measure associated to (L_t^n). Since $2/\delta > 1$, we deduce using successively Hölder's inequality and Burkholder's inequality

$$\begin{aligned}\mathbb{E} \sup_{s \leq 1} |\int_0^s a(X_{u-}^n, \sigma_0) dL_u^{n,1}|^\delta &\leq (\mathbb{E} \sup_{s \leq 1} |\int_0^s a(X_{u-}^n, \sigma_0) dL_u^{n,1}|^2)^{\delta/2} \\ &\leq C(\mathbb{E} \int_0^1 \int_{\{0<|z|\leq 1\}} a^2(X_{u-}^n, \sigma_0) z^2 \nu^n(dz) ds)^{\delta/2} \leq C,\end{aligned}$$

since a is bounded. Considering now the large jumps part and assuming moreover that $\delta < \min(1, \alpha_0)$ we have

$$\mathbb{E} \sup_{s \leq 1} |\int_0^s a(X_{u-}^n, \sigma_0) dL_u^{n,2}|^\delta \leq \mathbb{E} \int_0^1 \int_{\{|z|>1\}} |z|^\delta \mu^n(ds, dz) \leq C,$$

since $\delta < \alpha_0$, and i) follows. Observing that (4.1) implies $\mathbb{E}|n^{1/\alpha_0} L_{1/n}^{\alpha_0}|^\delta \leq C$ (taking $b = 0$ and $a = 1$), we obtain ii). □

From this proposition, we obtain a Central Limit Theorem for non bounded functions.

Theorem 4.1. *We assume H1, H2 and a bounded. Let* $h_i : \mathbb{R} \to \mathbb{R}$, $i = 1, 2, 3$ *be* $\mathcal{C}^1$ *functions such that*

$$\forall i,\ \forall x \in \mathbb{R}, \quad |h_i(x)| + |\partial_x h_i(x)| \leq C(1 + (\log(1 + |x|))^q),$$

for some constants $C > 0$ *and* $q > 0$ *and let* $f_i : \mathbb{R} \to \mathbb{R}$ *be continuous functions. We assume that* $\mathbb{E}h_i(S_1^{\alpha_0}) = 0$ *for* $i = 1, 2, 3$*. Then we have the stable convergence in law with respect to* $\sigma(L_s^{\alpha_0}, s \leq 1)$*:*

$$\frac{1}{n^{1/2}} \sum_{i=1}^{n} \begin{pmatrix} f_1(X_{\frac{i-1}{n}}) h_1(z_n(X_{\frac{i-1}{n}}, X_{\frac{i}{n}}, \beta_0)) \\ f_2(X_{\frac{i-1}{n}}) h_2(z_n(X_{\frac{i-1}{n}}, X_{\frac{i}{n}}, \beta_0)) \\ f_3(X_{\frac{i-1}{n}}) h_3(z_n(X_{\frac{i-1}{n}}, X_{\frac{i}{n}}, \beta_0)) \end{pmatrix} \xrightarrow{\mathcal{L}_s} \Sigma^{1/2}\mathcal{N},$$

where z_n *is defined by* (2.4), $\mathcal{N}$ *is a standard Gaussian variable independent of* Σ *and for* $1 \leq i, j \leq 3$

$$\Sigma_{i,j} = \int_0^1 (f_i f_j)(X_s) ds \ \mathbb{E}(h_i h_j)(S_1^{\alpha}).$$

Proof. Using Proposition 4.1 and following the proof of Corollary 3.1 in [6], we obtain the convergence in probability for $j = 1, 2, 3$

$$\frac{1}{n^{1/2}} \sum_{i=1}^{n} f_j(X_{\frac{i-1}{n}}) \left(h_j(z_n(X_{\frac{i-1}{n}}, X_{\frac{i}{n}}, \beta_0)) - h_j(n^{1/\alpha_0} \Delta L_i) \right) \to 0,$$

where $\Delta L_i = L_{\frac{i}{n}} - L_{\frac{i-1}{n}}$. Now we can extend the proof of Theorem 3.2 in [6] to non bounded functions h_j with logarithmic growth and we obtain the stable convergence in law

$$\frac{1}{n^{1/2}} \sum_{i=1}^{n} \begin{pmatrix} f_1(X_{\frac{i-1}{n}}) h_1(n^{1/\alpha_0} \Delta L_i) \\ f_2(X_{\frac{i-1}{n}}) h_2(n^{1/\alpha_0} \Delta L_i) \\ f_3(X_{\frac{i-1}{n}}) h_3(n^{1/\alpha_0} \Delta L_i) \end{pmatrix} \xrightarrow{\mathcal{L}_s} \Sigma^{1/2}\mathcal{N}. \qquad \square$$

An immediate consequence of Theorem 4.1 is the following convergence in probability.

Corollary 4.1. *We assume H1, H2 and a bounded. Let* $h : \mathbb{R} \to \mathbb{R}$ *be a* $\mathcal{C}^1$ *function such that*

$$\forall x \in \mathbb{R} \quad |h(x)| + |\partial_x h(x)| \leq C(1 + (\log(1 + |x|))^q),$$

for some constants $C > 0$ *and* $q > 0$ *and* $\mathbb{E}h(S_1^{\alpha_0}) = 0$*. Then we have the convergence in probability*

$$\frac{1}{\sqrt{n}\log n} \sum_{i=1}^{n} h(z_n(X_{\frac{i-1}{n}}, X_{\frac{i}{n}}, \theta_0, \sigma_0, \alpha_0)) \to 0,$$

with z_n *defined by* (2.4).

We finally establish some uniform convergence results that extend Theorem 3.1 in [6].

Theorem 4.2. *Assume H1, H2 and a bounded. Let f be a continuous function such that*

$$\sup_{(\theta,\sigma)\in K_0}(|f(x,\theta,\sigma)|+|\partial_\theta f(x,\theta,\sigma)|+|\partial_\sigma f(x,\theta,\sigma)|)\leq C(1+|x|^p),$$

where K_0 is a neighborhood of (θ_0,σ_0) and let $(z,\alpha)\mapsto g_\alpha(z)$ be a $\mathcal{C}^1$ function (with respect to (z,α)) such that $\partial_z g_\alpha$ is bounded (uniformly in α on compact subset of $(0,2)$) and such that

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

Then we have the convergence in probability

$$\sup_{\beta\in V_n^{(\eta)}(\beta_0)}|\frac{1}{n}\sum_{i=1}^n f(X_{\frac{i-1}{n}},\theta,\sigma)g_\alpha(z_n^i(\beta))-\int_0^1 f(X_s,\theta_0,\sigma_0)ds\mathbb{E}g_{\alpha_0}(S_1^{\alpha_0})|\to 0, \tag{4.3}$$

where $V_n^{(\eta)}(\beta_0)$ and $z_n^i(\beta)$ are defined respectively by (3.4) and (3.5).

Moreover if $\mathbb{E}g_{\alpha_0}(S_1^{\alpha_0})=0$, the following convergences in probability hold

$$\sup_{\beta\in V_n^{(\eta)}(\beta_0)}|\frac{1}{n^{1/\alpha_0}}\sum_{i=1}^n f(X_{\frac{i-1}{n}},\theta,\sigma)g_\alpha(z_n^i(\beta))|\to 0, \tag{4.4}$$

$$\sup_{\beta\in V_n^{(\eta)}(\beta_0)}|\frac{(\log n)^q}{n}\sum_{i=1}^n f(X_{\frac{i-1}{n}},\theta,\sigma)g_\alpha(z_n^i(\beta))|\to 0,\quad \forall q>0. \tag{4.5}$$

Before proving this result we remark that for $\beta\in V_n^{(\eta)}(\beta_0)$ we have: $|\theta-\theta_0|\leq \eta/n^{1/\alpha_0-1/2}$, $|\sigma-\sigma_0|\leq C\eta\log(n)/\sqrt{n}$ and $|\alpha-\alpha_0|\leq C\eta\log(n)/\sqrt{n}$.

The proof of Theorem 4.2 relies on Proposition 4.1 and on the following two Lemmas.

Lemma 4.1 (Lemma 4.2 in [6])**.** *Assuming H1 and H2, there exists $p>0$ such that $\forall\varepsilon>0$ and $\forall\delta\in(0,1)$*

$$\mathbb{P}\left(|z_n(x_0,X_{1/n},\beta_0)-n^{1/\alpha_0}L_{1/n}^{\alpha_0}|>\varepsilon\right)\leq\begin{cases} C(\varepsilon)(1+|x_0|^p)\frac{\log n}{n^{\alpha_0}} & \text{if } \alpha_0<1,\\ C(\varepsilon)(1+|x_0|^p)\frac{1}{n^{1-\delta}}, & \text{if } \alpha_0\geq 1,\end{cases}$$

where $C(\varepsilon)$ is a positive constant.

Lemma 4.2. *Assuming H1 and H2, there exists $p,q>0$ such that*

$$\forall\varepsilon>0,\ \mathbb{P}_{|\mathcal{F}_{\frac{i-1}{n}}}\left(\sup_{\beta\in V_n^{(\eta)}(\beta_0)}|z_n^i(\beta)-n^{1/\alpha_0}\Delta L_i|>\varepsilon\right)\leq C(\varepsilon)(1+|X_{\frac{i-1}{n}}|^p)/n^q, \tag{4.6}$$

where $C(\varepsilon)$ is a positive constant, $\Delta L_i=L_{\frac{i}{n}}^{\alpha_0}-L_{\frac{i-1}{n}}^{\alpha_0}$ and $z_n^i(\beta)$ is given in (3.5).

Proof. We have the decomposition

$$z_n^i(\beta) - n^{1/\alpha_0}\Delta L_i = \frac{n^{1/\alpha}}{n^{1/\alpha_0}}\frac{a(X_{\frac{i-1}{n}},\sigma_0)}{a(X_{\frac{i-1}{n}},\sigma)}(z_n^i(\beta_0) - n^{1/\alpha_0}\Delta L_i)$$
$$+(\frac{n^{1/\alpha}}{n^{1/\alpha_0}}\frac{a(X_{\frac{i-1}{n}},\sigma_0)}{a(X_{\frac{i-1}{n}},\sigma)} - 1)n^{1/\alpha_0}\Delta L_i$$
$$+\frac{n^{1/\alpha}}{n^{1/\alpha_0}}\frac{1}{a(X_{\frac{i-1}{n}},\sigma)}n^{1/\alpha_0}(\xi_{1/n}^i(\theta_0) - \xi_{1/n}^i(\theta)).$$

The proof follows then the same lines as the proof of Lemma 5.1 in [6] using Lemma 4.1, $\beta \in V_n(\beta_0)$ and observing that $|\frac{n^{1/\alpha}}{n^{1/\alpha_0}} - 1| \leq C\log(n)^2/\sqrt{n}$. ☐

Proof of Theorem 4.2. We recall the following useful result to prove convergence in probability of triangular arrays (see [11]).

Let (ζ_i^n) be a triangular array such that ζ_i^n is $\mathcal{F}_{\frac{i}{n}}$-measurable then the two following conditions imply the convergence in probability $\sum_{i=1}^n \zeta_i^n \to 0$:

$$\sum_{i=1}^n |\mathbb{E}_{|\mathcal{F}_{\frac{i-1}{n}}}\zeta_i^n| \to 0 \quad \text{in probability,}$$
$$\sum_{i=1}^n \mathbb{E}_{|\mathcal{F}_{\frac{i-1}{n}}}|\zeta_i^n|^2 \to 0 \quad \text{in probability.}$$

From this result, the methodology is similar to the proof of Theorem 3.1 in [6] and we just outline the main steps. The difference with [6] is that α varies and the function g_α is not bounded. Note however that $\partial_z g_\alpha$ is bounded. We check the following convergences in probability, $\forall q > 0$,

$$\sup_{\beta\in V_n^{(\eta)}(\beta_0)} |\frac{(\log n)^q}{n}\sum_{i=1}^n (f(X_{\frac{i-1}{n}},\theta,\sigma) - f(X_{\frac{i-1}{n}},\theta_0,\sigma_0))g_\alpha(z_n^i(\beta))| \to 0, \tag{4.7}$$

$$\sup_{\beta\in V_n^{(\eta)}(\beta_0)} |\frac{(\log n)^q}{n}\sum_{i=1}^n f(X_{\frac{i-1}{n}},\theta_0,\sigma_0)(g_\alpha(z_n^i(\beta)) - g_{\alpha_0}(n^{1/\alpha_0}\Delta L_i))| \to 0, \tag{4.8}$$

$$|\frac{(\log n)^q}{n}\sum_{i=1}^n f(X_{\frac{i-1}{n}},\theta_0,\sigma_0)(g_{\alpha_0}(n^{1/\alpha_0}\Delta L_i) - \mathbb{E}g_{\alpha_0}(S_1^{\alpha_0}))| \to 0, \tag{4.9}$$

$$|\frac{1}{n}\sum_{i=1}^n f(X_{\frac{i-1}{n}},\theta_0,\sigma_0) - \int_0^1 f(X_s,\theta_0,\sigma_0)ds| \to 0. \tag{4.10}$$

The last convergence is immediate, (4.9) is a consequence of Proposition 4.1. We check (4.7) observing that

$$\mathbb{E}_{|\mathcal{F}_{\frac{i-1}{n}}}[\sup_{\beta\in V_n^{(\eta)}(\beta_0)}(\log(1+|z_n^i(\beta)|))^p] \leq C(1+|X_{\frac{i-1}{n}}|^{p'}), \tag{4.11}$$

for some $p' > 0$. Indeed from (4.1) we have

$$\mathbb{E}_{|\mathcal{F}_{\frac{i-1}{n}}} |z_n^i(\beta_0)|^\delta \leq C,$$

for any $\delta < \alpha_0$. Furthermore a straightforward computation gives

$$z_n^i(\beta) - z_n^i(\beta_0) = (\frac{n^{1/\alpha}}{n^{1/\alpha_0}} \frac{a(X_{\frac{i-1}{n}}, \sigma_0)}{a(X_{\frac{i-1}{n}}, \sigma)} - 1) z_n^i(\beta_0)$$
$$+ \frac{n^{1/\alpha}}{n^{1/\alpha_0}} \frac{1}{a(X_{\frac{i-1}{n}}, \sigma)} n^{1/\alpha_0} (\xi_{1/n}^i(\theta_0) - \xi_{1/n}^i(\theta))$$

and then

$$\sup_{\beta \in V_n^{(\eta)}(\beta_0)} |z_n^i(\beta) - z_n^i(\beta_0)| \leq C(1 + |X_{\frac{i-1}{n}}|^p)(1 + |z_n^i(\beta_0)|)(\log n)^2/\sqrt{n}.$$

This permits to deduce (4.11) and at last, (4.8) follows from (4.11) and Lemma 4.2. The convergences (4.7)–(4.10) permit to obtain (4.3) and (4.5) in Theorem 4.2.

To prove (4.4), combining (4.11), Lemma 4.1 and Lemma 4.2, we check the convergences in probability for $\alpha_0 > 1$

$$\sup_{\beta \in V_n^{(\eta)}(\beta_0)} |\frac{1}{n^{1/\alpha_0}} \sum_{i=1}^n f(X_{\frac{i-1}{n}}, \theta, \sigma) 1_{\{|X_{\frac{i-1}{n}}| \leq K\}} [g_\alpha(z_n^i(\beta)) - g_{\alpha_0}(z_n^i(\beta_0))]| \to 0,$$
$$\sup_{\beta \in V_n^{(\eta)}(\beta_0)} |\frac{1}{n^{1/\alpha_0}} \sum_{i=1}^n f(X_{\frac{i-1}{n}}, \theta, \sigma) 1_{\{|X_{\frac{i-1}{n}}| \leq K\}} \Big[g_{\alpha_0}(z_n^i(\beta_0))$$
$$- \mathbb{E}_{|\mathcal{F}_{\frac{i-1}{n}}} g_{\alpha_0}(z_n^i(\beta_0)) \Big] | \to 0,$$
$$\sup_{\beta \in V_n^{(\eta)}(\beta_0)} |\frac{1}{n^{1/\alpha_0}} \sum_{i=1}^n f(X_{\frac{i-1}{n}}, \theta, \sigma) 1_{\{|X_{\frac{i-1}{n}}| \leq K\}} \mathbb{E}_{|\mathcal{F}_{\frac{i-1}{n}}} g_{\alpha_0}(z_n^i(\beta_0))| \to 0.$$

These convergences are obtained with similar computations than the one used in the proof of Theorem 3.1 in [6] and we omit the details. □

5. Numerical simulations

In this section, we make numerical simulations. We aim to show that the joint estimation of the three parameters (θ, σ, α) is feasible in practice in several models. We also want to illustrate that the asymptotic behavior of the estimator is different whether the model satisfies the condition NDM or NDNM, which is the main finding of Section 3.1.

5.1. A multiplicative model driven by an α-stable process

We consider the process $(X_t)_{t \in [0,1]}$ solution of

$$\mathrm{d}X_t = \theta X_t \mathrm{d}t + \sigma \sqrt{1 + X_t^2} \mathrm{d}S_t^\alpha,$$

where $(S_t^\alpha)_t$ is a symmetric α-stable process with characteristic function $u \mapsto e^{-|u|^\alpha}$. Assumption NDM holds true and we can apply results of Section 3.1. As matrix rate, we choose $v_n^{1,1} = \sigma_0$, $v_n^{1,2} = -\frac{\sigma_0}{\alpha_0^2}\log(n)$, $v_n^{2,1} = 0$, $v_n^{2,2} = 1$. This choice is such that (3.2) holds true with $\overline{v} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Let us denote by Z a Gaussian random vector with law $\mathcal{N}(0_{\mathbb{R}^3}, K_{\alpha_0})$ where

$$K_{\alpha_0} = \begin{pmatrix} \mathbb{E}[h_{\alpha_0}^2(S_1^{\alpha_0})] & 0 & 0 \\ 0 & \mathbb{E}[k_{\alpha_0}^2(S_1^{\alpha_0})] & -\mathbb{E}[(k_{\alpha_0} f_{\alpha_0})(S_1^{\alpha_0})] \\ 0 & -\mathbb{E}[(k_{\alpha_0} f_{\alpha_0})(S_1^{\alpha_0})] & \mathbb{E}[f_{\alpha_0}^2(S_1^{\alpha_0})] \end{pmatrix}^{-1}. \tag{5.1}$$

Then, from the stable convergence result of Theorem 3.2, one has the convergence to Z of the vector

$$\begin{pmatrix} (\int_0^1 \frac{\partial_\theta b(X_s,\theta_0)^2}{a(X_s,\sigma_0)^2} ds)^{1/2} n^{1/\alpha_0 - 1/2}(\hat{\theta}_n - \theta_0) \\ \sqrt{n}(\frac{\hat{\sigma}_n - \sigma_0}{\sigma_0}) + \frac{\log(n)}{\alpha_0^2}\sqrt{n}(\hat{\alpha}_n - \alpha_0) \\ \sqrt{n}(\hat{\alpha}_n - \alpha_0) \end{pmatrix}.$$

Thus, the rate of estimation is $n^{1/\alpha_0 - 1/2}$ for θ_0, and $\sqrt{n}$ for α_0. Moreover, we get that $\frac{\sqrt{n}}{\log(n)}(\frac{\hat{\sigma}_n - \sigma_0}{\sigma_0}) + \frac{1}{\alpha_0^2}\sqrt{n}(\hat{\alpha}_n - \alpha_0) \xrightarrow{n\to\infty} 0$. This implies that $\frac{\sqrt{n}}{\log(n)}(\hat{\sigma}_n - \sigma_0) = -\frac{\sigma_0}{\alpha_0^2}\sqrt{n}(\hat{\alpha}_n - \alpha_0) + o_P(1) \xrightarrow{n\to\infty} -\frac{\sigma_0}{\alpha_0^2} Z_3$. Hence, the rate of estimation for σ_0 is $\frac{\sqrt{n}}{\log(n)}$ and asymptotically the estimation errors for the parameters σ_0 and α_0 are proportional and have a correlation tending to -1. Comparing with the situation of non-multiplicative model, addressed in Theorem 3.1, we see that both parameters α_0 and σ_0 are estimated with rate slower by a $\log(n)$ factor in the multiplicative case.

5.1.1. Numerical results

For numerical simulations, we choose $\theta_0 = 0.5$, $\sigma_0 = 1$, $\alpha_0 \in \{0.7, 1.3, 1.7\}$. We let the number of data n ranges in the set $\{128, 256, 512, 1024, 2048\}$. We simulate the process (X_t) with an Euler scheme with step $(1000n)^{-1}$. In Tables 1–3, we give an estimation by Monte-Carlo of the mean of the estimators together with their standard deviations. In these Monte-Carlo experiments, we used 1000 replications.

From Table 1, we see that for $\alpha_0 = 0.7$ the joint estimation of the three parameters works well. Especially, the estimator of the drift parameter performs extremely well for $\alpha_0 = 0.7$, which is expected, since the rate of estimation is $n^{1/0.7-1/2} \simeq n^{0.93}$. For $\alpha_0 = 1.3$ (Table 2), the estimation of σ_0 and α_0 works well while the estimation of θ_0 has some bias which reduces slowly as n increases. For $\alpha_0 = 1.7$, we found that the estimation of the drift parameter θ_0 has both a very large bias and standard deviation. Actually, the convergence of the estimator $\hat{\theta}_n$ occurs with the extremely slow rate $n^{1/1.7-1/2} \simeq n^{0.0882}$, and it seems impossible, in practice, to get a correct estimate of the drift parameter when $\alpha_0 = 1.7$.

On the other hand, we see that the estimation of σ_0 and α_0 works well again. It means that the impossibility to estimate correctly the drift parameter for $\alpha_0 = 1.7$ has no negative impact on the estimation of the other parameters.

In Tables 4–6, we give an estimation of the standard deviation of the error of estimation rescaled in a way that it theoretically converges to a Gaussian variable whose variance can be computed using (5.1). Let us stress that, as the asymptotic law of $\hat{\theta}_n$ is mixed normal, the estimation error $\hat{\theta}_n - \theta_0$ is rescaled by a factor involving the random quantity

$$V_{\theta_0} = \left(\int_0^1 \frac{\partial_\theta b(X_s, \theta_0)^2}{a(X_s, \sigma_0)^2} ds \right)^{-1},$$

that we approximate, in practice, by a Riemann sum based on the simulated observations $(X_{i/n})_{i=0,\dots,n}$. As the entries of the matrix K_{α_0} given in (5.1) are not explicit, the theoretical asymptotic standard deviations for these rescaled errors are computed using numerical integration. These theoretical standard deviations are reported in the last line of each tables 4–6.

In Tables 4–6, we see that the asymptotic behavior of the estimator is exactly as predicted from the theoretical study: the rate of estimation for θ_0, σ_0, and α_0 are exactly $n^{1/\alpha_0 - 1/2}$, $n^{1/2}/\log(n)$ and $n^{1/2}$. Moreover, the asymptotic rescaled standard deviations are close to the theoretical one.

In Figures 1–3, we plot the histograms of the distribution of the rescaled errors of estimation, together with the density of their Gaussian limits. For the sake of shortness, we only plot the results for $n = 2048$ and $\alpha_0 \in \{0.7, 1.3, 1.7\}$. It appears that the empirical distributions are close to their theoretical limits, in all cases.

In Table 7, we display the empirical correlation between the estimators $\hat{\sigma}_n$ and $\hat{\alpha}_n$ for different values of α_0 and n. As expected from the theory, in the multiplicative case this correlation tends to -1 as $n \to \infty$.

Our last numerical experiment in the multiplicative case is related to Remark 3.2, where we state that for $\alpha_0 > 2/3$, one can replace in the contrast function, the quantity $\xi^x_{1/n}(\theta)$ by its one step Euler approximation $\xi^x_{1/n}(\theta) \simeq x + b(x,\theta)/n$. We see, by comparison of Table 8 with Table 1, that the quality of estimation is the same when one uses the approximation of $\xi^x_{1/n}(\theta)$ as when one uses its true value.

TABLE 1
Estimation: Multiplicative case $\alpha_0 = 0.7$

n	Mean $\hat{\theta}_n$	Std $\hat{\theta}_n$	Mean $\hat{\sigma}_n$	Std $\hat{\sigma}_n$	Mean $\hat{\alpha}_n$	Std $\hat{\alpha}_n$
128	0.498	$6.7 * 10^{-2}$	1.184	$8.18 * 10^{-1}$	0.709	$6.76 * 10^{-2}$
256	0.500	$3.30 * 10^{-2}$	1.157	$6.55 * 10^{-1}$	0.702	$4.70 * 10^{-2}$
512	0.500	$1.65 * 10^{-2}$	1.110	$4.75 * 10^{-1}$	0.700	$3.19 * 10^{-2}$
1024	0.500	$7.74 * 10^{-3}$	1.062	$3.70 * 10^{-1}$	0.700	$2.31 * 10^{-2}$
2048	0.500	$4.36 * 10^{-3}$	1.045	$2.75 * 10^{-1}$	0.700	$1.64 * 10^{-2}$

TABLE 2
Estimation: Multiplicative case $\alpha_0 = 1.3$

n	Mean $\hat{\theta}_n$	Std $\hat{\theta}_n$	Mean $\hat{\sigma}_n$	Std $\hat{\sigma}_n$	Mean $\hat{\alpha}_n$	Std $\hat{\alpha}_n$
128	0.443	$6.85 * 10^{-1}$	1.048	$3.66 * 10^{-1}$	1.316	$1.27 * 10^{-1}$
256	0.386	$6.31 * 10^{-1}$	1.031	$3.01 * 10^{-1}$	1.310	$9.35 * 10^{-2}$
512	0.424	$5.77 * 10^{-1}$	1.014	$2.39 * 10^{-1}$	1.306	$6.47 * 10^{-2}$
1024	0.429	$5.06 * 10^{-1}$	1.014	$1.78 * 10^{-1}$	1.302	$4.41 * 10^{-2}$
2048	0.457	$4.26 * 10^{-1}$	1.011	$1.38 * 10^{-1}$	1.300	$3.16 * 10^{-2}$

TABLE 3
Estimation: Multiplicative case $\alpha_0 = 1.7$

n	Mean $\hat{\theta}_n$	Std $\hat{\theta}_n$	Mean $\hat{\sigma}_n$	Std $\hat{\sigma}_n$	Mean $\hat{\alpha}_n$	Std $\hat{\alpha}_n$
128	-0.549	2.46	1.018	$2.20 * 10^{-1}$	1.706	$1.13 * 10^{-1}$
256	-0.493	2.39	1.012	$1.76 * 10^{-1}$	1.704	$9.20 * 10^{-2}$
512	-0.299	2.02	1.008	$1.41 * 10^{-1}$	1.701	$6.67 * 10^{-2}$
1024	-0.299	2.02	1.011	$1.07 * 10^{-1}$	1.700	$4.75 * 10^{-2}$
2048	-0.184	1.92	1.004	$8.06 * 10^{-2}$	1.700	$3.20 * 10^{-2}$

TABLE 4
Std of rescaled errors: Multiplicative case $\alpha_0 = 0.7$

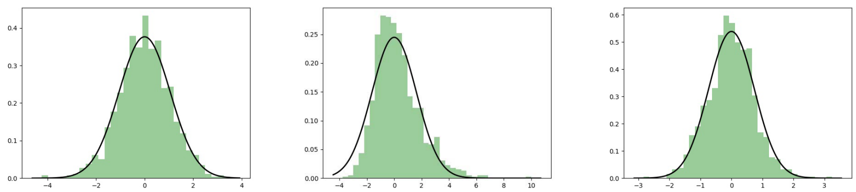
n	$V_{\theta_0}^{-\frac{1}{2}} n^{\frac{1}{\alpha_0}-\frac{1}{2}}(\hat{\theta}_n - \theta_0)$	$\frac{\sqrt{n}}{\log(n)}(\hat{\sigma}_n - \sigma_0)$	$\sqrt{n}(\hat{\alpha}_n - \alpha_0)$
128	1.15	1.91	0.76
256	1.11	1.89	0.75
512	1.08	1.73	0.72
1024	1.05	1.71	0.74
2048	1.06	1.63	0.74
Theoretical limit	1.10	1.55	0.76

TABLE 5
Std of rescaled errors: Multiplicative case $\alpha_0 = 1.3$

n	$V_{\theta_0}^{-\frac{1}{2}} n^{\frac{1}{\alpha_0}-\frac{1}{2}}(\hat{\theta}_n - \theta_0)$	$\frac{\sqrt{n}}{\log(n)}(\hat{\sigma}_n - \sigma_0)$	$\sqrt{n}(\hat{\alpha}_n - \alpha_0)$
128	1.27	0.85	1.44
256	1.35	0.86	1.50
512	1.39	0.87	1.46
1024	1.46	0.82	1.41
2048	1.48	0.81	1.43
Theoretical limit	1.52	0.84	1.42

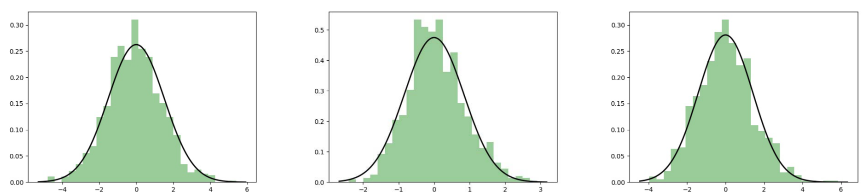
TABLE 6
Std of rescaled errors: Multiplicative case $\alpha_0 = 1.7$

n	$V_{\theta_0}^{-\frac{1}{2}} n^{\frac{1}{\alpha_0}-\frac{1}{2}}(\hat{\theta}_n - \theta_0)$	$\frac{\sqrt{n}}{\log(n)}(\hat{\sigma}_n - \sigma_0)$	$\sqrt{n}(\hat{\alpha}_n - \alpha_0)$
128	1.51	0.53	1.42
256	1.49	0.51	1.47
512	1.45	0.51	1.51
1024	1.49	0.49	1.52
2048	1.50	0.48	1.545
Theoretical limit	1.50	0.51	1.50



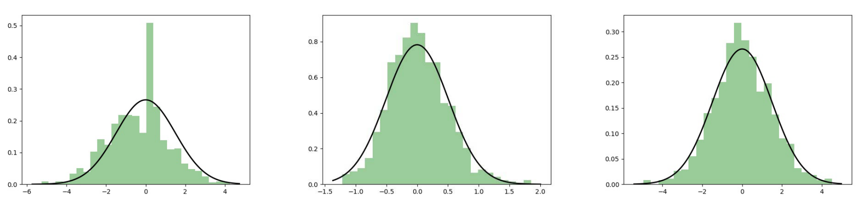
(a) $V_{\theta_0}^{-\frac{1}{2}} n^{\frac{1}{\alpha_0}-\frac{1}{2}}(\hat{\theta}_n - \theta_0)$ (b) $\frac{\sqrt{n}}{\log(n)}(\hat{\sigma}_n - \sigma_0)$ (c) $\sqrt{n}(\hat{\alpha}_n - \alpha_0)$

FIG 1. *Distribution of the rescaled errors of estimation and comparison with their theoretical Gaussian limits ($\theta_0 = 0.5$, $\sigma_0 = 1$, $\alpha_0 = 0.7$, $n = 2048$, multiplicative model).*



(a) $V_{\theta_0}^{-\frac{1}{2}} n^{\frac{1}{\alpha_0}-\frac{1}{2}}(\hat{\theta}_n - \theta_0)$ (b) $\frac{\sqrt{n}}{\log(n)}(\hat{\sigma}_n - \sigma_0)$ (c) $\sqrt{n}(\hat{\alpha}_n - \alpha_0)$

FIG 2. *Distribution of the rescaled errors of estimation and comparison with their theoretical Gaussian limits ($\theta_0 = 0.5$, $\sigma_0 = 1$, $\alpha_0 = 1.3$, $n = 2048$, multiplicative model).*



(a) $V_{\theta_0}^{-\frac{1}{2}} n^{\frac{1}{\alpha_0}-\frac{1}{2}}(\hat{\theta}_n - \theta_0)$ (b) $\frac{\sqrt{n}}{\log(n)}(\hat{\sigma}_n - \sigma_0)$ (c) $\sqrt{n}(\hat{\alpha}_n - \alpha_0)$

FIG 3. *Distribution of the rescaled errors of estimation and comparison with their theoretical Gaussian limits ($\theta_0 = 0.5$, $\sigma_0 = 1$, $\alpha_0 = 1.7$, $n = 2048$, multiplicative model).*

TABLE 7

Correlation between $\hat{\sigma}_n$ and $\hat{\alpha}_n$ (multiplicative model)

α \ n	128	256	512	1024	2048
0.7	-0.85	-0.89	-0.93	-0.94	-0.96
1.3	-0.91	-0.93	-0.95	-0.97	-0.97
1.7	-0.90	-0.93	-0.94	-0.96	-0.96

TABLE 8
Estimation: Multiplicative case $\alpha_0 = 0.7$. Euler approximation for $\xi^x_{1/n}(\theta)$

n	Mean $\hat{\theta}_n$	Std $\hat{\theta}_n$	Mean $\hat{\sigma}_n$	Std $\hat{\sigma}_n$	Mean $\hat{\alpha}_n$	Std $\hat{\alpha}_n$
128	0.501	$7.10 * 10^{-2}$	1.235	$8.74 * 10^{-1}$	0.707	$6.72 * 10^{-2}$
256	0.500	$3.11 * 10^{-2}$	1.146	$6.46 * 10^{-1}$	0.704	$4.74 * 10^{-2}$
512	0.499	$1.71 * 10^{-2}$	1.107	$4.62 * 10^{-1}$	0.700	$3.21 * 10^{-2}$
1024	0.501	$8.06 * 10^{-3}$	1.035	$3.50 * 10^{-1}$	0.702	$2.28 * 10^{-2}$
2048	0.500	$4.25 * 10^{-3}$	1.031	$2.61 * 10^{-1}$	0.700	$1.59 * 10^{-2}$

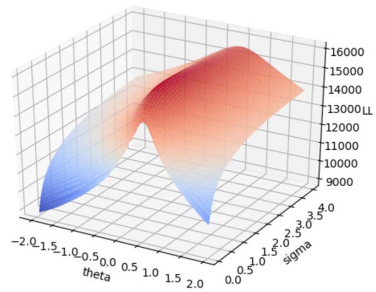
5.1.2. Discussion about the implementation

The minimization of the contrast function (2.3) was conducted using quasi-Newton methods implemented in Python Numpy package. It necessitates to compute numerically the values of the contrast function and of its derivatives, and thus involves numerous evaluations of the functions φ_α, $\partial_z \varphi_\alpha$ and $\partial_\alpha \varphi_\alpha$. These three functions are computed using their integral representations given in [20] and [21], that can be numerically intensive. However, numerical evaluation of the quantities $\varphi_\alpha(z_n(X_{\frac{i-1}{n}}, X_{\frac{i}{n}}, \beta))$, $\partial_z \varphi_\alpha(z_n(X_{\frac{i-1}{n}}, X_{\frac{i}{n}}, \beta))$ and $\partial_\alpha \varphi_\alpha(z_n(X_{\frac{i-1}{n}}, X_{\frac{i}{n}}, \beta))$ for different values of $i = 1, \ldots, n$ can be computed in parallel, using different threads for different values of i. In our numerical simulation, we used CUDA programming language, to implement the computation of the contrast function, and its derivative, with a multi-threaded code on GPU. Using a Nvidia GTX1080 GPU, the Monte-Carlo experiments presented in Table 1–3 with $n = 2048$ and 1000 iterations take around 2 hours each. Hence, searching the values of the parameter for one observation of length $n = 2048$ takes a few seconds, showing that our contrast method is implementable, and fast, in practice.

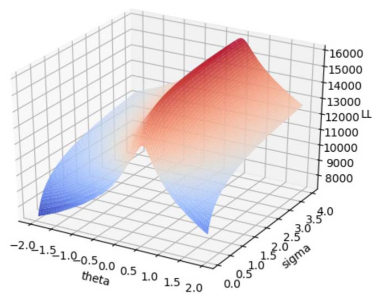
Theorem 3.2 states existence of some zero of the gradient of the contrast function, that yields a consistent estimator. However it does not prevent existence of other zeros that would not be a convergent estimator. Nevertheless, in practice the maximization algorithms always find a consistent estimator, and do not seem to be trapped on local maximum, or non consistent maximum, of the quasi-likelihood function. Searching directly the zeros of the gradient of the contrast function provides convergent estimators as well. This suggests that the zero of the gradient function might be unique for most simulations and reaches the global maximum of the contrast function. To support this, we draw one sample path of observations $(X_{i/n})_{i=0,\ldots,n}$ with $n = 2048$, $\theta_0 = 0.5$, $\sigma_0 = 1$, $\alpha_0 = 0.7$, and explore the shape of the contrast function (2.3). In Figure 4, we plot the graph of

$$\begin{aligned} [-2,2] \times [0.1, 3.961] &\to \mathbb{R} \\ (\theta, \sigma) &\mapsto \log L_n(\theta, \sigma, \alpha) \end{aligned}$$

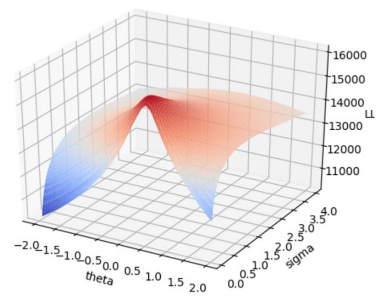
for $\alpha \in \{0.5, 0.6, 0.7, 0.8, 0.9\}$. Figure 4a) plots the cross section for the true value $\alpha_0 = 0.7$, and we see that the maximum in (θ, σ) is reached in a unique point close to the true value (θ_0, σ_0). In Figures 4b)–e), we see that the maximization



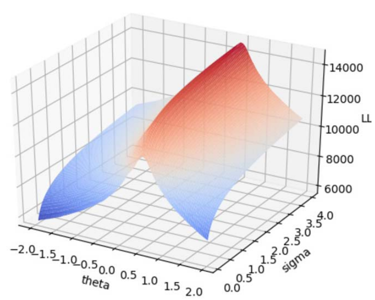
(a) Section of the contrast at $\alpha = 0.7$
Max is reached at $\theta = 0.480$ and
$\sigma = 0.958$ and its value is 16154



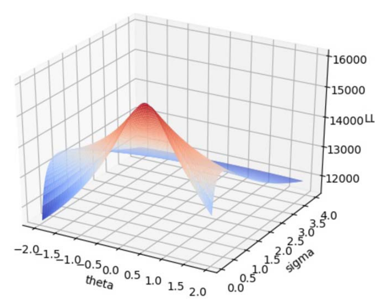
(b) Section of the contrast at $\alpha = 0.6$
Max is reached at $\theta = 0.480$ and
$\sigma = 3.961$ and its value is 16088



(c) Section of the contrast at $\alpha = 0.8$
Max is reached at $\theta = 0.480$ and
$\sigma = 0.256$ and its value is 16141



(d) Section of the contrast at $\alpha = 0.5$
Max is reached at $\theta = 0.480$ and
$\sigma = 3.961$ and its value is 14726



(e) Section of the contrast at $\alpha = 0.9$
Max is reached at $\theta = 0.480$ and
$\sigma = 0.1$ and its value is 16099

FIG 4. *Plot of the cross section at different values of α*

of the contrast function at cross section with values of α far from α_0 yields to a correct estimation of θ, while estimation of σ is far from its true value. However, the values of the maximum in Figures 4b)–e) are lower than the one for $\alpha = 0.7$,

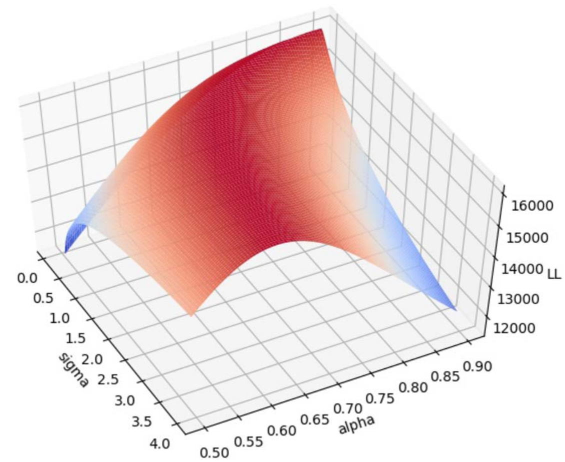


FIG 5. *Plot of the cross section at* $\theta = \theta_0 = 0.5$. *Max is reached at* $\sigma = 0.802$ *and* $\alpha = 0.712$ *and its value is* 16157

showing that when maximizing the contrast function with respect to (θ, σ, α), the maximum will be reached for α close to $\alpha_0 = 0.7$. Figure 5 shows the cross section of the contrast function at $\theta = \theta_0 = 0.5$. We see that the maximum in (α, σ) is reached near the true value (α_0, σ_0). Eventually, maximizing with respect to the three parameters, by Python Numpy package, yields to $\hat{\beta}_n = (0.495, 0.849, 0.709)$ and the quasi-Newton maximization algorithm converges after 18 steps.

5.2. A non multiplicative model driven by an α-stable process

We consider $(X_t)_{t\in[0,1]}$ solution of

$$\mathrm{d}X_t = \theta X_t \mathrm{d}t + \exp(\sigma \sin(X_t))\mathrm{d}S_t^{\alpha},$$

where $(S_t^{\alpha})_t$ is a symmetric α-stable process. The assumption NDNM holds true, and thus we can apply Theorem 3.1. As a consequence the rate of estimation is $n^{1/\alpha_0 - 1/2}$ for θ_0, $\sqrt{n}$ for σ_0 and $\sqrt{n}\log(n)$ for α_0. Comparing to multiplicative case, the rate of estimation is $\log(n)$ faster for both parameters σ_0 and α_0. We make numerical simulations to see if the rate is indeed faster, in practice, in the non-multiplicative case than in the multiplicative one. The asymptotic law of the estimation error is mixed Gaussian by Theorem 3.1, and we define rescaled errors of estimation that have Gaussian laws. Let us define

$$V_{\theta_0} = \left(\int_0^1 \frac{\partial_\theta b(X_s, \theta_0)^2}{a(X_s, \sigma_0)^2} ds \right)^{-1}$$

$$V_{\sigma_0} = \left(\int_0^1 \frac{\partial_\sigma a(X_s, \sigma_0)^2}{a(X_s, \sigma_0)^2} ds - (\int_0^1 \frac{\partial_\sigma a(X_s, \sigma_0)}{a(X_s, \sigma_0)} ds)^2 \right)^{-1}$$

Table 9
Estimation: Non multiplicative case $\alpha_0 = 0.7$

n	Mean $\hat{\theta}_n$	Std $\hat{\theta}_n$	Mean $\hat{\sigma}_n$	Std $\hat{\sigma}_n$	Mean $\hat{\alpha}_n$	Std $\hat{\alpha}_n$
128	0.500	$5.85 * 10^{-2}$	1.170	$8.55 * 10^{-1}$	0.703	$2.58 * 10^{-2}$
256	0.500	$2.92 * 10^{-2}$	1.056	$6.09 * 10^{-1}$	0.702	$1.69 * 10^{-2}$
512	0.500	$1.13 * 10^{-2}$	1.038	$4.63 * 10^{-1}$	0.701	$9.88 * 10^{-3}$
1024	0.500	$6.50 * 10^{-3}$	1.031	$3.69 * 10^{-1}$	0.700	$6.55 * 10^{-3}$
2048	0.500	$3.86 * 10^{-3}$	1.023	$2.51 * 10^{-1}$	0.700	$4.60 * 10^{-3}$

Table 10
Estimation: Non multiplicative case $\alpha_0 = 1.3$

n	Mean $\hat{\theta}_n$	Std $\hat{\theta}_n$	Mean $\hat{\sigma}_n$	Std $\hat{\sigma}_n$	Mean $\hat{\alpha}_n$	Std $\hat{\alpha}_n$
128	0.458	$3.82 * 10^{-1}$	1.046	$3.51 * 10^{-1}$	1.308	$5.94 * 10^{-2}$
256	0.448	$3.44 * 10^{-1}$	1.019	$2.54 * 10^{-1}$	1.302	$3.67 * 10^{-2}$
512	0.476	$3.54 * 10^{-1}$	1.008	$1.81 * 10^{-1}$	1.302	$2.35 * 10^{-2}$
1024	0.465	$2.91 * 10^{-1}$	1.002	$1.34 * 10^{-1}$	1.301	$1.59 * 10^{-2}$
2048	0.496	$2.81 * 10^{-1}$	1.003	$9.51 * 10^{-2}$	1.300	$9.82 * 10^{-3}$

Table 11
Estimation: Non multiplicative case $\alpha_0 = 1.7$

n	Mean $\hat{\theta}_n$	Std $\hat{\theta}_n$	Mean $\hat{\sigma}_n$	Std $\hat{\sigma}_n$	Mean $\hat{\alpha}_n$	Std $\hat{\alpha}_n$
128	-0.358	2.04	1.015	$2.23 * 10^{-1}$	1.702	$7.42 * 10^{-2}$
256	-0.289	1.98	1.001	$1.75 * 10^{-1}$	1.700	$5.08 * 10^{-2}$
512	0.033	1.44	1.003	$1.28 * 10^{-1}$	1.702	$3.36 * 10^{-2}$
1024	0.202	1.22	1.003	$9.17 * 10^{-2}$	1.701	$2.52 * 10^{-2}$
2048	0.245	1.06	1.004	$6.66 * 10^{-2}$	1.701	$1.55 * 10^{-2}$

$$V_{\alpha_0} = \alpha_0^4 \int_0^1 \frac{\partial_\sigma a(X_s,\sigma_0)^2}{a(X_s,\sigma_0)^2} ds V_{\sigma_0}.$$

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

$$V_{\theta_0}^{-1/2} n^{1/\alpha_0 - 1/2}(\hat{\theta}_n - \theta_0) \xrightarrow{n\to\infty} \mathcal{N}(0, (\mathbb{E}h_{\alpha_0}(S_1^{\alpha_0}))^{-1}) \tag{5.2}$$

$$V_{\sigma_0}^{-1/2} \sqrt{n}(\hat{\sigma}_n - \sigma_0) \xrightarrow{n\to\infty} \mathcal{N}(0, (\mathbb{E}k_{\alpha_0}(S_1^{\alpha_0}))^{-1}) \tag{5.3}$$

$$V_{\alpha_0}^{-1/2} \sqrt{n}\log(n)(\hat{\alpha}_n - \theta_0) \xrightarrow{n\to\infty} \mathcal{N}(0, (\mathbb{E}k_{\alpha_0}(S_1^{\alpha_0}))^{-1}) \tag{5.4}$$

In Tables 9–14, we present results of numerical simulations conducted with the true values of the parameters $\theta_0 = 0.5$, $\sigma_0 = 1$ and $\alpha_0 \in \{0.7, 1.3, 1.7\}$. We show a Monte-Carlo evaluation, based on 1000 replications, for the mean and standard deviation of these estimators. Moreover, we evaluate the standard deviation of the rescaled errors of these estimators defined as on the left hand-side of (5.2)–(5.4). We compare these standard deviations with the theoretical limit given by the standard deviation of the variables appearing on the right hand-side of (5.2)–(5.4).

From the results in Tables 9–11, we see that the estimation of the three parameters performs well for $\alpha_0 = 0.7$ and $\alpha_0 = 1.3$, and the parameters σ_0 and α_0 are well estimated for $\alpha_0 = 1.7$ as well. Moreover, from Tables 12–14, we see that the asymptotic behavior of the estimator is in practice very close to the description given by theoretical results (5.2)–(5.4). In Figures 6–8, we

TABLE 12
Std of rescaled errors: Non multiplicative case $\alpha_0 = 0.7$

n	$V_{\theta_0}^{-\frac{1}{2}} n^{\frac{1}{\alpha_0}-\frac{1}{2}}(\hat{\theta}_n - \theta_0)$	$V_{\sigma_0}^{-\frac{1}{2}} \sqrt{n}(\hat{\sigma}_n - \sigma_0)$	$V_{\alpha_0}^{-\frac{1}{2}} \sqrt{n}\log(n)(\hat{\alpha}_n - \alpha_0)$
128	1.23	1.90	1.85
256	1.36	1.93	1.82
512	1.17	2.48	1.87
1024	1.05	3.12	1.84
2048	1.06	2.29	1.86
Theoretical limit	1.05	1.86	1.86

TABLE 13
Std of rescaled errors: Non multiplicative case $\alpha_0 = 1.3$

n	$V_{\theta_0}^{-\frac{1}{2}} n^{\frac{1}{\alpha_0}-\frac{1}{2}}(\hat{\theta}_n - \theta_0)$	$V_{\sigma_0}^{-\frac{1}{2}} \sqrt{n}(\hat{\sigma}_n - \sigma_0)$	$V_{\alpha_0}^{-\frac{1}{2}} \sqrt{n}\log(n)(\hat{\alpha}_n - \alpha_0)$
128	1.31	1.10	1.16
256	1.29	1.09	1.15
512	1.34	1.15	1.15
1024	1.34	1.14	1.17
2048	1.38	1.16	1.18
Theoretical limit	1.52	1.15	1.15

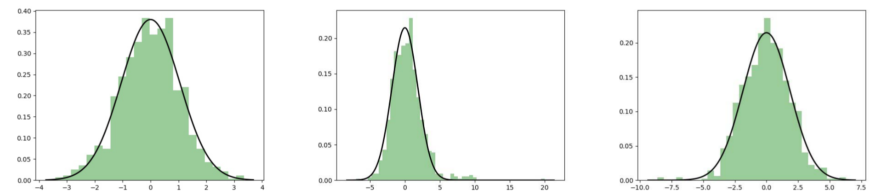
TABLE 14
Std of rescaled errors: Non multiplicative case $\alpha_0 = 1.7$

n	$V_{\theta_0}^{-\frac{1}{2}} n^{\frac{1}{\alpha_0}-\frac{1}{2}}(\hat{\theta}_n - \theta_0)$	$V_{\sigma_0}^{-\frac{1}{2}} \sqrt{n}(\hat{\sigma}_n - \sigma_0)$	$V_{\alpha_0}^{-\frac{1}{2}} \sqrt{n}\log(n)(\hat{\alpha}_n - \alpha_0)$
128	1.86	0.83	0.89
256	2.01	0.84	0.89
512	1.67	0.86	0.90
1024	1.94	0.90	1.03
2048	2.11	0.89	0.93
Theoretical limit	1.50	0.91	0.91

TABLE 15
Correlation between $\hat{\sigma}_n$ and $\hat{\alpha}_n$ (non multiplicative model)

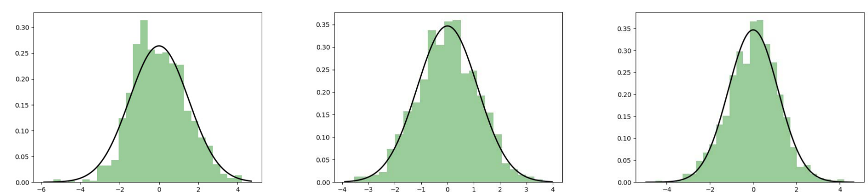
α \ n	128	256	512	1024	2048
0.7	0.23	0.12	0.12	0.093	0.13
1.3	0.34	0.41	0.26	0.33	0.31
1.7	0.34	0.46	0.49	0.60	0.50

plot the distributions of the rescaled errors of estimation given by the left hand sides of (5.2)–(5.4) together with their Gaussian limits, when $n = 2048$ and $\alpha_0 \in \{0.7, 1.3, 1.7\}$. From these figures, we see again that the law of the estimator is very close to its theoretical description. Especially, we observe numerically that the rate of estimation of σ_0, α_0 is different in this non-multiplicative model than for the multiplicative model of Section 5.1. Another difference is that the estimation errors of σ_0 and α_0 are no longer asymptotically proportional in the non-multiplicative case, which is consistent with the numerical evaluation of the correlation between these two estimators given in Table 15.



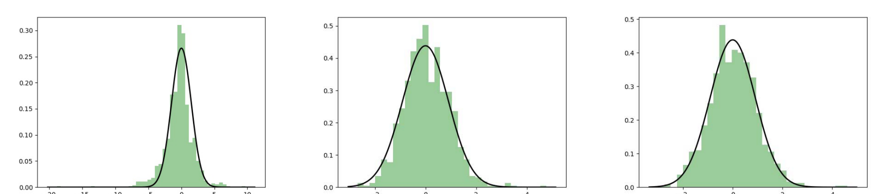
(a) $V_{\theta_0}^{-\frac{1}{2}} n^{\frac{1}{\alpha_0}-\frac{1}{2}} (\hat{\theta}_n - \theta_0)$ (b) $V_{\sigma_0}^{-\frac{1}{2}} \sqrt{n}(\hat{\sigma}_n - \sigma_0)$ (c) $V_{\alpha_0}^{-\frac{1}{2}} \sqrt{n}\ln(n)(\hat{\alpha}_n - \alpha_0)$

FIG 6. *Distribution of the rescaled errors of estimation and comparison with their theoretical Gaussian limits ($\theta_0 = 0.5$, $\sigma_0 = 1$, $\alpha_0 = 0.7$, $n = 2048$, non multiplicative model).*



(a) $V_{\theta_0}^{-\frac{1}{2}} n^{\frac{1}{\alpha_0}-\frac{1}{2}} (\hat{\theta}_n - \theta_0)$ (b) $V_{\sigma_0}^{-\frac{1}{2}} \sqrt{n}(\hat{\sigma}_n - \sigma_0)$ (c) $V_{\alpha_0}^{-\frac{1}{2}} \sqrt{n}\ln(n)(\hat{\alpha}_n - \alpha_0)$

FIG 7. *Distribution of the rescaled errors of estimation and comparison with their theoretical Gaussian limits ($\theta_0 = 0.5$, $\sigma_0 = 1$, $\alpha_0 = 1.3$, $n = 2048$, non multiplicative model).*



(a) $V_{\theta_0}^{-\frac{1}{2}} n^{\frac{1}{\alpha_0}-\frac{1}{2}} (\hat{\theta}_n - \theta_0)$ (b) $V_{\sigma_0}^{-\frac{1}{2}} \sqrt{n}(\hat{\sigma}_n - \sigma_0)$ (c) $V_{\alpha_0}^{-\frac{1}{2}} \sqrt{n}\ln(n)(\hat{\alpha}_n - \alpha_0)$

FIG 8. *Distribution of the rescaled errors of estimation and comparison with their theoretical Gaussian limits ($\theta_0 = 0.5$, $\sigma_0 = 1$, $\alpha_0 = 1.7$, $n = 2048$, non multiplicative model).*

5.3. Non linear drift S.D.E.

In this section, we consider a model with non linear drift:

$$\mathrm{d}X_t = (X_t - \frac{\theta}{1+X_t^2})\mathrm{d}t + \exp(\sigma\sin(X_t))\mathrm{d}L_t^\alpha.$$

Here, the quantity $\xi_{1/n}^x(\theta)$ can not be explicitly computed and we use instead the Euler approximation $\xi_{1/n}^x(\theta) = x + b(x,\theta)/n$. We focus on the case $\alpha_0 = 0.7$ and from Remark 3.2, this Euler approximation is valid as $\alpha_0 > 2/3$. We compare

TABLE 16
Estimation: Non multiplicative case $\alpha_0 = 0.7$, non linear drift, stable process

n	Mean $\hat{\theta}_n$	Std $\hat{\theta}_n$	Mean $\hat{\sigma}_n$	Std $\hat{\sigma}_n$	Mean $\hat{\alpha}_n$	Std $\hat{\alpha}_n$
128	1.015	$1.13 * 10^{-1}$	0.951	$4.84 * 10^{-1}$	0.712	$3.76 * 10^{-2}$
256	1.010	$1.61 * 10^{-1}$	0.960	$3.67 * 10^{-1}$	0.707	$2.93 * 10^{-2}$
512	1.003	$4.09 * 10^{-2}$	0.966	$2.70 * 10^{-1}$	0.703	$1.97 * 10^{-2}$
1024	1.000	$4.72 * 10^{-3}$	0.976	$2.09 * 10^{-1}$	0.703	$1.74 * 10^{-2}$
2048	1.006	$1.34 * 10^{-1}$	0.979	$1.67 * 10^{-1}$	0.703	$1.47 * 10^{-2}$

the results for two different driving Lévy processes, one being exactly α-stable, and the other one being locally α-stable.

TABLE 17
Estimation: Non multiplicative case $\alpha_0 = 0.7$, non linear drift, tempered stable process

n	Mean $\hat{\theta}_n$	Std $\hat{\theta}_n$	Mean $\hat{\sigma}_n$	Std $\hat{\sigma}_n$	Mean $\hat{\alpha}_n$	Std $\hat{\alpha}_n$
128	1.004	$1.25 * 10^{-2}$	1.108	$4.99 * 10^{-1}$	0.706	$3.08 * 10^{-2}$
256	1.001	$1.91 * 10^{-2}$	1.065	$3.87 * 10^{-1}$	0.704	$1.99 * 10^{-2}$
512	1.001	$3.49 * 10^{-3}$	1.022	$2.68 * 10^{-1}$	0.701	$1.29 * 10^{-2}$
1024	1.000	$1.67 * 10^{-3}$	1.010	$1.96 * 10^{-1}$	0.700	$7.78 * 10^{-3}$
2048	1.000	$9.13 * 10^{-4}$	1.010	$1.14 * 10^{-1}$	0.700	$5.19 * 10^{-3}$

5.3.1. Process driven by an α-stable process

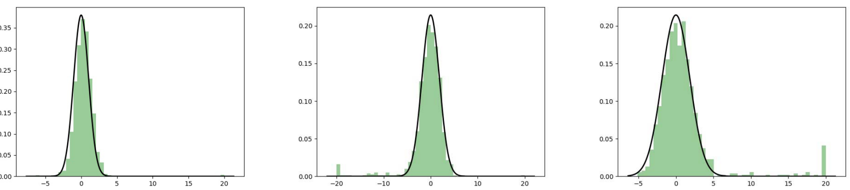
Here, we assume that the Lévy process $(L_t^\alpha)_t = (S_t^\alpha)_t$ is a symmetric α-stable process, as in Sections 5.1–5.2. Its Lévy measure is thus given by $\nu(\mathrm{d}z) = \frac{c_\alpha}{|z|^{1+\alpha}} 1_{\mathbb{R}\setminus\{0\}}(z)\mathrm{d}z$ where $c_\alpha = (-2\Gamma(-\alpha)\cos(\pi\alpha/2))^{-1}$.

The empirical means and standard deviations of the estimators are given in Table 16, for $\theta_0 = 1$, $\sigma_0 = 1$ and $\alpha_0 = 0.7$. In practice, we see that the estimators work well. However, the empirical standard deviation of $\hat{\theta}_n$ seems unstable, as it is not perfectly decreasing with n, and we found rather different values, for different runs of simulations (each with 1000 replications). In Figure 9, we plot the distributions of the rescaled errors together with their Gaussian limits. All errors outside of the interval $[-20, 20]$ are clipped to the interval borders. We see that the empirical distributions are close to their Gaussian limits, except for several larger values that fall outside the interval $[-20, 20]$. It explains why the estimators work well in practice, while the estimation of their variances can be unstable, due to a few extreme values.

5.3.2. Process driven by a tempered α-stable process

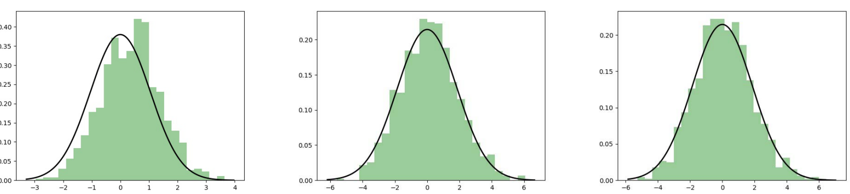
Here, we assume that the Lévy process $(L_t^\alpha)_t$ is a tempered α-stable process, whose Lévy measure is given by $\nu(\mathrm{d}z) = \frac{c_\alpha}{|z|^{1+\alpha}} e^{-|z|} 1_{\mathbb{R}\setminus\{0\}}(z)\mathrm{d}z$. To simulate tempered stable random variables, we use the rejection based method proposed in [15].

The empirical means and standard deviations of the estimator are given in Table 17, for $\theta_0 = 1$, $\sigma_0 = 1$ and $\alpha_0 = 0.7$.



(a) $V_{\theta_0}^{-\frac{1}{2}} n^{\frac{1}{\alpha_0}-\frac{1}{2}} (\hat{\theta}_n - \theta_0)$ (b) $V_{\sigma_0}^{-\frac{1}{2}} \sqrt{n}(\hat{\sigma}_n - \sigma_0)$ (c) $V_{\alpha_0}^{-\frac{1}{2}} \sqrt{n}\ln(n)(\hat{\alpha}_n - \alpha_0)$

FIG 9. *Distribution of the rescaled errors of estimation and comparison with their theoretical Gaussian limits ($\theta_0 = 1$, $\sigma_0 = 1$, $\alpha_0 = 0.7$, $n = 2048$, non linear drift, stable case).*



(a) $V_{\theta_0}^{-\frac{1}{2}} n^{\frac{1}{\alpha_0}-\frac{1}{2}} (\hat{\theta}_n - \theta_0)$ (b) $V_{\sigma_0}^{-\frac{1}{2}} \sqrt{n}(\hat{\sigma}_n - \sigma_0)$ (c) $V_{\alpha_0}^{-\frac{1}{2}} \sqrt{n}\ln(n)(\hat{\alpha}_n - \alpha_0)$

FIG 10. *Distribution of the rescaled errors of estimation and comparison with their theoretical Gaussian limits ($\theta_0 = 1$, $\sigma_0 = 1$, $\alpha_0 = 0.7$, $n = 2048$, non linear drift, tempered stable case).*

We see that the estimator works very well in practice. Especially, the estimation of the drift parameter has a smaller standard deviation than in the stable case. In Figure 10, we plot the distributions of the rescaled errors of estimation together with their Gaussian limits. We see that the empirical distributions fit very well the theoretical ones. Especially, we do not observe the presence of extreme values, as it is the case when the model is driven by a stable process.

References

[1] AÏT-SAHALIA, Y. and JACOD, J. (2007). Volatility estimators for discretely sampled Lévy processes. *Ann. Statist.* **35** 355–392. MR2332279

[2] AÏT-SAHALIA, Y. and JACOD, J. (2008). Fisher's information for discretely sampled Lévy processes. *Econometrica* **76** 727–761. MR2433480 (2010g:60088)

[3] AÏT-SAHALIA, Y. and JACOD, J. (2009). Estimating the degree of activity of jumps in high frequency data. *Ann. Statist.* **37** 2202–2244. MR2543690

[4] BICHTELER, K. and JACOD, J. (1983). Calcul de Malliavin pour les diffusions avec sauts: existence d'une densité dans le cas unidimensionnel. In *Seminar on probability, XVII. Lecture Notes in Math.* **986** 132–157. Springer, Berlin. MR770406

[5] Brouste, A. and Masuda, H. (2018). Efficient estimation of stable Lévy process with symmetric jumps. *Stat. Inference Stoch. Process.* **21** 289–307. MR3824969

[6] Clément, E. and Gloter, A. (2019). Estimating functions for SDE driven by stable Lévy processes. *Ann. Inst. Henri Poincaré Probab. Stat.* **55** 1316–1348. MR4010937

[7] 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 Probab. Stat.* **23** 136–175. MR3945580

[8] DuMouchel, W. H. (1973). On the asymptotic normality of the maximum-likelihood estimate when sampling from a stable distribution. *Ann. Statist.* **1** 948–957. MR0339376

[9] Fournier, N. and Printems, J. (2010). Absolute continuity for some one-dimensional processes. *Bernoulli* **16** 343–360. MR2668905

[10] Ivanenko, D., Kulik, A. M. and Masuda, H. (2015). Uniform LAN property of locally stable Lévy process observed at high frequency. *ALEA Lat. Am. J. Probab. Math. Stat.* **12** 835–862. MR3453298

[11] Jacod, J. and Protter, P. (2012). *Discretization of processes. Stochastic Modelling and Applied Probability* **67**. Springer, Heidelberg. MR2859096

[12] Jacod, J. and Sørensen, M. (2018). A review of asymptotic theory of estimating functions. *Stat. Inference Stoch. Process.*

[13] Jing, B.-Y., Kong, X.-B. and Liu, Z. (2012). Modeling high-frequency financial data by pure jump processes. *Ann. Statist.* **40** 759–784. MR2933665

[14] Kawai, R. and Masuda, H. (2011a). On the local asymptotic behavior of the likelihood function for Meixner Lévy processes under high-frequency sampling. *Statist. Probab. Lett.* **81** 460–469. MR2765166

[15] Kawai, R. and Masuda, H. (2011b). On simulation of tempered stable random variates. *J. Comput. Appl. Math.* **235** 2873–2887. MR2763192

[16] Kawai, R. and Masuda, H. (2013). Local asymptotic normality for normal inverse Gaussian Lévy processes with high-frequency sampling. *ESAIM Probab. Stat.* **17** 13–32. MR3002994

[17] Kong, X.-B., Liu, Z. and Jing, B.-Y. (2015). Testing for pure-jump processes for high-frequency data. *Ann. Statist.* **43** 847–877. MR3325712

[18] Masuda, H. (2009). Joint estimation of discretely observed stable Lévy processes with symmetric Lévy density. *J. Japan Statist. Soc.* **39** 49–75. MR2571802

[19] Masuda, H. (2019). Non-Gaussian quasi-likelihood estimation of SDE driven by locally stable Lévy process. *Stochastic Process. Appl.* **129** 1013–1059. MR3913278

[20] Matsui, M. and Takemura, A. (2006). Some improvements in numerical evaluation of symmetric stable density and its derivatives. *Comm. Statist. Theory Methods* **35** 149–172. MR2274041

[21] Nolan, J. P. (1997). Numerical calculation of stable densities and distribution functions. *Comm. Statist. Stochastic Models* **13** 759–774. Heavy tails and highly volatile phenomena. MR1482292

[22] Sørensen, M. (1999). On asymptotics of estimating functions. *Braz. J.*

Probab. Stat. **13** 111–136. MR1803041

[23] Sweeting, T. J. (1980). Uniform asymptotic normality of the maximum likelihood estimator. *Ann. Statist.* **8** 1375–1381. MR594652

[24] Todorov, V. and Tauchen, G. (2011). Limit theorems for power variations of pure-jump processes with application to activity estimation. *Ann. Appl. Probab.* **21** 546–588. MR2807966