Parameter estimation for a class of stable driven stochastic differential equations

Solym Manou-Abi

PARAMETER ESTIMATION FOR A CLASS OF STABLE DRIVEN STOCHASTIC DIFFERENTIAL EQUATIONS

Solym M. Manou-Abi [1,2,3]

[1] Institut Montpelliérain Alexander Grothendieck, UMR CNRS 5149, Montpellier
solym-mawaki.manou-abi@umontpellier.fr
[2] Centre Universitaire de Formation et de Recherche, Mayotte
Département Sciences et Technologies
solym.manou-abi@mayotte.fr
[3] Laboratoire de Mathématiques et Applications, UMR CNRS 7348, Poitiers
solym.manou.abi@math.univ-poitiers.fr

Abstract. In this paper, we first study the estimation of diffusion parameters for one-dimensional, ergodic stochastic processes observed at some discrete times, that is a solution of a given class of stochastic differential equations driven by α-stable processes, $\alpha \in (1, 2)$. After recalling the non-parametric estimation framework of the drift functionn namely the Nadaraya-Watson estimation, we discuss an estimation method of the diffusion parameters (the scaling and the driving stable process parameters) based on the Euler–Maruyama scheme. The novelty of this paper which is our baseline is the combination of a characteristic sample function and the Least Squares Estimation (LSE) method. Secondly, we apply the diffusion parameters results for stable driven Ornstein-Uhlenbeck (OU), Cox–Ingersoll–Ross (CIR) and Lotka–Volterra (LV) processes. We also consider the estimation of the drift coefficients in the linear case namely, the stable driven OU and CIR processes. Using the Itô formula and the linear statistical regression derived from the LSE method, we establish under certain conditions, the consistency of their drift coefficients estimators. We efficiently discuss our result with numerical simulations using synthetic data. A real data in finance, such as exchange rates is used to fit the parameters of a justified model amoung the stable driven OU and CIR processes. As a forthcoming work, we intend to study the rate of convergence of the estimators and to create a package on R software to handle this kind of estimation problem.

Key words and phrases: parameter estimation, Nadaraya-Watson estimation, stable process, characteristic function, regression method, CIR, OU and Lotka–Volterra processes.

1. Introduction

Due to an increase of powerful computation of statistical methods, there has been a great interest in parameter estimation for Stochastic Differential Equation (SDE). Such models which are mathematical tools for modelling

the time evolution of several natural phenomena in many fields like epidemiology, biology and finance. For instance, the Cox–Ingersoll–Ross (CIR) and the Ornstein-Uhlenbeck (OU) stochastic models [31, 10], have been broadly used in finance. Furthermore, it is worth noting that select stochastic differential equation (SDE) models derived from continuous-time branching processes have been demonstrated to serve as valuable tools for the analysis of population dynamics [2, 39]. Specifically, in the context of epidemic dynamics, when the infectious population is of small magnitude relative to the overall population size, it becomes feasible to approximate the continuous-time dynamics using a nonlinear Stochastic Differential Equation (SDE) model, such as a stochastic logistic model that incorporates variations arising from birth and death processes [2]. In the realm of classical estimation methods, the challenges of devising procedures that strike a balance between computational efficiency and achieving optimal statistical performance have reached a well-established level of understanding.

Let us consider the parameter estimation problem of the following SDE driven by a standard strictly stable process $(Z_t)_{t\geq 0}$ defined in a given filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$:

$$\begin{cases} dX_t = f(X_t)dt + \rho\, g(X_{t-})dZ_t, & t \in [0,T] \\ X_0 = x_0, \end{cases} \tag{1}$$

where x_0 is a starting point, $T > 0$ stands for time horizon. Notably, in our investigation, we assume the knowledge of the function $g : \mathbb{R} \to \mathbb{R}$. The function $f : \mathbb{R} \to \mathbb{R}$, constant $\rho > 0$ and the stable process $(Z_t)_{t\in[0,T]}$ (which is considered as a nuisance parameter) are unknown. The random variable Z_1 which is a standard strictly stable distribution with parameters $\alpha \in (1, 2)$ (the index of stability) and $\beta \in [-1, 1]$ (the skewness parameter) will be defined in the sequel. The existence and uniqueness of solutions to the SDE (1) under Lipschitz conditions are standard results in stochastic calculus [4]. Nevertheless, we recall some findings:

- For $f = 0$ and g is Hölder continuous with exponent $\frac{1}{\alpha}$; if Z is a symmetric stable process then it is proved in [5] that there exists a strong solution for $\alpha \in (1, 2)$.
- If f and g have at most linear growth it is proved in [14] that there exists a weak solution for $\alpha \in (1, 2)$ for which pathwise uniqueness holds whenever the function g is Hölder continuous with exponent lying in $[1 - \frac{1}{\alpha}, \frac{1}{\alpha}]$.
- When the drift term f is Hölder continuous of order lying in $(2 - \alpha, 1)$ and g is Lipshitz continuous and bounded, then existence and uniqueness of a strong solution is derived in [35], [38].

For further recent studies of existence and uniqueness, we refer to [32] and the references therein.

The main focus of this paper is the estimation of diffusion parameters for one-dimensional ergodic stochastic processes X observed at some discrete times, that is a solution of the SDE (1) driven by α-stable processes, $\alpha \in (1, 2)$. We assume that, the process X is observed at discrete time points $\{t_i = i\Delta_n,\ i = 0, 1, 2, ..., n\}$ with Δ_n a time frequency of the observation and n is the sample size. Firstly, we recall not only the non-parametric estimation framework of the drift coefficient but also the link between ergodicity and mixing conditions.

The case study of the parameter estimation theory for SDE by Brownian motions are well known in the literature. Nevertheless, one can mention some traditional methods such as the maximum likelihood estimator (MLE), the least squares estimator (LSE) techniques [29, 13, 24, 11]; the consistency and the asymptotic distribution [13, 24], [19], [40], [1]. For further recent investigation, we refer to [23] and the references therein. Significant advancements have been achieved in the realm of parameter estimation for Stochastic Differential Equations (SDEs) driven by Lévy processes with finite moments. The investigation conducted in [34] centered on the consistency and asymptotic normality when the driving process manifests as a zero-mean adapted Lévy process with finite moments. The study of the asymptotic normality of the Least Squares Estimator (LSE) and Maximum Likelihood Estimator (MLE) for the pure jump case is extensively explored in [46, 47]. It is noteworthy that research has been undertaken in instances where the driving process Z takes the form of a stable Lévy process, characterized by its distinctive infinite variance property. However, in this scenario, the MLE loses its validity as the explicit density function is not universally available, and the Girsanov measure transformation encounters challenges for α-stable processes with $\alpha \in (1, 2)$. Key contributions in this domain are highlighted in the following well-established results:

a) In [15, 16], authors use the trajectory fitting method in conjunction with the weighted least squares technique for the drift coefficient of an α-stable driven Ornstein–Uhlenbeck (OU) process, where $\alpha \in (1, 2)$. This approach is applicable when the process is observed at discrete time instants, encompassing both ergodic and non-ergodic cases. The work also delves into the consistency and asymptotic distribution of the estimator, showcasing a higher order of convergence compared to the classical Gaussian case.
b) The investigation conducted in [26, 51] revolves around the drift parameter estimation of a stable driven Cox–Ingersoll–Ross (CIR) model, a special subcritical continuous-state branching process with immigration. The authors derive the consistency and central limit theorems of the conditional least squares estimators and the weighted conditional least squares estimators of the drift parameters based on low-frequency observations.

c) [12] explores Lasso and Slope drift estimators for Lévy-driven Ornstein–Uhlenbeck processes.
d) Furthermore, [53] focuses on a maximum likelihood-type estimation of the drift and volatility constant coefficient parameters in a stable-driven CIR model.
e) In a recent contribution [6], the author addresses the joint parameter estimation of the drift parameters, scaling parameter for the diffusion coefficient, and the stability index α for the jump activity parameter. This estimation is conducted from high-frequency observations of the stable CIR process over a fixed time period. The methodology employed is grounded in the approximation of the conditional distribution.

Now, we shift our focus to non-parametric estimation, a pivotal domain in statistics that involves the estimation of an unknown function from a data sample without predefining a formula. Numerous authors have delved into the non-parametric estimation of the drift function, denoted as f, within the framework of diffusions driven by Brownian motions, considering various conditions. The fundamental statistical properties, including the consistency and rate of convergence, have been extensively investigated for non-parametric methods, particularly the Nadaraya-Watson (N-W) estimators [36, 50]. These investigations span scenarios under both independence with identically distributed data and weak dependence conditions, such as mixing conditions. A comprehensive overview of non-parametric methods for diffusion processes can be found in the survey paper authored by [17]. In the stable and non-parametric context, additional insights are explored in [52, 30], as well as [41, 55, 28]. However, tackling the non-parametric estimation of the diffusion part (ρ, g and Z) proves to be considerably more challenging [30].
In this paper, we operate under the assumption that the function g is known and provided. If the solution of the Stochastic Differential Equation (SDE) represented by (1) is stationary, and the stationary distribution possesses a continuous density and is strongly mixing, the authors in [30] establish a non-parametric estimation of the drift function utilizing the Nadaraya-Watson (N-W) estimator. The statistical properties, such as consistency and the rate of convergence of the N-W estimators, under dependence conditions like mixing, are systematically developed in [30]. In instances where the drift function is linear, the estimation for both discrete and continuous observations has been extensively investigated in previous works, including [16, 15, 27, 6, 53], encompassing stable-driven Cox–Ingersoll–Ross (CIR) and Ornstein–Uhlenbeck (OU) model types. All methods developed and applied within the framework of stable-driven CIR and OU models are conditionally dependent on the observation of the noise $Z = (Z_t)_{t\in[0,T]}$ parameter. Their performance is strongly influenced by the time frequency Δ_n and the time

horizon T. While parameter estimation for stable-driven Stochastic Differential Equations (SDE's) has seen significant development in recent years, few works, such as [6] in the CIR case, delve into joint parameter estimation for the drift coefficient, scaling parameter for the diffusion coefficient and, the jump activity diffusion parameters. These parameters include the stability index α and the scaling diffusion parameter ρ.

In this article, our primary objective is to address the estimation of diffusion parameters for discretely observed stable driven stochastic differential equations (1), assuming that only the function g is known. Our framework presupposes that the solution X is stationary and strongly mixing, with a connection emphasized with ergodic conditions. The proposed estimation procedure introduces a novel methodology that utilizes the N-W estimator for the drift function. Specifically, using the Euler–Maruyama scheme, we derive estimators for the parameters associated with the diffusion part (scaling parameter ρ and the stable noise parameters α and β) from a given sample characteristic function and employing linear (or weighted) regression techniques. Secondly, in cases where the drift function is linear, such as in stable-driven Cox–Ingersoll–Ross (CIR) and Ornstein–Uhlenbeck (OU) processes, we establish estimators for the drift coefficients using Itô formula and employing again linear regression techniques, we derived unbiased and consistent estimators of their drift coefficients.

The organizational structure of this paper is delineated as follows. In Section 2, we revisit fundamental facts about stable processes and introduce relevant assumptions for the classical N-W estimation framework. We present our main theorems in Section 3, which articulate explicit formulas for the estimators of the diffusion parameters (scaling parameter ρ and the stable noise parameters α and β). Additionally, we provide estimators for the drift coefficients in cases where the drift is linear, specifically for stable-driven CIR and OU processes. Section 4 is dedicated to presenting the proofs of our main results and Section 5 delves into the numerical performance evaluation of the estimators using synthetic data. We apply the best performance results to real financial data, such as exchange rates of the Canadian dollar against the US dollar over a fixed period. Finally, we outline all material such that tables and figures in the last section as well as the connection between ergodicity and mixing conditions for Markov processes for a better comprehension of this paper.

2. Preliminaries: Stable processes and non-parametric estimation of the drift

Throughout this paper we assume that $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P})$ is a filtered probability space satisfying the usual conditions, i.e.,

(1) $(\Omega, \mathcal{F}_t, \mathbb{P})$ is complete for all $t \in \mathbb{R}_+$, $\mathcal{F}_0$ contains all the P-null sets in $\mathcal{F}$ for all $t \in \mathbb{R}_+$.
(2) $\mathcal{F}_t = \mathcal{F}_{t+}$ where $\mathcal{F}_{t+} = \cap_{s>t}\mathcal{F}_s$, for all $t \geq 0$, i.e. the filtration is right-continuous.

In this section, our primary objectives are twofold. Firstly, we aim to introduce stable laws and processes, emphasizing a method for simulating stable alpha laws and processes. Secondly, we present key assumptions concerning the non-parametric estimation of the drift function via the N-W estimator's. Stable laws, introduced by Paul Lévy in 1925, emerge as the limit of normalized sums. The definitions and properties provided below are drawn from sources such as [44] and [37]. It's important to note that there exist multiple characterizations of stable distributions, each with distinct advantages depending on the intended applications.

Definition 2.1. *A random variable X follows an α-stable distribution with $\alpha \in (0, 2]$ and we write $X \sim S_\alpha(\sigma, \beta, \omega; 1)$ if it is uniquely determined by its characteristic function :*

$$\Psi(t) = \mathbb{E}(\exp(itX)) = \begin{cases} \exp\left(-\sigma^\alpha |t|^\alpha \left[1 - i\beta(\tan(\frac{\pi\alpha}{2}))sign(t)\right] + i\omega t\right) & \text{if } \alpha \neq 1. \\ \exp(-\sigma|t|[1 + i\beta\frac{2}{\pi}sign(t)\log(|t|)] + i\omega t) & \text{if } \alpha = 1 \end{cases}$$

where $\alpha \in (0, 2]$ is the index of stability, $\beta \in [-1, 1]$ the skewness parameter; $\sigma > 0$ the scale parameter and $\omega \in \mathbb{R}$ the location or shift parameter and $sign(t)$ the sign function.

The aforementioned definition corresponds to parametrization 1 in [37]. The following parametrization, denoted as parametrization 0, proves to be particularly useful for computer processing.

Definition 2.2. *A random variable $X \sim S_\alpha(\sigma, \beta, \omega; 0)$ if,*

$$\Psi(t) = \begin{cases} \exp\left(-\sigma^\alpha |t|^\alpha \left[1 + i\beta(\tan(\frac{\pi\alpha}{2}))sign(t)(|\sigma t|^{1-\alpha} - 1)\right] + i\omega t\right) & \text{if } \alpha \neq 1. \\ \exp(-\sigma|t|[1 + i\beta\frac{2}{\pi}sign(t)\log(\gamma|t|)] + i\omega t) & \text{if } \alpha = 1. \end{cases}$$

For a more in-depth exploration of stable distributions, please refer to [37]. When $\sigma = 1$, we designate X as a standard α-stable random variable. For $\sigma \neq 1$ and $\omega = 0$, X is referred to as a strictly α-stable random variable, and in this case, it is known that $\mathbb{E}(X) = 0$ whenever $\alpha \in (1, 2)$. Additionally, if $\beta = 0$, it is characterized as symmetric. In general $\mathbb{E}(|X|^p) < \infty$ for any $p < \alpha$ and $\mathbb{E}(|X|^p) = +\infty$ for any $p \geq \alpha$ so that the variance is infinite and in the case $\alpha \in (0, 1)$ the mean is infinite. Now, let's introduce α-stable processes, which can be regarded as a specific class of Lévy processes. For further insights, one can consult established literature such as [20].

Definition 2.3. *An $\mathcal{F}_t$-adapted stochastic process $Z = \{Z_t\}_{t\geq 0}$ is called an α-stable process if*

(1) *$Z_0 = 0$, a.s.;*
(2) *Z has α-stable stationary increments distributions $Z_t - Z_s \sim Z_{t-s} \sim S_\alpha(\sigma(t-s)^{1/\alpha}, \beta, \omega; k)$, $t > s \geq 0$ and $k = 0$ or 1.*

(3) *For any time points* $0 \leq s_0 < \ldots < s_m < \infty$*, the random variables* $Z_{s_0}, Z_{s_1} - Z_{s_0}, \ldots, Z_{s_m} - Z_{s_{m-1}}$ *are independent.*

For insights into the case of standard α-stable processes, [30] is a recommended reference. Set $\Delta Z_t = Z_t - Z_{t-}$. The jumps part of the stochastic process Z can be described by its Poisson random measure (jump measure of Z on interval $[0,t]$) defined as $N(t,A) = \sum_{0\leq s\leq t} \mathbb{I}_A(\Delta Z_s), \quad A \in \mathcal{B}(\mathbb{R}^*)$, the number of jumps of Z on the interval $[0,t]$ whose size lies in the set A bounded below. For such A, the process $N(,A)$ is a Poisson process and the Lévy measure $\nu(A) := \mathbb{E}(N(1,A))$ defined on $\mathbb{R}\setminus\{0\}$ has the following explicit form: $\nu(dx) := \frac{dx}{|x|^{\alpha+1}}\left(c_+\, 1_{\{x>0\}} + c_-\, 1_{\{x<0\}}\right)$. In the sequel, we'll denote by $\tilde{N}(t,A) = N(t,A) - t\nu(A)$ the compensated Poisson measure. Note that in the case $\alpha \in (1,2)$, the characteristic function of a strictly α-stable process is reduced into the form:

$$\Psi_{Z_t}(u) = \exp t\left(\int_{-\infty}^{+\infty}(e^{iuy} - 1 - iuy)\nu(dy)\right), \quad t \geq 0.$$

The parameters c_+, c_- mentioned above are non-negative, with the additional condition that $c_+ + c_- > 0$. There exists a connection between jump measurement coefficients c_+, c_- and the skewness parameter, for instance, $\beta = \frac{c_+ - c_-}{c_+ + c_-}$. If $\beta > 0$, we will say that Z admits a positive jump activity. We recall here a random walk approximation method based on the work of [8, 18], proves useful for simulating such a stable process. Naturally, other methods are also available, such as the series approximation of Lévy processes, see [18] for more details.

Algorithm 1 Discretized trajectory simulation for a strictly α-stable process

1: Set $Z = (Z_t)_{t\in[0,T]}$ un processus strictement α-stable.
2: **Step 1:**
   (1) Simulate n independent, uniformly distributed random variables Φ on $[-\pi/2, \pi/2]$ and n independent and identically distributed random variable W as an exponentially with parameter 1.
3: **Step 2:** Compute ΔZ_i for $i = 1, ...n$ as follows.
   (1) If $\alpha \neq 1$

$$\Delta Z_i = \sigma \left(\frac{T}{n}\right)^{1/\alpha} \frac{\sin(\alpha(\Phi - \phi_0))}{\cos(\Phi)^{1/\alpha}} \left(\frac{\cos(\Phi - \alpha(\Phi - \phi_0))}{W}\right)^{\frac{1-\alpha}{\alpha}}.$$

   (2) If $\alpha = 1$

$$\Delta Z_i = \sigma \left(\frac{T}{n}\right)^{1/\alpha} \frac{2}{\pi} \left(\left(\frac{\pi}{2} + \beta\Phi\right) \tan(\Phi) - \beta \log \left(\frac{\frac{1}{2}\pi W \cos(\Phi)}{\frac{1}{2}\pi + \beta\Phi} \right) \right)$$

   where

$$\phi_0 = -\frac{\beta\pi}{2} \frac{1 - |1 - \alpha|}{\alpha}.$$

4: **Step 3:**The discretized trajectory of Z is given by

$$Z(t_i) = \sum_{k=1}^{i} \Delta Z_k.$$

From a practical point of view, we consider the linear interpolation between instants t_i and t_{i+1} for graphical representations. Algebraic transformations can be used to extend the above algorithm to the general case.

Now, we revisit the non-parametric estimation framework for the drift function of the above stable driven SDE equation. Let X be a solution of the above SDE (1). Note that Lipschitz and Hölder (or bounded) conditions are typical conditions to show the existence of the solution. For more recent existence conditions, we refer to [32] and references therein.

The Nadaraya-Watson (N-W) estimator is a classical method to estimate the drift function f. It is in this way that, we make the following assumptions following [30]. Similar results on the drift function estimation is also presented in [28] where authors study the local polynomial estimation under regular conditions. We consider kernel function $K(\cdot)$ which is a symmetric and non negative probability density function satisfying $\sup(1 \vee |u|)K(u) < M_0 < +\infty$ and

$$\int_{-\infty}^{+\infty} u^2 K(u)du < +\infty, \quad \int_{-\infty}^{+\infty} K^2(u)du < +\infty.$$

($\mathbf{A_1}$). The solution X_t is stationary and admits a unique invariant distribution π which is geometrically strong mixing.
($\mathbf{A_2}$). The density function f(x) of the stationary distribution π is continuous.
($\mathbf{A_3}$). As $n \to \infty : \ h \to 0, \ \Delta_n \to 0,$

$$n\Delta_n \to \infty, \text{and} \frac{n\Delta_n h}{(\log(n\Delta_n))^2} \to +\infty.$$

Assume that, the process X is observed at some discrete time points $\{t_k = k\Delta_n, \quad k = 0, 1, 2, ..., n-1\}$ with Δ_n a time frequency of the observation and n is the sample size. The Euler-Maruyuma scheme of the above SDE is written as follows.

$$X^n_{t_{k+1}} = X^n_{t_k} + f(X_{t_k})\Delta_n + \rho g(X_{t_k})\Delta Z_k, \quad X^n_0 = x \in \mathbb{R},$$

where $\Delta Z_k = Z^n_{t_{k+1}} - Z^n_{t_k}$. Set $Y_k = X^n_{t_{k+1}} - X^n_{t_k}$. The main idea of N-W estimator is to minimize the following object function:

$$\sum_{k=0}^{n-1} W_{n,k}(x)(Y_k - b\Delta_n)^2$$

over the parameter space of b with certain weights functions $W_{n,k}(x)$ given by:

$$W_{n,k}(x) = \frac{K_h(\frac{X_{t_k}-x}{h})}{\sum_{k=0}^{n-1} K_h(\frac{X_{t_k}-x}{h})}$$

for all $x \in \mathbb{R}$, where $K_h(x) = (K(x/h))/h$ with K the Kernel and h the bandwidth parameter. Thus the N-W estimator of the drift function f is given by the following expression :

$$\hat{f}_n(x) = \sum_{k=0}^{n-1} W_{n,k}(x)Y_k, \ \ \forall x \in \mathbb{R}$$

so that

$$\hat{f}_n(x) = \sum_{i=1}^{n} \frac{K_h(\frac{X_{t_i}-x}{h})(X_{t_i} - X_{t_{i-1}})}{n^{-1}\sum_{i=1}^{n} K_h(\frac{X_{t_i}-x}{h})} \quad \text{for all } x \in \mathbb{R}. \tag{2}$$

In this study, we investigate several kernel functions; however, we specifically adopt the Gaussian kernel function for simulation purposes:

$$K(u) = \frac{1}{\sqrt{2\pi}} e^{-u^2/2}.$$

3. Main results

Using the N-W estimator as a foundation, and the Euler–Maruyama scheme, the primary goal in this section is twofold. Firstly, it involves the simultaneous estimation of the scaling parameter ρ and the parameters α and β characterizing the standard strictly stable process in the diffusion component of the SDE (1). Secondly, we focus on the estimation of common drift coefficients θ and μ for the Stable driven CIR and OU processes in equations (6) or (8).

Theorem 3.1. *Let $X = (X_t)_{t\in[0,1]}$ be a real valued stationnary and strongly mixing process with unique invariant distribution which is solution of the SDE (1). We assume that X is observed on a sample of discrete time points t_i with step Δ_n of size n such $g(X_{t_i}) \neq 0$. Assume that assumptions $(\mathbf{A_2})$ and $(\mathbf{A_3})$ holds. Let $(\hat{f}_n(X_{t_i}))_{i=1}^n$ be the sequence of the N-W estimate of the drift coefficient of the SDE (1). We define the following sample characteristic function*

$$\hat{\varphi}_n(z_k) = \frac{1}{n}\sum_{i=0}^{n-1} \exp\left(jz_k\Big(\frac{\Delta X_{t_i} - \hat{f}_n(X_{t_i})\Delta_n}{g(X_{t_i})}\Big)\right) \tag{3}$$

for all $k \in [[1, m]]$, $m > 0$, and $j^2 = -1$. Set

$$\begin{cases} \hat{\alpha}_m = \dfrac{\displaystyle\sum_{k=1}^{m} W_kV_k - \frac{1}{m}\sum_{k=1}^{m} V_k \sum_{k=1}^{m} W_k}{\big(\sum_{k=1}^{m} W_k^2\big) - \frac{1}{m}\big(\displaystyle\sum_{k=1}^{m} W_k\big)^2}, \\ \hat{\lambda}_m = \frac{1}{m}\displaystyle\sum_{k=1}^{m} V_k - \frac{\hat{\alpha}_m}{m}\sum_{k=1}^{m} W_k \\ \hat{\rho}_m = \Delta_n^{-\frac{1}{\hat{\alpha}_m}} \exp\frac{\hat{\lambda}_m}{\hat{\alpha}_m} \\ \hat{\beta}_m = \dfrac{\displaystyle\sum_{k=1}^{m} z_kS_k}{\displaystyle\sum_{k=1}^{m} z_kB_k}, \end{cases} \tag{4}$$

where

$$S_k = \arg(\hat{\varphi}_n(z_k)), \quad B_k = \tan(\frac{\pi\hat{\alpha}}{2})sign(z_k)\big(z_k - z_k^{\hat{\alpha}_m}\big)\Delta_n^{1/\hat{\alpha}_m},$$

and

$$V_k = \log(-\log|\hat{\varphi}_n(z_k)|) \quad \textit{and} \quad W_k = \log(|z_k|).$$

For n large enough and good choice of m values z_k, $\hat{\alpha}_m$, $\hat{\rho}_m$ and $\hat{\beta}_m$ are least squares estimators of α, ρ and β.

If we consider some weights p_k, for instance $p_k = \frac{1}{\delta_k^2}$ the inverse of the variance of the k-th observation, a weighted least squares method lead to the following estimators:

$$\begin{cases} \hat{\alpha}_m = \dfrac{\displaystyle\sum_{k=1}^{m} p_k W_k V_k - \frac{1}{\sum_{k=1}^{m} p_k} \sum_{k=1}^{m} V_k \sum_{k=1}^{m} W_k}{\left(\sum_{k=1}^{m} p_k W_k^2\right) - \frac{1}{m}\left(\sum_{k=1}^{m} p_k\right)\left(\displaystyle\sum_{k=1}^{m} W_k\right)^2}, \\ \text{Set } \hat{\lambda}_m = \frac{1}{\sum_{k=1}^{m} p_k} \displaystyle\sum_{k=1}^{m} p_k V_k - \frac{\hat{\alpha}_m}{\sum_{k=1}^{m} p_k} \sum_{k=1}^{m} p_k W_k \\ \hat{\beta}_m = \dfrac{\displaystyle\sum_{k=1}^{m} p_k z_k S_k}{\displaystyle\sum_{k=1}^{m} p_k z_k B_k}. \end{cases} \tag{5}$$

The proof of Theorem 3.1 is presented in Section 5, encompassing the scenarios of Stable driven CIR, OU and Lotka-Volterra processes.

Let's discuss about some formulations in order to check some consistency properties of these linear least squares estimators through a linear statistical regression model. Since (V_k) is a random sequence and $|\hat{\varphi}_n(z_k)| \in [0,1]$ one may say that $V = (V_k)_{k=1}^m$ for m well-chosen values of z_k can be obtained from a truncated uniform random variable U so that $-\log(U)$ follows a truncated exponential random variable T where occurrences is limited to finite positive values $[0, A]$ with $A > 0$. Hence V is a truncated distribution of $\log(T)$ such that $\mathbb{E}(\log(T)\mathbb{I}_{[0,A]})$ and $\mathrm{Var}(\log(T)\mathbb{I}_{[0,A]})$ are finite. When the number of observations m is sufficiently large, a Gaussian distribution can be assumed for the error $\epsilon_k = V_k - \alpha \log(|z_k|) - \log(\rho^\alpha \Delta_n)$ that can be considered to be uncorrelated with the (deterministic) regressors $\log(|z_k|)$. Hence the above estimators become Maximum Likelihood Estimators. When the normality assumption is no longer valid, we can try to reduce the data to Gaussian distributions by means of data transformations, and the symmetrization and standardization (since $\mathbb{E}(\log(T)\mathbb{I}_{[0,A]})$ and $\mathrm{Var}(\log(T)\mathbb{I}_{[0,A]})$ are finite) offers an advantage. The Min-Max Scaling can be applied since the data may vary in different scales, reducing the effect of outliers.

The optimal value of m is suggested in [21], advocating the selection of points z_k within the interval $[0.1, 1]$.

Theorem 3.1 applies particularly to the following α-stable-driven Ornstein Uhlenbeck (OU), Cox–Ingersoll–Ross (CIR) and Lotka–Volterra processes. Such models are popular in stochastic modelling for description of interest rates in finance and population dynamics. Assume that Z is a standard

strictly stable process with parameters $\alpha \in (1,2)$ and $\beta \in [-1,1]$. A generalized Ornstein-Uhlenbeck (OU) process driven by a strictly standard α-stable process $(Z_t)_{t\geq 0}$ is defined to be the solution to the following linear stochastic differential equation

Definition 3.1. *(Stable OU process)*

$$dX_t = \theta(\mu - X_t)dt + \rho dZ_t, \quad X_0 = x \in \mathbb{R}. \tag{6}$$

where θ and μ are constants.

We can apply Theorem 3.1 to this model to estimate parameters ρ, α and β.

Definition 3.2. *(Stable CIR process) A stable driven Cox–Ingersoll–Ross (Stable CIR* [53]*) is defined by:*

$$dX_t = \theta(\mu - X_t)dt + \rho|X_t|^{1/q}dZ_t, \quad X_0 \geq 0, \tag{7}$$

where $q > 0$ $\theta > 0$, $\rho > 0$, and $\mu \in \mathbb{R}$ are constants and Z is for example a strictly standard α-stable process with positive jump activity.

Note that if Z is symmetric and since $x \to |x|^{1/q}$ is Hölder continuous for $q \geq 1$, there exists a solution according to [5]. For $q = 2$ the Stable CIR is studied in [51] for symmetric stable process. According to [27], there is a pathwise unique positive solution for the above Stable CIR Strong as $q^{-1} + \alpha^{-1} \geq 1$ when ρ is small whenever the process Z have only positive jumps.We can apply Theorem 3.1 to this model to estimate parameters ρ, α and β when the parameter q is known.

For $q = \alpha \in (1,2)$ and in the case where Z is a pure-jump α-stable Lévy process Z with positive jumps the following stable driven CIR is introduced in the literature (see [6] or [26]) as a particular form of the continuous-state branching processes with immigration, which emerge as scaling limits of Galton–Watson branching processes with immigration (CBI-processes, [25], [26] and [39]).

Definition 3.3. *(SCIR process) The SCIR process is defined by:*

$$dX_t = \theta(\mu - X_t)dt + \rho|X_t|^{1/\alpha}dZ_t, \quad X_0 \geq 0, \tag{8}$$

where $\theta > 0$, $\rho > 0$, and $\mu \in \mathbb{R}$ are constants and Z is a standard and positive strictly α-stable process.

It is shown whenever that the solution is positive that is $\mathbb{P}\Big(X_t > 0\Big) = 1$, whenever $\mu > 0$, $\theta > 0$, $\rho > 0$ and $x_0 > 0$. Unfortunately, we cannot estimate α with this model, as q is a function of α. Only the parameters ρ and β must be unknown to use Theorem 3.1.

The following stable driven Lotka and Volterra extension was studied in [56] for $q = 1$:

Definition 3.4. *(A Stable driven Lotka–Volterra process)*
A Stable driven Lotka–Volterra process can be defined as follows

$$dX_t = X_t(\lambda - \theta X_t)dt + \rho|X_t|^{1/q}dZ_t, \quad X_0 \geq 0, \tag{9}$$

where λ, ρ, θ are real constants.

The author proved in [56] for $q = 1$, the existence of a unique global positive solution of the above stable driven Lotka–Volterra process, if Z is a spectrally positive strictly α-stable process for any $\alpha \in (0, 2)$.

A simulation study is conducted to evaluate the performance of the proposed estimation of the diffusion parameters in Section 5 from data generated by ergodic stationary processes. We also revisit the link between ergodicity and mixing conditions in Section 6. In what follows, we discuss the parameter estimation of the drift coefficients θ and μ of the above stable driven OU and the SCIR processes.

Theorem 3.2. *Let $X = (X_t)_{t\in[0,T]}$ be a real valued stationary and strongly mixing (ergodic) process satisfying (6) or (8). We assume that X is observed on a sample of discrete time points t_i with time frequency Δ_n of size n. Assume that assumptions $(\mathbf{A_2})$ and $(\mathbf{A_3})$ holds. Let $(\hat{f}_n(X_{t_i}))_{i=1}^n$ be the sequence of the N-W estimate of the linear drift function in (6) or (8) and let $a_n = \theta e^{-\theta\Delta_n}$. Set*

$$\begin{cases} \hat{a}_n &= \dfrac{\sum_{i=0}^{n-1} X_{t_i}\hat{f}_n(X_{t_{i+1}}) - \frac{1}{n}\sum_{i=0}^{n-1} X_{t_i}\sum_{i=1}^{n}\hat{f}_n(X_{t_{i+1}})}{\frac{1}{n}\big(\sum_{i=0}^{n-1} X_{t_i}\big)^2 - \sum_{i=0}^{n-1} X_{t_i}^2} \\ \hat{\mu}_n &= \frac{1}{n}\sum_{i=0}^{n-1}\hat{f}_n(X_{t_{i+1}}) - \hat{a}_n\frac{1}{n}\sum_{i=0}^{n-1} X_{t_i} \\ \hat{\theta}_n &= -\frac{1}{\Delta_n}W(-\hat{a}_n\Delta_n), \end{cases} \tag{10}$$

where W is the Lambert function.

- *We assume moreover that $\alpha \in (\sqrt{2}, 2)$ for the driving process Z in the SCIR process (8).*

For n large enough, $\hat{\theta}_n$ and $\hat{\mu}_n$ are unbiased and consistent estimators of θ and μ.

4. Proofs of the main Theorems

In this section, we provide the proofs for our main results.

4.1. **Proof of Theorem 3.1.**

Proof. We begin by revisiting the Euler approximation of the SDE (1):

$$X_{t_i}^n = X_{t_{i-1}}^n + f(X_{t_i}^n)\Delta_n + g(X_{t_i}^n)Y_i^n$$

where $Y_i^n = \rho\Delta Z_i^n = \rho(Z_{i+1}^n - Z_i^n)$, $i = 0, \ldots, n-1$, are identically sequence distributed as $Z_{\Delta_n} \sim S_\alpha\Big(\Delta_n^{1/\alpha}\rho, \beta, 0\Big)$. We have

$$Y_i^n = \frac{\Delta X_{t_i} - \hat{f}_n(X_{t_i})\Delta_n}{g(X_{t_i})}, \quad \forall i = 0, \ldots, n-1.$$

We can define an approximate sample characteristic function as follows:

$$\hat{\varphi}_n(u) = \frac{1}{n}\sum_{i=0}^{n-1} \exp\{ju\Big(\frac{\Delta X_{t_i} - \hat{f}_n(X_{t_i})\Delta_n}{g(X_{t_i})}\Big)\}, \quad \forall u \in \mathbb{R}, \tag{11}$$

which shall be asymptotically equal to the characteristic function $\varphi_{Z_{\Delta_n}}(u)$ of Z_{Δ_n}. Now, using Definition 2.1 or 2.2 and, utilizing the following formula:

$$\log(-\log|\varphi_{Z_{\Delta_n}}(u)|) = \log(\rho^\alpha\,\Delta_n) + \alpha\log(|u|),$$

we can establish an ordinary least squares regression method to estimate α and ρ. To achieve this, it is desired to find the vector parameter (α, λ) such that the underlined linear function fits best the given data in the least squares sense, that is, the following sum of squares is minimized:

$$G(\lambda, \alpha) = \sum_{k=1}^{m} (V_k - \lambda - \alpha W_k)^2$$
$$(\hat{\lambda}, \hat{\alpha}) = \arg\min_{(\lambda,\alpha)} G(\lambda, \alpha),$$

where

$$V_k = \log(-\log|\hat{\varphi}_n(u_k)|), \quad W_k = \log(|u_k|), \qquad \lambda = \log(\rho^\alpha\,\Delta_n).$$

The minimum value of $G(\lambda, \alpha)$ occurs when the gradient is zero. Since the model contains two parameters, there are two gradient equations:

$$\begin{cases} \frac{\partial G(\lambda,\alpha)}{\partial \alpha} = -2\sum_{k=1}^{m} W_k(V_k - \lambda - \alpha W_k) \\ \frac{\partial G(\lambda,\alpha)}{\partial \lambda} = -2\sum_{k=1}^{m}(V_k - \lambda - \alpha W_k), \end{cases}$$

and these gradient equations have a closed solution given by:

$$\begin{cases} \frac{\partial G(\lambda,\alpha)}{\partial \alpha} = 0 \\ \frac{\partial G(\lambda,\alpha)}{\partial \lambda} = 0 \end{cases} \Longrightarrow \begin{cases} \hat{\alpha}_m = \dfrac{\sum_{k=1}^{m} W_k V_k - \frac{1}{m}\sum_{k=1}^{m} V_k \sum_{k=1}^{m} W_k}{\sum_{k=1}^{m} W_k^2 - \frac{1}{m}\big(\sum_{k=1}^{m} W_k\big)^2}, \\ \hat{\lambda}_m = \frac{1}{m}\sum_{k=1}^{m} V_k - \frac{\hat{\alpha}_m}{m}\sum_{k=1}^{m} W_k. \end{cases} \tag{12}$$

To estimate β, we use Definition 2.2 and the following formula:

$$S_k = \beta B_k + \omega u_k \quad \text{here we have, } \omega = 0$$

where

$$S_k = \arg(\hat{\varphi}_n(u_k)), \quad B_k = \tan(\frac{\pi\hat{\alpha}}{2}) sign(u_i)\big(u_k - u_k^{\hat{\alpha}}\big)\Delta_n^{1/\hat{\alpha}}.$$

Note that we set $\omega = 0$ since we consider in this work strictly α-stable process Z. We employ once again the aforementioned least squares method to minimize the following sum of squares:

$$T(\beta,\omega) = \sum_{k=1}^{m} \big(S_k - \beta B_k - \omega u_k\big)^2.$$

From the gradient equations:

$$\begin{cases} \frac{\partial T(\beta,\omega)}{\partial \beta} = -2\sum_{k=1}^{m} B_k(S_k - \omega u_k - \beta B_k) \\ \frac{\partial T(\beta,\omega)}{\partial \omega} = -2\sum_{k=1}^{m} u_k(S_k - \omega u_k - \beta B_k), \end{cases}$$

we have the following closed solution when $\omega = 0$:

$$\hat{\beta}_m = \frac{\sum_{k=1}^{m} u_k S_k}{\sum_{k=1}^{m} u_k B_k}.$$

□

4.2. **Proof of Theorem 3.2.**

Proof. We have

$$\hat{f}(X_{t_{i+1}}) = \theta(\mu - X_{t_{i+1}}).$$

If X satisfy (6) we have the following integral representation:

$$X_t = e^{-\theta t}X_0 + \theta\mu\int_0^t e^{-\theta(t-s)}ds + \rho\int_0^t e^{-\theta(t-s)}dZ_s, \quad t \geq 0.$$

so that for any $t \geq r \geq 0$:

$$X_t = e^{-\theta(t-r)}X_r + \theta\mu \int_r^t e^{-\theta(t-s)}ds + \rho \int_r^t e^{-\theta(t-s)}dZ_s, \quad t \geq 0. \tag{13}$$

We refer to [20], Chapter 3 for the case when $\mu = 0$ and to [15] for the case $\mu \neq 0$.
If X satisfy (8), note that for $\alpha \in (1, 2)$ it is well known that the positive stable process Z admits the following representation:

$$Z_t = \int_0^t \int_0^{+\infty} z\, \tilde{N}(ds, dz), \quad t \geq 0.$$

and applying Itô's formula for Jump processes [42], we have the following integral representation (see also [26]) :

$$X_t = e^{-\theta(t-r)}X_r + \theta\mu \int_r^t e^{-\theta(t-s)}ds + \rho \int_r^t e^{-\theta(t-s)}|X_{s-}|^{1/\alpha}dZ_s, \quad t \geq 0. \tag{14}$$

We have

$$X_{t_i} = e^{-\theta(t_i - t_{i-1})}X_{t_{i-1}} + \theta\mu \int_{t_{i-1}}^{t_i} e^{-\theta(t_i - s)}ds + \epsilon_i, \tag{15}$$

where

$$\epsilon_i = \begin{cases} \rho \int_{t_{i-1}}^{t_i} e^{-\theta(t_i - s)}|X_{s-}|^{1/\alpha}dZ_s & \text{if X is the stable CIR process in (7)} \\ \rho \int_{t_{i-1}}^{t_i} e^{-\theta(t_i - s)}dZ_s & \text{if X is the stable OU process in (6).} \end{cases}$$

Note that if X is the stable OU process in (6), by basic properties of α-stable stochastic integrals for deterministic integrands [48, 43], the sequence (ϵ_i) is a α-stable random variable with distribution $S_\alpha\big(\rho\tau_n^{1/\alpha}, \beta, 0; 1\big)$ with $\tau_n = \frac{1-e^{-\theta\alpha\Delta_n}}{\theta\alpha}$ so that it is centered random sequence (also martingales difference). In the case where X is SCIR process in (8), according to the martingale property of a compensated Poisson stochastic integral it is a martingale differences if $\mathbb{E}|X_s|^{2/\alpha}$ is finite; which is the case whenever $\frac{2}{\alpha} < \alpha$ i.e. $\alpha \in (\sqrt{2}, 2)$. Therefore,

$$X_{t_i} = e^{-\theta\Delta_n}X_{t_{i-1}} + \mu\Big(1 - e^{-\theta\Delta_n}\Big) + \epsilon_i,$$

where (ϵ_i) is a centered random sequence (martingale differences). Finally,

$$\hat{f}(X_{t_{i+1}}) = \theta(\mu - X_{t_{i+1}}) = a\mu - aX_{t_i} + \eta_i, \quad a = \theta e^{-\theta\Delta_n}, \quad i \geq 0,$$

where $(\eta_i)_{i\geq 0}$ is again a centered random sequence (martingale differences). We can set up a Linear Least Square Method which consists in minimizing the following objective function:

$$\mathbb{G}(a, \mu) = \sum_{i=0}^{n-1} \big(\hat{f}(X_{t_{i+1}}) - a\mu + aX_{t_{i+1}}\big)^2.$$

Using, the gradient equations:

$$\begin{cases} \frac{\partial \mathbb{G}(a,\mu)}{\partial a} = 2(-\mu + X_{t_i}) \sum_{i=0}^{n-1} \Big(\hat{f}(X_{t_{i+1}}) - a\mu + aX_{t_i} \Big) \\ \frac{\partial \mathbb{G}(a,\mu)}{\partial \mu} = -2a \sum_{i=0}^{n-1} \Big(\hat{f}(X_{t_{i+1}}) - a\mu + aX_{t_i} \Big), \end{cases}$$

it is straightforward to obtain the following least square estimators :

$$\begin{cases} \hat{a}_n &= \dfrac{\sum_{i=0}^{n-1} X_{t_i} \hat{f}_n(X_{t_{i+1}}) - \frac{1}{n} \sum_{i=0}^{n-1} X_{t_i} \sum_{i=1}^{n} \hat{f}_n(X_{t_{i+1}})}{\frac{1}{n} \big(\sum_{i=0}^{n-1} X_{t_i} \big)^2 - \sum_{i=0}^{n-1} X_{t_i}^2} \\ \hat{\mu}_n &= \hat{a}_n^{-1} \frac{1}{n} \sum_{i=0}^{n-1} \hat{f}_n(X_{t_{i+1}}) + \dfrac{1}{n} \sum_{i=0}^{n-1} X_{t_i} \\ \hat{\theta}_n &= -\frac{1}{\Delta_n} W(-\hat{a}_n \Delta_n), \end{cases} \tag{16}$$

where W is the Lambert function [9]. Now, let us prove the consistency of these estimators. Set $U_i = \hat{f}_n(X_{t_{i+1}})$ and note that since we assume that the process X is ergodic and $\alpha \in (1,2)$ then $\bar{X} = \frac{1}{n} \sum_{i=0}^{n-1} X_{t_i}$ converge to μ and $\bar{U} = \frac{1}{n} \sum_{i=0}^{n-1} U_i$ also converge. We have

$$\hat{a}_n = \frac{\sum_{i=1}^{n} X_{t_i} U_i - \frac{1}{n} \sum_{i=1}^{n} U_i \sum_{i=1}^{n} X_{t_i}}{\frac{1}{n} \sum_{i=1}^{n} X_{t_i} \sum_{i=1}^{n} X_{t_i} - \sum_{i=1}^{n} X_{t_i}^2} = \frac{\sum_{i=1}^{n} X_{t_i} \Big(U_i - \bar{U} \Big)}{\sum_{i=1}^{n} X_{t_i} \Big(\bar{X} - X_{t_i} \Big)} \tag{17}$$

$$\text{where} \quad \bar{U} = \frac{1}{n} \sum_{i=1}^{n} U_i \text{ and } \bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_{t_i}, \tag{18}$$

so that

$$= \frac{\sum_{i=1}^{n} \Big(X_{t_i} - \bar{X} \Big) \Big(\bar{U} - U_i \Big)}{\sum_{i=1}^{n} \Big(X_{t_i} - \bar{X} \Big)^2} \text{since} \quad \sum_{i=1}^{n} \bar{X} (U_i - \bar{U}) = 0 \text{ and } \sum_{i=1}^{n} \bar{X} (X_{t_i} - \bar{X}) = 0. \tag{19}$$

From the relation

$$U_i - U = a(\bar{X} - X_{t_i}) + \varepsilon_i,$$

where (ε_i) is again a centered random sequence, we have

$$\hat{a}_n = \frac{\sum_{i=1}^{n}\Big(X_{t_i}-\bar{X}\Big)\Big(a(X_{t_i}-\bar{X})-\varepsilon_i\Big)}{\sum_{i=1}^{n}\Big(\bar{X}-X_{t_i}\Big)^2} = a + \frac{\sum_{i=1}^{n}\Big(X_{t_i}-\bar{X}\Big)\varepsilon_i}{\sum_{i=1}^{n}\Big(\bar{X}_t-X_{t_i}\Big)^2} \underset{n\to+\infty}{\longrightarrow} 0, \tag{20}$$

since the process $X = \{X_t, t \geq 0\}$ is ergodic, stationary and have infinite square variation. We deduce the consistency of $\hat{\theta}_n$ and $\hat{\mu}_n$ by the use of continuous mapping theorem. □

5. Numerical Application with ergodic stochastic models

We now scrutinize the numerical performance of our estimation procedure on synthetic data. Subsequently, we apply these procedures to real-world data. All the implemented codes were developed using the R software. For practical applications, we determine the optimal bandwidth $h = (h_n)$ using the method proposed by [45]. The R function `h.amize` facilitates this process. It is worth noting that, in many instances, people commonly opt for $h = h_n = n^{-1/5}$. A simulation study is conducted to evaluate the performance of the proposed estimation from data generated by ergodic stationary processes. We discuss in Appendix 6 the link between ergodicity and mixing conditions for Markov processes. We summarize the discussion as follows.

The exponential ergodicity of a Lévy driven OU process is established in [49]. The result implies that if $\theta > 0$ then the α-Stable OU process (6) has a unique invariant measure π and is strongly ergodic (mixing). More generally, let's consider the case where the drift function f and diffusion function g satisfy sufficient conditions for the solution of (1) to exist and be unique [5, 14]. According to [22], if $f(\cdot)$ is locally Lipschitz, and $\limsup_{|x|\to+\infty} \frac{f(x)}{x} < 0$, then, for any $\alpha \in (1,2)$ and bounded function g, the solution of (1) is exponentially ergodic, and its invariant distribution exists and is unique. This result applies to the case of the stable OU model (6) under the condition $\theta > 0$.

In the case of stable driven CIR-models, some ergodicity conditions are well known when the driving process has only positive jumps. For $\theta > 0$ and $\mu \geq 0$, this model can be seen as a subcritical CBI process with an immigration rate μ. Thus, using the result of [26], we conclude that the SCIR process is exponentially ergodic and hence strongly mixing.

For a more general SDE in the form (1), it was shown in [54], that the exponential ergodicity holds under some dissipative and non-degenerate assumptions on the drift f and diffusion function g. In terms of forthcoming

study, we are currently interested in ergodicity properties for SDE in the form (1) under more realistic conditions.

5.1. **Simulated data.** We simulate and approximate the solution $X = (X_t)_{t\geq 0}$ by using the Euler scheme on the interval $[0, T]$ with sample size $n = 500, 1000, 2000$ within a period $T = 1$. We employ the exponential ergodicity results when choosing the model parameters. We consider the following parameters:

$$x_0 = 1, \quad \theta = 0.3, \quad \mu = 1.2 \quad \alpha = 1.6 \quad \beta = 0 \,\text{or}\, 0.1 \quad \text{and} \quad \rho = 1.$$

Thus, Assumption $(\mathbf{A_1})$ is verified for model (6) for general strictly driving α-stable process Z. For model (8), we consider positive jump driving stable process Z so that Assumption $(\mathbf{A_1})$ is once again verified for $\beta > 0$. We choose $q = 2$ for model (7). For model (9) and according to [56], we choose (in order to have strongly mixing solution)

$$x_0 = 1, \quad \theta = 0.5, \quad \mu = 0.5 \quad \alpha = 1.5 \quad \beta = 0.1 \quad \text{and} \quad \rho = 0.5.$$

We generate data from models (6), (7) or (8) and (9) based on the previous configuration. To numerically validate the regularity condition on the density function of the stationary distribution π, assumed to be continuous in Assumption $(\mathbf{A_2})$, we plot the kernel density estimate of a realization of X alongside the histogram. Choosing a suitable time frequency (e.g., $\Delta_n = \frac{1}{\sqrt{n}}$), ensures that the models satisfy all the necessary conditions.

We assess the performance of the N-W estimator for the drift function and the estimated parameters of the diffusion part (α, β, ρ) in the context of the stable driven CIR, OU and Lotka-Volterra models on simulated data. Figures 1, 2 and 6 present the graphical performance of the N-W estimator concerning the true drift functions. The comparisons are made across three sample sizes $n = 500$, $n = 1000$ and $n = 2000$ with time interval $T = 1$ or $T = 10$. It is observed that varying T for a fixed n does not significantly improve the estimates of the drift function in the SCIR model. This confirms that the drift function cannot be precisely identified within a fixed time interval using the N-W estimator, regardless of how frequently the observations are sampled. Additionally, for a fixed length of the observation interval T, as the sample size increases, the N-W estimator does not exhibit better performance, aligning with the asymptotic theory of the N-W estimator for stochastic processes driven by Lévy motions. Figures 3 and 4 visually confirm the regularity condition on the density function of the stationary distribution. The performance of the estimation in 5 for $q = 2$ seems to imply that the model in (8) is ergodic. We have not found any references proving this. An actual forthcoming work is the study of the erogidicity of SDE 1 under various conditions. To assess the performance, we employ not only visual illustrations through figures but also quantitative measures such

as the Square Root of Average Square Errors, defined by:

$$\text{RMSE} = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{f}_n(x_i) - f(x_i)\right)^2$$

where $f(x_i)$ and $\hat{f}_n(x_i)$ are transformed (Min-Max scaling) to cover the range of sample paths of X in a common scale, reducing the effect of outliers.

We summarized the performance in Tables 1, 5 and 4, which reports the results on the diffusion parameters (α, β, ρ) as well as the RMSE on n replicates with three sample sizes $n = 500,\ n = 1000$ and $n = 2000$, respectively for time interval $T = 1$ and $T = 10$. We can see that varying T for a fixed n slightly changes the estimates of the stable process parameters α, β, except for the scaling parameter ρ.

5.2. **Real data: financial exchanges rates.** In this section, we shift our focus to the estimation using real data. We analyze financial data, specifically indexes or exchanges of the Canadian dollar against the US dollar, spanning a fixed period from May 2018 to June 2022. The analysis is conducted with a sample size of $n = 1000$.

Figures 7 (a), (b), and (c) present the time series variations of the exchange rates. To assess the stationarity of the time series, we examine the autocorrelation function (ACF) diagram [3]. A rapidly decreasing ACF indicates stationarity, and the Dickey Fuller test confirms this with a p-value of 0.01. The ACF diagram demonstrates a quick decline, indicating weak dependence and implying the data's mixing properties, specifically -mixing, which implies strong mixing.

Figures 5(a), (b) and (c) present the time series variations of the exchange rates. To assess the stationarity of the time series, we examine the autocorrelation function (ACF) diagram [3]. A rapidly decreasing ACF indicates stationarity, and the Dickey Fuller test confirms this with a p-value of 0.01. The ACF diagram demonstrates a quick decline, indicating weak dependence and implying the data's mixing properties, specifically ρ-mixing, which implies strong mixing. The density estimation of the stationary distribution is given also in Figures 5 (a'), 5(b') and 5(c') so that the regularity condition is satisfied.

In Figures 7, we present the graphical N-W estimation of the observed drift function. To summarize the parameter estimation of the diffusion part, Tables 3 and 6 provide detailed results. Regarding the drift part, Tables 2 summarize the estimated drift coefficients based on real data. Additionally, in Figure 8, we compare the prediction results with the stable OU model

since the N-W estimator of the drift function seems to be linear and match better than the Stable CIR process.

Funding

No funding was obtained for this study.

Competing Interest

The authors declare that they have no competing interests.

Declaration statement

The author certify that the submission is an original work and is not under review at any other publication.

Data availability statements

The real dataset used in this work is available online and free of charge. The R codes for the simulation study are available from the author on reasonable request.

References

[1] Yacine Aït-Sahalia. Maximum likelihood estimation of discretely sampled diffusions: a closed-form approximation approach. *Econometrica*, 70(1):223–262, 2002.

[2] Linda JS Allen. Stochastic population and epidemic models. *Mathematical biosciences lecture series, stochastics in biological systems*, page 128, 2015.

[3] Alexander Alvarez and Pablo Olivares. Méthodes d'estimation pour des lois stables avec des applications en finance. *Journal de la société française de statistique*, 146(4):23–54, 2005.

[4] David Applebaum. *Lévy processes and stochastic calculus*. Cambridge university press, 2009.

[5] Richard F Bass, Krzysztof Burdzy, and Zhen-Qing Chen. Stochastic differential equations driven by stable processes for which pathwise uniqueness fails. *Stochastic processes and their applications*, 111(1):1–15, 2004.

[6] Elise Bayraktar and Emmanuelle Clément. Estimation of a pure-jump stable cox-ingersoll-ross process. *arXiv preprint arXiv:2304.02386*, 2023.

[7] Patrick Cattiaux and S. Mawaki Manou-Abi. Limit theorems for some functionals with heavy tails of a discrete time markov chain. *ESAIM: Probability and Statistics*, 18:468–482, 2014.

[8] John M Chambers, Colin L Mallows, and BW4159820341 Stuck. A method for simulating stable random variables. *Journal of the american statistical association*, 71(354):340–344, 1976.

[9] Robert M Corless, Gaston H Gonnet, David EG Hare, David J Jeffrey, and Donald E Knuth. On the lambert w function. *Advances in Computational mathematics*, 5:329–359, 1996.

[10] John C Cox, Jonathan E Ingersoll Jr, and Stephen A Ross. A theory of the term structure of interest rates. In *Theory of valuation*, pages 129–164. World Scientific, 2005.

[11] Peter Craigmile, Radu Herbei, Ge Liu, and Grant Schneider. Statistical inference for stochastic differential equations. *Wiley Interdisciplinary Reviews: Computational Statistics*, 15(2):e1585, 2023.

[12] Niklas Dexheimer and Claudia Strauch. On lasso and slope drift estimators for l\'evy-driven ornstein–uhlenbeck processes. *arXiv preprint arXiv:2205.07813*, 2022.

[13] A Ja Dorogovcev. The consistency of an estimate of a parameter of a stochastic differential equation. *Theory Probab. Math. Stat.*, 10:73–82, 1976.

[14] Nicolas Fournier. On pathwise uniqueness for stochastic differential equations driven by stable lévy processes. In *Annales de l'IHP Probabilités et statistiques*, volume 49, pages 138–159, 2013.

[15] Yaozhong Hu and Hongwei Long. Parameter estimation for ornstein-uhlenbeck processes driven by stable lévy motions. *Communications on Stochastic Analysis*, 1(2):1, 2007.

[16] Yaozhong Hu and Hongwei Long. Least squares estimator for ornstein–uhlenbeck processes driven by α-stable motions. *Stochastic Processes and their applications*, 119(8):2465–2480, 2009.

[17] Stefano M Iacus et al. *Simulation and inference for stochastic differential equations: with R examples*, volume 486. Springer, 2008.

[18] Aleksander Janicki, Zbigniew Michna, and Aleksander Weron. Approximation of stochastic differential equations driven by stable lévy motion. *Applicationes Mathematicae*, 24(2):149–168, 1997.

[19] RA Kasonga. The consistency of a non-linear least squares estimator from diffusion processes. *Stochastic processes and their applications*, 30(2):263–275, 1988.

[20] Sato Ken-Iti. *Lévy processes and infinitely divisible distributions*, volume 68. Cambridge university press, 1999.

[21] Stephen M Kogon and Douglas B Williams. Characteristic function based estimation of stable distribution parameters. *A practical guide to heavy tails: statistical techniques and applications*, pages 311–338, 1998.

[22] Alexey M Kulik. Exponential ergodicity of the solutions to sde's with a jump noise. *Stochastic Processes and their Applications*, 119(2):602–632, 2009.

[23] Yu A Kutoyants. *Statistical inference for ergodic diffusion processes*. Springer Science & Business Media, 2004.

[24] A Le Breton. On continuous and discrete sampling for parameter estimation in diffusion type processes. In *Stochastic Systems: Modeling, Identification and Optimization, I*, pages 124–144. Springer, 2009.

[25] Pei-Sen Li, Zenghu Li, Jian Wang, and Xiaowen Zhou. Exponential ergodicity of branching processes with immigration and competition. *arXiv preprint arXiv:2205.15499*, 2022.

[26] Zenghu Li and Chunhua Ma. Asymptotic properties of estimators in a stable cox–ingersoll–ross model. *Stochastic Processes and their Applications*, 125(8):3196–3233, 2015.

[27] Zenghu Li and Leonid Mytnik. Strong solutions for stochastic differential equations with jumps. In *Annales de l'IHP Probabilités et statistiques*, volume 47, pages 1055–1067, 2011.

[28] ZhengYan Lin, YuPing Song, and JiangSheng Yi. Local linear estimator for stochastic differential equations driven by stable lévy motions. *Science China Mathematics*, 57:609–626, 2014.

[29] Robert Sevilevic Lipcer, Robert S Liptser, Albert Nikolaevic Shiraev, Albert N Shiraev, et al. *Statistics of Random Processes II: II. Applications*, volume 2. Springer Science and Business Media, 2001.

[30] Hongwei Long and Lianfen Qian. Nadaraya-Watson estimator for stochastic processes driven by stable Lévy motions. *Electronic Journal of Statistics*, 7(none):1387 – 1418, 2013.

[31] Ross A Maller, Gernot Müller, and Alex Szimayer. Ornstein–uhlenbeck processes and extensions. *Handbook of financial time series*, pages 421–437, 2009.
[32] Solym Manou-Abi. Approximate solution for a class of stochastic differential equation driven by stable processes. *Preprint hal-04217398*, 2023.
[33] Solym Mawaki Manou-Abi. *Théorèmes limites et ordres stochastiques relatifs aux lois et processus stables*. PhD thesis, Universite Toulouse 3 Paul Sabatier (UT3 Paul Sabatier), 2015.
[34] H Masuda. Simple estimators for non-linear markovian trend from sampled data: I. ergodic cases. *MHF Preprint Series*, 7, 2005.
[35] Remigijus Mikulevičius and Fanhui Xu. On the rate of convergence of strong euler approximation for sdes driven by levy processes. *Stochastics*, 90(4):569–604, 2018.
[36] Elizbar A Nadaraya. On estimating regression. *Theory of Probability & Its Applications*, 9(1):141–142, 1964.
[37] John P Nolan. *Univariate stable distributions*. Springer, 2020.
[38] Olivier Menoukeu Pamen and Dai Taguchi. Strong rate of convergence for the euler–maruyama approximation of sdes with hölder continuous drift coefficient. *Stochastic Processes and their Applications*, 127(8):2542–2559, 2017.
[39] Étienne Pardoux. *Probabilistic Models of Population Evolution: Scaling Limits, Genealogies and Interactions*, volume 1. Springer, 2016.
[40] BLS Prakasa Rao. Asymptotic theory for non-linear least squares estimator for diffusion processes. *Statistics: A Journal of Theoretical and Applied Statistics*, 14(2):195–209, 1983.
[41] BLS Prakasa Rao. Nonparametric estimation of linear multiplier in stochastic differential equations driven by stable noise. *arXiv e-prints*, pages arXiv–2109, 2021.
[42] N Privault. Stochastic calculus for jump processes. *Unpublished working paper. Nanyang Technological University (http://www. ntu. edu. sg/home/nprivault/indext. html)*, 2016.
[43] Jan Rosinski and WA Woyczynski. On itô stochastic integration with respect to p-stable motion: inner clock, integrability of sample paths, double and multiple integrals. *The Annals of Probability*, pages 271–286, 1986.
[44] Gennady Samorodnitsky, Murad S Taqqu, and RW Linde. Stable non-gaussian random processes: stochastic models with infinite variance. *Bulletin of the London Mathematical Society*, 28(134):554–555, 1996.
[45] Simon J Sheather. Density estimation. *Statistical science*, pages 588–597, 2004.
[46] Yasutaka Shimizu. M-estimation for discretely observed ergodic diffusion processes with infinitely many jumps. *Statistical Inference for Stochastic Processes*, 9:179–225, 2006.
[47] Yasutaka Shimizu and Nakahiro Yoshida. Estimation of parameters for diffusion processes with jumps from discrete observations. *Statistical Inference for Stochastic Processes*, 9:227–277, 2006.
[48] Murad S Taqqu. *Stable non-Gaussian random processes: stochastic models with infinite variance*. Chapman & Hall, 1994.
[49] Jian Wang. On the exponential ergodicity of lévy-driven ornstein–uhlenbeck processes. *Journal of Applied Probability*, 49(4):990–1004, 2012.
[50] Geoffrey S Watson. Smooth regression analysis. *Sankhyā: The Indian Journal of Statistics, Series A*, pages 359–372, 1964.
[51] Chao Wei. Estimation for the discretely observed cox–ingersoll–ross model driven by small symmetrical stable noises. *Symmetry*, 12(3):327, 2020.
[52] Wei Biao Wu. Nonparametric estimation for stationary processes. *University of Chicago. Technical Report*, 536, 2003.
[53] Xu Yang. Maximum likelihood type estimation for discretely observed cir model with small α-stable noises. *Statistics & Probability Letters*, 120:18–27, 2017.

[54] Xiaolong Zhang and Xicheng Zhang. Ergodicity of supercritical sdes driven by α-stable processes and heavy-tailed sampling. *Bernoulli*, 29(3):1933–1958, 2023.

[55] Xuekang Zhang, Haoran Yi, and Huisheng Shu. Nonparametric estimation of the trend for stochastic differential equations driven by small α-stable noises. *Statistics & Probability Letters*, 151:8–16, 2019.

[56] Zhenzhong Zhang, Xuekang Zhang, and Jinying Tong. Exponential ergodicity for population dynamics driven by α-stable processes. *Statistics & Probability Letters*, 125:149–159, 2017.

6. Appendix, Figures and Tables

Ergodic Markov processes and mixing conditions. We present in this section, some results concerning the link between ergodicity and mixing conditions for Markov processes. Let $(X_t, t \geq 0)$ be an ergodic Markov process with unique invariant measure π. It is well known that solution of some classical SDE driven by Lévy processes are Markov processes. Consider the Markov semigroup $(P_t)_{t\geq 0}$ associated to $(X_t, t \geq 0)$ and defined by $P_t f(x) = \mathbb{E}\big(f(X_t)/X_0 = x\big)$ for all f in $L^p(\pi)$ or measurable and positive functions f. Recall that P is a bounded operator in all $L^p(\pi)$, $p \geq 1$ with operator norm equal to 1 (i.e. a contraction). The adjoint operator P^* is defined by $\int f\, P_t g\, d\pi = \int g\, P_t^* f\, d\pi$. For functions f and g that are square integrable with respect to π, this operator is once more a contraction. The subsequent definition introduces a method for regulating the ergodic decay to equilibrium.

Definition 6.1 (Ergodic rates of convergence [7])**.** *For any $r \geq p \geq 1$ and $t \geq 0$ we define the following ergodic rates*

$$\eta_{p,r}(t) = \sup_{||f||_{L^r(\pi)} \int f d\pi = 0} ||P_t f||_{L^p(\pi)}.$$

The process X is said to be uniformly ergodic if $\lim_{t\to+\infty} \eta_{2,\infty}(t) = 0$.

The following definitions can be found in [56].

Definition 6.2 (Exponentially or strongly ergodic process)**.** *Assume that $X = (X_t, t \geq 0)$ is an ergodic Markov process with unique invariant measure π and $X_0 = x$.*

1. *The process X with is called exponentially ergodic if there exist a constant $k > 0$ and a positive measurable function $c(x)$ such that for all $t > 0$ we have*
$$||P_t(x, .) - \pi||_{var} \leq c(x)e^{-kt},$$
2. *The process X with is called strongly ergodic if there exist two constants $k, C > 0$ such that for all $t > 0$ we have*
$$||P_t(x, .) - \pi||_{var} \leq Ce^{-kt}.$$

where $||.||_{var}$ denotes the total variation norm on the space of signed probability measures defined by

$$\begin{aligned} ||P_t(x, .) - \pi||_{Var} &= \sup_{A\in\mathcal{F}} |P_t(x, A) - \mu(A)| \\ &= \sup_{||f||_\infty \leq 1 \text{ and } Law(Y)=\pi} |\mathbb{E}f(X_t) - \mathbb{E}f(Y)|. \end{aligned}$$

From this definition, one can state that the process X is considered strongly or exponentially ergodic iflim$_{t\to+\infty} \eta_{1,\infty}(t) = 0$, where $\eta_{1,\infty}$ is termed the

strong or exponential ergodic decay rate. Now, let's introduce some conventional mixing coefficients to compare mixing properties and ergodic decay properties.

Definition 6.3 (Mixing conditions, see [33])**.** . *Let $\mathcal{F}_s$ (resp.$\mathcal{G}_s$) be respectively the backward (or the past) and the forward (or the future) σ-fields generated by X_u for $0 \leq u \leq s$ (resp. $u \leq s$).*

1. *The strong mixing coefficient $\alpha_{mix}(t)$ is defined as :*

$$\begin{aligned}\alpha_{mix}(t) &= \sup_s \sup_{A\in A\in\mathcal{F}_s,\, B\in\mathcal{G}_{s+t}} |\mathbb{P}(A\cup B)-\mathbb{P}(A)\mathbb{P}(B)| \\ &= \frac{1}{4}\sup_s \left\{ \sup_{F,G}\ Cov(F,G)\,,\ F\ \mathcal{F}_s(resp.\, G\ \mathcal{G}_{t+s})\ measurable\ and\ bounded\ by\ 1. \right\}\end{aligned}$$

   If $\lim_{t\to\infty}\alpha_{mix}(t)=0$, the process is strongly mixing.

2. *The β-mixing coefficient $\varphi(t)$ is defined as:*

$$\beta_{mix}(t)=\sup_s\left\{\sup_{A,B}\left(\mathbb{P}(B|A)-\mathbb{P}(B)\right)\ ,\ A\in\mathcal{F}_s\,,\ B\in\mathcal{G}_{s+t}\right\}.$$

   If $\lim_{t\to\infty}\beta_{mix}(t)=0$, the process is β-mixing or uniformly mixing.

3. *The ρ-mixing coefficient $\rho_{mix}(t)$ is defined as the maximal correlation coefficient, i.e.*

$$\rho_{mix}(t)=\sup_s\left\{\sup_{F,G}\ Corr(F,G)\,,\ F\in L^2(\mathcal{F}_s)\,,\ G\in L^2(\mathcal{G}_{t+s})\right\}.$$

   If $\lim_{t\to\infty}\rho_{mix}(t)=0$ the process is ρ-mixing.

The following lemma, as contained in [7], enables a connection between ergodicity and mixing conditions.

Lemma 6.1. *For all $t\geq 0$, we have*

(1) $\eta^2_{\infty,2}(t)\vee(\eta^*)^2_{\infty,2}(t)\leq\alpha_{mix}(t)\leq\eta_{\infty,2}([t/2])\,\eta^*_{\infty,2}(t/2)$.

(2) *Either $\eta_{2,2}(t)=1$ for all t or $\eta_{2,2}(t)\leq c\,e^{-\lambda t}$ for some $\lambda>0$. In the second case*

$$\eta^2_{2,2}(t)=(\eta^*)^2_{2,2}(t)\leq\rho_{mix}(t)\leq\ c\,\eta_{2,2}(t)\,.$$

(3) $\beta_{mix}(t)\leq\eta^2_{1,\infty}(t/2)\,.$

From the above Lemma one easily derive the following result.

Theorem 6.2. *Assume that $X=(X_t,t\geq 0)$ is an ergodic Markov process with unique invariant measure π.*

a) *If X is strongly or exponentially ergodic then it is β-mixing.*
b) *If X is uniformly ergodic, then it is strongly mixing.*
c) *Any kind of exponential ergodic decay rate in L^2 imply the ρ-mixing.*

TABLE 1. Performance of the estimated diffusion and scaling parameters $(\hat{\alpha}, \hat{\beta}, \hat{\rho})$ with a standard symmetric 1.6-stable driven OU process

n and T	**True parameters**	**Estimate Parameters**	RMSE
500 and T=1	$\alpha = 1.6, \quad \beta = 0, \quad \rho = 1$	$\hat{\alpha}_n = 1.70, \quad \hat{\beta}_n = 0.08, \quad \hat{\rho}_n = 0.80$	1.55
500 and T=10	$\alpha = 1.6, \quad \beta = 0, \quad \rho = 1$	$\hat{\alpha}_n = 1.69, \quad \hat{\beta}_n = 0.015, \quad \hat{\rho}_n = 3.38$	8.34
1000 and T=1	$\alpha = 1.6, \quad \beta = 0, \quad \rho = 1$	$\hat{\alpha}_n = 1.55, \quad \hat{\beta}_n = 0.11, \quad \hat{\rho}_n = 1.2$	1.10
1000 and T=10	$\alpha = 1.6, \quad \beta = 0, \quad \rho = 1$	$\hat{\alpha}_n = 1.56, \quad \hat{\beta}_n = 0.13, \quad \hat{\rho}_n = 4.87$	8.43
2000 and T=1	$\alpha = 1.6, \quad \beta = 0, \quad \rho = 1$	$\hat{\alpha}_n = 1.6346, \quad \hat{\beta}_n = 0.1146, \quad \hat{\rho}_n = 0.9335$	0.77
2000 and T=10	$\alpha = 1.6, \quad \beta = 0, \quad \rho = 1$	$\hat{\alpha}_n = 1.6029, \quad \hat{\beta}_n = 0.064, \quad \hat{\rho}_n = 4.33$	6.91

TABLE 2. Estimated drift coefficients $(\hat{\theta}, \hat{\mu})$ with the exchange rates data

n	**Estimate Parameters**
1000	$\hat{\theta}_n = 1.08, \quad \hat{\mu}_n = 0.93$
500	$\hat{\theta}_n = 1.06, \quad \hat{\mu}_n = 0.95$
200	$\hat{\theta}_n = 1.028493, \quad \hat{\mu}_n = 0.9674548$

TABLE 3. Estimated diffusion and scaling parameters $(\hat{\alpha}, \hat{\beta}, \hat{\rho})$ with the exchange rates data using an unknown strictly standard stable driven OU process

n	**Estimated diffusion and scaling parameters**
1000	$\hat{\alpha}_n = 1.98, \quad \hat{\beta}_n = 1, \quad \hat{\rho}_n = 13.21$
500	$\hat{\alpha}_n = 1.99, \quad \hat{\beta}_n = 1, \quad \hat{\rho}_n = 10.11$
200	$\hat{\alpha}_n = 1.99, \quad \hat{\beta}_n = 1, \quad \hat{\rho}_n = 6.17$

TABLE 4. Performance of the estimated diffusion and scaling parameters $(\hat{\alpha}, \hat{\beta}, \hat{\rho})$ with a standard strictly 1.5-stable Lotka Volterra process with positive jump activity $\beta > 0$ and $q = 1$

n and T	**True parameters**	**Estimate Parameters**	RMSE
500 and T=1	$\alpha = 1.5, \quad \beta = 0.1, \quad \rho = 0.5$	$\hat{\alpha}_n = 1.51, \quad \hat{\beta}_n = 0.38, \quad \hat{\rho}_n = 0.51$	0.9699
1000 and T=1	$\alpha = 1.5, \quad \beta = 0.1, \quad \rho = 0.5$	$\hat{\alpha}_n = 1.59, \quad \hat{\beta}_n = 0.12, \quad \hat{\rho}_n = 0.43$	0.8033
2000 and T=1	$\alpha = 1.5, \quad \beta = 0.1, \quad \rho = 0.5$	$\hat{\alpha}_n = 1.54, \quad \hat{\beta}_n = 0.11, \quad \hat{\rho}_n = 0.45$	0.7980

TABLE 5. Performance of the estimated diffusion and scaling parameters $(\hat{\alpha}, \hat{\beta}, \hat{\rho})$ with a standard strictly 1.6-stable CIR process with positive jump activity $\beta > 0$ and $q = 2$

n and T	**True parameters**	**Estimate Parameters**	RMSE
500 and T=1	$\alpha = 1.6, \quad \beta = 0.1, \quad \rho = 1$	$\hat{\alpha}_n = 1.66, \quad \hat{\beta}_n = 0.13, \quad \hat{\rho}_n = 0.88$	1.17
500 and T=10	$\alpha = 1.6, \quad \beta = 0.1, \quad \rho = 1$	$\hat{\alpha}_n = 1.51, \quad \hat{\beta}_n = 0.08, \quad \hat{\rho}_n = 5.95$	8.24
1000 and T=1	$\alpha = 1.6, \quad \beta = 0.1, \quad \rho = 1$	$\hat{\alpha}_n = 1.63, \quad \hat{\beta}_n = 0.217, \quad \hat{\rho}_n = 0.98$	3.72
1000 and T=10	$\alpha = 1.6, \quad \beta = 0.1, \quad \rho = 1$	$\hat{\alpha}_n = 1.53, \quad \hat{\beta}_n = 0.135, \quad \hat{\rho}_n = 6.32$	15.36
2000 and T=1	$\alpha = 1.6, \quad \beta = 0.1, \quad \rho = 1$	$\hat{\alpha}_n = 1.61, \quad \hat{\beta}_n = 0.07, \quad \hat{\rho}_n = 1.03$	2.07
2000 and T=10	$\alpha = 1.6, \quad \beta = 0.1, \quad \rho = 1$	$\hat{\alpha}_n = 1.64, \quad \hat{\beta}_n = 0.04, \quad \hat{\rho}_n = 4$	11.13

TABLE 6. Estimated diffusion and scaling parameters $(\hat{\alpha}, \hat{\beta}, \hat{\rho})$ with the exchange rates data using an unknown strictly standard SCIR process with given parameter q.

n	q	**Estimate Parameters**
1000	q=2	$\hat{\alpha}_n = 1.80, \quad \hat{\beta}_n = -0.33, \quad \hat{\rho}_n = 67.89$
1000	q=1.99	$\hat{\alpha}_n = 1.78, \quad \hat{\beta}_n = -0.31, \quad \hat{\rho}_n = 70.57$
1000	q=1.95	$\hat{\alpha}_n = 1.78, \quad \hat{\beta}_n = -0.33, \quad \hat{\rho}_n = 71.59$
1000	q=1.9	$\hat{\alpha}_n = 1.77, \quad \hat{\beta}_n = -0.35, \quad \hat{\rho}_n = 75.83$
1000	q=1.8	$\hat{\alpha}_n = 1.72, \quad \hat{\beta}_n = -0.37, \quad \hat{\rho}_n = 8736$
1000	q=1.6	$\hat{\alpha}_n = 1.59, \quad \hat{\beta}_n = -0.28, \quad \hat{\rho}_n = 130.70$
1000	q=1.5	$\hat{\alpha}_n = 1.53, \quad \hat{\beta}_n = -0.20, \quad \hat{\rho}_n = 163.02$
500	q=2	$\hat{\alpha}_n = 1.79, \quad \hat{\beta}_n = -0.12, \quad \hat{\rho}_n = 39.61$
500	q=1.99	$\hat{\alpha}_n = 1.78, \quad \hat{\beta}_n = -0.12, \quad \hat{\rho}_n = 39.99$
500	q=1.95	$\hat{\alpha}_n = 1.77, \quad \hat{\beta}_n = -0.093, \quad \hat{\rho}_n = 41.56$
500	q=1.9	$\hat{\alpha}_n = 1.75, \quad \hat{\beta}_n = -0.06, \quad \hat{\rho}_n = 43.74$
500	q=1.8	$\hat{\alpha}_n = 1.72, \quad \hat{\beta}_n = -0.009, \quad \hat{\rho}_n = 48.98$
500	q=1.6	$\hat{\alpha}_n = 1.59, \quad \hat{\beta}_n = 0.08, \quad \hat{\rho}_n = 70.13$
500	q=1.5	$\hat{\alpha}_n = 1.53, \quad \hat{\beta}_n = 0.073, \quad \hat{\rho}_n = 86.03$
200	q=2	$\hat{\alpha}_n = 1.69, \quad \hat{\beta}_n = -0.31, \quad \hat{\rho}_n = 25.56$
200	q=1.99	$\hat{\alpha}_n = 1.69, \quad \hat{\beta}_n = -0.30, \quad \hat{\rho}_n = 25.83$
200	q=1.95	$\hat{\alpha}_n = 1.67, \quad \hat{\beta}_n = -0.25, \quad \hat{\rho}_n = 26.98$
200	q=1.9	$\hat{\alpha}_n = 1.65, \quad \hat{\beta}_n = -0.19, \quad \hat{\rho}_n = 28.59$
200	q=1.8	$\hat{\alpha}_n = 1.60, \quad \hat{\beta}_n = -0.008, \quad \hat{\rho}_n = 32.42$
200	q=1.6	$\hat{\alpha}_n = 1.51, \quad \hat{\beta}_n = 0.04, \quad \hat{\rho}_n = 43.64$
200	q=1.5	$\hat{\alpha}_n = 1.44, \quad \hat{\beta}_n = 0.08, \quad \hat{\rho}_n = 5386$

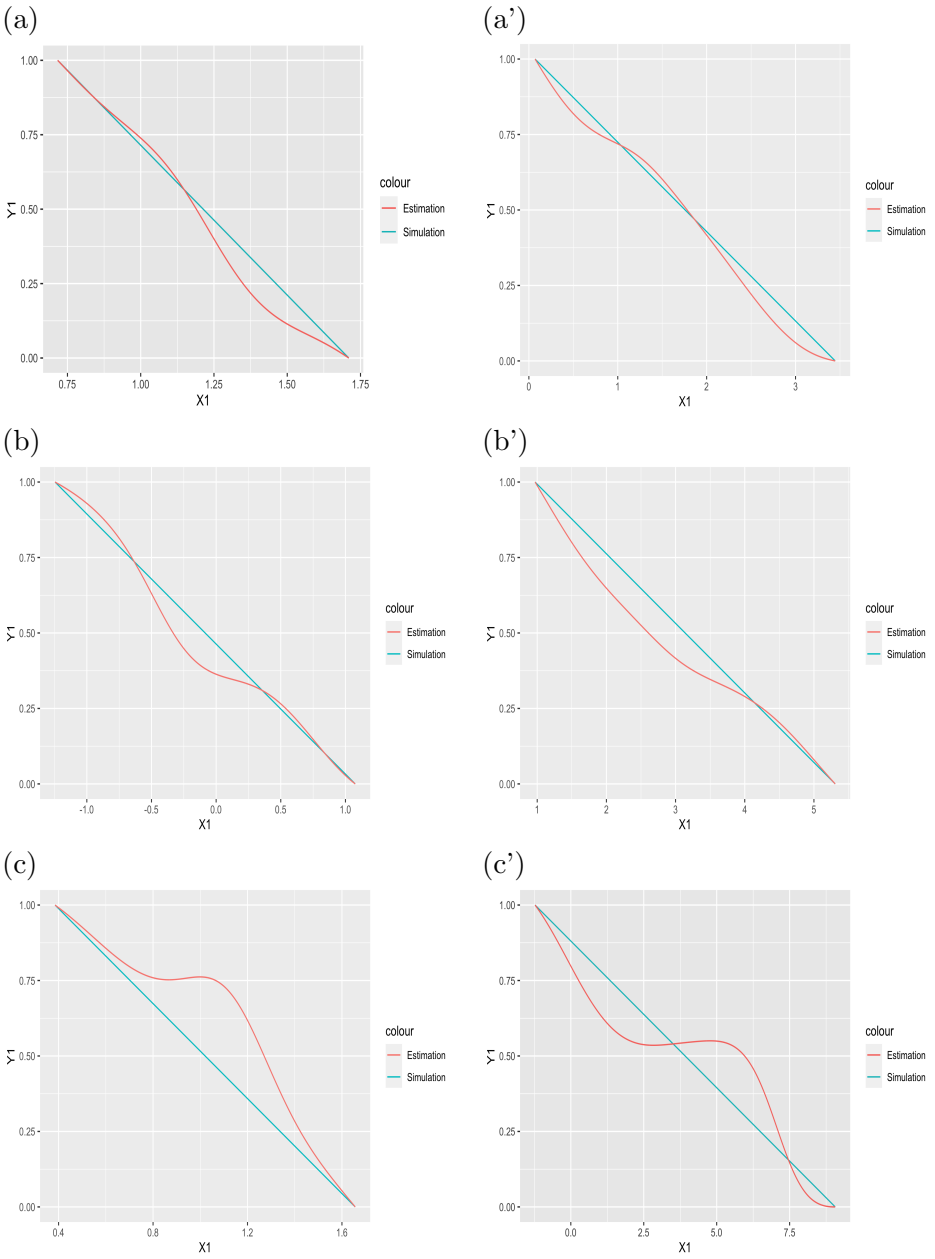


FIGURE 1. Performance of the Nadaraya-Watson kernel estimator with respect to the true drift in the case of a standard symmetric 1.6-stable driven OU process with sample size $n = 500$ and $t = 1$ in (a), sample size $n = 500$ and $t = 10$ in (a'), sample size $n = 1000$ and $t = 1$ in (b), sample size $n = 1000$ and $t = 10$ in (b'), sample size $n = 2000$ and $t = 1$ in (c) and sample size $n = 2000$ and $t = 10$ in (c').

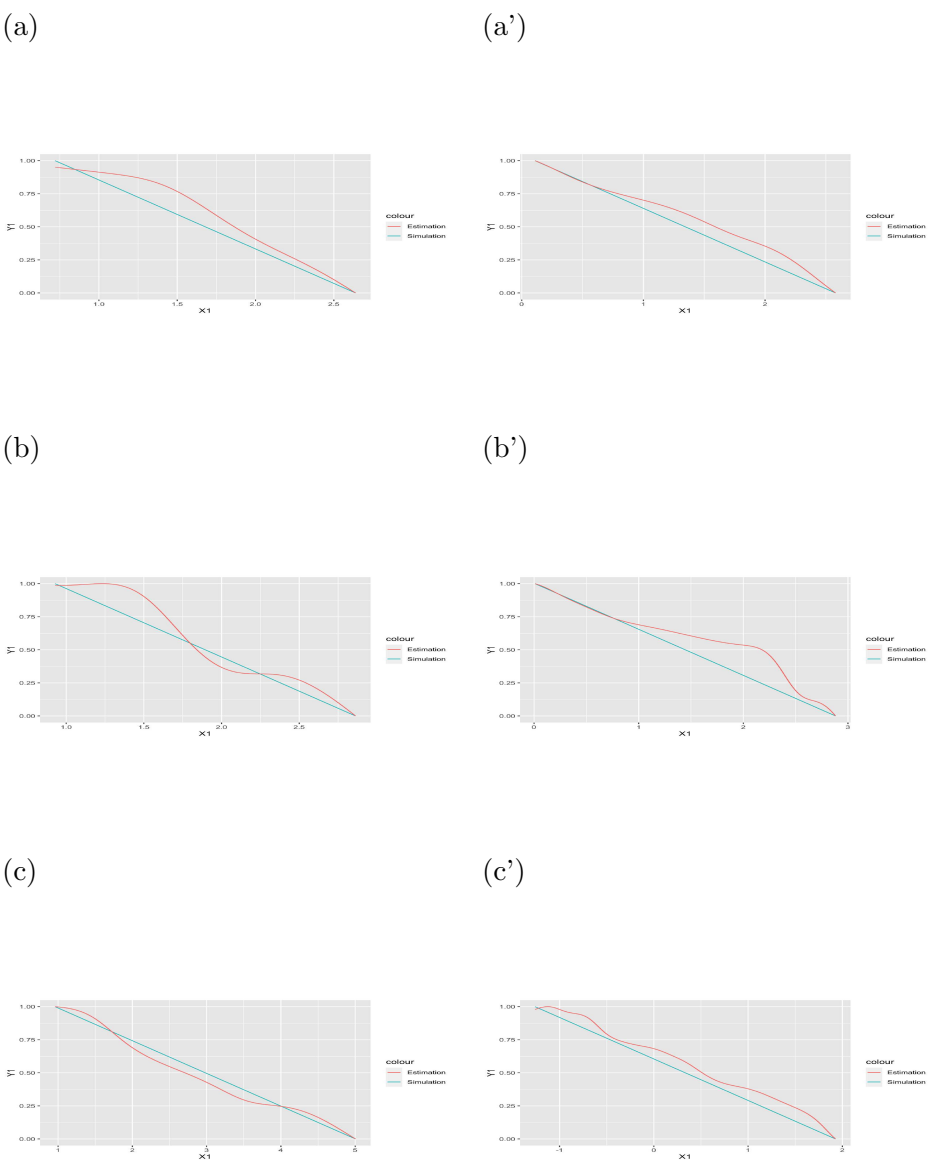
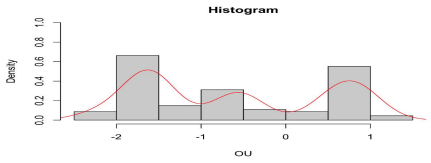
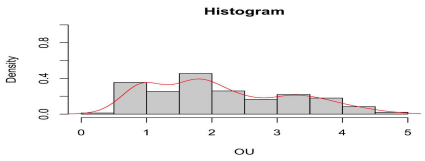


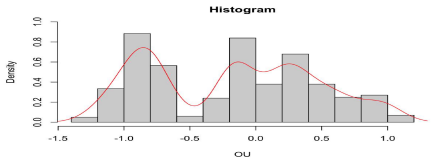
FIGURE 2. Performance of the Nadaraya–Watson kernel estimator with respect to the true drift in the case of a standard strictly 1.6-stable CIR process with positive jump activity and sample size $n = 500$ and $t = 1$ in (a), sample size $n = 500$ and $t = 10$ in (a'), sample size $n = 1000$ and $t = 1$ in (b), sample size $n = 1000$ and $t = 10$ in (b'), sample size $n = 2000$ and $t = 1$ in (c) and sample size $n = 2000$ and $t = 10$ in (c').

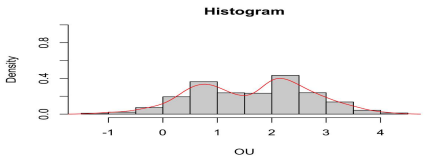
(a) (a')



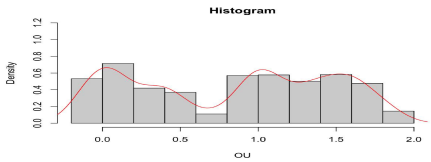


(b) (b')





(c) (c')



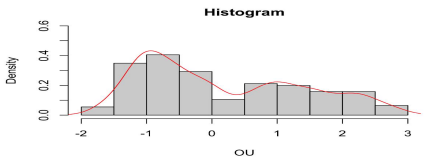
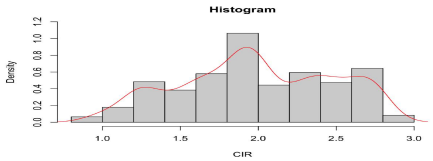
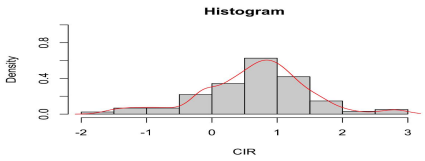


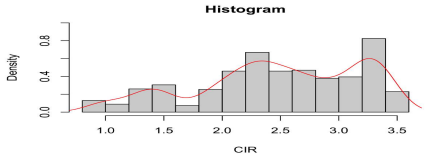
FIGURE 3. Kernel density estimation of a standard symmetric 1.6-stable driven OU process with sample size $n = 500$ and $t = 1$ in (a), sample size $n = 500$ and $t = 10$ in (a'), sample size $n = 1000$ and $t = 1$ in (b), sample size $n = 1000$ and $t = 10$ in (b'), sample size $n = 2000$ and $t = 1$ in (c) and sample size $n = 2000$ and $t = 10$ in (c').

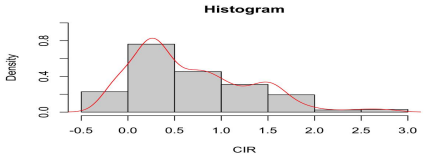
(a) (a')



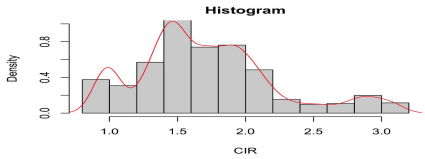


(b) (b')





(c) (c')



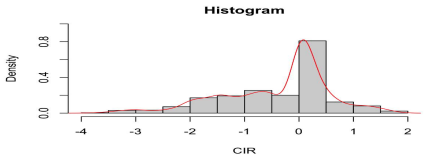
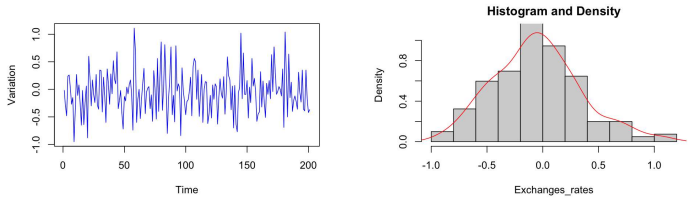
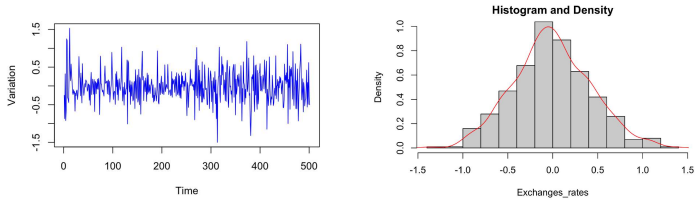


FIGURE 4. Kernel density estimation of a standard strictly 1.6-stable CIR process with positive jump activity and sample size $n = 500$ and $t = 1$ in (a), sample size $n = 500$ and $t = 10$ in (a'), sample size $n = 1000$ and $t = 1$ in (b), sample size $n = 1000$ and $t = 10$ in (b'), sample size $n = 2000$ and $t = 1$ in (c) and sample size $n = 2000$ and $t = 10$ in (c').

(a) (a')



(b) (b')



(c) (c')

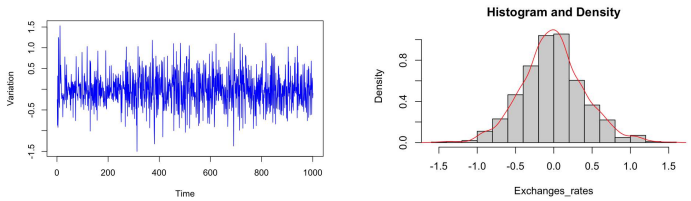


FIGURE 5. The exchange rates of the Canadian dollar against the US dollar global variation and density estimation with sample size $n = 1000$ (a); local sample size $n = 500$ (b) and sample size $n = 200$ (c).

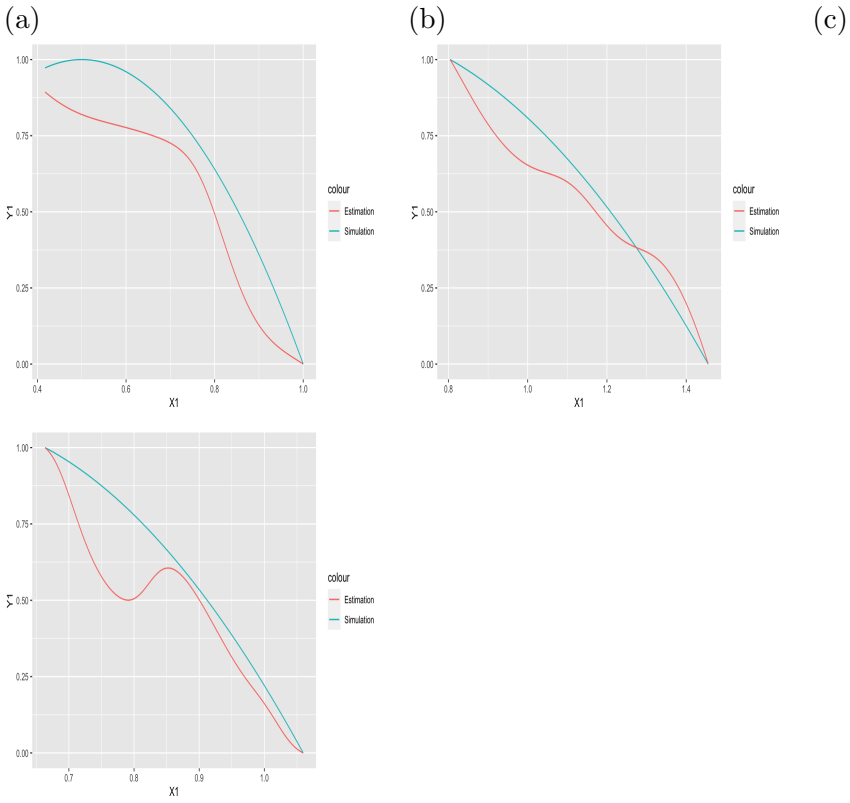
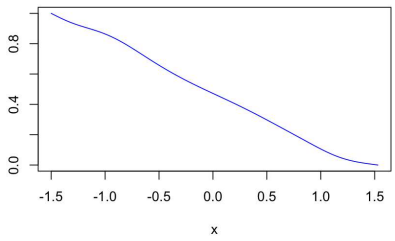
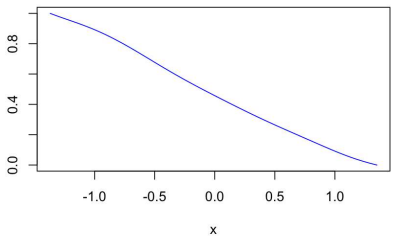


FIGURE 6. Performance of the Nadaraya-Watson kernel estimator with respect to the true drift in the case of a standard and positive strictly 1.5-stable driven Lotka-Volterra process with sample size $n = 500$ and $t = 1$ in (a), sample size $n = 1000$ and $t = 1$ in (b), sample size $n = 2000$ and $t = 1$ in (c).

(a)



(b)



(c)

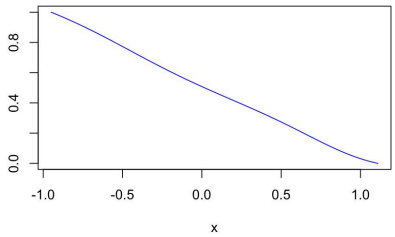
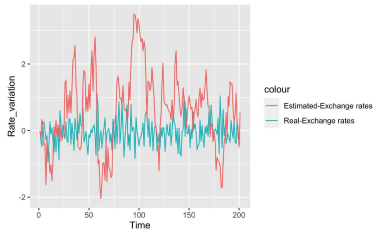
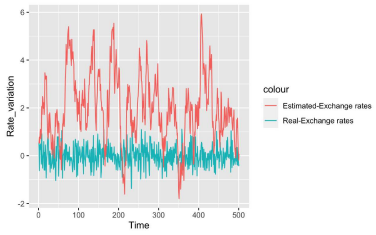


FIGURE 7. Nadaraya-Watson kernel estimator of the drift function with the exchange rates of the Canadian dollar against the US dollar variation with the global sample size $n = 1000$ (a); local sample size $n = 500$ (b) and sample size $n = 200$ (c).

(a)



(b)



(c)

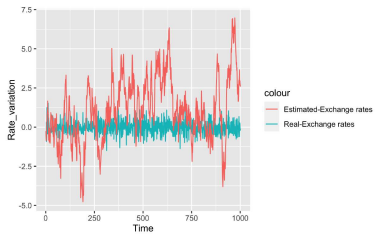


FIGURE 8. Comparison between exchange rates prediction with Stable driven OU process with sample size $n = 200$ (a); sample size $n = 500$ (b) and sample size $n = 1000$ (c).