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 7031, Poitiers
solym.manou.abi@math.univ-poitiers.fr

Abstract. In this paper, we consider parameter estimation for a stochastic process observed at some discrete times that is a solution of a given class of stochastic differential equations driven by α-stable processes, $\alpha \in (1, 2)$. Firstly, we consider the diffusion coefficients parameters estimation problem such as the scaling parameter and the driving stable process parameters. Secondly, we address the question of the joint estimation with the drift coefficients for Stable driven Cox–Ingersoll–Ross and Ornstein-Uhlenbeck processes. Our methodology which is based on the use of Nadaraya-Watson estimator is new according to the literature of such estimation problem. Indeed it is based on a combination of the characteristic sample function and the (linear or weighted) regression methods. We discuss the validity and efficiency of the numeric implementation of the estimators on synthetic data and real financial data as indexes and exchange rates. As a forthcoming work, we intend to provide a package (in the R programming language) for this type of estimation problem.

Key words and phrases: stochastic differential equation, stable process, parameter inference, characteristic function, regression method, Cox–Ingersoll–Ross and Ornstein-Uhlenbeck processes, Nadaraya-Watson estimation.

1. Introduction

Due to the increasing of computational powerfull statistical methods, there has been a great interest in parameter estimation for Stochastic Differential Equation (SDE). Such models are mathematical tools for modeling the time evolution of many natural phenomena in many disciplines such as epidemiology, biology and finance. As examples, the Cox–Ingersoll–Ross (CIR) and the Ornstein-Uhlenbeck (OU) stochastic models [35, 10], has been used widely in finance. Some equivalent SDE models from continuous-time branching processes are also shown to be useful models to study dynamic

population [2, 43]. More precisely in epidemic dynamics, when an infectious population size is small and the population size is large, on can approximate the continuous time dynamic by a nonlinear SDE (stochastic logistic model that includes variability due to births and deaths [2]). For classical estimation methods, issues such as how to construct procedures which are both computationally efficient and show optimal statistical performance are now well understood. 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 $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$:

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

where x_0 is a starting point, $T > 0$ a given time horizon and the given function $g : \mathbb{R} \to \mathbb{R}$ is assumed to be known. We assume that, 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 $\alpha \in (0, 2)$ (the index of stability) and $\beta \in [1, 1]$ (the skewness parameter) will be defined in the sequel.

The existence and uniqueness of the solution to the SDE (2.4) under Lipschitz conditions are standard results in stochastic calculus, [4]. For non Lipchtiz coefficients, some results have been studied. Let us recall some few existence results. 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 [15] 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 [39], [42]. For more recent existence conditions, we refer to [37] and references therein.

The main focus of this paper is the joint parameter estimation problem for a given solution X of the SDE (2.4) not only for the drift coefficient f but also for the scaling diffusion parameter ρ and the stable noise parameters appearing in the model equation. We assume that, the process X is observed at discrete time points $\{t_i = i\Delta_n, \quad i = 0, 1, 2, ..., n\}$ with Δ_n a time frequency of the observation and n is the sample size. The parameter estimation theory for SDE driven by Brownian motions are well know in the literature. Some traditional methods are the maximum likelihood estimator (MLE) or or the least squares estimator (LSE) techniques, [33, 13, 27, 11], based on the Girsanov density. The consitency and the asymptotic distribution are well studied, see for instance [13, 27], [21], [45], [1]. For a more recent comprehensive discussion, we refer to [26] and the references therein.

Substantial progress has been made in parameter estimation for SDE driven by Lévy processes with finite moments. The work in [38]] dealt with the consistency and asymptotic normality when the driving process is a zero-mean adapted Lévy process with finite moments. The asymptotic normality of the LSE and MLE for the pure jump case is studied in [49, 50]. Note also that, some work has been carried out, when the driving process Z is a stable Lévy process which is really particular, due to it's infinite variance property. However, the MLE is no longer valid in this setting because, the explicit density function is not always available and the Girsanov measure transformation is not well defined for α-stable processes, $\alpha \in (1, 2)$. Let us pointed out some well-known results. In [16, 17], authors use the trajectory fitting method combined with the weighted least squares technique of the drift coefficient for an α-stable driven OU process, $\alpha \in (1, 2)$, when the process is observed at discrete time instants, both for the ergodic and non ergodic case. They also discuss the consistency and asymptotic distribution of the estimator which has a higher order of convergence than in the classical Gaussian case. In [30, 53], authors study the drift parameter estimation of a stable driven Cox–Ingersoll–Ross (CIR) model (which is a special subcritical continuous state branching process with immigration). They 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. In [12], a Lasso and Slope drift estimators for Lévy-driven Ornstein–Uhlenbeck processes were considered. A maximum likelihood type estimation of the drift and volatility constant coefficient parameters in stable driven CIR is studied in [56]. In the recent paper [6], author address the joint parameter estimation of the drift parameters, scaling parameter for the diffusion coefficient and the jump activity parameter such as the index of stability α and skewness parameter $\beta \in [-1, 1]$ from high-frequency observations of the stable CIR process on a fixed time period. Their methodology is based on the approximation of the conditional distribution.

Now let us turn to non parametric estimation, an important field of statistics which consists of estimating an unkown function from a sample of data without first giving a formula. Many authors have investigated the non parametric estimation of the drift function f in the setting of diffusion's driven by Brownian motions under various conditions. The major statistical properties such as the consistency and rate of convergence of non parametric methods for the Nadaraya-Watson (N-W) estimators [40, 52], under independence with identically distribution and also weak dependence conditions such as mixing conditions. For a complete review of non parametric methods for diffusion processes; see the survey paper by [18]. In the stable and non parametric setting, some results are also studied in [54, 34] and [46, 58, 32]. But the case of the dispersion function g is much harder. In this paper, we assume that the function g is known and given. We recall from [34] that,

if the solution of SDE (2.4) is stationnary (and the stationnary distribution has continuous density) and strongly mixing, the authors establish a non parametric estimation using the N-W estimator. When the drift function is linear, the estimation both for discrete and continuous observations was studied in the above mentioned papers [17, 16, 31, 6, 56] including the stable driven CIR and OU models type. The statistical properties such as the consistency and rate of convergence of the N-W estimators under dependence conditions such as mixing are developed in [34]. All this method developed and applied within the framework of the stable driven CIR and OU models are conditional dependent on the observation of the noise $Z = (Z_t)_{t\in[0,T]}$ parameter and their performance is strongly dependent on the time frequency Δ_n and the time horizon T. Although the parameter estimation for Stable driven SDE have been developed in recent years, few works (for instance [6] in the CIR case) deal with the joint parameter estimation both for the drift coefficient, scaling parameter for the diffusion coefficient and the jump activity diffusion parameters such as the index of stability α, the sekewness β of the Stable CIR and OU processes.

In this article, our aim is the joint parameter estimation problem of the parameters appearing in the model equation (2.4) assuming that, only the function g is known. We assume that the solution X is stationary and strongly mixing but we highlight the link with some sets of ergodic conditions. Our joint parameter estimation procedure is a new methodology using the N-W estimator of the drift coefficient. Indeed, from a given sample characteristic function, we derive consistent estimators of the diffusion part parameters (scaling parameter ρ and the stable noise parameters α and β). In the case where the drift coefficient is linear (stable driven CIR and OU processes) we establish the drift coefficient estimators using linear (or weighted) regression. This paper is organized as follows. In Section 2, we recall basic fact on stable processes as well relevant stochastic models for the numerical application. We recall the connection between ergodicity and mixing conditions for Markov processes and the classical N-W estimation framework including relevant assumptions. In section 3, we state our main theorems, which consists on the explicit estimators formulas. We also give estimators of the drift coefficients when it is linear, namely for the Stable driven CIR and OU processes. Section 4, investigate the numerical performance of the estimators on synthetic data. We applied the result with real financial data such as exchanges rates of the Canadian dollar against the US dollar in a given fixed period. The last section is dedicated to the proofs of our main results.

2. Preliminaries

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

(1) $(\Omega, \mathcal{F}_t, P)$ 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 objective is firstly to present the stable laws and processes on the one hand as well as a method for estimating the stable alpha laws that we will use later, in particular this one by the characteristic functions. Secondly, we present some results concerning the links between the ergodicity and mixing conditions of a Markov process.

2.1. **Stable distributions and processes.** We will start this section by a quick reminder about the stable laws and processes. Stable laws were introduced by Paul Lévy in 1925 and arise as the limit of normalized sums . The definitions and properties given below are taken from [47] and [41]. Note that there are multiple characterizations of stable distributions and each of them have some advantages depending on the purposes.

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 :*

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

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 above definition is the parametrization 1 in [41]. The following is the parametrization 0 and is usefull 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 more detailed discussion on stable distributions, see [41]. When $\sigma = 1$ we say X is a standard α-stable random variable. For $\alpha \neq 1$ and $\omega = 0$ we say that X is a strictly α-stable random variable. Furthermore if $\beta = 0$ it is said to be symmetric. Now let us introduce α-stable processes that can be seen as a particular class of Lévy processes and we refer to the well-known standard literature like [23].

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 $Z_t - Z_s \sim Z_{t-s} \sim S_\alpha((t-s)^{1/\alpha}\sigma, \beta, \omega)$, $t > s \geq 0$;*

(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.*

We also refer to [34] for the case of standard α-stable process. Note that strictly α-stable process has the following self-similarity (scaling) property

$$(Z_{kt})_{t\in[0,T]} \overset{d}{=} (k^{1/\alpha}\, Z_t)_{t\in[0,T]},$$

where $k > 0$ and the equality $\overset{d}{=}$ is understood in the sense of finite dimensional distributions. Every α-stable process has a right continuous with left limits modification and we will henceforth assume in the sequel that we are using this modification. When $0 < \alpha < 1$, Z_t is a pure jump process with locally finite variation while for $\alpha \in (1, 2)$ a strictly α-stable process is a purely discontinuous martingale. For more path properties of α-stable process, we refer to Sato [23]. Note that in the case $\alpha \in (1, 2)$, the characteristic function of a strictly α-stable process Z_t is reduced into the form

$$\varphi_{Z_t}(u) = \exp\left(t \int_{-\infty}^{+\infty} (e^{iuy} - 1 - iuy)\nu(dy)\right), \quad t \in [0, T], \tag{1}$$

where ν stands for the Lévy measure defined on $\mathbb{R} \setminus \{0\}$ by

$$\nu(dx) := \frac{dx}{|x|^{\alpha+1}} \left(c_+ 1_{\{x>0\}} + c_- 1_{\{x<0\}}\right).$$

The parameters c_+, c_- above are non-negative with furthermore $c_+ + c_- > 0$. There is a connection between the coefficients c_+, c_- and parameters α and β; for instance $\beta = \frac{c_+ - c_+}{c_+ + c_+}$. Note that, due to the above self-similarity property and the property of stationnary independent increments, random walk approximation is usefful if one wants to simulate such process. The following algorithm ([9],[19]) is useful for simulating α-stable processes. Of course, others methods are also available like the series approximation of Lévy processes [19].

Proposition 2.1 (Discretized trajectory for a strictly and standard α-stable process)**.** *Let* $Z = (Z_t)_{t\in[0,T]}$ *be a strictly and standard-*α *stable process.*

Step 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.

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

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

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

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

To extend the algorithm to the non-standard case, we can choose between the two previous parameterisations. Note that, the parametrization 0 is important for numerical implementation because the characteristic function, density and cumulative distribution function are continuous with respect to the above four parameters α, β, γ and ω. There exists some package in R to perform the simulation (for instance the package *stabledist*) in [55]. Figure 2 show typical sample paths for some strictly standard α-processes. When α is closed to 2 and β to 0 the sample paths resemble to that of Brownian motions. For $\alpha = 1$ and $\beta = 0$, the sample paths corresponds to the that of Cauchy process.

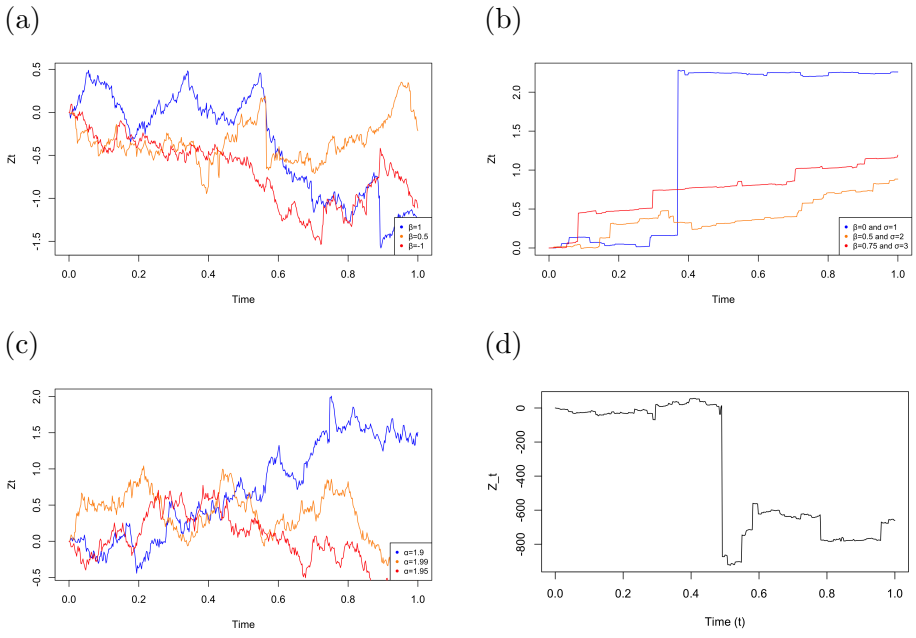

FIGURE 1. Paths of a standard strictly 1.7-stable process with different skewness (a), 0.7-stable process with different scale and skewness (b), the symmetric and standard α-stable process with different index of stability (c) and the case of standard 1-stable process (d).

2.2. **Stable driven Stochastic models.** We now present some α-stable-driven stochastic differential equations. We are particularly interested to the Ornstein Uhlenbeck (OU) and Cox–Ingersoll–Ross (CIR) models driven by an α-stable process. Such models are popular in stochastic modelling for description of interest rates in finance and population dynamics. In the remainder of this section we assume that Z is a standard strictly α-stable process with $\alpha \in (1, 2)$.

Definition 2.4. *(Stable driven OU process) A generalized Ornstein-Uhlenbeck (OU) process driven by an α-stable process $(Z_t)_{t\geq 0}$ is defined to be the solution to the following linear stochastic differential equation*

$$dX_t = -\theta X_t dt + \rho dZ_t, \quad X_0 = x \in \mathbb{R}. \tag{2}$$

and has the integral representation

$$X_t(\omega) = e^{-\theta t}\Big(x + \rho \int_0^t e^{\theta s} dZ_s(\omega)\Big), \quad for \quad t \in [0, T].$$

An additional drift term is sometimes added:

$$dX_t = \theta(\mu - X_t)dt + \rho dZ_t, \quad X_0 = x \in \mathbb{R}. \tag{3}$$

where μ is a constant. The classical Cox–Ingersoll–Ross process denoted by $(Y_t, t \geq 0)$ is given by the following SDE

$$dX_t = \theta(\mu - X_t)dt + \rho\sqrt{|X_t|}dW_t, \quad X_0 \geq 0,$$

where θ, μ and μ are positive constants and $W = (W_t)$ is the one-dimensional standard Brownian motion. It is natural to replace the driving Wiener process W by an α-stable process Z; see [7, 20, 44]. Some extensions of the classical CIR model to α-stable processes are given in the literature. We have the following stable driven CIR model.

Definition 2.5. *(Stable CIR process,* [56]*) An α-stable CIR process is defined by:*

$$dX_t = \theta(\mu - X_t)dt + \rho|X_t|^{1/q}dZ_t, \quad X_0 \geq 0, \tag{4}$$

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

For example, in [56], the authors consider the case of a pure-jump α-stable Lévy process with positive jumps Z with index $\alpha \in (1, 2)$. It was proven in [31] that, there is a pathwise unique positive strong solution whenever $q^{-1} + \alpha^{-1} > 1$. In particular for $q = 2$, this model is consider in [53] for $\alpha \in (0, 2)$. Another extension is introduced in the literature and named SCIR process (see [6]) given as follows for $q = \alpha$ when $\alpha \in (1, 2)$.

Definition 2.6. *(SCIR process) A SCIR process is defined by:*

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

where $\theta > 0$, $\rho > 0$, and $\mu \in \mathbb{R}$ are constants.

Similarly, it is shown in [20] that for $\mu > 0$, $\theta > 0$, $\rho > 0$ and $x_0 > 0$ that the solution is positive that is

$$\mathbb{P}\big(X_t > 0\Big) = 1.$$

This stable driven CIR model is a particular form of the so-called continuous-state branching processes with immigration (CBI-processes, [28], [30] and [43]), which arise as scaling limits of Galton–Watson branching processes with immigration (GWI-processes); see, e.g., [22]. A Markov process with state space $\mathbb{R}_+$ is called a continuous-state branching process (CB-process) with branching mechanism φ if it has transition semigroup $(Q_t)_{t\geq 0}$ given by

$$\int_0^{+\infty} e^{-\lambda y} Q_t(x, dy) = e^{-x v_t(\lambda)},$$

where $t \to v_t(\lambda)$ is the unique strictly positive solution to the differential equation

$$\frac{\partial}{\partial t} v_t(\lambda) = -\varphi(v_t(\lambda)), \qquad v_0(\lambda) = \lambda,$$

and with the Lévy–Khintchine representation:

$$\varphi(z) = \theta z + \frac{1}{2}\rho^2 z^2 + \int_0^{+\infty} \big(e^{-zu} - 1 + zu\big) m(du)$$

with $\rho \geq 0$; θ constant and m a finite measure on $(0, \infty)$. The CB-processes involve rich mathematical structures and have been applied to the research in several important areas. The CB-process is called critical, subcritical or supercritical as $\theta = 0$, $\theta > 0$ or $\theta < 0$ respectively. A Markov process with state space $\mathbb{R}_+$ is called a CBI-process with branching mechanism φ and immigration rate $\mu \geq 0$ if it has transition semigroup $(P_t)_{t\geq 0}$ given by

$$\int_0^{+\infty} e^{-\lambda y} P_t(x, dy) = e^{-x v_t(y) - \mu \int_0^t v_s(\lambda) ds}.$$

By a Theorem of [29], for any $\alpha \in (1, 2]$, we can identify the SCIR-model as a subcritical CBI-process with immigration rate $\mu \geq 0$ and branching mechanism

$$\varphi(z) = \theta z + \frac{\rho^\alpha}{\alpha} z^\alpha, \quad z \geq 0.$$

2.3. **The connection between ergodicity and mixing conditions.** Let $(X_t, t \geq 0)$ be a continuous time ergodic Markov process with unique invariant measure π. Denote by $P_t(x, dz)$ be the transition probability of the process X_t starting from $X_0 = x$. 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 measurable and positive functions or in $L^p(\pi)$. 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). Let us also recall the following adjoint operator P^*, defined by

$$\int f\, P_t g\, d\pi = \int g\, P_t^* f\, d\pi$$

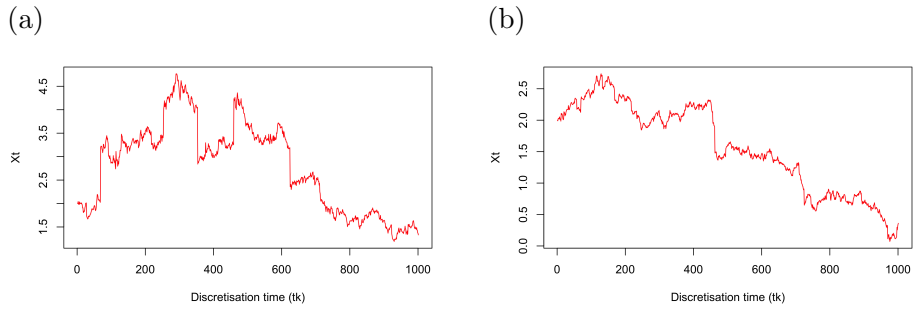

FIGURE 2. Paths of a standard strictly 1.7-stable OU process with $\beta = 0.5$, $\theta = 1.2$, $\rho = 1$ and sample size $n = 1000$ (a), standard strictly 1.7-stable CIR process with $\beta = 0.5$, $\theta = 1.2$, $\rho = 1$, $\mu = 0$, $q = 1.7$ and sample size $n = 1000$ (b).

for f and g square integrable w.r.t.π, which is again a contraction. The following definition introduces some way to control the ergodic decay to equilibrium.

Definition 2.7 (Ergodic rates of convergence [8]). *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 $\lim_{t\to+\infty} \eta_{2,\infty}(t) = 0$ and the we call $\eta_{2,\infty}$ the uniform decay rate.

The following definitions can be found in [59].

Definition 2.8 (Exponentially or strongly ergodic process). *Assume that $X = (X_t, t \geq 0)$ is a continuous time 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 say that the process X is said to be strongly or exponentially ergodic if $\lim_{t\to+\infty}\eta_{1,\infty}(t)=0$ and the we call $\eta_{1,\infty}$ the strong or exponential ergodic decay rate. Now, let us define some usual mixing coefficients in order to compare mixing properties and ergodic decay properties.

Definition 2.9 (Mixing conditions, see [36])**.** . *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 :*

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

   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.*

Note that if X is a stationary process (i.e. such that, for all $n\geq 0$, the law of X_{t+s} is the same as the one of X_s), the supremum on s is irrelevant. The following Lemma is contained in our past paper [8] and allow to make a connection between ergodicity and mixing conditions.

Lemma 2.1. *For all* $t\geq 0$*, we have*

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

(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 2.2. *Assume that* $X = (X_t, t \geq 0)$ *is a continuous time 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.*

Now let us discuss about the exponential ergodicity of the Stable OU and SCIR-models. Based on the explicit coupling method, the exponential ergodicity of a Lévy driven OU process is established in [51]. The result entails that if $\theta > 0$ then the α-Stable OU process (2) has a unique invariant measure π and it is strongly ergodic (mixing). More generally, let us consider the case where the drift f and diffusion coefficient g are supposed to satisfy sufficient conditions for the solution of (2.4) to exist and to be unique [5, 15]. According to [25], if $f(.)$ is locally Lipschitz and $\limsup_{|x|\to+\infty} \frac{f(x)}{x} < 0$ then, for any $\alpha \in (1, 2)$, the solution of (2.4) is exponentially ergodic and its invariant distribution exists and is unique. This therefore applies to the case of the stable OU model (3) under the condition $\theta > 0$. This result is also generalized in [14] to the case when $g = 1$, $\rho = 1$ and f is regularly varying at infinity with regularly varying coefficient greater than one. In the case of the SCIR-model, some results exists in [30]. Note that for $\theta > 0$ and $\mu \geq 0$ this model can be seen as a subcritical CBI process with immigration rate μ. Thus, using the result of [30], we conclude that the SCIR-model is exponentially ergodic and hence strongly mixing that is, as $t \to +\infty$, $\alpha_{mix}(t)$ decays to zero exponentially. Moreover, the Laplace transform of the unique stationary distribution π is given by

$$L_\pi(\lambda) = \int_0^{+\infty} e^{-\lambda x} \pi(dx) = e^{-\int_0^\lambda \frac{\alpha\mu}{\alpha\theta + \rho^\alpha z^{\alpha-1}} dz}, \quad z \geq 0.$$

For a more general SDE in the form (2.4), it was shown in [57], that the exponential ergodicity holds under some dissipative and non-degenerate assumptions on the drift f and diffusion function g. Their proofs are mainly based on heat kernel estimates, the strong Feller property and the irreducibility of the associated semi group.

2.4. **Non parametric estimation of the drift function.** In this section, we recall the non-parametric estimation framework of the drift coefficient of the following Stable driven SDE equation:

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

where f is an unknown function, $\rho > 0$ an unkown diffusion coefficient and the driving standard strictly stable process Z is unknown too. We assume in this paper that only the function g is known.
Let X be a solution of the above SDE. Note that Lipschitz and Hölder (or bounded) conditions are typical conditions to show the existence of a solution. Let us recall some existence result. Here are just some few existence results.

- If g is Hölder continuous with exponent $\frac{1}{\alpha}$ and Z is a symmetric stable process, it is proved in [5] that there exists a strong solution whenever $\alpha \in (1,2)$.
- If f and g have at most linear growth it is proved in [15] that there exists a weak solution for $\alpha \in (1,2)$ for which pathwise uniqueness holds whenever the function g is particularly Hölder continuous with exponent lying in $[1-\frac{1}{\alpha}, \frac{1}{\alpha}]$.

For more recent existence conditions, we refer to [37] 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 [34]. 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 set of parameters 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 :

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

so that

$$\hat{f}_n(x) = \sum_{i=1}^{n} \frac{K_h(\frac{X_{t_i}-x}{h})(X_{t_i} - X_{t_{i-1}})}{n^{-1}\sum_{i=1}^{n} K_h(\frac{X_{t_i}-x}{h})} \quad \text{for all } x \in \mathbb{R}. \tag{5}$$

We explore various Kernel functions in this paper but we choose only the one which perform the simulation. For instance we'll consider the Gaussian kernel function:

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

For pratical setting, we choose the optimal bandwidth h by considering the method proposed by [48]. The R function **h.amize** allow to do it. Note also that, in many cases people usually choose $h = h_n = n^{-1/5}$.

3. Main results

Based on the N-W estimator, the main objective in this section is firstly the joint estimation of the scaling parameter ρ and the standard strictly stable process parameters α and β of the diffusion component in the SDE in (2.4). Secondly, we particularly consider the drift common coefficients estimation θ and μ of the Stable driven CIR and OU processes in (3) or (4).

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 (2.4). 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 $(\mathbf{A_1})$, $(\mathbf{A_2})$ 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 (2.4).*
We define the following sample characteristic function

$$\hat{\phi}_n(u_k) = \frac{1}{n}\sum_{i=0}^{n-1} \exp\left(ju_k\Big(\frac{\Delta X_{t_i} - \hat{f}_n(X_{t_i})\Delta_n}{g(X_{t_i})}\Big)\right) \tag{6}$$

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

$$\begin{cases} \hat{\alpha}_m = \dfrac{\sum_{k=1}^{m} W_k V_k - \frac{1}{m}\sum_{k=1}^{m} V_k \sum_{k=1}^{m} W_k}{\sum_{k=1}^{m} W_k^2 - \frac{1}{m}\big(\sum_{k=1}^{m} W_k\big)^2}, \\ \hat{\lambda}_m = \frac{1}{m}\sum_{k=1}^{m} V_k - \frac{\hat{\alpha}_m}{m}\sum_{k=1}^{m} U_k \\ \hat{\rho}_m = \Delta_n^{-\frac{1}{\hat{\alpha}_m}} \exp \frac{\hat{\lambda}_m}{\hat{\alpha}_m} \\ \hat{\beta}_m = \dfrac{\sum_{k=1}^{m} u_k S_k}{\sum_{k=1}^{m} u_k B_k}, \end{cases} \tag{7}$$

where

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

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

Then $\hat{\alpha}_m$, $\hat{\rho}_m$ *and* $\hat{\beta}_m$ *are consistent estimators of* α, ρ *and* β.

A simulation study is conducted to evaluate the performance of the proposed estimation. The choice of optimal value of m was proposed in [24], which suggest selecting points u_i in the interval $[0.1, 1]$. We give the proof of Theorem 3.1 in section 5. Theorem 3.1 contains the case of the Stable driven CIR and OU processes mentioned in the preceding section. Concerning the latter, the following result discusses the estimation of the drift coefficient parameters θ and μ with $T=1$.

Theorem 3.2. *Let* $X=(X_t)_{t\in[0,1]}$ *be a real valued stationnary and strongly mixing process satisfying (3) or (4). We assume that* X *is observed on a sample of discrete time points* t_i *with time frequency* Δ_n *of size* n. *Let* $(\hat{f}_n(X_{t_i}))_{i=1}^n$ *be the sequence of the N-W estimate of the linear drift function in (3) or (4). set*

$$\begin{cases} \hat{\theta}_n &= \dfrac{\sum_{i=1}^{n} X_{t_i}\hat{f}_n(X_{t_i}) - \frac{1}{n}\sum_{i=1}^{n} X_{t_i}\sum_{i=1}^{n}\hat{f}_n(X_{t_i})}{\frac{1}{n}\big(\sum_{i=1}^{n} X_{t_i}\big)^2 - \sum_{i=1}^{n} X_{t_i}^2} \\ \hat{\mu}_n &= \hat{\theta}_n^{-1}\frac{1}{n}\sum_{i=1}^{n}\hat{f}_n(X_{t_i}) + \dfrac{1}{n}\sum_{i=1}^{n} X_{t_i}. \end{cases} \tag{8}$$

Then $\hat{\theta}_n$ *and* $\hat{\mu}_n$ *are consistent estimators of* θ *and* μ *as* $n \to +\infty$.

Similarly, the above consistency result is is again applicable using the weighted least squares method with given weight $K(i)$. For instance, one can consider the following estimator:

$$\tilde{\theta}_n = \frac{\sum_{i=1}^n K(i) X_{t_i} \hat{f}_n(X_{t_i}) - \sum_{i=1}^n K(i) \hat{f}_n(X_{t_i}) \sum_{i=1}^n K(i) X_{t_i}}{(\sum_{i=1}^n K(i) X_{t_i})^2 - \sum_{i=1}^n K(i) X_{t_i}^2}.$$

4. Numerical Application

We now investigate the numerical performance of our estimation procedure on synthetic. After that, we apply them to real data. All codes were developed in R software.

4.1. **Simulated data.** We simulate and approximate X by using the Euler scheme on the interval $[0, T]$ with sample size $n = 500, 1000,\ 2000$ and within a period $T = 1$. In order to ensure that the Stable driven OU and CIR processes satisfies Assumption $(\mathbf{A_1})$, we shall employ the exponential ergodicity results we discuss in subsection 2.3 when choosing the model parameters. In the sequel, we consider the following common parameters:

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

We proceed by generating data from models 3 and 4 according to the previous setting. We plot the kernel density estimate of a realization of X with the histogram in order to check numerically the regularity condition on the density function of the stationary distribution π, assumed to be continuous in Assumption $(\mathbf{A_2})$. Therefore choosing a suitable time frequency Δ_n (for instance $\Delta_n = \frac{1}{\sqrt{n}}$), the models satisfies all the required conditions. We explore the performance of the Nadayaro Watson estimator of the drift function and the estimate parameters of the diffusion part, namely α, β and ρ in the context of the stable driven CIR and OU models on simulated data. Figures 3 and 4 display the graphical performance of the N-W estimator with respect to the true drift function considered ($\theta = 0.3, \quad \mu = 1.2$). The pictures show the comparison according to three sample sizes $n = 500,\ n = 1000$ and $n = 2000$ with time interval $T = 1$ or $T = 10$. We can see that varying T for a fixed n slightly seems not improve the estimates of the drift function in the SCIR model. This confirms that the drift function can not be identified in a fixed time interval in the framework of N-W estimator, no matter how frequently the observations are sampled. One notices that for a fixed length of observation interval T, as the sample size gets larger, the N-W estimator does not behave better, which is consistent with the asymptotic theory of the N-W estimator for stochastic processes driven by Lévy motions. The regularity condition on the density function of the stationary distribution is graphically checked throughout Figures 5 and 6.

To evaluate the performance, we make not only illustration by picture but also, via the square-Root of Average Square Errors defined by

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

where $x = (x_i)$ are chosen uniformly to cover the range of sample path of X. We summarized the performance in Tables 1 and 5, which reports the results on the diffusion parameters (the standard and strictly stable process parameters α and β) and the scaling parameter 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 improves the estimates of the stable process parameters α, β, except for the scaling parameter ρ. But for T fixed to be for instance one, varying n preserves the consistency of the scaling parameter ρ.

4.2. **Financial exchanges rates data.** We now turn to the estimation using real data. We analyse real financial data such as indexes or exchanges of the Canadian dollar against the US dollar in a given fixed period varying from May 2018 to June 2022 are analysed with a sample size $n = 1000$.

Figures 7 (a), (b) and (c), show the time series variation of the exchanges rates. We also look at the speed of decreasing of the Covariance using the autocorrelation function (ACF) diagram [3]. Indeed for a stationary time series, the ACF will fall to zero quickly whereas the ACF of non-stationary time series decreases slowly. The Dickey Fuller test returns a p-criticism of 0.01, which confirms the stationarity of the times series. The decreasing of the covariance using the ACF diagram test shows that the data sequence is weakly dependent. The data can be considered as mixing in the sense of the ρ-mixing (which implies strong mixing). The density estimation of the stationnary distribution is given also in Figures 7 (a'), (b') and (c') so that the regularity condition is satisfied.

Figures in 8 display the graphical NW estimation of the observed drift function. Tables 3 and 4 summarise the parameter estimation of the diffusion part. As concerned the drift part, Tables 2 summarise the estimated drift coefficients with the real data. We compare the prediction result with the stable OU model in Figure 9 since it seems to match better than the Stable CIR process.

5. Proofs of the main Theorems

In this section, we derive the proofs of our main results.

5.1. **Proof of Theorem 3.1.**

Proof. We recall first the Euler approximation of SDE (2.4):

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

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

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

as an estimate value of Y_i^n for all $i = 0, \ldots, n-1$; so that

$$Y_i^n \overset{d}{=} \frac{\Delta X_{t_i} - \hat{f}_n(X_{t_i})\Delta_n}{g(X_{t_i})},$$

where the equality $\overset{d}{=}$ is understood in the sense of law. We can define an approximate sample characteristic function by

$$\hat{\varphi}_n(u) = \frac{1}{n}\sum_{i=0}^{n-1} \exp\{ju\Big(\frac{\Delta X_{t_i} - \hat{f}_n(X_{t_i})\Delta_n}{g(X_{t_i})}\Big)\}, \quad \forall u \in \mathbb{R}, \tag{9}$$

which is asymptotically equal to the characteristic function $\varphi_{Z_{\Delta_n}}(u)$ of Z_{Δ_n} when n is large enough. Now, from the following formula:

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

we can set up a regression model in order to estimate α and ρ. To this end, we make use of the least squares method to minimize

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

where

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

To estimate β, we start from the following formula:

$$S_k = \beta B_k + \omega u_k,$$

and we use again the above least squares method to obtain the desired estimators and deduce their consistency in the classical framework of regression estimation method. □

5.2. **Proof of Theorem 3.4.**

Proof. Since the drift term in (4) or (3) is linear, we can set up the following linear regression model :

$$U_i = \mu\theta - \theta X_{t_i} + \varepsilon_i, \;\; i \geq 1$$

where ε_i is a zero mean IID sequence of given Gaussian random variable with variance say δ^2. We consider the Least Square Estimation (LSE) which consists in minimizing the following objective function:

$$\mathbb{G} = \sum_{i=1}^{n} \left(U_i - \theta\mu + \theta X_i\right)^2.$$

Thus it is straightforward to obtain the desired estimators. Now, let us prove the consistency of these estimators. Firstly, remark that

$$\begin{aligned}\sum_{i=1}^{n}\left(\frac{X_{t_i} - \frac{1}{n}\sum_{i=1}^{n} X_{t_i}}{\frac{1}{n}\left(\sum_{i=1}^{n} X_{t_i}\right)^2 - \sum_{i=1}^{n} X_{t_i}^2}\right)\mu\theta &= \frac{\sum_{i=1}^{n}\left(X_{t_i} - \frac{1}{n}\sum_{i=1}^{n} X_{t_i}\right)}{\frac{1}{n}\left(\sum_{i=1}^{n} X_{t_i}\right)^2 - \sum_{i=1}^{n} X_{t_i}^2}\mu\theta \\ &= \frac{\sum_{i=1}^{n} X_{t_i} - n \times \frac{1}{n}\sum_{i=1}^{n} X_{t_i}}{\frac{1}{n}\left(\sum_{i=1}^{n} X_{t_i}\right)^2 - \sum_{i=1}^{n} X_{t_i}^2}\mu\theta = 0\end{aligned}$$

and

$$\begin{aligned}-\sum_{i=1}^{n}\left(\frac{X_{t_i} - \frac{1}{n}\sum_{i=1}^{n} X_{t_i}}{\frac{1}{n}\left(\sum_{i=1}^{n} X_{t_i}\right)^2 - \sum_{i=1}^{n} X_{t_i}^2}\right)\theta X_{t_i} &= -\theta\sum_{i=1}^{n}\left(\frac{X_{t_i} - \frac{1}{n}\sum_{i=1}^{n} X_{t_i}}{\frac{1}{n}\left(\sum_{i=1}^{n} X_{t_i}\right)^2 - \sum_{i=1}^{n} X_{t_i}^2}\right)X_{t_i} \\ &= -\theta\sum_{i=1}^{n}\left(\frac{X_{t_i}^2 - \frac{1}{n}\left(\sum_{i=1}^{n} X_{t_i}\right)X_{t_i}}{\frac{1}{n}\left(\sum_{i=1}^{n} X_{t_i}\right)^2 - \sum_{i=1}^{n} X_{t_i}^2}\right) \\ &= -\theta\frac{\sum_{i=1}^{n}\left(X_{t_i}^2 - \frac{1}{n}\left(\sum_{i=1}^{n} X_{t_i}\right)\right)X_{t_i}}{\frac{1}{n}\left(\sum_{i=1}^{n} X_{t_i}\right)^2 - \sum_{i=1}^{n} X_{t_i}^2} \\ &= -\theta\frac{\sum_{i=1}^{n} X_{t_i}^2 - \frac{1}{n}\left(\sum_{1}^{n} X_{t_i}\right)\sum_{i=1}^{n} X_{t_i}}{\frac{1}{n}\left(\sum_{i=1}^{n} X_{t_i}\right)^2 - \sum_{i=1}^{n} X_{t_i}^2} \\ &= -\theta\frac{\sum_{i=1}^{n} X_{t_i}^2 - \frac{1}{n}\left(\sum_{i=1}^{n} X_{t_i}\right)^2}{\frac{1}{n}\left(\sum_{i=1}^{n} X_{t_i}\right)^2 - \sum_{i=1}^{n} X_{t_i}^2} \\ &= \theta.\end{aligned}$$

Consequently, $\hat{\theta}_n$ is rewritten as follows

$$\hat{\theta}_n = \theta + \sum_{i=1}^{n}\left(\frac{X_{t_i} - \frac{1}{n}\sum_{i=1}^{n} X_{t_i}}{\frac{1}{n}\left(\sum_{i=1}^{n} X_{t_i}\right)^2 - \sum_{i=1}^{n} X_{t_i}^2}\right)\varepsilon_i.$$

This show that $\hat{\theta}_n$ is an unbiased estimator of θ. Now, from the following relation,

$$\hat{\mu}_n = \hat{\theta}_n^{-1}\frac{1}{n}\sum_{i=1}^{n} U_i + \frac{1}{n}\hat{\theta}_n^{-1}\theta\sum_{i=1}^{n} X_{t_i},$$

we deduce that $\hat{\mu}_n$ is also an unbiased estimator of μ. To check the consistency, note that,

$$\hat{\theta}_n = \frac{\sum_{i=1}^{n} X_{t_i}U_i - \frac{1}{n}\sum_{i=1}^{n} U_i \sum_{i=1}^{n} X_{t_i}}{\frac{1}{n}\sum_{i=1}^{n} X_{t_i}\sum_{i=1}^{n} X_{t_i} - \sum_{i=1}^{n} X_{t_i}^2} = \frac{\sum_{i=1}^{n} X_{t_i}\Big(U_i - \bar{U}\Big)}{\sum_{i=1}^{n} X_{t_i}\Big(\bar{X} - X_{t_i}\Big)} \tag{10}$$

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

$$= \frac{\sum_{i=1}^{n}\Big(X_{t_i} - \bar{X}\Big)\Big(\bar{U} - U_i\Big)}{\sum_{i=1}^{n}\Big(X_{t_i} - \bar{X}\Big)^2} \text{since} \quad \sum_{i=1}^{n}\bar{X}(U_i - \bar{U}) = 0 \text{ and } \sum_{i=1}^{n}\bar{X}(X_{t_i} - \bar{X}) = 0. \tag{11}$$

Now, from the fact that $U_i - U = \theta(\bar{X} - X_{t_i}) + \epsilon_i$, we have

$$\hat{\theta}_n = \frac{\sum_{i=1}^{n}\Big(X_{t_i} - \bar{X}\Big)\Big(\theta(X_{t_i} - \bar{X}) - \varepsilon_i\Big)}{\sum_{i=1}^{n}\Big(\bar{X} - X_{t_i}\Big)^2} = \theta + \frac{\sum_{i=1}^{n}\Big(X_{t_i} - \bar{X}\Big)\varepsilon_i}{\sum_{i=1}^{n}\Big(\bar{X}_t - X_{t_i}\Big)^2}. \tag{12}$$

Recall that, the errors ε_i are uncorrelated and have the same variance δ^2. As a result, the variance of the sum is the sum of the variance. Hence we have:

$$\mathbb{V}ar[\hat{\theta}_n] = \mathbb{E}(\hat{\theta}_n - \theta)^2 = \delta^2\frac{\sum_{i=1}^{n}\big(X_{t_i} - \bar{X}\big)^2}{\Big(\sum_{i=1}^{n}(X_{t_i} - \bar{X})^2\Big)^2} = \frac{\delta^2}{\sum_{i=1}^{n}(X_{t_i} - \bar{X})^2} \underset{n\to+\infty}{\longrightarrow} 0, \tag{13}$$

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

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] Peter Carr and Liuren Wu. The finite moment log stable process and option pricing. *The journal of finance*, 58(2):753–777, 2003.

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

[9] 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.

[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] Richard Eon and Mihai Gradinaru. Distribution tails for solutions of sde driven by an asymmetric stable lévy process. *Probability and Mathematical Statistics*, 40(2):317–330, 2020.

[15] 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.

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

[17] 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.

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

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

[20] Ying Jiao, Chunhua Ma, and Simone Scotti. Alpha-cir model with branching processes in sovereign interest rate modeling. *Finance and Stochastics*, 21(3):789–813, 2017.

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

[22] Kiyoshi Kawazu and Shinzo Watanabe. Branching processes with immigration and related limit theorems. *Theory of Probability & Its Applications*, 16(1):36–54, 1971.

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

[24] 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.

[25] 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.

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

[27] 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.

[28] 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.

[29] Zenghu Li and Zenghu Li. Entrance laws and excursion laws. *Measure-Valued Branching Markov Processes*, pages 179–206, 2011.

[30] 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.

[31] 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.

[32] 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.

[33] Robert Ševilevič Lipcer, Robert S Liptser, Albert Nikolaevič Širâev, Albert N Shiryaev, et al. *Statistics of Random Processes II: II. Applications*, volume 2. Springer Science & Business Media, 2001.

[34] Hongwei Long and Lianfen Qian. Nadaraya-watson estimator for stochastic processes driven by stable lévy motions. 2013.

[35] Ross A Maller, Gernot Müller, and Alex Szimayer. Ornstein–uhlenbeck processes and extensions. *Handbook of financial time series*, pages 421–437, 2009.

[36] Solym Manou-Abi. *Théorèmes limites et ordres stochastiques relatifs aux lois et processus stables*. PhD thesis, Universite Toulouse 3 Paul Sabatier (UT3 Paul Sabatier), 2015.
[37] Solym Manou-Abi. Approximate solution for a class of stochastic differential equation driven by stable processes. 2023.
[38] H Masuda. Simple estimators for non-linear markovian trend from sampled data: I. ergodic cases. *MHF Preprint Series*, 7, 2005.
[39] 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.
[40] Elizbar A Nadaraya. On estimating regression. *Theory of Probability & Its Applications*, 9(1):141–142, 1964.
[41] John P Nolan. *Univariate stable distributions*. Springer, 2020.
[42] 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.
[43] Étienne Pardoux. *Probabilistic Models of Population Evolution: Scaling Limits, Genealogies and Interactions*, volume 1. Springer, 2016.
[44] Jun Peng. A stable cox–ingersoll–ross model with restart. *Journal of Mathematical Analysis and Applications*, 444(2):1185–1194, 2016.
[45] 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.
[46] BLS Prakasa Rao. Nonparametric estimation of linear multiplier in stochastic differential equations driven by stable noise. *arXiv e-prints*, pages arXiv–2109, 2021.
[47] 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.
[48] Simon J Sheather. Density estimation. *Statistical science*, pages 588–597, 2004.
[49] Yasutaka Shimizu. M-estimation for discretely observed ergodic diffusion processes with infinitely many jumps. *Statistical Inference for Stochastic Processes*, 9:179–225, 2006.
[50] 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.
[51] Jian Wang. On the exponential ergodicity of lévy-driven ornstein–uhlenbeck processes. *Journal of Applied Probability*, 49(4):990–1004, 2012.
[52] Geoffrey S Watson. Smooth regression analysis. *Sankhyā: The Indian Journal of Statistics, Series A*, pages 359–372, 1964.
[53] Chao Wei. Estimation for the discretely observed cox–ingersoll–ross model driven by small symmetrical stable noises. *Symmetry*, 12(3):327, 2020.
[54] Wei Biao Wu. Nonparametric estimation for stationary processes. *University of Chicago. Technical Report*, 536, 2003.
[55] Diethelm Wuertz, Martin Maechler, and Maintainer Martin Maechler. Package 'stabledist'. *URL https://cran. r-project. org/web/packages/stabledist/stabledist. pdf*, 2016.
[56] Xu Yang. Maximum likelihood type estimation for discretely observed cir model with small α-stable noises. *Statistics & Probability Letters*, 120:18–27, 2017.
[57] Xiaolong Zhang and Xicheng Zhang. Ergodicity of supercritical sdes driven by α-stable processes and heavy-tailed sampling. *Bernoulli*, 29(3):1933–1958, 2023.
[58] 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.

[59] 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

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 with the exchange rates

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 parameters with the exchange rates using an unknown strictly standard stable driven OU process

n	**Estimated drift coefficients**
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. Estimated diffusion parameters with the exchange rates using an unknown strictly standard stable driven CIR 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	$\hat{\alpha}_n = 1.44, \quad \hat{\beta}_n = 0.08, \quad \hat{\rho}_n = 5386$

TABLE 5. Performance of the estimated diffusion and scaling parameters $(\hat{\alpha}, \hat{\beta}, \hat{\rho})$ with a standard symmetric 1.6-stable driven CIR process

n and T	**True parameters**	**Estimate Parameters**	RMSE
500 and T=1	$\alpha = 1.6, \quad \beta = 0, \quad \rho = 1$	$\hat{\alpha}_n = 1.66, \quad \hat{\beta}_n = 0.13, \quad \hat{\rho}_n = 0.88$	1.17
500 and T=10	$\alpha = 1.6, \quad \beta = 0, \quad \rho = 1$	$\hat{\alpha}_n = 1.51, \quad \hat{\beta}_n = 0.08, \quad \hat{\rho}_n = 5.95$	8.24
1000 and T=1	$\alpha = 1.6, \quad \beta = 0, \quad \rho = 1$	$\hat{\alpha}_n = 1.63, \quad \hat{\beta}_n = 0.217, \quad \hat{\rho}_n = 0.98$	3.72
1000 and T=10	$\alpha = 1.6, \quad \beta = 0, \quad \rho = 1$	$\hat{\alpha}_n = 1.53, \quad \hat{\beta}_n = 0.35, \quad \hat{\rho}_n = 6.32$	15.36
2000 and T=1	$\alpha = 1.6, \quad \beta = 0, \quad \rho = 1$	$\hat{\alpha}_n = 1.61, \quad \hat{\beta}_n = 0.07, \quad \hat{\rho}_n = 1.03$	2.07
2000 and T=10	$\alpha = 1.6, \quad \beta = 0, \quad \rho = 1$	$\hat{\alpha}_n = 1.64, \quad \hat{\beta}_n = 0.04, \quad \hat{\rho}_n = 4$	11.13

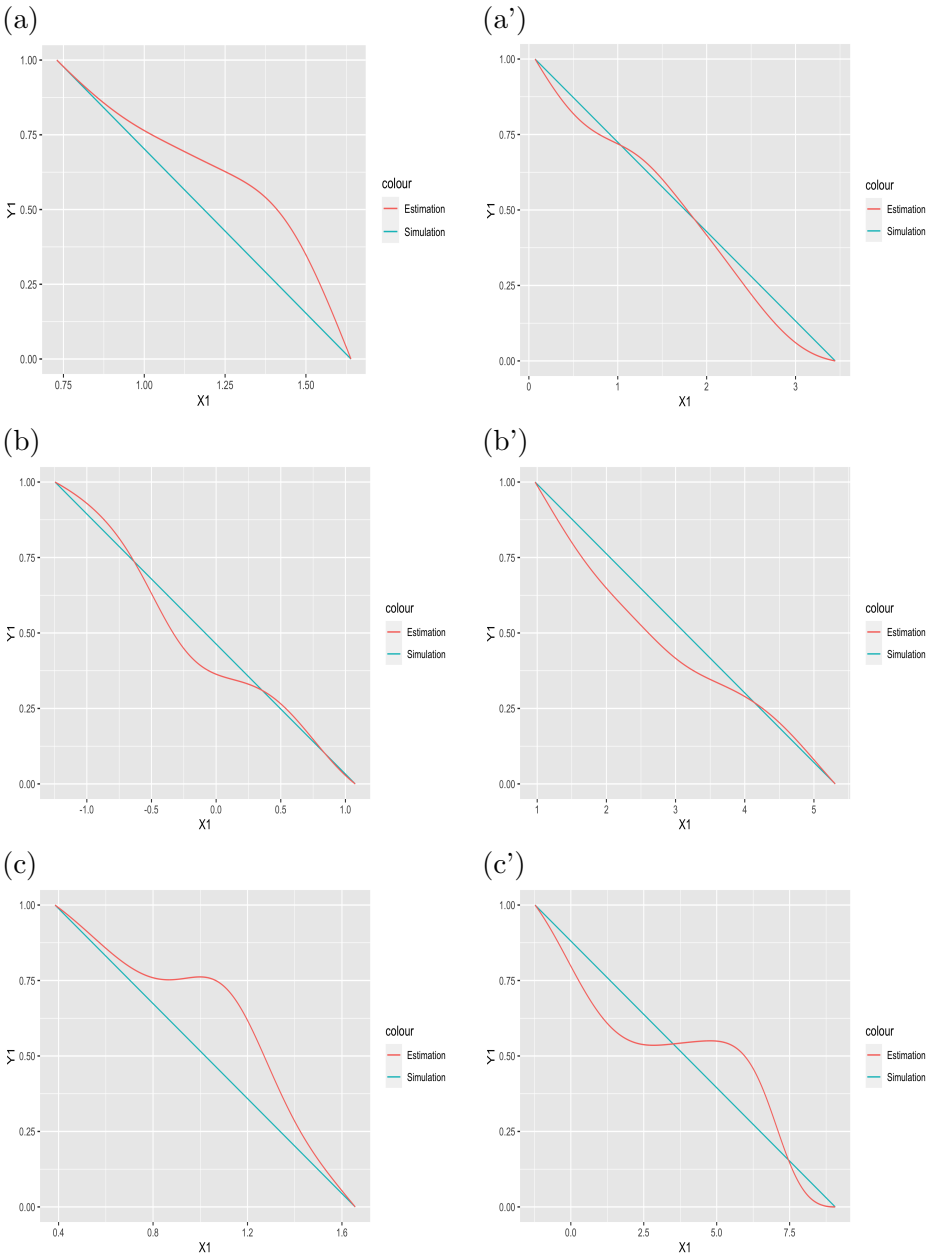

FIGURE 3. Performance of the Nadayaro Watson estimator with respect to the true drift in the case of a standard symmetric 1.6-stable driven OU process with sample size $n = 500$ and $t = 1$ in (a), sample size $n = 500$ and $t = 10$ in (a'), sample size $n = 1000$ and $t = 1$ in (b), sample size $n = 1000$ and $t = 10$ in (b'), sample size $n = 2000$ and $t = 1$ in (c) and sample size $n = 2000$ and $t = 10$ in (c').

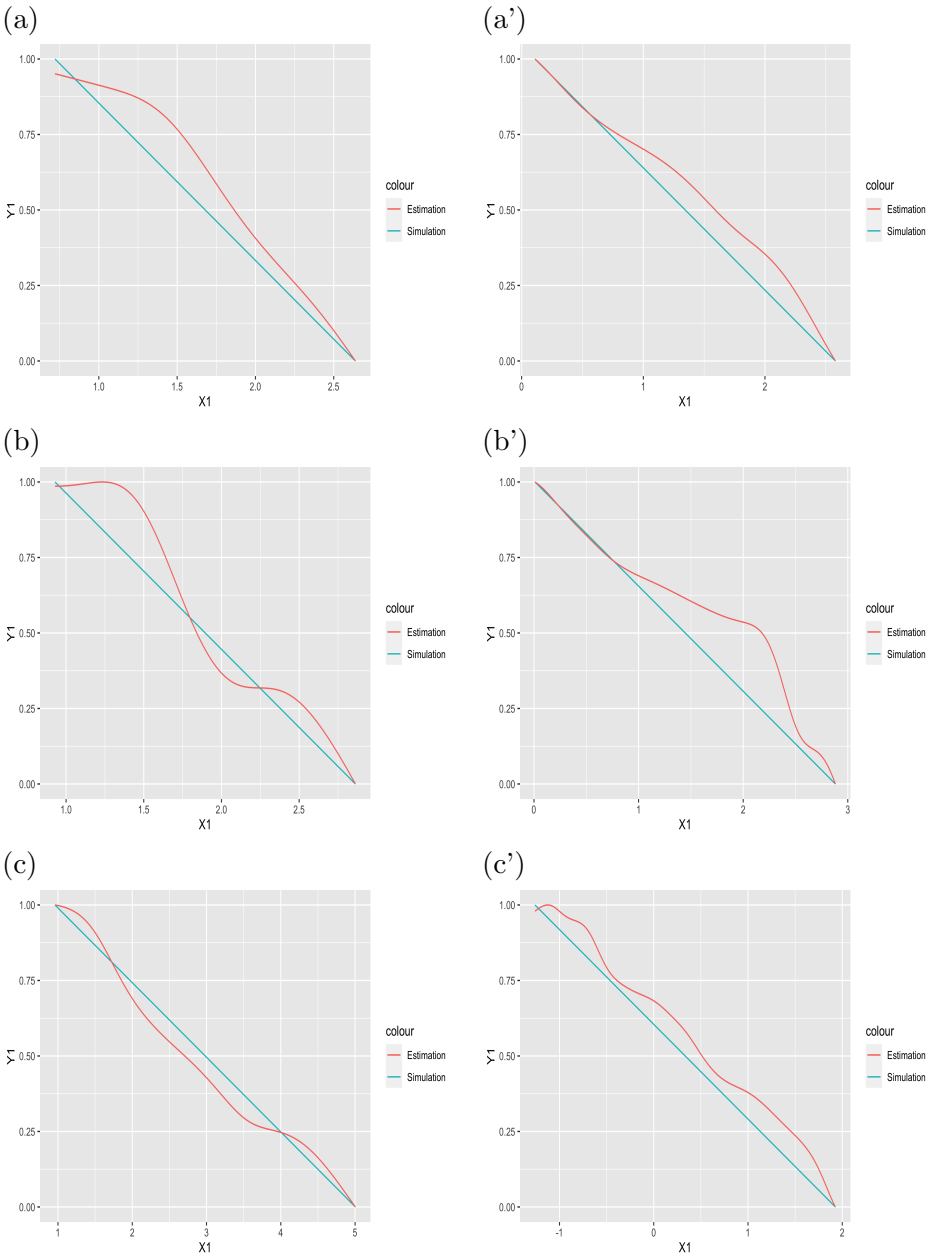

FIGURE 4. Performance of the Nadayaro Watson estimator with respect to the true drift in the case of a standard symmetric 1.6-stable driven CIR process with sample size $n = 500$ and $t = 1$ in (a), sample size $n = 500$ and $t = 10$ in (a'), sample size $n = 1000$ and $t = 1$ in (b), sample size $n = 1000$ and $t = 10$ in (b'), sample size $n = 2000$ and $t = 1$ in (c) and sample size $n = 2000$ and $t = 10$ in (c').

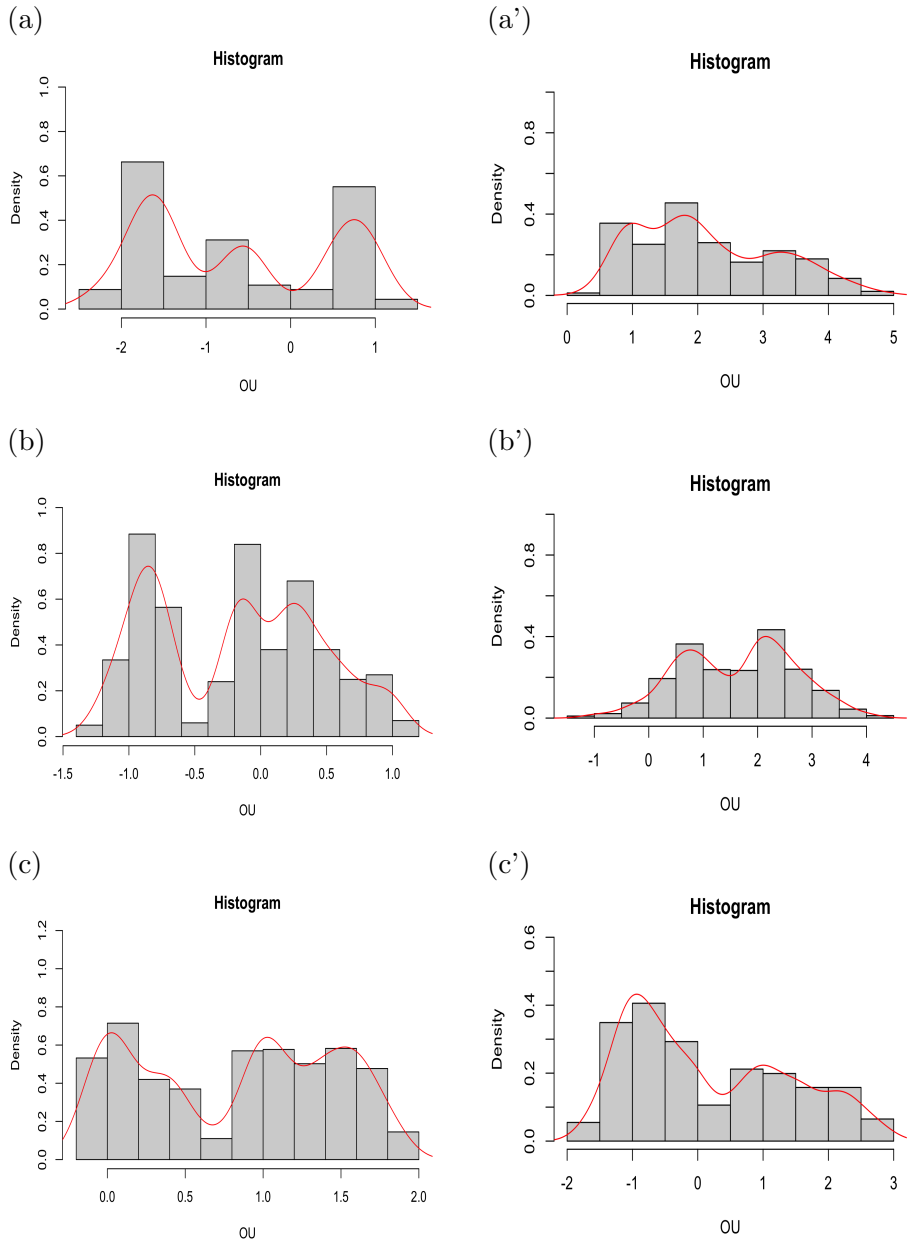

FIGURE 5. Kernel density estimate of a standard symmetric 1.6-stable driven OU process with sample size $n = 500$ and $t = 1$ in (a), sample size $n = 500$ and $t = 10$ in (a'), sample size $n = 1000$ and $t = 1$ in (b), sample size $n = 1000$ and $t = 10$ in (b'), sample size $n = 2000$ and $t = 1$ in (c) and sample size $n = 2000$ and $t = 10$ in (c').

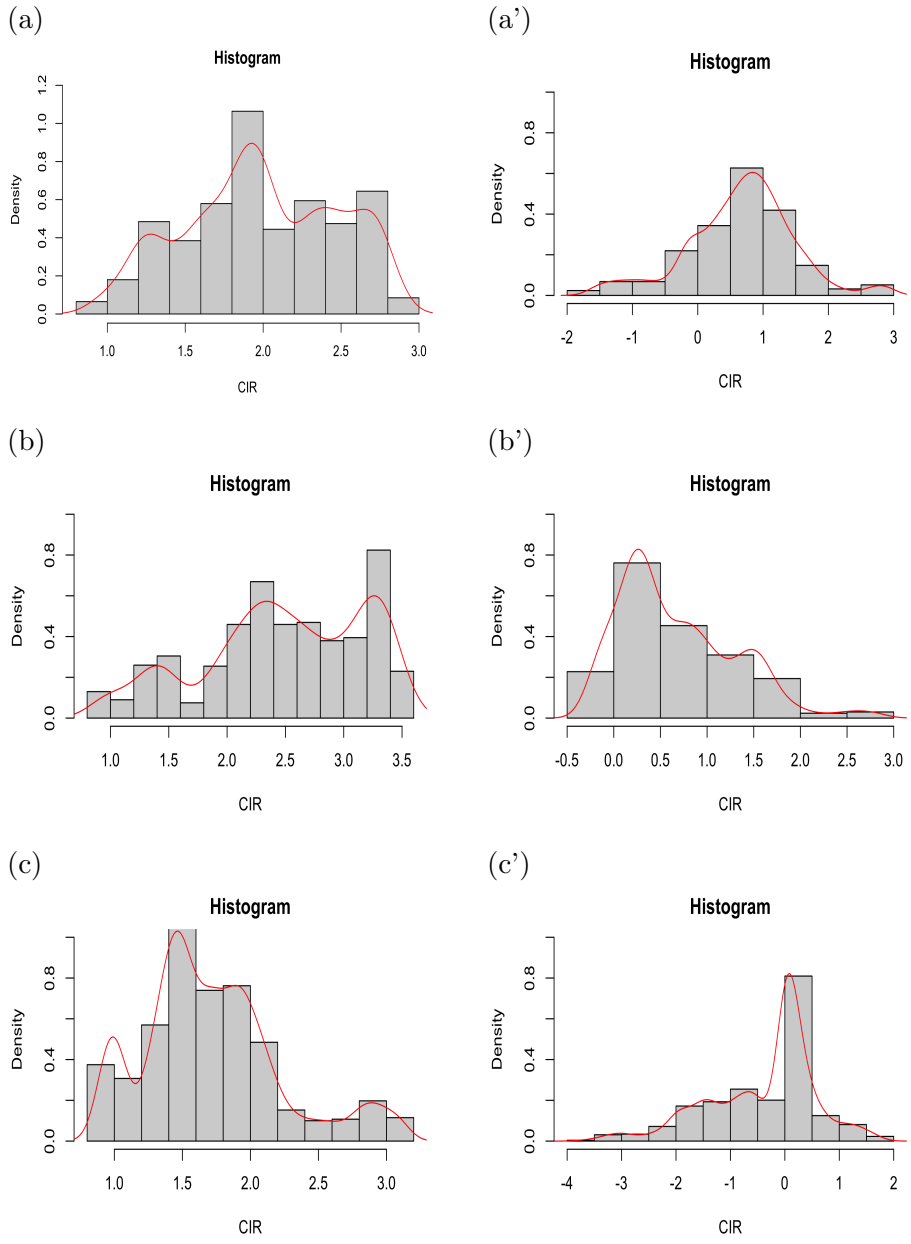

FIGURE 6. Kernel density estimate of a standard symmetric 1.6-stable driven CIR process with sample size $n = 500$ and $t = 1$ in (a), sample size $n = 500$ and $t = 10$ in (a'), sample size $n = 1000$ and $t = 1$ in (b), sample size $n = 1000$ and $t = 10$ in (b'), sample size $n = 2000$ and $t = 1$ in (c) and sample size $n = 2000$ and $t = 10$ in (c').

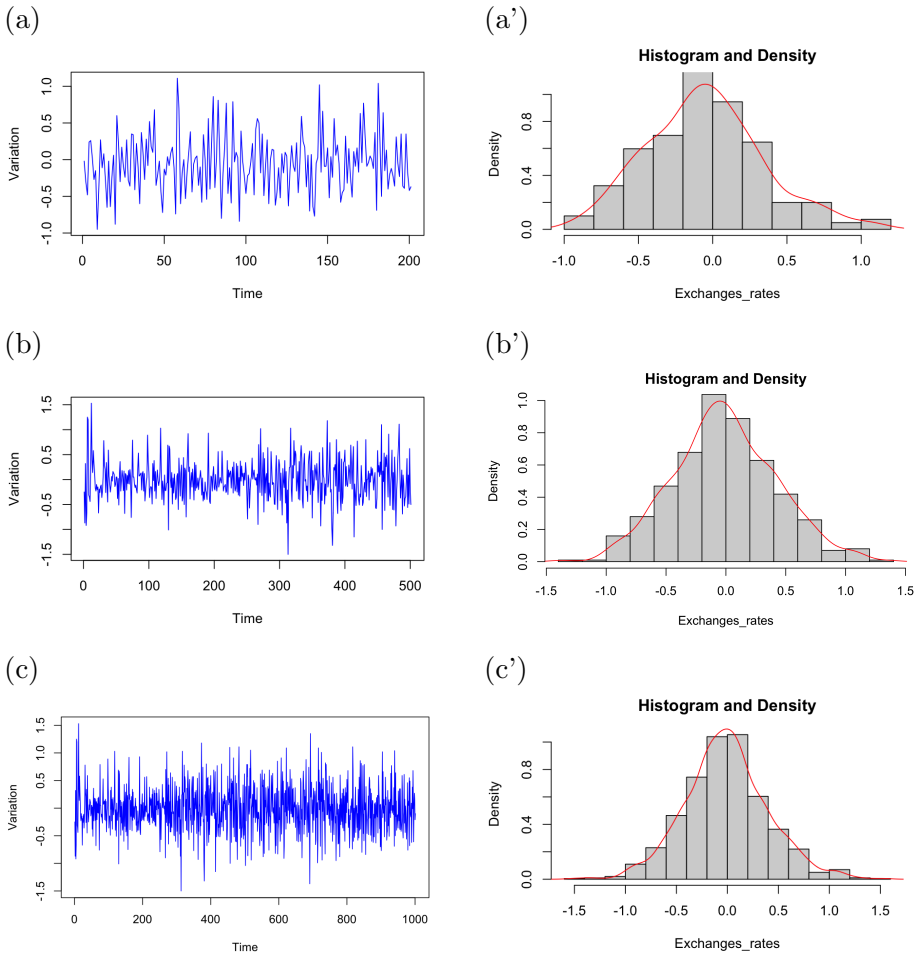

FIGURE 7. The exchange rates of the Canadian dollar against the US dollar global variation and density estimation with t sample size $n = 1000$ (a); local sample size $n = 500$ (b) and sample size $n = 200$ (c).

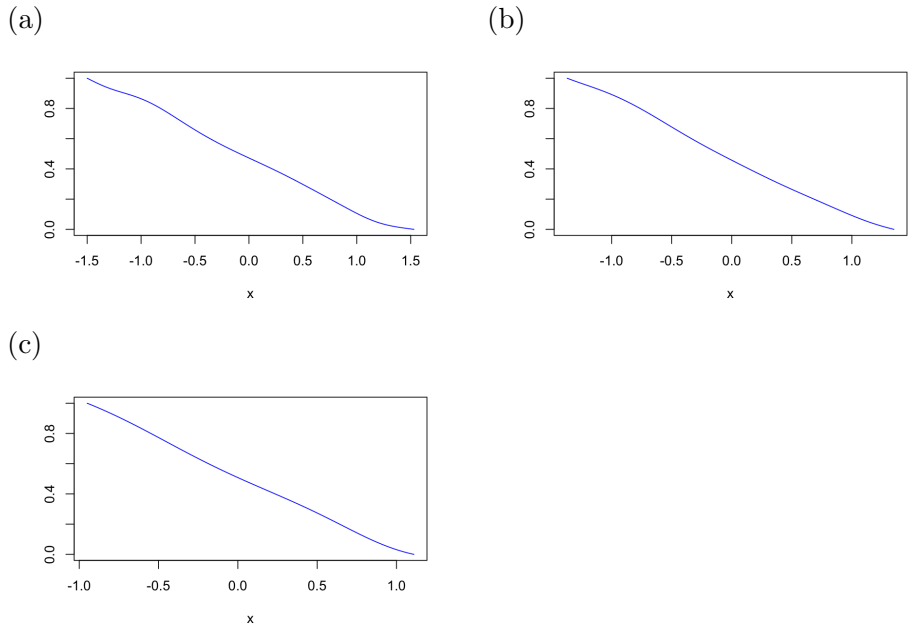

FIGURE 8. NW drift estimated function with the exchange rates of the Canadian dollar against the US dollar variation with the global sample size $n = 1000$ (a); local sample size $n = 500$ (b) and sample size $n = 200$ (c).

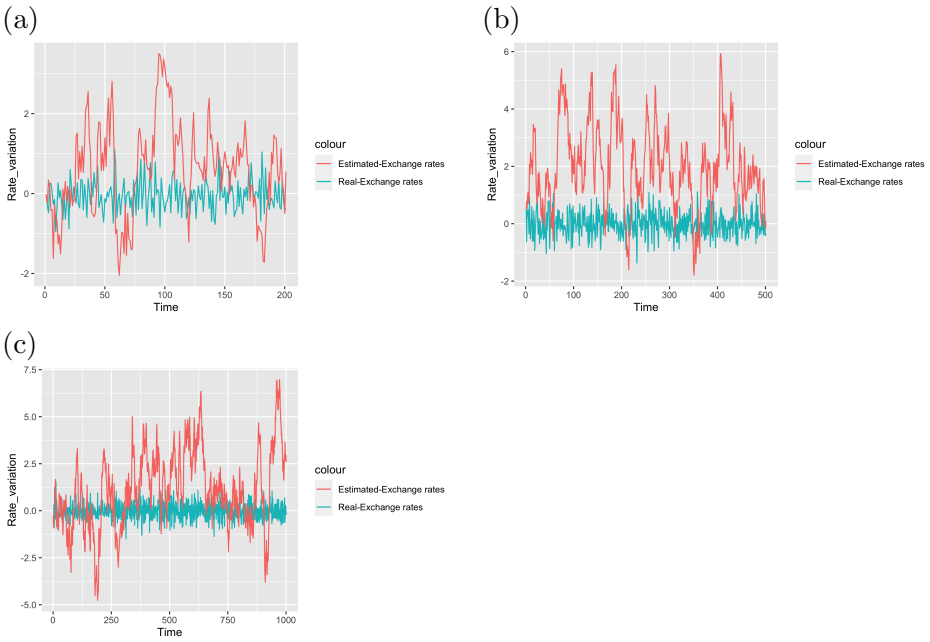

FIGURE 9. Comparison between exchange rates prediction with Stable driven OUP with sample size $n = 200$ (a); sample size $n = 500$ (b) and sample size $n = 1000$ (c).