Weierstraß-Institut
für Angewandte Analysis und Stochastik
Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 2198-5855

Maximum likelihood drift estimation for a threshold diffusion

Antoine Lejay[1,2], Paolo Pigato[3]

submitted: March 22, 2018

[1] Université de Lorraine, CNRS
IECL
UMR 7502
54600 Vandœuvre-lès-Nancy
France
E-Mail: antoine.lejay@univ-lorraine.fr

[2] Inria
54600 Villers-lès-Nancy
France
E-Mail: antoine.lejay@inria.fr

[3] Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: paolo.pigato@wias-berlin.de

No. 2497
Berlin 2018

2010 *Mathematics Subject Classification.* 62M05, 62F12, 60J60.

Key words and phrases. Threshold diffusion, oscillating Brownian motion, maximum likelihood estimator, null recurrent process, ergodic process, transient process, mixed normal distribution.

P. Pigato gratefully acknowledges financial support from ERC via Grant CoG-683164.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Maximum likelihood drift estimation for a threshold diffusion

Antoine Lejay , Paolo Pigato

Abstract

We study the maximum likelihood estimator of the drift parameters of a stochastic differential equation, with both drift and diffusion coefficients constant on the positive and negative axis, yet discontinuous at zero. This threshold diffusion is called the drifted Oscillating Brownian motion. The asymptotic behaviors of the positive and negative occupation times rule the ones of the estimators. Differently from most known results in the literature, we do not restrict ourselves to the ergodic framework: indeed, depending on the signs of the drift, the process may be ergodic, transient or null recurrent. For each regime, we establish whether or not the estimators are consistent; if they are, we prove the convergence in long time of the properly rescaled difference of the estimators towards a normal or mixed normal distribution. These theoretical results are backed by numerical simulations.

1 Introduction

We consider the process, called a *drifted Oscillating Brownian motion* (DOBM), which is the solution to the Stochastic Differential equation (SDE)

$$\xi_t = \xi_0 + \int_0^t \sigma(\xi_s)\,\mathrm{d}W_s + \int_0^t b(\xi_s)\,\mathrm{d}s, \tag{1}$$

with

$$\sigma(x) = \begin{cases} \sigma_+ > 0 & \text{if } x \geq 0, \\ \sigma_- > 0 & \text{if } x < 0 \end{cases} \quad \text{and} \quad b(x) = \begin{cases} b_+ \in \mathbb{R} & \text{if } x \geq 0, \\ b_- \in \mathbb{R} & \text{if } x < 0. \end{cases} \tag{2}$$

The strong existence to (1) follows for example from the results of [28]. Separately on $\mathbb{R}_+$ and $\mathbb{R}_-$, the dynamics of such process is the one of a Brownian motion with drift, with threshold and regime-switch at 0, consequence of the discontinuity of the coefficients.

This model can be seen as an alternative to the model studied in [36], which is a continuous time version of the Self-Exciting Threshold Autoregressive models (SETAR), a subclass of the TAR models [45, 46].

The practical interest of such processes are numerous. In finance, we show in [31] that an exponential form of this process generalizes the Black & Scholes model in a way to model leverage effects. Moreover, the introduction of a piecewise constant drift such as the one in (2) is a straightforward way to produce a mean-reverting process, if $b_+ < 0$ and $b_- > 0$. In [31], we find some evidence on empirical financial data that this may be the case. This corroborates other studies with different models [35, 38, 44].

Still in finance, the solution to (1) models other quantities than stocks. In [13], Eq. (1) with constant volatility serves as a model for the surplus of a company after the payment of dividends, which are payed only if the profits of the company are higher than a certain threshold. Similar threshold dividend

DOI 10.20347/WIAS.PREPRINT.2497

pay-out strategies are considered in [3]. In these works, the behavior of the process at the discontinuity is referred to as “refraction”. SETAR models have also applications to deal with transaction costs or regulator interventions [49], to interest and exchange rates [6, 9], ...

More general discontinuous drifts and volatilities arise in presence of Atlas models and other ranks based models [17]. SDE with discontinuous coefficients have also numerous applications in physics [39, 42], meteorology [15] and many other domains.

In [43, 44], F. Su and K.-S. Chan study the asymptotic behavior of the quasi-likelihood estimator of a diffusion with piecewise regular diffusivity and piecewise affine drift with an unknown threshold. The quasi-likelihood they use is based on the Girsanov density where the diffusivity is replaced by 1. In particular, they construct some hypothesis test to decide whether or not the drift is affine or piecewise affine in the ergodic situation.

In [26], Y. Kutoyants consider the estimation of a threshold r of a diffusion with a known or unknown drift switching at r. His results are then specialized to Ornstein-Uhlenbeck type processes. Also this framework assumes that the diffusion is ergodic.

In the present paper, we derive some maximum likelihood estimators for the drift parameters b_- and b_+ from continuous observations. We study their asymptotic behavior as the time tends to infinity in order to derive some confidence intervals when available. This article completes [30], where we estimate (σ_-, σ_+) for high-frequency data. We use our estimators on financial historical data in [31].

Our estimators of $b_T^\pm$ are

$$\beta_T^\pm = \pm \frac{(\pm\xi_T) \vee 0 - (\pm\xi_0) \vee 0 - L_T(\xi)/2}{Q_T^\pm},$$

where Q_T^+ (resp. Q_T^-) is the occupation time of the positive (resp. negative) side of the real axis, and $L_T(\xi)$ is the symmetric local time of ξ at 0. Estimators for $Q^\pm$ and $L_T(\xi)$ are quite straightforward to implement from discrete observations of a trajectory of ξ, and so are estimators for $b_\pm$. As for the estimators of (σ_-, σ_+) in [30], the local time and the occupation times play a central role in the study of the estimators of (b_-, b_+).

The long time asymptotic regime of the process depends on the respective signs of the coefficients (b_-, b_+). Using symmetries, this leads 5 different cases in which the process may be ergodic, null recurrent or transient and the estimators have different asymptotic behaviors. In some situations, the estimators are not convergent. In others, we establish consistency as well as Central Limit Theorems, with speed $T^{1/2}$ or $T^{1/4}$, depending again on the signs of $b_\pm$. We summarize in Table 1 the various asymptotic behaviors. We are in a situation close to the one encountered by M. Ben Alaya and A. Kebaier in [1] for estimating square-root diffusions, where several situations shall be treated. The work [26, 43, 44] mentioned above only consider ergodic situations. Non-parametric estimation of the drift in the recurrent case is considered in [4].

Finally, we develop in Section 7.1 a hypothesis test for the value of the drift which is based on the Wilk's theorem, which relates asymptotically the log-likelihood to a χ^2 distribution with 2 degrees of freedom. Besides, we show in Section 7.2 the Local Asymptotic Normality (LAN [27, 29]) and the Local Asymptotic Mixed Normality (LAMN [22]) in the ergodic case and the null recurrent case with non vanishing drift. These LAN/LAMN properties are related to the efficiency of the operators. The Wilk's as well as the LAN/LAMN properties are proved by combining the quadratic nature of the log-likelihood with our martingale central limit theorems.

Outline. In Section 2, we present the maximum likelihood estimator, which is based on the Girsanov

			$\beta_T^+ - b_+$	$\beta_T^- - b_-$
(E)	$b_+ < 0, b_- > 0$	ergodic	$\approx \frac{1}{\sqrt{T}}\sqrt{1-\frac{b_+}{b_-}}\sigma_+\mathcal{N}$	$\approx \frac{1}{\sqrt{T}}\sqrt{1-\frac{b_-}{b_+}}\sigma_-\mathcal{N}$
(N0)	$b_+ = 0, b_- = 0$	null recurrent	$\frac{1}{\sqrt{T}}\beta_1^+$	$\frac{1}{\sqrt{T}}\beta_1^-$
(N1)	$b_+ = 0, b_- > 0$	null recurrent	$\approx \frac{1}{\sqrt{T}}\sigma_+\mathcal{N}^+$	$\approx \frac{1}{T^{1/4}}\sigma_-\frac{\sqrt{b_-}}{\sqrt{\sigma_+}}\frac{\mathcal{N}^-}{\sqrt{\lvert\mathcal{N}\rvert}}$
(T0)	$b_+ > 0, b_- \geq 0$	transient	$\approx \frac{1}{\sqrt{T}}\sigma_+\mathcal{N}$	$\mathcal{R}_{\mathbf{T0}}$ as $T \to \infty$
(T1)	$b_+ > 0, b_- < 0$	transient	$\approx \frac{1}{\sqrt{T}}\sigma_+\mathcal{N}$	$\mathcal{R}_{\mathbf{T1}}^+$ as $T \to \infty$

Table 1: Asymptotic behavior of estimators, where $\mathcal{N}$, $\mathcal{N}^+$ and $\mathcal{N}^-$ are independent, unit Gaussian variables. The law of (β_1^-, β_1^+) in case **(N0)** is given in (25). The r.v.s $\mathcal{R}_{\mathbf{T0}}$ and $\mathcal{R}_{\mathbf{T1}}^+$ follow the law in (22). Results of both sides in **(T1)** are wrt to $\mathbb{P}_+$ (cf. Proposition 6), which intuitively can be thought as conditioning to the process diverging towards postive infinity.

transform. In Section 3, we characterize the different regimes of the process accordingly to the signs of the drifts. Our main results are presented in Section 4. The limit theorems that we use are presented in Section 5. The proofs for each cases are detailed in Section 6. We present the Wilk theorem and the LAN/LAMN property in Section 7. Finally, in Section 8, we conclude this article with numerical experiments.

2 The maximum likelihood estimator

In this section, we propose and discuss an estimator for the parameters (b_-, b_+) of the drift coefficient of ξ from continuous time observations

Data 1. *We observe of a path $(\xi_t)_{t\in[0,T]}$ on the time interval $[0,T]$ of the solution to* (1)*, together with its negative and positive occupation times*

$$Q_T^- = \int_0^T \mathbf{1}_{\xi_s<0}\,\mathrm{d}s \text{ and } Q_T^+ = \int_0^T \mathbf{1}_{\xi_s>0}\,\mathrm{d}s,$$

as well as its symmetric local time

$$L_T(\xi) = \lim_{\epsilon\to 0}\frac{1}{2\epsilon}\int_0^T \mathbf{1}_{-\epsilon\leq\xi_s\leq\epsilon}\,\mathrm{d}s.$$

The coefficients (σ_-, σ_+) are known.

Remark 1. Approximations of $(Q_T^-, Q_T^+, L_T(\xi))$ are easy to construct from $(\xi_t)_{t\in[0,T]}$ so that observing $(\xi_t)_{t\in[0,T]}$ is sufficient to build approximations of our estimator. This is detailed in Section 8.1.

The Girsanov weight of the distribution of (1), with respect to the one of the solution to $\mathrm{d}\xi_t = \sigma(\xi_t)\,\mathrm{d}\widetilde{W}$, for a Brownian motion $\widetilde{W}$, is

$$G(b_-, b_+) = \exp\left(\int_0^T \frac{b(\xi_s)}{\sigma(\xi_s)}\,\mathrm{d}\widetilde{W}_s - \frac{1}{2}\int_0^T \frac{b^2(\xi_s)}{\sigma^2(\xi_s)}\,\mathrm{d}s\right). \tag{3}$$

A reasonable way to set up an estimator of (b_-, b_+) is to consider $G(b_-, b_+)$ as a likelihood and to optimize this quantity. This is how estimators for the drift are classically constructed [25, 33].

Notation 1. To avoid confusion with the $+$ and $-$ used as indices, we write $(\!|x|\!)^+ := \max\{x,0\}$ for the positive part of x and $(\!|x|\!)^- := \max\{-x,0\} \geq 0$ for the negative part.

Proposition 1. *The likelihood $G(b_-,b_+)$ is maximal at (β_T^-,β_T^+) given by*

$$\beta_T^\pm = \pm\frac{(\!|\xi_T|\!)^\pm - (\!|\xi_0|\!)^\pm - L_T(\xi)/2}{Q_T^\pm}. \tag{4}$$

Proof. Let us denote by ξ the canonical process. Let us consider the measure $\mathbb{P}$ such that ξ is the unique solution to $\mathrm{d}\xi_t = \sigma(\xi_t)\,\mathrm{d}\widetilde{W}_t$, $t\in[0,T]$ for a Brownian motion $\widetilde{W}$.

Under the distribution $\mathbb{Q}$, with density $G(b_-,b_+)$ given by (3), with respect to $\mathbb{P}$, the process ξ is solution to

$$\mathrm{d}\xi_t = \sigma(\xi_t)\,\mathrm{d}W_t + b(\xi_t)\,\mathrm{d}t$$

for W defined by $W_t = \widetilde{W}_t - \int_0^t \frac{b}{\sigma}(\xi_s)\,\mathrm{d}s$. Under $\mathbb{Q}$, the process W is a Brownian motion.

We define $\gamma(x) = b(x)/\sigma^2(x)$ and $F(x) = \gamma(x)x$ for $x\in\mathbb{R}$. The function F is piecewise linear with

$$F'(x) = \gamma(x) \text{ for } x \neq 0 \text{ and } \beta = \gamma(0+) - \gamma(0-) = \frac{b_+}{\sigma_+^2} - \frac{b_-}{\sigma_-^2}.$$

From the Itô-Tanaka formula[1] [23, Theorem 7.1, p. 218],

$$F(\xi_t) - F(\xi_0) = \int_0^t F'_-(\xi_s)\,\mathrm{d}\xi_s + \frac{\beta}{2}L_t(\xi) = \int_0^t \gamma(\xi_s)\,\mathrm{d}\xi_s + \frac{\beta}{2}L_t(\xi),\ t\in[0,T].$$

Since $\gamma(\xi_s)\,\mathrm{d}\xi_s = b(\xi_s)/\sigma(\xi_s)\,\mathrm{d}\widetilde{W}_s$,

$$\int_0^T \frac{b(\xi_s)}{\sigma(\xi_s)}\,\mathrm{d}\widetilde{W}_s = F(\xi_T) - F(\xi_0) - \frac{\beta}{2}L_T(\xi).$$

Injecting this in the formula (3),

$$\begin{aligned}\log G(b_+,b_-) &= F(\xi_T) - F(\xi_0) - \frac{\beta}{2}L_T(\xi) - \frac{1}{2}\int_0^T \frac{b^2(\xi_s)}{\sigma^2(\xi_s)}\,\mathrm{d}s \\ &= F(\xi_T) - F(\xi_0) - \frac{\beta}{2}L_T(\xi) - \frac{b_+^2}{2\sigma_+^2}Q_T^+ - \frac{b_-^2}{2\sigma_-^2}Q_T^-. \end{aligned}\tag{5}$$

Maximizing (5) over b_+ and b_- leads to (4). □

The proof of the next lemma is a direct consequence of the Itô-Tanaka formula. It is the key to study the asymptotic behavior of $\beta_\pm$.

Lemma 1. *For any $T\geq 0$,*

$$\beta_T^\pm = b_\pm + \frac{M_T^\pm}{Q_T^\pm},$$

where $M^\pm := \pm\int_0^\cdot \sigma_\pm \mathbf{1}_{\pm\xi_s\geq 0}\,\mathrm{d}B_s$ are continuous time martingales with $\langle M^\pm\rangle = \sigma_\pm^2 Q^\pm$ and $\langle M^+,M^-\rangle = 0$.

The occupation time is non decreasing. For the sake of simplicity, let us write

$$Q_\infty^\pm = \lim_{T\to\infty} Q_T^\pm \in \mathbb{R}_+\cup\{\infty\}.$$

[1] Our local time $L(\xi)$ is twice the one that appear in Theorem 7.1 of [23].

3 Analytic characterization of the regime of the process

3.1 Scale function and speed measure

A well known fact [19, 23, 41] states that the infinitesimal generator $(\mathcal{L}, \mathrm{Dom}(\mathcal{L}))$ of the process ξ solution to (1) may be written as

$$\mathcal{L}f = \frac{1}{2}\sigma^2(x)e^{-h(x)}\frac{\mathrm{d}}{\mathrm{d}x}\left(e^{h(x)}\frac{\mathrm{d}f(x)}{\mathrm{d}x}\right) \text{ with } h(x) = \int_0^x \frac{2b(y)}{\sigma^2(y)}\,\mathrm{d}y$$
$$\text{for all } f \in \mathrm{Dom}(\mathcal{L}) = \{f \in \mathcal{C}_0(\mathbb{R}) \mid \mathcal{L}f \in \mathcal{C}_0(\mathbb{R})\}.$$

The process X is fully characterized by its *speed measure* M with a density m and its *scale function* S with

$$m(x) := \frac{2}{\sigma(x)^2}\exp(h(x)) \text{ and } S(x) := \int_0^x \exp(-h(y))\,\mathrm{d}y. \tag{6}$$

3.2 The regimes of the process

The diffusion X is either recurrent or transient. If $\lim_{x\to+\infty} S(x) = +\infty$ and $\lim_{x\to-\infty} S(x) = -\infty$, then the process is (positively or null) *recurrent*. Otherwise, it is *transient* [19, 23].

When $b(x) = b_+$ for $x \geq 0$,

$$S(x) = \begin{cases} x & \text{if } b_+ = 0, \\ \dfrac{\sigma_+^2}{2b_+}\left(1 - \exp\left(-\dfrac{2b_+x}{\sigma_+^2}\right)\right) & \text{if } b_+ > 0, \\ \dfrac{\sigma_+^2}{2|b_+|}\left(\exp\left(\dfrac{2|b_+|x}{\sigma_+^2}\right) - 1\right) & \text{if } b_+ < 0. \end{cases}$$

Similar formulas hold for b_-. Hence, the process ξ is transient if only if $b_+ > 0$ or $b_- < 0$.

A recurrent process is either *null recurrent* or *positive recurrent*. The process is positive recurrent if and only if $M(\mathbb{R}) := \int_{\mathbb{R}} m(x)\,\mathrm{d}x < +\infty$, in which case it is actually *ergodic*. Therefore, the process ξ is ergodic if and only if $b_+ < 0$ and $b_- > 0$. Otherwise, the process ξ is only null recurrent.

When the process is ergodic ($b_+ < 0, b_- > 0$), its invariant measure is

$$\frac{m(x)}{M(\mathbb{R})}\,\mathrm{d}x = \begin{cases} \dfrac{b_-}{b_- + |b_+|}e^{\frac{-2x|b_+|}{\sigma_+^2}} & \text{if } x \geq 0, \\ \dfrac{|b_+|}{b_- + |b_+|}e^{\frac{2xb_-}{\sigma_-^2}} & \text{if } x < 0. \end{cases}$$

Therefore, the regimes of ξ depends only on the respective signs of b_+ and b_-. Nine combinations are possible. As some cases are symmetric, we actually consider five cases exhibiting different asymptotic behaviors of $Q_T^\pm$, hence of the estimators. This is summarized in Table 2.

These cases are:

- **E)** Ergodic case $b_+ < 0, b_- > 0$.
- **N0)** Null recurrent case $b_+ = 0, b_- = 0$.
- **N1)** Null recurrent case $b_+ = 0, b_- > 0$.

	$b_+ > 0$	$b_+ = 0$	$b_+ < 0$
$b_- > 0$	transient **T0**	null recurrent **N1**	ergodic **E**
$b_- = 0$	transient **T0**	null recurrent **N0**	null recurrent **N1**
$b_- < 0$	transient **T1**	transient **T0**	transient **T0**

Table 2: Recurrence and transience properties of ξ.

T0) Transient case $b_+ > 0, b_- \geq 0$.
T1) Transient case $b_+ > 0, b_- < 0$.

Case **T0** corresponds to two entries of table 2. The case $b_+ < 0, b_- = 0$ is symmetric to **N1**. Case $b_+ \leq 0, b_- < 0$ is symmetric to **T0**.

4 Asymptotic behavior of the estimators

In this section, we state our main results on the asymptotic behavior of the occupation times of the process and the corresponding ones of the estimators, for each of the 5 cases.

Proposition 2 (Ergodic case **E**). *If $b_+ < 0, b_- > 0$, then*

$$\left(\frac{Q_T^+}{T}, \frac{Q_T^-}{T}\right) \xrightarrow[T\to\infty]{\text{a.s.}} \left(\frac{|b_-|}{|b_-| + |b_+|}, \frac{|b_+|}{|b_-| + |b_+|}\right). \tag{7}$$

In addition,

$$(\beta_T^+, \beta_T^-) \xrightarrow[T\to\infty]{\text{a.s.}} (b_+, b_-)$$

$$\text{and } \frac{\sqrt{T}}{\sqrt{|b_-| + |b_+|}}(\beta_T^+ - b_+, \beta_T^- - b_-) \xrightarrow[T\to\infty]{\text{law}} \left(\frac{\sigma_+}{\sqrt{|b_-|}}\mathcal{N}^+, \frac{\sigma_-}{\sqrt{|b_+|}}\mathcal{N}^-\right),$$

where $\mathcal{N}^+$ and $\mathcal{N}^-$ are two independent, unit Gaussian random variables.

Proposition 3 (Null recurrent case with vanishing drift **N0**). *Assume $b_+ = b_- = 0$. Assume $\xi_0 = 0$. Then*

$$\left(\frac{Q_T^+}{T}, \frac{Q_T^-}{T}\right) \overset{\text{law}}{=} (\Lambda, 1 - \Lambda) \text{ for all } T > 0,$$

where Λ follows a law of arcsine type with density

$$p_\Lambda(u) := \frac{1}{\pi}\frac{1}{\sqrt{u(1-u)}}\frac{\sigma_+/\sigma_-}{1 - (1 - (\sigma_+/\sigma_-)^2)u} \quad \text{for } 0 < u < 1.$$

Besides,

$$\sqrt{T}(\beta_T^+, \beta_T^-) \overset{\text{law}}{=} (\beta_1^+, \beta_1^-) \tag{8}$$

where the explicit joint density of (β_1^+, β_1^-) is given by (25) *below. In particular, (β_T^+, β_T^-) converges almost surely to $(b_+, b_-) = (0, 0)$.*

Proposition 4 (Null recurrent case with non-vanishing drift **N1**). *Assume $b_+ = 0$, $b_- > 0$. Then*

$$\frac{Q_T^+}{T} \xrightarrow[T\to\infty]{\text{a.s.}} 1 \text{ and } (\beta_T^+, \beta_T^-) \xrightarrow[T\to\infty]{\text{a.s.}} (b_+, b_-).$$

In addition, there exists three independent unit Gaussian random variables $\mathcal{N}^-$, $\mathcal{N}^+$ and $\mathcal{N}$ such that

$$\left(\frac{Q_T^-}{\sqrt{T}}, \sqrt{T}(\beta_T^+ - b_+), T^{1/4}(\beta_T^- - b_-)\right) \xrightarrow[T\to\infty]{\text{law}} \left(\frac{\sigma_+}{b_-}|\mathcal{N}|, \sigma_+\mathcal{N}^+, \sigma_- \frac{\sqrt{b_-}}{\sqrt{\sigma_+}} \cdot \frac{\mathcal{N}^-}{\sqrt{|\mathcal{N}|}}\right). \quad (9)$$

Proposition 5 (Transient case for upward drift **T0**). *Assume $b_+ > 0$, $b_- \geq 0$ so that the process ξ is transient and $\lim_{T\to\infty} \xi_T = +\infty$. Then Q_T^+/T converges almost surely to 1 as $T \to \infty$ and*

$$\beta_T^+ \xrightarrow[T\to\infty]{\text{a.s.}} b_+ \text{ and } \sqrt{T}(\beta_T^+ - b_+) \xrightarrow[T\to\infty]{\text{law}} \sigma_+\mathcal{N}^+ \quad (10)$$

for a unit Gaussian random variable $\mathcal{N}^+$. Let ℓ_0 be the last passage time to 0, which is almost surely finite. Assume $\xi_0 = 0$. We have

$$\beta_T^- \mathbf{1}_{T>\ell_0} = \mathcal{R}_{\mathbf{T0}} \mathbf{1}_{T>\ell_0} \text{ and } \lim_{T\to\infty} \beta_T^- = \mathcal{R}_{\mathbf{T0}} \text{ a.s. with } \mathcal{R}_{\mathbf{T0}} := \frac{L_\infty(\xi)}{2Q_{\ell_0}^-} = \frac{L_\infty(\xi)}{2Q_\infty^-}. \quad (11)$$

The density of $\mathcal{R}_{\mathbf{T0}}$ is given by (22) *below. The case $b_+ \leq 0$, $b_- < 0$ is treated by symmetry.*

Proposition 6 (Transient case for diverging drift **T1**). *Assume $b_+ > 0$, $b_- < 0$ so that the process ξ is transient. Assume that $\xi_0 = 0$. Then there exists a Bernoulli random variable $\mathcal{B} \in \{0, 1\}$ such that*

$$\mathbb{P}(\mathcal{B} = 1) = 1 - \mathbb{P}(\mathcal{B} = 0) = \frac{\sigma_- b_+}{\sigma_+ b_- + \sigma_- b_+},$$
$$\mathbb{P}_+\left(\frac{Q_T^+}{T} \xrightarrow[T\to\infty]{} 1\right) = 1 \text{ and } \mathbb{P}_-\left(\frac{Q_T^-}{T} \xrightarrow[T\to\infty]{} 1\right) = 1$$
$$\text{with } \mathbb{P}_+(\cdot) = \mathbb{P}(\cdot \mid \mathcal{B} = 1) \text{ and } \mathbb{P}_-(\cdot) = \mathbb{P}(\cdot \mid \mathcal{B} = 0).$$

On the event $\{\mathcal{B} = 1\}$ (resp. $\{\mathcal{B} = 0\}$), β_T^+ (resp. β_T^-) converges almost surely to b_+ (resp. b_-) while $\beta_T^- - b_-$ (resp. $\beta_T^+ - b_+$) is the ratio of two a.s. finite random variables.

In addition, for unit Gaussian random variables $\mathcal{N}^+$ and $\mathcal{N}^-$,

$$\sqrt{T}(\beta_T^+ - b_+) \xrightarrow[T\to\infty]{\text{law}} \sigma_+\mathcal{N}^+ \text{ under } \mathbb{P}_+, \quad (12)$$

$$\sqrt{T}(\beta_T^- - b_-) \xrightarrow[T\to\infty]{\text{law}} \sigma_-\mathcal{N}^- \text{ under } \mathbb{P}_-. \quad (13)$$

In addition,

$$\lim_{T\to\infty} \beta_T^- = \mathcal{R}_{\mathbf{T1}}^- \text{ a.s. under } \mathbb{P}_+ \text{ with } \mathcal{R}_{\mathbf{T1}}^- = \frac{L_\infty(\xi)}{2Q_\infty^-}, \quad (14)$$

$$\lim_{T\to\infty} \beta_T^+ = -\mathcal{R}_{\mathbf{T1}}^+ \text{ a.s. under } \mathbb{P}_- \text{ with } \mathcal{R}_{\mathbf{T1}}^+ = \frac{L_\infty(\xi)}{2Q_\infty^+}. \quad (15)$$

The distribution of $\mathcal{R}_{\mathbf{T1}}^-$ is that of of $\mathcal{R}$ given (22) *below. That of $\mathcal{R}_{\mathbf{T1}}^+$ is found by symmetry.*

5 Auxiliary tools

In this section, we give first some results on a martingale central limit theorem that will be used constantly. To deal with the transient or null recurrent cases, we make use of some analytic properties of one-dimensional diffusions.

5.1 Limit theorems on martingales

The following result follows immediately from [32, Proposition 1, p. 148; Theorem 1, p. 150].

Proposition 7 (A criterion for convergence)**.** *Under the true probability* $\mathbb{P}$*,*

(i) *as* $T \to \infty$*,* β_T^+ *(resp.* β_T^-*) converges a.s. to* b_+ *(resp.* b_-*) on the event* $\{Q_\infty^+ = +\infty\}$ *(resp.* $\{Q_\infty^- = +\infty\}$*).*

(ii) *as* $T \to \infty$*,* M_T^+ *(resp.* M_T^-*) converges a.s. to a finite value on the event* $\{Q_\infty^+ < +\infty\}$ *(resp.* $\{Q_\infty^- < +\infty\}$*). In other words,* $\beta_T^\pm$ *is not a consistent estimator on* $\{Q_\infty^\pm < +\infty\}$*.*

We now state an instance of a Central Limit theorem for martingales which follows from [7]. This theorem will be used to deal with the cases **E**, **N1** and **T1**. Let us start by recalling the notion of stable convergence introduced by A. Rényi [21, 40].

Definition 1 (Stable convergence)**.** A sequence $(X_n)_{n\in\mathbb{N}}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is said to *converge stably* with respect to a σ-algebra $\mathcal{G} \subset \mathcal{F}$ if for any bounded, continuous function f and any bounded, $\mathcal{G}$-measurable random variable Y,

$$\mathbb{E}(f(X_n)Y) \xrightarrow[n\to\infty]{} \mathbb{E}(f(X)Y).$$

Proposition 8 (A central limit theorem for martingales)**.** *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be the underlying probability space of the process* ξ *with a filtration* $(\mathcal{F}_t)_{t\geq 0}$*. If for some constants* $c_+, c_- > 0$*,*

$$\frac{Q_T^+}{T} \xrightarrow[T\to+\infty]{\mathbb{P}} c_+ \text{ and } \frac{Q_T^-}{T} \xrightarrow[T\to+\infty]{\mathbb{P}} c_-,$$

then for the martingales $M^\pm$ *defined in Lemma 1,*

$$\left(\frac{M_T^+}{\sqrt{T}}, \frac{M_T^-}{\sqrt{T}}\right) \xrightarrow[T\to\infty]{\mathcal{F}_\infty-\text{stably}} (\sigma_+\sqrt{c_+}\mathcal{N}^+, \sigma_-\sqrt{c_-}\mathcal{N}^-), \tag{16}$$

on a probability space $(\Omega', \mathcal{F}', \mathbb{P}')$ *extending* $(\Omega, \mathcal{F}, \mathbb{P})$ *and containing two independent unit Gaussian random variables* $\mathcal{N}^+, \mathcal{N}^-$*, themselves independent from* ξ*. In addition,*

$$\sqrt{T}\left(\frac{M_T^+}{Q_T^+}, \frac{M_T^-}{Q_T^-}\right) \xrightarrow[T\to\infty]{\mathcal{F}_\infty-\text{stably}} \left(\frac{\sigma_+}{\sqrt{c_+}}\mathcal{N}^+, \frac{\sigma_-}{\sqrt{c_-}}\mathcal{N}^-\right), \tag{17}$$

Proof. Set

$$a_T := \begin{bmatrix} 1/\sqrt{T} & 0 \\ 0 & 1/\sqrt{T} \end{bmatrix} \text{ and } q_T = \langle M, M\rangle_T = \begin{bmatrix} \sigma_+^2 Q_T^+ & 0 \\ 0 & \sigma_-^2 Q_T^- \end{bmatrix}.$$

Thus,

$$a_T q_T a_T' = \begin{bmatrix} \sigma_+^2 \frac{Q_T^+}{T} & 0 \\ 0 & \sigma_-^2 \frac{Q_T^-}{T} \end{bmatrix} \xrightarrow[T\to\infty]{\mathbb{P}} \begin{bmatrix} c_+ \\ c_- \end{bmatrix}. \tag{18}$$

Theorem 2.2 in [7] yields (16). Besides,

$$\sqrt{T}\frac{M_T^\pm}{Q_T^\pm} = \frac{T}{Q_T^\pm} \times \frac{M_T^\pm}{\sqrt{T}}. \tag{19}$$

If a sequence $(X_n)_n$ converges $\mathcal{F}_\infty$-stably and a sequence $(Y_n)_n$ of $\mathcal{F}_\infty$-measurable random variables converges in probability, then $(X_n, Y_n)_n$ converges $\mathcal{F}_\infty$-stably. Using the property in (19) and (17) yields (17). □

5.2 The fundamental system

Along with the characterization through the scale function and the speed measure, much information on the process can be read from the so-called *fundamental system* [11,19,41]: For any $\lambda > 0$, there exists some functions ϕ_λ and ψ_λ such that

- ψ_λ and ϕ_λ are continuous, positive from $\mathbb{R}$ to $\mathbb{R}$ with $\phi_\lambda(0) = \psi_\lambda(0) = 1$.
- ψ_λ is increasing with $\lim_{x\to-\infty} \psi_\lambda(x) = 0$, $\lim_{x\to\infty} \psi_\lambda(x) = +\infty$.
- ϕ_λ is decreasing with $\lim_{x\to-\infty} \phi_\lambda(x) = +\infty$, $\lim_{x\to\infty} \phi_\lambda(x) = 0$.
- ϕ_λ and ψ_λ are solutions to $\mathcal{L}f = \lambda f$.

In the case of piecewise constant coefficients with one discontinuity at 0, these solutions may be computed as linear combinations of the minimal functions for constant coefficients. Using the fact that ϕ, ϕ', ψ and ψ' are continuous at 0,

$$\psi_\lambda(x) = \begin{cases} \exp\left(x\frac{-b_-+\sqrt{b_-^2+2\sigma_-^2\lambda}}{\sigma_-^2}\right) & \text{if } x < 0 \\ \kappa_+ \exp\left(x\frac{-b_++\sqrt{b_+^2+2\sigma_+^2\lambda}}{\sigma_+^2}\right) + \delta_+ \exp\left(x\frac{-b_+-\sqrt{b_+^2+2\sigma_+^2\lambda}}{\sigma_+^2}\right) & \text{if } x \geq 0, \end{cases} \tag{20}$$

$$\phi_\lambda(x) = \begin{cases} \kappa_- \exp\left(x\frac{-b_--\sqrt{b_-^2+2\sigma_-^2\lambda}}{\sigma_-^2}\right) + \delta_- \exp\left(x\frac{-b_-+\sqrt{b_-^2+2\sigma_-^2\lambda}}{\sigma_-^2}\right) & \text{if } x < 0, \\ \exp\left(x\frac{-b_+-\sqrt{b_+^2+2\sigma_+^2\lambda}}{\sigma_+^2}\right) & \text{if } x \geq 0 \end{cases} \tag{21}$$

with

$$\begin{aligned}
\kappa_+ &:= \frac{-b_-\sigma_+^2 + b_+\sigma_-^2 + \sigma_-^2\sqrt{b_+^2 + 2\lambda\sigma_+^2} + \sigma_+^2\sqrt{b_-^2 + 2\lambda\sigma_-^2}}{2\sigma_-^2\sqrt{b_+^2 + 2\lambda\sigma_+^2}},\\
\delta_+ &:= \frac{b_-\sigma_+^2 - b_+\sigma_-^2 + \sigma_-^2\sqrt{b_+^2 + 2\lambda\sigma_+^2} - \sigma_+^2\sqrt{b_-^2 + 2\lambda\sigma_-^2}}{2\sigma_-^2\sqrt{b_+^2 + 2\lambda\sigma_+^2}},\\
\kappa_- &:= \frac{-b_-\sigma_+^2 + b_+\sigma_-^2 - \sigma_-^2\sqrt{b_+^2 + 2\lambda\sigma_+^2} + \sigma_+^2\sqrt{b_-^2 + 2\lambda\sigma_-^2}}{2\sigma_+^2\sqrt{b_-^2 + 2\lambda\sigma_-^2}},\\
\delta_- &:= \frac{b_-\sigma_+^2 - b_+\sigma_-^2 + \sigma_-^2\sqrt{b_+^2 + 2\lambda\sigma_+^2} + \sigma_+^2\sqrt{b_-^2 + 2\lambda\sigma_-^2}}{2\sigma_+^2\sqrt{b_-^2 + 2\lambda\sigma_-^2}}.
\end{aligned}$$

We also define the quantities [37]

$$\widehat{\psi}(\lambda) := \frac{1}{2}\frac{\psi_\lambda'(0)}{\psi_\lambda(0)} = \frac{-b_- + \sqrt{b_-^2 + 2\sigma_-^2\lambda}}{2\sigma_-^2} \geq 0$$

$$\text{and } \widehat{\phi}(\lambda) := -\frac{1}{2}\frac{\phi_\lambda'(0)}{\phi_\lambda(0)} = \frac{b_+ + \sqrt{b_+^2 + 2\sigma_+^2\lambda}}{2\sigma_+^2} \geq 0.$$

In particular,

$$\widehat{\psi}(0) = 0 \text{ and } \widehat{\phi}(0) = \frac{b_+}{\sigma_+^2} \text{ when } b_- \geq 0 \text{ and } b_+ \geq 0.$$

5.3 Last passage time and occupation time for the transient process

When ξ is a transient process, the last passage time $\ell_0 = \sup\{t \geq 0 \mid \xi_t = 0\}$ of ξ at 0 is almost surely finite. Its Laplace transform is (See (53) in [37]):

$$\mathbb{E}_0[\exp(-\lambda\ell_0)] = \frac{\widehat{\psi}(0) + \widehat{\phi}(0)}{\widehat{\psi}(\lambda) + \widehat{\phi}(\lambda)}.$$

Let us now assume that $b_+ > 0$ and $b_- \geq 0$. This is the transient case **T0** where the process ends up almost surely in the positive semi-axis. Thus, $Q^-_{\ell_0} = Q^-_\infty$ and $L_{\ell_0}(\xi) = L_\infty(\xi)$.

Let us write $\widehat{b}_\pm := b_\pm/\sigma_\pm^2$. From Corollary 5 in [37],

$$\mathbb{E}_0[\exp(-\alpha L_\infty(\xi) - \lambda Q^-_\infty)] = \frac{\widehat{\psi}(0) + \widehat{\phi}(0)}{\alpha + \widehat{\phi}(0) + \widehat{\psi}(\lambda)} = \frac{\widehat{b}_+}{\alpha + \widehat{b}_+ - \frac{\widehat{b}_-}{2} + \frac{1}{2\sigma_-^2}\sqrt{b_-^2 + 2\sigma_-^2\lambda}}.$$

Let $p_{L_\infty(\xi)}(t)$ be the density of $L_\infty(\xi)$. Setting $\lambda = 0$, we see that $L_\infty(\xi)$ is distributed according to an exponential distribution of rate $\widehat{b}_+$. Thus, $p_{L_\infty(\xi)}(t) = \widehat{b}_+ \exp(-\widehat{b}_+ t)$. A conditioning shows that

$$\mathbb{E}_0[\exp(-\alpha L_\infty(\xi) - \lambda Q^-_\infty)] = \int_0^{+\infty} \exp(-\alpha t)\mathbb{E}_0\big(\exp(-\lambda Q^-_\infty) \bigm| L_\infty(\xi) = t\big) p_{L_\infty}(t)\,\mathrm{d}t.$$

By inverting the Laplace transform with respect to α, since $p_{L_\infty}(t) = \widehat{b}_+ \exp(-\widehat{b}_+ t)$,

$$\mathbb{E}_0\big(\exp(-\lambda Q^-_\infty) \bigm| L_\infty(\xi) = t\big) = \exp\left(-t\left(-\frac{\widehat{b}_-}{2} + \frac{1}{\sqrt{2}\sigma_-}\sqrt{\frac{b_-^2}{2\sigma_-^2} + \lambda}\right)\right).$$

Inverting the latter Laplace transform with respect to λ, the density $p_{Q^-_\infty}(s|t)$ of Q^-_∞ given $\{L_\infty(\xi) = t\}$ is

$$p_{Q^-_\infty}(s|t) = \frac{t}{\sigma_- 2\sqrt{2\pi}s^{3/2}} \exp\left(\frac{tb_-}{2\sigma_-^2} - \frac{b_-^2 s}{2\sigma_-^2} - \frac{t^2}{8\sigma_-^2 s}\right).$$

Hence, the distribution $p_{(Q^-_\infty, L_\infty(\xi))}(s,t)$ of the pair $(Q^-_\infty, L_\infty(\xi))$ is

$$p_{(Q^-_\infty, L_\infty(\xi))}(s,t) = \frac{tb_+}{\sigma_+^2\sigma_- 2\sqrt{2\pi}s^{3/2}} \exp\left(\left(\frac{b_-}{2\sigma_-^2} - \frac{b_+}{\sigma_+^2}\right)t - \frac{b_-^2 s}{2\sigma_-^2} - \frac{t^2}{8\sigma_-^2 s}\right).$$

The density $p_{\mathcal{R}_{\mathbf{T0}}}(r)$ of the random variable $\mathcal{R}_{\mathbf{T0}} := L_\infty(\xi)/2Q^-_\infty$ is then

$$p_{\mathcal{R}_{\mathbf{T0}}}(r) = 2\int_0^{+\infty} s \cdot p_{(Q^-_\infty, L_\infty(\xi))}(s, 2rs)\;\mathrm{d}s$$
$$= \frac{rb_+}{\sigma_+^2\sigma_-\sqrt{2}}\left(\frac{2rb_+}{\sigma_+^2} + \frac{(r-b_-)^2}{2\sigma_-^2}\right)^{-3/2},\ r > 0.$$

The distribution of $L_\infty(\xi)/2Q^+_\infty$ when $b_- < 0, b_+ \leq 0$ is found by symmetric arguments.

Let us now assume that $b_+ > 0$ and $b_- < 0$. This is the transient case **T1** where the process can end up in both semi-axis. The Laplace transform is

$$\mathbb{E}_0[\exp(-\alpha L_\infty(\xi) - \lambda Q^-_\infty)] = \frac{\widehat{\psi}(0) + \widehat{\phi}(0)}{\alpha + \widehat{\phi}(0) + \widehat{\psi}(\lambda)} = \frac{\widehat{b}_+ - \widehat{b}_-}{\alpha + \widehat{b}_+ - \frac{\widehat{b}_-}{2} + \frac{1}{2\sigma_-^2}\sqrt{b_-^2 + 2\sigma_-^2\lambda}}.$$

With analogous computations as before we get that the density $p_{\mathcal{R}^-_{\mathbf{T1}}}(r)$ of $\mathcal{R}^-_{\mathbf{T1}}$ is

$$p_{\mathcal{R}^-_{\mathbf{T1}}}(r) = \frac{r}{\sigma_-\sqrt{2}}\left(\frac{b_+}{\sigma_+^2} - \frac{b_-}{\sigma_-^2}\right)\left(\frac{2rb_+}{\sigma_+^2} + \frac{(r-b_-)^2}{2\sigma_-^2}\right)^{-3/2}, \ r > 0.$$

Considering also the previous case, we can write the following formula, holding for the density of $\mathcal{R} = \mathcal{R}_{\mathbf{T0}}$ or $\mathcal{R} = \mathcal{R}^-_{\mathbf{T1}}$ in both cases **T0** and **T1**:

$$p_{\mathcal{R}}(r) = \frac{r}{\sigma_-\sqrt{2}}\left(\frac{b_+}{\sigma_+^2} + \frac{(\!(b_-)\!)^-}{\sigma_-^2}\right)\left(\frac{2rb_+}{\sigma_+^2} + \frac{(r-b_-)^2}{2\sigma_-^2}\right)^{-3/2}, \ r > 0. \tag{22}$$

Notice that this is the density of a positive random variable which is not integrable. This gives the limit behavior of the estimator β^-_T of b_-. The behavior of β^+_T in the corresponding cases can be found by symmetric arguments.

6 Proofs of the asymptotic behavior of the estimators

6.1 Asymptotic behavior for the ergodic case (E)

The ergodic case is the most favorable one. The process ξ is ergodic, so that for any bounded, measurable function f, $\frac{1}{T}\int_0^T f(\xi_s)\,\mathrm{d}s$ converges almost surely to $\int f(x)\frac{m(x)}{M(\mathbb{R})}\,\mathrm{d}x$.

With the explicit expression of M that follows from (6),

$$M(\mathbb{R}_+) = \frac{-1}{b_+},\ M(\mathbb{R}_-) = \frac{1}{b_-} \text{ and } M(\mathbb{R}) = \frac{|b_+b_-|}{|b_-| + |b_+|}.$$

From the ergodic theorem, since $Q^\pm_t = \int_0^t \mathbf{1}_{\pm\xi_s \geq 0}\,\mathrm{d}s$,

$$\frac{Q^\pm_T}{T} \xrightarrow[T\to\infty]{\text{a.s.}} \frac{M(\mathbb{R}_\pm)}{M(\mathbb{R})}$$

so that

$$\frac{Q^+_T}{T} \xrightarrow[T\to\infty]{\text{a.s.}} \frac{|b_-|}{|b_-| + |b_+|} \text{ and } \frac{Q^-_T}{T} \xrightarrow[T\to\infty]{\text{a.s.}} \frac{|b_+|}{|b_-| + |b_+|}. \tag{23}$$

The rate of convergence follows from Proposition 8 and (23).

6.2 Asymptotic behavior for the null recurrent case with vanishing drift (N0)

When $b_- = b_+ = 0$, the process ξ is an Oscillating Brownian motion (OBM, introduced first in [24], see also [30]). Supposing $\xi_0 = 0$, using the scaling relation [30, Remark 3.7], for any $T > 0$,

$$\left(\frac{(\!|\xi_T|\!)^+}{\sqrt{T}}, \frac{(\!|\xi_T|\!)^-}{\sqrt{T}}, \frac{L_T(\xi)}{\sqrt{T}}, \frac{Q_T^+}{T}\right) \stackrel{\text{law}}{=} ((\!|\xi_1|\!)^+, (\!|\xi_1|\!)^-, L_1(\xi), Q_1^+).$$

Therefore,

$$\sqrt{T}\begin{pmatrix}\beta_T^+\\ \beta_T^-\end{pmatrix} = \begin{pmatrix}\dfrac{(\!|\xi_T|\!)^+/\sqrt{T} - \frac{1}{2}L_T(\xi)/\sqrt{T}}{Q_T^+/T}\\ \dfrac{\frac{1}{2}L_T(\xi)/\sqrt{T} - (\!|\xi_T|\!)^-/\sqrt{T}}{Q_T^-/T}\end{pmatrix} \stackrel{\text{law}}{=} \begin{pmatrix}\beta_1^+\\ \beta_1^-\end{pmatrix} = \begin{pmatrix}\dfrac{(\!|\xi_1|\!)^+ - \frac{1}{2}L_1(\xi)}{Q_1^+}\\ \dfrac{\frac{1}{2}L_1(\xi) - (\!|\xi_1|\!)^-}{Q_1^-}\end{pmatrix}.$$

We recall now that $X = \Phi(\xi) := \xi/\sigma(\xi)$ is a Skew Brownian motion [10, 30]. An explicit formula for the density for the position a Skew Brownian motion, its local time and its occupation time is known [2, 12]. Since the transform Φ is piecewise linear, one easily recover the one of an OBM, its local and occupation times. Hence, the density of $(\xi_1, L_1(\xi), Q_1^+(\xi))$ is

$$\begin{aligned} &p_{(\xi_1, L_1(\xi), Q_1^+)}(\rho, \lambda, \tau)\\ &\quad= \begin{cases} \dfrac{(\lambda/2+\rho)\lambda/2}{2\pi\sigma_-\sigma_+^3(1-\tau)^{3/2}\tau^{3/2}} \exp\left(-\dfrac{(\lambda/2)^2}{2\sigma_-^2(1-\tau)} - \dfrac{(\lambda/2+\rho)^2}{2\sigma_+^2\tau}\right) & \text{for } \rho \geq 0,\\ \dfrac{(\lambda/2-\rho)\lambda/2}{2\pi\sigma_+\sigma_-^3(1-\tau)^{3/2}\tau^{3/2}} \exp\left(-\dfrac{(\lambda/2)^2}{2\sigma_+^2\tau} - \dfrac{(\lambda/2-\rho)^2}{2\sigma_-^2(1-\tau)}\right) & \text{for } \rho < 0. \end{cases} \end{aligned} \tag{24}$$

The change of variable in the density suggested by

$$(\beta_1^+, \beta_1^-, Q_1^+) = \left(\frac{(\!|\xi_1|\!)^+ - L_1(\xi)/2}{Q_1^+}, \frac{L_1(\xi)/2 - (\!|\xi_1|\!)^-}{1 - Q_1^+}, Q_1^+\right)$$

gives

$$\begin{aligned} &p_{(\beta_1^+, \beta_1^-, Q_1^+)}(a, b, \delta)\\ &\qquad= 2\delta(1-\delta)p_{(\xi_1, L_1(\xi), Q_1^+)}(a\delta + b(1-\delta), |a\delta + b(1-\delta)| - a\delta + b(1-\delta), \delta) \end{aligned}$$

and then, since $Q_1^+ \in [0, 1]$,

$$\begin{aligned} &p_{(\beta_1^+, \beta_1^-)}(a, b)\\ &\quad= \int_0^1 2\delta(1-\delta)p_{(\xi_1, L_1(\xi), Q_1^+)}(a\delta + b(1-\delta), |a\delta + b(1-\delta)| - a\delta + b(1-\delta), \delta)\,\mathrm{d}\delta. \end{aligned} \tag{25}$$

6.3 Asymptotic behavior for the null recurrent case with non-vanishing drift (N1)

We consider $b_+ = 0$, $b_- > 0$. The particle is then pushed upward when its position is negative. Yet the process is only null recurrent. The measure M satisfies

$$M(\mathbb{R}_-) = \frac{1}{b_-} \text{ and } M(\mathbb{R}_+) = +\infty. \tag{26}$$

Using 9) in [19, Section 6.8, p. 228] or [34, 47], with (26),

$$\frac{Q_T^-}{T} \xrightarrow[T\to\infty]{\text{a.s.}} \frac{M(\mathbb{R}_-)}{M(\mathbb{R})} = 0.$$

Since $Q_T^+ + Q_T^- = T$, it holds that Q_T^+/T converges almost surely to 1.

Using Proposition 8 on M^+ only, we obtain that

$$\frac{M_T^+}{\sqrt{T}} \xrightarrow[T\to\infty]{\text{law}} \sigma^+\mathcal{N}^+ \tag{27}$$

for a Gaussian random variable $\mathcal{N}^+ \sim \mathcal{N}(0,1)$ and then that $\sqrt{T}(\beta_T^+ - b_+) \xrightarrow[T\to\infty]{\text{law}} \sigma^+\mathcal{N}^+$.

The asymptotic behavior of Q_T^- is more delicate to deal with as the process ξ is only null recurrent. For this, we use the results of [14] which extends the one of D.A. Darling and M. Kac [8] on additive and martingale additive functionals.

The Green kernel with respect to the invariant measure M of $\mathcal{L}$ is given by [11, 19, 41]

$$g_\lambda(x,y) := \frac{1}{W_\lambda}\begin{cases}\psi_\lambda(x)\phi_\lambda(y) & \text{if } x < y,\\ \phi_\lambda(x)\psi_\lambda(y) & \text{if } x \geq y\end{cases} \text{ with } W_\lambda = \frac{\psi_\lambda'(0)\phi_\lambda(0) - \psi_\lambda(0)\phi_\lambda'(0)}{S'(0)}.$$

The Wronskian W_λ is then equal to

$$W_\lambda = \frac{\sqrt{2\lambda}}{\sigma_+} + \frac{\sqrt{b_-^2 + 2\lambda\sigma_-^2}}{\sigma_-^2} - \frac{b_-}{\sigma_-^2}.$$

In particular, $W_\lambda/\sqrt{\lambda}$ converges to $\sqrt{2}/\sigma_+$ as λ converges to 0.

On the other hand, it follows from (20) and (21) that

$$\psi_\lambda(x) \xrightarrow[\lambda\to 0]{} \psi_0(x) := 1 \text{ and } \phi_\lambda(x) \xrightarrow[\lambda\to 0]{} \phi_0(x) := 1 \text{ when } x > 0$$

while

$$\psi_\lambda(x) = \exp(x\sqrt{2\lambda}) \xrightarrow[\lambda\to 0]{} \psi_0(x) := 1 \text{ and } \phi_\lambda(x) = \phi_0(x) := 1 \text{ when } x \geq 0.$$

For a measurable function $f : \mathbb{R} \to \mathbb{R}_+$ such that $\int_{\mathbb{R}} f\,\mathrm{d}M < +\infty$, the above convergence results imply that

$$\sqrt{\lambda}\int_{\mathbb{R}} g_\lambda(x,y)f(x)m(y)\,\mathrm{d}y \xrightarrow[\lambda\to 0]{} \frac{\sigma_+}{\sqrt{2}}\int_{\mathbb{R}} f(y)m(y)\,\mathrm{d}y,\ \forall x \in \mathbb{R}.$$

We then define $\ell(\lambda) := \sqrt{2}/\sigma_+$ which is a constant function, and $\alpha := 1/2$, the exponent of λ.

Let $(\mathcal{M}_t)_{t\geq 0}$ be a Mittag-Leffler process of index $1/2$ (it is the inverse of an increasing stable process of index $1/2$). The process $2^{-1/2}\mathcal{M}$ is equal in distribution to the running maximum of a Brownian motion, or equivalently, to the local time of a Brownian motion [14, Remark 2.9, p. 21].

From Theorem 3.1 and Corollary 3.2 in [14, p. 26], since $\langle M^-\rangle_t = \sigma_-^2 Q_t^-$, $t \geq 0$ and M is continuous,

$$\left(\sqrt{\ell(n)}\frac{M_{nt}^-}{n^{1/4}}, \ell(n)\frac{Q_{nt}^-}{\sqrt{n}}\right)_{t\in[0,1]} \xrightarrow[n\to\infty]{\text{law}} \left(\sigma_-\sqrt{\nu}B^-(\mathcal{M}_t), \nu\mathcal{M}_t\right)_{t\in[0,1]} \tag{28}$$

with respect to the uniform topology, where B is a Brownian motion independent from $\mathcal{M}$ and

$$\nu = \int_{\mathbb{R}} m(y)\mathbb{E}_y\left(\int_0^1 \mathbf{1}_{\xi_s\leq 0}\,\mathrm{d}s\right)\mathrm{d}y = M(\mathbb{R}_-) = \frac{1}{b_-}.$$

From the reflection principle, the distribution of $\mathcal{M}_1$ is the same as the one of a truncated normal distribution $\sqrt{2}\mathcal{T}$ where $\mathcal{T} := |\mathcal{G}|$ with $\mathcal{G} \sim \mathcal{N}(0,1)$. Setting $t = 1$ in (28) and using the scaling property of the Brownian motion B^-,

$$\left(\frac{M_T^-}{T^{1/4}}, \frac{Q_T^-}{T^{1/2}}\right) \xrightarrow[T\to\infty]{\text{law}} \left(\sigma_-\frac{\sqrt{\sigma_+}}{\sqrt{b_-}}\sqrt{\mathcal{T}}\cdot\mathcal{N}^-, \frac{\sigma_+}{b_-}\mathcal{T}\right)$$

for a Gaussian random variable $\mathcal{N}^- \sim \mathcal{N}(0,1)$ independent from $\mathcal{T}$.

It remains to show the independence of $\mathcal{N}^+, \mathcal{N}^-$ and $\mathcal{T}$.

Since $\langle M^\pm\rangle_t = \sigma_\pm^2 Q_t^\pm$ and $\langle M^+, M^-\rangle_t = 0$ for any $t \geq 0$, the Knight theorem [18, Theorem 7.3', p. 92] implies that there exists on an extension of $(\Omega, \mathcal{F}, \mathbb{P})$ a 2-dimensional Brownian motion (B^+, B^-) such that $M_t^\pm = B^\pm(\sigma_\pm^2 Q_t^\pm)$ for any $t \geq 0$. Let us set $B_n^+(t) = n^{-1/2}B^+(nt)$ and $B_n^-(t) = n^{-1/4}B^-(\sqrt{n}t)$ for any integer n and any $t \geq 0$. From the scaling property, (B_n^+, B_n^-) is still a 2-dimensional Brownian motion which converges in distribution to a 2-dimensional Brownian motion (B_∞^+, B_∞^-) in the space $\mathcal{C}([0,1],\mathbb{R}^2)$ of continuous functions.

For any $0 \leq s \leq t \leq 1$, $Q_{nt}^+ - Q_{ns}^+ \leq n(t-s)$ so that $(Q_{nt}^+/n)_{t\in[0,1]}$ is tight in the space of continuous functions. Hence, $(Q_{nt}^+/n)_{t\in[0,1]}$ converges in probability to the identity map $t \mapsto t$ in the space of continuous function $\mathcal{C}([0,1],\mathbb{R})$.

Combining this result with (28), it holds that $\left(B_n^+(t), B_n^-(t), n^{-1}Q_{nt}^+, n^{-1/2}Q_{nt}^-\right)_{t\in[0,1]}$ is tight in $\mathcal{C}([0,1],\mathbb{R}^4)$ and then necessarily

$$\left(B_n^+(t), B_n^-(t), \frac{Q_{nt}^+}{n}, \frac{Q_{nt}^-}{\sqrt{n}}\right)_{t\in[0,1]} \xrightarrow[n\to\infty]{\text{law}} (B_\infty^+(t), B_\infty^-(t), t, \nu\mathcal{M}_t)_{t\in[0,1]}$$

in the space of continuous functions $\mathcal{C}([0,1],\mathbb{R}^4)$. Being the inverse of a $1/2$-stable process, hence a pure jump process, $\mathcal{M}$ is independent from (B_∞^+, B_∞^-) for the arguments presented in [14, p. 38] or [18, Theorem 6.3, p. 77].

For any $t \in [0,1]$, it holds that

$$\frac{M_{nt}^+}{\sqrt{n}} = \frac{1}{\sqrt{n}}B^+(\sigma_+^2 Q_{nt}^+) = B_n^+\left(\sigma_+^2\frac{Q_{nt}^+}{n}\right) \text{ and } \frac{M_{nt}^-}{n^{1/4}} = \frac{1}{\sqrt{n}}B^-(\sigma_-^2 Q_{nt}^-) = B_n^+\left(\sigma_-^2\frac{Q_{nt}^+}{\sqrt{n}}\right).$$

Using $(n^{-1}Q_{nt}^+)_{t\in[0,1]}$ and $(n^{-1/2}Q_{nt}^-)_{t\in[0,1]}$ as random time changes, we deduce from the results in [5, p. 144] that

$$\left(\frac{M_{nt}^+}{n^{1/2}}, \frac{M_{nt}^-}{n^{1/4}}, \frac{Q_{nt}^+}{n}, \frac{Q_{nt}^-}{\sqrt{n}}\right)_{t\in[0,1]} \xrightarrow[n\to\infty]{\text{law}} (B^+(\sigma_+^2 t), B^-(\sigma_-^2\nu\mathcal{M}_t), t, \nu\mathcal{M}_t)_{t\in[0,1]}.$$

By setting $\mathcal{N}^+ := B^+(\sigma_+^2)/\sigma_+$ in (27), $\mathcal{T} := \mathcal{M}_1/\sqrt{2}$ and $\mathcal{N}^- := B^-(\sigma_-^2\nu)/\sigma_-\sqrt{\nu}$, this proves (9) using $t = 1$ in the above limit.

6.4 Asymptotic behavior for the transient case (T0)

We recall that we consider only $b_+ > 0$ and $b_- \geq 0$.

It is known that the last passage time ℓ_0 to 0 is almost surely finite so that $Q^-_\infty < \infty$ almost surely. Since $Q^+_T = T - Q^-_T$, we obtain that Q^+_T/T converges almost surely to 1. The convergence results regarding β^+_T follows from Proposition 8 applied only to one component.

The asymptotic behavior of β^+_T follows from Proposition 7(ii).

When $T > \ell_0$ and $\xi_0 = 0$, then $\xi_T > 0$ and thus $\beta^-_T = L_T(\xi)/2Q^-_T$. Yet the local time $L_T(\xi)$ and the occupation times Q^-_T are constant when $T > \ell_0$. The result follows by the computations of Section 5.3.

6.5 Asymptotic behavior for the transient case generated by diverging drift (T1)

When $b_- < 0$ and $b_+ > 0$, the process is also transient as $S(+\infty) < +\infty$ and $S(-\infty) > -\infty$. The scale function S map $\mathbb{R}$ to $(-\gamma_-, \gamma_+)$ with $\gamma_\pm = |\sigma^2_\pm|2b_\pm$. From the Feller test [23, Theorem 5.29], the process does not explode. Thus, as $\xi_0 = 0$, it follows from [48] or [23, Proposition 5.5.22, p. 354] that

$$p := \mathbb{P}\Big(\lim_{T\to\infty} \xi_T = +\infty\Big) = 1 - \mathbb{P}\Big(\lim_{T\to\infty} \xi_T = -\infty\Big) = \frac{\gamma_-}{\gamma_- + \gamma_+} = \frac{\sigma^2_- b_+}{\sigma^2_+ b_- + \sigma^2_- b_+}.$$

Then event that $\{\lim_{T\to\infty}|\xi_T| = +\infty\}$ arise when the process starts an excursion with infinite lifetime, thus after the last passage time to 0. We denote by $A^\pm$ the event $\{\lim_{T\to\infty} Q^\pm_T/T = 1\}$, so that $A^+ \cap A^- = \emptyset$. Hence, $\mathbb{P}(A^+) = 1 - \mathbb{P}(A^-) = p$.

Using the same arguments as in the case **T0**, given $A^\pm$, $\beta^\pm_T$ converges almost surely to $b_\pm$ while $\beta^\mp_T = \pm L_\infty(\xi)/Q^\mp_{\ell_0}$.

For the Central Limit Theorem, we apply Corollary 2.3 in [7] on $M^+_T/\sqrt{T}$. As $Q^+_T > 0$ a.s. as soon as $T > 0$ since $\xi_0 = 0$ and

$$\{\mathcal{B} = p\} = A^+, \text{ a.s. for } p = 0, 1,$$

it follows that for a normal distribution $\mathcal{N}$,

$$\sqrt{\frac{T}{\sigma^2_+ Q^+_T}} \cdot \frac{M_T}{\sqrt{T}} \xrightarrow[N\to\infty]{\mathcal{F}_\infty-\text{stably}} \mathcal{N} \text{ under } \mathbb{P}(\cdot \mid \mathcal{B} = 1).$$

It follows that

$$\beta^+_T - b_+ = \sqrt{T}\frac{M_T}{Q^+_T} = \sigma_+\sqrt{\frac{T}{Q^+_T}}\sqrt{\frac{T}{\sigma^2_+ Q^+_T}}\frac{M_T}{\sqrt{T}} \xrightarrow[N\to\infty]{\mathcal{F}_\infty-\text{stably}} \sigma_+\mathcal{N} \text{ under } \mathbb{P}(\cdot \mid \mathcal{B} = 1).$$

Hence the result.

7 Wilk's theorem and LAN property

Owing to the quadratic nature of the log-likelihood, we easily deduce both a Wilk theorem, on which a hypothesis test may be developed, as well as the Local Asymptotic Normality (LAN) property, which proves that our estimators are asymptotically efficient.

7.1 Wilk's theorem and a hypothesis test

The log-likelihood $\log G(b_+, b_-)$ can be computed from the data using (5). Moreover, this function is quadratic in b_+ and b_-. We have then

$$D(b_+, b_-) := \nabla \log G(b_+, b_-) = \begin{bmatrix} \frac{(\!|\xi_T|\!)^+}{\sigma_+^2} - \frac{(\!|\xi_0|\!)^+}{\sigma_+^2} - \frac{1}{2\sigma_+^2} L_T(\xi) - \frac{b_+}{\sigma_+^2} Q_T^+ \\ -\frac{(\!|\xi_T|\!)^-}{\sigma_-^2} + \frac{(\!|\xi_0|\!)^-}{\sigma_-^2} + \frac{1}{2\sigma_-^2} L_T(\xi) - \frac{b_-}{\sigma_-^2} Q_T^- \end{bmatrix}$$

$$\text{and } H(b_+, b_-) := \operatorname{Hess} \log G(b_+, b_-) = \begin{bmatrix} -\frac{Q_T^+}{\sigma_+^2} & 0 \\ 0 & -\frac{Q_T^-}{\sigma_-^2} \end{bmatrix}.$$

Therefore, around any point (b_+, b_-),

$$\log G(b_+ + \Delta b_+, b_- + \Delta b_-) = \log G(b_+, b_-) + D(b_+, b_-) \cdot \begin{bmatrix} \Delta b_+ \\ \Delta b_- \end{bmatrix} - \frac{Q_T^+}{2\sigma_+^2} \Delta b_+^2 - \frac{Q_T^-}{2\sigma_-^2} \Delta b_-^2. \quad (29)$$

In particular, we prove a result in the asymptotic behavior of the log-likelihood for the ergodic case **E** or the null recurrent case **N1**. A similar result can be given for the null recurrent cases **N0** with a different limit distribution that can be identified with the density given in Section 6.2.

Proposition 9 (Wilk's theorem; Ergodic case **E** or null recurrent case **N1**). *Denote by $(b_+^{\mathrm{true}}, b_-^{\mathrm{true}})$ be the real parameters. Then*

$$-2 \log G(b_+^{\mathrm{true}}, b_-^{\mathrm{true}}) \xrightarrow[T\to\infty]{\text{law}} \chi_2 := (\mathcal{N}^+)^2 + (\mathcal{N}^-)^2 \text{ under } \mathbb{P}_{(b_+^{\mathrm{true}}, b_-^{\mathrm{true}})} \quad (30)$$

for two independent, unit Gaussian random variables $\mathcal{N}^+$ and $\mathcal{N}^-$. Besides, when $(b_+, b_-) \neq (b_+^{\mathrm{true}}, b_-^{\mathrm{true}})$, then

$$-\log G(b_+, b_-) \xrightarrow[T\to\infty]{\text{a.s.}} +\infty \text{ under } \mathbb{P}_{(b_+^{\mathrm{true}}, b_-^{\mathrm{true}})}. \quad (31)$$

Proof. Considering (29) at $(b_+, b_-) = (\beta_T^+, \beta_T^-)$ since $D(\beta_T^+, \beta_T^-) = (0, 0)$, for any parameter (b_+, b_-) and any $\alpha_+, \alpha_- > 0$,

$$\log G(b_+, b_-) = -\frac{Q_T^+}{2T^{\alpha_+}\sigma_+^2} \left(T^{\frac{\alpha_+}{2}}(b_+ - \beta_T^+)\right)^2 - \frac{Q_T^-}{2T^{\alpha_-}\sigma_-^2} \left(T^{\frac{\alpha_-}{2}}(b_- - \beta_T^-)\right)^2.$$

When the process is ergodic (Case **E**), we set $(\alpha_+, \alpha_-) = (1, 1)$. It follows from Proposition 2 that (30) holds. When $(b_+, b_-) \neq (b_+^{\mathrm{true}}, b_-^{\mathrm{true}})$, then $\beta_T^\pm - b_\pm$ does not converge to 0 while Q_T^+ converges a.s. to infinity. This proves (31). The result is similar in the case **N1** with $(\alpha_+, \alpha_-) = (1, 1/2)$. □

A *hypothesis test* can be developed from Proposition 9. The null hypothesis is $(b_+, b_-) = (b_+^0, b_-^0)$ for a given drift (b_+^0, b_-^0) while the alternative hypothesis is $(b_+, b_-) \neq (b_+^0, b_-^0)$.

Using (5), we compute $\log G(b_+^0, b_-^0)$. The null hypothesis is rejected with a confidence level α if $-\log G(b_+^0, b_-^0) > q_\alpha$ where q_α is the α-quantile $\mathbb{P}[\chi_2 \leq q_\alpha] = \alpha$ while χ_2 follows a χ^2 distribution with 2 degrees of freedom.

7.2 The LAN property

The LAN (local asymptotic normal) property, introduced by L. Le Cam in [29], characterizes the efficiency of the estimator (See also [16,27] among many other references). It was extended as the LAMN (local asymptotic mixed normal) to deal with a mixed normal limits by P. Jeganathan in [22].

The quadratic nature of the log-likelihood as well as our limit theorems implies that the LAN (resp. LAMN) property is verified in the ergodic case **E** (resp. null recurrent case **N1**).

Proposition 10 (LAN property; Ergodic case **E**). *In the ergodic case **E**, the LAN property holds for the likelihood at $(b_+^{\mathrm{true}}, b_-^{\mathrm{true}})$ with rate of convergence $(\sigma_+^2/\sqrt{T}, \sigma_-^2/\sqrt{T})$ and asymptotic Fisher information*

$$\Gamma := \frac{1}{|b_-| + |b_+|} \begin{bmatrix} \sigma_+^2 |b-| & 0 \\ 0 & \sigma_-^2 |b+| \end{bmatrix}.$$

Proof. At $(b_+^{\mathrm{true}}, b_-^{\mathrm{true}})$, the gradient may be written

$$D(b_+^{\mathrm{true}}, b_-^{\mathrm{true}}) = \begin{bmatrix} \frac{Q_T^+}{\sqrt{T}\sigma_+^2} \sqrt{T}(\beta_T^+ - b_+^{\mathrm{true}}) \\ \frac{Q_T^-}{\sqrt{T}\sigma_-^2} \sqrt{T}(\beta_T^- - b_-^{\mathrm{true}}) \end{bmatrix}.$$

Using (29),

$$\begin{aligned} R(T) &:= \log \frac{G\left(b_+^{\mathrm{true}} + \sigma_+^2 \frac{\Delta b_+}{\sqrt{T}}, b_-^{\mathrm{true}} + \sigma_-^2 \frac{\Delta b_-}{\sqrt{T}}\right)}{G(b_+^{\mathrm{true}}, b_-^{\mathrm{true}})} \\ &= \begin{bmatrix} Q_T^+(\beta_T^+ - b_+^{\mathrm{true}}) \\ Q_T^-(\beta_T^- - b_-^{\mathrm{true}}). \end{bmatrix} \cdot \begin{bmatrix} \Delta b_+ \\ \Delta b_- \end{bmatrix} - \frac{\sigma_+^2 Q_T^+}{2T} \Delta b_+^2 - \frac{\sigma_-^2 Q_T^-}{2T} \Delta b_-^2 \\ &= \frac{1}{\sqrt{T}} \begin{bmatrix} M_T^+ \\ M_T^- \end{bmatrix} \cdot \begin{bmatrix} \Delta b_+ \\ \Delta b_- \end{bmatrix} - \frac{1}{2T} \begin{bmatrix} \Delta b_+ \\ \Delta b_- \end{bmatrix} \cdot \langle M^+, M^- \rangle_T \begin{bmatrix} \Delta b_+ \\ \Delta b_- \end{bmatrix} \end{aligned} \quad (32)$$

With $c = |b_+| + |b_-|$, Proposition 2 implies that $(T^{-1/2} M^+, ^{-1/2} M^-)$ converges in distribution to $\mathcal{G} \sim \mathcal{N}(0, \Gamma)$ and $T^{-1}\langle M^+, M^- \rangle_T$ converges to the diagonal, definite positive matrix Γ. This proves the LAN property. □

Proposition 11 (LAMN property; Null recurrent case **N1**). *In the null recurrent case **N1**, the LAMN property holds for the likelihood at $(b_+^{\mathrm{true}}, b_-^{\mathrm{true}})$ with rate of convergence $(\sigma_+^2/T^{1/2}, \sigma_-^2/T^{1/4})$ and asymptotic (random) Fisher information*

$$\Gamma := \begin{bmatrix} \sigma_+^2 & 0 \\ 0 & \sigma_-^2 \frac{\sigma_+}{b_-} |\mathcal{N}| \end{bmatrix}$$

for a normal random variable $\mathcal{N} \sim \mathcal{N}(0, 1)$.

8 Simulation study

8.1 From continuous to discrete data

In this section, we apply the estimator (4) to simulated processes. We test whether the results are good or not depending on sign and magnitude of involved quantities (cf. [36, Section 4.1]).

Estimator (4) involves (ξ_0, ξ_T) which are observable, as well as $(L_T(\xi), Q_T^-, Q_T^+)$ which are not.

Imagine that we observe ξ_t on a discrete time grid $\{kT/n; k = 0, \dots, n\}$. The time step between two observations is $\Delta t = T/n$. In [30] we discuss, in the case of the Oscillating Brownian Motion, the estimator for the occupation time Q_T^+ given the Riemann sums:

$$\widehat{Q}_{T,n}^+ = \Delta t \sum_{k=1}^{n} \mathbf{1}_{\xi_{kT/n} \geq 0}.$$

We proved that the speed of convergence is strictly better than $\sqrt{n}$, meaning that $\sqrt{n}(\widehat{Q}_{T,n}^+ - Q_T^+) \xrightarrow[n\to\infty]{\mathbb{P}} 0$. This result can be extended to the drifted process ξ via the Girsanov theorem.

Again in [30], we propose the following approximation of the local time of ξ at 0

$$\widehat{L}_{T,n} = \frac{-3}{2}\sqrt{\frac{\pi}{2\Delta t}}\frac{\sigma_+ + \sigma-}{\sigma_+\sigma_-}\sum_{k=1}^{n}\left((\!|\xi_{kT/n}|\!)^+ - (\!|\xi_{(k-1)T/n}|\!)^+\right)\cdot\left((\!|\xi_{kT/n}|\!)^- - (\!|\xi_{(k-1)T/n}|\!)^-\right). \quad (33)$$

Up to some constants, this is essentially an approximation of the covariation between $(\!|\xi_t|\!)^+$ and $(\!|\xi_t|\!)^-$. In [30] we also proved the consistence of such estimator. For the Brownian motion, this estimator converges at speed $n^{1/4}$. The proof has not been adapted to our case because of the technical difficulties due to the discontinuity of the coefficients in 0. Anyway we conjecture a rate of $1/4$. An alternative estimator can be constructed counting the number of crossings, as in [20]. We use (33) for empirical reasons, since it performs better on simulated trajectory.

At this point, we also notice that estimator (33) involves $\sigma_\pm$, which are quantities not directly observable on the trajectory; let us mention that the same is true for the classic estimator of local time for SDEs with differentiable coefficients [20]. Therefore, if $\sigma_\pm$ are known a priori, (33) can be implemented using such quantities. Otherwise, $\sigma_\pm$ must be estimated on discrete time observations of ξ as well, a problem which has been thoroughly investigated in [31].

We do not push further in the present paper the theoretical discussion on the quality of these discrete time approximations. Some more insights, based on numerical results, are given in the following section.

8.2 Implementation and simulation

The aim of the following section is to show on figures the numerical evidence of the central limit theorems stated in Section 4. We also mean to say something more on the choice of the step of the time grid in relation to the quality of the estimation of the local time. The code used for the following simulations have been implemented using the software R. We will consider time grids of the type $\{0, T/N, 2T/N, \dots, T\} \subset \mathbb{N}$, so with $N \in \mathbb{N}$, and use as approximation of local and occupation time the following:

$$\widehat{L}_{T,N} \text{ and } \widehat{Q}_T^+ := \widehat{Q}_{T,N}^+.$$

Summing up, as estimators of $b_\pm$ we use the approximation

$$\widehat{b}_{T,N}^\pm = \pm\frac{(\!|\xi_T|\!)^\pm - (\!|\xi_0|\!)^\pm - \widehat{L}_{T,N}/2}{\widehat{Q}_{T,N}^\pm}. \quad (34)$$

These are our discrete times approximations of the continuous time estimators for which the convergence results have been proved. In practice, in some cases the estimator does not really depend on the local time, but is essentially determined by the final value of the process and the occupation times. In those cases, the quality of the discrete time approximation of the local time does not really matter, and therefore we can take $N = 1$. In most cases, anyway, a good approximation of the local time is needed in order to observe on simulations the theoretical central limit behavior expected from Section 4. In these cases $N \in \mathbb{N}$ must be taken large enough.

In what follows, the choice of the parameters is detailed for every figure. The diffusion parameter is taken constant $\sigma_+ = \sigma_- = 0.01$, the same one for all the different simulations, and supposed known a priori. In such manner the CLTs that we want to test can be better observed. To use estimator (34) on real data, with $\sigma_\pm$ not known, one must first use the estimators for $\sigma_\pm$ in [31] and then use such estimations in the estimator (33) for the local time. We indicate with $[+]$ and $[-]$ estimation on positive and negative semiaxis, e.g. **(N1)**$[+]$ stands for "estimation of b_+" in case **N1**.

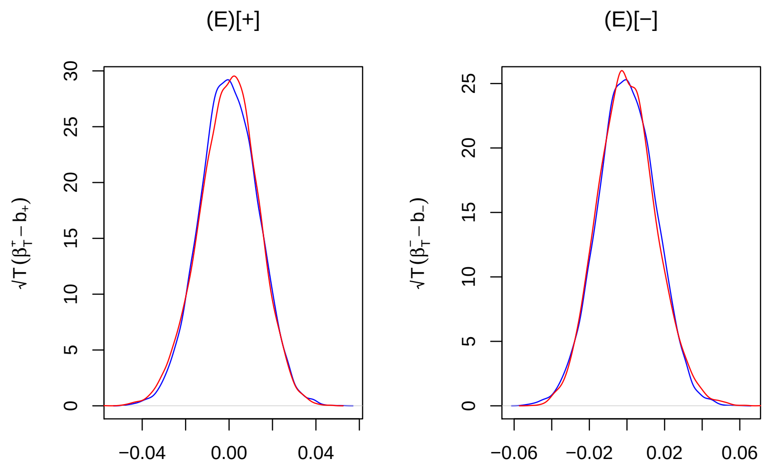

Figure 1: **(E)**. SDE parameters: $\sigma_\pm = 0.01$, $b_- = 0.004$, $b_+ = -0.003$. Simulation parameters: $T = 1000$, $N = 100\,000$. We show both sides (positive and negative) of the estimation, displaying the density of $\sqrt{T}(\beta_T^\pm - b_\pm)$. The CLT in Proposition (7) is accurate for large T and time step T/N small, since the quality of the estimation of the local time is key in this case. The limit behavior is Gaussian.

DOI 10.20347/WIAS.PREPRINT.2497

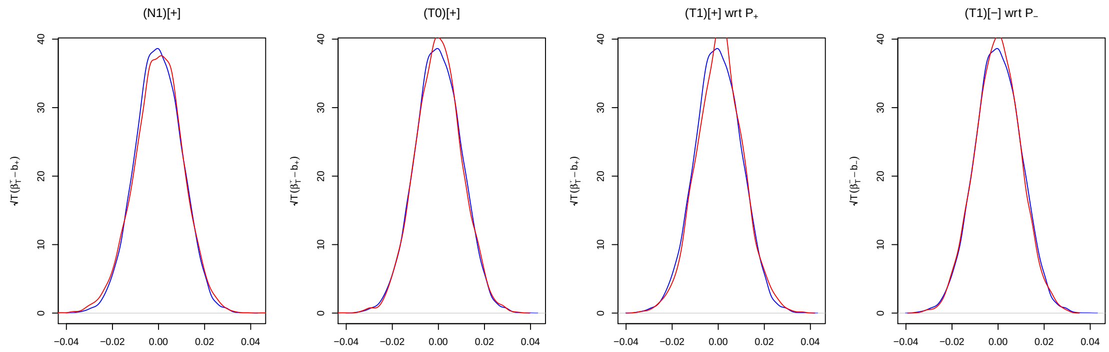

Figure 2: **(N1)**$[+]$, **(T0)**$[+]$, **(T1)**$[+]$ w.r.t $\mathbb{P}_+$, **(T1)**$[-]$ w.r.t $\mathbb{P}_-$. SDE parameters: $\sigma_\pm = 0.01$; in case **N1**: $b_- = 0.004$, $b_+ = 0$; in case **T0**: $b_- = 0.004$, $b_+ = 0.006$; in case **T1**: $b_- = -0.004$, $b_+ = 0.003$. Simulation parameters: $T = 1000$, $N = 1000$. We display the density of $\sqrt{T}(\beta_T^+ - b_+)$ in cases **N1** and **T0**. We also show $\sqrt{T}(\beta_T^\pm - b_\pm)$ in case **T1**, but the density is w.r.t $\mathbb{P}_\pm$ (cf. (6)). This is approximated computing the estimator on trajectories such that ξ_T is larger or respectively smaller than 0. In all these cases the CLT is Gaussian and we do not need to have a fine discretisation/time grid, since the local time is asymptotically negligible and the quantities which matter in the estimator are ξ_T and the occupation times. This accounts of (9)-positive part, (10), (12) and (13).

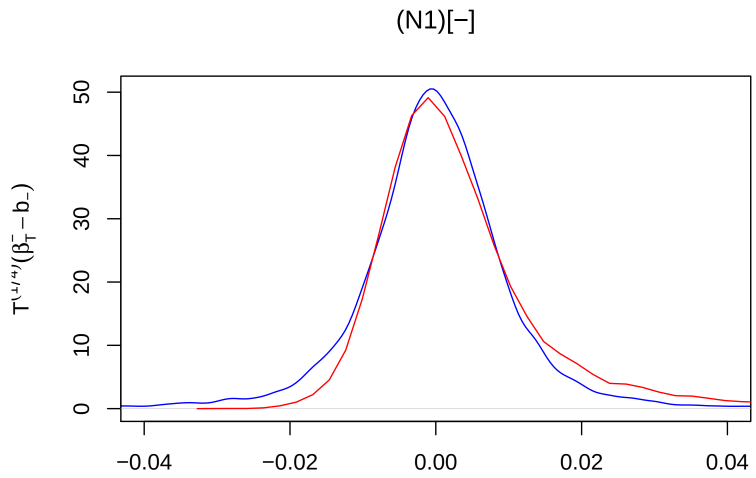

Figure 3: **(N1)**$[-]$. SDE parameters: $\sigma_\pm = 0.01$; in case **N1**: $b_- = 0.004; b_+ = 0$. Simulation parameters: $T = 1000$, $N = 100\,000$. The CLT in (9)-negative part, is accurate for large T and time step T/N small, since the quality of the estimation of the local time is key in this case. Remark that in this case (null recurrent), the CLT has speed of convergence $T^{1/4}$ and the limit law is not Gaussian. This accounts of (9)-negative part.

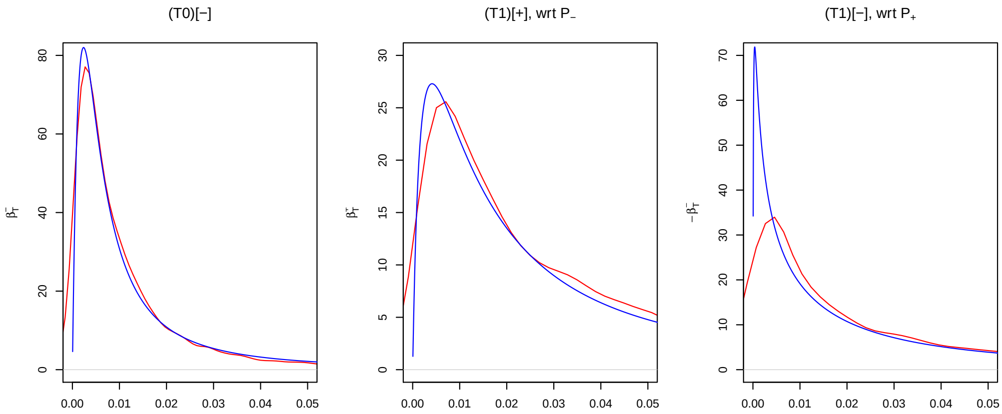

Figure 4: **(T0)**$[-]$, **(T1)**$[+]$ w.r.t to $\mathbb{P}_-$ and **(T1)**$[-]$ w.r.t to $\mathbb{P}_+$. SDE parameters: $\sigma_\pm = 0.01$; in case **T0**: $b_- = 0.004$, $b_+ = 0.003$; in case **T1**: $b_- = -0.003$, $b_+ = 0.01$. In case **T0**: simulation parameters: $T = 20$, $N = 1000$. We display the density of β_T^- . In case **T1**: simulation parameters: $T = 20$, $N = 3 \times 10^7$. We display the density of $\beta_T^\pm$ w.r.t $\mathbb{P}_\mp$ (cf. (6)). This is approximated computing the estimator on trajectories such that ξ_T is smaller or respectively larger than 0. In these cases the estimator is not consistent, so what we show is not actually a CLT but the convergence of the estimators towards the law (22) (cf. results (11), (14), (15)). This convergence is accurate for large T but also depends on the time step T/N. Moreover, we see that the theoretical distribution of β_T^- in case **T1** is almost singular at the origin, and therefore the exact behavior near the origin is hard to catch on simulated trajectories. This can be improved using different kernels (instead of the Gaussian one) in the estimation of the density. This can be easily done with the function “density” in $\mathbb{R}$. The limit behavior is better approximated when b_+ and b_- have similar magnitude. We chose here to display the case $b_- = -0.003$; $b_+ = 0.01$ to mention this critical behavior. Anyway, this feature does not really matter in statistical application, because in this case the estimator not only is not consistent, but does not even guess the correct sign of the parameter. Indeed, we are here in the very critical case of a transient process generated by diverging drift **T1**.

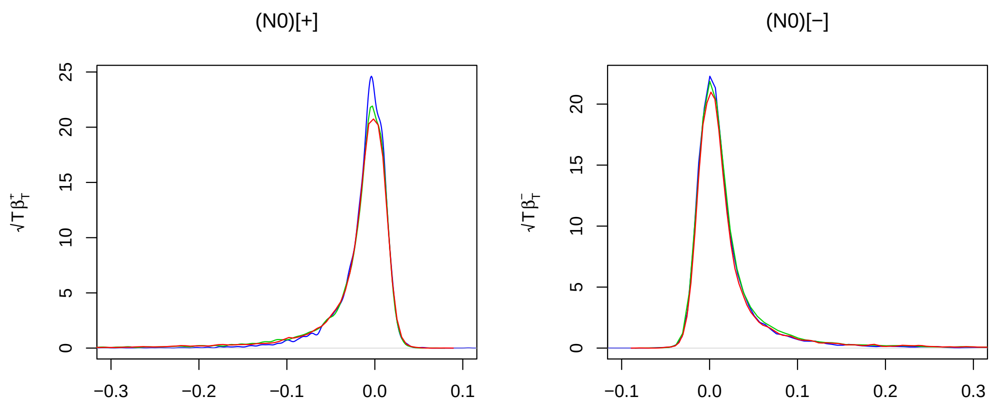

Figure 5: **(N0)**. SDE parameters: $\sigma_{\pm} = 0.01,\ b_{-} = 0\, b_{+} = 0;$. Simulation parameters: $T = 10,\ N =$ 1000; $T = 100,\ N =$ 10 000; $T =$ 1000, $N =$ 100 000. Differently from before, we do not show the convergence to the scaled limit law (25), but the scaling relation (8), for three different final times. We show it on both positive and negative semiaxes. Because of the estimation of the local time, this also depends on the choice of N.

References

[1] Mohamed Alaya Ben and Ahmed Kebaier. Parameter estimation for the square-root diffusions: ergodic and nonergodic cases. *Stoch. Models*, 28(4):609–634, 2012.

[2] T. Appuhamillage, V. Bokil, E. Thomann, E. Waymire, and B. Wood. Occupation and local times for skew Brownian motion with applications to dispersion across an interface. *Ann. Appl. Probab.*, 21(1):183–214, 2011.

[3] Soren Asmussen and Michael Taksar. Controlled diffusion models for optimal dividend pay-out. *Insurance: Mathematics and Economics*, 20(1):1–15, 1997.

[4] Federico M. Bandi and Peter C. B. Phillips. Fully nonparametric estimation of scalar diffusion models. *Econometrica*, 71(1):241–283, 2003.

[5] Patrick Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1999. A Wiley-Interscience Publication.

[6] Michael P Clements and Jeremy Smith. A Monte Carlo study of the forecasting performance of empirical SETAR models. *Journal of Applied Econometrics*, 14(2):123–141, March 1999.

[7] Irene Crimaldi and Luca Pratelli. Convergence results for multivariate martingales. *Stochastic Process. Appl.*, 115(4):571–577, 2005.

[8] D. A. Darling and M. Kac. On occupation times for Markoff processes. *Transactions of the American Mathematical Society*, 84(2):444–458, 1957.

[9] Marc Decamps, Marc Goovaerts, and Wim Schoutens. Self exciting threshold interest rates models. *Int. J. Theor. Appl. Finance*, 9(7):1093–1122, 2006.

[10] Pierre Étoré. On random walk simulation of one-dimensional diffusion processes with discontinuous coefficients. *Electron. J. Probab.*, 11:no. 9, 249–275, 2006.

[11] William Feller. Generalized second order differential operators and their lateral conditions. *Illinois J. Math.*, 1:459–504, 1957.

[12] A. Gairat and V. Shcherbakov. Density of skew Brownian motion and its functionals with application in finance. *Mathematical Finance*, 26(4):1069–1088, 2016.

[13] Hans U. Gerber Gerber and Elias S.W. Shiu. On optimal dividends: From reflection to refraction. *Journal of Computational and Applied Mathematics*, 186(1):4 – 22, 2006. Special Issue: Jef Teugels.

[14] R. Höpfner and E. Löcherbach. Limit theorems for null recurrent Markov processes. *Mem. Amer. Math. Soc.*, 161(768), 2003.

[15] Scott Hottovy and Samuel N. Stechmann. Threshold models for rainfall and convection: Deterministic versus stochastic triggers. *SIAM Journal on Applied Mathematics*, 75(2):861–884, 2015.

[16] I. A. Ibragimov and R. Z. Has'minskiĭ. *Statistical estimation*, volume 16 of *Applications of Mathematics*. Springer-Verlag, New York-Berlin, 1981.

[17] Tomoyuki Ichiba, Ioannis Karatzas, and Mykhaylo Shkolnikov. Strong solutions of stochastic equations with rank-based coefficients. *Probab. Theory Related Fields*, 156(1-2):229–248, 2013.

[18] Nobuyuki Ikeda and Shinzo Watanabe. *Stochastic differential equations and diffusion processes*, volume 24 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam; Kodansha, Ltd., Tokyo, second edition, 1989.

[19] K. Itô and H.P.J. McKean. *Diffusion Processes and their Sample Paths: Reprint of the 1974 Edition*. Classics in Mathematics. Springer Berlin Heidelberg, 1996.

[20] Jean Jacod. Rates of convergence to the local time of a diffusion. *Ann. Inst. H. Poincaré Probab. Statist.*, 34(4):505–544, 1998.

[21] Jean Jacod and Albert N. Shiryaev. *Limit theorems for stochastic processes*, volume 288 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, second edition, 2003.

[22] P. Jeganathan. On the asymptotic theory of estimation when the limit of the log-likelihood ratios is mixed normal. *Sankhyā Ser. A*, 44(2):173–212, 1982.

[23] Ioannis Karatzas and Steven E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, second edition, 1991.

[24] Julian Keilson and Jon A. Wellner. Oscillating Brownian motion. *J. Appl. Probability*, 15(2):300–310, 1978.

[25] Yu. A. Kutoyants. Efficient density estimation for ergodic diffusion processes. *Statistical Inference for Stochastic Processes*, 1(2):131–155, 1998.

[26] Yury A. Kutoyants. On identification of the threshold diffusion processes. *Ann. Inst. Statist. Math.*, 64(2):383–413, 2012.

[27] Lucien Le Cam and Grace Lo Yang. *Asymptotics in statistics*. Springer Series in Statistics. Springer-Verlag, New York, second edition, 2000.

[28] J.-F. Le Gall. One-dimensional stochastic differential equations involving the local times of the unknown process. *Stochastic Analysis. Lecture Notes Math.*, 1095:51–82, 1985.

[29] Lucien LeCam. On some asymptotic properties of maximum likelihood estimates and related Bayes' estimates. *Univ. California Publ. Statist.*, 1:277–329, 1953.

[30] A. Lejay and P. Pigato. Statistical estimation of the Oscillating Brownian Motion. *Bernoulli*, 2017. To appear.

[31] A. Lejay and P. Pigato. A threshold model for local volatility: evidence of leverage and mean reversion effects on historical data, 2018. Preprint.

[32] D. Lépingle. Sur le comportement asymptotique des martingales locales. In *Séminaire de Probabilités, XII (Univ. Strasbourg, Strasbourg, 1976/1977)*, volume 649 of *Lecture Notes in Math.*, pages 148–161. Springer, Berlin, 1978.

[33] R.S. Lipster and A.N. Shiryaev. *Statistics of random processes. II. Applications*. Springer, 2001.

[34] Gisirô Maruyama and Hiroshi Tanaka. Some properties of one-dimensional diffusion processes. *Mem. Fac. Sci. Kyusyu Univ. Ser. A. Math.*, 11:117–141, 1957.

[35] Michael Monoyios and Lucio Sarno. Mean reversion in stock index futures markets: A nonlinear analysis. *Journal of Futures Markets*, 22(4):285–30, April 2002.

[36] Pedro P. Mota and Manuel L. Esquível. On a continuous time stock price model with regime switching, delay, and threshold. *Quant. Finance*, 14(8):1479–1488, 2014.

[37] Jim Pitman and Marc Yor. Hitting, occupation and inverse local times of one-dimensional diffusions: martingale and excursion approaches. *Bernoulli*, 9(1):1–24, 02 2003.

[38] James M. Poterba and Lawrence H. Summers. Mean reversion in stock prices. *Journal of Financial Economics*, 22(1):27–59, 1988.

[39] J. M. Ramirez, E. A. Thomann, and E. C. Waymire. Advection–dispersion across interfaces. *Statist. Sci.*, 28(4):487–509, 2013.

[40] Alfréd Rényi. On stable sequences of events. *Sankhyā Ser. A*, 25:293–302, 1963.

[41] L. C. G. Rogers and David Williams. *Diffusions, Markov processes, and martingales. Vol. 2*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2000. Itô calculus, Reprint of the second (1994) edition.

[42] F. Sattin. Fick's law and Fokker–Planck equation in inhomogeneous environments. *Physics Letters A*, 37(22):3941–3945, 2008.

[43] Fei Su and Kung-Sik Chan. Quasi-likelihood estimation of a threshold diffusion process. *J. Econometrics*, 189(2):473–484, 2015.

[44] Fei Su and Kung-Sik Chan. Testing for threshold diffusion. *J. Bus. Econom. Statist.*, 35(2):218–227, 2017.

[45] Howell Tong. *Threshold models in nonlinear time series analysis*, volume 21 of *Lecture Notes in Statistics*. Springer-Verlag, New York, 1983.

[46] Howell Tong. Threshold models in time series analysis—30 years on. *Stat. Interface*, 4(2):107–118, 2011.

[47] Hisao Watanabe and Minoru Motoo. Ergodic property of recurrent diffusion processes. *J. Math. Soc. Japan*, 10:272–286, 1958.

[48] Shinzo Watanabe. Generalized arc-sine laws for one-dimensional diffusion processes and random walks. In *Stochastic analysis (Ithaca, NY, 1993)*, volume 57 of *Proc. Sympos. Pure Math.*, pages 157–172. Amer. Math. Soc., Providence, RI, 1995.

[49] Pradeep K Yadav, Peter F Pope, and Krishna Paudyal. Threshold autoregressive modeling in finance: the price differences of equivalent assets. *Mathematical Finance*, 4(2):205–221, April 1994.