Two consistent estimators for the Skew Brownian motion

Antoine Lejay, Ernesto Mordecki, Soledad Torres

Two consistent estimators for the Skew Brownian motion

Antoine Lejay* Ernesto Mordecki† Soledad Torres ‡

May 19, 2017

Abstract

The Skew Brownian motion is of primary importance in modeling diffusion in media with interfaces which arise in many domains ranging from population ecology to geophysics and finance. We show that the maximum likelihood estimator provides a consistent estimator of the parameter of a Skew Brownian motion observed at discrete times. The difficulties are that this process is only null recurrent and has a singular distribution with respect to the one of the Brownian motion. Finally, using the idea of the Expectation-Maximization algorithm, we show that the maximum likelihood estimator can be naturally interpreted as the expected number of positive excursions divided by the expected number of excursions.

Keywords. Skew Brownian motion; maximum likelihood estimator (MLE) null recurrent process; Expectation-Maximization (EM) algorithm; excursion theory.

1. Introduction

Introduced in the 70's as a variation of the Brownian motion [8, 10, 24], the Skew Brownian motion (SBm) has attracted a lot of interest as a "basic

*Université de Lorraine, IECL, UMR 7502, Vandœuvre-lès-Nancy, F-54600, France
CNRS, IECL, UMR 7502, Vandœuvre-lès-Nancy, F-54600, France
Inria, Villers-lès-Nancy, F-54600, France
E-mail: Antoine.Lejay@univ-lorraine.fr

†Facultad de Ciencias, Centro de Matemática. Iguá 4225, CP 11400, Montevideo, Uruguay
E-mail: mordecki@cmat.edu.uy

‡CIMFAV- Facultad de Ingeniería, Universidad de Valparaíso, Chile.
E-mail: Soledad.torres@uv.cl

brick" to model diffusion phenomena in presence of permeable barriers. Actually, the SBm is related through suitable changes of variables to the more general class of stochastic differential equations (SDE) with local times [14, 15]. The need for such models arise in many domains (see [17, 19] for more references): population ecology [23], finance [1, 5, 25], geophysics [2], electro-encephalography [21], molecular chemistry [4], meteorology [27], oceanography [26], and astrophysics [29].

As such, the SBm $X = (X_t)_{t\geq 0}$ is the strong solution to

$$X_t = x + B_t + \theta L_t,$$

where $B = (B_t)_{t\geq 0}$ is a Brownian motion, $L = (L_t)_{t\geq 0}$ is the symmetric local time of X at 0, and θ a parameter in $[-1, 1]$.

When $\theta = 1$ (resp. $\theta = -1$), then X is a positively (resp. negatively) reflected Brownian motion. When $\theta = 0$, X is a Brownian motion starting at x. The parameter θ reflects the trend of the process to move upward (when $\theta > 0$) or downward (when $\theta < 0$) when it crosses 0. Since the process comes back immediately to 0, this is only a heuristic description that could be expressed rigorously with the help of the excursion theory [10]. Actually, there are many ways to construct the SBm (the article [17] gives ten of them, but more are possible).

The problem we consider in this paper is the following: given observations $(X_{t_i})_{i=1,\dots,n}$ at times $t_i = iT/n$ of the SBm, how to estimate θ?

At the best of our knowledge, this problem was dealt first in [3] for a SBm living in a finite domain in which ergodicity is used, while the SBm in the free space is null recurrent. As pointed out in [20], the occupation time of positive or negative axis could not be used to construct estimators as their variance do not decrease to 0 whatever the number of observations.

The density transition function of the SBm can be expressed in closed form [17, 28] by

$$p_\theta(t, x, y) = p(t, x - y) + \operatorname{sgn}(y)\theta p(t, |x| + |y|). \tag{1}$$

where $p(t, x) = (2\pi t)^{-1/2} \exp(-x^2/2t)$ is the Gaussian density with mean zero and variance t.

Hence, the Maximum Likelihood Estimator (MLE) is easy to construct from the observations and can be dealt with standard numerical optimization problems. This article is devoted to show the consistency of the MLE, a point left open in [20].

More precisely, in [20], we have studied the asymptotic behavior of the maximum likelihood estimator θ_n when the true distribution is the one of the Brownian motion, the case $\theta = 0$, proving its consistency. An asymptotic development of θ_n is also given, and shown that θ_n converges to θ at rate $n^{1/4}$, unlike the "standard" theory. The limit of $n^{1/4}(\theta_n - \theta)$, of mixed normal type, involves the local time.

Actually, this slow rate of convergence is due to the fact the parameter estimation depends mostly of the behavior of the process close to 0. When away from 0, the SBm behaves like a Brownian motion whatever θ. The SBm is a null recurrent process, for which the known results, such as the ones presented in [9], cannot be applied.

In this article, we show that the MLE is *consistent*, which means that θ_n converges in probability to θ for any $\theta \in [-1, 1]$ when the true distribution is the one of a SBm of parameter θ. This completes the results of [20], where only the convergence around $\theta = 0$ was considered. In the present paper we also provide another estimator, which rely on counting the number of up- and down-crossings through the level $x = 0$.

For this, we adapt some results of [12] to the case of a SBm, whose distribution is singular with respect to the one of the Brownian motion when $\theta \neq 0$ [14].

However, although it seems natural in this situation to think that the rate of convergence is of order $n^{1/4}$, the proof of this seems really intricate.

In [18], the MLE for a biased random walk is considered. The latter is an approximation of the SBm under a proper scaling. There, the excursions can be observed. Therefore, the MLE is nothing more than the ratio of the upward excursions to all the excursions. The analysis can be performed with elementary tools due to this possibility.

Using the relationship with the Expectation-Maximization (EM) algorithm [6, 22], we show that the MLE can be understood in a simple manner as a way of counting upward and downward (unknown) excursions given the observations. This shows that the situation between the biased random walk and the one of the SBm are rather similar in spirit.

As pointed out at the beginning of the introduction, the SBm is also strongly related to processes with discontinuous coefficients. In [16], an explicit estimation of the volatilities is provided, for an *oscillating Brownian motion* [13], a solution to a SDE whose diffusion coefficients takes two values according to the sign of its position. For this, we use a reduction to the SBm through a simple transform and some of the results given here regarding the convergence of the estimators. We then believe that this work could be applied to deal with other problems related to the estimation of process with discontinuous coefficients, a subject which is largely open.

Outline. In Section 2, we extend some of the results of J. Jacod [12] to the distribution of the Skew Brownian motion. In Section 3, we study the limit of ratios of estimators. The consistency of the ratio of up- and down-crossings as well as the MLE are studied in Section 4. Finally, in Section 5, we give an alternative expression for the MLE which relies on counted the (unknown) number of excursions straddling the sample's times.

2. Some convergence results

As in (1), we denote by $p(t,x)$ the Gaussian density with mean 0 and variance t. For a measurable, bounded function f, set $P_t f(x) = \int p(t, y - x) f(y)\,\mathrm{d}y$.

Set $f_+(x) = f(x)\mathbf{1}_{\{x\geq 0\}}$ and $f_-(x) = f(x)\mathbf{1}_{\{x<0\}}$.

Acting on the space of continuous functions that vanish at infinity, the semi-group $P^\theta = (P_t^\theta)_{t\geq 0}$ of the SBm of parameter θ is

$$P_t^\theta f(x) = \int p_\theta(t,x,y) f(y)\,\mathrm{d}y = P_t f(x) + \theta P_t f_+(-|x|) - \theta P_t f_-(|x|), \quad (2)$$

where p_θ is given by (1). For $f\colon \mathbb{R}\to\mathbb{R}$ and $\gamma \geq 0$, we define

$$\begin{aligned}\beta_\gamma(f) &= \int_{-\infty}^{+\infty} |x|^\gamma |f(x)|\,\mathrm{d}x,\\ \beta_\gamma^\theta(f) &= \beta_\gamma(f) + |\theta|\left(\beta_\gamma(f_+) + \beta_\gamma(f_-)\right).\end{aligned}$$

For any $\theta \in (-1,1)$,

$$\beta_\gamma^\theta(f) \leq 2\beta_\gamma(f). \quad (3)$$

When the integrals are well defined, we set $\lambda(f) = \int_{\mathbb{R}} f(x)\,\mathrm{d}x$, and

$$\begin{aligned}\lambda_\theta(f) &= (1+\theta)\int_0^{+\infty} f(x)\,\mathrm{d}x + (1-\theta)\int_{-\infty}^0 f(x)\,\mathrm{d}x\\ &= \lambda(f) + \theta(\lambda(f_+) - \lambda(f_-)).\end{aligned}$$

If f is even then $\lambda_\theta(f) = \lambda(f)$ for all $\theta\in[-1,1]$. In addition, $\lambda_\theta(f) \leq 2\beta_0(f)$.

We adapt from [12, Lemma 3.1, p. 518] the following result which is the key to identify the limit.

Lemma 1. *If $\beta_\gamma(f) < \infty$ for $\gamma = 0,1,2$, then for some constant K,*

$$|P_t^\theta f(x)| \leq \frac{\sqrt{2}}{\sqrt{\pi t}}\beta_0(f), \quad (4)$$

$$|P_t^\theta f(x) - \lambda_\theta(f) p(t,x)| \leq \frac{K}{t^{3/2}}(\beta_2^\theta(f) + \beta_1^\theta(f)|x|). \quad (5)$$

Proof. In this proof, the constant K may vary from line to line. Inequality (4) is a direct consequence of $p_\theta(t,x,y) \le (1+|\theta|)/\sqrt{2\pi t}$. Using (2) and (3.3) in Lemma 3.1 in [12],

$$\begin{aligned}
&|P_t^\theta f(x) - \lambda_\theta(f)p(t,x)| \\
&\quad \le |P_t f(x) - \lambda(f)p(t,x)| + |\theta|\cdot|P_t f_+(-|x|) - \lambda(f_+)p(t,x)| \\
&\qquad\qquad + |\theta|\cdot|P_t f_-(|x|) - \lambda(f_-)p(t,x)| \\
&\quad \le \frac{K}{t^{3/2}}\left(\beta_2(f) + \beta_1(f)|x|\right) + \frac{K|\theta|}{t^{3/2}}\left(\beta_2(f_+) + \beta_1(f_+)|x|\right) \\
&\qquad\qquad + \frac{K|\theta|}{t^{3/2}}\left(\beta_2(f_-) + \beta_1(f_-)|x|\right) \\
&\qquad\qquad\qquad = \frac{K}{t^{3/2}}\left(\beta_2^\theta(f) + \beta_1^\theta(f)|x|\right).
\end{aligned}$$

This shows (5). □

Hypothesis 1. The function f is bounded and $\beta_\gamma(f) < +\infty$ for $\gamma = 0,1,2$.

Fix $T > 0$ and set $t_i^n = iT/n$. Let B be a Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. This Brownian motion B generates a filtration $(\mathcal{F}_t)_{t\ge 0}$ which is completed and augmented. With the result of [8], there exists a unique strong solution of the stochastic differential equation with local time $X_t = x + B_t + \theta L_t$ for any $\theta \in [-1,1]$. This process is the SBm of parameter θ. We denote its distribution by $\mathbb{P}^\theta$.

Lemma 2. *Let f satisfy Hypothesis 1 and $\lambda_\theta(f) = 0$ for some $\theta \in [-1,1]$. Then, for some constant K which depends on T, we have*

$$\begin{aligned}
&\mathbb{E}^\theta\left[\sup_{t\in[0,T]}\left|\frac{1}{\sqrt{n}}\sum_{i=0}^{\lfloor (n-1)t\rfloor} f(X_{t_i^n}\sqrt{n})\right|^2\right] \\
&\qquad\qquad \le \frac{K}{\sqrt{n}}\left(\beta_0(f)\sup|f| + \beta_0(f)\beta_2(f) + \beta_1(f)^2\right).
\end{aligned}$$

Proof. As the SBm presents the same scaling property as the Brownian motion, we set for the sake of simplicity $T = 1$.

Set $f_n(x) = f(x\sqrt{n})$ so that $\beta_\gamma(f_n) = n^{-(\gamma+1)/2}\beta_\gamma(f)$ for $\gamma = 0,1,2$. Also $\lambda_\theta(f_n) = 0$.

For $i > 1$, with (5), (3) and since $\lambda_\theta(f_n) = 0$ and $\sum_{i=1}^n i^{-3/2} \le \zeta(3/2) \le 3$,

$$\begin{aligned}\left|\sum_{j=i+1}^{n} \mathbb{E}^\theta[f(X_{t_j^n}\sqrt{n})|X_{t_i^n} = x]\right| &\le \sum_{j=i+1}^{n} |P^\theta_{(j-i)/n} f_n(x)| \\ &\le \sum_{j=i+1}^{n} \frac{K}{(j-i)^{3/2}}(\beta_2(f) + \beta_1(f)\sqrt{n}|x|) \\ &\le K'(\beta_2(f) + \beta_1(f)\sqrt{n}|x|) \text{ with } K' = 3K.\end{aligned}$$

With $0 \le t \le 1$, (4), the scaling property of β_0 and the controls (3) of β^θ_γ by β_γ,

$$\begin{aligned}&\left|\sum_{i=1}^{\lfloor nt\rfloor}\sum_{j=i+1}^{\lfloor nt\rfloor} \mathbb{E}^\theta[f(X_{t_i^n}\sqrt{n})f(X_{t_j^n}\sqrt{n})]\right| \\ &\le \sum_{i=1}^{n} \mathbb{E}^\theta\left[|f(X_{t_i^n}\sqrt{n})|(\beta_2(f) + \beta_1(f)\sqrt{n}|X_{t_i^n}|)K'\right] \\ &\le \sum_{i=1}^{n} \frac{\sqrt{2}K'}{\sqrt{\pi}\sqrt{i}}\beta_0(|f(x)|(\beta_2(f) + \beta_1(f)|x|)) \\ &\le K''\beta_0(|f(x)|(\beta_2(f) + \beta_1(f)|x|))(1 + 2\sqrt{n}) \text{ with } K'' = \frac{\sqrt{2}K'}{\sqrt{\pi}}.\end{aligned}$$

The right-hand side is finite for each n since

$$\beta_0(g) = \beta_1(f) \text{ for } g(x) = f(x)|x|. \tag{6}$$

Hence,

$$\begin{aligned}&\mathbb{E}^\theta\left[\sup_{t\in[0,1]}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{\lfloor nt\rfloor} f(X_{t_i^n}\sqrt{n})\right)^2\right] \\ &\le \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[f(X_{t_i^n}\sqrt{n})^2] + \frac{2}{n}\left|\sum_{i=1}^{\lfloor nt\rfloor}\sum_{j=i+1}^{\lfloor nt\rfloor}\mathbb{E}[f(X_{t_i^n}\sqrt{n})f(X_{t_j^n}\sqrt{n})]\right| \\ &\le \beta_0(f^2)\frac{K}{n^{3/2}}\sum_{i=1}^{n}\frac{n^{3/2}}{i^{3/2}} + \frac{K''(1+2\sqrt{n})}{n}(\beta_0(|f(x)|(\beta_2(f) + \beta_1(f)|x|)) \\ &\le \beta_0(f^2)\frac{K}{\sqrt{n}} + \frac{2K''}{\sqrt{n}}\big(\beta_0(f)\beta_2(f) + \beta_1(f)^2\big),\end{aligned}$$

where (6) is used to established the last inequality. Since f is bounded, $\beta_0(f^2) < \beta_0(f) \sup_{x\in\mathbb{R}} |f(x)|$. This proves the result. $\square$

Proposition 1. *Let f be a function satisfying Hypothesis 1. Then for any $a > 0$,*

$$\sup_{\theta\in[-1,1]} \mathbb{P}^\theta \left[\sup_{t\in[0,T]} \left| \frac{1}{\sqrt{n}} \sum_{i=0}^{\lfloor (n-1)t \rfloor} f(X_{t_i^n}\sqrt{n}) - \lambda_\theta(f) L_t \right| > a \right] \xrightarrow[n\to\infty]{} 0. \tag{7}$$

Proof. Whatever the value of θ, $|X|$ is distributed like a reflected Brownian motion under $\mathbb{P}^\theta$. The function

$$g(x) = \mathbb{E}^\theta[|X_1| - |x|] = \int (|x+y| - |x|) p(1,y) dy$$

is an even function that does not depend on θ and that satisfies $\lambda_\theta(g) = \lambda(g) = 1$ for any $\theta \in [-1,1]$ (see 1.14 in [12]). By the scaling property and using the Itô-Tanaka formula,

$$\frac{g(X_{t_i^n}\sqrt{n})}{\sqrt{n}} = \mathbb{E}^\theta[|X_{t_{i+1}^n}| - |X_{t_i^n}|\, |\mathcal{F}_{t_i^n}] = \mathbb{E}^\theta[L_{t_{i+1}^n} - L_{t_i^n} | \mathcal{F}_{t_i^n}].$$

From [11, Lemma 2.14], since the distribution of the local time do not depend on θ, for any $a > 0$,

$$\sup_{\theta\in[-1,1]} \mathbb{P}^\theta \left[\sup_{t\in[0,T]} \left| \sum_{i=0}^{\lfloor (n-1)t \rfloor} \mathbb{E}^\theta[L_{t_{i+1}^n} - L_{t_i^n} | \mathcal{F}_{t_i^n}] - L_t \right| \geq a \right] \xrightarrow[n\to\infty]{} 0. \tag{8}$$

Set $h(x) = f(x) - \lambda_\theta(f) g(x)$, so that $\lambda_\theta(h) = 0$. Using Lemma 2 and (8), one gets that (7) holds. $\square$

Hypothesis 2. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a function such that for

$$F_k(x) = \int_{\mathbb{R}} p_\theta(1,x,y) f(x,y)^\gamma \, \mathrm{d}y \text{ for } \gamma = 1, 2,$$

the functions F_1, $(F_1)^2$ and F_2 satisfy Hypothesis 1.

Since $p_\theta(1/n, x, y) = \sqrt{n} p_\theta(1, x\sqrt{n}, y\sqrt{n})$,

$$F_k(X_{t_i^n}\sqrt{n}) = \mathbb{E}^\theta[f(X_{t_i^n}\sqrt{n}, X_{t_{i+1}^n}\sqrt{n})^\gamma | \mathcal{F}_{t_i^n}] \text{ with } \gamma = 1, 2,$$

a relation we have already used in the proof of Proposition 1.

Proposition 2. *Let f be a function satisfying Hypothesis 2. Then, for any $a > 0$,*

$$\sup_{\theta\in[-1,1]} \mathbb{P}^\theta \left[\sup_{t\in[0,T]} \left| \frac{1}{\sqrt{n}} \sum_{i=0}^{\lfloor (n-1)t\rfloor} f(X_{t_i^n}\sqrt{n}, X_{t_{i+1}^n}\sqrt{n}) - \lambda_\theta(F_1) L_t \right| > a \right] \xrightarrow[n\to\infty]{} 0. \tag{9}$$

Proof. Applying Doob's inequality

$$\mathbb{E}^\theta \left[\sup_{t\in[0,T]} \left| \frac{1}{\sqrt{n}} \sum_{i=0}^{\lfloor (n-1)t\rfloor} (f(X_{t_i^n}\sqrt{n}, X_{t_{i+1}^n}\sqrt{n}) - F_1(X_{t_i^n}\sqrt{n})) \right| \right]^2 \\ \leq \frac{K}{n} \mathbb{E}^\theta \left[\sum_{i=0}^{n-1} \big(f(X_{t_i^n}\sqrt{n}, X_{t_{i+1}^n}\sqrt{n}) - F_1(X_{t_i^n}\sqrt{n}) \big) \right]^2 .$$

Using a conditional expectation type argument, we obtain that the cross terms in the following sum vanish so that

$$\frac{1}{n} \mathbb{E}^\theta \left(\sum_{i=0}^{n-1} (f(X_{t_i^n}\sqrt{n}, X_{t_{i+1}^n}\sqrt{n}) - F_1(X_{t_i^n}\sqrt{n})) \right)^2 \\ = \frac{1}{n} \sum_{i=0}^{n-1} (\mathbb{E}^\theta[F_2(X_{t_i^n}\sqrt{n})] - \mathbb{E}^\theta[F_1(X_{t_i^n}\sqrt{n})^2]).$$

Hence, this quantity converges in probability to 0 as $n \to \infty$ by Proposition 1 uniformly in $\theta \in [-1, 1]$. One deduces that (9) holds. □

3. Limit theorems for ratios

In this section we study the asymptotic behavior of some estimators given by ratios. For this purpose we introduce two new hypotheses.

Hypothesis 3. Let f and g be functions satisfying Hypothesis 2 with $g \geq 0$ and $\lambda_\theta(G_1) > 0$.

As the estimators we construct are sensitive to the behavior of the process around 0 and involve the local time, we could wait an observable time $\tau_0 \geq \tau$ where τ is the first hitting time from 0. The observation window is then reduced from $[0, T]$ to $[\tau_0, T]$. Under Hypothesis 4, $\tau_0 = 0$ and $L_T > 0$ almost surely for any $T > 0$.

Hypothesis 4. The starting point $X_0 = x$ is 0.

Let us consider the asymptotic behavior of

$$R_n(f,g) = \frac{Y_n}{Z_n} \text{ with } \begin{cases} Y_n := \sum_{i=1}^{\lfloor (n-1)T \rfloor} f(X_{t_i^n}\sqrt{n}, X_{t_{i+1}^n}\sqrt{n}), \\ Z_n := \sum_{i=1}^{\lfloor (n-1)T \rfloor} g(X_{t_i^n}\sqrt{n}, X_{t_{i+1}^n}\sqrt{n}), \end{cases}$$

under the assumption that $Z_n > 0$ almost surely.

Corollary 1. *Under Hypotheses 3 and 4, under $\mathbb{P}^\theta$, we have*

$$R_n(f,g) \xrightarrow[n\to\infty]{proba} \frac{\lambda_\theta(F_1)}{\lambda_\theta(G_1)}.$$

Proof. The convergence of $R_n(f,g)$ to $\lambda_\theta(F_1)/\lambda_\theta(G_1)$ follows from Proposition 2, standard computations and the facts that almost surely, $0 < L_T < +\infty$ and $\lambda_\theta(G_1) > 0$. $\square$

4. Estimation of the parameter of the SBm

We assume that we observe the discretization $\{X_{t_i^n}\}_{i=0,\dots,n}$ of a path of a SBm of parameter $\widetilde{\theta}$.

We are then willing to estimate the parameter $\widetilde{\theta}$ from these observations.

4.1. An estimator by the number of crossings

Our first estimator records the upward passages of the discrete observations among their passage to 0. Of course, if the observations stay positive (resp. negative), then set $\theta = 1$ (resp. $\theta = -1$). This corresponds intuitively to the best observation that one may perform.

Proposition 3. *Under the probability $\mathbb{P}^{\widetilde{\theta}}$ when $\widetilde{\theta} \in (-1,1)$*

$$\mathbf{1}_{C_n} \frac{\sum_{i=0}^{n-1} \mathbf{1}_{\{X_{t_i^n} < 0, X_{t_{i+1}^n} > 0\}}}{\sum_{i=0}^{n-1} \mathbf{1}_{\{X_{t_i^n} X_{t_{i+1}^n} < 0\}}} \xrightarrow[n\to\infty]{proba} \frac{1+\widetilde{\theta}}{2} \mathbf{1}_{L_T > 0},$$

with $C_n = \{\exists i, X_{t_i^n} X_{t_{i+1}^n} < 0\}$, the event that a crossing occurs.

Proof. Let us note that $C_n \subset \{L_T > 0\}$ and $\limsup_{n\to\infty} C_n = \{L_T > 0\}$. Set $f(x,y) = \mathbf{1}_{\{x<0,y>0\}}$. Then, for $k = 1,2$, we have

$$\mathbb{E}[f(x, X_{T/n})^k] = \int_0^{+\infty} p_{\widetilde{\theta}}(T/n, x, y)\,\mathrm{d}y = (1+\widetilde{\theta})\Psi(|x|\sqrt{n/T}),$$

with $\Psi(x) = \int_x^{+\infty} p(1,y)\,\mathrm{d}y$. With the Mills ratio, $\Psi(x) = \mathrm{O}(\exp(-|x|^2/2)/x)$, Proposition 2 applies, so that

$$\frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} \mathbf{1}_{\{X_{t_i^n}<0, X_{t_{i+1}^n}>0\}} \xrightarrow[n\to\infty]{\text{proba}} (1+\widetilde{\theta})\lambda(\Psi(|\cdot|/\sqrt{T}))L_T,$$
$$\frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} \mathbf{1}_{\{X_{t_i^n}>0, X_{t_{i+1}^n}<0\}} \xrightarrow[n\to\infty]{\text{proba}} (1-\widetilde{\theta})\lambda(\Psi(|\cdot|/\sqrt{T}))L_T.$$

Hence the result. □

4.2. Consistency of the Maximum Likelihood Estimator

Under $\mathbb{P}^\theta$, the *likelihood* of these observations is

$$\Lambda_n(\theta) = \prod_{i=0}^{n-1} p_\theta(\Delta t, X_{t_i^n}, X_{t_{i+1}^n}). \tag{10}$$

Definition 1. The *maximum likelihood estimator* (MLE) is defined by $\theta_n = \operatorname{argmax}_{\theta\in[-1,1]} \Lambda_n(\theta)$.

Our aim is to show that θ_n converges to the true parameter $\widetilde{\theta}$ under $\mathbb{P}^{\widetilde{\theta}}$ for any value of $\widetilde{\theta} \in [-1,1]$.

We start by considering particular situations.

Lemma 3. *If all the $X_{t_i^n}$ have the same sign, say positive (resp. negative), then $\theta \mapsto \Lambda_n(\theta)$ is maximal at $\theta = 1$ (resp. $\theta = -1$).*

Proof. If all the $X_{t_i^n}$ are positive (resp. negative), then $\theta \mapsto p_\theta(\Delta t, X_{t_i^n}, X_{t_{i+1}^n})$ is strictly increasing (resp. strictly decreasing). Hence the result. □

In particular, if $\widetilde{\theta} = \pm 1$, then necessarily $\theta_n = \widetilde{\theta}$.

On the other hand, if the observations keep the same sign, one cannot conclude that the true parameter $\widetilde{\theta}$ is ± 1. Simply, there is not enough information to draw any inference on $\widetilde{\theta}$ since there is no way to distinguish the observations of $\{X_{t_i^n}\}_{i=0,\dots,n}$ from the one of a reflected Brownian motion.

For $\theta \in (-1,1)$, the *score* is

$$S_n(\theta) = \partial_\theta \log \Lambda_n(\theta) = \sum_{i=0}^{n-1} \frac{\partial_\theta p_\theta(\Delta t, X_{t_i^n}, X_{t_{i+1}^n})}{p_\theta(\Delta t, X_{t_i^n}, X_{t_{i+1}^n})}$$
$$= \sum_{i=0}^{n-1} k_\theta\left(X_{t_i^n}\sqrt{n}, X_{t_{i+1}^n}\sqrt{n}\right), \tag{11}$$

where

$$k_\theta(x,y) = \frac{\partial_\theta p_\theta(\Delta t, x, y)}{p_\theta(\Delta t, x, y)} = \frac{\operatorname{sgn}(y) p(\Delta t, |x| + |y|)}{p_\theta(\Delta t, x, y)} = \frac{\operatorname{sgn}(y)}{\operatorname{sgn}(y)\theta + e^{2(xy)^+/\Delta t}}.$$

A point θ_n which maximizes the likelihood Λ_n defined by (10) also maximizes $\log \Lambda_n$. If $\theta_n \in (-1,1)$, it is also characterized as a solution to $S_n(\theta_n) = 0$. The next result shows that θ_n is uniquely defined.

Lemma 4. *The log-likelihood* $\log \Lambda_n(\theta)$ *is strictly concave in* $\theta \in (-1,1)$.

Proof. Since $\partial_\theta p_\theta(\Delta t, x, y) = \operatorname{sgn}(y) p(\Delta t, |x| + |y|)$ does not depend on θ,

$$\partial_\theta k_\theta(x,y) = \frac{-(\partial_\theta p_\theta(\Delta t, |x| + |y|))^2}{p_\theta(\Delta t, x, y)^2} < 0,$$

hence $\partial_\theta S_n(\theta) = \partial^2_{\theta\theta} \Lambda_n(\theta) < 0$. □

If for some $i \in \{0, \dots, n-1\}$, $\operatorname{sgn}(X_{t_i^n} \cdot X_{t_{i+1}^n}) = -1$, then

$$k_\theta(X_{t_i^n}\sqrt{n}, X_{t_{i+1}^n}\sqrt{n}) \xrightarrow[\theta \to \pm 1]{} \mp\infty.$$

If the sign of $i \mapsto \operatorname{sgn}(X_{t_i^n})$ changes twice, then necessarily the maximum of $\log \Lambda_n$ is reached inside the interval $(-1,1)$.

With only one change of sign, it may be possible that $\theta_n = -1$. For example, when $X_0 \geq 0$ and $X_{t_i^n} < 0$ for $i = 1, \dots, n$. Then $k_\theta(X_{t_i^n}\sqrt{n}, X_{t_{i+1}^n}\sqrt{n}) < 0$ so that $\Lambda_n(\theta)$ is decreasing and $\theta_n = -1$.

Remark 1. The sign of the observations conveys some information, but the MLE also takes into account the probability of crossing.

To avoid technical complications, we rule out the situations where possibly $|\theta_n| = 1$ or $|\widetilde{\theta}| = 1$.

Hypothesis 5. The true parameter $\widetilde{\theta}$ belongs to $(-1,1)$ and $i \mapsto \operatorname{sgn} X_{t_i^n}$ changes its value at least twice.

Theorem 1 (Consistency of the MLE)**.** *Under Hypothesis 5,* θ_n *converges to* $\widetilde{\theta}$ *in probability under* $\mathbb{P}^{\widetilde{\theta}}$.

For $k = 1, 2$, define

$$K_k(x; \theta, \widetilde{\theta}) = \int p_{\widetilde{\theta}}(T, x, y) \left(\frac{\partial_\theta p_\theta(T, x, y)}{p_\theta(T, x, y)} \right)^k \mathrm{d}y,$$

so that, since $n^{1/2} p_\theta(T/n, x/\sqrt{n}, y/\sqrt{n}) = p_\theta(T, x, y)$,

$$K_k(X_{t_i^n}\sqrt{n}; \theta, \widetilde{\theta}) = \mathbb{E}_{\widetilde{\theta}}[k_\theta(X_{t_i^n}, X_{t_{i+1}^n})^k | \mathcal{F}_{t_i^n}].$$

The proof relies on the following technical lemma.

Lemma 5. *Fix $|\widetilde{\theta}| < 1$ and $|\theta| < 1$. Then $K_1(\cdot : \theta, \widetilde{\theta})$, $K_2(\cdot : \theta, \widetilde{\theta})$ and $K_1^2(\cdot : \theta, \widetilde{\theta})$ satisfy Hypothesis 1. In addition,*

$$\frac{1}{\sqrt{n}} S_n(\theta) \xrightarrow[n\to\infty]{} s(\theta, \widetilde{\theta}) L_T$$

in probability under $\mathbb{P}^{\widetilde{\theta}}$ with

$$s(\theta, \widetilde{\theta}) \begin{cases} < 0 & \text{if } \theta > \widetilde{\theta}, \\ = 0 & \text{if } \theta = \widetilde{\theta}, \\ > 0 & \text{if } \theta < \widetilde{\theta}, \end{cases}$$

where the coefficient $s(\theta, \widetilde{\theta})$ is defined in (12) *and* (13).

Proof. To simplify the computations, we set $T = 1$ (the general case may be dealt with a scaling argument). Since $p(1, y-x)/p(1, |x|+|y|) = \exp(2(xy)^+)$,

$$p_{\widetilde{\theta}}(1, x, y) \frac{\partial_\theta p_\theta(1, x, y)}{p_\theta(1, x, y)} = \frac{\exp(2(xy)^+) + \widetilde{\theta} \operatorname{sgn}(y)}{\exp(2(xy)^+) + \theta \operatorname{sgn}(y)} \operatorname{sgn}(y) p(1, |x| + |y|).$$

Thus, for $x \geq 0$,

$$\begin{aligned} K_1(x; \theta, \widetilde{\theta}) &= \int_0^{+\infty} p(1, x+y) \frac{\exp(2xy) + \widetilde{\theta}}{\exp(2xy) + \theta} \, \mathrm{d}y - \int_{-\infty}^0 p(1, x-y) \frac{1 - \widetilde{\theta}}{1 - \theta} \, \mathrm{d}y \\ &= \int_0^{+\infty} p(1, x+y) \frac{\exp(2xy) + \widetilde{\theta}}{\exp(2xy) + \theta} \, \mathrm{d}y - \int_0^{\infty} p(1, x+y) \frac{1 - \widetilde{\theta}}{1 - \theta} \, \mathrm{d}y. \end{aligned}$$

For some constant C depending only on θ and $\widetilde{\theta}$,

$$|K_1(x; \theta, \widetilde{\theta})| \leq C \int_x^{+\infty} p(1, z) \, \mathrm{d}z = C\Psi(x).$$

Similarly, for $x < 0$, after a change of variable

$$K_1(x; \theta, \widetilde{\theta}) = \int_0^{+\infty} p(1, y-x) \frac{1 + \widetilde{\theta}}{1 + \theta} \, \mathrm{d}y - \int_0^{+\infty} p(1, y-x) \frac{\exp(-2xy) - \widetilde{\theta}}{\exp(-2xy) - \theta} \, \mathrm{d}y.$$

For some constant C depending only on θ and $\widetilde{\theta}$,

$$|K_1(x; \theta, \widetilde{\theta})| \leq C \int_{-x}^{+\infty} p(1, z) \, \mathrm{d}z = C\Psi(-x).$$

Thus $|K_1(x;\theta,\widetilde{\theta})| \leq C\Psi(|x|)$. Using for example the Mills ratio, it is easily established that K_1 and K_1^2 satisfy Hypothesis 1. Similar computations show that K_2 satisfies Hypothesis 1.

Let us recall that $S_n(\theta)$ is the score defined by (11). With Proposition 2, $n^{-1/2}S_n(\theta)$ converges in probability, uniformly in θ and $\widetilde{\theta}$ when $\theta,\widetilde{\theta} \in (-1+\kappa, 1+\kappa)$ for $\kappa \in (0,1)$, to $s(\theta,\widetilde{\theta})L_t$ with $s(\theta,\widetilde{\theta}) = \lambda_{\widetilde{\theta}}(K_1(\cdot,\theta,\widetilde{\theta}))$.

Under $\mathbb{P}^{\widetilde{\theta}}$,

$$s(\theta,\widetilde{\theta}) = (1+\widetilde{\theta})\int_0^{+\infty} K_1(x;\theta,\widetilde{\theta})\,\mathrm{d}x + (1-\widetilde{\theta})\int_{-\infty}^0 K_1(x;\theta,\widetilde{\theta})\,\mathrm{d}x.$$

We deduce that

$$s(\theta,\widetilde{\theta}) = \int_0^{+\infty}\int_0^{+\infty} p(1,x+y)\kappa(x,y,\theta,\widetilde{\theta})\,\mathrm{d}x\,\mathrm{d}y \tag{12}$$

with

$$\kappa(x,y,\theta,\widetilde{\theta}) = (1+\widetilde{\theta})\frac{\exp(2xy)+\widetilde{\theta}}{\exp(2xy)+\theta} - (1+\widetilde{\theta})\frac{1-\widetilde{\theta}}{1-\theta} + (1-\widetilde{\theta})\frac{1+\widetilde{\theta}}{1+\theta} - (1-\widetilde{\theta})\frac{\exp(2xy)-\widetilde{\theta}}{\exp(2xy)-\theta}. \tag{13}$$

It is then easily checked that both $\kappa(x,y,\theta,\widetilde{\theta})$ and $\partial_\theta\kappa(x,y,\theta,\widetilde{\theta})$ are bounded in $x,y \geq 0$ and that

$$\kappa(\widetilde{\theta},\widetilde{\theta},x,y) = 0 \text{ and } \partial_\theta\kappa(\widetilde{\theta},\widetilde{\theta},x,y) < 0.$$

Then $\partial_\theta s(\theta,\widetilde{\theta}) < 0$. In particular, $\partial_\theta s(\widetilde{\theta},\widetilde{\theta}) < 0$, which proves the result. □

Proof of Theorem 1. For any $\epsilon > 0$ small enough, $n^{-1/2}S_n(\widetilde{\theta}\pm\epsilon)$ converges in probability under $\mathbb{P}^{\widetilde{\theta}}$ to $s(\widetilde{\theta}\pm\epsilon,\widetilde{\theta})L_T$, and $\operatorname{sgn} s(\widetilde{\theta}\pm\epsilon,\widetilde{\theta}) = \mp 1$ when $L_T > 0$. Since θ_n is solution to $S_n(\theta_n) = 0$, $\theta_n \in [\widetilde{\theta}-\epsilon,\widetilde{\theta}+\epsilon]$ on a set whose probability increases to $\mathbb{P}[L_T > 0]$ as $n \to \infty$. □

5. An alternative expression for the Maximum Likelihood Estimator

The parameter θ depends essentially on what happens when the process is around 0. We have first proposed an estimator which clearly compares the number of observed upward transitions to the number of observed crossing. Following the results from D. Florens [7] and J. Jacod [12] for diffusions, the number of observed crossings is related to the local time, a result we have seen in Section 4.1.

It is less clear why the asymptotic behavior of the MLE is also related to the local time. We present here an alternative expression for the MLE which is based on the idea of the Expectation-Maximization (EM) algorithm [6, 22]. We point out that there is no need to use the EM algorithm for numerical purposes. Here, we only use the core idea of the EM algorithm which relies on the use of *latent* or *hidden* variables.

Since n is fixed, we set $X_i = X_{t_i^n}$. We define

$$Z_i = \begin{cases} 1 & \text{if } X(t) = 0 \text{ for some } t \in [t_i^n, t_{i+1}^n], \\ 0 & \text{otherwise.} \end{cases}$$

Of course, the Z_i's cannot be observed and serve as the latent variables. We set $\mathbf{X} = (X_0, \dots, X_n)$, $\mathbf{Z} = (Z_1, \dots, Z_n)$, $x = (x_0, \dots, x_n)$ and $z = (z_0, \dots, z_n)$. For $\tau, \theta \in (-1, 1)$, we define

$$Q_n(\tau, \theta) = \mathbb{E}^\theta \left[\log p_\tau(\mathbf{X}, \mathbf{Z}) | \mathbf{X} = x \right].$$

According to the fundamental principles of the EM algorithm [6], if $\theta = \operatorname{argmax}_\tau Q(\tau, \theta)$, then θ is a stationary point for the likelihood.

Proposition 4 ([6]). *The MLE θ_n is a solution to $\partial_\tau Q_n(\tau, \theta_n)_{|\tau=\theta_n} = 0$.*

The main result of this section is the following one. Let us denote by U the number of times $z_i = 1$ and $x_i \geq 0$, by D the number of times $z_i = 1$ and $x_i < 0$, and define $N = U + D$, the number of crossings.

The next proposition is the central result of this section. The EM algorithm is not used as a numerical tool, as a root finding algorithm is sufficient to evaluate θ_n from the observations. We use it to relate the MLE to the latent variables which are the Z_i's through Proposition 4.

Proposition 5. *The MLE θ_n is solution to*

$$\frac{1 + \theta_n}{2} = \frac{\mathbb{E}^{\theta_n}[U | \mathbf{X} = x]}{\mathbb{E}^{\theta_n}[N | \mathbf{X} = x]}. \tag{14}$$

The interpretation of this result is very natural: $(1 + \theta_n)/2$, the estimated probability that an excursion of the SBm, is given by the ratio of expected upward excursions straddling the t_i^n's on the expectation excursions straddling the t_i^n's, when the expectation is considered under $\mathbb{P}^{\theta_n}$ given the observations. It is then a generalization of the result obtained in [18].

Proof. Let $p_\tau(x,z)$ be the density of $(\mathbf{X},\mathbf{Z})$, $p_\tau(z)$ the density of $\mathbf{Z}$ and $p_\tau(x|z)$ the density of $\mathbf{X}$ given $\mathbf{Z}$. From the Markov property,

$$p_\tau(x,z)=\prod_{i=1}^{n}p_\tau(x_i,z_i|x_{i-1}).$$

From the construction of the density of the SBm proposed by J.B. Walsh [28], in which the independence of the sign and the position of X_{i+1} given $Z_i=1$ is used,

$$p_\tau(x_i,z_i=1|x_{i-1})=\begin{cases}(1+\tau)p(\Delta t,|x_i|+|x_{i-1}|) & \text{if } x_i\geq 0,\\ (1-\tau)p(\Delta t,|x_i|+|x_{i-1}|) & \text{if } x_i<0,\end{cases}$$

while

$$p_\tau(x_i,z_i=0|x_{i-1})=\begin{cases}p(\Delta t,|x_{i-1}|,|x_i|)-p(\Delta t,|x_{i-1}|,-|x_i|) & \text{if } x_ix_{i-1}>0,\\ 0 & \text{if } x_ix_{i-1}\leq 0.\end{cases}$$

Thus,

$$\begin{aligned}&\mathbb{E}^\theta[\log p_\tau(x,z)|\mathbf{X}=x]\\ &\qquad=\mathbb{E}^\tau[U|\mathbf{X}=x]\log(1+\tau)+\mathbb{E}^\tau[D|\mathbf{X}=x]\log(1-\tau)+C(x),\end{aligned}$$

where $C(x)$ is a function that does not depend on τ. Maximizing over τ yields that the fixed point of $\theta=\operatorname{argmax}_{\tau\in(-1,1)}Q_n(\tau,\theta)$ is

$$\frac{1+\theta}{2}=\frac{\mathbb{E}^\theta[U|\mathbf{X}=x]}{\mathbb{E}^\theta[N|\mathbf{X}=x]}.$$

According to Proposition 4, the parameter θ is nothing more that the MLE. □

Unfortunately, the equation (14) is not easy to deal with and does not simplify the computations. An iteration method — either for a fixed point or to solve an optimization problem — has to be used.

Proposition 6. *The MLE θ_n is solution to the equation*

$$\frac{1+\theta_n}{2}\left(\sum_{i=1}^{n}\psi_i\right)=\sum_{i=1}^{n}\left(\mathbf{1}_{x_{i-1}<0,x_i>0}+\psi_i\frac{1+\theta_n}{1+\psi_i\theta_n}\mathbf{1}_{x_{i-1}>0,x_i>0}\right),$$

with $x_i=X_{t_i^n}$ and $\psi_i=\exp(-2(x_ix_{i+1})^+/\Delta t)$.

Remark 2. Since $\psi_i = 1$ when $x_{i-1}x_i > 0$, the MLE is in some sense close to $\widehat{\theta}_n$ constructed by

$$\frac{1+\widehat{\theta}_n}{2} = \frac{\sum_{i=1}^n \mathbf{1}_{x_{i-1}<0, x_i>0}}{\sum_{i=1}^n \mathbf{1}_{x_{i-1}x_i<0}}$$

seen in Section 4.1. Yet it also takes into account the probability that the path has crossed 0 when x_{i-1} and x_i have the same sign. More precisely,

$$(\theta_n - \widehat{\theta}_n)\sum_{i=1}^n \mathbf{1}_{x_{i-1}x_i<0} = (1+\theta_n)\sum_{i=1}^n \psi_i \left(\frac{1-\psi_i\theta_n}{1+\psi_i\theta_n}\mathbf{1}_{x_{i-1}>0, x_i>0} - \mathbf{1}_{x_{i-1}<0, x_i<0}\right),$$

and both θ_n and $\widehat{\theta}_n$ have the same limit in probability.

Proof. Conditionally to $\mathbf{X} = x$, the Markov property implies that the distribution of Z_i depends only on the values of (x_i, x_{i+1}) and is independent from Z_j for $j \neq i$. Hence,

$$\mathbb{E}^\theta[P|\mathbf{X} = x] = \sum_{i=1}^n \mathbb{P}^\theta[Z_i = 1|X_{i-1} = x_{i-1}, X_i = x_i]\mathbf{1}_{x_i>0},$$
$$\mathbb{E}^\theta[N|\mathbf{X} = x] = \sum_{i=1}^n \mathbb{P}^\theta[Z_i = 1|X_{i-1} = x_{i-1}, X_i = x_i].$$

Clearly,

$$\mathbb{P}^\theta[Z_i = 1|X_{i-1} = x_{i-1}, X_i = x_i] = \exp(-2(x_i x_{i+1})^+/\Delta t),$$

as the probability that X_t crosses 0 given X_i and X_{i+1} is $\exp(-2(x_i x_{i+1})^+/\Delta t)$ when $x_i x_{i-1} > 0$, and 1 as $x_i x_{i-1} < 0$. For $x_i, x_{i+1} > 0$, we have

$$\mathbb{P}^\theta[Z_i = 1|X_{i-1} = x_{i-1}, X_i = x_i] = \frac{p_\theta(x_i, z_i = 1|x_{i-1})}{p_\theta(x_{i-1}, x_i)} = \frac{1+\theta}{\exp\left(\frac{x_i x_{i+1}}{\Delta t}\right) + \theta},$$

while for $x_i < 0$ and $x_{i+1} > 0$, $\mathbb{P}^\theta[Z_i = 1|X_{i-1} = x_{i-1}, X_i = x_i] = 1$. The result is obtained by combining these results. □

Acknowledgement. This work is supported by the Inria's Équipe Associée ANESTOC-TOSCA between France and Chile. S. Torres is supported by the grant FONDECYT 1130586.

References

[1] L. H. R. Alvarez & P. Salminen. Timing in the Presence of Directional Predictability: Optimal Stopping of Skew Brownian Motion (2016). Preprint arxiv:1608.04537.

[2] T. Appuhamillage, V. Bokil, E. Thomann, E. Waymire & B. D. Wood. Occupation and local times for Skew Brownian motion with application to dispersion accross an interface. *Ann. Appl. Probab.* **21**.1 (2011), 183–214. DOI: 10.1214/10-AAP691.

[3] O. Bardou & M. Martinez. Statistical estimation for reflected skew processes. *Stat. Inference Stoch. Process.* **13**.3 (2010), 231–248. DOI: 10.1007/s11203-010-9047-6.

[4] M. Bossy, N. Champagnat, S. Maire & D. Talay. Probabilistic interpretation and random walk on spheres algorithms for the Poisson-Boltzmann equation in molecular dynamics. *M2AN Math. Model. Numer. Anal.* **44**.5 (2010), 997–1048. DOI: 10.1051/m2an/2010050.

[5] M. Decamps, M. Goovaerts & W. Schoutens. Self exciting threshold interest rates models. *Int. J. Theor. Appl. Finance* **9**.7 (2006), 1093–1122. DOI: 10.1142/S0219024906003937.

[6] A. P. Dempster, N. M. Laird & D. B. Rubin. Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)* **39**.1 (Jan. 1977), 1–38.

[7] D. Florens. Estimation of the diffusion coefficient from crossings. *Stat. Inference Stoch. Process.* **1**.2 (1998), 175–195. DOI: 10.1023/A:1009927813898.

[8] J. M. Harrison & L. A. Shepp. On skew Brownian motion. *Ann. Probab.* **9**.2 (1981), 309–313.

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

[10] K. Itô & H. P. McKean Jr. *Diffusion processes and their sample paths.* 2nd ed. Springer-Verlag, Berlin-New York, 1974.

[11] J. Jacod. Une généralisation des semimartingales: les processus admettant un processus à accroissements indépendants tangent. In: *Seminaire de probabilités XVIII.* Vol. 1059. Lecture Notes in Math. Berlin: Springer, 1984, 91–118. DOI: 10.1007/BFb0100035.

[12] J. Jacod. Rates of convergence to the local time of a diffusion. *Ann. Inst. H. Poincaré Probab. Statist.* **34**.4 (1998), 505–544. DOI: 10.1016/S0246-0203(98)80026-5.

[13] J. Keilson & J. A. Wellner. Oscillating Brownian motion. *J. Appl. Probability* **15**.2 (1978), 300–310.

[14] J.-F. Le Gall. One-dimensional stochastic differential equations involving the local times of the unknown process. In: *Stochastic analysis and applications, Swansea, 1983*. Vol. 1095. Lecture Notes in Math. Springer, 1984, 51–82.

[15] J.-F. Le Gall. One-Dimensional Stochastic Differential Equations Involving the Local Times of the Unknown Process. In: *Stochastic Analysis and Applications*. Vol. 1095. Lecture Notes in Mathematics. Springer Verlag, 1985, 51–82.

[16] A. Lejay & P. Pigato. Statistical estimation of the Oscillating Brownian Motion (2017). Preprint `arxiv:1701.02129`.

[17] A. Lejay. On the constructions of the skew Brownian motion. *Probab. Surv.* **3** (2006), 413–466. DOI: `10.1214/154957807000000013`.

[18] A. Lejay. Estimation of the biais parameter of the Skew Random Walk and application to the Skew Brownian Motion. *Statistical Inference for Stochastic Processes* (2017). DOI: `10.1007/s11203-017-9161-9`.

[19] A. Lejay & G. Pichot. Simulating Diffusion Processes in Discontinuous Media: Benchmark Tests. *J. Comput. Phys.* **314** (2016), 348–413. DOI: `10.1016/j.jcp.2016.03.003`.

[20] A. Lejay, E. Mordecki & S. Torres. Is a Brownian motion skew? *Scand. J. Stat.* **41**.2 (2014), 346–364. DOI: `10.1111/sjos.12033`.

[21] M. Martinez. Interprétations probabilistes d'opérateurs sous forme divergence et analyse de méthodes numériques associées. Ph.D. thesis. Université de Provence / INRIA Sophia-Antipolis, 2004.

[22] G. J. McLachlan & T. Krishnan. *The EM algorithm and extensions*. 2nd ed. Wiley Series in Probability and Statistics. Wiley-Interscience, 2008. DOI: `10.1002/9780470191613`.

[23] O. Ovaskainen & S. J. Cornell. Biased movement at a boundary and conditional occupancy times for diffusion processes. *J. Appl. Probab.* **40**.3 (2003), 557–580.

[24] N. I. Portenko. Diffusion processes with a generalized drift coefficient. *Teor. Veroyatnost. i Primenen.* **24**.1 (1979), 62–77.

[25] D. Rossello. Arbitrage in skew Brownian motion models. *Insurance Math. Econom.* **50**.1 (2012), 50–56. DOI: `10.1016/j.insmatheco.2011.10.004`.

[26] D. Spivakovsakaya, A. Heemink & E. Deleersnijder. The backward Ito method for the Lagrangian simulation of transport processes with large space variations of the diffusivity. *Ocean Sci.* **3**.4 (2007), 525–535.

[27] D. Thomson, W. Physick & R. Maryon. Treatment of Interfaces in Random Walk Dispersion Models. *J. Appl. Meteorol.* **36** (1997), 1284–1295.

[28] J. Walsh. A diffusion with discontinuous local time. In: *Temps locaux.* Vol. 52–53. Astérisques. Société Mathématique de France, 1978, 37–45.

[29] M. Zhang. Calculation of diffusive shock acceleration of charged particles by skew Brownian motion. *The Astrophysical Journal* **541** (2000), 428–435.