ESAIM: PS 23 (2019) 567–583
https://doi.org/10.1051/ps/2018018

ESAIM: Probability and Statistics
www.esaim-ps.org

TWO CONSISTENT ESTIMATORS FOR THE SKEW BROWNIAN MOTION

Antoine Lejay[1,2,3,*], Ernesto Mordecki[4] and Soledad Torres[5]

Abstract. The skew Brownian motion (SBm) 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 procedure estimates consistently the parameter of an SBm observed at discrete times. The difficulties arise because the observed 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 total number of positive excursions divided by the expected number of excursions given the observations. The theoretical results are illustrated by numerical simulations.

Mathematics Subject Classification. 60H10, 62F12.

Received May 19, 2017. Accepted July 18, 2018.

1. Introduction

Introduced in the 70s as a variation of the Brownian motion [8, 10, 27], the skew Brownian motion (SBm) has attracted a lot of interest as a "basic 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 (SDEs) with local times [14, 15]. The need for such models arises in many domains (see [16, 19] for more references): population ecology [26], finance [1, 5, 28], geophysics [2], electro-encephalography [24], molecular chemistry [4], meteorology [30], oceanography [29], and astrophysics [32].

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

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

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 is a parameter in $[-1, 1]$.

Keywords and phrases: Skew Brownian motion, maximum likelihood estimator (MLE) null recurrent process, expectation–maximization (EM) algorithm, excursion theory.

[1] Université de Lorraine, IECL, UMR 7502, Vandœuvre-lès-Nancy, 54600, France.
[2] CNRS, IECL, UMR 7502, Vandœuvre-lés-Nancy, 54600, France.
[3] Inria, Villers-lés-Nancy, 54600, France.
[4] Facultad de Ciencias, Centro de Matemática. Iguá 4225, CP 11400, Montevideo, Uruguay.
[5] CIMFAV, Facultad de Ingeniería, Universidad de Valparaíso, Valparaíso, Chile.

* Corresponding author: antoine.lejay@inria.fr

Article published by EDP Sciences

© EDP Sciences, SMAI 2019

When $\theta = 1$ (resp. $\theta = -1$) the process X is a positively (resp. negatively) reflected Brownian motion. When $\theta = 0$ the process 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 [16] 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 an 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 [22], the occupation times of positive or negative axis are not convergent estimators: their variances do not decrease to 0 whatever the number of observations or the time horizon.

The density transition function of the SBm can be expressed in closed form [16, 31] by

$$p_\theta(t,x,y) := p(t,x-y) + \operatorname{sgn}(y)\theta p(t,|x|+|y|), \tag{1.1}$$

where

$$p(t,x) := (2\pi t)^{-1/2}\exp(-x^2/2t) \tag{1.2}$$

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 computed through standard numerical optimization algorithms. This article is devoted to show the consistency of the MLE, a point left open in [22].

More precisely, in [22], we have studied the asymptotic behavior of the MLE θ_n when the true distribution is the one of the Brownian motion, the case $\theta = 0$, proving its consistency. In this paper, 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.

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 the present 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. Numerical experiments are satisfactory.

To prove this consistency, we adapt some results of [12] to the case of an 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 seems really intricate. Numerical experiments back this conjecture.

In [17], 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 number of upward excursions to the total number of the excursions. The analysis can be performed with elementary tools due to the simplicity of the random walk setting.

Using the relationship with the expectation–maximization (EM) algorithm [6, 25], we show that the MLE is linked to the expected number of upward and downward (unknown) excursions given the observations. This shows that the situation between the biased random walk and the one of the SBm is 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 [20], an explicit estimation of the volatilities is provided, for an *oscillating Brownian motion* [13], a solution to an SDE whose diffusion coefficients takes two values according to the sign of its position. For this, a reduction to the SBm through a simple transform is used, and some of the results given here are related to the convergence of these estimators. The problem of estimating a piecewise constant drift with a

piecewise constant volatility is developed. In [21], the two above works are applied on financial data to detect leverage and mean-reverting effects. We then believe that the present work could be applied to deal with other problems related to the estimation of processes with discontinuous coefficients, a subject which is largely open.

Outline. In Section 2, we extend some of the results of Jacod [12] to the distribution of the SBm. 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 counting the (unknown) number of excursions straddling the sample's times. Finally, we give some numerical examples in Section 6.

2. Some convergence results

We denote by $p(t,x)$ the Gaussian density with mean 0 and variance t as in (1.2). For a measurable, bounded function f, set $P_t f(x) := \int p(t,y-x)f(y)\,\mathrm{d}y$.

Set $f_\oplus(x) := f(x)\mathbf{1}_{\{x\geq 0\}}$ and $f_\ominus(x) = f(x)\mathbf{1}_{\{x<0\}}$.

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

$$P_t^\theta f(x) := \int_{-\infty}^{+\infty} p_\theta(t,x,y)f(y)\,\mathrm{d}y = P_t f(x) + \theta P_t f_\oplus(-|x|) - \theta P_t f_\ominus(|x|), \tag{2.1}$$

where p_θ is given by (1.1). For $f\colon \mathbb{R}\to\mathbb{R}$ and $\gamma\geq 0$, we define

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

For any $\theta\in(-1,1)$, $\gamma\geq 0$,

$$\beta_\gamma^\theta(f) \leq 2\beta_\gamma(f). \tag{2.2}$$

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

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

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], Lem. 3.1, p. 518) the following result which is the key to identify the limit.

Lemma 2.1. *If $\beta_\gamma(f)<\infty$ for $\gamma=0,1,2$, then for some constant K,*

$$|P_t^\theta f(x)| \leq \frac{\sqrt{2}}{\sqrt{\pi t}}\beta_0(f), \tag{2.3}$$

$$|P_t^\theta f(x) - \lambda_\theta(f)p(t,x)| \leq \frac{K}{t^{3/2}}(\beta_2^\theta(f) + \beta_1^\theta(f)|x|). \tag{2.4}$$

Proof. In this proof, the constant K may vary from line to line. Inequality (2.3) is a direct consequence of $p_\theta(t,x,y) \le (1+|\theta|)/\sqrt{2\pi t}$. Using (2.1) and (3.3) in Lemma 3.1 from [12],

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

This shows (2.4) and concludes the proof of the lemma. □

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. Let us consider $\theta \in [-1,1]$. With the result of [8], the SDE with local time $X_t = x + B_t + \theta L_t$, $t \ge 0$ has a unique strong solution. This process is the SBm of parameter θ. We denote its distribution by $\mathbb{P}^\theta$.

Hypothesis 2.2. The function f is bounded and $\beta_\gamma(f) < +\infty$ for $\gamma = 0,1,2$.

Lemma 2.3. *Let f satisfy Hypothesis 2.2 and $\lambda_\theta(f) = 0$ for some $\theta \in [-1,1]$. Then, for some constant K which depends on T,*

$$\mathbb{E}^\theta\left[\left|\frac{1}{\sqrt{n}}\sum_{i=0}^{\lfloor nt/T\rfloor} f(X_{t_i^n}\sqrt{n})\right|^2\right] \le \frac{K}{\sqrt{n}}\big(\beta_0(f)\sup|f| + \beta_0(f)\beta_2(f) + \beta_1(f)^2\big). \tag{2.5}$$

Proof. As the SBm presents the same scaling property as the Brownian motion, we set for the sake of simplicity $T = 1$. By scaling, this changes the value of the constant K in (2.5).

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 $i > 1$, using (2.4) with $\lambda_\theta(f_n) = 0$, the inequality (2.2), and taking into account that $\sum_{i=1}^n i^{-3/2} \le \zeta(3/2) \le 3$,

$$\begin{aligned}
\left|\sum_{j=i+1}^n \mathbb{E}^\theta[f(X_{t_j^n}\sqrt{n})|X_{t_i^n} = x]\right| &\le \sum_{j=i+1}^n |P^\theta_{(j-i)/n} f_n(x)| \\
&\le \sum_{j=i+1}^n \frac{2K}{(j-i)^{3/2}}(\beta_2(f) + \beta_1(f)\sqrt{n}|x|) \\
&\le K'(\beta_2(f) + \beta_1(f)\sqrt{n}|x|) \text{ with } K' = 6K.
\end{aligned}$$

For $0 \le t \le 1$, based on the Markov property, inequality (2.3), the scaling property of β_0 and $\sum_{i=1}^n 1/\sqrt{i} \le 2\sqrt{n}$,

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

The right-hand side is finite for each n since $\beta_0(f(x)|x|) = \beta_1(f)$. Using again this equality,

$$\beta_0(|f(x)|(\beta_2(f)+\beta_1(f)|x|)) = \beta_0(f)\beta_2(f)+\beta_1(f)^2.$$

Applying (2.3),

$$\begin{aligned}\mathbb{E}^\theta\left[\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{\lfloor nt\rfloor}f(X_{t_i^n}\sqrt{n})\right)^2\right] &\le \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}^\theta[f(X_{t_i^n}\sqrt{n})^2] + \frac{2}{n}\mathbb{E}^\theta\left[\sum_{i=1}^{\lfloor nt\rfloor}\sum_{j=i+1}^{\lfloor nt\rfloor}f(X_{t_i^n}\sqrt{n})f(X_{t_j^n}\sqrt{n})\right]\\ &\le \frac{\sqrt{2}}{\sqrt{\pi}}\beta_0(f^2)\frac{K}{n^{3/2}}\sum_{i=1}^{n}\frac{\sqrt{n}}{\sqrt{i}} + \frac{2K''\sqrt{n}}{n}(\beta_0(f)\beta_2(f)+\beta_1(f)^2)\\ &\le \beta_0(f^2)\frac{2\sqrt{2}}{\sqrt{n\pi}} + \frac{2K''}{\sqrt{n}}(\beta_0(f)\beta_2(f)+\beta_1(f)^2).\end{aligned}$$

Since f is bounded, $\beta_0(f^2) \le \beta_0(f)\sup_{x\in\mathbb{R}}|f(x)|$. This proves the result. □

Proposition 2.4. *Let f be a function satisfying Hypothesis 2.2. Then for any $a > 0$,*

$$\sup_{\theta\in[-1,1]}\mathbb{P}^\theta\left[\sup_{t\in[0,T]}\left|\frac{1}{\sqrt{n}}\sum_{i=0}^{\lfloor nt/T\rfloor-1}f(X_{t_i^n}\sqrt{n}) - \lambda_\theta(f)L_t\right| > a\right] \xrightarrow[n\to\infty]{} 0. \tag{2.6}$$

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

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

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 (Lem. 2.14 of [11]), since the distribution of the local time does not depend on θ, for any $a > 0$,

$$\sup_{\theta\in[-1,1]} \mathbb{P}^\theta \left[\sup_{t\in[0,T]} \left| \sum_{i=0}^{\lfloor nt/T\rfloor -1} \mathbb{E}^\theta[L_{t_{i+1}^n} - L_{t_i^n} | \mathcal{F}_{t_i^n}] - L_t \right| \geq a \right] \xrightarrow[n\to\infty]{} 0. \tag{2.7}$$

Set $h(x) := f(x) - \lambda_\theta(f)g(x)$, so that $\lambda_\theta(h) = 0$. Using Lemma 2.3 on h and (2.7), one obtains that for any $t \in [0,1]$,

$$\sup_{\theta\in[-1,1]} \mathbb{P}^\theta \left[|Y_n(t) - \lambda_\theta(f)L_t| > a \right] \xrightarrow[n\to\infty]{} 0, \tag{2.8}$$

with

$$Y_n(t) := \frac{1}{\sqrt{n}} \sum_{i=0}^{\lfloor nt/T\rfloor -1} f(X_{t_i^n}\sqrt{n}).$$

Assume now that f is non-negative, so that Y_n is non-decreasing as a function of t. Therefore, for a partition $\{t_i\}_{i=0,\dots,m}$ of $[0,T]$

$$\begin{aligned}
\mathbb{P}^\theta \left[\sup_{t\in[0,T]} |Y_n(t) - \lambda_\theta(f)L_t| > a \right] &\leq \mathbb{P}^\theta \left[\sup_{i=0,\dots,m-1} \sup_{t\in[t_i,t_{i+1}]} |Y_n(t) - \lambda_\theta(f)L_t| > a \right] \\
&\leq \sum_{i=0}^{m-1} \mathbb{P}^\theta \left[\sup_{t\in[t_i,t_{i+1}]} |Y_n(t) - Y_n(t_i) - \lambda_\theta(f)(L_t - L_{t_i})| > \frac{a}{2} \right] \\
&\quad + \sum_{i=0}^{m-1} \mathbb{P}^\theta \left[|Y_n(t_i) - \lambda_\theta(f)L_{t_i}| > \frac{a}{2} \right] \\
&\leq \sum_{i=0}^{m-1} \mathbb{P}^\theta \left[\sup_{t\in[t_i,t_{i+1}]} |Y_n(t) - Y_n(t_i)| > \frac{a}{4} \right] \\
&\quad + \sum_{i=0}^{m-1} \mathbb{P}^\theta \left[\sup_{t\in[t_i,t_{i+1}]} |\lambda_\theta(f)(L_t - L_{t_i})| > \frac{a}{4} \right] \\
&\quad + \sum_{i=0}^{m-1} \mathbb{P}^\theta \left[|Y_n(t_i) - \lambda_\theta(f)L_{t_i}| > \frac{a}{2} \right].
\end{aligned}$$

Since Y_n is non-decreasing, for $t_i \leq t \leq t_{i+1}$,

$$|Y_n(t) - Y_n(t_i)| \leq |Y_n(t_{i+1}) - Y_n(t_i)| \leq |Y_n(t_{i+1}) - \lambda_\theta(f)L_{t_{i+1}}| + |Y_n(t_i) - \lambda_\theta(f)L_{t_i}| + \lambda_\theta(f)|L_{t_{i+1}} - L_{t_i}|.$$

Recall that $\lambda_\theta(f) \leq 2\beta_0(f)$. Since the local time L is continuous and does not depend on θ, for any $\epsilon > 0$ and any $a > 0$, there exists n_0 large enough and $\delta > 0$ such that

$$\sup_{\theta\in[-1,1]} \mathbb{P}^\theta \left[\sup_{|t-s|\leq\delta} |L_t - L_s| > a \right] \leq \epsilon, \ \forall n \geq n_0. \tag{2.9}$$

Combining (2.9) with (2.8) applied to the times t_i, we then obtain (2.6).

For a general function f, it is sufficient to apply (2.6) to its positive and its negative part. □

Remark 2.5. Let $\{f_\sigma\}_{\sigma\in S}$ be a parametric family of functions that satisfy Hypothesis 2.2 with $\sup_{\sigma\in S}\beta_\gamma(f_\sigma) < +\infty$. As the $\mathrm{L}^2(\mathbb{P}^\theta)$-convergence in Lemma 2.3 is uniform in the σ, it is easily seen that the convergence in probability of $\frac{1}{\sqrt{n}}\sum_{i=0}^{\lfloor nt/T\rfloor-1} f_\sigma(X_{t_i^n}\sqrt{n})$ toward $\lambda_\theta(f_\sigma)L_t$ is also uniform in σ.

Hypothesis 2.6. Let $f:\mathbb{R}^2\to\mathbb{R}$ be a function such that for

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

the functions F_1, $(F_1)^2$ and F_2 satisfy Hypothesis 2.2.

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

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

a relation we have already used in the proof of Proposition 2.4.

Proposition 2.7. *Let f be a function satisfying Hypothesis 2.6. Then, for any $a>0$,*

$$\sup_{\theta\in[-1,1]}\mathbb{P}^\theta\left[\sup_{t\in[0,T]}\left|\frac{1}{\sqrt{n}}\sum_{i=0}^{\lfloor nt/T\rfloor-1} f(X_{t_i^n}\sqrt{n},X_{t_{i+1}^n}\sqrt{n}) - \lambda_\theta(F_1)L_t\right| > a\right] \xrightarrow[n\to\infty]{} 0. \tag{2.10}$$

Proof. Applying Doob's inequality

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

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

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

Hence, this quantity converges in probability to 0 as $n\to\infty$ by Proposition 2.4 uniformly in $\theta\in[-1,1]$. One obtains that (2.10) holds. □

Remark 2.8. Following Remark 2.5, for a parametric family of functions $\{f_\sigma\}_{\sigma\in S}$ that satisfies Hypothesis 2.6 uniformly in $\sigma\in S$, the convergence in probability in Proposition 2.7 is also uniform in σ.

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 3.1. Let f and g be functions satisfying Hypothesis 2.6 with $g \geq 0$ and $\lambda_\theta(G_1) > 0$, where $G_1 := \int_{\mathbb{R}} p_\theta(1,x,y) g(x,y)^\gamma \, \mathrm{d}y$.

Hypothesis 3.2. The starting point $X_0 = x$ is 0.

When $X_0 \neq 0$, we could wait an observable time τ_0 where τ_0 is the first hitting time of 0. The observation window is then reduced from $[0,T]$ to the random one $[\tau_0, T]$. Under Hypothesis 3.2, $\tau_0 = 0$ and $L_T > 0$ almost surely for any $T > 0$. Otherwise, $L_T = 0$ on $\{\tau < T\}$. Hypothesis 3.2 is there for the sake of simplicity.

Let us consider the asymptotic behavior of

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

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

Corollary 3.3. *Under Hypotheses 3.1 and 3.2,*

$$R_n(f,g) \xrightarrow[n\to\infty]{proba} \frac{\lambda_\theta(F_1)}{\lambda_\theta(G_1)} \text{ under } \mathbb{P}^\theta$$

for any $\theta \in [-1,1]$.

Proof. According to Proposition 2.7,

$$\frac{1}{\sqrt{n}} Y_n \xrightarrow[n\to\infty]{\text{proba}} \lambda_\theta(F_1) L_T \text{ and } \frac{1}{\sqrt{n}} Z_n \xrightarrow[n\to\infty]{\text{proba}} \lambda_\theta(G_1) L_T.$$

Therefore, $n^{-1/2}(Y_n, Z_n)$ converges in probability to $(\lambda_\theta(F_1)L_T, \lambda_\theta(G_1)L_T)$ under $\mathbb{P}^\theta$. The result follows from standard computations and the fact that almost surely, $0 < L_T < +\infty$ and $\lambda_\theta(G_1) > 0$ under Hypothesis 3.2. □

4. Estimation of the parameter of the SBm

We assume that we observe the discretization $\{X_{t_i^n}\}_{i=0,\dots,n}$ at discrete time steps $t_i^n = iT/n$ of a path of an SBm with an unknown parameter $\widetilde{\theta}$. We are willing to estimate the local time L_T as well as $\widetilde{\theta}$.

4.1. An estimator of the local time based on the number of crossings

Our first estimator records the upward passages of the discrete observations among their passage to 0. This estimator converges to the local time.

Proposition 4.1. *Under the probability* $\mathbb{P}^{\widetilde{\theta}}$ *when* $\widetilde{\theta} \in (-1,1)$,

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

and

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

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 $k = 1, 2$, for $x < 0$,

$$F_1(x) = F_2(x) = \int_0^{+\infty} p_{\widetilde{\theta}}(T/n, x, y)\,\mathrm{d}y = (1+\widetilde{\theta})\Psi(|x|),$$

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 2.7 applies, so that under $\mathbb{P}^{\widetilde{\theta}}$,

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

Hence the result after a straightforward computation. □

Remark 4.2. The ratio estimator $\mathbf{1}_{C_n}\sum_{i=0}^{n-1}\mathbf{1}_{\{X_{t_i^n}<0, X_{t_{i+1}^n}>0\}} / \sum_{i=0}^{n-1}\mathbf{1}_{\{X_{t_i^n}X_{t_{i+1}^n}<0\}}$ converges to $1/2\mathbf{1}_{L_T>0}$. Since upward and downward crossings alternate, this ratio estimator is meaningless to estimate θ. However, the MLE estimator may be thought as a variant of this estimator by weighting the upward and the downward crossings.

4.2. Consistency of the maximum likelihood estimator

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

$$\Lambda_n(\theta) := \prod_{i=0}^{n-1} p_\theta(\Delta t, X_{t_i^n}, X_{t_{i+1}^n}). \tag{4.1}$$

Definition 4.3. The *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 4.4. *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

$$S_n(\theta) := \partial_\theta \log \Lambda_n(\theta) = \sum_{i=0}^{n-1} \frac{\partial_\theta p_\theta(\Delta t, X_{t_i^n}, X_{t_{i+1}^n})}{p_\theta(\Delta t, X_{t_i^n}, X_{t_{i+1}^n})} = \sum_{i=0}^{n-1} k_\theta\left(X_{t_i^n}\sqrt{n}, X_{t_{i+1}^n}\sqrt{n}\right), \tag{4.2}$$

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 (4.1) also maximizes $\log \Lambda_n$. If $\theta_n \in (-1,1)$, this point is also characterized as a solution to $S_n(\theta) = 0$. The next result shows that θ_n is uniquely defined.

Lemma 4.5. *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 values 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 4.6. The signs of the observations convey most of the information, but the MLE also takes into account the probability of crossing. This explains the difference of performance between the MLE and the estimator based on ratio observed in the numerical simulations of Section 6.

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

Hypothesis 4.7. 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 4.8 (Consistency of the MLE)**.** *Under Hypothesis 4.7, the MLE* 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 4.9. *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 2.2. 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 (4.3) and (4.4) below.

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 2.2 uniformly in $\theta \in (-1 + \epsilon, 1 - \epsilon)$ for some $\epsilon \in (0, 1)$. Similar computations show that K_2 satisfies Hypothesis 2.2 also uniformly in $\theta \in (-1 + \epsilon, 1 - \epsilon)$.

Let us recall that $S_n(\theta)$ is the score defined by (4.2). We set $s(\theta,\widetilde{\theta}) := \lambda_{\widetilde{\theta}}(K_1(\cdot,\theta,\widetilde{\theta}))$.

With Proposition 2.7 and Remark 2.8, $n^{-1/2}S_n(\theta)$ converges to $s(\theta,\widetilde{\theta})L_t$ in probability uniformly in θ and $\widetilde{\theta}$ when $\theta,\widetilde{\theta} \in (-1+\epsilon, 1-\epsilon)$ for $\epsilon \in (0,1)$.

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

$$s(\theta,\widetilde{\theta}) = \int_0^{+\infty}\int_0^{+\infty} p(1,x+y)\kappa(x,y,\theta,\widetilde{\theta})\,\mathrm{d}x\,\mathrm{d}y \tag{4.3}$$

with

$$\kappa(x,y,\theta,\widetilde{\theta}) = (1+\widetilde{\theta})\frac{\exp(2xy)+\widetilde{\theta}}{\exp(2xy)+\theta} - (1+\widetilde{\theta})\frac{1-\widetilde{\theta}}{1-\theta} + (1-\widetilde{\theta})\frac{1+\widetilde{\theta}}{1+\theta} - (1-\widetilde{\theta})\frac{\exp(2xy)-\widetilde{\theta}}{\exp(2xy)-\theta}. \tag{4.4}$$

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 4.8. 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 following the results from Florens [7] and 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 EM algorithm [6, 25]. 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 $X_i = X_{t_i^n}$, and $X_0 = x_0$. We define

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

Of course, the Z_i's cannot be observed and serve as the latent variables. We set $\mathbf{X} = (X_1,\dots,X_n)$, $\mathbf{Z} = (Z_1,\dots,Z_n)$, $x = (x_1,\dots,x_n)$ and $z = (z_1,\dots,z_n)$. For $\theta \in (-1,1)$, we define $p_{(\theta)}(x,z)$ as the density of the pair $(\mathbf{X},\mathbf{Z}) \in \mathbb{R}^{2n}$ under $\mathbb{P}^\theta$. 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.

For the convenience of the reader we recall a key result from [6].

Proposition 5.1 ([6]). *The MLE θ_n is a solution to $\partial_\tau Q_n(\tau, \theta)_{|\tau=\theta} = 0$.*

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 5.1.

Proposition 5.2. *The MLE θ_n is solution to*

$$\frac{1+\theta}{2} = \frac{\mathbb{E}^\theta[U|\mathbf{X} = x]}{\mathbb{E}^\theta[N|\mathbf{X} = x]}. \tag{5.1}$$

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 [17].

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,

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

Recall that $(t, x) \mapsto p(t, x)$ is a centered Gaussian with variance t. From the construction of the density of the SBm proposed by Walsh [31], in which the independence of the sign and the position of X_i 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

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

Thus,

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

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

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

According to Proposition 5.1, the parameter $\theta^\star$ is nothing more than the MLE. □

Unfortunately, the equation (5.1) 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 5.3. *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 $\psi_i = \exp(-2(x_i x_{i-1})^+/\Delta t)$.

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

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

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 $x_i, x_{i-1} > 0$,

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

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. □

6. Numerical simulations

We present some numerical simulations on the estimator. For this, we simulate an SBm using the method proposed in [18].

We now focus on the MLE estimator θ_n of $\theta \in (-1, 1)$.

In Table 1 and Figure 1, we see that even for relatively small values of n, *e.g.* $n = 1024$, the estimator θ_n is close to θ. Besides, the estimator seems to be unbiased.

We conjecture that the rate of convergence of θ_n toward θ is 1/4. This was actually proved in [22] for the Brownian motion, that is $\theta = 0$. In Figure 2, we plot $n^{1/4}(\theta_n - \theta)$ for large values of n. We see that the density of $n^{1/4}(\theta_n - \theta)$ seems to stabilize.

Based on [22] and [17], we then conjecture the following convergence.

Conjecture 6.1. *Under $\mathbb{P}^\theta$,*

$$n^{1/4}(\theta_n - \theta) \xrightarrow[n\to\infty]{dist} \kappa\sqrt{1-\theta^2}\Upsilon \text{ with } \Upsilon := \frac{G}{\sqrt{L_1}}$$

for a constant κ, where G is a unit Gaussian distribution and L_1 the local time at time 1 of a Brownian motion independent from G.

TABLE 1. Statistics of the MLE θ_n with $\Delta t = 0.1$, $T = 102.4$, $\theta \in \{0, 0.3, 0.7, 0.9\}$ and $n = 2^{10} = 1024$ observations of $\{X_{k\Delta t}\}$. The statistics are based on $N = 1000$ simulated paths.

	$\theta = 0$	$\theta = 0.3$	$\theta = 0.7$	$\theta = 0.9$
Mean	0.007	0.294	0.672	0.860
Median	−0.001	0.299	0.696	0.883
2.5%-Quantile	−0.482	−0.130	0.321	0.631
10%-Quantile	−0.275	0.056	0.500	0.759
90%-Quantile	0.290	0.523	0.832	0.946
97.5%-Quantile	0.512	0.664	0.885	0.967

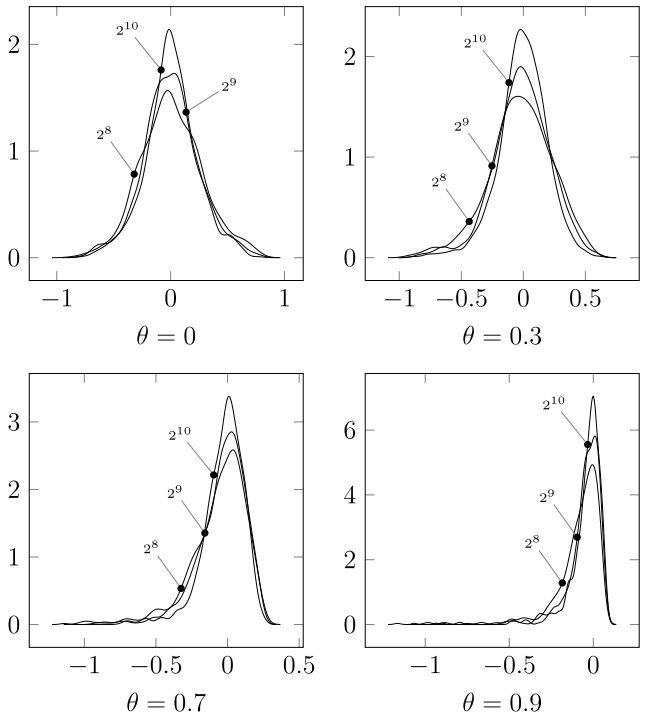


FIGURE 1. Density of the difference $\theta_n - \theta$ with $\Delta t = 0.1$, $T = 102.4$, $\theta \in \{0, 0.3, 0.7, 0.9\}$ and $n = 2^k$ points sampled at equal distance from the observations $\{X_{k\Delta t}\}$. The densities are constructed from $N = 1000$ paths.

An analytic expression is given for the constant κ, yet it is difficult to compute. This conjecture is also true for $\theta = \pm 1$. The shape for $\sqrt{1-\theta^2}$ is conjectured from the results in [17] on skew random walks.

The local time L_1 in Conjecture 6.1 is easily simulated by $L_1 \overset{\text{dist}}{=} \frac{1}{2}(U + \sqrt{V + U^2})$ where U is uniformly distributed on $[0, 1]$ and V is independent from U and follows an exponential distribution of parameter $1/2$ [22, 23].

To empirically check the validity of this conjecture, we simulate for $N = 400$ paths for $\theta = 0, 0.1, \ldots, 0.9$ and various horizons to get $n = 2^{13}, \ldots, 2^{18}$ observations (the horizon is $T = n/\delta t$). For each value of θ, we adjust a constant c as the one which minimize the Kolmogorov–Smirnov distance between the empirical distribution function of $\Theta_n := n^{1/4}(\theta_n - \theta)$ and the one of $c\Upsilon$. We then set $\widehat{\kappa} := c/\sqrt{1-\theta^2}$.

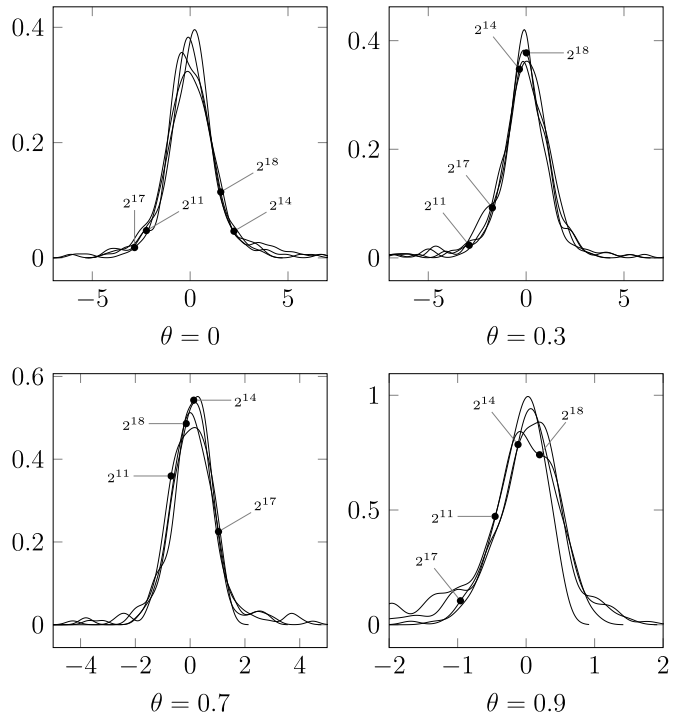


FIGURE 2. Density of the rescaled difference $n^{1/4}(\theta_n - \theta)$ with $\Delta t = 0.1$, $T = 28\,000$, $\theta \in \{0, 0.3, 0.7, 0.9\}$ and $n = 2^k$ points sampled at equal distance from the observations $\{X_{k\Delta t}\}$. The densities are constructed from $N = 200$ paths.

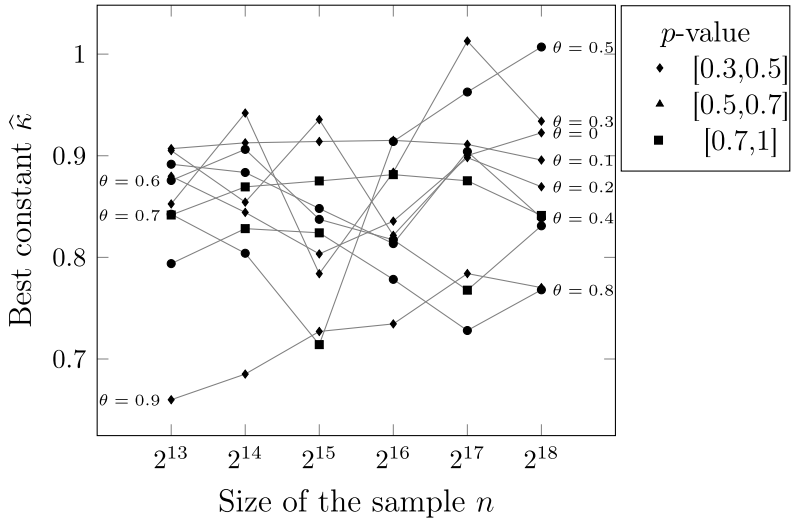


FIGURE 3. Estimation of $\widehat{\kappa}$ obtained by fitting the distribution of $n^{1/4}(\theta_n - \theta)$ against the one of $\widehat{\kappa}\Upsilon$ using the Kolmogorov–Smirnov statistics for $\theta \in \{0, 0.1, \ldots, 0.9\}$.

In Figure 3, we plot the value of $\widehat{\kappa}$ in function of n, with an indication of the p-value. We see that $\widehat{\kappa}$ varies between 0.6 and 0.9. Besides, the p-value is always greater than 0.3, and sometimes close to 1. This indicates that our conjecture is empirically sound and at last that the convergence holds at rate 1/4.

Acknowledgements. This work is supported by the Inria's Équipe Associée ANESTOC-TOSCA between France and Chile. S. Torres is supported by the grant FONDECYT 1171335. We are also grateful to Christian Y. Robert who pointed out an error in the previous version of this work.

References

[1] L.H.R. Alvarez and P. Salminen, Timing in the presence of directional predictability: Optimal stopping of skew brownian motion. *Math. Methods Oper. Res.* **82** (2017) 377–400.
[2] T.A. Appuhamillage, V.A. Bokil, E. Thomann, E. Waymire and B.D. Wood, Occupation and local times for skew brownian motion with application to dispersion accross an interface. *Ann. Appl. Probab.* **21** (2011) 183–214.
[3] O. Bardou and M. Martinez, Statistical estimation for reflected skew processes. *Stat. Inference Stoch. Process.* **13** (2010) 231–248.
[4] M. Bossy, N. Champagnat, S. Maire and D. Talay, Probabilistic interpretation and random walk on spheres algorithms for the Poisson-Boltzmann equation in molecular dynamics. *Math. Model. Numer. Anal.* **44** (2010) 997–1048.
[5] M. Decamps, M. Goovaerts and W. Schoutens, Self exciting threshold interest rates models. *Int. J. Theor. Appl. Finance* **9** (2006) 1093–1122.
[6] A.P. Dempster, N.M. Laird and D.B. Rubin, Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. Ser. B (Methodol.)* **39** (1977) 1–38.
[7] D. Florens, Estimation of the diffusion coefficient from crossings. *Stat. Inference Stoch. Process.* **1** (1998) 175–195.
[8] J.M. Harrison and L.A. Shepp, On skew Brownian motion. *Ann. Probab.* **9** (1981) 309–313.
[9] R. Höpfner and E. Löcherbach, Limit theorems for null recurrent Markov processes. *Mem. Am. Math. Soc.* **161** (2003).
[10] K. Itô and H.P. McKean, Jr., Diffusion Processes and Their Sample Paths, 2nd edn. 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 of *Lecture Notes in Mathematics*. Springer, Berlin (1984) 91–118.
[12] J. Jacod, Rates of convergence to the local time of a diffusion. *Ann. Inst. Henri Poincaré Probab. Stat.* **34** (1998) 505–544.
[13] J. Keilson and J.A. Wellner, Oscillating Brownian motion. *J. Appl. Probab.* **15** (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 of *Lecture Notes in Mathematics*. Springer, Berlin, Heidelberg (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 of *Lecture Notes in Mathematics*. Springer Verlag, Berlin (1985) 51–82.
[16] A. Lejay, On the constructions of the skew Brownian motion. *Probab. Surv.* **3** (2006) 413–466.
[17] A. Lejay, Estimation of the biais parameter of the Skew Random Walk and application to the Skew Brownian Motion. *Stat. Inference Stoch. Process.* **21** (2018) 539–551.
[18] A. Lejay and G. Pichot, Simulating diffusion processes in discontinuous media: a numerical scheme with constant time steps. *J. Comput. Phys.* **231** (2012) 7299–7314.
[19] A. Lejay and G. Pichot, Simulating diffusion processes in discontinuous media: Benchmark tests. *J. Comput. Phys.* **314** (2016) 348–413.
[20] A. Lejay and P. Pigato, Statistical estimation of the oscillating brownian motion. *Bernoulli* **24** (2018) 3568–3602.
[21] A. Lejay and P. Pigato, A threshold model for local volatility: evidence of leverage and mean reversion effects on historical data. 2018. Preprint.
[22] A. Lejay, E. Mordecki and S. Torres, Is a Brownian motion skew? *Scand. J. Stat.* **41** (2014) 346–364.
[23] D. Lépingle, Un schéma d'Euler pour équations différentielles stochastiques réfléchies. *C. R. Acad. Sci. Paris, Sér. I Math.* **316** (1993) 601–605.
[24] 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).
[25] G.J. McLachlan and T. Krishnan, The EM Algorithm and Extensions, 2nd edn. *Wiley Series in Probability and Statistics*. Wiley-Interscience, NJ (2008).
[26] O. Ovaskainen and S.J. Cornell, Biased movement at a boundary and conditional occupancy times for diffusion processes. *J. Appl. Probab.* **40** (2003) 557–580.
[27] N.I. Portenko, Diffusion processes with a generalized drift coefficient. *Teor. Veroyatnost. i Primenen.* **24** (1979) 62–77.
[28] D. Rossello, Arbitrage in skew Brownian motion models. *Insurance Math. Econom.* **50** (2012) 50–56.
[29] D. Spivakovsakaya, A.W. Heemink and E. Deleersnijder. The backward ito method for the lagrangian simulation of transport processes with large space variations of the diffusivity. *Ocean Sci.* **3** (2007) 525–535.
[30] D.J. Thomson, W.L. Physick and R.H. Maryon, Treatment of interfaces in random walk dispersion models. *J. Appl. Meteorol.* **36** (1997) 1284–1295.
[31] J.B. Walsh, A diffusion with discontinuous local time, in Temps Locaux. Vol. 52–53 of *Astérisques*. Société Mathématique de France, Marseille (1978) 37–45.
[32] M. Zhang, Calculation of diffusive shock acceleration of charged particles by skew Brownian motion. *Astrophys. J.* **541** (2000) 428–435.