Scandinavian Journal of Statistics Theory and Applications

Scandinavian Journal of Statistics, Vol. 41: 346–364, 2014
doi: 10.1111/sjos.12033
© 2013 Board of the Foundation of the Scandinavian Journal of Statistics. Published by Blackwell Publishing Ltd.

Is a Brownian Motion Skew?

ANTOINE LEJAY
IECN, Université de Lorraine
CNRS, IECL
Inria, Tosca

ERNESTO MORDECKI
Centro de Matemática, Facultad de Ciencias, Universidad de la República

SOLEDAD TORRES
CIMFAV, Facultad de Ingeniería, Universidad de Valparaíso

ABSTRACT. We study the asymptotic behaviour of the maximum likelihood estimator corresponding to the observation of a trajectory of a skew Brownian motion, through a uniform time discretization. We characterize the speed of convergence and the limiting distribution when the step size goes to zero, which in this case are non-classical, under the null hypothesis of the skew Brownian motion being an usual Brownian motion. This allows to design a test on the skewness parameter. We show that numerical simulations can be easily performed to estimate the skewness parameter and provide an application in Biology.

Key words: maximum likelihood, skew Brownian motion, statistical estimation

1. Introduction

The skew Brownian motion (SBm) has attracted interest within other facts, due to its relations with diffusions with discontinuous coefficients or to media with permeable barriers, being the first example of the solution of a stochastic differential equation (SDE) with the local time of the solution as drift (Harrison & Shepp, 1981): the SBm $X = \{X_t : 0 \le t \le T\}$ can be defined as the strong solution of the SDE

$$X_t = x + B_t + \theta \ell_t, \tag{1}$$

where $B = \{B_t : 0 \le t \le T\}$ is a standard Brownian motion defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, the initial condition is $x \ge 0$ (the case $x < 0$ is symmetrical), $\theta \in [-1, 1]$ is the *skewness parameter* and $\ell = \{\ell_t : 0 \le t \le T\}$ is the local time at level zero of the (unknown) solution X of the equation departing from x, defined by

$$\ell_t = \lim_{\epsilon \to 0} \frac{1}{2\epsilon} \int_0^t \mathbf{1}_{(-\epsilon,\epsilon)}(X_s)\, \mathrm{d}s.$$

In the literature, the skewness parameter is sometimes defined as $p = (\theta + 1)/2$; this second parametrization being more convenient for an alternative construction of the SBm: depart from the reflected Brownian motion and choose, independently with probability $p \in [0, 1]$, whether each particular excursion of the reflected Brownian motion remains positive.

In the special case $\theta = 1$ ($p = 1$), the solution to (1) is the reflected Brownian motion. The case $\theta = 0$ ($p = 1/2$) corresponds to the standard Brownian motion.

Recently, several papers have considered the SBm in modelling or simulation issues, as well as some optimization problems. Thanks to the Itô-Tanaka formula, the SBm is strongly related

to diffusion processes whose infinitesimal generator, both divergence-form or non-divergence-form, has discontinuous coefficients (Lejay & Martinez, 2006). Hence, the SBm appears naturally for example in geophysics as a model of the diffusion of a pollutant (Ramirez *et al.*, 2006), in brain imaging by diffusion MRI or in electro/encephalography (Martinez, 2004), in population ecology (Cantrell & Cosner, 1999; Ovaskainen & Cornell, 2003), in astrophysics (Zhang, 2000). See the review by Lejay (2006) for references on the subject, as well as a survey of the various possible constructions and applications of the SBm. In Lejay & Pichot (2012), it is shown how to simulate exactly the position of the SBm after a constant time step.

In this paper, we are interested in the statistical estimation of θ, the skewness parameter, when we observe a trajectory of the process through an equally spaced time grid. From the statistical point of view, we find this problem interesting because it is in certain sense intermediate between the classical problem of drift estimation in a diffusion, where the measures generated by the trajectories of the process for different values of the parameter are equivalent (Kutoyants, 1984; Lipster & Shiryaev, 2001), and the estimation of the variance (the volatility in financial terms) of a diffusion (see for instance, Florens-Zmirou (1993), or Jacod (2006) and the references therein), where the probability measures generated by the trajectories are singular for different values of the parameter. At the best of our knowledge, the only estimator of θ is the one constructed in Bardou & Martinez (2010), where the authors assume that the SBm is reflected at levels 1 and -1 to ensure ergodicity, considering a different scheme of observation of the trajectory.

The strategy proposed by O. Bardou and M. Martinez does not apply here, as it is strongly related to ergodicity, whereas the SBm is only null recurrent. The distribution of the renormalized occupation time on the positive side $A^+(t)/t$ of the SBm up to time t is computed in Lamperti (1958) (for $\theta = 0$, this is the famous Arc-Sine distribution). Despite the mean of $A^+(t)/t$ is equal to $(1 + \theta)/2$, it has a positive variance, and it may not be used to set up an estimator of θ, nor its discretized version (see Fujihara *et al.* (2007), e.g.).

Indeed, only the behaviour of the SBm close to 0 brings some information for estimating θ. The expansion the log-likelihood provides indeed some weighted version of the difference between upcrossings and downcrossings.

The main results of this article follows. First, we prove that under the distribution of the Brownian motion, that is with $\theta = 0$, the maximum likelihood estimator (MLE) θ_n from n equally spaced observations of $X_{kT/n}$ on $[0, T]$ converges in probability to the true value $\theta = 0$. Second, we prove that the scaled estimator $n^{1/4}\theta_n$ converges stably to $cT^{1/4}W(\ell_T)/\ell_T$, when the trajectory hits zero where c is a constant and W is a Brownian motion independent from X. The convergence of $n^{1/4}\theta_n$ is the local asymptotically mixed normality property at the point $\theta = 0$ (Le Cam & Yang, 2000), but with a rate $1/4$, whereas in general, the rate is $1/2$. This situation is typical for null recurrent estimation problems (Höpfner & Löcherbach, 2003). Third, we obtain an expansion of θ_n as $\theta_n^{(1)}/n^{1/4} + \theta_n^{(2)}/n^{1/2} + \theta_n^{(3)}/n^{3/4} + \cdots$ where the $\theta_n^{(i)}$ are computed recursively.

Numerical simulations show us that it is plausible that θ_n converges in probability to θ under the distribution of the SBm of parameter θ. However, we have not been able to prove that θ^n converges to θ in probability in such a situation. The key argument we used for $\theta = 0$, which are contained in some limit theorem studied in Jacod (1998), are difficult to extend here, as the distribution of the SBm for $\theta = 0$ is singular with respect to the one of the Brownian motion (Le Gall, 1984). However, if all the observations are positive (resp. negative), then the MLE estimator gives 1 (resp. -1). Hence, under the distribution of the reflected Brownian motion, the estimator always gives the right value of θ, but this corresponds to a degenerate case.

© 2013 Board of the Foundation of the Scandinavian Journal of Statistics.

The rest of the paper is organized as follows. Section 2 describes the maximum likelihood methodology and the convergence results. In Section 3, we describe the limit distribution. Section 4 presents a statistical test and some numerical simulations on the likelihood function. Section 5 presents an application to diffusion of species in two different habitats, and Section 6 presents our conclusions. Finally, in the Appendix, we provide the theorems from Jacod (1998) used in the proof of our main results in Section 2.

2. The maximum likelihood estimator

Consider the SBm X with parameter $\theta \in [-1,1]$ defined in (1) and the sampling scheme denoted by $X_i := X_{iT/n}$ $(i = 0,\dots,n)$, and $\Delta = T/n$. In this section, we derive the asymptotic behaviour of MLE θ_n of the parameter θ when we observe the sample $X_1,\dots,X_n$. As shown in Walsh (1978), the transition density of the SBm of parameter $\theta \in [-1,1]$ is given by:

$$q_\theta(t,x,y) := p(t,y-x) + \operatorname{sgn}(y)\theta p(t,|x|+|y|), \tag{2}$$

where

$$p(t,x) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{x^2}{2t}\right)$$

is the density of a Gaussian random variable with variance t and mean 0 and $\operatorname{sgn}(x) = 1$ if $x > 0$, $\operatorname{sgn}(x) = -1$ if $x < 0$ and $\operatorname{sgn}(x) = 0$ if $x = 0$. The likelihood of the sample is given by

$$\Lambda_n(\theta) := \prod_{i=0}^{n-1} q_\theta(\Delta, X_i, X_{i+1}).$$

The *maximum likelihood estimator* (MLE) is

$$\theta_n := \operatorname{argmax}_{\theta\in[-1,1]} \Lambda_n(\theta).$$

The uniqueness of the MLE follows from simple computations.

Lemma 1. *For each n, there exists a unique MLE θ_n.*

Proof. The derivatives of $L_n(\theta) = \log(\Lambda_n(\theta))$ with respect to θ are easily computed thanks to the simple expression of the density q_θ in (2). As seen in (3), $\partial^2_\theta L_n(\theta) < 0$. This proves that $\theta \mapsto L_n(\theta)$ is strictly concave so that θ_n is unique. □

Remark 1. The SBm has the same scaling property as the Brownian motion. If X is a SBm of parameter θ, so is cX_{t/c^2} for $c > 0$. Hence, for fixed n and T, $Y = \left\{\sqrt{n/T}X_{tT/n} : t \in [0,n]\right\}$ with $X_0 = 0$ is a SBm of parameter θ for a SBm X of parameter θ. It is then easily seen that the MLE applied to the sequence $\{Y_i\}_{i=0,\dots,n}$ with a time-step $\Delta = 1$ is distributed as the one for the sequence $\{X_{iT/n}\}_{i=0,\dots,n}$ with a time-step $\Delta = T/n$. In this case, high-frequency estimation and long-time estimation are equivalent. In any case, we are in the situation of a statistical estimation problem for a null-recurrent Markov process.

The difference between a SBm and a Brownian motion lies at $x = 0$, the origin, so that the path should cross 0 to obtain a meaningful estimator.

Lemma 2. *If all the X_i's are positive (resp. negative) for $i = 1,\dots,n$, then $\theta_n = 1$ (resp. $\theta_n = -1$). This means that the estimator sees a positively (resp. negatively) reflected Brownian motion.*

© 2013 Board of the Foundation of the Scandinavian Journal of Statistics.

Proof. From (2), for fixed $x \in \mathbb{R}$ and $y \neq 0$, $\operatorname{argmax} q_\theta(\Delta, x, y) = \operatorname{sgn}(y)$. Besides, if $y = 0$, then $p_\theta(\Delta, x, 0)$ does not depend on θ. Hence, if all the observations $X_1, \ldots, X_n$ are non-negative (resp. non-positive), then $\theta_n = 1$ (resp. $\theta_n = -1$) because $\Lambda_n(\theta)$ is the product of the $q_\theta(\Delta, X_i, X_{i+1})$'s. We exclude the degenerated case that all the X_i's are equal to 0, which is almost surely never observed whatever the value of θ. □

Lemma 3. *If θ_n is the MLE associated to $\{X_i\}_{i=0,\ldots,n}$, then $-\theta_n$ is the MLE associated to $\{-X_i\}_{i=0,\ldots,n}$.*

Proof. This is an easy consequence of (2): The likelihood is left unchanged when the signs of θ and of the X_i's are changed. □

2.1. The log-likelihood around $\theta = 0$

Instead of using the likelihood $\Lambda_n(\theta)$, it is often more convenient to use the *log-likelihood* defined by

$$L_n(\theta) := \log \Lambda_n(\theta) = \sum_{i=0}^{n-1} \log q_\theta(\Delta, X_i, X_{i+1}).$$

The MLE is the point θ_n at which $\theta \mapsto L_n(\theta)$ reaches its maximum. The MLE θ_n may equivalently be defined as the solution to $L_n^{(1)}(\theta_n) = 0$.

For $k \geq 1$, its scaled (for notational convenience) k-th derivatives are

$$L_n^{(k)}(\theta) = \frac{1}{(k-1)!} \frac{\partial^k}{\partial \theta^k} L_n(\theta),$$

that are computed as

$$L_n^{(k)}(\theta) = (-1)^{k-1} \sum_{i=0}^{n-1} \frac{\operatorname{sgn}(X_{i+1})^k \, p(\Delta, |X_i| + |X_{i+1}|)^k}{q_\theta(\Delta, X_i, X_{i+1})^k}. \tag{3}$$

An analytic development of $L_n^{(1)}(\theta)$ holds around 0:

$$L_n^{(1)}(\theta) = \sum_{k=0}^{+\infty} \theta^k L_n^{(k+1)}(0). \tag{4}$$

The inequality

$$\frac{p(\Delta, |x| + |y|)}{q_0(\Delta, x, y)} = \exp\left(-\frac{|xy| + xy}{\Delta}\right) = \exp\left(\frac{-2(xy)^+}{\Delta}\right) \leq 1, \ x, y \in \mathbb{R} \tag{5}$$

implies that $|L_n^{(k)}(0)| \leq n$ and thus the series in (4) is absolutely convergent for $|\theta| < 1$.

Introduce, for $k = 1, 2, \ldots$, the sequence of functions

$$h_k(x, y) = h(x, y)^k \text{ with } h(x, y) = \operatorname{sgn}(x + y) \exp(-(2/T)(x(x + y))^+).$$

We can then rewrite $L_n^{(k)}(0)$, for $k = 1, 2, \ldots$, as

$$L_n^{(k)}(0) = (-1)^{k-1} \sum_{i=0}^{n-1} h_k\left(\sqrt{n} X_i, \sqrt{n}(X_{i+1} - X_i)\right).$$

© 2013 Board of the Foundation of the Scandinavian Journal of Statistics.

We then see that the study of the limit behaviour of this type of sums, presented in the next proposition, can be directly obtained from results in Jacod (1998) (for convenience, we present these results in an Appendix).

2.2. Asymptotic behaviour of the maximum likelihood estimator under the Brownian distribution

Let us set

$$\alpha_n := -n^{1/4}\frac{L_n^{(1)}(0)}{L_n^{(2)}(0)} = -n^{1/4}\frac{\sum_{i=0}^{n-1}\operatorname{sgn}(X_{i+1})h\left(\sqrt{n}X_i, \sqrt{n}(X_{i+1}-X_i)\right)}{\sum_{i=0}^{n-1}h\left(\sqrt{n}X_i, \sqrt{n}(X_{i+1}-X_i)\right)^2}.$$

Unlike the MLE, α_n is pretty simple to compute from the X_i's.

For a non-negative starting point $x = X_0$, we consider the events

$$A_n := \left\{\omega : \inf_{1\le i\le n} X_i(\omega) < 0 \text{ and } \sup_{1\le i\le n} X_i(\omega) > 0\right\},$$
$$A := \left\{\omega : \inf_{0\le t\le T} X_t(\omega) < 0\right\}.$$

Let $\{\ell_t : 0 \le t \le T\}$ be the local time at 0 of the process X. The event A means that the path X crossed zero between 0 and T. Hence, $A = \{\ell_T > 0\}$. If $x = 0$, then $\mathbb{P}[A] = 1$. Of course, $A_n \subset A$ and $\mathbb{P}[A \setminus \cup_{n\ge} A_n] = 0$. On the complementary event $\overline{A}_n$ of A_n, lemma 2 asserts that the MLE θ_n is equal to 1.

Let X be a SBm defined as the strong solution to (1) on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ for a canonical Brownian motion W with a filtration $\{\mathcal{F}_t : t \ge 0\}$ which satisfy the usual hypotheses.

Our main theorem is the following. It proves the convergence of the MLE θ_n to 0 and specify its speed of convergence. In addition, this show that the MLE may be expanded as a series in terms of $1/n^{1/4}$, where α_n plays a fundamental role.

Before stating our theorem, we review the notion stable convergence due to A. Renyi (see Rényi (1963), Aldous & Eagleson (1978), & Jacod & Shiryaev (1987)).

Definition 1 (Stable convergence)**.** Consider a sequence of random variables $Y, Y_1, Y_2, \ldots$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and a σ-algebra $\mathcal{G} \subset \mathcal{F}$.

We say that the sequence of random variables $Y_1, Y_2 \ldots$ converge $\mathcal{G}$-stably in distribution to a random variable Y defined on an extension $(\Omega', \mathcal{F}', \mathbb{P}')$ of $(\Omega, \mathcal{F}, \mathbb{P})$, and denote

$$Y_n \xrightarrow[n\to\infty]{\mathcal{G}\text{-stably}} Y$$

when

$$\mathbb{E}(Zf(Y_n)) \xrightarrow[n\to\infty]{} \mathbb{E}(Zf(Y))$$

for any bounded $\mathcal{G}$ measurable random variable Z, and any bounded and continuous function f.

Theorem 1. *Assume that $\theta = 0$, that is let X be a Brownian motion on $[0, T]$ departing from x. Let μ_k for $k = 1, 2, 3, \ldots$ be the constants given by (8) below. Then*

(i) For $\lambda = -\mu_1/\mu_2$, $\alpha_n \mathbf{1}_{A_n} \xrightarrow[n\to\infty]{\mathcal{F}\text{-stably}} \lambda\Upsilon_T$

© 2013 Board of the Foundation of the Scandinavian Journal of Statistics.

with

$$\Upsilon_T = \begin{cases} T^{1/4} W(\ell_T)/\ell_T & \text{if } \ell_T > 0, \\ 0 & \text{if } \ell_T = 0, \end{cases}$$

where W is a Brownian motion independent from B.

(ii) Let $d_n^{(1)}, d_n^{(2)}, \ldots$ be a sequence of random variables given the recursive relation $d_n^{(1)} = 1$ and

$$d_n^{(m+1)} = -\sum_{k=1}^{m+1} \frac{L_n^{(k+1)}(0)}{L_n^{(2)}(0)} \sum_{\substack{1 \le i_1, \ldots, i_k \le m \\ i_1 + \cdots + i_k = m+1}} d_n^{(i_1)} \cdots d_n^{(i_k)}. \tag{6}$$

Let $d^{(1)}, d^{(2)}, \ldots$ be a sequence of random variables defined recursively by $d^{(1)} = 1$, $d^{(2)} = 0$ and for $m \ge 2$,

$$d^{(m+1)} = -\sum_{k=1}^{m} \lambda^{(k)} \sum_{\substack{1 \le i_1, \ldots, i_k \le m \\ i_1 + \cdots + i_k = m+1}} d_n^{(i_1)} \cdots d_n^{(i_k)}$$

with $\lambda^{(2k+1)} = \dfrac{T^{1/4} \mu_{2k+1} W(\ell_T)}{\mu_2 \ell_T}$ *and* $\lambda^{(2k+2)} = 0$.

For any integer $p \ge 0$, the vector $\left(\mathbf{1}_{A_n} d_n^{(1)}, \ldots, \mathbf{1}_A d_n^{(p+1)}\right)$ converges $\mathcal{F}$-stably to $(d^{(1)}, \ldots, d^{(p+1)})$ depending only on ℓ_T and $W(\ell_T)$. Besides, for any $\epsilon > 0$, there exists some integer n_0 large enough and some K such that

$$\mathbb{P}\left[n^{\frac{p}{4}+\frac{1}{2}} |\theta_n - \Theta_n| \ge K; A_n\right] \le \epsilon \text{ for any } n \ge n_0,$$

where

$$\Theta_n = \frac{\alpha_n}{n^{1/4}} + d_n^{(2)} \frac{\alpha_n^2}{n^{1/2}} + \cdots + d_n^{(p+1)} \frac{\alpha_n^{p+1}}{n^{(p+1)/4}}.$$

In addition, $d_n^{(2)}$ converges in probability to 0 and $n^{1/4} d_n^{(2)}$ is bounded.

(iii) It holds that $\theta_n \mathbf{1}_{A_n}$ converges in probability to 0 under the Brownian distribution and

$$n^{1/4} \theta_n \mathbf{1}_{A_n} \xrightarrow[n \to \infty]{\mathcal{F}\text{-stably}} \lambda \Upsilon_T. \tag{7}$$

Remark 2. Using the scaling property of the Brownian motion and then of the local time, and adjusting the initial condition of the process, Υ_T is equal in distribution to Υ_1 for any $T > 0$.

3. Proof of the theorem

3.1. Convergence of the derivatives of the log-likelihood

The main points in the proof of theorem 1 follows from the asymptotic behaviour of the derivative $L_n^{(k)}(0)$ of the log-likelihood. Let Φ be the cumulative distribution function of the standard normal distribution.

Let us note that if X is a SBm of parameter θ then from the scaling property, one obtains easily that the process Y defined by $Y_t = T^{-1/2} X_{tT}$ is also a SBm of parameter θ and $\ell_t(Y) = T^{-1/2} \ell_{tT}$, where $\ell_t(Y)$ is the symmetric local time at zero of Y. With this remark, it is possible to apply the results from Jacod (1998) to Y with $T = 1$ and then to transform then back on results on X.

© 2013 Board of the Foundation of the Scandinavian Journal of Statistics.

Proposition 1. *Assume that $\theta = 0$ in (1), that is let X be a Brownian motion on $[0, T]$ departing from x, and let ℓ denote its symmetric local time at zero.*

(i) *Assume that $k = 2, 4, \ldots$. Denote*

$$\mu_k = -2\int_0^\infty \left[1 + \frac{1}{2k-1}\exp\left(\frac{2k(k-1)x^2}{(2k-1)^2}\right)\right]\Phi(-x)\,\mathrm{d}x. \tag{8}$$

Then,

$$\frac{L_n^{(k)}(0)}{n^{1/2}} \xrightarrow[n\to\infty]{\text{prob.}} \frac{\mu_k}{\sqrt{T}}\ell_T. \tag{9}$$

(ii) *Assume that $k = 1, 3, \ldots$. Then, for $\mu_k = C(h^k)$, where C has been defined by (20) in Appendix and $h(x, y) := sgn(x + y)\exp\left(-(2/T)(x(x + y))^+\right)$, there exists a Brownian motion $W^{(k)}$ independent from B such that*

$$\frac{L_n^{(k)}(0)}{n^{1/4}} \xrightarrow[n\to\infty]{\mathcal{F}-stably} \frac{\mu_k}{T^{1/4}}W^{(k)}(\ell_T). \tag{10}$$

Remark 3. The results of J. Jacod are given for multidimensional statistics. For any integer k, the joint convergence in probability of $n^{-1/2}\left(L_n^{(2)}(0), \ldots, L_n^{(2k)}(0)\right)$ holds as well as the joint $\mathcal{F}$-stable convergence of $n^{-1/4}\left(L_n^{(1)}(0), \ldots, L_n^{(2k+1)}(0)\right)$. Hence, using the property of stable convergence (see Aldous & Eagleson (1978, Theorem 1' and subsequent Remark)), this implies the joint $\mathcal{F}$-stable convergence of

$$\left(n^{-1/4}L_n^{(1)}(0), n^{-1/2}L_n^{(2)}(0), \ldots, n^{-1/4}L_n^{(2k+1)}(0), n^{-1/2}L_n^{(2k+2)}(0)\right) \xrightarrow[n\to\infty]{\mathcal{F}-\text{stably}}$$
$$\left(T^{-1/4}\mu_1 W^{(1)}(\ell_T), T^{-1/2}\mu_2\ell_T, \ldots, T^{-1/4}\mu_{2k+1}W^{(2k+1)}(\ell_T), T^{-1/2}\mu_{2+2}\ell_T\right).$$

The correlations between the Brownian motions $W^{(k)}$ may be computed also thanks to Theorem 1.2 in Jacod (1998), but we do not use them here.

Proof. We apply theorem 2 in the Appendix to the process $Y_t = T^{-1/2}X_{tT}$, and then we assume that $T = 1$. For $T \neq 1$, the scaling property has then to be used again. Observe that

$$h_k(x, y) \le \exp(-(x(x + y))^+) \le \exp(|y| - |x \wedge x^2|).$$

We then have that (17) in Appendix holds with $a = 1$, $\hat{h}(x) = \exp(-|x \wedge x^2|)$ and $r = 0$, then it holds for any $r > 0$. In consequence, by the after mentioned theorem, the convergence in (18) holds for $h = h_k$ with $k = 2, 4, \ldots$. It remains to compute the constant in (19). We have

© 2013 Board of the Foundation of the Scandinavian Journal of Statistics.

$$
\begin{aligned}
c(h_k) &= \iint_{\mathbb{R}^2} h_k(x,y)p(1,y)\,\mathrm{d}x\,\mathrm{d}y = 2\int_0^\infty \mathrm{d}x \int_{-\infty}^{-x} p(1,y)\,\mathrm{d}y \\
&\quad + \frac{2}{\sqrt{2\pi}}\int_0^\infty \mathrm{d}x \int_{-x}^\infty \exp\left(-\frac{1}{2}y^2 - 2kxy - 2kx^2\right)\mathrm{d}y \\
&= 2\int_0^\infty \Phi(-x)\,\mathrm{d}x \\
&\quad + \frac{2}{\sqrt{2\pi}}\int_0^\infty \mathrm{d}x \exp(2k(k-1)x^2)\int_{-x}^\infty \exp\left(-\frac{1}{2}(y+2kx)^2\right)\mathrm{d}y \\
&= 2\int_0^\infty \Phi(-x)\,\mathrm{d}x + 2\int_0^\infty \exp(2k(k-1)x^2)\Phi(-(2k-1)x)\,\mathrm{d}x \\
&= 2\int_0^\infty \Phi(-x)\,\mathrm{d}x + \frac{2}{2k-1}\int_0^\infty \exp\left(\frac{2k(k-1)x^2}{(2k-1)^2}\right)\Phi(-x)\,\mathrm{d}x.
\end{aligned}
$$

Taking into account that $\mu_k = -c(h_k)$, we conclude that (9) holds with μ_k given in (8).

To prove (i) we rely on theorem 3 in the Appendix. Observe then that $c(h_k) = 0$ for odd k due to the property

$$h_k(-x,-y) = -h_k(x,y) \quad \text{for odd } k.$$

In view of the fact that (17) holds for all h_k with $r = 4$, we conclude the proof. □

The proof of the following corollary is immediate from proposition 1 because $A = \{\ell_T > 0\}$, Remark 3 and the continuous mapping theorem.

Corollary 1. *Given the conditions of proposition 1, for $k = 0, 1, 2, \ldots$, we have*

$$
\frac{L_n^{(2k+1)}(0)}{L_n^{(2)}(0)}\mathbf{1}_{A_n} \xrightarrow[n\to\infty]{prob.} 0, \qquad \frac{L_n^{(2k+2)}(0)}{L_n^{(2)}(0)}\mathbf{1}_{A_n} \xrightarrow[n\to\infty]{prob.} \frac{\mu_{2k+2}}{\mu_2}\mathbf{1}_A,
$$

$$
n^{1/4}\frac{L_n^{(2k+1)}(0)}{L_n^{(2)}(0)}\mathbf{1}_{A_n} \xrightarrow[n\to\infty]{\mathcal{F}-stably} T^{1/4}\frac{\mu_{2k+1}W(\ell_T)}{\mu_2\ell_T}\mathbf{1}_A.
$$

In particular, for $\lambda = -\mu_1/\mu_2$

$$
-n^{1/4}\frac{L_n^{(1)}(0)}{L_n^{(2)}(0)}\mathbf{1}_{A_n} \xrightarrow[n\to\infty]{\mathcal{F}-stably} T^{1/4}\lambda\frac{W(\ell_T)}{\ell_T}\mathbf{1}_A.
$$

Hence, part (i) of theorem 1 is immediate from this corollary.

3.2. Asymptotic development of the maximum likelihood estimator

We now study the development of the MLE θ_n as a series in $1/n^{1/4}$. This will prove point (ii) of theorem 1. Point (iii) is an immediate consequence of point (ii).

Using the result of Remark 3, we consider the asymptotic behaviour of the vector

$$\left(n^{-1/4}L_n^{(1)}(0), n^{-1/2}L_n^{(2)}(0), \ldots, n^{-1/2}L_n^{(2k)}(0)\right)$$

for some $k \geq 1$. We may then consider a probability space $(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{\mathbb{P}})$ such that this sequence is equal in distribution to a sequence converging almost surely to $\left(T^{-1/4}\mu_1 W^{(1)}(\ell_T), T^{-1/2}\mu_2\ell_T, \ldots, T^{-1/2}\mu_{2k}\ell_T\right)$. We now consider some point in this probability space such that $\ell_T > 0$. If the starting point is 0, then the event $\{\ell_T > 0\}$ is of full measure.

© 2013 Board of the Foundation of the Scandinavian Journal of Statistics.

Proposition 2. *On the probability space* $(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{\mathbb{P}})$ *earlier, the random sequences* $d_n^{(i)}$ *given by (6) are convergent and bounded in n. Besides, for* $m = 1, 2, 3, \ldots$,

$$\theta_n = \frac{\alpha_n}{n^{1/4}} + d_n^{(2)} \frac{\alpha_n^2}{n^{2/4}} + d_n^{(3)} \frac{\alpha_n^3}{n^{3/4}} + \cdots + d_n^{(m)} \frac{\alpha_n^m}{n^{m/4}} + \mathrm{O}\left(\frac{1}{n^{(m+1)/4}}\right)$$

almost surely in the event $\{\ell_T > 0\}$.

Let us start by a simple lemma to obtain a control over the finite Taylor expansion of $L_n^{(1)}(\theta)$.

Lemma 4. *For any* θ *and any integer* $m \geq 1$*, there exists a random constant C such that*

$$\left| L_n^{(1)}(\theta) - \sum_{k=0}^{m} L_n^{(k+1)}(0)\theta^k \right| \leq \sup_{|\xi| \leq |\theta|} |L_n^{(m+2)}(\xi)| \cdot |\theta|^{m+1} \leq C n^{1/2} \frac{|\theta|^{m+1}}{(1-|\theta|)^{m+2}}. \tag{11}$$

Proof. With (5) and for $\theta \in (-1, 1)$,

$$1 - |\theta| \leq \frac{q_\theta(\Delta, X_i, X_{i+1})}{q_0(\Delta, X_i, X_{i+1})} \leq 1 + |\theta|. \tag{12}$$

With (3), because $L_n^{(k)}(0)/n^{1/2}$ converges in probability (either to $\mu_k \ell_T$ or to 0 depending if k is even or odd), there exists a random constant C such that

$$|L_n^{(k)}(\theta)| \leq \frac{n^{1/2} C}{(1-|\theta|)^k}.$$

Hence,

$$\left| L_n^{(1)}(\theta) - \sum_{k=0}^{m} L_n^{(k+1)}(0)\theta^k \right| \leq \sup_{|\xi| \leq |\theta|} |L_n^{(m+2)}(\xi)| \cdot |\theta|^{m+1}.$$

With (9) and (10), this gives (11) because $L_n^{(k)}(0)/n^{1/2}$ is bounded in n. □

Lemma 5. *For n large enough, the function* $L_n^{(1)}(\theta)$ *is invertible. Besides, for n large enough, the function* $\left(L_n^{(1)}(\theta)\right)^{-1}$ *is Lipschitz in* θ *with a constant* $8/n^{1/2}\mu_2\ell_T$ *on the event* A_n.

Proof. With (12),

$$-L_n^{(2)}(\theta) \geq \left(\sum_{i=1}^{n-1} \frac{p(\Delta, |X_i| + |X_{i+1}|)^2}{q_0(\Delta, X_i, X_{i+1})^2} \right) \frac{1}{(1+|\theta|)^2} \geq \frac{L_n^{(2)}(0)}{4}.$$

Because $L_n^{(2)}(0) < 0$ for n large enough as $n^{-1/2} L_n^{(2)}(\theta)$ converges in probability to some negative random variable (see (9)), we obtain that $L_n^{(1)}(\theta)$ is one-to-one. With the formula $\partial_\theta \left(L_n^{(1)}(\theta)\right)^{-1} = 1/L_n^{(2)}\left(L_n^{(1)}(\theta)\right)$, $\left(L_n^{(1)}(\theta)\right)^{-1}$ is Lipschitz in θ with constant $4/L_n^{(2)}(0)$. □

The idea of the proof is then the following: We construct an estimator Θ_n such that for some constant C and $p \geq 0$,

$$\sup_{n \in \mathbb{N}} n^{p/4} |L_n^{(1)}(\Theta_n)| \leq C.$$

© 2013 Board of the Foundation of the Scandinavian Journal of Statistics.

Because $L_n^{(1)}(\theta_n) = 0$, for n large enough,

$$|\Theta_n - \theta_n| = |\left(L_n^{(1)}\right)^{-1}\left(L_n^{(1)}(\Theta_n)\right) - \left(L_n^{(1)}\right)^{-1}\left(L_n^{(1)}(\theta_n)\right)| \leq \frac{8T^{1/2}}{n^{1/2}\mu_2\ell_T}|L_n^{(1)}(\Theta_n)| \leq \frac{8CT^{1/2}}{n^{p/4+1/2}\mu_2\ell_T}. \tag{13}$$

Proof of proposition 2. For the sake of simplicity, let us set $q = n^{1/4}$.

Set $\Theta_n = \alpha_n q + \beta_n q^2 + \gamma_n q^3 + \xi_n q^4$ for some β_n, γ_n and ξ_n to be carefully chosen. Here, we consider only the first terms in the development of θ_n. It is easily to convince oneself that this method may be applied to any order and that the involved terms $\beta_n, \gamma_n, \xi_n, \ldots$ may be computed recursively and give rise to (6).

With (11) and $m = 4$, there exists a constant C such that

$$|L_n^{(1)}(\Theta_n) - L_n^{(1)}(0) - L_n^{(2)}(0)\Theta_n - L_n^{(3)}(0)\Theta_n^2 - L_n^{(4)}(0)\Theta_n^3 - L_n^{(5)}(0)\Theta_n^4| \leq Cn^{1/2}\frac{|\Theta_n|^5}{(1-|\Theta_n|)^6}.$$

Remark that $L_n^{(1)}(0) - L_n^{(2)}(0)\alpha_n q = 0$. To get rid of the terms in q^2, set

$$\beta_n = \frac{-L_n^{(3)}(0)}{L_n^{(2)}(0)}\alpha_n^2$$

Because the joint distribution of α_n and $n^{1/4}L_n^{(3)}(0)/L_n^{(2)}(0)$ converges stably, $n^{1/4}\beta_n$ converges stably. Also, β_n converges to 0.

To get rid of the terms in q^3, set

$$\gamma_n = -\frac{L_n^{(4)}(0)}{L_n^{(2)}(0)}\alpha_n^3 - \frac{L_n^{(3)}(0)}{L_n^{(2)}(0)}\alpha_n\beta_n.$$

From Corollary 1, γ_n converges stably because α_n and β_n converges stably.

To get rid of the terms in q^4, set

$$\xi_n = -2\frac{L_n^{(3)}(0)}{L_n^{(2)}(0)}\left(\alpha_n\gamma_n + \beta_n^2\right) - 4\frac{L_n^{(4)}(0)}{L_n^{(2)}(0)}\alpha_n^2\beta_n.$$

Again, ξ_n converges thanks to Corollary 1.

Hence,

$$L_n^{(1)}(\Theta_n) = \sum_{r=5}^{20} q^r B_n^{(r)} + R_n(\Theta_n),$$

where $R_n(\Theta_n) \leq n^{1/2}|\Theta_n|^5/(1-|\Theta_n|)^6$ and $B_n^{(r)}$ are terms that depend linearly on $L_n^{(k)}(0)$ and on the power of the α_n, β_n, γ_n and ξ_n. Because the $L_n^{(k)}(0)/n^{1/2}$ are bounded, we obtain that the $n^{3/4}B_n^{(r)}$ are bounded.

In addition, $n^{1/4}\Theta_n$ is bounded in n, so that $n^{3/4}R_n(\Theta_n)$ is bounded in n. With (13), this proves that for some constant K,

$$|\Theta_n - \theta_n| \leq \frac{K}{n^{5/4}}.$$

Finally, let us note that $\beta_n = \alpha_n^2 d_n^{(2)}$ with $d_n^{(2)} = -L_n^{(3)}(0)/L_n^{(2)}(0)$. With (1), $d_n^{(2)}$ converges in probability to 0 and $n^{1/4}d_n^{(2)}$ is bounded in n. Remark that $\gamma_n = \alpha_n^3 d_n^{(3)}$ and $\xi_n = \alpha_n^4 d_n^{(4)}$ where $d_n^{(3)}$ and $d_n^{(4)}$ are bounded in n. This result may be generalized to any order. □

© 2013 Board of the Foundation of the Scandinavian Journal of Statistics.

Proof of theorem 1 point (ii). On the event $\{\ell_T > 0\}$, the local time has a density (see lemma 6 later). For each $\epsilon > 0$, one may find a set $\Omega(\epsilon)$ as well as some values $0 < a' < b'$ and c' such that $\omega \in \Omega(\epsilon)$ implies that $\ell_T^x \in (a', b')$ and $|W(\ell_T)/\ell_T| \leq c'$ and

$$\widehat{\mathbb{P}}[\Omega(\epsilon)|\{\ell_T > 0\}] \geq 1 - \epsilon/2.$$

From the joint convergence of the $n^{1/4}L_n^{(2k+1)}(0)$ to $\mu_{2k+1}W(\ell_T)$ and the joint convergence of the $n^{1/2}L_n^{(2k)}(0)$ to $\mu_{2k}\ell_T$, we obtain for any $\epsilon > 0$, there exists $0 < k < a'$ and $K > b'$ as well as a measurable set $\Omega'(\epsilon, n) \subset \Omega(\epsilon)$ such that

$$\begin{aligned} &L_n^{(k)}(0) \leq K\sqrt{n} \text{ on } \Omega'(\epsilon, n)\\ &L_n^{(2)}(0) \geq k\sqrt{n} \text{ on } \Omega'(\epsilon, n)\\ \text{and } \forall n \geq n_0,\ &\widehat{\mathbb{P}}[\Omega'(n, \epsilon)|\{\ell_T > 0\}] \geq 1 - \epsilon. \end{aligned}$$

In the proof of proposition 2, we constructed some estimator Θ_n such that for some $p \geq 0$, $n^p L_n^{(1)}(\Theta_n)$ is bounded by some constant depending the upper bounds of the $n^{1/2}L_n^{(k)}(0)$. Besides, we use the Lipschitz constant of $\left(L_n^{(1)}\right)^{-1}$ which depends on the lower bound of $n^{1/2}L_n^{(2)}(0)$. Thus, on $\Omega'(\epsilon, n)$, we obtain that $|\theta_n - \Theta_n| \leq C/n^{p+1/2}$, where C depends only on K and k, assuming that $n \geq n_0$. This means that

$$n^{p+1/2}|\theta_n - \Theta_n| \leq C.$$

Thus, for any $\epsilon > 0$, there exists n_0 large enough such that

$$\forall n \geq n_0,\ \widehat{\mathbb{P}}[n^{p+1/2}|\theta_n - \Theta_n| \geq C] \leq \epsilon.$$

which yields the result. □

3.3. The contrast function

To study the maximum likelihood, it is also possible to consider the contrast function

$$Z_n(\theta) = \frac{\exp(L_n(\theta))}{\exp(L_n(0))}.$$

Using the asymptotic development of $L_n(u)$ around 0, we obtain that

$$\log Z_n(\theta) = \theta L_n^{(1)}(0) + \frac{\theta^2}{2}L_n^{(2)}(0) + \mathrm{O}(\theta^3).$$

Thus, with the result of proposition 1, we see that

$$\log Z_n(\theta/n^{1/4}) \xrightarrow[n\to\infty]{\mathcal{F}-\text{stably}} -\mu_2\left(T^{-1/4}\lambda\theta W(\ell_T) - T^{-1/2}\frac{\theta^2}{2}\ell_T\right). \tag{14}$$

From this convergence, we can intuitively check our result in (7), based in the theory of convergence of statistical experiments and the local asymptotically mixed normality property in (14). The theory states (under certain stringent conditions that we do not verify) that the MLE of the prelimit experiments converges stably to the MLE of the limit experiment (Ibragimov & Has'minskii, 1981). It is direct, differentiating with respect to θ in the r.h.s. of (14), to obtain, when $\ell_T > 0$, that the MLE in the limit experiment is $T^{1/4}\lambda W(\ell_T)/\ell_T$. We then obtain (7).

© 2013 Board of the Foundation of the Scandinavian Journal of Statistics.

4. The limit distribution

In this section, we consider that the starting point is 0 and then that $\mathbb{P}[A] = 1$. Otherwise, on the event A, one may use the strong Markov property at the first time the process hit zero.

As $n^{1/4}\theta_n$ and α_n converge to $\lambda\Upsilon_T$ with $\Upsilon_T = T^{1/4}W(\ell_T)/\ell_T$, we give the main characteristics of this random variable. As noted in Remark 2, thanks to the scaling property of the Brownian motion, we see that Υ is equal in distribution to $W(\ell_1)/\ell_1$, whatever the value of T.

Indeed, this random variable is easy to simulate.

Lemma 6. *Under $\mathbb{P}_0$, the distribution of Υ is symmetric. Besides, its density is*

$$f_\Upsilon(x) = \frac{\mathrm{d}F_\Upsilon(x)}{\mathrm{d}x} = \int_0^{+\infty} \mathrm{d}y \int_0^1 \frac{\sqrt{y}}{2\pi\sqrt{t^3}} \exp\left(\frac{-xy}{2} - \frac{y^2}{2t}\right) \mathrm{d}t \tag{15}$$

and it is equal in distribution to

$$\Upsilon = \frac{G(H)}{H} \text{ with } H = \frac{1}{2}\left(U + \sqrt{V + U^2}\right), \tag{16}$$

where $G(H)$, U and V are independent random variables, $G(H) \sim \mathcal{N}(0, H)$, $U \sim \mathcal{N}(0,1)$ and $V \sim \exp(1/2)$.

Proof. It is well known that the local time ℓ_1 at time 1 is equal in distribution to the supremum of the Brownian motion $\sup_{r\in[0,1]} B_r$ on $[0,1]$. It follows that

$$F_{\ell_1}(y) = \mathbb{P}_0[\ell_1 < y] = \mathbb{P}_0\left[\sup_{r\in[0,1]} B_r < y\right] = \mathbb{P}_0[\tau_y > 1],$$

where $\tau_y = \inf\{t > 0 \mid B_t = y\}$. The density $\upsilon(t; y)$ of τ_y is equal to

$$\upsilon(t; y) = \frac{1}{\sqrt{2\pi t^3}} \exp\left(-\frac{y^2}{2t}\right),$$

so that

$$F_{\ell_1}(y) = 1 - \int_0^1 \frac{1}{\sqrt{2\pi t^3}} \exp\left(-\frac{y^2}{2t}\right) \mathrm{d}t,$$

and the density $f_{\ell_1}(y)$ of ℓ_1 is then equal to

$$f_{\ell_1}(y) = \int_0^1 \frac{2y}{\sqrt{2\pi t^3}} \exp\left(-\frac{y^2}{2t}\right) \mathrm{d}t.$$

Thus, conditioning with respect to the value of ℓ_1,

$$F_\Upsilon(x) = \mathbb{P}[\Upsilon < x] = \int_0^{+\infty} \mathbb{P}[W(y) < xy] f(y)\, \mathrm{d}y$$

and this leads to (15).

Expression (16) follows from the equality in distribution of ℓ_1 and $\frac{1}{2}(U + \sqrt{V + U^2})$. This expression has been used in Lépingle (1993) and Lépingle (1995) to simulate the reflected Brownian motion. □

The variance of Υ is 3.16. We see in Fig. 1 that the density of Υ is close to that of the normal distribution, yet narrower.

© 2013 Board of the Foundation of the Scandinavian Journal of Statistics.

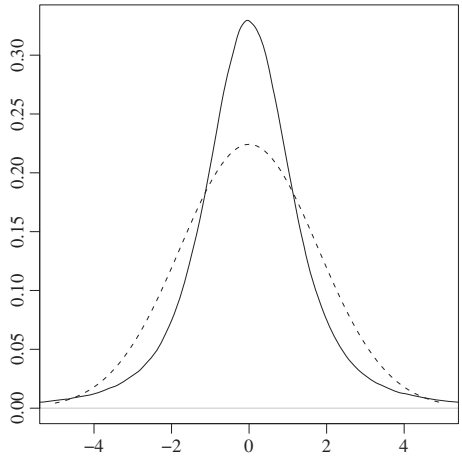


Fig. 1. Density of Υ (solid) and density of the normal distribution with variance $\mathrm{Var}(\Upsilon)$ (dashed).

5. Numerical tests and observations on the estimators

To simplify the notations, let us set

$$\widetilde{\alpha}_n = \frac{\alpha_n}{n^{1/4}} = -\frac{L_n^{(1)}(0)}{L_n^{(2)}(0)} \xrightarrow[n\to\infty]{\text{proba.}} 0.$$

Numerical tests are easy to perform, as all the formulae are easy to implement. For example, the SBm is easily simulated: it could be approximated by random walks (Harrison & Shepp, 1981) or by the exact simulation algorithm proposed in Lejay & Pichot (2012). As pointed out, the estimator $\widetilde{\alpha}_n$ is easily computed from the data, whereas the MLE may be found by a numerical optimization procedure because the density (2) of the SBm has a simple, analytic expression. In Section 4, we saw how to simulate random variates distributed as Υ.

5.1. Numerical behaviour of the estimators for the skew Brownian motion with $\theta \neq 0$

Of course, a natural question that arise is to know the behaviour of the MLE θ_n and of $\widetilde{\alpha}_n$ when the observed points $\{X_i\}_{i=0,\dots,n}$ arise from a SBm of parameter $\theta \neq 0$.

As noted in the introduction of this section, numerical tests could easily be performed for any value of θ, so that one may for example construct some hypothesis test with thresholds given numerical estimation of quantiles. This could indeed give better values than the one obtained by simply using an asymptotic theorem.

In Fig. 2, we plot for values of the skewness parameter θ in the set $\{0, \frac{5}{100}, \dots, \frac{95}{100}\}$ the mean values of θ_n and $\widetilde{\alpha}_n$ over 10,000 paths with $n = 10,000$ samples points. We also plot its $^1/_{10}$, $^1/_4$, $^3/_4$ and $^9/_{10}$ quantiles.

In Fig. 3, we plot the density of θ_n and $\widetilde{\alpha}_n$ for $\theta \in \{0, \frac{5}{10}, \frac{9}{10}\}$ for $n = 10,000$ sample points and constructed from $N = 10,000$ paths. We see that the density of θ_n is skewed for $\theta \neq 0$, as one could expect, and that the variance of θ_n seems to decrease as θ increase to one. This is not the case with $\widetilde{\alpha}_n$.

We see that the average value of θ_n is close θ and that the MLE seems to be a consistent estimator of θ. On the other hand, the mean value of $\widetilde{\alpha}_n$ is close to θ only for small values of θ (say $\theta < {}^5/_{10}$) and overestimates θ for larger values.

© 2013 Board of the Foundation of the Scandinavian Journal of Statistics.

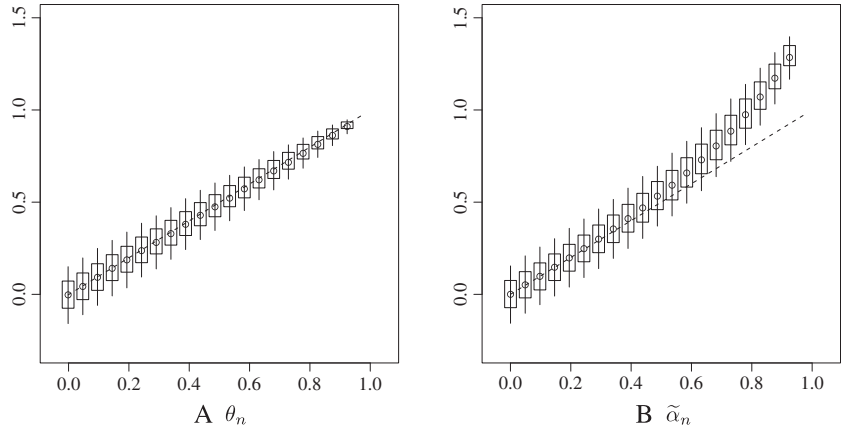


Fig. 2. Statistics of θ_n and $\widetilde{\alpha}_n$: mean value (points), $^1/_4$ and $^3/_4$ quantiles (the upper and lower boundaries of the rectangles) as well as the $^1/_{10}$ and $^9/_{10}$ quantiles (the extremities of the segments) for $n = 10{,}000$ over the 10,000 realizations of paths.

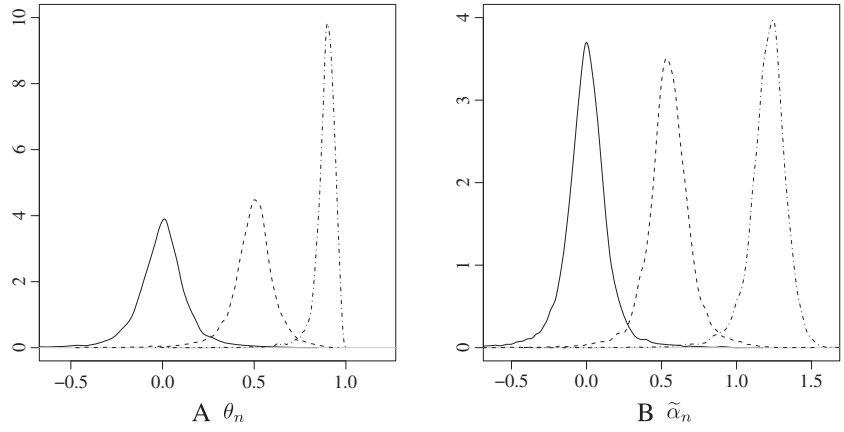


Fig. 3. Density of θ_n and $\widetilde{\alpha}_n$ for $\theta = 0$, $\theta = {}^5/_{10}$ and $\theta = {}^9/_{10}$ and $n = 10,000$ sample points.

5.2. Study of the estimator for small values of n and $\theta = 0$

Theorem 1 asserts that $\theta_n - \widetilde{\alpha}_n$ is of order $O(1/n^{3/4})$. In Table 1, we give the $^1/_{10}$ and $^9/_{10}$ quantiles of $n^{3/4}(\theta_n - \widetilde{\alpha}_n)$. for small values of n. However, the standard deviation increases with n.

We also see in Fig. 4A that the density of $\alpha_n = n^{1/4}\widetilde{\alpha}_n$ remains stable with values of n, even for small values of n. Figure 4B presents Quantile–Quantile plot for α_{100} against $\alpha_{250,000}$, which confirms the closeness of these two distributions. In Fig. 4C, we see that the distribution of $\alpha_{250,000}$ is also close to the one of $\Upsilon = cW(\ell_1)/\ell_1$ with c chosen to obtain equality of the variances.

6. An example of application: diffusion of species

As endowed in the introduction, the SBm is a fundamental tool when one has to model a permeable barrier. In addition, it appears when one writes down the processes generated by

© 2013 Board of the Foundation of the Scandinavian Journal of Statistics.

Table 1. *Quantiles and standard deviation for of* $n^{3/4}(\theta_n - \widetilde{\alpha}_n)$ *over* $N = 10,000$ *paths*

n	100	250	500	1000	2500	5000	10,000	50,000	100,000	250,000
$^1/_{10}$ quantiles	−0.59	−0.76	−1.01	−1.06	−1.15	−1.20	−1.17	−1.34	−0.94	−1.48
$^9/_{10}$ quantiles	0.65	0.89	0.91	0.98	1.11	1.13	1.35	1.34	1.23	1.33
Standard deviation	1.20	1.94	2.89	4.38	6.38	9.95	13.00	31.95	30.80	80.47

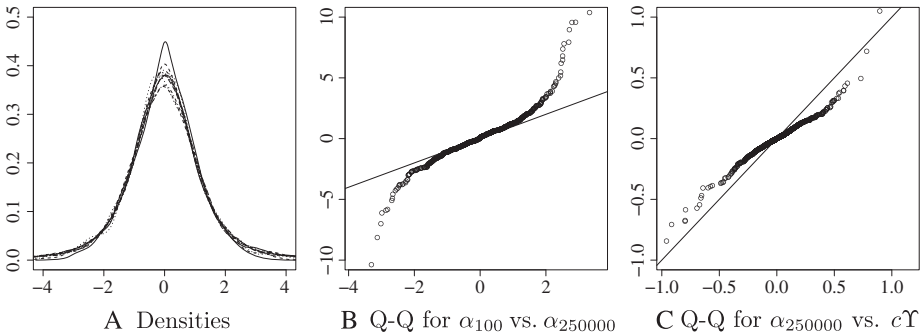


Fig. 4. Densities of $\alpha_n = n^{1/4}\widetilde{\alpha}_n$ for the values of n give in Table 1 and Quantile–Quantile plot of α_n for $n = 250,000$ against $n = 100$ and α_n against $c\Upsilon$ with c such that $\mathrm{Var}(\alpha_n) = \mathrm{Var}(c\Upsilon)$. The straight lines are defined by $y = x$ and show the concordance.

diffusion equations with discontinuous coefficients in a one-dimensional media: this issue is presented in the survey article Lejay (2006) with references to the articles where the SBm arises and covering various fields, such as ecology, finance, astrophysics, geophysics and so on.

We present here a possible application to ecology of our hypothesis test, which can be surely applied to other fields.

6.1. Has a boundary between two habitats an effect?

Diffusions are commonly used in ecology to explain the spread of a specie, at the level of individual cells (see, e.g. the book Berg (1993)) or the level of an animal in a wild environment.

Several authors have proposed the use of biased diffusions to model the behaviour of a specie at the boundary between two habitats, such as Cantrell & Cosner (1999) & Ovaskainen & Cornell (2003), when the species diffuse with different species at speed in each habitat.

Now, consider a situation where the dispersion of a specie in two different habitats is well modelled by a diffusion process, and that the measurement of the diffusion coefficient give the same value. Does it means that the boundary has no effect on the displacement of the individuals?

Let us apply this in a one-dimensional world, where one habitat is $[0, +\infty)$, and the other is $(-\infty, 0]$. We assume that we may track the position of an individual, whose displacement in each of the habitat is given by $x + \sigma B_t$.

Then, we may apply our hypothesis test to determine whether or not the position shall be modelled by

$$(\mathrm{H}_0) \quad X_t = x + \sigma B_t$$

or by

$$(\mathrm{H}_1) \quad X_t = x + \sigma B_t + \theta \ell_t^0(X).$$

© 2013 Board of the Foundation of the Scandinavian Journal of Statistics.

Under hypothesis (H_0), the boundary has no effect and is not seen. Under hypothesis (H_1), the individual is more likely to go in one of the two habitat, depending on the sign of θ.

6.2. What is the underlying operator?

Now, let us consider that we have a measurement of the diffusion coefficients that gives two different values a_+ on $\mathbb{R}_+$ and a_- on $\mathbb{R}_-$.

One may then wonder which differential operator shall be used to model the diffusive behavior. For $a = a_+\mathbf{1}_{[0,+\infty)} + a_-\mathbf{1}_{(-\infty,0)}$, is it

$$L = \frac{1}{2}\nabla(a\nabla\cdot) \text{ or } A = \frac{1}{2}a\Delta \quad ?$$

On $(0, +\infty)$ and $(-\infty, 0)$, there is no difference between these two operators, which means that the local dynamic of the particle/individual is not affected by the choice of L and A. However, the difference arises at 0: the process X generated by L is solution to

$$X_t = x + \int_0^t \sqrt{a(X_s)}\,\mathrm{d}B_s + \frac{a_+ - a_-}{a_+ + a_-}\ell_t^0(X)$$

whereas the process Y generated by A is solution to

$$Y_t = x + \int_0^t \sqrt{a(X_s)}\,\mathrm{d}B_s,$$

for a Brownian motion B (see, e.g. Lejay & Martinez (2006) & Lejay (2006)). From the analytic point of view, the domain $\mathrm{Dom}(A)$ of A contains the functions of class $\mathcal{C}^2(\mathbb{R})$ which are bounded with bounded, first and second order derivatives. The domain $\mathrm{Dom}(L)$ of L contains functions of class $\mathcal{C}^2(\mathbb{R}\setminus\{0\})$ with bounded first and second order derivatives which are furthermore continuous at 0, and such that $a_+\nabla f(0+) = a_-\nabla f(0-)$. This condition is called the *flux condition*. In many physical situations, it is assumed that the flux $a\nabla u$ is continuous and this is why divergence-form operators of type L arise.

Remark 4. Both L and A can be embedded in a single class of operators of type $\frac{\rho}{2}\nabla(a\nabla\cdot)$. If ρ and a are constant on $(0, +\infty)$ and $(-\infty, 0)$, then we may use the following characterization: let us consider

$$C = \frac{1}{2}\nabla(a\nabla\cdot) \text{ with } a = a_+\mathbf{1}_{[0,+\infty)} + a_-\mathbf{1}_{(-\infty,0)}$$

and

$$\mathrm{Dom}(C) = \left\{ f \in \mathcal{C}^2(\mathbb{R}\setminus\{0\}) \;\middle|\; \begin{array}{l} f, f', f'' \text{ are bounded on } \mathbb{R}\setminus\{0\} \\ f(0-) = f(0+) \\ (1+\lambda)f'(0+) = (1-\lambda)f'(0-),\ \lambda \in (-1,1) \end{array} \right\}.$$

This class of operators is then specified by three parameters, $a_+ > 0$, $a_- > 0$ and $\lambda \in (-1, 1)$. The operator A corresponds to $\lambda = 0$, whereas L corresponds to $\lambda = (a_+ - a_-)/(a_+ + a_-)$.

For $\Phi(x) = \int_0^x \mathrm{d}x/\sqrt{a(x)}$, $\widehat{X} = \Phi(X)$ is solution to the SDE (Lejay & Martinez, 2006; Lejay, 2006)

$$\widehat{X}_t = \Phi(x) + B_t + \frac{\sqrt{a_+} - \sqrt{a_-}}{\sqrt{a_+} + \sqrt{a_-}}\ell_t^0(\widehat{X}),$$

© 2013 Board of the Foundation of the Scandinavian Journal of Statistics.

while $\widehat{Y} = \Phi(Y)$ is solution to the SDE

$$\widehat{Y}_t = \Phi(x) + B_t + \frac{\sqrt{a_-} - \sqrt{a_+}}{\sqrt{a_+} + \sqrt{a_-}} \ell_t^0(\widehat{Y}).$$

We then see that both $\widehat{X}$ and $\widehat{Y}$ are SBms, but the coefficients in front of their local time have opposite signs.

Even if we have not studied the asymptotic behavior of the MLE for the SBm with skewness parameter different from 0, numerical experiments back the following hypotheses test:

(i) Given an observed X, estimate the diffusion coefficient for the process on each side of 0.
(ii) Apply the function Φ to the observed process.
(iii) Compute the MLE θ_n of the Skewness parameter. If $a_+ > a_-$ (resp. $a_+ < a_-$) and $\theta_n > 0$ then decided that the infinitesimal generator of X is L (resp. A). Otherwise, decide that it is A (resp. L).

7. Conclusion

In this article, we have studied the behavior of the maximum likelihood for the SBm when the parameter to estimate is 0.

In particular, we have shown that the rate of convergence of the estimator θ_n is $n^{1/4}$ and not $n^{1/2}$ as in the classical case. This should not be surprising: indeed, away from 0, the SBm behaves like a Brownian motion, and only its dynamic close to 0 allows one to see the difference between a SBm with a parameter $\theta \neq 0$ and a Brownian motion. It is also not surprising that the local time enters in the limit distribution.

The case $\theta \neq 0$ remains open. One needs to prove results similar to the one of Jacod (1998), when the Brownian motion is replaced by the SBm (its distribution with respect to the Wiener measure is singular). Of course, one cannot expect the limit law to be symmetric. Yet, it is pretty easy to simulate the SBm and to estimate the maximum likelihood, so that numerical studies are easy to perform.

Appendix

In this Appendix, we provide the theorems given in Jacod (1998) used for the proofs of the main results in Section 2. We slightly change the notation and present the results in the particular cases that are relevant to us in the present work.

Denote by $X = \{X_t : 0 \leq t \leq 1\}$ a Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Introduce a Borel function $h : \mathbb{R}^2 \to \mathbb{R}$ such that there exist $a \in \mathbb{R}$ and $\hat{h} : \mathbb{R} \to \mathbb{R}$ such that

$$h(x, y) \leq e^{a|y|} \hat{h}(x) \quad \text{and} \quad \int |x|^r |\hat{h}(x)| \, \mathrm{d}x < \infty. \tag{17}$$

Theorem 2 (Jacod (1998, Theorem 1.1, p. 508)). *Consider h as earlier, satisfying (17) with $r = 0$. Then,*

$$\frac{1}{n^{1/2}} \sum_{i=0}^{n-1} h\left(\sqrt{n} X_{i/n}, \sqrt{n}(X_{(i+1)/n} - X_{i/n})\right) \xrightarrow[n\to\infty]{\text{prob.}} c(h)\ell_1, \tag{18}$$

where

$$c(h) = \iint_{\mathbb{R}^2} h(x, y) p(1, y) \, \mathrm{d}x \, \mathrm{d}y, \tag{19}$$

and $(\ell_t)_{t\geq 0}$ denotes the local time of X at level zero.

© 2013 Board of the Foundation of the Scandinavian Journal of Statistics.

Remark 5. It must be noticed that the convergence in (18), as stated in Jacod (1998), is stronger, in the sense that both terms in (18) are processes (i.e. depend on t) and the convergence is locally uniformly in time, in probability. Recall that a sequence $(Z^n)_{n\geq 1}$ of processes is said to *converges locally uniformly in time, in probability*, to a limiting processes Z if for any $t \in \mathbb{R}^+$ the sequence $\sup_{s\leq t} |Z^n_s - Z_s|$ goes to 0 in probability.

Let us assume now that h is such that $c(h) = 0$. Then, set

$$C(h) = c(h^2) + 2\int_{\mathbb{R}}\int_{\mathbb{R}} p(1, y)h(x, y)H(x + y)\,\mathrm{d}y\,\mathrm{d}x \tag{20}$$

with

$$H(x) = \sum_{k=0}^{+\infty} \int_{\mathbb{R}} p(k, y) \int_{\mathbb{R}} p(1, z)h(x + y, z)\,\mathrm{d}z\,\mathrm{d}y.$$

Theorem 3 (Jacod (1998, Theorem 1.2, p. 511))**.** *Consider h as earlier, satisfying (17) with some $r > 3$, and assume that $c(h) = 0$ (see (19)). Then,*

$$\frac{1}{n^{1/4}} \sum_{i=0}^{n-1} h\left(\sqrt{n}X_{i/n}, \sqrt{n}(X_{(i+1)/n} - X_{i/n})\right) \xrightarrow[n\to\infty]{\mathcal{F}-\text{stable in dist.}} \sqrt{C(h)}W(\ell_1), \tag{21}$$

where $W = \{W_t : t \geq 0\}$ is a Brownian motion independent of X, and $(\ell_t)_{t\geq 0}$ is the local time of X at level zero. The constant $C(h)$ is given by (20).

Remark 6. As in the previous remark, the theorem stated in Jacod (1998) is stronger, now in the sense that both terms in (21) are processes, and the processes converge stably in distribution in the Skorokhod space of càdlàg functions.

Acknowledgements

Antoine Lejay has been supported by the MathAmSud programme and the Inria Equipe Associée Anestoc-Tosca. Ernesto Mordecki's research is supported by CMAT-UDELAR, and by project *SAMP Network*, of the MathAmSud programme. Soledad Torres was supported by Dipuv grant 26/2009, Fondecyt Grant N. 1130586, ACT1112, CONICYT-PIA: Stochastic Analysis Research Network.

References

Aldous, D. J. & Eagleson, G. K. (1978). On mixing and stability of limit theorems. *Ann. Probability* **6**, (2), 325–331.

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

Berg, H. (1993). *Random walks in biology*, Princeton University Press.

Cantrell, R. & Cosner, C. (1999). Diffusion models for population dynamics incorporating individual behavior at boundaries: Applications to refuge design. *Theoretical Population Biology* **55**, (2), 189–207.

Florens-Zmirou, D. (1993). On estimating the diffusion coefficient from discrete observations. *J. Appl. Probab.* **30**, (4), 790–804.

Fujihara, E., Kawamura, Y. & Yano, Y. (2007). Functional limit theorems for occupation times of Lamperti's stochastic processes in discrete time. *J. Math. Kyoto Univ.* **47**, (2), 429–440.

© 2013 Board of the Foundation of the Scandinavian Journal of Statistics.

Harrison, J. M. & Shepp, L. A. (1981). On skew Brownian motion. *Ann. Probab.* **9**, (2), 309–313.
Höpfner, R. & Löcherbach, E. (2003). Limit theorems for null recurrent Markov processes. *Mem. Amer. Math. Soc.* **161**, (768).
Ibragimov, I. & Has'minskii, R. (1981). *Statistical estimation asymptotic theory*, Springer.
Jacod, J. & Shiryaev, A. N. (1987). *Limit theorems for stochastic processes*, Springer-Verlag.
Jacod, J. (1998). Rates of convergence to the local time of a diffusion. *Ann. Inst. H. Poincaré Probab. Statist.* **34**, (4), 505–544, DOI 10.1016/S0246-0203(98)80026-5.
Jacod, J. (2006). Parametric inference for discretely observed non-ergodic diffusions. *Bernoulli* **12**, (3), 383–401, DOI 10.3150/bj/1151525127.
Kutoyants, Y. (1984). *Parameter estimation for stochastic processes*, Heldermann.
Lamperti, J. (1958). An occupation time theorem for a class of stochastic processes. *Trans. Amer. Math. Soc.* **88**, 380–387.
Le Cam, L. & Yang, G. L. (2000). *Asymptotics in statistics*, (2nd ed.)., Springer Series in Statistics, Springer-Verlag.
Le Gall, J. -F. (1984). 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; 51–82.
Lejay, A. (2006). On the constructions of the skew Brownian motion. *Probab. Surv.* **3**, 413–466, DOI 10.1214/154957807000000013.
Lejay, A. & Martinez, M. (2006). A scheme for simulating one-dimensional diffusion processes with discontinuous coefficients. *Ann. Appl. Probab.* **16**, (1), 107–139, DOI 10.1214/105051605000000656.
Lejay, A. & Pichot, G. (2012). Simulating diffusion processes in discontinuous media: a numerical scheme with constant time steps. *J. Comput. Phys.* **231**, (21), 7299–7314, DOI 10.1016/j.jcp.2012.07.011.
Lépingle, D. (1995). Euler scheme for reflected stochastic differential equations. *Math. Comput. Simulation* **38**, (1-3), 119–126, DOI 10.1016/0378-4754(93)E0074-F.
Lépingle, D. (1993). Un schéma d'Euler pour équations différentielles stochastiques réfléchies. *C. R. Acad. Sci. Paris Sér. I Math.* **316**, (6), 601–605.
Lipster, R. & Shiryaev, A. (2001). *Statistics of random processes. II. Applications*, Springer.
Martinez, M. (2004). *Interprétations probabilistes d'opérateurs sous forme divergence et analyse de méthodes numériques associées*, Ph.D. thesis, Université de Provence / INRIA Sophia-Antipolis.
Ovaskainen, O. & Cornell, S. J. (2003). Biased movement at a boundary and conditional occupancy times for diffusion processes. *J. Appl. Probab.* **40**, (3), 557–580.
Ramirez, J. M., Thomann, E. A., Waymire, E. C., Haggerty, R. & Wood, B. (2006). A generalized Taylor-Aris formula and skew diffusion. *Multiscale Model. Simul.* **5**, (3), 786–801, DOI 10.1137/050642770.
Rényi, A. (1963). On stable sequences of events. *Sankhyā Ser. A* **25**, 293–302.
Walsh, J. (1978). A diffusion with discontinuous local time. In *Temps locaux*, vol. 52–53, Astérisques Société Mathématique de France; 37–45.
Zhang, M. (2000). Calculation of diffusive shock acceleration of charged particles by skew Brownian motion. *The Astrophysical Journal* **541**, 428–435.

Received December 2010, in final form April 2013

Antoine Lejay, Inria, Tosca, Villers-lès-Nancy, F-54600, France.
E-mail: Antoine.Lejay@univ-lorraine.fr

© 2013 Board of the Foundation of the Scandinavian Journal of Statistics.

Copyright of Scandinavian Journal of Statistics is the property of Wiley-Blackwell and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.