



**HAL**  
open science

# Estimation of the biais parameter of the Skew Random Walk and application to the Skew Brownian motion

Antoine Lejay

► **To cite this version:**

Antoine Lejay. Estimation of the biais parameter of the Skew Random Walk and application to the Skew Brownian motion. 2016. hal-01319319v1

**HAL Id: hal-01319319**

**<https://inria.hal.science/hal-01319319v1>**

Preprint submitted on 20 May 2016 (v1), last revised 12 May 2017 (v4)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Estimation of the bias parameter of the Skew Random Walk and application to the Skew Brownian motion

Antoine Lejay\*

May 20, 2016

## Abstract

We study the asymptotic property of simple estimator of the parameter of a Skew Brownian motion when one observes its positions on a fixed grid — or equivalently of a simple random walk with a bias at 0. This estimator, nothing more than the Maximum Likelihood Estimator, is based only on the number of passages of the random walk at 0. It is very simple to set up, is consistent and satisfies the local asymptotic mixed normality. We believe that this simplified framework is helpful to understand the asymptotic behavior of the maximum likelihood of the SBM observed at discrete time which is studied in a companion paper.

## Keywords

skew Brownian motion, skew random walk, maximum likelihood estimator, local asymptotic mixed normality, local time

A flourishing literature is devoted to the stochastic diffusion processes with discontinuous coefficients or involving the local time. This covers many fields of application, ranging from geophysics [2] to population ecology [1, 8] through molecular biology [6], or to study rank based statistics [16] as well as dealing with the tiled volatility model [33] in mathematical finance.

---

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

In this direction of research, the Skew Brownian motion (SBM) is a process of utter importance. It plays the role of the Brownian motion for diffusions with continuous coefficients. Informally, the SBM is similar to the Brownian motion, excepted that when it hits 0, its behavior depends on a parameter  $\alpha \in (0, 1)$  which gives it a trend to go on the positive or negative axis according to  $\alpha > 1/2$  or  $\alpha < 1/2$ . As pointed out for example in [30], Lamperti's type transforms allow to reduce one-dimensional diffusion processes generated by either  $a\Delta$  or  $\nabla(a\nabla\cdot)$  with  $a$  piecewise constant to processes which behave locally as SBM.

Formally, the SBM of parameter  $\alpha$  is the strong solution to the stochastic differential equation

$$X_t = B_t + \theta L_t, \quad \theta = \frac{1 + \alpha}{2} \in (-1, 1), \quad (1)$$

where  $B$  is a Brownian motion and  $L_t$  the symmetric local time of  $X$  at 0. Introduced through its infinitesimal generator by K. Itô and H.P. McKean (See *e.g.* [20, p. 220] or [21, Problem 1, p. 115]), there exists many ways to describe this process (See [29] for a survey of some possible constructions) The existence of a strong solution to (1) has been proved first by J.M. Harrison and L.A. Shepp in [18].

In the same way the Brownian motion is approximated by a simple random walk properly renormalized, the SBM is approximated by a Skew Random Walk (SRW) under the same scaling in space and time [18] (the SBM and the Brownian motion shares the same scaling property). A SRW is a simple random walk  $\{\zeta_k\}_{k \geq 0}$  with a bias at 0:

$$\mathbb{P}[\zeta_{k+1} = i | \zeta_k = j] = \begin{cases} 1/2 & \text{if } j \neq 0, i = j \pm 1, \\ \alpha = (1 + \theta)/2 & \text{if } j = 0, i = 1, \\ 1 - \alpha = (1 - \theta)/2 & \text{if } j = 0, i = -1, \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

Let us fix a parameter  $h > 0$  which represents the size of a grid. The successive positions  $\xi_k(h)$  at which a SBM  $X$  passes through the grid  $h\mathbb{Z}$ , up to time  $T$ , are observed and recorded (the number of recorded points is random  $n^*(h)$ , yet close to  $T/h^2$ ). By doing so, the embedded Markov chain  $\{\xi_k(h)\}_{k \geq 0}$  is equal in distribution to the rescaled SRW  $\{h\zeta_k\}_{k \geq 0}$ .

Our main concern is to estimate  $\alpha = (1 + \theta)/2$  from these observations.

From now, we denote by  $\mathbb{P}_\alpha$  the distribution of the SBM of parameter  $\alpha \in (0, 1)$ .

Although  $\mathbb{P}_\alpha[X_t \geq 0] = \alpha$  for any  $t > 0$ , the mean occupation time above 0 up to time  $t$  follows a variant of the Arc Sine distribution [26, Theorem 1].

Its variance remains constant with  $t$ , leading to a meaningless estimator. As the SBM is a null recurrent process, ergodic theory could not be used.

The sole information useful for our estimators are the successive points  $(\xi_k(h), \xi_{k+1}(h))$  such that  $\xi_k(h) = 0$ . The excursions of the SRW behave like the ones of a simple random walk. They carry no information about  $\alpha$ , excepted from their sign.

Our estimator of  $\alpha$  is then simply

$$\alpha(h) = \frac{\#\text{observed upcrossings before } T}{\#\text{observed passages to 0 before } T}, \quad (3)$$

where a passage to 0 counts to 1 when  $\xi_k(h) = 0$  and an upcrossing counts to 1 when  $\xi_k(h) = 0$  and  $\xi_{k+1}(h) = 1$ .

This estimator (3) is also the mean number of positive excursions. It is the Maximum Likelihood Estimator (MLE) of  $\alpha$  for the SRW.

First, we show that  $\alpha(h)$  is consistent, *i.e.* converges in probability to  $\alpha$  as  $h \rightarrow 0$  under  $\mathbb{P}_\alpha$  for any  $\alpha \in (0, 1)$ , and that

$$h^{-1/2}(\alpha(h) - \alpha) \text{ converges in distribution to } \sqrt{\alpha(1-\alpha)} \frac{W(L_T)}{L_T},$$

where  $W$  is a Brownian motion independent from  $B$ . Asymptotically,  $\alpha(h)$  behaves like a mixed normal distribution. For  $n$  observations, the rate of convergence is not  $n^{-1/2}$ , but  $n^{-1/4}$ . The reason is that the number of passages to 0 is roughly of order  $\sqrt{n}$ . We then show that the distribution of the  $\{\xi_k(h)\}_{k=0, \dots, n_s}$  satisfies the Local Asymptotic Mixed Normal (LAMN) property [23, 24] — a generalization of the Local Asymptotic Normal (LAN) property [19, 28] — which carries some information about any asymptotic efficient estimator.

We are in a situation close to the one studied by D. Florens in [17] where non-parametric estimation of the diffusion coefficient of the solution of a stochastic differential equation observed at discrete time is performed through the number of crossings of levels. The kind of limit theorem we obtain here is typical for of null recurrent “regular” diffusion processes [13], although the approach shall be adapted to deal with distributions which are singular with respect to the Wiener measure [27].

In [31], we have studied the MLE estimator for the SBM observed at frequency  $T/n$  when  $\alpha = 1/2$  (or  $\theta = 0$ ). By using results from J. Jacod in [22], we obtained similar results, both to identify the limit and the rate of convergence. The consistency of the estimator when  $\alpha \neq 0$ , which was left open, is now treated in the subsequent article [32]. However, following [22], these articles use fine estimates on the semi-groups and martingales limit theorems.

As shown in [32], the MLE estimator found there could be seen as a “weighted” version of the computation of the expected number of crossings when one sees the actual crossing between two observations time as hidden data. With the record of passages, the situation becomes rather simple as the convergence of the estimator follows the convergence of the binomial estimator with a random number of samples, with the CLT-like results steaming from the Berry-Esseen theorem and the P. Lévy downcrossing theorem, which relates the local time as a limit of the renormalized number of crossings of the SRW.

We then believe that the results in this article explain with simple computations and elementary tools the results in [31, 32]. Also, the approach developed here could be used for other statistics involving excursions.

**Outline.** Section 1 gives a quick account on the SBM and the SRW. The asymptotic of the estimator  $\alpha(h)$  in (3) is studied Section 2. In Section 3, the LAMN property is proved.

## 1. Approximation of the Skew Brownian motion by the Skew Random Walk

Let  $X$  be SBM of parameter  $\alpha \in (0, 1)$  with distribution  $\mathbb{P}_\alpha$  and  $X_0 = 0$ , which is the strong solution to (1) [18]. The *symmetric* local time of  $X$  at 0 is defined as  $L = \mathcal{L}(Y)$  when for a semi-martingale  $Y$ ,

$$\mathcal{L}_t(Y) = \lim_{\epsilon \rightarrow 0} \frac{1}{2\epsilon} \int_0^t \mathbf{1}_{Y_s \in [-\epsilon, \epsilon]} d\langle Y \rangle_s, \quad t \geq 0. \quad (4)$$

All the results we need here on the local time could be found for example<sup>1</sup> in [21] or [37].

For the SBM,  $\langle X \rangle_t = t$  and  $L = \mathcal{L}(|X|)$  whatever  $\alpha \in (0, 1)$ .

**Proposition 1** (Scaling property). *For any  $c > 0$ , with  $X^c = cX_{./c^2}$ ,  $(X^c, \mathcal{L}(X^c))$  is equal in distribution to  $(X, L)$  when  $X_0 = 0$ .*

*Proof.* With (4),  $\mathcal{L}(X^c) \stackrel{\text{law}}{=} cL_{./c^2}$ , so that  $X_{t/c^2}^c = cB_{t/c^2} + \theta cL_{t/c^2}$  and then and then  $X_{t/c^2}^c = W_t + \theta \mathcal{L}_t(X^c)$  for the Brownian motion  $W = cB_{./c^2}$ . The equality in distribution of  $X^c$  to  $X$  follows from the uniqueness to the solution to (1).  $\square$

*Hypothesis 1.* We consider that  $X_0 = 0$  so that  $\xi_0(h) = 0$ .

<sup>1</sup>The reader shall be warned that different normalizations are used in different textbooks. For example, in [37, Chapter VI], the right local time is used. For the SBM, this distinction is crucial since  $a \mapsto L_t^a(X)$  is discontinuous at 0 [39].

This hypothesis aims at simplifying the presentation of the results. Otherwise, we have to wait the random time  $\tau$  at which the SBM hits first 0. The observation timespan is then reduced from  $[0, T]$  to  $[\tau, T]$ .

For  $h > 0$ , define iteratively  $\tau_0(h) = 0$  and

$$\tau_{k+1}(h) = \inf\{t > \tau_k(h); |X_t - X_{\tau_k(h)}| = h\},$$

the successive passage times and positions of  $X$  to the grid  $h\mathbb{Z}$ . Also, set  $\xi_k(h) = X_{\tau_k(h)}$ , the normalized passage positions. Since the SBM is a strong Markov process,  $(\xi_k(h))_{k \geq 0}$  is a Markov chain with values in  $\mathbb{Z}$ .

The following result is an easy consequence of the results in [18, 39] on the scale function and the speed measure of the SBM, or from the construction from K. Itô and H.P. McKean [21] through its excursions.

**Proposition 2** (Embedding a Markov chain into a SBM). *The discrete process  $(\xi_k(h))_{k \geq 0}$  is a recurrent Markov chain with values in  $h\mathbb{Z}$ , and  $(h^{-1}\xi_k(h))_{k \geq 0}$  has the same probability transition as the SRW  $(\zeta_k)_{k \geq 0}$  whose transition kernel is defined by (2).*

*In addition,  $\tau_{k+1}(h) - \tau_k(h)$  is independent from the  $(\xi_k)_{k \geq 0}$  and  $(\tau_{k+1}(h) - \tau_k(h))_{k \geq 0}$  is a sequence of independent and identically distributed random variables, with  $\mathbb{E}_\alpha[\tau_{k+1}(h) - \tau_k(h)] = h^2$ .*

*Proof.* The SBM is a recurrent, strong Markov process so that  $(\xi_k(h))_{k \geq 0}$  is a recurrent Markov chain. Besides, if  $\xi_k(h) \neq 0$ , then the process behaves like a Brownian motion between the times  $\tau_k(h)$  and  $\tau_{k+1}(h)$ . On the other hand,  $\mathbb{P}_\alpha[\xi_{k+1}(h) = h | \xi_k(h) = 0]$  is the probability that the first excursion of  $X$  whose height is greater than  $k$  is positive, and this probability is equal to  $\alpha$ , an easy consequence from the result from K. Itô and H.P. McKean on the SBM [21, p. 115].  $\square$

Generalizing the Donsker theorem, J.M. Harrison and L.A. Shepp [18] proved first that the SBM is the limit of a SWR (See also [15, 27]). Hence, the piecewise linear interpolation of  $(\xi_k(h))_{k \geq 0}$  along the times  $(\tau_k(h))_{k \geq 0}$  converges in probability to  $X$ .

## 2. An excursions based estimator

**The data.** We consider that given a path of a SBM  $X$ , we record the positions on the grid of the embedded chain  $\boldsymbol{\xi}(h) = (\xi_k(h))_{k=0, \dots, n^*(h)}$ . From these observations, we construct an estimator  $\alpha(h)$  of  $\alpha \in (0, 1)$ .

Indeed, using the scaling property, we may assume that we record either  $(\xi_k(1))_{k=0, \dots, T/h^2}$  (long time horizon) or  $\boldsymbol{\xi}(h)$  for  $\tau_k \leq T$  (fixed time horizon). We state only our result for the latter.

---

$X_t$	$\stackrel{\text{law}}{=} hX_{t/h^2}$	SBM, same scaling of for Brownian motion
$L_t$	$\stackrel{\text{law}}{=} hL_{t/h^2}$	Local time at 0, same scaling as for Brownian local time
$\tau_k(h)$	$\stackrel{\text{law}}{=} h^2\tau_k(1)$	Sum of iid rv, $\mathbb{E}_\alpha[\tau_1(1)] = \mathbb{E}_\alpha[\tau_1(1)^2] = 1$
$\xi_k(h)$	$\stackrel{\text{law}}{=} h\xi_k(1)$	Biased random walk
$N_n(h)$	$\stackrel{\text{law}}{=} N_n(1)$	Number of zeros in $n$ steps, $N_n(h)/\sqrt{n} \xrightarrow[n \rightarrow \infty]{\text{dist.}} Y \stackrel{\text{law}}{=} L_1$
$N_n^+(h)$	$\stackrel{\text{law}}{=} N_n^+(1)$	Number of upcrossings in $n$ steps
$n^*(h)$	$\approx Th^{-2}$	Number of steps of the walk before $T$

---

Table 1: Main variables and their scaling properties.

Embedding the family SRW into a SBM allows one to study the limit of  $\alpha(h)$  as the parameter  $h$  decreases to 0. For this, the underlying probability space is the one of the SBM. Of course, our estimator is also suitable when one observes only a SRW. In this case, the estimation in long time horizon has to be considered, and the results are similar in nature (the local time has to be replaced by the limit of  $n^{-1/2} \sum_{k=0}^n \mathbf{1}_{\xi_k(1)=0}$ , which has the same distribution as the local time. See [36, Theorem 9.11, p. 99] and also Lemma 3 by noting that  $L_n/\sqrt{n} \stackrel{\text{law}}{=} L_1$  and Remark 6).

**The estimator.** For  $n \geq 0$ , we set

$$N_n(h) = \sum_{k=0}^{n-1} \mathbf{1}_{\xi_k(h)=0} \text{ and } N_n^+(h) = \sum_{k=0}^{n-1} \mathbf{1}_{\xi_k(h)=0, \xi_{k+1}(h)=1}.$$

As the random walk  $(\xi_k(h))_{k \geq 0}$  is recurrent,  $N_n(h) \rightarrow +\infty$  as  $n \rightarrow \infty$  almost surely.

Finally, for  $T > 0$ , we set  $n^*(h) = \sup\{k \geq 0 \mid \tau_k(h) \leq T\}$ . From the Strong Law of Large Numbers,  $\lim_{n \rightarrow \infty} \tau_n(h)/n = \mathbb{E}_\alpha[\tau_1(h)] = h^2$  almost surely. As  $\tau_{n^*(h)}(h) \leq T < \tau_{n^*(h)+1}(h)$ ,  $h^2 n^*(h)$  converges almost surely to  $T$  as  $h \rightarrow 0$ .

Table 1 summarizes the main variables and their scaling properties, which is justified by Proposition 1.

Our estimator of  $\alpha$  is

$$\alpha(h) = \frac{N_{n^*(h)}^+(h)}{N_{n^*(h)}(h)}, \tag{5}$$

which is formalization of (3).

*Remark 1.* The likelihood is

$$\text{Lik}_\alpha(\boldsymbol{\xi}(h)) = \prod_{k=0}^{n^*(h)-1} p_\alpha(\xi_k(h), \xi_{k+1}(h)) = \alpha^{N_n^+(h)} (1 - \alpha)^{N_n(h) - N_n^+(h)}$$

and the log-likelihood is

$$\text{Log-Lik}_\alpha(\boldsymbol{\xi}(h)) = N_n^+(h) \log(\alpha) + (N_n(h) - N_n^+(h)) \log(1 - \alpha).$$

Differentiating the log-likelihood with respect to  $\alpha$  to solve  $\partial_\alpha \text{Log-Lik}_{\alpha(h)}(\boldsymbol{\xi}(h)) = 0$  implies that  $\alpha(h)$  given by (5) is the maximum likelihood estimator (MLE)

$$\alpha(h) = \operatorname{argmax}_\alpha \text{Log-Lik}_\alpha(\boldsymbol{\xi}) = \frac{N_{n^*(h)}^+(h)}{N_{n^*(h)}(h)}.$$

**Theorem 1** (Convergence of the estimator  $\alpha(h)$ ). *For any  $\alpha \in (0, 1)$ , under  $\mathbb{P}_\alpha$ ,*

$$\alpha(h) \xrightarrow[h \rightarrow 0]{a.s.} \alpha, \tag{6}$$

*implying in particular that  $\alpha(h)$  is a consistent estimator of  $\alpha$ . In addition, there exists a Brownian motion  $W$  independent from  $B$  such that*

$$\frac{1}{\sqrt{h}} (\alpha(h) - \alpha) \xrightarrow[h \rightarrow 0]{dist.} \Upsilon_T = \sqrt{\alpha(1 - \alpha)} \frac{W(L_T)}{L_T}.$$

*Moreover, for some constant  $C$ ,*

$$\sup_{x \in \mathbb{R}} |\mathbb{P}[h^{-1/2}(\alpha_h - \alpha) \leq x] - \mathbb{P}[\Upsilon_T \leq x]| \leq Ch^{\frac{2}{9}}.$$

*Remark 2.* Using the scaling property on both the Brownian motion and the local time,  $\Upsilon_T \stackrel{\text{law}}{=} T^{-1/4} \Upsilon_1$  for any  $T > 0$ . See [31] for more results about this random variable.

*Remark 3.* Since our estimator relies on the convergence of a binomial random variate toward the Gaussian distribution, one could make use of the large literature on confidence intervals, even for small samples, which are better than the Wald confidence interval based on the Gaussian approximation: See *e.g.* [5, 7, 38] by randomizing the number of samples. In particular, for  $\alpha \neq 0$ , these specific approaches provide non symmetric confidence interval around the estimator for a finite number of samples, although the limiting distribution of  $h^{-1/2}(\alpha(h) - \alpha)$  is symmetric. This was observed empirically on the numerical tests in [30] on the MLE for the SBM.



*Remark 4.* In this estimator, we record around  $n \approx T/h^2$  points of the SBM, we really use  $N_{n^*(h)}(h)$  points for the estimation, as the  $\xi_k(h)$ 's bring no information when the walk is away from 0. The number of points used for the estimation is of order  $\sqrt{n} \approx \sqrt{T}/h$ , and the rate of convergence in our CLT like result is  $h^{1/2}$ , that is of order  $n^{-1/4}$ . We are then not in the “usual” situation where the rate of convergence is  $n^{-1/2}$ . These are the same phenomena as the one of D. Florens [17] for estimators for stochastic differential equations based on crossings, and the results of J. Jacod [22].

The convergence of  $\alpha(h)$  in (6) depends on the next lemma and the almost sure convergence of  $n^*(h)(h)$  to infinity.

**Lemma 1** (Distribution of  $N_n^+(h)$  given  $N_n(h)$ ). *Under  $\mathbb{P}_\alpha$ , the distribution of  $N_n(h)$ ,  $N_n^+(h)$  given  $N_n(h)$  is a binomial distribution with parameters  $(\alpha, N_n(h))$ , so that*

$$\frac{N_n^+(h)}{N_n(h)} \xrightarrow[n \rightarrow \infty]{} \alpha \text{ almost surely.} \quad (7)$$

The convergence of  $h^{-1/2}(\alpha(h) - \alpha)$  follows from the Berry-Esseen theorem, whose application is immediate because of the conditional independence of  $N_n^+(h)$  with respect to  $N_n(h)$ .

**Lemma 2** (Application of the Berry-Esseen inequality [4, 14]). *Let  $\Phi$  be the cumulative distribution of the Gaussian distribution  $\mathcal{N}(0, 1)$ . Then with  $\sigma = \sqrt{\alpha(1 - \alpha)}$ ,*

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P}_\alpha \left[ \frac{N_n^+(h) - N_n(h)\alpha}{\sigma \sqrt{N_n(h)}} \leq x \mid N_n(h) \right] - \Phi(x) \right| \leq \frac{C\alpha}{\sigma^3 \sqrt{N_n(h)}}, \quad (8)$$

where  $C$  is a constant.

*Remark 5.* Many articles have been devoted to exhibit the best constant  $C$  in the convergence of the Berry-Esseen theorem for binomial distribution (See e.g. [34, 35] and the references within).

From now, the result stem from two ingredients: (i) the convergence of  $hN_{n^*(h)}$  toward  $L_T$ , which is granted by the P. Lévy's downcrossing theorem (See e.g. [10, 21, 25] or [37, Exercice VI.1.19] and references within for refinements of the statement). (ii) a control on  $\mathbb{E}_\alpha[1/\sqrt{N_{n^*(h)}(h)}]$  which follows from the convergence of  $N_n(h)/\sqrt{n}$  (more precise statements are given in [9] which contains some almost sure control of  $N_n(h)/\sqrt{n}$  for  $n$  large enough).

**Lemma 3** (P. Lévy downcrossing theorem). *Almost surely,  $hN_{n^*(h)}(h)$  converges to  $L_T$  as  $h \rightarrow 0$ .*

*Besides, for some constant  $C$ ,*

$$\mathbb{E}_\alpha \left[ \left| \frac{N_n(1)}{\sqrt{n}} - \frac{L_n}{\sqrt{n}} \right|^2 \right]^{1/2} \leq \frac{C}{n^{1/4}}. \quad (9)$$

*Remark 6.* The cumulative distribution function of the local time  $L_1$  is  $\mathbb{P}[L_1 \leq x] = \sqrt{2/\pi} \int_0^x \exp(-y^2/2) dy$  (See *e.g.* [36, Theorem 9.1, p. 101]), a consequence of the reflection principle and the identity in distribution between  $L_1$  and the supremum of a Brownian motion on  $[0, 1]$ .

*Remark 7.* It is proved in [11] (see also [36, Consequence 10.3, p. 110]) that for any  $\epsilon > 0$ ,  $N_n(1) - L_n = o(n^{1/4+\epsilon})$  (See [3, 12] for more precise bounds). Thus, the rate of convergence of the  $L^2$ -norm of  $1/n^{1/4}$  in (9) is “nearly” optimal.

*Proof.* The first statement is the P. Lévy’s downcrossing theorem, so that we skip its proof. It is valid for the SBM as it depends only on the absolute value of  $X$ , which is a Brownian motion in distribution.

Let us now consider that  $h = 1$  and we drop reference to  $h$ .

The Itô-Tanaka formula yields that

$$|X_t| = |X_s| + \int_s^t \operatorname{sgn}(X_r) dX_r + L_t - L_s.$$

Since  $|a| - |b| \leq |a - b|$ ,

$$|L_n - L_{\tau_n}| \leq |X_n - X_{\tau_n}| + \int_{\tau_n}^n \operatorname{sgn}(X_r) dX_r.$$

Thus,

$$\mathbb{E}_\alpha[|L_n - L_{\tau_n}|^2] \leq 2\mathbb{E}_\alpha[|X_n - X_{\tau_n}|^2] + 2\mathbb{E}_\alpha[|\tau_n - n|]$$

Since  $(X_t - X_r)^2 = \int_s^t X_r dX_r + (t - s)$  and  $\int_s^t X_r dL_r = \int_s^t X_r \mathbf{1}_{X_r=0} dL_r = 0$ ,

$$\frac{1}{n} \mathbb{E}_\alpha[|L_n - L_{\tau_n}|^2] \leq 4\mathbb{E}_\alpha \left[ \left| \frac{\tau_n}{n} - 1 \right| \right] \leq \frac{4}{n},$$

as  $\tau_n$  is the sum of  $n$  independent random variables of mean 1 and variance 1.

On the other hand, it is easily seen that

$$|\xi_{k+1}| = |\xi_k| + \operatorname{sgn}(\xi_k)(\xi_{k+1} - \xi_k) + \mathbf{1}_{\xi_k=0}. \quad (10)$$

Comparing (2) with (10), as  $\xi_k = X_{\tau_k}$ ,

$$N_n = L_{\tau_n} + M_n \text{ with } M_n = \int_0^{\tau_n} (\operatorname{sgn}(X_r) - \operatorname{sgn}(X_{\pi(r)})) dX_r$$

where  $\pi(s) = \tau_k$  when  $\tau_k \leq s < \tau_{k+1}$ . It follows that

$$\begin{aligned} \frac{1}{n} \mathbb{E}_\alpha[M_n^2] &= \frac{1}{n} \mathbb{E}_\alpha \left[ \int_0^{\tau_n} (\operatorname{sgn}(X_r) - \operatorname{sgn}(X_{\pi(r)})) \, dr \right] \leq \frac{1}{n} \mathbb{E}_\alpha \left[ \int_0^{\tau_n} \mathbf{1}_{X_r \in [-1, 1]} \, dr \right] \\ &\leq \frac{1}{n} \mathbb{E}_\alpha \left[ \int_0^n \mathbf{1}_{X_r \in [-1, 1]} \, dr \right] + \frac{1}{n} \mathbb{E}_\alpha[|\tau_n - n|]. \end{aligned}$$

Since  $|X| \stackrel{\text{law}}{=} |B|$ ,  $\mathbb{P}_\alpha[X_r \in [-1, 1]] = \mathbb{P}_\alpha[|B_r| \in [-1, 1]] \leq 2/\sqrt{2\pi r}$ , so that

$$\frac{1}{n} \mathbb{E}_\alpha[M_n^2] \leq \frac{1}{n} + \frac{4}{\sqrt{2\pi n}}.$$

Combining these two results proves the  $L^2$ -convergence of  $n^{-1/2}(N_n - L_n)$  to 0.  $\square$

**Lemma 4.** *For a constant  $C$ ,*

$$\mathbb{E}_\alpha \left[ \frac{1}{\sqrt{N_{n^*(h)}(h)}} \right] \leq Ch^{\frac{2}{9}} \xrightarrow{h \rightarrow \infty} 0. \quad (11)$$

*Proof.* Since  $N_n(h) \geq 1$ ,  $\mathbb{E}_\alpha[N_n(h)^{-1/2}] \leq 1$  for any  $n \in \mathbb{N}$ ,  $h > 0$ .

From the scaling property,  $L_n/\sqrt{n} \stackrel{\text{law}}{=} L_1$ , and  $L_1$  has the density  $2 \exp(-y^2/2)/\sqrt{2\pi}$  (See Remark 6).

We write  $N_n$  for  $N_n(1)$ . It holds that

$$\begin{aligned} \mathbb{P}_\alpha \left[ \frac{N_n}{\sqrt{n}} \leq x \right] &\leq \mathbb{P}_\alpha \left[ \frac{N_n}{\sqrt{n}} \leq x; \left| \frac{N_n}{\sqrt{n}} - \frac{L_n}{\sqrt{n}} \right| \leq \delta \right] + \mathbb{P}_\alpha \left[ \frac{N_n}{\sqrt{n}} \leq x; \left| \frac{N_n}{\sqrt{n}} - \frac{L_n}{\sqrt{n}} \right| \geq \delta \right] \\ &\leq \mathbb{P}_\alpha \left[ \frac{L_n}{\sqrt{n}} \leq x + \delta \right] + \frac{1}{\delta} \mathbb{E}_\alpha \left[ \left| \frac{N_n}{\sqrt{n}} - \frac{L_n}{\sqrt{n}} \right|^2 \right]^{1/2} \\ &\leq \sqrt{\frac{2}{\pi}}(x + \delta) + \frac{C}{\delta n^{1/4}} \quad (12) \end{aligned}$$

Let  $\phi : \mathbb{N} \rightarrow \mathbb{R}_+$  such that

$$\phi(n) \underset{n \rightarrow \infty}{\searrow} 0 \text{ and } n^{1/4}\phi(n) \underset{n \rightarrow \infty}{\nearrow} +\infty.$$

Since  $N_n \geq 1$ ,

$$\begin{aligned} \mathbb{E}_\alpha \left[ \frac{1}{\sqrt{N_n}} \right] &= \mathbb{E}_\alpha \left[ \frac{1}{\sqrt{N_n}}; N_n < \sqrt{n}\phi(n) \right] + \mathbb{E}_\alpha \left[ \frac{1}{\sqrt{N_n}}; N_n \geq \sqrt{n}\phi(n) \right] \\ &\leq \mathbb{P}_\alpha[N_n < \sqrt{n}\phi(n)] + \frac{1}{\sqrt{n^{1/2}\phi(n)}}. \end{aligned}$$

Choosing  $x = \delta = \phi(n)$  in (12), for some constant  $C$ ,

$$\mathbb{P}_\alpha \left[ \frac{N_n}{\sqrt{n}} \leq \phi(n) \right] \leq C\phi(n) + \frac{C}{n^{1/4}\phi(n)}.$$

By balancing the expressions,  $\phi(n) = 1/n^{1/8}$  is a suitable choice which ensures that

$$\mathbb{E}_\alpha \left[ \frac{1}{\sqrt{N_n}} \right] \leq \frac{C}{n^{1/8}} \xrightarrow{n \rightarrow \infty} 0. \quad (13)$$

For  $\kappa(h) \nearrow_{h \rightarrow 0} +\infty$  and  $h^2\kappa(h) \searrow_{h \rightarrow 0} 0$ ,

$$\begin{aligned} \mathbb{E}_\alpha \left[ \frac{1}{\sqrt{N_{n^*(h)}(h)}} \right] &= \mathbb{E}_\alpha \left[ \frac{1}{\sqrt{N_{n^*(h)}(h)}}; n^*(h) \leq \kappa(h) \right] + \mathbb{E}_\alpha \left[ \frac{1}{\sqrt{N_{\kappa(h)}(h)}} \right] \\ &\leq \mathbb{P}_\alpha[n^*(h) \geq \kappa(h)] + \mathbb{E}_\alpha \left[ \frac{1}{\sqrt{N_{\kappa(h)}(1)}} \right]. \end{aligned} \quad (14)$$

Using the Bienaymé-Tchebychev inequality,

$$\mathbb{P}_\alpha[n^*(h) \geq \kappa(h)] = \mathbb{P}_\alpha[\tau_{\kappa(h)-1}(h) \geq T] \leq \frac{1}{T} \mathbb{E}_\alpha[\tau_{\kappa(h)-1}(h)] \leq \frac{\kappa(h) - 1}{T} h^2.$$

Balancing the choice of  $\kappa(h)$  using (13), we take  $\kappa(h) = h^{-16/9}$  to balance the rate of convergence in the two terms in the right-hand side of (14). This gives (11).  $\square$

*Proof of Theorem 1.* Since  $X_0 = 0$ ,  $N_n(h) > 0$  and  $L_T > 0$  almost surely.

The convergence of  $\alpha(h) = N_{n^*(h)}^+(h)/N_{n^*(h)}(h)$  to  $\alpha$  is a direct consequence of Lemma 1 and the convergence of  $h^{-2}n^*(h)$  to  $T$ .

From the Berry-Esseen theorem (See Lemma 2),

$$\begin{aligned} F_n(x) &= \mathbb{P}_\alpha \left[ \frac{T^{1/4}}{\sqrt{h}} \left( \frac{N_n^+(h)}{N_n(h)} - \alpha \right) \leq x \right] \\ &= \mathbb{E}_\alpha \left[ \mathbb{P}_\alpha \left[ \frac{N_n^+(h) - \alpha N_n(h)}{\sigma \sqrt{N_n(h)}} \leq \frac{x \sqrt{h N_n(h)}}{\sigma T^{1/4}} \middle| N_n(h) \right] \right] \\ &= \mathbb{E}_\alpha \left[ \Phi \left( \frac{x \sqrt{h N_n(h)}}{\sigma T^{1/4}} \right) \right] + \epsilon_n(h) \text{ with } |\epsilon_n(h)| \leq \mathbb{E}_\alpha \left[ \frac{C\alpha}{\sigma^3 \sqrt{N_n(h)}} \right]. \end{aligned}$$

Hence,

$$\mathbb{E}_\alpha[F_{n^*(h)}(x)] = \mathbb{E}_\alpha \left[ \Phi \left( \frac{x \sqrt{h N_{n^*(h)}(h)}}{\sigma T^{1/4}} \right) \right] + \mathbb{E}_\alpha[\epsilon_{n^*(h)}(h)].$$

The convergence of  $hN_{n^*(h)}(h)$  to  $L_T$  and Lemma 4 implies the convergence in distribution of  $h^{-1/2}(\alpha(h) - \alpha)$  to  $G/L_T$  for  $G \sim \mathcal{N}(0, 1)$ , where  $G$  is independent from  $L_T$ . The rate of convergence of the cumulative distribution function is given by Lemma 4.

Let us consider a Brownian motion  $W$  independent from  $B$ . Using the scaling property of the Brownian motion,

$$\mathbb{E}_\alpha \left[ \Phi \left( \frac{x\sqrt{L_T}}{\sigma T^{1/4}} \right) \right] = \mathbb{P}_\alpha \left[ \frac{\sigma T^{1/4}W(1)}{\sqrt{L_T}} \leq x \right] = \mathbb{P}_\alpha \left[ \frac{\sigma T^{1/4}W(L_T)}{L_T} \leq x \right].$$

Thus, the limit in distribution of  $h^{-1/2}(\alpha(h) - \alpha)$  could be written  $\sigma W(L_T)/L_T$ .  $\square$

### 3. The Local Asymptotic Mixed Normal property

The Local Asymptotic Mixed Normal (LAMN) property [23] is an extension of the Local Asymptotic Normal property introduced by L. Le Cam [19, 28] when the limits are mixed normal distributions. Among other properties, one may deduce results on the convergence of estimators constructed from  $\boldsymbol{\xi}(h)$  (See *e.g.* [24]).

**Proposition 3** (The LAMN property). *For any  $\gamma \in \mathbb{R}$  and  $h > 0$ ,*

$$\log \frac{\text{Lik}_{\alpha+\gamma\sqrt{h}}(\boldsymbol{\xi}(h))}{\text{Lik}_\alpha(\boldsymbol{\xi}(h))} - \gamma\sqrt{Q(h)}G(h) - \frac{\gamma^2}{2}Q(h) + \epsilon(h, \gamma) \xrightarrow[h \rightarrow 0]{\mathbb{P}_\alpha} 0$$

where  $G(h)$  and  $Q(h)$  are  $\mathcal{F}_{n^*(h)h}$ -measurable,

$$(G(h), Q(h)) \xrightarrow[h \rightarrow 0]{\text{dist.}} (H, \sigma^{-2}L_T) \text{ with } H \sim \mathcal{N}(0, 1),$$

and  $H$  and  $L_T$  are independent. In other words,  $(\mathbb{P}_\alpha \circ \boldsymbol{\xi}(h)^{-1})_{h>0}$  satisfies the LAMN property.

*Proof.* To simplify the notations, let us set  $N(h) = N_{n^*(h)}$ ,  $N^+(h) = N_{n^*(h)}$  and  $N^-(h) = N(h) - N^+(h)$ .

The logarithm of the ratio of the likelihood is, for any  $\gamma \in \mathbb{R}$ ,

$$\begin{aligned} \log \frac{\text{Lik}_{\alpha+\gamma\sqrt{h}}(\boldsymbol{\xi})}{\text{Lik}_\alpha(\boldsymbol{\xi})} &= N^+(h) \log \left( 1 + \frac{\gamma\sqrt{h}}{\alpha} \right) + N^-(h) \log \left( 1 + \frac{\gamma\sqrt{h}}{1-\alpha} \right) \\ &= \gamma\sqrt{h} \left( \frac{N^+(h)}{\alpha} - \frac{N^-(h)}{1-\alpha} \right) - \frac{\gamma^2 h}{2} \left( \frac{N^+(h)}{\alpha^2} + \frac{N^-(h)}{(1-\alpha)^2} \right) + \epsilon(h, \gamma) \end{aligned}$$

with, from the Taylor development of the logarithm,

$$\epsilon(h, \gamma) \leq C\gamma^3 N(h)h^{3/2}$$

for some constant  $C$ . Since  $hN(h)$  converges almost surely to  $L_T$  which is almost surely finite,  $\epsilon(h, \gamma)$  converges almost surely to 0 for any  $\gamma > 0$ , hence in probability. Thanks to Lemma 1,

$$A(h) = \frac{1}{N(h)} \left( \frac{N^+(h)}{\alpha} - \frac{N^-(h)}{1-\alpha} \right) \xrightarrow[h \rightarrow 0]{\text{a.s.}} 0$$

and  $B(h) = \frac{1}{N(h)} \left( \frac{N^+(h)}{\alpha^2} + \frac{N^-(h)}{(1-\alpha)^2} \right) \xrightarrow[h \rightarrow 0]{\text{a.s.}} \frac{1}{\sigma^2}$  with  $\sigma = \sqrt{\alpha(1-\alpha)}$ .

Theorem 1 implies that

$$\frac{A(h)}{\sqrt{h}} = \frac{1}{\sigma^2} \left( \frac{N^+(h)}{N(h)} - \alpha \right) \xrightarrow[h \rightarrow 0]{\text{dist.}} \frac{W(L_T)}{\sigma L_T}.$$

On the other hand, since  $hN(h)$  converges almost surely to  $L_T$ ,

$$\sqrt{h}N(h)A(h) = hN(h) \frac{A(h)}{\sqrt{h}} \xrightarrow[h \rightarrow 0]{\text{dist.}} \sigma^{-1}W(L_T) \stackrel{\text{law}}{=} \sigma^{-1}\sqrt{L_T}W(1).$$

To summarize,

$$\log \frac{\text{Lik}_{\alpha+\gamma\sqrt{h}}(\boldsymbol{\xi})}{\text{Lik}_{\alpha}(\boldsymbol{\xi})} = \gamma\sqrt{h}N(h)A(h) - \gamma^2 hN(h)B(h) + \epsilon(h, \gamma).$$

With

$$Q(h) = hN(h)B(h) \xrightarrow[h \rightarrow 0]{\text{a.s.}} \sigma^{-2}L_T$$

and  $G(h) = A(h) \frac{\sqrt{N(h)}}{\sqrt{B(h)}} \xrightarrow[h \rightarrow 0]{\text{dist.}} W(1) \sim \mathcal{N}(0, 1)$ .

This proves the LAMN property.  $\square$

**Acknowledgement.** This work has been developed within the framework of the Inria's Équipe Associée ANESTOC-TOSCA between France and Chile. It is associated to a joint work with E. Mordecki and S. Torres on the estimation of the parameter of the Skew Brownian motion. The author wishes to thank them for interesting discussion on this topic.

## References

- [1] M. A. J. D. Reeve, M. Xiao & D. Xu. Identification of Diffusion Coefficient in Nonhomogeneous Landscapes. In: *Neural Information Processing*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, 290–297. DOI: 10.1007/978-3-642-34481-7\_36.
- [2] T. Appuhamillage, V. Bokil, E. Thomann, E. Waymire & B. Wood. Occupation and local times for skew Brownian motion with applications to dispersion across an interface. *Ann. Appl. Probab.* **21.1** (2011), 183–214. DOI: 10.1214/10-AAP691.
- [3] R. F. Bass & D. Khoshnevisan. Rates of convergence to Brownian local time. *Stochastic Process. Appl.* **47.2** (1993), 197–213. DOI: 10.1016/0304-4149(93)90014-U.
- [4] A. C. Berry. The accuracy of the Gaussian approximation to the sum of independent variates. *Trans. Amer. Math. Soc.* **49** (1941), 122–136.
- [5] C. R. Blyth. Approximate binomial confidence limits. *J. Amer. Statist. Assoc.* **81.395** (1986), 843–855.
- [6] M. Bossy, N. Champagnat, S. Maire & D. Talay. Probabilistic interpretation and random walk on spheres algorithms for the Poisson-Boltzmann equation in molecular dynamics. *M2AN Math. Model. Numer. Anal.* **44.5** (2010), 997–1048. DOI: 10.1051/m2an/2010050.
- [7] L. D. Brown, T. T. Cai & A. DasGupta. Confidence intervals for a binomial proportion and asymptotic expansions. *Ann. Statist.* **30.1** (2002), 160–201. DOI: 10.1214/aos/1015362189.
- [8] R. S. Cantrell & C Cosner. Diffusion Models for Population Dynamics Incorporating Individual Behavior at Boundaries: Applications to Refuge Design. *Theoretical Population Biology* **55.2** (1999), 189–207. DOI: 10.1006/tpbi.1998.1397.
- [9] K. L. Chung & G. A. Hunt. On the zeros of  $\sum_1^n \pm 1$ . *Ann. of Math. (2)* **50** (1949), 385–400.
- [10] K. L. Chung & R. Durrett. Downcrossings and local time. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **35.2** (1976), 147–149.
- [11] E. Csáki & P. Révész. A combinatorial proof of a theorem of P. Lévy on the local time. *Acta Sci. Math. (Szeged)* **45.1-4** (1983), 119–129.
- [12] M. Csörgő & L. Horváth. On best possible approximations of local time. *Statist. Probab. Lett.* **8.4** (1989), 301–306. DOI: 10.1016/0167-7152(89)90036-9.

- [13] S. Delattre & M. Hoffmann. Asymptotic equivalence for a null recurrent diffusion. *Bernoulli* **8.2** (2002), 139–174.
- [14] C.-G. Esseen. On the Liapounoff limit of error in the theory of probability. *Ark. Mat. Astr. Fys.* **28A.9** (1942), 19.
- [15] P. Étoré. On random walk simulation of one-dimensional diffusion processes with discontinuous coefficients. *Electron. J. Probab.* **11.9** (2006), 249–275. DOI: 10.1214/EJP.v11-311.
- [16] E. R. Fernholz, T. Ichiba & I. Karatzas. Two Brownian particles with rank-based characteristics and skew-elastic collisions. *Stochastic Process. Appl.* **123.8** (2013), 2999–3026. DOI: 10.1016/j.spa.2013.03.019.
- [17] D. Florens. Estimation of the diffusion coefficient from crossings. *Stat. Inference Stoch. Process.* **1.2** (1998), 175–195. DOI: 10.1023/A:1009927813898.
- [18] J. M. Harrison & L. A. Shepp. On skew Brownian motion. *Ann. Probab.* **9.2** (1981), 309–313.
- [19] I. A. Ibragimov & R. Z. Has'minskiĭ. *Statistical estimation. Asymptotic theory*. Vol. 16. Applications of Mathematics. Springer-Verlag, New York-Berlin, 1981.
- [20] K. Itô & H. P. McKean Jr. Brownian motions on a half line. *Illinois J. Math.* **7** (1963), 181–231.
- [21] K. Itô & H. P. McKean Jr. *Diffusion processes and their sample paths*. 2nd ed. Springer-Verlag, Berlin-New York, 1974.
- [22] J. Jacod. Rates of convergence to the local time of a diffusion. *Ann. Inst. H. Poincaré Probab. Statist.* **34.4** (1998), 505–544. DOI: 10.1016/S0246-0203(98)80026-5.
- [23] P. Jeganathan. On the asymptotic theory of estimation when the limit of the log-likelihood ratios is mixed normal. *Sankhyā Ser. A* **44.2** (1982), 173–212.
- [24] P. Jeganathan. Some asymptotic properties of risk functions when the limit of the experiment is mixed normal. *Sankhyā Ser. A* **45.1** (1983), 66–87.
- [25] D. Khoshnevisan. Exact rates of convergence to Brownian local time. *Ann. Probab.* **22.3** (1994), 1295–1330.
- [26] J. Lamperti. An occupation time theorem for a class of stochastic processes. *Trans. Amer. Math. Soc.* **88** (1958), 380–387.



- [27] J.-F. Le Gall. One-Dimensional Stochastic Differential Equations Involving the Local Times of the Unknown Process. In: *Stochastic Analysis and Applications*. Vol. 1095. Lecture Notes in Mathematics. Springer Verlag, 1985, 51–82.
- [28] L. LeCam. On some asymptotic properties of maximum likelihood estimates and related Bayes' estimates. *Univ. California Publ. Statist.* **1** (1953), 277–329.
- [29] A. Lejay. On the constructions of the skew Brownian motion. *Probab. Surv.* **3** (2006), 413–466. DOI: 10.1214/154957807000000013.
- [30] A. Lejay & G. Pichot. Simulating diffusion processes in discontinuous media: a numerical scheme with constant time steps. *J. Comput. Phys.* **231.21** (2012), 7299–7314. DOI: 10.1016/j.jcp.2012.07.011.
- [31] A. Lejay, E. Mordecki & S. Torres. Is a Brownian motion skew? *Scand. J. Stat.* **41.2** (2014), 346–364. DOI: 10.1111/sjos.12033.
- [32] A. Lejay, E. Mordecki & S. Torres. *Maximum Likelihood estimation of the parameter of the Skew Brownian motion*. In preparation. 2016.
- [33] A. Lipton & A. Sepp. Filling the Gap. *Risk* (Oct. 2011), 86–91.
- [34] S. V. Nagaev & V. I. Chebotarev. On an estimate for the closeness of the binomial distribution to the normal distribution. *Dokl. Akad. Nauk* **436.1** (2011), 26–28. DOI: 10.1134/S1064562411010030.
- [35] S. V. Nagaev & V. I. Chebotarev. On the estimation of the closeness of the binomial distribution to the normal distribution. *Theory Probabl. Appl.* **56.2** (2012), 213–239.
- [36] P. Révész. *Random walk in random and nonrandom environments*. World Scientific Publishing Co., Inc., Teaneck, NJ, 1990. DOI: 10.1142/1107.
- [37] D. Revuz & M. Yor. *Continuous martingales and Brownian motion*. 3rd ed. Vol. 293. Grundlehren der Mathematischen Wissenschaften. Springer-Verlag, Berlin, 1999. DOI: 10.1007/978-3-662-06400-9.
- [38] T. D. Ross. Accurate confidence intervals for binomial proportion and Poisson rate estimation. *Computers in Biology and Medicine* **33.6** (Nov. 2003), 509–531.
- [39] J. Walsh. A diffusion with discontinuous local time. In: *Temps locaux*. Vol. 52–53. Astérisques. Société Mathématique de France, 1978, 37–45.