Parameter estimation for skew (and sticky) BM

Maximum likelihood estimation : an asymptotic expansion.

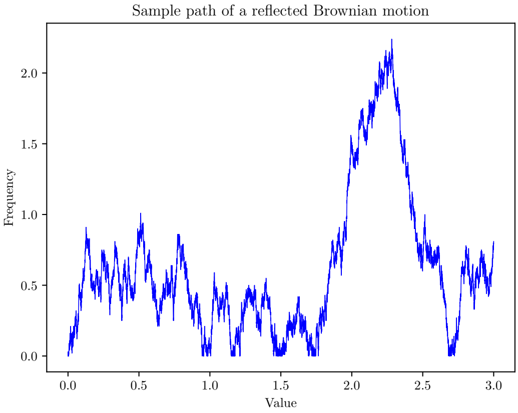
Sara Mazzonetto (IECL, Nancy)

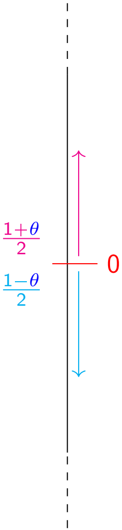
joint works in progress with
Alexis Anagnostakis (LJK, Grenoble) and *Antoine Lejay* (IECL/Inria Nancy)

40 years of reflected Brownian motion and related topics

Roscoff, April 28th 2023

What is **Skew** BM? Let us start with reflected BM





Trajectorial construction flipping excursions [Itô and McKean (1974)].

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

What is **Skew** BM?

It is a BM perturbed by a semi-permeable barrier and semi-reflecting barrier point, say at 0, with skewness parameter $\theta \in (-1,1)$,

is the unique strong solution the stochastic differential equation (SDE):

$$X_t = X_0 + W_t + \theta\, L_t^0(X); \qquad L_t^0(X) = \lim_{\epsilon \to 0} \frac{1}{2\epsilon} \int_0^t \mathbf{1}_{\{-\epsilon \leqslant X_s \leqslant \epsilon\}}\, \mathrm{d}s.$$

where $L_t^0(X)$ is the symmetric local time at 0 [Harrison and Shepp (1981)].

Survey on Skew BM, several constructions and its applications see [Lejay (2004)].
Several authors, one-dimensional case: Le Gall, Zeitseva, Weinryb, Bass and Chen, Burdzy and Kaspi, Étoré and Martinez...
Multi-dimensional: Portenko, Aryasova and Pilipenko, Atar and Budhiraja, Trutnau...

⚠ strongly related to SDEs with discontinuous coefficients: $X_t = \int_0^t b(X_s)\,\mathrm{d}s + \int_0^t \sigma(X_s)\,\mathrm{d}W_s$.

What is this seminar about?

Skew BM of unknown parameter $\theta \in [-1, 1]$
observed on $[0, T]$ ($T = 1$ for simplicity) at high frequency times: $\mathbf{X} = (X_0, X_{1/n}, X_{2/n}, \ldots, X_1)$.

Goal: Estimation, its asymptotics, and hypothesis testing for $\theta \in (-1, 1)$ from discrete observations.

Related results

- Skew-RW [Lejay 2018].
- skew two-sided-reflected BM (ergodicity) [Bardou and Martinez (2010)] from continuous observations in long time.

  SBM is null recurrent and has “singular” distribution.
- Oscillating BM $X_t = \int_0^t (\sigma_- \mathbf{1}_{X_s<0} + \int_0^t \sigma_+ \mathbf{1}_{X_s \geq 0}) \, \mathrm{d}W_s$. [Lejay and Pigato (2018)]

What is this seminar about?

Skew BM of unknown parameter $\theta \in [-1, 1]$
observed on $[0, T]$ ($T = 1$ for simplicity) at high frequency times: $\mathbf{X} = (X_0, X_{1/n}, X_{2/n}, \dots, X_1)$.

Goal: Estimation, its asymptotics, and hypothesis testing for $\theta \in (-1, 1)$ from discrete observations.

Related results

- Skew-RW [Lejay 2018].
- skew two-sided-reflected BM (ergodicity) [Bardou and Martinez (2010)] from continuous observations in long time.

  SBM is null recurrent and has “singular” distribution.
- Oscillating BM $X_t = \int_0^t (\sigma_- \mathbf{1}_{X_s<0} + \int_0^t \sigma_+ \mathbf{1}_{X_s \geq 0})\,\mathrm{d}W_s$. [Lejay and Pigato (2018)]

✓ Classical method: *Maximum Likelihood Estimator* (MLE).

✓ Local time approximation.

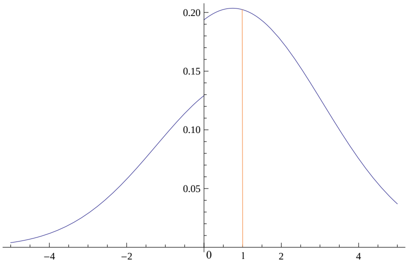
Skew BM: the transition density

$$p_\theta(t,x,y) := \frac{1}{\sqrt{2\pi t}} e^{-\frac{(y-x)^2}{2t}} + \theta \,\mathrm{sgn}(y) \frac{1}{\sqrt{2\pi t}} e^{-\frac{(|y|+|x|)^2}{2t}}$$

is the density of X_{t+s} the θ-skew BM with $X_s = x \in \mathbb{R}$. Established by [Walsh (1978)].

$$(1-\theta)\; p_\theta(t,x,0^+) = (1+\theta)\; p_\theta(t,x,0^-), \quad x \in \mathbb{R}$$

Example: Let $\theta = 0.2$, $t = 5$, $x = 1$, then $y \mapsto p_\theta(t,x,y)$ is



MLE for Skew BM

$\mathbf{X} = (X_0, X_{1/n}, X_{2/n}, \ldots, X_1)$.
$\widehat{\theta}(n)$ should maximize the likelihood $\theta \mapsto \mathcal{L}_n(\theta; \mathbf{X})$
where $\theta \mapsto \mathcal{L}_n(\theta; \cdot)$ is the density function of $\mathbf{X}$ under $\mathbb{P}_\theta$.

⚠ $\widehat{\theta}(n)$ is the zero of the score

$$S_n(\theta) := \partial_\theta \ln \mathcal{L}(\theta; \mathbf{X}) = \sum_{i=0}^{n-1} \partial_\theta \log p_\theta(1/n, X_{i1/n}, X_{(i+1)1/n})$$

where

$$\partial_\theta \log p_\theta(t, x, y) = \frac{\operatorname{sgn}(y)}{\operatorname{sgn}(y)\theta + \exp(2\max(xy, 0)/t)}$$ ← change sign.

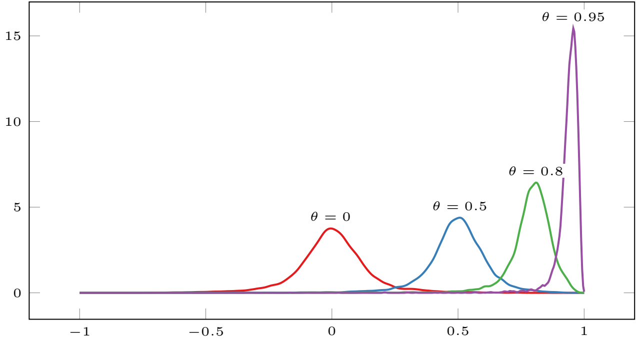
score asymptotics

Results on MLE estimator for Skew BM

[Lejay, Mordecki, and Torres (2014, 2019)]
Show **consistency*** of MLE and conjecture asymptotic law.
If $\theta = 0$ convergence rates and *polynomial approximation* of the MLE.

In [M. (2019+)], new estimators of θ with **convergence rates***.
Allows to treat MLE in the case $\theta \neq 0$

* rely on generalizations of [Jacod (1998)] for local time approximation.



MLE empirical density,
show more *skewness*, the more θ is close to 1.
Simulations of SBM from [Lejay and Pichot (2012)].

Asymptotic mixed-normality

Theorem: MLE asymptotic mixed normality [Lejay and M. (2023+)]

Under $\mathbb{P}_\theta$, the MLE $\widehat{\theta}(n)$ is asymptotically mixed gaussian:

$$n^{1/4}(\widehat{\theta}(n) - \theta) \quad \text{converges in } \textit{distribution} \text{ to} \quad s(\theta)\frac{G}{\sqrt{L_T^0(X)}},$$

where $G \sim \mathcal{N}(0,1)$ independent of X, $s(\theta)$ easily numerically computable and

$$s(\theta) \approx \frac{\sqrt{1-\theta^2}}{\sqrt{1.3 + 0.23\,\theta^2 + 0.07\,\theta^4}}.$$

Asymptotic mixed-normality

Theorem: MLE asymptotic mixed normality [Lejay and M. (2023+)]

Under $\mathbb{P}_\theta$, the MLE $\widehat{\theta}(n)$ is asymptotically mixed gaussian:

$$n^{1/4}(\widehat{\theta}(n)-\theta) \quad \text{converges in } \textit{distribution} \text{ to} \quad s(\theta)\frac{G}{\sqrt{L_T^0(X)}},$$

where $G \sim \mathcal{N}(0,1)$ independent of X, $s(\theta)$ easily numerically computable and

$$s(\theta) \approx \frac{\sqrt{1-\theta^2}}{\sqrt{1.3+0.23\,\theta^2+0.07\,\theta^4}}.$$

Different $s(\theta)$ for other consistent estimators obtained in [M. (2019+)] via local time approximation:

$$\hat{\theta}_n = 1-\frac{2\sum_{i=0}^{n-1}|X_{i+1}|\mathbf{1}_{\{X_i X_{i+1}<0\}}}{\frac{1}{\sqrt{n}}\sum_{k=0}^{n-1}\mathbf{1}_{(0,1)}(\sqrt{n}X_k)}. \qquad \text{Recall} \quad L_T^0(X) = \lim_{\epsilon\to 0}\frac{1}{2\epsilon}\int_0^T \mathbf{1}_{\{-\epsilon\leqslant X_s\leqslant\epsilon\}}\,\mathrm{d}\langle X\rangle_s.$$

Proof: the Local time approximation ingredient

$S_n(\theta)/\sqrt{n}$ and its derivatives $\partial_\theta S_n(\theta)/\sqrt{n}$ are of the form

$$\varepsilon_n^f := \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} f(\sqrt{n} X_{k/n}, \sqrt{n} X_{(k+1)/n}) \qquad \left(\text{example } \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} \mathbf{1}_{X_{k/n} X_{(k+1)/n} < 0} \right).$$

Theorem (Convergence rates, cf. [M. (2019+)], generalization of [Jacod (1998)])

Let f satisfying suitable integrability conditions.
Then there exists a real constant K (integral expression)

$$n^{1/4} \left(\varepsilon_n^f - L_1^0(X) \int_{\mathbb{R}} \mathbb{E}_x[f(x, X_1)](1 + \operatorname{sgn}(x)\theta) \, \mathrm{d}x \right) \xrightarrow[n\to\infty]{\text{law}} K \sqrt{L_1^0(X)} B$$

where $B \sim \mathcal{N}(0,1)$ on an extension of the probability space, independent of X.

The result is more general. LLN in [Lejay, Mordecki, and Torres (2019)].
Special case $\int_{\mathbb{R}} \mathbb{E}_x[f(x, X_1)](1 + \operatorname{sgn}(x)\theta) \, \mathrm{d}x = 0$ for skew jump processes in [Robert (2022)].

Proof: score derivatives behavior

LLN If $k \geqslant 1$, then $\partial_\theta^k S_n(\theta)/\sqrt{n}$ converges in probability to $\xi_k(\theta) L_1(X)$; $\xi_1(\theta) = s(\theta)^{-2}$, (recall $s(\theta)$ goes like $\sqrt{1-\theta^2}$).

CLT When $\xi_k(\theta) = 0$ ($k = 0$ for all θ and k even for $\theta = 0$) then $n^{-\frac{1}{4}} \partial_\theta^k S_n(\theta)$ converges to a mixed Gaussian law $\sqrt{L_1(X)} G$ up to a constant depending on θ. If $k = 0$, then the constant is $1/s(\theta)$.

Analogy to MLE of Bernoulli random variables of parameter $\theta \in (0,1)$.

$$S_n(\theta) = \frac{1}{(1-\theta)\theta} \left(\sum_{j=1}^{n} X_j - n\theta \right)$$

Use of CLT for even derivatives if $\theta = 1/2$.

Proof: local time approximation and polynomial expansion

⚠ The MLE is implicit: zero of an (analytic) random function $S_n(\cdot)$.

✓ The score and its derivative are related to local time approximation.

Goal: Polynomial approximation or expansion of $\widehat{\theta}(n)$

$$\theta_p(n) = \theta + d_{0,n}(\theta) + \sum_{k=2}^{p} D_{k,n}(\theta) d_{0,n}(\theta)^k$$

($p = \infty$) such that, as $n \to \infty$, we have joint convergence of

- ⋆ $\varphi(n)^{-1} d_{0,n}(\theta)$ (provides the asymptotic behavior) for $\varphi(n)$ which vanishes;
- ⋆ $(D_{k,n}(\theta))_{k=2}^{p}$ (in some sense: $\mathbb{P}_\theta$ or law) to $(D_{k,\infty}(\theta))_{k=2}^{p}$. They depend on the true θ.

This approximation is not unique in general.

Proof: polynomial expansion

Taylor expansion around θ gives

$$0 = S_n(\widehat{\theta}(n)) = \sum_{k=0}^{\infty} \frac{1}{k!} \partial_\theta^k S_n(\theta)(\widehat{\theta}(n) - \theta)^k.$$

Polynomial approximation or expansion of $\widehat{\theta}(n)$

$$\widehat{\theta}(n) = \theta + d_{0,n}(\theta) + \sum_{k=2}^{\infty} D_{k,n}(\theta) d_{0,n}(\theta)^k$$

- ⋆ $d_{0,n}(\theta) := -\frac{S_n(\theta)}{\partial_\theta S_n(\theta)}$ and $n^{1/4} d_{0,n}(\theta)$ converges in law to $G/\sqrt{L_1(X)}$ times $s(\theta)$;
- ⋆ $D_{1,n}(\theta) = 1$ and $\quad D_{q,n}(\theta) = -\sum_{m=2}^{q} \frac{1}{m!} \frac{\partial_\theta^m S_n(\theta)}{\partial_\theta S_n(\theta)} \sum_{k_1+\cdots+k_m=q} D_{k_1,n}(\theta) \cdots D_{k_m,n}(\theta)$

  which converges in probability towards a constant.

For $\theta = 0$, the above expansion was already given in [Lejay, Mordecki, and Torres (2019)].

Remarks

We have shown the asymptotic mixed normality

$$n^{1/4}(\widehat{\theta}(n) - \theta) = n^{1/4}\left(d_{0,n}(\theta) + \sum_{k=2}^{\infty} D_{k,n}(\theta) d_{0,n}(\theta)^k\right) \quad \text{converges in } \textit{distribution} \text{ to} \quad s(\theta)\frac{G}{\sqrt{L_T^0(X)}},$$

- $n^{1/4} d_{0,n}(\theta) := -n^{1/4} \frac{S_n(\theta)}{\partial_\theta S_n(\theta)}$ asymptotic behavior $\rightarrow$ Hypothesis testing.
- Alternative expansion when $\theta = 0$ because of CLT, sort of phase transition.

Remarks

We have shown the asymptotic mixed normality

$$n^{1/4}(\widehat{\theta}(n) - \theta) = n^{1/4}\left(d_{0,n}(\theta) + \sum_{k=2}^{\infty} D_{k,n}(\theta) d_{0,n}(\theta)^k\right) \quad \text{converges in } \textit{distribution} \text{ to} \quad s(\theta)\frac{G}{\sqrt{L_T^0(X)}},$$

- $n^{1/4} d_{0,n}(\theta) := -n^{1/4}\frac{S_n(\theta)}{\partial_\theta S_n(\theta)}$ asymptotic behavior → Hypothesis testing.
- Alternative expansion when $\theta = 0$ because of CLT, sort of phase transition.

⋆ Expansion capture *skewness* of MLE, despite limit symmetric.

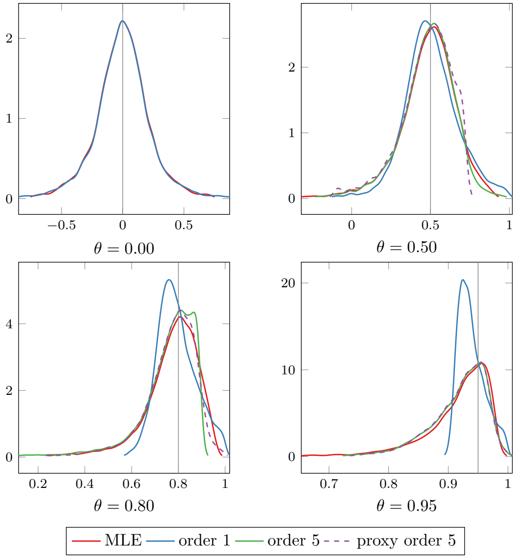
⋆ Five terms of semi-asymptotic expansion suffices. Proxy

$$\widehat{\theta}(n) \approx \theta + d_{0,n}(\theta) + \sum_{k=2}^{\infty} D_{k,\infty}(\theta) d_{0,n}(\theta)^k.$$

Study of $D_{k,\infty}(\theta)$ numerically: $D_{2,\infty}(\theta) = -\frac{1}{2}\frac{\xi_2(\theta)}{\xi_1(\theta)}$ and $D_{3,\infty}(\theta) = \frac{1}{2}\left(\left(\frac{\xi_2(\theta)}{\xi_1(\theta)}\right)^2 - \frac{\xi_3(\theta)}{3\xi_1(\theta)}\right)$.

⋆ Boundary layer effect.

Numerics for proxy and truncated expansion



Empirical density MLE $\widehat{\theta}(n)$,

$\theta + d_{0,n}(\theta)$,

5th order expansion

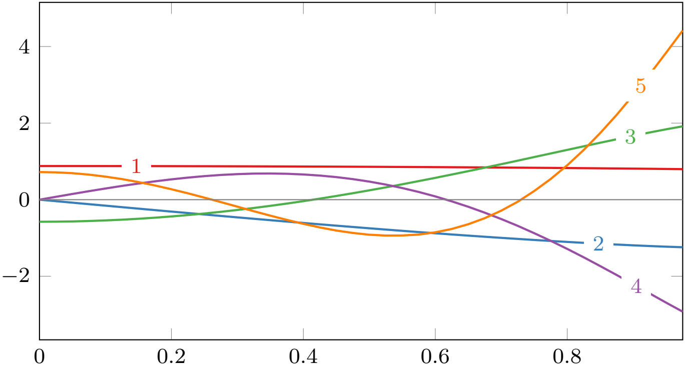
$$\theta + d_{0,n}(\theta) + \sum_{k=2}^{5} D_{k,n}(\theta) d_{0,n}(\theta)^k$$

and 5th order proxy

$$\theta + d_{0,n}(\theta) + \sum_{k=2}^{5} D_{k,\infty}(\theta) d_{0,n}(\theta)^k$$

$n = 10^3, 10^4$ samples.

Behavior of $D_{k,\infty}(\theta)$



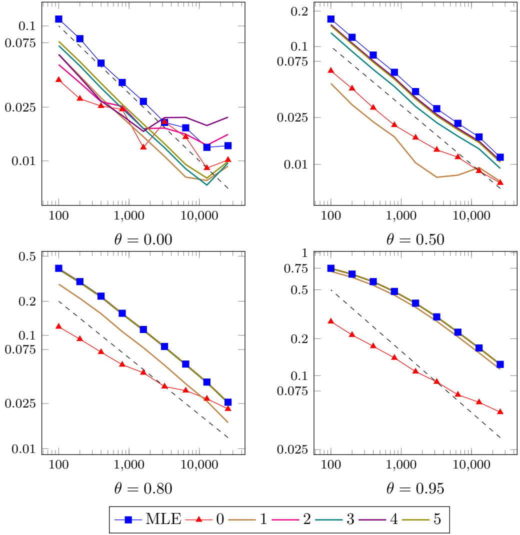
Functions
$\theta \mapsto D_{k,\infty}(\theta)s(\theta)^k(1-\theta^2)^{k/2-1}$.
$D_{k,\infty}(\theta)s(\theta)^k$ goes like $(1-\theta^2)^{1-k/2}$

Boundary layer effect from $k \geqslant 3$:

$$\theta + d_{0,n}(\theta) + \sum_{k=2}^{5} D_{k,\infty}(\theta) d_{0,n}(\theta)^k$$

$d_{0,n}(\theta)$ converges to $s(\theta)G/\sqrt{L_1^0(X)}$.

Kolmogorov-Smirnov distances between distributions

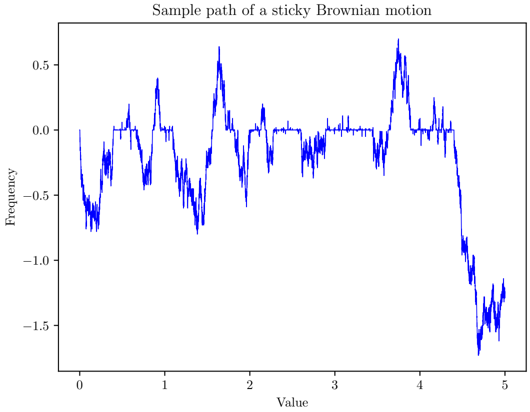


10^4 samples of SBM,
Comparison with $C/\sqrt{n}$ (dashed line), in log-log scale, for statistics

- $s(\theta)^{-1} n^{1/4}(\widehat{\theta}(n) - \theta)$ and $G/\sqrt{L_T}$.
- $n^{-1/2}\partial_\theta^k S_n(\theta)/\Xi_{k,k}(\theta)$ and $\sqrt{L_T}G$ for $k = 0$ (and when $\theta = 0$, $k = 2, 4$)
- $n^{-1/2}\partial_\theta^k S_n(\theta)/\xi_k(\theta)$ and L_T otherwise

Future? Berry-Esseen type results.

Estimation and local time approximation of Sticky BM



"It is a slowed down BM."
One-dimensional diffusions

The speed measure gives the "slow down":

$$m_\rho(\mathrm{d}x) = \rho\,\delta_0(\mathrm{d}x) + \mathrm{d}x.$$

(Skew)-Sticky threshold: "the process spends more time at around zero"

We seek for unified results on local time approximation for these singular processes.

Let $z \in \mathbb{R}$, $\rho > 0$, $\theta \in [-1, 1]$

$$\begin{cases} Z_t = \int_0^t \mathbf{1}_{\{Z_s \neq 0\}} \sigma(Z_s)\,\mathrm{d}W_s + \theta L_t^0(Z) + \int_0^t b(Z_s)\,\mathrm{d}s \\ \int_0^t \mathbf{1}_{\{Z_s = 0\}}\,\mathrm{d}s = \frac{\rho}{2} L_t^0(Z) \quad \longleftarrow \quad \text{this is not 0!} \end{cases}$$

for $\sigma(x) = \sigma_- \mathbf{1}_{(-\infty,0)}(x) + \sigma_+ \mathbf{1}_{(0,+\infty)}(x) > 0$.

Pathwise description sticky BM [Bass (2014)], [Engelbert and Peskir (2014)] and sticky-skew diffusion [Salin and Spiliopoulos (2017)].

Sticky parameter estimation

$\alpha \in (0,1)$.

$$\hat{\rho}_N := 2 \int g(x) m(\mathrm{d}x) \, \frac{\sum_{k=0}^{N-1} \mathbf{1}_{\{X_{k/N}=0\}}}{N^\alpha \sum_{k=0}^{N-1} g(N^\alpha X_{k/N})}.$$

This statistic compares

- how many observations are located at 0
- with how many are located around 0.

$$\frac{1}{N} \sum_{k=0}^{N-1} \mathbf{1}_{\{X_{k/N}=0\}} \to \int_0^1 \mathbf{1}_{\{X_s=0\}} \, \mathrm{d}s = \frac{\rho}{2} L_1^0(X).$$

$$m(\mathrm{d}x) = \left(\frac{1+\beta}{\sigma_+^2} \mathbf{1}_{(0,+\infty)}(x) + \frac{1-\beta}{\sigma_-^2} \mathbf{1}_{(-\infty,0)}(x) \right) \mathrm{d}x$$

Think of Lebesgue measure, for simplicity.

Local time approximation of singular diffusions: spot functions

Theorem ([Anagnostakis and M.])

Let X be a sticky OSBM, $h_N = N^{-\alpha}$ $\alpha \in (0,1)$, m be the speed measure of an OSBM, and g bounded and m-integrable satisfying $g(0) = 0$.
Let a_N such that $N^{\alpha} a_N \to 2\lambda \in \mathbb{R}$. Then

$$\frac{N^{\alpha}}{N} \sum_{k=0}^{N-1} \left(g(N^{\alpha} X_{k/N}) + a_N \mathbf{1}_{\{X_{k/N}=0\}} \right) \xrightarrow[N\to\infty]{\mathbb{P}} \left(\int_{\mathbb{R}} g(x) m(\mathrm{d}x) + \rho\lambda \right) L_1^0(X)$$

For BM follows from the already cited results (even rate of convergence known).
For skew-BM and oscillating-BM, cited results for $\alpha = 1/2$ (and rate known).
For sticky BM, $\alpha \in (0, 1/2)$, $g(x) = 0$ around 0, and particular a_N [Anagnostakis (2022+)].

Perspectives and remarks

- Rates of convergence for the estimators of the parameter of sticky BM.
- Extend the results for cross functionals (with different scaling) in the most general case (mix of singularities and diffusions).
- The parameter estimation is considered only in the case of “pure” behavior (except in some cases). Joint estimation of the coefficients of sticky-OS diffusions is a must.
- Estimating jointly the threshold would be crucial in applications for diffusions.

Acknowledgments and the references the talk is based on

- A. Lejay, S. M., *Maximum likelihood estimator for skew Brownian motion: the convergence rate*, preprint, 2023.
- A. Lejay, S. M., *Beyond the delta method*, preprint, 2022.
- S. M., *Rates of convergence to the local time of Oscillating and Skew Brownian Motions*, preprint 2019.

Thank you very much for your attention

Merci beaucoup pour l'attention.