ESAIM: PS 28 (2024) 292–328
https://doi.org/10.1051/ps/2024010

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

A HYPOTHESIS TEST FOR THE DOMAIN OF ATTRACTION OF A RANDOM VARIABLE

Héctor Olivero[1,*] and Denis Talay[2]

Abstract. In this work, we address the problem of detecting whether a sampled probability distribution of a random variable V has infinite first moment. This issue is notably important when the sample results from complex numerical simulation methods. For example, such a situation occurs when one simulates stochastic particle systems with complex and singular McKean–Vlasov interaction kernels. As stated, the detection problem is ill-posed. We thus propose and analyze an asymptotic hypothesis test for independent copies of a given random variable which is supposed to belong to an unknown domain of attraction of a stable law. The null hypothesis $\mathbf{H_0}$ is: '$X = \sqrt{V}$ is in the domain of attraction of the Normal law' and the alternative hypothesis is $\mathbf{H_1}$: 'X is in the domain of attraction of a stable law with index smaller than 2'. Our key observation is that X cannot have a finite second moment when $\mathbf{H_0}$ is rejected (and therefore $\mathbf{H_1}$ is accepted). Surprisingly, we find it useful to derive our test from the statistics of random processes. More precisely, our hypothesis test is based on a statistic which is inspired by methodologies to determine whether a semimartingale has jumps from the observation of one single path at discrete times. We justify our test by proving asymptotic properties of discrete time functionals of Brownian bridges.

Mathematics Subject Classification. 62F05, 62G10, 60J65.

Received 11 December 2023. Accepted 5 July 2024.

1. Introduction

In this paper we consider situations where one observes a sample of an unknown probability distributions and wonder if that probability distribution has finite moments.

Let us give an important example of such a situation, which actually has motivated our research. A now standard stochastic numerical method to solve non-linear McKean-Vlasov-Fokker-Planck equations consists in simulating stochastic particle systems with McKean-Vlasov interactions of the type

$$X_t^{i,N} = X_0^{i,N} + \int_0^t \frac{1}{N}\sum_{j=1}^N \beta(X_s^{i,N}, X_s^{j,N})\,\mathrm{d}s + \int_0^t \frac{1}{N}\sum_{j=1}^N \gamma(X_s^{i,N}, X_s^{j,N})\,\mathrm{d}W_s^i, \quad 1 \le i \le N, \tag{1.1}$$

Keywords and phrases: Hypothesis test, domains of attraction of stable laws.

[1] CIMFAV – Instituto de Ingeniería Matemática, Universidad de Valparaíso. Chile.
[2] INRIA Saclay, France.

[*] Corresponding author: hector.olivero@uv.cl

where the functions β and γ from $\mathbb{R}^d$ to $\mathbb{R}^d$ are the interaction kernels, the $X_0^{i,N}$'s are independent copies of a random variable X_0 and the W^i are independent Brownian motions. Under appropriate conditions, the preceding system propagates chaos: The measure-valued process $\frac{1}{N}\sum_{i=1}^N \delta_{X_\bullet^{i,N}}$ weakly converges to a unique probability distribution on the space of continuous functions $\mathcal{C}$ from $\mathbb{R}^d$ to $\mathbb{R}^d$ which is the probability law of the solution to the following McKean-Vlasov stochastic differential equation:

$$\begin{cases} X_t = X_0 + \int_0^t \int_{\mathcal{C}} \beta(X_s, \omega_s)\ \mu(\mathrm{d}\omega)\ \mathrm{d}s + \int_0^t \int_{\mathcal{C}} \gamma(X_s, \omega_s)\ \mu(\mathrm{d}\omega)\ \mathrm{d}W_s, \\ \mu\ :=\ \text{Probability distribution of the process } (X_t). \end{cases} \tag{1.2}$$

The time marginal distributions of this mean-field limit probability measure solve the McKean-Vlasov-Fokker-Planck equation under consideration in the sense of the distributions. See the seminal survey by Sznitman [1]. We emphasize that, within the context of singular interactions, most often it seems inaccessible to prove or even intuit accurate tail estimates on the probability distributions μ_s.

When the interactions kernels β and γ are so singular than no known theorem guarantees the well-posedness of the equation (1.2), numerical simulations may help to study it. In particular, simulations may give intuition on the (possibly local in time) finiteness of the expectations $\int_{\mathcal{C}} |\beta(x, \omega_s)|\ \mu(\mathrm{d}\omega)$ and $\int_{\mathcal{C}} |\gamma(x, \omega_s)|\ \mu(\mathrm{d}\omega)$ with $x \in \mathbb{R}^d$. Of course, this property is crucial to get existence of a solution. Similarly, we are interested in knowing whether $|\beta(x, \omega_s)|^2$ and $|\beta(x, \omega_s)|^2$ are μ-integrable and therefore whether we can apply a central-limit theorem to get the weak convergence rate of $(X_t^{1,N})$ to (X_t) in terms of the number N of particles. See Talay and Tomasevic [2] for an example of numerical simulations within a highly singular context.

Supposing that the system (1.1) is well-posed and propagates chaos for small times, considering that the empirical measure $\frac{1}{N}\sum_{i=1}^N \delta_{X_s^{i,N}}$ is a good approximation of the probability distribution of X_s, the numerical strategy the naturally consists in using it to test the finiteness of the expectations or variances of $|\beta(x, X_s)|$ and $|\gamma(x, X_s)|$.

In [3] Hawkins examines a related question in the simpler case of i.i.d. sequences: 'Does there exist a test, which makes the right decision with arbitrarily high probability if given sufficient data, of the hypothesis that a given random variable V has finite expectation?'. To address that question, the author introduces the set $\mathcal{G}$ (resp. $\mathcal{H}$) of densities with finite (resp. infinite) means. He also introduces the class $\mathcal{T}$ of sequential tests which terminate in finite time, whatever is the density of the data. The sets $\mathcal{G}$ and $\mathcal{H}$ would be distinguishable in $\mathcal{T}$ if: $\forall \epsilon$ it would exist a test ϕ in $\mathcal{T}$ s.t.

$$\begin{cases} \mathbb{E}_F(\phi) < \epsilon & \text{if } \ F \in \mathcal{G}, \\ \mathbb{E}_F(\phi) > 1 - \epsilon & \text{if } \ F \in \mathcal{H}. \end{cases}$$

It is proven in [3] that $\mathcal{G}$ and $\mathcal{H}$ are not distinguishable.

Hawkins' nice result led us to develop a test with a more restrictive objective than testing the finiteness of $\mathbb{E}|V|$. We actually consider the related question: 'How to determine from samples of a random variable V whether its probability distribution has heavy tails?'

In our Supplementary Material [4] we present some numerical experiments which illustrate Hawkins' Theorem [3] and the irrelevance of naive statistical procedures to determine the heaviness of the tails of a probability distribution.

In this article we construct and analize an asymptotic statistical test under the additional assumption that $X := \sqrt{|V|}$ belongs to some domain of attraction DA(α) of a stable law of index $0 < \alpha \le 2$ (see Sect. 2 below for reminders on domains of attractions and stable laws). More precisely, we develop an hypothesis test for which the null and alternative hypotheses respectively are:

$$\mathbf{H_0} : X \in \mathrm{DA}(2)$$

and

$$\mathbf{H_1} : \exists 0 < \alpha < 2, \quad X \in \mathrm{DA}(\alpha).$$

Our key observation is that X cannot have a finite second moment when $\mathbf{H_0}$ is rejected (and therefore $\mathbf{H_1}$ is accepted).

The construction of an effective test is not obvious because, first, the hypothesis test concerns tail properties of the unknown distribution and, second, when α is close to 2 the probability density of the symmetric α-stable law can hardly be distinguished from the standard Gaussian density on large finite intervals (see for example [5], Figs. 2 and 3). Our construction of an effective hypothesis test is original. Unexpectedly, it is based on fine properties of bivariations of semimartingales and a test for jumps which allows one to discriminate between discontinuous stable processes and Brownian motions.

Here is our main result.

Theorem 1.1. *Assume that X belongs to some domain of attraction. Consider and i.i.d. sample $X_1, \dots, X_m$ of X, and the statistic*

$$\widehat{\mathscr{S}_n^m} = \frac{\sum_{i=1}^{n-1} \left| \sum_{j=\lfloor \frac{m(i-1)}{n} \rfloor +1}^{\lfloor \frac{mi}{n} \rfloor} (X_j - \overline{X}_m) \right| \left| \sum_{j=\lfloor \frac{mi}{n} \rfloor +1}^{\lfloor \frac{m(i+1)}{n} \rfloor} (X_j - \overline{X}_m) \right|}{\sum_{i=1}^{n} \left| \sum_{j=\lfloor \frac{m(i-1)}{n} \rfloor +1}^{\lfloor \frac{mi}{n} \rfloor} (X_j - \overline{X}_m) \right|^2},$$

where $\overline{X}_m$ stands for the sample mean. Let z_q denote the q-quantile of a standard normal random variable and let $\sigma_\pi^2 := 1 + \frac{4}{\pi} - \frac{20}{\pi^2}$. The rejection region

$$C_{n,m} := \left\{ \left| \widehat{\mathscr{S}_n^m} - \frac{2}{\pi} \right| > z_{1-q/2} \sqrt{\frac{\sigma_\pi^2}{n}} \right\}$$

satisfies:

1. $\limsup_{n\to\infty} \limsup_{m\to\infty} \mathbb{P}\left(C_{n,m} | \mathbf{H_0}\right) \le q$.
2. $\lim_{n\to\infty} \lim_{m\to\infty} \mathbb{P}\left(C_{n,m} | \mathbf{H_1}\right) = 1$.

Plan of the paper: The plan of the paper is as follows.

In Section 2, we recall important results on stable laws and domains of attraction.

In Section 3, we introduce the statistic $\widehat{\mathscr{S}_n^m}$ on which is based our test.

In Section 4, we present our hypothesis test for domains of attraction of stable laws.

In Section 5, we prove the consistency of the statistic $\widehat{\mathscr{S}_n^m}$.

In Section 6, we prove a Central Limit theorem for $\widehat{\mathscr{S}_n^m}$.

In Section 7, we discuss some numerical experiments.

In our conclusion we comment on the performance and limitations of our test.

Finally, in the Appendix A we prove few intermediate technical results.

2. A few reminders on stable laws and domains of attraction

In this section we gather the few results on stable laws and domains of attraction which we need in the sequel. For their proofs and further information, see *e.g.* Feller [6], Section 8, Chapter 9, Embrechts *et al.* [7], Section 2, Chapter 2 or Whitt [8], Section 5, Chapter 4. We also reformulate a standard functional limit theorem under a form which prepares our hypothesis test to reject or accept $\mathbf{H_0}$.

Let $X, X_1, X_2, \dots$ be a sequence of non-degenerate i.i.d. random variables and let

$$S_m := \sum_{j=1}^{m} X_j, \quad m = 1, 2, \dots$$

One says that X, or the law of X, belongs to the domain of attraction of a given law $\mathcal{L}$ if there exist centering constants μ_m and positive normalizing constants c_m such that $(S_m - \mu_m)/c_m$ converges in distribution to $\mathcal{L}$.

The only probability laws which have a non-empty domain of attraction are the stable laws.

The probability distributions of stable laws are fully characterized by the following theorem (see *e.g.* Feller [6], Thm. 1a, Sect. 8, Chap. 9). We recall that a function $H : \mathbb{R}_+ \to \mathbb{R}$ is said to be slowly varying at infinity if $\lim_{s\to\infty} H(sx)/H(s) = 1$ for any $x > 0$. The logarithm is an example of such a function.

Theorem 2.1. *Let F be a probability distribution function. For $R > 0$ denote the truncated second moment of F by*

$$U(R) := \int_{-R}^{R} x^2 F(\mathrm{d}x).$$

(i) The probability distribution F belongs to the domain of attraction DA(2) *of the Gaussian distribution if and only if the function U is slowly varying at infinity.*

(ii) It belongs to some other domain of attraction DA(α) *with $0 < \alpha < 2$ if there exist a slowly varying at infinity function $H : \mathbb{R}_+ \to \mathbb{R}$ and positive numbers p and q such that*

$$\begin{cases} 1 - F(x) + F(-x) \sim \frac{2-\alpha}{\alpha}\ x^{-\alpha}\ H(x), \quad x \to \infty, \\ \lim_{x\to\infty} \frac{1-F(x)}{1-F(x)+F(-x)} = p, \quad \lim_{x\to\infty} \frac{F(-x)}{1-F(x)+F(-x)} = q. \end{cases}$$

The parameter α is called the characteristic exponent or the index of the stable law. Given $0 < \alpha \le 2$, every stable law with index α is called an α-stable distribution and denoted by $\mathcal{L}_\alpha$.

For specific values of α the stable law $\mathcal{L}_\alpha$ is simple. For $\alpha = 2$, $\mathcal{L}_\alpha$ is Gaussian (and possibly degenerate). For $\alpha = 1$, $\mathcal{L}_\alpha$ is Cauchy. For $\alpha = 1/2$, $\mathcal{L}_\alpha$ is Lévy. For $\alpha < 2$, $\mathcal{L}_\alpha$ has a density.

The preceding theorem implies the following moment properties (see *e.g.* Embrechts *et al.* [7], Cor. 2.2.10).

Theorem 2.2. *If the random variable X belongs to $D(\alpha)$ then*

$$\mathbb{E}(|X|^\delta) < \infty \quad \textit{for} \quad 0 < \delta < \alpha,$$
$$\mathbb{E}(|X|^\delta) = \infty \quad \textit{for} \quad \delta > \alpha \quad \textit{and} \quad \alpha < 2.$$

In particular, $\mathbb{E}(X^2) = \infty$ for $\alpha < 2$.

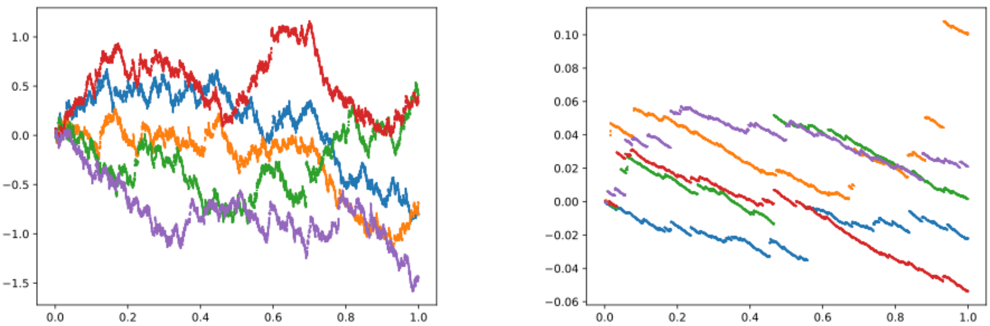
Remark 2.3. In view of Theorem 2.1, if $\mathbb{E}(X^2) < \infty$ then X belongs to DA(2). The converse is not true. For example, $X = \sqrt{|Y|}$ with Y Cauchy belongs to DA(2) and has an infinite second moment. This observation combined with Theorem 2.2 explains our choice of the null hypothesis.

In the Introduction we explained why often one cannot get a priori theoretical tail estimates on the probability distribution of particles with complex dynamics. Consequently, even if one a priori knows that this probability distribution belongs to some domain of attraction, to determine the parameter α is a hard statistical question. That was a motivation to build a specific hypothesis test which is based on the following limit theorems.

We start with Donsker's invariance principle for i.i.d. random variables. See *e.g.* Whitt [8], Theorems 4.5.2–4.5.3.

Theorem 2.4. *Let $X, X_1, X_2, \ldots$ be a sequence of non-degenerate i.i.d. random variables such that $X \in$* DA(α)*. Then there exist centering constants μ_m and normalizing constants c_m such that*

$$L^m := \left(\frac{S_{\lfloor mt \rfloor} - \mu_m t}{c_m}, t \ge 0 \right) \xrightarrow[m\to\infty]{\mathcal{D}} L,$$



(A) $r = 0.2$ (B) $r = 0.8$

FIGURE 1: Trajectories of L^m when the subjacent random variable $X \sim |G|^{-r}$, where $G \sim \mathcal{N}(0, 1)$. In the left column $r = 0.2$, therefore the limit of L^m is a Brownian motion and the trajectories of L^m seem to be continuous. In the right column, $r = 0.8$, hence the limit of L^m is a Lévy process and the trajectories of L^m seem to have jumps.

where L is a standard α-stable Lévy process if $\alpha < 2$, whereas for $\alpha = 2$ L is a Brownian motion. Here, one considers weak convergence in the Skorohod space endowed with J_1 topology.

For $\alpha < 2$, the trajectories of α-stable processes are a.s. discontinuous, whereas for $\alpha = 2$ the trajectories of the Brownian motion are a.s. continuous. In addition, one expects that for m large enough, the trajectories of $(S_{\lfloor mt \rfloor} - \mu_m t)/c_m$ resemble the trajectories of the limit process (see Fig. 1). Therefore, testing for jumps in the trajectories of $(S_{\lfloor mt \rfloor} - \mu_m t)/c_m$ should allow to discriminate between $X \in \mathrm{DA}(2)$ and $X \in \mathrm{DA}(\alpha)$. This is illustrated by Figure 1. Simulations have been run with $X = |G|^{-r}$, with G a standard normal random variable and $m = 10000$. For $r = 0.2$ we are under $\mathbf{H_0}$ and we can compute the explicit value of $\mathbb{E}(X)$ and $\mathbb{E}(X^2)$. For $r = 0.8$ we know that $|G|^{-r}$ is in the normal domain of attraction of stable distribution with index $1/r$. Hence, the normalizing constant is $c_m = m^{1/r}$. Indeed, respectively denoting by Φ and ϕ the cumulative distribution function and the density function of a standard normal distribution we have

$$\mathbb{P}(|G|^{-r} > x) = 2\Phi\left(\frac{1}{x^{1/r}}\right) - 1.$$

From L'Hôpital's Theorem it follows that

$$\lim_{x\to\infty} \frac{\mathbb{P}(|G|^{-r} > x)}{1/x^{1/r}} = \lim_{x\to\infty} \frac{2\Phi\left(1/x^{1/r}\right) - 1}{1/x^{1/r}} = \lim_{x\to\infty} \frac{2\phi\left(1/x^{1/r}\right)\left(1/x^{1/r}\right)'}{\left(1/x^{1/r}\right)'} = \sqrt{\frac{2}{\pi}}.$$

We thus deduce the desired result from [8], Theorem 4.5.2.

This heuristic approach has a severe drawback: we do not know the values of μ_m and c_m, in particular because we do not know α. Nevertheless, as Proposition 2.7 below shows that one can use the Mapping Theorem to bypass the fact that μ_m is unknown.

Remark 2.5. The conclusion of Theorem 2.4 still holds true in some cases when the X_i's are weakly dependent. Sufficient conditions to respectively obtain Brownian or stable Lévy limits can be found in Shao [9] or Tyran-Kamińska [10].

The following lemma will allow us to reformulate Theorem 2.4 as a statement where μ_m does not need to be known.

Lemma 2.6. *Denote by $D[0,1]$ the space of cadlag functions defined on $[0,1]$ and consider the mapping $\Psi : D[0,1] \to D[0,1]$ defined as*

$$\Psi(x)(t) := x(t) - tx(1).$$

The mapping Ψ is continuous for the the J_1-Skorokhod topology.

Proof. The addition is continuous in the J_1-Skorokhod topology if one of the summands is continuous. (See *e.g.* Jacod and Shiryaev [11], Prop. 1.23, Sect. 1b, Chap. VI). □

From the previous lemma and the Mapping Theorem we deduce:

Proposition 2.7. *Let*

$$Z^m := \left(\frac{S_{\lfloor mt \rfloor} - tS_m}{c_m}, t \geq 0 \right). \tag{2.1}$$

If $X \in \mathrm{DA}(\alpha)$, then

$$Z^m \xrightarrow[m\to\infty]{\mathcal{D}} Z := \Psi(L),$$

where L is a standard α-stable Lévy process for $\alpha < 2$ (and therefore $\Psi(L)$ is a discontinuous process), whereas for $\alpha = 2$ L is a Brownian motion (and therefore $\Psi(L)$ is a Brownian bridge).

To conclude, testing for jumps the trajectories of the limit process of Z^m defined in (2.1) should allow one to discriminate between $X \in \mathrm{DA}(2)$ and $X \in \mathrm{DA}(\alpha)$, $0 < \alpha < 2$. It now remains to construct an hypothesis test for the continuity of $\Psi(L)$ which does not suppose that c_m is known.

3. A bivariation based statistic

Determining whether a stochastic process Y is continuous or not from the observation of one single path at discrete times, is an important modelling issue in many fields, notably in economics and financial mathematics. It has been addressed by several authors in the last years. See for example Ait-Sahalia and Jacod [12] and the references therein.

Our situation is somehow different since we are not observing a trajectory of a given process. We rather are constructing one discrete time path by means of a normalization procedure of our (observed or simulated) data and this constructed trajectory is always discontinuous. We therefore aim to construct a statistic whose asymptotic properties will allow us to apply detection of jumps of semimartingale methods.

Following Barndorff-Nielsen and Shephard [13], for any stochastic process $(Y_t)_{0\leq t\leq 1}$ we set $\Delta_i^n Y = Y_{i/n} - Y_{(i-1)/n}$ and consider the realized bivariation, the realized quadratic variation and the normalized realized bivariation respectively defined by

$$\widehat{B}(Y,n) := \sum_{i=1}^{n-1} |\Delta_i^n Y||\Delta_{i+1}^n Y|, \quad \widehat{Q}(Y,n) := \sum_{i=1}^{n} |\Delta_i^n Y|^2, \quad \widehat{\mathscr{S}_n}(Y) := \frac{\widehat{B}(Y,n)}{\widehat{Q}(Y,n)}. \tag{3.1}$$

In [13], the authors consider processes (Y_t) of the form

$$Y_t = Y_0 + \int_0^t a_s \, \mathrm{d}s + \int_0^t \sigma_s \, \mathrm{d}W_s + \sum_{j=1}^{N_t} c_j, \tag{3.2}$$

where W is a Brownian motion and (N_t) is a counting process. They assume that a and σ are càdlàg processes, c_j are non-zero random variable, σ is pathwise bounded away from zero and the joint process (a, σ) is independent of the Brownian motion (W_t). Barndorff-Nielsen and Shephard provided a test to decide '$\mathcal{H}_0$: $(N_t) \equiv 0$' against '$\mathcal{H}_1$: $(N_t) \not\equiv 0$'.

Barndorff-Nielsen and Shephard's test is based on the statistic

$$\frac{1}{\sqrt{n}}\left(\frac{\frac{2}{\pi}\widehat{B}(Y,n)}{\widehat{Q}(Y,n)}-1\right)\frac{\int_0^t \sigma_s^2 \,\mathrm{d}s}{\sqrt{\int_0^t \sigma_s^4 \,\mathrm{d}s}}$$

or to a similar statistic depending on $\widehat{B}(Y,n)$, $\widehat{Q}(Y,n)$ and $\int_0^t \sigma_s^4 \,\mathrm{d}s$ only.

Barndorff-Nielsen and Shephard's analysis of the test (see [13], Thm. 1) does not apply in our context since Brownian bridges are not of the type (3.2) with (σ_s) satisfying the requested constraints.

The non-degeneracy of the integrands is also needed by Ait-Sahalia and Jacod (see Asm. 1-(f) in [12], p. 187). In addition, they use the p-variation of semimartingales with $p > 2$ to test continuous paths against jumps. We cannot follow the same way since we test 'Brownian bridge' against 'a transformed alpha-stable process' (recall Prop. 2.7) and the semimartingale representation of a Brownian bridge implies that the p-variation is asymptotically infinite in both cases.

We therefore need to introduce a new statistic which is adapted to our specific situation.

Definition 3.1. For any $n \in \mathbb{N}$ consider the real-valued functional $\widehat{\mathscr{S}_n}$ defined as follows: for $z \in D[0,1]$,

$$\widehat{\mathscr{S}_n}(z) := \frac{\sum_{i=1}^{n-1} |z(i/n) - z((i-1)/n)||z((i+1)/n) - z(i/n)|}{\sum_{i=1}^{n} |z(i/n) - z((i-1)/n)|^2}. \tag{3.3}$$

Let $m \in \mathbb{N}$, given an i.i.d. sample of $X : X_1, \ldots, X_m$ and the corresponding process Z^m as in (2.1) we define our statistic $\widehat{\mathscr{S}_n^m}$ as the normalized bivariation of Z^m, that is,

$$\widehat{\mathscr{S}_n^m} := \widehat{\mathscr{S}_n}(Z^m). \tag{3.4}$$

The next result provides a key property of the statistic $\widehat{\mathscr{S}_n^m}$.

Proposition 3.2. *The statistic $\widehat{\mathscr{S}_n^m}$ is scale-free and satisfies*

$$\widehat{\mathscr{S}_n^m} = \frac{\sum_{i=1}^{n-1}\left|\sum_{j=\lfloor\frac{m(i-1)}{n}\rfloor+1}^{\lfloor\frac{mi}{n}\rfloor}(X_j - \overline{X}_m)\right|\left|\sum_{j=\lfloor\frac{mi}{n}\rfloor+1}^{\lfloor\frac{m(i+1)}{n}\rfloor}(X_j - \overline{X}_m)\right|}{\sum_{i=1}^{n}\left|\sum_{j=\lfloor\frac{m(i-1)}{n}\rfloor+1}^{\lfloor\frac{mi}{n}\rfloor}(X_j - \overline{X}_m)\right|^2}.$$

Proof. Observe that

$$\begin{aligned}\widehat{\mathscr{S}_n^m} &:= \widehat{\mathscr{S}_n}(Z^m)\\ &= \frac{\sum_{i=1}^{n-1} |Z^m_{i/n} - Z^m_{(i-1)/n}||Z^m_{(i+1)/n} - Z^m_{i/n}|}{\sum_{i=1}^{n} |Z^m_{i/n} - Z^m_{(i-1)/n}|^2}\end{aligned}$$

$$= \frac{\sum_{i=1}^{n-1} \left| \frac{S_{\lfloor \frac{mi}{n} \rfloor} - \frac{i}{n} S_m}{c_m} - \frac{S_{\lfloor \frac{m(i-1)}{n} \rfloor} - \frac{(i-1)}{n} S_m}{c_m} \right| \left| \frac{S_{\lfloor \frac{m(i+1)}{n} \rfloor} - \frac{(i+1)}{n} S_m}{c_m} - \frac{S_{\lfloor \frac{mi}{n} \rfloor} - \frac{i}{n} S_m}{c_m} \right|}{\sum_{i=1}^{n} \left| \frac{S_{\lfloor \frac{mi}{n} \rfloor} - \frac{i}{n} S_m}{c_m} - \frac{S_{\lfloor \frac{m(i-1)}{n} \rfloor} - \frac{i-1}{n} S_m}{c_m} \right|^2}$$

$$= \frac{\sum_{i=1}^{n-1} \left| S_{\lfloor \frac{mi}{n} \rfloor} - S_{\lfloor \frac{m(i-1)}{n} \rfloor} - \frac{1}{n} S_m \right| \left| S_{\lfloor \frac{m(i+1)}{n} \rfloor} - S_{\lfloor \frac{mi}{n} \rfloor} - \frac{1}{n} S_m \right|}{\sum_{i=1}^{n} \left| S_{\lfloor \frac{mi}{n} \rfloor} - S_{\lfloor \frac{m(i-1)}{n} \rfloor} - \frac{1}{n} S_m \right|^2}$$

$$= \frac{\sum_{i=1}^{n-1} \left| \sum_{j=\lfloor \frac{m(i-1)}{n} \rfloor + 1}^{\lfloor \frac{mi}{n} \rfloor} (X_j - \overline{X}_m) \right| \left| \sum_{j=\lfloor \frac{mi}{n} \rfloor + 1}^{\lfloor \frac{m(i+1)}{n} \rfloor} (X_j - \overline{X}_m) \right|}{\sum_{i=1}^{n} \left| \sum_{j=\lfloor \frac{m(i-1)}{n} \rfloor + 1}^{\lfloor \frac{mi}{n} \rfloor} (X_j - \overline{X}_m) \right|^2},$$

which is the desired result. □

Remark 3.3. Notice that computing $\widehat{\mathscr{S}_n^m}$ only needs the values of the sample. In particular, one does not need the unknown centering and normalizing factors μ_m and c_m in Theorem 2.4.

4. Our hypothesis test for domains of attraction of stable laws

In this section we present our hypothesis test for $\mathbf{H_0}$ against $\mathbf{H_1}$. It is based on the following consistency property and Central Limit Theorem for the statistic $\widehat{\mathscr{S}_n^m}$.

The consistency of $\widehat{\mathscr{S}_n^m}$ is provided by the following proposition whose lenghthy proof is postponed to Section 5.

Proposition 4.1. *For any $\epsilon > 0$ one has*

$$\lim_{n\to\infty} \lim_{m\to\infty} \mathbb{P}\left(|\widehat{\mathscr{S}_n^m} - \kappa| \le \epsilon \right) = 1, \tag{4.1}$$

with

$$\kappa := \begin{cases} \frac{2}{\pi} & \text{when} \quad X \in \mathrm{DA}(2), \\ 0 & \text{when} \quad X \in \mathrm{DA}(\alpha),\ \alpha < 2. \end{cases}$$

The next proposition provides a Central Limit Theorem for the statistic $\widehat{\mathscr{S}_n^m}$. We postpone its proof to Section 6.

Proposition 4.2. *If the subjacent random variable X belongs to the domain of attraction of the normal law, for any bounded and continuous function $\varphi : \mathbb{R} \to \mathbb{R}$ we have*

$$\lim_{n\to\infty} \lim_{m\to\infty} \mathbb{E}\left[\varphi\left(\frac{\sqrt{n}}{\sigma_\pi} \left(\widehat{\mathscr{S}_n^m} - \frac{2}{\pi} \right) \right) \right] = \mathbb{E}\left[\varphi(\mathcal{N})\right],$$

where $\sigma_\pi^2 = 1 + \frac{4}{\pi} - \frac{20}{\pi^2}$ and $\mathcal{N}$ is a standard Gaussian random variable.

Propositions 4.1 and 4.2 allow us to propose and analyse the following hypothesis test:

Theorem 4.3. *Assume that X belongs to some domain of attraction. Consider and i.i.d. sample $X_1, \dots, X_m$ of a r.v. X. We consider the test hypotheses*

$\mathbf{H_0}$*:* $X \in \mathrm{DA}(2)$
and
$\mathbf{H_1}$*:* $\exists 0 < \alpha < 2,\ X \in \mathrm{DA}(\alpha)$.

Let z_q *denote the* q*-quantile of a standard normal random variable and* $\sigma_\pi^2 = 1 + \frac{4}{\pi} - \frac{20}{\pi^2}$*. The rejection region*

$$C_{n,m} := \left\{ \left| \widehat{\mathscr{S}_n^m} - \frac{2}{\pi} \right| > z_{1-q/2} \frac{\sigma_\pi}{\sqrt{n}} \right\} = \left\{ \frac{\sqrt{n}}{\sigma_\pi} \ (\widehat{\mathscr{S}_n^m} - \frac{2}{\pi}) < z_{q/2} \right\} \bigcup \left\{ \frac{\sqrt{n}}{\sigma_\pi} \ (\widehat{\mathscr{S}_n^m} - \frac{2}{\pi}) > z_{1-q/2} \right\}$$

satisfies:

1. $\limsup_{n\to\infty} \limsup_{m\to\infty} \mathbb{P}\left(C_{n,m} | \mathbf{H_0}\right) \le q$.
2. $\lim_{n\to\infty} \lim_{m\to\infty} \mathbb{P}\left(C_{n,m} | \mathbf{H_1}\right) = 1$.

Proof. To prove the first claim, we fix $\delta > 0$ and consider the function $\psi_\delta : \mathbb{R} \to \mathbb{R}$ given by:

$$\psi_\delta(x) = \begin{cases} 1 & \text{for} \quad x < z_{q/2}, \\ 1 - \frac{1}{\delta}(x - z_{q/2}) & \text{for} \quad z_{q/2} \le x \le z_{q/2} + \delta, \\ 0 & \text{for} \quad z_{q/2} + \delta < x < z_{1-q/2} - \delta, \\ \frac{1}{\delta}(x - z_{1-q/2}) & \text{for} \quad z_{1-q/2} \le x \le z_{1-q/2}, \\ 1 & \text{for} \quad z_{1-q/2} < x. \end{cases}$$

Therefore,

$$\mathbb{P}\left(C_{n,m} | \mathbf{H_0}\right) \le \mathbb{E}\left(\psi_\delta(\widehat{\mathscr{S}_n^m}) | \mathbf{H_0}\right).$$

In view of Proposition 4.2 we obtain

$$\limsup_{n\to\infty} \limsup_{m\to\infty} \mathbb{P}\left(C_{n,m} | \mathbf{H_0}\right) \le \mathbb{E}\left[\psi_\delta(\mathcal{N})\right],$$

where $\mathcal{N}$ is a standard Gaussian random variable. Thanks to the dominated convergence theorem we have

$$\lim_{\delta\to 0} \mathbb{E}\left[\psi_\delta(\mathcal{N})\right] = \mathbb{P}\left(\mathcal{N} < z_{q/2}\right) + \mathbb{P}\left(\mathcal{N} > z_{1-q/2}\right) = q.$$

We thus can conclude that

$$\limsup_{n\to\infty} \limsup_{m\to\infty} \mathbb{P}\left(C_{n,m} | \mathbf{H_0}\right) \le q.$$

As for the second claim, it suffices to use Proposition 4.1 since under $\mathbf{H_1}$ it holds that

$$\forall \epsilon > 0, \quad \lim_{n\to\infty} \lim_{m\to\infty} \mathbb{P}\left(\widehat{\mathscr{S}_n^m} \le \epsilon\right) = 1.$$

□

Remark 4.4. Other tests in the literature related to ours: In [14] Jurečková and Picek develop a statistical test for the heaviness of the tail of a distribution function F assuming that F is absolutely continuous and strictly increasing on the set $\{x : F(x) > 0\}$. In their case, for m_0 given, the null hypothesis is

$$H_{m_0}: \ x^{m_0}(1 - F(x)) \ge 1, \ \ \forall x > x_0 \text{ for some } x_0 \ge 0,$$

whereas the alternative hypothesis is

$$K_{m_0}:\ \limsup_{x\to\infty} x^{m_0}(1-F(x)) < 1.$$

Notice that for $m_0 = 2$ the non-rejection H_{m_0} implies that F has infinite second moment. Although this test has a very good behavior even for small samples, it is only applicable to absolutely continuous distributions whereas our test does not need such a condition, which is essential for applications to samples produced by complex simulations of random processes. Moreover, Jurečková and Picek's test applies to I.I.D. random variables, whereas our methodology can be extended to some weak dependence cases (see Sect. 7.4 below). This potentially allows applications to interacting particles which propagate chaos. Finally, we refer to [15] for other nonparametric techniques in the context of heavy tailed distributions.

5. Consistency of the statistic $\widehat{\mathscr{S}}_n^m$: Proof of Prop. 4.1

The objective of this section is to establish the consistency proposition 4.1.

To prove this proposition we need two preliminary results.

5.1. Two consistency properties of $\widehat{\mathscr{S}_n}$

Proposition 5.1. *Recall that we have set $Z = \Psi(L)$. For any $n \in \mathbb{N}$ it holds that*

$$\widehat{\mathscr{S}}_n^m \xrightarrow[m\to\infty]{\mathcal{D}} \widehat{\mathscr{S}_n}(Z). \tag{5.1}$$

Proof. For any positive integer n recall the functional $\widehat{\mathscr{S}_n} : D[0,1] \to \mathbb{R}$ defined in (3.3). This functional is not continuous on $D[0,1]$ for the Skorokhod topology. Its discontinuity set is

$$\begin{aligned} J_n :=& \{z \in D[0,1] : \exists i = 1,\dots,n-1 : \Delta z(i/n) > 0\} \\ &\cup \{z \in D[0,1] : \forall i = 0,\dots,n-1 : z(i/n) = z((i-1)/n)\}. \end{aligned}$$

Therefore, as soon as $\mathbb{P}(Z \in J_n) = 0$ we can apply the Continuous Mapping Theorem to get

$$\widehat{\mathscr{S}_n}(Z^m) \xrightarrow[m\to\infty]{\mathcal{D}} \widehat{\mathscr{S}_n}(Z).$$

On the one hand, if the subjacent random variable X belongs to DA(2), then Z is a Brownian Bridge and therefore $\mathbb{P}(Z \in J_n) = 0$.

On the other hand, if X belongs to DA(α) with $o < \alpha < 2$, then the limit process Z is equal to $\Psi(L)$ where L is a α-stable process. Notice that Z and L hav the same jumps. Since the probability of L having a jump at any fixed time is zero, it follows that $\mathbb{P}\,(Z$ has a jump at $i/n) = 0$. In addition, we have

$$\mathbb{P}(Z \in J_n) \le \sum_{i=1}^{n-1} \mathbb{P}\,(Z \text{ has a jump at } i/n) + \mathbb{P}\left(Z_{i/n} = 0,\ i = 1,\dots,n-1\right) = 0,$$

since since $Z_0 = Z_1 = 0$, each summand in the first term of the right-hand side is zero, and the second term in the right-hand side is zero because law(Z_t) is absolutely continuous with respect to the Lebesgue measure (see Bertoin [16], p. 218).

To summarize, in both cases we have $\mathbb{P}(Z \in J_n) = 0$ and (5.1) holds true. $\square$

Proposition 5.2. *Let Ψ be defined as in Lemma 2.6 and $\widehat{\mathscr{S}_n}$ be defined as in (3.1). Then,* *1. If $(Y_t)_{0\leq t\leq 1}$ is a standard Brownian motion one has*

$$\widehat{\mathscr{S}_n}(\Psi(Y)) \xrightarrow[n\to\infty]{a.s.} \frac{2}{\pi}.$$

2. If $(Y_t)_{0\leq t\leq 1}$ is an α-stable process starting from 0 one has

$$\forall 1<\alpha\leq 2, \quad \widehat{\mathscr{S}_n}(\Psi(Y)) \xrightarrow[n\to\infty]{a.s.} 0.$$

For $\alpha\leq 1$ the convergence holds in probability only.

Proof. In view of (3.1) we succesively consider $\widehat{Q}(\Psi(Y),n)$ and $\widehat{B}(\Psi(Y),n)$.

A straightforward computation leads to

$$\widehat{Q}(\Psi(Y),n) = \widehat{Q}(Y,n) - \frac{1}{n}Y_1^2. \tag{5.2}$$

The second term in the right-hand side converges to zero a.s. We thus have to prove the a.s. convergence of $\widehat{Q}(Y,n)$.

When Y is a standard Brownian motion we have

$$\widehat{Q}(Y,n) = \frac{1}{n}\sum_{i=1}^{n}\left(\sqrt{n}\Delta_i^n Y\right)^2 \xrightarrow[n\to\infty]{\text{a.s.}} 1,$$

thanks to the Strong Law of Large Numbers, from which

$$\widehat{Q}(\Psi(Y),n) \xrightarrow[n\to\infty]{\text{a.s.}} 1 \tag{5.3}$$

and

$$\lim_{n\to\infty}\widehat{\mathscr{S}_n}(\Psi(Y)) = \lim_{n\to\infty}\frac{\widehat{B}(\Psi(Y),n)}{\widehat{Q}(\Psi(Y),n)} = \lim_{n\to\infty}\widehat{B}(\Psi(Y),n).$$

Consequently, the first statement of the proposition results from the two following results which will be proven in the next subsection (see Lem. 5.3 and 5.4): For any Brownian motion Y,

$$\widehat{B}(Y,n) \xrightarrow[n\to\infty]{\text{a.s.}} \frac{2}{\pi}$$

and

$$\widehat{B}(Y,n) - \widehat{B}(\Psi(Y),n) \xrightarrow[n\to\infty]{\text{a.s.}} 0.$$

Now, when Y is an α-stable process with $\alpha<2$ we have

$$\widehat{Q}(Y,n) = \sum_{i=1}^{n}(\Delta_i^n Y)^2 = \frac{1}{n^{2/\alpha}}\sum_{i=1}^{n}\left(n^{1/\alpha}\Delta_i^n Y\right)^2$$

and the positive random variable $U_i = \left(n^{1/\alpha}\Delta_i^n Y\right)^2$ satisfies $\mathbb{P}(U_i>t)\sim t^{-\alpha/2}$. Consequently, as n goes to infinity, $\widehat{Q}(Y,n)$ converges in distribution to a stable law Q with index $\frac{\alpha}{2}<1$ (see Thm. 4.5.2 in [8], Sect. 5, Chap. 4), from which $\mathbb{P}(Q=0)=0$. From (5.2) it follows that $\widehat{Q}(\Psi(Y),n)$ also converges in distribution to Q. We now use the two following results which will be proven in the next subsection: If Y is an α-stable process starting from 0, then

$$\widehat{B}(Y,n) \xrightarrow[n\to\infty]{\text{a.s.}} 0$$

and

$$\lim_{n\to\infty} \widehat{B}(\Psi(Y),n) = \lim_{n\to\infty} \widehat{B}(Y,n),$$

where the limit is in probability.

Consequently, in view of Slutsky's lemma,

$$\widehat{\mathscr{S}_n}(\Psi(Y)) = \frac{\widehat{B}(\Psi(Y),n)}{\widehat{Q}(\Psi(Y),n)} \xrightarrow[n\to\infty]{\mathbb{P}} 0.$$

□

In the proof of Proposition 5.2 we have used the two following lemmas.

5.2. Two key lemmas

The first point of the first Lemma in this subsection is contained in Barndorff-Nielsen and Shephard [13], Theorem 4. We here give its easy proof for the sake of completeness. The second point requires less obvious arguments.

Lemma 5.3. *We have:*

1. *Let $(Y_t)_{0\le t\le 1}$ be a standard Brownian motion, then*

$$\widehat{B}(Y,n) \xrightarrow[n\to\infty]{a.s.} \frac{2}{\pi}.$$

2. *Let $(Y_t)_{0\le t\le 1}$ be an α-stable process starting from 0. For $\alpha>1$, we have*

$$\widehat{B}(Y,n) \xrightarrow[n\to\infty]{a.s.} 0,$$

whereas for $\alpha \le 1$ this last convergence holds in probability only.

Proof. 1. Let $(Y_t)_{0\le t\le 1}$ be a Brownian motion. Then,

$$\begin{aligned}\sum_{i=1}^{n-1} |\Delta_i^n Y||\Delta_{i+1}^n Y| &= \sum_{i=1}^{\lfloor n/2\rfloor} |\Delta_{2i-1}Y||\Delta_{2i}Y| + \sum_{i=1}^{\lfloor n/2\rfloor -1} |\Delta_{2i}Y||\Delta_{2i+1}Y| \\ &\quad + 2\left(\frac{n}{2} - \lfloor\frac{n}{2}\rfloor\right)|\Delta_{n-1}^n Y||\Delta_n^n Y| \\ &= \frac{1}{2}\frac{2}{n}\sum_{i=1}^{\lfloor n/2\rfloor} |\sqrt{n}\Delta_{2i-1}Y||\sqrt{n}\Delta_{2i}Y| + \frac{1}{2}\frac{2}{n}\sum_{i=1}^{\lfloor n/2\rfloor -1} |\sqrt{n}\Delta_{2i}Y||\sqrt{n}\Delta_{2i+1}Y| \\ &\quad + 2\left(\frac{n}{2} - \lfloor\frac{n}{2}\rfloor\right)|\Delta_{n-1}^n Y||\Delta_n^n Y|.\end{aligned}$$

Notice that $\sqrt{n}\Delta_i Y \sim \mathcal{N}(0,1)$. The Law of the Large Numbers imples

$$\frac{2}{n}\sum_{i=1}^{\lfloor n/2\rfloor} |\sqrt{n}\Delta_{2i-1}Y||\sqrt{n}\Delta_{2i}Y| \xrightarrow[n\to\infty]{\text{a.s.}} \frac{2}{\pi},$$

$$\frac{2}{n}\sum_{i=1}^{\lfloor n/2\rfloor} |\sqrt{n}\Delta_{2i}Y||\sqrt{n}\Delta_{2i+1}Y| \xrightarrow[n\to\infty]{\text{a.s.}} \frac{2}{\pi}.$$

In addition, the continuity of the trajectories of the Brownian motion implies that

$$\left|2\left(\frac{n}{2}-\lfloor\frac{n}{2}\rfloor\right)|\Delta^n_{n-1}Y||\Delta^n_n Y|\right| \leq 2|\Delta^n_{n-1}Y||\Delta^n_n Y| \xrightarrow[n\to\infty]{\text{a.s.}} 0,$$

Putting all together we get

$$\widehat{B}(Y,n) \xrightarrow[n\to\infty]{\text{a.s.}} \frac{2}{\pi}.$$

2. Let $(Y_t)_{0\leq t\leq 1}$ be an α-stable process starting from 0. In this case we have

$$\begin{aligned}\widehat{B}(Y,n) = {}& \frac{1}{2}\frac{2}{n^{2/\alpha}}\sum_{i=1}^{\lfloor n/2\rfloor} |n^{1/\alpha}\Delta_{2i-1}Y||n^{1/\alpha}\Delta_{2i}Y| + \frac{1}{2}\frac{2}{n^{2/\alpha}}\sum_{i=1}^{\lfloor n/2\rfloor-1} |n^{1/\alpha}\Delta_{2i}Y||n^{1/\alpha}\Delta_{2i+1}Y| \\ &+ 2\left(\frac{n}{2}-\lfloor\frac{n}{2}\rfloor\right)|\Delta^n_{n-1}Y||\Delta^n_n Y|,\end{aligned}$$

and for any $i=1,\dots,n$, $n^{1/\alpha}\Delta_i Y$ are independent α-stable random variables. We proceed by cases:

- $\alpha\in(1,2)$: Notice that

$$\frac{2}{n^{2/\alpha}}\sum_{i=1}^{\lfloor n/2\rfloor} |n^{1/\alpha}\Delta_{2i-1}Y||n^{1/\alpha}\Delta_{2i}Y| = \frac{1}{n^{2/\alpha-1}}\left(\frac{2}{n}\sum_{i=1}^{\lfloor n/2\rfloor} |n^{1/\alpha}\Delta_{2i-1}Y||n^{1/\alpha}\Delta_{2i}Y|\right).$$

The Law of Large Numbers implies that the term inside the parenthesis in the right-hand side a.s. converges to a finite quantity. Therefore the whole right-hand side a.s. tends to 0. Similarly,

$$\frac{2}{n^{2/\alpha}}\sum_{i=1}^{\lfloor n/2\rfloor-1} |n^{1/\alpha}\Delta_{2i}Y||n^{1/\alpha}\Delta_{2i+1}Y| \xrightarrow[n\to\infty]{\text{a.s.}} 0.$$

We finally observe that

$$\limsup_{n\to\infty}\left|2\left(\frac{n}{2}-\lfloor\frac{n}{2}\rfloor\right)|\Delta^n_{n-1}Y||\Delta^n_n Y|\right| \leq 2\lim_{n\to\infty}\left|Y_1 - Y_{1-\frac{1}{n}}\right|\left|Y_{1-\frac{1}{n}} - Y_{1-\frac{2}{n}}\right| = 0 \text{ a.s.}$$

since the paths of Y have left limits. Therefore, for $\alpha\in(1,2)$ we have

$$\widehat{B}(Y,n) \xrightarrow[n\to\infty]{\text{a.s.}} 0.$$

- $\alpha \in (0,1)$: In view of Embrechts and Goldie [17], corollary of Theorem 3 we have that

$$|n^{1/\alpha}\Delta_{2i-1}Y||n^{1/\alpha}\Delta_{2i}Y| \in \text{ DA}(\alpha).$$

Therefore, there exists a slowly varying function ℓ_0 such that

$$\frac{1}{n^{1/\alpha}\ell_0(n)}\sum_{i=1}^{\lfloor n/2\rfloor}|n^{1/\alpha}\Delta_{2i-1}Y||n^{1/\alpha}\Delta_{2i}Y|,$$

converges in distribution to an α-stable random variable. Hence

$$\frac{2}{n^{2/\alpha}}\sum_{i=1}^{\lfloor n/2\rfloor}|n^{1/\alpha}\Delta_{2i-1}Y||n^{1/\alpha}\Delta_{2i}Y| = \frac{\ell_0(n)}{n^{1/\alpha}}\frac{1}{n^{1/\alpha}\ell_0(n)}\sum_{i=1}^{\lfloor n/2\rfloor}|n^{1/\alpha}\Delta_{2i-1}Y||n^{1/\alpha}\Delta_{2i}Y|$$

tends to zero in probability, because is the product of a sequence of random variables converging in distribution and a real sequence converging to zero. For the other terms we can proceed as before. We thus are in a position to conclude that

$$\widehat{B}(Y,n)\xrightarrow[n\to\infty]{\mathbb{P}}0.$$

- $\alpha = 1$: We need to study normalized sums of the type

$$\frac{1}{2}\frac{2}{n^2}\sum_{i=1}^{\lfloor n/2\rfloor}|n\Delta_{2i-1}Y||n\Delta_{2i}Y|.$$

Denote by A such a quantity. Notice that for each i, $n\Delta_i Y$ has a 1-stable law. By again using Embrechts and Goldie [17], corollary of Theorem 3 we get that $|n\Delta_{2i-1}Y|n\Delta_{2i}Y|$ belongs to DA(1). Hence, we need to introduce centering factors to obtain the convergence of A.
In view of Whitt [8], Theorem 4.5.1 the centered and normalized sums

$$\frac{1}{(n/2)\ell_0(n/2)}\Bigg(\sum_{i=1}^{\lfloor n/2\rfloor}|n\Delta_{2i-1}Y||n\Delta_{2i}Y| \\ -(n/2)^2\ell_0(n/2)\mathbb{E}\left[\sin\left(\frac{|n\Delta_{2i-1}Y||n\Delta_{2i}Y|}{(n/2)\ell_0(n/2)}\right)\right]\Bigg)$$

converge in distribution to a 1-stable random variable. Then

$$A = \frac{\ell_0(n/2)}{2n}\Big(\text{Converging in Distribution Sequence}\Big) \\ +\ell_0(n/2)\mathbb{E}\left[\sin\left(\frac{|n\Delta_1 Y||n\Delta_2 Y|}{(n/2)\ell_0(n/2)}\right)\right],$$

The first term tends to 0 in probability because is the product of a deterministic sequence converging to 0 and a weakly convergent sequence of random variables. As for the second one, there exist 1-stable random variables S_1 and S_2 such that

$$\ell_0(n/2)\mathbb{E}\left[\sin\left(\frac{|n\Delta_{2i-1}Y||n\Delta_{2i}Y|}{(n/2)\ell_0(n/2)}\right)\right] = \ell_0(n/2)\mathbb{E}\left[\sin\left(\frac{|S_1 S_2|}{(n/2)\ell_0(n/2)}\right)\right] \leq \mathbb{E}\left[\sqrt{|S_1|}\right]^2 \frac{\sqrt{\ell_0(n/2)}}{\sqrt{n/2}},$$

where in the last inequality we have used $|\sin(x)| \leq \sqrt{|x|}$. Hence the right-hand side converges to 0. We thus conclude that

$$\widehat{B}(Y,n) \xrightarrow[n\to\infty]{\mathbb{P}} 0.$$

□

The second lemma in this subsection shows that $\widehat{B}(Y,n)$ and $\widehat{B}(\Psi(Y),n)$ have a similar asymptotic behaviour.

Lemma 5.4. *If $(Y_t)_{0\leq t\leq 1}$ is, either a Brownian motion or an α-stable process starting from 0 with $\alpha > 1$, then*

$$\widehat{B}(Y,n) - \widehat{B}(\Psi(Y),n) \xrightarrow[n\to\infty]{a.s.} 0.$$

Moreover, if $(Y_t)_{0\leq t\leq 1}$ is an α-stable process starting from 0 with $\alpha \leq 1$ this last convergence holds in probability only.

Proof of Lemma 5.4 . Since for all $a, b \in \mathbb{R}$ $||a| - |b|| \leq |a-b|$, we have

$$\left|\sum_{i=1}^{n-1} |\Delta_i^n \Psi(Y)||\Delta_{i+1}^n \Psi(Y)| - \sum_{i=1}^{n-1} |\Delta_i^n Y||\Delta_{i+1}^n Y|\right| \leq \sum_{i=1}^{n-1} \left|\Delta_i^n \Psi(Y)\Delta_{i+1}^n \Psi(Y) - \Delta_i^n Y \Delta_{i+1}^n Y\right|,$$

and then

$$\left|\sum_{i=1}^{n-1} |\Delta_i^n \Psi(Y)||\Delta_{i+1}^n \Psi(Y)| - \sum_{i=1}^{n-1} |\Delta_i^n Y||\Delta_{i+1}^n Y|\right| \leq \sum_{i=1}^{n-1} \left|\frac{Y_1}{n}\Delta_i^n Y\right| + \sum_{i=1}^{n-1} \left|\frac{Y_1}{n}\Delta_{i+1}^n Y\right| + \sum_{i=1}^{n-1} \left|\frac{1}{n^2}Y_1^2\right|. \tag{5.4}$$

We now analyze each term in the right-hand side of (5.4). We start with the third one:

$$\sum_{i=1}^{n-1} \left|\frac{1}{n^2}Y_1^2\right| = \frac{(n-1)}{n^2}Y_1^2 \xrightarrow[n\to\infty]{\text{a.s.}} 0.$$

As for the second term in the right-hand side of (5.4), we have

$$\sum_{i=1}^{n-1} \left|\frac{Y_1}{n}\Delta_i^n Y\right| = \frac{|Y_1|}{n^{1/\alpha}} \frac{1}{n} \sum_{i=1}^{n-1} \left|n^{1/\alpha}\Delta_i^n Y\right|.$$

Notice that, for $i = 1, \ldots, n-1$, $|n^{1/\alpha}\Delta_i^n Y|$ are i.i.d. random variables with stable law of index α.

When Y is a Brownian motion or when $\alpha > 1$, there exists $y \in \mathbb{R}$ such that $\mathbb{E}\left[\left|n^{1/\alpha}\Delta_i^n Y\right|\right] = y$ for any $n \geq 1$. Then, the Law of Large Numbers implies

$$\frac{1}{n}\sum_{i=1}^{n-1}\left|n^{1/\alpha}\Delta_i^n Y\right| \xrightarrow[n\to\infty]{\text{a.s.}} y,$$

from which

$$\sum_{i=1}^{n-1}\left|\frac{Y_1}{n}\Delta_i^n Y\right| = \frac{|Y_1|}{n^{1/\alpha}}\frac{1}{n}\sum_{i=1}^{n-1}\left|n^{1/\alpha}\Delta_i^n Y\right| \xrightarrow[n\to\infty]{\text{a.s.}} 0.$$

When $\alpha < 1$ we use Embrechts *et al.* [7], Theorem 2.2.15 to get that there exists a slowly varying function ℓ_0 such that

$$\frac{1}{\ell_0(n)}\sum_{i=1}^{n-1}|\Delta_i^n Y|$$

converges in distribution as $n \to \infty$ to a stable random variable. Therefore,

$$\sum_{i=1}^{n-1}\left|\frac{Y_1}{n}\Delta_i^n Y\right| = \frac{|Y_1\ell_0(n)|}{n}\frac{1}{\ell_0(n)}\sum_{i=1}^{n-1}|\Delta_i^n Y|\,,$$

is the product of a random variable which a.s. converges to 0 when n goes to infinity and of a random variable which converges in distribution, and thus converges to 0 in probability.

Finally, if $\alpha = 1$, we have that $n\Delta_i^n Y \in \mathrm{DA}(1)$. Therefore, in view of Whitt [8], Theorem 4.5.1, there exists a slowly varying function ℓ_0 such that

$$\frac{1}{n\ell_0(n)}\left(\sum_{i=1}^{n}|n\Delta_i^n Y| - n^2\ell_0(n)\mathbb{E}\left[\sin\left(\frac{n\Delta_1^n Y}{n\ell_0(n)}\right)\right]\right)$$

converges in distribution to a stable distribution. We conclude that

$$\begin{aligned}\sum_{i=1}^{n-1}\left|\frac{Y_1}{n}\Delta_i^n Y\right| = \frac{\ell_0(n)|Y_1|}{n}\frac{1}{n\ell_0(n)}&\left(\sum_{i=1}^{n}|n\Delta_i^n Y| - n^2\ell_0(n)\mathbb{E}\left[\sin\left(\frac{n\Delta_1^n Y}{n\ell_0(n)}\right)\right]\right)\\ &+ \ell_0(n)\mathbb{E}\left[\sin\left(\frac{n\Delta_1^n Y}{n\ell_0(n)}\right)\right]\end{aligned}$$

converges to zero in probability since the first term in the right-hand side is the product of a random variable which converges to zero a.s. when n goes to infinity and a a random variables whcih converges in distribution. As for the second term in the right-hand side, one can check that it goes to zero by proceeding as in the end of the proof of the previous proposition. To summarize, we have

$$\sum_{i=1}^{n-1}\left|\frac{Y_1}{n}\Delta_i^n Y\right| \xrightarrow[n\to\infty]{\mathbb{P}} 0.$$

We thus have obtained that

$$\sum_{i=1}^{n-1}\left|\frac{Y_1}{n}\Delta_{i+1}^n Y\right| \xrightarrow[n\to\infty]{\mathbb{P}} 0$$

and

$$\left|\sum_{i=1}^{n-1}|\Delta_i^n\Psi(Y)||\Delta_{i+1}^n\Psi(Y)| - \sum_{i=1}^{n-1}|\Delta_i^n Y|\;|\Delta_{i+1}^n Y|\right| \xrightarrow[n\to\infty]{\mathbb{P}} 0.$$

□

5.3. End of the proof of Proposition 4.1

We now are in a position to prove the proposition 4.1.

We start with proving that for any bounded and continuous function $\varphi:\mathbb{R}\to\mathbb{R}$ we have

$$\lim_{n\to\infty}\lim_{m\to\infty}\mathbb{E}\left[\varphi(\widehat{\mathscr{S}_n^m})\right] = \varphi(\kappa). \tag{5.5}$$

In view of Proposition 5.1 we have:

$$\lim_{m\to\infty}\left|\mathbb{E}\left[\varphi(\widehat{\mathscr{S}_n^m})\right] - \mathbb{E}\left[\varphi(\widehat{\mathscr{S}_n}(Z))\right]\right| = 0.$$

In addition, in view of Proposition 5.2 we have

$$\lim_{n\to\infty}\left|\mathbb{E}\left[\varphi(\widehat{\mathscr{S}_n}(Z))\right] - \varphi(\kappa)\right| = 0.$$

Therefore, the desired limit in (5.5) is obtained by letting m and then n tend to infinity in the right-hand side of

$$\left|\mathbb{E}\left[\varphi(\widehat{\mathscr{S}_n^m})\right] - \varphi(\kappa)\right| \le \left|\mathbb{E}\left[\varphi(\widehat{\mathscr{S}_n^m})\right] - \mathbb{E}\left[\varphi(\widehat{\mathscr{S}_n}(Z))\right]\right| + \left|\mathbb{E}\left[\varphi(\widehat{\mathscr{S}_n}(Z))\right] - \varphi(\kappa)\right|.$$

Now, fix $\epsilon>0$ and define the function ψ as follows: ψ takes the value 0 outside the closed interval $[\kappa-\epsilon,\kappa+\epsilon]$, growths linearly from 0 to 1 on $[\kappa-\epsilon,\kappa]$ and decreases linearly from 1 to 0 on $[\kappa,\kappa+\epsilon]$. From

$$\mathbb{P}\left(|\widehat{\mathscr{S}_n^m}-\kappa|\le\epsilon\right) = \mathbb{E}\left[\mathbb{1}_{[\kappa-\epsilon,\kappa+\epsilon]}(\widehat{\mathscr{S}_n^m})\right] \ge \mathbb{E}\left[\varphi(\widehat{\mathscr{S}_n^m})\right],$$

it results that

$$\lim_{n\to\infty}\lim_{m\to\infty}\mathbb{P}\left(|\widehat{\mathscr{S}_n^m}-\kappa|\le\epsilon\right)] \ge \lim_{n\to\infty}\lim_{m\to\infty}\mathbb{E}\left[\varphi(\widehat{\mathscr{S}_n^m})\right] = \varphi(\kappa) = 1.$$

6. A CLT for $\widehat{\mathscr{S}_n^m}$: Proof of Prop. 4.2

In this section we prove a Central Limit theorem for $\widehat{\mathscr{S}_n^m}$ defined as in Proposition 4.2.

We start with proving the CLT for the pair $(\widehat{B}(Y,n),\widehat{Q}(Y,n))$, with $(Y_t)_{0\le t\le 1}$ being a Brownian Motion.

6.1. Three CLT for $(\widehat{B}(Y,n), \widehat{Q}(Y,n))$

The first result in this subsection concerns the case where Y is a Brownian motion. It is a particular case of Barndorff-Nielsen and Shephard [13], Theorem 3. For the sake of completeness we here provide a straightforward proof adapted to the specific Brownian case.

Lemma 6.1. *Let $(Y_t)_{0\le t\le 1}$be a Brownian motion. One then has*

$$\left(\sqrt{n}\left[\widehat{B}(Y,n)-\frac{2}{\pi}\right], \sqrt{n}\left[\widehat{Q}(Y,n)-1\right]\right) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}_2\left(0,\Sigma\right),$$

where

$$\Sigma = \begin{pmatrix} 1+\frac{4}{\pi}-\frac{12}{\pi^2} & \frac{4}{\pi} \\ \frac{4}{\pi} & 2 \end{pmatrix}. \tag{6.1}$$

Proof. Notice that

$$\begin{aligned}&\left(\sqrt{n}\left[\widehat{B}(Y,n)-\frac{2}{\pi}\right], \sqrt{n}\left[\widehat{Q}(Y,n)-1\right]\right)\\ &\qquad = \frac{1}{\sqrt{n}}\left(\sum_{i=1}^{n-1}|\sqrt{n}\Delta_i^n Y||\sqrt{n}\Delta_{i+1}^n Y|-\frac{2}{\pi}, \sum_{i=1}^{n-}|\sqrt{n}\Delta_i^n Y|^2-1\right).\end{aligned}$$

The right-hand side has the same law as

$$\frac{1}{\sqrt{n}}\left(\sum_{i=1}^{n-1}\left[|G_i||G_{i+1}|-\frac{2}{\pi}\right], \sum_{i=1}^{n}\left[G_i^2-1\right]\right),$$

where G_i is a standard Normal random variable.

The vectors $(|G_i||G_{i+1}|-2/\pi, G_i^2-1)$ are identically distributed, centered, and such that $(|G_i||G_{i+1}|-2/\pi, G_i^2-1)$ is independent of $(|G_j||G_{j+1}|-2/\pi, G_j^2-1)$ if $|j-i|>1$. Hence, we can apply the Central Limit Theorem for 1-dependent sequences of vectors (see Hoeffding and Robbins [18], Thm. 3). The asymptotic covariance matrix is given by

$$\Sigma_{11} := \mathbb{E}\left[\left(|G_{i+1}||G_{i+2}|-\frac{2}{\pi}\right)^2\right]+2\mathbb{E}\left[\left(|G_i||G_{i+1}|-\frac{2}{\pi}\right)\left(|G_{i+1}||G_{i+2}|-\frac{2}{\pi}\right)\right] = 1+\frac{4}{\pi}-\frac{12}{\pi^2},$$

$$\Sigma_{12} := \mathbb{E}\left[\left(|G_1||G_2|-\frac{2}{\pi}\right)\left(G_1^2-1\right)\right]+\mathbb{E}\left[\left(|G_1||G_2|-\frac{2}{\pi}\right)\left(G_2^2-1\right)\right]+\mathbb{E}\left[\left(|G_2||G_3|-\frac{2}{\pi}\right)\left(G_1^2-1\right)\right] = \frac{4}{\pi},$$

and

$$\Sigma_{22} := \mathbb{E}\left[\left(G_1^2-1\right)^2\right]+2\mathbb{E}\left[\left(G_1^2-1\right)\left(G_2^2-1\right)\right] = 2.$$

$\square$

In the next Lemma we extend the previous result to Brownian Bridges.

Lemma 6.2. *Let $(Y_t)_{0\le t\le 1}$ be a Brownian motion. Then*

$$\mathbb{E}\left[\left(\sqrt{n}\left[\widehat{B}(\Psi(Y),n)-\widehat{B}(Y,n)\right]\right)^2\right]+\mathbb{E}\left[\left(\sqrt{n}\left[\widehat{Q}(\Psi(Y),n)-\widehat{Q}(Y,n)\right]\right)^2\right]\le\frac{C}{n}. \tag{6.2}$$

In addition, one has

$$\left(\sqrt{n}\left[\widehat{B}(\Psi(Y),n)-\frac{2}{\pi}\right],\sqrt{n}\left[\widehat{Q}(\Psi(Y),n)-1\right]\right)\xrightarrow[n\to\infty]{\mathcal{D}}\mathcal{N}_2\left(0,\Sigma\right), \tag{6.3}$$

where Σ is defined as in (6.1).

Proof. A straightforward computation shows that

$$\widehat{Q}(\Psi(Y),n)=\widehat{Q}(Y,n)-\frac{1}{n}Y_1^2.$$

Therefore, we have

$$\mathbb{E}\left[\left(\sqrt{n}\left[\widehat{Q}(\Psi(Y),n)-\widehat{Q}(Y,n)\right]\right)^2\right]=\mathbb{E}\left[\frac{1}{n}Y_1^4\right]=\frac{3}{n}.$$

Now consider a family of independent random variables $G_i\sim\mathcal{N}(0,1)$ and set $\bar{G}^n:=n^{-1}\sum_{i=1}^n G_i$. It holds that

$$\begin{aligned}\mathbb{E}\left[\left(\sqrt{n}\left[\widehat{B}(\Psi(Y),n)-\widehat{B}(Y,n)\right]\right)^2\right]&=n\sum_{i=1}^{n-1}\sum_{k=1}^{n-1}\mathbb{E}\Big[\left(|\Delta_i^n\Psi(Y)||\Delta_{i+1}^n\Psi(Y)|-|\Delta_i^nY||\Delta_{i+1}^nY|\right)\\&\qquad\times\left(|\Delta_k^n\Psi(Y)||\Delta_{k+1}^n\Psi(Y)|-|\Delta_k^nY||\Delta_{k+1}^nY|\right)\Big]\\&=\frac{1}{n}\sum_{i=1}^{n-1}\sum_{k=1}^{n-1}\mathbb{E}\Big[\left(|G_i-\bar{G}^n||G_{i+1}-\bar{G}^n|-|G_i||G_{i+1}|\right)\\&\qquad\times\left(|G_k-\bar{G}^n||G_{k+1}-\bar{G}^n|-|G_k||G_{k+1}|\right)\Big]\\&=:S_1+S_2,\end{aligned}$$

where we have split the sum according to $k=i,i+1$ in S_1 and $k\notin\{i,i+1\}$ in S_2.

As for S_1, notice that

$$|G_i-\bar{G}^n||G_{i+1}-\bar{G}^n|-|G_i||G_{i+1}|\le|\bar{G}^n||\bar{G}^n-G_i-G_{i+1}|,$$

and therefore

$$\begin{aligned}S_1&=\frac{1}{n}\sum_{i=1}^{n-2}\sum_{k=i}^{i+1}\mathbb{E}\left[\left(|G_i-\bar{G}^n||G_{i+1}-\bar{G}^n|-|G_i||G_{i+1}|\right)\left(|G_k-\bar{G}^n||G_{k+1}-\bar{G}^n|-|G_k||G_{k+1}|\right)\right]\\&\le\frac{1}{n}\sum_{i=1}^{n-2}\sum_{k=i}^{i+1}\sqrt{\mathbb{E}\left[\left(|G_i-\bar{G}^n||G_{i+1}-\bar{G}^n|-|G_i||G_{i+1}|\right)^2\right]}\\&\qquad\times\sqrt{\mathbb{E}\left[\left(|G_k-\bar{G}^n||G_{k+1}-\bar{G}^n|-|G_k||G_{k+1}|\right)^2\right]}\end{aligned}$$

$$
\begin{aligned}
&\leq \frac{1}{n}\sum_{i=1}^{n-2}\sum_{k=i}^{i+1}\sqrt{\mathbb{E}\left[|\bar{G}^n|^2|\bar{G}^n - G_i - G_{i+1}|^2\right]\mathbb{E}\left[|\bar{G}^n|^2|\bar{G}^n - G_k - G_{k+1}|^2\right]} \\
&\leq \frac{1}{n}\sum_{i=1}^{n-2}\sum_{k=i}^{i+1}\sqrt[4]{\mathbb{E}\left[|\bar{G}^n|^4\right]\mathbb{E}\left[|\bar{G}^n - G_i - G_{i+1}|^4\right]\mathbb{E}\left[|\bar{G}^n|^4\right]\mathbb{E}\left[|\bar{G}^n - G_k - G_{k+1}|^4\right]} \\
&\leq \frac{C}{n}.
\end{aligned}
$$

As for S_2, we use the following inequality which we admit for a while (see Sect. 6.4 for its proof):

$$
\mathbb{E}\Bigg[\left(|G_i - \bar{G}^n||G_{i+1} - \bar{G}^n| - |G_i||G_{i+1}|\right)\left(|G_k - \bar{G}^n||G_{k+1} - \bar{G}^n| - |G_k||G_{k+1}|\right)\Bigg] \leq \frac{C}{n^2}. \tag{6.4}
$$

This ends the proof of (6.2).

From (6.2) it follows that $(\sqrt{n}[\widehat{B}(\Psi(Y), n) - \widehat{B}(Y, n)], \sqrt{n}[\widehat{Q}(\Psi(Y), n)) - \widehat{Q}(Y, n))]$ converges in probability to zero. Therefore, Slutsky's Theorem implies that

$$
\begin{aligned}
&\left(\sqrt{n}\left[\widehat{B}(\Psi(Y), n) - \frac{2}{\pi}\right], \sqrt{n}\left[\widehat{Q}(\Psi(Y), n) - 1\right]\right) \\
&\qquad = \left(\sqrt{n}\left[\widehat{B}(\Psi(Y), n) - \widehat{B}(Y, n)\right], \sqrt{n}\left[\widehat{Q}(\Psi(Y), n)) - \widehat{Q}(Y, n))\right]\right) \\
&\qquad\quad + \left(\sqrt{n}\left[\widehat{B}(Y, n) - \frac{2}{\pi}\right], \sqrt{n}\left[\widehat{Q}(Y, n) - 1\right]\right)
\end{aligned}
$$

converges in distribution to $\mathcal{N}_2(0, \Sigma)$.

□

We thus now in a position to obtain the following result.

Proposition 6.3. *Let $(Y_t)_{0\leq t\leq 1}$ be a Brownian motion. One then has:*

$$
\sqrt{n}\left[\widehat{\mathscr{S}_n}(\Psi(Y)) - \frac{2}{\pi}\right] \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}\left(0, 1 + \frac{4}{\pi} - \frac{20}{\pi^2}\right).
$$

Proof. Notice that

$$
\begin{aligned}
\sqrt{n}\left[\frac{\widehat{B}(\Psi(Y), n)}{\widehat{Q}(\Psi(Y), n)} - \frac{2}{\pi}\right] &= \frac{\sqrt{n}}{\widehat{Q}(\Psi(Y), n)}\left[\widehat{B}(\Psi(Y), n) - \frac{2}{\pi}\widehat{Q}(\Psi(Y), n)\right] \\
&= \frac{1}{\widehat{Q}(\Psi(Y), n)}\left(\sqrt{n}\left[\widehat{B}(\Psi(Y), n) - \frac{2}{\pi}\right] - \frac{2}{\pi}\sqrt{n}\left[\widehat{Q}(\Psi(Y), n) - 1\right]\right).
\end{aligned}
$$

In view of Lemma 6.2 the term in the parenthesis in the right-hand side converges to a centered Gaussian random variable with variance

$$
\Sigma_{11} - 2\frac{2}{\pi}\Sigma_{12} + \frac{4}{\pi^2}\Sigma_{22} = 1 + \frac{4}{\pi} - \frac{12}{\pi^2} - 2\frac{2}{\pi} \times \frac{4}{\pi} + \frac{4}{\pi^2} \times 2 = 1 + \frac{4}{\pi} - \frac{20}{\pi^2}.
$$

In addition, $\widehat{Q}(\Psi(Y), n) \xrightarrow[n\to\infty]{\text{a.s.}} 1$ in view of (5.3). Therefore, Slutsky's Theorem allows us to conclude. □

6.2. End of the proof of Prop.4.2

We now are in a position to end the proof of Prop.4.2. Observe that for Y a Brownian motion

$$\begin{aligned}\left|\mathbb{E}\left[\varphi\left(\frac{\sqrt{n}}{\sigma_\pi}\left(\widehat{\mathscr{S}_n^m}-\frac{2}{\pi}\right)\right)\right]-\mathbb{E}\left[\varphi(\mathcal{N})\right]\right| & \\ \leq \left|\mathbb{E}\left[\varphi\left(\frac{\sqrt{n}}{\sigma_\pi}\left(\widehat{\mathscr{S}_n^m}-\frac{2}{\pi}\right)\right)\right]-\mathbb{E}\left[\varphi\left(\frac{\sqrt{n}}{\sigma_\pi}\left(\widehat{\mathscr{S}_n}(\Psi(Y))-\frac{2}{\pi}\right)\right)\right]\right| & \\ +\left|\mathbb{E}\left[\varphi\left(\frac{\sqrt{n}}{\sigma_\pi}\left(\widehat{\mathscr{S}_n}(\Psi(Y))-\frac{2}{\pi}\right)\right)\right]-\mathbb{E}\left[\varphi(\mathcal{N})\right]\right|. &\end{aligned}$$

Thanks to the Mapping Theorem and Proposition 2.7, we get that the first term in the right-hand side goes to zero when m tends to infinity. Thanks to Proposition 6.3, we get that the second term in the right-hand side also tends to zero when n tends to infinity.

6.3. A corollary to Tanaka's formula

Before proving Inequality (6.4) we need to deduce a key estimate from Tanaka's formula.

Proposition 6.4. *Let $(W_s, s \geq 0)$ a standard Brownian motion, $G \in \mathbb{R}^{2d}$ a random vector with density f_G with respect to the Lebesgue measure and independent of W and $f : \mathbb{R} \times \mathbb{R}^{2d} \to \mathbb{R}$. Let us denote $\underline{z}^i = (z_1, \dots, z_{i-1}, z_{i+1}, \dots, z_{2d}) \in \mathbb{R}^{2d-1}$, $\underline{z}^i(x) = (z_1, \dots, z_{i-1}, x, z_{i+1}, \dots, z_{2d}) \in \mathbb{R}^{2d}$ and*

$$f_z : \mathbb{R} \to \mathbb{R}, \;\; x \to f_z(x) := f(x, z).$$

Suppose:

1. *For any $z \in \mathbb{R}^d$, f_z is regular enough to apply the Itô-Tanaka formula*

$$f_z(W_t) = f_z(0) + \int_0^t D_- f_z(W_s)\, \mathrm{d}W_s + \frac{1}{2}\int_{-\infty}^{\infty} L_t^y(W)\, D^2 f_z(\mathrm{d}y), \tag{6.5}$$

   where $D_- f_z$ is the left derivative of f_z, $D^2 f_z$ is the second derivative of f in the distributional sense and $L_t^y(W)$ is the Local time of W.
2. *For any $z \in \mathbb{R}^d$, $f_z(0) = 0$.*
3. *For any $z \in \mathbb{R}^d$, $D_- f_z$ has polynomial growth such that the stochastic integral in (6.5) is a martingale.*
4. *For any $z \in \mathbb{R}^d$, the measure $D^2 f_z$ can be decomposed as*

$$D^2 f_z(\mathrm{d}y) = f_z''(y)\, \mathrm{d}y + \sum_{i=1}^{2d} \gamma_z^i\, \delta_{z_i}(\mathrm{d}y), \tag{6.6}$$

   where δ_a is de Dirac distribution at point $a \in \mathbb{R}$.
5. *$\mathbb{E}\left[f_G''(W_t)\right]$ is bounded in a neighborhood of $t = 0$.*
6. *For any $i = 1, \dots, 2d$, $\mathbb{E}\left[h_i(W_t)\right]$ is bounded in a neighborhood of $t = 0$, where*

$$h_i(x) = \left(\int_{\mathbb{R}^{2d-1}} \gamma^i_{\underline{z}^i(x)}\, f_G(\underline{z}^i|x)\, \mathrm{d}\underline{z}^i\right) f_{G_i}(x),$$

   where the $\gamma^i_\cdot$'s are given by (6.6), $f_G(\underline{z}^i|x)$ is the conditional density of the vector G knowing that $G_i = x$ and f_{G_i} is the Gaussian density of G_i.

Then one has $\mathbb{E}\left[f(W_t,G)\right]=O(t)$ for any t in a neighborhood of 0^+.
Moreover, assume now that conditions 5 and 6 are respectively replaced by

5'. *The function $x\to h_0(x):=\mathbb{E}\left[f''_G(x)\right]$ is null at zero, twice differentiable with first and second derivative that have at most polynomial growth.*
6'. *For any $i=1,\dots,2d$, $h_i(0)=0$, $h_i\in\mathcal{C}^2$, and h'_i and h''_i have at most polynomial growth.*

Then, it holds that $\mathbb{E}\left[f(W_t,G)\right]=O(t^2)$ for any t in a neighborhood of 0^+.

Proof. From Conditions 1, 2 and 3 if follows that

$$\mathbb{E}\left[f_z(W_t)\right]=\frac{1}{2}\mathbb{E}\left[\int_{-\infty}^{\infty}D^2f_z(\mathrm{d}y)L_t^y(W)\right].$$

From condition 4 it then results that

$$\mathbb{E}\left[f_z(W_t)\right]=\frac{1}{2}\mathbb{E}\left[\int_{-\infty}^{\infty}f''_z(y)L_t^y(W)\mathrm{d}y\right]+\frac{1}{2}\mathbb{E}\left[\sum_{i=1}^{2d}\gamma_z^iL_t^{z_i}(W)\right].$$

By using the Occupation Time Formula we thus get

$$\mathbb{E}\left[f_z(W_t)\right]=\frac{1}{2}\int_0^t\mathbb{E}\left[f''_z(W_s)\right]\mathrm{d}s+\frac{1}{2}\mathbb{E}\left[\sum_{i=1}^{2d}\gamma_z^iL_t^{z_i}(W)\right].$$

In addition, as G and W are independent we have

$$\begin{aligned}\mathbb{E}\left[f(W_t,G)\right]&=\mathbb{E}\left[\mathbb{E}\left[f(W_t,z)\right]|_{z=G}\right]\\&=\mathbb{E}\left[\mathbb{E}\left[f_z(W_t)\right]|_{z=G}\right]\\&=\frac{1}{2}\int_0^t\mathbb{E}\left[f''_G(W_s)\right]\mathrm{d}s+\frac{1}{2}\mathbb{E}\left[\sum_{i=1}^{2d}\gamma_G^iL_t^{G_i}(W)\right].\end{aligned}$$

Now observe that

$$\begin{aligned}\mathbb{E}\left[\gamma_G^iL_t^{G_i}(W)\right]&=\mathbb{E}\int_{\mathbb{R}^{2d}}\gamma_z^iL_t^{z_i}(W)f_G(z)dz\\&=\mathbb{E}\int_{-\infty}^{\infty}\int_{\mathbb{R}^{2d-1}}\gamma^i_{\underline{z}^i(z_i)}L_t^{z_i}(W)f_G(\underline{z}^i|z_i)f_{G_i}(z_i)\mathrm{d}\underline{z}^idz_i\\&=\mathbb{E}\int_{-\infty}^{\infty}\left(\int_{\mathbb{R}^{2d-1}}\gamma^i_{\underline{z}^i(z_i)}f_G(\underline{z}^i|z_i)\mathrm{d}\underline{z}^i\right)f_{G_i}(z_i)L_t^{z_i}(W)dz_i\\&=\mathbb{E}\int_{-\infty}^{\infty}h_i(z_i)L_t^{z_i}(W)dz_i\\&=\mathbb{E}\int_0^th_i(W_s)\mathrm{d}s,\end{aligned}$$

where in the last equality we have again used the Occupation Time formula. Consequently, we have

$$\mathbb{E}\left[f(W_t,G)\right]=\frac{1}{2}\int_0^t\mathbb{E}\left[f''_G(W_s)\right]\mathrm{d}s+\frac{1}{2}\sum_{i=1}^{2d}\int_0^t\mathbb{E}\left[h_i(W_s)\right]\mathrm{d}s.$$

If conditions 5 and 6 hold true, we have just obtained

$$|\mathbb{E}\left[f(W_t,G)\right]| \le C_1 t + C_d t.$$

If conditions 5' and 6' hold, we apply Itô's formula to $f''_G(W_s)$ and get

$$\mathbb{E}\left[f(W_t,G)\right] = \frac{1}{4}\int_0^t\int_0^s \mathbb{E}\left[h_0''(W_u)\right]\mathrm{d}u\mathrm{d}s + \frac{1}{4}\sum_{i=1}^{2d}\int_0^t\int_0^s \mathbb{E}\left[h_i''(W_u)\right]\mathrm{d}u\mathrm{d}s,$$

from which the desired result follows. □

6.4. Proof of Inequality (6.4)

In this subsection we prove the inequality (6.4) which we have used in the proof of Lemma 6.2.

Consider a family of independent random variables $G_i \sim \mathcal{N}(0,1)$ and set $\bar{G}^n := n^{-1}\sum_{i=1}^n G_i$. Define the function $q(x,a,b) : \mathbb{R}^3 \to \mathbb{R}$ by

$$q(x,a,b) := |x-a||x-b| - |a||b|.$$

We aim to prove: For any $k \notin \{i, i+1\}$ one has

$$\mathbb{E}\left[q(\bar{G}^n, G_i, G_{i+1})q(\bar{G}^n, G_k, G_{k+1})\right] \le \frac{C}{n^2}. \tag{6.7}$$

Notice that $\bar{G}^n$ can be expressed as the sum of a random variable independent of $(G_i, G_{i+1}, G_k, G_{k+1})$ and a random variable which a.s. converges to 0:

$$\bar{G}^n = \frac{1}{n}\sum_{j\notin\{i,i+1,k,k+1\}} G_i + \frac{G_i + G_{i+1} + G_k + G_{k+1}}{n} =: \bar{G}^{n\#} + \frac{H_{i,k}}{n}.$$

In addition, as

$$q(x+y,a,b) \le q(x,a,b) + |y||2x-a-b| + |y|^2,$$

for any $y \in \mathbb{R}$, we have that

$$\begin{aligned}
&\mathbb{E}\left[q(\bar{G}^n, G_i, G_{i+1})q(\bar{G}^n, G_k, G_{k+1})\right] \\
&\quad = \mathbb{E}\Bigg[q(\bar{G}^{n\#} + \frac{H_{i,k}}{n}, G_i, G_{i+1})q(\bar{G}^{n\#} + \frac{H_{i,k}}{n}, G_k, G_{k+1})\ \Bigg] \\
&\quad \le \mathbb{E}\Bigg[\ \left(q(\bar{G}^{n\#}, G_i, G_{i+1}) + |\frac{H_{i,k}}{n}||G_i + G_{i+1} - 2\bar{G}^{n\#}| + |\frac{H_{i,k}}{n}|^2\right) \\
&\qquad\quad \times \left(q(\bar{G}^{n\#}, G_k, G_{k+1}) + |\frac{H_{i,k}}{n}||G_k + G_{k+1} - 2\bar{G}^{n\#}| + |\frac{H_{i,k}}{n}|^2\right)\Bigg] \\
&\quad = \mathbb{E}\Bigg[q(\bar{G}^{n\#}, G_i, G_{i+1})q(\bar{G}^{n\#}, G_k, G_{k+1})\Bigg] \\
&\qquad + \mathbb{E}\Bigg[q(\bar{G}^{n\#}, G_i, G_{i+1})\left(|\frac{H_{i,k}}{n}||G_k + G_{k+1} - 2\bar{G}^{n\#}|\right)\Bigg]
\end{aligned}$$

$$+\mathbb{E}\bigg[\bigg(|\frac{H_{i,k}}{n}||G_i+G_{i+1}-2\bar{G}^{n\#}|\bigg)\,q(\bar{G}^{n\#},G_k,G_{k+1})\bigg]+\mathbb{E}\bigg[\frac{H_{i,k}^2}{n^2}A\bigg],$$

where A is random variable with finite second moment. An easy computation shows that

$$\mathbb{E}\bigg[\frac{H_{i,k}^2}{n^2}A\bigg]\le\frac{C}{n^2}.$$

Since $G_i, G_{i+1}, G_k, G_{k+1}$ are i.i.d. we have

$$\begin{aligned}\mathbb{E}\bigg[q(\bar{G}^{n\#},G_i,G_{i+1})\bigg(|\frac{H_{i,k}}{n}||G_k+G_{k+1}-2\bar{G}^{n\#}|\bigg)\bigg]&\\ =\mathbb{E}\bigg[\bigg(|\frac{H_{i,k}}{n}||G_i+G_{i+1}-2\bar{G}^{n\#}|\bigg)\,q(\bar{G}^{n\#},G_k,G_{k+1})\bigg]&.\end{aligned}$$

Hence, to prove (6.7) it is enough to prove

$$\mathbb{E}\bigg[q(\bar{G}^{n\#},G_i,G_{i+1})q(\bar{G}^{n\#},G_k,G_{k+1})\bigg]\le\frac{C}{n^2} \tag{6.8}$$

and

$$\mathbb{E}\bigg[q(\bar{G}^{n\#},G_i,G_{i+1})\bigg(|\frac{H_{i,k}}{n}||G_k+G_{k+1}-2\bar{G}^{n\#}|\bigg)\bigg]\le\frac{C}{n^2}. \tag{6.9}$$

Notice that $\bar{G}^{n\#}$ is independent of $(G_i, G_{i+1}, G_k, G_{k+1})$ and has the same probability distribution as a Brownian motion W at time $(n-4)/n^2$. Therefore, to prove the two preceding inequalities we aim to use Proposition 6.4. This step requires easy calculations only. We postpone it to the Appendix.

7. Numerical experiments

In this section we present some numerical experiments which illustrate our main theorem 4.3 and its limitations when applied to synthetic data. For two families of examples we examine the practical choice of the parameters m and n and their effects on the Empirical Rejection Rates and Type II errors. We also numerically compare the power of the test and the significance level in a neighborhood of the null hypothesis.

The interested reader can find an extended version of this section in our Supplementary Material [4].

7.1. Experiments under $\mathbf{H_0}$

In this subsection we consider the r.v. $X=|G|^{-r}$ with $G\sim\mathcal{N}(0,1)$ and $r>0$. Notice that X has finite moments of order smaller than $1/r$ and that $X\in\mathrm{DA}(2)$ when $r\in(0,1/2]$.

We have performed simulations for the parameter r taking values in $\{0.1,0.2,0.3,0.4,0.45\}$ and for levels of confidence $q=0.1$ and $q=0.05$. For each value of r, we have generated 10^4 independent samples with sizes from $m=10^5$ up to $m=10^8$. In the cases $r=0.4$ and $r=0.45$ we have also considered $m=10^9$. To study the respective effects of m and n, for each sample size m we have made the number of time steps n vary from $n=10^1$ up to $n=10^4$.

According to Theorem 4.3 one expects that the larger m and n are, the closer to q the empirical rejection rate is. The experiments show that increasing m improves the approximation of the empirical rejection rates to their expected values, whereas increasing n does not have the same effect. One actually observes that the

TABLE 1: $X = |G|^{-r}$. Values of the parameters m and n for which the best empirical rejection rate (ERR) of the $\widehat{\mathscr{S}_n^m}$-test is attained under $\mathbf{H_0}$ for different values of r and q.

r	m	n	ERR
0.1	10^8	$10^1, 10^2$	0.100
0.2	10^8	10^3	0.100
0.3	10^7	10^1	0.100
0.4	$10^7, 10^8$	10^1	0.101
0.45	$10^8, 10^9$	10^1	0.109

(a) $q = 0.1$.

r	m	n	ERR
0.1	10^8	10^4	0.050
0.2	10^8	10^4	0.050
0.3	10^8	10^1	0.050
0.4	10^8	10^1	0.053
0.45	10^8	10^1	0.061

(b) $q = 0.05$.

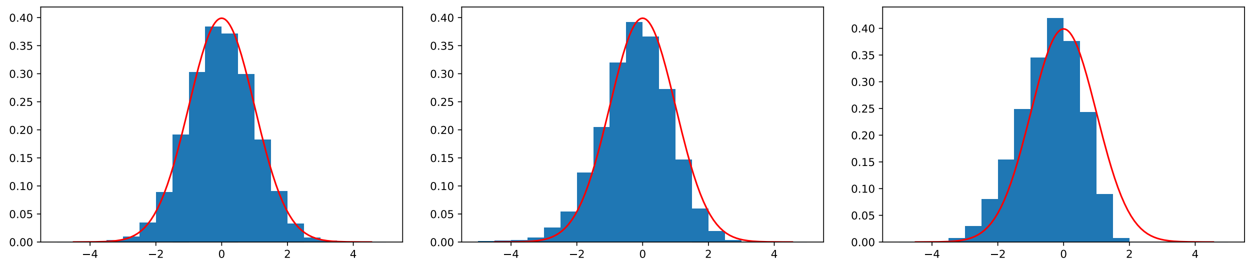


FIGURE 2: $X = |G|^{-r}$. The empirical distribution of standardized $\widehat{\mathscr{S}_n^m}$ for different values of r, m and n. The red line corresponds to the density of the standard normal distribution. In the left-hand side $r = 0.1$, $m = 10^8$ and $n = 10^4$. In the center $r = 0.4$, $m = 10^9$ and $n = 10^2$. In the right-hand side $r = 0.45$, $m = 10^9$ and $n = 10^2$.

empirical rejection rates become close to 1 when n becomes too large. Moreover, it seems that m being fixed the optimal choice of n depends on the value of r.

In other words, even if our theoretical result is valid when n goes to infinity, in practice, we should use n significantly smaller than m. In addition, as m increases, the convergence of the process Z^m can be very slow. Therefore, if n is chosen too large, one is zooming in too much and sees the discontinuities that Z^m has by construction. We continue the discussion on that issue in Sections 7.2 and 7.3.

In Tables 1a and 1b we report the experimental values of m and n for which, under $\mathbf{H_0}$, the best empirical rejection rate is attained. Notice that the best value for n decreases when the parameter r tends to the critical value $r = 0.5$ (that is, when the subjacent random variable has lower and lower finite moments). In particular, it does not seem to exist an optimal selection for the parameters m and n which is satisfying for any random variable in the domain of attraction of the Gaussian law.

In Figure 2 appear the empirical distribution of standardized $\widehat{\mathscr{S}_n^m}$ for different values of r, m and n. Notice that the adjust to the standard normal distribution is worse when r is closer to the critical value.

7.2. Experiments under $\mathbf{H_1}$

In this section we present empirical results which concern the Type II Error of our test, that is, the probability of failing to reject the Gaussian hypothesis $\mathbf{H_0}$ when the sample belongs to the domain of attraction $\mathrm{DA}(\alpha)$, with $\alpha < 2$, notably when α is close to 2. Under $\mathbf{H_1}$ the empirical rejection rate (ERR) is expected to be close to 1.

We consider two cases. First, the case where $X = |G|^{-r}$ with $G \sim \mathcal{N}(0,1)$ and $r > 1/2$, so that $X \in \mathrm{DA}(1/r)$. Second, the case where X has an α-stable distribution.

Concerning the case where $X = |G|^{-r}$, we have performed simulations for the parameter r taking values in $\{0.6, 0.75, 0.9\}$ and for levels of confidence $q = 0.1$ and $q = 0.05$. For each value of r, we have generated 10^4

TABLE 2: $X = |G|^{-r}$. First row: Empirical rejection rates (ERR) of the $\widehat{\mathscr{S}_n^m}$-test under $\mathbf{H_1}$ for $m = 10^8$ and different values of r, q. Second row: Empirical Type II Errors of the $\widehat{\mathscr{S}_n^m}$-test under $\mathbf{H_1}$ for $m = 10^8$ and different values of r, q. Third row: Empirical Type II Errors of the $\widehat{\mathscr{S}_n^m}$-test under $\mathbf{H_1}$ for $r = 0.55$ and different values of m, q, n.

r	$n = 10^1$	$n = 10^2$	$n = 10^3$
0.55	0.157	0.589	1.000
0.6	0.194	0.761	1.000
0.75	0.268	0.950	1.000
0.9	0.340	0.987	1.000

(a) $q = 0.1$.

r	$n = 10^1$	$n = 10^2$	$n = 10^3$
0.55	0.102	0.526	1.000
0.6	0.123	0.712	1.000
0.75	0.189	0.934	1.000
0.9	0.251	0.984	1.000

(b) $q = 0.05$.

r	$n = 10^1$	$n = 10^2$	$n = 10^3$
0.55	0.843	0.411	0.000
0.6	0.806	0.239	0.000
0.75	0.732	0.050	0.000
0.9	0.660	0.013	0.000

(c) $q = 0.1$.

r	$n = 10^1$	$n = 10^2$	$n = 10^3$
0.55	0.898	0.474	0.000
0.6	0.887	0.288	0.000
0.75	0.811	0.066	0.000
0.9	0.749	0.016	0.000

(d) $q = 0.05$.

	n		
m	200	250	500
10^5	0.077	0.042	0.001
10^6	0.126	0.076	0.004
10^7	0.164	0.109	0.011
10^8	0.192	0.130	0.019

(e) $q = 0.1$.

	n		
m	200	250	500
10^5	0.105	0.058	0.002
10^6	0.160	0.103	0.007
10^7	0.206	0.139	0.017
10^8	0.233	0.163	0.028

(f) $q = 0.05$.

independent samples with sizes from $m = 10^5$ up to $m = 10^8$, and to study the respective effects of m and n, for each sample size m we have made the number of time steps n vary from $n = 10^1$ up to $n = 10^4$. To evaluate the behavior of the test when the parameters are close to the limit case for moderate values of n, we also consider $r = 0.55$ with $n \in \{200, 250, 500\}$.

In Tables 2a and 2b we report the empirical rejection rates for different values of r and q when $n \leq 10^3$. Notice that the rates change between $n = 10^1$ and $n = 10^2$. Tables 2c and 2d show that the Type II Error increases when r decreases to the limit case $1/2$. Tables 2e and 2f show the satisfying behaviour of our test for moderate values of n and m in the case where $r = 0.55$ for which the subjacent random variable belongs to DA(20/11).

Concerning the case where X has an α-stable distribution, in our simulations the parameter α takes the values $\{0.9, 1.2, 1.5, 1.8\}$. The levels of confidence, sample sizes and numbers of steps are as above.

Similarly to the preceding tables, Tables 3a, 3b, 3c and 3d respectively concern empirical rejection rates and Type II Errors for large or moderate values of m.

In both cases, we observe that, if one increases m by letting n fixed, the quality of the test decreases. The situation seems symmetric of what we noticed above under H_0, although in this case increasing m for n fixed is like zooming out the constructed trajectory and leads the test to interpret a small jump in the limit trajectory as a side effect of the discretization and thus fails to reject the null hypothesis.

Moreover, Tables 2 and 3 allow one to deduce the Empirical Type II Error of the Test (**ETIIE**) for different fixed values of m. We observe that it decreases, approximately, exponentially fast with n. For example, when $X = |G|^{-r}$ with $r = 0.45$, one observes that **ETIIE** $\approx 0,9\exp(-0,007n)$. One thus may consider that **ETIIE** is

TABLE 3: α-stable X. First row: Empirical rejection rates (ERR) of the $\widehat{\mathscr{S}_n^m}$-test under $\mathbf{H_1}$ for $m = 10^7$ and different values of α, q. Second row: Empirical Type II Errors of the $\widehat{\mathscr{S}_n^m}$-test under $\mathbf{H_1}$ for $m = 10^7$ and different values of α, q. Third row: Empirical Type II Errors of the $\widehat{\mathscr{S}_n^m}$-test under $\mathbf{H_1}$ for $\alpha = 1.8$ and different values of m, q, n.

α	$n = 10^1$	$n = 10^2$	$n = 10^3$
0.9	0.396	0.997	1.000
1.2	0.300	0.977	1.000
1.5	0.224	0.883	1.000
1.8	0.162	0.611	1.000

(a) $q = 0.1$.

α	$n = 10^1$	$n = 10^2$	$n = 10^3$
0.9	0.302	0.996	1.000
1.2	0.218	0.968	1.000
1.5	0.153	0.854	1.000
1.8	0.102	0.552	1.000

(b) $q = 0.05$.

α	$n = 10^1$	$n = 10^2$	$n = 10^3$
0.9	0.604	0.003	0.000
1.2	0.700	0.023	0.000
1.5	0.776	0.117	0.000
1.8	0.838	0.389	0.000

(c) $q = 0.1$.

α	$n = 10^1$	$n = 10^2$	$n = 10^3$
0.9	0.698	0.004	0.000
1.2	0.782	0.032	0.000
1.5	0.847	0.146	0.000
1.8	0.898	0.448	0.000

(d) $q = 0.05$.

	n		
m	200	250	500
10^5	0.120	0.067	0.005
10^6	0.149	0.095	0.010
10^7	0.177	0.118	0.015

(e) $q = 0.1$.

	n		
m	200	250	500
10^5	0.154	0.091	0.007
10^6	0.187	0.126	0.016
10^7	0.220	0.150	0.023

(f) $q = 0.05$.

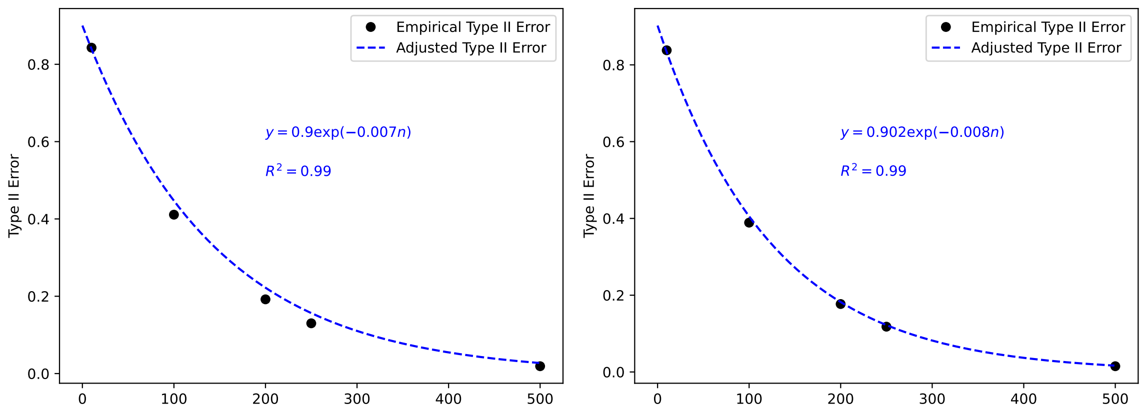


FIGURE 3: Empirical Type II Error as a function of n. Left: $X = |G|^{-r}$ with $r = 0.45$ and sample size $m = 10^8$. Right: α-stable X with $\alpha = 1.8$ and sample size $m = 10^7$. Both figures exhibit a satisfying Determination Coefficient R^2 larger than 0.99.

less than 0.049 when $n \geq 415$. Similarly, when X is 1.8-stable, one observes that **ETIIE** $\approx 0,902 \exp(-0,008n)$ and one can expect that **ETIIE** is less than 0.049 when $n \geq 364$. See Figure 3.

To summarize the tables in this section, in the cases where $X = |G|^{-r}$ and X has an α-stable distribution, one can observe that for the moderate value $n = 500$ and different levels of confidence and different values of m, the Type II error is much smaller than the level of confidence q, even when r is close to the critical value $r = 0.5$ (respectively, when α is close to 2).

7.3. Some heuristic computations towards convergence rate estimates and finite-sample unbiasedness

In this subsection we discuss the important issue of the convergence rate of our statistics when m and n tend to infinity. Accurate convergence rate estimates would allow to prove the finite-sample unbiasedness of our hypothesis test. As explained in Jureckova and Kalina [19], under $\mathbf{H_1}$ one actually expects that the rejection probability to be at least the significance level of the test. See [19] for proofs of this property for several rank tests under suitable fairly general conditions.

In our context, to accomplish this programme we have to face the complex algebraic expression of $\widehat{\mathscr{S}_n^m}$. So far, we have not succeeded to overcome the algebraic complexity of the necessary exact calculations. We thus limit ourselves to the particular case of a symmetric α-stable random variable X and an analysis which is not fully rigorous because of the use of a couple of approximations. However, our estimates below provide interesting insights to the respective roles of the parameters m and n and on the rate of convergence of $\lim_{m\to\infty} \mathbb{P}\left(C_{n,m}|\mathbf{H_1}\right)$ to 1 when n goes to infinity. Although asymptotic, these estimates also provide insights on the behaviour of our test for finite samples in the neighborhood of the hypothesis and the necessary order of magnitude of the parameters m and n to ensure under $\mathbf{H_1}$ a rejection probability larger that the significance level of the test.

Recall that

$$\widehat{\mathscr{S}_n^m} = \frac{\sum_{i=1}^{n-1} \left|\sum_{j=\lfloor \frac{m(i-1)}{n}\rfloor+1}^{\lfloor \frac{mi}{n}\rfloor} (X_j - \overline{X}_m)\right| \left|\sum_{j=\lfloor \frac{mi}{n}\rfloor+1}^{\lfloor \frac{m(i+1)}{n}\rfloor} (X_j - \overline{X}_m)\right|}{\sum_{i=1}^{n} \left|\sum_{j=\lfloor \frac{m(i-1)}{n}\rfloor+1}^{\lfloor \frac{mi}{n}\rfloor} (X_j - \overline{X}_m)\right|^2}.$$

We are interested in cases where $\alpha \approx 2$ for which the expectation of $|X|$ is finite . Thus the Strong Law of Large numbers applies. Assume that m is large enough to have the empirical mean $\bar{X}_m$ close to the mean of X, that is, to have $\bar{X}_m \approx 0$ a.s., and set

$$\widetilde{\mathscr{S}_n^m} := \frac{\sum_{i=1}^{n-1} \left|\sum_{j=\lfloor \frac{m(i-1)}{n}\rfloor+1}^{\lfloor \frac{mi}{n}\rfloor} X_j\right| \left|\sum_{j=\lfloor \frac{mi}{n}\rfloor+1}^{\lfloor \frac{m(i+1)}{n}\rfloor} X_j\right|}{\sum_{i=1}^{n} \left|\sum_{j=\lfloor \frac{m(i-1)}{n}\rfloor+1}^{\lfloor \frac{mi}{n}\rfloor} X_j\right|^2}.$$

Setting

$$W_i := \left(\frac{n}{m}\right)^{1/\alpha} \sum_{j=\lfloor \frac{mi}{n}\rfloor+1}^{\lfloor \frac{m(i+1)}{n}\rfloor} X_j$$

we have

$$\widetilde{\mathscr{S}_n^m} = \frac{\sum_{i=1}^{n-1} |W_{i-1}||W_i|}{\sum_{i=1}^{n} W_i^2}$$

Notice that for any integer i the random variable W_i has the same distribution symmetric α-stable as X.

Proceeding as in the proof of Lemma 5.3 we can show that the Strong Law of the Large numbers holds for the numerator. Therefore, for any n large enough,

$$\frac{1}{n}\sum_{i=1}^{n1}|W_{i-1}||W_i| \approx (\mathbb{E}|W_1|)^2.$$

In addition, we observe that

$$\mathbb{P}\left(|W_i|^2 > t\right) = 2\mathbb{P}\left(W_i > t^{1/2}\right) = 2\mathbb{P}\left(X_i > t^{1/2}\right) \sim t^{\alpha/2}.$$

We thus can proceed as in the proof of Proposition 5.2 (see p. 11) and apply Theorem 4.5.2 in [8], Section 5, Chapter 4 to the weighted sum

$$\frac{1}{n^{2/\alpha}}\sum_{i=1}^{n-1}|W_i|^2$$

to conclude that it converges in distribution to an $\alpha/2$-stable positive random variable $\mathcal{W}$.

Consequently,

$$\widetilde{\mathscr{S}}_n^m = \frac{1}{n^{2/\alpha-1}}\frac{\frac{1}{n}\sum_{i=1}^{n-1}|W_i||W_{i+1}|}{\frac{1}{n^{2/\alpha}}\sum_{i=1}^{n-1}|W_i|^2} \approx \frac{1}{n^{2/\alpha-1}}\frac{(\mathbb{E}|W_1|)^2}{\mathcal{W}},$$

where $\mathcal{W}$ is $\alpha/2$-stable. It comes:

$$\begin{aligned}
\mathbb{P}\left(\widehat{\mathscr{S}}_n^m < \frac{2}{\pi} - z_{1-q/2}\sqrt{\frac{\sigma_\pi^2}{n}}\right) &\approx \mathbb{P}\left(\widetilde{\mathscr{S}}_n^m < \frac{2}{\pi} - z_{1-q/2}\sqrt{\frac{\sigma_\pi^2}{n}}\right)\\
&\approx \mathbb{P}\left(\frac{1}{n^{2/\alpha-1}}\frac{(\mathbb{E}|W_1|)^2}{\mathcal{W}} < \frac{2}{\pi} - z_{1-q/2}\sqrt{\frac{\sigma_\pi^2}{n}}\right)\\
&\approx \mathbb{P}\left(\mathcal{W} > \frac{1}{n^{2/\alpha-1}}\frac{(\mathbb{E}|W_1|)^2}{\frac{2}{\pi} - z_{1-q/2}\sqrt{\frac{\sigma_\pi^2}{n}}}\right)\\
&\approx 1 - \mathbb{P}\left(\mathcal{W} \le \frac{1}{n^{2/\alpha-1}}\frac{(\mathbb{E}|W_1|)^2}{\frac{2}{\pi} - z_{1-q/2}\sqrt{\frac{\sigma_\pi^2}{n}}}\right).
\end{aligned}$$

Similarly,

$$\begin{aligned}
\mathbb{P}\left(\widehat{\mathscr{S}}_n^m > \frac{2}{\pi} + z_{1-q/2}\sqrt{\frac{\sigma_\pi^2}{n}}\right) &\approx \mathbb{P}\left(\frac{1}{n^{2/\alpha-1}}\frac{(\mathbb{E}|W_1|)^2}{\mathcal{W}} > \frac{2}{\pi} + z_{1-q/2}\sqrt{\frac{\sigma_\pi^2}{n}}\right)\\
&\approx \mathbb{P}\left(\mathcal{W} < \frac{1}{n^{2/\alpha-1}}\frac{(\mathbb{E}|W_1|)^2}{\frac{2}{\pi} + z_{1-q/2}\sqrt{\frac{\sigma_\pi^2}{n}}}\right).
\end{aligned}$$

We now use a bound for the density of $\mathcal{W}$. Let ξ be a $\overline{\alpha}$-stable random variable. Its characteristic function is of the form

$$\exp(i\ a\ \lambda - c^{\overline{\alpha}}\ |\lambda|^{\overline{\alpha}}(1 - i\ \beta\ \mathrm{sign}(\lambda)\ w(\lambda,\overline{\alpha})))$$

for some parameters $0 < \overline{\alpha} \leq 2$, $c > 0$ and $-1 \leq \beta \leq 1$, where

$$w(\lambda,\overline{\alpha}) = \begin{cases} \tan(\frac{\pi\overline{\alpha}}{2}) & \text{if} \quad \overline{\alpha} \neq 1, \\ -\frac{2}{\pi}\ \log(|\lambda|) & \text{if} \quad \overline{\alpha} = 1. \end{cases}$$

In our case, $\mathcal{W}$ is positive and thus $\beta = 1$ (see *e.g.* Nolan [20], Lem. 1.1, Chap. 1, Sect. 1.4).

We now use that the density of any α-stable random variable with skewness parameter $\beta = 1$ and scale parameter c is bounded by $\Gamma(1 + 1/\bar{\alpha})/c\pi$ (see Lem. 3.11 in [20]), so

$$\mathbb{P}\left(\mathcal{W} < \frac{1}{n^{2/\alpha-1}} \frac{(\mathbb{E}|W_1|)^2}{\frac{2}{\pi} + z_{1-q/2}\sqrt{\frac{\sigma_\pi^2}{n}}}\right) \leq \frac{\Gamma(1+2/\alpha)}{c\pi} \frac{1}{n^{2/\alpha-1}} \frac{(\mathbb{E}|W_1|)^2}{\frac{2}{\pi} + z_{1-q/2}\sqrt{\frac{\sigma_\pi^2}{n}}},$$

$$\mathbb{P}\left(\mathcal{W} \leq \frac{1}{n^{2/\alpha-1}} \frac{(\mathbb{E}|W_1|)^2}{\frac{2}{\pi} - z_{1-q/2}\sqrt{\frac{\sigma_\pi^2}{n}}}\right) \leq \frac{\Gamma(1+2/\alpha)}{c\pi} \frac{1}{n^{2/\alpha-1}} \frac{(\mathbb{E}|W_1|)^2}{\frac{2}{\pi} - z_{1-q/2}\sqrt{\frac{\sigma_\pi^2}{n}}}.$$

Finally,

$$\begin{aligned} \mathbb{P}(C_{n,m}|\mathbf{H}_1) &= \mathbb{P}\left(\widehat{\mathscr{S}_n^m} < \frac{2}{\pi} - z_{1-q/2}\sqrt{\frac{\sigma_\pi^2}{n}}\right) + \mathbb{P}\left(\widehat{\mathscr{S}_n^m} > \frac{2}{\pi} + z_{1-q/2}\sqrt{\frac{\sigma_\pi^2}{n}}\right) \\ &\geq 1 - \frac{C(\alpha,q)}{n^{\frac{2}{\alpha}-1}}, \end{aligned} \tag{7.1}$$

where the constant $C(\alpha, q)$ remains uniformly bounded for $\alpha \in (1, 2)$. Therefore, up to the couple of approximations made in the above calculation and the fact that $\alpha \in (1, 2)$, $\mathbb{P}(C_{n,m}|\mathbf{H}_1)$ increases to 1 at a rate close to $n^{2/\alpha-1}$. Notice that the parameter m is needed large enough only to ensure that $\bar{X}_m \approx 0$ a.s. The following useful convergence rate estimate w.r.t. m can be found in Basu *et al.* [21]:

Theorem 7.1. *Let (X_j) be a sequence of i.i.d. randonm variables with common probability density v_1. Suppose that they are centered and belong to the domain of attraction of a stable law of index $1 < \alpha < 2$ whose probability density is denoted by v_α. Suppose that their characteristic function belongs to $L^r(\mathbb{R})$ with $r \geq 1$. Finally, suppose that $\int_{\mathbb{R}} x^2\ |v_1(x) - v_\alpha(x)|\ \mathrm{d}x < \infty$. Then, for some positive number A, for any m large enough, the probability density v_m of the normalized sum $Z_m := \frac{A}{m^{1/\alpha}}\ \sum_{i=1}^m X_i$ satisfies:*

$$\sup_x (1 + |x|^\alpha)\ |v_m(x) - v_\alpha(x)| = \mathcal{O}(\frac{1}{m^{\frac{2}{\alpha}-1}}).$$

In the case of our simulations where the common law of the X_j's is stable we deduce that

$$\begin{aligned} \mathbb{P}(|\overline{X}_m| > \epsilon) &= \mathbb{P}(|Z_m| > A\epsilon\ m^{1-\frac{1}{\alpha}}) \\ &\leq \int_{|x|>A\epsilon m^{1-\frac{1}{\alpha}}} v_\alpha(x)\ \mathrm{d}x + \mathcal{O}(\frac{1}{m^{\frac{2}{\alpha}-1}}), \end{aligned}$$

We now use tail estimates from Gairing and Imkeller [22], Section 4.1. Let ξ be a centered $\overline{\alpha}$-stable random variable with characteristic function as above. Then, For any $x > 0$ it holds that

$$\mathbb{P}(\xi > x) = \frac{1}{\pi}\ \frac{\overline{\alpha}(\overline{\alpha}+1)}{\overline{\alpha}}\ \sin((\overline{\alpha}+\delta)\tfrac{\pi}{2})\ \frac{c^{\overline{\alpha}}}{\cos(\delta\frac{\pi}{2})}\ x^{-\overline{\alpha}} + \mathcal{O}(x^{-2\overline{\alpha}})$$

with $\delta := \frac{2}{\pi}\ \arctan(\beta \tan(\overline{\alpha}\frac{\pi}{2}))$. Similarly,

$$\mathbb{P}(\xi < -x) = \frac{1}{\pi}\ \frac{\overline{\alpha}(\overline{\alpha}+1)}{\overline{\alpha}}\ \sin((\overline{\alpha}-\delta)\tfrac{\pi}{2})\ \frac{c^{\overline{\alpha}}}{\cos(\delta\frac{\pi}{2})}\ x^{-\overline{\alpha}} + \mathcal{O}(x^{-2\overline{\alpha}}).$$

Here, $\overline{\alpha} = \alpha$ and we deduce:

$$\mathbb{P}(|\overline{X}_m| > \epsilon) \leq \frac{C}{\epsilon^{\alpha}\ m^{\alpha-1}} + \frac{C}{m^{\frac{2}{\alpha}-1}}. \tag{7.2}$$

In the preceding calculations, the approximation of $\widehat{\mathscr{S}_n^m}$ by the more tractable quantity $\widetilde{\mathscr{S}_n^m}$ is based on the hypothesis that the common law of the X_j's is stable and symmetric. In contrast, the result in [21] do not suppose that the X_i's have a stable law: They actually apply to the case of random variables in the domain of attraction of a stable law. Anyway, our estimates are certainly sub-optimal. For example, for $\alpha = 1.8$ the heuristic convergence rate is $n^{-0.1}$ whereas the results in Table 3 tend to show an exponential decay.

However, whilst not perfect, these estimates give some insight on the reason for which m needs to be chosen much larger than n, especially when α is close to 2. Actually, $\overline{X}_m$ can be seen as a random perturbation term in $\widehat{\mathscr{S}_n^m}$ whose expectation needs to be much smaller than the right-hand side of (7.1).

Finally, notice that, up to the approximation of $\bar{X}_m$ made in the above calculation, the inequalities (7.1) and (7.2) show that, for any small ϵ and any m and n large enough we have $\mathbb{P}(C_{n,m}|\mathbf{H}_1) > 1-\epsilon$. which suggests that our test is finite-sample unbiased.

7.4. Numerical experiments under weak dependence

The objective of this section is to show we can extend our test to some cases of weakly dependent data. Certainly, extra assumptions are necessary to obtain functional limit theorems by using the J_1-topology as in Theorem 2.4: For example, Avram and Taqqu [23] show that the J_1-weak convergence of normalized sums cannot hold for the classes of dependent random variables they consider.

Let us exhibit some sufficient conditions to replace the independency condition in Theorem 4.3. Assume that the sample is stationary, r-dependent and satisfy the two following conditions:

$$\sup\left\{|\operatorname{Corr}(f,g)| : \text{real } f \in L^2(\sigma(X_1)), \text{real } g \in L^2(\sigma(X_2, X_3, \ldots))\right\} < 1 \tag{7.3}$$

and

$$(\forall \epsilon > 0,\ \forall j = 2, \ldots, k) \quad \lim_{m\to\infty} \mathbb{P}\left(|X_j| > \epsilon c_m / |X_1| > \epsilon c_m\right) = 0. \tag{7.4}$$

Let us start with adapting the proof of Proposition 4.1 to this setting. First, if X is in the domain of attraction of the normal law, in view of (7.3) Corollary 1.1 in Shao [9] implies that $(S_{\lfloor mt \rfloor} - \mu_n)/c_m$ converges to a scaled Brownian motion, from which Z^m converges to a Brownian bridge. Second, if X belongs to the domain of attraction of a stable law, under (7.4) Corollary 1.4 in Tyran-Kamińska [10] states that $(S_{\lfloor mt \rfloor} - \mu_m)/c_m$ converges in distribution to a α-stable random process, from which Z^m converges in distribution to a α-stable bridge. The rest of the arguments in the proof of Proposition 4.1 do not rely on the independency hypothesis.

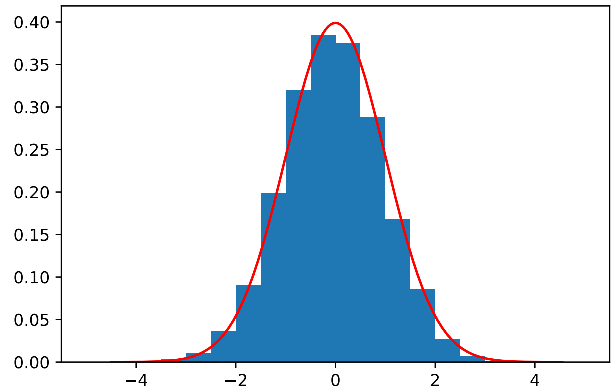


FIGURE 4: Empirical distribution of standardized $\widehat{\mathscr{S}}_n^m$ for the variables Y_k defined in (7.5). The red line corresponds to the density of the standard normal distribution.

Let us now turn to Proposition 4.2. Assume that X is in the domain of attraction of the normal law. In view of (7.3) Z^m converges to a Brownian bridge. Therefore, under $\mathbf{H_0}$,

$$\left|\mathbb{E}\left[\psi\left(\frac{\sqrt{n}}{\sigma_\pi}(\widehat{\mathscr{S}}_n^m - \frac{2}{\pi})\right)\right] - \mathbb{E}\left[\psi\left(\frac{\sqrt{n}}{\sigma_\pi}(\widehat{\mathscr{S}}_n(Z) - \frac{2}{\pi})\right)\right]\right|,$$

converges to zero, for every fixed n, as m goes to infinity. All the other arguments in the proof of Proposition 4.2 remain valid in this weak dependence setting.

We thus have shown one possible extension for Theorem 4.3 to samples of non-independent random variables.

To illustrate the behavior of the test for non-independent random variables, we consider a sequence $Y_2, \dots, Y_m$ of 1-dependent random variables constructed as follows. First, we generate a sample $X_1, \dots, X_m$ of independent copies of $|G|^{-r}$ where $G \sim \mathcal{N}(0,1)$, and then

$$Y_k = \frac{X_{k-1}}{X_{k-1}+1} X_k. \tag{7.5}$$

Second, we apply our methodology to $Y_2, \dots, Y_m$ for $r = 0.3$ and $r = 0.75$. Just as above we consider 10^4 scenarios.

For $r = 0.3$, $\mathbf{H_0}$ holds and the experiment shows that as m increases the empirical rejections rates tend to the theoretical ones, for example for $q = 0.1$, $m = 10^8$ and $n = 10^3$, the empirical rejection rate is 0.11. As in the i.i.d. case we also observe that the performance of the methodology decreases when n becomes too large.

For $r = 0.75$, that is, under $\mathbf{H_1}$, smaller values of n leads to bigger 'type II' error. Nevertheless, for $n = 10^2$ we observe empirical rejection rates close to 1.

8. Conclusion

Inspired by Barndorff-Nielsen and Shephard's methodology to test the existence of jumps in the trajectories of certain semimartingales, we have built an statistical test to determine if a given random variable X belongs to the domain of attraction of the Gaussian law or of another stable law. This statistical test allows us to give a partial answer to the ill-posed (but important in practice) problem of testing the finiteness of moments of an observed random variable.

So far, our methodology is based on asymptotic results. For the purpose of practical applications it would be interesting to obtain a non-asymptotic version of our test or, at least, to obtain theoretical results on suitable dependences between the parameters m and n which guarantee that our hypothesis test is reliable.

To obtain accurate estimates w.r.t. m and n for $\mathbb{P}(C_{n,m}|H_0)$ and $\mathbb{P}(C_{n,m}|H_1) = 1$ (which supposes to obtain non asymptotic estimates precising the CLT for $\widehat{\mathscr{S}_n^m}$) the calculations appear to be heavy, lengthy and very technical. We hope to be able to obtain satisfying results in the next future.

Appendix A. An elementary lemma on Gaussian distributions

In the proof below of the inequalities (6.8) and (6.9) we use the following elementary property of standard Gaussian laws. Recall that at the beginning of Section 6.4 the function q was defined as

$$q(x, a, b) = |x - a||x - b| - |a||b|.$$

Let ϕ, Φ respectively denote the density and the cumulative distribution function of a standard Gaussian random variable. In the computations below we will use the following elementary equalities that hold for any $x \in \mathbb{R}$:

$$\int_{-\infty}^{\infty} \operatorname{sgn}(x - z_1)\ \phi(z_1)\ dz_1 = 2\Phi(x) - 1, \tag{A.1}$$

$$\int_{-\infty}^{\infty} |x - z_2|\ \phi(z_2)\ dz_2 = 2x\Phi(x) - x + 2\phi(x), \tag{A.2}$$

and

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} q(x, z_3, z_4)\ \phi(z_3)\ \phi(z_4)\ dz_3\ dz_4 = (2x\Phi(x) - x + 2\phi(x))^2 - \frac{2}{\pi}. \tag{A.3}$$

A.1 Proof of Inequality (6.8)

Let $f(x, z_1, z_2, z_3, z_4)$ be defined as

$$f(x, z_1, z_2, z_3, z_4) := q(x, z_1, z_2)q(x, z_3, z_4).$$

Then

$$D_- f(x, z_1, z_2, z_3, z_4) = D_- q(x, z_1, z_2)q(x, z_3, z_4) + q(x, z_1, z_2)D_- q(x, z_3, z_4),$$

and

$$\begin{aligned} D^2 f(\mathrm{d}x, z_1, z_2, z_3, z_4) = f''(x, z_1, z_2, z_3, z_4)\mathrm{d}x \\ + 2|z_1 - z_2|q(z_1, z_3, z_4)\delta_{z_1}(\mathrm{d}x) + 2|z_2 - z_1|q(z_2, z_3, z_4)\delta_{z_2}(\mathrm{d}x) \\ + 2|z_3 - z_4|q(z_3, z_1, z_2)\delta_{z_3}(\mathrm{d}x) + 2|z_4 - z_3|q(z_4, z_1, z_2)\delta_{z_4}(\mathrm{d}x), \end{aligned}$$

where

$$\begin{aligned} f''(x, z_1, z_2, z_3, z_4) = 2\left[\operatorname{sgn}(x - z_1)|x - z_2| + |x - z_1| \operatorname{sgn}(x - z_2)\right] \\ \times \left[\operatorname{sgn}(x - z_3)|x - z_4| + |x - z_3| \operatorname{sgn}(x - z_4)\right] \\ + 2 \operatorname{sgn}(x - z_1) \operatorname{sgn}(x - z_2)q(x, z_3, z_4) \\ + 2 \operatorname{sgn}(x - z_3) \operatorname{sgn}(x - z_4)q(x, z_1, z_2). \end{aligned}$$

It is then easy to check that the conditions 1-4 of Proposition 6.4 hold true.

It remains to check that conditions 5' and 6' of Proposition 6.4 also hold true. We use the same notation as in the aforementioned proposition. Observe that

$$\begin{aligned}
h_0(x) &= \mathbb{E}\left[f''(x, G_i, G_{i+1}, G_k, G_{k+1})\right] \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f''(x, z_1, z_2, z_3, z_4)\phi(z_1)\phi(z_2)\phi(z_3)\phi(z_4)dz_1dz_2dz_3dz_4 \\
&= 8\left(\int_{-\infty}^{\infty} \operatorname{sgn}(x - z_1)\phi(z_1)dz_1\right)^2 \left(\int_{-\infty}^{\infty} |x - z_2|\phi(z_2)dz_2\right)^2 \\
&\quad + 4\left(\int_{-\infty}^{\infty} \operatorname{sgn}(x - z_1)\phi(z_1)dz_1\right)^2 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} q(x, z_3, z_4)\phi(z_3)\phi(z_4)dz_3dz_4.
\end{aligned}$$

In view of equalities (A.1), (A.2) and (A.3) a straightforward computation leads to

$$h_0(x) = 12\left(2\Phi(x) - 1\right)^2\left(2x\Phi(x) - x + 2\phi(x)\right)^2 - \frac{8}{\pi}\left(2\Phi(x) - 1\right)^2.$$

It is clear that $h_0(0) = 0$ and also that h is smooth enough to satisfy Condition 5' in Proposition 6.4. On the other hand, the weights of the singular part of $D^2 f(\mathrm{d}x, z_1, z_2, z_3, z_4)$ are given by

$$\begin{aligned}
\gamma_z^1 &= 2|z_1 - z_2|q(z_1, z_3, z_4), & \gamma_z^2 &= 2|z_2 - z_1|q(z_2, z_3, z_4), \\
\gamma_z^3 &= 2|z_3 - z_4|q(z_3, z_1, z_2), & \gamma_z^4 &= 2|z_4 - z_3|q(z_4, z_1, z_2).
\end{aligned}$$

Since $G_i, G_{i+1}, G_k, G_{k+1}$ are i.i.d. it follows

$$\begin{aligned}
h_1(x) &= \left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |x - z_2|q(x, z_3, z_4)\phi(z_2)\phi(z_3)\phi(z_4)dz_2dz_3dz_4\right)\phi(x) \\
&= \left(\int_{-\infty}^{\infty} |x - z_2|\phi(z_2)dz_2\right)\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} q(x, z_3, z_4)\phi(z_3)\phi(z_4)dz_3dz_4\right)\phi(x) \\
&= \left(2x\Phi(x) - x + 2\phi(x)\right)\left(\left(2x\Phi(x) - x + 2\phi(x)\right)^2 - \frac{2}{\pi}\right)\phi(x).
\end{aligned}$$

Notice that h_1 is $\mathcal{C}^\infty$, and its derivatives have polynomial growth, therefore h_1 satisfies Condition 6' in Proposition 6.4.

Since $G_i, G_{i+1}, G_k, G_{k+1}$ are identically distributed, one has that $h_2(x) = h_3(x) = h_4(x) = h_1(x)$, and therefore h_2, h_3 and h_4 also satisfy Condition 6' in Proposition 6.4.

Appendix B. Proof of Inequality (6.9)

Notice that

$$\begin{aligned}
&\mathbb{E}\left[q(\bar{G}^{n\#}, G_i, G_{i+1})\left(|\frac{H_{i,k}}{n}||G_k + G_{k+1} - 2\bar{G}^{n\#}|\right)\right] \\
&\leq \mathbb{E}\left[q(\bar{G}^{n\#}, G_i, G_{i+1})\left(|\frac{G_i + G_{i+1}}{n}| + |\frac{G_k + G_{k+1}}{n}|\right)\left(|G_k + G_{k+1}| + 2|\bar{G}^{n\#}|\right)\right] \\
&= \mathbb{E}\left[q(\bar{G}^{n\#}, G_i, G_{i+1})|\frac{G_i + G_{i+1}}{n}||G_k + G_{k+1}|\right] + 2\mathbb{E}\left[q(\bar{G}^{n\#}, G_i, G_{i+1})|\frac{G_i + G_{i+1}}{n}||\bar{G}^{n\#}|\right]
\end{aligned}$$

$$
\begin{aligned}
&+\mathbb{E}\Bigg[q(\bar{G}^{n\#},G_i,G_{i+1})\frac{|G_k+G_{k+1}|^2}{n}\Bigg]+2\mathbb{E}\Bigg[q(\bar{G}^{n\#},G_i,G_{i+1})|\frac{G_k+G_{k+1}}{n}||\bar{G}^{n\#}|\Bigg]\\
&=: E_1+E_2+E_3+E_4.
\end{aligned}
$$

We are now going to bound from above each E_i.

Again notice that $\bar{G}^{n\#}$ is equal in law to a Brownian motion W at time $(n-4)/n^2$, from which

$$
E_2=\frac{2}{n}\mathbb{E}\Bigg[|G_i+G_{i+1}|q(W_{\frac{n-4}{n^2}},G_i,G_{i+1})|W_{\frac{n-4}{n^2}}|\Bigg],
$$

We aim to apply the proposition 6.4 to the function

$$
f(x,z_1,z_2):=|z_1+z_2|\ q(x,z_1,z_2)\ |x|.
$$

In this case, $z=(z_1,z_2)$. We have:

$$
D_-f_z(x)=|z_1+z_2|D_-q(x,z_1,z_2)|x|+|z_1+z_2|q(x,z_1,z_2)\operatorname{sgn}(x),
$$

and

$$
D^2f_z(\mathrm{d}x)=f_z''(x)\mathrm{d}x+2|z_1+z_2||z_1-z_2|\delta_{z_1}(\mathrm{d}x)+2|z_1+z_2||z_2-z_1|\delta_{z_2}(\mathrm{d}x),
$$

where

$$
\begin{aligned}
f_z''(x)=&\,2|z_1+z_2|\left(\operatorname{sgn}(x-z_1)|x-z_2|+|x-z_1|\operatorname{sgn}(x-z_2)\right)\operatorname{sgn}(x)\\
&+2|z_1+z_2|\operatorname{sgn}(x-z_1)\operatorname{sgn}(x-z_2).
\end{aligned}
$$

We now show that for $i=0,1,2$, the functions $t\to\mathbb{E}\left[h_i(W_t)\right]$ are bounded for t in a neighborhood of 0. Indeed,

$$
\begin{aligned}
|h_0(x)|&=|\mathbb{E}\left[f_G''(x)\right]|\\
&\le\int_{\mathbb{R}^2}\left(2|z_1+z_2|\left(|x-z_2|+|x-z_1|\right)+2|z_1+z_2|\right)\phi(z_1)\phi(z_2)dz_1dz_2\\
&\le\int_{\mathbb{R}^2}\Big((z_1+z_2)^2+\left(|x-z_2|+|x-z_1|\right)^2+2|z_1+z_2|\Big)\phi(z_1)\phi(z_2)dz_1dz_2\\
&=2+\int_{\mathbb{R}^2}\Big((x-z_2)^2+2|x-z_2||x-z_1|+(x-z_1)^2\Big)\phi(z_1)\phi(z_2)dz_1dz_2+\frac{4}{\sqrt{\pi}}\\
&=2+2+2x^2+2\left(\int_{-\infty}^{\infty}|x-z_1|\phi(z_1)dz_1\right)\left(\int_{-\infty}^{\infty}|x-z_2|\phi(z_2)dz_1dz_2\right)+\frac{4}{\sqrt{\pi}}\\
&=4+2x^2+2\left(2x\Phi(x)-x+2\phi(x)\right)^2+\frac{4}{\sqrt{\pi}}\\
&\le C_1+C_2|x|+C_3x^2.
\end{aligned}
$$

Hence,

$$
\mathbb{E}\left[h_0(W_t)\right]\le C_1+C_2\mathbb{E}\left[|W_t|\right]+C_3\mathbb{E}\left[W_t^2\right]=C_1+C_2\sqrt{t}+C_3t.
$$

In particular, Condition 5 in Proposition 6.4 holds true.

Similarly,

$$|h_1(x)| = \left|2\left(\int_{-\infty}^{\infty} |x^2 - z_2^2|\phi(z_2)dz_2\right)\phi(x)\right| \leq 2\left(x^2+1\right).$$

Due to the symmetries of the weights $h_1 = h_2$, and therefore for $i = 1, 2$

$$\mathbb{E}\left[h_i(W_t)\right] \leq C_1 + C_2\mathbb{E}\left[W_t^2\right] = C_1 + C_2 t,$$

and Condition 6 in 6.4 also holds true. We therefore are in a position to use Proposition 6.4 and get

$$\mathbb{E}\left[|G_i + G_{i+1}|q(W_{\frac{n-4}{n^2}}, G_i, G_{i+1})|W_{\frac{n-4}{n^2}}|\right] \leq C\frac{n-4}{n^2},$$

from which

$$E_2 \leq \frac{C(n-4)}{n^3}.$$

Similar arguments allow us to show that $E_i \leq \frac{C(n-4)}{n^3}$ for $i = 1, 3, 4$. That ends the proof of (6.9).

Acknowledgements

The two authors thank the referees for their useful suggestions to improve the overall organisation of the manuscript and to clarify some mathematical issues. The first author is also grateful to Karine Bertin and Joaquin Fontbona for some helpfull discussions and comments on a previous version of this work. Partially supported by FONDECYT POSTDOCTORADO 3180777.

References

[1] A.-S. Sznitman, Topics in propagation of chaos, in *École d'Été de Probabilités de Saint-Flour XIX-1989*, Vol. 1464 of *Lecture Notes in Mathematics*. Springer, Berlin (1991) 165–251.

[2] D. Talay and M. Tomasevic, A stochastic particle method to solve the parabolic-parabolic Keller–Segel system: an algorithmic and numerical study. (2022) submitted for publication.

[3] D.L. Hawkins, Can the finiteness of a mean be tested? *Stat. Probab. Lett.* **32** (1997) 273–277.

[4] H. Olivero and D. Talay, Supplementary material to "a hypothesis test for the domain of attraction of a random variable". https://hal.science/hal-04266438 (2023).

[5] D.R. Holt and E.L. Crow, Tables and graphs of the stable probability density functions. *J. Res. Natl. Bureau Standards B* **77** (1973).

[6] W. Feller, An Introduction to Probability Theory and Its Applications. Vol. 2 of *An Introduction to Probability Theory and Its Applications*. Wiley (1971).

[7] P. Embrechts, C. Klüppelberg and T. Mikosch, Modelling Extremal Events: for Insurance and Finance. Stochastic Modelling and Applied Probability. Springer Berlin Heidelberg (1997).

[8] W. Whitt, Stochastic-process Limits: An Introduction to Stochastic-Process Limits and their Application to Queues. Springer Science & Business Media (2002).

[9] S. Qiman, An invariance principle for stationary p-mixing sequences with infinite variance. *Chinese Ann. Math. B* **14** (1993) 27–42.

[10] M. Tyran-Kamińska, Convergence to lévy stable processes under some weak dependence conditions. *Stochastic Processes Appl.* **120** (2010) 1629–1650.

[11] J. Jacod and A. Shiryaev, Limit Theorems for Stochastic Processes, Vol. 288. Springer Science & Business Media (2013).

[12] Y. Aït-Sahalia and J. Jacod, Testing for jumps in a discretely observed process. *Ann. Stat.* **37** (2009) 184–222.

[13] O.E Barndorff-Nielsen and N. Shephard, Econometrics of testing for jumps in financial economics using bipower variation. *J. Financ. Econ.* **4** (2006) 1–30.

[14] J. Jurečková and J. Picek, A class of tests on the tail index. *Extremes* **4** (2001) 165–183.

[15] N. Markovich, Nonparametric Analysis of Univariate Heavy-tailed Data: Research and Practice. John Wiley & Sons (2008).

[16] J. Bertoin, Lévy Processes, Vol. 121. Cambridge University Press (1998).

[17] P. Embrechts and C.M. Goldie, On closure and factorization properties of subexponential and related distributions. *J. Austr. Math. Soc.* **29** (1980) 243–256.

[18] W. Hoeffding and H. Robbins, The central limit theorem for dependent random variables. *Duke Math. J.* **15** (1948) 773–780.

[19] J. Jurečková and J. Kalina, Nonparametric multivariate rank tests and their unbiasedness. *Bernoulli* **18** (2012) 229–251.

[20] J.P. Nolan, Univariate Stable Distributions, Vol. 10. Springer (2020).

[21] S.K. Basu, M. Maejima and N.K. Patra, A non-uniform rate of convergence in a local limit theorem concerning variables in the domain of normal attraction of a stable law. *Yokohama Math. J.* **27** (1979) 63–72.

[22] J.M. Gairing and P. Imkeller, Stable clts and rates for power variation of α-stable Lévy processes. *Methodol. Comput. Appl. Probab.* **17** (2015) 73–90.

[23] F. Avram and M.S. Taqqu, Weak convergence of sums of moving averages in the α-stable domain of attraction. *Ann. Probab.* **20** (1992) 483–503.