Hölder norm test statistics for epidemic change*

Alfredas Račkauskas

Vilnius University and Institute of Mathematics and Informatics

Department of Mathematics, Vilnius University

Naugarduko 24, Lt-2006 Vilnius

Lithuania

Charles Suquet

Université des Sciences et Technologies de Lille

Mathématiques Appliquées, F.R.E. CNRS 2222

Bât. M2, U.F.R. de Mathématiques

F-59655 Villeneuve d'Ascq Cedex France

Abstract

To detect epidemic change in the mean of a sample of size n, we introduce new test statistics UI and DI based on weighted increments of partial sums. We obtain their limit distributions under the null hypothesis of no change in the mean. Under alternative hypothesis our statistics can detect very short epidemics of length $\log^{\gamma} n$, $\gamma > 1$. Using self-normalization and adaptiveness to modify UI and DI, allows us to prove the same results under very relaxed moment assumptions. Trimmed versions of UI and DI are also studied.

Keywords: change point, epidemic alternative, functional central limit theorem, Hölder norm, partial sums processes, selfnormalization.
Mathematics Subject Classifications (2000): 62E20, 62G10, 60F17.

*Research supported by a cooperation agreement CNRS/LITHUANIA (4714)

1 Introduction

An important question in the large area of change point problems involves testing the null hypothesis of no parameter change in a sample versus the alternative that parameter changes do take place at an unknown time. For a survey we refer to the books by Brodsky and Darkhovsky [4] or Csörgő and Horváth [5]. In this paper we suggest a class of new statistics for testing change in the mean under so called epidemic alternative. More precisely, given a sample $X_1, X_2, \dots, X_n$, we want to test the standard null hypothesis of constant mean

(H_0): $X_1, \dots, X_n$ *all have the same mean denoted by* μ_0,

against the epidemic alternative

(H_A): *there are integers* $1 < k^* < m^* < n$ *and a constant* $\mu_1 \neq \mu_0$ *such that* $\operatorname{E} X_i = \mu_0 + (\mu_1 - \mu_0)\mathbf{1}_{\{k^* < i \leq m^*\}}$, $i = 1, 2, \dots, n$.

Writing $l^* := m^* - k^*$ for the length of the epidemic, we assume throughout the paper that both l^* and $n - l^*$ go to infinity with n.

In this paper we follow the classical methodology to build test statistics by using continuous functionals of a partial sums process. Set

$$S(0) = 0, \quad S(t) = \sum_{k \leq t} X_k, \quad 0 < t \leq n.$$

When A is a set of integers, $S(A)$ will denote $\sum_{i \in A} X_i$. For instance we suggest to use with $0 < \alpha < 1/2$

$$\operatorname{UI}(n, \alpha) := \max_{1 \leq i < j \leq n} \frac{\left|S(j) - S(i) - S(n)(j/n - i/n)\right|}{\left[(j/n - i/n)\bigl(1 - (j - i)/n\bigr)\right]^{\alpha}}.$$

This is a functional, continuous in some Hölder topology, of the classical Donsker-Prokhorov polygonal line process. The statistics $\operatorname{UI}(n, 0)$ was suggested by Levin and Kline (1985), see [5].

To motivate the definition of such test statistics, let us assume just for a moment that the changes times k^* and m^* are known. Suppose moreover under (H_0) that the $(X_i,\ 1 \leq i \leq n)$ are independent identically distributed with finite variance σ^2. Under (H_A), suppose that the $(X_i,\ i \in \mathbb{I}_n)$ and the $(X_i,\ i \in \mathbb{I}_n^c)$ are separately independent identically distributed with the same finite variance σ^2, where

$$\mathbb{I}_n := \bigl\{i \in \{1, \dots, n\};\ k^* < i \leq m^*\bigr\}, \quad \mathbb{I}_n^c := \{1, \dots, n\} \setminus \mathbb{I}_n.$$

Then we simply have a two sample problem with known variances. It is then natural to accept (H_0) for small values of the statistics $|Q|$ and to reject it for large ones, where

$$Q := \frac{S(\mathbb{I}_n) - l^* S(n)/n}{(l^*)^{1/2}} - \frac{S(\mathbb{I}_n^c) - (n-l^*)S(n)/n}{(n-l^*)^{1/2}}.$$

After some algebra, Q may be recast as

$$Q = \frac{(n)^{1/2}}{\big(l^*(n-l^*)\big)^{1/2}} \Big(S(\mathbb{I}_n) - \frac{l^*}{n} S(n) \Big) \Big[(1 - l^*/n)^{1/2} + (l^*/n)^{1/2} \Big]$$

As the last factor into square brackets ranges between 1 and $2^{1/2}$, we may drop it and so replace $|Q|$ by the statistics

$$R := \frac{(n)^{1/2}}{\big(l^*(n-l^*)\big)^{1/2}} \Big| S(\mathbb{I}_n) - \frac{l^*}{n} S(n) \Big| = n^{-1/2} \frac{\big| S(\mathbb{I}_n) - \frac{l^*}{n} S(n) \big|}{\big[\frac{l^*}{n} \big(1 - \frac{l^*}{n} \big) \big]^{1/2}}.$$

Introducing now the notation

$$t_k = t_{n,k} := \frac{k}{n}, \quad 0 \le k \le n,$$

enables us to rewrite R as

$$R = n^{-1/2} \frac{\big| S(m^*) - S(k^*) - S(n)(t_{m^*} - t_{k^*}) \big|}{\big[(t_{m^*} - t_{k^*}) \big(1 - (t_{m^*} - t_{k^*}) \big) \big]^{1/2}}.$$

Now in the more realistic situation where k^* and m^* are unknown it is reasonnable to replace R by taking the maximum over all possible indexes for k^* and m^*. This leads to consider

$$\mathrm{UI}(n, 1/2) := \max_{1 \le i < j \le n} \frac{\big| S(j) - S(i) - S(n)(t_j - t_i) \big|}{\big[(t_j - t_i) \big(1 - (t_j - t_i) \big) \big]^{1/2}}.$$

It is worth to note that the same statistics arises from likelihood arguments in the special case where the observations X_i are Gaussian, see [14]. The asymptotic distribution of $\mathrm{UI}(n, 1/2)$ is unknown, due to difficulties caused by the denominator (for historical remarks see [5, p.183]).

In our setting, the X_i's are not supposed to be Gaussian. Moreover it seems fair to pay something in terms of normalization when passing from R to $n^{-1/2}\mathrm{UI}(n, 1/2)$. Intuitively the cost should depend on the moment assumptions made about the X_i's. To discuss this, let us introduce the polygonal partial sums process ξ_n defined by linear interpolation between the points $\big(t_k, S(k)\big)$. Then $\mathrm{UI}(n,1/2)$ appears as the discretization through the grid $(t_k,\, 0\le k\le n)$ of the functional $T_{1/2}(\xi_n)$ where

$$T_{1/2}(x) := \sup_{0<s<t<1} \frac{\big|x(t)-x(s)-\big(x(1)-x(0)\big)(t-s)\big|}{\big[(t-s)\big(1-(t-s)\big)\big]^{1/2}}. \tag{1}$$

This functional is continuous in the Hölder space $\mathcal{H}^o_{1/2}$ of functions $x:[0,1]\to\mathbb{R}$ such that $|x(t+h)-x(t)| = o(h^{1/2})$, uniformly in t. Obviously finite dimensional distributions of $n^{-1/2}\sigma^{-1}\xi_n$ converge to those of a standard Brownian motion W. However Lvy's theorem on the modulus of uniform continuity of W implies that $\mathcal{H}^o_{1/2}$ has too strong a topology to support a version of W. So $n^{-1/2}\xi_n$ cannot converge in distribution to σW in the space $\mathcal{H}^o_{1/2}$. This forbid us to obtain limiting distribution for $T_{1/2}(n^{-1/2}\xi_n)$ by invariance principle in $\mathcal{H}^o_{1/2}$ via continuous mapping. Fortunately, Hölderian invariances principles do exist for, roughly speaking, all the scale of Hölder spaces $\mathcal{H}^o_\rho$ of functions x such that $|x(t+h)-x(t)| = o\big(\rho(h)\big)$, uniformly in t, provided that the weight function ρ satisfies $\lim_{h\downarrow 0}\rho(h)(h\log|h|)^{-1/2} = \infty$. This type of invariance principles goes back to Lamperti [6] who studied the case $\rho(h) = h^\alpha$ $(0<\alpha<1/2)$. A complete characterization in terms of moments assumptions on X_1 was obtained recently by the authors in the general case (see Theorem 11 below). This leads us to replace $T_{1/2}(n^{-1/2}\xi_n)$ by $T_\alpha(n^{-1/2}\xi_n)$ obtained substituting the denominator in (1) by $(t-s)^\alpha\big(1-(t-s)\big)^\alpha$. Going back to the discretization we finally suggest the class UI (uniform increments) of statistics which includes particularly $\mathrm{UI}(n,\alpha)$ and similar ones $\mathrm{UI}(n,\rho)$ built with a general weight $\rho(h)$ instead of h^α. Together with UI we consider the class of DI (dyadic increments) statistics, which includes particularly

$$\mathrm{DI}(n,\alpha) = \max_{1\le j\le \log_2 n} 2^{j\alpha} \max_{r\in \mathrm{D}_j} \Big| S(nr) - \frac{1}{2}S(nr+n2^{-j}) - \frac{1}{2}S(nr-n2^{-j})\Big|,$$

where D_j is the set of dyadic numbers of the level j and $\log_2$ denotes the logarithm with basis 2. $\mathrm{DI}(n,\alpha)$ and $\mathrm{UI}(n,\alpha)$ have similar asymptotic behaviors. Moreover, dyadic increments statistics are of particular interest since their limiting distributions are completely specified (see Theorem 10 below).

Due to the independence of the X_i's, it is easy to see that even stochastic boundedness of either of $n^{-1/2}\mathrm{UI}(n,\alpha)$ or $n^{-1/2}\mathrm{DI}(n,\alpha)$ yields that of $n^{-1/2+\alpha}$ $\max_{1\le i\le n}|X_i|$. Hence, necessarily $\mathrm{P}(|X_1|>t) = O\big(t^{-p(\alpha)}\big)$, where $p(\alpha) = (1/2-\alpha)^{-1}$. So, heavier is the weight $\rho(h)$, stronger are the required moment assumptions to obtain the convergence of $n^{-1/2}\mathrm{UI}(n,\rho)$. On the other hand, the interest of a heavy weight is in the detection of short epidemics. Indeed (see Theorem 4 below), $\mathrm{UI}(n,0)$ can detect only epidemics whose the length l^* is such that $n^{1/2}=o(l^*)$. For $0<\alpha<1/2$, $\mathrm{UI}(n,\alpha)$ detects epidemics with $n^\delta=o(l^*)$ where $\delta=(1-2\alpha)/(2-2\alpha)$. With the weight $\rho(h)=h^{1/2}\log^\beta(c/h)$, $\beta>1/2$, $\mathrm{UI}(n,\rho)$ detects epidemics such that $\log^{2\beta}n=o(l^*)$.

We consider two ways to preserve the sensitiveness to short epidemics while relaxing moments assumptions. One is trimming and the other one is self-normalization and adaptive selection of partial sums increments. Trimming leads to a class of statistics, which includes

$$\mathrm{DI}(n,\alpha,\gamma)=\max_{1\le j\le\gamma\log_2 n}2^{j\alpha}\max_{r\in \mathrm{D}_j}\Big|S(nr)-\frac{1}{2}S(nr+n2^{-j})-\frac{1}{2}S(nr-n2^{-j})\Big|,$$

where $0<\gamma<1$. Adaptive construction of partial sums process and the corresponding functional central limit theorem proved in [10] allows to deal with a class of statistics which includes e.g. $\mathrm{SUI}(n,\alpha)$ obtained by replacing in $\mathrm{UI}(n,\alpha)$ the deterministic points t_k by the random $v_k:=V_k^2/V_n^2$, where $V_k^2=X_1^2+\cdots+X_k^2$, $k=1,\dots,n$,

$$\mathrm{SUI}(n,\alpha)=\max_{1\le i<j\le n}\frac{\big|S(j)-S(i)-S(n)(v_j-v_i)\big|}{\big[(v_j-v_i)\big(1-(v_j-v_i)\big)\big]^\alpha}.$$

When X_1 is symmetric, the convergence in distribution of $V_n^{-1}\mathrm{SUI}(n,\alpha)$ requires only the membership of X_1 in the domain of attraction of normal law, otherwise the existence of $\mathrm{E}\,|X_1|^{2+\varepsilon}$ for some $\varepsilon>0$ is sufficient.

The paper is organized as follows. In Section 2, definitions of classes of statistics are presented and limiting distributions are given. All proofs are deferred to Section 3.

2 UI and DI statistics and their asymptotics

All the test statistics studied in this paper may be viewed as discretizations of some Hölder norms or semi-norms. The following subsection contains the relevant background.

2.1 The Hölderian framework

Let $\rho : [0,1] \to \mathbb{R}_+$ be a weight function. Membership of a continuous function $x : [0,1] \to \mathbb{R}$ in the Hölder space $\mathcal{H}_\rho^o$ means roughly that $|x(t+h)-x(t)| = o\big(\rho(|h|)\big)$ uniformly in t. The classical scale of Hölder spaces uses $\rho(h) = h^\alpha$, $0 < \alpha < 1$. Lévy's theorem on the modulus of continuity of the Brownian motion restricts the investigation of invariance principles in this scale to the range $0 < \alpha < 1/2$. The same result leads naturally to consider the Hölder spaces built with the functions $\rho(h) = h^{1/2}\log^\beta(c/h)$ for $\beta > 1/2$. We include these two cases of practical interest in the following rather general class $\mathcal{R}$.

Definition 1. *We denote by $\mathcal{R}$ the class of non decreasing functions $\rho : [0,1] \to \mathbb{R}_+$ satisfying*

i) for some $0 < \alpha \le 1/2$, and some function L, positive on $[1,\infty)$ and normalized slowly varying at infinity,

$$\rho(h) = h^\alpha L(1/h), \quad 0 < h \le 1;$$

ii) $\theta(t) = t^{1/2}\rho(1/t)$ is C^1 on $[1,\infty)$;

iii) there is a $\beta > 1/2$ and some $a > 0$, such that $\theta(t)\log^{-\beta}(t)$ is non decreasing on $[a,\infty)$.

Let us recall that L is normalized slowly varying at infinity if and only if for every $\delta > 0$, $t^\delta L(t)$ is ultimately increasing and $t^{-\delta}L(t)$ is ultimately decreasing [3, Th. 1.5.5]. The main practical examples we have in mind may be parametrized by

$$\rho(h) = \rho(h,\alpha,\beta) := h^\alpha \log^\beta(c/h).$$

We write $C[0,1]$ for the Banach space of continuous functions $x : [0,1] \to \mathbb{R}$ endowed with the supremum norm $\|x\|_\infty := \sup\{|x(t)|;\ t \in [0,1]\}$.

Definition 2. *Let ρ be a real valued non decreasing function on $[0,1]$, null and right continuous at 0. We denote by $\mathcal{H}_\rho^o$ the Hölder space*

$$\mathcal{H}_\rho^o := \{x \in C[0,1];\ \lim_{\delta\to 0}\omega_\rho(x,\delta) = 0\},$$

where

$$\omega_\rho(x,\delta) := \sup_{\substack{s,t\in[0,1],\\ 0<t-s<\delta}} \frac{|x(t)-x(s)|}{\rho(t-s)}.$$

$\mathcal{H}^o_\rho$ is a separable Banach space for its native norm

$$\|x\|_\rho := |x(0)| + \omega_\rho(x,1).$$

Under technical assumptions, satisfied by any ρ in $\mathcal{R}$, the space $\mathcal{H}^o_\rho$ may be endowed with an equivalent norm $\|x\|^{\mathrm{seq}}_\rho$ built on weighted dyadic increments of x. Let us denote by D_j the set of dyadic numbers in $[0,1]$ of level j, i.e.

$$\mathrm{D}_0 = \{0,1\}, \qquad \mathrm{D}_j = \big\{(2l-1)2^{-j};\ 1\le l\le 2^{j-1}\big\}, \quad j\ge 1.$$

We write D (resp. D^*) for the sets of dyadic numbers in $[0,1]$ (resp. $(0,1]$)

$$\mathrm{D} := \bigcup_{j=0}^{\infty} \mathrm{D}_j, \quad \mathrm{D}^* := \mathrm{D}\setminus\{0\}.$$

Put for $r\in\mathrm{D}_j$, $j\ge 0$,

$$r^- := r-2^{-j}, \quad r^+ := r+2^{-j}.$$

For any function $x:[0,1]\to\mathbb{R}$, define its Schauder coefficients $\lambda_r(x)$ by

$$\lambda_r(x) := x(r) - \frac{x(r^+)+x(r^-)}{2}, \quad r\in\mathrm{D}_j,\ j\ge 1 \tag{2}$$

and in the special case $j=0$ by $\lambda_0(x) := x(0)$, $\lambda_1(x) := x(1)$. When ρ belongs to $\mathcal{R}$, we have on the space $\mathcal{H}^o_\rho$ the equivalence of norms (see [11] and [12])

$$\|x\|_\rho \sim \|x\|^{\mathrm{seq}}_\rho := \sup_{j\ge 0} \frac{1}{\rho(2^{-j})} \max_{r\in\mathrm{D}_j} |\lambda_r(x)|.$$

Let $W=(W(t), t\in[0,1])$ be a standard Wiener process and $B=(B(t), t\in[0,1])$ the corresponding Brownian bridge $B(t)=W(t)-tW(1)$, $t\in[0,1]$. Consider for ρ in $\mathcal{R}$, the following random variables

$$\mathrm{UI}(\rho) := \sup_{0<t-s<1} \frac{|B(t)-B(s)|}{\rho\big((t-s)(1-(t-s))\big)} \tag{3}$$

and

$$\mathrm{DI}(\rho) = \sup_{j\ge 1} \frac{1}{\rho(2^{-j})} \max_{r\in\mathrm{D}_j} \Big| W(r) - \frac{1}{2}W(r^+) - \frac{1}{2}W(r^-)\Big| = \|B\|^{\mathrm{seq}}_\rho. \tag{4}$$

These variables serve as limiting for uniform increment (UI) and dyadic increment (DI) statistics respectivly. No analytical form seems to be known for the distribution function of $\mathrm{UI}(\rho)$, whereas the distribution of $\mathrm{DI}(\rho)$ is completely specified by Theorem 10 below.

Remark 1. *We do not treat in this paper the problem of low regularity Hölder norms associated to the weights $\rho(h) = L(1/h)$ with L slowly varying. In this case nothing seems known about equivalence of the norms $\| \ \|_\rho$ and $\| \ \|_\rho^{\mathrm{seq}}$, which plays a key rôle in our proofs of Hölderian invariance principles. The extension of our results to this boundary case would require other technics and seems to be of limited practical scope.*

2.2 Statistics $\mathrm{UI}(n, \rho)$ and $\mathrm{DI}(n, \rho)$

To simplify notation put

$$\varrho(h) := \rho\big(h(1-h)\big), \quad 0 \le h \le 1.$$

For $\rho \in \mathcal{R}$, define (recalling that $t_k = k/n$, $0 \le k \le n$),

$$\mathrm{UI}(n, \rho) = \max_{1 \le i < j \le n} \frac{\big|S(j) - S(i) - S(n)(t_j - t_i)\big|}{\varrho(t_j - t_i)}$$

and

$$\mathrm{DI}(n, \rho) = \max_{1 < 2^j \le n} \frac{1}{\rho(2^{-j})} \max_{r \in \mathrm{D}_j} \Big| S(nr) - \frac{1}{2} S(nr^+) - \frac{1}{2} S(nr^-) \Big|.$$

To obtain limiting distribution for these statistics we shall work with a stronger null hypothesis, namely

(H_0'): *$X_1, \dots, X_n$ are independent identically distributed random variables with mean denoted μ_0.*

Theorem 3. *Under (H_0'), assume that $\rho \in \mathcal{R}$ and for every $A > 0$,*

$$\lim_{t \to \infty} t \, \mathrm{P}(|X_1| > A\theta(t)) = 0. \tag{5}$$

Then

$$\sigma^{-1} n^{-1/2} \mathrm{UI}(n, \rho) \xrightarrow[n \to \infty]{\mathcal{D}} \mathrm{UI}(\rho) \tag{6}$$

and

$$\sigma^{-1} n^{-1/2} \mathrm{DI}(n, \rho) \xrightarrow[n \to \infty]{\mathcal{D}} \mathrm{DI}(\rho), \tag{7}$$

where $\sigma^2 = \mathrm{var}(X_1)$ and $\mathrm{UI}(\rho)$, $\mathrm{DI}(\rho)$ are defined by (3) and (4) respectively.

When $\alpha < 1/2$, it is enough to take $A = 1$ in (5). In the special case $\rho(h) = \rho(h, \alpha, 0) = h^\alpha$, (5) may be recast as

$$\mathrm{P}(|X_1| > t) = o(t^{-p}), \quad \text{with} \quad p := 1/(1/2 - \alpha).$$

In the other special case $\rho(h) = \rho(h, 1/2, \beta) = h^{1/2} \log^\beta(c/h)$, where $\beta > 1/2$, it is not possible to drop the constant A in Condition (5) which is easily seen to be equivalent to

$$\mathrm{E} \exp\{\lambda |X_1|^{1/\beta}\} < \infty \ \text{ for each } \ \lambda > 0. \tag{8}$$

Let us note that either of (6) or (7) yields stochastic boundedness of the random variable $\max_{1\le k\le n} |X_k|/\theta(n)$ which on its turn gives (8).

Remark 2. *In the case where the variance σ^2 is unknown the results remain valid if σ^2 is substituted by its standard estimator $\widehat{\sigma}^2$.*

Remark 3. *Under (H_0), the statistics $n^{-1/2}\mathrm{UI}(n, \rho)$ keeps the same value when each X_i is substituted by $X_i' := X_i - \mathrm{E}\, X_i$. This property is only asymptotically true for $n^{-1/2}\mathrm{DI}(n, \rho)$, see the proof of Theorem 3. For practical use of $\mathrm{DI}(n, \rho)$, it is preferable to replace X_i by X_i' if μ_0 is known or else by $X_i - \overline{X}$ where $\overline{X} = n^{-1}(X_1 + \cdots + X_n)$. This will avoid a bias term which may be of the order of $|\mathrm{E}\, X_1| \ln^{-\beta} n$ in the worst cases.*

To see a consistency of tests to reject null hypothesis versus epidemic alternative (H_A) for large values of $n^{-1/2}\mathrm{UI}(n, \rho)$, we naturally assume that the numbers of observations k^*, $m^* - k^*$, $n - m^*$ before, during and after the epidemic go to infinity with n. Write $l^* := m^* - k^*$ for the length of the epidemic.

Theorem 4. *Let $\rho \in \mathcal{R}$. Assume under (H_A) that the X_i's are independent and $\sigma_0^2 := \sup_{k\ge 1} \mathrm{var}(X_k)$ is finite. If*

$$\lim_{n\to\infty} n^{1/2} \frac{h_n}{\rho(h_n)} |\mu_1 - \mu_0| = \infty, \quad \textit{where} \quad h_n := \frac{l^*}{n}\Big(1 - \frac{l^*}{n}\Big), \tag{9}$$

then

$$n^{-1/2}\mathrm{UI}(n, \rho) \ \xrightarrow[n\to\infty]{\mathrm{P}} \ \infty, \tag{10}$$

$$n^{-1/2}\mathrm{DI}(n, \rho) \ \xrightarrow[n\to\infty]{\mathrm{P}} \ \infty.$$

In our setting μ_0 and μ_1 are constants, but the proof of Theorem 9 remains valid when μ_1 and μ_0 are allowed to depend on n. This explains the presence of the factor $|\mu_1 - \mu_0|$ in (9). This dependence on n of μ_0 and μ_1 is discarded in the following exemples.

When $\rho(h) = h^\alpha$, (9) is equivalent to

$$\lim_{n\to\infty} \frac{h_n}{n^{1/(2-2\alpha)}} = \infty. \tag{11}$$

In this case one can detect short epidemics such that $n^{(1-2\alpha)/(2-2\alpha)} = o(l^*)$ as well as long epidemics such that $n^{(1-2\alpha)/(2-2\alpha)} = o(n - l^*)$.

When $\rho(h) = h^{1/2}\log^\beta(c/h)$ with $\beta > 1/2$, (9) is satisfied provided that $h_n = n^{-1}\log^\gamma n$, with $\gamma > 2\beta$. This leads to detection of short epidemics such that $\log^\gamma n = o(l^*)$ as well as of long ones verifying $\log^\gamma n = o(n - l^*)$.

2.3 Statistics SUI and SDI

To introduce the adaptive selfnormalized statistics, we shall restrict the null hypothesis assuming that $\mu_0 = 0$. Practically the reduction to this case by centering requires the knowledge of μ_0. Although this seems a quite reasonnable assumption, it is fair to point out that this knowledge was not required for the UI and DI statistics.

For any index set $A \subset \{1, \dots, n\}$, define

$$V^2(A) := \sum_{i\in A} X_i^2.$$

Then $V_k^2 = V^2(\{1, \dots, k\})$, $V_0^2 = 0$. To simplify notation we write v_k for the random points of $[0, 1]$

$$v_k := \frac{V_k^2}{V_n^2}, \quad k = 0, 1, \dots, n.$$

For $\rho \in \mathcal{R}$ define

$$\mathrm{SUI}(n, \rho) = \max_{1\le i<j\le n} \frac{\left|S(j) - S(i) - S(n)(v_j - v_i)\right|}{\varrho(v_j - v_i)}.$$

Introduce for any $t \in [0, 1]$,

$$\tau_n(t) := \max\{i \le n;\ V_i^2 \le tV_n^2\}$$

and denoting by $\log_2$ the logarithm with basis 2 ($\log_2(2^t) = t = \log(\mathrm{e}^t)$),

$$J_n := \log_2(V_n^2/X_{n:n}^2) \quad \text{where} \quad X_{n:n} := \max_{1\le k\le n} |X_k|.$$

Now define

$$\mathrm{SDI}(n,\rho) := \max_{1\le j\le J_n} \frac{1}{\rho(2^{-j})} \max_{r\in \mathrm{D}_j} \left| S\big(\tau_n(r)\big) - \frac{1}{2} S\big(\tau_n(r^+)\big) - \frac{1}{2} S\big(\tau_n(r^-)\big) \right|.$$

Theorem 5. *Assume that under (H_0'), $\mu_0 = 0$ and the random variables $X_1, \dots, X_n$ are either symmetric and belong to the domain of attraction of normal law or $\mathrm{E}\,|X_1|^{2+\varepsilon} < \infty$ for some $\varepsilon > 0$. Then for every $\rho \in \mathcal{R}$,*

$$V_n^{-1}\mathrm{SUI}(n,\rho) \xrightarrow[n\to\infty]{\mathcal{D}} \mathrm{UI}(\rho) \tag{12}$$

and

$$V_n^{-1}\mathrm{SDI}(n,\rho) \xrightarrow[n\to\infty]{\mathcal{D}} \mathrm{DI}(\rho). \tag{13}$$

Theorem 6. *Let $\rho \in \mathcal{R}$. Under (H_A) with $\mu_0 = 0$, assume that the X_i's satisfy*

$$S(\mathbb{I}_n) = l^*\mu_1 + O_{\mathrm{P}}\big(l^{*1/2}\big), \quad S(\mathbb{I}_n^c) = O_{\mathrm{P}}\big((n-l^*)^{1/2}\big) \tag{14}$$

and

$$\frac{V^2(\mathbb{I}_n)}{l^*} \xrightarrow[n\to\infty]{\mathrm{P}} b_1, \quad \frac{V^2(\mathbb{I}_n^c)}{n-l^*} \xrightarrow[n\to\infty]{\mathrm{P}} b_0, \tag{15}$$

for some finite constants b_0 and b_1. Assume that l^*/n converges to a limit $c \in [0,1]$ and when $c = 0$ or 1, assume moreover that

$$\lim_{n\to\infty} n^{1/2} \frac{h_n}{\rho(h_n)} = \infty, \quad \textit{where} \quad h_n := \frac{l^*}{n}\Big(1 - \frac{l^*}{n}\Big). \tag{16}$$

Then

$$V_n^{-1}\mathrm{SUI}(n,\rho) \xrightarrow[n\to\infty]{\mathrm{P}} \infty. \tag{17}$$

Note that in Theorem 6, no assumption is made about the dependence structure of the X_i's. It is easy to verify the general hypotheses (14) and (15) in various situations of weak dependence like mixing or association. Under independence of the X_i's (without assuming identical distributions inside each block $\mathbb{I}_n$ and $\mathbb{I}_n^c$), it is enough to have $\sup_{k\ge1} \mathrm{E}\,|X_k|^{2+\varepsilon} < \infty$ and $\mathrm{E}\,V^2(\mathbb{I}_n)/l^* \to b_1$, $\mathrm{E}\,V^2(\mathbb{I}_n^c)/(n-l^*) \to b_0$. This follows easily from Lindebergh's condition for the central limit theorem (giving (14)) and of Theorem 5 p.261 in Petrov [9] giving the weak law of large numbers required for the X_i's.

Theorem 7. *Let $\rho \in \mathcal{R}$ and set $X_i' := X_i - \mathrm{E}(X_i)$. Under (H_A) with $\mu_0 = 0$, assume that the random variables X_i' are independent identically distributed and that $\mathrm{E}\,|X_1'|^{2+\varepsilon}$ is finite for some $\varepsilon > 0$. Then under (16),*

$$V_n^{-1}\mathrm{SDI}(n,\rho) \xrightarrow[n\to\infty]{\mathrm{P}} \infty. \tag{18}$$

Remark 4. *Theorems 5 and 7 are proved using the* Hölderian *FCLT for self-normalized partial sums processes (see [10]). As far as we know, the removing of the assumption of finite $2+\varepsilon$ moment in the non symmetric case for the Hölderian FCLT remains an open problem. Of course when there is no weight ($\rho(h) = 1$), we can apply the weak invariance principle in $C[0,1]$ under self-normalization. In this special case, Theorems 5 and 7 are valid with X_1 (resp. X_1') in the domain of attraction of the normal distribution.*

2.4 Trimmed statistics

Let $\rho \in \mathcal{R}$ and let the sequence $m_n \to \infty$ be such that $m_n \le \log_2 n$. Define

$$\mathrm{DI}(n,\rho,m_n) := \max_{1\le j\le m_n} \frac{1}{\rho(2^{-j})} \max_{r\in \mathrm{D}_j} \Big| S(nr) - \frac{1}{2}\big(S(nr^+) + S(nr^-)\big)\Big|.$$

Theorem 8. *If under (H_0'), $\mathrm{E}\,|X_1|^{2+\tau} < \infty$ for some $0 < \tau \le 1$ and*

$$\lim_{n\to\infty} m_n^{1+\tau} n^{-\tau/2} \rho^{-2-\tau}(2^{-m_n}) = 0 \tag{19}$$

then

$$\sigma^{-1} n^{-1/2}\mathrm{DI}(n,\rho,m_n) \xrightarrow[n\to\infty]{\mathcal{D}} \mathrm{DI}(\rho).$$

Let (ϵ_n) and (ϵ_n') be two sequences of positive real numbers, both converging to 0. Define with $t_k = k/n$,

$$\mathrm{UI}(n,\rho,\epsilon_n,\epsilon_n') := \max_{n\epsilon_n \le i < j \le n(1-\epsilon_n')} \frac{\big|S(j) - S(i) - S(n)(t_j - t_i)\big|}{\varrho(t_j - t_i)}.$$

Theorem 9. *If under (H_0'), $\mathrm{E}\,|X_1|^{2+\tau} < \infty$ for some $0 < \tau \le 1$ and*

$$\lim_{n\to\infty} n^{-\tau/2} \rho^{-2-\tau}(\epsilon_n \epsilon_n') \log^{1+\tau} n = 0,$$

then

$$\sigma^{-1} n^{-1/2}\mathrm{UI}(n,\rho,\epsilon_n,\epsilon_n') \xrightarrow[n\to\infty]{\mathcal{D}} \mathrm{UI}(\rho).$$

2.5 The distribution of $\mathrm{DI}(\rho)$

The distribution function of $\mathrm{DI}(\rho)$ may be conveniently expressed in terms of the error function:

$$\operatorname{erf} x = \frac{2}{\pi^{1/2}} \int_0^x \exp(-s^2)\,\mathrm{d}s.$$

The following asymptotic expansion will be useful

$$\operatorname{erf} x = 1 - \frac{1}{x\pi^{1/2}} \exp(-x^2)\bigl(1 + O(x^{-2})\bigr), \quad x \to \infty.$$

Theorem 10. *Let* $c = \limsup_{j\to\infty} j^{1/2}/\theta(2^j)$.

i) If $c = \infty$ *then* $\mathrm{DI}(\rho) = \infty$ *almost surely.*

ii) If $0 \le c < \infty$, *then* $\mathrm{DI}(\rho)$ *is almost surely finite and its distribution functions is given by*

$$\mathrm{P}\bigl(\mathrm{DI}(\rho) \le x\bigr) = \prod_{j=1}^{\infty} \bigl\{\operatorname{erf}\bigl(\theta(2^j)x\bigr)\bigr\}^{2^{j-1}}, \quad x > 0.$$

The distribution function of $\mathrm{DI}(\rho)$ *is continuous with support* $[c\,(\log 2)^{1/2}, \infty)$.

Remark 5. *The exact distribution of* $\mathrm{UI}(\rho) = \|B\|_\rho$ *is not known. To obtain critical values for* $\mathrm{UI}(n,\rho)$, *one can use a general result on large deviations for Gaussian norms (see e.g. [7]) which reads*

$$\mathrm{P}\bigl(\mathrm{UI}(\rho) \ge x\bigr) \le 4 \exp\left(\frac{-x^2}{8\mathrm{E}\,\|B\|_\rho^2}\right), \quad x > 0. \tag{20}$$

Another approach is to use the equivalence of the norms $\|B\|_\rho$ *and* $\|B\|_\rho^{\mathrm{seq}}$ *which provides constants* c_ρ *and* C_ρ *such that*

$$c_\rho \mathrm{DI}(\rho) \le \mathrm{UI}(\rho) \le C_\rho \mathrm{DI}(\rho). \tag{21}$$

Both methods give only rough estimates of the critical value for $\mathrm{UI}(n,\rho)$.

3 Proofs

3.1 Tools

The proofs of Theorems 3 and 5 are based on invariance principles in Hölder spaces, for the partial sums processes $\xi_n = (\xi_n(t), t \in [0,1])$ and $\zeta_n = (\zeta_n(t), t \in [0,1])$

defined by

$$\xi_n \text{ is the polygonal line with vertices } \quad \Big(\frac{k}{n}, S(k)\Big), \quad 0 \le k \le n; \tag{22}$$

$$\zeta_n \text{ is the polygonal line with vertices } \quad \Big(\frac{V_k^2}{V_n^2}, S(k)\Big), \quad 0 \le k \le n. \tag{23}$$

The relevant invariance principles are proved in [11] and [10] respectively and may be stated as follows.

Theorem 11. *Let ρ be in $\mathcal{R}$. Then under (H_0') with $\mu_0 = 0$, the convergence*

$$\sigma^{-1} n^{-1/2} \xi_n \xrightarrow{\mathcal{D}} W$$

takes place in $\mathcal{H}_\rho^o$ if and only if the condition (5) holds.

Theorem 12. *Under the conditions of Theorem 5*

$$V_n^{-1} \zeta_n \xrightarrow{\mathcal{D}} W$$

in $\mathcal{H}_\rho^o$ for any $\rho \in \mathcal{R}$.

The two next lemmas are useful to simplify and unify the proofs of Theorems 3 and 5.

Lemma 13. *Let $(\eta_n)_{n \ge 1}$ be a tight sequence of random elements in the separable Banach space B and g_n, g be continuous functionals $B \to \mathbb{R}$. Assume that g_n converges pointwise to g on B and that $(g_n)_{n \ge 1}$ is equicontinuous. Then*

$$g_n(\eta_n) = g(\eta_n) + o_{\mathrm{P}}(1).$$

Proof. By the tightness assumption, there is for every $\varepsilon > 0$, a compact subset K in B such that for every $n \ge 1$, $\mathrm{P}(\eta_n \notin K) < \varepsilon$. Now by a classical corollary of Ascoli's theorem, the equicontinuity and pointwise convergence of (g_n) give its uniform convergence to g on the compact K. Then for every $\delta > 0$, there is some $n_0 = n_0(\delta, K)$, such that

$$\sup_{x \in K} |g_n(x) - g(x)| < \delta, \quad n \ge n_0.$$

Therefore we have for $n \ge n_0$,

$$\mathrm{P}\big(|g_n(\eta_n) - g(\eta_n)| \ge \delta\big) \le \mathrm{P}(\eta_n \notin K) < \varepsilon,$$

which is the expected conclusion. □

Lemma 14. *Let $(B, \| \|)$ be a vector normed space and $q : B \to \mathbb{R}$ such that*

a) q is subadditive: $q(x+y) \le q(x) + q(y)$, $x, y \in B$;

b) q is symmetric: $q(-x) = q(x)$, $x \in B$;

c) For some constant C, $q(x) \le C \|x\|$, $x \in B$.

Then q satisfies the Lipschitz condition

$$|q(x+y) - q(x)| \le C \|y\|, \quad x, y \in B. \tag{24}$$

If $\mathcal{F}$ is any set of functionals q fulfilling a), b), c) with the same constant C, then a), b), c) are inherited by $g(x) := \sup\{q(x);\ q \in \mathcal{F}\}$ which therefore satisfies (24).

The proof is elementary and will be omitted.

3.2 Proof of Theorem 3

Convergence of $\mathrm{UI}(n,\rho)$. Consider the functionals g_n, g, defined on $\mathcal{H}^o_\rho$ by

$$g_n(x) := \max_{1 \le i < j \le n} I(x, i/n, j/n), \quad g(x) := \sup_{0<s<t<1} I(x, s, t), \tag{25}$$

where

$$I(x,s,t) := \frac{|x(t) - x(s) - (t-s)x(1)|}{\varrho(t-s)}, \quad 0 < t-s < 1 \tag{26}$$

and $\varrho(h) = \rho(h(1-h))$, $h \in [0,1]$. Observe that

$$\mathrm{UI}(n,\rho) = g_n(\xi_n), \quad \mathrm{UI}(\rho) = g(W). \tag{27}$$

Clearly the functional $q = I(\,.\,, s, t)$ satisfies Conditions a) and b) of Lemma 14. Let us check Condition c). From the Definition 1, it is clear that for $\rho \in \mathcal{R}$, the ratio $\rho(h)/\rho(h/2)$ is continuous on $(0,1]$ and has the finite limit 2^α at zero. Hence it is bounded on $[0,1]$ by a constant $a = a(\rho) < \infty$. Similarly the ratio $\rho^\star(h) := h/\rho(h)$ vanishes at 0 and and is bounded on $[0,1]$ by some $b = b(\rho) < \infty$.

If $0 < t-s \le 1/2$, then $\varrho(t-s) \ge \rho\big((t-s)/2\big) \ge a^{-1}\rho(t-s)$ and

$$\begin{aligned} I(x,s,t) &\le a\frac{|x(t)-x(s)|}{\rho(t-s)} + a\frac{t-s}{\rho(t-s)}|x(1)| \qquad (28)\\ &\le a(1+b)\|x\|_\rho. \qquad (29) \end{aligned}$$

If $1/2 < t - s < 1$, then $\varrho(t-s) \geq \rho\big((1-t+s)/2\big) \geq a^{-1}\rho(1-t+s)$ and $\rho(1-t+s)$ is bigger than both $\rho(s)$ and $\rho(1-t)$. Assuming that $x(0) = 0$, we write $x(t) - x(s) - (t-s)x(1) = x(t) - x(1) + x(0) - x(s) + (1-t+s)x(1)$ and obtain

$$\begin{aligned} I(x,s,t) &\leq a\frac{|x(1)-x(t)|}{\rho(1-t)} + a\frac{|x(s)-x(0)|}{\rho(s)} + a\frac{1-t+s}{\rho(1-t+s)}|x(1)| && (30) \\ &\leq a(2+b)\|x\|_\rho. && (31) \end{aligned}$$

Introduce the closed subspace $B := \{x \in \mathcal{H}^o_\rho;\ x(0) = 0\}$. From (29) and (31) we see that the functionals $q = I(\,.\,,s,t)$ satisfy on B the Condition c) of Lemma 14 with the same constant $C = C(\rho) = a(2+b)$. It follows by Lemma 14 that g_n as well as g are Lipschitz on B with the same constant C. In particular the sequence $(g_n)_{n\geq 2}$ is equicontinuous on B.

Now we shall apply Lemma 13 with $\eta_n := \sigma^{-1}n^{-1/2}\xi_n$ with ξ_n defined by (22). By Theorem 11, $(\eta_n)_{n\geq 2}$ is tight on B. To check the pointwise convergence on B of g_n to g, it is enough to show that for each $x \in B$, the function $(s,t) \mapsto I(x,s,t)$ can be extended by continuity to the compact set $T = \{(s,t) \in [0,1]^2;\ 0 \leq s \leq t \leq 1\}$. From (28) we get $0 \leq I(x,s,t) \leq a\omega_\rho(x, t-s) + a|x(1)|\rho^\star(t-s)$, which allows the continuous extension along the diagonal puting $I(x,s,s) := 0$. From (30) we get $0 \leq I(x,s,t) \leq 2a\omega_\rho(x, 1+t-s) + a|x(1)|\rho^\star(1+t-s)$ which allows the continuous extension at the point $(0,1)$ puting $I(x,0,1) := 0$.

The pointwise convergence of (g_n) being now established, Lemma 13 gives

$$g_n(\sigma^{-1}n^{-1/2}\xi_n) = g(\sigma^{-1}n^{-1/2}\xi_n) + o_{\mathrm{P}}(1) \qquad (32)$$

and the convergence of $\sigma^{-1}n^{-1/2}\mathrm{UI}(n,\rho)$ to $\mathrm{UI}(\rho)$ follows from Theorem 11, (27) and (32). □

Convergence of $\mathrm{DI}(n,\rho)$. Let $\mathrm{DI}'(n,\rho)$ be defined like $\mathrm{DI}(n,\rho)$, replacing X_i by $X'_i = X_i - \mathrm{E}\,X_i$. As

$$\sum_{nr^- < i \leq nr} X_i - \sum_{nr < i \leq nr^+} X_i = \sum_{nr^- < i \leq nr} X'_i - \sum_{nr < i \leq nr^+} X'_i + c(r)\mathrm{E}\,X_1,$$

with $|c(r)| \leq 1$, we clearly have $n^{-1/2}\big(\mathrm{DI}(n,\rho) - \mathrm{DI}'(n,\rho)\big) = o_{\mathrm{P}}(1)$, so we may and do assume without loss of generality that $\mu_0 = 0$ in the proof. First we observe that

$$\mathrm{DI}(n,\rho) = \max_{1<2^j\leq n} \frac{1}{\rho(2^{-j})} \max_{r\in \mathrm{D}_j} |\lambda_r(S(n\,.))|, \qquad (33)$$

where the λ_r are defined by (2) and $S(n\,.)$ is the discontinuous process $(S(nt), 0 \le t \le 1)$. This process may also be written as

$$S(nt) = \xi_n(t) - (nt - [nt])X_{[nt]+1}, \quad t \in [0,1],$$

from which we see that

$$|\lambda_r(S(n\,.)) - \lambda_r(\xi_n)| \le 2 \max_{1\le i\le n} |X_i|.$$

Plugging this estimate into (33) leads to the representation

$$\sigma^{-1}n^{-1/2}\mathrm{DI}(n,\rho) = g_n(\sigma^{-1}n^{-1/2}\xi_n) + Z_n \tag{34}$$

where

$$g_n(x) := \max_{1<2^j\le n} \max_{r\in \mathrm{D}_j} \frac{|\lambda_r(x)|}{\rho(2^{-j})}, \quad x \in \mathcal{H}_\rho^o$$

and the random variable Z_n satisfies

$$|Z_n| \le \frac{2}{\sigma n^{1/2}\rho(1/n)} \max_{1\le i\le n} |X_i| = \frac{2}{\sigma\theta(n)} \max_{1\le i\le n} |X_i|. \tag{35}$$

Hypothesis (5) in Theorem 3 gives immediately the convergence in probability to zero of $\max_{1\le i\le n} |X_i|/\theta(n)$. The functionals $q_r(x) := |\lambda_r(x)|/\rho(r - r^-)$ satisfy obviously the hypotheses of Lemma 14 with the same constant $C = 1$. The same holds for $g_n = \max_{1<2^j\le n} \max_{r\in\mathrm{D}_j} q_r$ and for

$$g(x) := \sup_{j\ge 1} \max_{r\in \mathrm{D}_j} q_r(x).$$

Accounting (34) and (35), Lemma 13 leads to the representation

$$\sigma^{-1}n^{-1/2}\mathrm{DI}(n,\rho) = g(\sigma^{-1}n^{-1/2}\xi_n) + o_{\mathrm{P}}(1),$$

from which the conclusion follows by Theorem 11 and continuous maping. □

3.3 Proof of Theorem 4

Divergence of $\mathrm{UI}(n,\rho)$. Define the random variables

$$X_i' := \begin{cases} X_i - \mu_0 & \text{if } i \in \mathbb{I}_n^c, \\ X_i - \mu_1 & \text{if } i \in \mathbb{I}_n. \end{cases}$$

In order to find a lower bound for $\mathrm{UI}(n,\rho)$, let us consider the contribution to its numerator of the increment corresponding to the full length of epidemics. It may be writen as

$$\begin{aligned} S(m^*) - S(k^*) - S(n)(t_{m^*} - t_{k^*}) &= \Big(1 - \frac{l^*}{n}\Big) S(\mathbb{I}_n) - \frac{l^*}{n} S(\mathbb{I}_n^c) \\ &= l^* \Big(1 - \frac{l^*}{n}\Big)(\mu_1 - \mu_0) + R_n, \qquad (36) \end{aligned}$$

where

$$R_n := -\frac{l^*}{n} \sum_{i \in \mathbb{I}_n^c} X_i' + \Big(1 - \frac{l^*}{n}\Big) \sum_{i \in \mathbb{I}_n} X_i'.$$

It is easy to see that $n^{-1/2} R_n = O_{\mathrm{P}}(h_n^{1/2})$. Indeed

$$\mathrm{var}(n^{-1/2} R_n) \le \frac{1}{n} \Big(\frac{l^*}{n}\Big)^2 (n - l^*) \sigma_0^2 + \frac{1}{n} \Big(1 - \frac{l^*}{n}\Big)^2 l^* \sigma_0^2 = \sigma_0^2 h_n.$$

This estimate together with (36) leads to the lower bound

$$n^{-1/2} \mathrm{UI}(n,\rho) \ge n^{1/2} \frac{h_n}{\rho(h_n)} |\mu_1 - \mu_0| - O_{\mathrm{P}} \Big(\frac{h_n{}^{1/2}}{\rho(h_n)} \Big). \qquad (37)$$

From Definition 1 iii), it is easily seen that $\lim_{h_n \to 0} h_n^{1/2}/\rho(h_n) = 0$, so (10) follows from Condition (9) via (37). □

Divergence of $\mathrm{DI}(n,\rho)$. Let us estimate first $\lambda_r(S(n\,.))$ under the special configuration $r^- \le t_{k^*} < t_{m^*} \le r$ (then necessarily, $l^*/n \le 1/2$). For notational simplification, define for $0 \le s < t \le 1$ and $\mu \in \mathbb{R}$,

$$S_n(s,t) := \sum_{ns < k \le nt} X_k \quad \text{and} \quad S_n'(s,t,\mu) := \sum_{ns < k \le nt} (X_k' + \mu).$$

Then we have

$$\begin{aligned} 2\lambda_r(S(n\,.)) &= S_n(r^-, r) - S_n(r, r^+) \\ &= S_n'(r^-, t_{k^*}, \mu_0) + S_n'(t_{k^*}, t_{m^*}, \mu_1) + S_n'(t_{m^*}, r, \mu_0) \\ &\quad - S_n'(r, r^+, \mu_0) \\ &= (\mu_0 - \mu_1) l^* + O(1) + \sum_{nr^- < k \le nr^+} \varepsilon_k X_k', \end{aligned}$$

where $O(1)$ and the ε_k's are deterministic and $\varepsilon_k = \pm 1$. If the level j of r ($r \in \mathrm{D}_j$) is such that $2^{-j-1} < l^*/n \le 2^{-j}$, this gives for $\mathrm{DI}(n,\rho)$ the same lower bound as the right hand side of (37). Clearly the same result holds true when $[t_{k^*}, t_{m^*}]$ is included in $[r, r^+]$. Denoting by θ the middle of $[t_{k^*}, t_{m^*}]$, we obtain the same lower bound (up to multiplicative constants) under the configurations where $[\theta, t_{m^*}] \subset [r^-, r]$ or $[t_{k^*}, \theta] \subset [r, r^+]$.

Assume still that $l^*/n \le 1/2$ and fix the level j by $2^{-j-1} < l^*/n \le 2^{-j}$. There is a unique $r_0 \in \mathrm{D}_j$ such that $r_0^- \le \theta < r_0^+$. Then the divergence of $n^{-1/2}\mathrm{DI}(n, \rho)$ follows from Hypothesis (9) by applying the above remarks in each of the following cases.

a) $r_0^- \le \theta < r_0^- + 2^{-j-1}$. Then $\theta + l^*/(2n) \le r_0^- + 2^{-j-1} + 2^{-j-1} = r_0$, so $[\theta, t_{m^*}] \subset [r_0^-, r_0]$.

b) $r_0 + 2^{-j-1} \le \theta < r_0^+$. Then $\theta - l^*/(2n) \ge r_0$, so $[t_{k^*}, \theta] \subset [r_0, r_0^+]$.

c) $r_0 - 2^{-j-1} \le \theta < r_0 + 2^{-j-1}$. Then $r_0^- \le t_{k^*} < t_{m^*} \le r_0^+$. Only one of both dyadics r_0^- and r_0^+ has the level $j-1$. If $r_0^- \in \mathrm{D}_{j-1}$, writing $r_1 := r_0^-$ we have $r_1^+ = r_0^+$ and $[t_{k^*}, t_{m^*}] \subset [r_1, r_1^+]$. Else $r_0^+ \in \mathrm{D}_{j-1}$ and with $r_1 := r_0^+$, $r_1^- = r_0^-$ so that $[t_{k^*}, t_{m^*}] \subset [r_1^-, r_1]$.

It remains to consider the case where $l^*/n > 1/2$. Assume first that $t_{k^*} \ge 1 - t_{m^*}$, so that $t_{k^*} \ge (1 - l^*/n)/2$. Then there is a unique j such that $0 < 2^{-j-1} < t_{k^*} \le 2^{-j} \le 1/2 < t_{m^*}$. Choosing $r_0 := 2^{-j} \in \mathrm{D}_j$, we easily obtain

$$2\lambda_{r_0}(S(n\,.)) = (\mu_0 - \mu_1)(nt_{k^*}) + O(1) + \sum_{nr^- < k \le nr^+} \varepsilon_k X'_k.$$

Now the divergence of $n^{-1/2}\mathrm{DI}(n, \rho)$ follows from (9) in view of the lower bound

$$\left|\lambda_{r_0}(S(n\,.))\right| \ge |\mu_0 - \mu_1| \frac{n - l^*}{4} - O_{\mathrm{P}}\big((n - l^*)^{1/2}\big).$$

When $t_{k^*} < 1 - t_{m^*}$, fixing j by $1 - 2^{-j} \le t_{m^*} < 1 - 2^{-j-1}$ and choosing $r_0 := 1 - 2^{-j} \in \mathrm{D}_j$, leads clearly to the same lower bound. The proof is complete. □

3.4 Proof of Theorem 5

Convergence of $\mathrm{SUI}(n, \rho)$. We use the straightforward representation

$$\mathrm{SUI}(n, \rho) = \max_{1 \le i < j \le n} I\left(\zeta_n, \frac{V_i^2}{V_k^2}, \frac{V_j^2}{V_k^2}\right), \tag{38}$$

where ζ_n and I are defined by (23) and (26) respectively. By Lemma 9 in [10], if X_1 is in the domain of attraction of the normal distribution then

$$\max_{0\le k\le n}\left|\frac{V_k^2}{V_n^2}-\frac{k}{n}\right| \xrightarrow[n\to\infty]{\mathrm{P}} 0. \tag{39}$$

Theorem 12 and the same arguments as in the proof of Theorem 3 give

$$g_n(V_n^{-1}\zeta_n)=\max_{1\le i<j\le n} I\Big(\zeta_n,\frac{i}{n},\frac{j}{n}\Big)=g\big(V_n^{-1}\zeta_n\big)+o_{\mathrm{P}}(1), \tag{40}$$

with g defined by (25). By continuous mapping, the right hand side of (40) converges in distribution to $\mathrm{UI}(\rho)$. Hence the convergence of $\mathrm{SUI}(n,\rho)$ to $\mathrm{UI}(\rho)$ will follow from the estimate

$$\mathrm{SUI}(n,\rho)-g_n(V_n^{-1}\zeta_n)=o_{\mathrm{P}}(1).$$

Due to the tightness in $\mathcal{H}_\rho^o$ of $(V_n^{-1}\zeta_n)_{n\ge1}$, this follows in turn from (38), (39) and the purely analytical next lemma. □

Lemma 15. *For any $x\in\mathcal{H}_\rho^o$ such that $x(0)=0$, denote by I_x the continuous function*

$$I_x: T\to\mathbb{R},\quad (s,t)\mapsto I(x,s,t),$$

where $T:=\{(s,t)\in[0,1]^2;\ 0\le s\le t\le 1\}$ and I is defined by (26) and its continuous extension to the boundary of T puting $I(x,s,s):=0$ and $I(x,0,1):=0$. Let K be a compact subset in $\mathcal{H}_\rho^o$ of functions x such that $x(0)=0$. Then the set of functions $\mathcal{E}_K:=\{I_x;\ x\in K\}$ is equicontinuous.

Proof. Put for notational simplification $y(t):=x(t)-tx(1)$ and define

$$\dot{y}_\varrho(s,t):=\frac{y(t)-y(s)}{\varrho(|t-s|)},\quad 0<|t-s|<1,$$

with $\dot{y}_\varrho(s,t):=0$ when $|t-s|=0$ or 1. This reduces the problem to the proof of the equicontinuity of $\mathcal{E}_{K'}:=\{\dot{y}_\varrho: T\to\mathbb{R};\ y\in K'\}$ where K' is a compact subset in $\mathcal{H}_\rho^o$ of functions vanishing at 0 and 1.

It is convenient to estimate the increments of y in terms of the function $\varrho(h)=\rho\big(h(1-h)\big)$. First we always have for $0\le s\le s+h\le 1$,

$$|y(s+h)-y(s)|\le\rho(h)\omega_\rho(y,h). \tag{41}$$

Because $y(0) = y(1) = 0$ and ρ and $\omega_\rho(y, .)$ are non decreasing, we also have for $0 \le s \le s + h \le 1$,

$$\begin{aligned} |y(s+h) - y(s)| &\le |y(s+h) - y(1)| + |y(s) - y(0)| \\ &\le \rho(1-s-h)\omega_\rho(y, 1-s-h) + \rho(s)\omega_\rho(y, s) \\ &\le 2\rho(1-h)\omega_\rho(y, 1-h). \end{aligned} \tag{42}$$

Recalling the constant $a = a(\rho) = \sup\{\rho(h)/\rho(h/2);\, 0 < h \le 1\} < \infty$, we see that

$$\rho(h) \wedge \rho(1-h) \le a\rho\big(h(1-h)\big) = a\varrho(h). \tag{43}$$

By (41), (42) and (43) there is a constant $c = c(\rho)$ such that

$$|y(s+h) - y(s)| \le c\varrho(h)\omega_\rho\big(y, h \wedge (1-h)\big), \quad 0 \le s \le s + h \le 1.$$

Writing for simplicity

$$\tilde{\omega}_\rho(y, h) := c\omega_\rho\big(y, h \wedge (1-h)\big),$$

the above estimate leads to

$$y(t) = y(s) + \dot{y}_\varrho(s,t)\varrho(t-s), \quad (s,t) \in [0,1]^2, \tag{44}$$

where

$$|\dot{y}_\varrho(s,t)| \le \tilde{\omega}_\rho(y, |t-s|). \tag{45}$$

Using the expansion (44), we get for any (s,t) and (s',t') in the interior of T (writing $h' := t' - s'$ and $h = t - s$)

$$\begin{aligned} \dot{y}_\varrho(s',t') - \dot{y}_\varrho(s,t) &= \frac{\big(y(t) - y(s)\big)\big(\varrho(h) - \varrho(h')\big)}{\varrho(h)\varrho(h')} \\ &\quad + \frac{\dot{y}_\varrho(t,t')\varrho(|t'-t|) - \dot{y}_\varrho(s,s')\varrho(|s'-s|)}{\varrho(h')} \\ &= \dot{y}_\varrho(s,t)\frac{\varrho(h) - \varrho(h')}{\varrho(h')} \\ &\quad + \frac{\dot{y}_\varrho(t,t')\varrho(|t'-t|) - \dot{y}_\varrho(s,s')\varrho(|s'-s|)}{\varrho(h')}. \end{aligned}$$

Due to (45), this gives

$$|\dot{y}_\varrho(s',t') - \dot{y}_\varrho(s,t)| \le c\|y\|_\rho \frac{|\varrho(h) - \varrho(h')| + \varrho(|t'-t|) + \varrho(|s'-s|)}{\rho(h')} \tag{46}$$

To deal with the points close to the border of T, we complete this estimate with

$$|\dot{y}_\varrho(s',t') - \dot{y}_\varrho(s,t)| \le c\big(\tilde{\omega}_\rho(y,h) + \tilde{\omega}_\rho(y,h')\big). \tag{47}$$

Now the equicontinuity of $\mathcal{E}_{K'}$ follows easily from (46), (47), the continuity of ρ, the boundedness in $\mathcal{H}^o_\rho$ of the compact K' and the uniform convergence to zero over K' of $\omega_\rho(y,\delta)$ (when δ goes to zero). This last property is again an application of Ascoli's theorem or of Dini's theorem. □

Convergence of $\mathrm{SDI}(n,\rho)$. Recall that $\tau_n(t) = \max\{i \le n;\ V_i^2 \le tV_n^2\}$, $t \in [0,1]$ and $J_n = \log_2(V_n^2/X_{n:n}^2)$ where $X_{n:n} := \max_{1\le k\le n} |X_k|$. Note also that $\mathrm{SDI}(n,\rho)$ may be recast as

$$\mathrm{SDI}(n,\rho) = \max_{1\le j\le J_n} \frac{1}{\rho(2^{-j})} \max_{r\in \mathrm{D}_j} |\lambda_r(S\circ\tau_n)|.$$

By O'Brien [8], X_1 is in the domain of attraction of the normal distribution if and only if

$$V_n^{-1} \max_{1\le k\le n} |X_k| = o_{\mathrm{P}}(1). \tag{48}$$

The polygonal random line ζ_n defined by (23) may be written as

$$\zeta_n(t) = S\big(\tau_n(t)\big) + \frac{t - V^2_{\tau_n(t)}V_n^{-2}}{X^2_{\tau_n(t)+1}V_n^{-2}} X_{\tau_n(t)+1} = S\big(\tau_n(t)\big) + \psi_n(t).$$

Noting that $|\psi_n(t)| \le X_{n:n}$, we have $|\lambda_r(\psi_n)| \le 2X_{n:n}$ for every dyadic r. This provides us the representation

$$V_n^{-1}\mathrm{SDI}(n,\rho) = \max_{1\le j\le J_n} \frac{1}{\rho(2^{-j})} \max_{r\in \mathrm{D}_j} \Big|\lambda_r\Big(\frac{\zeta_n}{V_n}\Big)\Big| + Z_n',$$

where

$$|Z_n'| \le \frac{2}{V_n\rho\big(X_{n:n}^2V_n^{-2}\big)} = \frac{2}{\theta\big(V_n^2X_{n:n}^{-2}\big)},$$

with $\theta(t) = t^{1/2}\rho(1/t)$. Recalling that for $\rho\in\mathcal{R}$, $\lim_{t\to\infty}\theta(t) = \infty$, we see from (48) that $Z_n' = o_{\mathrm{P}}(1)$. Introducing again the functionals $q_r(x) := |\lambda_r(x)|/\rho(r-r^-)$, $g_n = \max_{1\le j\le \log_2 n}\max_{r\in\mathrm{D}_j} q_r$ and $g(x) := \sup_{j\ge 1}\max_{r\in\mathrm{D}_j} q_r(x)$ which satisfy the hypotheses of Lemma 14 with the same constant $C = 1$, we have

$$V_n^{-1}\mathrm{SDI}(n,\rho) = g_{J_n}(V_n^{-1}\zeta_n) + Z_n'.$$

By (48), J_n goes to infinity in probability, so an obvious extension of Lemma 13 leads to the representation

$$V_n^{-1}\mathrm{SDI}(n,\rho) = g(V_n^{-1}\zeta_n) + o_{\mathrm{P}}(1)$$

and the convergence of $V_n^{-1}\mathrm{SDI}(n,\rho)$ to $\mathrm{DI}(\rho)$ follows from Theorem 12 by continous mapping. □

3.5 Proof of Theorem 6

Divergence of $\mathrm{SUI}(n,\rho)$*.* We shall exploit the lower bound

$$\mathrm{SUI}(n,\rho) \geq \frac{\left|S(\mathbb{I}_n) - S(n)(v_{m^*} - v_{k^*})\right|}{\varrho(v_{m^*} - v_{k^*})}. \tag{49}$$

Writing again $X_i' := X_i - \mathrm{E}\,(X_i)$, we get

$$S(\mathbb{I}_n) - S(n)(v_{m^*} - v_{k^*}) = \frac{V^2(\mathbb{I}_n^c)}{V_n^2} l^* \mu_1 + R_n, \tag{50}$$

where

$$R_n := \frac{V^2(\mathbb{I}_n^c)}{V_n^2} \sum_{i\in\mathbb{I}_n} X_i' - \frac{V^2(\mathbb{I}_n)}{V_n^2} \sum_{i\in\mathbb{I}_n^c} X_i'.$$

From (15) it follows that

$$\frac{V_n^2}{n} = cb_1 + (1-c)b_0 + o_{\mathrm{P}}(1), \tag{51}$$

so by (14) and (15) we easily obtain

$$n^{-1/2} R_n = O_{\mathrm{P}}\big(h_n^{1/2}\big). \tag{52}$$

Rewriting the principal term in (50) as

$$\frac{V^2(\mathbb{I}_n^c)}{V_n^2} l^* \mu_1 = \frac{(n-l^*)^{-1} V^2(\mathbb{I}_n^c)}{n^{-1} V_n^2} \frac{l^*}{n} (n-l^*)\mu_1$$

leads by (15) to the equivalence in probability

$$n^{-1/2} \frac{V^2(\mathbb{I}_n^c)}{V_n^2} l^* \mu_1 \overset{\mathrm{P}}{\sim} \frac{b_1\mu_1}{cb_1 + (1-c)b_0} n^{1/2} h_n. \tag{53}$$

Finally we also have from (14)

$$\frac{V^2(\mathbb{I}_n)}{V_n^2} \frac{V^2(\mathbb{I}_n^c)}{V_n^2} \overset{\mathrm{P}}{\sim} \frac{b_0 b_1}{\big(cb_1 + (1-c)b_0\big)^2} h_n,$$

from which (recalling that $\rho(h) = h^\alpha L(1/h)$ with L slowly varying) we deduce that for some constant $d > 0$,

$$\varrho(v_{m^*} - v_{k^*}) \le d\rho(h_n)\big(1 + o_{\mathrm{P}}(1)\big). \tag{54}$$

Going back to (49) with the estimates (52), (53) and (54) provides the lower bound

$$n^{-1/2}\mathrm{SUI}(n,\rho) \ge C\frac{n^{1/2}h_n}{\rho(h_n)} - o_{\mathrm{P}}(1) - O_{\mathrm{P}}\Big(\frac{h_n^{1/2}}{\rho(h_n)}\Big), \tag{55}$$

with a positive constant C depending on μ_1, b_0, b_1, c and ρ. As $\rho \in \mathcal{R}$, $h_n^{1/2} = o\big(\rho(h_n)\big)$ when h_n goes to zero, so (55) gives the divergence of $n^{-1/2}\mathrm{SUI}(n,\rho)$, accounting Hypothesis (16). Due to (51), the divergence of $V_n^{-1}\mathrm{SUI}(n,\rho)$ follows.

□

3.6 Proof of Theorem 7

Divergence of $\mathrm{SDI}(n,\rho)$. For $A \subset \{1,\dots,n\}$ and $1 \le k \le n$, define

$$S'(A) := \sum_{i\in A} X'_i, \quad S'(0) := 0, \quad S'(k) = S'(\{1,\dots,k\})$$

and

$$V'^2(A) := \sum_{i\in A} X'^2_i, \quad V'^2_k := V'^2(\{1,\dots,k\}), \quad v'_0 := 0, \quad v'_k := \frac{V'^2_k}{V'^2_n}.$$

Denote by ζ'_n the polygonal line with vertices $\big(v'_k, S'(k)\big)$, $0 \le k \le n$.

As a preliminary step, we look for some control on the ratios V_n/V'_n and $(v_{m*} - v_{k*})/(v'_{m*} - v'_{k*})$. Clearly

$$\frac{V_n^2}{V'^2_n} = 1 + \frac{l^*\mu_1^2}{V'^2_n} + \frac{2\mu_1 S'(\mathbb{I}_n)}{V'^2_n}.$$

Under the assumptions made on the sequence (X'_i), $V'^{-1}_n S'(\mathbb{I}_n) = O_{\mathrm{P}}(1)$ and $n^{-1}V'^2_n$ converges in probability to $\sigma^2 := \mathrm{E}\,(X'^2_1)$, so

$$1 \le \liminf_{n\to+\infty} \frac{V_n^2}{V'^2_n} \le \limsup_{n\to+\infty} \frac{V_n^2}{V'^2_n} \le 1 + \frac{\mu_1^2}{\sigma^2} \quad \text{in probability.} \tag{56}$$

Similarly we have

$$\begin{aligned} \frac{v_{m*} - v_{k*}}{v'_{m*} - v'_{k*}} &= \frac{V'^2_n\big(V'^2(\mathbb{I}_n) + l^*\mu_1^2 + 2\mu_1 S'(\mathbb{I}_n)\big)}{V_n^2 V'^2(\mathbb{I}_n)} \\ &= \frac{V'^2_n}{V_n^2}\left\{1 + \frac{l^*\mu_1^2}{V'^2(\mathbb{I}_n)}\Big(1 + \frac{2S'(\mathbb{I}_n)}{\mu_1 l^*}\Big)\right\}. \end{aligned}$$

Applying the weak law of large numbers to ${V'}^2(\mathbb{I}_n)/l^*$ and $S'(\mathbb{I}_n)/l^*$, we see that in probability,

$$\Big(1+\frac{\mu_1^2}{\sigma^2}\Big)^{-1} \le \liminf_{n\to+\infty} \frac{v_{m*}-v_{k*}}{v'_{m*}-v'_{k*}} \le \limsup_{n\to+\infty} \frac{v_{m*}-v_{k*}}{v'_{m*}-v'_{k*}} \le 1+\frac{\mu_1^2}{\sigma^2}. \tag{57}$$

The statistics $\mathrm{SDI}(n,\rho)$ may be expressed as

$$\mathrm{SDI}(n,\rho) = \max_{1\le j\le J_n} \frac{1}{2\rho(2^{-j})} \max_{r\in \mathrm{D}_j} \Bigg| \sum_{r^-<v_k\le r} X_k - \sum_{r<v_k\le r^+} X_k \Bigg|$$

Define the random index J by $2^{-J} \ge v_{m^*} - v_{k^*} > 2^{-J-1}$ and consider first the configuration where $r^- \le v_{k^*} < v_{m^*} \le r$ where $r \in D_J$. Then

$$V_n^{-1}\mathrm{SDI}(n,\rho) \ge \frac{l^*|\mu_1|}{2V_n\rho(2^{-J})} - \frac{1}{2}R_n,$$

where

$$R_n := \frac{1}{V_n\rho(2^{-J})} \Bigg| \sum_{r^-<v_k\le r} X'_k - \sum_{r<v_k\le r^+} X'_k \Bigg|.$$

Due to (56), (57), the definition of J and the form of ρ, there are some positive random variables K_i, not depending on n such that

$$\frac{l^*|\mu_1|}{V_n\rho(2^{-J})} \ge \frac{K_1 l^*}{V_n\rho(v_{m*}-v_{k*})} \ge \frac{K_2\Big(\frac{l^*}{{V'_n}^2}\Big)V'_n}{\rho\Big(\frac{{V'}^2(\mathbb{I}_n)}{{V'_n}^2}\Big)} \ge K_3 \frac{n^{1/2}h_n}{\rho(h_n)}$$

and this lower bound goes to infinity in probability by Condition (16). So it remains to check that $R_n = O_{\mathrm{P}}(1)$.

To this end, we split R_n in four blocks $R_{n,1},\dots,R_{n,4}$ indexed respectively by $r^- < v_k$ with $k \le k^*$, $k \in \mathbb{I}_n$, $v_k \le r$ with $k > m^*$, $r < v_k \le r^+$. Recall that $\tau_n(t) := \max\{i \le n;\ v_i \le t\}$. To bound $R_{n,1}$, note that if $\tau_n(r^-) > k^*$, then $R_{n,1} = 0$. Assuming now that $\tau_n(r^-) \le k^*$, we get

$$\begin{aligned} R_{n,1} &= \frac{1}{\rho(2^{-J})}\frac{V'_n}{V_n}\Bigg|\frac{S'\big((\tau_n(r^-),k^*]\big)}{V'_n}\Bigg| \\ &\le \frac{V'_n}{V_n}\Big\|\frac{\zeta'_n}{V'_n}\Big\|_\rho \frac{\rho\big(v'_{k^*}-v'_{\tau_n(r^-)}\big)}{\rho(2^{-J})} \\ &\le \frac{V'_n}{V_n}\Big\|\frac{\zeta'_n}{V'_n}\Big\|_\rho \frac{\rho\big(v'_{k^*}-v'_{\tau_n(r^-)}\big)}{\rho\big(v_{k^*}-v_{\tau_n(r^-)}\big)}. \end{aligned} \tag{58}$$

In the bound (58), $V_n'/V_n = O_{\mathrm{P}}(1)$ by (56) and $\|V_n'^{-1}\zeta_n'\|_\rho = O_{\mathrm{P}}(1)$ by Theorem 12. To bound the last fraction, we observe that

$$\frac{v'_{k^*} - v'_{\tau_n(r^-)}}{v_{k^*} - v_{\tau_n(r^-)}} = \frac{V_n'^{-2} V'^2\big((\tau_n(r^-), k^*]\big)}{V_n^{-2} V^2\big((\tau_n(r^-), k^*]\big)} = \frac{V_n^{\,2}}{V_n'^{2}},$$

since $X_i' = X_i$ for $i \in (\tau_n(r^-), k^*]$. As ρ is of the form $\rho(h) = h^\alpha L(1/h)$ with L slowly varying, it easily follows from (56) that

$$\frac{\rho\big(v'_{k^*} - v'_{\tau_n(r^-)}\big)}{\rho\big(v_{k^*} - v_{\tau_n(r^-)}\big)} = O_{\mathrm{P}}(1).$$

Similarly the control of $R_{n,2}$ reduces to establishing the stochastic boudedness of the ratio $\rho(v'_{m^*} - v'_{k^*})/\rho(v_{m^*} - v_{k^*})$, which in turn follows from (57).

The control of $R_{n,3}$ and $R_{n,4}$ is similar to the one of $R_{n,1}$. Finally the proof is completed by a discussion of the configurations like in the proof of Theorem 4. □

3.7 Proof of Theorem 8

Like in the proof of the convergence of $\mathrm{DI}(n,\rho)$ (Theorem 3), it is enough to consider the case $\mu_0 = 0$. We shall assume $\sigma^2 = 1$. Denoting

$$\delta_{ni}(j,t) = \frac{1}{\rho(2^{-j})} n^{-1/2} \Big[\mathbf{1}\{nt^- < i \le nt\} - \mathbf{1}\{nt < i \le nt^+\}\Big],$$

where $j = 1, \dots, m_n$; $t \in \mathrm{D}_j$ and $i = 1, \dots, n$, we easily find the following representation

$$n^{-1/2}\mathrm{DI}(n,\rho,m_n) = \frac{1}{2} \max_{1\le j\le m_n} \max_{t\in \mathrm{D}_j} \Big| \sum_{i=1}^n X_i \delta_{ni}(j,t)\Big|.$$

Let $(\gamma_i, i = 1, \dots, n)$ be a collection of independent standard normal random variables. Consider

$$V_{n,\rho}(m_n) = \frac{1}{2} \max_{1\le j\le m_n} \max_{t\in \mathrm{D}_j} \Big| \sum_{i=1}^n \gamma_i \delta_{ni}(j,t)\Big|.$$

Set $k_n = 1 + 2 + \dots + 2^{m_n}$. For $x = (x_1, \dots, x_{k_n}) \in \mathbb{R}^{k_n}$, write

$$||x||_\infty := \max_{1\le l\le k_n} |x_l|.$$

Consider for each $\varepsilon > 0$ and $r \ge \varepsilon$ the function $f_{r,\varepsilon} : \mathbb{R}^{k_n} \to \mathbb{R}$ which possesses the following properties

i) for each $x \in \mathbb{R}^{k_n}$

$$\mathbf{1}\{||x||_\infty \le r - \varepsilon\} \le f_{r,\varepsilon}(x) \le \mathbf{1}\{||x||_\infty \le r + \varepsilon\}; \tag{59}$$

ii) the function f_ε is infinitely many times differentiable and for each $\ell \geq 1$ there exists an absolute constant $C_\ell > 0$ such that

$$\begin{aligned} ||f_{r,\varepsilon}^{(\ell)}|| := \quad & \sup\{|f_{r,\varepsilon}^{(\ell)}(x)(h)^\ell| : x, h \in \mathbb{R}^{k_n}, ||h||_\infty \leq 1\} \\ & \leq C_\ell \varepsilon^{-\ell} \log^{\ell-1} k_n. \end{aligned} \tag{60}$$

Here $f_{r,\varepsilon}^{(\ell)}$ denotes the ℓ-th derivative of the function $f_{r,\varepsilon}$ and $f_{r,\varepsilon}^{(\ell)}(x)(h)^\ell$ the corresponding differential. Such functions are constructed in [1]. Denoting for $i = 1, \ldots, n$, $\mathbf{X}_{ni} = (X_i \delta_{ni}(j,t),\ j = 1, \ldots, m_n;\ t \in \mathrm{D}_j)$ and $\mathbf{Y}_{ni} = (\gamma_i \delta_{ni}(j,t),\ t \in \mathrm{D}_j,\ j = 1, \ldots, m_n)$ we have

$$\mathrm{DI}(n, \rho, m_n) = \Big\| \sum_{i=1}^n \mathbf{X}_{ni} \Big\|_\infty, \quad V_{n,\rho}(m_n) = \Big\| \sum_{i=1}^n \mathbf{Y}_{ni} \Big\|_\infty.$$

Applying the property (59) of the functions $f_\varepsilon = f_{r+\varepsilon,\varepsilon}$ and $f_\varepsilon = f_{r-\varepsilon,\varepsilon}$ we derive for $r \geq \varepsilon$

$$\begin{aligned} & \Delta_n(r) := |\mathrm{P}(\mathrm{DI}(n, \rho, m_n) \leq r) - \mathrm{P}(V_{n,\rho}(m_n) \leq r)| = \\ & \leq \max \Big| \mathrm{E}\, f_\varepsilon\Big(\sum_{i=1}^n \mathbf{X}_{ni}\Big) - \mathrm{E}\, f_\varepsilon\Big(\sum_{i=1}^n \mathbf{Y}_{ni}\Big) \Big| + \mathrm{P}(r - \varepsilon \leq V_{n,\rho}(m_n) \leq r + \varepsilon), \end{aligned}$$

where maximum extends over two functions $f_\varepsilon = f_{r-\varepsilon,\varepsilon}$ and $f_\varepsilon = f_{r+\varepsilon,\varepsilon}$.

Writing $\mathbf{Z}_{nk} = \sum_{i=1}^{k-1} \mathbf{X}_{ni} + \sum_{i=k+1}^n \mathbf{Y}_{ni}$ and noting that $\mathbf{Z}_{nk} + \mathbf{X}_{nk} = \mathbf{Z}_{n,k+1} + \mathbf{Y}_{nk}$ we have

$$\begin{aligned} I := \quad & \mathrm{E}\, f_\varepsilon\Big(\sum_{i=1}^n \mathbf{X}_{ni}\Big) - \mathrm{E}\, f_\varepsilon\Big(\sum_{i=1}^n \mathbf{Y}_{ni}\Big) = \\ & \sum_{k=1}^n \mathrm{E}\, f_\varepsilon(\mathbf{Z}_{nk} + \mathbf{X}_{nk}) - \mathrm{E}\, f_\varepsilon(\mathbf{Z}_{nk} + \mathbf{Y}_{nk}). \end{aligned}$$

Next we shall use Taylor's expansion $f_\varepsilon(x+h) = f_\varepsilon(x) + f_\varepsilon'(x)h + 2^{-1} f_\varepsilon''(x)h^2 + R$ with interpolated remainder $|R| \leq 2^{1-\tau} ||f_\varepsilon''||^{1-\tau} ||f_\varepsilon'''||^\tau ||h||_\infty^{2+\tau}$ where $0 < \tau \leq 1$. Noting that for each $k = 1, \ldots, n$

$$\mathrm{E}\, f_\varepsilon'(\mathbf{Z}_{nk})(\mathbf{X}_{nk}) = \mathrm{E}\, f_\varepsilon'(\mathbf{Z}_{nk})(\mathbf{Y}_{nk}) = 0$$

and

$$\mathrm{E}\, f_\varepsilon''(\mathbf{Z}_{nk})(\mathbf{X}_{nk})^2 = \mathrm{E}\, f_\varepsilon''(\mathbf{Z}_{nk})(\mathbf{Y}_{nk})^2$$

we obtain $I \le 2^{1-\gamma}\varepsilon^{-3}\sum_{i=1}^n \left[\mathrm{E}\,||\mathbf{X}_{ni}||_\infty^{2+\tau} + \mathrm{E}\,||\mathbf{Y}_{ni}||_\infty^{2+\tau}\right]||f_\varepsilon''||^{1-\tau}||f_\varepsilon'''||^\tau$. Clearly

$$\mathrm{E}\,||\mathbf{X}_{ni}||_\infty^{2+\tau} \le \mathrm{E}\,|X_1|^{2+\tau} n^{-(2+\tau)/2}\rho^{-2-\tau}(2^{-m_n})$$

and

$$\mathrm{E}\,||\mathbf{Y}_{ni}||_\infty^{2+\tau} \le c_\tau n^{-(2+\tau)/2}\rho^{-2-\tau}(2^{-m_n}),$$

where c_τ depends on τ only. Finally, accounting (60) we obtain

$$|I| \le c_\tau \varepsilon^{-3} n^{-\tau/2}\rho^{-2-\tau}(2^{-m_n}) m_n^{1+\tau}.$$

Therefore, under the condition (19) we have for each $r \ge \varepsilon$

$$\limsup_{n\to\infty} \Delta_n(r) \le \lim_{n\to\infty} \mathrm{P}(r-\varepsilon \le V_{n,\rho}(m_n) \le r+\varepsilon).$$

If $r < \varepsilon$ we have evidently

$$\Delta_n(r) \le \Delta_n(\varepsilon) + 2\mathrm{P}(V_{n,\rho}(m_n) \le \varepsilon),$$

hence

$$\limsup_{n\to\infty} \Delta_n(r) \le 3 \lim_{n\to\infty} \mathrm{P}(r-\varepsilon \le V_{n,\rho}(m_n) \le r+\varepsilon) \tag{61}$$

for all $r > 0$. Since for each $a > 0$

$$\lim_{n\to\infty} \mathrm{P}(V_{n,\rho}(m_n) \le a) = F_\rho^{(1)}(a) = \mathrm{P}(\mathrm{DI}(\rho) \le a)$$

and the distribution function of $\mathrm{DI}(\rho)$ is continuous we complete the proof by letting $\varepsilon \to 0$ in (61). □

The proof of Theorem 9 is similar to the proof of Theorem 8. The only difference is in the definition of random vectors $\mathbf{X}_{ni}$ (the modification being evident) and the dimension k_n which now is $O(n^2)$.

3.8 Proof of Theorem 10

The result relies essentially on the representation of a standard Brownian motion as a series of triangular functions. For $r \in \mathrm{D}_j$, $j \ge 1$, let H_r be the $L^2[0,1]$ normalized Haar function, defined on $[0,1]$ by

$$H_r(t) = \begin{cases} +(r^+ - r^-)^{-1/2} = +2^{(j-1)/2} & \text{if } t \in (r^-, r]; \\ -(r^+ - r^-)^{-1/2} = -2^{(j-1)/2} & \text{if } t \in (r, r^+]; \\ 0 & \text{else.} \end{cases}$$

Define moreover $H_1(t) := \mathbf{1}_{[0,1]}(t)$. For $r \in \mathrm{D}_j$, $j \geq 1$, the triangular Faber-Schauder functions Λ_r are defined as

$$\Lambda_r(t) := \begin{cases} (t-r^-)/(r-r^-) = 2^j(t-r^-) & \text{if } t \in (r^-, r]; \\ (r^+-t)/(r^+-r) = 2^j(r^+-t) & \text{if } t \in (r, r^+]; \\ 0 & \text{else.} \end{cases}$$

In the special case $j = 0$, we set $\Lambda_1(t) := t$. The Λ_r's are linked to the H_r's in the general case $r \in \mathrm{D}_j$, $j \geq 1$ by

$$\Lambda_r(t) = 2(r^+ - r^-)^{-1/2} \int_0^t H_r(s)\,\mathrm{d}s = 2^{(j+1)/2} \int_0^t H_r(s)\,\mathrm{d}s \tag{62}$$

and in the special case $r = 1$ by

$$\Lambda_1(t) = \int_0^t H_1(s)\,\mathrm{d}s.$$

It is well known that any function $x \in C[0,1]$ such that $x(0) = 0$ may be expanded in the uniformly convergent series

$$x = \lambda_1(x)\Lambda_1 + \sum_{j=1}^{\infty} \sum_{r \in \mathrm{D}_j} \lambda_r(x)\Lambda_r,$$

where the $\lambda_r(x)$'s are given by (2).

Let us recall the Lvy - Kamp de Feriet representation of the Brownian motion, which was generalized by Shepp [13] with any orthonormal basis. Let $\{Y_1, Y_r; r \in \mathrm{D}_j, j \geq 1\}$ be a collection of independent standard normal random variables. Then the series

$$W(t) := Y_1 t + \sum_{j=1}^{\infty} \sum_{r \in \mathrm{D}_j} Y_r \int_0^t H_r(s)\,\mathrm{d}s, \quad t \in [0,1], \tag{63}$$

converges almost surely uniformly on $[0,1]$ and W is a Brownian motion started at 0. Removing the first term in (63) gives a Brownian bridge

$$B(t) := W(t) - tW(1) = \sum_{j=1}^{\infty} \sum_{r \in \mathrm{D}_j} Y_r \int_0^t H_r(s)\,\mathrm{d}s, \quad t \in [0,1]. \tag{64}$$

Now from (62) to (64), it is clear that the sequential Hölder norms of B may be writen as

$$\|B\|_\rho^{\mathrm{seq}} = \sup_{j \geq 1} \frac{1}{(2)^{1/2}\,\theta(2^j)} \max_{r \in D_j} |Y_r|. \tag{65}$$

Recalling that $\mathrm{DI}(\rho)$ has the same distribution as $\|B\|_\rho^{\mathrm{seq}}$, the explicit formula for its distribution function is easily derived from (65) using standard techniques for the maxima of independent identically distributed Gaussian variables. □

References

[1] Bentkus, V. Yu. (1984), Differentiable functions defined in the spaces c_0 and R^k, *Lithuan. Math. J.*, vol. **23**, No **2**, 146–154.

[2] Billingsley, P. (1999). *Convergence of Probability measures.* John Wiley & Sons, New York.

[3] Bingham, N. H., Goldie, C. M. & Teugels, J. L. (1987). Regular variation. Encyclopaedia of Mathematics and its Applications. Cambridge University Press.

[4] Brodsky, B., E. & Darkhovsky, B., S. (1993). *Non parametric methods in change point problems.* Kluwer.

[5] Csörgő, M. & Horváth, L. (1997). *Limit Theorems in Change-Point Analysis.* John Wiley & Sons, New York.

[6] Lamperti, J. (1962). On convergence of stochastic processes. *Trans. Amer. Math. Soc.* **104**, 430–435.

[7] Ledoux, M. and Talagrand, M. (1991). *Probability in Banach Spaces.* Springer-Verlag, Berlin, Heidelberg.

[8] O'Brien, G. L. (1980) A limit theorem for sample maxima and heavy branches in Galton-Watson trees. *J. Appl. Probab.* **17**, 539–545.

[9] Petrov, V. V. (1975) *Sums of independent random variables*, Springer.

[10] Račkauskas, A. & Suquet, Ch. (2001). Invariance principles for adaptive self-normalized partial sums processes. *Stochastic Process. Appl.*, **95**, 63–81.

[11] Račkauskas, A. & Suquet, Ch. (2001). Necessary and sufficient condition for the Hölderian functional central limit theorem. *Pub. IRMA Lille* **57-I**, preprint.

[12] Semadeni, Z. (1982). *Schauder bases in Banach spaces of continuous functions.* Lecture Notes in Math. **918**. Springer Verlag, Berlin, Heidelberg, New York.

[13] SHEPP, L.A. (1966). Radon-Nikodym derivatives of Gaussian measures. *Annals of Mathematical Statistics* **37**, 321–354.

[14] YAO, Q. (1993). Tests for change-points with epidemic alternatives. *Biometrika*, **80**, 179–191.