Bernoulli **23**(4B), 2017, 3412–3436
DOI: 10.3150/16-BEJ851

Integrated empirical processes in L^p with applications to estimate probability metrics

JAVIER CÁRCAMO

Departamento de Matemáticas, Universidad Autónoma de Madrid, 28049 Madrid, Spain.
E-mail: javier.carcamo@uam.es

We discuss the convergence in distribution of the r-fold (reverse) integrated empirical process in the space L^p, for $1 \le p \le \infty$. In the case $1 \le p < \infty$, we find the necessary and sufficient condition on a positive random variable X so that this process converges weakly in L^p. This condition defines a Lorentz space and can be also characterized in terms of several integrability conditions related to the process $\{(X-t)_+^r : t \ge 0\}$. For $p = \infty$, we obtain an integrability requirement on X guaranteeing the convergence of the integrated empirical process. In particular, these results imply a limit theorem for the stop-loss distance between the empirical and the true distribution. As an application, we derive the asymptotic distribution of an estimator of the Zolotarev distance between two probability distributions. The connections of the involved processes with equilibrium distributions and stochastic integrals with respect to the Brownian bridge are also briefly explained.

Keywords: distributional limit theorems; integrated Brownian bridge; integrated empirical process; Lorentz spaces; probability metrics; stochastic integral; stop-loss distance; Zolotarev metric

1. Introduction

Empirical processes play a central role in asymptotic statistics. Applications of the widely developed theory of empirical processes continuously arise in non-parametric statistics. For instance, many estimators are constructed by a plug-in approach, that is, the true (and unknown in practice) distribution of the underlying random variable is replaced with the empirical distribution corresponding to a random sample. The asymptotic properties of the estimator are often derived by analyzing the behavior of the empirical process in a suitable metric space and the help of other mathematical tools, such as invariance principles or continuous mapping theorems (see van der Vaart and Wellner [32]). The development of central limit theorems (CLT) for Banach valued random variables (see, for instance, Araujo and Giné [2] and Ledoux and Talagrand [19]), and the analysis of the convergence of the empirical process in such spaces, has also been essential for understanding the asymptotic behavior of important elements in statistics, such as Cramér-von Mises type statistics (see Shorack and Wellner [30] or del Barrio *et al.* [6]). As apparent in the recent book by Giné and Nickl [10], empirical processes are also in the core of the probabilistic foundations of infinite-dimensional statistical models.

Throughout this paper, unless noted otherwise, X is a non-degenerate and positive random variable with (cumulative) distribution function F, and $X_1, \ldots, X_n$ ($n \in \mathbb{N}$) is a random sample from X. We denote by $\mathbb{F}_n$ the classical *empirical distribution function* of the sample, i.e.,

$$\mathbb{F}_n(t) := \frac{1}{n}\sum_{i=1}^{n} I_{\{X_i \le t\}}, \qquad n \in \mathbb{N}, t \ge 0,$$

where I_A stands for the indicator function of the set A. In the sequel, $\mathbb{E}_n$ is the *empirical process* associated with X, that is,

$$\mathbb{E}_n(t) := \sqrt{n}\big(\mathbb{F}_n(t) - F(t)\big), \qquad n \in \mathbb{N}, t \ge 0. \tag{1}$$

For $1 \le p \le \infty$, we consider the space $L^p \equiv L^p([0,\infty), \mathcal{A}, m)$, where $\mathcal{A}$ and m are the Lebesgue σ-algebra and measure, respectively, of equivalence classes of measurable functions $f : [0,\infty) \longrightarrow \mathbb{R}$ endowed with the norm

$$\|f\|_p := \begin{cases} \left(\displaystyle\int_0^\infty \big|f(x)\big|^p \,\mathrm{d}x\right)^{1/p}, & \text{if } 1 \le p < \infty, \\ \operatorname*{ess\,sup}_{x\in[0,\infty)} \big|f(x)\big|, & \text{if } p = \infty. \end{cases}$$

In this work, we discuss the asymptotic behavior in L^p of the *r-fold (reverse) integrated empirical process*, that is, the process recursively defined by

$$\mathbb{I}_n^{[1]}(t) := \int_t^\infty \mathbb{E}_n(x)\,\mathrm{d}x \quad \text{and} \quad \mathbb{I}_n^{[r]}(t) := \int_t^\infty \mathbb{I}_n^{[r-1]}(x)\,\mathrm{d}x, \qquad r = 2, 3, \ldots.$$

Upon integration by parts, this process can be expressed as

$$\mathbb{I}_n^{[r]}(t) = \frac{1}{\Gamma(r)} \int_t^\infty (x-t)^{r-1}\mathbb{E}_n(x)\,\mathrm{d}x, \qquad n \in \mathbb{N}, t \ge 0, r \ge 1, \tag{2}$$

where $\Gamma(\cdot)$ is the Euler gamma function. This alternative representation of $\mathbb{I}_n^{[r]}$ makes sense for non-integer values of r. We therefore consider real values of $r \in [1,\infty)$. The process $\mathbb{I}_n^{[1]}$ will be simply denoted by $\mathbb{I}_n$.

The asymptotic behavior of $\mathbb{E}_n$ in L^1 is well-understood and it has been deeply discussed in del Barrio *et al.* [7]. For instance, it is known that $\mathbb{E}_n$ converges in distribution to $\mathbb{B}_F$ (the F-Brownian bridge) in L^1 if and only if X belongs to the Lorentz space $\mathcal{L}^{2,1}$ (see Section 2 for the definition of this space). We recall that the *F-Brownian bridge* is the process $\mathbb{B}_F := \mathbb{B} \circ F$, where $\mathbb{B}$ is a standard Brownian bridge on [0, 1], that is, a centered Gaussian process with covariance function $\gamma(s,t) = s \wedge t - st$ and continuous paths, with probability 1. Nevertheless, it should be remarked that the result that asserts the convergence of $\mathbb{E}_n$ in L^1 cannot be directly applied to derive the limiting distribution of $\mathbb{I}_n$ in L^1 because the linear mapping

$$\xi(f)(t) := \int_t^\infty f(x)\,\mathrm{d}x, \qquad t \ge 0,$$

is *not* continuous, not even well-defined, from L^1 to L^1.

Obviously, as $\mathbb{E}_n$ converges to $\mathbb{B}_F$ in several metric spaces, the natural candidate for being the (weak) limit of $\mathbb{I}_n^{[r]}$ in (2) is the *r-fold (reverse) integrated F-Brownian bridge*,

$$\mathbb{I}_F^{[r]}(t) := \frac{1}{\Gamma(r)} \int_t^{\infty} (x-t)^{r-1} \mathbb{B}_F(x)\,\mathrm{d}x, \qquad t \geq 0, r \geq 1. \tag{3}$$

Again, for simplicity, $\mathbb{I}_F^{[1]}$ is denoted by $\mathbb{I}_F$. The main goal of this work is to find conditions on the random variable X so that $\mathbb{I}_n^{[r]}$ converges weakly to $\mathbb{I}_F^{[r]}$ in L^p.

Apart from the possible independent theoretical interest of the results in this paper, the main ideas of this work can also be potentially useful to determine the asymptotic distribution of functionals in which integrals of empirical processes are involved. This may occur when considering the empirical counterpart of quantities related to integrals of distribution functions. We briefly mention three contexts in which this frequently happens.

1. *Reliability theory*: One of the main topics in reliability theory is the notion of the *mean residual life function* of a positive random variable X (see, for instance, Lai and Xie [18]), defined by

$$\mu(t) := \mathrm{E}(X-t|X>t) = \frac{\mathrm{E}(X-t)_+}{1-F(t)} = \frac{1}{1-F(t)} \int_t^{\infty} \bigl(1-F(x)\bigr)\,\mathrm{d}x, \qquad t \geq 0,$$

where $(a)_+ := \max\{a, 0\}$ is the positive part of the real number a. The function $\mu(t)$ characterizes a lifetime distribution and there are several ageing notions defined in terms of the behavior of $\mu(t)$.

Additionally, if $0 < \mu := \mathrm{E}X < \infty$, X is said to be *harmonic new better than used in expectation* (HNBUE) if

$$\int_t^{\infty} \bigl(1-F(x)\bigr)\,\mathrm{d}x \leq \mu \exp(-t/\mu) \qquad \text{for all } t \geq 0. \tag{4}$$

Analogously, X is *harmonic new worse than used in expectation* (HNWUE) if the reverse inequality in (4) holds. The HNBUE (respectively, HNWUE) class is fairly large and includes all the usual ageing (respectively, anti-ageing) classes of life distributions (see Lai and Xie [18] or Marshall and Olkin [21]).

2. *Probability metrics*: Let X and Y be two random variables with distribution functions F and G, respectively. For $r \in \mathbb{N}$ and $1 \leq p \leq \infty$, the *stop-loss metric* of order r in L^p (see Rachev *et al.* [26]) is defined by

$$d_{r,p}(F, G) := \frac{1}{r!} \bigl\| \pi_X^{[r]} - \pi_Y^{[r]} \bigr\|_p, \tag{5}$$

where

$$\pi_X^{[r]}(t) := \mathrm{E}(X-t)_+^r \quad \text{and} \quad \pi_Y^{[r]}(t) := \mathrm{E}(Y-t)_+^r, \qquad t \in \mathbb{R}.$$

In risk theory (see Denuit *et al.* [8]), the metrics $d_{1,\infty}$ and $d_{1,1}$ are respectively called the *stop-loss distance* and the *integrated stop-loss distance*. By Fubini's theorem, it can be readily checked that

$$d_{r,p}(F,G)=\begin{cases}\dfrac{1}{\Gamma(r)}\left(\displaystyle\int_{-\infty}^{\infty}\left|\int_t^{\infty}(x-t)^{r-1}\big(F(x)-G(x)\big)\,\mathrm{d}x\right|^p\mathrm{d}t\right)^{1/p}, & \text{if } p<\infty,\\[2ex] \dfrac{1}{\Gamma(r)}\sup_{t\in\mathbb{R}}\left|\displaystyle\int_t^{\infty}(x-t)^{r-1}\big(F(x)-G(x)\big)\,\mathrm{d}x\right|, & \text{if } p=\infty.\end{cases}\tag{6}$$

This metric is closely related to the so-called *Zolotarev metric* of order r in L^p (see Section 5 for details).

From (6), we see that $\sqrt{n}d_{r,p}(\mathbb{F}_n,F)=\|\mathbb{I}_n^{[r]}\|_p$. Hence, the convergence of the process $\mathbb{I}_n^{[r]}$ in L^p immediately translates into a limit theorem for the stop-loss distance between the empirical and the true distribution (see Remarks 2 and 7 in Section 3).

3. *Stochastic orderings*: Let X and Y be two random variables with distribution functions F and G, respectively. It is said that X is less than or equal to Y in the *stop-loss order* (see, for instance, Denuit *et al.* [8]) if

$$\int_t^{\infty}\big(F(x)-G(x)\big)\,\mathrm{d}x\geq 0,\qquad \text{for all } t\in\mathbb{R}.$$

This relation is also called the *risk-seeking stochastic dominance rule* in risk theory (see Levy [20]) and it is closely related to other important stochastic orders such as the (increasing) convex order and the Lorenz order (see Shaked and Shanthikumar [29]). Higher order stochastic dominance rules can be defined analogously by integrating recursively. Many inequality and poverty measures, such as (generalized) Gini indices and Lorenz curves, can also be expressed by means of integrals of distribution functions.

The results in this paper differ substantially from those obtained by Henze and Nikitin [14, 15], where the authors considered integrals of the empirical process with respect to the empirical measure. In this way, the integrability problems that arise in an infinite measure space are avoided. Integrated empirical distribution functions and integrated Brownian motion processes also appear in the context of the asymptotic theory of nonparametric estimates of convex functions (see Groeneboom *et al.* [12,13]). In this set of problems, the convergence issues come down to the local behavior of the integrated processes and the limiting process is a (drifted) version of integrated Brownian motion rather than Brownian bridge. Other integrated processes have been considered in different fields. For instance, there is a substantial interest in the integrated Brownian motion in connection with "small ball probabilities" (see Chen and Li [5]). In the same direction, Nazarov and Nikitin [22] deal with integrated Gaussian processes.

In the next section, we introduce the necessary definitions and some preliminary technical results. We recall some basic facts and properties of Lorentz spaces. We also concisely summarize previos well-known results on the weak convergence of random variables taking values in L^p. In Section 3, we establish the most significant contributions of this paper. For $1\leq p<\infty$, we characterize the convergence in distribution of $\mathbb{I}_n^{[r]}$ in L^p by means of several equivalent integrability conditions, such as the membership of X to a certain Lorentz space and others related to the process $\{(X-t)_+^r : t\geq 0\}$. These conditions also amount to the convergence of the empirical process in a weighted L^p space. We also obtain a sufficient condition that guarantees the weak convergence of $\mathbb{I}_n^{[r]}$ in L^∞. The techniques in the proofs differ depending on the value of p. In the case $1 \le p < \infty$, one of the ingredients of the proof of the main result (Theorem 1) is the (functional) CLT in L^p spaces. However, when $p = \infty$, we show that, under appropriate assumptions, the process $\mathbb{I}_n^{[r]}$ is asymptotically tight (Theorem 2). In Section 4, we point out some properties of the Gaussian process $\mathbb{I}_F$. Section 5 illustrates the usefulness of the ideas in this paper by deriving the asymptotic distribution of an estimator of the Zolotarev distance between two probability distributions. Finally, some concluding remarks, in which we discuss the relationships of the involved processes with equilibrium distributions and stochastic integrals with respect to the Brownian bridge, close the paper.

2. Definitions and auxiliary results

For $0 < p, q < \infty$, the *Lorentz space* $\mathcal{L}^{p,q} \equiv \mathcal{L}^{p,q}(\Omega, \mathcal{F}, \mathrm{P})$ of real-valued random variables is defined by

$$\mathcal{L}^{p,q} := \left\{ X : \|X\|_{p,q} := \left(q \int_0^\infty \bigl(t^p \mathrm{P}\bigl(|X| > t\bigr)\bigr)^{q/p} \frac{\mathrm{d}t}{t} \right)^{1/q} < \infty \right\},$$

where $(\Omega, \mathcal{F}, \mathrm{P})$ is the underlying probability space (see Ledoux and Talagrand [19], page 10). Note that $X \in \mathcal{L}^{p,q}$ if and only if X fulfills the integrability condition $\Lambda_{p,q}(X) < \infty$, where $\Lambda_{p,q}$ is the functional given by

$$\Lambda_{p,q}(X) := \int_0^\infty t^{q-1} \bigl(\mathrm{P}\bigl(|X| > t\bigr)\bigr)^{q/p} \, \mathrm{d}t.$$

In particular, $\mathcal{L}^{p,p} \equiv \mathcal{L}^p$ is the usual space of random variables on $(\Omega, \mathcal{F}, \mathrm{P})$ such that $\mathrm{E}|X|^p < \infty$. It can also be shown that $\mathcal{L}^{p,q_1} \subset \mathcal{L}^{p,q_2}$, whenever $q_1 \le q_2$ (see Grafakos [11], Section 1.4). Further, it is easy to check that $\mathcal{L}^{p+\varepsilon} \subset \mathcal{L}^{p,q}$, for all $\varepsilon > 0$. Observe additionally that $X \in \mathcal{L}^{pr,qr}$ if and only if $|X|^r \in \mathcal{L}^{p,q}$, for all $r > 0$.

For $1 \le p < \infty$ and $r \ge 0$, we introduce the following space of equivalence classes of measurable functions $f : [0, \infty) \longrightarrow \mathbb{R}$ given by

$$W^{p,r} := \left\{ f : \|f\|_{W^{p,r}} := \left(\int_0^\infty t^{pr} \bigl|f(t)\bigr|^p \, \mathrm{d}t \right)^{1/p} < \infty \right\}. \tag{7}$$

Observe that $W^{p,r}$ is a weighted L^p space and $W^{p,0} \equiv L^p$.

We define the linear map $\xi_r : W^{p,r} \longrightarrow L^p$ given by

$$\xi_r(f)(t) := \frac{1}{\Gamma(r)} \int_t^\infty (x - t)^{r-1} f(x) \, \mathrm{d}x, \qquad t \ge 0, r \ge 1. \tag{8}$$

The map ξ_r connects $\mathbb{I}_n^{[r]}$ with the empirical process as $\xi_r(\mathbb{E}_n) = \mathbb{I}_n^{[r]}$ (and $\xi_r(\mathbb{B}_F) = \mathbb{I}_F^{[r]}$). It is easy to see that ξ_r is a continuous mapping from $W^{p,r}$ to L^p using the following lemma, a direct consequence of an inequality by G.H. Hardy (see Bennett and Sharpley [3], equation (3.19), page 124).

Lemma 1. *For $1 \le p < \infty$, the following inequality holds*

$$\int_0^\infty \left(\int_t^\infty \left| f(x) \right| \mathrm{d}x \right)^p \mathrm{d}t \le p^p \int_0^\infty x^p \left| f(x) \right|^p \mathrm{d}x. \tag{9}$$

We also specify that if $\mathbb{S}$ and $\mathbb{S}_n$ ($n \in \mathbb{N}$) are stochastic processes taking values in L^p (with $1 \le p < \infty$), it is said that $\mathbb{S}_n$ *converges in distribution to* $\mathbb{S}$ *in* L^p if $\lim_{n\to\infty} \mathrm{E} f(\mathbb{S}_n) = \mathrm{E} f(\mathbb{S})$, for all continuous and bounded functions $f : L^p \longrightarrow \mathbb{R}$. Observe that if $\mathbb{S}$ and $\mathbb{S}_n$ are jointly measurable and have almost all their trajectories in L^p, then they can be identified with Borel-measurable random elements in L^p (see Byczkowski [4]). Therefore, the previous expectations are well-defined.

The case $p = \infty$ is different because L^∞ is *not* separable. The separability assumption is convenient to avoid a number of measurability problems. Therefore, in general, the expectations have to be substituted by outer expectations in the definition of weak convergence in L^∞ (see, for instance, van der Vaart and Wellner [32]). However, note that $\mathbb{I}_n^{[r]}$ and $\mathbb{I}_F^{[r]}$ actually take values in the separable space C_0, the space of all continuous functions $f : [0,\infty) \longrightarrow \mathbb{R}$ such that $\lim_{t\to\infty} f(t) = 0$. The weak convergence of $\mathbb{I}_n^{[r]}$ in L^∞ is hence equivalent to the convergence in C_0 (with the sup-norm) (see van der Vaart [31], Lemma 18.13), and the previous definition of convergence in distribution in L^p is still valid for $p = \infty$ and the processes $\mathbb{I}_n^{[r]}$ and $\mathbb{I}_F^{[r]}$.

In the following, we denote the weak convergence of probability measures in the space L^p by $\mathbb{S}_n \to_{\mathrm{w}} \mathbb{S}$ in L^p ($1 \le p \le \infty$).

Remark 1. For the process $\mathbb{I}_n^{[r]}$ in (2), by inequality (9), and for $1 \le p < \infty$, we obtain that

$$\begin{aligned}
\left\| \mathbb{I}_n^{[r]} \right\|_p^p &\le \frac{n^{p/2}}{\Gamma(r)^p} \left[\int_0^\infty \left(\int_t^\infty x^{r-1} \bigl(1 - F(x)\bigr) \mathrm{d}x \right)^p \mathrm{d}t \right. \\
&\qquad \left. + \int_0^\infty \left(\int_t^\infty x^{r-1} \bigl(1 - \mathbb{F}_n(x)\bigr) \mathrm{d}x \right)^p \mathrm{d}t \right] \\
&\le \frac{n^{p/2} p^p}{\Gamma(r)^p} \left[\Lambda_{r+1/p, rp+1}(X) + \frac{1}{(rp+1)} \frac{1}{n} \sum_{i=1}^n X_i^{rp+1} \right].
\end{aligned}$$

Hence, if $X \in \mathcal{L}^{r+1/p, rp+1}$, the trajectories of $\mathbb{I}_n^{[r]}$ belong to L^p a.s. Actually, it is readily checked that $X \in \mathcal{L}^{r+1}$ if and only if $\mathbb{I}_n^{[r]} \in L^1$ a.s. (and, consequently, $\mathbb{I}_n^{[r]} \in L^p$ a.s., for all $p \ge 1$). Analogously, it can be seen that $X \in \mathcal{L}^r$ if and only if $\mathbb{I}_n^{[r]} \in L^\infty$ a.s.

To analyze the convergence of $\mathbb{I}_n^{[r]}$ in L^p in the case $1 \le p < \infty$, we first need to discuss the convergence of $\mathbb{E}_n$ in the space $W^{p,r}$ defined in (7). Observe that $W^{p,r} = L^p([0,\infty), \mathcal{A}, \mu_{p,r})$, where $\mathrm{d}\mu_{p,r}(t) = t^{pr}\,\mathrm{d}m(t)$, with $\mathcal{A}$ and m the Lebesgue σ-algebra and measure on $[0,\infty)$, respectively. For this reason, next we briefly review previous well-known results on the CLT in L^p spaces.

We recall that a Borel random variable $\mathbb{X}$ taking values in a separable Banach space B is said to satisfy the CLT if the sequence $(\mathbb{S}_n/\sqrt{n})$ converges weakly in B to a (necessarily) Gaussian random variable, where $\mathbb{S}_n := \mathbb{X}_1 + \cdots + \mathbb{X}_n$ and $\mathbb{X}_1, \ldots, \mathbb{X}_n$ are n independent copies of $\mathbb{X}$. For $1 \le p < \infty$, the CLT in L^p-spaces is completely understood. For $1 \le p \le 2$, the Banach space L^p has cotype 2 (and type p), and, for $2 < p < \infty$, L^p has type 2 (and cotype p) (see Albiac and Kalton [1]). The case $p = 2$ is special because L^2 is both type 2 and cotype 2 (Hilbert space), and the results in this paper adopt a simpler form in L^2. The different geometric properties of cotype 2 and type 2 spaces lead to different characterizations of the CLT in L^p, according to $1 \le p \le 2$ or $2 < p < \infty$. The next lemma summarizes previous results in this direction. Part (i) can be found in Araujo and Giné [2], page 205, whereas part (ii) can be traced back to Pisier and Zinn [25], Theorem 5.1 (see also Ledoux and Talagrand [19], Theorem 10.10).

Lemma 2. *Let $(S, \mathcal{F}, \mu)$ be a σ-finite measure space and let $\mathbb{X}$ be a centered process taking values in $L^p(\mu) \equiv L^p(S, \mathcal{F}, \mu)$.*

(i) *For $1 \le p \le 2$, $\mathbb{X}$ satisfies the CLT in $L^p(S, \mathcal{F}, \mu)$ if and only if*

$$\int_S \bigl(\mathrm{E}\mathbb{X}^2(t)\bigr)^{p/2}\,\mathrm{d}\mu(t) < \infty. \tag{10}$$

(ii) *For $2 < p < \infty$, $\mathbb{X}$ satisfies the CLT in $L^p(S, \mathcal{F}, \mu)$ if and only if (10) holds and*

$$\lim_{t\to\infty} t^2 \mathrm{P}\bigl(\|\mathbb{X}\|_{L^p(\mu)} \ge t\bigr) = 0. \tag{11}$$

Condition (10) amounts to saying that $\mathbb{X}$ is *pre-Gaussian*, that is, there is a Gaussian Radon measure with the same covariance as $\mathbb{X}$. Obviously, this is always a necessary condition.

To discuss the weak convergence in L^∞, we will use the following result which shows that weak convergence in L^∞ can be characterized as asymptotic tightness plus convergence of marginals (see van der Vaart and Wellner [32], Theorems 1.5.4 and 1.5.6, or van der Vaart [31], Theorem 18.14). For any set T, the space $L^\infty(T)$ is defined as the set of all uniformly bounded, real functions on T with the sup-norm.

Lemma 3. *Let T be an arbitrary set and let $\mathbb{X}_n : \Omega_n \to L^\infty(T)$ be a sequence of maps.*

(i) *If $\mathbb{X}_n$ is asymptotically tight and for every finite set of points $t_1, \ldots, t_k$ in T, the marginals $(\mathbb{X}_n(t_1), \ldots, \mathbb{X}_n(t_k))$ converge in distribution in $\mathbb{R}^k$ to the marginals $(\mathbb{X}(t_1), \ldots, \mathbb{X}(t_k))$ of a stochastic process $\mathbb{X}$, then there exists a version of $\mathbb{X}$ with uniformly bounded sample paths and $\mathbb{X}_n \to_{\mathrm{w}} \mathbb{X}$ in $L^\infty(T)$.*

(ii) *The sequence $\mathbb{X}_n$ is asymptotically tight if and only if for every $\varepsilon, \delta > 0$, there exists a partition of T into finitely many sets $T_1, \ldots, T_k$ such that*

$$\limsup_{n\to\infty} \mathrm{P}^*\Bigl(\sup_i \sup_{s,t\in T_i} \bigl|X_n(s) - X_n(t)\bigr| \ge \varepsilon\Bigr) \le \delta,$$

where P^* stands for the outer probability.

3. Integrated empirical processes in L^p

As mentioned in the Introduction, the results in this section are divided in two groups according to $1 \le p < \infty$ or $p = \infty$. If p is finite, we will apply the CLT in L^p spaces summarized in Lemma 2. When $p = \infty$, we will show the asymptotic tightness of $\mathbb{I}_n^{[r]}$ using Lemma 3.

3.1. The case $1 \le p < \infty$

The following theorem characterizes the weak convergence of $\mathbb{I}_n^{[r]}$ in L^p when $1 \le p < \infty$ by means of several equivalent integrability conditions. We recall that the space $W^{p,r}$ is defined in (7).

Theorem 1. *For $1 \le p < \infty$ and $r \ge 1$, the following assertions are mutually equivalent.*

(a) $X \in \mathcal{L}^{2(r+1/p),pr+1}$.
(b) $\mathbb{E}_n \to_{\mathrm{w}} \mathbb{B}_F$ *in* $W^{p,r}$.
(c) $\mathbb{I}_n^{[r]} \to_{\mathrm{w}} \mathbb{I}_F^{[r]}$ *in* L^p.
(d) $\int_0^\infty (\operatorname{Var}(X-t)_+^r)^{p/2}\,\mathrm{d}t < \infty$.
(e) $\int_0^\infty (\mathrm{E}(X-t)_+^{2r}))^{p/2}\,\mathrm{d}t < \infty$.

Moreover, $\mathbb{I}_F^{[r]}$ is a centered Gaussian process.

Proof. Let us fix $1 \le p < \infty$ and $r \ge 1$. We assume first that (a) is satisfied. To begin, we can write

$$\mathbb{E}_n = \frac{1}{\sqrt{n}}\sum_{i=1}^n \mathbb{X}_i,$$

where $\mathbb{X}_1, \dots, \mathbb{X}_n$ are independent copies of the process

$$\mathbb{X}(t) := \mathrm{P}(X > t) - I_{\{X>t\}}, \qquad t \ge 0. \tag{12}$$

As $W^{p,r} = L^p([0,\infty), \mathcal{A}, \mu_{p,r})$, where $\mathrm{d}\mu_{p,r}(t) = t^{pr}\,\mathrm{d}m(t)$, with m and $\mathcal{A}$ the Lebesgue measure and σ-algebra on $[0,\infty)$, respectively, we can apply Lemma 2 to the process $\mathbb{X}$ in (12). When $1 \le p \le 2$, from Lemma 2(i), we have that $\mathbb{X}$ satisfies the CLT in $W^{p,r}$ if and only if

$$\int_0^\infty \big(\mathrm{E}\mathbb{X}^2(t)\big)^{p/2}\,\mathrm{d}\mu_{p,r}(t)$$
$$= \int_0^\infty t^{pr}\big(F(t)\big(1-F(t)\big)\big)^{p/2}\,\mathrm{d}t < \infty,$$

which is equivalent to $\Lambda_{2(r+1/p),pr+1}(X) < \infty$. Further, the limiting Gaussian process of $\mathbb{E}_n$ is $\mathbb{B}_F$ because they have the same covariance. We thus see that (a) is equivalent to (b).

When $2 < p < \infty$, by Lemma 2(ii), it suffices to show that (a) implies that $\mathbb{X}$ in (12) satisfies (11) (for the norm of the space $W^{p,r}$). We have that

$$\begin{aligned}\|\mathbb{X}\|_{W^{p,r}}^p &= \int_0^X t^{pr}\big(\mathrm{P}(X \le t)\big)^p\,\mathrm{d}t + \int_X^\infty t^{pr}\big(\mathrm{P}(X > t)\big)^p\,\mathrm{d}t \\ &\le \frac{X^{pr+1}}{pr+1} + \Lambda_{r+1/p,pr+1}(X) \\ &\le X^{pr+1} + \Lambda_{2(r+1/p),pr+1}(X).\end{aligned} \tag{13}$$

From (13), and taking into account that (a) means that $\Lambda_{2(r+1/p),pr+1}(X) < \infty$, we obtain that

$$\limsup_{t\to\infty} t^2\mathrm{P}\big(\|\mathbb{X}\|_{W^{p,r}} \ge t\big) \le \limsup_{t\to\infty} t^2\mathrm{P}\big(X^{pr+1} \ge t^p\big) = \limsup_{t\to\infty} t^{2(r+1/p)}\mathrm{P}(X \ge t). \tag{14}$$

Finally, condition $\Lambda_{2(r+1/p),pr+1}(X) < \infty$ implies that

$$t^{pr+1}\big(\mathrm{P}(X \ge t)\big)^{p/2} = \big(t^{2(r+1/p)}\mathrm{P}(X \ge t)\big)^{p/2} \to 0 \qquad \text{as } t \to \infty.$$

We thus obtain that $\lim_{t\to\infty} t^{2(r+1/p)}\mathrm{P}(X > t) = 0$ and, from (14) and Lemma 2(ii), we conclude that (a) and (b) are equivalent.

Next, we assume that (b) holds. By inequality (9), the linear mapping $\xi_r : W^{p,r} \longrightarrow L^p$ defined in (8) satisfies

$$\big\|\xi_r(f)\big\|_p \le \frac{p}{\Gamma(r)}\|f\|_{W^{p,r}},$$

and it is therefore continuous. Thus, by (b) and the continuous mapping theorem (see for instance van der Vaart [31], Theorem 18.11), we obtain that

$$\mathbb{I}_n^{[r]} = \xi_r(\mathbb{E}_n) \to_{\mathrm{w}} \xi_r(\mathbb{B}_F) = \mathbb{I}_F^{[r]} \qquad \text{in } L^p,$$

and (c) holds.

Assume now that (c) is fulfilled. Taking into account the equality

$$\mathrm{E}_{F_Y}(Y-t)_+^r = r\int_t^\infty (x-t)^{r-1}\big(1-F_Y(x)\big)\,\mathrm{d}x, \qquad t \ge 0, \tag{15}$$

which holds for any random variable Y with distribution function F_Y, we have that

$$\begin{aligned}\mathbb{I}_n^{[r]}(t) &= \frac{\sqrt{n}}{\Gamma(r)}\left[\int_t^\infty (x-t)^{r-1}\big(1-F(x)\big)\,\mathrm{d}x - \int_t^\infty (x-t)^{r-1}\big(1-\mathbb{F}_n(x)\big)\,\mathrm{d}x\right] \\ &= \frac{\sqrt{n}}{\Gamma(r+1)}\big[\mathrm{E}_F(X-t)_+^r - \mathrm{E}_{\mathbb{F}_n}(X-t)_+^r\big] \\ &= \frac{1}{\Gamma(r+1)}\frac{1}{\sqrt{n}}\sum_{i=1}^n\big[\mathrm{E}(X_i-t)_+^r - (X_i-t)_+^r\big].\end{aligned} \tag{16}$$

Therefore, the centered process $\mathbb{Y}(t) := (X-t)_+^r - \mathrm{E}(X-t)_+^r$ ($t \geq 0$) satisfies the CLT in L^p. In particular (see Lemma 2), $\mathbb{Y}$ fulfills (10), and (d) is therefore satisfied. It also follows that $\mathbb{I}_F^{[r]}$ is a centered Gaussian process.

Let us assume that (d) holds. Let us consider first the case $r = 1$. From equality (15), it is readily checked that

$$\frac{\mathrm{d}}{\mathrm{d}t} \operatorname{Var}(X-t)_+ = -2F(t)\mathrm{E}(X-t)_+, \qquad t > 0. \tag{17}$$

Therefore, by (17), (15), and Fubini's theorem, we have that

$$\begin{aligned}\operatorname{Var}(X-t)_+ &= \int_t^\infty 2F(x)\mathrm{E}(X-x)_+ \mathrm{d}x \\ &= \int_t^\infty \bigl(1-F(s)\bigr)\left(\int_t^s 2F(x)\,\mathrm{d}x\right)\mathrm{d}s.\end{aligned}$$

In particular, for $t \geq \mathrm{Me}(X)$ (the median of X), we obtain

$$\operatorname{Var}(X-t)_+ \geq \int_t^\infty (s-t)\bigl(1-F(s)\bigr)\,\mathrm{d}s = \frac{1}{2}\mathrm{E}(X-t)_+^2.$$

This shows that (d) implies (e) when $r = 1$.

Now, we fix $r > 1$. Condition (d) implies that $X \in \mathcal{L}^{2r}$ (i.e., $\mathrm{E}X^{2r} < \infty$). By dominated convergence, we see that $\mathrm{E}(X-t)_+^r \to 0$, as $t \to \infty$. In particular, there exists $T > 0$ such that, for all $t \geq T$,

$$\begin{aligned}\operatorname{Var}(X-t)_+^r &\geq \mathrm{E}(X-t)_+^{2r} - \mathrm{E}(X-t)_+^r \\ &= \int_t^\infty r\bigl(2(x-t)^{2r-1} - (x-t)^{r-1}\bigr)\bigl(1-F(x)\bigr)\,\mathrm{d}x \\ &\geq \int_{t+2}^\infty 2r(x-t-2)^{2r-1}\bigl(1-F(x)\bigr)\,\mathrm{d}x - \int_t^{t+2} r\bigl(1-F(x)\bigr)\,\mathrm{d}x,\end{aligned}$$

where for the second inequality we have used the fact that $2x^{2r-1} - x^{r-1} \geq 2(x-2)^{2r-1}$, for $x \geq 2$ and $r \geq 1$, and $2x^{2r-1} - x^{r-1} \geq -1$, for $0 \leq x \leq 2$ and $r \geq 1$. Therefore, we conclude that

$$\mathrm{E}\bigl(X-(t+2)\bigr)_+^{2r} \leq \operatorname{Var}(X-t)_+^r + 2r\mathrm{P}(X > t), \qquad t \geq T. \tag{18}$$

Now, when $1 \leq p \leq 2$, the function given by $x \mapsto x^{p/2}$ is subadditive. Hence, from (18), we have that

$$\int_{T+2}^\infty \bigl(\mathrm{E}(X-t)_+^{2r}\bigr)^{p/2}\,\mathrm{d}t \leq \int_T^\infty \bigl(\operatorname{Var}(X-t)_+^r\bigr)^{p/2}\,\mathrm{d}t + (2r)^{p/2}\int_T^\infty \bigl(\mathrm{P}(X>t)\bigr)^{p/2}\,\mathrm{d}t. \tag{19}$$

In the case $2 < p < \infty$, we use Minkowski's inequality, and, from (18), we obtain that

$$\begin{aligned}\left(\int_{T+2}^{\infty}\left(\mathrm{E}(X-t)_+^{2r}\right)^{p/2}\mathrm{d}t\right)^{2/p} &\leq \left(\int_{T}^{\infty}\left(\mathrm{Var}(X-t)_+^{r}\right)^{p/2}\mathrm{d}t\right)^{2/p}\\ &\quad + 2r\left(\int_{T}^{\infty}\left(\mathrm{P}(X>t)\right)^{p/2}\mathrm{d}t\right)^{2/p}.\end{aligned} \tag{20}$$

To finish, it is enough to note that $\mathcal{L}^{2r} \subset \mathcal{L}^{2/p,1}$, and hence, if (d) holds, the integrals in the right-hand side of (19)–(20) are finite. We thus see that (e) is satisfied.

Finally, it holds that

$$\mathrm{E}(X-t)_+^{2r} \geq \int_{\{X-t>t\}}(X-t)^{2r}\,\mathrm{dP} \geq t^{2r}\mathrm{P}(X>2t), \qquad t \geq 0.$$

Therefore, we conclude that (e) implies (a), and the proof of the theorem is complete. □

Remark 2. Under the conditions in Theorem 1, we directly obtain that $\|\mathbb{I}_n^{[r]}\|_p \to_{\mathrm{d}} \|\mathbb{I}_F^{[r]}\|_p$, where "$\to_{\mathrm{d}}$" stands for the usual convergence in distribution of random variables. In other words, whenever $X \in \mathcal{L}^{2(r+1/p),pr+1}$, we conclude that

$$\sqrt{n}d_{r,p}(\mathbb{F}_n, F) \to_{\mathrm{d}} \left\|\mathbb{I}_F^{[r]}\right\|_p,$$

where $d_{r,p}$ is the stop-loss metric of order r in L^p defined in (5)–(6).

Remark 3. As mentioned in the Introduction, $\mathbb{E}_n \to_{\mathrm{w}} \mathbb{B}_F$ in L^1 is equivalent to the finiteness of $\Lambda_{2,1}(X)$, which defines $\mathcal{L}^{2,1}$ (see del Barrio *et al.* [7]). The integrability requirement to obtain $\mathbb{I}_n^{[r]} \to_{\mathrm{w}} \mathbb{I}_F^{[r]}$ in L^1 is exactly $(r+1)$ times more demanding, as we need that $X \in \mathcal{L}^{2(r+1),r+1}$, which amounts to saying that $X^{r+1} \in \mathcal{L}^{2,1}$.

We believe that the case $p = 2$ should be emphasized because of its neatness and simplicity:

$$\mathbb{I}_n^{[r]} \to_{\mathrm{w}} \mathbb{I}_F^{[r]} \text{ in } L^2 \text{ if and only if } X \in \mathcal{L}^{2r+1}, \text{ i.e., } \mathrm{E}X^{2r+1} < \infty.$$

Observe also that for $1 \leq p < 2$ and $\varepsilon > 0$, it holds that

$$\mathcal{L}^{2(r+1/p)+\varepsilon} \subset \mathcal{L}^{2(r+1/p),pr+1} \subset \mathcal{L}^{2(r+1/p)}$$

(see Section 2). Therefore, condition (a) in Theorem 1, that is,

$$\Lambda_{2(r+1/p),pr+1}(X) = \int_0^{\infty} t^{pr}\left(\mathrm{P}(X>t)\right)^{p/2}\mathrm{d}t < \infty,$$

is slightly stronger than $\mathrm{E}X^{2(r+1/p)} < \infty$.

However, when $2 < p < \infty$, we have that

$$\mathcal{L}^{2(r+1/p)} \subset \mathcal{L}^{2(r+1/p),pr+1}.$$

Condition (a) in Theorem 1 is hence slightly weaker than $\mathrm{E}X^{2(r+1/p)} < \infty$.

Remark 4. Theorem 1 provides a new characterization of Lorentz spaces of (positive) random variables. For $2 < p < \infty$ and $2 \le q < \infty$, we actually have that $X \in \mathcal{L}^{p,q}$ if and only if

$$\int_0^\infty \bigl(\mathrm{E}(X-t)_+^{p(1-1/q)}\bigr)^{q/p}\,\mathrm{d}t < \infty,$$

or, alternatively, if and only if

$$\int_0^\infty \bigl(\mathrm{Var}(X-t)_+^{p(1-1/q)/2}\bigr)^{q/p}\,\mathrm{d}t < \infty.$$

Remark 5. Let us consider the process

$$\mathbb{J}_n(t) := \int_0^t \mathbb{E}_n(x)\,\mathrm{d}x, \qquad t \ge 0.$$

We obviously have that $\lim_{t\to\infty}\mathbb{J}_n(t) = \sqrt{n}(\mu - \hat{\mu})$, where μ and $\hat{\mu}$ are the population and sample means, respectively. Hence $\mathbb{J}_n \notin L^p$ $(1 \le p < \infty)$ unless $\hat{\mu} = \mu$. Moreover, taking into account that $\mathbb{J}_n = \sqrt{n}(\mu - \hat{\mu}) - \mathbb{I}_n$, by Remark 1 and the standard CLT, we actually obtain that when $X \in \mathcal{L}^2$, $\lim_{n\to\infty}\int_0^\infty \mathbb{J}_n(t)\,\mathrm{d}t$ takes the values $\pm\infty$ with probability $1/2$.

3.2. The case $p = \infty$

The following theorem provides a sufficient condition so that $\mathbb{I}_n^{[r]}$ converges weakly to $\mathbb{I}_F^{[r]}$ in L^∞.

Theorem 2. *Let $r \ge 1$ and assume that $X \in \mathcal{L}^{2r,r}$. We have that $\mathbb{I}_n^{[r]} \to_{\mathrm{w}} \mathbb{I}_F^{[r]}$ in L^∞.*

Proof. We fix $r \ge 1$. As the finite-dimensional distributions of $\mathbb{I}_n^{[r]}$ converge in distribution to the corresponding marginals of $\mathbb{I}_F^{[r]}$, by Lemma 3(i), it suffices to show that $\mathbb{I}_n^{[r]}$ is asymptotically tight. We will show the equivalent condition of Lemma 3(ii). For $\eta > 0$ and $m \in \mathbb{N}$, we consider the finite partition of $[0,\infty)$ given by $P(\eta, m) := \{T_i : i = 1,\dots,m+1\}$ with

$$T_i := \bigl[(i-1)\eta, i\eta\bigr) \qquad (i = 1,\dots,m) \quad \text{and} \quad T_{m+1} := [m\eta, \infty).$$

We also denote

$$S_n^{[r]}\bigl(P(\eta,m)\bigr) := \sup_{1\le i\le m+1}\sup_{s,t\in T_i}\bigl|\mathbb{I}_n^{[r]}(s) - \mathbb{I}_n^{[r]}(t)\bigr|. \tag{21}$$

If $1 \le i \le m$ and $s, t \in T_i$ with $s < t$, we have that

$$\begin{aligned}
\bigl|\mathbb{I}_n^{[r]}(s) - \mathbb{I}_n^{[r]}(t)\bigr| &\le \frac{1}{\Gamma(r)}\int_s^t (x-s)^{r-1}\bigl|\mathbb{E}_n(x)\bigr|\,\mathrm{d}x \\
&\quad + \frac{1}{\Gamma(r)}\int_t^\infty \bigl[(x-s)^{r-1} - (x-t)^{r-1}\bigr]\bigl|\mathbb{E}_n(x)\bigr|\,\mathrm{d}x \\
&\le \frac{\eta^r}{\Gamma(r+1)}\|\mathbb{E}_n\|_\infty + \frac{1}{\Gamma(r)}\int_t^\infty \bigl[(x-s)^{r-1} - (x-t)^{r-1}\bigr]\bigl|\mathbb{E}_n(x)\bigr|\,\mathrm{d}x.
\end{aligned} \tag{22}$$

Now, to control the last integral in (22), we use the following inequality for $x, \gamma \geq 0$,

$$(x+\gamma)^{r-1} - x^{r-1} \leq \begin{cases} \gamma^{r-1}, & \text{if } 1 < r < 2, \\ \gamma(r-1)(x+\gamma)^{r-2}, & \text{if } r \geq 2. \end{cases} \tag{23}$$

(Inequality (23) follows by subadditivity if $1 < r < 2$, and by the mean value theorem when $r \geq 2$.) From (22)–(23), it is easy to see that

$$\left|\mathbb{I}_n^{[r]}(s) - \mathbb{I}_n^{[r]}(t)\right| \leq \begin{cases} \eta\|\mathbb{E}_n\|_\infty, & \text{if } r = 1, \\ \dfrac{\eta^r}{\Gamma(r+1)}\|\mathbb{E}_n\|_\infty + \dfrac{\eta^{r-1}}{\Gamma(r)}\|\mathbb{E}_n\|_1, & \text{if } 1 < r < 2, \\ \dfrac{\eta^r}{\Gamma(r+1)}\|\mathbb{E}_n\|_\infty + \dfrac{\eta}{\Gamma(r-1)}\|\mathbb{E}_n\|_{W^{1,r-2}}, & \text{if } r \geq 2. \end{cases} \tag{24}$$

(Recall the definition of the norm of $W^{p,r}$ in (7).)

If $s, t \in T_{m+1}$, we have that

$$\begin{aligned} \left|\mathbb{I}_n^{[r]}(s) - \mathbb{I}_n^{[r]}(t)\right| &\leq \frac{1}{\Gamma(r)}\int_s^\infty (x-s)^{r-1}\left|\mathbb{E}_n(x)\right|\mathrm{d}x + \frac{1}{\Gamma(r)}\int_t^\infty (x-t)^{r-1}\left|\mathbb{E}_n(x)\right|\mathrm{d}x \\ &\leq \frac{2}{\Gamma(r)}\int_{m\eta}^\infty x^{r-1}\left|\mathbb{E}_n(x)\right|\mathrm{d}x \\ &= \frac{2}{\Gamma(r)}\left\|\mathbb{E}_n I_{(m\eta,\infty)}\right\|_{W^{1,r-1}}. \end{aligned} \tag{25}$$

From (24)–(25), we conclude that $S_n^{[r]}(P(\eta, m))$ in (21) can be bounded above by

$$\begin{cases} \eta\|\mathbb{E}_n\|_\infty + \|\mathbb{E}_n I_{(m\eta,\infty)}\|_1, & \text{if } r = 1, \\ \dfrac{\eta^r}{\Gamma(r+1)}\|\mathbb{E}_n\|_\infty + \dfrac{\eta^{r-1}}{\Gamma(r)}\|\mathbb{E}_n\|_1 + \dfrac{2}{\Gamma(r)}\|\mathbb{E}_n I_{(m\eta,\infty)}\|_{W^{1,r-1}}, & \text{if } 1 < r < 2, \\ \dfrac{\eta^r}{\Gamma(r+1)}\|\mathbb{E}_n\|_\infty + \dfrac{\eta}{\Gamma(r-1)}\|\mathbb{E}_n\|_{W^{1,r-2}} & \\ \qquad + \dfrac{2}{\Gamma(r)}\|\mathbb{E}_n I_{(m\eta,\infty)}\|_{W^{1,r-1}}, & \text{if } r \geq 2. \end{cases} \tag{26}$$

Finally, let us fix $\varepsilon, \delta > 0$. We assume first that $r = 1$. If $X \in \mathcal{L}^{2,1}$, we have that $\mathbb{E}_n \to_{\mathrm{w}} \mathbb{B}_F$ in L^1. Therefore, from (26), we obtain that

$$\limsup_{n\to\infty} \mathrm{P}^*\left(S_n^{[1]}\left(P(\eta, m)\right) \geq \varepsilon\right) \leq \mathrm{P}\left(\eta\|\mathbb{B}_F\|_\infty \geq \varepsilon/2\right) + \mathrm{P}\left(\|\mathbb{B}_F I_{(m\eta,\infty)}\|_1 \geq \varepsilon/2\right). \tag{27}$$

We can take $\eta_0 = \eta_0(\varepsilon, \delta) > 0$ small enough so that

$$\mathrm{P}\left(\eta_0\|\mathbb{B}_F\|_\infty \geq \varepsilon/2\right) \leq \delta/2. \tag{28}$$

Also, by dominated convergence, we have that $\|\mathbb{B}_F I_{(m\eta_0,\infty)}\|_1 \to 0$ a.s., as $m \to \infty$. Hence, we can select a large enough $m_0 = m_0(\varepsilon, \delta, \eta_0) \in \mathbb{N}$ so that

$$\mathrm{P}\big(\|\mathbb{B}_F I_{(m_0\eta_0,\infty)}\|_1 \geq \varepsilon/2\big) \leq \delta/2. \tag{29}$$

For such values η_0 and m_0, and from (27), (28) and (29), we obtain that

$$\limsup_{n\to\infty} \mathrm{P}^*\big(S_n^{[1]}\big(P(\eta_0, m_0)\big) \geq \varepsilon\big) \leq \delta.$$

By Lemma 3(ii), we conclude that $\mathbb{I}_n^{[1]}$ is asymptotically tight.

When $r > 1$, by Theorem 1, we obtain that $X \in \mathcal{L}^{2r,r}$ is equivalent to $\mathbb{E}_n \to_{\mathrm{w}} \mathbb{B}_F$ in $W^{1,r-1}$ (and additionally, when $r \geq 2$, $X \in \mathcal{L}^{2r,r}$ also implies that $\mathbb{E}_n \to_{\mathrm{w}} \mathbb{B}_F$ in $W^{1,r-2}$). Therefore, the bounds in (26), and a similar reasoning as for the case $r = 1$, show that $\mathbb{I}_n^{[r]}$ is asymptotically tight, and the proof of the theorem is complete. □

Remark 6. In contrast to Theorem 1, Theorem 2 provides only a sufficient condition so that $\mathbb{I}_n^{[r]}$ converges in L^∞. Therefore, the requirement $X \in \mathcal{L}^{2r,r}$ in the statement of the previous theorem may not be optimal. Note however that, from (16), the condition $X \in \mathcal{L}^{2r}$ is necessary for $\mathbb{I}_n^{[r]}$ to converge weakly in L^∞.

It is also interesting to observe that by letting $p \to \infty$ in Theorem 1(a), we obtain $X \in \mathcal{L}^{2r,\infty}$, that is, X belongs to the weak-$\mathcal{L}^{2r}$ space, the space of all random variables such that $\sup_{t\geq 0} t^{2r}\mathrm{P}(X > t) < \infty$ (see Ledoux and Talagrand [19], page 9). Obviously, $\mathcal{L}^{2r,r} \subset \mathcal{L}^{2r} \subset \mathcal{L}^{2r,\infty}$.

Remark 7. Theorem 2 implies that Remark 2 is still valid for $p = \infty$ (with the obvious modifications).

4. The integrated F-Brownian bridge

In this section, we briefly enumerate some properties of the process $\mathbb{I}_F$ in (3). The corresponding properties of $\mathbb{I}_F^{[r]}$, for $r > 1$, are analogous.

We first point out that the trajectories of the centered Gaussian process $\mathbb{I}_F$ belong to $C^1([0,\infty))$ a.s. By Fubini's theorem, the covariance function of $\mathbb{I}_F$ is given by

$$\gamma(s,t) = \int_s^\infty \int_t^\infty \big[F(x) \wedge F(y) - F(x)F(y)\big]\,\mathrm{d}x\,\mathrm{d}y, \qquad s,t > 0. \tag{30}$$

By (16), we also have that

$$\gamma(s,t) = \mathrm{E}\big[(X-s)_+(X-t)_+\big] - \mathrm{E}(X-s)_+\mathrm{E}(X-t)_+. \tag{31}$$

From (17), we see that $\mathrm{Var}(\mathbb{I}_F(t)) = \mathrm{Var}[(X-t)_+]$ is a decreasing function of $t \geq 0$. In fact, it is also easy to check that $\gamma(s,t)$ is decreasing in each argument.

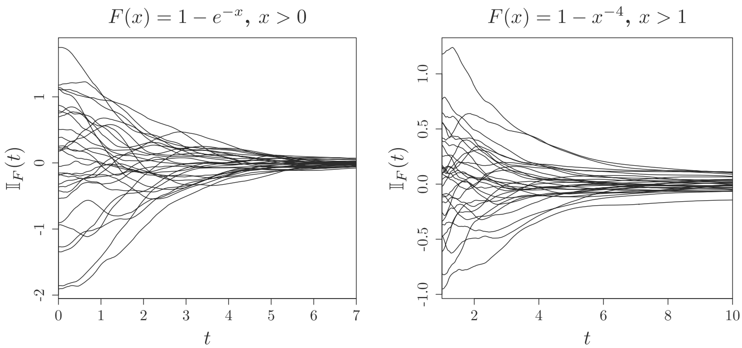

Figure 1. 30 trajectories of $\mathbb{I}_F$ for two distribution functions.

Using (30) or (31), we can compute $\gamma(s,t)$ in the usual examples. For instance, if X has uniform distribution on the interval [0, 1], we have that

$$\gamma(s,t) = \frac{1}{12}(1 - s \vee t)^2\big[1 + 2(s \vee t) - 3(s \wedge t)^2\big], \qquad s,t \in [0,1].$$

In the case that X has the exponential distribution of mean $\mu > 0$, we obtain that

$$\gamma(s,t) = \mu e^{-(s+t)/\mu}\big[e^{(s\wedge t)/\mu}\big(2\mu + |t-s|\big) - \mu\big], \qquad s,t \geq 0.$$

The previous simple computations show that, in general, and as it is expected, the process $\mathbb{I}_F$ has no stationary, neither independent increments.

In Figure 1, we have displayed 30 trajectories of $\mathbb{I}_F$ for two distribution functions. Note that, from Theorems 1 and 2, when $F(x) = 1 - x^{-4}$ $(x \geq 1)$, $\mathbb{I}_n \rightarrow_{\mathrm{w}} \mathbb{I}_F$ in L^p for $1 < p \leq \infty$, but *not* in L^1. In Figure 2, 30 trajectories of $\mathbb{I}_n$ are also displayed for different values of n.

5. An application to estimate Zolotarev metrics

Here we show how Theorems 1 and 2 can be applied to derive the asymptotic distribution of an estimator of the Zolotarev distance between two probability distributions.

Let X and Y be two random variables with distribution functions F and G, respectively. For $r \in \mathbb{N}$, and $1 \leq p \leq \infty$, the *Zolotarev metric* of order r in L^p is defined by

$$\zeta_{r,p}(F,G) := \sup\big\{\big|\mathrm{E}f(X) - \mathrm{E}f(Y)\big| : f \in \mathcal{F}_{r,p}\big\},$$

where $\mathcal{F}_{r,p}$ is the set of functions $f : \mathbb{R} \longrightarrow \mathbb{R}$ for which $f^{(r+1)}$ exists and satisfies $\|f^{(r+1)}\|_q \leq 1$ (q being the conjugate of p, that is, q is such that $1/p + 1/q = 1$). We will indistinctly use the notation $\zeta_{r,p}(X,Y) \equiv \zeta_{r,p}(F,G)$. The metric $\zeta_{r,p}$ is ideal of order $r + 1/p$, that is, for Z independent of (X,Y), it holds that:

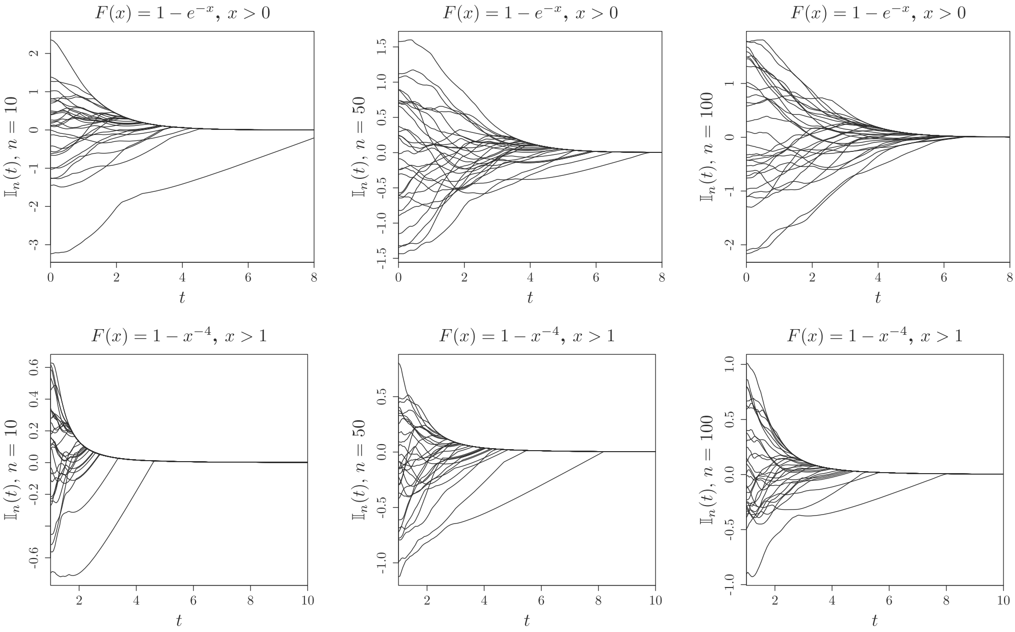

Figure 2. 30 trajectories of $\mathbb{I}_n$ for different values of n and two distribution functions.

(a) $\zeta_{r,p}(X+Z, Y+Z) \leq \zeta_{r,p}(X, Y)$.
(b) $\zeta_{r,p}(cX, cY) = |c|^{r+1/p}\zeta_{r,p}(X, Y)$, for $c \in \mathbb{R}$.

Note that the finiteness of $\zeta_{r,p}(X, Y)$ implies that X and Y have identical moments up to order r. Further, when $\mathrm{E}|X|^{r+1} + \mathrm{E}|Y|^{r+1} < \infty$, we also have that $\zeta_{r,p}(X, Y) < \infty$. For a general reference and properties of $\zeta_{r,p}(X, Y)$, we refer to Rachev *et al.* [26].

The most important case corresponds to $p = 1$. In this situation, $\zeta_{r+1} \equiv \zeta_{r,1}$ is the usual Zolotarev ideal metric of order $r+1$. It can be seen that convergence in ζ_{r+1}-metric implies weak convergence plus convergence of the $(r+1)$th absolute moment. Zolotarev metrics have been used in Rao [27] to obtain a CLT for independent, non-identically distributed random variables. The metric ζ_3 has also been applied in the context of distributional recurrences in Neininger and Rüschendorf [23,24]. Further, several bounds (from above and below) with the Zolotarev metric can be given for other probability metrics such as Kantorovich, Kolmogorov, Wasserstein and Lévy metrics (see Rachev *et al.* [26]).

In the following, we assume that X and Y are positive. Whenever $\zeta_{r,p}(X, Y)$ is finite, it can be expressed as

$$\zeta_{r,p}(X, Y) = \|H_r\|_p, \tag{32}$$

where

$$H_r(t) := \frac{1}{\Gamma(r)} \int_t^\infty (x-t)^{r-1}\big(F(x) - G(x)\big)\,\mathrm{d}x, \qquad t \geq 0. \tag{33}$$

(As mentioned in the Introduction, this quantity also corresponds to the stop-loss distance of order r in L^p.) Therefore, given $X_1,\dots,X_n$ and $Y_1,\dots,Y_m$ two random samples from X and Y, respectively, we consider the plug-in estimator of (32), that is,

$$\hat{\zeta}_{r,p}(X,Y) := \begin{cases} \dfrac{1}{\Gamma(r)}\left(\displaystyle\int_0^\infty \left|\int_t^\infty (x-t)^{r-1}\big(\mathbb{F}_n(x)-\mathbb{G}_m(x)\big)\,\mathrm{d}x\right|^p \mathrm{d}t\right)^{1/p}, & \text{if } 1\le p<\infty, \\ \dfrac{1}{\Gamma(r)}\sup_{t\ge 0}\left|\displaystyle\int_t^\infty (x-t)^{r-1}\big(\mathbb{F}_n(x)-\mathbb{G}_m(x)\big)\,\mathrm{d}x\right|, & \text{if } p=\infty, \end{cases}$$

where $\mathbb{F}_n$ and $\mathbb{G}_m$ are the empirical distribution functions of X and Y, respectively. Note however that in general $\zeta_{r,p}(\mathbb{F}_n,\mathbb{G}_m)=\infty$ a.s., as the corresponding sample moments are different.

We aim to analyze the asymptotic behavior of the normalized estimator of the Zolotarev distance of order r in L^p between X and Y, that is, the statistic given by

$$Z_{n,m}(r,p) := \sqrt{\frac{nm}{n+m}}\big(\hat{\zeta}_{r,p}(X,Y)-\zeta_{r,p}(X,Y)\big). \tag{34}$$

The following representation of $Z_{n,m}(r,p)$ will be crucial:

$$Z_{n,m}(r,p) = \left\|\left(\sqrt{\frac{m}{n+m}}\mathbb{I}_n^{[r]} - \sqrt{\frac{n}{n+m}}\tilde{\mathbb{I}}_m^{[r]}\right) + \sqrt{\frac{nm}{n+m}}H_r\right\|_p - \sqrt{\frac{nm}{n+m}}\|H_r\|_p, \tag{35}$$

where $\mathbb{I}_n^{[r]}$ and $\tilde{\mathbb{I}}_m^{[r]}$ are (independent) r-fold integrated empirical processes (defined in (2)) associated with X and Y, respectively, and H_r is given in (33).

From (35), to derive the asymptotic distribution of $Z_{n,m}(r,p)$, we first need to analyze the continuity of the functional

$$\Delta_p(f,g,\lambda) := \|f+\lambda g\|_p - \lambda\|g\|_p, \qquad 1\le p\le\infty, \tag{36}$$

where $f,g\in L^p$ and $\lambda\to\infty$. As in the previous section, there are substantial differences between the cases $1\le p<\infty$ and $p=\infty$.

In the following lemma, we determine the behavior (as $\lambda\to\infty$) of $\Delta_p(f,g,\lambda)$ in (36) for $1\le p<\infty$. We use the notation $\mathrm{sgn}(\cdot)$ for the sign function and A^c is the complement of the set A.

Lemma 4. *Let $1\le p<\infty$ and consider the functional $\Delta_p : L^p\times L^p\times[0,\infty)\longrightarrow\mathbb{R}$ defined in (36). For $f,g\in L^p$, we have that $\lim_{\lambda\to\infty}\Delta_p(f,g,\lambda)=\Delta_p(f,g)$, with*

$$\Delta_1(f,g) := \int_{I_g}|f| + \int_{I_g^c} f\,\mathrm{sgn}(g), \tag{37}$$

where $I_g := \{t : g(t)=0\}$, and for $1<p<\infty$,

$$\Delta_p(f,g) := \begin{cases} \|f\|_p, & \text{if } g=0 \text{ a.e.}, \\ \dfrac{1}{\|g\|_p^{p-1}}\displaystyle\int f|g|^{p-1}\,\mathrm{sgn}(g), & \text{otherwise}. \end{cases} \tag{38}$$

Moreover, if $f_n \to f$ in L^p and $\lambda_n \to \infty$, then $\Delta_p(f_n, g, \lambda_n) \to \Delta_p(f, g)$.

Proof. Observe that we can always assume that $g \geq 0$. If this is not the case, it is enough to write $g = |g| \operatorname{sgn}(g)$.

We first consider the case $p = 1$. We have that

$$\begin{aligned}\Delta_1(f, g, \lambda) &= \int_{I_g} |f| + \int_{I_g^c} \big(|f + \lambda g| - \lambda g\big) \\ &= \int_{I_g} |f| + \int_{I_g^c \cap \{f+\lambda g \geq 0\}} f - \int_{I_g^c \cap \{f+\lambda g<0\}} (f + 2\lambda g).\end{aligned} \tag{39}$$

By dominated convergence, the second integral in (39) converges to $\int_{I_g^c} f$, as $\lambda \to \infty$. For the third integral in (39), we obtain that

$$\begin{aligned}\int_{I_g^c \cap \{f+\lambda g<0\}} |f + 2\lambda g| &\leq 2 \int_{I_g^c \cap \{f+\lambda g<0\}} |f + \lambda g| + \int_{I_g^c \cap \{f+\lambda g<0\}} |f| \\ &\leq 3 \int_{I_g^c \cap \{f+\lambda g<0\}} |f|.\end{aligned}$$

Hence, by dominated convergence, we conclude that this integral goes to 0, as $\lambda \to \infty$. We therefore obtain that $\Delta_1(f, g, \lambda) \to \Delta_1(f, g)$, when $\lambda \to \infty$.

Now, let us fix $1 < p < \infty$, and we assume that $\|g\|_p \neq 0$. We can write

$$\begin{aligned}\Delta_p(f, g, \lambda) &= \left(\int_{I_g} |f|^p + \int_{I_g^c} (\lambda g)^p \left|1 + \frac{f}{\lambda g}\right|^p\right)^{1/p} - \lambda \|g\|_p \\ &= \left(\int_{I_g} |f|^p + \int_{I_g^c} (\lambda g)^p \left(1 + p\frac{f}{\lambda g} + r_p\left(\frac{f}{\lambda g}\right)\right)\right)^{1/p} - \lambda \|g\|_p,\end{aligned}$$

where

$$r_p(t) := |1 + t|^p - (1 + pt), \qquad t \in \mathbb{R}. \tag{40}$$

We conclude that

$$\Delta_p(f, g, \lambda) = \lambda \|g\|_p \big(1 + \delta_p(f, g, \lambda)\big)^{1/p} - \lambda \|g\|_p, \tag{41}$$

with

$$\delta_p(f, g, \lambda) := \frac{1}{\lambda^p \|g\|_p^p} \left(\int_{I_g} |f|^p + p\lambda^{p-1} \big\| f g^{p-1} \big\|_1 + \lambda^p \int_{I_g^c} g^p r_p\left(\frac{f}{\lambda g}\right)\right). \tag{42}$$

Using the integral form of the remainder in Taylor's theorem, it is easy to check that the function r_p in (40) satisfies the bound

$$\big|r_p(t)\big| \leq \begin{cases} C_p t^2, & \text{if } |t| \leq 1, \\ C_p |t|^p, & \text{if } |t| > 1, \end{cases} \tag{43}$$

where $C_p > 0$ is a constant depending only on p. When $1 \le p \le 2$, from (43), we have that $|r_p(t)| \le C_p|t|^p$ ($t \in \mathbb{R}$). Using this inequality, it can be readily checked that

$$\delta_p(f, g, \lambda) \to 0 \qquad \text{as } \lambda \to \infty. \tag{44}$$

If $2 < p < \infty$, we use (43) and Hölder's inequality to obtain

$$\begin{aligned} \left| \int_{I_g^c} g^p r_p\left(\frac{f}{\lambda g}\right) \right| &\le \int_{I_g^c \cap \{|f/(\lambda g)| \le 1\}} g^p \left| r_p\left(\frac{f}{\lambda g}\right) \right| + \int_{I_g^c \cap \{|f/(\lambda g)| > 1\}} g^p \left| r_p\left(\frac{f}{\lambda g}\right) \right| \\ &\le C_p \left(\frac{1}{\lambda^2} \int_{I_g^c} f^2 g^{p-2} + \frac{1}{\lambda^p} \|f\|_p^p \right) \\ &\le C_p \left(\frac{1}{\lambda^2} \|f\|_p^2 \|g\|_p^{p-2} + \frac{1}{\lambda^p} \|f\|_p^p \right). \end{aligned} \tag{45}$$

From (42) and (45), we conclude that (44) also holds for $2 < p < \infty$. Now, from (41) and (44), for $1 \le p < \infty$, we obtain that

$$\Delta_p(f, g, \lambda) \sim \lambda \|g\|_p \left(1 + \frac{1}{p} \delta_p(f, g, \lambda) \right) - \lambda \|g\|_p, \qquad \lambda \to \infty,$$

where "$\sim$" means asymptotically equivalent. Using again (43), and following a similar reasoning as for the bounds in (45) for the case $2 < p < \infty$, we can show that

$$\Delta_p(f, g, \lambda) \sim \frac{\lambda \|g\|_p}{p} \delta_p(f, g, \lambda) \to \Delta_p(f, g), \qquad \lambda \to \infty.$$

Finally, if $f_n \to f$ in L^p (for $1 \le p < \infty$) and $\lambda_n \to \infty$, by Minkowski's inequality, we have that

$$\begin{aligned} \left| \Delta_p(f_n, g, \lambda_n) - \Delta_p(f, g) \right| &\le \left| \Delta_p(f_n, g, \lambda_n) - \Delta_p(f, g, \lambda_n) \right| + \left| \Delta_p(f, g, \lambda_n) - \Delta_p(f, g) \right| \\ &\le \|f_n - f\|_p + \left| \Delta_p(f, g, \lambda_n) - \Delta_p(f, g) \right|. \end{aligned}$$

Hence, we see that $\Delta_p(f_n, g, \lambda_n) \to \Delta_p(f, g)$, as $n \to \infty$, and the proof is complete. □

Remark 8. It is interesting to note that $\Delta_p(f, g)$ is *not* continuous at $p = 1$, as in general $\Delta_1(f, g) \ne \lim_{p \to 1} \Delta_p(f, g)$. Also, the limit for $p = 2$ is simply

$$\Delta_2(f, g) = \frac{1}{\|g\|_2} \int fg.$$

We now consider the case $p = \infty$. First, it is important to observe that for $f, g \in L^\infty$ the functional $\Delta_\infty(f, g, \lambda)$ in (36) may have an oscillatory character when $\lambda \to \infty$. Hence, there is no hope to obtain an analogous result as Lemma 4 for $p = \infty$. This problem can be dodged by noting that both $\mathbb{I}_n^{[r]}$ and $\mathbb{I}_F^{[r]}$ take values in the C_0, the space of all continuous functions

$f : [0, \infty) \longrightarrow \mathbb{R}$ such that $\lim_{t\to\infty} f(t) = 0$. The behavior of $\Delta_\infty(f, g, \lambda)$ when $f, g \in C_0$ and $\lambda \to \infty$ is explained in the next lemma.

Lemma 5. *Let us consider the space C_0 with the norm $\|\cdot\|_\infty$ and let $\Delta_\infty : C_0 \times C_0 \times [0, \infty) \longrightarrow \mathbb{R}$ be the functional defined in (36). For $f, g \in C_0$, we have that $\lim_{\lambda\to\infty} \Delta_\infty(f, g, \lambda) = \Delta_\infty(f, g)$, with*

$$\Delta_\infty(f, g) := \begin{cases} \|f\|_\infty, & \text{if } g \equiv 0, \\ \sup_{M_g} f \operatorname{sgn}(g), & \text{otherwise}, \end{cases} \tag{46}$$

where $M_g := \{t : |g(t)| = \|g\|_\infty\}$.

Moreover, if $f_n \to f$ in C_0 and $\lambda_n \to \infty$, then $\Delta_\infty(f_n, g, \lambda_n) \to \Delta_\infty(f, g)$.

Proof. We can assume that $\|g\|_\infty > 0$. As $g \in C_0$, $M_g \neq \varnothing$. Further, if $t \notin M_g$, for each $t_0 \in M_g$, we have that $|g(t)| < |g(t_0)|$. Hence, for λ large enough, we obtain that

$$\big|f(t) + \lambda g(t)\big| \leq \|f\|_\infty + \lambda\big|g(t)\big| \leq \big|f(t_0) + \lambda g(t_0)\big|.$$

Consequently, for λ large enough, we conclude that

$$\begin{aligned} \|f + \lambda g\|_\infty &= \sup_{M_g} |f + \lambda g| \\ &= \sup_{M_g} \big|f \operatorname{sgn}(g) + \lambda\|g\|_\infty\big| \\ &= \lambda\|g\|_\infty + \sup_{M_g} f \operatorname{sgn}(g). \end{aligned} \tag{47}$$

Equalities in (47) imply that $\Delta_\infty(f, g)$ in (46) is the limit of $\Delta_\infty(f, g, \lambda)$ (as $\lambda \to \infty$). The last part of the lemma follows by the same argument as in the proof of Lemma 4. Details are omitted. □

In the following result, we recall that the usual convergence in distribution of random variables is denoted by "$\to_{\mathrm{d}}$."

Theorem 3. *Let us consider a sampling scheme such that as $n, m \to \infty$, $n/(n + m) \to \lambda$, with $0 < \lambda < 1$. For $r \geq 1$ and $1 \leq p \leq \infty$, let us assume that $\zeta_{r,p}(X, Y) < \infty$ and $X, Y \in \mathcal{L}^{2(r+1/p),pr+1}$ or $X, Y \in \mathcal{L}^{2r,r}$, according to $1 \leq p < \infty$ or $p = \infty$. For $Z_{n,m}(r, p)$ in (34), we have that*

$$Z_{n,m}(r, p) \to_{\mathrm{d}} \Delta_p\big(\mathbb{D}^{[r]}_{F,G}, H_r\big), \qquad \text{as } n, m \to \infty, \tag{48}$$

where Δ_p is defined in (37), (38) and (46), $\mathbb{D}^{[r]}_{F,G}$ is a centered Gaussian process given by

$$\mathbb{D}^{[r]}_{F,G} := \sqrt{1-\lambda}\,\mathbb{I}^{[r]}_F - \sqrt{\lambda}\,\tilde{\mathbb{I}}^{[r]}_G \tag{49}$$

($\mathbb{I}_F^{[r]}$ *and* $\tilde{\mathbb{I}}_G^{[r]}$ *being two independent* r*-fold integrated Brownian bridges*), *and* H_r *is defined in* (33).

Proof. From (34)–(36), we see that

$$Z_{n,m}(r,p) = \Delta_p\left(\sqrt{\frac{m}{n+m}}\mathbb{I}_n^{[r]} - \sqrt{\frac{n}{n+m}}\tilde{\mathbb{I}}_m^{[r]}, H_r, \sqrt{\frac{nm}{n+m}}\right),$$

where $\mathbb{I}_n^{[r]}$ and $\tilde{\mathbb{I}}_m^{[r]}$ are r-fold integrated empirical processes associated with X and Y, respectively. To obtain the result, it is enough to use Theorems 1 and 2, Lemmas 4 and 5, the independence of the processes $\mathbb{I}_n^{[r]}$ and $\tilde{\mathbb{I}}_m^{[r]}$, and the extended continuous mapping theorem (see van der Vaart and Wellner [32], Theorem 1.11.1). (For $p=\infty$, note that $\mathbb{I}_n^{[r]}$ and $\mathbb{I}_F^{[r]}$ take values in the C_0 and the weak convergence in L^∞ is equivalent to the weak convergence in C_0 (see, for instance, van der Vaart [31], Lemma 18.13).) □

Remark 9. The random variables X and Y have the same distribution if and only if $H_r \equiv 0$, where H_r is defined in (33). In such a case, the limit of $Z_{n,m}(r,p)$ in (34) obtained in the previous theorem is just $\|\mathbb{D}_{F,G}^{[r]}\|_p$. Also, note that the function H_r can be rewritten as

$$H_r(t) = \frac{1}{\Gamma(r+1)}\left(\mathrm{E}(Y-t)_+^r - \mathrm{E}(X-t)_+^r\right), \qquad t \ge 0.$$

Taking into account the limiting distribution in Theorem 3, it can be guessed that a direct approach to obtain the previous asymptotic result could be a difficult task. By contrast, using the convergence of $\mathbb{I}_n^{[r]}$ in L^p, the problem reduces to analyze the continuity of the linking functional.

In the following corollary, we show that in some cases the limit of (48) is actually normally distributed.

Corollary 1. *In the conditions of Theorem* 3, *let us further assume that*:

(i) *If* $p=1$, *the Lebesgue measure of the set* $I_{H_r} = \{t : H_r(t) = 0\}$ *is zero.*
(ii) *If* $1 < p < \infty$, X *and* Y *are* not *equally distributed.*
(iii) *If* $p=\infty$, *the set* $M_{H_r} = \{t : |H_r(t)| = \|H_r\|_\infty\}$ *contains exactly one point.*

We have that, as $n, m \to \infty$, $Z_{n,m}(r,p)$ *in* (34) *converges in distribution to a zero mean normal random variable.*

Proof. Under the assumptions (i) or (ii), from Theorem 3 and for $1 \le p < \infty$, we have that

$$\Delta_p\left(\mathbb{D}_{F,G}^{[r]}, H_r\right) = \frac{1}{\|H_r\|_p^{p-1}} \int \mathbb{D}_{F,G}^{[r]} |H_r|^{p-1}\operatorname{sgn}(H_r),$$

where $\mathbb{D}_{F,G}^{[r]}$ and H_r are given in (49) and (33), respectively. The conclusion follows as $\mathbb{D}_{F,G}^{[r]}$ is a centered Gaussian process and H_r is smooth and nonrandom.

Finally, if $p=\infty$ and $M_{H_r} = \{t_0\}$, observe that $\Delta_\infty(\mathbb{D}_{F,G}^{[r]}, H_r) = \mathbb{D}_{F,G}^{[r]}(t_0)\operatorname{sgn}(H_r(t_0))$ is a zero mean normal variable. □

Remark 10. Corollary 1 reflects a different behavior of the limit in (48) for the cases $p = 1$, $1 < p < \infty$ and $p = \infty$. When $p = 1$ or $p = \infty$, it is possible to obtain a non Gaussian limit if the set I_{H_r} has a strictly positive Lebesgue measure or the set M_{H_r} has more than one point. However, this cannot happen when $1 < p < \infty$ if X and Y do not have the same distribution.

6. Concluding remarks

The results in the previous sections can be connected with equilibrium distributions and stochastic integrals.

Equilibrium distributions: Given a positive random variable X such that $0 < \mathrm{E}X < \infty$, the function $G(t) := \mathrm{E}(X - t)_+/\mathrm{E}X$ $(t \geq 0)$ is the survival function of a probability distribution called the *equilibrium distribution* of X. This distribution is important in renewal theory (see, for instance, Feller [9]) and in reliability theory (see Lai and Xie [18]). Higher order equilibrium distributions can be defined recursively by

$$\bar{F}_1(t) := 1 - F(t), \qquad \bar{F}_r(t) := \frac{1}{\mu_r}\int_t^\infty \bar{F}_{r-1}(x)\,\mathrm{d}x, \qquad t \geq 0, r \geq 2,$$

with $\mu_r := \int_0^\infty \bar{F}_{r-1}(x)\,\mathrm{d}x$. The function $\bar{F}_r$ is the survival function of a random variable and it is called the *rth equilibrium distribution* of X. Observe that when $\mathrm{E}X^{r-1} < \infty$, we have that

$$\bar{F}_r(t) = \frac{\mathrm{E}(X - t)_+^{r-1}}{\mathrm{E}X^{r-1}}, \qquad t \geq 0, r = 2, 3, \ldots.$$

Therefore, for $r \geq 1$, condition (e) in Theorem 1 means that the random variable E_{2r+1} (with the $(2r + 1)$th equilibrium distribution of X), belongs to $\mathcal{L}^{2/p,1}$. For $p = 2$, we obtain that $\mathbb{I}_n^{[r]} \to_{\mathrm{w}} \mathbb{I}_F^{[r]}$ in L^2 if and only if E_{2r+1} is integrable.

Stochastic integrals: Although $\mathbb{I}_F$ is *not* a stochastic integral, it can be expressed in terms of a stochastic integral. First, note that the standard Brownian bridge on [0, 1], $\mathbb{B}$, is a semimartingale with respect to the induced filtration $\mathcal{F}_t$ $(0 \leq t \leq 1)$ (see Kallenberg [17], page 294). As semimartingales are closed under changes of time (see Jacod [16], Corollary 10.12), $\mathbb{B}_F$ is also a semimartingale with respect to $\mathcal{G}_t := \mathcal{F}_{F(t)}$ $(t \geq 0)$ and the integral with respect to $\mathbb{B}_F$ is well defined. Moreover, if F is continuous, then $\mathbb{B}_F$ is a continuous semimartingale and hence Itô's formula is satisfied (see, for instance, Revuz and Yor [28], Theorem 3.3). That is, if $f(s, x)$ is a twice continuously differentiable function, for $t \geq 0$, we have that

$$\begin{aligned} f\bigl(t, \mathbb{B}_F(t)\bigr) = f\bigl(t, \mathbb{B}_F(0)\bigr) &+ \int_0^t \frac{\partial f}{\partial s}\bigl(s, \mathbb{B}_F(s)\bigr)\,\mathrm{d}s + \int_0^t \frac{\partial f}{\partial x}\bigl(s, \mathbb{B}_F(s)\bigr)\,\mathrm{d}\mathbb{B}_F(s) \\ &+ \int_0^t \frac{\partial^2 f}{\partial x^2}\bigl(s, \mathbb{B}_F(s)\bigr)\,\mathrm{d}[\mathbb{B}_F, \mathbb{B}_F](s), \end{aligned}$$

where $[\mathbb{B}_F, \mathbb{B}_F]$ is the quadratic variation process. (If F is not continuous, additional terms appear in the previous formula to take into account the jumps of the process, $\mathbb{B}_F(t) - \mathbb{B}_F(t^-)$.) In particular, by choosing the function $f(s,x)=sx$, we obtain the next integration by parts formula

$$\int_0^t \mathbb{B}_F(x)\,\mathrm{d}x = t\mathbb{B}_F(t) - \int_0^t x\,\mathrm{d}\mathbb{B}_F(x), \qquad t \geq 0.$$

From this equality, we directly obtain that

$$\mathbb{I}_F(t) = -t\mathbb{B}_F(t) - \int_t^\infty x\,\mathrm{d}\mathbb{B}_F(x), \qquad t \geq 0. \tag{50}$$

This representation allows us to apply Theorem 1 to the stochastic integral $\int_t^\infty x\,\mathrm{d}\mathbb{B}_F(x)$. Observe that the mapping $f \mapsto tf(t) + \int_t^\infty f(x)\,\mathrm{d}x$ $(t \geq 0)$ is continuous from $W^{p,1}$ to L^p. Therefore, by (50) and Theorem 1, whenever $X \in \mathcal{L}^{2+2/p,p+1}$ (with $1 \leq p < \infty$), we have that

$$t\mathbb{E}_n(t) + \mathbb{I}_n(t) \to_{\mathrm{w}} -\int_t^\infty x\,\mathrm{d}\mathbb{B}_F(x) \qquad \text{in } L^p.$$

Acknowledgments

I really want to thank the Editor and an Associate Editor for the careful review of the first version of this manuscript. They encouraged the author to extend the earlier result to the whole spectrum of values of $p \in [1,\infty]$. I also thank the reviewers for some valuable suggestions and for pointing to me the references Groeneboom *et al.* [12,13]. I am grateful to various colleagues of my Department of Mathematics: Eugenio Hernández for the reference Bennett and Sharpley [3], Amparo Baíllo for Figures 1–2, and Carlos Mora-Corral and Pedro Balodis for the illuminating conversations on the continuity of the key functional considered in Lemma 4.

Research supported by the Spanish MCyT Grants MTM2011-27248 and MTM2013-44045-P.

References

[1] Albiac, F. and Kalton, N.J. (2006). *Topics in Banach Space Theory*. *Graduate Texts in Mathematics* **233**. New York: Springer. MR2192298

[2] Araujo, A. and Giné, E. (1980). *The Central Limit Theorem for Real and Banach Valued Random Variables*. New York: Wiley. MR0576407

[3] Bennett, C. and Sharpley, R. (1988). *Interpolation of Operators*. *Pure and Applied Mathematics* **129**. Boston, MA: Academic Press. MR0928802

[4] Byczkowski, T. (1976). Gaussian measures on L_p spaces, $0 \leq p < \infty$. *Studia Math.* **59** 249–261. MR0448471

[5] Chen, X. and Li, W.V. (2003). Quadratic functionals and small ball probabilities for the m-fold integrated Brownian motion. *Ann. Probab.* **31** 1052–1077. MR1964958

[6] del Barrio, E., Deheuvels, P. and van de Geer, S. (2007). *Lectures on Empirical Processes*: *Theory and Statistical Applications*. *EMS Series of Lectures in Mathematics*. Zürich: EMS. MR2284824

[7] del Barrio, E., Giné, E. and Matrán, C. (1999). Central limit theorems for the Wasserstein distance between the empirical and the true distributions. *Ann. Probab.* **27** 1009–1071. MR1698999

[8] Denuit, M., Dhaene, J., Goovaerts, M. and Kaas, R. (2005). *Actuarial Theory for Dependent Risks*: *Measures*, *Orders and Models*. Chichester: Wiley.

[9] Feller, W. (1971). *An Introduction to Probability Theory and Its Applications*. *Vol*. *II*, 2nd ed. New York: Wiley. MR0270403

[10] Giné, E. and Nickl, R. (2016). *Mathematical Foundations of Infinite-Dimensional Statistical Models*. Cambridge: Cambridge Univ. Press.

[11] Grafakos, L. (2008). *Classical Fourier Analysis*, 2nd ed. *Graduate Texts in Mathematics* **249**. New York: Springer. MR2445437

[12] Groeneboom, P., Jongbloed, G. and Wellner, J.A. (2001). A canonical process for estimation of convex functions: The "invelope" of integrated Brownian motion $+t^4$. *Ann*. *Statist*. **29** 1620–1652. MR1891741

[13] Groeneboom, P., Jongbloed, G. and Wellner, J.A. (2001). Estimation of a convex function: Characterizations and asymptotic theory. *Ann*. *Statist*. **29** 1653–1698. MR1891742

[14] Henze, N. and Nikitin, Y.Y. (2000). A new approach to goodness-of-fit testing based on the integrated empirical process. *J*. *Nonparametr*. *Stat*. **12** 391–416. MR1760715

[15] Henze, N. and Nikitin, Y.Y. (2003). Two-sample tests based on the integrated empirical process. *Comm*. *Statist*. *Theory Methods* **32** 1767–1788. MR1998981

[16] Jacod, J. (1979). *Calcul Stochastique et Problèmes de Martingales*. *Lecture Notes in Math*. **714**. Berlin: Springer. MR0542115

[17] Kallenberg, O. (1997). *Foundations of Modern Probability*. *Probability and Its Applications* (*New York*). New York: Springer. MR1464694

[18] Lai, C.-D. and Xie, M. (2006). *Stochastic Ageing and Dependence for Reliability*. New York: Springer. MR2223811

[19] Ledoux, M. and Talagrand, M. (2011). *Probability in Banach Spaces*: *Isoperimetry and Processes*. Berlin: Springer.

[20] Levy, H. (2006). *Stochastic Dominance*: *Investment Decision Making Under Uncertainty*, 2nd ed. *Studies in Risk and Uncertainty* **12**. New York: Springer. MR2239375

[21] Marshall, A.W. and Olkin, I. (2007). *Life Distributions*. *Springer Series in Statistics*. New York: Springer. MR2344835

[22] Nazarov, A.I. and Nikitin, Y.Y. (2004). Exact L_2-small ball behavior of integrated Gaussian processes and spectral asymptotics of boundary value problems. *Probab*. *Theory Related Fields* **129** 469–494. MR2078979

[23] Neininger, R. and Rüschendorf, L. (2004). A general limit theorem for recursive algorithms and combinatorial structures. *Ann*. *Appl*. *Probab*. **14** 378–418. MR2023025

[24] Neininger, R. and Rüschendorf, L. (2004). On the contraction method with degenerate limit equation. *Ann*. *Probab*. **32** 2838–2856. MR2078559

[25] Pisier, G. and Zinn, J. (1978). On the limit theorems for random variables with values in L_p $(2 \le p < \infty)$. *Z*. *Wahrsch*. *Verw*. *Gebiete* **41** 289–304. MR0471010

[26] Rachev, S.T., Klebanov, L.B., Stoyanov, S.V. and Fabozzi, F.J. (2013). *The Methods of Distances in the Theory of Probability and Statistics*. New York: Springer. MR3024835

[27] Rao, M.M. (1997). *Real and Stochastic Analysis*. *Probability and Stochastics Series*. Boca Raton, FL: CRC Press. MR1464220

[28] Revuz, D. and Yor, M. (1991). *Continuous Martingales and Brownian Motion*. *Grundlehren der Mathematischen Wissenschaften* **293**. Berlin: Springer. MR1083357

[29] Shaked, M. and Shanthikumar, J.G. (2007). *Stochastic Orders*. *Springer Series in Statistics*. New York: Springer. MR2265633

[30] Shorack, G.R. and Wellner, J.A. (1986). *Empirical Processes with Applications to Statistics*. *Wiley Series in Probability and Mathematical Statistics*: *Probability and Mathematical Statistics*. New York: Wiley. MR0838963

[31] van der Vaart, A.W. (1998). *Asymptotic Statistics*. *Cambridge Series in Statistical and Probabilistic Mathematics* **3**. Cambridge: Cambridge Univ. Press. MR1652247

[32] van der Vaart, A.W. and Wellner, J.A. (1996). *Weak Convergence and Empirical Processes*. *Springer Series in Statistics*. New York: Springer. MR1385671

Received September 2015 and revised March 2016