An elementary proof of the weak convergence of empirical processes

Dragan Radulović

Department of Mathematics, Florida Atlantic University

Marten Wegkamp

Department of Mathematics & Department of Statistical Science, Cornell University

February 2016

Abstract

This paper develops a simple technique for proving the weak convergence of a stochastic process $\bar{\mathbb{Z}}_n(g) := \int g \, \mathrm{d}\mathbb{Z}_n$, indexed by functions g in some class $\mathcal{G}$. The main novelty is a decoupling argument that allows to derive asymptotic equicontinuity of the process $\{\bar{\mathbb{Z}}_n(g), g \in \mathcal{G}\}$ from that of the basic process $\{\mathbb{Z}_n(t),\ t \in \mathbb{R}\}$, with $\mathbb{Z}_n(t) = \bar{\mathbb{Z}}_n(f_t)$ and $f_t(x) = 1_{(-\infty,t]}(x)$. The method leads to novel results for empirical processes based on stationary processes and its bootstrap versions.

Running title: Weak convergence of empirical processes indexed by functions of bounded variation.

MSC2000 Subject classification: Primary 60F17 ; secondary 60G99.

Keywords and phrases: Integration by parts, bounded variation, weak convergence.

1 Introduction

This paper describes a simple, yet very general technique for proving the weak convergence of a stochastic process $\bar{\mathbb{Z}}_n(g) := \int g \, \mathrm{d}\mathbb{Z}_n$, indexed by functions g belonging to some class $\mathcal{G}$. The canonical situation we have in mind is when $\mathbb{Z}_n$ equals the standard empirical process

$$\mathbb{G}_n(t) \quad = \quad \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (1_{(-\infty,t]}(X_i) - F(t)), \quad t \in \mathbb{R} \tag{1}$$

based on a stationary sequence of random variables X_i from a distribution function F.

The traditional proof of weak convergence of the empirical process $\{\int g \, \mathrm{d}\mathbb{G}_n, \, g \in \mathcal{G}\}$ in $\ell^\infty(\mathcal{G})$ based on i.i.d. observations $X_1, X_2, \ldots$ uses maximal inequalities, which in turn rely on bounds involving metric entropy numbers of the class $\mathcal{G}$, see, for instance, Van der Vaart & Wellner (1996, Chapter 2). Instead, we show that stochastic equicontinuity of the process $\{\bar{\mathbb{Z}}_n(g), \, g \in \mathcal{G}\}$, follows from weak convergence of the (arbitrary) process $\{\mathbb{Z}_n(t), \, t \in \mathbb{R}\}$. The heart of our argument is Lemma 3 which decouples the quantity $|\bar{\mathbb{Z}}_n(g)|$ into two parts; one that is controlled by the process $\{\mathbb{Z}_n(t), \, t \in \mathbb{R}\}$, and another, non-probabilistic part, that is bounded by L_1 and total variation norms of the function g. The proof uses a simple integration by parts trick and a finite dimensional approximation technique. It does not involve any metric entropy computations and relies only on elementary measure theory.

The maximal inequalities are well known in the i.i.d. setting, in particular for uniformly bounded classes of functions of bounded variation, see, for instance, Van der Vaart & Wellner (1996, p.149, Example 2.6.21 or p.159, Theorem 2.7.5). Thus, for standard empirical processes $\mathbb{G}_n$ based on i.i.d. variables, the novelty of our results are purely aesthetic in nature. However, the results presented in this paper are far more general and they apply to a variety of interesting and novel situations. In particular, we are able to greatly enhance the existing results on stationary dependent sequences and the bootstrap. Only the case of the rather restrictive beta-mixing is known to give comparable results to the i.i.d. case, partly due to the fact that good maximal inequalities for $\bar{\mathbb{Z}}_n$ exist in this case, see, for instance Arcones and Yu (1994), Radulović

(1996) and Rio (1998).
Moreover, the standard argument of enlarging a P-Donsker class (in our case $1_{(-\infty,t]}$ to its convex hull does not seem to apply to stationary sequences. To the best of our knowledge the proof of such an extension (See van der Vaart and Wellner (1996, page 191) works for i.i.d. sequences only.

The paper is organized as follows. Section 2 contains our main results (Lemma 3, Theorem 5) with their short, elementary proofs, followed by a discussion in Section 3. The important extension to distributions in $\mathbb{R}^d$, $d \geq 2$, will be considered separately in another paper.

2 Main result

We use notation

$$||g||_{TV} = \sup_{\Pi} \sum_{x_i \in \Pi} |g(x_i) - g(x_{i-1})|$$

for the total variation norm of a function $g : \mathbb{R} \to \mathbb{R}$. Here the supremum is taken over all countable partitions $\Pi = \{x_1 < x_2 < \ldots\}$ of $\mathbb{R}$. We set

$$BV_T := \{g : \mathbb{R} \to \mathbb{R} : \ \|g\|_{TV} \leq T\}$$

for $T > 0$. Throughout this paper, we make the blanket assumption that $\{\mathbb{Z}_n(t),\ t \in \mathbb{R}\}$ is an arbitrary stochastic process such that

(i) $\lim_{|t|\to\infty} \mathbb{Z}_n(t) = 0$

(ii) the sample paths of $\mathbb{Z}_n$ are right-continuous and of bounded variation.

Clearly, both requirements (i) and (ii) are met for the canonical empirical process $\mathbb{G}_n(t)$ in (1) based on (not necessarily independent) observations from a distribution F.

In this paper, we study the limit distribution of the process

$$\bar{\mathbb{Z}}_n(g) := \int g(x)\, \mathrm{d}\mathbb{Z}_n(x), \quad g \in \mathcal{G}$$

for some class $\mathcal{G} \subseteq BV_T$, for some finite T. Recall that, see, for instance, Van der Vaart & Wellner (1996, Chapter 1.5), the process $\{\bar{\mathbb{Z}}_n(g),\, g \in \mathcal{G}\}$ converges weakly to a tight limit in $\ell^\infty(\mathcal{G})$, provided

(a) the marginals $(\bar{\mathbb{Z}}_n(g_1), \ldots, \bar{\mathbb{Z}}_n(g_k))$ converge weakly for every finite subset $g_1, \ldots, g_k \in \mathcal{G}$, and

(b) there exists a semi-metric ρ on $\mathcal{G}$ such that $(\mathcal{G}, \rho)$ is totally bounded and $\bar{\mathbb{Z}}_n(g)$ is ρ-stochastically equicontinuous, that is,

$$\lim_{\delta \downarrow 0} \limsup_{n \to \infty} \mathbb{P}\left\{ \sup_{\rho(f,g) \leq \delta} |\bar{\mathbb{Z}}_n(f) - \bar{\mathbb{Z}}_n(g)| > \varepsilon \right\} = 0$$

for all $\varepsilon > 0$. The supremum is taken over all $f, g \in \mathcal{G}$ with $\rho(f,g) \leq \delta$.

Point (a) will be addressed in Section 2.1, while point (b) will be addressed in Section 2.2. For ρ we will take any $L_p(F_0)$ semi-metric based on an arbitrary cumulative distribution function (c.d.f.) F_0 and we say that $\bar{Z}(f)$ is $L_1(F_0)$-continuous if its sample paths are continuous with respect to the $L_1(F_0)$ semi-norm.

First, we note that we may consider only right continuous functions g. This is a consequence of the following lemma.

Lemma 1. *Let $BV_T' \subset BV_T$ be the class of all right-continuous functions $g : \mathbb{R} \to \mathbb{R}$ with $\|g\|_{TV} \leq T$. We have*

$$\sup_{g \in BV_T} \inf_{h \in BV_T'} \left| \bar{\mathbb{Z}}_n(g) - \bar{\mathbb{Z}}_n(h) \right| \;\leq\; T \sup_x |\mathbb{Z}_n(x) - \mathbb{Z}_n(x^-)|.$$

Proof. Let g be an arbitrary function in BV_T. Recall that a function of bounded variation on the real line only has countably many discontinuities. We denote the points of discontinuity of g by a_i. Let $\bar{g}$ be the right-continuous version of g, that is, $\bar{g}(x) = g(x)$ for all $x \neq a_i$; and $\bar{g}(a_i) = g(a_i^+)$ for all i. Then

$$\begin{aligned} \left| \int g \,\mathrm{d}\mathbb{Z}_n - \int \bar{g} \,\mathrm{d}\mathbb{Z}_n \right| &\leq \sum_i |g(a_i) - \bar{g}(a_i)| |\mathbb{Z}_n(a_i) - \mathbb{Z}_n(a_i^-)| \\ &\leq \|g\|_{TV} \sup_x |\mathbb{Z}_n(x) - \mathbb{Z}_n(x^-)|. \end{aligned}$$

The conclusion follows easily. ∎

In case $\mathbb{Z}_n$ is the ordinary empirical process $\mathbb{G}_n$ in (1), the term on the right in the bound of Lemma 1 is of order $O(Tn^{-1/2})$.

2.1 Finite dimensional convergence

First we address the finite dimensional convergence. In many cases, the limiting distribution is Gaussian, and the weak convergence of $(\bar{\mathbb{Z}}_n(g_1), \ldots, \bar{\mathbb{Z}}_n(g_k))$ for every finite $g_1, \ldots, g_k \in \mathcal{G}$ follows from that of $\bar{\mathbb{Z}}_n(g)$ for each $g \in \mathcal{G}$ by the Cramér-Wold device.

Lemma 2. *Assume that the stochastic process $\mathbb{Z}_n(t)$ converges weakly to a Gaussian process $\mathbb{Z}(t)$. Then, for any right-continuous function $g : \mathbb{R} \to \mathbb{R}$ of bounded variation, $\bar{\mathbb{Z}}_n(g) := \int g \, \mathrm{d}\mathbb{Z}_n$ is well defined, and converges to a normal distribution on $\mathbb{R}$.*

Proof. Let g be an arbitrary right-continuous function of bounded variation. First, we notice that $\int g \, \mathrm{d}\mathbb{Z}_n$ and $\int \mathbb{Z}_n \, \mathrm{d}g$ are indeed well-defined as Lebesgue-Stieltjes integrals. Recall that a function of bounded variation on the real line only has countably many discontinuities. We denote the points of discontinuity of g by a_i. By the integration by parts formula, Lemma 12 in the appendix, we have

$$\int g(x) \, \mathrm{d}\mathbb{Z}_n(x) \quad = \quad T_1(\mathbb{Z}_n) + T_2(\mathbb{Z}_n)$$

with operators $T_1, T_2 : \ell^\infty(\mathbb{R}) \to \mathbb{R}$

$$\begin{aligned} T_1(\mathbb{Z}_n) \quad &:= \quad -\int \mathbb{Z}_n(x) \, \mathrm{d}g(x) \\ T_2(\mathbb{Z}_n) \quad &:= \quad \int\int 1_{x=y} \, \mathrm{d}g(x) \, \mathrm{d}\mathbb{Z}_n(y) \\ &= \quad \sum_i \left(g(a_i) - g(a_i^-)\right) \left(\mathbb{Z}_n(a_i) - \mathbb{Z}_n(a_i^-)\right). \end{aligned}$$

Since g has finite variation, it is bounded and, for $\alpha_i := g(a_i) - g(a_i^-)$, $\sum_i |\alpha_i| < \infty$ and $\sum_i \alpha_i^2 < \infty$. Hence, the operator

$$T_2(\mathbb{Z}_n) = \sum_i \alpha_i \left(\mathbb{Z}_n(a_i) - \mathbb{Z}_n(a_i^-)\right)$$

is linear. We conclude the proof by observing that the linearity of the operators T_1 and T_2 and the weak convergence of $\mathbb{Z}_n$ to a (Gaussian) process $\mathbb{Z}$, ensure that the sequence of random variables $Y_n := \int g \, \mathrm{d}\mathbb{Z}_n = T_1(\mathbb{Z}_n) + T_2(\mathbb{Z}_n)$ converges weakly (to a normal distribution) by the continuous mapping theorem. $\square$

2.2 Asymptotic tightness

We now address the asymptotic equicontinuity of the process $\{\bar{\mathbb{Z}}_n(g),\, g \in \mathcal{G}\}$. Lemma 3 below plays the key role in our analysis. It relates the process $\{\bar{\mathbb{Z}}_n(g),\, g \in \mathcal{G}\}$ directly to the process $\{\mathbb{Z}_n(t),\, t \in \mathbb{R}\}$, using the following quantities:
For any c.d.f. F_0 and any $\beta > 0$, we define

$$\begin{aligned} \Psi_{\beta,F_0}(\mathbb{Z}_n) &:= \sup_{|F_0(s)-F_0(t)|\le\beta} |\mathbb{Z}_n(s) - \mathbb{Z}_n(t)|\,, \\ \eth_{\beta,F_0}(\mathbb{Z}_n) &:= \sup_{F_0(s)-F_0(s^-)>\beta} \left|\mathbb{Z}_n(s) - \mathbb{Z}_n(s^-)\right|. \end{aligned}$$

Clearly, $\eth_{\beta,F_0}(\mathbb{Z}_n) = 0$ for all $\beta > 0$ if F_0 is continuous. In general, for arbitrary F_0, the quantity $\eth_{\beta,F_0}(\mathbb{Z}_n)$ is bounded in probability, for all $\beta > 0$, by the continuous mapping theorem, as long as $\mathbb{Z}_n$ converges weakly. The latter bound, $\eth_{\beta,F_0} = O_p(1)$, suffices in our application of Lemma 3 in Theorem 5 below.

Lemma 3 (Decoupling lemma)**.** *For any distribution function F_0 on $\mathbb{R}$, any right-continuous function $g : \mathbb{R} \to \mathbb{R}$ of bounded variation, and any $\beta > 0$, we have*

$$|\bar{\mathbb{Z}}_n(g)| \;\le\; \left\{2\beta^{-1}||g||_{L_1(F_0)} + 6||g||_{TV}\right\} \Psi_{2\beta,F_0}(\mathbb{Z}_n) + \beta^{-1}\|g\|_{L_1(F_0)} \eth_{\beta,F_0}(\mathbb{Z}_n).$$

Here $\|g\|_{L_1(F_0)} = \int |g|\, dF_0$.

Proof. Without loss of generality we can assume that $||g||_{TV} + ||g||_{L_1(F_0)} < \infty$. Since F_0 is a distribution function, we can construct, for any $0 < \beta < 1$, a finite grid $-\infty = s_0 < s_1 < \cdots < s_{M-1} < s_M < \infty$ such that

$$F_0(s_j) - F_0(s_{j-1}) \ge \beta,\;\; F_0(s_M) < 1 - 2\beta$$

and

$$F_0(s_j^-) - F_0(s_{j-1}) \le 2\beta,$$

leaving the possible jump $F_0(s_j) - F_0(s_j^-)$ unspecified. Based on this grid, we approximate $\mathbb{Z}_n(t)$ by

$$\widetilde{\mathbb{Z}}_n(t) = \sum_{j=1}^{M} \mathbb{Z}_n(s_{j-1}) 1_{[s_{j-1},s_j)}(t)$$

and we set $\widetilde{\mathbb{Z}}_n(\pm\infty) = 0$. We observe that by construction

$$\begin{aligned}
\sup_x \left|\mathbb{Z}_n(x) - \widetilde{\mathbb{Z}}_n(x)\right| &\leq \max_{1\leq j\leq M} \sup_{x\in[s_{j-1},s_j)} |\mathbb{Z}_n(x) - \mathbb{Z}_n(s_{j-1})| + \sup_{x\in[s_M,\infty)} |\mathbb{Z}_n(x)| \\
&\leq \sup_{|F_0(x)-F_0(y)|\leq 2\beta} |\mathbb{Z}_n(x) - \mathbb{Z}_n(y)| \\
&= \Psi_{2\beta,F_0}.
\end{aligned}$$

The process $\widetilde{\mathbb{Z}}_n$ inherits the bounded variation property from $\mathbb{Z}_n$ and

$$\int g(s)\,\mathrm{d}\mathbb{Z}_n(s) = \int g(s)\,\mathrm{d}\widetilde{\mathbb{Z}}_n(s) + \int g(s)\,\mathrm{d}(\mathbb{Z}_n - \widetilde{\mathbb{Z}}_n)(s) \tag{2}$$

is well defined for any function g of bounded variation. Using the integration by parts formula (Lemma 12 in the appendix), we obtain for the last term on the right,

$$\int g(s)\,\mathrm{d}(\mathbb{Z}_n - \widetilde{\mathbb{Z}}_n)(s) = -\int (\mathbb{Z}_n - \widetilde{\mathbb{Z}}_n)(s)\,\mathrm{d}g(s) + \int\int 1_{x=y}\,\mathrm{d}g(x)\,\mathrm{d}(\mathbb{Z}_n - \widetilde{\mathbb{Z}}_n)(y).$$

Since g is of bounded variation, it has countably many discontinuities a_i, and we can write

$$\begin{aligned}
&\left|\int\int 1_{x=y}\,\mathrm{d}g(x)\,d(\mathbb{Z}_n - \widetilde{\mathbb{Z}}_n)(y)\right| \\
&\leq \sum_i \left|g(a_i) - g(a_i^-)\right| \left|(\mathbb{Z}_n - \widetilde{\mathbb{Z}}_n)(a_i) - (\mathbb{Z}_n - \widetilde{\mathbb{Z}}_n)(a_i^-)\right| \\
&\leq \|g\|_{TV} \sup_i \left|(\mathbb{Z}_n - \widetilde{\mathbb{Z}}_n)(a_i)\right| + \|g\|_{TV} \sup_i \left|(\mathbb{Z}_n - \widetilde{\mathbb{Z}}_n)(a_i^-)\right| \\
&\leq 2||g||_{TV} \Psi_{2\beta,F_0}(\mathbb{Z}_n).
\end{aligned}$$

Consequently,

$$\begin{aligned}
\left|\int g(s)\,\mathrm{d}(\mathbb{Z}_n - \widetilde{\mathbb{Z}}_n)(s)\right| &\leq \left|\int (\mathbb{Z}_n - \widetilde{\mathbb{Z}}_n)(s)\,\mathrm{d}g(s)\right| + \left|\int\int 1_{x=y}\,\mathrm{d}g(x)\,\mathrm{d}(\mathbb{Z}_n - \widetilde{\mathbb{Z}}_n)(y)\right| \\
&\leq \left|\int (\mathbb{Z}_n - \widetilde{\mathbb{Z}}_n)(s)\,\mathrm{d}g(s)\right| + 2||g||_{TV} \Psi_{2\beta,F_0}(\mathbb{Z}_n) \\
&\leq 3||g||_{TV} \Psi_{2\beta,F_0}(\mathbb{Z}_n).
\end{aligned} \tag{3}$$

Next, we deal with the finite dimensional approximation

$$\int g \, \mathrm{d}\widetilde{\mathbb{Z}}_n = \sum_{j=1}^{M} g(s_j) \left(\mathbb{Z}_n(s_j) \right) - \mathbb{Z}_n(s_{j-1})) \, .$$

By the triangle inequality, we have

$$\begin{aligned} & \left| \int g \, \mathrm{d}\widetilde{\mathbb{Z}}_n \right| \\ & \leq \left| \sum_{j=1}^{M} g(s_j) \left(\mathbb{Z}_n(s_j^-)) - \mathbb{Z}_n(s_{j-1}) \right) \right| + \left| \sum_{j=1}^{M} g(s_j) \left(\mathbb{Z}_n(s_j)) - \mathbb{Z}_n(s_j^-) \right) \right| \end{aligned} \tag{4}$$

and we address each term on the right separately. For the first term on the right, we introduce the step function

$$g^*(t) = \sum_{j=1}^{M} 1_{(s_{j-1}, s_j]}(t) \inf_{s_{j-1} < s \leq s_j} |g(s)|$$

and we observe that

$$\begin{aligned} \sum_{j=1}^{M} g^*(s_j) & \leq \beta^{-1} \sum_{j=1}^{M} g^*(s_j)(F_0(s_j) - F_0(s_{j-1})) \\ & = \beta^{-1} \int g^* \, \mathrm{dF}_0 \\ & \leq \beta^{-1} \int |g| \, \mathrm{d}F_0 \\ & = \beta^{-1} \|g\|_{L_1(F_0)}. \end{aligned}$$

Hence, we have

$$\begin{aligned} \left| \sum_{j=1}^{M} g(s_j) \left(\mathbb{Z}_n(s_j^-)) - \mathbb{Z}_n(s_{j-1}) \right) \right| & \leq \Psi_{2\beta, F_0}(\mathbb{Z}_n) \left\{ \sum_{j=1}^{M} (|g(s_j)| - g^*(s_j)) + \sum_{j=1}^{M} g^*(s_j) \right\} \\ & \leq \Psi_{2\beta, F_0}(\mathbb{Z}_n) \left(||g||_{TV} + \beta^{-1} ||g||_{L_1(F_0)} \right) . \end{aligned} \tag{5}$$

For the second term in (4), we have

$$\begin{aligned}
&\left|\sum_{j=1}^{M} g(s_j)\left(\mathbb{Z}_n(s_j)) - \mathbb{Z}_n(s_j^-)\right)\right| \\
&\leq \left|\sum_{j:\ F_0(s_j)-F_0(s_j^-)\leq\beta} g(s_j)\left(\mathbb{Z}_n(s_j)) - \mathbb{Z}_n(s_j^-)\right)\right| + \left|\sum_{j:\ F_0(s_j)-F_0(s_j^-)>\beta} g(s_j)\left(\mathbb{Z}_n(s_j)) - \mathbb{Z}_n(s_j^-)\right)\right| \\
&\leq \Psi_{2\beta,F_0}(\mathbb{Z}_n)\left(2||g||_{TV} + \beta^{-1}||g||_{L_1(F_0)}\right) + \eth_{\beta,F_0}(\mathbb{Z}_n)\beta^{-1}\|g\|_{L_1(F_0)}, \qquad (6)
\end{aligned}$$

using for the last inequality

$$\sum_{j=1}^{M} |g(s_j)| \leq 2\|g\|_{TV} + \beta^{-1}\|g\|_{L_1(F_0)}$$

(by repeating the same computation in (5) before), and

$$\begin{aligned}
\sum_{j=1}^{M} |g(s_j)| 1\{F_0(s_j) - F_0(s_j^-) > \beta\} &\leq \beta^{-1}\sum_{j=1}^{M} |g(s_j)||(F_0(s_j) - F_0(s_j^-)| \\
&\leq \beta^{-1}\|g\|_{L_1(F_0)}.
\end{aligned}$$

The proof now follows easily after collecting all the bounds (2), (3), (4), (5) and (6). □

An immediate corollary is the following result:

Corollary 4. *For any c.d.f. F_0, and for all $T < \infty$, $\delta > 0$ and $p \geq 1$, we have*

$$\sup_{\|g\|_{L_p(F_0)}\leq\delta,\ \|g\|_{TV}\leq T} \left|\int g\,\mathrm{d}\mathbb{Z}_n\right| \leq (2\sqrt{\delta} + 6T)\Psi_{2\sqrt{\delta},F_0}(\mathbb{Z}_n) + \eth_{\sqrt{\delta},F_0}(\mathbb{Z}_n)\sqrt{\delta},$$

where $\|g\|_{L_p(F_0)} = \left(\int |g|^p \mathrm{d}F_0\right)^{1/p}$.

Proof. The proof follows trivially from Lemma 3 taking $\beta = \sqrt{\delta}$ and using $\|g\|_{L_1(F_0)} \leq \|g\|_{L_p(F_0)} \leq \delta$ for all $p \geq 1$. □

2.3 Main result

Throughout the papr we view $\mathbb{Z}_n$ as a random element in $\ell^\infty(\mathbb{R})$ and its weak convergence to a tight Gaussian process is understood in the sense of Hoffmann-Jorgensen (see Van der Vaart and Wellner, 1996).

Theorem 5. *Assume that the stochastic process* $\{\mathbb{Z}_n(t),\ t \in \mathbb{R}\}$ *converges weakly to a Gaussian process* $\{\mathbb{Z}(t),\ t \in \mathbb{R}\}$, *that is continuous with respect to the distance* $d(s,t) = |F_0(s) - F_0(t)|$ *for some c.d.f.* F_0. *Then, for any* $T < \infty$ *and* $\mathcal{G} \subseteq BV_T'$, *the process* $\{\bar{\mathbb{Z}}_n(g),\ g \in \mathcal{G}\}$ *converges weakly to a* $L_1(F_0)$*-continuous Gaussian process in* $\ell^\infty(\mathcal{G})$.

Proof. The finite dimensional convergence follows trivially from Lemma 2, invoking the weak convergence of $\mathbb{Z}_n(t)$ to a Gaussian limit. As for the stochastic equicontinuity of $\bar{\mathbb{Z}}_n(g)$, it suffices to show that, for all $\varepsilon > 0$,

$$\lim_{\delta \downarrow 0} \limsup_{n \to \infty} \mathbb{P} \left\{ \sup_{\|h\|_{L_1(F_0)} \leq \delta} \left| \int h \, \mathrm{d}\mathbb{Z}_n \right| > \varepsilon \right\} = 0.$$

Here the supremum is taken over all differences $h = g - g'$ with $g, g' \in \mathcal{G}$ and $\|h\|_{L_1(F_0)} = \int |h| \, \mathrm{d}F_0 \leq \delta$. Since $\|h\|_{TV} \leq 2T$, we find by Corollary 4 that

$$\sup_{\|h\|_{L_1(F_0)} \leq \delta} \left| \int h \, d\mathbb{Z}_n \right| \quad \leq \quad \Psi_{2\sqrt{\delta}, F_0}(\mathbb{Z}_n)(12T + 2\sqrt{\delta}) + \eth_{\sqrt{\delta}, F_0}(\mathbb{Z}_n)\sqrt{\delta}.$$

Let $f_t(x) = 1\{x \leq t\}$, so that $\mathbb{Z}_n(t) = \bar{\mathbb{Z}}_n(f_t)$ and

$$d(s,t) = |F_0(s) - F_0(t)| = \int |f_s - f_t| \, \mathrm{d}F_0,$$

and observe that

$$\sup_{d(s,t) \leq \delta} |\mathbb{Z}_n(s) - \mathbb{Z}_n(t)| = \Psi_{\delta, F_0}(\mathbb{Z}_n).$$

The weak convergence of $\{\mathbb{Z}_n(t), t \in \mathbb{R}\}$ (or, equivalently, $\bar{\mathbb{Z}}_n(f_t)$) to a continuous (with respect to d) process $\mathbb{Z}(t)$ implies that

$$\Psi_{2\sqrt{\delta}, F_0}(\mathbb{Z}_n) \xrightarrow{P} 0 \quad \text{as } \delta \to 0 \text{ and } n \to \infty.$$

Moreover, the weak convergence of $\{\mathbb{Z}_n(t),\ t \in \mathbb{R}\}$ implies that $\eth_{\delta,F_0}(\mathbb{Z}_n)$ is bounded in probability, so that $\eth_{\sqrt{\delta},F_0}(\mathbb{Z}_n)\sqrt{\delta}\xrightarrow{P}0$ as $\delta \to 0$ and $n \to \infty$.

Summarizing, the process $\bar{\mathbb{Z}}_n(g)$ converges for each $g \in \mathcal{G}$ to a Gaussian limit, $\bar{\mathbb{Z}}_n(g)$ is uniformly $L_1(F_0)$-equicontinuous, in probability, and $(\mathcal{G}, L_1(F_0))$ is totally bounded (this is true for any c.d.f. F_0). This implies the weak convergence of $\bar{\mathbb{Z}}_n(g)$ to a $L_1(F_0)$-continuous Gaussian process, see Theorems 1.5.4 and 1.5.7 in Van der Vaart & Wellner (1996). □

3 Discussion

3.1 Alternative proof if $F_0(t) = t$

The proof of Theorem 5 can be shortened considerably if the limit $\mathbb{Z}$(t) of $\mathbb{Z}_n(t)$ is continuous. Let d_{BL} be the bounded Lipschitz metric that metrics weak convergence, see, e.g., Van der Vaart & Wellner (1996, page 73) for the definition. Hence, $d_{BL}(\mathbb{Z}_n, \mathbb{Z}) \to 0$ as $n \to \infty$. Set $\bar{\mathbb{Z}}_n(g) = \int g \,\mathrm{d}\mathbb{Z}_n$ for any $g \in BV_T$. By Lemma 1 and the fact that $\mathbb{Z}$ is continuous, we only need to prove weak convergence in $\ell^\infty(\mathcal{G})$ with $\mathcal{G} \subseteq BV_T'$. For this class, the Lebesgue Stieltjes integrals

$$\widetilde{\mathbb{Z}}_n(g) = -\int \mathbb{Z}_n \,\mathrm{d}g$$

are well defined. Next, by the integration by parts formula in Lemma 12, we have

$$\bar{\mathbb{Z}}_n(g) = \tilde{\mathbb{Z}}_n(g) + R_n(g)$$

with

$$R_n(g) \quad \le \quad T \sup_x |\mathbb{Z}_n(x) - \mathbb{Z}_n(x^-)| \to 0,$$

in probability, as $n \to \infty$, as $\mathbb{Z}$ is continuous. For fixed g, $\tilde{\mathbb{Z}}_n(g)$ converges weakly to $\tilde{\mathbb{Z}}(g) := -\int \mathbb{Z} \,\mathrm{d}g$ by the continuous mapping theorem and weak convergence of $\mathbb{Z}_n$. The continuous mapping theorem also guarantees that the limit $\tilde{\mathbb{Z}}$ is tight in $\ell^\infty(\mathcal{G})$ as the map $\Gamma : \ell^\infty(\mathbb{R}) \to \ell^\infty(\mathcal{G})$ defined as $\Gamma_f(X) = -\int X \,\mathrm{d}g$, $g \in \mathcal{G}$, is continuous. By the triangle inequality,

$$
\begin{aligned}
d_{BL}(\bar{\mathbb{Z}}_n, \tilde{\mathbb{Z}}) &\le d_{BL}(\bar{\mathbb{Z}}_n, \tilde{\mathbb{Z}}_n) + d_{BL}(\tilde{\mathbb{Z}}_n, \tilde{\mathbb{Z}}) \\
&= \sup_{H \in BL_1} |\mathbb{E}H(\bar{\mathbb{Z}}_n) - \mathbb{E}H(\tilde{\mathbb{Z}}_n)| + \sup_{H \in BL_1} |\mathbb{E}H(\tilde{\mathbb{Z}}_n) - \mathbb{E}H(\tilde{\mathbb{Z}})| \\
&\le T\mathbb{E}[\sup_x |\mathbb{Z}_n(x) - \mathbb{Z}_n(x^-)|] + T \sup_{H' \in BL_1} |\mathbb{E}H'(\mathbb{Z}_n) - \mathbb{E}H'(\mathbb{Z})| \\
&= T\mathbb{E}[\sup_x |\mathbb{Z}_n(x) - \mathbb{Z}_n(x^-)|] + T d_{BL}(\mathbb{Z}_n, \mathbb{Z})
\end{aligned}
$$

The second term follows since the map $\Gamma_f(X) := \int X \ \mathrm{d}f$ is Lipschitz with Lipschitz constant $\int |\mathrm{d}f| \le T$ and the suprema are taken over all Lipschitz functionals $H : \ell^\infty(\mathcal{G}) \to \mathbb{R}$ with $\|H\|_\infty \le 1$ and $|H(X) - H(Y)| \le \|X - Y\|_\infty$ and $H' : \ell^\infty(\mathbb{R}) \to \mathbb{R}$ with $\|H'\|_\infty \le 1$ and $|H'(X) - H'(Y)| \le \|X - Y\|_\infty$, respectively Together with the tightness of the limit $\tilde{\mathbb{Z}}$, this implies that the empirical process $\bar{\mathbb{Z}}_n(g)$ indexed by $g \in \mathcal{G} \subset BV_T$ converges weakly.

3.2 Standard empirical processes based on i.i.d. sequences

Although there are examples of Donsker classes of infinite variation, such as $f : [0,1] \to [0,1]$ with $|f(x) - f(y)| \le |x - y|^\alpha$, $1/2 < \alpha < 1$, such cases are rather the exception than the norm. The majority of examples of bounded Donsker classes that are given in the literature are subsets of BV_T. Nevertheless, since any function of bounded variation can be written as a difference of two monotone functions, Theorem 5, for $\mathbb{Z}_n(t)$ equal to the standard empirical process $\mathbb{G}_n(t)$ in (1), based on i.i.d. $X_1, X_2, \cdots$, is certainly not novel, except for its mathematical proof. For stationary, dependent sequences X_i, the situation is radically different.

3.3 Stationary sequences

Theorem 5 allows us to argue weak convergence of $\bar{\mathbb{Z}}_n(g)$ via $\mathbb{Z}_n(t)$, regardless of the structure of the latter process. For instance, taking $\mathbb{Z}_n$ as the standard empirical processes $\mathbb{G}_n$ based on stationary sequences X_i, we obtain the following corollary as an immediate consequence of Theorem 5.

Corollary 6. *Let X_k be a stationary sequence of random variables with distiribution F and alpha-mixing coefficients α_n satisfying $\alpha_n = O(n^{-r})$, $n \geq 1$, for some $r > 1$. Then, for any $\mathcal{G} \subseteq BV_T'$, the empirical process $\{\int g \, \mathrm{d}\mathbb{G}_n,\, g \in \mathcal{G}\}$ converges weakly to a $L_1(F)$-continuous Gaussian process in $\ell^\infty(\mathcal{G})$.*

Proof. It is well known, see Theorem 7.2, page 96 in Rio (2000), that $\alpha_n = O(n^{-r})$ implies that the standard empirical process $\mathbb{G}_n$ converges weakly to a Brownian bridge process with continuous paths with respect to the distance $d(s,t) = |F(s) - F(t)|$ for the stationary distribution F of X_k. We apply Theorem 5 to obtain the result. □

The limit $\bar{\mathbb{Z}}$ of $\mathbb{Z}_n$ in Corollary 6 is a mean zero Gaussian process with covariance structure

$$\begin{aligned}\mathbb{E}[\bar{\mathbb{Z}}(f)\bar{\mathbb{Z}}(g)] &= \mathrm{Cov}(f(X_0), g(X_0)) + \sum_{k=1}^{\infty} \mathrm{Cov}(f(X_0), g(X_k)) \\ &\quad + \sum_{k=1}^{\infty} \mathrm{Cov}(f(X_k), g(X_0)).\end{aligned}$$

We would like to point out that alpha-mixing is the least restrictive form of available mixing assumptions in the literature. To the best of our knowledge, there are actually very few results that treat processes $\{\int g \, \mathrm{d}\mathbb{G}_n,\, g \in \mathcal{G}\}$, indexed by functions, and they all require very stringent conditions on the entropy numbers of $\mathcal{G}$ and on the rate of decay for α_k. See, for instance, Andrews and Pollard (1994). This is due to fact that alpha-mixing does not allow for sharp exponential inequalities for partial sums. Consequently, the only known cases for which we have sharp conditions are under more restrictive, beta-mixing dependence. Beta-mixing allows for decoupling and it does yield exponential inequalities not unlike the i.i.d. case. The current state-of-the art results, Arcones & Yu (1994), Doukhan, Massart & Rio (1995), applied to bounded sequences, require $\sum_n \beta_n < \infty$.

However, Theorem 5 goes beyond dependence defined via mixing conditions. For example, they allow for short memory casual linear sequences. These sequences are defined by

$$X_i = \sum_{j=0}^{\infty} a_j \xi_{i-j}$$

based on i.i.d. random variables ξ_i and constants a_i. While the X_i form a stationary sequence, they do not necessarily satisfy any mixing condition. Weak convergence of the empirical processes $\{\mathbb{G}_n(t),\, t \in \mathbb{R}\}$ was established under sharp conditions (Doukhan and Surgailis, 1998). To the best of our knowledge, there are no extensions to the more general processes $\bar{\mathbb{Z}}_n(g)$. Theorem 5 and the Doukhan and Surgailis (1998) result combined imply the following:

Corollary 7. *Let $X_i = \sum_{j \geq 0} a_j \xi_{i-j}$ be such that conditions of Doukhan and Surgailis (1998, pp 87–88) are satisfied and let F be the stationary disitribution of X_i. Then, for any $\mathcal{G} \subseteq BV_T$, the empirical process $\bar{\mathbb{Z}}_n = \{\int g \, \mathrm{d}\mathbb{G}_n,\ g \in \mathcal{G}\}$ converges weakly to a $L_1(F)$-continuous Gaussian process in $\ell^\infty(\mathcal{G})$.*

Proof. Doukhan and Surgailis (1998) prove that $\mathbb{G}_n(t)$ converges weakly to a Gaussian process in the Skorohod space. Since F is continuous under their assumptions, see Doukhan and Surgailis (1998, p 88), the limiting process of $\mathbb{G}_n$ is continuous with respect to the distance $d(s,t) = |F(s) - F(t)|$. The result follows from Theorem 5. □

The recent papers by Dehling et al. (2009 and 2014) offer yet another, clever way to prove the weak limit of the standard empirical processes $\mathbb{G}_n$ based a stationary sequences that are not necessarily mixing. Their technique uses finite dimensional convergence coupled with a bound on the higher moments of partial sums, which in turn controls the dependence structure. Dehling et al. (2009) establishes the weak convergence of $\mathbb{G}_n$, while Dehling et al. (2014) extends this idea to more general classes of functions. However, the authors impose cumbersome entropy conditions and only manage to marginally extend the classes. For example, they manage to prove weak convergence of the process $\int f_t \, \mathrm{d}\mathbb{G}_n$, indexed by functions $f_t(x)$ in the one-dimensional monotone class (with restrictive requirement that $s \leq t \Rightarrow f_s \leq f_t$). Theorem 5 applied in their setting, yields a more general result.

Corollary 8. *Let $\mathcal{G} \subseteq BV_T$. Under assumptions (i) and (ii) in Section 1 of Dehling et al. (2009), the empirical process $\{\int g \, \mathrm{d}\mathbb{G}_n,\ g \in \mathcal{G}\}$ converges weakly to a Gaussian process in $\ell^\infty(\mathcal{G})$.*

Proof. The underlying distribution function F of X_i in Dehling et al (2009) is continuous, see their display (3) at p 3702. Consequently, the limiting process of $\mathbb{G}_n$ is continuous with respect to the distance $d(s,t) = |F(s) - F(t)|$ and Theorem 5 yields the result. $\square$

3.4 Bootstrap

Given the sample $X_1, \ldots, X_n$, we let $\mathbb{G}_n^*$ be bootstrap empirical process

$$\mathbb{G}_n^*(t) \quad = \quad \sqrt{m_n}\left(\frac{1}{m_n}\sum_{i=1}^{m_n} 1_{(-\infty,t]}(X_i^*) - \frac{1}{n}\sum_{i=1}^{n} 1_{(-\infty,t]}(X_i)\right), \ t \in \mathbb{R},$$

based on a bootstrap sample $X_1^*, \ldots, X_{m_n}^*$. We stress that no additional assumption on the structure of the variables $X_{i,n}^*$ is required. Analogous to $\bar{\mathbb{Z}}_n(g) = \int g \, \mathrm{d}\mathbb{G}_n$, we define $\bar{\mathbb{Z}}_n^*(g) := \int g \, \mathrm{d}\mathbb{G}_n^*$ for any $g \in BV_T$ with $T < \infty$. Recall that the bounded Lipschitz distance

$$d_{BL}(\bar{\mathbb{Z}}_n^*, \bar{\mathbb{Z}}) = \sup_{h \in BL_1} \left|\mathbb{E}^*[h(\bar{\mathbb{Z}}_n^*)] - \mathbb{E}[h(\bar{\mathbb{Z}})]\right|$$

between two processes $\bar{\mathbb{Z}}_n^*$ and $\bar{\mathbb{Z}}$ metrizes weak convergence. Here $\mathbb{E}^*$ is the expectation over the randomness of the bootstrap sample $X_1^*, \ldots, X_{m_n}^*$, conditionally given the original sample $X_1, \ldots, X_n$, and BL_1 is the space of Lipschitz functionals $h : \ell^\infty(\mathcal{G}) \to \mathbb{R}$ with $|h(x)| \leq 1$ and $|h(x) - h(y)| \leq \|x - y\|_\infty$ for all $x, y \in \mathcal{G} \subset BV_T$. Customary in the literature, if the random variable $d_{BL}(\bar{\mathbb{Z}}_n^*, \bar{\mathbb{Z}})$ converges to zero in probability, we speak of weak convergence in probability; if it converges to zero almost surely, we speak of weak convergence almost surely.

Theorem 9. *Let $\mathcal{G} \subseteq BV_T'$. Assume that, conditionally on $X_1, \ldots, X_n$, in probability, $\{\mathbb{G}_n^*(t),\ t \in \mathbb{R}\}$ converges weakly to a Gaussian process that is continuous with respect to the the distance $d(s,t) = |F(s) - F(t)|$ based on the stationary distribution F of X_i. Then, the process $\bar{\mathbb{Z}}_n^* = \{\int g \, \mathrm{d}\mathbb{G}_n^*,\ g \in \mathcal{G}\}$ converges weakly to a $L_1(F)$-continuous Gaussian process in $\ell^\infty(\mathcal{G})$.*
If the weak convergence of $\mathbb{G}_n^*$ holds almost surely, then the conclusion that $\bar{\mathbb{Z}}_n^*$ converges weakly in $\ell^\infty(\mathcal{G})$ holds almost surely.

Proof. We only prove the "in probability" statement. The almost sure statement follows after straightforward changes in the proof. A simple modification of Lemma 2 yields

$$\int g(t)d\mathbb{G}_n^*(t) = T_1(\mathbb{G}_n^*) + T_2(\mathbb{G}_n^*)$$

for the same operators T_1 and T_2 defined in the proof of Lemma 1. Since $\mathbb{G}_n^*$ converges to a Gaussian process, the finite dimensional convergence of $\bar{\mathbb{Z}}_n^*$ follows.

As for stochastic equicontinuity of $\bar{\mathbb{Z}}_n^*$, we find, analogous to Corollary 4, that, for all $\delta > 0$,

$$\sup_{\|h\|_{TV}\leq 2T,\ \|h\|_{L_1(F)}\leq\delta} \left|\int h\, d\mathbb{G}_n^*\right| \leq \Psi_{2\sqrt{\delta},F}(\mathbb{G}_n^*)(12T + 2\sqrt{\delta}) + \eth_{\sqrt{\delta},F}(\mathbb{G}_n^*)\sqrt{\delta}.$$

Now use the weak convergence of $\mathbb{G}_n^*$, as in the proof of Theorem 5, to conclude the result. □

If $\mathbb{G}_n$ and $\mathbb{G}_n^*$ converge to the same limit, then $\bar{\mathbb{Z}}_n$ and $\bar{\mathbb{Z}}_n^*$ converge to the same limit in $\ell^\infty(\mathcal{G})$. More precisely, if $d_{BL}(\bar{\mathbb{G}}_n^*, \bar{\mathbb{G}}_n) \to 0$, in probability (or almost surely), then the bootstrap works in that

$$d_{BL}(\bar{\mathbb{Z}}_n^*, \bar{\mathbb{Z}}_n) \to 0$$

in probability (or almost surely).

The literature offers numerous bootstrapping techniques for stationary data (moving block bootstrap, stationary bootstrap, sieved bootstrap, Markov chain bootstrap, etc.), but, their validity is proved for specific cases/statistics only. Due to complications with entropy calculations for dependent triangular arrays, almost all results are done for standard empirical processes $\mathbb{G}_n^*$ with few notable exceptions. The moving block bootstrap was justified for VC-type classes, but only under rather restrictive beta-mixing conditions on X_i (Radulović, 1996). Bracketing classes were considered by Bühlmann (1995), but his conditions are even more restrictive.
In contrast, the process $\{\mathbb{G}_n(t),\, t \in \mathbb{R}\}$ is rather easy to bootstrap. This coupled with Theorem 9 offers the following result.

Corollary 10. *Let X_j be a stationary sequence of random variables with continuous stationary c.d.f. F and alpha-mixing coefficients satisfying $\sum_{k\geq n}\alpha_k = O(n^{-\gamma})$, for some $0<\gamma<1/3$. Let $\mathbb{G}_n^*$ be the bootstrapped standard empirical process based on the moving block bootstrap, with block sizes b_n, specified in Peligrad (1998, p 882). Then, for $\mathcal{G}\subset BV_T$, the bootstrap empirical process $\bar{\mathbb{Z}}_n^* = \{\int g\,\mathrm{d}\mathbb{G}_n^*,\ g\in\mathcal{G}\}$ converges weakly to a $L_1(F)$-continuous Gaussian process, almost surely.*

Proof. Theorem 2.3 of Peligrad (1998) establishes the convergence of $\mathbb{G}_n^*$ to a Gaussian process that is continuous with respect to the distance $d(s,t) = |F(s) - F(t)|$ and Theorem 9 extends it to $\bar{\mathbb{Z}}_n^*$. □

Just as with the weak convergence of the empirical process based on stationary sequences, there are numerous results that treat stationary bootstrap for non-mixing sequences. For example, Ktaibi et al. (2014) study short memory casual linear sequences, and prove weak convergence of $\mathbb{G}_n^*$ under conditions akin to the ones required for its non-bootstrap counterpart $\mathbb{G}_n$ (Doukhan and Surgailis, 1998). Again, Theorem 9 easily extends this result.

Corollary 11. *Let $X_i = \sum_{j\geq 0} a_j\xi_{i-j}$ be a sequence of random variables with stationary distribution F such that conditions of Ktaibi et al (2014) are satisfied. Then, for any $\mathcal{G}\subset BV_T$, the process $\mathbb{Z}_n^* = \{\int g\,\mathrm{d}\mathbb{G}_n^*,\ g\in\mathcal{G}\}$ converges weakly a.s. to a $L_1(F)$-continuous Gaussian limit.*

Proof. Ktaibi et al (2014) prove the convergence of $\mathbb{G}_n^*$ in the Skorohod space, for continuous F, which in turn implies that the limiting process $Z(t)$ is continuous with respect to $d(s,t) = |F(s) - F(t)|$. Theorem 9 extends it to $\bar{\mathbb{Z}}_n^*$. □

A Integration by parts formula

For completeness, we state the following classical result and give a simple elementary proof which was communicated to us by David Pollard.

Lemma 12 (Integration by parts)**.** *Let f and g be right-continuous functions of bounded variation and define measures μ and v as $\mu(-\infty, x] = f(x) - f(-\infty)$ and $\nu(-\infty, y] = g(y) - g(-\infty)$. Then*

$$\int f(x)\,\mathrm{d}g(x) + \int g(x)\,\mathrm{d}f(x) \;=\; (fg)(\infty) - (fg)(-\infty) + \int\int 1_{x=y} d\mu\,\mathrm{d}\nu..$$

Moreover, if either $f(\pm\infty) = 0$ of $g(\pm\infty) = 0$, then

$$\int f(x)\,\mathrm{d}g(x) + \int g(x)\,\mathrm{d}f(x) \;=\; \int\int 1_{x=y}\,\mathrm{d}\mu\,\mathrm{d}\nu.$$

Proof. Set $H(x, y) = 1\{x \le y\}$ and observe that by the very definition of Lebesgue integral

$$f(y) \;=\; \int H(x, y)\,\mathrm{d}\mu(x) + f(-\infty)$$

and

$$g(x) \;=\; \int H(y, x)\,\mathrm{d}\nu(y) + g(-\infty).$$

Hence

$$\begin{aligned}
&\int f(y)\,\mathrm{d}\nu(y) + \int g(x)\,\mathrm{d}\mu(x) \\
&= \int \Big(\int H(x, y)\,\mathrm{d}\mu(x)\Big)\,\mathrm{d}\nu(y) + \int \Big(\int H(y, x)\,\mathrm{d}\nu(y)\Big)\,\mathrm{d}\mu(x) \\
&\quad + f(-\infty)(g(\infty) - g(-\infty)) + g(-\infty)(f(\infty) - f(-\infty))
\end{aligned}$$

Next we apply Fubini

$$\begin{aligned}
&\int \Big(\int H(x, y)\,\mathrm{d}\mu(x)\Big)\,\mathrm{d}\nu(y) + \int \Big(\int H(y, x)\,\mathrm{d}v(y)\Big)\,\mathrm{d}\mu(x) \\
&= \int\int (H(x, y) + H(y, x))\,\mathrm{d}\mu(x)\,\mathrm{d}\nu(y) = \int\int (1_{x\le y} + 1_{y\le x})\,\mathrm{d}\mu(x)\,\mathrm{d}\nu(y) \\
&= \int\int \mathrm{d}\mu(x)\,\mathrm{d}\nu(y) + \int\int 1_{x=y}\,\mathrm{d}\mu(x)\,\mathrm{d}\nu(y) \\
&= \{f(\infty) - f(-\infty)\}\{g(\infty) - g(-\infty)\} + \int\int 1_{x=y}\,\mathrm{d}\mu(x)\,\mathrm{d}\nu(y).
\end{aligned}$$

Combine the two previous displays to prove the lemma. □

Acknowledgement: We thank David Pollard for the elegant proof of integration by parts.

References

[1] D.W.K. Andrews and D.P. Pollard (1994). An Introduction to Functional Central Limit Theorems for Dependent Stochastic Processes. *International Statistical Review*, **62**(1), 119–132.

[2] M.A. Arcones, and B. Yu (1994). Central limit theorems for empirical and U-processes of stationary mixing sequences. *Journal of Theoretical Probability* **7**, 47–71.

[3] P. Bühlmann (1995). The blockwise bootstrap for general empirical processes of stationary sequences. *Stochastic Processes and their Applications*, **58**, 247–265.

[4] H. Dehling, O. Durieu and D. Volny (2009). New techniques for empirical processes of dependent data. *Stochastic Processes and their Applications*, **119**, 3699–3718.

[5] H. Dehling, O. Durieu and M. Tusche (2014). Approximating class approach for empirical processes of dependent sequences indexed by functions. *Bernoulli*, **20**, 1372–1403.

[6] P. Doukhan, and D. Surgailis (1998). Functional central limit theorem for the empirical process of short memory linear processes. *Comptes Rendus de l'Académie des Sciences, Mathématique*, **326**, 87–92.

[7] F. El Ktaibi, B.G. Ivanoff, N.C. Weber (2014). Bootstrapping the empirical distribution of a linear process. *Statistics & Probability Letters*, **93**, 134–142

[8] T.H. Hildebrandt (1963). *Introduction to the Theory of Integration.* Academic Press, New York, London.

[9] M. Peligrad (1998). On the blockwise bootstrap for empirical processes for stationary sequences. *Annals of Probability*, **26**, 887–901.

[10] D. Radulović (1996). The bootstrap for empirical processes based on stationary observations. *Stochastic Processes and Their Applications*, **65**, 259–279.

[11] E. Rio (1998). Processus empiriques absolument réguillets et entropie universelle. *Probability Thoery and Related Fields*, **111**, 585–608.

[12] E. Rio (2000). *Théorie asymptotique des processus aléatoires faiblement dependants*, Collection Mathématiques and Applications, **31**. Springer.

[13] G.R. Shorack and J.A. Wellner (2009). *Empirical Processes with Applications to Statistics*, 2nd Edition, SIAM.

[14] A.W. van der Vaart and J. Wellner (2000). *Weak convergence and empirical processes*. Springer Series in Statistics. Springer.

[15] D. Viennet (1997). Inequalities for absolutely regular sequences: application to density estimation. *Probability Thoery and Related Fields*, **107**, 467–492.