Hilbertian invariance principles for the empirical process under association *

Bruno Morel Charles Suquet

Statistique et Probabilités F.R.E. CNRS 2222

Bât. M2, Université des Sciences et Technologies de Lille

F-59655 Villeneuve d'Ascq Cedex, FRANCE

Abstract

We prove that a *necessary and sufficient* condition for the weak $\mathrm{L}^2[0,1]$ convergence of the empirical process of a stationary *associated* sequence (X_k) of uniform random variables is simply $\sum(2/3 - \mathbb{E}\max(X_0, X_k)) < \infty$. This condition is more natural in this setting than the classical covariance summability conditions. In the same spirit, we discuss the weak convergence in the Besov spaces $\mathrm{B}_2^{s,2}$ $(s < 1/2)$.

Keywords: association, Besov space, Brownian bridge, empirical process, invariance principles, positive dependence, quadratic statistics, Sobolev statistics, test for uniformity on the circle, tightness.

AMS classifications: 60F17, 60F25, 62G30.

---

*`Revised version`, July 7, 2004

1 Introduction

Let $(X_n,\ n \in \mathbb{Z})$ be a stationary sequence of uniform random variables on $[0,1]$. We denote by ξ_n the corresponding uniform empirical process:

$$\xi_n(t) = n^{-1/2} \sum_{i=1}^{n} \big(\mathbf{1}(X_i \le t) - t\big), \qquad t \in [0,1], \tag{1}$$

where $\mathbf{1}(A)$ is the indicator function of the event A. When the X_i are independent, $(\xi_n,\ n \ge 1)$ is well known to converge weakly in the Skorohod space $D(0,1)$ to the Brownian bridge. The theoretical importance of this result has motivated many investigations to extend it to the case of non independent observations X_i. Roughly speaking, the extensions have the following form. If the dependence structure of the sequence $(X_n,\ n \in \mathbb{Z})$ is close enough to independence, then ξ_n converges weakly in $D(0,1)$ to a centered Gaussian process ξ with covariance:

$$\Gamma(s,t) = \sum_{k \in \mathbb{Z}} \big(P(X_0 \le s, X_k \le t) - st\big), \quad 0 \le s,t \le 1. \tag{2}$$

Observe that the term indexed by $k = 0$ in the above series is equal to $s \wedge t - st$, which is the covariance of the classical Brownian bridge. The sum of the other terms represents the perturbation of the limiting process induced by the lack of independence. The closeness to independence is expressed by the rate of convergence to zero of some quantity like the various mixing coefficients (see for instance [17], [2] and the references therein) or the non negative covariances $\mathrm{Cov}(X_0, X_k)$ when $(X_n, n \in \mathbb{Z})$ is an associated sequence. In this paper we shall focus on this type of positive dependence. Recall that the sequence $(X_n, n \in \mathbb{Z})$ is said *associated* if for each finite choice of indexes $i_1, \dots, i_m$ in $\mathbb{Z}$ and for each pair of coordinatewise nondecreasing functions f, g defined on $\mathbb{R}^m$, we have

$$\mathrm{Cov}\big(f(X_{i_1}, \dots, X_{i_m}), g(X_{i_1}, \dots, X_{i_m})\big) \ge 0,$$

whenever this covariance exists. For the basic properties of association, we refer to Esary, Proschan and Walkup [3]. The first results on asymptotic behavior of empirical distribution function of an associated sequence were obtained by Yu [22]. He proved the weak $D(0,1)$ convergence of ξ_n, assuming the convergence of $\sum_{n\geq 1} n^{13/2+\varepsilon}\operatorname{Cov}(X_0,X_n)$. This assumption has been much improved by Shao and Yu [17] into $\operatorname{Cov}(X_0,X_n)=O(n^{-a})$ with $a>(3+\sqrt{33})/2\simeq 4.373$ and more relaxed by Louhichi [9] who only requires $a>4$. The relative severity of these conditions is due to the necessity to have an uniform bound of $H_{i,j}(s,t)=\operatorname{Cov}\bigl(\mathbf{1}(X_i\leq s),\mathbf{1}(X_j\leq t)\bigr)$ in terms of $\operatorname{Cov}(X_i,X_j)$ (the estimate $\|H_{i,j}\|_\infty=O\bigl(\operatorname{Cov}^{1/3}(X_i,X_j)\bigr)$ is the best known up today) and to the use of some generalized Rosenthal inequality of order 4 to control the increments of ξ_n.

However, the weak convergence of some useful functionals of paths of ξ_n does not require the $D(0,1)$-continuity of these functionals. An important example is the case of quadratic statistics like Cramér-von Mises test statistics, Watson statistics, which only need the $\mathrm{L}^2[0,1]$ functional framework (for more examples, see [13] and the references therein). With this motivation, Oliveira and Suquet [12] obtained the weak $\mathrm{L}^2[0,1]$ convergence of ξ_n to the Gaussian process ξ under the condition

$$\sum_{n=1}^{\infty}\operatorname{Cov}^{1/3}(X_0,X_n)<\infty. \tag{3}$$

A statistical application of this result was proposed by Suquet and Viano [20] in a problem of change point detection for the marginal distribution of an associated sequence.

The present contribution improves on [12] by establishing that a *necessary* and sufficient condition for the weak $\mathrm{L}^2[0,1]$ convergence of ξ_n to a Gaussian random element ξ in $\mathrm{L}^2[0,1]$ with covariance kernel given by (2) is

$$\sum_{n=1}^{\infty}\Bigl(\frac{2}{3}-\mathbb{E}\max(X_0,X_n)\Bigr)<\infty. \tag{4}$$

According to Lemma 11 below, equivalent conditions are obtained replacing the general term of the above series by $\mathbb{E}\min(X_0, X_n) - 1/3$ or $1/3 - \mathbb{E}\,|X_0 - X_n|$. Condition (4) is intermediate between the convergences of $\sum_{n\geq 1}\mathrm{Cov}(X_0, X_n)$ and $\sum_{n\geq 1}\mathrm{Cov}^{1/2}(X_0, X_n)$, see Lemmas 12 and 14. We think that this unusual coefficient $2/3 - \mathbb{E}\max(X_0, X_k)$ is the natural one in the $\mathrm{L}^2[0,1]$ framework. Indeed let us assume, just for heuristic convenience, that Γ is continuous. Then obviously the finiteness of $\int_0^1 \Gamma(t,t)\,\mathrm{d}t = \int_0^1 \mathbb{E}\,\xi(t)^2\,\mathrm{d}t = \mathbb{E}\parallel \xi \parallel_2^2$ is necessary. This leads to (4), via the following elementary computation where X, Y, denote uniform $[0,1]$ distributed random variables:

$$\begin{aligned}\int_0^1 \mathrm{Cov}\bigl(\mathbf{1}(X\leq t), \mathbf{1}(Y\leq t)\bigr)\,\mathrm{d}t &= \int_0^1 \bigl\{P\bigl(\max(X,Y)\leq t\bigr) - t^2\bigr\}\,\mathrm{d}t\\ &= \int_0^1 \bigl\{1 - P\bigl(\max(X,Y) > t\bigr)\bigr\}\,\mathrm{d}t - \frac{1}{3}\\ &= \frac{2}{3} - \mathbb{E}\max(X,Y). \qquad (5)\end{aligned}$$

The above optimal result in $\mathrm{L}^2[0,1]$ rely on variance estimates while in the $D[0,1]$ setting, controlling the fourth moments of the ξ_n's increments is the main difficulty. It was then rather natural to investigate some stronger topologies than the $\mathrm{L}^2[0,1]$ one, allowing the use of variance estimate as basic tool to obtain tightness. With this motivation we consider the scale of Besov spaces $\mathrm{B}_2^{s,2}[0,1]$ obtained (roughly speaking) by controlling the $\mathrm{L}^2[0,1]$ modulus of smoothness. For $0 < s < 1/2$, the corresponding Besov space supports steps functions, while for $s > 1/2$, $\mathrm{B}_2^{s,2}[0,1]$ is continuously embedded in some Hölder space. We prove that when $0 < s < 1/2$, a sufficient condition for the $\mathrm{B}_2^{s,2}[0,1]$ weak convergence of ξ_n is

$$\sum_{k\geq 1}\mathrm{Cov}^{1/2-s}(X_0, X_k) < \infty. \qquad (6)$$

Due to the classical embeddings in the general scale of Besov spaces $\mathrm{B}_p^{s,q}[0,1]$, this result allows us to relax the sufficient condition in [13] for $\mathrm{L}^p[0,1]$ weak convergence of ξ_n under association.

Finally we note that in concrete situations, the covariance Γ is often unknown and has to be estimated. First steps in this direction may be found in [4] and [6].

2 Invariance principle in $\mathrm{L}^2[0,1]$

Let ξ be a square integrable random element in $\mathrm{L}^2[0,1]$ ($\mathbb{E} \| \xi \|_2^2$ is finite). With the usual identification of $\mathrm{L}^2[0,1]$ and its dual, the covariance of ξ is the continuous bilinear form defined by

$$\mathrm{C}_\xi(f_1, f_2) := \mathbb{E}\big(\langle f_1, \xi\rangle \langle f_2, \xi\rangle\big), \quad f_1, f_2 \in \mathrm{L}^2[0,1].$$

Clearly C_ξ is symmetric and non negative. Since ξ is square integrable, we have $\sum_i \mathrm{C}_\xi(e_i, e_i) < \infty$ for any Hilbertian basis $(e_i,\ i \in \mathbb{N})$ of $\mathrm{L}^2[0,1]$. Then there is a unique element K of $\mathrm{L}^2\big([0,1]^2\big)$ such that

$$\mathrm{C}_\xi(f_1, f_2) = \int_{[0,1]^2} K(s,t) f_1(s) f_2(t) \,\mathrm{d}s \,\mathrm{d}t, \quad f_1, f_2 \in \mathrm{L}^2[0,1].$$

We call K the covariance integral kernel of ξ.

The Haar basis $(e_n, n \geq 0)$ is a very convenient tool to prove the main results of this section. Let us recall its definition and fix some notations. Put $e_0 := \mathbf{1}_{[0,1]}$ and write each $n \geq 1$ under the form $n = 2^j + k$, with $0 \leq k < 2^j$, $j \geq 0$. Let $I_{j,k} := \big[k 2^{-j-1}, (k+1) 2^{-j-1}\big)$ and

$$e_n = e_{j,k} := 2^{j/2}(\chi_{j,k} - \chi_{j,k+1}),$$

where $\chi_{j,k}$ is the indicator function of $I_{j,k}$. Denote by E_j the operator of orthogonal projection onto the subspace spanned by $\{e_n,\ 0 \leq n < 2^{j+1}\}$. For $f \in \mathrm{L}^2[0,1]$, $E_j f$ is simply the approximation of f by a step function equal to its mean value over each $I_{j,k}$. Let us write $E_j(s,t)$ for the integral kernel of E_j, that is:

$$\big(E_j f\big)(s) = \int_0^1 E_j(s,t) f(t) \,\mathrm{d}t, \quad s \in [0,1].$$

It is easily verified that $E_j(s,t)$ is the uniform probability density on the union of diagonal squares $C_{j,k} = I_{j,k}^2$ $(0 \le k < 2^j)$. It follows that for any function g *continuous* on $[0,1]^2$,

$$\lim_{j\to\infty} \int_{[0,1]^2} g(s,t) E_j(s,t) \,\mathrm{d}s \,\mathrm{d}t = \int_0^1 g(s,s) \,\mathrm{d}s. \tag{7}$$

Proposition 1. *Let $(X_k,\ k \in \mathbb{Z})$ be a stationary and associated sequence of uniform random variables on $[0,1]$ and consider the series*

$$\Gamma(s,t) = \sum_{k\in\mathbb{Z}} \operatorname{Cov}\big(\mathbf{1}(X_0 \le s), \mathbf{1}(X_k \le t)\big), \quad 0 \le s,t \le 1. \tag{8}$$

a) Assume that

$$\sum_{k\ge 1} \Big(\frac{2}{3} - \mathbb{E}\max(X_0, X_k)\Big) < \infty. \tag{9}$$

Then $\int_0^1 \Gamma(s,s)\,\mathrm{d}s$ is finite and the series in (8) converges almost everywhere on $[0,1]^2$ and on its diagonal. Moreover the sequence of kernels

$$\Gamma_n(s,t) := \mathbb{E}\,\xi_n(s)\xi_n(t) = \sum_{|k|<n} \Big(1 - \frac{|k|}{n}\Big) \operatorname{Cov}\big(\mathbf{1}(X_0 \le s), \mathbf{1}(X_k \le t)\big),$$

converges in $\mathrm{L}^2([0,1]^2)$ to Γ.

b) If Γ defined by (8) is the covariance integral kernel of some square integrable random element in $\mathrm{L}^2[0,1]$, then (9) holds.

Proof of a). First observe that, due to the association of (X_k), Γ defined by (8) is the limit of a non decreasing sequence of continuous functions and hence is a nonnegative measurable function (with possibly infinite values). The same holds true for the restriction of Γ to the diagonal of $[0,1]^2$. Now recalling (5), the finiteness of $\int_0^1 \Gamma(s,s)\,\mathrm{d}s$ follows clearly from (9) by Beppo Levi theorem. Hence Γ is finite almost everywhere on the diagonal of $[0,1]^2$.

Next, we note that $\Gamma_n(s,t)^2 \leq \Gamma_n(s,s)\Gamma_n(t,t)$ by Cauchy-Schwarz inequality in $L^2(\Omega)$, whence

$$\int_{[0,1]^2} \Gamma_n(s,t)^2 \,\mathrm{d}s\,\mathrm{d}t \leq \Big(\int_0^1 \Gamma_n(s,s)\,\mathrm{d}s\Big)^2 \leq \Big(\int_0^1 \Gamma(s,s)\,\mathrm{d}s\Big)^2 < \infty.$$

By Beppo Levi theorem, it follows that

$$\int_{[0,1]^2} \Gamma(s,t)^2 \,\mathrm{d}s\,\mathrm{d}t \leq \Big(\int_0^1 \Gamma(s,s)\,\mathrm{d}s\Big)^2 < \infty.$$

Hence Γ belongs to $\mathrm{L}^2([0,1]^2)$, so $\Gamma(s,t)$ is finite almost everywhere on $[0,1]^2$. By monotonicity of Γ_n, it follows that $\Gamma_n(s,t)$ converges almost everywhere to $\Gamma(s,t)$ on $[0,1]^2$. Since $\int_{[0,1]^2} |\Gamma - \Gamma_1|^2 \,\mathrm{d}s\,\mathrm{d}t$ is bounded by $4\big(\int_0^1 \Gamma(s,s)\,\mathrm{d}s\big)^2$, we can apply the monotone convergence theorem to obtain the convergence of $\int_{[0,1]^2} |\Gamma - \Gamma_n|^2 \,\mathrm{d}s\,\mathrm{d}t$ to zero. $\square$

Proof of b). If Γ is the integral covariance kernel of some square integrable random element ξ in $\mathrm{L}^2[0,1]$, then Γ is in $\mathrm{L}^2\big([0,1]^2\big)$ and

$$\mathbb{E} \| E_j\xi \|_2^2 = \int_{[0,1]^2} \Gamma(s,t)E_j(s,t)\,\mathrm{d}s\,\mathrm{d}t.$$

This together with the inequalities $\| \xi \|_2^2 \geq \| E_j\xi \|_2^2$, $E_j(s,t) \geq 0$ and $0 \leq \Gamma_n(s,t) \leq \Gamma(s,t)$ gives for any integers j, n:

$$\infty > \mathbb{E} \| \xi \|_2^2 \geq \int_{[0,1]^2} \Gamma_n(s,t)E_j(s,t)\,\mathrm{d}s\,\mathrm{d}t.$$

Fix n and recall that Γ_n is continuous. Letting j increase to infinity, we have by (7): $\mathbb{E} \| \xi \|_2^2 \geq \int_0^1 \Gamma_n(s,s)\,\mathrm{d}s$. When n in turn, goes to infinity, this gives by monotone convergence

$$\int_0^1 \Gamma(s,s)\,\mathrm{d}s \leq \mathbb{E} \| \xi \|_2^2 < \infty,$$

from which (9) clearly follows. $\square$

Theorem 2. *Let ξ_n defined by (1) be the empirical process of a stationary and associated sequence $(X_k,\ k \in \mathbb{Z})$ of uniform random variables on $[0,1]$. Then ξ_n converges in distribution in $\mathrm{L}^2[0,1]$ to a Gaussian random element ξ with covariance integral kernel Γ given by (8) if and only if*

$$\sum_{k \geq 1} \Big(\frac{2}{3} - \mathbb{E} \max(X_0, X_k) \Big) < \infty.$$

Proof. The necessity of Condition (9) is clear from Proposition 1 b). To prove the convergence in distribution of ξ_n to ξ in $\mathrm{L}^2[0,1]$, we have to check the convergence in distribution of the random variables $\int_0^1 f(t)\xi_n(t)\,\mathrm{d}t$ to $\int_0^1 f(t)\xi(t)\,\mathrm{d}t$ for every $f \in \mathrm{L}^2[0,1]$ and the tightness of (ξ_n). For the first point, we write

$$\int_0^1 f(t)\xi_n(t)\,\mathrm{d}t = \frac{g(X_1) + \cdots + g(X_n)}{\sqrt{n}},$$

where $g(X_i) = \int_{X_i}^1 f(t)\,\mathrm{d}t - \mathbb{E}\int_{X_i}^1 f(t)\,\mathrm{d}t$. Then the requested convergence follows from Newman's central limit theorem [10, 11] for absolutely continuous mappings of associated variables, observing that

$$\mathbb{E}\Big(\int_0^1 f(t)\xi_n(t)\,\mathrm{d}t \Big)^2 = \int_{[0,1]^2} f(s)f(t)\Gamma_n(s,t)\,\mathrm{d}s\,\mathrm{d}t$$

and using the $\mathrm{L}^2\big([0,1]^2\big)$ convergence of Γ_n to Γ.

To establish the tightness of (ξ_n), we use the following adaptation of Prohorov's theorem [15, Th. 1.13]. Consider an orthonormal basis $(f_i, i \geq 1)$ of $\mathrm{L}^2[0,1]$ and define $r_p^2(f) := \sum_{i>p} |\langle f, f_i\rangle|^2$. Then the sequence (ξ_n) is tight in $\mathrm{L}^2[0,1]$ if it satisfies

i) $\sup_{n \geq 1} \mathbb{E}\, \|\xi_n\|_2^2 < \infty$;

ii) $\lim_{p\to\infty} \sup_{n\geq 1} \mathbb{E}\, r_p^2(\xi_n) = 0$.

In the original statement [15, Th. 1.13], Condition i) was missing. It is added here to avoid situations where the projection of (ξ_n) on the first vectors of the basis is not tight, which obviously can not be detected by ii), see also the remark after Theorem 5 in [19]. It is easy to check that in Condition ii), $\sup_{n\geq 1}$ may be replaced by $\limsup_{n\to\infty}$ and that it suffices to take the limit in p along some subsequence. So choosing for orthonormal basis the Haar basis and the subsequence of r_p^2 indexed by $p = 2^{j+1} - 1$, we have to check i) and

ii') $\displaystyle\lim_{j\to\infty} \limsup_{n\to\infty} \big(\mathbb{E} \parallel \xi_n \parallel_2^2 - \mathbb{E} \parallel E_j \xi_n \parallel_2^2\big) = 0.$

Condition i) is easily satisfied, noting that for *every* $(s,t) \in [0,1]^2$ we have $0 \leq \Gamma_n(s,t) \leq \Gamma(s,t) \leq \infty$, whence

$$\mathbb{E}\, \|\xi_n\|_2^2 = \int_0^1 \Gamma_n(t,t)\,\mathrm{d}t \leq \int_0^1 \Gamma(t,t)\,\mathrm{d}t.$$

Recall that by Proposition 1 a), this last integral is finite under (9).

To prove ii'), observe first that

$$\mathbb{E} \parallel \xi_n \parallel_2^2 - \mathbb{E} \parallel E_j \xi_n \parallel_2^2 = \int_0^1 \Gamma_n(s,s)\,\mathrm{d}s - \int_{[0,1]^2} \Gamma_n(s,t) E_j(s,t)\,\mathrm{d}s\,\mathrm{d}t =: A_{n,j}.$$

Let us fix for a moment an integer N. As for any $(s,t) \in [0,1]^2$, the sequence $(\Gamma_n(s,t), n \geq 1)$ is non decreasing and $E_j(s,t)$ is non negative, we get

$$\limsup_{n\to\infty} A_{n,j} \leq \int_0^1 \Gamma(s,s)\,\mathrm{d}s - \int_{[0,1]^2} \Gamma_N(s,t) E_j(s,t)\,\mathrm{d}s\,\mathrm{d}t. \tag{10}$$

Due to the continuity of Γ_N, we have by (7)

$$\limsup_{j\to\infty} \limsup_{n\to\infty} A_{n,j} \leq \int_0^1 \Gamma(s,s)\,\mathrm{d}s - \int_0^1 \Gamma_N(s,s)\,\mathrm{d}s. \tag{11}$$

Finally letting N go to infinity we get by monotone convergence and finiteness of $\int_0^1 \Gamma(s,s)\,\mathrm{d}s$,

$$\limsup_{j\to\infty} \limsup_{n\to\infty} A_{n,j} = 0,$$

which establishes ii') and completes the proof. $\square$

Remark. It follows from Proposition 1b) and Theorem 2 that Γ is the covariance integral kernel of a Gaussian random element in $\mathrm{L}^2[0,1]$ if and only if (9) holds.

3 Invariance principle in $B_2^{s,2}[0,1]$

For a function $f:[0,1]\to\mathbb{R}$, and $0\le h\le 1$, let

$$\Delta_h f(t) := \begin{cases} f(t+h)-f(t) & \text{if } 0\le t\le 1-h,\\ 0 & \text{else.}\end{cases}$$

Definition 3. *Let $s\in(0,1)$, p and $q\in[1,\infty)$. $\mathrm{B}_p^{s,q}[0,1]$ is the subspace of $\mathrm{L}^p[0,1]$ induced by the norm:*

$$\|f\|_{p,s,q} := \|f\|_p + \omega_{p,s,q}(f,1),$$

where $\omega_{p,s,q}(f,a)$ is the modulus of smoothness defined by:

$$\omega_{p,s,q}(f,a) := \left(\int_0^a \frac{\|\Delta_h f\|_p^q}{|h|^{sq}}\frac{\mathrm{d}h}{|h|}\right)^{1/q}, \quad 0<a\le 1.$$

The Besov space $\mathrm{B}_p^{s,q}[0,1]$ is a separable Banach space (separability holds because p and q are both finite). Step functions are in (separable) Besov spaces for $p,q\in[1,\infty)$ and $0<s<1/p$.

It is not difficult to construct a sequence of step functions which converges in each $\mathrm{B}_p^{s,q}[0,1]$ for $s<1/p$, but not for Skorohod topology and an other sequence of step functions which converges in Skorohod sense (and even uniformly), but for none $\mathrm{B}_p^{s,q}[0,1]$ with $0<s<1/p$. Consequently, we can not compare Skorohod topology with any Besovian topology with $0<s<1/p$.

In the spirit of Theorem 2, we now give a sufficient condition for membership in Besov spaces for the limiting process ξ.

Proposition 4. *Let $(X_k,\ k\in\mathbb{Z})$ be a stationary and associated sequence of uniform random variables on $[0,1]$ satisfying (9). As in Proposition 1, denote by ξ the Gaussian random element of $\mathrm{L}^2[0,1]$ whose covariance is given by (8). Assume moreover* *that for a given $s \in (0, 1/2)$,*

$$\sum_{k \geq 1} \Big(\frac{2}{3 - 2s} - \mathbb{E} \max(X_0, X_k)^{1-2s} \Big) < \infty, \tag{12}$$

$$\sum_{k \geq 1} \Big(\frac{2}{3 - 2s} - \mathbb{E} \big(1 - \min(X_0, X_k)\big)^{1-2s} \Big) < \infty, \tag{13}$$

$$\sum_{k \geq 1} \Big| \frac{1}{(3 - 2s)(1 - s)} - \mathbb{E} \left| X_0 - X_k \right|^{1-2s} \Big| < \infty. \tag{14}$$

Then,

$$\mathbb{E}\, \omega_{2,s,2}^2(\xi, 1) < \infty, \tag{15}$$

and for all $a \in (0, 1)$

$$\lim_{n \to \infty} \mathbb{E}\, \omega_{2,s,2}^2(\xi_n, a) = \mathbb{E}\, \omega_{2,s,2}^2(\xi, a). \tag{16}$$

Proof. By Fubini's theorem,

$$\mathbb{E}\, \omega_{2,s,2}(\xi, 1)^2 = \int_0^1 \frac{\mathrm{d}h}{h^{1+2s}} \int_0^{1-h} \mathbb{E} \big| \xi(t + h) - \xi(t) \big|^2 \, \mathrm{d}t.$$

By reporting (8), we get

$$\mathbb{E}\, \omega_{2,s,2}(\xi, 1)^2 = \int_0^1 \frac{\mathrm{d}h}{h^{1+2s}} \int_0^{1-h} \sum_{k \in \mathbb{Z}} \mathbb{E} \left(\mathbf{1}(t < X_0, X_k \leq t + h) - h^2 \right) \, \mathrm{d}t. \tag{17}$$

Our aim is to exchange summations. For fixed h, let $Y_k(h)$ be the random variable defined by:

$$Y_k(h) := \int_0^{1-h} \mathbf{1}(t < X_0, X_k \leq t + h) \, \mathrm{d}t.$$

Notice first that

$$Y_k(h) := \Big[\min\big(X_0, X_k, (1 - h)\big) - \big(\max(X_0, X_k) - h\big)_+ \Big]_+,$$

where $a_+ := \max(a, 0)$. To describe explicitly the random variable $Y_k(h)$ let us introduce the spacing random variables U_k, V_k and W_k:

$$\begin{aligned} U_k &:= \min(X_0, X_k), \\ V_k &:= |X_0 - X_k|, \\ W_k &:= 1 - \max(X_0, X_k). \end{aligned}$$

These quantities are linked by $U_k + V_k + W_k = 1$. Now observe that:

$$Y_k(h) = \begin{cases} 0 & \text{when } 0 \le h \le V_k, \\ h - V_k & \text{when } V_k \le h \le V_k + \min(U_k, W_k), \\ \min(U_k, W_k) & \text{when } V_k + \min(U_k, W_k) \le h \le V_k + \max(U_k, W_k), \\ 1 - h & \text{when } V_k + \max(U_k, W_k) \le h \le 1. \end{cases}$$

Let us define J_k by

$$J_k := \int_0^1 Y_k(h) \frac{\mathrm{d}h}{h^{1+2s}},$$

and observe that

$$J_k = \frac{1}{2s(1-2s)} \left[(U_k + V_k)^{1-2s} + (V_k + W_k)^{1-2s} - V_k^{1-2s} - 1 \right].$$

Let us now consider $I_k := J_k - C(s)$, where

$$C(s) = \int_0^1 h^2(1-h) \frac{\mathrm{d}h}{h^{1+2s}} = \frac{1}{(2-2s)(3-2s)},$$

is a centering term such that $\mathbb{E}\, I_k = 0$ when the X_j's are i.i.d. Now, we can rewrite I_k as

$$I_k = c_s \left(\max(X_0, X_k)^{1-2s} + \big(1 - \min(X_0, X_k)\big)\big)^{1-2s} - |X_0 - X_k|^{1-2s} - c_s' \right),$$

with $c_s = 1/(1-2s)(2-2s)$ and $c_s' = (2-3s)/(1-2s)(1-s)$. Now observe that $\mathbb{E}\, I_k$ is a linear combination of the terms of rank k of the series (12), (13), and (14). Therefore $\sum_k \mathbb{E}\, I_k$ converges absolutely. By exchanging summations in (17), we obtain

$$\mathbb{E}\, \omega_{2,s,2}(\xi, 1)^2 = \sum_{k \in \mathbb{Z}} \mathbb{E}\, I_k < \infty.$$

In order to prove (16), consider

$$\mathbb{E}\,\omega_{2,s,2}^2(\xi_n,a) = \int_0^a \frac{\mathrm{d}h}{h^{1+2s}} \int_0^{1-h} \mathbb{E}\big|\xi_n(t+h)-\xi_n(t)\big|^2\,\mathrm{d}t.$$

By stationarity of the sequence (X_k),

$$\mathbb{E}\,\omega_{2,s,2}^2(\xi_n,a) = \int_0^a \frac{\mathrm{d}h}{h^{1+2s}} \int_0^{1-h} \sum_{|k|<n} \left(1-\frac{|k|}{n}\right) \mathbb{E}\left(\mathbf{1}(t<X_0,X_k\le t+h)-h^2\right)\,\mathrm{d}t,$$

whence (16) follows by dominated convergence. □

Theorem 5. *Let ξ_n defined by (1) be the empirical process of a stationary and associated sequence $(X_k,\ k\in\mathbb{Z})$ of uniform random variables on $[0,1]$. Then ξ_n converges in distribution in $\mathrm{B}_2^{s,2}[0,1]$, with $0<s<1/2$, to a Gaussian random element ξ with covariance integral kernel Γ given by (8) if*

$$\sum_{k\ge 1} \mathrm{Cov}^{1/2-s}(X_0,X_k) < \infty. \tag{18}$$

Proof. By applying Lemma 14 in the case $\beta=1$ and by taking account of Lemma 11, we can deduce from (18) that

$$\sum_{k\ge 1}\Big(\frac{2}{3}-\mathbb{E}\max(X_0,X_k)\Big) < \infty.$$

Therefore, Theorem 2 ensures that ξ_n converges in distribution in $\mathrm{L}^2[0,1]$ to ξ.

Observe now that the conclusion of Proposition 4 holds under (18). By Lemma 13, (18) implies (12). If (X,Y) is PQD with uniform marginals, so is $(1-X,1-Y)$ and we get (13) by another application of Lemma 13. Moreover we can deduce (14) from (18) via Lemma 14.

Then, we apply Lemma 8 to show the tightness of (ξ_n) in $\mathrm{B}_2^{s,2}[0,1]$. Since we already know that (ξ_n) is tight in $\mathrm{L}^2[0,1]$, *i)* is fulfilled. In order to check *ii)*, let us consider $\varepsilon>0$. By Proposition 4,

$$\limsup_{n\to+\infty} \mathbb{P}\{\omega_{2,s,2}(\xi_n,a)>\varepsilon\} \le \frac{1}{\varepsilon^2}\,\mathbb{E}\,\omega_{2,s,2}(\xi,a)^2,$$

and the result follows since the quantity on the right goes to 0 with a. $\square$

Corollary 6. *Let ξ_n defined by (1) be the empirical process of a stationary and associated sequence $(X_k,\ k \in \mathbb{Z})$ of uniform random variables on $[0,1]$. Then ξ_n converges in distribution in $\mathrm{L}^p[0,1]$, with $p \in [2,\infty)$, to a Gaussian random element ξ with covariance integral kernel Γ given by (8) if $\mathrm{Cov}(X_0, X_n) = O(n^{-\tau})$, with $\tau > p$.*

Proof. We use the following embedding of Besov spaces (see e.g. [14] or [21]) which allows to exchange "smoothness" (s) for "integrability" (p). Let $1 \le p_1 \le p_2 < \infty$, $0 < s_2 \le s_1$, $1 \le q < \infty$, with $s_1 - 1/p_1 = s_2 - 1/p_2$. Then

$$\mathrm{B}_{p_1}^{s_1,q}[0,1] \hookrightarrow \mathrm{B}_{p_2}^{s_2,q}[0,1].$$

Now choose $p_1 = 2$, $q = 2$, $p_2 = p \in (2,\tau)$ and $\varepsilon > 0$ such that $1/p = (1+\varepsilon)/\tau + \varepsilon$. Finally choose $s_1 = s$ with $1/2 - s = (1+\varepsilon)/\tau$ and $s_2 = \varepsilon$. Then the conclusion follows from the embeddings

$$\mathrm{B}_2^{s,2}[0,1] \hookrightarrow \mathrm{B}_p^{\varepsilon,2}[0,1] \hookrightarrow \mathrm{L}^p[0,1].$$

$\square$

Corollary 6 improves on the similar result in [13], obtained for $\tau > 3p/2$. This result has an interest only if $\tau \le 4$, since Louhichi [9] recently obtains the convergence in $D(0,1)$ for $\tau > 4$.

4 Some applications to test statistics

One motivation for the study of the weak convergence of the empirical process in the topology of some function space, is the application to the convergence of statistics which are continuous functionals of ξ_n in the same topology. Some applications of the $\mathrm{L}^2[0,1]$ or $\mathrm{L}^p[0,1]$ weak convergence are presented in [12] and [13]. We propose here two applications of the $\mathrm{B}_2^{s,2}[0,1]$ weak convergence of ξ_n to the problem of testing uniformity on a circle. In all this section, we parametrize the unit circle by $t \mapsto \exp(2i\pi t)$, the observations X_j are interpreted as angular data, the functions $f \in \mathrm{L}^2[0,1]$ are implicitly extended by 1-periodicity and $\Delta_h f(t) := f(t+h) - f(t)$ for every $t \in \mathbb{R}$. The question is to decide whether the given observations $\{\exp(2i\pi X_j), 1 \leq j \leq n\}$ indicate any preferred direction or whether these data come from a uniform distribution on the circumference. Let us observe that the problem of the goodness of fit of angular data Y_j to a completely specified distribution with continuous distribution function F may be reduced to the test of uniformity for the $F(Y_j)$'s.

Giné [5, Section 6] considered the class of Sobolev statistics

$$T_n(\{a_k\}) := \frac{2}{n} \sum_{k=1}^{\infty} a_k^2 \sum_{j,l=1}^{n} \cos\bigl(2\pi k(X_j - X_l)\bigr), \tag{19}$$

where the sequence $(a_k)_{k\geq 1}$ satisfies some summability condition. The simplest example of statistics of this type is the Rayleigh statistics given by

$$\frac{1}{n}\Bigl[\Bigl(\sum_{j=1}^{n} \cos(2\pi X_j)\Bigr)^2 + \Bigl(\sum_{j=1}^{n} \sin(2\pi X_j)\Bigr)^2\Bigr] = \frac{1}{n} \sum_{j,l=1}^{n} \cos\bigl(2\pi(X_j - X_l)\bigr),$$

which corresponds to the choice $a_1 = 1$ and $a_k = 0$ for $k \geq 2$. The well known Watson's statistics

$$U_n^2 = \int_0^1 \Bigl(\xi_n(t) - \int_0^1 \xi_n(u)\,\mathrm{d}u\Bigr)^2 \mathrm{d}t,$$

corresponds to the choice $a_k = c/k$ $(k \geq 1)$ for some constant c. Another classical example is Ajne's statistics

$$A_n = \frac{1}{n} \int_0^1 \Bigl|N(t) - \frac{n}{2}\Bigr|^2 \mathrm{d}t,$$

where $N(t)$ is the number of observations located on the half circle centered at $\exp(2\pi i t)$. The corresponding sequence (a_k) is of the form $a_{2k} = 0$ and $a_{2k+1} = c(-1)^k(2k+1)^{-1}$. Some generalization of Ajne's statistics is investigated in [7].

After integration by parts, the Sobolev statistics (19) may be recast as

$$T_n(\{a_k\}) = 8\pi^2 \sum_{k=1}^{\infty} k^2 a_k^2 \Big| \int_0^1 \xi_n(t) \mathrm{e}^{2i\pi kt} \, \mathrm{d}t \Big|^2 = 8\pi^2 \sum_{k=1}^{\infty} k^2 a_k^2 |c_k(\xi_n)|^2,$$

where we have defined the Fourier coefficients $c_k(f)$ of $f \in \mathrm{L}^2[0,1]$ by

$$c_k(f) := \int_0^1 f(t) \mathrm{e}^{-2i\pi kt} \, \mathrm{d}t, \quad k \in \mathbb{Z}.$$

As $|c_k(\Delta_h f)|^2 = 4\sin^2(\pi kh)|c_k(f)|^2$, the Plancherel identity leads to

$$A \sum_{k\in\mathbb{Z}} |k|^{2s} |c_k(f)|^2 \leq \omega_{2,s,2}(f)^2 \leq A' \sum_{k\in\mathbb{Z}} |k|^{2s} |c_k(f)|^2,$$

with $A = 4\int_0^1 u^{-1-2s} \sin^2(\pi u) \, \mathrm{d}u$ and $A' = 4\int_0^\infty u^{-1-2s} \sin^2(\pi u) \, \mathrm{d}u$ and $0 < s < 1$. Hence we have the equivalence of norms

$$\|f\|_{2,s,2}^2 \sim \sum_{k\in\mathbb{Z}} (1 + |k|^{2s}) |c_k(f)|^2.$$

This shows that the statistics (19) are functionals of ξ_n, continuous in the $\mathrm{B}_2^{s,2}[0,1]$ topology provided that $a_k = O(k^{s-1})$. Here the range of validity for s is of course $0 < s < 1/2$, as the Fourier coefficients of ξ_n are exactly of asymptotic order $1/k$.

Another approach using χ^2 methods was proposed by Rao [16] to test uniformity on the circle. The statistic considered by Rao may be defined as follows. First *fix* the number of classes m and choose some direction $\exp(2\pi i\alpha)$. Next divide the unit-circle in m segments of the same width, parametrized by the intervals $I_j(\alpha) = \big(\alpha + (j-1)/m, \alpha + j/m\big]$ $(1 \leq j \leq m)$. Denote by $n_j(\alpha)$ the number of observations X_k that fall in $I_j(\alpha)$. The corresponding usual χ^2 statistic is then

$$\chi_n^2(\alpha) = \sum_{j=1}^{m} \big(n_j(\alpha) - n/m\big)^2 / (n/m).$$

To have a statistic independent of the starting point α, Rao proposes to use

$$R_{n,m} := \int_0^1 \chi_n^2(\alpha) \, \mathrm{d}\alpha = \frac{m}{n} \sum_{j=1}^{m} \int_0^1 \Big| n_j(\alpha) - \frac{n}{m} \Big|^2 \, \mathrm{d}\alpha. \tag{20}$$

Recalling that ξ_n has been extended by 1-periodicity we get

$$n_j(\alpha) - \frac{n}{m} = \sqrt{n}\Big[\xi_n\Big(\alpha + \frac{j}{m}\Big) - \xi_n\Big(\alpha + \frac{j-1}{m}\Big)\Big],$$

whence

$$R_{n,m} = m^2 \int_0^1 \Big|\xi_n\Big(\alpha + \frac{1}{m}\Big) - \xi_n(\alpha)\Big|^2 \,\mathrm{d}\alpha = m^2 \| \Delta_{1/m}\xi_n \|_2^2.$$

So $R_{n,m}$ appears as a functional of ξ_n, continuous in the $\mathrm{L}^2[0,1]$ topology and this continuity is sufficient to deduce from the mild assumption of Theorem 2 the weak convergence of this statistic when n goes to infinity and m remains fixed.

To allow more flexibility with respect to m, we may consider the weighted sums

$$R_n := \sum_{m=2}^{M} b_m R_{n,m}, \tag{21}$$

where $M = M(n) \uparrow \infty$ with n. If we choose the weights b_m such that $0 \le b_m \le cm^{2s-3}$, then comparison between series and integral gives

$$R_n \le C \int_{1/M}^1 \| \Delta_h \xi_n \|_2^2 \frac{\mathrm{d}h}{h^{1+2s}}, \tag{22}$$

where c, C are positive constants. Now the space $\mathrm{B}_2^{s,2}[0,1]$ is the relevant framework to investigate the existence of a limiting distribution for R_n.

Proposition 7. *Let the weights b_m $(m \ge 2)$ satisfy for some constant c, $0 \le b_m \le cm^{2s-3}$. Assume that the sequence (X_j) fulfills the conditions of Theorem 5. Then R_n defined by (21) converges weakly to*

$$R := \sum_{m=2}^{\infty} m^2 b_m \| \Delta_{1/m}\xi \|_2^2,$$

where ξ is a Gaussian random element in $\mathrm{B}_2^{s,2}[0,1]$ with covariance kernel given by (8).

Proof. Consider the functionals

$$T_n(f) := \sum_{m=2}^{M(n)} m^2 b_m \| \Delta_{1/m} f \|_2^2, \qquad T(f) := \sum_{m=2}^{\infty} m^2 b_m \| \Delta_{1/m} f \|_2^2.$$

Both are continuous on $\mathrm{B}_2^{s,2}[0,1]$ as squares of seminorms dominated by $\omega_{2,s,2}(f,1)$. Write

$$R_n = T_n(\xi_n) = (T_n - T)(\xi_n) + T(\xi_n).$$

The weak convergence of $T(\xi_n)$ to $T(\xi) = R$ follows from Theorem 5 and the continuity of T. So it suffices to check that $(T_n - T)(\xi_n)$ goes to zero in probability. For this convergence, a comparison between series and integral gives

$$\big|(T_n - T)(\xi_n)\big| \leq \sum_{m=M(n)+1}^{\infty} cm^{2s-1} \| \Delta_{1/m} \xi_n \|_2^2 \leq C\omega_{2,s,2}^2\Big(\xi_n, \frac{1}{M(n)}\Big).$$

Now

$$\mathrm{P}\Big\{\omega_{2,s,2}^2\Big(\xi_n, \frac{1}{M(n)}\Big) \geq \varepsilon\Big\} \leq \sup_{k\geq 1} \mathrm{P}\Big\{\omega_{2,s,2}^2\Big(\xi_k, \frac{1}{M(n)}\Big) \geq \varepsilon\Big\}$$

and this upper bound goes to zero due to the tightness of $(\xi_k)_{k\geq 1}$, according to the remark following the proof of Lemma 8 below. □

5 Tools and auxiliary results

We give here a sufficient condition for the tightness in Besov spaces, involving the moduli of smoothness $\omega_{p,s,q}$.

Lemma 8. *Let $p, q \in [1, \infty)$ and $s \in (0,1)$. Assume the sequence $(\xi_n, n \geq 1)$ of random elements in $\mathrm{L}^p[0,1]$ satisfies*

i) (ξ_n) *is tight in* $\mathrm{L}^p[0,1]$,

ii) $\forall \varepsilon > 0, \quad \lim_{a\to 0} \limsup_{n\to\infty} \mathrm{P}\{\omega_{p,s,q}(\xi_n, a) > \varepsilon\} = 0.$

Then $(\xi_n, n \geq 1)$ *is tight in* $\mathrm{B}^{s,q}_p[0,1]$.

Proof. The proof relies on the following characterization of relative compactness in $\mathrm{B}^{s,q}_p[0,1]$: a subset K of $\mathrm{B}^{s,q}_p[0,1]$ is relatively compact if and only if it is relatively compact in $\mathrm{L}^p[0,1]$ and satisfies:

$$\lim_{a\to 0} \sup_{f\in K} \omega_{p,s,q}(f,a) = 0. \tag{23}$$

The necessity of relative compactness in $\mathrm{L}^p[0,1]$ is obvious. To establish the necessity of (23), suppose K relatively compact and fix $\varepsilon > 0$. Then we have some finite ε-net $(g_1, \dots, g_m)$ in K for $\| \ \|_{p,s,q}$. As (23) is obvious for a finite set of functions, the elementary inequality

$$\omega_{p,s,q}(f,a) \leq \omega_{p,s,q}(g,a) + \|f-g\|_{p,s,q},$$

gives the bound $\omega_{p,s,q}(f,a) \leq 2\varepsilon$ uniformly in $f \in K$, for a small enough.

Conversely, assume that K is relatively compact in $\mathrm{L}^p[0,1]$ and satisfies (23). Let $\varepsilon > 0$ and $b > 0$ be such that $\omega_{p,s,q}(f,b) \leq \varepsilon$ uniformly in $f \in K$. Put

$$c_b := \left(\int_{b\leq |h|<1} \frac{2^q}{|h|^{sq}} \frac{\mathrm{d}h}{|h|^d} \right)^{1/q}.$$

We can cover K with a finite number of balls for $\| \ \|_p$, with radius less than $\varepsilon/(1+c_b)$. The elementary estimates

$$\|\Delta_h f\|_p \leq 2\|f\|_p,$$

$$\omega_{p,s,q}(f-g,a) \leq \omega_{p,s,q}(f,a) + \omega_{p,s,q}(g,a),$$

lead to

$$\|f-g\|_{p,s,q} \leq \|f-g\|_p + \left(\int_{b\leq|h|<1} \frac{2^{mq}\|f-g\|_p^q}{|h|^{sq}} \frac{\mathrm{d}h}{|h|^d} \right)^{1/q} + \omega_{p,s,q}(f,b) + \omega_{p,s,q}(g,b).$$

So K can be covered by a finite number of balls with radius 3ε for $\| \ \|_{p,s,q}$.

Finally by Prohorov theorem and usual techniques (see for instance Billingsley [1, p. 55] for analogous result in $C[0,1]$), we obtain a tightness criterion in $\mathrm{B}^{s,q}_p[0,1]$ with *ii′*) $\forall \varepsilon > 0, \quad \lim_{a \to 0} \sup_{n \geq 1} \mathbb{P}\{\omega_{p,s,q}(\xi_n, a) > \varepsilon\} = 0,$

instead of *ii*). A standard argument allows to change the sup into a lim sup to obtain *ii*). □

Remark. Condition *ii′*) is necessary for the tightness of $(\xi_n)_{n\geq 1}$.

Proof. Indeed by monotony, it suffices to prove that if $(\xi_n)_{n\geq 1}$ is tight, then for every positive ε,

$$\lim_{l\to\infty} \sup_{n\geq 1} \mathbb{P}\{\omega_{p,s,q}(\xi_n, 1/l) > \varepsilon\} = 0.$$

As $p, q < \infty$, we have $\omega_{p,s,q}(f, 1/l) \to 0$ for any $f \in \mathrm{B}_p^{s,q}[0,1]$ by dominated convergence. So the sequence of closed sets

$$F_l := \big\{ f \in \mathrm{B}_p^{s,q}[0,1];\ \omega_{p,s,q}(f, 1/l) \geq \varepsilon \big\}$$

decreases to the empty set when l increases to infinity. Now the result follows from the relative compactness of the sequence of distributions $\mathbb{P}\,\xi_n^{-1}$, see e.g. [19, Lem. 1 p. 206]. □

The following integration by parts formulas for covariances are basic tools throughout the paper. For notational simplicity, let us write for random variables X, Y.

$$H(u,v) = \mathrm{Cov}\big(\mathbf{1}(X \leq u), \mathbf{1}(Y \leq v)\big) = \mathrm{Cov}\big(\mathbf{1}(X > u), \mathbf{1}(Y > v)\big).$$

Similarly we write $H_{i,j}$ when X and Y are replaced by X_i, X_j respectively in the above formula.

Lemma 9. *If X and Y are square integrable random variables, then*

$$\mathrm{Cov}(X,Y) = \int_{\mathbb{R}^2} H(x,y)\,\mathrm{d}x\,\mathrm{d}y. \tag{24}$$

For the proof, we refer to [8], see also [22] for an interesting generalization.

Lemma 10. *Let f be a complex valued function defined on $\mathbb{R}^2$ with second order partial derivatives continuous and bounded on $\mathbb{R}^2$. Let X and Y be square integrable random variables. Then*

$$\int_{\mathbb{R}^2} f(x,y)\big(P_{X,Y} - P_X \otimes P_Y\big)(\,\mathrm{d}x,\,\mathrm{d}y) = \int_{\mathbb{R}^2} \frac{\partial^2 f}{\partial u \partial v}(u,v) H(u,v)\,\mathrm{d}u\,\mathrm{d}v.$$

A detailed proof can be found for instance in [18]. The adaptation of these results to the case of random vectors (X,Y) whose distribution is supported by $[0,1]^2$ instead of $\mathbb{R}^2$ is straightforward.

The three following lemmas involve the coefficients $2/3 - \mathbb{E}\max(X,Y)$ used in Section 2. Since the full strength of the association property is not required for their proofs, we prefer to state them in the more general framework of positive quadrant dependence. Recall that the sequence (X_n) is said *pairwise positive quadrant dependent* (pairwise PQD) if, given any reals s, t, for $i \neq j$,

$$H_{i,j}(s,t) \geq 0.$$

Lemma 11. *For each random vector (X,Y) with uniform marginals on $[0,1]$,*

$$\mathbb{E}\min(X,Y) - \frac{1}{3} = \frac{2}{3} - \mathbb{E}\max(X,Y) = \frac{1}{2}\Big(\frac{1}{3} - \mathbb{E}\,|X-Y|\Big). \tag{25}$$

If (X,Y) is PQD, all these quantities are non negative.

Proof. The proof relies on integral representations of $\mathbb{E}\min(X,Y) - 1/3$ and $2/3 - \mathbb{E}\max(X,Y)$ which give also the sign of these quantities when (X,Y) is PQD. Recalling (5), we have

$$\begin{aligned}
\frac{2}{3} - \mathbb{E}\max(X,Y) &= \int_0^1 H(t,t)\,\mathrm{d}t \\
&= \int_0^1 \operatorname{Cov}\big(\mathbf{1}(X>t), \mathbf{1}(Y>t)\,\mathrm{d}t \\
&= \int_0^1 \big\{P\big(\min(X,Y)>t\big) - P(X>t)^2\big\}\,\mathrm{d}t \\
&= \mathbb{E}\min(X,Y) - \frac{1}{3}.
\end{aligned}$$

To obtain the last equality in (25), it now suffices to write:

$$\begin{aligned}\frac{2}{3} - \mathbb{E}\max(X,Y) &= \frac{1}{2}\Big(\frac{2}{3} - \mathbb{E}\max(X,Y) + \mathbb{E}\min(X,Y) - \frac{1}{3}\Big)\\ &= \frac{1}{2}\Big(\frac{1}{3} - \mathbb{E}\,|X-Y|\Big).\end{aligned}$$

□

Lemma 12. *Let $(X_k, k\in\mathbb{Z})$ be a stationary and pairwise PQD sequence of random variables with X_k uniformly $[0,1]$ distributed. Then*

$$0 \le \sum_{k\in\mathbb{Z}} \operatorname{Cov}(X_0,X_k) \le \sum_{k\in\mathbb{Z}}\Big(\frac{2}{3} - \mathbb{E}\max(X_0,X_k)\Big). \tag{26}$$

Proof. Integrating on $[0,1]^2$ the relation

$$\mathbb{E}\,\xi_n(s)\xi_n(t) = \sum_{|k|<n}\Big(1-\frac{|k|}{n}\Big)H_{0,k}(s,t), \tag{27}$$

gives

$$\int_{[0,1]^2} \mathbb{E}\,\xi_n(s)\xi_n(t)\,\mathrm{d}s\,\mathrm{d}t = \sum_{|k|<n}\Big(1-\frac{|k|}{n}\Big)\operatorname{Cov}(X_0,X_k). \tag{28}$$

On the other hand, by Fubini's theorem and comparison of norms we have

$$\int_{[0,1]^2} \mathbb{E}\,\xi_n(s)\xi_n(t)\,\mathrm{d}s\,\mathrm{d}t = \mathbb{E}\Big(\int_0^1 \xi_n(s)\,\mathrm{d}s\Big)^2 \le \mathbb{E}\int_0^1 \xi_n(s)^2\,\mathrm{d}s. \tag{29}$$

Now it follows from (27) that

$$\begin{aligned}\int_0^1 \mathbb{E}\,\xi_n(s)^2\,\mathrm{d}s &= \sum_{|k|<n}\Big(1-\frac{|k|}{n}\Big)\int_0^1 H_{0,k}(s,s)\,\mathrm{d}s\\ &= \sum_{|k|<n}\Big(1-\frac{|k|}{n}\Big)\Big(\frac{2}{3} - \mathbb{E}\max(X_0,X_k)\Big). \end{aligned}\tag{30}$$

Under pairwise PQD, $\operatorname{Cov}(X_0,X_k)$ and $2/3 - \mathbb{E}\max(X_0,X_k)$ are non negative, so (28), (29) and (30) give for every $n\ge 1$:

$$0 \le \sum_{|k|<n}\Big(1-\frac{|k|}{n}\Big)\operatorname{Cov}(X_0,X_k) \le \sum_{|k|<n}\Big(1-\frac{|k|}{n}\Big)\Big(\frac{2}{3} - \mathbb{E}\max(X_0,X_k)\Big). \tag{31}$$

The conclusion follows by letting n increase to infinity in (31). □

Lemma 13. *Let $\beta \in [0,1]$. If (X,Y) is a positively quadrant dependent vector with marginals uniform on $[0,1]$, then*

$$0 \leq \frac{2}{\beta+2} - \mathbb{E}\max(X,Y)^{\beta} \leq \beta\Big(\frac{2}{3} - \mathbb{E}\max(X,Y)\Big)^{\beta}. \tag{32}$$

Proof. When $\beta = 0$ or $\beta = 1$, it is obvious. Assume $\beta \in (0,1)$. Observe first that $0 \leq H(u,u) \leq u$. By integrating by parts, then by Jensen inequality we obtain successively:

$$\begin{aligned}
\frac{2}{\beta+2} - \mathbb{E}\max(X,Y)^{\beta} &= \int_0^1 \big(\mathbb{P}(X \leq u, Y \leq u) - u^2\big)\beta u^{\beta-1}\,\mathrm{d}u \\
&= \int_0^1 H(u,u)^{\beta}\beta u^{\beta-1} H(u,u)^{1-\beta}\,\mathrm{d}u \\
&\leq \beta\int_0^1 \big(\mathbb{P}(X \leq u, Y \leq u) - u^2\big)^{\beta}\,\mathrm{d}u \\
&\leq \beta\Big(\int_0^1 \big(\mathbb{P}(X \leq u, Y \leq u) - u^2\big)\,\mathrm{d}u\Big)^{\beta} \\
&= \beta\Big(\frac{2}{3} - \mathbb{E}\max(X,Y)\Big)^{\beta}.
\end{aligned}$$

□

Lemma 14. *Let $\beta \in (0,1]$. If (X,Y) is a positively quadrant dependent vector with marginals uniform on $[0,1]$, then there exists a constant $C = C(\beta)$ such that:*

$$\Big|\mathbb{E}\,|X-Y|^{\beta} - \frac{2}{(\beta+1)(\beta+2)}\Big| \leq C\,\mathrm{Cov}^{\beta/2}(X,Y). \tag{33}$$

Proof. Let Z be a uniform $[0,1]$ distributed random variable, independent of X. An elementary conputation gives

$$\frac{2}{(\beta+1)(\beta+2)} = \mathbb{E}\,|X-Z|^{\beta}.$$

Consider the function

$$g_{\varepsilon}(u) = \begin{cases} |u|^{\beta}, & \text{when } |u| \geq \varepsilon \\ f_{\varepsilon}(u), & \text{when } |u| < \varepsilon, \end{cases}$$

with
$$f_\varepsilon(u) := \varepsilon^\beta \Big\{ \Big(\frac{\beta^2}{8} - \frac{\beta}{4}\Big)\Big(\frac{u}{\varepsilon}\Big)^4 + \Big(\beta - \frac{\beta^2}{4}\Big)\Big(\frac{u}{\varepsilon}\Big)^2 + 1 - \frac{3\beta}{4} + \frac{\beta^2}{8} \Big\}.$$
The function g_ε is of class C^2 with a second derivative bounded and continuous on $\mathbb{R}$. It is easy to see that
$$\sup_{|u|\le\varepsilon} |f_\varepsilon(u)| \le \Big(1 + \frac{\beta}{2}\Big)\varepsilon^\beta,$$
so that $\Big||u|^\beta - g_\varepsilon(u)\Big| \le 5\varepsilon^\beta/2$. Hence, the left hand side of (33) differs from
$$|\,\mathbb{E}\, g_\varepsilon(X-Y) - \mathbb{E}\, g_\varepsilon(X-Z)| =: A_\varepsilon$$
no more than by $5\varepsilon^\beta$. We estimate A_ε by Lemma 10. As $\|g_\varepsilon''\|_\infty = \beta(1-\beta)\varepsilon^{\beta-2}$,
$$A_\varepsilon \le \beta(1-\beta)\varepsilon^{\beta-2} \int_{[0,1]^2} H(s,t)\,\mathrm{d}s\,\mathrm{d}t = \beta(1-\beta)\varepsilon^{\beta-2}\,\mathrm{Cov}(X,Y).$$
Using $\varepsilon = \sqrt{\mathrm{Cov}(X,Y)}$, inequality (33) is achieved. □

Acknowledgment: The authors are grateful to the anonymous referee for his very careful reading of the manuscript and many useful suggestions.

References

[1] Billingsley, P., Convergence of probability measures, Wiley, New-York, 1968.

[2] Doukhan, P., Massart, P., Rio, E., Invariance principles for the empirical measure of a weakly dependent process. Ann. Inst. H. Poincaré, 31 (1995), 393–427.

[3] Esary, J., Proschan, F., Walkup, D., Association of random variables with applications. Ann. Math. Statist. 38 (1967), 1466–1474.

[4] Franche, A., Estimation des valeurs propres de la covariance du processus empirique sous dépendance faible. Application à la statistique de test de Cramér-von Mises. Pub. IRMA Lille 50-V (1999), preprint.

[5] Giné, M.E., Invariant tests for uniformity on compact Riemannian manifolds based on Sobolev norms. Ann. Statist. 3 (1975), 1243–1266.

[6] Henriques, C., Oliveira, P.E., Covariance estimation for associated random variables, Universidade de Coimbra, Departamento de Matemática, pré-publicaçoes 99-03 (1999).

[7] Hermans, M., Rasson J.-P., A new Sobolev test for uniformity on the circle. Biometrika 72 (1985), 698–702.

[8] Lehman, E., Some concepts of dependence. Ann. Math. Statist. 37 (1966), 1137–1153.

[9] Louhichi, S., Weak convergence for empirical processes of associated sequences, preprint Université Paris Sud Orsay 98.36 (1998).

[10] Newman, C.M., A general central limit theorem for FKG systems. Comm. Math. Phys. 91 (1983), 75–80.

[11] Newman, C., M. Asymptotic independence and limit theorems for positively and negatively dependent random variables, in *Inequalities in Statistics and Probability*, IMS Lect. Notes - Monograph Series 5 (1984), 127–140.

[12] Oliveira, P.E., Suquet, Ch., $L^2(0,1)$ weak convergence of the empirical process for dependent variables. Lecture Notes in Statistics 103 (1995), A. Antoniadis and G. Oppenheim (Eds) Wavelets and Statistics, 331-344.

[13] Oliveira, P.E., Suquet, Ch., Weak convergence in $L^p[0,1]$ of the uniform empirical process under dependence, Stat. Probab. Letters 39 (1998) 363–370.

[14] Peetre J., New thoughts on Besov spaces. Duke University Mathematics Series 1, Durham, 1976.

[15] Prohorov, Y.V., Convergence of random processes and limit theorems in probability theory. Theor. Prob. Appl. 1 (1956), 157–214.

[16] Rao, J.S., Some variants of Chi-square for testing uniformity on the circle. Z. Wahrscheinlichkeitstheorie verw. Geb. 22 (1972), 33–44.

[17] Shao, Q.-M., Yu, H., Weak convergence for weighted empirical process of dependent sequences. Ann. Probab. 24 (1996), 2052-2078.

[18] Suquet, Ch., Introduction à l'association. Pub. IRMA Lille, 34-XIII (1994).

[19] Suquet, Ch., Tightness in Schauder decomposable Banach spaces. Amer. Math. Soc. Transl. 193 (1999), Proceedings of the St Petersburg Math. Soc. V, 201–224.

[20] Suquet, Ch., Viano, M.-C., Change point detection in dependent sequences: invariance principles for some quadratic statistics, Math. Methods of Stat. 7 (1998) 157–191.

[21] Triebel H., Theory of function spaces. Birkhäuser, Basel, 1983.

[22] Yu, H., A Glivenko-Cantelli lemma and weak convergence for empirical processes of associated sequences, Probab. Theory Relat. Fields 95(1993), 357-370.