ESAIM: PS
March 2007, Vol. 11, p. 102–114
DOI: 10.1051/ps:2007009

ESAIM: Probability and Statistics
www.edpsciences.org/ps

THE EMPIRICAL DISTRIBUTION FUNCTION FOR DEPENDENT VARIABLES: ASYMPTOTIC AND NONASYMPTOTIC RESULTS IN $\mathbb{L}^p$

Jérôme Dedecker[1] and Florence Merlevède[2]

Abstract. Considering the centered empirical distribution function $F_n - F$ as a variable in $\mathbb{L}^p(\mu)$, we derive non asymptotic upper bounds for the deviation of the $\mathbb{L}^p(\mu)$-norms of $F_n - F$ as well as central limit theorems for the empirical process indexed by the elements of generalized Sobolev balls. These results are valid for a large class of dependent sequences, including non-mixing processes and some dynamical systems.

Mathematics Subject Classification. 60F10, 62G30.

Received September 20, 2005. Revised May 22, 2006.

1. Introduction

Let $(X_i)_{1\leq i\leq n}$ be equidistributed real-valued random variables with common distribution function F. Let F_n be the empirical distribution function $F_n(t) = n^{-1}\sum_{i=1}^n \mathbb{1}_{X_i\leq t}$. Let $1 \leq p < \infty$ and μ be a σ-finite measure on $\mathbb{R}$. Suppose that F satisfies

$$\int_{\mathbb{R}_-} (F(t))^p \mu(\mathrm{d}t) + \int_{\mathbb{R}_+} (1 - F(t))^p \mu(\mathrm{d}t) < \infty. \tag{1.1}$$

Under this assumption, the process $\{t \to F_n(t) - F(t), t \in \mathbb{R}\}$ may be viewed as a random variable with values in the space $\mathbb{L}^p(\mu)$. Let $\|.\|_{p,\mu}$ be the $\mathbb{L}^p$-norm with respect to μ, and define

$$D_{p,n}(\mu) = \Big(\int |F_n(t) - F(t)|^p \mu(\mathrm{d}t)\Big)^{1/p} = \|F_n - F\|_{p,\mu}.$$

When $p = 2$ and $\mu = \mathrm{d}F$, $D^2_{2,n}(\mu)$ is known as the Cramér-von Mises statistics, and is commonly used for testing goodness of fit. When $p = 1$ and λ is the Lebesgue measure on the real line, then $D_{1,n}(\lambda)$ is the $\mathbb{L}^1(\lambda)$-minimal distance between F_n and F, denoted in what follows by $K(F_n, F)$.

It is interesting to write $D_{p,n}(\mu)$ as the supremum of the empirical process over a particular class of functions. The proof of the following lemma is given in the appendix.

Keywords and phrases. Deviation inequalities, weak dependence, Cramér-von Mises statistics, empirical process, expanding maps.

[1] Laboratoire de Statistique Théorique et Appliquée, Université Paris 6, 175 rue du Chevaleret, 75013 Paris, France.
dedecker@ccr.jussieu.fr

[2] Laboratoire de probabilités et modèles aléatoires, UMR 7599, Université Paris 6, 175 rue du Chevaleret, 75013 Paris, France.
merleve@ccr.jussieu.fr

Lemma 1. *Let F and G be two distributions functions satisfying (1.1). If p' is the conjugate exponent of p, let $W_{p',1}(\mu)$ be the set of functions*

$$\Big\{f : f(t) = f(0) + \Big(\int_{[0,t[} f'(x)\mu(\mathrm{d}x)\Big)\mathbb{1}_{t>0} - \Big(\int_{[t,0[} f'(x)\mu(\mathrm{d}x)\Big)\mathbb{1}_{t\le 0},\ \|f'\|_{p',\mu} \le 1\Big\}.$$

For any f in $W_{p',1}(\mu)$, we have

$$\int f\mathrm{d}F - \int f\mathrm{d}G = \mu(f'(F-G)) \quad \textit{so that} \quad \|F-G\|_{p,\mu} = \sup_{f\in W_{p',1}(\mu)} \Big|\int f\mathrm{d}F - \int f\mathrm{d}G\Big|.$$

According to Lemma 1, since $D_{p,n}(\mu) = \|F_n - F\|_{p,\mu}$, we have that

$$D_{p,n}(\mu) = \sup_{f\in W_{p',1}(\mu)} \Big|\frac{1}{n}\sum_{i=1}^{n} f(X_i) - \mathbb{E}(f(X_i))\Big|.$$

If λ_I is the Lebesgue measure on the interval $I \subset \mathbb{R}$, $W_{p',1}(\lambda_I)$ is the space of absolutely continuous functions f such that $\lambda_I(|f'|^{p'}) \le 1$ (hence it contains the unit ball of the Sobolev space of order 1 with respect to $\mathbb{L}^{p'}(\lambda_I)$). Let us now recall what is known about the entropy properties of the class: $W^0_{p',1}(\lambda_I) = \{f - f(0), f \in W_{p',1}(\lambda_I)\}$. If I is compact and if $1 < p' \le \infty$ (or equivalently $1 \le p < \infty$), then, according to Corollary 1 in Birgé and Massart (2000) [3], the space $W^0_{p',1}(\lambda_I)$ is compact with respect to the $\mathbb{L}^\infty$-norm with ε-entropy of order ε^{-1}. Of course this is no longer true if I is not compact, and as far as we know, the entropy properties of these classes have not been studied. However, arguing as in van der Vaart (1994) [17], one can prove that the ε-$\mathbb{L}^r(P)$ bracketing entropy with respect to a probability P is of order ε^{-1} provided $\sum_{n\in\mathbb{Z}}(|n|^r P([n, n+1]))^{1/(1+r)} < \infty$. To our knowledge, no entropy bounds are available for the class $W^0_{p',1}(\mu)$ when μ is not the Lebesgue measure on some interval.

These entropy bounds can be used to prove uniform central limit theorems and maximal inequalities for the empirical process indexed by elements of $W^0_{p',1}(\lambda_I)$ in the iid case (see again van der Vaart (1994) [17], Sect. 4). Some of these results can be extended to the dependent context, but the general results based on entropy methods are only available in a mixing context (see Rem. 6 below). Our aim in this paper is to show that we can obtain asymptotic and non-asymptotic results for the empirical process indexed by elements of $W_{p',1}(\mu)$, for some dependence coefficients which are perfectly adapted to the class $W_{p',1}(\mu)$.

In Section 2, we give a nonasymptotic upper bound for the deviation of $D_{p,n}(\mu)$ when $2 \le p < \infty$. In Section 3 we study the weak convergence of $\sqrt{n}(F_n - F)$ in the spaces $\mathbb{L}^p(\mu)$ when $2 \le p < \infty$, which in turn is equivalent to the weak convergence of the normalized empirical process indexed by elements of the class $W_{p',1}(\mu)$ (see Lems. 2 and 3). In both cases, the conditions are expressed in terms of some natural dependence coefficients, which can be viewed as mixing coefficients restricted to the class $W_{p',1}(\mu)$ (see Def. 1 below). In Section 4, we compare these coefficients to other well known measures of dependence, and we show how they can be computed for two large classes of examples, including many non-mixing processes. In Section 5, we apply our results to the case of iterates of expanding maps of the unit interval, and to the simple example of linear processes.

2. Exponential bounds

We first define the dependence coefficients which naturally appear in this context.

Definition 1. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, let X be a real-valued random variable and let $\mathcal{M}$ be a σ-algebra of $\mathcal{A}$. Let $\|.\|_p$ be the the $\mathbb{L}^p$-norm with respect to $\mathbb{P}$. Denote by $\mathbb{P}_X$ the distribution of X and by

$\mathbb{P}_{X|\mathcal{M}}$ a regular distribution of X given $\mathcal{M}$. Let $F_X(t) = \mathbb{P}_X(]-\infty, t])$ and $F_{X|\mathcal{M}}(t) = \mathbb{P}_{X|\mathcal{M}}(]-\infty, t])$. For $1 \le p \le \infty$ and $1 \le q \le \infty$, define the coefficient

$$\tau_{\mu,p,q}(\mathcal{M}, X) = \Big\| \Big(\int |F_{X|\mathcal{M}}(t) - F_X(t)|^p \mu(\mathrm{d}t) \Big)^{1/p} \Big\|_q = \big\| \, \|F_{X|\mathcal{M}} - F_X\|_{p,\mu} \big\|_q.$$

From Lemma 1, we see that an equivalent definition is

$$\tau_{\mu,p,q}(\mathcal{M}, X) = \Big\| \sup_{f \in W_{p',1}(\mu)} \Big| \int f \mathrm{d}F_{X|\mathcal{M}} - \int f \mathrm{d}F_X \Big| \Big\|_q .$$

Theorem 1. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, and $(X_i)_{1\le i\le n}$ be equidistributed real-valued random variables with common distribution function F. Let $\mathcal{M}_0 = \{\emptyset, \Omega\}$ and let $(\mathcal{M}_k)_{1\le k\le n}$ be an increasing sequence of σ-algebra such that $\sigma(X_i, 1 \le i \le k) \subseteq \mathcal{M}_k$. For any $2 \le p < \infty$, any finite measure μ and any positive x, we have the upper bound*

$$\mathbb{P}\big(\sqrt{n} D_{p,n}(\mu) \ge x\big) \le 2 \exp\left(- \frac{n x^2}{2(p-1) \sum_{i=1}^n \Big(\sum_{k=i}^n \big\| \, \|F_{X_k|\mathcal{M}_i} - F_{X_k|\mathcal{M}_{i-1}}\|_{p,\mu} \big\|_\infty \Big)^2} \right). \tag{2.1}$$

In particular, if

$$C(p, n, \mu) = \sum_{i=1}^n \left(\sum_{k=i}^n \Big(\tau_{\mu,p,\infty}(\mathcal{M}_i, X_k) + \tau_{\mu,p,\infty}(\mathcal{M}_{i-1}, X_k) \Big) \right)^2 ,$$

we have the upper bound

$$\mathbb{P}\big(\sqrt{n} D_{p,n}(\mu) \ge x\big) \le 2 \exp\Big(- \frac{n x^2}{2(p-1) C(p, n, \mu)} \Big). \tag{2.2}$$

Remark 1. Let $\tau_{\mu,p,q}(i) = \max_{1\le k\le n-i} \tau_{\mu,p,q}(\mathcal{M}_k, X_{k+i})$ and $Z_i = \{t \to \mathbb{1}_{X_i \le t} - F(t)\}$. Since

$$\begin{aligned} C(p, n, \mu) &\le \sum_{i=1}^n \Big(\big\| \|Z_1\|_{p,\mu} \big\|_\infty + \sum_{k=1}^{n-i} \tau_{\mu,p,\infty}(k) + \Big(\sum_{k=1}^{n-i+1} \tau_{\mu,p,\infty}(k) \Big) \mathbb{1}_{i>1} \Big)^2 \\ &\le n \Big(\big\| \|Z_1\|_{p,\mu} \big\|_\infty + 2 \sum_{k=1}^{n-1} \tau_{\mu,p,\infty}(k) \Big)^2 , \end{aligned}$$

we obtain from (2.2) that

$$\mathbb{P}\big(\sqrt{n} D_{p,n}(\mu) \ge x\big) \le 2 \exp\left(- \frac{x^2}{2(p-1) \big(\big\| \|Z_1\|_{p,\mu} \big\|_\infty + 2 \sum_{k=1}^{n-1} \tau_{\mu,p,\infty}(k) \big)^2} \right). \tag{2.3}$$

Proof of Theorem 1. Let Z_i be as in Remark 1 and $S_n = \sum_{i=1}^n Z_i$. Clearly, we get that

$$\sqrt{n} D_{p,n}(\mu) = \frac{\|S_n\|_{p,\mu}}{\sqrt{n}} \cdot \tag{2.4}$$

We apply the method of martingale differences, as done in Yurinskii (1974) [19]. Since $\mathbb{E}(S_n|\mathcal{M}_0) = \mathbb{E}(S_n) = 0$, we have that $S_n = \sum_{i=1}^n (\mathbb{E}(S_n|\mathcal{M}_i) - \mathbb{E}(S_n|\mathcal{M}_{i-1}))$. For all $1 \le i \le n$, let $d_{i,n} = \mathbb{E}(S_n|\mathcal{M}_i) - \mathbb{E}(S_n|\mathcal{M}_{i-1})$.

Clearly $d_{i,n}$ is an $\mathcal{M}_i$-measurable random variable with values in $\mathbb{L}^p(\mu)$ such that $\mathbb{E}(d_{i,n}|\mathcal{M}_{i-1}) = 0$ almost surely. From (2.4) and Theorem 3 in Pinelis (1992) [15], we infer that for any positive real x,

$$\mathbb{P}(\sqrt{n}D_{p,n}(\mu) \geq x) \leq 2\exp\Big(-\frac{nx^2}{2(p-1)\,b_n^2}\Big), \quad \text{where } b_n^2 = \sum_{i=1}^n \big\|\|d_{i,n}\|_{p,\mu}\big\|_\infty^2. \tag{2.5}$$

The inequality (2.1) follows by noting that

$$d_{i,n} = \sum_{k=i}^n F_{X_k|\mathcal{M}_i} - F_{X_k|\mathcal{M}_{i-1}}. \tag{2.6}$$

The inequality (2.2) follows from (2.1) by noting that

$$\big\|\|F_{X_k|\mathcal{M}_i} - F_{X_k|\mathcal{M}_{i-1}}\|_{p,\mu}\big\|_\infty \leq \tau_{\mu,p,\infty}(\mathcal{M}_i, X_k) + \tau_{\mu,p,\infty}(\mathcal{M}_{i-1}, X_k).$$ □

Remark 2. The bounds of Theorem 1 are valid for $2 \leq p < \infty$. If p belongs to $[1,2[$, the space $\mathbb{L}^p(\mu)$ is no longer smooth, so that the method of martingale differences in Banach spaces does not work. Hence a reasonable question is: does (2.2) still holds (with possibly different constants) if p belongs to $[1,2[$? This would give a bound in terms of the coefficients $\tau_{\mu,1,\infty}$ which are the weakest among the coefficients $\tau_{\mu,p,\infty}$.

Of course, since $D_{p,n}(\mu) \leq D_{2,n}(\mu)$ for any probability μ and any $1 \leq p \leq 2$, Theorem 1 provides an upper bound for the deviation of $D_{p,n}(\mu)$ in terms of the coefficient $\tau_{\mu,2,\infty}$ (and hence in terms of $\tau_{\mu,1,\infty}$ since $\tau_{\mu,2,\infty}(\mathcal{M}, X) \leq (\tau_{\mu,1,\infty}(\mathcal{M}, X))^{1/2}$). If the X_i' s are in $[0,1]$ and λ_1 is the Lebesgue measure on $[0,1]$, we obtain a bound for the deviation of the Kantorovitch distance $K(F_n, F) = \|F_n - F\|_{1,\lambda_1}$: for any positive x,

$$\mathbb{P}\big(\sqrt{n}K(F_n,F) \geq x\big) \leq \mathbb{P}\big(\sqrt{n}D_{2,n}(\lambda_1) \geq x\big) \leq 2\exp\Big(-\frac{nx^2}{2C(2,n,\lambda_1)}\Big). \tag{2.7}$$

From Remark 1, we also have, for a sequence $(X_i)_{i\geq 1}$ of variables with values in $[0,1]$,

$$\mathbb{P}\big(\sqrt{n}K(F_n,F) \geq x\big) \leq \mathbb{P}\big(\sqrt{n}D_{2,n}(\lambda_1) \geq x\big) \leq 2\exp\left(-\frac{x^2}{2\big(\big\|\|Z_1\|_{2,\lambda_1}\big\|_\infty + 2\sum_{k=1}^{n-1}\tau_{\lambda_1,2,\infty}(k)\big)^2}\right). \tag{2.8}$$

Remark 3. If $\mu = \delta_t$, we are looking for the deviation of $\sqrt{n}|F_n(t) - F(t)|$. Starting from (2.6), we see that $d_{i,n}(t)$ belongs to the interval $[A_i, B_i]$, where A_i and B_i are the $\mathcal{M}_{i-1}$-measurable random variables

$$\begin{aligned} A_i &= -F(t) - \sum_{k=i+1}^n \|\mathbb{E}(Z_k(t)|\mathcal{M}_i)\|_\infty - \sum_{k=i}^n \mathbb{E}(Z_k(t)|\mathcal{M}_{i-1}) \\ B_i &= (1 - F(t)) + \sum_{k=i+1}^n \|\mathbb{E}(Z_k(t)|\mathcal{M}_i)\|_\infty - \sum_{k=i}^n \mathbb{E}(Z_k(t)|\mathcal{M}_{i-1}). \end{aligned}$$

The sum of the lengths $(B_i - A_i)^2$ is then $L(t,n) = \sum_{i=1}^n (1 + 2\sum_{k=i+1}^n \|\mathbb{E}(Z_k(t)|\mathcal{M}_i)\|_\infty)^2$. Applying Azuma's inequality (1967) [1], we obtain

$$\mathbb{P}(\sqrt{n}|F_n(t) - F(t)| > x) \leq 2\exp\Big(-\frac{2nx^2}{L(t,n)}\Big). \tag{2.9}$$

Note that, for $\mu = \delta_t$, $\|\mathbb{E}(Z_k(t)|\mathcal{M}_i)\|_\infty = \tau_{\mu,p,\infty}(\mathcal{M}_i, X_k)$, for any $1 \le p \le \infty$. For this choice of μ, the bound (2.9) is much more precise than (2.2). In view of (2.9) one can wonder if (2.2) holds for any probability μ with

$$L(p,n,\mu) = \sum_{i=1}^{n}\Big(1 + 2\sum_{k=i+1}^{n} \tau_{\mu,p,\infty}(X_k, \mathcal{M}_i)\Big)^2$$

instead of $4(p-1)C(p,n,\mu)$. We have no idea of how to prove such a bound. It is probably delicate, for if it is true then $\mathbb{P}(\sqrt{n}D_{p,n}(\mathrm{d}F) \ge x) \le 2\exp(-2x^2)$ for iid variables. If this bound holds for any $2 \le p < \infty$, then it must hold for $p = \infty$, which is Massart's bound (1990) [11] for the deviation of Kolmogorov-Smirnov statistics.

3. Empirical process indexed in Sobolev balls

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $T : \Omega \mapsto \Omega$ be a bijective bimeasurable transformation preserving the probability $\mathbb{P}$. An element A of $\mathcal{A}$ is said to be invariant if $T(A) = A$. We denote by $\mathcal{I}$ the σ-algebra of all invariant sets.

Let $\mathcal{M}_0$ be a σ-algebra of $\mathcal{A}$ satisfying $\mathcal{M}_0 \subseteq T^{-1}(\mathcal{M}_0)$ and define the nondecreasing filtration $(\mathcal{M}_i)_{i\in\mathbb{Z}}$ by $\mathcal{M}_i = T^{-i}(\mathcal{M}_0)$. Let X_0 be an $\mathcal{M}_0$-measurable real-valued random variable and define the sequence $(X_i)_{i\in\mathbb{Z}}$ by $X_i = X_0 \circ T^i$. Define the coefficient $\tau_{\mu,p,q}(i)$ of the sequence $(X_i)_{i\in\mathbb{Z}}$ by

$$\tau_{\mu,p,q}(i) = \tau_{\mu,p,q}(\mathcal{M}_0, X_i).$$

Let F be the distribution function of X_0 and F_n be the empirical distribution function. Let G_n be the centered and normalized empirical measure $G_n = \sqrt{n}(\mathrm{d}F_n - \mathrm{d}F)$, and $\mathcal{F}$ be a class of measurable functions from $\mathbb{R}$ to $\mathbb{R}$. The space $\ell^\infty(\mathcal{F})$ is the space of all functions z from $\mathcal{F}$ to $\mathbb{R}$ such that $\|z\|_{\mathcal{F}} = \sup_{f\in\mathcal{F}} |z(f)|$ is finite.

Let $1 \le p < \infty$, let p' be the conjugate exponent of p, and let μ be any σ-finite measure such that (1.1) holds. We shall see that the convergence in distribution of G_n indexed by the elements of $W_{p',1}(\mu)$ is equivalent to the convergence in distribution of $\sqrt{n}(F_n - F)$ in $\mathbb{L}^p(\mu)$. We first define an appropriate isometry between $\mathbb{L}^p(\mu)$ and a subspace of $\ell^\infty(W_{p',1}(\mu))$.

Lemma 2. *Given f in $W_{p',1}(\mu)$, let f' be as in Lemma 1. Let $h : \mathbb{L}^p(\mu) \mapsto \ell^\infty(W_{p',1}(\mu))$ be such that $h(g) = \{\mu(f'g), f \in W_{p',1}(\mu)\}$. The function h is an isometry from $\mathbb{L}^p(\mu)$ to $h(W_{p',1}(\mu)) = G_{p'}(\mu)$. In particular, $G_{p'}(\mu)$ is a separable Banach space.*

Proof of Lemma 2. By duality $\|h(g_1)-h(g_2)\|_{W_{p',1}(\mu)} = \|g_1-g_2\|_{p,\mu}$. □

Now, by Lemma 1, $\{G_n(f), f \in W_{p',1}(\mu)\} = h(\sqrt{n}(F_n - F))$. Consequently, under (1.1), the empirical process G_n indexed by the elements of $W_{p',1}(\mu)$ is a random variable with values in $G_{p'}(\mu)$. In addition, if $\gamma : G_{p'}(\mu) \mapsto \mathbb{R}$ is continuous, then $\gamma \circ h : \mathbb{L}^p(\mu) \mapsto \mathbb{R}$ is also continuous. It follows immediately that

Lemma 3. *If (1.1) holds, then the sequence $\{G_n(f), f \in W_{p',1}(\mu)\}$ converges in distribution in the space $G_{p'}(\mu)$ if and only if the sequence $\sqrt{n}(F_n - F)$ converges in distribution in the space $\mathbb{L}^p(\mu)$.*

Consequently, central limit theorems for the empirical process indexed by the elements of $W_{p',1}(\mu)$ can be deduced from central limit theorems for Banach-valued random variables. This approach leads to Theorem 2 below.

Theorem 2. *Define the function F_μ by: $F_\mu(x) = \mu([0, x[)$ if $x \ge 0$ and $F_\mu(x) = -\mu([x, 0[)$ if $x \le 0$. Define also the nonnegative random variable $Y_{p,\mu} = |F_\mu(X_0)|^{1/p}$ and assume that $\|Y_{p,\mu}\|_2 < \infty$. Consider the three conditions*

1. *p belongs to $[2, \infty[$, and* $\sum_{k>0} \tau_{\mu,p,2}(k) < \infty$.
2. *$p = 2$, $\mu(\mathbb{R}) < \infty$, and* $\sum_{k>0} \tau_{\mu,2,1}(k) < \infty$.

3. $p = 2$, $F_{X_0|\mathcal{M}_{-\infty}} = F$, *and* $\sum_{k>0} \big\| \|F_{X_k|\mathcal{M}_0} - F_{X_k|\mathcal{M}_{-1}}\|_{2,\mu}\big\|_2 < \infty$.

If one of these conditions holds, then the sequence $\{G_n(f), f \in W_{p',1}(\mu)\}$ *converges in distribution in the space* $G_{p'}(\mu)$ *to a random variable whose conditional distribution with respect to* $\mathcal{I}$ *is that of a zero-mean Gaussian process with covariance function*

$$\Gamma(f,g) = \sum_{k\in\mathbb{Z}} \mathrm{Cov}(f(X_0), g(X_k)|\mathcal{I}).$$

Remark 4. For $p = 2$, 1 implies 3. Note also that, if $\mu(\mathbb{R}) < \infty$, then $Y_{p,\mu}$ is bounded by $\mu(\mathbb{R})^{1/p}$.

Remark 5. According to Remark 6 in Dedecker and Merlevède (2003) [5], by noting that $\| \|\mathbb{E}(Z_k|\mathcal{M}_0)\|_{2,\mu}\|_2 = \tau_{\mu,2,2}(k)$ with Z_k defined in Remark 1, the condition 3 is realized if there exists a sequence $(L_k)_{k>0}$ of positive numbers such that

$$\sum_{i>0}\Big(\sum_{k=1}^{i} L_k\Big)^{-1} < \infty \quad \text{and} \quad \sum_{k>0} L_k \tau^2_{\mu,2,2}(k) < \infty. \tag{3.1}$$

In particular, (3.1) (hence 3) holds if

$$\sum_{k>0} \frac{\tau_{\mu,2,2}(k)}{\sqrt{k}} < \infty. \tag{3.2}$$

In addition, since $\tau_{\mu,2,2}(k) \le (\tau_{\mu,1,1}(k))^{1/2}$, it follows that (3.2) holds as soon as

$$\sum_{k>0} \frac{\sqrt{\tau_{\mu,1,1}(k)}}{\sqrt{k}} < \infty.$$

Proof of Theorem 2. We first prove that any of the conditions of Theorem 2 implies the weak convergence of $\sqrt{n}(F_n - F)$ in $\mathbb{B} = \mathbb{L}^p(\mu)$. A random variable Z is in $\mathbb{L}^2(\mathbb{B})$ if $\big\|\|Z\|_{p,\mu}\big\|_2 < \infty$. Let $Z_i = \{t \to \mathbf{1}_{X_i \le t} - F(t)\}$. Clearly,

$$\|Z_i\|_{p,\mu} \le \Big(\int_{]-\infty,0[} (\mathbf{1}_{X_i\le t})^p \mu(\mathrm{d}t) + \int_{[0,\infty[} (1-\mathbf{1}_{X_i\le t})^p \mu(\mathrm{d}t)\Big)^{1/p} + \Big(\int_{]-\infty,0[} (F(t))^p \mu(\mathrm{d}t) + \int_{[0,\infty[} (1-F(t))^p \mu(\mathrm{d}t)\Big)^{1/p},$$

so that $\|Z_i\|_{p,\mu} \le |F_\mu(X_i)|^{1/p} + \mathbb{E}(|F_\mu(X_i)|^{1/p})$ and $\big\|\|Z_i\|_{p,\mu}\big\|_2 \le 2\|Y_{p,\mu}\|_2$.

Case 1. From the non ergodic version of Woyczyński's result (1975) [18], we know that if an $\mathcal{M}_0$-measurable r.v. M_0 in $\mathbb{L}^2(\mathbb{B})$ is such that $\mathbb{E}(M_0|\mathcal{M}_{-1}) = 0$, then $n^{-1/2}(M_0 \circ T + \cdots + M_0 \circ T^n)$ converges in distribution to a random variable M. Now if

$$U_0 = M_0 + N_0 - N_0 \circ T \quad \text{and} \quad U_i = U_0 \circ T^i \tag{3.3}$$

with N_0 in $\mathbb{L}^2(\mathbb{B})$, we easily infer that $n^{-1/2}(U_1 + \cdots + U_n)$ also converges in distribution to M. Assume that $\sum_{k>0} \| \|\mathbb{E}(U_k|\mathcal{M}_0)\|_{p,\mu}\|_2 < \infty$. From a well known decomposition of Gordin (1969) [10], we have

$$U_0 = \sum_{i=0}^{\infty} \mathbb{E}(U_i|\mathcal{M}_0) - \mathbb{E}(U_i|\mathcal{M}_{-1}) + \sum_{i=0}^{\infty} \mathbb{E}(U_i|\mathcal{M}_{-1}) - \sum_{i=1}^{\infty} \mathbb{E}(U_i|\mathcal{M}_0), \tag{3.4}$$

provided U_0 is $\mathcal{M}_0$-measurable. Hence, (3.3) holds with $M_0 = \sum_{i=0}^{\infty} \mathbb{E}(U_i|\mathcal{M}_0) - \mathbb{E}(U_i|\mathcal{M}_{-1})$ and the coboundary $N_0 = \sum_{i=0}^{\infty} \mathbb{E}(U_i|\mathcal{M}_{-1})$. Applying the preceding remarks to the random variable $U_i = Z_i$, we infer that $\sqrt{n}(F_n - F) = n^{-1/2}(Z_1 + \cdots + Z_n)$ converges in distribution in $\mathbb{B}$ as soon as $\sum_{k\ge 0} \| \|\mathbb{E}(Z_k|\mathcal{M}_0)\|_{p,\mu}\|_2$ is finite. To conclude, note that $\| \|\mathbb{E}(Z_k|\mathcal{M}_0)\|_{p,\mu}\|_2 = \tau_{\mu,p,2}(k)$.

Case 2. The fact that $\sqrt{n}(F_n - F)$ converges in distribution in $\mathbb{B}$ follows from Corollary 2(δ) in Dedecker and Merlevède (2003) [5], by noting that $\|Z_0\|_{2,\mu} \leq 2\mu(\mathbb{R})^{1/2}$ and that $\| \|\mathbb{E}(Z_k|\mathcal{M}_0)\|_{2,\mu}\|_1 = \tau_{\mu,2,1}(k)$.

Case 3. The fact that $\sqrt{n}(F_n - F)$ converges in distribution in $\mathbb{B}$ follows from Corollary 3 in Dedecker and Merlevède (2003) [5] (the condition $F_{X_0|\mathcal{M}_{-\infty}} = F$ means that $\mathbb{E}(Z_0|\mathcal{M}_{-\infty}) = 0$ a.s.).

The operator Γ. It remains to identify Γ. Let $(f_i)_{1\leq i\leq k}$ be some functions of $W_{p',1}(\mu)$ and $(\alpha_i)_{1\leq i\leq k}$ some real numbers. We shall prove that $G_n(\alpha_1 f_1 + \cdots + \alpha_k f_k)$ converges to a random variable whose conditional distribution with respect to $\mathcal{I}$ is that of a Gaussian random variable with variance $\sum_{i=1}^n \sum_{j=1}^n \alpha_i\alpha_j\Gamma(f_i, f_j)$.

Let $h = \alpha_1 f_1 + \cdots + \alpha_k f_k$. In case 1, note that, by Lemma 1

$$\|\mathbb{E}(h(X_k)|\mathcal{M}_0) - \mathbb{E}(h(X_k))\|_2 \leq \Big(\sum_{i=1}^n |\alpha_i|\Big) \| \|\mathbb{E}(Z_k|\mathcal{M}_0)\|_{p,\mu}\|_2 \tag{3.5}$$

and the result follows from Theorem 1 in Dedecker and Rio (2000) [8].

In case 2, note that

$$\|(h(X_0) - \mathbb{E}(h(X_0)))\mathbb{E}(h(X_k) - \mathbb{E}(h(X_k))|\mathcal{M}_0)\|_1 \leq \Big(\sum_{i=1}^n |\alpha_i|\Big)^2 \mathbb{E}(\|Z_0\|_{2,\mu}\|\mathbb{E}(Z_k|\mathcal{M}_0)\|_{2,\mu}).$$

Since $\|Z_0\|_{2,\mu} \leq 2\mu(\mathbb{R})^{1/2}$ and $\| \|\mathbb{E}(Z_k|\mathcal{M}_0)\|_{2,\mu}\|_1 = \tau_{\mu,2,1}(k)$, we infer that

$$\sum_{k\geq 1} \|(h(X_0) - \mathbb{E}(h(X_0)))\mathbb{E}(h(X_k) - \mathbb{E}(h(X_k))|\mathcal{M}_0)\|_1 < \infty,$$

and the result follows from Theorem 1 in Dedecker and Rio (2000) [8].

In case 3, the result follows from (3.5) by applying Corollary 3 and Remark 6 in Dedecker and Merlevède (2003) [5] with $\mathbb{H} = \mathbb{R}$ (see Rem. 5 of the same paper for the expression of the covariance). □

4. Comparison of coefficients and examples

4.1. Some upper bounds

Let λ be the Lebesgue measure on $\mathbb{R}$. The four coefficients

$$\tau(\mathcal{M}, X) = \tau_{\lambda,1,1}(\mathcal{M}, X) \qquad \varphi(\mathcal{M}, X) = \tau_{\lambda,1,\infty}(\mathcal{M}, X)$$
$$\beta(\mathcal{M}, X) = \tau_{\lambda,\infty,1}(\mathcal{M}, X) \qquad \phi(\mathcal{M}, X) = \tau_{\lambda,\infty,\infty}(\mathcal{M}, X)$$

have been introduced and studied in Dedecker and Prieur (2004) [6] and (2005) [7]. The authors have shown that these coefficents can be easily computed in many situations. The coefficients $\beta(\mathcal{M}, X)$ and $\phi(\mathcal{M}, X)$ are weaker than the usual β-mixing and ϕ-mixing coefficients, and they can be computed also for many non mixing models. The following lemma shows how to compare $\tau_{\mu,p,q}(\mathcal{M}, X)$ with $\tau(\mathcal{M}, X), \varphi(\mathcal{M}, X), \beta(\mathcal{M}, X)$ and $\phi(\mathcal{M}, X)$.

Lemma 4. *Let X be a real-valued random variable and $\mathcal{M}$ a σ-algebra of $\mathcal{A}$.*

1. *For any p, q in $[1, \infty]$ and any finite measure μ, $\tau_{\mu,p,q}(\mathcal{M}, X) \leq \mu(\mathbb{R})^{1/p}\phi(\mathcal{M}, X)$.*
2. *For any p, q in $[1, \infty]$ and any finite measure μ, $\tau_{\mu,p,q}(\mathcal{M}, X) \leq \mu(\mathbb{R})^{1/p}\beta(\mathcal{M}, X)^{1/q}$.*
3. *If $t \to \mu(] - \infty, t])$ is K-lipschitz, then for any p, q in $[1, \infty]$, $\tau_{\mu,p,q}(\mathcal{M}, X) \leq (K\varphi(\mathcal{M}, X))^{1/p}$.*
4. *If $t \to \mu(] - \infty, t])$ is K-lipschitz, then for any p in $[1, \infty]$ and $q \leq p$, $\tau_{\mu,p,q}(\mathcal{M}, X) \leq (K\tau(\mathcal{M}, X))^{1/p}$.*

Remark 6. Let $(X_i)_{i\in\mathbb{Z}}$ and $\mathcal{M}_0$ be as in Section 3, and let μ be some finite measure. From Lemma 4 and Theorem 2, we infer that $\sqrt{n}(F_n - F)$ converges in distribution in $\mathbb{L}^p(\mu)$ as soon as

1. p belongs to $[2, \infty[$ and $\sum_{k>0} \phi(\mathcal{M}_0, X_k) < \infty$;

2. p belongs to $[2, \infty[$ and $\sum_{k>0} \sqrt{\beta(\mathcal{M}_0, X_k)} < \infty$;
3. $p = 2$ and $\sum_{k>0} \beta(\mathcal{M}_0, X_k) < \infty$;
4. $p = 2$ and $\displaystyle\sum_{k>0} \frac{\phi(\mathcal{M}_0, X_k)}{\sqrt{k}} < \infty$.

Arguing as in Dedecker and Merlevède (2003, p. 250) [5], one can prove that

$$\tau_{\mu,2,1}(k) \leq 18\mu(\mathbb{R})\alpha(\mathcal{M}_0, \sigma(X_k)), \tag{4.1}$$

where $\alpha(\mathcal{A}, \mathcal{B})$ is the strong mixing coefficient of Rosenblatt between two σ-algebras $\mathcal{A}$ and $\mathcal{B}$. Hence, we obtain from the condition 2 of Theorem 2 that $\sqrt{n}(F_n - F)$ converges in distribution in $\mathbb{L}^2(\mu)$ as soon as $\sum_{k>0} \alpha(\mathcal{M}_0, \sigma(X_k)) < \infty$. Note that, when $\mu = \lambda_1$ where λ_1 is the Lebesgue measure on $[0, 1]$, Oliveira and Suquet (1995) [13] obtained the same result under the slightly stronger condition $\sum_{k>0} \alpha(\mathcal{M}_0, \sigma(X_i, i \geq k)) < \infty$ (this condition implies that the sequence $(X_i)_{i\in\mathbb{Z}}$ is ergodic, so that the limiting process is Gaussian). In Oliveira and Suquet (1998) [14], they proved the convergence of $\sqrt{n}(F_n - F)$ in $\mathbb{L}^p(\lambda_1)$ for $p > 2$ under the condition $\alpha(\mathcal{M}_0, \sigma(X_i, i \geq k)) = O(n^{-p/2-\epsilon})$ for some positive ϵ (in both papers, the authors also consider the case of associated random variables). In this case, since $\tau_{\lambda_1,p,2}(k) \leq (\tau_{\lambda_1,2,1}(k))^{2/p}$, we infer from the condition 1 of Theorem 2 and (4.1), that the convergence in $\mathbb{L}^p(\lambda_1)$ for $p > 2$ holds as soon as $\sum_{k>0} \big(\alpha(\mathcal{M}_0, \sigma(X_k))\big)^{2/p} < \infty$ which improves on the condition obtained in Oliveira and Suquet. Of course, if $\mu = \lambda_1$, one can also apply the result of Doukhan *et al.* (1995) [9]: since for $1 < p' \leq \infty$ the ε-entropy of the class $W^0_{p',1}(\lambda_1)$ with respect to the $\mathbb{L}^\infty$-norm is of order ε^{-1}, it follows that $\sqrt{n}(F_n - F)$ converges in distribution in $\mathbb{L}^p(\lambda_1)$ for any $1 \leq p < \infty$ as soon as $\sum_{k>0} \beta(\mathcal{M}_0, \sigma(X_i, i \geq k)) < \infty$, where $\beta(\mathcal{A}, \mathcal{B})$ is the β-mixing coefficient of Rozanov and Volkonskii between two σ-algebras $\mathcal{A}$ and $\mathcal{B}$ (recall that $2\alpha(\mathcal{A}, \mathcal{B}) \leq \beta(\mathcal{A}, \mathcal{B})$).

Proof of Lemma 4. Item 1 is clear. Item 2 follows from the inequality

$$\big\| \|F_{X|\mathcal{M}} - F_X\|_{p,\mu} \big\|_q \leq \mu(\mathbb{R})^{1/p} \Big\| \sup_{t\in\mathbb{R}} |F_{X|\mathcal{M}}(t) - F_X(t)| \Big\|_1^{1/q}.$$

To prove the items 3 and 4, note that, if $t \to \mu(] - \infty, t])$ is K-lipschitz,

$$\big\| \|F_{X|\mathcal{M}} - F_X\|_{p,\mu} \big\|_q \leq \Big(\mathbb{E}\Big(\Big(\int |F_{X|\mathcal{M}}(t) - F_X(t)|^p K \mathrm{d}t\Big)^{q/p}\Big)\Big)^{1/q}.$$

For $q = \infty$, we obtain

$$\big\| \|F_{X|\mathcal{M}} - F_X\|_{p,\mu} \big\|_\infty \leq \Big(K \Big\| \int |F_{X|\mathcal{M}}(t) - F_X(t)|^p \mathrm{d}t \Big\|_\infty\Big)^{1/p} \leq (K\varphi(\mathcal{M}, X))^{1/p}.$$

If $q \leq p$, point 4 follows from Jensen's inequality and Fubini:

$$\big\| \|F_{X|\mathcal{M}} - F_X\|_{p,\mu} \big\|_q \leq \Big(\int \mathbb{E}(|F_{X|\mathcal{M}}(t) - F_X(t)|^p) K \mathrm{d}t\Big)^{1/p} \leq \Big(K \int \|F_{X|\mathcal{M}}(t) - F_X(t)\|_1 \mathrm{d}t\Big)^{1/p}. \qquad \square$$

Before giving detailed examples, we state the following useful bounds for $\tau_{\mu,p,q}(\mathcal{M}, X))$ when μ is a probability measure and X has a continuous distribution function F.

Lemma 5. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, X a real-valued random variable and $\mathcal{M}$ a σ-algebra of $\mathcal{A}$. Assume that X has a continuous distribution function F. If X^* is a random variable distributed as X and independent of $\mathcal{M}$ then, for any $y \in [0, 1]$, any probability measure μ and any p, q in $[1, \infty]$, we have that*

$$\tau_{\mu,p,q}(\mathcal{M}, X) \leq y + \|\mathbb{E}(\mathbf{1}_{|F(X)-F(X^*)|>y} | \mathcal{M})\|_q. \tag{4.2}$$

In particular, taking $y = \|F(X) - F(X^*)\|_\infty$ in the previous inequality, we obtain that

$$\tau_{\mu,p,q}(\mathcal{M}, X)) \le \|F(X) - F(X^*)\|_\infty.$$

Now if w is the modulus of continuity of F, then , for any any $x > 0$ and any $r < \infty$ such that $rq \in [1, \infty]$,

$$\tau_{\mu,p,q}(\mathcal{M}, X)) \le w(x) + \left(\frac{\|X - X^*\|_{qr}}{x}\right)^r \quad \text{and} \quad \tau_{\mu,p,q}(\mathcal{M}, X)) \le w(\|X - X^*\|_\infty). \tag{4.3}$$

In particular, if X has a density bounded by K, we have that

$$\tau_{\mu,p,q}(\mathcal{M}, X)) \le C(r)(K\|X - X^*\|_{qr})^{r/(r+1)} \quad \text{and} \quad \tau_{\mu,p,q}(\mathcal{M}, X)) \le K\|X - X^*\|_\infty, \tag{4.4}$$

with $C(r) = r^{1/(r+1)} + r^{-r/(r+1)}$ (note that $C(r) \le 2$ and $C(\infty) = 1$).

Proof of Lemma 5. Note first that, since μ is a probability measure,

$$\tau_{\mu,p,q}(\mathcal{M}, X)) \le \Big\| \sup_{t\in\mathbb{R}} |F_{X|\mathcal{M}}(t) - F(t)| \Big\|_q.$$

Let $Y = F(X)$ and $Y^* = F(X^*)$. According to Lemma 3 in Dedecker and Prieur (2005) [7], we have that

$$\sup_{t\in\mathbb{R}} |F_{X|\mathcal{M}}(t) - F(t)| = \sup_{t\in\mathbb{R}} |F_{Y|\mathcal{M}}(t) - F_Y(t)|.$$

Since F is continuous, Y is uniformly distributed over $[0, 1]$. From inequality (3.3) in Dedecker and Prieur (2005) [7], we obtain for any $y \in [0, 1]$,

$$|F_{Y|\mathcal{M}}(t) - t| \le y + \mathbb{E}(\mathbf{I}_{|Y-Y^*|>y}|\mathcal{M}),$$

and (4.2) follows. Now, since $|F(X) - F(X^*)| \le w(|X - X^*|)$ we obtain that

$$\tau_{\mu,p,q}(\mathcal{M}, X)) \le w(x) + \|\mathbb{E}(\mathbf{I}_{w(|F(X)-F(X^*)|)>w(x)}|\mathcal{M})\|_q \le w(x) + \|\mathbb{E}(\mathbf{I}_{|X-X^*|>x}|\mathcal{M})\|_q.$$

Applying Markov inequality at order r, we obtain (4.3). Finally, we prove (4.4) by noting that $w(x) \le Kx$ and by minimizing in x. □

4.2. **Example 1: causal functions of stationary sequences**

Let $(\xi_i)_{i\in\mathbb{Z}}$ be a stationary sequence of random variables with values in a measurable space $\mathcal{X}$. Assume that there exists a function H defined on a subset of $\mathcal{X}^{\mathbb{N}}$, with values in $\mathbb{R}$ and such that $H(\xi_0, \xi_{-1}, \xi_{-2}, \ldots,)$ is defined almost surely. The stationary sequence $(X_n)_{n\in\mathbb{Z}}$ defined by $X_n = H(\xi_n, \xi_{n-1}, \xi_{n-2}, \ldots)$ is called a causal function of $(\xi_i)_{i\in\mathbb{Z}}$.

Assume that there exists a stationary sequence $(\xi'_i)_{i\in\mathbb{Z}}$ distributed as $(\xi_i)_{i\in\mathbb{Z}}$ and independent of $(\xi_i)_{i\le 0}$. Define $X^*_n = H(\xi'_n, \xi'_{n-1}, \xi'_{n-2}, \ldots)$. Clearly X^*_n is independent of $\mathcal{M}_0 = \sigma(\xi_i, i \le 0)$ and distributed as X_n. For any $p \ge 1$ (p may be infinite) define the sequence $(\delta_{i,p})_{i>0}$ by

$$\|X_i - X^*_i\|_p = \delta_{i,p}. \tag{4.5}$$

From Lemma 5, we infer that, if μ is a probability measure and X_0 has a density bounded by K, then for any p, q in $[1, \infty]$ and any r such that $rq \in [1, \infty]$,

$$\tau_{\mu,p,q}(\mathcal{M}_0, X_k) \le C(r)(K\delta_{k,qr})^{r/(r+1)} \quad \text{and} \quad \tau_{\mu,p,q}(\mathcal{M}_0, X_k) \le K\delta_{k,\infty}. \tag{4.6}$$

In particular, these results apply to the case where the sequence $(\xi_i)_{i\in\mathbb{Z}}$ is β-mixing in the sense of Rozanov and Volkonskii. Let $\widetilde{\xi}_k = (\xi_k, \xi_{k+1}, \ldots)$, $\mathcal{F}_0 = \sigma(\xi_i, i \le 0)$ and $\mathcal{G}_k = \sigma(\xi_i, i \ge k)$. Let $\mathbb{P}_{\widetilde{\xi}_k}$ be the distribution of $\widetilde{\xi}_k$, and let $\mathbb{P}_{\widetilde{\xi}_k|\mathcal{F}_0}$ be a conditional distribution of $\widetilde{\xi}_k$ given $\mathcal{F}_0$. According to Theorem 4.4.7 in Berbee (1979) [2], if Ω is rich enough, there exists $(\xi'_i)_{i\in\mathbb{Z}}$ distributed as $(\xi_i)_{i\in\mathbb{Z}}$ and independent of $(\xi_i)_{i\le 0}$ such that

$$\mathbb{P}(\xi_i \ne \xi'_i \text{ for some } i \ge k) = \frac{1}{2}\mathbb{E}(\|\mathbb{P}_{\widetilde{\xi}_k|\mathcal{F}_0} - \mathbb{P}_{\widetilde{\xi}_k}\|_v) = \beta(\mathcal{F}_0, \mathcal{G}_k),$$

where $\|\cdot\|_v$ is the variation norm. If the sequence $(\xi_i)_{i\in\mathbb{Z}}$ is iid, it suffices to take $\xi'_i = \xi_i$ for $i > 0$ and $\xi'_i = \xi''_i$ for $i \le 0$, where $(\xi''_i)_{i\in\mathbb{Z}}$ is an independent copy of $(\xi_i)_{i\in\mathbb{Z}}$.

Application: causal linear processes. In that case $X_n = \sum_{j\ge 0} a_j \xi_{n-j}$. For any $p \ge 1$, we have that

$$\delta_{i,p} \le \sum_{j\ge 0} |a_j| \|\xi_{i-j} - \xi'_{i-j}\|_p \le \|\xi_0 - \xi'_0\|_p \sum_{j\ge i} |a_j| + \sum_{j=0}^{i-1} |a_j| \|\xi_{i-j} - \xi'_{i-j}\|_p.$$

From Proposition 2.3 in Merlevède and Peligrad (2002) [12], we obtain that

$$\delta_{i,p} \le \|\xi_0 - \xi'_0\|_p \sum_{j\ge i} |a_j| + \sum_{j=0}^{i-1} |a_j| \Big(2^p \int_0^{\beta(\mathcal{F}_0, \mathcal{G}_{i-j})} Q^p_{\xi_0}(u)\Big)^{1/p} \mathrm{d}u.$$

where Q_{ξ_0} is the cadlag inverse of the tail function $x \to \mathbb{P}(|\xi_0| > x)$.

If the sequence $(\xi_i)_{i\in\mathbb{Z}}$ is iid, it follows that $\delta_{i,p} \le \|\xi_0 - \xi'_0\|_p \sum_{j\ge i} |a_j|$. Moreover, for $p = 2$ we have exactly $\delta_{i,2} = (2\mathrm{Var}(\xi_0) \sum_{j\ge i} a_j^2)^{1/2}$. For instance, if $a_i = 2^{-i-1}$ and $\xi_0 \sim \mathcal{B}(1/2)$, then $\delta_{i,\infty} \le 2^{-i}$. Since X_0 is uniformly distributed over $[0,1]$, we have $\tau_{\mu,p,q}(\mathcal{M}_0, X_i) \le 2^{-i}$. Recall that this sequence is not strongly mixing in the sense of Rosenblatt.

4.3. **Example 2: iterated random functions**

Let $(X_n)_{n\ge 0}$ be a real-valued stationary Markov chain, such that $X_n = F(X_{n-1}, \xi_n)$ for some measurable function F and some iid sequence $(\xi_i)_{i>0}$ independent of X_0. Let X_0^* be a random variable distributed as X_0 and independent of $(X_0, (\xi_i)_{i>0})$. Define $X_n^* = F(X_{n-1}^*, \xi_n)$. The sequence $(X_n^*)_{n\ge 0}$ is distributed as $(X_n)_{n\ge 0}$ and independent of X_0. Let $\mathcal{M}_i = \sigma(X_j, 0 \le j \le i)$. As in Example 1, define the sequence $(\delta_{i,p})_{i>0}$ by (4.5). From Lemma 5, we infer that, if μ is a probability measure and X_0 has a density bounded by K, then for any p, q in $[1, \infty]$ and any r such that $rq \in [1, \infty]$, the bounds given in (4.6) hold.

Let ν be the distribution of X_0 and $(X_n^x)_{n\ge 0}$ the chain starting from $X_0^x = x$. With these notations, we have

$$\delta^p_{i,p} = \iint \|X_i^x - X_i^y\|_p^p \mu(\mathrm{d}x)\mu(\mathrm{d}y).$$

For instance, if there exists a sequence $(d_{i,p})_{i\ge 0}$ of positive numbers such that

$$\|X_i^x - X_i^y\|_p \le d_{i,p}|x - y|,$$

then $\delta_{i,p} \le d_{i,p}\|X_0 - X_0^*\|_p$. In the usual case where $\|F(x, \xi_0) - F(y, \xi_0)\|_p \le \kappa|x - y|$ for some $\kappa < 1$, we can take $d_{i,p} = \kappa^i$.

An important example is $X_n = f(X_{n-1}) + \xi_n$ for some κ-lipschitz function f. If X_0 has a moment of order p, then $\delta_{i,p} \le \kappa^i \|X_0 - X_0^*\|_p$. In particular, if X_0 is bounded and has a density bounded by K then $\tau_{\mu,p,q}(\mathcal{M}_0, X_i) \le 2K\|X_0\|_\infty \kappa^i$.

5. Applications

5.1. Iterates of expanding maps

Let $I = [0,1]$, T be a map from I to I and define $X_i = T^i$. If the probability π is invariant by T, the sequence $(X_i)_{i\geq 0}$ of random variables from (I,π) to I is strictly stationary. Denote by $\|g\|_{1,\lambda_1}$ the $\mathbb{L}^1$-norm with respect to the Lebesgue measure λ_1 on $[0,1]$ and by $\|\nu\| = |\nu|(I)$ the total variation of a signed measure ν. Let $\mathcal{M}_n = \sigma(X_i, i \geq n)$. In many interesting cases (see Sect. 4.4 in Dedecker and Prieur (2005) [7]), one can prove that, for any bounded variation function h on I and any integrable $\mathcal{M}_n$-measurable random variable Y,

$$|\mathrm{Cov}(h(X_0), Y)| \leq a_n \|Y\|_1 (\|h\|_{1,\lambda_1} + \|\mathrm{d}h\|), \tag{5.1}$$

for some nonincreasing sequence a_n tending to zero as n tends to infinity. Note that if (5.1) holds, then $|\mathrm{Cov}(h(X_0), Y)| = |\mathrm{Cov}(h(X_0) - h(0), Y)| \leq a_n \|Y\|_1 (\|h - h(0)\|_{1,\lambda_1} + \|\mathrm{d}h\|)$. Since $\|h - h(0)\|_{1,\lambda_1} \leq \|\mathrm{d}h\|$, we obtain that

$$|\mathrm{Cov}(h(X_0), Y)| \leq 2a_n \|Y\|_1 \|\mathrm{d}h\|. \tag{5.2}$$

From Lemma 4 in Dedecker and Prieur (2005) [7], we infer that if the inequality (5.2) holds, then $\phi(\mathcal{M}_n, X_0) \leq 2a_n$. Applying Lemma 4, we infer that $\tau_{\lambda_1,p,\infty}(\mathcal{M}_n, X_0) \leq 2a_n$ for any p in $[1,\infty]$. In particular, it follows from (2.8) that

$$\mathbb{P}\big(\sqrt{n}K(F_n, F) \geq x\big) \leq \mathbb{P}\big(\sqrt{n}D_{2,n}(\lambda_1) \geq x\big) \leq 2\exp\left(-\frac{x^2}{2(1 + 4(a_1 + \cdots + a_{n-1}))^2}\right). \tag{5.3}$$

In a recent paper, Collet, Martinez and Schmitt (2002) [4] studied a class of expanding maps for which (5.1) (and hence (5.2)) holds with $a_n = C\rho^n$ for some $\rho < 1$. Using a concentration inequality for Lipschitz functions they prove in Theorem III.1 that there exist a number $x_0 > 0$ and a constant $R > 0$ (both depending on T) such that, for any $x > x_0$ and any integer n,

$$\mathbb{P}\big(\sqrt{n}K(F_n, F) \geq x\big) \leq \exp(-Rx^2). \tag{5.4}$$

Clearly, (5.3) is more precise than (5.4), for it holds for any positive x. Moreover, we do not require that a_n decreases geometrically, and we are able to give an expression for R in terms of the coefficients $(a_i)_{1\leq i\leq n-1}$.

5.2. Causal linear processes

The main interest of the coefficients $\tau_{\mu,p,q}$ is that they are very easy to evaluate in many situations. Let us focus on the stationary sequence

$$X_k = \sum_{j\geq 0} a_j \xi_{k-j} \tag{5.5}$$

where $(\xi_i)_{i\in\mathbb{Z}}$ is a sequence of iid random variables and $\sum_{j\geq 0} |a_j| < \infty$. From (4.6) and the application of Section 4.2, we know that if μ is a probability measure and if X_0 has a density bounded by K, then for any p, q in $[1,\infty]$ and any r such that $rq \in [1,\infty]$,

$$\tau_{\mu,p,q}(\mathcal{M}_0, X_n) \leq C(r)\Big(K\|\xi_0 - \xi_1\|_{rq} \sum_{j\geq n} |a_j|\Big)^{r/(r+1)} \quad \text{and} \quad \tau_{\mu,p,\infty}(\mathcal{M}_0, X_n) \leq K\|\xi_0 - \xi_1\|_\infty \sum_{j\geq n} |a_j| \tag{5.6}$$

where $\mathcal{M}_0 = \sigma(\xi_k, k \leq 0)$. For instance, if we use the last bound in (5.6), we obtain from (2.3) that

$$\mathbb{P}(\sqrt{n}D_{p,n}(\mu) \geq x) \leq 2\exp\left(-\frac{x^2}{2(p-1)\big(1 + 4K\|\xi_0\|_\infty \sum_{j=1}^\infty (n \wedge j)|a_j|\big)^2}\right). \tag{5.7}$$

Applying this result in the particular case $\mu = \delta_t$, we infer that

$$\sup_{t\in\mathbb{R}} \mathbb{P}(\sqrt{n}|F_n(t) - F(t)| \ge x) \le 2\exp\left(-\frac{x^2}{2\big(1 + 4K\|\xi_0\|_\infty \sum_{j=1}^\infty (n\wedge j)|a_j|\big)^2}\right), \tag{5.8}$$

which is similar (up to numerical constants) to the upper bound obtained in Corollary 2 in Rio (2000) [16]. In that case, the bound (5.8) can be improved by using the inequality (2.9).

From Condition 1 of Theorem 2, we infer that $\sqrt{n}(F_n - F)$ converges in distribution in any $\mathbb{L}^p(\mu)$, where μ is a probability measure and $2 \le p < \infty$, as soon as $\|\xi_0\|_{2r} < \infty$ and $\sum_{k>0}(\sum_{j\ge k}|a_j|)^{r/(r+1)} < \infty$ for some r in $[1/2, \infty]$. From Condition (3.2), we infer that $\sqrt{n}(F_n - F)$ converges in distribution in $\mathbb{L}^2(\mu)$ as soon as $\|\xi_0\|_{2r} < \infty$ and $\sum_{k>0} k^{-1/2}(\sum_{j\ge k}|a_j|)^{r/(r+1)} < \infty$. In particular, the latter condition is realized for bounded innovations provided $|a_i| = O(i^{-3/2-\epsilon})$ for some positive ϵ.

Note that all the results mentioned above are valid without assuming that the innovations have a density : we have only assumed that X_0 has a density, which is much weaker (think of the well know example where $a_i = 2^{-i-1}$ and $\xi_0 \sim \mathcal{B}(1/2)$). Now, if we assume that ξ_0 has a density bounded by C, we can also evaluate the quantities $\big\|\|F_{X_k|\mathcal{M}_i} - F_{X_k|\mathcal{M}_{i-1}}\|_{p,\mu}\big\|_q$ which appear in the inequality (2.1) and in Condition 3 of Theorem 2.

Lemma 6. *Let $(X_k)_{k>0}$ be defined by (5.5) and let $\mathcal{M}_i = \sigma(\xi_k, k \le i)$. If ξ_0 has a density bounded by C, then, for any probability measure μ and any p, q in $[1,\infty]$,*

$$\big\|\|F_{X_k|\mathcal{M}_0} - F_{X_k|\mathcal{M}_{-1}}\|_{p,\mu}\big\|_q \le C\|\xi_1 - \xi_0\|_q \frac{|a_k|}{|a_0|}.$$

Applying Lemma 6, and arguing as for the proof of Inequality (2.1), we infer that,

$$\mathbb{P}(\sqrt{n}D_{p,n}(\mu) \ge x) \le 2\exp\left(-\frac{x^2}{2(p-1)\big(1 + 2|a_0|^{-1}C\|\xi_0\|_\infty \sum_{k=1}^\infty |a_k|\big)^2}\right) \tag{5.9}$$

which is more precise than (5.7), since it gives $\mathbb{P}(\sqrt{n}D_{p,n}(\mu) \ge x) \le 2\exp(-Rx^2)$ for some positive constant R as soon as $\sum_{i>0}|a_i| < \infty$. In the same way, we infer from Condition 3 of Theorem 2 that $\sqrt{n}(F_n - F)$ converges in distribution in $\mathbb{L}^2(\mu)$ as soon as $\|\xi_0\|_2 < \infty$ and $\sum_{i>0}|a_i| < \infty$.

Proof of Lemma 6. Let $Y_k = \sum_{i=0}^{k-1} a_i\xi_{k-i}$. Clearly

$$F_{X_k|\mathcal{M}_0}(t) = F_{Y_k}(t - (X_k - Y_k)) \quad \text{and} \quad F_{X_k|\mathcal{M}_{-1}}(t) = \int F_{Y_k}(t - a_k x - (X_k - Y_{k+1}))P_{\xi_0}(\mathrm{d}x).$$

Let f_ξ be the density of ξ_0. The density of Y_k is given by $f_{Y_k} = |a_0|^{-1}f_\xi(\cdot/a_0) * \cdots * |a_{k-1}|^{-1}f_\xi(\cdot/a_{k-1})$, and hence it is bounded by $C|a_0|^{-1}$. Consequently

$$|F_{X_k|\mathcal{M}_0}(t) - F_{X_k|\mathcal{M}_{-1}}(t)| \le C|a_0|^{-1}|a_k| \int |x - \xi_0| P_{\xi_0}(\mathrm{d}x).$$

The result follows by taking the $\|.\|_q$-norm and applying Jensen's inequality. □

6. Appendix: proof of Lemma 1.

Let $f \in W_{p',1}(\mu)$. We first check that under (1.1), $|f|$ is integrable with respect to $\mathrm{d}F$. Without loss of generality, assume that $f(0) = 0$. Clearly

$$\int |f(t)|\mathrm{d}F(t) \le \int_{\mathbb{R}^+}\Big(\int_{[0,t[} |f'(x)|\mu(\mathrm{d}x)\Big)\mathrm{d}F(t) + \int_{\mathbb{R}^-}\Big(\int_{[t,0[} |f'(x)|\mu(\mathrm{d}x)\Big)\mathrm{d}F(t).$$

Applying Fubini, we obtain that

$$\int |f(t)|\mathrm{d}F(t) \leq \int_{\mathbb{R}^+} |f'(x)|(1-F(x))\mu(\mathrm{d}x) + \int_{\mathbb{R}^-} |f'(x)|F(x)\mu(\mathrm{d}x).$$

Since f' belongs to $\mathbb{L}^{p'}(\mu)$, the right hand side is finite as soon as (1.1) holds. In the same way, we have both

$$\begin{aligned}\int f(t)\mathrm{d}F(t) &= \int_{\mathbb{R}^+} f'(x)(1-F(x))\mu(\mathrm{d}x) - \int_{\mathbb{R}^-} f'(x)F(x)\mu(\mathrm{d}x)\\ \int f(t)\mathrm{d}G(t) &= \int_{\mathbb{R}^+} f'(x)(1-G(x))\mu(\mathrm{d}x) - \int_{\mathbb{R}^-} f'(x)G(x)\mu(\mathrm{d}x).\end{aligned}$$

Consequently

$$\int f(t)\mathrm{d}F(t) - \int f(t)\mathrm{d}G(t) = \int f'(x)(G(x)-F(x))\mu(\mathrm{d}x).$$

The second equality follows by noting that

$$\|F-G\|_{p,\mu} = \sup_{\|g\|_{p',\mu}\leq 1} \Big| \int_{\mathbb{R}} g(x)(F(x)-G(x))\mathrm{d}\mu(x)\Big|.$$

Acknowledgements. The authors are indebted to the referees for carefully reading the manuscript and for helpful comments.

References

[1] K. Azuma, Weighted sums of certain dependent random variables. *Tôkohu Math. J.* **19** (1967) 357–367.
[2] H.C.P. Berbee, *Random walks with stationary increments and renewal theory.* Mathematical Centre Tracts **112**, Mathematisch Centrum, Amsterdam (1979).
[3] L. Birgé and P. Massart, An adaptive compression algorithm in Besov Spaces. *Constr. Approx.* **16** (2000) 1–36.
[4] P. Collet, S. Martinez and B. Schmitt, Exponential inequalities for dynamical measures of expanding maps of the interval. *Probab. Theory Relat. Fields* **123** (2002) 301–322.
[5] J. Dedecker and F. Merlevède, The conditional central limit theorem in Hilbert spaces. *Stoch. Processes Appl.* **108** (2003) 229–262.
[6] J. Dedecker and C. Prieur, Coupling for τ-dependent sequences and applications. *J. Theoret. Probab.* **17** (2004) 861–885.
[7] J. Dedecker and C. Prieur, New dependence coefficients. Examples and applications to statistics. *Probab. Theory Relat. Fields* **132** (2005) 203–236.
[8] J. Dedecker and E. Rio, On the functional central limit theorem for stationary processes. *Ann. Inst. H. Poincaré Probab. Statist.* **36** (2000) 1–34.
[9] P. Doukhan, P. Massart and E. Rio, Invariance principle for absolutely regular empirical processes. *Ann. Inst. H. Poincaré Probab. Statist.* **31** (1995) 393–427.
[10] M.I. Gordin, The central limit theorem for stationary processes. *Dokl. Akad. Nauk SSSR* **188** (1969) 739–741.
[11] P. Massart, The tight constant in the Dvoretzky-Kiefer-Wolfowitz inequality. *Ann. Probab.* **18** (1990) 1269–1283.
[12] F. Merlevède and M. Peligrad, On the coupling of dependent random variables and applications, in *Empirical process techniques for dependent data*, Birkhäuser (2002) 171–193.
[13] P. Oliveira and C. Suquet, $\mathbb{L}^2([0,1])$ weak convergence of the empirical process for dependent variables, in *Wavelets and statistics* (Villard de Lans 1994), *Lect. Notes Statist.* **103** (1995) 331–344.
[14] P. Oliveira and C. Suquet, Weak convergence in $\mathbb{L}^p([0,1])$ of the uniform empirical process under dependence. *Statist. Probab. Lett.* **39** (1998) 363–370.
[15] I.F. Pinelis, An approach to inequalities for the distributions of infinite-dimensional martingales, in *Probability in Banach spaces*, *Proc. Eight Internat. Conf.* **8** (1992) 128–134.
[16] E. Rio, Inégalités de Hoeffding pour les fonctions lipschitziennes de suites dépendantes. *C. R. Acad. Sci. Paris Série I* **330** (2000) 905–908.
[17] A.W. van der Vaart, Bracketing smooth functions. *Stoch. Processes Appl.* **52** (1994) 93–105.
[18] W.A. Woyczyński, A central limit theorem for martingales in Banach spaces. *Bull. Acad. Polon. Sci. Sr. Sci. Math. Astronom. Phys.* **23** (1975) 917–920.
[19] V.V. Yurinskii, Exponential bounds for large deviations. *Theory Prob. Appl.* **19** (1974) 154–155.