SFB
823

Empirical processes of Markov chains and dynamical systems indexed by classes of functions

Discussion Paper

Herold Dehling, Olivier Durieu, Marco Tusche

Nr. 3/2012

SFB
823

EMPIRICAL PROCESSES OF MARKOV CHAINS AND DYNAMICAL SYSTEMS INDEXED BY CLASSES OF FUNCTIONS

HEROLD DEHLING, OLIVIER DURIEU, AND MARCO TUSCHE

Abstract. We study weak convergence of empirical processes of dependent data, indexed by classes of functions. We obtain results that are especially suitable for data arising from dynamical systems and Markov chains, where the Central Limit Theorem for partial sums is commonly derived via the spectral gap technique. Our results apply, e.g. to the empirical process of ergodic torus automorphisms.

Contents

1. Introduction and Main Results

Empirical process central limit theorems for dependent data have been studied by many authors. One of the earliest results is Billingsley (1968), where functions of mixing processes are being studied. Philipp (1982) studied the multivariate empirical process in the case of mixing process. Doukhan, Massart and Rio (1995) study empirical processes for absolutely regular data. Andrews and Pollard (1994) study strongly mixing processes. Borovkova, Burton and Dehling (2001) investigate the emprical process and the empirical U-process for functionals of absolutely regular processes. For many further results, see the survey article of Dehling and Philipp (2002), the book by Dedecker, Doukhan, Lang, León, Louhichi and Prieur (2007) as well as the paper by Dedecker and Prieur (2007).

Recently, Dehling, Durieu and Volny (2009) have investigated empirical processes of Markov chains and dynamical systems for which the partial sum central limit theorem is

Date: January 9, 2012.

Key words and phrases. Empirical processes indexed by classes of functions, dependent data, Markov chains, dynamical systems, ergodic torus automorphism, weak convergence.

Research partially supported by German Research Foundation grant DE 370-4 *New Techniques for Empirical Processes of Dependent Data*, and the Collaborative Research Grant SFB 823 *Statistical Modelling of Nonlinear Dynamic Processes.*

usually established by the spectral gap method. In this case, one faces the challenge that the CLT as well as moment bounds for partial sums are not available for the class of functions that one wants to consider in the empirical process CLT, but only for a different class of functions such as Lipschitz or Hölder functions. Dehling et al. (2009) developed techniques to handle this problem. Dehling and Durieu (2011) have extended these techniques to the case of multidimensional random variables that satisfy a multiple mixing condition. As an example, they could prove the empirical process central limit theorem for ergodic torus automorphisms and random iterative Lipschitz models. Generalizations of these results concerning the dependence structure of the underlying processes can be found in the paper of Durieu and Tusche (2011). It is the goal of the present paper to extend the techniques developed by Dehling et al (2009) to empirical processes indexed by classes of functions. For the theory of empirical processes of i.i.d. data, indexed by classes of functions, see the book by van der Vaart and Wellner (1996).

Let $(X_i)_{i\geq 1}$ be an $\mathcal{X}$-valued stationary random process, where $(\mathcal{X},\mathcal{A})$ is a measurable space. Let μ be the common distribution of the X_i. Let $\mathcal{F}$ be a uniformly bounded class of measurable functions from $\mathcal{X}$ to $\mathbb{R}$. If Q is a signed measure on $(\mathcal{X},\mathcal{A})$, we use the notation $Qf=\int f dQ$. We consider the map F_n from $\mathcal{F}$ to $\mathbb{R}$ induced by the empirical measure of order n given by

$$F_n(f)=\frac{1}{n}\sum_{i=1}^{n} f(X_i), \quad f\in\mathcal{F}.$$

The $\mathcal{F}$-indexed empirical process of order n is given by

$$U_n(f)=\sqrt{n}(F_n(f)-\mu f)=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(f(X_i)-\mu f), \quad f\in\mathcal{F}.$$

We regard the empirical process $(U_n(f))_{f\in\mathcal{F}}$ as random element of $\ell^\infty(\mathcal{F})$; this holds as $\mathcal{F}$ is supposed to be uniformly bounded. $\ell^\infty(\mathcal{F})$ is equipped with the supremum norm and the Borel σ-field which is generated by the open sets. It is well known that, in general, $(U_n(f))_{f\in\mathcal{F}}$ is not measurable and thus the usual theory of weak convergence of random variables does not apply. We use here the theory which is based on convergence of outer expectations; see van der Vaart and Wellner (1996). Given a Borel probability measure L on $\ell^\infty(\mathcal{F})$, we say that $(U_n)_{n\geq 1}$ converges in distribution to L if

$$\mathbf{E}^*(\varphi(U_n))\to\int\varphi(x)dL(x),$$

for all $\varphi:\ell^\infty(\mathcal{F})\to\mathbb{R}$ that are bounded and continuous. Here $\mathbf{E}^*$ denotes the outer integral. Note that $E^*(X)=E(X^*)$, where X^* denotes the measurable cover function of X; see Lemma 1.2.1 in van der Vaart and Wellner (1996).

In what follows, we will frequently require two assumptions concerning the processes $(f(X_i))_{i\geq 1}$, where $f:\mathcal{X}\to\mathbb{R}$ belongs to some Banach space $(\mathcal{B},\|\cdot\|_{\mathcal{B}})$ of measurable functions from $\mathcal{X}$ to $\mathbb{R}$. The precise choice of this Banach space will depend on the specific example. Common examples are spaces of Lipschitz- or Hölder-continuous functions.

We assume that for some subset $\mathcal{G}\subset\mathcal{B}$ the following two properties hold.

(1) Central Limit Theorem: For all $f\in\mathcal{G}$,

$$\frac{1}{n}\sum_{i=1}^{n}(f(X_i)-\mu f)\xrightarrow{\mathcal{D}}\mathcal{N}(0,\sigma^2), \tag{1.1}$$

where $N(0,\sigma^2)$ denotes a normal law with mean zero and variance

$$\sigma^2 = \mathbf{E}(f(X_0)-\mu f)^2 + 2\sum_{i=1}^{\infty} \mathrm{Cov}(f(X_0), f(X_i));$$

(2) Bound on the $2p$-th moments: For any $p \geq 1$, there exist a constant $C > 0$, such that for all $f \in \mathcal{G}$

$$\mathbf{E}\left[\left(\sum_{i=1}^{n}(f(X_i)-\mu f)\right)^{2p}\right] \leq C\sum_{i=1}^{p} n^i \|f\|_1^i \log^{2p-i}(\|f\|_{\mathcal{B}}+1), \tag{1.2}$$

where $\|f\|_1 := \int_{\mathcal{X}} |f| \, d\mu$ is the $L^1(\mu)$-norm of f.

In most applications, we will take $\mathcal{G}$ to be the intersection of $\mathcal{B}$ with the ball of a fixed radius $M > 0$ in $\ell^\infty(\mathcal{X})$, i.e.

$$\mathcal{G} = \{f \in \mathcal{B} : \|f\|_\infty \leq M\}.$$

The Central Limit Theorem (1.1) has been established for many stationary processes, see e.g. Bradley (2007) for mixing processes, Dedecker et al. (2007) for weakly dependent processes and Hennion and Hervé for many examples of Markov chains and dynamical systems. Moment bounds have also been obtained for many stochastic processes, see e.g. Bradley (2007), Dedecker et al (2007). Durieu (2009) proved moment bounds for $\mathcal{B}$-geometrically ergodic Markov chains. Dehling and Durieu (2011) give moment bounds for dynamical systems that satisfy a multiple mixing condition.

Empirical process central limit theorems require bounds on the size of the class of functions $\mathcal{F}$, usually measured by the number of ε-balls required to cover $\mathcal{F}$. Here we will introduce a covering number adapted to the fact that (1.1) and (1.2) hold only for $f \in \mathcal{G}$, and that both the $\mathcal{B}$-norm as well as the $L^1(\mu)$-norm occur on the right-hand side of the bound (1.2).

Definition. (i) Let $l, u : \mathcal{X} \to \mathbb{R}$ be two functions satisfying $l(x) \leq u(x)$ for all $x \in \mathcal{X}$. We define the bracket

$$[l,u] := \{f : \mathcal{X} \to \mathbb{R} : l(x) \leq f(x) \leq u(x), \text{ for all } x \in \mathcal{X}\}.$$

Given $\varepsilon, A > 0$, $\mathcal{G} \subset \mathcal{B}$ and a probability law μ, we call $[l,u]$ an $(\varepsilon, A, \mathcal{G}, L^1(\mu))$-bracket if $l, u \in \mathcal{G}$ and

$$\|u-l\|_1 \leq \varepsilon$$
$$\|u\|_{\mathcal{B}} \leq A, \|l\|_{\mathcal{B}} \leq A,$$

where $\|\cdot\|_1$ denotes the $L^1(\mu)$-norm.

(ii) For a class of measurable functions $\mathcal{F}$ and a subset $\mathcal{G} \subset \mathcal{B}$, we define the bracketing number $N(\varepsilon, A, \mathcal{F}, \mathcal{G}, L^1(\mu))$ as the smallest number of $(\varepsilon, A, \mathcal{G}, L^1(\mu))$-brackets needed to cover $\mathcal{F}$.

We can now state the main theorem of the present paper. The proof will be given in Section 3

Theorem 1.1. *Assume (1.1) and (1.2) and that there exist constants $r \geq 1$, $\gamma > 1$ and $C > 0$ such that*

$$\int_0^1 \varepsilon^{r-1} \sup_{\delta \geq \varepsilon} N\left(\delta, \exp\left(C\delta^{-\frac{1}{\gamma}}\right), \mathcal{F}, \mathcal{G}, L^1(\mu)\right) d\varepsilon < \infty \tag{1.3}$$

holds. Then the empirical process $(U_n(f))_{f\in\mathcal{F}}$ converges in distribution in $\ell^\infty(\mathcal{F})$ to a tight gaussian process $(W(f))_{f\in\mathcal{F}}$.

Remark 1.2. The proof of Theorem 1.1 shows (cf. page 9) that the statement also holds, if condition (1.2) is only satisfied for some $p \geq 1$ such that

$$p > \frac{2r\gamma}{\gamma - 1}.$$

Remark 1.3. A sufficient condition for (1.3) is

$$N(\varepsilon, \exp(C\varepsilon^{-\frac{1}{\gamma}}), \mathcal{F}, \mathcal{G}, L^1(\mu)) = O(\varepsilon^{-s}), \quad \text{as } \varepsilon \to 0$$

for some $s < r$.

Examples of classes of functions satisfying condition (1.3) such as indicators of multi-dimensional rectangles and ellipsoids or centered balls of arbitrary norm will be given in Section 4.

We can apply Theorem 1.1 to the empirical process of ergodic torus automorphisms. Let $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ be the torus of dimension $d > 1$. If M is a square matrix of dimension d with integer coefficients and determiant ± 1, then the transformation $T : \mathbb{T}^d \longrightarrow \mathbb{T}^d$ defined by

$$Tx = Mx \mod 1$$

is an automorphism of $\mathbb{T}^d$ and preserves the Lebesgue measure λ. Thus $(\mathbb{T}^d, \mathcal{B}(\mathbb{T}^d), \lambda, T)$ is a dynamical system. It is ergodic if and only if the matrix M has no eigenvalue which is a root of the unity. A result of Kronecker shows that in this case, M always has at least one eignevalue which has modulus different than 1. The hyperbolic automorphisms (i.e. no eigenvalue of modulus 1) are particular cases of Anosov diffeomorphisms. Their property are better understood than in the general case. However, the general case of ergodic automorphism is an example of a partially hyperbolic system for which strong result can be proved. The central limit theorem for regular observables has been proved by Leonov (1960), see also Le Borgne (1999) for refinements. Other limit theorems can be found in Dolgopyat (2004). The one-dimensional empirical process, for $\mathbb{R}$-valued regular observables, has been studied by Durieu and Jouan (2008). Dehling and Durieu (2011) proved weak convergence of the classical empirical process (indexed by indicators of left infinite rectangles). We can now generalize this result to empirical process indexed by different classes of functions. In particular, we can deal with the class of indicators of balls and we can get the following proposition, as a corollary of the preceding theorem.

Theorem 1.4. *Let T be an ergodic d-torus automorphism and $\mathcal{F}$ be the class of indicator functions of balls of $\mathbb{T}^d$. Then the empirical process*

$$U_n(f) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(f \circ T^i - \lambda f \right), \quad f \in \mathcal{F}$$

converges in distribution in $\ell^\infty(\mathcal{F})$ to a tight Gaussian process $(W(f))_{f \in \mathcal{F}}$.

The proof of the theorem will be given in section 4.3.

2. Preliminary Result

In the proof of Theorem 1.1, we need a generalization of Theorem 4.2 of Billingsley (1968). Billingsley considers random variables X_n, $X_n^{(m)}$, $X^{(m)}$, X, $m, n \geq 1$, with values in the separable metric space (S, ρ) satisfying (a) $X_n^{(m)} \xrightarrow{\mathcal{D}} X^{(m)}$ as $n \to \infty$, for all $m \geq 1$, (b) $X^{(m)} \xrightarrow{\mathcal{D}} X$ as $m \to \infty$ and (c) $\limsup_{n\to\infty} P(\rho(X_n^{(m)}, X_n) \geq \delta) \to 0$ as $n \to \infty$. Theorem 4.2 of Billingsley (1968) states that then $X_n \xrightarrow{\mathcal{D}} X$. Dehling, Durieu and Volny (2009) proved that this result holds without condition (b), provided that S is a complete separable metric space. More precisely, they could show that in this situation (a) and (c) together imply the existence of a random variable X satisfying (b), and thus by Billingsley's theorem $X_n \xrightarrow{\mathcal{D}} X$. Here we will generalize this theorem to possibly non-measurable random elements with values in non-separable spaces. Regarding convergence in distribution of non-measurable random elements, we use the notation of van der Vaart and Wellner (1996). In accordance with the terms van der Vaart and Wellner use, we will call a not necessarily measurable function with values in a measurable space a random element, while a random variable shall still be as a measurable random element.

Theorem 2.1. *Let $X_n, X_n^{(m)}, X^{(m)}$, $m, n \geq 1$ be random elements with values in the complete metric space (S, ρ), and suppose that $X^{(m)}$ is measurable and separable, i.e. there is a separable set $S^{(m)} \subset S$ such that $\mathbf{P}(X^{(m)} \in S^{(m)}) = 1$. If the conditions*

$$X_n^{(m)} \xrightarrow{\mathcal{D}} X^{(m)}, \text{ as } n \to \infty, \text{ for all } m \geq 1, \tag{2.1}$$

$$\limsup_{n\to\infty} \mathbf{P}^*\big(\rho(X_n, X_n^{(m)}) \geq \delta\big) \longrightarrow 0, \text{ as } m \to \infty, \text{ for all } \delta > 0 \tag{2.2}$$

are satisfied, then there exists an S-valued, separable random variable X (i.e. X is measurable) such that $X^{(m)} \xrightarrow{\mathcal{D}} X$, as $m \to \infty$, and

$$X_n \xrightarrow{\mathcal{D}} X, \text{ as } n \to \infty. \tag{2.3}$$

Proof. (i) We will first show that $X^{(m)}$ converges in distribution to some random variable X. We denote by $L^{(m)}$ the distribution of $X^{(m)}$; this is defined since $X^{(m)}$ is measurable. Moreover, $L^{(m)}$ is a separable Borel probability measure on S. By Theorem 1.12.4. of Van der Vaart and Wellner (1996), weak convergence of separable Borel measures on a metric space S can be metrized by the bounded Lipschitz metric, defined by

$$d_{BL_1}(L_1, L_2) = \sup_{f \in BL_1} \left| \int f(x) dL_1(x) - \int f(x) dL_2(x) \right|, \tag{2.4}$$

for any Borel measures L_1, L_2 on S. Here $BL_1 := \{f : S \longrightarrow \mathbb{R} : \|f\|_{BL_1} \leq 1\}$, where

$$\|f\|_{BL_1} := \max\left\{ \sup_{x\in S} |f(x)|, \sup_{x \neq y \in S} \frac{f(x) - f(y)}{\rho(x, y)} \right\}.$$

In addition, the theorem states that the space of all separable Borel measures on a complete space is complete with respect to the bounded Lipschitz metric. Thus it suffices to show that $L^{(m)}$ is a d_{BL_1}-Cauchy sequence. We obtain

$$\begin{aligned} d_{BL_1}(L^{(m)}, L^{(l)}) &= \sup_{f\in BL_1} |\mathbf{E} f(X^{(m)}) - \mathbf{E} f(X^{(l)})| \\ &\leq \sup_{f \in BL_1} \Big\{ |\mathbf{E} f(X^{(m)}) - \mathbf{E}^* f(X_n^{(m)})| + |\mathbf{E}^* f(X_n^{(m)}) - \mathbf{E}^* f(X_n)| \\ &\quad + |\mathbf{E}^* f(X_n) - \mathbf{E}^* f(X_n^{(l)})| + |\mathbf{E}^* f(X_n^{(l)}) - \mathbf{E} f(X^{(l)})| \Big\} \end{aligned}$$

for all $n \in \mathbb{N}$. For a Borel measurable, separable random elements $X^{(m)}$ weak convergence $X_n^{(m)} \xrightarrow{\mathcal{D}} X^{(m)}$ as $n \to \infty$ is equivalent to $\sup_{f\in BL_1} |\mathbf{E} f(X^{(m)}) - \mathbf{E}^* f(X_n^{(m)})| \longrightarrow 0$; see van der Vaart and Wellner (1996, p.73). Hence by (2.1) we obtain

$$d_{BL_1}(X^{(m)}, X^{(l)}) \leq \liminf_{n\to\infty} \sup_{f\in BL_1} |\mathbf{E}^* f(X_n^{(m)}) - \mathbf{E}^* f(X_n)| + |\mathbf{E}^* f(X_n) - \mathbf{E}^* f(X_n^{(l)})|$$

Using Lemma 1.2.2 in van der Vaart and Wellner (1996), we obtain

$$|\mathbf{E}^* f(X_n^{(m)}) - \mathbf{E}^* f(X_n)| \leq \mathbf{E}(|f(X_n) - f(X_n^{(m)})|^*)$$

and therefore

$$\sup_{f\in BL_1} |\mathbf{E}^* f(X_n^{(m)}) - \mathbf{E}^* f(X_n)| \le \mathbf{E}\big(\rho(X_n, X_n^{(m)})^* \wedge 2\big) = \int_0^\infty \mathbf{P}^*\Big(\big(\rho(X_n, X_n^{(m)}) \wedge 2\big) \ge t\Big)\, dt. \tag{2.5}$$

Now, let $\varepsilon > 0$ be given. By (2.2), there exists $m_0 \in \mathbb{N}$ such that for every $m \ge m_0$ there is some $n_0 \in \mathbb{N}$ such that for every $n \ge n_0$ we have $\mathbf{P}^*\big(\rho(X_n, X_n^{(m)}) \ge \varepsilon/3\big) \le \varepsilon/3$. Therefore

$$\mathbf{P}^*\Big(\rho(X_n, X_n^{(m)}) \wedge 2 \ge t\Big) \le \begin{cases} 1, & \text{if } t < \frac{\varepsilon}{3} \\ \frac{\varepsilon}{3}, & \text{if } \frac{\varepsilon}{3} \le t \le 2 \\ 0, & \text{if } 2 < t. \end{cases}$$

Applying this inequality to (2.5), we obtain

$$\liminf_{n\to\infty} \sup_{f\in BL_1} |\mathbf{E}^* f(X_n^{(m)}) - \mathbf{E}^* f(X_n)| \le \int_0^2 \frac{\varepsilon}{3} + 1_{\{t<\frac{\varepsilon}{3}\}}\, dt = \varepsilon$$

for all $m \ge m_0$. Hence for $l, m \ge m_0$ we have $d_{BL_1}(L^{(m)}, L^{(l)}) \le 2\varepsilon$; i.e. $(L^{(m)})_{m\in\mathbb{N}}$ is a d_{BL_1}-Cauchy sequence in a complete metric space.

(ii) The remaining part of the proof follows closely the proof of Theorem 4.2 in Billingsley (1968), replacing the probability measure $\mathbf{P}$ by the outer measure $\mathbf{P}^*$ where necessary and making use of the Portmanteau theorem; see van der Vaart and Wellner (1996), Theorem 1.3.4, and the subadditivity of outer measures. From part (i), we already know that there is some measurable X such that $X^{(m)} \xrightarrow{\mathcal{D}} X$. Let $F \subset S$ be closed. Given $\varepsilon > 0$, we define the ε-neighborhood $F_\varepsilon := \{s \in S\ : \inf_{x\in F} \rho(s, x) \le \varepsilon\}$, and observe that F_ε is also closed. Since $\{X_n \in F\} \subset \{X_n^{(m)} \in F_\varepsilon\} \cup \{\rho(X_n^{(m)}, X_n) \ge \varepsilon\}$, we obtain

$$\mathbf{P}^*(X_n \in F) \le \mathbf{P}^*(X_n^{(m)} \in F_\varepsilon) + \mathbf{P}^*(\rho(X_n^{(m)}, X_n) \ge \varepsilon),$$

for all $m \in \mathbb{N}$. By (2.2) we may choose m_0 so large that for all $m \ge m_0$

$$\limsup_{n\to\infty} \mathbf{P}^*(\rho(X_n^{(m)}, X_n) \ge \varepsilon) \le \varepsilon/2.$$

As $X^{(m)} \xrightarrow{\mathcal{D}} X$, by the Portmanteau theorem we may choose m_1 so large that for all $m \ge m_1$

$$\mathbf{P}(X^{(m)} \in F_\varepsilon) \le \mathbf{P}(X \in F_\varepsilon) + \varepsilon/2.$$

We now fix $m \ge \max(m_0, m_1)$. By (2.1) we have $X_n^{(m)} \xrightarrow{\mathcal{D}} X^{(m)}$ as $n \to \infty$. Thus an application of the Portmanteau theorem yields

$$\limsup_{n\to\infty} \mathbf{P}^*(X_n^{(m)} \in F_\varepsilon) \le \mathbf{P}(X^{(m)} \in F_\varepsilon),$$

$$\limsup_{n\to\infty} \mathbf{P}^*(X_n \in F) \le \mathbf{P}(X \in F_\varepsilon) + \varepsilon.$$

Since this holds for any $\varepsilon > 0$ and $\lim_{\varepsilon\to 0} \mathbf{P}(X \in F_\varepsilon) = \mathbf{P}(X \in F)$, we get

$$\limsup_{n\to\infty} \mathbf{P}^*(X_n \in F) \le \mathbf{P}(X \in F),$$

for all closed sets $F \subset S$. By a final application of the Portmanteau theorem we infer $X_n \xrightarrow{\mathcal{D}} X$. □

3. Proof of Theorem 1.1

For all $q \geq 1$, there exist two sets of $N_q := N(2^{-q}, \exp(C2^{\frac{q}{\gamma}}), \mathcal{F}, \mathcal{G}, L^1(\mu))$ functions $\{g_{q,1}, \ldots, g_{q,N_q}\} \subset \mathcal{G}$ and $\{g'_{q,1}, \ldots, g'_{q,N_q}\} \subset \mathcal{G}$, such that $\|g_{q,i} - g'_{q,i}\|_1 \leq 2^{-q}$, $\|g_{q,i}\| \leq \exp(C2^{\frac{q}{\gamma}})$, $\|g'_{q,i}\| \leq \exp(C2^{\frac{q}{\gamma}})$ and for all $f \in \mathcal{F}$, there exists i such that $g_{q,i} \leq f \leq g'_{q,i}$. Further, by (1.3),

$$\sum_{q\geq 1} 2^{-rq} N_q < \infty. \tag{3.1}$$

For all $q \geq 1$, we can build a partition $\mathcal{F} = \bigcup_{i=1}^{N_q} \mathcal{F}_{q,i}$ of the class $\mathcal{F}$ into N_q subsets such that for all $f \in \mathcal{F}_{q,i}$, $g_{q,i} \leq f \leq g'_{q,i}$. To see this define $\mathcal{F}_{q,1} = [g_{q,1}, g'_{q,1}]$ and $\mathcal{F}_{q,i} = [g_{q,i}, g'_{q,i}] \backslash (\cup_{j=1}^{i-1} \mathcal{F}_j)$.

In the sequel, we will use the notation $\pi_q f = g_{q,i}$ and $\pi'_q f = g'_{q,i}$ if $f \in \mathcal{F}_{q,i}$.

For each $q \geq 1$, we introduce the process

$$F_n^{(q)}(f) := F_n(\pi_q f) = \frac{1}{n} \sum_{i=1}^{n} \pi_q f(X_i); \quad f \in \mathcal{F}$$

which is constant on each $\mathcal{F}_{q,i}$. Further, if $f \in \mathcal{F}_{q,i}$, we have

$$F_n^{(q)}(f) \leq F_n(f) \leq F_n(\pi'_q f)$$

We introduce

$$U_n^{(q)}(f) := U_n(\pi_q f) = \sqrt{n}(F_n^{(q)}(f) - \mu(\pi_q f)); \quad f \in \mathcal{F}.$$

Proposition 3.1. *For all $q \geq 1$, the sequence $(U_n^{(q)}(f))_{f\in\mathcal{F}}$ converges in distribution in $\ell^\infty(\mathcal{F})$ to a piecewise constant Gaussian process $(U^{(q)}(f))_{f\in\mathcal{F}}$ as $n \to \infty$.*

Proof. Since $\pi_q f \in \mathcal{G}$ and $\mathcal{G}$ is a subset of $\mathcal{B}$, by assumption (1.1), the CLT holds and $U_n^{(q)}(f)$ converges to a Gaussian law for all $f \in \mathcal{F}$. We can apply the Cramér-Wold device to get the finite dimensional convergence : for all $k \geq 1$, for all $f_1, \ldots, f_k \in \mathcal{F}$, $(U_n^{(q)}(f_1), \ldots, U_n^{(q)}(f_k))$ converges in distribution to a Gaussian vector in $\mathbb{R}^k$. Since $U_n^{(q)}$ is constant on each element $\mathcal{F}_{q,i}$ of the partition, the finite dimensional convergence implies the weak convergence of the process. □

Proposition 3.2. *For all $\varepsilon > 0$, $\eta > 0$ there exists q_0 such that for all $q \geq q_0$*

$$\limsup_{n\to\infty} \mathbf{P}^*(\sup_{f\in\mathcal{F}} |U_n(f) - U_n^{(q)}(f)| > \varepsilon) \leq \eta.$$

Proof. For a random variable Y let $\overline{Y}$ denote its centering $\overline{Y} := Y - \mathbf{E}Y$.

If for arbitrary random variables Y_l, Y, Y_u we have $Y_l \leq Y \leq Y_u$ then

$$|\overline{Y} - \overline{Y_l}| \leq |\overline{Y_u} - \overline{Y_l}| + \mathbf{E}|\overline{Y_u} - \overline{Y_l}|.$$

Using $F_n^{(q+K)}(f) \leq F_n(f) \leq F_n(\pi'_{q+K})$ and $\mathbf{E}|F_n(\pi'_{q+K} f) - F_n^{(q+K)}(f)| \leq 2^{-(q+K)}$ for all $f \in \mathcal{F}$, we obtain

$$|U_n(f) - U_n^{(q)}(f)| = \left| \left\{ \sum_{k=1}^{K} U_n^{(q+k)}(f) - U_n^{(q+k-1)}(f) \right\} + U_n(f) - U_n^{(q+K)}(f) \right|$$

$$\leq \left\{ \sum_{k=1}^{K} \left| U_n^{(q+k)}(f) - U_n^{(q+k-1)}(f) \right| + \left| U_n(\pi'_{q+K} f) - U_n^{(q+K)}(f) \right| \right\} + \sqrt{n} 2^{-(q+K)}.$$

In order to assure $\frac{\varepsilon}{4} \leq 2^{-(q+K)}\sqrt{n} \leq \frac{\varepsilon}{2}$, for fixed n and q, choose $K = K_{n,q}$, where

$$K_{n,q} := \lfloor \log(\frac{4\sqrt{n}}{2^q \varepsilon}) \log(2)^{-1} \rfloor.$$

For each $i \in \{1, \ldots, N_q\}$, we obtain

$$\begin{aligned}\sup_{f\in\mathcal{F}_{q,i}} |U_n(f) - U_n^{(q)}(f)| \quad &\leq \quad \sum_{k=1}^{K_{n,q}} \sup_{f\in\mathcal{F}_{q,i}} |U_n^{(q+k)}(f) - U_n^{(q+k-1)}(f)| \\ &\quad + \sup_{f\in\mathcal{F}_{q,i}} |U_n(\pi'_{q+K_{n,q}} f) - U_n^{(q+K_{n,q})}(f)| + \frac{\varepsilon}{2}.\end{aligned}$$

By taking $\varepsilon_k = \frac{\varepsilon}{4k(k+1)}$, $\sum_{k\geq 1} \varepsilon_k = \frac{\varepsilon}{4}$ and we get for each $i \in \{1, \ldots, N_q\}$,

$$\begin{aligned}\mathbf{P}^*\left(\sup_{f\in\mathcal{F}_{q,i}} |U_n(f) - U_n^{(q)}(f)| \geq \varepsilon\right) \quad &\leq \quad \sum_{k=1}^{K_{n,q}} \mathbf{P}^*\left(\sup_{f\in\mathcal{F}_{q,i}} |U_n^{(q+k)}(f) - U_n^{(q+k-1)}(f)| \geq \varepsilon_k\right) \\ &\quad + \mathbf{P}^*\left(\sup_{f\in\mathcal{F}_{q,i}} |U_n(\pi'_{q+K_{n,q}} f) - U_n^{(q+K_{n,q})}(f)| \geq \frac{\varepsilon}{4}\right).\end{aligned}$$

Notice that the suprema in the r.h.s. are in fact maxima over finite numbers of functions, since the functionals π_q and π'_q (and thus $U_n^{(q)}$) are constant on the $\mathcal{F}_{q,i}$. Therefore the outer probabilities can be replaced by usual probabilities. For each k, choose (by the axiom of choice) a set F_k composed by at most $N_{k-1}N_k$ functions of $\mathcal{F}$ in such a way that F_k contains one function in each non empty $\mathcal{F}_{k-1,i} \cap \mathcal{F}_{k,j}$, $i = 1, \ldots, N_{k-1}$, $j = 1, \ldots, N_k$. Then, for each $i \in \{1, \ldots, N_q\}$, we have

$$\begin{aligned}\mathbf{P}^*\left(\sup_{f\in\mathcal{F}_{q,i}} |U_n(f) - U_n^{(q)}(f)| \geq \varepsilon\right) \quad &\leq \quad \sum_{k=1}^{K_{n,q}} \sum_{f\in\mathcal{F}_{q,i}\cap F_{q+k}} \mathbf{P}\left(|U_n^{(q+k)}(f) - U_n^{(q+k-1)}(f)| \geq \varepsilon_k\right) \\ &\quad + \sum_{f\in\mathcal{F}_{q,i}\cap F_{q+K_{n,q}}} \mathbf{P}\left(|U_n(\pi'_{q+K_{n,q}} f) - U_n^{(q+K_{n,q})}(f)| \geq \frac{\varepsilon}{4}\right).\end{aligned}$$

Now using Markov's inequality at the order $2p$ (p will be chosen later) and assumption (1.2), we infer

$$\begin{aligned}&\mathbf{P}^*\left(\sup_{f\in\mathcal{F}_{q,i}} |U_n(f) - U_n^{(q)}(f)| \geq \varepsilon\right) \\ &\leq \sum_{k=1}^{K_{n,q}} \sum_{f\in\mathcal{F}_{q,i}\cap F_{q+k}} \frac{1}{\varepsilon_k^{2p}} \sum_{j=1}^{p} n^{j-p} \|\pi_{q+k} f - \pi_{q+k-1} f\|_1^j \log^{2p-j}(\|\pi_{q+k} f - \pi_{q+k-1} f\|_{\mathcal{B}} + 1) \\ &\quad + \sum_{f\in\mathcal{F}_{q,i}\cap F_{q+K_{n,q}}} \left(\frac{4}{\varepsilon}\right)^{2p} \sum_{j=1}^{p} n^{j-p} \|\pi_{q+K_{n,q}} f - \pi'_{q+K_{n,q}} f\|_1^j \log^{2p-j}(\|\pi_{q+K_{n,q}} f - \pi'_{q+K_{n,q}} f\|_{\mathcal{B}} + 1).\end{aligned}$$

Note that by construction, for each $k \geq 1$,

$$\begin{aligned}\|\pi_{q+k} f - \pi_{q+k-1} f\|_1 \quad &\leq \quad \|\pi_{q+k} f - f\|_1 + \|\pi_{q+k-1} f - f\|_1 \leq 3 \cdot 2^{-(q+k)} \\ \|\pi_{q+k} f - \pi'_{q+k} f\|_1 \quad &\leq \quad 2^{-(q+k)} \\ \|\pi_{q+k} f - \pi_{q+k-1} f\|_{\mathcal{B}} \quad &\leq \quad 2\exp(2^{\frac{q+k}{\gamma}}) \\ \|\pi_{q+k} f - \pi'_{q+k} f\|_{\mathcal{B}} \quad &\leq \quad 2\exp(2^{\frac{q+k}{\gamma}}).\end{aligned}$$

Thus,

$$\mathbf{P}^*\left(\sup_{f\in\mathcal{F}_{q,i}} |U_n(f)-U_n^{(q)}(f)|\geq\varepsilon\right)$$
$$\leq C\sum_{j=1}^{p}\sum_{k=1}^{K_{n,q}}\#(\mathcal{F}_{q,i}\cap F_{q+k})\frac{(k(k+1))^{2p}}{\varepsilon^{2p}}n^{j-p}2^{-j(q+k)}\log^{2p-j}(2\exp(2^{\frac{q+k}{\gamma}})+1),$$

and if q is large enough, there exists a new positive constant C such that

$$\mathbf{P}^*\left(\sup_{f\in\mathcal{F}}|U_n(f)-U_n^{(q)}(f)|\geq\varepsilon\right)$$
$$\leq\sum_{i=1}^{N_q}\mathbf{P}^*\left(\sup_{f\in\mathcal{F}_{q,i}}|U_n(f)-U_n^{(q)}(f)|\geq\varepsilon\right)$$
$$\leq C\sum_{i=1}^{N_q}\sum_{j=1}^{p}\sum_{k=1}^{K_{n,q}}\#(\mathcal{F}_{q,i}\cap F_{q+k})\frac{(k(k+1))^{2p}}{\varepsilon^{2p}}n^{j-p}2^{-j(q+k)}2^{(2p-j)\frac{q+k}{\gamma}}.$$

By construction, $\sum_{i=1}^{N_q}\#(\mathcal{F}_{q,i}\cap F_{q+k})\leq N_{q+k}^2$, thus we have

$$\begin{aligned}
\mathbf{P}^*\left(\sup_{f\in\mathcal{F}}|U_n(f)-U_n^{(q)}(f)|\geq\varepsilon\right) &\leq C\sum_{j=1}^{p}\frac{n^{j-p}}{\varepsilon^{2p}}\sum_{k=1}^{K_{n,q}}N_{q+k}^2k^{4p}2^{(2p-j(\gamma+1))\frac{q+k}{\gamma}}\\
&\leq C\sum_{j=1}^{p-1}\frac{n^{j-p}}{\varepsilon^{2p}}2^{(p-j)(\gamma+1)\frac{q+K_{n,q}}{\gamma}}\sum_{k=1}^{K_{n,q}}N_{q+k}^2k^{4p}2^{p(1-\gamma)\frac{q+k}{\gamma}}\\
&\quad+\frac{1}{\varepsilon^{2p}}\sum_{k=1}^{K_{n,q}}N_{q+k}^2k^{4p}2^{p(1-\gamma)\frac{q+k}{\gamma}}\\
&\leq C\sum_{j=1}^{p-1}\frac{n^{(p-j)(\frac{\gamma+1}{2\gamma}-1)}}{\varepsilon^{2p+(p-j)\frac{\gamma+1}{\gamma}}}\sum_{k=1}^{\infty}N_{q+k}^2k^{4p}2^{p(1-\gamma)\frac{q+k}{\gamma}}\\
&\quad+\frac{1}{\varepsilon^{2p}}\sum_{k=1}^{\infty}N_{q+k}^2k^{4p}2^{p(1-\gamma)\frac{q+k}{\gamma}}
\end{aligned}$$

where the value of the constant C (which depends on p) may vary from line to line. As $p\frac{1-\gamma}{\gamma}\to-\infty$ when p tends to infinity, there exists $p>1$, such that $p\frac{1-\gamma}{\gamma}<-2r$ and thus

$$\sum_{k=1}^{\infty}N_k^2k^{4p}2^{p(1-\gamma)\frac{k}{\gamma}}<+\infty$$

by (3.1) Then the first summand goes to zero as n goes to infinity and the second summand goes to zero as q goes to infinity. □

4. Examples of Classes of Functions

In the sequel, we fix $\alpha\in(0,1]$ and we choose $\mathcal{B}=\mathcal{H}_\alpha$, the space of bounded α-hölder continuous functions, equipped with the norm

$$\|f\|_\alpha=\sup_{x\in\mathbb{R}^d}|f(x)|+\sup_{x\neq y\in\mathbb{R}^d}\frac{|f(x)-f(y)|}{\|x-y\|^\alpha}.$$

We denote by ω_F the modulus of continuity of a function F.

4.1. **Example 1: Indicators of Rectangles.** In its classical form, the empirical process is defined by the class of indicator function of left infinite rectangles, i.e. by considering the class $\{1_{(-\infty,t]} : t \in \mathbb{R}^d\}$, where $(-\infty, t]$ denotes the set of points x such that[1] $x \leq t$. Under similar assumptions than in the present paper, this case was treated by Dehling and Durieu (2011). We will see that Theorem 1.1 covers the results of that paper.

The following proposition gives an upper bound to the bracketing number associated with the more general class of indicator functions of rectangles. A rectangle is here defined by two points. If $s, t \in [-\infty, +\infty]^d$ with $s \leq t$, the rectangle $(s, t]$ is composed by points x such that $s < x$ and $x \leq t$, that is the Cartesian product of the intervals $(s_i, t_i]$.

Proposition 4.1. *Let μ be a probability distribution on $\mathbb{R}^d$ whose distribution function F satisfies*

$$\omega_F(x) = O(|\log(x)|^{-\gamma}) \text{ for some } \gamma > 1. \tag{4.1}$$

If $\mathcal{F}$ is the class of indicator functions of rectangles in $\mathbb{R}^d$, i.e.

$$\mathcal{F} = \{1_{(s,t]} : s, t \in [-\infty, +\infty]^d, s \leq t\}$$

and $\mathcal{G}$ is the class of functions of $\mathcal{H}_\alpha$ bounded by 1, then there exists $C > 0$ such that

$$N(\varepsilon, \exp(C\varepsilon^{-\frac{1}{\gamma}}), \mathcal{F}, \mathcal{G}, L^1(\mu)) = O\left(\varepsilon^{-2d}\right)$$

as $\varepsilon \to 0$.

Proof. Let $\varepsilon \in (0, 1)$ and $m = \lfloor 6d\varepsilon^{-1} + 1 \rfloor$. For all $i \in \{1, \ldots, d\}$ and $j \in \{0, \ldots, m\}$, we define the quantiles

$$t_{i,j} := F_i^{-1}\left(\frac{j}{m}\right)$$

where $F_i^{-1}(t) := \sup\{s \in \mathbb{R} : F_i(s) \leq t\}$ is the pseudo-inverse of the marginal distribution function[2] F_i. Now, if $j = (j_1, \ldots, j_d) \in \{0, \ldots, m\}^d$, we write

$$t_j = (t_{1,j_1}, \ldots, t_{d,j_d}).$$

In the following definitions, for convenience, we will also denote by $t_{i,-1}$ or $t_{i,-2}$ the points $t_{i,0}$ and by $t_{i,m+1}$ the points $t_{i,m}$. We introduce the brackets $[l_{k,j}, u_{k,j}]$, $k \in \{0, \ldots, m\}^d$, $j \in \{0, \ldots, m\}^d$, $k \leq j$, given by the α-Hölder functions

$$l_{k,j}(x) := \begin{cases} 1 & \text{if } x \in [t_{k+1}, t_{j-2}] \\ 0 & \text{if } x \in \mathbb{R}^d \setminus [t_k, t_{j-1}] \\ \text{affine interpolation} & \text{if } x \in [t_k, t_{j-1}] \setminus [t_{k+1}, t_{j-2}] \end{cases}$$

$$u_{k,j}(x) := \begin{cases} 1 & \text{if } x \in [t_{k-1}, t_j] \\ 0 & \text{if } x \in \mathbb{R}^d \setminus [t_{k-2}, t_{j+1}] \\ \text{affine interpolation} & \text{if } x \in [t_{k-2}, t_{j+1}] \setminus [t_{k-1}, t_j] \end{cases}$$

where we have used the convention that $[s, t] = \emptyset$ if $s \not\leq t$.

[1]On $\mathbb{R}^d$, we use the partial order : $x \leq t$ if and only if $x_i \leq t_i$ for all $i = 1, \ldots, d$.

[2]$F_i(t) = \mu(\mathbb{R} \times \cdots \times \mathbb{R} \times (-\infty, t] \times \mathbb{R} \times \cdots \times \mathbb{R})$

For each $k \leq j$, we have

$$\begin{aligned}\|l_{k,j} - u_{k,j}\|_1 &\leq \mu\left([t_{k-2}, t_{j+1}] \setminus [t_{k+1}, t_{j-2}]\right) \\ &\leq \sum_{i=1}^{d} |F_i(t_{i,k_i+1}) - F_i(t_{i,k_i-2})| + |F_i(t_{i,j_i+1}) - F_i(t_{i,j_i-2})| \\ &\leq 2\frac{3d}{m} \leq \varepsilon.\end{aligned}$$

Using (4.1), we also have

$$\begin{aligned}\|l_{k,j}\|_\alpha &\leq 1 + d \max_{i=1,\ldots,d} \max\left\{\frac{1}{|t_{i,k_i} - t_{i,k_i+1}|^\alpha}, \frac{1}{|t_{i,j_i-1} - t_{i,j_i-2}|^\alpha}\right\} \\ &\leq 1 + d\left[\inf\left\{s > 0 : \exists i \in \{1,\ldots,d\}, \exists t, F_i(t+s) - F_i(t) \geq \frac{1}{m}\right\}\right]^{-\alpha} \\ &\leq 1 + d\left[\inf\left\{s > 0 : C|log(s)|^{-\gamma} \geq \frac{1}{m}\right\}\right]^{-\alpha} \\ &\leq 1 + d\exp\left(\alpha(Cm)^{\frac{1}{\gamma}}\right).\end{aligned}$$

and the same bound for $\|u_{k,j}\|_\alpha$. Thus, there exists a new constant $C > 0$ such that for all $k \leq j \in \{0,\ldots,m\}^d$, $[l_{k,j}, u_{k,j}]$ is an $(\varepsilon, \exp(C\varepsilon^{-\frac{1}{\gamma}}), \mathcal{G}, L^1(\mu))$-brackets.

It is clear that for each function $f \in \mathcal{F}$ there exist a bracket of the form $[l_{k,j}, u_{k,j}]$ which contain f. Further, we have less than $(m+1)^{2d}$ such brackets. Thus the proposition is proved. □

Notice that, under assumptions of the proposition, condition (1.3) is satisfied and therefore Theorem 1.1 applies to empirical processes indexed by the class of indicators of rectangles, regarding the class of bounded Hölder functions.

Corollary 4.2. *Let $(X_i)_{i\leq 0}$ be a $\mathbb{R}^d$-valued stationary process. Let $\mathcal{F}$ be the class of indicator functions of rectangles in $\mathbb{R}^d$ and $\mathcal{G}$ be the class of α-Hölder functions bounded by 1. Assume that (1.1), (1.2) hold, and that the distribution function of the X_i satisfies (4.1). Then the empirical process $(U_n(f))_{f\in\mathcal{F}}$ converges in distribution in $\ell^\infty(\mathcal{F})$ to a tight Gaussian process.*

Remark 4.3. By regarding the class of indicator functions of left infinite rectangles as a sub-class of $\mathcal{F}$, we obtain Theorem 1 of Dehling and Durieu (2011) as a particular case of the preceding corollary.

4.2. **Example 2: Indicators of Multidimensional Balls in the Unit Cube.** Here, we consider the class $\mathcal{F}$ of indicator functions of balls on $\mathcal{X} = [0,1]^d$, i.e.

$$\mathcal{F} := \{1_{B(x,r)} : x \in [0,1]^d, r \geq 0\}$$

where $B(x,r) = \{y \in \mathcal{X} : \|x - y\| < r\}$ and $\|\cdot\|$ denotes the Euclidian norm on $\mathbb{R}^d$. Let $\mathcal{G}$ be the class of α-Hölder functions on $[0,1]^d$ which take values in $[0,1]$. We have the following upper bound.

Proposition 4.4. *Let μ be a probability distribution on $\mathcal{X}$ with bounded density. Then there exists $C > 0$ such that*

$$N(\varepsilon, C\varepsilon^{-\alpha}, \mathcal{F}, \mathcal{G}, L^1(\mu)) = O\left(\varepsilon^{-(d+1)}\right).$$

Remark 4.5. The reader can check in the proof, that Proposition 4.4 still holds if we consider balls of the torus $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$. We will use this result to prove Theorem 1.4:

Proof of Proposition 4.4. Let $\varepsilon > 0$ be fixed and $m = \lfloor \frac{1}{\varepsilon} \rfloor$. For all $i = (i_1, \dots, i_d) \in \{0, \dots, m\}^d$, we denote by c_i the center of the rectangle $[\frac{i_1-1}{m}, \frac{i_1}{m}] \times \cdots \times [\frac{i_d-1}{m}, \frac{i_d}{m}]$. Then we define, for $i \in \{1, \dots, m\}^d$ and $j \in \{0, \dots, m\}$,

$$l_{i,j}(x) := \begin{cases} 1 & \text{if } x \in B\left(c_i, \frac{j-2}{m}\sqrt{d}\right) \\ 0 & \text{if } x \in \mathcal{X} \setminus B\left(c_i, \frac{j-1}{m}\sqrt{d}\right) \\ \text{affine interpolation} & \text{if } x \in B\left(c_i, \frac{j-1}{m}\sqrt{d}\right) \setminus B\left(c_i, \frac{j-2}{m}\sqrt{d}\right) \end{cases}$$

$$u_{i,j}(x) := \begin{cases} 1 & \text{if } x \in B\left(c_i, \frac{j+2}{m}\sqrt{d}\right) \\ 0 & \text{if } x \in \mathcal{X} \setminus B\left(c_i, \frac{j+3}{m}\sqrt{d}\right) \\ \text{affine interpolation} & \text{if } x \in B\left(c_i, \frac{j+3}{m}\sqrt{d}\right) \setminus B\left(c_i, \frac{j+2}{m}\sqrt{d}\right) \end{cases}$$

where we use the convention that a ball with negative radius is the empty set. These functions are α-Hölder (for any $\alpha \in (0, 1]$) and satisfy

$$\|l_{i,j}\|_\alpha = \|u_{i,j}\|_\alpha = 1 + \left(\frac{m}{\sqrt{d}}\right)^\alpha .$$

Further, since μ has a bounded density with respect to Lebesgue measure, we have

$$\begin{aligned} \|l_{i,j} - u_{i,j}\|_1 &\le \mu\left(B\left(c_i, \frac{j+3}{m}\sqrt{d}\right) \setminus B\left(c_i, \frac{j-2}{m}\sqrt{d}\right)\right) \\ &\le C\left(\left(\frac{j+3}{m}\sqrt{d}\right)^d - \left(\frac{j-2}{m}\sqrt{d}\right)^d\right) \\ &= O\left(\frac{1}{m}\right). \end{aligned}$$

Further we have $(m+1)m^d$ of these couple of functions.

Now, if f belongs to $\mathcal{F}$, $f = 1_{B(x,r)}$ for some $x \in \mathcal{X}$, and $0 \le r \le \sqrt{d}$. Thus there exist $i = (i_1, \dots, i_d) \in \{0, \dots, m\}^d$ and $j \in \{0, \dots, m\}$ such that

$$x \in \left[\frac{i_1 - 1}{m}, \frac{i_1}{m}\right) \times \cdots \times \left[\frac{i_d - 1}{m}, \frac{i_d}{m}\right)$$

and

$$\frac{j}{m}\sqrt{d} \le r \le \frac{j+1}{m}\sqrt{d}.$$

We then have $l_{i,j} \le f \le u_{i,j}$.

Thus the $(m+1)m^d$ brackets $[l_{i,j}, u_{i,j}]$, $i \in \{1, \dots, m\}^d$ and $j \in \{0, \dots, m\}$, cover the class $\mathcal{F}$. Therefore, there exist $C > 0$, such that $N(\varepsilon, C\varepsilon^{-\alpha}, \mathcal{F}, \mathcal{G}, L^1(\mu)) = O(\frac{1}{\varepsilon^{d+1}})$. □

4.3. **Proof of Theorem 1.4.** Let $\mathcal{F}$ be the class of indicators of balls on $\mathbb{T}^d$ and $\mathcal{G}$ be the class of α-Hölder functions bounded by 1 for some $\alpha > 1/2$. We consider the $\mathbb{T}^d$-valued stationary process $X_i = id \circ T^i$. Since, the distribution of X_0 is the Lebesgue measure on $\mathbb{T}^d$, Proposition 4.4 shows that Condition (1.3) holds. For all $f \in \mathcal{G}$, the central limit theorem (1.1) holds; see Leonov (1960) and Le Borgne (1999). Dehling and Durieu (2011), Proposition 3, show that ergodic automorphisms of the torus satisfy a multiple mixing property. By Theorem 4 of Dehling and Durieu (2011), this implies the $2p$-th moment bound (1.2) for Hölder functions. Thus Theorem 1.1 can be applied. □

4.4. **Example 3: Indicators of Uniformly Bounded Multidimensional Ellipsoids, Centered in the Unit Cube.** Set $\mathcal{X} = [0,1]^d$ (we can also work with any other bounded subspace of $\mathbb{R}^d$) and let $\|\cdot\|$ denotes the euclidean norm on $\mathbb{R}^d$. Here, we consider the class of ellipsoids which are aligned with coordinate axes, have their center in $\mathcal{X}$ and their parameters bounded by some constant $D \in \mathbb{N}$. For all $x = (x_1, \dots, x_d) \in \mathcal{X}$ and all $r = (r_1, \dots, r_d) \in [0, D]^d$, we set

$$\mathcal{E}(x,r) := \{y \in \mathbb{R}^d : \sum_{i=1}^d \frac{(y_i - x_i)^2}{r_i^2} \le 1\}.$$

We denote by $\mathcal{F}$ the class of indicator functions of these ellipsoids, i.e.

$$\mathcal{F} := \{1_{\mathcal{E}(x,r)} : x \in \mathcal{X}, r \in [0,D]^d\}.$$

Let $\mathcal{G}$ be the class of α-Hölder functions on $[0,1]^d$ which take values in $[0,1]$. We have the following upper bound.

Proposition 4.6. *Let μ be a probability distribution on $\mathbb{R}^d$ with bounded density. Then there exists $C > 0$ (which depends on d and D) such that*

$$N(\varepsilon, C\varepsilon^{-\alpha}, \mathcal{F}, \mathcal{G}, L^1(\mu)) = O\left(\varepsilon^{-2d}\right).$$

Proof. Let $\varepsilon > 0$ be fixed and $m = \lfloor \frac{1}{\varepsilon} \rfloor$. For all $i = (i_1, \dots, i_d) \in \{0, \dots, m\}^d$, we denote by I_i the rectangle $[\frac{i_1-1}{m}, \frac{i_1}{m}] \times \dots \times [\frac{i_d-1}{m}, \frac{i_d}{m}]$. Then, for $i \in \{1, \dots, m\}^d$ and $j = (j_1, \dots, j_d) \in \{0, \dots, Dm-1\}^d$, we define the sets

$$U_{i,j} = \bigcup_{x \in I_i} \mathcal{E}\left(x, \frac{j}{m}\right) = \left\{y \in \mathbb{R}^d : \min_{x \in C_i} \sum_{k=1}^d \frac{(y_k - x_k)^2}{j_k^2} \le 1\right\}$$

and

$$L_{i,j} = \bigcap_{x \in I_i} \mathcal{E}\left(x, \frac{j}{m}\right) = \left\{y \in \mathbb{R}^d : \max_{x \in C_i} \sum_{k=1}^d \frac{(y_k - x_k)^2}{j_k^2} \le 1\right\}.$$

We set

$$l_{i,j}(x) := \begin{cases} 1 & \text{if } x \in L_{i,j-1} \\ 0 & \text{if } x \in \mathbb{R}^d \setminus L_{i,j} \\ \text{affine interpolation} & \text{if } x \in L_{i,j} \setminus L_{i,j-1} \end{cases}$$

$$u_{i,j}(x) := \begin{cases} 1 & \text{if } x \in U_{i,j+1} \\ 0 & \text{if } x \in \mathbb{R}^d \setminus U_{i,j+2} \\ \text{affine interpolation} & \text{if } x \in U_{i,j+2} \setminus U_{i,j+1} \end{cases}$$

where we use the convention that an ellipsoid with a negative parameter is the empty set. Since the distance between $U_{i,j}$ and $\mathbb{R}^d \setminus U_{i,j+1}$ is $\frac{1}{m}$ and the distance between $L_{i,j}$ and $\mathbb{R}^d \setminus L_{i,j+1}$ is $\frac{1}{m}$, these functions are α-Hölder (for any $\alpha \in (0,1]$) with

$$\|l_{i,j}\|_\alpha = \|u_{i,j}\|_\alpha = 1 + m^\alpha.$$

Recall that, if $j = (j_1, \dots, j_d) \in \mathbb{R}_+^d$ and $x \in \mathbb{R}^d$, the Lebesgue measure of the ellipsoid $\mathcal{E}(x,j)$ is given by

$$\lambda(\mathcal{E}(x,j) = C_d \prod_{k=1}^d j_k,$$

where C_d is the constant $\frac{\pi^{d/2}}{\Gamma(d/2+1)}$ (Γ is the gamma function).

By construction, $U_{i,j}$ is an ellipsoid (of parameters j/m) that is cut along its hyperplanes of symmetry and stretched orthogonally to them of a thickness $1/m$. Then its volume can be computed by the formula

$$\begin{aligned}\lambda(U_{i,j}) &= C_d \prod_{k=1}^{d} \frac{j_k}{m} + \frac{C_{d-1}}{m} \sum_{k=1}^{d} \prod_{l \neq k} \frac{j_l}{m} + \dots \\ &\quad \dots + \frac{C_{d-p}}{m^p} \sum_{\{k_1,\dots,k_p\} \subset \{1,\dots,d\}} \prod_{l \notin \{k_1,\dots,k_p\}} \frac{j_l}{m} + \dots \\ &\quad \dots + \frac{C_2}{m^{d-1}} \sum_{k=1}^{d} \frac{j_k}{m} + \frac{1}{m^d}.\end{aligned}$$

We also have

$$\begin{aligned}\lambda(L_{i,j}) &\geq C_d \prod_{k=1}^{d} \frac{j_k}{m} - \frac{C_{d-1}}{m} \sum_{k=1}^{d} \prod_{l \neq k} \frac{j_l}{m} - \dots \\ &\quad \dots - \frac{C_{d-p}}{m^p} \sum_{\{k_1,\dots,k_p\} \subset \{1,\dots,d\}} \prod_{l \notin \{k_1,\dots,k_p\}} \frac{j_l}{m} - \dots \\ &\quad \dots - \frac{C_2}{m^{d-1}} \sum_{k=1}^{d} \frac{j_k}{m} - \frac{1}{m^d}.\end{aligned}$$

Since μ has a bounded density with respect to Lebesgue measure (say bounded by some constant K), we have

$$\begin{aligned}\|l_{i,j} - u_{i,j}\|_1 &\leq \mu\left(U_{i,j+2} \setminus L_{i,j-1}\right) \\ &\leq K\lambda(U_{i,j+2}) - K\lambda(L_{i,j-1}) \\ &= O\left(\frac{1}{m}\right).\end{aligned}$$

Now, if f belongs to $\mathcal{F}$, $f = 1_{\mathcal{E}(x,r)}$ for some $x \in \mathcal{X}$, and $r \in [0, D]^d$. Thus there exist $i = (i_1, \dots, i_d) \in \{0, \dots, m\}^d$ and $j \in \{0, \dots, Dm - 1\}^d$ such that

$$x \in \left[\frac{i_1 - 1}{m}, \frac{i_1}{m}\right) \times \dots \times \left[\frac{i_d - 1}{m}, \frac{i_d}{m}\right)$$

and for each $k = 1, \dots, d$,

$$\frac{j_k}{m} \leq r_k \leq \frac{j_k + 1}{m}.$$

We then have $l_{i,j} \leq f \leq u_{i,j}$.

Thus the $D^d m^{2d}$ brackets $[l_{i,j}, u_{i,j}]$, $i \in \{1, \dots, m\}^d$ and $j \in \{0, \dots, Dm - 1\}^d$, cover the class $\mathcal{F}$. Since we have $m^d \cdot (Dm)^d$ of these couples of functions, there exist $C > 0$, such that $N(\varepsilon, C\varepsilon^{-\alpha}, \mathcal{F}, \mathcal{G}, L^1(\mu)) = O(\frac{1}{\varepsilon^{2d}})$. □

4.5. **Example 4: Indicators of Uniformly Bounded, Multidimensional Ellipsoids.** In Example 3 we only considered indicators of ellipsoids centered in a compact subset of $\mathbb{R}^d$ (the unit square). The following lemma will allow us to extend such results to statements for indicators of sets in the whole $\mathbb{R}^d$ on the cost of a moderate additional assumption and a (marginal) increase of the bracketing numbers.

Lemma 4.7. *Let $\mathcal{X} = \mathbb{R}^d$ equipped with the Euclidian norm and a measure μ which is absolutely continuous with respect to the Lebesgue measure. Let $\mathcal{G}$ be the class of α-Hölder functions on $\mathbb{R}^d$ which take values in $[0,1]$ and $\mathcal{F} := \{1_S : S \in \mathcal{S}\}$, where $\mathcal{S}$ is a class of measurable sets of diameter not larger than D.*

Assume that there are universal constants $p, q \in \mathbb{N}$ and $C', C'' > 0$, such that for any $K > 0$ we have

$$N(\varepsilon, \varphi(\varepsilon), \mathcal{F}_K, \mathcal{G}, L^1(\mu)) \leq C'' K^p \varepsilon^{-q} \tag{4.2}$$

where $\mathcal{F}_K := \{1_S : S \in \mathcal{S},\ S \subset [-K,K]^d\}$ and $\varphi(\varepsilon) \leq \exp(C'\varepsilon^{-1/\gamma})$ for all sufficiently small $\varepsilon > 0$. If furthermore

$$\omega_F(x) = O(|log(x)|^{-\gamma}) \tag{4.3}$$

and

$$\overline{H}(t) := \mu(\{x \in \mathbb{R}^d : |x| > t\}) = O(t^{-\frac{1}{\beta}}) \tag{4.4}$$

for some norm $|\cdot|$ on $\mathbb{R}^d$, then there exists a $C \geq C'$ such that

$$N(\varepsilon, \exp(C\varepsilon^{-\frac{1}{\gamma}}), \mathcal{F}, \mathcal{G}, L^1(\mu)) = O(\varepsilon^{-(\beta p + q)}).$$

Proof. Note that, if there is an $\varepsilon_0 > 0$ such that $\varphi(\varepsilon) \leq \psi(\varepsilon)$ for all sufficiently small $\varepsilon \in (0, \varepsilon)$, for these ε every $(\varepsilon, \varphi(\varepsilon), \mathcal{G}, L^1(\mu))$-bracket is also an $(\varepsilon, \psi(\varepsilon), \mathcal{G}, L^1(\mu))$-bracket and thus

$$N(\varepsilon, \psi(\varepsilon), \mathcal{G}, L^1(\mu)) \leq N(\varepsilon, \varphi(\varepsilon), \mathcal{G}, L^1(\mu)) \text{ for all } \varepsilon \in (0, \varepsilon_0). \tag{4.5}$$

Set $\psi(\varepsilon) = \exp(C'\varepsilon^{-1/\gamma})$ and choose $\varepsilon_0 > 0$ such that condition (4.2) holds for all $\varepsilon \in (0, \varepsilon_0)$. Then (4.5) implies that for all these ε

$$N(\varepsilon, \exp(C'\varepsilon^{-\frac{1}{\gamma}}), \mathcal{F}_{K_{\varepsilon/2}+D}, \mathcal{G}, L^1(\mu)) \leq C''(K_{\varepsilon/2} + D)^p \varepsilon^{-q}.$$

Fix $\varepsilon \in (0, \varepsilon_0)$ for the time and set $K_\varepsilon = \sup\{K > 0 : \mu([-K,K]^d) \leq 1 - \varepsilon\}$. We will denote the function $(0, \varepsilon_0) \to \mathbb{R}^+$, $\varepsilon \mapsto K_\varepsilon$ by $K_\bullet$. Now introduce the bracket $[U_\varepsilon, L]$, given by $L \equiv 0$

$$U_\varepsilon := \begin{cases} 1, & \text{if } x \notin [-K_{\varepsilon/2}, K_{\varepsilon/2}]^d \\ 0, & \text{if } x \in [-K_\varepsilon, K_\varepsilon]^d \\ \text{affine interpolation}, & \text{if } x \in [-K_{\varepsilon/2}, K_{\varepsilon/2}]^d \backslash [-K_\varepsilon, K_\varepsilon]^d. \end{cases}$$

Obviously we have $\|U_\varepsilon - L\|_1 \leq \varepsilon$.

To get a boundary for the α-norm of U_ε consider the distribution function

$$G(t) := \mu\big(\{x \in \mathbb{R}^d : |x|_{\max} \leq t\}\big)$$

on $\mathbb{R}$. Observe that the pseudoinverse G^{-1} of G, given by $G^{-1}(s) := \sup\{t \in \mathbb{R} : G(t) \leq s\}$, is linked to $K_\bullet$ by the equality

$$K_\varepsilon = G^{-1}(1 - \varepsilon).$$

With geometrical arguments we infer

$$G(t) = \sum_{j \in \{-1,1\}^d} \sigma(j) F(tj),$$

where $\sigma(j) := \prod_{i=1}^{d} j_i \in \{-1,1\}$. Therefore

$$\begin{aligned}\omega_G(s) &= \sup_{t\in\mathbb{R}}\{G(t+s)-G(t)\} = \sup_{t\in\mathbb{R}} \sum_{j\in\{-1,1\}^d} \sigma(j)\big(F((t+s)j)-F(tj)\big)\\ &\le \sum_{j\in\{-1,1\}^d} \sup_{t\in\mathbb{R}}\big|F((t+s)j)-F(tj)\big| \le \sum_{j\in\{-1,1\}^d} \omega_F(\sqrt{d}s)\\ &\le 2^d\omega_F(\sqrt{d}s).\end{aligned}$$

By assumption this implies that

$$\omega_G(s) = O(|log(\sqrt{d}s)|^{-\gamma}).$$

Now with the same arguments in the proof of Proposition 4.1 we obtain

$$\begin{aligned}\|U_\varepsilon\|_\alpha &\le 1 + \frac{1}{|G^{-1}(1-\frac{\varepsilon}{2}) - G^{-1}(1-\varepsilon)|^\alpha}\\ &\le 1 + \inf\{s>0 : \exists t\in\mathbb{R} \text{ such that } G(t+s)-G(t) \ge \frac{\varepsilon}{2}\}^{-\alpha}\\ &\le 1 + \inf\{s>0 : \exists t\in\mathbb{R} \text{ such that } \omega_G(s) \ge \frac{\varepsilon}{2}\}^{-\alpha}\\ &\le 1 + \frac{1}{\sqrt{d}}\exp(C'''\varepsilon^{-\frac{1}{\gamma}})\\ &= \exp(C'''\varepsilon^{-\frac{1}{\gamma}}),\end{aligned}$$

where the constant C''' may change from line to line.

Thus $[U_\varepsilon, L]$ is a $(\varepsilon, \exp(C'''\varepsilon^{-1/\gamma}), \mathcal{G}, L^1(\mu))$-bracket. Since it contains any $f \in \mathcal{F}\backslash\mathcal{F}_{K_{\varepsilon/2}+D}$ we obtain the bound

$$N(\varepsilon, \exp(C\varepsilon^{-\frac{1}{\gamma}}), \mathcal{F}, \mathcal{G}, L^1(\mu)) \le C(K_{\varepsilon/2}+D)^p\varepsilon^{-q} + 1$$

where $C = \max\{C', C'''\}$.

Let us finally consider the growing rate of $K_{\varepsilon/2}$ as $\varepsilon \to \infty$. Since every norm in $\mathbb{R}^d$ is equivalent there is a constant $a > 0$ such that $|x|/a > |x|_{\max}$ for every $x \in \mathbb{R}^d$. Furthermore, by assumption there is a $b > 0$ such that $\overline{H}(t) \le bt^{-1/\beta}$. Therefore

$$\begin{aligned}\mu\Big(\big\{x\in\mathbb{R}^d : |x|_{\max} < a^{-1}(2b)^\beta\varepsilon^{-\beta}\big\}\Big) &= 1 - \mu\Big(\big\{x\in\mathbb{R}^d : |x|_{\max} \ge a^{-1}(2b)^\beta\varepsilon^{-\beta}\big\}\Big)\\ &\ge 1 - \mu\Big(\big\{x\in\mathbb{R}^d : |x| \ge (2b)^\beta\varepsilon^{-\beta}\big\}\Big)\\ &= 1 - \overline{H}\big((2b)^\beta\varepsilon^{-\beta}\big)\\ &\ge 1 - \frac{\varepsilon}{2}.\end{aligned}$$

By the definition of $K_\bullet$ we therefore obtain that $K_{\varepsilon/2} \le a^{-1}(2b)^\beta\varepsilon^{-\beta} = O(\varepsilon^{-\beta})$ and thus

$$N(\varepsilon, \exp(C\varepsilon^{-\frac{1}{\gamma}}), \mathcal{F}, \mathcal{G}, L^1(\mu)) = O(\varepsilon^{-(\beta p+q)}).$$

□

In the situation of Example 3 change $\mathcal{X}$ to $[-K,K]^d$. Since we have a bounded density, condition (4.3) is alway satisfied. Following the proof of Proposition 4.6 we can easily see that condition (4.2) holds for $p = d$ and $q = 2d$.

Thus Proposition 4.6 can be extended to

Corollary 4.8. *Let $\mathcal{F}$ denote the class of indicators of ellipsoids of uniformly bounded diameter in the whole space $\mathbb{R}^d$.*

If (4.4) holds there exists a constant $C > 0$ such that

$$N(\varepsilon, C\varepsilon^{-\alpha}, \mathcal{F}, \mathcal{G}, L^1(\mu)) = O\left(\varepsilon^{-(\beta+2)d}\right).$$

Remark 4.9. Under the assumption that (4.4) holds, for the class of indicators of balls in $\mathbb{R}^d$ with uniformly bounded diameter, we can obtain the slightly sharper bound

$$N(\varepsilon, C\varepsilon^{-\alpha}, \mathcal{F}, \mathcal{G}, L^1(\mu)) = O(\varepsilon^{-(\beta+1)d+1})$$

for some $C > 0$ by applying Lemma 4.7 directly to the situation in Example 3 and using the same arguments as in the previous example.

4.6. **Example 5: Indicators of Zero-Centered Balls of an Arbitrary Norm.** In $\mathcal{X} = \mathbb{R}^d$ a zero centered ball of the norm $\|\cdot\|_*$ is given by

$$B(t) := \{x \in \mathbb{R}^d : \|x\|_* \leq t\}.$$

Let $\mathcal{G}$ be the class of α-Hölder functions on $\mathbb{R}^d$ with values in $[0,1]$. We have the following bound on the bracketing numbers of the class $\mathcal{F} := \{1_{B(t)} : t > 0\}$:

Proposition 4.10. *If for the probability measure μ on $\mathbb{R}^d$ the modulus of continuity ω_G of the function*

$$G(t) := \mu(B(t))$$

satisfies

$$\omega_G(s) = O(|\log s|^{-\gamma}), \tag{4.6}$$

then there is a constant $C > 0$ such that

$$N(\varepsilon, \exp(C\varepsilon^{-\frac{1}{\gamma}}), \mathcal{F}, \mathcal{G}, L^1(\mu)) = O(\varepsilon^{-1}).$$

Remark 4.11. Note that in the case that $\mathcal{X} = \mathbb{R}^2$, $d\mu(t) = \rho(t)dt$ and $\|\cdot\|_*$ denotes the usual Euclidian norm, a equivalent condition to (4.6) is

$$\sup_{r\geq 0} \int_r^{r+s} t \int_0^{2\pi} \rho(te^{i\varphi})\, d\varphi\, dt = O(|\log s|^{-\gamma}).$$

Proof of Proposition 4.10. Let G^{-1} denote the pseudoinverse of G given by

$$G^{-1}(s) := \sup\{t \in \mathbb{R} : G(t) \leq s\}$$

and set for $i \in \{1, \ldots, m\}$

$$r_i := G^{-1}\Big(\frac{1}{m}\Big), \qquad B_i := B(r_i).$$

For convenience set $B_{-1}, B_0 := \emptyset$ and $B_{m+1} = \mathbb{R}^d$. Define

$$l_i := \begin{cases} 1, & \text{if } x \in B_{i-2} \\ 0, & \text{if } x \in (B_{i-1})^c \\ \text{affine interpolation}, & \text{if } x \in B_{i-1} \backslash B_{i-2} \end{cases}$$

and

$$u_i := \begin{cases} 1, & \text{if } x \in B_i \\ 0, & \text{if } x \in (B_{i+1})^c \\ \text{affine interpolation}, & \text{if } x \in B_{i+1} \backslash B_i. \end{cases}$$

The system $\{[l_i, u_i] : i \in \{1, \dots, m\}\}$ is a covering for $\mathcal{F}$. Obviously

$$\|u_i - l_i\|_1 = O(m^{-1}).$$

By the equivalence of all norms on $\mathbb{R}^d$ there must be some constant $c > 0$ (depending only on the choice of the norm $\|\cdot\|_*$ and the norm $\|\cdot\|$ in the definition of the α-norm) such that

$$\|u_i\|_\alpha \le 1 + \frac{1}{c|r_{i+1} - r_i|^\alpha}$$

Since by condition (4.6)

$$\begin{aligned} r_{i+1} - r_i &\ge \inf\Big\{ s > 0 : \exists t \in \mathbb{R} \text{ such that } G(t+s) - G(t) \ge \frac{1}{m} \Big\} \\ &\ge \inf\Big\{ s > 0 : \exists t \in \mathbb{R} \text{ such that } \omega_G(s) \ge \frac{1}{m} \Big\} \\ &\ge \exp(-C' m^{\frac{1}{\gamma}}) \end{aligned}$$

for some constant $C' > 0$, there is a constant $C > 0$ such that

$$\|u_i\|_\alpha \le 1 + \frac{1}{c|r_{i+1} - r_i|^\alpha} \le 1 + c^{-1} \exp(\alpha C' m^{\frac{1}{\gamma}}) \le \exp(C m^{\frac{1}{\gamma}})$$

Analogously we can show that $\|l_i\|_\alpha \le \exp(\alpha C m^{\frac{1}{\gamma}})$. This implies that all $[l_i, u_i]$ are $(\varepsilon, \exp(C\varepsilon^{-\frac{1}{\gamma}}), \mathcal{F}, \mathcal{G}, L^1(\mu$ brackets and thus

$$N(m^{-1}, \exp(C m^{\frac{1}{\gamma}}), \mathcal{F}, \mathcal{G}, L^1(\mu)) = O(m).$$

□

References

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

[2] Patrick Billingsley (1968): *Convergence of Probability Measures.* J. Wiley & Sons, New York.

[3] Svetlana Borovkova, Robert Burton and Herold Dehling (2001): Limit theorems for functionals of mixing processes with applications to U-statistics and dimension estimation. *Transactions of the American Mathematical Society* **353**, 4261–4318.

[4] Richard Bradley (2007): *Introduction to Strong Mixing Conditions*, Vol. 1-3, Kendrick Press, Heber City, Utah.

[5] Jérôme Dedecker, Paul Doukhan, Gabriel Lang, José Rafael León, Sana Louhichi and Clémentine Prieur (2007): *Weak Dependence*. Springer Lecture Notes in Statistics **170**.

[6] Jérôme Dedecker and Clémentine Prieur (2007): An Empirical Process Central Limit Theorem for Dependent Sequences. *Stochastic Processes and Their Applications* **117**, 121–142.

[7] Herold Dehling and Walter Philipp (2002): Empirical Process Techniques for Dependent Data. in: H. G. Dehling, T. Mikosch and M. Sorensen (eds.), *Empirical Process Techniques for Dependent Data*. Birkhäuser, Boston, 3–111.

[8] Herold Dehling, Olivier Durieu and Dalibor Volny (2009): New Techniques for Empirical Processes of Dependent Data. *Stochastic Processes and their Applications* **119**, 3699–3718.

[9] Herold Dehling and Olivier Durieu (2011): Empirical Processes of Multidimensional Systems with Multiple Mixing Properties. *Stochastic Processes and their Applications* **121**, 1076–1096.

[10] Paul Doukhan, Pascal Massart and Emmanuel Rio (1995): Invariance Principle for Absolutely Regular Empirical Processes. *Annales de l'Institut Henri Poincaré - Probabilités et Statistiques* 31, 393–427.

[11] Olivier Durieu (2009): A Fourth Moment Inequality for Functionals of Stationary Processes. *Journal of Applied Probability* **45**, 1086–1096.

[12] Olivier Durieu and Philippe Jouan (2008): Empirical invariance principle for ergodic torus automorphisms; genericity. *Stoch. Dyn.*, 8(2):173–195.

[13] Olivier Durieu and Marco Tusche. An Empirical Process Central Limit Theorem for Multidimensional Dependent Data. Preprint, Online on http://arxiv.org/abs/1110.0963 [State: 06.10.2011]

[14] Dmitry Dolgopyat (2004): Limit theorems for partially hyperbolic systems. *Trans. Amer. Math. Soc.*, 356(4):1637–1689 (electronic).

[15] Hubert Hennion and Loîc Hervé (2001): *Limit Theorems for Markov Chains and Stochastic Properties of Dynamical Systems by Quasi-Compactness.* Lecture Notes in Mathematics **1766**, Springer Verlag.

[16] Stéphane Le Borgne (1999): Limit theorems for non-hyperbolic automorphisms of the torus. *Israel J. Math.*, 109:61–73.

[17] V. P. Leonov. (1960): On the central limit theorem for ergodic endomorphisms of compact commutative groups. *Dokl. Akad. Nauk SSSR*, 135:258–261.

[18] Walter Philipp (1982): Invariance Principles for Sums of Mixing Random Elements and the Multivariate Empirical Process. *Colloquia Mathematica Societatis Janos Bolyai* **36**, 843–873.

[19] Aad W. van der Vaart and Jon A. Wellner (1996): *Weak Convergence and Empirical Processes.* Springer Verlag, New York.

(H. Dehling) Fakultät für Mathematik, Ruhr-Universität Bochum, 44780 Bochum, Germany
E-mail address: herold.dehling@rub.de

(O. Durieu) Laboratoire de Mathématiques et Physique Théorique, UMR 6083 CNRS, Université Francois Rabelais de Tours, Parc de Grandmont, 37200 Tours, France
E-mail address: olivier.durieu@lmpt.univ-tours.fr

(M. Tusche) Fakultät für Mathematik, Ruhr-Universität Bochum, 44780 Bochum, Germany
E-mail address: marco.tusche@rub.de