

EMPIRICAL PROCESSES OF MULTIDIMENSIONAL SYSTEMS WITH MULTIPLE MIXING PROPERTIES

Herold Dehling* Olivier Durieu†

January 26, 2011

Abstract

We establish a multivariate empirical process central limit theorem for stationary \mathbb{R}^d -valued stochastic processes $(X_i)_{i \geq 1}$ under very weak conditions concerning the dependence structure of the process. As an application we can prove the empirical process CLT for ergodic torus automorphisms. Our results also apply to Markov chains and dynamical systems having a spectral gap on some Banach space of functions. Our proof uses a multivariate extension of the techniques introduced by Dehling, Durieu & Volný (2009) in the univariate case. As an important technical ingredient, we prove a $2p$ -th moment bound for partial sums in multiple mixing systems.

Keywords: Multivariate Empirical Processes, Multiple Mixing Property, Dynamical Systems, Spectral Gap Property.

AMS classification: 60F17; 60G10; 62G30.

1 Introduction and Statement of Main Results

Let $(X_i)_{i \geq 1}$ be an \mathbb{R}^d -valued stationary stochastic process with multivariate marginal distribution function $F(t) = P(X_1 \leq t)$, $t \in \mathbb{R}^d$. We define the empirical distribution function and the empirical process by

$$\begin{aligned} F_n(t) &:= \frac{1}{n} \#\{1 \leq i \leq n : X_i \leq t\} \\ U_n(t) &:= \sqrt{n}(F_n(t) - F(t)), \end{aligned}$$

$t \in \mathbb{R}^d$. Here " \leq " denotes the coordinate-wise ordering, i.e. $(t_1, \dots, t_d) \leq (s_1, \dots, s_d)$ if and only if $t_i \leq s_i$ for all $i \in \{1, \dots, d\}$. In this paper, we study weak convergence of the empirical process towards a Gaussian process in the space $D([-\infty, \infty]^d)$. We make very weak assumptions concerning the dependence structure of the underlying process $(X_i)_{i \geq 1}$. Effectively, we require a multiple mixing condition, the central limit theorem for partial sums of a restricted class of functions and a condition on the modulus of continuity of the distribution function F . Our results apply to Markov chains and dynamical systems whose transfer operator has

*Fakultät für Mathematik, Ruhr-Universität Bochum, Universitätsstraße 150, 44780 Bochum, Germany; e-mail: herold.dehling@ruhr-uni-bochum.de; Research supported by the German Science Foundation, Sonderforschungsbereich 823

†Laboratoire de Mathématiques et Physique Théorique, UMR 6083 CNRS, Université François Rabelais de Tours, Parc de Grandmont, 37200 Tours, France; e-mail: olivier.durieu@lmpt.univ-tours.fr

a spectral gap on some Banach space of functions. As most significant application, we can establish the multivariate empirical process CLT for non-hyperbolic ergodic torus automorphisms, a system that does not have a spectral gap on common spaces of functions.

The study of empirical processes was initiated by Donsker's empirical process invariance principle (see Donsker 1952), which covered the case of i.i.d. \mathbb{R} -valued observations. Donsker's original theorem has been generalized to dependent variables by a number of authors, starting with work by Billingsley (1968) who could establish the empirical process invariance principle for functionals of uniformly mixing processes. Billingsley applied this result in his investigations of statistical properties of the continued fraction expansion. Berkes and Philipp (1977/78) were able to treat the empirical process of strongly mixing sequences. Borovkova, Burton and Dehling (2001) could treat functionals of absolutely regular processes. Dehling and Taqqu (1989) proved an empirical process invariance principle for long-range dependent data. Dedecker and Prieur (2007) proved empirical process invariance principles for processes that satisfy one of the weak dependence conditions introduced earlier by the same authors, see Dedecker and Prieur (2005). These are generalizations of classical mixing coefficients, in a different way than Doukhan and Louhichi (1999); see also the recent book by Dedecker et al (2007). Wu (2008) studied empirical processes in the case that the underlying process can be represented as a functional of an i.i.d. process $(\epsilon_i)_{i \in \mathbb{Z}}$, i.e. $X_i = f((\epsilon_{i-k})_{k \geq 0})$. Wu and Shao (2004) investigated the empirical process for certain classes of Markov chains.

In recent years, a lot of research has been devoted to the study of statistical properties of data arising from dynamical systems. Given a measure preserving dynamical system $(\Omega, \mathcal{F}, T, P)$, consider the process $X_n := T(X_{n-1})$, $n \geq 1$. When T is a uniformly expanding map of the unit interval, this process can be represented as a functional of an absolutely regular process; see Hofbauer and Keller (1982). Denker and Keller (1986) used this representation in their investigations of the asymptotic behavior of U -statistics when the underlying data arise from a dynamical system. By combining this representation with coupling ideas, Borovkova, Burton and Dehling (2001) were able to study the empirical process and more generally, U -processes.

The spectral gap technique is a very powerful technique that allows to study much larger classes of dynamical systems. Let Q denote the Perron-Frobenius operator, defined by $\int f \circ T g dP = \int f Q g dP$, where $f \in L_\infty$, $g \in L_1$. The spectral gap technique studies the dynamical system via spectral properties of the Perron-Frobenius operator, viewed as operator on a suitable invariant subspace of L_1 . The spectral gap technique has been very successfully applied to the study of central limit theorems and large deviations properties; a survey and a large number of examples can be found in the monograph by Hennion and Hervé (2001). It is possible to treat empirical processes within the framework of the spectral gap technique if the Perron-Frobenius operator has a spectral gap on the space of functions of bounded variation, see Collet, Martinez and Schmitt (2004). In this case one can directly apply the standard proof of the empirical process invariance principle, i.e. establish finite dimensional convergence and tightness. This is essentially due to the fact that the indicator functions are functions of bounded variation. The situation is different when the spectral gap property can only be established on a smaller space of functions, such as Lipschitz functions. Recently, Gouëzel (2009) has given a one-dimensional example of such a dynamical system.

Dehling, Durieu and Volný (2009) have developed a new technique that is particularly useful when handling data from dynamical systems whose Perron-Frobenius operator has a spectral gap in the space of Lipschitz functions. Instead of trying to deduce inequalities for bounded variation functions from those on Lipschitz functions, the authors proposed a proof

which only involves Lipschitz functions. In principle, this technique can also be applied to other spaces of functions, not just Lipschitz functions. The technique can also be applied to the study of Markov processes whose Markov operator has a spectral gap in the space of Lipschitz functions. The new technique uses classical chaining ideas, but replaces the indicator functions that are commonly used by Lipschitz functions. Dehling, Durieu and Volný (2009) make two assumptions concerning the process $(X_i)_{i \geq 0}$. For any Lipschitz functions $f : \mathbb{R} \rightarrow \mathbb{R}$, the partial sums $\sum_{i=1}^n f(X_i)$ satisfy the central limit theorem and a suitable bound on the 4-th central moments. Under these two assumptions and a mild additional assumption on the modulus of continuity of the distribution function F , Dehling, Durieu and Volný (2009) could establish the empirical process invariance principle.

In the present paper, we extend the techniques of Dehling, Durieu and Volný (2009) to multidimensional systems satisfying a multiple mixing condition. In the multidimensional case, the 4th moment bounds have to be replaced by bounds on higher order moments. We establish such a bound for multiple mixing systems. The multiple mixing condition has been used in the study of the statistical properties of dynamical systems, e.g. in the work of Le Borgne (1999) and Durieu and Jouan (2008) on non-hyperbolic torus automorphisms.

Definition. Let $(\mathcal{B}, \|\cdot\|)$ be a Banach space of measurable functions $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}$. We say that the process $(X_n)_{n \geq 0}$ has a multiple mixing property with respect to \mathcal{B} if there exist constants $0 < \theta < 1$ and $r \geq 1$ such that for any $p \in \mathbb{N} \setminus \{0\}$, there exist a positive constant C and an integer ℓ such that the following assertions hold: for any $i_1, \dots, i_p \in \mathbb{N}$ and $q \in \{1, \dots, p\}$, for any φ in \mathcal{B} such that $\mathbb{E}_\nu(\varphi) = 0$ and $\|\varphi\|_\infty \leq 1$,

$$|\text{Cov}(\varphi(X_0)\varphi(X_{i_1^*}) \dots \varphi(X_{i_{q-1}^*}), \varphi(X_{i_q^*}) \dots \varphi(X_{i_p^*}))| \leq C \|\varphi(X_0)\|_r \|\varphi\|^\ell P(i_1, \dots, i_p) \theta^{i_q}, \quad (1)$$

where P is a polynomial function of p variables which does not depend on φ and where we use the following notation: if $(a_n)_{n \geq 1}$ is a sequence of real number, a_n^* is the sum $\sum_{i=1}^n a_i$.

In this paper, we will work mostly with the Banach space \mathcal{H}^α of bounded α -Hölder continuous functions, $0 < \alpha \leq 1$. We define the norm

$$\|g\| = \|g\|_\infty + \sup_{x \neq y \in \mathbb{R}^d} \frac{|g(x) - g(y)|}{|x - y|^\alpha}. \quad (2)$$

Theorem 1. Let $(X_i)_{i \geq 1}$ be an \mathbb{R}^d -valued stationary stochastic process satisfying the multiple mixing property with respect to the Banach space \mathcal{H}^α . Assume that for all $\phi \in \mathcal{H}^\alpha$, the partial sums $\sum_{i=1}^n \phi(X_i)$ satisfy the central limit theorem, i.e. that

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n (\phi(X_i) - E\phi(X_1)) \rightarrow N(0, \sigma^2)$$

where $\sigma^2 = \text{Var}(\phi(X_1)) + 2 \sum_{i=2}^\infty \text{Cov}(\phi(X_i), \phi(X_1))$. If the modulus of continuity ω of the distribution function F satisfies the condition

$$\omega(\delta) = O(|\log(\delta)|^{-\gamma}) \text{ for some } \gamma > r, \quad (3)$$

where r is given by (1), then the empirical process central limit theorem holds, i.e.

$$(U_n(t))_{t \in [-\infty, \infty]^d} \xrightarrow{\mathcal{D}} (W(t))_{t \in [-\infty, \infty]^d},$$

where “ $\xrightarrow{\mathcal{D}}$ ” denotes the weak convergence in the Skorohod space $D([-\infty, \infty]^d)$. Here $(W(t))_{t \in [-\infty, \infty]^d}$ is a mean zero Gaussian process with covariance structure

$$\begin{aligned} EW(s) \cdot W(t) &= \text{Cov}(1_{(-\infty, s]}(X_0), 1_{(-\infty, t]}(X_0)) \\ &+ \sum_{k=1}^{\infty} \text{Cov}(1_{(-\infty, s]}(X_0), 1_{(-\infty, t]}(X_k)) \\ &+ \sum_{k=1}^{\infty} \text{Cov}(1_{(-\infty, s]}(X_k), 1_{(-\infty, t]}(X_0)). \end{aligned}$$

Further, almost surely, $(W(t))_{t \in [-\infty, \infty]^d}$ has continuous sample paths.

Notice that the assumption on the modulus of continuity of F plays a key role in the proof of the theorem. It allows us to control the indicator functions of the process from the control on the Hölder observables. As mentioned by the referee of this paper, there are good reasons to believe that condition (3) is necessary if one only has assumptions on Hölder functions.

The proof of Theorem 1 will be given in two parts. In Section 2 we will establish a general empirical process CLT under the conditions of Theorem 1, but with the multiple mixing replaced by a $2p$ -th moment bound. In Section 3 we will show that multiple mixing implies this $2p$ -th moment bound.

As an application of Theorem 1 we can establish the empirical process invariance principle for ergodic torus automorphism. We consider the torus \mathbb{T}^d , identified to $[0, 1]^d$ and equipped with the Lebesgue measure, and define the automorphism $T : \mathbb{T}^d \rightarrow \mathbb{T}^d$ by

$$T(x) = Mx \pmod{1}.$$

Here M is a $d \times d$ matrix with integer entries and $|\det M| = 1$. We assume that the matrix M has no eigenvalue which is a root of unity which is equivalent to ergodicity. Such torus automorphisms always have at least one eigenvalue of modulus strictly bigger than 1 and then another one of modulus strictly smaller than 1. Thus a part of the action of the automorphism on the torus has some hyperbolicity. The automorphism is called hyperbolic if it has no eigenvalue of modulus one, and quasi-hyperbolic if it has eigenvalues of modulus one. For more details on torus automorphisms see Lind (1982). Le Borgne (1999) established the central limit theorem for quasi-hyperbolic torus automorphisms. Durieu and Jouan (2008) proved the empirical process central limit theorem for certain univariate functionals of quasi-hyperbolic torus automorphisms, i.e. they considered the empirical distribution of the sequence $(f(T^k x))_{k \geq 1}$, for $f : \mathbb{T} \rightarrow \mathbb{R}$. In this paper we can establish the full empirical process invariance principle in the same case.

Theorem 2. *Let T be an ergodic automorphism of the d -dimensional torus. Then the empirical process*

$$U_n(t) = \sqrt{n}(F_n(t) - \prod_{i=1}^d t_i), \quad t = (t_1, \dots, t_d) \in [0, 1]^d$$

converges in distribution to a centered Gaussian process $(W_t)_{t \in [0, 1]^d}$ which has almost surely continuous sample paths.

2 An invariance principle for the multivariate empirical process

Let $(X_n)_{n \geq 0}$ be a stationary process with values in \mathbb{R}^d , $d \geq 1$. For $t = (t_1, \dots, t_d) \in \mathbb{R}^d$ and $s = (s_1, \dots, s_d) \in \mathbb{R}^d$, we use the notations

$$\begin{aligned} t + s &= (t_1 + s_1, \dots, t_d + s_d), \\ s \leq t &\Leftrightarrow s_i \leq t_i \text{ for all } i = 1, \dots, d, \\ s < t &\Leftrightarrow s_i < t_i \text{ for all } i = 1, \dots, d \end{aligned}$$

and for $s < t$,

$$[s, t] = \prod_{i=1}^d [s_i, t_i].$$

As mentioned in the introduction, to generalize the result of Dehling et al. (2009) to multivariate processes, we need higher moment bounds. In this section, we will assume that the process $(X_n)_{n \geq 0}$ satisfies moment bounds on a given Banach space. The technique which is developed here is useful in cases where the Banach space does not contain the indicators functions. Our technique will work if the Banach space \mathcal{B} is enough well adapted to approximate indicator functions. Typically, we could work with space of regular functions as the spaces of Lipschitz continuous functions, Hölder continuous functions, or \mathcal{C}^k functions. Here, to have a link with the applications, we will work with the space \mathcal{H}^α of bounded α -Hölder functions, for some fixed $\alpha \in (0, 1]$. This space is equipped with the norm $\|\cdot\| = \|\cdot\|_{\mathcal{H}^\alpha}$ defined in (2)

Then, we make two assumptions concerning the process $(X_i)_{i \geq 0}$,

1. For any function $f \in \mathcal{H}^\alpha$, the CLT holds, i.e. $\sum_{i=0}^{\infty} \text{Cov}(f(X_0), f(X_i))$ converges and

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n (f(X_i) - E f(X_i)) \xrightarrow{\mathcal{D}} N(0, \sigma^2), \quad (4)$$

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

$$\sigma^2 = E(f(X_0) - E f(X_0))^2 + 2 \sum_{i=1}^{\infty} \text{Cov}(f(X_0), f(X_i)).$$

2. For any $p \geq 1$, a bound on the $2p$ central moments of partial sums of $(f(X_i))_{i \geq 0}$, $f \in \mathcal{H}^\alpha$ with $E(f(X_0)) = 0$ and $\|f\|_\infty \leq 1$, of the type

$$\mathbb{E}_\nu \left[\left(\sum_{i=1}^n f(X_i) \right)^{2p} \right] \leq K \sum_{i=1}^p n^i \|f(X_0)\|_r^i \log^{2p-i}(\|f\| + a) \quad (5)$$

where K is some universal constant, $a > 1$ and $r \geq 1$.

In Section 3, we will show how this condition is implied by the multiple mixing property.

Recall that the empirical distribution function $(F_n(t), t \in \mathbb{R}^d)$ and the empirical process $(U_n(t), t \in \mathbb{R}^d)$ are defined by

$$\begin{aligned} F_n(t) &= \frac{1}{n} \sum_{i=1}^n 1_{(-\infty, t]}(X_i), \quad t \in \mathbb{R}^d, \\ U_n(t) &= \sqrt{n}(F_n(t) - F(t)), \quad t \in \mathbb{R}^d, \end{aligned}$$

where F is the distribution function of X_0 and the modulus of continuity of a function $f : \mathbb{R}^d \rightarrow \mathbb{R}$ by

$$\omega_f(\delta) = \sup \left\{ |f(s) - f(t)| : s, t \in \mathbb{R}^d, |s - t| < \delta \right\}.$$

Theorem 3. *Let $(X_i)_{i \geq 0}$ be an \mathbb{R}^d -valued stationary random process such that the conditions (4) and (5) hold. Assume that X_0 has a distribution function F satisfying the following condition,*

$$\omega_F(\delta) = O(|\log(\delta)|^{-\gamma}) \text{ for some } \gamma > r, \quad (6)$$

where r is given by (5). Then

$$(U_n(t))_{t \in [-\infty, \infty]^d} \xrightarrow{\mathcal{D}} (W(t))_{t \in [-\infty, \infty]^d},$$

where $W(t)$ is a mean-zero Gaussian process with covariances

$$\begin{aligned} EW(s) \cdot W(t) &= \text{Cov}(1_{(-\infty, s]}(X_0), 1_{(-\infty, t]}(X_0)) \\ &\quad + \sum_{k=1}^{\infty} \text{Cov}(1_{(-\infty, s]}(X_0), 1_{(-\infty, t]}(X_k)) \\ &\quad + \sum_{k=1}^{\infty} \text{Cov}(1_{(-\infty, s]}(X_k), 1_{(-\infty, t]}(X_0)). \end{aligned}$$

Further, almost surely, $(W(t))_{t \in \mathbb{R}^d}$ has continuous sample paths.

Remark 1. *Here, if (5) holds for $r = 1$ as in Dehling et al.(2009), then in assumption (6) we can consider every $\gamma > 1$. This is an improvement of the corresponding theorem for dimension 1 of Dehling et al.(2009). This is a consequence of the fact that we consider $2p$ -th moment inequalities ($p \geq 1$) instead of only a 4-th moment bound.*

In this paper, we work with the Skorohod topology on the function space $D([-\infty, \infty]^d)$, as introduced by Neuhaus (1971) and Straf (1972). In fact, these authors considered the space $D([0, 1]^d)$, but we can easily extend their definitions since $[0, 1]^d$ and $[-\infty, \infty]^d$ are homeomorphic. Take any homeomorphism $\phi : [-\infty, \infty] \rightarrow [0, 1]$, e.g.

$$\phi(t) = \frac{1}{2} + \frac{1}{\pi} \arctan(t),$$

and define $\Phi : [-\infty, \infty]^d \rightarrow [0, 1]^d$ by

$$\Phi(t_1, \dots, t_d) = (\phi(t_1), \dots, \phi(t_d)).$$

The map Φ induces a map that associates to any function $f \in D([0, 1]^d)$ the function $f \circ \Phi : [-\infty, \infty]^d \rightarrow \mathbb{R}$. We define $D([-\infty, \infty]^d)$ as the image of $D([0, 1]^d)$ under this map. Neuhaus (1971) and Straf (1972) introduced a metric d_0 on $D([0, 1]^d)$ that generates the Skorohod topology and such that $(D([0, 1]^d), d_0)$ is a complete separable metric space. We can naturally extend d_0 to $D([-\infty, \infty]^d)$ by defining for $g_1, g_2 \in D([-\infty, \infty]^d)$

$$\tilde{d}_0(g_1, g_2) = d_0(g_1 \circ \Phi^{-1}, g_2 \circ \Phi^{-1}).$$

In what follows, we will denote the metric on $D([-\infty, \infty]^d)$ also by d_0 . Note that d_0 is bounded by the supremum distance, i.e.

$$d_0(g_1, g_2) \leq \sup_{t \in [-\infty, \infty]^d} |g_1(t) - g_2(t)|,$$

and that $(D([-\infty, \infty]^d), d_0)$ is a complete separable metric space. Note that the sample paths of the processes arising in this paper are elements of the function space $D([-\infty, \infty]^d)$, since their limits as any of the arguments approach ∞ exist. Alternatively one can see this by observing that e.g. the empirical distribution function F_n of the process $(X_i)_{i \geq 1}$ is the image under the map defined above of the empirical distribution function of the process $(\Phi(X_i))_{i \geq 1}$.

Proof of Theorem 3.

To prove Theorem 3, we shall adapt the technique introduced by Dehling et al. (2009). The idea is to replace the indicator functions $1_{(-\infty, t]}(x)$ by approximations in the space \mathcal{H}^α .

For each $i = 1, \dots, d$, we denote by F_i the marginal distribution functions of X_0 corresponding to the i -th coordinate. Note that the F_i also verify condition (6).

Given a partition of $[0, 1]$,

$$0 = r_0 < \dots < r_m = 1$$

we define

$$t_{i,j} = F_i^{-1}(r_j)$$

where F_i^{-1} is given by

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

Thus, by continuity of the F_i , we have subdivisions

$$-\infty \leq t_{i,0} < \dots < t_{i,m} = +\infty.$$

If $j = (j_1, \dots, j_d) \in \{0, \dots, m\}^d$, we set

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

We introduce the functions $\varphi_j : \mathbb{R}^d \rightarrow \mathbb{R}$, $j = (j_1, \dots, j_d) \in \{1, \dots, m\}^d$ defined by

$$\varphi_j(x) = \begin{cases} \prod_{i=1}^d \varphi\left(\frac{x_i - t_{i,j_i-1}}{|t_{i,j_i-1} - t_{i,j_i-2}|}\right) & \text{if } (2, \dots, 2) \leq j \\ 0 & \text{otherwise} \end{cases}$$

with

$$\varphi(x) = 1_{(-\infty, -1]}(x) - x 1_{(-1, 0]}(x) \tag{7}$$

and where we eventually used the convention that $\frac{1}{\infty} = 0$. Note that φ is a α -Hölder function on \mathbb{R} (for all $\alpha \in (0, 1]$). The function φ_j will serve as a \mathcal{H}^α -approximation to the indicator function $1_{(-\infty, t_{j-1}]}$.

Now, we introduce the process

$$\begin{aligned} F_n^{(m)}(t) &= \frac{1}{n} \sum_{i=1}^n \sum_{j \in \{1, \dots, m\}^d} 1_{[t_{j-1}, t_j)}(t) \varphi_j(X_i) \\ &= \sum_{j \in \{1, \dots, m\}^d} \left(\frac{1}{n} \sum_{i=1}^n \varphi_j(X_i) \right) 1_{[t_{j-1}, t_j)}(t). \end{aligned}$$

Note that $F_n^{(m)}(t)$ is a piecewise constant approximation to the empirical distribution function $F_n(t)$. For $t \in [t_{j-1}, t_j)$, we have the inequality

$$F_n(t_{j-2}) \leq F_n^{(m)}(t) \leq F_n(t_{j-1}).$$

We define further

$$F^{(m)}(t) = E \left(F_n^{(m)}(t) \right) = \sum_{j=1}^m E \left(\varphi_j(X_0) \right) 1_{[t_{j-1}, t_j)}(t),$$

and finally the centered and normalized process

$$U_n^{(m)}(t) = \sqrt{n} \left(F_n^{(m)}(t) - F^{(m)}(t) \right), \quad t \in \mathbb{R}^d. \quad (8)$$

Theorem 3 will follow by application of the following Theorem which is proved in Dehling et al. (2009).

Theorem. *Let (S, ρ) be a complete separable metric space and let $X_n, X_n^{(m)}$ and $X^{(m)}$, $n, m \geq 1$ be S -valued random variables satisfying*

$$X_n^{(m)} \xrightarrow{\mathcal{D}} X^{(m)} \text{ as } n \rightarrow \infty, \forall m \quad (9)$$

$$\lim_{m \rightarrow \infty} \limsup_{n \rightarrow \infty} P(\rho(X_n, X_n^{(m)}) \geq \varepsilon) = 0, \forall \varepsilon > 0. \quad (10)$$

Then there exists an S -valued random variable X such that

$$X_n \xrightarrow{\mathcal{D}} X \text{ as } n \rightarrow \infty.$$

Here we work in the complete separable metric space $(D([-\infty, \infty]^d), d_0)$. We shall prove separately that (9) and (10) hold for the $D([-\infty, \infty]^d)$ -valued random variables $U_n^{(m)}$ in Proposition 1 and Proposition 2.

Proposition 1. *For any partition $0 = r_0 < \dots < r_m = 1$, there exists a piecewise constant Gaussian process $(W^{(m)}(t))_{t \in \mathbb{R}^d}$ such that*

$$\left(U_n^{(m)}(t) \right)_{t \in \mathbb{R}^d} \xrightarrow{\mathcal{D}} \left(W^{(m)}(t) \right)_{t \in \mathbb{R}^d}.$$

The sample paths of the processes $(W^{(m)}(t))_{t \in \mathbb{R}^d}$ are constant on each of the rectangles $[t_{j-1}, t_j)$, $1 \leq j \leq m$, and $W^{(m)}(0) = 0$. The vector $(W^{(m)}(t_1), \dots, W^{(m)}(t_m))$ has a multivariate normal distribution with mean zero and covariances

$$\begin{aligned} \text{Cov}(W^{(m)}(t_{i-1}), W^{(m)}(t_{j-1})) &= \text{Cov}(\varphi_i(X_0), \varphi_j(X_0)) \\ &+ \sum_{k=1}^{\infty} \text{Cov}(\varphi_i(X_0), \varphi_j(X_k)) \\ &+ \sum_{k=1}^{\infty} \text{Cov}(\varphi_i(X_k), \varphi_j(X_0)) \end{aligned}$$

Proof.

Use (4) and the Cramér-Wold device. \square

Proposition 2. For any $\varepsilon, \eta > 0$ there exists a partition $0 = r_0 < \dots < r_m = 1$ such that

$$\limsup_{n \rightarrow \infty} P \left(\sup_{t \in \mathbb{R}^d} |U_n(t) - U_n^{(m)}(t)| > \varepsilon \right) \leq \eta.$$

Proof.

From here, we assume the partition $0 = r_0 < \dots < r_m = 1$ is a regular partition of step $h = m^{-1}$. Let $j = 1, \dots, m$. On the interval $[r_{j-1}, r_j]$ we introduce a sequence of refining partitions

$$r_{j-1} = s_{j,0}^{(k)} < s_{j,1}^{(k)} < \dots < s_{j,2^k}^{(k)} = r_j$$

by

$$s_{j,l}^{(k)} = r_{j-1} + l \cdot \frac{h}{2^k}, \quad 0 \leq l \leq 2^k.$$

For each $i \in \{1, \dots, d\}$, let us define

$$s_{i,j,l}^{(k)} = F_i^{-1}(s_{j,l}^{(k)}), \quad 0 \leq l \leq 2^k.$$

We now have partitions of $[t_{i,j-1}, t_{i,j}]$,

$$t_{i,j-1} = s_{i,j,0}^{(k)} < s_{i,j,1}^{(k)} < \dots < s_{i,j,2^k}^{(k)} = t_{i,j}.$$

For convenience, for $j > 1$, we also consider the points

$$s_{i,j,-1}^{(k)} = F_i^{-1} \left(r_{j-1} - \frac{h}{2^k} \right) = s_{i,j-1,2^k-1}^{(k)}$$

and for $j < m$, we consider the points

$$s_{i,j,2^k+1}^{(k)} = F_i^{-1} \left(r_{j-1} + (2^k + 1) \frac{h}{2^k} \right) = s_{i,j+1,1}^{(k)}.$$

For any $t \in [t_{i,j-1}, t_{i,j})$ and $k \geq 0$ we define the index

$$l_{i,j}(k, t) = \max \left\{ l : s_{i,j,l}^{(k)} \leq t \right\}.$$

Now fix $j \in \{1, \dots, m\}^d$ (then the index related to j will be forgotten). For $l = (l_1, \dots, l_d)$ and $t \in [t_{j-1}, t_j)$, we write

$$s_l^{(k)} = (s_{1,j_1,l_1}^{(k)}, \dots, s_{d,j_d,l_d}^{(k)})$$

and

$$l(k, t) = (l_{1,j_1}(k, t_1), \dots, l_{d,j_d}(k, t_d)).$$

In this way we obtain a chain,

$$t_{j-1} = s_{l(0,t)}^{(0)} \leq s_{l(1,t)}^{(1)} \leq \dots \leq s_{l(k,t)}^{(k)} \leq t \leq s_{l(k,t)+1}^{(k)},$$

linking the point t_{j-1} to t . Note that for $t \in [t_{j-1}, t_j)$ we have by definition $U_n^{(m)}(t) = U_n^{(m)}(t_{j-1})$.

We define the functions $\psi_l^{(k)}$, $k \geq 0$, $l \in \{0, \dots, 2^k + 1\}^d$, in the following way : We first define, for $i = 1, \dots, d$ and $l \in \{0, \dots, 2^k + 1\}$,

$$\psi_{i,l}^{(k)}(x_i) = \begin{cases} 0 & \text{if } j_i = 1 \text{ and } l = 0 \\ 1 & \text{if } j_i = m \text{ and } l \geq 2^k \\ \varphi \left(\frac{x_i - s_{i,j_i,l}^{(k)}}{|s_{i,j_i,l}^{(k)} - s_{i,j_i,l-1}^{(k)}|} \right) & \text{otherwise} \end{cases}$$

where φ is defined as in (7) (eventually, we use $\frac{1}{\infty} = 0$). Then we set, for $l = (l_1, \dots, l_d) \in \{0, \dots, 2^k + 1\}^d$,

$$\psi_l^{(k)}(x) = \prod_{i=1}^d \psi_{i,l_i}^{(k)}(x).$$

Observe that by definition of $s_{l(k,t)}^{(k)}$ and of $\psi_l^{(k)}$, $\psi_{l(0,t)}^{(0)}(x) = \varphi_j(x)$ and

$$\varphi_j(x) \leq \psi_{l(1,t)}^{(1)}(x) \leq \dots \leq \psi_{l(k,t)}^{(k)}(x) \leq 1_{(-\infty, t]}(x) \leq \psi_{l(k,t)+2}^{(k)}(x).$$

In this way we get

$$\begin{aligned} F_n(t) - F_n^{(m)}(t) &= \sum_{k=1}^K \frac{1}{n} \sum_{i=1}^n \left(\psi_{l(k,t)}^{(k)}(X_i) - \psi_{l(k-1,t)}^{(k-1)}(X_i) \right) \\ &\quad + \frac{1}{n} \sum_{i=1}^n \left(1_{(-\infty, t]}(X_i) - \psi_{l(K,t)}^{(K)}(X_i) \right) \end{aligned} \quad (11)$$

where K is some integer to be chosen later.

From (11) we get by centering and normalization

$$\begin{aligned} U_n(t) - U_n^{(m)}(t) &= \sum_{k=1}^K \frac{1}{\sqrt{n}} \sum_{i=1}^n \left\{ \left(\psi_{l(k,t)}^{(k)}(X_i) - E\psi_{l(k,t)}^{(k)}(X_i) \right) \right. \\ &\quad \left. - \left(\psi_{l(k-1,t)}^{(k-1)}(X_i) - E\psi_{l(k-1,t)}^{(k-1)}(X_i) \right) \right\} \\ &\quad + \frac{1}{\sqrt{n}} \sum_{i=1}^n \left\{ \left(1_{(-\infty, t]}(X_i) - F(t) \right) \right. \\ &\quad \left. - \left(\psi_{l(K,t)}^{(K)}(X_i) - E\psi_{l(K,t)}^{(K)}(X_i) \right) \right\}. \end{aligned}$$

For the last term on the r.h.s. we have the following upper and lower bounds,

$$\begin{aligned} & \frac{1}{\sqrt{n}} \sum_{i=1}^n \left\{ (1_{(-\infty, t]}(X_i) - F(t)) - \left(\psi_{l(K, t)}^{(K)}(X_i) - E\psi_{l(K, t)}^{(K)}(X_i) \right) \right\} \\ & \leq \frac{1}{\sqrt{n}} \sum_{i=1}^n \left\{ \left(\psi_{l(K, t)+2}^{(K)}(X_i) - E\psi_{l(K, t)+2}^{(K)}(X_i) \right) - \left(\psi_{l(K, t)}^{(K)}(X_i) - E\psi_{l(K, t)}^{(K)}(X_i) \right) \right\} \\ & \quad + \sqrt{n} \left(E\psi_{l(K, t)+2}^{(K)}(X_i) - F(t) \right) \end{aligned}$$

and

$$\begin{aligned} & \frac{1}{\sqrt{n}} \sum_{i=1}^n \left\{ (1_{(-\infty, t]}(X_i) - F(t)) - \left(\psi_{l(K, t)}^{(K)}(X_i) - E\psi_{l(K, t)}^{(K)}(X_i) \right) \right\} \\ & \geq -\sqrt{n} \left(F(t) - E\psi_{l(K, t)}^{(K)}(X_i) \right). \end{aligned}$$

Now choose $K = 4 + \left\lceil \log \left(d \frac{\sqrt{nh}}{\varepsilon} \right) \log^{-1}(2) \right\rceil$ and note that

$$\frac{\varepsilon}{2^4} \leq d\sqrt{n} \frac{h}{2^K} \leq \frac{\varepsilon}{2^3}.$$

We thus have

$$\begin{aligned} \sqrt{n} \left| E\psi_{l(K, t)+2}^{(K)}(X_i) - E\psi_{l(K, t)}^{(K)}(X_i) \right| & \leq \sqrt{n} \sum_{i=1}^d \left| F_i(s_{l_i, j_i}^{(K)}(K, t)+2) - F_i(s_{l_i, j_i}^{(K)}(K, t)-1) \right| \\ & \leq \frac{\varepsilon}{2}. \end{aligned}$$

Therefore, since

$$\psi_{l(K, t)}^{(K)}(X_i) \leq 1_{(-\infty, t]}(X_i) \leq \psi_{l(K, t)+2}^{(K)}(X_i),$$

we get for all $t \in [t_{j-1}, t_j]$,

$$\begin{aligned} \left| U_n(t) - U_n^{(m)}(t) \right| & \leq \sum_{k=1}^K \frac{1}{\sqrt{n}} \left| \sum_{i=1}^n \left\{ \left(\psi_{l(k, t)}^{(k)}(X_i) - E\psi_{l(k, t)}^{(k)}(X_i) \right) \right. \right. \\ & \quad \left. \left. - \left(\psi_{l(k-1, t)}^{(k-1)}(X_i) - E\psi_{l(k-1, t)}^{(k-1)}(X_i) \right) \right\} \right| \\ & \quad + \frac{1}{\sqrt{n}} \left| \sum_{i=1}^n \left\{ \left(\psi_{l(K, t)+2}^{(K)}(X_i) - E\psi_{l(K, t)+2}^{(K)}(X_i) \right) \right. \right. \\ & \quad \left. \left. - \left(\psi_{l(K, t)}^{(K)}(X_i) - E\psi_{l(K, t)}^{(K)}(X_i) \right) \right\} \right| \\ & \quad + \frac{\varepsilon}{2}. \end{aligned}$$

Note that by definition of $l(k, t)$ and of $s_l^{(k)}$, we have

$$l(k-1, t) = \left\lfloor \frac{l(k, t)}{2} \right\rfloor$$

where the integer part $\lfloor \cdot \rfloor$ is taken on each coordinate. We infer

$$\begin{aligned}
\sup_{t_{j-1} \leq t \leq t_j} \left| U_n(t) - U_n^{(m)}(t) \right| &\leq \sum_{k=1}^K \frac{1}{\sqrt{n}} \max_{l \in \{0, \dots, 2^k\}^d} \left| \sum_{i=1}^n \left((\psi_l^{(k)}(X_i) - E\psi_l^{(k)}(X_i)) \right. \right. \\
&\quad \left. \left. - (\psi_{\lfloor \frac{l}{2} \rfloor}^{(k-1)}(X_i) - E\psi_{\lfloor \frac{l}{2} \rfloor}^{(k-1)}(X_i)) \right) \right| \\
&\quad + \frac{1}{\sqrt{n}} \max_{l \in \{0, \dots, 2^K\}^d} \left| \sum_{i=1}^n \left((\psi_{l+2}^{(K)}(X_i) - E\psi_{l+2}^{(K)}(X_i)) \right. \right. \\
&\quad \left. \left. - (\psi_l^{(K)}(X_i) - E\psi_l^{(K)}(X_i)) \right) \right| \\
&\quad + \frac{\varepsilon}{2}.
\end{aligned}$$

Now, taking $\varepsilon_k = \frac{\varepsilon}{4k(k+1)}$, we obtain

$$\begin{aligned}
&P \left(\sup_{t_{j-1} \leq t \leq t_j} \left| U_n(t) - U_n^{(m)}(t) \right| \geq \varepsilon \right) \\
&\leq \sum_{k=1}^K \sum_{l \in \{0, \dots, 2^k\}^d} P \left(\frac{1}{\sqrt{n}} \left| \sum_{i=1}^n \left\{ (\psi_l^{(k)}(X_i) - E\psi_l^{(k)}(X_i)) \right. \right. \right. \\
&\quad \left. \left. - (\psi_{\lfloor \frac{l}{2} \rfloor}^{(k-1)}(X_i) - E\psi_{\lfloor \frac{l}{2} \rfloor}^{(k-1)}(X_i)) \right\} \right| \geq \varepsilon_k \right) \\
&\quad + \sum_{l \in \{0, \dots, 2^K\}^d} P \left(\frac{1}{\sqrt{n}} \left| \sum_{i=1}^n \left\{ (\psi_{l+2}^{(K)}(X_i) - E\psi_{l+2}^{(K)}(X_i)) \right. \right. \right. \\
&\quad \left. \left. - (\psi_l^{(K)}(X_i) - E\psi_l^{(K)}(X_i)) \right\} \right| \geq \frac{\varepsilon}{4} \right).
\end{aligned}$$

At this point we shall use Markov's inequality at the order $2p$ together with the $2p$ -th moment bound (5) for an integer p such that

$$p > d \frac{r\gamma}{\gamma - r}. \quad (12)$$

First, remark that these following bounds also hold. Since $\psi_{i, l_i}^{(k)} - \psi_{i, \lfloor \frac{l_i}{2} \rfloor}^{(k-1)}$ vanishes outside $[s_{i, j_i, \lfloor \frac{l_i}{2} \rfloor - 1}^{(k-1)}, s_{i, j_i, l_i}^{(k)}]$, we have

$$\begin{aligned}
\left\| \psi_l^{(k)}(X_0) - \psi_{\lfloor \frac{l}{2} \rfloor}^{(k-1)}(X_0) \right\|_r &\leq \left(\sum_{i=1}^d \left| F_i(s_{i, j_i, l_i}^{(k)}) - F_i(s_{i, j_i, \lfloor \frac{l_i}{2} \rfloor - 1}^{(k-1)}) \right| \right)^{\frac{1}{r}} \\
&\leq d \max_{i=1}^d \left| F_i(s_{i, j_i, l_i}^{(k)}) - F_i(s_{i, j_i, l_i - 3}^{(k)}) \right|^{\frac{1}{r}} \\
&= \left(\frac{3dh}{2^k} \right)^{\frac{1}{r}}
\end{aligned}$$

and in the same way

$$\left\| \psi_{l+2}^{(K)}(X_0) - \psi_l^{(K)}(X_0) \right\|_r \leq \left(\frac{3dh}{2^K} \right)^{\frac{1}{r}}.$$

Now by (6), if k is big enough, we have

$$\begin{aligned} \left\| \psi_l^{(k)} \right\| &\leq 1 + d \max_{i=1}^d \frac{1}{|s_{i,j_i,l_i}^{(k)} - s_{i,j_i,l_i-1}^{(k)}|^\alpha} \\ &\leq 1 + d \left[\inf \left\{ s > 0 : \exists i \in \{1, \dots, d\}, \forall t, F_i(t+s) - F_i(t) \geq \frac{h}{2^k} \right\} \right]^{-\alpha} \\ &\leq 1 + d \left[\inf \left\{ s > 0 : D |\log(s)|^{-\gamma} \geq \frac{h}{2^k} \right\} \right]^{-\alpha} \\ &= 1 + d \exp \left(\alpha \left(\frac{D2^k}{h} \right)^{\frac{1}{\gamma}} \right). \end{aligned}$$

Thus, there is a positive constant B such that for arbitrary $k \geq 0$,

$$\left\| \psi_l^{(k)} \right\| \leq B \exp \left(\alpha \left(\frac{D2^k}{h} \right)^{\frac{1}{\gamma}} \right).$$

Therefore, applying successively Markov's inequality at the order $2p$, the $2p$ -th moment bound (5) and the preceding inequalities, we get

$$\begin{aligned} &P \left(\sup_{t_{j-1} \leq t \leq t_j} |U_n(t) - U_n(t_j)| \geq \varepsilon \right) \\ &\leq C \sum_{i=1}^p \sum_{k=1}^K 2^{dk} \frac{(k(k+1))^{2p}}{\varepsilon^{2p}} \frac{1}{n^p} n^i \left(\frac{h^i}{2^{ik}} \right)^{\frac{1}{r}} \log^{2p-i} \left(a + B \exp \left(\alpha \left(\frac{D2^k}{h} \right)^{\frac{1}{\gamma}} \right) \right) \\ &\leq C \sum_{i=1}^{p-1} \frac{h^d}{\varepsilon^{2p} n^{p-i}} \sum_{k=1}^K \left(\frac{2^k}{h} \right)^{d-\frac{i}{r}} k^{4p} \left(\frac{2^k}{h} \right)^{\frac{2p-i}{\gamma}} \\ &\quad + C \frac{h^d}{\varepsilon^{2p}} \sum_{k=1}^K \left(\frac{2^k}{h} \right)^{d-\frac{p}{r}} k^{4p} \left(\frac{2^k}{h} \right)^{\frac{p}{\gamma}} \\ &\leq C \sum_{i=1}^{p-1} \frac{h^d}{\varepsilon^{2p+d-i+\frac{2p-i}{\gamma}}} (\sqrt{n})^{d-\frac{i}{r}+\frac{2p-i}{\gamma}-2(p-i)} K^{4p+1} \\ &\quad + C \frac{h^{\frac{p}{r}-\frac{p}{\gamma}}}{\varepsilon^{2p}} \sum_{k=1}^{\infty} 2^{(d-\frac{p}{r}+\frac{p}{\gamma})k} k^{4p}, \end{aligned}$$

where C always denotes a positive constant, but its value changes from line to line.

By Condition (12), the series $\sum_{k=1}^{\infty} 2^{(d-\frac{p}{r}+\frac{p}{\gamma})k} k^{4p}$ converges and there exists an $A > 0$ such that for all $i \in \{1, \dots, p-1\}$,

$$d - \frac{i}{r} + \frac{2p-i}{\gamma} - 2(p-i) < -A.$$

Finally, using $mh = 1$, we have

$$\begin{aligned}
& P \left(\sup_{0 \leq t \leq 1} \left| U_n(t) - U_n^{(m)}(t) \right| \geq \varepsilon \right) \\
& \leq \sum_{j \in \{1, \dots, m\}^d} P \left(\sup_{t_{j-1} \leq t \leq t_j} \left| U_n(t) - U_n^{(m)}(t) \right| \geq \varepsilon \right) \\
& \leq C(p-2) \frac{1}{\varepsilon^{4p+d}} (\sqrt{n})^{-A} \left(4 + \log \left(d \frac{\sqrt{nh}}{\varepsilon} \right) \right)^{4p+1} + C \frac{h^{\frac{p}{r} - d - \frac{p}{\gamma}}}{\varepsilon^{2p}}.
\end{aligned}$$

The first summand converges to zero as $n \rightarrow \infty$ and, since $\frac{p}{r} - d - \frac{p}{\gamma} > 0$, the second can be made arbitrarily small by choosing a partition that is fine enough (i.e. h small). \square

Remark 2. *The tightness of the empirical process can be proved using exactly the same proof than the one of Proposition 2.3 in Dehling et al. (2009). The almost sure continuity of the limit process follows.*

Remark 3. *Assume that conditions (4) and (5) hold for a space of C^k -functions instead of a space of Hölder continuous functions. It is clear that the same technique works (in the proof, take for example $\varphi(x) = 1_{(-\infty, -1]}(x) - \sin\left(\pi x + \frac{\pi}{2}\right) 1_{(-1, 0]}(x)$).*

3 Moment bounds for partial sums

For a function $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}$, we consider the partial sum

$$S_n(\varphi) = \sum_{i=0}^{n-1} \varphi(X_i).$$

Multiple mixing properties allow us to obtain some useful moment inequalities. In this section, we show how the multiple mixing property implies the $2p$ -th moment bound which is required in Theorem 3.

Theorem 4. *Let $(X_n)_{n \geq 0}$ be a stationary process having a multiple mixing property on \mathcal{B} and $\varphi \in \mathcal{B}$ such that $\mathbb{E}_\nu(\varphi) = 0$ and $\sup_x |\varphi(x)| \leq 1$. Then for all $p \geq 1$,*

$$\mathbb{E}_\nu \left[S_n(\varphi)^{2p} \right] \leq K \sum_{i=1}^p n^i \|\varphi(X_0)\|_r^i \log^{2p-i}(\|\varphi\| + \theta^{-1})$$

and

$$\left| \mathbb{E}_\nu \left[S_n(\varphi)^{2p+1} \right] \right| \leq K \sum_{i=1}^p n^i \|\varphi(X_0)\|_r^i \log^{2p-i+1}(\|\varphi\| + \theta^{-1})$$

where K is a constant which does not depend on n or φ and $r \geq 1$ is given by (1).

Proof of Theorem 4. Let us consider the assumptions of Theorem 4 hold and let $\varphi \in \mathcal{B}$ with $\mathbb{E}(\varphi) = 0$ and $\sup_x |\varphi(x)| \leq 1$ be fixed. We use the notation $a_n^* = \sum_{i=1}^n a_i$.

Notations: For all $p \geq 1$, we define

$$I_n(p) = \sum_{\substack{0 \leq i_1, \dots, i_p \leq n-1 \\ i_p^* \leq n-1}} |\mathbb{E}(\varphi(X_0)\varphi(X_{i_1^*}) \dots \varphi(X_{i_p^*}))|$$

and $I_n(0) = |\mathbb{E}(\varphi(X_0))| = 0$.

As the process is stationary, for $p \geq 1$, we have

$$|\mathbb{E}[S_n(\varphi)^p]| \leq p!nI_n(p-1).$$

So, to prove Theorem 4, it is sufficient to prove the following lemma.

Lemma 1. For all $p \geq 1$,

$$I_n(2p-1) \leq K \sum_{i=1}^p n^{i-1} \|\varphi(X_0)\|_r^i \log^{2p-i}(\|\varphi\| + \theta^{-1})$$

and

$$I_n(2p) \leq K \sum_{i=1}^p n^{i-1} \|\varphi(X_0)\|_r^i \log^{2p-i+1}(\|\varphi\| + \theta^{-1})$$

where K is a constant which does not depend on n or φ .

Notations: For all $p \geq 1$ and $q \in \{1, \dots, p\}$, we define

$$J_n(p, q) = \sum_{i_q=0}^{n-1} \sum_{\substack{0 \leq i_1, \dots, i_{q-1}, i_{q+1}, \dots, i_p \leq i_q \\ i_p^* \leq n-1}} |\mathbb{E}(\varphi(X_0)\varphi(X_{i_1^*}) \dots \varphi(X_{i_p^*}))|.$$

We have

$$I_n(p) \leq \sum_{q=1}^p J_n(p, q).$$

To prove Lemma 1, we will use the following lemma.

Lemma 2. For all $p \in \mathbb{N}^*$ and $q \in \{1, \dots, p\}$,

$$J_n(p, q) \leq C \|\varphi(X_0)\|_r \log^p(\|\varphi\| + \theta^{-1}) + nI_n(q-1)I_n(p-q),$$

where C is a constant which does not depend on n or φ .

Proof of Lemma 2.

Let n_0 be a positive integer such that

$$\frac{\log(\|\varphi\| + \theta^{-1})}{-\log \theta} < n_0 \leq \frac{\log(\|\varphi\| + \theta^{-1})}{-\log \theta} + 1.$$

We thus have the inequality $\theta^{n_0} \|\varphi\| \leq 1$ and $n_0 \geq 2$.

We have

$$J_n(p, q) \leq \sum_{i_q=0}^{n-1} \sum_{\substack{0 \leq i_1, \dots, i_{q-1}, i_{q+1}, \dots, i_p \leq i_q \\ i_p^* \leq n-1}} [A_{i_1, \dots, i_p} + B_{i_1, \dots, i_p}]$$

where

$$A_{i_1, \dots, i_p} = |\text{Cov}(\varphi(X_0)\varphi(X_{i_1^*}) \dots \varphi(X_{i_{q-1}^*}), \varphi(X_{i_q^*})\varphi(X_{i_{q+1}^*}) \dots \varphi(X_{i_p^*}))|$$

and

$$B_{i_1, \dots, i_p} = |\mathbb{E}(\varphi(X_0)\varphi(X_{i_1^*}) \dots \varphi(X_{i_{q-1}^*}))| |\mathbb{E}(\varphi(X_0)\varphi(X_{i_{q+1}^*}) \dots \varphi(X_{i_p^*}))|.$$

Using Hölder inequality for $i_q = 0$ to $n_0\ell - 2$ and multiple mixing property (1) for $i_q \geq n_0\ell$ (where ℓ comes from (1)), we obtain

$$\begin{aligned} & \sum_{i_q=0}^{n-1} \sum_{\substack{0 \leq i_1, \dots, i_{q-1}, i_{q+1}, \dots, i_p \leq i_q \\ i_p^* \leq n-1}} A_{i_1, \dots, i_p} \\ & \leq C \sum_{i_q=0}^{n_0\ell-2} (i_q + 1)^{p-1} \|\varphi(X_0)^{p+1}\|_1 + C \sum_{i_q=n_0\ell-1}^n (i_q + 1)^{p-1} \kappa \|\varphi(X_0)\|_r \theta^{i_q} \|\varphi\|^\ell Q(i_q) \\ & \leq C(n_0 - 1)^p \|\varphi(X_0)\|_r + C\kappa \|\varphi(X_0)\|_r \sum_{i_q=n_0\ell-1}^{\infty} (i_q + 1)^{p-1} \theta^{i_q - n_0\ell} Q(i_q) \end{aligned}$$

where $Q(i_q) = \sum_{0 \leq i_1, \dots, i_{q-1}, i_{q+1}, \dots, i_p \leq i_q} P(i_1, \dots, i_p)$ is a polynomial function of i_q . Then $\sum_{i_q=0}^{\infty} \theta^{i_q} Q(i_q)$ converges and we deduce that

$$\sum_{i_q=n_0\ell-1}^{\infty} (i_q + 1)^{p-1} \theta^{i_q - n_0\ell} Q_n(i_q) \leq C(n_0 - 1)^{p-1}$$

where C is independent of n_0 . Thus, since $n_0 - 1 \leq C \log(\|\varphi\| + \theta^{-1})$,

$$\sum_{i_q=0}^{n-1} \sum_{\substack{0 \leq i_1, \dots, i_{q-1}, i_{q+1}, \dots, i_p \leq i_q \\ i_p^* \leq n-1}} A_{i_1, \dots, i_p} \leq C\kappa \|\varphi(X_0)\|_r \log^p(\|\varphi\| + \theta^{-1}).$$

On the other hand,

$$\sum_{\substack{0 \leq i_1, \dots, i_{q-1}, i_{q+1}, \dots, i_p \leq i_q \\ i_p^* \leq n-1}} B_{i_1, \dots, i_p} \leq I_n(q-1) I_n(p-q).$$

Therefore,

$$J_n(p, q) \leq C\kappa \|\varphi(X_0)\|_r \log^p(\|\varphi\| + \theta^{-1}) + n I_n(q-1) I_n(p-q).$$

□

Proof of Lemma 1.

We proceed by induction. We have, $I_n(1) = J_n(1, 1)$. Then, by Lemma 2,

$$I_n(1) \leq C \|\varphi(X_0)\|_r \log(\|\varphi\| + \theta^{-1}).$$

In the same way, $I_n(2) \leq J_n(2, 1) + J_n(2, 2)$. Then, by Lemma 2,

$$I_n(2) \leq C \|\varphi(X_0)\|_r \log^2(\|\varphi\| + \theta^{-1}).$$

In the general case, by Lemma 2,

$$\begin{aligned} I_n(p) &\leq \sum_{q=1}^p J_n(p, q) \\ &\leq Cp \|\varphi(X_0)\|_r \log^p(\|\varphi\| + \theta^{-1}) + n \sum_{q=2}^{p-1} I_n(q-1) I_n(p-q). \end{aligned}$$

Studying $\sum_{q=2}^{p-1} I_n(q-1) I_n(p-q)$ according to the parity of p , we deduce the inequalities of Lemma 1. □

4 Markov chains and dynamical systems with a spectral gap

Let $(X_n)_{n \geq 0}$ be a homogeneous Markov chain with a stationary measure ν . Denote by P the associated Markov operator and E the state space. Consider a Banach algebra $(\mathcal{B}, \|\cdot\|)$ of ν -measurable functions from E to \mathbb{R} , which contains the function $\mathbf{1} = \mathbf{1}_E$ and which is continuously included in $(\mathcal{L}^s(\nu), \|\cdot\|_s)$ for some $s \in [1, +\infty]$,

i.e. $\exists C > 0$ such that $\forall f \in \mathcal{B}$,

$$\|f\|_s \leq C \|f\|. \quad (13)$$

We say that the Markov chain $(X_n)_{n \geq 0}$ is \mathcal{B} -geometrically ergodic or strongly ergodic (with respect to \mathcal{B}) if there exist $\kappa > 0$ and $0 < \theta < 1$ such that for all $f \in \mathcal{B}$,

$$\|P^n f - \Pi f\| \leq \kappa \theta^n \|f\| \quad (14)$$

where $\Pi f = \mathbb{E}_\nu(f) \mathbf{1}$.

Strong ergodicity corresponds to the fact that the Markov transition operator acting on \mathcal{B} has 1 as simple eigenvalue and the rest of the spectrum is included in a closed ball of radius strictly smaller than 1 (see Hennion & Hervé 2001).

Lemma 3. *Let $(X_n)_{n \geq 0}$ be a \mathcal{B} -geometrically ergodic Markov chain, then it satisfies the multiple mixing property (1) on \mathcal{B} with $r = \frac{s}{s-1}$ ($r = 1$ if $s = \infty$).*

Proof of Lemma 3.

Let \mathcal{F}_i be the σ -algebra generated by the X_j , $j \leq i$ and let φ belongs to \mathcal{B} such that $\mathbb{E}_\nu(\varphi) = 0$ and $\|\varphi\|_\infty \leq 1$. Using the operator properties, we have

$$\begin{aligned}
& |\text{Cov}(\varphi(X_0)\varphi(X_{i_1^*}) \dots \varphi(X_{i_{q-1}^*}), \varphi(X_{i_q^*})\varphi(X_{i_{q+1}^*}) \dots \varphi(X_{i_p^*})| \\
&= |\mathbb{E} \left[\varphi(X_0)\varphi(X_{i_1^*}) \dots \varphi(X_{i_{q-1}^*}) \left[\mathbb{E}(\varphi(X_{i_q^*})\mathbb{E}(\varphi(X_{i_{q+1}^*}) \dots \mathbb{E}(\varphi(X_{i_p^*})|\mathcal{F}_{i_{p-1}^*}) \dots |\mathcal{F}_{i_q^*})|\mathcal{F}_{i_{q-1}^*}) \right. \right. \\
&\quad \left. \left. - \mathbb{E}(\varphi(X_{i_q^*})\varphi(X_{i_{q+1}^*}) \dots \varphi(X_{i_p^*})) \right] \right]| \\
&\leq \|\varphi(X_0)\varphi(X_{i_1^*}) \dots \varphi(X_{i_{q-1}^*})\|_r \\
&\quad \|\mathbb{E}(\varphi(X_{i_q^*})\mathbb{E}(\varphi(X_{i_{q+1}^*}) \dots \mathbb{E}(\varphi(X_{i_p^*})|\mathcal{F}_{i_{p-1}^*}) \dots |\mathcal{F}_{i_q^*})|\mathcal{F}_{i_{q-1}^*}) - \mathbb{E}(\varphi(X_{i_q^*})\varphi(X_{i_{q+1}^*}) \dots \varphi(X_{i_p^*}))\|_s \\
&\leq C\kappa\|\varphi(X_0)\|_r\theta^{i_q}\|\varphi\mathbb{E}(\varphi(X_{i_{q+1}^*}) \dots \mathbb{E}(\varphi(X_{i_p^*})|\mathcal{F}_{i_{p-1}^*}) \dots |\mathcal{F}_{i_q^*})|\mathcal{F}_{i_{q-1}^*})\|.
\end{aligned}$$

Further,

$$\|\varphi\mathbb{E}(\varphi(X_{i_{q+1}^*}) \dots \mathbb{E}(\varphi(X_{i_p^*})|\mathcal{F}_{i_{p-1}^*}) \dots |\mathcal{F}_{i_q^*})|\mathcal{F}_{i_{q-1}^*})\| \leq C\|\varphi\|^{p-q+1}$$

and the result follows. \square

The corresponding result holds in the setting of dynamical systems. Let $(\Omega, \mathcal{A}, \mu)$ be a probability space and T a measurable measure preserving transformation. Let us consider the Perron-Frobenius operator (or the transfer operator) of T , $P : \mathcal{L}^1(\mu) \rightarrow \mathcal{L}^1(\mu)$ defined by

$$\int_{\Omega} Pf(x)g(x)d\mu(x) = \int_{\Omega} f(x)g \circ T(x)d\mu(x)$$

for all $f \in \mathcal{L}^1(\mu)$ and $g \in \mathcal{L}^\infty(\mu)$.

We assume there exists a Banach algebra $(\mathcal{B}, \|\cdot\|)$ of μ -measurable functions from Ω to \mathbb{R} which contains $\mathbf{1}$ and satisfies (13) and that P verifies: there exist $\kappa > 0$ and $0 < \theta < 1$ such that for all $f \in \mathcal{B}$,

$$\|P^n f - \Pi f\| \leq \kappa\theta^n\|f\|$$

where $\Pi f = \mathbb{E}_\mu(f)\mathbf{1}$.

Lemma 4. *Let $f \in \mathcal{B}$ and $X_i = f \circ T^i$, $i \leq 0$. Then $(X_i)_{i \leq 0}$ satisfies the multiple mixing property (1) on \mathcal{B} with $r = \frac{s}{s-1}$ ($r = 1$ if $s = \infty$).*

For both setting, if the space \mathcal{B} is enough well adapted to approximate indicator functions, then Theorem 3 applies. Examples of Section 4 in Dehling et al. (2009) can be generalized in higher dimensions. In particular, we can state a result for multidimensional linear processes.

Linear processes. Let $(\xi_i)_{i \geq Z}$ be a sequence of i.i.d. random variables in \mathbb{R}^d with $\|\xi_0\|_\infty < \infty$. Let $(a_i)_{i \geq 0}$ be a sequence of endomorphisms of \mathbb{R}^d such that $\|a_i\|_\infty \leq \theta^i$ (for a $\theta < 1$). Define the linear process $X_k = \sum_{i \geq 0} a_i(\xi_{k-i})$ and assume that its distribution function F satisfies condition (6). Then we can show that the process $(X_k)_{k \geq 0}$ satisfies condition (14) for the space of bounded Lipschitz functions on \mathbb{R}^d . We deduce conditions (4) and (5) and Theorem 3 leads to the following corollary.

Corollary 1. *If the distribution function of X_0 verifies condition (6) with $r = 1$, then the multivariate empirical process associated to $(X_k)_{k \geq 0}$ converges in distribution to an almost surely continuous Gaussian process.*

This result was already proved by Dedecker (2010), using a different technique.

Random iterative Lipschitz models. Here, let us focus on application concerning random iterative Lipschitz models. This example has been investigated before by Wu and Shao (2004) or Dedecker and Priour (2007), with different techniques and under different conditions. Here we want to show that our technique also applies in this situation.

Let $g : \mathbb{R}^d \times \mathbb{R} \rightarrow \mathbb{R}^d$ be a measurable function and let $(Y_n)_{n \geq 1}$ be an \mathbb{R} -valued i.i.d. process. Let X_0 be an \mathbb{R}^d -valued random variable independent of $(Y_n)_{n \geq 1}$. Define the Markov chain $(X_n)_{n \in \mathbb{N}}$ by

$$X_n = g(X_{n-1}, Y_n), \quad n \geq 1.$$

Assume that for all $y \in \mathbb{R}$, $g(\cdot, y)$ is Lipschitz. Define the Lipschitz constant

$$K(y) := \sup_{x, x' \in E, x \neq x'} \frac{|g(x, y) - g(x', y)|}{|x - x'|}$$

and suppose that there exists $\gamma_0 > 1$ such that

$$E[(1 + K(Y_1) + |g(0, Y_1)|)^{\gamma_0+1}(1 + K(Y)))] < \infty$$

and

$$E[K(Y_1) \max\{K(Y_1), 1\}^{2\gamma_0}] < 1.$$

Let $\gamma \in (0, \gamma_0]$ and consider the Banach space \mathcal{B}_γ of functions from \mathbb{R}^d to \mathbb{R} satisfying

$$m_\gamma(f) = \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y| p(x)^\gamma p(y)^\gamma} < \infty$$

where $p(x) = 1 + |x|$. The associated norm is $\|\cdot\|_\gamma = N_\gamma(\cdot) + m_\gamma(\cdot)$ with $N_\gamma(f) = \sup_x \frac{|f(x)|}{p(x)^{\gamma+1}}$. According to Hennion and Hervé (2004) Theorem 5.5, the Markov chain has an invariant probability measure ν such that $\nu(|\cdot|^{\gamma_0+1}) < \infty$ and for all $\gamma \in (0, \gamma_0]$, the chain is \mathcal{B}_γ geometrically ergodic. Further, if $f \in \mathcal{B}_\gamma$, $\nu(|f|_0^\gamma)^{\frac{1}{\gamma_0}} \leq N_\gamma(f) \nu(p^{\gamma_0(\gamma+1)})^{\frac{1}{\gamma_0}}$. Then for $\gamma_1 = \frac{\gamma}{\gamma_0}$, \mathcal{B}_{γ_1} is continuously included in $L^{\gamma_0}(\nu)$. By Lemma 3, $(X_n)_{n \geq 0}$ satisfies a multiple mixing property on \mathcal{B}_{γ_1} , for $r = \frac{\gamma_0}{\gamma_0 - 1}$. Note that if f is a bounded Lipschitz function on \mathbb{R}^d then $f \in \mathcal{B}_{\gamma_1}$ and $\|f\|_{\gamma_1} \leq \|f\|$ (where $\|f\| = \sup |f| + m_0(f)$). Thus we have the multiple mixing property for the space of bounded Lipschitz functions. As a corollary of Theorem 3 we get:

Corollary 2. *Assume that the distribution function of ν satisfies (6) for $r = \frac{\gamma_0}{\gamma_0 - 1}$. If the distribution of X_0 is ν , then the empirical process associated to $(X_n)_{n \geq 0}$ converges in distribution to an almost surely continuous Gaussian process.*

5 Ergodic torus automorphisms

Let T be an ergodic automorphism of the torus of dimension d and μ the Lebesgue measure on \mathbb{T}^d . Then we have the following multiple mixing property.

Proposition 3. *There exist $C > 0$, $0 < \gamma < 1$, for all $m, p \in \mathbb{N}^*$, for all bounded α -Hölder function φ ($\alpha \in (0, 1]$) with $\|\varphi\|_\infty \leq 1$, for all $k_1 \leq \dots \leq k_m \leq 0 \leq l_1 \leq \dots \leq l_p$, for all $n \in \mathbb{N}$,*

$$\left| \text{Cov} \left(\prod_{j=1}^m \varphi \circ T^{k_j}, \prod_{j=1}^p \varphi \circ T^{l_j+n} \right) \right| \leq C \|\varphi\|_1 \|\varphi\|_{\mathcal{H}^\alpha} P(k_1, \dots, k_m) \gamma^n \quad (15)$$

where $P(k_1, \dots, k_m) = \sum_{i=1}^m |k_i|^r$ with r the size of the biggest Jordan's block of T restricted to its neutral subspace.

Almost the same proposition appears in Le Borgne and Pène (2005). The slightly modification here is that we keep the L^1 -norm appearing in the upper bound.

Proof.

Denote by E^s , E^u and E^c the T -stable subspaces of \mathbb{R}^d corresponding respectively to the stable, the unstable and the central directions of T (where T is identified to its representative matrix). We have

$$\mathbb{R}^d = E^s \oplus E^u \oplus E^c$$

and there exists $\lambda > 1$ such that for all $n \geq 1$,

$$|T^n v| \leq \lambda^{-n} |v| \text{ for all } v \in E^s,$$

$$|T^n v| \geq \lambda^n |v| \text{ for all } v \in E^u,$$

$$|T^n v| \leq n^r |v| \text{ for all } v \in E^c,$$

where $|\cdot|$ denotes the maximum norm on \mathbb{R}^d and r is the size of the greatest Jordan block of T restricted to the space E^c . Further μ can be written as the product measure of μ_s , μ_u and μ_c . Set

$$B_i(0, \rho) = \{x \in E^i / |x| \leq \rho\}, \quad \rho > 0, i = s, u, c.$$

Denote by $\|\cdot\|_{\mathcal{H}_s^\alpha}$ (resp. $\|\cdot\|_{\mathcal{H}_{u,c}^\alpha}$) the \mathcal{H}^α -norm is the stable direction (resp. unstable-central direction). Following ideas of the proof of Proposition III.3 in Le Borgne (1999) (see also Le Borgne & Pène 2005), one can prove a result concerning the good distribution of the stable leaves in the torus.

Lemma 5. *(good distribution of stable leaves) There exist $\theta < 1$ such that for all $\phi \in \mathcal{H}^\alpha$, $x \in \mathbb{T}^d$ and $\rho > 0$,*

$$\frac{1}{\mu_s(T^{-n}B_s(0, \rho))} \left| \int_{T^{-n}B_s(0, \rho)} \phi(T^{-n}x + s) ds \right| \leq C \|\phi\|_{\mathcal{H}_{u,c}^\alpha} \theta^n.$$

Let \mathcal{A}_0 be a sub- σ -algebra of the Borelian one for which the atoms are pieces of stable leaves and set $\mathcal{A}_n = T^{-n}\mathcal{A}_0$. Let ϕ and ψ be two \mathcal{C}^1 -function with zero mean.

$$\begin{aligned} \text{Cov}(\phi, \psi \circ T^n) &= \text{Cov}(\phi - \mathbb{E}(\phi | \mathcal{A}_{-\lfloor \frac{n}{2} \rfloor}), \psi \circ T^n) + \text{Cov}(\mathbb{E}(\phi | \mathcal{A}_{-\lfloor \frac{n}{2} \rfloor}), \psi \circ T^n) \\ &\leq \|\psi\|_1 \|\phi - \mathbb{E}(\phi | \mathcal{A}_{-\lfloor \frac{n}{2} \rfloor})\|_\infty + \|\phi\|_1 \|\mathbb{E}(\psi | \mathcal{A}_{\lceil \frac{n}{2} \rceil})\|_\infty. \end{aligned}$$

But, since the diameter of the atoms of \mathcal{A}_{-n} decreases exponentially fast,

$$\|\phi - \mathbb{E}(\phi | \mathcal{A}_k)\|_\infty \leq C \|\phi\|_{\mathcal{H}_s^\alpha} \lambda^{-\frac{n}{2}}$$

and, by Lemma 5,

$$\|\mathbb{E}(\psi | \mathcal{A}_{\lceil \frac{n}{2} \rceil})\|_\infty \leq C \|\psi\|_{\mathcal{H}_{u,c}^\alpha} \theta^{\frac{n}{2}}.$$

Thus, for $\gamma = \max\{\lambda, \theta\}^{\frac{1}{2}} < 1$, we get

$$|\text{Cov}(\phi, \psi \circ T^n)| \leq C (\|\psi\|_1 \|\phi\|_{\mathcal{H}_s^\alpha} + \|\phi\|_1 \|\psi\|_{\mathcal{H}_{u,c}^\alpha}) \gamma^n. \quad (16)$$

Further, for all $n \geq 0$ and for all $\phi \in \mathcal{H}^\alpha$, we have

$$\|\phi \circ T^n\|_{\mathcal{H}_s^\alpha} \leq C\|\phi\|_{\mathcal{H}_s^\alpha}. \quad (17)$$

Indeed, if $x \in \mathbb{T}^d$ and $v \in E^s$, by linearity of the map T ,

$$\begin{aligned} |\phi \circ T^n(x) - \phi \circ T^n(x+v)| &= |\phi \circ T^n(x) - \phi(T^n(x) + T^n(v))| \\ &\leq \|\phi\|_{\mathcal{H}_s^\alpha} |T^n(v)|^\alpha \\ &\leq \|\phi\|_{\mathcal{H}_s^\alpha} \lambda^{-n} |v|^\alpha. \end{aligned}$$

In the same way, we get

$$\|\phi \circ T^{-n}\|_{\mathcal{H}_{u,c}^\alpha} \leq Cn^r \|\phi\|_{\mathcal{H}_{u,c}^\alpha}. \quad (18)$$

Now, to prove the proposition, we apply what precedes to

$$\phi = \prod_{j=1}^m \varphi \circ T^{k_j} \quad \text{and} \quad \psi = \prod_{j=1}^p \varphi \circ T^{l_j}$$

with negative k_i and positive l_i . Using (17) and (18), the computation shows that

$$\|\phi\|_{\mathcal{H}_{u,c}^\alpha} \leq \sum_{j=1}^m \|\varphi\|_{\infty}^{m-1} \|\varphi\|_{\mathcal{H}_{u,c}^\alpha} |k_j|^r \leq \|\varphi\|_{\mathcal{H}^\alpha} \sum_{j=1}^m |k_j|^r$$

and

$$\|\psi\|_{\mathcal{H}_s^\alpha} \leq \sum_{j=1}^p \|\varphi\|_{\infty}^{p-1} \|\varphi\|_{\mathcal{H}_s^\alpha} \leq \|\varphi\|_{\mathcal{H}^\alpha}$$

Then, by (16), the proposition is proved. \square

Further the classical central limit theorem holds for any Hölder functions of the torus (see Leonov 1960, Le Borgne 1999)(See Leonov (1960); Le Borgne (1999)). Then Theorem 3 applies and we get Theorem 2.

Acknowledgement

The authors would like to thank the referee for her/his very careful reading of the manuscript and for several thoughtful comments that helped improving the paper.

References

- Berkes, I. & Philipp, W. (1977/78), ‘An almost sure invariance principle for the empirical distribution function of mixing random variables’, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **41**(2), 115–137.
- Billingsley, P. (1968), *Convergence of probability measures*, John Wiley & Sons Inc., New York.
- Borovkova, S., Burton, R. & Dehling, H. (2001), ‘Limit theorems for functionals of mixing processes with applications to U -statistics and dimension estimation’, *Trans. Amer. Math. Soc.* **353**(11), 4261–4318.
- Collet, P., Martinez, S. & Schmitt, B. (2004), ‘Asymptotic distribution of tests for expanding maps of the interval’, *Ergodic Theory Dynam. Systems* **24**(3), 707–722.

- Dedecker, J. (2010), ‘An empirical central limit theorem for intermittent maps’, *Probab. Theory Related Fields* **148**(1-2), 177–195.
- Dedecker, J., Doukhan, P., Lang, G., León R., J. R., Louhichi, S. & Prieur, C. (2007), *Weak dependence: with examples and applications*, Vol. 190 of *Lecture Notes in Statistics*, Springer, New York.
- Dedecker, J. & Prieur, C. (2005), ‘New dependence coefficients. Examples and applications to statistics’, *Probab. Theory Related Fields* **132**(2), 203–236.
- Dedecker, J. & Prieur, C. (2007), ‘An empirical central limit theorem for dependent sequences’, *Stochastic Process. Appl.* **117**(1), 121–142.
- Dehling, H., Durieu, O. & Volný, D. (2009), ‘New techniques for empirical processes of dependent data’, *Stochastic Process. Appl.* **119**(10), 3699–3718.
- Dehling, H. & Taqqu, M. S. (1989), ‘The empirical process of some long-range dependent sequences with an application to U -statistics’, *Ann. Statist.* **17**(4), 1767–1783.
- Denker, M. & Keller, G. (1986), ‘Rigorous statistical procedures for data from dynamical systems’, *J. Statist. Phys.* **44**(1-2), 67–93.
- Donsker, M. D. (1952), ‘Justification and extension of Doob’s heuristic approach to the Komogorov-Smirnov theorems’, *Ann. Math. Statistics* **23**, 277–281.
- Doukhan, P. & Louhichi, S. (1999), ‘A new weak dependence condition and applications to moment inequalities’, *Stochastic Process. Appl.* **84**(2), 313–342.
- Durieu, O. & Jouan, P. (2008), ‘Empirical invariance principle for ergodic torus automorphisms; genericity’, *Stoch. Dyn.* **8**(2), 173–195.
- Gouëzel, S. (2009), ‘An interval map with a spectral gap on Lipschitz functions, but not on bounded variation functions’, *Discrete Contin. Dyn. Syst.* **24**(4), 1205–1208.
- Hennion, H. & Hervé, L. (2001), *Limit theorems for Markov chains and stochastic properties of dynamical systems by quasi-compactness*, Vol. 1766 of *Lecture Notes in Mathematics*, Springer-Verlag, Berlin.
- Hennion, H. & Hervé, L. (2004), ‘Central limit theorems for iterated random Lipschitz mappings’, *Ann. Probab.* **32**(3A), 1934–1984.
- Hofbauer, F. & Keller, G. (1982), ‘Ergodic properties of invariant measures for piecewise monotonic transformations’, *Math. Z.* **180**(1), 119–140.
- Le Borgne, S. (1999), ‘Limit theorems for non-hyperbolic automorphisms of the torus’, *Israel J. Math.* **109**, 61–73.
- Le Borgne, S. & Pène, F. (2005), ‘Vitesse dans le théorème limite central pour certains systèmes dynamiques quasi-hyperboliques’, *Bull. Soc. Math. France* **133**(3), 395–417.
- Leonov, V. P. (1960), ‘On the central limit theorem for ergodic endomorphisms of compact commutative groups’, *Dokl. Akad. Nauk SSSR* **135**, 258–261.
- Lind, D. A. (1982), ‘Dynamical properties of quasihyperbolic toral automorphisms’, *Ergodic Theory Dynamical Systems* **2**(1), 49–68.
- Neuhaus, G. (1971), ‘On weak convergence of stochastic processes with multidimensional time parameter’, *Ann. Math. Statist.* **42**, 1285–1295.
- Straf, M. L. (1972), Weak convergence of stochastic processes with several parameters, in ‘Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability theory’, Univ. California Press, Berkeley, Calif., pp. 187–221.

Wu, W. B. (2008), 'Empirical processes of stationary sequences', *Statist. Sinica* **18**(1), 313–333.

Wu, W. B. & Shao, X. (2004), 'Limit theorems for iterated random functions', *J. Appl. Probab.* **41**(2), 425–436.