Empirical central limit theorems for ergodic automorphisms of the torus

Jérôme Dedecker, Florence Merlevède, Françoise Pene

▶ To cite this version:

HAL Id: hal-00741466

https://hal.archives-ouvertes.fr/hal-00741466

Submitted on 12 Oct 2012

Empirical central limit theorems for ergodic automorphisms of the torus

Jérôme Dedecker [a], Florence Merlevède [b] and Françoise Pène [c]

[a] Université Paris Descartes, Sorbonne Paris Cité, Laboratoire MAP5 and CNRS UMR 8145, 45 rue des Saints-Pères, F-75270 Paris cedex 06, France. E-mail: jerome.dedecker@parisdescartes.fr

[b] Université Paris-Est, LAMA (UMR 8050), UPEMLV, CNRS, UPEC, F-77454 Marne-La-Vallée, FRANCE. E-mail: florence.merlevede@univ-mlv.fr

[c] Université de Brest, Laboratoire de Mathématiques de Bretagne Atlantique UMR CNRS 6205. Supported by the french ANR project Perturbations (ANR 10-BLAN 0106). E-mail: francoise.pene@univ-brest.fr

Abstract

Let T be an ergodic automorphism of the d-dimensional torus $\mathbb{T}^d$, and f be a continuous function from $\mathbb{T}^d$ to $\mathbb{R}^\ell$. On the probability space $\mathbb{T}^d$ equipped with the Lebesgue-Haar measure, we prove the weak convergence of the sequential empirical process of the sequence $(f \circ T^i)_{i \geq 1}$ under some condition on the modulus of continuity of f. The proofs are based on new limit theorems and new inequalities for non-adapted sequences, and on new estimates of the conditional expectations of f with respect to a natural filtration.

1 Introduction

Let $d \geq 2$ and $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ be the d-dimensional torus. For every $x \in \mathbb{R}^d$, we write $\bar{x}$ its class in $\mathbb{T}^d$. We denote by λ the Lebesgue measure on $\mathbb{R}^d$, and by $\bar{\lambda}$ the Lebesgue measure on $\mathbb{T}^d$.

On the probability space $(\mathbb{T}^d, \bar{\lambda})$, we consider a group automorphism T of $\mathbb{T}^d$. We recall that T is the quotient map of a linear map $\tilde{T} : \mathbb{R}^d \to \mathbb{R}^d$ given by $\tilde{T}(x) = S \cdot x$, where S is a $d \times d$-matrix with integer entries and with determinant 1 or -1. The map $\tilde{T}$ preserves the infinite Lebesgue measure λ on $\mathbb{R}^d$ and T preserves the probability Lebesgue measure $\bar{\lambda}$.

We assume that T is ergodic, which is equivalent to the fact that no eigenvalue of S is a root of the unity. This hypothesis holds true in the case of hyperbolic automorphisms of the torus (i.e. in the case when no eigenvalue of S has modulus one) but is much weaker. Indeed, as mentionned in [9], the following matrix gives an example of an ergodic non-hyperbolic automorphism of $\mathbb{T}^4$:

$$S := \begin{pmatrix} 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \end{pmatrix}.$$

When T is ergodic but non-hyperbolic, the dynamical system $(\mathbb{T}^d, T, \bar{\lambda})$ has no Markov partition. However, it is possible to construct some measurable partition (see [11]), and to prove some decorrelation properties for regular functions (see [11, 10]).

Let ℓ be some positive integer, and let $f = (f_1, \ldots f_\ell)$ be a function from $\mathbb{T}^d$ to $\mathbb{R}^\ell$. On the probability space $(\mathbb{T}^d, \bar{\lambda})$, the sequence $(f \circ T^k)_{k \in \mathbb{Z}}$ is a stationary sequence of $\mathbb{R}^\ell$-valued random variables. When $\ell = 1$ and f is square integrable, Le Borgne [9] proved the functional central limit theorem and the Strassen strong invariance principle for the partial sums

$$\sum_{i=1}^{n} (f \circ T^i - \bar{\lambda}(f))$$

under weak hypotheses on the Fourier coefficients of f, thanks to Gordin's method and to the partitions studied by Lind in [11]. In the recent paper [4], we slightly improve on Le Borgne's conditions, and we show how to obtain rates of convergence in the strong invariance principle up to $n^{1/4} \log(n)$, by reinforcing the conditions on the Fourier coefficients of f.

Now, for any $s \in \mathbb{R}^\ell$, define the partial sum

$$S_n(s) = \sum_{k=1}^{n} (\mathbf{1}_{f \circ T^k \leq s} - F(s)), \tag{1.1}$$

where as usual $\mathbf{1}_{f \circ T^k \leq s} = \mathbf{1}_{f_1 \circ T^k \leq s_1} \times \cdots \times \mathbf{1}_{f_\ell \circ T^k \leq s_\ell}$, and $F(s) = \bar{\lambda}(\mathbf{1}_{f \circ T^k \leq s})$ is the multivariate distribution function of f.

In this paper, we give some conditions on the modulus of continuity of f for the weak convergence to a Gaussian process of the sequential empirical process

$$\left\{ \frac{S_{[nt]}(s)}{\sqrt{n}}, t \in [0, 1], s \in \mathbb{R}^\ell \right\}. \tag{1.2}$$

The paper is organized as follows. Our main results are given in Section 2 and proved in Section 5. The proofs require new probabilistic results established in Section 3 combined with a key estimate for toral automorphisms which is given in Section 4. Let us give now an overview of our results.

In Section 2.1, we consider the case where $\ell = 1$ and S_n is viewed as an $\mathbb{L}^p$-valued random variable for some $p \in [2, \infty[$ (this is possible because $\int |S_n(s)|^p ds < \infty$ for any $p \in [2, \infty[)$, so that the sequential empirical process is an element of $D_{\mathbb{L}^p}([0, 1])$, the space of $\mathbb{L}^p$-valued càdlàg functions. We prove the weak convergence on $D_{\mathbb{L}^p}([0, 1])$ equipped with the uniform metric to a $\mathbb{L}^p$-valued Wiener process, and we give the covariance operator of this Wiener process. The proof is based on a new central limit theorem for dependent sequences with values in smooth Banach spaces, which is given in Section 3.1.1.

In Section 2.2, we state the convergence of the sequential empirical process (1.2) in the space $\ell^\infty([0, 1] \times \mathbb{R}^\ell)$ of bounded functions from $[0, 1] \times \mathbb{R}^\ell$ to $\mathbb{R}$ equipped with the uniform metric. In that case, the limiting Gaussian process is a generalization of the Kiefer process introduced by Kiefer in [8] for the sequential empirical process of independent and identically distributed random variables. The proof is based on a new Rosenthal inequality for dependent sequences, which is given in Section 3.1.2. The weak convergence of the empirical process $\{n^{-1/2} S_n(s), s \in \mathbb{R}^\ell\}$ has also been treated in [7] and [6]. We shall be more precise on these two papers in Section 2.2.

To prove these results, we shall use a control of the conditional expectations of continuous observables with respect to the filtration introduced by Lind [11], involving the modulus of continuity of the observables (See Theorem 18 of Section 4). As far as we know, such controls were known for Hölder observables only (see [10]). The inequalities given in Theorem 18 can be used in many other situations. Let us give two examples of applications. Let f be a continuous function from $\mathbb{T}^d$ to $\mathbb{R}$ with modulus of continuity $\omega(f, \cdot)$ (see Section 4, equation (4.1), for the definition).

1. Weak invariance principle. If

$$\int_0^{1/2} \frac{\omega(f, t)}{t |\log t|^{1/2}} dt < \infty,$$

2

then the series
$$\sigma^2(f) = \bar{\lambda}((f - \bar{\lambda}(f))^2) + 2 \sum_{k>0} \bar{\lambda}((f - \bar{\lambda}(f)) \cdot f \circ T^k)$$

converges absolutely, and the process

$$\left\{ \frac{1}{\sqrt{n}} \sum_{k=1}^{[nt]} f \circ T^k, t \in [0,1] \right\}$$

converges to a Wiener process with variance $\sigma^2(f)$ in the space $D([0,1])$ of càdlàg function equipped with the uniform metric .

2. Rates of convergence in the strong invariance principle. Let $p \in]2,4]$, and assume that

$$\omega(f,x) \leq C|\log(x)|^{-a} \text{ in a neighborhood of 0 for some } a > \frac{1 + \sqrt{1 + 4p(p-2)}}{2p} + 1 - \frac{2}{p}.$$

Then, enlarging $\mathbb{T}^d$ if necessary, there exists a sequence $(Z_i)_{i\geq 1}$ of independent and identically distributed Gaussian random variables with mean zero and variance $\sigma^2(f)$ such that, for any $t > 2/p$,

$$\sup_{1 \leq k \leq n} \left| \sum_{i=1}^{k} f \circ T^i + \sum_{i=1}^{k} Z_i \right| = o\big(n^{1/p}(\log(n))^{(t+1)/2}\big) \quad \text{almost surely as } n \to \infty.$$

In particular, we obtain the rate of convergence $n^{1/4} \log(n)$ as soon as $a \geq 3/2$. This follows from Theorem 3.1 in [3].

2 Empirical central limit theorems

2.1 Empirical central limit theorem in $\mathbb{L}^p$

In this section, $\mathbb{L}^p$ is the space of Borel-measurable functions g from $\mathbb{R}$ to $\mathbb{R}$ such that $\lambda(|g|^p) < \infty$, λ being the Lebesgue measure on $\mathbb{R}$. If f is a bounded function, then, for any any $p \in [2,\infty[$, the random variable S_n defined in (1.1) is an $\mathbb{L}^p$-valued random variable, and the process $\{n^{-1/2}S_{[nt]}, t \in [0,1]\}$ is a random variable with values in $D_{\mathbb{L}^p}([0,1])$, the space of $\mathbb{L}^p$-valued càdlàg functions. In the next theorem, we give a condition on the modulus of continuity $\omega(f,\cdot)$ of f under which the process $\{n^{-1/2}S_{[nt]}, t \in [0,1]\}$ converges in distribution to an $\mathbb{L}^p$-valued Wiener process, in the space $D_{\mathbb{L}^p}([0,1])$ equipped with the uniform metric. We refer to Section 4, equation (4.1), for the precise definition of $\omega(f,\cdot)$.

By an $\mathbb{L}^p$-valued Wiener process with covariance operator Λ_p, we mean a centered Gaussian process $W = \{W_t, t \in [0,1]\}$ such that $\mathbb{E}(\|W_t\|^2_{\mathbb{L}^p}) < \infty$ for all $t \in [0,1]$ and, for any g, h in $\mathbb{L}^q$ (q being the conjugate exponent of p),

$$\text{Cov}\Big(\int_{\mathbb{R}} g(u)W_t(u)du, \int_{\mathbb{R}} h(u)W_s(u)du \Big) = \min(t,s)\Lambda_p(g,h).$$

Theorem 1. *Let* $f : \mathbb{T}^d \to \mathbb{R}$ *be a continuous function, with modulus of continuity* $\omega(f,\cdot)$. *Let* $p \in [2,\infty[$, *and let* q *be its conjugate exponent. Assume that*

$$\int_0^{1/2} \frac{\big(\omega(f,t)\big)^{1/p}}{t|\log t|^{1/p}} dt < \infty.$$

3

Then the process $\{n^{-1/2}S_{[nt]}, t \in [0,1]\}$ converges in distribution in the space $D_{\mathbb{L}^p}([0,1])$ to an $\mathbb{L}^p$-valued Wiener process W, with covariance operator Λ_p defined by

$$\Lambda_p(g,h) = \sum_{k \in \mathbb{Z}} \mathrm{Cov}\left(\int_{\mathbb{R}} g(s)\mathbf{1}_{f \leq s}ds, \int_{\mathbb{R}} h(s)\mathbf{1}_{f \circ T^k \leq s}ds \right), \quad \text{for any } g, h \text{ in } \mathbb{L}^q. \tag{2.1}$$

The proof of Theorem 1 is based on results of Sections 3 and 4 and is postponed to Section 5.

Remark 2. *In particular, if f is Hölder continuous, then the conclusion of Theorem 1 holds for any $p \in [2, \infty[$.*

Let us give an application of this theorem to the Kantorovich-Rubinstein distance between the empirical measure of $(f \circ T^i)_{1 \leq i \leq n}$ and the distribution μ of f. Let

$$\mu_n = \frac{1}{n} \sum_{i=1}^{n} \delta_{f \circ T^i} \quad \text{and} \quad \mu_{n,k} = \frac{1}{n}\left((n-k)\mu + \sum_{i=1}^{k} \delta_{f \circ T^i} \right).$$

The Kantorovich distance between two probability measures ν_1 and ν_2 is defined as

$$K(\nu_1, \nu_2) = \inf \left\{ \int |x - y|\lambda(dx, dy), \lambda \in \mathcal{M}(\nu_1, \nu_2) \right\},$$

where $\mathcal{M}(\nu_1, \nu_2)$ is the set of probability measures with margins ν_1 and ν_2.

Corollary 3. *Let $f : \mathbb{T}^d \to \mathbb{R}$ be a continuous function, with modulus of continuity $\omega(f, \cdot)$. Assume that*

$$\int_0^{1/2} \frac{\sqrt{\omega(f,t)}}{t\sqrt{|\log t|}} dt < \infty.$$

Then $\sqrt{n}K(\mu_n, \mu)$ converges in distribution to $\|W_1\|_{\mathbb{L}^1}$, and $\sup_{1 \leq k \leq n} \sqrt{n}K(\mu_{n,k}, \mu)$ converges in distribution to $\sup_{t \in [0,1]} \|W_t\|_{\mathbb{L}^1}$, where W is the $\mathbb{L}^2$-valued Wiener process with covariance operator Λ_2 defined by (2.1).

Proof of Corollary 3. Applying Theorem 1 with $p = 2$, we know that $\{n^{-1/2}S_{[nt]}, t \in [0,1]\}$ converges in distribution in the space $D_{\mathbb{L}^2}([0,1])$ to an $\mathbb{L}^2$-valued Wiener process W, with covariance operator Λ_2 defined by (2.1). Since f is continuous on $\mathbb{T}^d$, it follows that $|f| \leq M$ for some positive constant M, so that $S_{[nt]}(s) = 0$ and $W_t(s) = 0$ for any $t \in [0,1]$ and any $|s| > M$. Since $\|\cdot\|_{\mathbb{L}^1}$ is a continuous function on the space of functions in $\mathbb{L}^2$ with support in $[-M, M]$, it follows that $n^{-1/2}\|S_n\|_{\mathbb{L}^1}$ converges in distribution to $\|W_1\|_{\mathbb{L}^1}$, and that $\sup_{t \in [0,1]} n^{-1/2}\|S_{[nt]}\|_{\mathbb{L}^1}$ converges in distribution to $\sup_{t \in [0,1]} \|W_t\|_{\mathbb{L}^1}$. Now, if ν_1 and ν_2 are probability measures on the real line, with distribution functions F_{ν_1} and F_{ν_2} respectively,

$$K(\nu_1, \nu_2) = \int_{\mathbb{R}} |F_{\nu_1}(t) - F_{\nu_2}(t)|dt.$$

Hence $nK(\mu_n, \mu) = \|S_n\|_{\mathbb{L}^1}$ and $\sup_{1 \leq k \leq n} nK(\mu_{n,k}, \mu) = \sup_{t \in [0,1]} \|S_{[nt]}\|_{\mathbb{L}^1}$, and the result follows. $\square$

2.2 Weak convergence to the Kiefer process

Let ℓ be a positive integer. Let $f = (f_1, \ldots, f_\ell)$ be a continuous function from $\mathbb{T}^d$ to $\mathbb{R}^\ell$. The modulus of continuity $\omega(f, \cdot)$ of f is defined by

$$\omega(f, x) = \sup_{1 \leq i \leq \ell} \omega(f_i, x),$$

where we recall that $\omega(f_i, x)$ is defined by equation (4.1).

As usual, we denote by $\ell^\infty([0,1] \times \mathbb{R}^\ell)$ the space of bounded functions from $[0,1] \times \mathbb{R}^\ell$ to $\mathbb{R}$ equipped with the uniform norm. For details on weak convergence on the non separable space $\ell^\infty([0,1] \times \mathbb{R}^\ell)$, we refer to [17] (in particular, we shall not discuss any measurability problems, which can be handled by using the outer probability).

For any positive integer ℓ and any $\alpha \in]0,1]$, let

$$a(\ell, \alpha) = \min_{p \geq \max(\ell+2, 2\ell)} g_{\ell,\alpha}(p), \text{ where } g_{\ell,\alpha}(p) = \max\left(\frac{p}{\alpha(p-2\ell)}, \frac{(p-1)(2\alpha+p)}{p\alpha}\right). \qquad (2.2)$$

Note that this minimum is reached at $p_1 = \max(3, p_0)$, where p_0 is the unique solution in $]2\ell, 4\ell[$ of the equation

$$\frac{p}{(p-2\ell)} = \frac{(p-1)(p+2\alpha)}{p} \qquad (2.3)$$

(in particular, $p_1 = p_0$ if $\ell > 1$).

We are now in position to state the main result of this section.

Theorem 4. *Let $f = (f_1, \ldots, f_\ell) : \mathbb{T}^d \to \mathbb{R}^\ell$ be a continuous function, with modulus of continuity $\omega(f, \cdot)$. Assume that the distribution functions of the f_i's are Hölder continuous of order $\alpha \in]0,1]$. If*

$$\omega(f, x) \leq C|\log(x)|^{-a} \quad \text{for some} \quad a > a(\ell, \alpha),$$

then the process $\{n^{-1/2} S_{[nt]}(s), t \in [0,1], s \in \mathbb{R}^\ell\}$ converges in distribution in the space $\ell^\infty([0,1] \times \mathbb{R}^\ell)$ to a Gaussian process K with covariance function Γ defined by: for any $(t, t') \in [0,1]^2$ and any $(s, s') \in \mathbb{R}^\ell \times \mathbb{R}^\ell$,

$$\Gamma(t, t', s, s') = min(t, t')\Lambda(s, s') \quad with \quad \Lambda(s, s') = \sum_{k \in \mathbb{Z}} \text{Cov}(\mathbf{1}_{f \circ T \leq s}, \mathbf{1}_{f \circ T^k \leq s'}).$$

The proof of Theorem 4 is given in Section 5. It uses results of Sections 3 and 4.

Remark 5. *Using the Cardan formulas (see the appendix) to solve (2.3), we get*

$$p_0 = 2\frac{\ell + 1 - \alpha}{3} + 2\sqrt{-\frac{p'}{3}} \cos\left(\frac{1}{3} \arccos\left(-\frac{q}{2}\sqrt{\frac{27}{-(p')^3}}\right)\right),$$

with

$$p' := -4\alpha\ell + 2\ell - 2\alpha - \frac{1}{3}(-2\ell + 2\alpha - 2)^2 < 0$$

and

$$q := \frac{1}{27}(-2\ell + 2\alpha - 2)(2(-2\ell + 2\alpha - 2)^2 + 36\alpha\ell - 18\ell + 18\alpha) + 4\alpha\ell.$$

For example, for $\alpha = \ell = 1$, we get $p_0 \sim 2.903211927$ and finally $a(1,1) = 10/3$.

Recall that, by Theorem 1, if $\ell = 1$ and $p \in]2, \infty[$, the weak invariance principle holds in $D_{\mathbb{L}^p}([0,1])$ as soon as $a > p - 1$ without any condition on the distribution function of f.

The weak convergence of the (non sequential) empirical process $\{n^{-1/2} S_n(s), s \in \mathbb{R}^\ell\}$ has been studied in [7] and [6]. When $\ell = 1$, a consequence of the main result of the paper [7] is that the empirical process converges weakly to a Gaussian process for any Hölder continuous function f having an Hölder continuous distribution function. In the paper [6] this result is extended to any dimension ℓ, under the assumption that the moduli of continuity of the distribution functions of the f_i's are smaller than $C|\log(x)|^{-a}$ in a neighborhood of 0, for some $a > 1$.

Note that, in our case, one cannot apply Theorem 1 of [6]. Indeed, one cannot prove the multiple mixing for the sequence $(f \circ T^i)_{i \in \mathbb{Z}}$ by assuming only that $\omega(f, x) \leq C |\log(x)|^{-a}$ in a neighborhood of zero (in that case one can only prove that $|\mathrm{Cov}(f, f \circ T^n)|$ is $O(n^{-a})$). However, even if our condition on the regularity of f is much weaker than in [6], our result cannot be directly compared to that of [6], because we assume that the distribution functions of the f_i's are Hölder continuous of order α, which is a stronger assumption than the corresponding one in [6].

3 Probabilistic results

In this section, C is a positive constant which may vary from lines to lines, and the notation $a_n \ll b_n$ means that there exists a numerical constant C not depending on n such that an $a_n \leq C b_n$, for all positive integers n.

3.1 Limit theorems and inequalities for stationary sequences

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, and $T : \Omega \mapsto \Omega$ be a bijective bimeasurable transformation preserving the probability $\mathbb{P}$. For a σ-algebra $\mathcal{F}_0$ satisfying $\mathcal{F}_0 \subseteq T^{-1}(\mathcal{F}_0)$, we define the nondecreasing filtration $(\mathcal{F}_i)_{i \in \mathbb{Z}}$ by $\mathcal{F}_i = T^{-i}(\mathcal{F}_0)$. Let $\mathcal{F}_{-\infty} = \bigcap_{k \in \mathbb{Z}} \mathcal{F}_k$ and $\mathcal{F}_\infty = \bigvee_{k \in \mathbb{Z}} \mathcal{F}_k$. Let $\mathcal{I}$ be the σ-algebra of T-invariant sets. As usual, we say that $(T, \mathbb{P})$ is ergodic if each element A of $\mathcal{I}$ is such that $\mathbb{P}(A) = 0$ or 1.

Let $(\mathbb{B}, |\cdot|_{\mathbb{B}})$ be a separable Banach space. For a random variable X with values in $\mathbb{B}$, let $\|X\|_p = (\mathbb{E}(|X|_{\mathbb{B}}^p))^{1/p}$ and $\mathbb{L}^p(\mathbb{B})$ be the space of $\mathbb{B}$-valued random variables such that $\|X\|_p < \infty$. For $X \in \mathbb{L}^1(\mathbb{B})$, we shall use the notations $\mathbb{E}_k(X) = \mathbb{E}(X|\mathcal{F}_k)$, $\mathbb{E}_\infty(X) = \mathbb{E}(X|\mathcal{F}_\infty)$, $\mathbb{E}_{-\infty}(X) = \mathbb{E}(X|\mathcal{F}_{-\infty})$, and $P_k(X) = \mathbb{E}_k(X) - \mathbb{E}_{k-1}(X)$. Recall that $\mathbb{E}(X|\mathcal{F}_n) \circ T^m = \mathbb{E}(X \circ T^m|\mathcal{F}_{n+m})$.

Let X_0 be a random variable with values in $\mathbb{B}$. Define the stationary sequence $(X_i)_{i \in \mathbb{Z}}$ by $X_i = X_0 \circ T^i$, and the partial sum S_n by $S_n = X_1 + X_2 + \cdots + X_n$.

3.1.1 Weak invariance principle in smooth Banach spaces

Following Pisier [16], we say that a Banach space $(\mathbb{B}, |\cdot|_{\mathbb{B}})$ is 2-smooth if there exists an equivalent norm $\|\cdot\|$ such that

$$\sup_{t>0} \left\{ \frac{1}{t^2} \sup\{\|x + ty\| + \|x - ty\| - 2 : \|x\| = \|y\| = 1\} \right\} < \infty.$$

From [16], we know that if $\mathbb{B}$ is 2-smooth and separable, then there exists a constant K such that, for any sequence of $\mathbb{B}$-valued martingale differences $(D_i)_{i \geq 1}$,

$$\mathbb{E}(|D_1 + \cdots + D_n|_{\mathbb{B}}^2) \leq K \sum_{i=1}^{n} \mathbb{E}(|D_i|_{\mathbb{B}}^2). \tag{3.1}$$

From [16], we see that 2-smooth Banach spaces play the same role for martingales as space of type 2 do for sums of independent variables. Note that, for any measure space $(T, \mathcal{A}, \nu)$, $\mathbb{L}^p(T, \mathcal{A}, \nu)$ is 2-smooth with $K = p - 1$ for any $p \geq 2$, and that any separable Hilbert space is 2-smooth with $K = 2$.

Let $D_{\mathbb{B}}([0, 1])$ be the space of $\mathbb{B}$-valued càdlàg functions. In the next theorem, we give a condition under which the process $\{n^{-1/2} S_{[nt]}, t \in [0, 1]\}$ converges in distribution to an $\mathbb{B}$-valued Wiener process, in the space $D_{\mathbb{B}}([0, 1])$ equipped with the uniform metric.

By an $\mathbb{B}$-valued Wiener process with covariance operator $\Lambda_{\mathbb{B}}$, we mean a centered Gaussian process $W = \{W_t, t \in [0, 1]\}$ such that $\mathbb{E}(|W_t|_{\mathbb{B}}^2) < \infty$ for all $t \in [0, 1]$ and, for any g, h in the dual space $\mathbb{B}^*$,

$$\mathrm{Cov}(g(W_t), h(W_s)) = \min(t, s) \Lambda_{\mathbb{B}}(g, h).$$

Proposition 6. *Assume that $\mathbb{B}$ be a 2-smooth Banach space having a Schauder Basis, that $(T, \mathbb{P})$ is ergodic, that $\|X_0\|_2 < \infty$ and that $\mathbb{E}(X_0) = 0$. If*

$$\sum_{k \in \mathbb{Z}} \|P_0(X_i)\|_2 < \infty \tag{3.2}$$

then the process $\{n^{-1/2} S_{[nt]}, t \in [0,1]\}$ converges in distribution in the space $D_{\mathbb{B}}([0,1])$ equipped with the uniform metric to a $\mathbb{B}$-valued Wiener process $W_{\Lambda_{\mathbb{B}}}$, where $\Lambda_{\mathbb{B}}$ is the covariance operator defined by

$$\text{for any } g, h \text{ in } \mathbb{B}^*, \quad \Lambda_{\mathbb{B}}(g, h) = \sum_{k \in \mathbb{Z}} \text{Cov}(g(X_0), h(X_k)) \,.$$

Proof of Proposition 6. Let us prove first that the result holds if $\mathbb{E}(X_0 | \mathcal{F}_{-1}) = 0$, that is when $(X_k)_{k \in \mathbb{Z}}$ is a martingale difference sequence. As usual, it suffices to prove that:

1. for any $0 = t_0 < t_1 < \cdots < t_d = 1$

$$\frac{1}{\sqrt{n}} (S_{[nt_1]}, S_{[nt_2]} - S_{[nt_1]}, \cdots, S_{[nt_d]} - S_{[nt_{d-1}]})$$

   converges in distribution to a the Gaussian distribution μ on $\mathbb{B}^d$ defined by $\mu = \mu_1 \otimes \mu_2 \cdots \otimes \mu_d$, where μ_i is the Gaussian distribution on $\mathbb{B}$ with covariance operator C_i:

$$\text{for any } g, h \text{ in } \mathbb{B}^*, \quad C_i(g, h) = (t_i - t_{i-1}) \text{Cov}(g(X_0), h(X_0)) \,.$$

2. For any $\varepsilon > 0$,

$$\lim_{\delta \to 0} \limsup_{n \to \infty} \frac{1}{\delta} \mathbb{P} \Big(\max_{1 \le k \le [n\delta]} |S_k|_{\mathbb{B}} > 2\sqrt{n}\varepsilon \Big) = 0$$

The first point can be proved exactly as in [18], who proved the result only for $t_1 = 1$. Let us prove the second point. For any positive number M, let

$$X_i' = X_i \mathbf{1}_{|X_i|_{\mathbb{B}} \le M} - \mathbb{E}(X_i \mathbf{1}_{|X_i|_{\mathbb{B}} \le M} | \mathcal{F}_{i-1}) \quad \text{and} \quad X_i'' = X_i - X_i' \,.$$

Let also $S_n' = X_1' + \cdots + X_n'$ and $S_n'' = X_1'' + \cdots + X_n''$. Since $\mathbb{B}$ is 2-smooth, Burkholder's inequality holds (see for instance [15]), in such a way that $\mathbb{E}(\max_{1 \le k \le n} |S_k'|_{\mathbb{B}}^q) \le K_q M^q n^{q/2}$ for any $q \ge 2$. Hence, applying Markov's inequality at order $q > 2$,

$$\frac{1}{\delta} \mathbb{P} \Big(\max_{1 \le k \le [n\delta]} |S_k'|_{\mathbb{B}} > \sqrt{n}\varepsilon \Big) \le \frac{K_q M^q \delta^{(q-2)/2}}{\varepsilon^q} \,.$$

As a consequence, we get that

$$\lim_{\delta \to 0} \limsup_{n \to \infty} \frac{1}{\delta} \mathbb{P} \Big(\max_{1 \le k \le [n\delta]} |S_k'|_{\mathbb{B}} > \sqrt{n}\varepsilon \Big) = 0. \tag{3.3}$$

In the same way, applying Markov's inequality at order 2

$$\frac{1}{\delta} \mathbb{P} \Big(\max_{1 \le k \le [n\delta]} |S_k''|_{\mathbb{B}} > \sqrt{n}\varepsilon \Big) \le \frac{K_2}{\varepsilon^2} \mathbb{E}(|X_0|_{\mathbb{B}}^2 \mathbf{1}_{|X_0|_{\mathbb{B}} > M}) \,. \tag{3.4}$$

The term $\mathbb{E}(|X_0|_{\mathbb{B}}^2 \mathbf{1}_{|X_0|_{\mathbb{B}} > M})$ is as small as we wish by choosing M large enough. The point 2 follows from (3.3) and (3.4).

7

We now consider the general case. Since $\mathbb{B}$ is 2-smooth, Burkholder's inequality holds and so Proposition 4.1 in [4] applies: if (3.2) holds, then, setting $d_k = \sum_{i \in \mathbb{Z}} P_k(X_i)$, we have

$$\Big\| \max_{1 \le k \le n} \Big| \sum_{i=1}^{k} X_i - \sum_{i=1}^{k} d_i \Big|_{\mathbb{B}} \Big\|_2 = o(\sqrt{n}). \tag{3.5}$$

Since $(d_i)_{i \in \mathbb{Z}}$ is a stationary martingale differences sequence in $\mathbb{L}^2(\mathbb{B})$, we have just proved that it satisfies the conclusion of Proposition 6. From (3.5) it follows that the conclusion of Proposition 6 is also true for $(X_i)_{i \in \mathbb{Z}}$ with

$$\Lambda_{\mathbb{B}}(g, h) = \mathrm{Cov}(g(d_0), h(d_0)), \quad \text{for any } g, h \text{ in } \mathbb{B}^*.$$

It remains to see that this covariance function can also be written as in Proposition 6. Recall that, for any g and h in $\mathbb{B}^*$,

$$\sum_{k \in \mathbb{Z}} |\mathrm{Cov}(g(X_0), h(X_k))| \le \Big(\sum_{k \in \mathbb{Z}} \|P_0(g(X_k))\|_2 \Big) \Big(\sum_{k \in \mathbb{Z}} \|P_0(h(X_k))\|_2 \Big) < \infty.$$

(see the proof of item 1 of Theorem 4.1 in [4]). Hence, for any g in $\mathbb{B}^*$,

$$\lim_{n \to \infty} \frac{1}{n} \mathbb{E}\Big(\Big(\sum_{k=1}^{n} g(X_k) \Big)^2 \Big) = \sum_{k \in \mathbb{Z}} \mathrm{Cov}(g(X_0), g(X_k)). \tag{3.6}$$

Now, from (3.5), we also know that

$$\lim_{n \to \infty} \frac{1}{n} \mathbb{E}\Big(\Big(\sum_{k=1}^{n} g(X_k) \Big)^2 \Big) = \mathbb{E}((g(d_0))^2). \tag{3.7}$$

Applying (3.6) and (3.7) with g, h and $g + h$, we infer that

$$\mathrm{Cov}(g(d_0), h(d_0)) = \sum_{k \in \mathbb{Z}} \mathrm{Cov}(g(X_0), h(X_k)),$$

which completes the proof. $\qquad\qquad\square$

3.1.2 A Rosenthal inequality for non adapted sequences

We begin with a maximal inequality that is useful to compare the moment of order p of the maximum of the partial sums of a non necessarily adapted process to the corresponding moment of the partial sum. The adapted version of this inequality has been proven in the adapted case (that is when X_0 is $\mathcal{F}_0$-measurable) in [12]. Notice that Proposition 2 of [12] is stated for real valued random variables, but it holds also for variables taking values in a separable Banach space $(\mathbb{B}, |\cdot|_{\mathbb{B}})$.

Proposition 7. *Let $p > 1$ be a real number and q be its conjugate exponent. Let X_0 be a random variable in $\mathbb{L}^p(\mathbb{B})$ and $\mathcal{F}_0$ a σ-algebra satisfying $\mathcal{F}_0 \subseteq T^{-1}(\mathcal{F}_0)$. Then, for any integer r, the following inequality holds:*

$$\Big\| \max_{1 \le m \le 2^r} |S_m|_{\mathbb{B}} \Big\|_p \le q \|S_{2^r}\|_p + q 2^{r/p} \sum_{\ell=0}^{r-1} 2^{-\ell/p} \|\mathbb{E}_0(S_{2^\ell})\|_p + (q+1) 2^{r/p} \sum_{\ell=0}^{r} 2^{-\ell/p} \|S_{2^\ell} - \mathbb{E}_{2^\ell}(S_{2^\ell})\|_p. \tag{3.8}$$

8

Remark 8. *If we do not assume stationarity, so if we consider a sequence $(X_i)_{i \in \mathbb{Z}}$ in $\mathbb{L}^p(\mathbb{B})$ for a $p > 1$, and an increasing filtration $(\mathcal{F}_i)_{i \in \mathbb{Z}}$, our proof reveals that the following inequality holds true: for any integer r,*

$$\left\| \max_{1 \leq m \leq 2^r} |S_m|_{\mathbb{B}} \right\|_p \leq q\|S_{2^r}\|_p + q \sum_{l=0}^{r-1} \left(\sum_{k=1}^{2^{r-l}-1} \|\mathbb{E}_{k2^l}(S_{(k+1)2^l} - S_{k2^l})\|_p^p \right)^{1/p}$$

$$+ (q+1) \sum_{l=0}^{r} \left(\sum_{k=1}^{2^{r-l}} \|S_{k2^l} - S_{(k-1)2^l} - \mathbb{E}_{k2^l}(S_{k2^l} - S_{(k-1)2^l})\|_p^p \right)^{1/p}.$$

Remark 9. *Under the assumptions of Proposition 7, we also have that for any integer n,*

$$\left\| \max_{1 \leq k \leq n} |S_k|_{\mathbb{B}} \right\|_p \leq 2q \max_{1 \leq k \leq n} \|S_k\|_p + a_p n^{1/p} \sum_{\ell=1}^{n} \frac{\|\mathbb{E}_0(S_\ell)\|_p}{\ell^{1+1/p}} + b_p n^{1/p} \sum_{\ell=1}^{2n} \frac{\|S_\ell - \mathbb{E}_\ell(S_\ell)\|_p}{\ell^{1+1/p}}, \qquad (3.9)$$

where

$$a_p = \frac{2^{1+1/p}q}{1 - 2^{-1-1/p}} \quad and \quad b_p = 2(q+1)\frac{2^{1+1/p}}{1 - 2^{-1-1/p}}.$$

The proof of this remark will be done at the end of this section.

In the next results, we consider the case where $(\mathbb{B}, |\cdot|_{\mathbb{B}}) = (\mathbb{R}, |\cdot|)$. The next inequality is the non adapted version of the Rosenthal type inequality given in [12] (see their Theorem 6).

Theorem 10. *Let $p > 2$ be a real number and q be its conjugate exponent. Let X_0 be a real-valed random variable in $\mathbb{L}^p$ and $\mathcal{F}_0$ a σ-algebra satisfying $\mathcal{F}_0 \subseteq T^{-1}(\mathcal{F}_0)$. Then, for any positive integer r, the following inequality holds:*

$$\mathbb{E}\left(\max_{1 \leq j \leq 2^r} |S_j|^p \right) \ll 2^r \mathbb{E}(|X_0|)^p + 2^r \left(\sum_{k=0}^{r-1} \frac{\|\mathbb{E}_0(S_{2^k})\|_p}{2^{k/p}} \right)^p + 2^r \left(\sum_{k=0}^{r} \frac{\|S_{2^k} - \mathbb{E}_{2^k}(S_{2^k})\|_p}{2^{k/p}} \right)^p$$

$$+ 2^r \left(\sum_{k=0}^{r-1} \frac{\|\mathbb{E}_0(S_{2^k}^2)\|_{p/2}^\delta}{2^{2\delta k/p}} \right)^{p/(2\delta)}, \quad (3.10)$$

where $\delta = \min(1, 1/(p-2))$.

Remark 11. *The inequality in the above theorem implies that for any positive integer n,*

$$\mathbb{E}\left(\max_{1 \leq j \leq n} |S_j|^p \right) \ll n\mathbb{E}(|X_1|)^p + n \left(\sum_{k=1}^{n} \frac{1}{k^{1+1/p}} \|\mathbb{E}_0(S_k)\|_p \right)^p + n \left(\sum_{k=1}^{2n} \frac{1}{k^{1+1/p}} \|S_k - \mathbb{E}_k(S_k)\|_p \right)^p$$

$$+ n \left(\sum_{k=1}^{n} \frac{1}{k^{1+2\delta/p}} \|\mathbb{E}_0(S_k^2)\|_{p/2}^\delta \right)^{p/(2\delta)}.$$

To see this, it suffices to use the arguments developed in the proof of Remark 9 together with the following additional subadditivity property: for any integers i and j, and any $\delta \in]0, 1]$:

$$\|\mathbb{E}_0(S_{i+j}^2)\|_{p/2}^\delta \leq 2^\delta \|\mathbb{E}_0(S_i^2)\|_{p/2} + 2^\delta \|\mathbb{E}_0(S_j^2)\|_{p/2}.$$

So, according to the first item of Lemma 37 of [12], for any integer $n \in]2^{r-1}, 2^r]$,

$$\sum_{k=0}^{r-1} \frac{\|\mathbb{E}_0(S_{2^k}^2)\|_{p/2}^\delta}{2^{2\delta k/p}} \ll \sum_{k=1}^{n} \frac{1}{k^{1+2\delta/p}} \|\mathbb{E}_0(S_k^2)\|_{p/2}^\delta.$$

9

Remark 12. *Theorem 10 has been stated in the real case. Notice that if we assume X_0 to be in $\mathbb{L}^p(\mathbb{B})$ where $(\mathbb{B}, |\cdot|_{\mathbb{B}})$ is a separable Banach space and p is a real in $]2, \infty[$, then a Rosenthal-type inequality similar as (3.10) can be obtained but with a different δ for $2 < p < 4$. To be more precise, we get that*

$$\mathbb{E}\Big(\max_{1 \leq j \leq 2^r} |S_j|_{\mathbb{B}}^p\Big) \ll 2^r \mathbb{E}(|X_0|_{\mathbb{B}})^p + 2^r \left(\sum_{k=0}^{r} \frac{\|S_{2^k} - \mathbb{E}_{2^k}(S_{2^k})\|_p}{2^{k/p}}\right)^p + 2^r \left(\sum_{k=0}^{r-1} \frac{\|\mathbb{E}_0(|S_{2^k}|_{\mathbb{B}}^2)\|_{p/2}^{\delta}}{2^{2\delta k/p}}\right)^{p/(2\delta)},$$
$$(3.11)$$

where $\delta = \min(1/2, 1/(p-2))$. The proof of this inequality is given later.

As a consequence of (3.10), one can prove the following inequality. This inequality will be used to prove the tightness of the sequential empirical process (1.2) in the space $\ell^\infty([0, 1] \times \mathbb{R}^\ell)$ (see the proof of Theorem 4, Section 5).

Proposition 13. *Let $p > 2$. Let X_0 be a real-valed random variable in $\mathbb{L}^p$ and $\mathcal{F}_0$ a σ-algebra satisfying $\mathcal{F}_0 \subseteq T^{-1}(\mathcal{F}_0)$. For any $j \geq 1$, let*

$$A(X, j) = \max\left(2 \sup_{i \geq 0} \|\mathbb{E}_0(X_i X_{j+i})\|_{p/2}, \sup_{0 \leq i \leq j} \|\mathbb{E}_0(X_j X_{j+i}) - \mathbb{E}(X_j X_{j+i})\|_{p/2}\right). \quad (3.12)$$

Then for every positive integer n,

$$\Big\|\max_{1 \leq j \leq n} |S_j|\Big\|_p \ll n^{1/2}\Big(\sum_{k=0}^{n-1} |\mathbb{E}(X_0 X_k)|\Big)^{1/2} + n^{1/p}\|X_1\|_p + n^{1/p}\sum_{k=1}^{n} \frac{1}{k^{1/p}}\|\mathbb{E}_0(X_k)\|_p$$
$$+ n^{1/p}\sum_{k=1}^{2n} \frac{1}{k^{1/p}}\|X_0 - \mathbb{E}_k(X_0)\|_p + n^{1/p}\Big(\sum_{k=1}^{n} \frac{1}{k^{(2/p)-1}}(\log k)^\gamma A(X, k)\Big)^{1/2}.$$

where γ can be taken $\gamma = 0$ for $2 < p \leq 3$ and $\gamma > p - 3$ for $p > 3$. The constant that is implicitly involved in the notation $\ll$ depends on p and γ but it depends neither on n nor on the X_i's.

The proof of this proposition is left to the reader since it uses the same arguments as those developed for the proof of Proposition 20 in [12].

We would like also to pint out that Theorem 10 implies the following Burkholder-type inequality. This has been already mentioned in the adapted case in [12, Corollary 13].

Corollary 14. *Let $p > 2$ be a real number, X_0 be a real random variable in $\mathbb{L}^p$ and $\mathcal{F}_0$ a σ-algebra satisfying $\mathcal{F}_0 \subseteq T^{-1}(\mathcal{F}_0)$. Then, for any integer r, the following inequality holds*

$$\mathbb{E}\big(\max_{1 \leq j \leq 2^r} |S_j|^p\big) \ll 2^{rp/2}\mathbb{E}(|X_0|^p) + 2^{rp/2}\Big(\sum_{j=0}^{r-1} \frac{\|\mathbb{E}_0(S_{2^j})\|_p}{2^{j/2}}\Big)^p + 2^{rp/2}\Big(\sum_{j=1}^{r} \frac{\|S_{2^j} - \mathbb{E}_{2^j}(S_{2^j})\|_p}{2^{j/2}}\Big)^p.$$

The above corollary (up to constants) is then the non adapted version of [13, Theorem 1] when $p > 2$. Let us give an application of it for the partial sums associated to continuous functions of the iterates of an ergodic automorphism of the torus. Let T be a ergodic automorphism of $\mathbb{T}^d$ as defined in the introduction. Let f be a continuous function from $\mathbb{T}^d$ to $\mathbb{R}$ with modulus of continuity $\omega(f, \cdot)$ (see Section 4, equation (4.1), for the definition). Using Corollary 14 together with Theorem 18 (see also Remark 19), we infer that if

$$\int_0^{1/2} \frac{\omega(f, t)}{t|\log t|^{1/2}} dt < \infty,$$

then for any $p > 2$,

$$\Big\|\max_{1 \le k \le n} \Big| \sum_{i=1}^{k} (f \circ T^i - \bar{\lambda}(f)) \Big| \Big\|_p \ll n^{1/2} \,.$$

Proof of Proposition 7. For any $k \in \{1, \ldots, 2^r\}$, we have that

$$S_k = S_k - \mathbb{E}_k(S_k) + \mathbb{E}_k(S_{2^r}) - \mathbb{E}_k(S_{2^r} - S_k) \,.$$

Consequently

$$\Big\| \max_{1 \le k \le 2^r} |S_k|_{\mathbb{B}} \Big\|_p \le \Big\| \max_{1 \le k \le 2^r} |\mathbb{E}_k(S_{2^r})|_{\mathbb{B}} \Big\|_p + \Big\| \max_{1 \le m \le 2^r - 1} |\mathbb{E}_{2^r - m}(S_{2^r} - S_{2^r - m})|_{\mathbb{B}} \Big\|_p$$
$$+ \|S_{2^r} - \mathbb{E}_{2^r}(S_{2^r})\|_p + \Big\| \max_{1 \le m \le 2^r - 1} |S_m - \mathbb{E}_m(S_m)|_{\mathbb{B}} \Big\|_p \,. \qquad (3.13)$$

Following the proof of Proposition 2 in [12], we get that

$$\Big\| \max_{1 \le k \le 2^r} |\mathbb{E}_k(S_{2^r})|_{\mathbb{B}} \Big\|_p + \Big\| \max_{1 \le m \le 2^r - 1} |\mathbb{E}_{2^r - m}(S_{2^r} - S_{2^r - m})|_{\mathbb{B}} \Big\|_p$$
$$\le q \, \|\mathbb{E}(S_{2^r} | \mathcal{F}_{2^r})\|_p + q \sum_{\ell = 0}^{r-1} \Big(\sum_{k=1}^{2^{r-\ell} - 1} \|\mathbb{E}_{k2^\ell}(S_{(k+1)2^\ell} - S_{k2^\ell})\|_p^p \Big)^{1/p} \,.$$

So, by stationarity,

$$\Big\| \max_{1 \le k \le 2^r} |\mathbb{E}_k(S_{2^r})|_{\mathbb{B}} \Big\|_p + \Big\| \max_{1 \le m \le 2^r - 1} |\mathbb{E}_{2^r - m}(S_{2^r} - S_{2^r - m})|_{\mathbb{B}} \Big\|_p$$
$$\le q \, \|\mathbb{E}(S_{2^r} | \mathcal{F}_{2^r})\|_p + q 2^{r/p} \sum_{\ell = 0}^{r-1} 2^{-\ell/p} \|\mathbb{E}(S_{2^\ell} | \mathcal{F}_0)\|_p \,. \qquad (3.14)$$

We now bound the last term in the right hand side of (3.13). For any $m \in \{1, \ldots, 2^r - 1\}$, we consider its binary expansion:

$$m = \sum_{i=0}^{r-1} b_i(m) 2^i, \quad \text{where } b_i(m) = 0 \text{ or } b_i(m) = 1 \,.$$

Set $m_l = \sum_{i=l}^{r-1} b_i(m) 2^i$, and write that for any $p \ge 1$,

$$|S_m - \mathbb{E}_m(S_m)|_{\mathbb{B}} \le \sum_{l=0}^{r-1} |S_{m_l} - S_{m_{l+1}} - \mathbb{E}_m(S_{m_l} - S_{m_{l+1}})|_{\mathbb{B}} \,, \qquad (3.15)$$

since $S_0 = 0$ and $m_r = 0$. Now, since for any $l = 0, \ldots, r - 1$, $\mathcal{F}_{m_l} \subseteq \mathcal{F}_m$, the following decomposition holds:

$$|S_{m_l} - S_{m_{l+1}} - \mathbb{E}_m(S_{m_l} - S_{m_{l+1}})|_{\mathbb{B}} \le |S_{m_l} - S_{m_{l+1}} - \mathbb{E}_{m_l}(S_{m_l} - S_{m_{l+1}})|_{\mathbb{B}}$$
$$+ \big| \mathbb{E}\big(S_{m_l} - S_{m_{l+1}} - \mathbb{E}_{m_l}(S_{m_l} - S_{m_{l+1}}) | \mathcal{F}_m \big) \big|_{\mathbb{B}} \,.$$

Notice that $m_l \ne m_{l+1}$ only if $m_l = k_{m,l} 2^l$ with $k_{m,l}$ odd. Then, setting

$$B_{r,l} = \max_{1 \le k \le 2^{r-l}, k \text{ odd}} |S_{k2^l} - S_{(k-1)2^l} - \mathbb{E}_{k2^l}(S_{k2^l} - S_{(k-1)2^l})|_{\mathbb{B}} \,,$$

11

it follows that
$$|S_{m_l} - S_{m_{l+1}} - \mathbb{E}_m(S_{m_l} - S_{m_{l+1}})|_{\mathbb{B}} \leq B_{r,l} + |\mathbb{E}(B_{r,l}|\mathcal{F}_m)| \,.$$
Starting from (3.15), we then get that

$$\left\| \max_{1 \leq m \leq 2^r - 1} |S_m - \mathbb{E}_m(S_m)|_{\mathbb{B}} \right\|_p \leq \sum_{l=0}^{r-1} \|B_{r,l}\|_p + \sum_{l=0}^{r-1} \left\| \max_{1 \leq m \leq 2^r - 1} |\mathbb{E}(B_{r,l}|\mathcal{F}_m)| \right\|_p \,.$$

Since $(\mathbb{E}(B_{r,l}|\mathcal{F}_m))_{m \geq 1}$ is a martingale, by using Doob's maximal inequality, we get that

$$\left\| \max_{1 \leq m \leq 2^r - 1} |\mathbb{E}(B_{r,l}|\mathcal{F}_m)| \right\|_p \leq q\|\mathbb{E}(B_{r,l}|\mathcal{F}_{2^r - 1})\|_p \leq q\|B_{r,l}\|_p \,,$$

yielding to

$$\left\| \max_{1 \leq m \leq 2^r - 1} |S_m - \mathbb{E}_m(S_m)|_{\mathbb{B}} \right\|_p \leq (q+1) \sum_{l=0}^{r-1} \|B_{r,l}\|_p \,.$$

Since

$$B_{r,l} \leq \left(\sum_{k=1}^{2^{r-l}-1} |S_{k2^l} - S_{(k-1)2^l} - \mathbb{E}_{k2^l}(S_{k2^l} - S_{(k-1)2^l})|_{\mathbb{B}}^p \right)^{1/p},$$

we derive that

$$\|S_{2^r} - \mathbb{E}_{2^r}(S_{2^r})\|_p + \left\| \max_{1 \leq m \leq 2^r - 1} |S_m - \mathbb{E}_m(S_m)|_{\mathbb{B}} \right\|_p$$

$$\leq (q+1) \sum_{l=0}^{r-1} \left(\sum_{k=1}^{2^{r-l}-1} \|S_{k2^l} - S_{(k-1)2^l} - \mathbb{E}_{k2^l}(S_{k2^l} - S_{(k-1)2^l})\|_p^p \right)^{1/p} \,.$$

So, by stationarity,

$$\left\| \max_{1 \leq m \leq 2^r - 1} |S_m - \mathbb{E}_m(S_m)|_{\mathbb{B}} \right\|_q \leq (q+1) 2^{r/p} \sum_{l=0}^{r-1} 2^{-l/p} \|S_{2^l} - \mathbb{E}_{2^l}(S_{2^l})\|_p \,. \tag{3.16}$$

Starting from (3.13) and taking into account (3.14) and (3.16), the inequality (3.8) follows.
$\square$

Proof of Theorem 10. Thanks to Proposition 7, it suffices to prove that the inequality (3.10) is satisfied for $\mathbb{E}(|S_{2^r}|^p)$ instead of $\mathbb{E}(\max_{1 \leq j \leq 2^r} |S_j|^p)$. Let $a_n = \|S_n\|_p$. According to the proof of Lemma 11 in [12], the theorem will follow if we can prove the following recurrence formula: for any positive integer n,
$$a_{2n}^p \leq 2a_n^p + c_1 a_n^{p-1} \big(\|\mathbb{E}_0(S_n)\|_p + \|S_n - \mathbb{E}_n(S_n)\|_p \big) + c_2 a_n^{p-2\delta} \|\mathbb{E}_0(S_n^2)\|_{p/2}^\delta \,. \tag{3.17}$$
where c_1 and c_2 are positive constants depending only on p. To prove (3.17), we denote by $\bar{S}_n = X_{n+1} + \cdots + X_{2n}$, and we write that
$$S_{2n} = S_n - \mathbb{E}_n(S_n) + \mathbb{E}_n(S_n) + \bar{S}_n \,.$$

Recall first the following algebraic inequality: Let x and y be two positive real numbers and $p \geq 1$ any real number. Then
$$(x+y)^p \leq x^p + y^p + 4^p(x^{p-1}y + xy^{p-1}) \,. \tag{3.18}$$

12

(see Inequality (87) in [12]). The above inequality with $x = |\mathbb{E}_n(S_n) + \bar{S}_n|$ and $y = |S_n - \mathbb{E}_n(S_n)|$ gives

$$a_{2n}^p \leq \|\mathbb{E}_n(S_n) + \bar{S}_n\|_p^p + \|S_n - \mathbb{E}_n(S_n)\|_p^p$$
$$+ 4^p \mathbb{E}\big(|\mathbb{E}_n(S_n) + \bar{S}_n|^{p-1}|S_n - \mathbb{E}_n(S_n)|\big) + 4^p \mathbb{E}\big(|\mathbb{E}_n(S_n) + \bar{S}_n||S_n - \mathbb{E}_n(S_n)|^{p-1}\big),$$

which combined with Hölder's inequality and stationarity leads to

$$a_{2n}^p \leq \|\mathbb{E}_n(S_n) + \bar{S}_n\|_p^p + 2^{p-1}(1 + 2^{2p+1})a_n^{p-1}\|S_n - \mathbb{E}_n(S_n)\|_p. \qquad (3.19)$$

Starting from (3.19), (3.17) will follow if we can prove that there exist two positive constants c and c_2 depending only on p such that

$$\|\mathbb{E}_n(S_n) + \bar{S}_n\|_p^p \leq 2a_n^p + c\,a_n^{p-1}\|\mathbb{E}_0(S_n)\|_p + c_2 a_n^{p-2\delta}\|\mathbb{E}_0(S_n^2)\|_{p/2}^\delta.$$

This inequality can be proven by following the lines of the end of the proof of Theorem 6 in [12]. Indeed, it suffices to replace in their proof, $x = S_n$ by $x = \mathbb{E}_n(S_n)$, and to use the following estimates (coming from the proof of their lemma 34 and the stationarity assumption): for any reals p and u such that $0 \leq u \leq p - 2$,

$$\mathbb{E}(|\mathbb{E}_n(S_n)|^u|\bar{S}_n|^{p-u}) \leq a_n^{p-2u/(p-2)}\|\mathbb{E}_n(\bar{S}_n^2)\|_{p/2}^{u/(p-2)},$$

and

$$\mathbb{E}(|\mathbb{E}_n(S_n)|^{p-1}|\bar{S}_n|) \leq a_n^{p-1}\|\mathbb{E}_n(\bar{S}_n^2)\|_{p/2}^{1/2}.$$

$\qquad\qquad\square$

Proof of Remark 12. As it is pointed out in the proof of Theorem 10, the remark will be proven with the help of Proposition 7, if we can show that

$$a_{2n}^p \leq 2a_n^p + c_1 a_n^{p-1}\|S_n - \mathbb{E}_n(S_n)\|_p + c_2 a_n^{p-2\delta}\|\mathbb{E}_0(|S_n|_\mathbb{B}^2)\|_{p/2}^\delta,$$

where $a_n^p = \mathbb{E}(|S_n|_\mathbb{B}^p)$, c_1 and c_2 are positive constants depending only on p and $\delta = \min(1/2, 1/(p-2))$. Indeed, the second term in the right-hand side of (3.8) can be bounded by the last term in the right-hand side of (3.11). To see this it suffices to use Jensen's inequality and the fact that $\delta \leq 1/2$.

Starting from (3.19) (by replacing the absolute values by the norm $|\cdot|_\mathbb{B}$), we see that to prove the above recurrence formula it suffices to show that there exists a positive constant c depending only on p such that

$$\|\mathbb{E}_n(S_n) + \bar{S}_n\|_p^p \leq 2a_n^p + ca_n^{p-2\delta}\|\mathbb{E}_0(|S_n|_\mathbb{B}^2)\|_{p/2}^\delta.$$

The difference at this step with the proof of Theorem 10 is that the inequality (3.18) is used whatever $p \in]2, \infty[$ (in Theorem 10, so in the real case, when $p \in]2, 4[$ more precise inequalities can be used as done in the proof of Theorem 6 in [12]). $\qquad\square$

Proof of Corollary 14. To prove the corollary, it suffices to show that for any $0 < \delta \leq 1$ and any real $p > 2$,

$$2^r \left(\sum_{k=0}^{r-1} \frac{\|\mathbb{E}_0(S_{2^k}^2)\|_{p/2}^\delta}{2^{2\delta k/p}}\right)^{p/(2\delta)} \ll 2^{rp/2}\|\mathbb{E}_0(X_1^2)\|_{p/2}^{p/2} + 2^{rp/2}\Big(\sum_{j=0}^{r-1} \frac{\|\mathbb{E}_0(S_{2^j})\|_p}{2^{j/2}}\Big)^p, \qquad (3.20)$$

and to apply Theorem 10. Following the proof of Lemma 12 in [12] and setting $b_n = \|\mathbb{E}_0(S_n^2)\|_{p/2}$, we infer that (3.20) will follow if we can prove that, for any integer n,

$$b_{2n} \leq 2b_n + 2b_n^{1/2}\|\mathbb{E}_0(S_n)\|_p + 2b_n^{1/2}\|S_n - \mathbb{E}_n(S_n)\|_p. \qquad (3.21)$$

13

By using the notation $\bar{S}_n = X_{n+1} + \cdots + X_n$ and the fact that $S_{2n}^2 = S_n^2 + \bar{S}_n^2 + 2\mathbb{E}_n(S_n)\bar{S}_n + 2(S_n - \mathbb{E}_n(S_n))\bar{S}_n$, we get, by stationarity, that

$$b_{2n} \leq 2b_n + 2\|\mathbb{E}_0\big(\mathbb{E}_n(S_n)\mathbb{E}_n(\bar{S}_n)\big)\|_{p/2} + 2\|\mathbb{E}_0\big((S_n - \mathbb{E}_n(S_n))\bar{S}_n\big)\|_{p/2}\,.$$

Hence the inequality (3.21) follows from the following upper bounds: Applying Cauchy-Schwarz inequality twice and using stationarity, we get

$$\|\mathbb{E}_0\big(\mathbb{E}_n(S_n)\mathbb{E}_n(\bar{S}_n)\big)\|_{p/2} \leq \|\mathbb{E}_0(\mathbb{E}_n^2(S_n))\|_{p/2}^{1/2} \times \|\mathbb{E}_0(\mathbb{E}_n^2(\bar{S}_n))\|_{p/2}^{1/2}$$
$$\leq \|\mathbb{E}_0(S_n^2))\|_{p/2}^{1/2} \times \|\mathbb{E}_n^2(\bar{S}_n)\|_{p/2}^{1/2} \leq b_n^{1/2}\|\mathbb{E}_0(S_n)\|_p\,,$$

and

$$\|\mathbb{E}_0\big((S_n - \mathbb{E}_n(S_n))\bar{S}_n\big)\|_{p/2} \leq \|\mathbb{E}_0(((S_n - \mathbb{E}_n(S_n))^2)\|_{p/2}^{1/2}\|\mathbb{E}_0(\bar{S}_n^2)\|_{p/2}^{1/2}$$
$$\leq b_n^{1/2}\|S_n - \mathbb{E}_n(S_n)\|_p\,.$$

$\square$

Proof of Remark 9. Let n and r be integers such that $2^{r-1} \leq n < 2^r$. Notice first that

$$\Big\|\max_{1\leq k\leq n}|S_k|_\mathbb{B}\Big\|_p \leq \Big\|\max_{1\leq k\leq 2^r}|S_m|_\mathbb{B}\Big\|_p \quad \text{and} \quad \|S_{2^r}\|_p \leq 2\|S_{2^{r-1}}\|_p \leq 2\max_{1\leq k\leq n}\|S_k\|_p \tag{3.22}$$

(for the second inequality we use the stationarity). Now, setting $V_m = \|\mathbb{E}_0(S_m)\|_p$, we have by stationarity that for all $n, m \geq 0$, $V_{n+m} \leq V_n + V_m$ and then, according to the first item of Lemma 37 of [12],

$$2^{r/p}\sum_{\ell=0}^{r-1} 2^{-\ell/p}\|\mathbb{E}_0(S_{2^\ell})\|_p \leq n^{1/p}\frac{2^{1/p}2^{2+1/p}}{2^{1+1/p}-1}\sum_{k=1}^n \frac{\|\mathbb{E}_0(S_k)\|_p}{k^{1+1/p}}$$
$$\leq n^{1/p}\frac{2^{1+1/p}}{1-2^{-1/p-1}}\sum_{k=1}^n \frac{\|\mathbb{E}_0(S_k)\|_p}{k^{1+1/p}}\,. \tag{3.23}$$

On an other hand, let $W_m = \|S_m - \mathbb{E}_m(S_m)\|_p$, and note that the following claim is valid:

Claim 15. *If $\mathcal{F}$ and $\mathcal{G}$ are σ-algebras such that $\mathcal{G} \subset \mathcal{F}$, then for any X in $\mathbb{L}^p(\mathbb{B})$ where $p \geq 1$, $\|X - \mathbb{E}(X|\mathcal{F})\|_p \leq 2\|X - \mathbb{E}(X|\mathcal{G})\|_p$.*

The above claim together with the stationarity imply that for all $n, m \geq 0$, $W_{n+m} \leq 2(W_n + W_m)$. Therefore, using once again the first item of Lemma 37 of [12], we get that

$$2^{r/p}\sum_{\ell=0}^r 2^{-\ell/p}\|S_{2^\ell} - \mathbb{E}_{2^\ell}(S_{2^\ell})\|_p \leq 2n^{1/p}\frac{2^{1+1/p}}{1-2^{-1/p-1}}\sum_{\ell=1}^{2n} \frac{\|S_\ell - \mathbb{E}_\ell(S_\ell)\|_p}{\ell^{1+1/p}}\,. \tag{3.24}$$

The inequality (3.9) then follows from the inequality (3.8) by taking into account the upper bounds (3.22), (3.23) and (3.24). $\square$

3.2 A tightness criterion

We begin with the definition of the number of brackets of a family of functions.

Definition 16. *Let P be a probability measure on a measurable space $\mathcal{X}$. For any measurable function f from $\mathcal{X}$ to $\mathbb{R}$, let $\|f\|_{P,1} = P(|f|)$. If $\|f\|_{P,1}$ is finite, one says that f belongs to L_P^1. Let $\mathcal{F}$ be some subset of L_P^1. The number of brackets $\mathcal{N}_{P,1}(\varepsilon, \mathcal{F})$ is the smallest integer N for which there exist some functions $f_1^- \leq f_1, \ldots, f_N^- \leq f_N$ in $\mathcal{F}$ such that: for any integer $1 \leq i \leq N$ we have $\|f_i - f_i^-\|_{P,1} \leq \varepsilon$, and for any function f in $\mathcal{F}$ there exists an integer $1 \leq i \leq N$ such that $f_i^- \leq f \leq f_i$.*

The first step in the proof of Theorem 4 is the following proposition, whose proof is based on a decomposition given in [1] (see also [5]).

Proposition 17. *Let $(X_i)_{i \geq 1}$ be a sequence of identically distributed random variables with values in a measurable space $\mathcal{X}$, with common distribution P. Let P_n be the empirical measure $P_n = n^{-1} \sum_{i=1}^n \delta_{X_i}$, and let S_n be the empirical process $S_n = n(P_n - P)$. Let $\mathcal{F}$ be a class of functions from $\mathcal{X}$ to $\mathbb{R}$ and $\mathcal{G} = \{f - l, (f, l) \in \mathcal{F} \times \mathcal{F}\}$. Assume that there exist $r \geq 2$ and $p > 2$ such that for any function g of $\mathcal{G} \cup \mathcal{F}$, we have*

$$\left\| \max_{1 \leq k \leq n} |S_k(g)| \right\|_p \leq C(\sqrt{n} \|g\|_{P,1}^{1/r} + n^{1/p}), \tag{3.25}$$

where the constant C does not depend on g nor on n. If moreover

$$\int_0^1 x^{(1-r)/r} (\mathcal{N}_{P,1}(x, \mathcal{F}))^{1/p} dx < \infty \quad and \quad \lim_{x \to 0} x^{p-2} \mathcal{N}_{P,1}(x, \mathcal{F}) = 0,$$

then

$$\lim_{\delta \to 0} \limsup_{n \to \infty} \mathbb{E} \Big(\max_{1 \leq k \leq n} \sup_{g \in \mathcal{G}, \|g\|_{P,1} \leq \delta} n^{-p/2} |S_k(g)|^p \Big) = 0, \tag{3.26}$$

$$and \quad \lim_{\delta \to 0} \limsup_{n \to \infty} \frac{1}{\delta} \mathbb{E} \Big(\max_{1 \leq k \leq [n\delta]} \sup_{f \in \mathcal{F}} n^{-p/2} |S_k(f)|^p \Big) = 0. \tag{3.27}$$

Proof of Proposition 17. It is almost the same as that of Proposition 6 in [5]. Let us only give the main steps.

For any positive integer k, denote by $\mathcal{N}_k = \mathcal{N}_{P,1}(2^{-k}, \mathcal{F})$ and by $\mathcal{F}_k$ a family of functions $f_1^{k,-} \leq f_1^k, \ldots, f_{\mathcal{N}_k}^{k,-} \leq f_{\mathcal{N}_k}^k$ in $\mathcal{F}$ such that $\|f_i^k - f_i^{k,-}\|_{P,1} \leq 2^{-k}$, and for any f in $\mathcal{F}$, there exists an integer $1 \leq i \leq \mathcal{N}_k$ such that $f_i^{k,-} \leq f \leq f_i^k$.

We follow exactly the proof of Proposition 6 in [5]. One can prove that for any $\varepsilon > 0$, there exist $N(\varepsilon)$ and $m = m(\varepsilon)$ such that: for any $n \geq N(\varepsilon)$ there exists $f_{n,m}$ in $\mathcal{F}_m$ for which

$$\left\| \max_{1 \leq k \leq n} \sup_{f \in \mathcal{F}} n^{-1/2} |S_k(f) - S_k(f_{n,m})| \right\|_p \leq \varepsilon. \tag{3.28}$$

Now, (3.26) follows from (3.28) as in [1] (see the end of the proof of Proposition 6 in [5]).

Let us prove (3.27). We apply (3.28) with $\varepsilon = 1$: for $n \geq \delta^{-1} N(1)$, we infer from (3.28) that there exists $f_{[n\delta],m}$ in $\mathcal{F}_m$ for which

$$\left\| \max_{1 \leq k \leq [n\delta]} \sup_{f \in \mathcal{F}} n^{-1/2} |S_k(f) - S_k(f_{[n\delta],m})| \right\|_p \leq \sqrt{\delta}.$$

Hence

$$\left\| \max_{1 \leq k \leq [n\delta]} \sup_{f \in \mathcal{F}} n^{-1/2} |S_k(f)| \right\|_p \leq \sqrt{\delta} + \left\| \max_{1 \leq k \leq [n\delta]} \sup_{f \in \mathcal{F}} n^{-1/2} |S_k(f_{[n\delta],m})| \right\|_p. \tag{3.29}$$

15

Now, since $\mathcal{F}_m$ contains $2\mathcal{N}_m$ functions $(g_\ell)_{\ell \in 2\mathcal{N}_m}$ (each g_ℓ being one of the functions f_i^m or $f_i^{m,-}$ in $\mathcal{N}_m$), it follows that

$$\left\| \max_{1 \le k \le [n\delta]} \sup_{f \in \mathcal{F}} n^{-1/2} |S_k(f_{[n\delta],m})| \right\|_p \le \sum_{\ell=1}^{2\mathcal{N}_m} \frac{1}{\sqrt{n}} \left\| \max_{1 \le k \le [n\delta]} |S_k(g_\ell)| \right\|_p.$$

Let $K_m = \max_{f \in \mathcal{F}_m} \|f\|_{P,1}$. Applying (3.25), we infer that

$$\left\| \max_{1 \le k \le [n\delta]} \sup_{f \in \mathcal{F}} n^{-1/2} |S_k(f_{[n\delta],m})| \right\|_p \le 2C\mathcal{N}_m (K_m^{1/r} \sqrt{\delta} + n^{-(p-2)/2p} \delta^{1/p}). \tag{3.30}$$

Since $m = m(1)$ is fixed, (3.27) follows from (3.29) and (3.30) and the fact that $p > 2$. $\qquad\square$

4   Inequalities for ergodic torus automorphisms

In this section, we keep the same notations as in the introduction. Let us denote by E_u, E_e and E_s the S-stable vector spaces associated to the eigenvalues of S of modulus respectively larger than one, equal to one and smaller than one. Let d_u, d_e and d_s be their respective dimensions. Let $v_1, ..., v_d$ be a basis of $\mathbb{R}^d$ such that $v_1, ..., v_{d_u}$ are in E_u, $v_{d_u+1}, ..., v_{d_u+d_e}$ are in E_e and $v_{d_u+d_e+1}, ..., v_d$ are in E_s. We suppose moreover that $\det(v_1|v_2|\cdots|v_d) = 1$. Let $||\cdot||$ be the norm on $\mathbb{R}^d$ given by

$$\left\| \sum_{i=1}^d x_i v_i \right\| = \max_{i=1,...,d} |x_i|$$

and $d_0(\cdot, \cdot)$ the metric induced by $||\cdot||$ on $\mathbb{R}^d$. Let also d_1 be the metric induced by d_0 on $\mathbb{T}^d$ namely,

$$d_1(\bar{x}, \bar{y}) = \inf_{z \in \mathbb{Z}^d} d_0(x + z, y).$$

We define now $B_u(\delta) := \{y \in E_u \ : \ ||y|| \le \delta\}$, $B_e(\delta) := \{y \in E_e \ : \ ||y|| \le \delta\}$ and $B_s(\delta) = \{y \in E_s \ : \ ||y|| \le \delta\}$. For every $f : \mathbb{T}^d \to \mathbb{R}$, we consider the moduli of continuity defined, for every $\delta > 0$, by

$$\omega(f, \delta) := \sup_{\bar{x}, \bar{y} \in \mathbb{T}^d \, : \, d_1(\bar{x}, \bar{y}) \le \delta} |f(\bar{x}) - f(\bar{y})|, \tag{4.1}$$

$$\omega_{(s,e)}(f, \delta) = \sup\{|f(\bar{x}) - f(\bar{x} + \overline{h_s} + \overline{h_e})|, \ \bar{x} \in \mathbb{T}^d, \ h_s \in B_s(\delta), \ h_e \in B_e(\delta)\}$$

and

$$\omega_{(u)}(f, \delta) = \sup\{|f(\bar{x}) - f(\bar{x} + \overline{h_u})|, \ \bar{x} \in \mathbb{T}^d, \ h_u \in B_u(\delta)\}.$$

Let r_u be the spectral radius of $S_{|E_u}^{-1}$. For every $\rho_u \in (r_u, 1)$, there exists $K > 0$ such that, for every integer $n \ge 0$, we have

$$\forall h_u \in E_u, \ \ ||S^{-n} h_u|| \le K\rho_u^n ||h_u|| \tag{4.2}$$

and

$$\forall (h_e, h_s) \in E_e \times E_s, \ \ ||S^n(h_e + h_s)|| \le K n^{d_e} ||h_e + h_s||. \tag{4.3}$$

The following inequality is an extension to continuous functions of a result for Hölder functions established in [10].

16

Theorem 18. *Let $\rho_u \in (r_u, 1)$ and $\zeta \in (\rho_u^{1/(3(d+2)(d_e+d_s))}, 1)$. There exist $C > 0$, $N \geq 0$, $\xi \in (0, 1)$, a sequence of measurable sets $(\mathcal{V}_n)_{n \geq 0}$ and a σ-algebra $\mathcal{F}_0$ such that $\mathcal{F}_0 \subseteq T^{-1}\mathcal{F}_0$ and such that, for every bounded $\varphi : \mathbb{T}^d \to \mathbb{R}$ and every integer $n \geq N$, we have*

$$\|\mathbb{E}[\varphi|\mathcal{F}_n] - \varphi\|_\infty \leq C\omega_{(u)}(\varphi, \rho_u^n), \tag{4.4}$$

$$on \ \mathcal{V}_n, \quad |\mathbb{E}[\varphi|\mathcal{F}_{-n}] - \mathbb{E}[\varphi]| \leq C(\|\varphi\|_\infty \xi^n + \omega_{(s,e)}(\varphi, \zeta^n)) \tag{4.5}$$

and

$$\bar{\lambda}(\mathbb{T}^d \setminus \mathcal{V}_n) \leq C\xi^n, \tag{4.6}$$

where $\mathcal{F}_k := T^{-k}\mathcal{F}_0$ for every $k \in \mathbb{Z}$.

Remark 19. *With the notations of Theorem 18, (4.5) and (4.6) implies that, for every $p \geq 1$ and every (ρ_u, ζ) as in Theorem 18, there exists c_p such that, for every bounded $\varphi : \mathbb{T}^d \to \mathbb{R}$ and every integer $n \geq 0$, we have*

$$\forall n \geq 0, \quad \|\mathbb{E}[\varphi|\mathcal{F}_{-n}] - \mathbb{E}[\varphi]\|_p \leq c_p(\|\varphi\|_\infty \xi^{\frac{n}{p}} + \omega_{(s,e)}(\varphi, \zeta^n)). \tag{4.7}$$

The remainder of this section is devoted to the proof of Theorem 18 and to the statements and the proofs of some preliminary results. Let $\rho_u \in (r_u, 1)$ and K satisfying (4.2) and (4.3). Let m_u, m_e, m_s be the Lebesgue measure on E_u (in the basis $v_1, ..., v_{d_u}$), E_e (in the basis $v_{d_u+1}, ..., v_{d_u+d_e}$) and E_s (in the basis $v_{d_u+d_e+1}, ..., v_d$) respectively. We observe that $d\lambda(h_u + h_e + h_s) = dm_u(h_u)dm_e(h_e)dm_s(h_s)$.

The properties satisfied by the filtration considered in [11, 9] and enabling the use of Gordin's method will be crucial here. Given a finite partition $\mathcal{P}$ of $\mathbb{T}^d$, we define the measurable partition $\mathcal{P}_0^\infty$ by :

$$\forall \bar{x} \in \mathbb{T}^d, \quad \mathcal{P}_0^\infty(\bar{x}) := \bigcap_{k \geq 0} T^k \mathcal{P}(T^{-k}(\bar{x}))$$

and, for every integer n, the σ-algebra $\mathcal{F}_n$ generated by

$$\forall \bar{x} \in \mathbb{T}^d, \quad \mathcal{P}_{-n}^\infty(\bar{x}) := \bigcap_{k \geq -n} T^k \mathcal{P}(T^{-k}(\bar{x})) = T^{-n}(\mathcal{P}_0^\infty(T^n(\bar{x}))).$$

We obviously have $\mathcal{F}_n = T^{-n}\mathcal{F}_0 \subseteq \mathcal{F}_{n+1} = T^{-1}\mathcal{F}_n$. Let $r_0 > 0$ be such that $(h_u, h_e, h_s) \mapsto \overline{h_u + h_e + h_s}$ defines a diffeomorphism from $B_u(r_0) \times B_e(r_0) \times B_s(r_0)$ on its image in $\mathbb{T}^d$. Observe that, for every $\bar{x} \in \mathbb{T}^d$, on the set $\bar{x} + B_u(r_0) + B_e(r_0) + B_s(r_0)$, we have $d\bar{\lambda}(\bar{x} + \overline{h_u} + \overline{h_e} + \overline{h_s}) = dm_u(h_u)dm_e(h_e)dm_s(h_s)$.

Proposition 20 ([11, 9] applied to T^{-1}, see also [3]). *There exist some $Q > 0$ and some finite partition $\mathcal{P}$ of $\mathbb{T}^d$ whose elements are of the form $\sum_{i=1}^d I_i \overline{v_i}$ where the I_i are intervals with diameter smaller than $\min(r_0, K)$ such that, for almost every $\bar{x} \in \mathbb{T}^d$,*

- *the local leaf $\mathcal{P}_0^\infty(\bar{x})$ of $\mathcal{P}_0^\infty$ containing $\bar{x}$ is a bounded convex set $\bar{x} + \overline{F_{\bar{x}}}$, with $0 \in F_{\bar{x}} \subseteq E_u$, $F_{\bar{x}}$ having non-empty interior in E_u,*

- *we have, for all $n \in \mathbb{Z}$,*

$$\mathbb{E}[f|\mathcal{F}_n](\bar{x}) = \frac{1}{m_u(S^{-n}F_{T^n\bar{x}})} \int_{S^{-n}F_{T^n\bar{x}}} f(\bar{x} + \overline{h_u}) \, dm_u(h_u),$$

- *for every $\gamma > 0$, we have*

$$m_u(\partial(F_{\bar{x}})(\gamma)) \leq Q\gamma,$$

  where

$$\partial \mathcal{C}(\beta) := \{y \in F \ : \ d_0(y, \partial \mathcal{C}) \leq \beta\}.$$

Recall now an exponential decorrelation result for Lipschitz continuous functions.

Proposition 21 ([11] and also section 4.1 of [14]). *Let $\xi_0 \in (\rho_u^{1/3}, 1)$. There exists $C_0 > 0$ such that, for every nonnegative integer n and every Lipschitz continuous functions $f, g : \mathbb{T}^d \to \mathbb{C}$ with $\int_{\mathbb{T}^d} g \, d\bar{\lambda} = 0$, we have*

$$\left| \int_{\mathbb{T}^d} (f.g \circ T^n) \, d\bar{\lambda} \right| \leq C_0 (||f||_\infty ||g||_\infty + ||f||_\infty Lip(g) + ||g||_\infty Lip(f)) \xi_0^n \,,$$

where $Lip(h)$ is the Lipschitz constant of h.

Let Q be the constant appearing in Proposition 20. The following result is an adaptation of Proposition 1.3 of [10].

Proposition 22. *Let $\zeta_1 \in (\rho_u^{1/(3(d+2)(d_e+d_s))}, 1)$. There exist $C_1 > 0$, $N_1 \geq 1$ and $\xi_1 \in (0,1)$ such that, for every $\bar{\lambda}$-centered bounded function $\varphi : \mathbb{T}^d \to \mathbb{R}$, every $\bar{x} \in \mathbb{T}^d$, every $n \geq N_1$ and every bounded convex set $\mathcal{C} \subseteq E_u$ with diameter smaller than r_0, satisfying $m_u(\partial \mathcal{C}(\beta)) \leq Q\beta$ (for every $\beta > 0$), we have*

$$\left| \frac{1}{m_u(S^n \mathcal{C})} \int_{S^n \mathcal{C}} \varphi(\bar{x} + \overline{h_u}) \, dm_u(h_u) \right| \leq K_1 \left(\frac{||\varphi||_\infty \xi_1^n}{m_u(\mathcal{C})} + \omega(\varphi, \zeta_1^n) \right) \,.$$

Proof. Let ξ_0 be as in Proposition 21 with $\zeta_1 > \xi_0^{1/((d+2)(d_e+d_s))}$. Let $r := \xi_0^{-1/(d+2)}$. We take $\varepsilon_n = \alpha^n$ with $\alpha \in (0,1)$ such that $\zeta_1 > \alpha > \xi_0^{1/((d+2)(d_e+d_s))} \geq r^{-1}$. Let $U := T^{-n}\bar{x} + \overline{\mathcal{C} + B_s(\varepsilon_n) + B_e(\varepsilon_n)}$. We have $T^n(U) = \bar{x} + \overline{S^n \mathcal{C} + S^n B_s(\varepsilon_n) + S^n B_e(\varepsilon_n)}$. We have

$$\int_{\mathbb{T}^d} \mathbf{1}_{T^n U}.\varphi \, d\bar{\lambda} = \int_{\mathcal{C} \times B_e(\varepsilon_n) \times B_s(\varepsilon_n)} \varphi(T^n(T^{-n}\bar{x} + \overline{h_u} + \overline{h_e} + \overline{h_s})) \, dm_u(h_u) dm_e(h_e) dm_s(h_s)$$

$$= \int_{\mathcal{V}_n} \varphi(\bar{x} + \overline{h_u} + \overline{h_e} + \overline{h_s}) \, dm_u(h_u) dm_e(h_e) dm_s(h_s) \,,$$

with $\mathcal{V}_n := S^n \mathcal{C} \times S^n B_e(\varepsilon_n) \times S^n B_s(\varepsilon_n)$. Moreover we have

$$\int_{S^n \mathcal{C}} \varphi(\bar{x} + \overline{h_u}) \, dm_u(h_u) = \frac{1}{m_s(S^n(B_s(\varepsilon_n)) m_e(S^n(B_e(\varepsilon_n)))} \int_{\mathcal{V}_n} \varphi(\bar{x} + \overline{h_u}) \, dm_u(h_u) dm_e(h_e) dm_s(h_s) \,.$$

Hence, due to (4.3), we have

$$\left| \int_{\mathbb{T}^d} \mathbf{1}_{T^n U}.\varphi \, d\bar{\lambda} - m_s(S^n(B_s(\varepsilon_n)) m_e(S^n(B_e(\varepsilon_n))) \int_{S^n \mathcal{C}} \varphi(\bar{x} + \overline{h_u}) \, dm_u(h_u) \right| \leq \bar{\lambda}(U) \omega_{(s,e)}(\varphi, K n^{d_e} \varepsilon_n) \,.$$

Since $\bar{\lambda}(U) = m_u(S^n \mathcal{C}) m_s(S^n(B_s(\varepsilon_n)) m_e(S^n(B_e(\varepsilon_n)))$, we get, for n large enough (that is, such that $K n^{d_e} \varepsilon_n \leq \zeta_1^n$),

$$\left| \frac{1}{\bar{\lambda}(U)} \int_{\mathbb{T}^d} \mathbf{1}_{T^n U} \varphi \, d\bar{\lambda} - \frac{1}{m_u(S^n \mathcal{C})} \int_{S^n \mathcal{C}} \varphi(\bar{x} + \overline{h_u}) \, dm_u(h_u) \right| \leq \omega_{(s,e)}(\varphi, K n^{d_e} \varepsilon_n)$$

$$\leq \omega_{(s,e)}(\varphi, \zeta_1^n) \,.$$

For every $n \geq 0$ and $\bar{x} \in \mathbb{T}^d$, we define $\chi_n(\bar{x}) := (d+1) 2^{-d} r^{n(d+1)} d_1(\bar{x}, \mathbb{T}^d \setminus B(0, r^{-n}))$, where $B(0, r^{-n}) = \{\bar{x} \in \mathbb{T}^d, d_1(\bar{0}, \bar{x}) \leq r^{-n}\}$. Let us observe that χ_n is a nonnegative $(d+1) r^{n(d+1)} 2^{-d}$-Lipschitz continuous function supported in $B(0, r^{-n})$, uniformly bounded by $(d+1) 2^{-d} r^{nd}$ and such

that $\int_{\mathbb{T}^d} \chi_n \, d\bar{\lambda} = 1$. We will denote by $*$ the usual convolution product with respect to $\bar{\lambda}$. We will estimate

$$\left| \int_{\mathbb{T}^d} \mathbf{1}_U \circ T^{-n}.\varphi \, d\bar{\lambda} - \int_{\mathbb{T}^d} (\chi_n * \mathbf{1}_U) \circ T^{-n}.(\chi_n * \varphi)) \, d\bar{\lambda} \right|.$$

First observe that

$$\left| \int_{\mathbb{T}^d} (\chi_n * \mathbf{1}_U) \circ T^{-n}.(\chi_n * \varphi - \varphi) \, d\bar{\lambda} \right| \leq \omega(\varphi, r^{-n}) \bar{\lambda}(U). \tag{4.8}$$

Second, we have

$$\left| \int_{\mathbb{T}^d} (\chi_n * \mathbf{1}_U - \mathbf{1}_U) \circ T^{-n}.\varphi \, d\bar{\lambda} \right| \leq \|\varphi\|_\infty \int_{\mathbb{T}^d} |\chi_n * \mathbf{1}_U - \mathbf{1}_U| d\bar{\lambda}, \tag{4.9}$$

and let us prove that

$$\int_{\mathbb{T}^d} |\chi_n * \mathbf{1}_U - \mathbf{1}_U| d\bar{\lambda} \leq 3\bar{\lambda}(\partial U(r^{-n})). \tag{4.10}$$

To see this, observe that $\chi_n(\bar{t})\mathbf{1}_U(\bar{x} - \bar{t}) - \mathbf{1}_U(\bar{x}) = (\chi_n(\bar{t}) - 1)\mathbf{1}_U(\bar{x})$ except if $\mathbf{1}_U(\bar{x} - \bar{t}) \neq \mathbf{1}_U(\bar{x})$ and if $\bar{t} \in B(0, r^{-n})$. Hence $\chi_n * \mathbf{1}_U(\bar{x}) \neq \mathbf{1}_U(\bar{x})$ implies either that $\bar{x} \in \partial U(r^{-n})$ or that $\bar{x}$ belongs to the set U' of points such that $\bar{x} \notin U$ but there exists $\bar{t}_0 \in B(0, r^{-n})$ such that $\bar{x} - \bar{t}_0 \in U$.

On the one hand, we have

$$
\begin{aligned}
\int_{\partial U(r^{-n})} |\chi_n * \mathbf{1}_U - \mathbf{1}_U| d\bar{\lambda} &\leq \int_{\partial U(r^{-n})} \left(\int_{\mathbb{T}^d} \chi_n(\bar{t})\mathbf{1}_U(\bar{x} - \bar{t}) \, d\bar{\lambda}(t) \right) d\bar{\lambda}(x) + \bar{\lambda}(\partial U(r^{-n})) \\
&\leq \bar{\lambda}(\partial U(r^{-n})) \int_{\mathbb{T}^d} \chi_n(\bar{t}) d\bar{\lambda}(t) + \bar{\lambda}(\partial U(r^{-n})) \\
&\leq 2\bar{\lambda}(\partial U(r^{-n})),
\end{aligned}
\tag{4.11}
$$

using the fact that χ_n is nonnegative with unit integral. On the other hand, we have

$$
\begin{aligned}
\int_{U'} |\chi_n * \mathbf{1}_U - \mathbf{1}_U| d\bar{\lambda} &\leq \int_{U'} \left(\int_{\mathbb{T}^d} \chi_n(\bar{t})\mathbf{1}_U(\bar{x} - \bar{t}) \, d\bar{\lambda}(t) \right) d\bar{\lambda}(x) \\
&\leq \int_{\mathbb{T}^d \setminus U} \left(\int_{\bar{t}:\bar{x}-\bar{t} \in U} \chi_n(\bar{t}) \, d\bar{\lambda}(t) \right) d\bar{\lambda}(x) \\
&\leq \int_{\mathbb{T}^d} \left(\int_{\partial U(r^{-n})} \chi_n(\bar{x} - \bar{s}) \, d\bar{\lambda}(s) \right) d\bar{\lambda}(x) \\
&\leq \int_{\partial U(r^{-n})} \left(\int_{\mathbb{T}^d} \chi_n(\bar{x} - \bar{s}) \, d\bar{\lambda}(x) \right) d\bar{\lambda}(s) = \bar{\lambda}(\partial U(r^{-n})), \quad (4.12)
\end{aligned}
$$

using again the properties of χ_n. Now, (4.11) and (4.12) directly give (4.10). Due to (4.8), (4.9) and (4.10), we have

$$
\begin{aligned}
\frac{1}{\bar{\lambda}(U)} \left| \int_{\mathbb{T}^d} \mathbf{1}_U \circ T^{-n}.\varphi \, d\bar{\lambda} \right| \leq \frac{1}{\bar{\lambda}(U)} \Big(&\left| \int_{\mathbb{T}^d} (\chi_n * \mathbf{1}_U) \circ T^{-n}.(\chi_n * \varphi)) \, d\bar{\lambda} \right| \\
&+ \bar{\lambda}(U)\omega(\varphi, r^{-n}) + 3\|\varphi\|_\infty \bar{\lambda}(\partial U(r^{-n})) \Big).
\end{aligned}
$$

Now, the hypothesis on $m_u(\partial \mathcal{C}(\beta))$ implies that there exists Q_1 (depending on Q and on T) such that

$$\forall n \geq 0, \quad \bar{\lambda}(\partial U(r^{-n})) \leq Q_1 r^{-n}.$$

19

Moreover, applying Proposition 21 with $f = \chi_n * \varphi$ and $g = \chi_n * \mathbf{1}_U$ and using the following facts

$$\|\chi_n * \varphi\|_\infty \leq \|\varphi\|_\infty, \quad \|\chi_n * \mathbf{1}_U\|_\infty \leq 1, \ Lip(\chi_n * \mathbf{1}_U) \leq Lip(\chi_n) \ \text{ and } \ Lip(\chi_n * \varphi) \leq \|\varphi\|_\infty Lip(\chi_n),$$

we get the existence of $\tilde{C}_0$ (depending on C_0 and on Q) such that we have

$$\frac{1}{\bar{\lambda}(U)}\left|\int_{\mathbb{T}^d} \mathbf{1}_U \circ T^{-n}.\varphi\, d\bar{\lambda}\right| \leq \tilde{C}_0\|\varphi\|_\infty \frac{r^{-n} + (1 + r^{n(d+1)})\xi_0^n}{\varepsilon_n^{d_e+d_s} m_u(\mathcal{C})} + \omega(\varphi, r^{-n})$$

$$\leq 3\tilde{C}_0\|\varphi\|_\infty \frac{\xi_0^{n/(d+2)}}{\varepsilon_n^{d_e+d_s} m_u(\mathcal{C})} + \omega(\varphi, \zeta_1^n),$$

since $r^{-1} = r^{d+1}\xi_0 = \xi_0^{1/(d+2)}$. We conclude by taking $\xi_1 := \xi_0^{1/(d+2)}\alpha^{-(d_e+d_s)} < 1$. □

In the next result (which is an adaptation of Proposition 1.4 of [10]), we prove that Proposition 22 holds true with the stable-neutral continuity modulus $\omega_{(s,e)}$ instead of ω.

Proposition 23. *Let* $\zeta_1 \in (\rho_u^{1/(3(d+2)(d_e+d_s))}, 1)$. *There exist* $C_2 > 0$, $N_2 \geq 1$ *and* $\xi_2 \in (0,1)$ *such that, for every* $\bar{\lambda}$-*centered bounded function* $\varphi : \mathbb{T}^d \to \mathbb{R}$, *every* $\bar{x} \in \mathbb{T}^d$, *every* $n \geq N_2$ *and every bounded convex set* $\mathcal{C} \subseteq E_u$ *with diameter smaller than* r_0 *and satisfying* $m_u(\partial\mathcal{C}(\beta)) \leq Q\gamma$, *we have*

$$\left|\frac{1}{m_u(S^n(\mathcal{C}))}\int_{S^n\mathcal{C}} \varphi(\bar{x} + \overline{h_u})\, dm_u(h_u)\right| \leq K_2\left(\frac{\|\varphi\|_\infty}{m_u(\mathcal{C})}\xi_2^n + \omega_{(s,e)}(\varphi, \zeta_1^n)\right).$$

Proof. We consider a finite cover of $\mathbb{T}^d$ by sets $P_i = \bar{y}_i + \overline{B_u(r_0) + B_e(r_0) + B_s(r_0)}$ for $i = 1, ..., I$, $\bar{y}_i$ being fixed points of $\mathbb{T}^d$. We consider a partition of the unity $H_1, ..., H_I$ (i.e. $\sum_{i=1}^{I} H_i = 1$) such that each H_i is infinitely differentiable, with support in P_i. Let $\varphi : \mathbb{T}^d \to \mathbb{R}$ be a bounded centered function. For every $i = 1, ..., I$, we define $\varphi_i := H_i\varphi$. We have

$$\int_{S^n\mathcal{C}} \varphi(\bar{x} + \overline{h_u})\, dm_u(h_u) = \sum_{i=1}^{I}\int_{S^n\mathcal{C}} \varphi_i(\bar{x} + \overline{h_u})\, dm_u(h_u). \tag{4.13}$$

We also consider a continuously differentiable function $g : E_u \to [0, +\infty)$ with support in $B_u(r_0)$ and such that $\int_{E_u} g(h_u)\, dm_u(h_u) = 1$. We approximate now each φ_i by a regular function ψ_i by setting, for every $(h_u, h_e, h_s) \in B_u(r_0) \times B_e(r_0) \times B_s(r_0)$,

$$\psi_i(\bar{y}_i + \overline{h_u} + \overline{h_e} + \overline{h_s}) = g(h_u)\int_{B_u(r_0)} \varphi_i(\bar{y}_i + \overline{h'_u} + \overline{h_e} + \overline{h_s})\, dm_u(h'_u),$$

ψ_i being null outside of P_i. We observe that

$$\int_{P_i} \psi_i\, d\bar{\lambda} = \int_{P_i} \varphi_i\, d\bar{\lambda},$$

that $\|\psi_i\|_\infty \leq \|\varphi\|_\infty \|g\|_\infty m_u(B_u(r_0))$ and that, for every $\delta > 0$,

$$\omega(\psi_i, \delta) \leq m_u(B_u(r_0))\left[\|\varphi\|_\infty Lip(g)\delta + \|g\|_\infty \omega_{(s,e)}(\varphi_i, \delta)\right]$$

$$\leq m_u(B_u(r_0))\left[\|\varphi\|_\infty Lip(g)\delta + \|g\|_\infty \|\varphi\|_\infty Lip(H_i)\delta + \|g\|_\infty \omega_{(s,e)}(\varphi, \delta)\|H_i\|_\infty\right].$$

Now, applying Proposition 22 to ψ_i, for every $n \geq N_1$, we have

$$\left|\frac{1}{m_u(S^n\mathcal{C})}\int_{S^n\mathcal{C}} \psi_i(\bar{x} + \overline{h_u})\, dm_u(h_u)\right| \leq K'_1\left(\frac{\|\varphi\|_\infty \xi_1^n}{m_u(\mathcal{C})} + \omega_{(s,e)}(\varphi, \zeta_1^n) + \|\varphi\|_\infty \zeta_1^n\right). \tag{4.14}$$

20

We observe that the connected components of $(\bar{x}+\overline{S^n\mathcal{C}})\cap P_i$ are $\bar{x}+\overline{C_{i,j}}$, where $C_{i,j}$ are some connected subsets of E_u. We have

$$\int_{S^n\mathcal{C}} \varphi_i(\bar{x} + \overline{h_u})\, dm_u(h_u) = \sum_j \int_{C_{i,j}} \varphi_i(\bar{x} + \overline{h_u})\, dm_u(h_u)$$

and

$$\int_{S^n\mathcal{C}} \psi_i(\bar{x} + \overline{h_u})\, dm_u(h_u) = \sum_j \int_{C_{i,j}} \psi_i(\bar{x} + \overline{h_u})\, dm_u(h_u)\,.$$

Now, if $C_{i,j}$ does not contain any point of $\partial(S^n\mathcal{C})$, then there exists $h_e^{(j)} \in B_e(r_0)$ and $h_s^{(j)} \in B_s(r_0)$ such that

$$\bar{x} + \overline{C_{i,j}} = \left\{ \bar{y}_i + \overline{h_e^{(j)}} + \overline{h_s^{(j)}} + \overline{h_u};\ \ h_u \in B_u(r_0) \right\}\,.$$

Using the definition of ψ_i, we get

$$
\begin{aligned}
\int_{C_{i,j}} \psi_i(\bar{x} + \overline{h_u})\, dm_u(h_u) &= \int_{B_u(r_0)} \psi_i(\bar{y}_i + \overline{h_e^{(j)}} + \overline{h_s^{(j)}} + \overline{h_u})\, dm_u(h_u) \\
&= \int_{B_u(r_0)} \varphi_i(\bar{y}_i + \overline{h_e^{(j)}} + \overline{h_s^{(j)}} + \overline{h_u})\, dm_u(h_u),
\end{aligned}
$$

since $\int_{B_u(r_0)} g(h_u)\, dm_u(h_u) = 1$ and so

$$\int_{C_{i,j}} \psi_i(\bar{x} + \overline{h_u})\, dm_u(h_u) = \int_{C_{i,j}} \varphi_i(\bar{x} + \overline{h_u})\, dm_u(h_u).$$

Therefore we have

$$
\begin{aligned}
\left| \frac{1}{m_u(S^n\mathcal{C})} \int_{S^n\mathcal{C}} (\psi_i(\bar{x} + \overline{h_u}) - \varphi_i(\bar{x} + \overline{h_u}))\, dm_u(h_u) \right| &\leq 2\|\varphi\|_\infty \frac{m_u(\partial(S^n\mathcal{C})(r_0))}{m_u(S^n\mathcal{C})} \\
&\leq 2\|\varphi\|_\infty \frac{m_u(\partial\mathcal{C}(K\rho_u^n r_0))}{m_u(\mathcal{C})} \\
&\leq 2\|\varphi\|_\infty \frac{QK\rho_u^n r_0}{m_u(\mathcal{C})} \qquad (4.15)
\end{aligned}
$$

We conclude thanks to (4.13), (4.14) and (4.15), by taking $\xi_2 := \max(\xi_1, \zeta_1, \rho_u)$. $\qquad\square$

Proof of Theorem 18. The first point comes from the expression of $\mathbb{E}[\varphi|\mathcal{F}_n]$ given in Proposition 20 and from (4.2).

Let ζ_1, C_2, ξ_2 and N_2 as in Proposition 23 with $\zeta_1 < \zeta$. Let $\beta \in (\xi_2, 1)$ and $\mathcal{V}_n := \{m_u(F.) \geq \beta^n\}$. We take $\xi = \max(\xi_2/\beta, \beta^{\frac{1}{d_u}})$. To prove the second point, we use again the expression of $\mathbb{E}[\varphi|\mathcal{F}_{-n}]$ given in Proposition 20 and we apply Proposition 23 with $\mathcal{C} = F_{T^{-n}(\bar{x})}$ with the notation of Proposition 20.

Now, the last point comes from the fact (proved in Proposition II.1 of [9]) that

$$\exists L > 0,\ \ \forall n \geq 0,\ \bar{\lambda}(m_u(F.) < \beta^n) \leq L\beta^{\frac{n}{d_u}}\,.$$

$\qquad\square$

5 Proof of Theorems 1 and 4

In this section, C is a positive constant which may vary from lines to lines, and the notation $a_n \ll b_n$ means that there exists a numerical constant C not depending on n such that an $a_n \leq C b_n$, for all positive integers n.

Proof of Theorem 1. The proof is based on Proposition 6 of Section 3, which gives sufficient conditions for the weak invariance principle in 2-smooth Banach spaces.

Let $Y_i(s) = \mathbf{1}_{f \circ T^i \leq s} - F(s)$ and let $\mathcal{F}_i$ be the filtration introduced in Section 4. Note first that, for $2 \leq p < \infty$, the space $\mathbb{L}^p$ is 2-smooth and p-convex (see [16]). Moreover it has a Schauder basis (and even an unconditional basis).

Hence it suffices to check (3.2) of Proposition 6. As in [2], there exists a positive constant C such that

$$\sum_{k=1}^{\infty} \| \|P_{-k}(Y_0)\|_{\mathbb{L}^p} \|_2 \leq C \sum_{k=1}^{\infty} \Big(\frac{1}{k} \sum_{i=k}^{\infty} \| \|P_{-i}(Y_0)\|_{\mathbb{L}^p} \|_2^p \Big)^{1/p} \leq C \sum_{k=1}^{\infty} \Big(\frac{1}{k} \sum_{i=k}^{\infty} \| \|P_{-i}(Y_0)\|_{\mathbb{L}^p} \|_p^p \Big)^{1/p},$$

and $$\sum_{k=-\infty}^{0} \| \|P_{-k}(Y_0)\|_{\mathbb{L}^p} \|_2 \leq C \sum_{k=1}^{\infty} \Big(\frac{1}{k} \sum_{i=k}^{\infty} \| \|P_{i+1}(Y_0)\|_{\mathbb{L}^p} \|_2^p \Big)^{1/p} \leq C \sum_{k=1}^{\infty} \Big(\frac{1}{k} \sum_{i=k}^{\infty} \| \|P_{i+1}(Y_0)\|_{\mathbb{L}^p} \|_p^p \Big)^{1/p}.$$

Since $\mathbb{L}_p$ is p-convex, it follows that

$$\sum_{i=k}^{\infty} \| \|P_{-i}(Y_0)\|_{\mathbb{L}^p} \|_p^p \leq K \| \|\mathbb{E}(Y_k|\mathcal{F}_0)\|_{\mathbb{L}^p} \|_p^p \quad \text{and} \quad \sum_{i=k}^{\infty} \| \|P_{i+1}(Y_0)\|_{\mathbb{L}^p} \|_p^p \leq K \| \|Y_{-n} - \mathbb{E}(Y_{-n}|\mathcal{F}_0)\|_{\mathbb{L}^p} \|_p.$$

Hence (3.2) is true as soon as

$$\sum_{n \geq 1} \frac{1}{n^{1/p}} \| \|\mathbb{E}(Y_n|\mathcal{F}_0)\|_{\mathbb{L}^p} \|_p < \infty \quad \text{and} \quad \sum_{n \geq 1} \frac{1}{n^{1/p}} \| \|Y_{-n} - \mathbb{E}(Y_{-n}|\mathcal{F}_0)\|_{\mathbb{L}^p} \|_p < \infty.$$

Let us have a look to

$$\| \|\mathbb{E}(Y_n|\mathcal{F}_0)\|_{\mathbb{L}^p} \|_p = \Big(\mathbb{E} \int_{\mathbb{R}} |F_{f \circ T^n | \mathcal{F}_0}(t) - F(t)|^p dt \Big)^{1/p}$$

$$\leq \Big(\mathbb{E} \int_{\mathbb{R}} |F_{f \circ T^n | \mathcal{F}_0}(t) - F(t)| dt \Big)^{1/p}.$$

Now

$$\int_{\mathbb{R}} |F_{f \circ T^n | \mathcal{F}_0}(t) - F(t)| dt = \sup_{g \in \Lambda_1} \Big| \mathbb{E}(g \circ f \circ T^n | \mathcal{F}_0) - \mathbb{E}(g \circ f) \Big|,$$

where Λ_1 is the set of 1-lipschitz functions. Hence, since $\omega_{(s,e)}(g \circ f, \cdot)$ is smaller than $\omega_{(s,e)}(f, \cdot)$, it follows from (4.5) and (4.6) of Theorem 18 that

$$\| \|\mathbb{E}(Y_k|\mathcal{F}_0)\|_{\mathbb{L}^p} \|_p \leq \Big(\mathbb{E} \Big(\sup_{g \in \Lambda_1} \Big| \mathbb{E}(g \circ f \circ T^k | \mathcal{F}_0) - \mathbb{E}(g \circ f) \Big| \Big) \Big)^{1/p} \leq C((\omega_{(s,e)}(f, \zeta^n))^{1/p} + \|f\|_\infty^{1/p} \xi^{n/p}),$$

by noticing that we can replace Λ_1 by the set of $g \in \Lambda_1$ such that $g \circ f(0) = 0$. In the same way, due to (4.4) of Theorem 18, we have

$$\| \|Y_{-n} - \mathbb{E}(Y_{-n}|\mathcal{F}_0)\|_{\mathbb{L}^p} \|_p \leq C(\omega_{(u)}(f, \rho_u^n))^{1/p}.$$

The result follows. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

22

Proof of Theorem 4. Our aim is to apply the tightness criterion given in Proposition 17. Let $X_i = f \circ T^i$ and let $\mathcal{F}_i$ be the filtration defined in Section 4. We need the following upper bounds.

Lemma 24. *Let $g_{s,t}(v) = \mathbf{1}_{v \leq t} - \mathbf{1}_{v \leq s}$, and let P be the image measure of $\bar{\lambda}$ by f. Under the assumptions of Theorem 4, we have, for any $\beta > 1$,*

$$\sum_{k=0}^{n} |\mathrm{Cov}(g_{s,t}(X_0), g_{s,t}(X_k))| \ll \|g_{s,t}\|_{P,1}^{(\beta+\alpha-1)/(\beta+\alpha)} \sum_{k=0}^{n} \frac{1}{(k+1)^{a\alpha/(\beta+\alpha)}} .$$

Lemma 25. *Let $p > 2$. Under the assumptions of Theorem 4, we have*

$$\|\mathbb{E}_0(g_{s,t}(X_k)) - \mathbb{E}(g_{s,t}(X_k))\|_p \ll k^{-a\alpha/(\alpha+p)}$$

$$\|g_{s,t}(X_0) - \mathbb{E}_k(g_{s,t}(X_0))\|_p \ll k^{-a\alpha/(\alpha+p)} ,$$

and, for the coefficient $A(g_{s,t}(X) - \mathbb{E}[g_{s,t}(X)], j)$ defined in (3.12),

$$A(g_{s,t}(X) - \mathbb{E}[g_{s,t}(X)], j) \ll j^{-2a\alpha/(2\alpha+p)}$$

Let us continue the proof of Theorem 4 with the help of these lemmas. From Proposition 13 and Lemma 25, we derive that, for $p > 2$,

$$\left\| \max_{1 \leq k \leq n} |S_k(g_{s,t})| \right\|_p \ll n^{1/2} \left(\|g_{s,t}\|_{P,1}^{(\beta+\alpha-1)/(\beta+\alpha)} \sum_{k=1}^{n} \frac{1}{k^{a\alpha/(\beta+\alpha)}} \right)^{1/2} + n^{1/p} \left(\sum_{k=1}^{n} \frac{k^{-2a\alpha/(2\alpha+p)}}{k^{(2/p)-1}} (\log k)^{\gamma} \right)^{1/2} ,$$

where γ can be taken $\gamma = 0$ for $2 < p \leq 3$ and $\gamma > p - 3$ for $p > 3$. Therefore if

$$a > \max\left(1 + \frac{\beta}{\alpha}, \frac{(p-1)(2\alpha+p)}{p\alpha}\right),$$

then setting $r = 2(\beta+\alpha)/(\beta+\alpha-1)$, we get that

$$\left\| \max_{1 \leq k \leq n} |S_k(g_{s,t})| \right\|_p \ll n^{1/2} \|g_{s,t}\|_{P,1}^{1/r} + n^{1/p} .$$

We shall apply the tightness criterion given in Proposition 17. Since $\mathcal{N}_{P,1}(x, \mathcal{F}) \leq C x^{-\ell}$ for the class $\mathcal{F} = \{u \mapsto \mathbf{1}_{u \leq t}, t \in \mathbb{R}^{\ell}\}$, we get that

$$\int_0^1 x^{(1-r)/r} (\mathcal{N}_{P,1}(x, \mathcal{F}))^{1/p} dx \leq C \int_0^1 x^{(1-r)/r} x^{-\ell/p} dx < \infty, \tag{5.1}$$

as soon as $p > 2\ell(\beta+\alpha)/(\beta+\alpha-1)$. Moreover

$$\lim_{x \to 0} x^{p-2} \mathcal{N}_{P,1}(x, \mathcal{F}) = 0 \tag{5.2}$$

as soon as $p > 2 + \ell$.

Hence if $p \in]2, 2\ell(1+\alpha^{-1})]$, we take $\beta = (2\alpha\ell + (1-\alpha)p)/(p-2\ell) + \varepsilon$ for some positive and small enough ε (so that $\beta > 1$), and we infer that (5.1) and (5.2) hold provided that $p > \max(\ell+2, 2\ell)$ and

$$a > g_{\ell,\alpha}(p) = \max\left(\frac{p}{\alpha(p-2\ell)}, \frac{(p-1)(2\alpha+p)}{p\alpha}\right).$$

Taking the minimum in $p \geq \max(\ell+2, 2\ell)$ on the right hand, we obtain that (5.1) and (5.2) hold provided that $a > a(\ell, \alpha)$, where $a(\ell, \alpha)$ has been defined in (2.2).

We infer that (3.26) and (3.27) of Proposition 17 hold for this choice of a, which prove the tightness of the empirical process (see [17], page 227).

Note that the weak convergence of the finite dimensional distributions holds as soon as $a > (\alpha+2)/2\alpha$ (this can be proved as in [5] by using Lemma 25).

$\square$

23

Proof of Lemma 24. We prove the results for $\ell = 2$. The general case can be proved in the same way. For $u \in \mathbb{R}$, let $h_u(x) = \mathbf{1}_{x \leq u}$. By definition of $g_{s,t}$,

$$g_{s,t} = h_{t_1} \otimes h_{t_2} - h_{s_1} \otimes h_{s_2} \,,$$

with the notation $(G_1 \otimes G_2)(u_1, u_2) := G_1(u_1)G_2(u_2)$. For $\varepsilon > 0$, let $h_{u,\varepsilon}(x) = \mathbf{1}_{x \leq u} - \varepsilon^{-1}(x - u - \varepsilon)\mathbf{1}_{u < x \leq u+\varepsilon}$ and note that $h_{u,\varepsilon}$ is Lipschitz with Lipschitz constant ε^{-1}. We have the decomposition $h_{t_1} \otimes h_{t_2} = h_{t_1,\varepsilon} \otimes h_{t_2,\varepsilon} + R_{t,\varepsilon}$, where

$$R_{t,\varepsilon} = (h_{t_1} - h_{t_1,\varepsilon}) \otimes h_{t_2} + h_{t_1,\varepsilon} \otimes (h_{t_2} - h_{t_2,\varepsilon}) \,.$$

Setting

$$g_{s,t,\varepsilon} = h_{t_1,\varepsilon} \otimes h_{t_2,\varepsilon} - h_{s_1,\varepsilon} \otimes h_{s_2,\varepsilon} \,,$$

we obtain the decomposition

$$g_{s,t} = g_{s,t,\varepsilon} + H_{s,t,\varepsilon}, \quad \text{with} \quad H_{s,t,\varepsilon} = R_{t,\varepsilon} - R_{s,\varepsilon} \,. \tag{5.3}$$

On the other hand, we have that

$$\mathrm{Cov}(g_{s,t}(X_0), g_{s,t}(X_k)) = \mathbb{E}((g_{s,t}(X_0) - \mathbb{E}(g_{s,t}(X_0)|\mathcal{F}_{[k/2]}))g_{s,t}(X_k)) + \mathrm{Cov}(\mathbb{E}(g_{s,t}(X_0)|\mathcal{F}_{[k/2]}), g_{s,t}(X_k)) \,.$$

Using (5.3), we have that

$$\mathbb{E}((g_{s,t}(X_0) - \mathbb{E}(g_{s,t}(X_0)|\mathcal{F}_{[k/2]}))g_{s,t}(X_k)) = \mathbb{E}((g_{s,t,\varepsilon}(X_0) - \mathbb{E}(g_{s,t,\varepsilon}(X_0)|\mathcal{F}_{[k/2]}))g_{s,t}(X_k))$$
$$+ \mathbb{E}((H_{s,t,\varepsilon}(X_0) - \mathbb{E}(H_{s,t,\varepsilon}(X_0)|\mathcal{F}_{[k/2]}))g_{s,t}(X_k)) \,. \tag{5.4}$$

Applying (4.4) of Theorem 18, we infer that

$$|\mathbb{E}((g_{s,t,\varepsilon}(X_0) - \mathbb{E}(g_{s,t,\varepsilon}(X_0)|\mathcal{F}_{[k/2]}))g_{s,t}(X_k))| \leq C\|g_{s,t}\|_{P,1}\varepsilon^{-1}\omega_{(u)}(f, \rho_u^{[k/2]}) \,. \tag{5.5}$$

Applying Hölder's inequality, and using the fact that the distributions functions of f_1 and f_2 are Hölder continuous of order α, we get that

$$|\mathbb{E}((H_{s,t,\varepsilon}(X_0) - \mathbb{E}(H_{s,t,\varepsilon}(X_0)|\mathcal{F}_{[k/2]}))g_{s,t}(X_k))| \leq C\|g_{s,t}\|_{P,1}^{(\beta-1)/\beta}\varepsilon^{\alpha/\beta} \,. \tag{5.6}$$

Using (5.3) again, we also have that

$$\mathrm{Cov}(\mathbb{E}(g_{s,t}(X_0)|\mathcal{F}_{[k/2]}), g_{s,t}(X_k)) = \mathrm{Cov}(\mathbb{E}(g_{s,t}(X_0)|\mathcal{F}_{[k/2]}), g_{s,t,\varepsilon}(X_k))$$
$$+ \mathrm{Cov}(\mathbb{E}(g_{s,t}(X_0)|\mathcal{F}_{[k/2]}), H_{s,t,\varepsilon}(X_k)) \,. \tag{5.7}$$

Let $\mathcal{V}_n$ be the set introduced in Theorem 18. Applying (4.5) of Theorem 18, we have that

$$|\mathrm{Cov}(\mathbb{E}(g_{s,t}(X_0)|\mathcal{F}_{[k/2]}), g_{s,t,\varepsilon}(X_k))| \leq C\mathbb{E}(|\mathbb{E}(g_{s,t}(X_{-k})|\mathcal{F}_{[k/2]-k})|\mathbf{1}_{\mathcal{V}_{k-[k/2]}^c}) +$$
$$C\|g_{s,t}\|_{P,1}(\xi^{[k/2]} + \varepsilon^{-1}\omega_{(s,e)}(f, \zeta^{[k/2]})) \,.$$

Since $\bar{\lambda}(\mathcal{V}_{k-[k/2]}^c) \leq C\xi^{[k/2]}$, applying Hölder's inequality, we get that

$$\mathbb{E}(|\mathbb{E}(g_{s,t}(X_{-k})|\mathcal{F}_{[k/2]-k})|\mathbf{1}_{\mathcal{V}_{k-[k/2]}^c}) \leq C\|g_{s,t}\|_{P,1}^{(\beta+\alpha-1)/(\beta+\alpha)}\xi^{[k/2]/(\beta+\alpha)} \,. \tag{5.8}$$

Applying Hölder's inequality again, and using that the distributions functions of f_1 and f_2 are Hölder continuous of order α, we get that

$$|\mathrm{Cov}(\mathbb{E}(g_{s,t}(X_0)|\mathcal{F}_{[k/2]}), H_{s,t,\varepsilon}(X_k))| \leq C\|g_{s,t}\|_{P,1}^{(\beta-1)/\beta}\varepsilon^{\alpha/\beta} \,. \tag{5.9}$$

Gathering the bounds (5.4), (5.5), (5.6), (5.7), (5.8) and (5.9), we get that

$$|\text{Cov}(g_{s,t}(X_0), g_{s,t}(X_k))| \leq C\Big(\|g_{s,t}\|_{P,1}\frac{1}{\varepsilon k^a} + \|g_{s,t}\|_{P,1}^{(\beta-1)/\beta}\varepsilon^{\alpha/\beta} + \|g_{s,t}\|_{p,1}^{(\beta+\alpha-1)/(\beta+\alpha)}\xi^{[k/2]/(\beta+\alpha)}\Big).$$

Taking $\varepsilon = \|g_{s,t}\|_{P,1}^{1/(\alpha+\beta)}k^{-a\beta/(\alpha+\beta)}$, we get that

$$|\text{Cov}(g_{s,t}(X_0), g_{s,t}(X_k))| \leq C\|g_{s,t}\|_{P,1}^{(\beta+\alpha-1)/(\beta+\alpha)}\Big(\frac{1}{k^{a\alpha/(\alpha+\beta)}} + \xi^{[k/2]/(\beta+\alpha)}\Big).$$

The result follows by summing in k. $\hfill\square$

Proof of Lemma 25. Using the same notations as in the proof of Lemma 24, and using that the distribution functions of f_1 and f_2 are Hölder continuous of order α, we obtain that

$$\|\mathbb{E}_0(g_{s,t}(X_k) - \mathbb{E}(g_{s,t}(X_k))\|_p \leq \|\mathbb{E}_0(g_{s,t,\varepsilon}(X_k) - \mathbb{E}(g_{s,t,\varepsilon}(X_k))\|_p + C\varepsilon^{\alpha/p}.$$

Recall that the $\mathcal{V}_n$ introduced in Theorem 18 is such that $\bar{\lambda}(\mathcal{V}_n^c) \leq C\xi^n$. Applying Theorem 18 (see (4.7)), we obtain that

$$\|\mathbb{E}_0(g_{s,t,\varepsilon}(X_k) - \mathbb{E}(g_{s,t,\varepsilon}(X_k))\|_p \leq C(\varepsilon^{-1}\omega_{(s,e)}(f, \zeta^k) + \xi^{k/p}).$$

Consequently

$$\|\mathbb{E}_0(g_{s,t}(X_k) - \mathbb{E}(g_{s,t}(X_k))\|_p \leq C\Big(\frac{1}{\varepsilon k^a} + \varepsilon^{\alpha/p} + \xi^{k/p}\Big).$$

Choosing $\varepsilon = k^{-ap/(\alpha+p)}$, we obtain that

$$\|\mathbb{E}_0(g_{s,t}(X_k) - \mathbb{E}(g_{s,t}(X_k))\|_p \leq C\Big(\frac{1}{k^{a\alpha/(\alpha+p)}} + \xi^{k/p}\Big),$$

proving the first inequality.

In the same way

$$\|g_{s,t}(X_0) - \mathbb{E}_k(g_{s,t}(X_0))\|_p \leq \|g_{s,t,\varepsilon}(X_0) - \mathbb{E}_k(g_{s,t,\varepsilon}(X_0))\|_p + C\varepsilon^{\alpha/p}.$$

Applying (4.4) of Theorem 18, we obtain that

$$\|g_{s,t}(X_0) - \mathbb{E}_k(g_{s,t}(X_0))\|_p \leq C(\varepsilon^{-1}\omega_{(u)}(f, \rho_u^k) + \varepsilon^{\alpha/p}).$$

Since $\omega_{(u)}(f, \rho_u^k) \leq Ck^{-a}$, the choice $\varepsilon = k^{-ap/(\alpha+p)}$ gives the second inequality.

Let $h^{(0)}(X_i) = h(X_i) - \mathbb{E}(h(X_i))$. To prove the third inequality, we have to bound up

$$\sup_{i\geq 0}\|\mathbb{E}_0(g_{s,t}^{(0)}(X_i)g_{s,t}^{(0)}(X_{j+i}))\|_{p/2} \quad \text{and} \quad \sup_{0\leq i\leq j}\|\mathbb{E}_0(g_{s,t}^{(0)}(X_j)g_{s,t}^{(0)}(X_{j+i})) - \mathbb{E}(g_{s,t}^{(0)}(X_j)g_{s,t}^{(0)}(X_{j+i}))\|_{p/2}.$$

Using the decomposition (5.3), and the fact that the distibution functions of f_1 and f_2 are Hölder continuous of order α, we get that

$$\|\mathbb{E}_0(g_{s,t}^{(0)}(X_i)g_{s,t}^{(0)}(X_{j+i}))\|_{p/2} \leq \|\mathbb{E}_0(g_{s,t,\varepsilon}^{(0)}(X_i)g_{s,t,\varepsilon}^{(0)}(X_{j+i}))\|_{p/2} + C\varepsilon^{2\alpha/p}, \tag{5.10}$$

and

$$\|\mathbb{E}_0(g_{s,t}^{(0)}(X_j)g_{s,t}^{(0)}(X_{j+i})) - \mathbb{E}(g_{s,t}^{(0)}(X_j)g_{s,t}^{(0)}(X_{j+i}))\|_{p/2}$$
$$\leq \|\mathbb{E}_0(g_{s,t,\varepsilon}^{(0)}(X_j)g_{s,t,\varepsilon}^{(0)}(X_{j+i})) - \mathbb{E}(g_{s,t,\varepsilon}^{(0)}(X_j)g_{s,t,\varepsilon}^{(0)}(X_{j+i}))\|_{p/2} + C\varepsilon^{2\alpha/p}. \tag{5.11}$$

25

Writing

$$\|\mathbb{E}_0(g_{s,t,\varepsilon}^{(0)}(X_i)g_{s,t,\varepsilon}^{(0)}(X_{j+i}))\|_{p/2} \leq \|\mathbb{E}_0((g_{s,t,\varepsilon}(X_i) - \mathbb{E}(g_{s,t,\varepsilon}(X_i)|\mathcal{F}_{i+[j/2]}))g_{s,t,\varepsilon}^{(0)}(X_{j+i}))\|_{p/2}$$
$$+ \|\mathbb{E}_0(\mathbb{E}(g_{s,t,\varepsilon}(X_i)|\mathcal{F}_{i+[j/2]})g_{s,t,\varepsilon}^{(0)}(X_{j+i}))\|_{p/2}, \quad (5.12)$$

and arguing as in Lemma 24, we infer that

$$\|\mathbb{E}_0(g_{s,t,\varepsilon}^{(0)}(X_i)g_{s,t,\varepsilon}^{(0)}(X_{j+i}))\|_{p/2} \leq C\Big(\frac{1}{\varepsilon j^a} + \xi^{[j/2]}\Big). \quad (5.13)$$

From (5.10) and (5.13), we obtain the bound

$$\|\mathbb{E}_0(g_{s,t}^{(0)}(X_i)g_{s,t}^{(0)}(X_{j+i}))\|_{p/2} \leq C\Big(\frac{1}{\varepsilon j^a} + \varepsilon^{2\alpha/p} + \xi^{[j/2]}\Big).$$

Taking $\varepsilon = j^{-ap/(2\alpha+p)}$, we obtain that

$$\sup_{i\geq 0}\|\mathbb{E}_0(g_{s,t}^{(0)}(X_i)g_{s,t}^{(0)}(X_{j+i}))\|_{p/2} \leq Cj^{-2a\alpha/(2\alpha+p)}. \quad (5.14)$$

Let $\varphi := g_{s,t,\varepsilon} \circ f - \bar{\lambda}(g_{s,t,\varepsilon} \circ f)$. Applying Theorem 18 (see (4.7)), for $i \leq j$,

$$\|\mathbb{E}_0(g_{s,t,\varepsilon}^{(0)}(X_j)g_{s,t,\varepsilon}^{(0)}(X_{j+i})) - \mathbb{E}(g_{s,t,\varepsilon}^{(0)}(X_j)g_{s,t,\varepsilon}^{(0)}(X_{j+i}))\|_{p/2} = \|\mathbb{E}(\varphi.\varphi \circ T^i|\mathcal{F}_{-j}) - \mathbb{E}(\varphi.\varphi \circ T^i)\|_{p/2}$$
$$\leq C(\xi^{2j/p} + \omega_{(s,e)}(\varphi.\varphi \circ T^i, \zeta^j)).$$

By (4.3), $\omega_{(s,e)}(\varphi.\varphi \circ T^i, \zeta^j) \leq \omega_{(s,e)}(\varphi, K\zeta^j j^{d_e})) \leq \omega_{(s,e)}(\varphi, L\zeta_0^j))$, so that

$$\|\mathbb{E}_0(g_{s,t,\varepsilon}^{(0)}(X_j)g_{s,t,\varepsilon}^{(0)}(X_{j+i})) - \mathbb{E}(g_{s,t,\varepsilon}^{(0)}(X_j)g_{s,t,\varepsilon}^{(0)}(X_{j+i}))\|_{p/2} \leq C(\xi^{2j/p} + \omega_{(s,e)}(\varphi, L\zeta_0^j)). \quad (5.15)$$

Since $\omega_{(s,e)}(\varphi, L\zeta_0^j) \leq \varepsilon^{-1}\omega_{(s,e)}(f, L\zeta_0^j) \leq C\varepsilon^{-1}j^{-a}$, we obtain from (5.11) and (5.15) that

$$\|\mathbb{E}_0(g_{s,t}^{(0)}(X_j)g_{s,t}^{(0)}(X_{j+i})) - \mathbb{E}(g_{s,t}^{(0)}(X_j)g_{s,t}^{(0)}(X_{j+i}))\|_{p/2} \leq C\Big(\frac{1}{\varepsilon j^a} + \varepsilon^{2\alpha/p} + \xi^{2j/p}\Big).$$

Taking $\varepsilon = j^{-ap/(2\alpha+p)}$, we obtain that

$$\sup_{0\leq i\leq j}\|\mathbb{E}_0(g_{s,t}^{(0)}(X_j)g_{s,t}^{(0)}(X_{j+i})) - \mathbb{E}(g_{s,t}^{(0)}(X_j)g_{s,t}^{(0)}(X_{j+i}))\|_{p/2} \leq Cj^{-2a\alpha/(2\alpha+p)}. \quad (5.16)$$

The third inequality of Lemma 25 follows from (5.14), (5.16) and from the definition of $A(g_{s,t}(X) - \mathbb{E}[g_{s,t}(X)], j)$ given in Proposition 13. $\qquad\square$

6 Appendix

In this section, we prove Remark 5, so we give the solutions of the equation (2.3). We first write (2.3) under the following form $p^3 + bp^2 + cp + d = 0$. Following the classical Cardan method, we set $p' := -\frac{b^2}{3} + c$ and $q := \frac{b}{27}(2b^2 - 9c) + d$ (this leads to the formulas for p' and q as given in Remark 5). Observe that $p^3 + bp^2 + cp + d = 0$ means that $z = p + \frac{b}{3}$ satisfies $z^3 + p'z + q = 0$. We then compute as usual $\Delta := q^2 + \frac{4}{27}(p')^3$. We get

$$\Delta = ((64/27)\ell - (64/27)\ell^2 - 16/27)\alpha^4 + (-(128/27)\ell^3$$
$$+ (128/27)\ell^2 - (32/9)\ell)\alpha^3 + ((32/27)\ell - (64/27)\ell^4 + (16/27)\ell^2 - 16/27 - (128/27)\ell^3)\alpha^2$$
$$+ (-(32/9)\ell - (32/27)\ell^2 - (64/27)\ell^4 - (32/9)\ell^3)\alpha - (16/27)\ell^2 - (16/27)\ell^4 < 0.$$

26

Since Δ is negative, we use the usual expression of the solutions z with cos and arccos (to which we substract $b/3$). So the solutions are

$$p_k = 2\frac{\ell + 1 - \alpha}{3} + 2\sqrt{-\frac{p'}{3}} \cos\left(\frac{1}{3}\arccos\left(-\frac{q}{2}\sqrt{\frac{27}{-(p')^3}}\right) + \frac{2k\pi}{3}\right)$$

for $k \in \{0, 1, 2\}$. Clearly $p_1 < p_2 < p_0$. The unique solution in $]2\ell, 4\ell[$ is then p_0.

References

[1] Andrews, D. W. K. and Pollard, D. *An introduction to functional central limit theorems for dependent stochastic processes*. Int. Stat. Rev. 62 (1994), 119-132.

[2] Dedecker, J., Merlevède, F. and Peligrad, M. *Invariance principles for linear processes with application to isotonic regression*. Bernoulli 17 (2011), 88–113.

[3] Dedecker, J. Merlevède, F. and Pène, F. *Rates of convergence in the strong invariance principle for non-adapted sequences. Application to ergodic automorphisms of the torus*. To appear in Proc. Intern. Conf. on High Dim. Probab., VI.

[4] Dedecker, J. Merlevède, F. and Pène, F. *Rates in the strong invariance principle for ergodic automorphisms of the torus*. Preprint (2012). arXiv: 1206.4336

[5] Dedecker, J. and Prieur, C. *An empirical central limit theorem for dependent sequences*. Stochastic Process. Appl. 117 (2007), 121–142.

[6] Dehling, H. and Durieu, O. *Empirical processes of multidimensional systems with multiple mixing properties*. Stochastic Process. Appl. 121 (2011), 1076–1096.

[7] Durieu, O. and Jouan, P. *Empirical invariance principle for ergodic torus automorphisms; genericity*. Stoch. Dyn. 8 (2008), 173–195.

[8] Kiefer, J. *Skorohod embedding of multivariate random variables and the sample distribution function*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 24 (1972), 1-35.

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

[10] Le Borgne, S. and Pène, F. *Vitesse dans le théorème limite central pour certains systèmes dynamiques quasi-hyperboliques*, Bull. Soc. Math. France 133 (2005), no. 3, 395–417

[11] Lind, D. A. *Dynamical properties of quasihyperbolic toral automorphisms*, Ergodic Theory Dynamical Systems 2 (1982), 49–68.

[12] Merlevède, F. and Peligrad, M. *Rosenthal inequalities for martingales and stationary sequences and examples*. To appear in Ann. Probab. (2012). arXiv:1103.3242.

[13] Peligrad, M., Utev, S. and Wu, W.B. *A maximal $\mathbf{L}_p$-inequality for stationary sequences and its applications*. Proc. Amer. Math. Soc. 135 (2007) 541–550.

[14] Pène F. *Averaging method for differential equations perturbed by dynamical systems*, ESAIM, Probab. stat. 6 (2002), 33-88.

[15] Pinelis, I. *Optimum bounds for the distributions of martingales in Banach spaces*. Ann. Probab. 22 (1994), 1679-1706.

[16] Pisier, G. *Martingales with values in uniformly convex spaces*. Israel J. Math. 20 (1975), 326-350.

[17] van der Vaart, A. W. and Wellner, J. A. *Weak Convergence and Empirical Processes*. Springer, Berlin, 1996.

[18] Woyczyński, W.A. *A central limit theorem for martingales in Banach spaces*. Bull. Acad. Polon. Sci. Sr. Sci. Math. Astronom. Phys. 23 (1975), 917–920.