ALEA, Lat. Am. J. Probab. Math. Stat. 10 (2), 731–766 (2013)

Empirical central limit theorems for ergodic automorphisms of the torus

Jérôme Dedecker, Florence Merlevède and Françoise Pène

Université Paris Descartes, Sorbonne Paris Cité,
Laboratoire MAP5 and CNRS UMR 8145,
45 rue des Saints Pères,
75270 Paris Cedex 06, France.
E-mail address: jerome.dedecker@parisdescartes.fr
URL: http://w3.mi.parisdescartes.fr/~jdedecke/

Université Paris-Est,
LAMA (UMR 8050), UPEMLV, CNRS, UPEC,
5 Boulevard Descartes,
77454 Marne La Vallée, France.
E-mail address: florence.merlevede@univ-mlv.fr
URL: http://perso-math.univ-mlv.fr/users/merlevede.florence/

Université de Brest,
Laboratoire de Mathématiques de Bretagne Atlantique UMR CNRS 6205,
6 avenue Le Gorgeu,
29238 Brest cedex 3, France.
E-mail address: francoise.pene@univ-brest.fr
URL: http://lmba.math.univ-brest.fr/perso/francoise.pene/

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 mild conditions on the modulus of continuity of f. The proofs are based on new limit theorems, on new inequalities for non-adapted sequences, and on new estimates of the conditional expectations of f with respect to a natural filtration.

Received by the editors October 12, 2012; accepted September 19, 2013.

2010 *Mathematics Subject Classification.* 60F17, 37D30.

Key words and phrases. Empirical distribution function, Kiefer process, Ergodic automorphisms of the torus, Moment inequalities.

Françoise Pène is supported by the french ANR project Perturbations (ANR 10-BLAN 0106).

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 Le Borgne (1999), 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 Lind (1982)), and to prove some decorrelation properties for regular functions (see Lind (1982); Le Borgne and Pène (2005)).

Let ℓ be some positive integer, and let $f = (f_1, \dots f_\ell)$ 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 (1999) proved the functional central limit theorem and the Strassen strong invariance principle for the partial sums

$$\sum_{i=1}^{n} (f \circ T^i - \bar{\lambda}(f)) \tag{1.1}$$

under weak hypotheses on the Fourier coefficients of f, thanks to Gordin's method and to the partitions studied by Lind (1982). In the recent paper by Dedecker et al. (2013a), 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

$$S_n(s) = \sum_{k=1}^{n} (\mathbf{1}_{f\circ T^k \leq s} - F(s)) , \tag{1.2}$$

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 $F(s) = \bar{\lambda}(f \leq s)$ 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

$$\left\{ \frac{S_{[nt]}(s)}{\sqrt{n}}, t \in [0, 1], s \in \mathbb{R}^\ell \right\} . \tag{1.3}$$

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 of the process $\{n^{-1/2} S_{[nt]}, t \in [0,1]\}$ in $D_{\mathbb{L}^p}([0,1])$ 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 (1.3) 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 process introduced by Kiefer (1972) for the sequential empirical process of independent and identically distributed random variables. The proof is based on a new Rosenthal inequality for dependent sequences (possibly non adapted), 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 Durieu and Jouan (2008) and Dehling and Durieu (2011). 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 (1982), involving the modulus of continuity of the observables (See Theorem 4.1 of Section 4). As far as we know, such controls were known for Hölder observables only (see Le Borgne and Pène (2005)). Let us indicate that the inequalities given in Theorem 4.1 are interesting by themselves. For instance one can use them to establish weak invariance principle and rates of convergence in the strong invariance principle for the partial sums (1.1) (see Section 6).

In this paper, the conditions on a function f from $\mathbb{T}^d$ to $\mathbb{R}$ will be expressed in terms of its modulus of continuity $\omega(f, \cdot)$ defined as follows:

$$\text{for } \delta > 0, \quad \omega(f, \delta) := \sup_{\bar{x}, \bar{y} \in \mathbb{T}^d \,:\, d_1(\bar{x}, \bar{y}) \leq \delta} |f(\bar{x}) - f(\bar{y})|\,, \tag{1.4}$$

where $d_1(\bar{x}, \bar{y}) = \min_{k \in \mathbb{Z}^d} \|x - y + k\|$ for some norm $\|\cdot\|$ on $\mathbb{R}^d$.

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 $p \in [2, \infty[$, the random variable S_n defined in (1.2) 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. 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),

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

Theorem 2.1. *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*

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

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

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

The proof of Theorem 2.1 is based on results of Sections 3 and 4 and is postponed to Section 5.

Remark 2.2. In particular, if f is Hölder continuous, then the conclusion of Theorem 2.1 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 $(f\circ T^i)_{1\le i\le n}$ and the distribution μ of f. Let

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

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

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

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

Corollary 2.3. *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 $\sup_{1\le k\le n}\sqrt{n}K(\mu_{n,k},\mu)$ 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 2.3: Applying Theorem 2.1 with $p=2$, we know that the process $\{n^{-1/2}S_{[nt]},t\in[0,1]\}$ 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 $|f|\le M$ 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 $\sup_{1\le k\le n} nK(\mu_{n,k},\mu) = \sup_{t\in[0,1]} \|S_{[nt]}\|_{\mathbb{L}_1}$, and the result follows. □

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

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

where we recall that $\omega(f_i,x)$ is defined by equation (1.4).

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 van der Vaart and Wellner (1996) (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

$$a(\ell,\alpha) = \min_{p\ge\max(\ell+2,2\ell)} k_{\ell,\alpha}(p),\ \text{ where } k_{\ell,\alpha}(p) = \max\Big(\frac{p}{\alpha(p-2\ell)}, \frac{(p-1)(2\alpha+p)}{p\alpha}\Big)\,. \tag{2.2}$$

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

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

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

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

Theorem 2.4. *Let $f = (f_1,\dots,f_\ell) : \mathbb{T}^d \to \mathbb{R}^\ell$ 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*

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

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 $(s,s')\in\mathbb{R}^\ell\times\mathbb{R}^\ell$,

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

The proof of Theorem 2.4 is given in Section 5. It uses results of Sections 3 and 4.

Remark 2.5. 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 $p_0 \sim 2.9$ and finally $a(1,1) = 10/3$.

Recall that, by Theorem 2.1, 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 Durieu and Jouan (2008) and Dehling and Durieu (2011). When $\ell = 1$, a consequence of the main result of the paper by Durieu and Jouan (2008) 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 by Dehling and Durieu (2011) this result is extended to any dimension ℓ, under the assumptions that f is Hölder continuous and 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 Dehling and Durieu (2011). Indeed, one cannot prove the multiple mixing for the sequence $(f \circ T^i)_{i\in\mathbb{Z}}$ by assuming only that $\omega(f,x) \le C|\log(x)|^{-a}$ 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 Dehling and Durieu (2011), our result cannot be directly compared to that of Dehling and Durieu (2011), 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 Dehling and Durieu (2011).

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 $a_n \le Cb_n$, 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 (1975), we say that a Banach space $(\mathbb{B}, |\cdot|_{\mathbb{B}})$ is 2-smooth if there exists an equivalent norm $\|\cdot\|$ such that

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

From Pisier (1975), 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 Pisier (1975), we see that 2-smooth Banach spaces play the same role for martingales as spaces of type 2 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 a $\mathbb{B}$-valued Wiener process, in the space $D_{\mathbb{B}}([0,1])$ equipped with the uniform metric.

By a $\mathbb{B}$-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 3.1. *Assume that $\mathbb{B}$ is 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 $\mathbb{E}_{-\infty}(X_0)=0$ a.s., X_0 is $\mathcal{F}_\infty$-measurable, and*

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

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

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

Proof of Proposition 3.1: Let us prove first that the result holds if $\mathbb{E}_{-1}(X_0)=0$ almost surely, 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 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:

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

(2) for any $\varepsilon > 0$,

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

The first point can be proved exactly as in Woyczyński (1975), 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 Pinelis (1994)), in such a way that

$$\mathbb{E}\Big(\max_{1 \le k \le n} |S_k'|_{\mathbb{B}}^q \Big) \le K_q M^q n^{q/2} \quad \text{for any } q \ge 2.$$

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

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

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).

We now consider the general case. Since $\mathbb{B}$ is 2-smooth, Burkholder's inequality holds and so Proposition 3.1 in Dedecker et al. (2013a) (with $|\cdot|_{\mathbb{B}}$ instead of $|\cdot|_{\mathbb{H}}$) 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 3.1. From (3.5) it follows that the conclusion of Proposition 3.1 is also true for $(X_i)_{i \in \mathbb{Z}}$ with

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

It remains to see that this covariance function can also be written as in Proposition 3.1. Recall that since $\mathbb{E}_{-\infty}(X_0) = 0$ a.s. and X_0 is $\mathcal{F}_\infty$-measurable, for any g and h in $\mathbb{B}^*$,

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

(see the proof of Theorem 3.1 in Dedecker et al. (2013a)). Hence, for any g in $\mathbb{B}^*$,

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

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. □

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 Merlevède and Peligrad (2013). Notice that Proposition 2 of Merlevède and Peligrad (2013) 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 3.2. *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:*

$$\Big\| \max_{1\le m\le 2^r} |S_m|_{\mathbb{B}} \Big\|_p \le q\|S_{2^r}\|_p + q2^{r/p}\sum_{\ell=0}^{r-1} 2^{-\ell/p}\|\mathbb{E}_0(S_{2^\ell})\|_p + (q+1)2^{r/p}\sum_{\ell=0}^{r} 2^{-\ell/p}\|S_{2^\ell} - \mathbb{E}_{2^\ell}(S_{2^\ell})\|_p\,. \tag{3.8}$$

Remark 3.3. 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 some $p>1$, and an increasing filtration $(\mathcal{F}_i)_{i\in\mathbb{Z}}$, our proof reveals that the following inequality holds true: for any integer r,

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

Remark 3.4. Under the assumptions of Proposition 3.2, we also have that for any integer n,

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

where

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

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 Merlevède and Peligrad (2013) (see their Theorem 6).

Theorem 3.5. *Let $p > 2$ be a real number and q be its conjugate exponent. Let X_0 be a real-valued 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:*

$$\mathbb{E}\Big(\max_{1\le j\le 2^r} |S_j|^p\Big) \ll 2^r \mathbb{E}(|X_0|)^p + 2^r \left(\sum_{k=0}^{r-1} \frac{\|\mathbb{E}_0(S_{2^k})\|_p}{2^{k/p}}\right)^p + 2^r \left(\sum_{k=0}^{r} \frac{\|S_{2^k} - \mathbb{E}_{2^k}(S_{2^k})\|_p}{2^{k/p}}\right)^p + 2^r \left(\sum_{k=0}^{r-1} \frac{\|\mathbb{E}_0(S_{2^k}^2)\|_{p/2}^\delta}{2^{2\delta k/p}}\right)^{p/(2\delta)}, \quad (3.10)$$

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

Remark 3.6. The inequality in the above theorem implies that for any positive integer n,

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

To prove Remark 3.6, it suffices to use the arguments developed in the proof of Remark 3.4 together with the following additional subadditivity property: for any integers i and j, and any $\delta \in]0,1]$:

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

So, according to the first item of Lemma 37 of Merlevède and Peligrad (2013), for any integer $n \in]2^{r-1}, 2^r]$,

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

Remark 3.7. Theorem 3.5 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 number 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

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

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

As a consequence of (3.10), one can prove the following proposition which will be a key tool to prove the tightness of the sequential empirical process (1.3) in the space $\ell^\infty([0,1]\times\mathbb{R}^\ell)$ (see the proof of Theorem 2.4, Section 5).

Proposition 3.8. *Let $p>2$. Let X_0 be a real-valued 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*

$$A(X,j)=\max\Big(2\sup_{i\geq 0}\|\mathbb{E}_0(X_iX_{j+i})\|_{p/2},\ \sup_{0\leq i\leq j}\|\mathbb{E}_0(X_jX_{j+i})-\mathbb{E}(X_jX_{j+i})\|_{p/2}\Big). \tag{3.12}$$

Then, for every positive integer n,

$$\Big\|\max_{1\leq j\leq n}|S_j|\Big\|_p\ll n^{1/2}\Big(\sum_{k=0}^{n-1}|\mathbb{E}(X_0X_k)|\Big)^{1/2}+n^{1/p}\|X_1\|_p+n^{1/p}\sum_{k=1}^{n}\frac{1}{k^{1/p}}\|\mathbb{E}_0(X_k)\|_p$$
$$+\,n^{1/p}\sum_{k=1}^{2n}\frac{1}{k^{1/p}}\|X_0-\mathbb{E}_k(X_0)\|_p+n^{1/p}\Big(\sum_{k=1}^{n}\frac{1}{k^{(2/p)-1}}(\log k)^\gamma A(X,k)\Big)^{1/2}.$$

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 Merlevède and Peligrad (2013).

We would like also to point out that Theorem 3.5 implies the following Burkholder-type inequality. This has been already mentioned in the adapted case in Merlevède and Peligrad (2013, Corollary 13).

Corollary 3.9. *Let $p>2$ be a real number, X_0 be a real-valued 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:*

$$\mathbb{E}\Big(\max_{1\leq j\leq 2^r}|S_j|^p\Big)\ll 2^{rp/2}\mathbb{E}(|X_0|^p)+2^{rp/2}\Big(\sum_{j=0}^{r-1}\frac{\|\mathbb{E}_0(S_{2^j})\|_p}{2^{j/2}}\Big)^p$$
$$+\,2^{rp/2}\Big(\sum_{j=1}^{r}\frac{\|S_{2^j}-\mathbb{E}_{2^j}(S_{2^j})\|_p}{2^{j/2}}\Big)^p. \tag{3.13}$$

The above corollary (up to constants) is then the non adapted version of Peligrad et al. (2007, Theorem 1) when $p>2$.

We now give the proof of the results of this section.

Proof of Proposition 3.2: For any $k\in\{1,\dots,2^r\}$, we have

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

Consequently

$$\Big\|\max_{1\leq k\leq 2^r}|S_k|_{\mathbb{B}}\Big\|_p\leq\Big\|\max_{1\leq k\leq 2^r}|\mathbb{E}_k(S_{2^r})|_{\mathbb{B}}\Big\|_p+\Big\|\max_{1\leq m\leq 2^r-1}|\mathbb{E}_{2^r-m}(S_{2^r}-S_{2^r-m})|_{\mathbb{B}}\Big\|_p$$
$$+\,\|S_{2^r}-\mathbb{E}_{2^r}(S_{2^r})\|_p+\Big\|\max_{1\leq m\leq 2^r-1}|S_m-\mathbb{E}_m(S_m)|_{\mathbb{B}}\Big\|_p. \tag{3.14}$$

Following the proof of Proposition 2 in Merlevède and Peligrad (2013), we get

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

So, by stationarity,

$$\Big\| \max_{1\le k\le 2^r} |\mathbb{E}_k(S_{2^r})|_{\mathbb{B}} \Big\|_p + \Big\| \max_{1\le m\le 2^r-1} |\mathbb{E}_{2^r-m}(S_{2^r} - S_{2^r-m})|_{\mathbb{B}} \Big\|_p \\ \le q\, \|\mathbb{E}_{2^r}(S_{2^r})\|_p + q 2^{r/p} \sum_{\ell=0}^{r-1} 2^{-\ell/p} \|\mathbb{E}(S_{2^\ell}|\mathcal{F}_0)\|_p . \quad (3.15)$$

We now bound the last term in the right hand side of (3.14). 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

$$|S_m - \mathbb{E}_m(S_m)|_{\mathbb{B}} \le \sum_{l=0}^{r-1} |S_{m_l} - S_{m_{l+1}} - \mathbb{E}_m(S_{m_l} - S_{m_{l+1}})|_{\mathbb{B}} , \quad (3.16)$$

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

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

it follows that

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

Starting from (3.16), we then get

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

Since $(\mathbb{E}(B_{r,l}|\mathcal{F}_m))_{m\ge1}$ is a martingale, by using Doob's maximal inequality, we get

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

yielding to

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

Since

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

we derive that

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

So, by stationarity,

$$\Big\| \max_{1\le m\le 2^r-1} |S_m - \mathbb{E}_m(S_m)|_{\mathbb{B}} \Big\|_p \le (q+1) 2^{r/p} \sum_{l=0}^{r-1} 2^{-l/p} \|S_{2^l} - \mathbb{E}_{2^l}(S_{2^l})\|_p . \quad (3.17)$$

Starting from (3.14) and taking into account (3.15) and (3.17), the inequality (3.8) follows.

□

Proof of Theorem 3.5: Thanks to Proposition 3.2, it suffices to prove that the inequality (3.10) is satisfied for $\mathbb{E}\big(|S_{2^r}|^p\big)$ instead of $\mathbb{E}\big(\max_{1\le j\le 2^r} |S_j|^p\big)$. We shall use similar dyadic induction arguments as those developed in the proof of Theorem 6 in Merlevède and Peligrad (2013). With the notation $a_n = \|S_n\|_p$, we shall establish the following recurrence formula: for any positive integer n and any $p > 2$,

$$a_{2n}^p \le 2a_n^p + c_1 a_n^{p-1} \big(\|\mathbb{E}_0(S_n)\|_p + \|S_n - \mathbb{E}_n(S_n)\|_p \big) + c_2 a_n^{p-2\delta} \|\mathbb{E}_0(S_n^2)\|_{p/2}^{\delta} , \quad (3.18)$$

where c_1 and c_2 are positive constants depending only on p. Before proving it, let us show that (3.18) implies our result. With this aim, we give the following lemma which is a slight modification of Lemma 11 in Merlevède and Peligrad (2013).

Lemma 3.10. *Assume that for some* $0 < \delta \le 1$ *the recurrence formula (3.18) holds. Then, for any integer* r*,*

$$a_{2^r}^p \le 2^r \Big(4a_{2^0}^p + \Big(2c_1 \sum_{k=0}^{r-1} 2^{-k/p} \|\mathbb{E}_0(S_{2^k})\|_p \Big)^p + \Big(2c_1 \sum_{k=0}^{r-1} 2^{-k/p} \|S_{2^k} - \mathbb{E}_{2^k}(S_{2^k})\|_p \Big)^p + \Big(2c_2 \sum_{k=0}^{r-1} 2^{-2k\delta/p} \ \|\mathbb{E}_0(S_{2^k}^2)\|_{p/2}^{\delta} \Big)^{p/2\delta} \Big) . \quad (3.19)$$

Let us prove the lemma. From inequality (3.18), by recurrence on the first term, we obtain, for any positive integer r,

$$a_{2^r}^p \le 2^r \Big(a_{2^0}^p + c_1 \sum_{k=0}^{r-1} 2^{-k-1} a_{2^k}^{p-1} \|\mathbb{E}_0(S_{2^k})\|_p + c_1 \sum_{k=0}^{r-1} 2^{-k-1} a_{2^k}^{p-1} \|S_{2^k} - \mathbb{E}_{2^k}(S_{2^k})\|_p + c_2 \sum_{k=0}^{r-1} 2^{-k-1} a_{2^k}^{p-2\delta} \|\mathbb{E}_0(S_{2^k}^2)\|_{p/2}^{\delta} \Big) .$$

With the notation $B_r = \max_{0\le k\le r}(a^p_{2^k}/2^k)$, it follows that

$$B_r \le a^p_{2^0} + c_1 B_r^{1-1/p}\sum_{k=0}^{r-1} 2^{-1-k/p}\|\mathbb{E}_0(S_{2^k})\|_p$$

$$+c_1 B_r^{1-1/p}\sum_{k=0}^{r-1} 2^{-1-k/p}\|S_{2^k}-\mathbb{E}_{2^k}(S_{2^k})\|_p + c_2 B_r^{1-2\delta/p}\sum_{k=0}^{r-1} 2^{-1-2k\delta/p}\ \|\mathbb{E}_0(S^2_{2^k})\|^{\delta}_{p/2}\,.$$

Therefore, taking into account that either $B_r \le 4a^p_{2^0}$ or

$$B_r^{1/p} \le 4c_1\sum_{k=0}^{r-1} 2^{-1-k/p}\|\mathbb{E}_0(S_{2^k})\|_p \quad \text{or} \quad B_r^{1/p} \le 4c_1\sum_{k=0}^{r-1} 2^{-1-k/p}\|S_{2^k}-\mathbb{E}_{2^k}(S_{2^k})\|_p$$

or $B_r^{2\delta/p} \le 4c_2\sum_{k=0}^{r-1} 2^{-1-2k\delta/p}\ \|\mathbb{E}_0(S^2_{2^k})\|^{\delta}_{p/2}$, the inequality (3.19) follows.

To end the proof of Theorem 3.5, it remains to prove (3.18). With this aim, we denote by $\bar{S}_n = X_{n+1} + \cdots + X_{2n}$, and we write

$$S_{2n} = S_n - \mathbb{E}_n(S_n) + \mathbb{E}_n(S_n) + \bar{S}_n\,.$$

Recall now the following algebraic inequality: Let x and y be two positive real numbers and $p \ge 1$ any real number. Then

$$(x+y)^p \le x^p + y^p + 4^p(x^{p-1}y + xy^{p-1}) \tag{3.20}$$

(see Inequality (87) in Merlevède and Peligrad (2013)). The above inequality with $x = |\mathbb{E}_n(S_n) + \bar{S}_n|$ and $y = |S_n - \mathbb{E}_n(S_n)|$ gives

$$a^p_{2n} \le \|\mathbb{E}_n(S_n) + \bar{S}_n\|^p_p + \|S_n - \mathbb{E}_n(S_n)\|^p_p$$
$$+4^p\mathbb{E}\big(|\mathbb{E}_n(S_n)+\bar{S}_n|^{p-1}\times|S_n-\mathbb{E}_n(S_n)|\big)+4^p\mathbb{E}\big(|\mathbb{E}_n(S_n)+\bar{S}_n|\times|S_n-\mathbb{E}_n(S_n)|^{p-1}\big)\,.$$

Next using Hölder's inequality and stationarity, we derive that, for any $p \ge 2$,

$$a^p_{2n} \le \|\mathbb{E}_n(S_n) + \bar{S}_n\|^p_p + 2^{p-1}(1+2^{2p+1})a_n^{p-1}\|S_n - \mathbb{E}_n(S_n)\|_p\,. \tag{3.21}$$

Starting from (3.21), (3.18) will follow if we can prove that there exist two positive constants c and c_2 depending only on p such that

$$\|\mathbb{E}_n(S_n) + \bar{S}_n\|^p_p \le 2a^p_n + c\,a_n^{p-1}\|\mathbb{E}_0(S_n)\|_p + c_2 a_n^{p-2\delta}\|\mathbb{E}_0(S^2_n)\|^{\delta}_{p/2}\,. \tag{3.22}$$

This inequality can be proven by following the lines of the end of the proof of Theorem 6 in Merlevède and Peligrad (2013) replacing in their proof $x = S_n$ by $x = \mathbb{E}_n(S_n)$. However, for reader's convenience we shall give the details. The proof is divided in three cases according to the values of p.

Assume first that $2 < p \le 3$. Inequality (85) in Merlevède and Peligrad (2013) applied with $x = \mathbb{E}_n(S_n)$ and $y = \bar{S}_n$, gives

$$|\mathbb{E}_n(S_n) + \bar{S}_n|^p \le |\mathbb{E}_n(S_n)|^p + |\bar{S}_n|^p + p|\mathbb{E}_n(S_n)|^{p-1}\text{sign}(\mathbb{E}_n(S_n))\bar{S}_n$$
$$+ \frac{p(p-1)}{2}|\mathbb{E}_n(S_n)|^{p-2}\bar{S}^2_n\,.$$

But $\mathbb{E}\big(|\mathbb{E}_n(S_n)|^p\big) \le a^p_n$ and, by stationarity, $\mathbb{E}\big(|\bar{S}_n|^p\big) = a^p_n$. Moreover, Hölder's inequality combined with stationarity gives

$$\mathbb{E}\big(|\mathbb{E}_n(S_n)|^{p-1}\text{sign}(\mathbb{E}_n(S_n))\bar{S}_n\big) = \mathbb{E}\big(|\mathbb{E}_n(S_n)|^{p-1}\text{sign}(\mathbb{E}_n(S_n))\mathbb{E}_n(\bar{S}_n)\big)$$
$$\le a_n^{p-1}\|\mathbb{E}_0(S_n)\|_p\,,$$

and

$$\mathbb{E}\big(|\mathbb{E}_n(S_n)|^{p-2}\bar{S}_n^2\big) = \mathbb{E}\big(|\mathbb{E}_n(S_n)|^{p-2}\mathbb{E}_n(\bar{S}_n^2)\big) \le a_n^{p-2}\|\mathbb{E}_0(S_n^2)\|_{p/2}\,.$$

So, overall, we get

$$\|\mathbb{E}_n(S_n)+\bar{S}_n\|_p^p \le 2a_n^p + p\,a_n^{p-1}\|\mathbb{E}_0(S_n)\|_p + \frac{p(p-1)}{2}a_n^{p-2}\|\mathbb{E}_0(S_n^2)\|_{p/2}\,,$$

proving (3.22) with $\delta=1$, $c=p$ and $c_2=p(p-1)/2$.

Assume now that $p\in]3,4[$. Inequality (86) in Merlevède and Peligrad (2013) (applied with $x=\mathbb{E}_n(S_n)$ and $y=\bar{S}_n$) together with stationarity lead to

$$\|\mathbb{E}_n(S_n)+\bar{S}_n\|_p^p \le 2a_n^p + pa_n^{p-1}\|\mathbb{E}_0(S_n)\|_p + \frac{p(p-1)}{2}a_n^{p-2}\|\mathbb{E}_0(S_n^2)\|_{p/2} + 2p(p-2)^{-1}\mathbb{E}\big(|\mathbb{E}_n(S_n)||\bar{S}_n|^{p-1}\big)\,.$$

To handle the last term in the right-hand side, we notice that for any $p\ge 3$ and any positive random variables Y_0 and Y_1 such that $\mathbb{E}(Y_0^p)\le a^p$ and $\mathbb{E}(Y_1^p)\le a^p$,

$$\mathbb{E}(Y_0Y_1^{p-1}) \le a^{p-2/(p-2)}\|\mathbb{E}(Y_1|Y_0)\|_{p/2}^{1/(p-2)} \tag{3.23}$$

(see the proof of inequality (83) in Merlevède and Peligrad (2013)). Using stationarity and applying (3.23) with $Y_0=|\mathbb{E}_n(S_n)|$ and $Y_1=|\bar{S}_n|$, we get, for any $p\ge 3$,

$$\mathbb{E}\big(|\mathbb{E}_n(S_n)||\bar{S}_n|^{p-1}\big) \le a_n^{p-2/(p-2)}\|\mathbb{E}_0(S_n)\|_{p/2}^{1/(p-2)}\,. \tag{3.24}$$

So, overall, for any $p\in]3,4[$,

$$\|\mathbb{E}_n(S_n)+\bar{S}_n\|_p^p \le 2a_n^p + pa_n^{p-1}\|\mathbb{E}_0(S_n)\|_p + \frac{p(p-1)}{2}a_n^{p-2}\|\mathbb{E}_0(S_n^2)\|_{p/2} + 2p(p-2)^{-1}a_n^{p-2/(p-2)}\|\mathbb{E}_0(S_n)\|_{p/2}^{1/(p-2)}\,. \tag{3.25}$$

But, for $p\ge 3$, $\|\mathbb{E}_0(S_n^2)\|_{p/2}\le a_n^{2-2/(p-2)}\|\mathbb{E}_0(S_n^2)\|_{p/2}^{1/(p-2)}$ which together with (3.25) show that (3.22) holds with $\delta=1/(p-2)$, $c=p$ and $c_2=p(p-1)/2+2p/(p-2)$.

It remains to prove the inequality (3.22) for $p\ge 4$. Inequality (3.20) (applied with $x=\mathbb{E}_n(S_n)$ and $y=\bar{S}_n$) together with stationarity lead to

$$\|\mathbb{E}_n(S_n)+\bar{S}_n\|_p^p \le 2a_n^p + 4^p\mathbb{E}\big(|\mathbb{E}_n(S_n)|^{p-1}|\bar{S}_n|\big) + 4^p\mathbb{E}\big(|\mathbb{E}_n(S_n)||\bar{S}_n|^{p-1}\big)\,. \tag{3.26}$$

Notice that Hölder's inequality combined with stationarity entails that

$$\mathbb{E}\big(|\mathbb{E}_n(S_n)|^{p-1}|\bar{S}_n|\big) = \mathbb{E}\big(|\mathbb{E}_n(S_n)|^{p-1}\mathbb{E}_n(|\bar{S}_n|)\big) \le a_n^{p-1}\|\mathbb{E}_0(|S_n|)\|_p\,.$$

But, by Jensen's inequality, $\|\mathbb{E}_0(|S_n|)\|_p\le\|\mathbb{E}_0(S_n^2)\|_{p/2}^{1/2}$. Hence, since $p\ge 4$, by using stationarity, we derive that

$$\mathbb{E}\big(|\mathbb{E}_n(S_n)|^{p-1}|\bar{S}_n|\big) \le a_n^{p-1}\|\mathbb{E}_0(S_n^2)\|_{p/2}^{1/2} \le a_n^{p-2/(p-2)}\|\mathbb{E}_0(S_n^2)\|_{p/2}^{1/(p-2)}\,. \tag{3.27}$$

Therefore, starting from (3.26) and using the bounds (3.24) and (3.27), we get

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

proving (3.22) with $\delta=1/(p-2)$, $c=0$ and $c_2=2^{2p+1}$. □

Proof of Remark 3.7: As it is pointed out in the proof of Theorem 3.5, the remark will be proven with the help of Proposition 3.2, if we can show that

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

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 $\delta \le 1/2$.

Starting from (3.21) (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

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

The difference at this step with the proof of Theorem 3.5 is that the inequality (3.20) is used whatever $p > 2$ (in the case of real-valued random variables, we have used more precise inequalities when $p \in]2, 4[$). □

Proof of Corollary 3.9: To prove the corollary, it suffices to show that for any $0 < \delta \le 1$ and any real $p > 2$,

$$2^r \left(\sum_{k=0}^{r-1} \frac{\|\mathbb{E}_0(S_{2^k}^2)\|_{p/2}^{\delta}}{2^{2\delta k/p}}\right)^{p/(2\delta)} \ll 2^{rp/2}\|\mathbb{E}_0(X_1^2)\|_{p/2}^{p/2} + 2^{rp/2}\Big(\sum_{j=0}^{r-1} \frac{\|\mathbb{E}_0(S_{2^j})\|_p + \|S_{2^j} - \mathbb{E}_{2^j}(S_{2^j})\|_p}{2^{j/2}}\Big)^p\,, \quad (3.28)$$

and to apply Theorem 3.5.

To prove (3.28), we shall use similar arguments as those developed in the proof of Lemma 12 in Merlevède and Peligrad (2013). Setting $b_n = \|\mathbb{E}_0(S_n^2)\|_{p/2}$, assume that we can prove that, for any integer n,

$$b_{2n} \le 2b_n + 2b_n^{1/2}(\|\mathbb{E}_0(S_n)\|_p + \|S_n - \mathbb{E}_n(S_n)\|_p)\,. \quad (3.29)$$

Then, by recurrence on the first term, the above inequality will entail that for any positive integer k,

$$b_{2^k} \le 2^k b_1 + \sum_{j=0}^{k-1} 2^{k-j} b_{2^j}^{1/2}\big(\|\mathbb{E}_0(S_{2^j})\|_p + \|S_{2^j} - \mathbb{E}_{2^j}(S_{2^j})\|_p\big)\,.$$

Next, with the notation $B_k = \max_{0\le j\le k} 2^{-j} b_{2^j}$, it will follow that

$$B_k \le 2\max\Big(b_1, B_k^{1/2}\sum_{j=0}^{k-1} 2^{-j/2}\big(\|\mathbb{E}_0(S_{2^j})\|_p + \|S_{2^j} - \mathbb{E}_{2^j}(S_{2^j})\|_p\big)\Big)\,,$$

implying that

$$2^{-k} b_{2^k} \le B_k \le 2b_1 + 2^2\Big(\sum_{j=0}^{k-1} 2^{-j/2}\big(\|\mathbb{E}_0(S_{2^j})\|_p + \|S_{2^j} - \mathbb{E}_{2^j}(S_{2^j})\|_p\big)\Big)^2\,.$$

Since the above inequality clearly entails (3.28), to prove the corollary it then suffices to prove (3.29). With this aim, by using the notation $\bar{S}_n = X_{n+1} + \cdots + X_n$, we first write that $S_{2n}^2 = S_n^2 + \bar{S}_n^2 + 2\mathbb{E}_n(S_n)\bar{S}_n + 2(S_n - \mathbb{E}_n(S_n))\bar{S}_n$. Hence, by stationarity,

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

Therefore the inequality (3.29) follows from the following upper bounds: applying Cauchy-Schwarz inequality twice and using stationarity, we get

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

and

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

□

Proof of Remark 3.4: Let n and r be integers such that $2^{r-1} \le n < 2^r$. Notice first that

$$\Big\|\max_{1\le k\le n}|S_k|_{\mathbb{B}}\Big\|_p \le \Big\|\max_{1\le k\le 2^r}|S_m|_{\mathbb{B}}\Big\|_p \text{ and } \|S_{2^r}\|_p \le 2\|S_{2^{r-1}}\|_p \le 2\max_{1\le k\le n}\|S_k\|_p \tag{3.30}$$

(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 $n, m \ge 0$, $V_{n+m} \le V_n + V_m$ and then, according to the first item of Lemma 37 of Merlevède and Peligrad (2013),

$$\begin{aligned}2^{r/p}\sum_{\ell=0}^{r-1} 2^{-\ell/p}\|\mathbb{E}_0(S_{2^\ell})\|_p &\le n^{1/p}\frac{2^{1/p}2^{2+1/p}}{2^{1+1/p}-1}\sum_{k=1}^{n}\frac{\|\mathbb{E}_0(S_k)\|_p}{k^{1+1/p}} \\ &\le n^{1/p}\frac{2^{1+1/p}}{1-2^{-1/p-1}}\sum_{k=1}^{n}\frac{\|\mathbb{E}_0(S_k)\|_p}{k^{1+1/p}}\,.\end{aligned} \tag{3.31}$$

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

Claim 3.11. *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 $p \ge 1$, $\|X - \mathbb{E}(X|\mathcal{F})\|_p \le 2\|X - \mathbb{E}(X|\mathcal{G})\|_p$.*

The above claim together with the stationarity imply that for all $n, m \ge 0$, $W_{n+m} \le 2(W_n + W_m)$. Therefore, using once again the first item of Lemma 37 of Merlevède and Peligrad (2013), we get

$$2^{r/p}\sum_{\ell=0}^{r} 2^{-\ell/p}\|S_{2^\ell} - \mathbb{E}_{2^\ell}(S_{2^\ell})\|_p \le 2n^{1/p}\frac{2^{1+1/p}}{1-2^{-1/p-1}}\sum_{\ell=1}^{2n}\frac{\|S_\ell - \mathbb{E}_\ell(S_\ell)\|_p}{\ell^{1+1/p}}\,. \tag{3.32}$$

The inequality (3.9) then follows from the inequality (3.8) by taking into account the upper bounds (3.30), (3.31) and (3.32). □

3.2. *A tightness criterion.* We begin with the definition of the number of brackets of a family of functions.

Definition 3.12. 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^1_P. Let $\mathcal{F}$ be some subset of L^1_P. The number of brackets $\mathcal{N}_{P,1}(\varepsilon, \mathcal{F})$ is the smallest integer N for which there exist some functions $f_1^- \le f_1, \dots, f_N^- \le f_N$ in $\mathcal{F}$ such that: for any integer $1 \le i \le N$ we have $\|f_i - f_i^-\|_{P,1} \le \varepsilon$, and for any function f in $\mathcal{F}$ there exists an integer $1 \le i \le N$ such that $f_i^- \le f \le f_i$.

Proposition 3.13 below gives a general tightness criterion for empirical processes. Its proof is based on a decomposition given in Andrews and Pollard (1994) (see also Dedecker and Prieur (2007)). Under the setting and conditions of Theorem 2.4, the criterion (3.33) will be shown to hold with the help of Proposition 3.8 (see the proof of Theorem 2.4 in Section 5).

Proposition 3.13. *Let $(X_i)_{i\ge 1}$ 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 $r \ge 2$, $p > 2$ and $C > 0$ such that for any function g of $\mathcal{G}\cup\mathcal{F}$ and any positive integer n, we have*

$$\Big\| \max_{1\le k\le n} |S_k(g)| \Big\|_p \le C(\sqrt{n}\|g\|_{P,1}^{1/r} + n^{1/p}), \tag{3.33}$$

where $S_k(g) := \sum_{i=1}^k (g(X_i) - P(g))$. If moreover

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

then

$$\lim_{\delta\to 0}\limsup_{n\to\infty} \mathbb{E}\Big(\max_{1\le k\le n}\sup_{g\in\mathcal{G}, \|g\|_{P,1}\le\delta} n^{-p/2}|S_k(g)|^p\Big) = 0, \tag{3.34}$$

$$\text{and} \quad \lim_{\delta\to 0}\limsup_{n\to\infty} \frac{1}{\delta}\mathbb{E}\Big(\max_{1\le k\le [n\delta]}\sup_{f\in\mathcal{F}} n^{-p/2}|S_k(f)|^p\Big) = 0. \tag{3.35}$$

Proof of Proposition 3.13: It is almost the same as that of Proposition 6 in Dedecker and Prieur (2007). 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 $f_1^{k,-} \le f_1^k, \dots, f_{\mathcal{N}_k}^{k,-} \le f_{\mathcal{N}_k}^k$ in $\mathcal{F}$ such that $\|f_i^k - f_i^{k,-}\|_{P,1} \le 2^{-k}$, and for any f in $\mathcal{F}$, there exists an integer $1 \le i \le \mathcal{N}_k$ such that $f_i^{k,-} \le f \le f_i^k$.

We follow exactly the proof of Proposition 6 in Dedecker and Prieur (2007). For reader's convenience, we give the key details. For any f in $\mathcal{F}$, there exist two functions g_k^- and g_k^+ in $\mathcal{F}_k$ such that $g_k^- \le f \le g_k^+$ and $\|g_k^+ - g_k^-\|_{P,1} \le 2^{-k}$. Hence, for any $1 \le j \le n$,

$$S_j(f) - S_j(g_k^-) \le S_j(g_k^+) - S_j(g_k^-) + \sum_{i=1}^j \mathbb{E}((g_k^+ - f)(X_i)) \le |S_j(g_k^+) - S_j(g_k^-)| + j2^{-k}.$$

Since $g_k^- \le f$, we also have that $S_j(g_k^-) - S_j(f) \le j2^{-k}$, which enables us to conclude that

$$|S_j(f) - S_j(g_k^-)| \le |S_j(g_k^+) - S_j(g_k^-)| + j2^{-k}.$$

Consequently

$$\sup_{f\in\mathcal{F}} |S_j(f) - S_j(g_k^-)| \le \max_{1\le i\le \mathcal{N}_k} |S_j(f_i^k) - S_j(f_i^{k,-})| + j2^{-k} \,. \tag{3.36}$$

Notice now the following elementary fact: given N real-valued random variables $Z_1, \ldots, Z_N$, we have

$$\| \max_{1\le i\le N} |Z_i| \|_p \le N^{1/p} \max_{1\le i\le N} \|Z_i\|_p \,. \tag{3.37}$$

Combining (3.37) and (3.36), we obtain

$$\Big\| \max_{1\le j\le n} \sup_{f\in\mathcal{F}} |S_j(f) - S_j(g_k^-)| \Big\|_p \le \mathcal{N}_k^{1/p} \max_{1\le i\le \mathcal{N}_k} \| \max_{1\le j\le n} |S_j(f_i^k) - S_j(f_i^{k,-})| \|_p + n2^{-k} \,. \tag{3.38}$$

Starting from (3.38) and applying (3.33), we obtain

$$\Big\| \max_{1\le j\le n} \sup_{f\in\mathcal{F}} n^{-1/2} |S_j(f) - S_j(g_k^-)| \Big\|_p \le C(\mathcal{N}_k^{1/p} 2^{-k/r} + \mathcal{N}_k^{1/p} n^{1/p-1/2}) + \sqrt{n} 2^{-k} \,. \tag{3.39}$$

By the arguments developed right after the inequality (4.6) in Dedecker and Prieur (2007), we infer that there exists a sequence $h_{k(n)}(f)$ belonging to $\mathcal{F}_{k(n)}$ such that

$$\lim_{n\to\infty} \Big\| \max_{1\le j\le n} \sup_{f\in\mathcal{F}} n^{-1/2} |S_j(f) - S_j(h_{k(n)}(f))| \Big\|_p = 0 \,. \tag{3.40}$$

We prove now that for any $\varepsilon > 0$, there exist $N(\varepsilon)$ and $m = m(\varepsilon)$ such that : for any $n \ge N(\varepsilon)$ there exists a function $f_{n,m}$ in $\mathcal{F}_m$ such that

$$\Big\| \max_{1\le j\le n} \sup_{f\in\mathcal{F}} n^{-1/2} |S_j(f_{n,m}) - S_j(h_{k(n)}(f))| \Big\|_p \le \varepsilon \,. \tag{3.41}$$

Given h in $\mathcal{F}_k$, choose a function $T_{k-1}(h)$ in $\mathcal{F}_{k-1}$ such that $\|h - T_{k-1}(h)\|_{P,1} \le 2^{-k+1}$. Denote by $\pi_{k,k} = Id$ and for $l < k$, $\pi_{l,k}(h) = T_l \circ \cdots \circ T_{k-1}(h)$. We consider the function $f_{n,m} = \pi_{m,k(n)}(h_{k(n)}(f))$. For the sake of brevity, we write $h_{k(n)}$ instead of $h_{k(n)}(f)$. We have that

$$\Big\| \max_{1\le j\le n} \sup_{f\in\mathcal{F}} |S_j(f_{n,m}) - S_j(h_{k(n)})| \Big\|_p \le \sum_{l=m+1}^{k(n)} \Big\| \max_{1\le j\le n} \sup_{f\in\mathcal{F}} |S_j(\pi_{l,k(n)}(h_{k(n)})) - S_j(\pi_{l-1,k(n)}(h_{k(n)}))| \Big\|_p \,. \tag{3.42}$$

Clearly

$$\Big\| \max_{1\le j\le n} \sup_{f\in\mathcal{F}} |S_j(\pi_{l,k(n)}(h_{k(n)})) - S_j(\pi_{l-1,k(n)}(h_{k(n)}))| \Big\|_p \le \Big\| \max_{1\le j\le n} \max_{f\in\mathcal{F}_l} |S_j(f) - S_j(T_{l-1}(f))| \Big\|_p \,.$$

Using then (3.33) combined with (3.37), it follows that

$$\Big\| \max_{1\le j\le n} \sup_{f\in\mathcal{F}} n^{-1/2} |S_j(f_{n,m}) - S_j(h_{k(n)})| \Big\|_p \le C \sum_{l=m+1}^{k(n)} (2^{1/r} \mathcal{N}_l^{1/p} 2^{-l/r} + \mathcal{N}_l^{1/p} n^{1/p-1/2}) \,.$$

To complete the proof of (3.41) we use the same arguments as in Dedecker and Prieur (2007), page 130.

Combining (3.40) and (3.41), it follows 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

$$\Big\| \max_{1\leq k\leq n} \sup_{f\in\mathcal{F}} n^{-1/2}|S_k(f) - S_k(f_{n,m})| \Big\|_p \leq 2\varepsilon \,. \tag{3.43}$$

Using the same argument as in Andrews and Pollard (1994) (see the paragraph "Comparison of pairs" page 124), we obtain

$$\Big\| \max_{1\leq k\leq n} \sup_{\substack{f,g\in\mathcal{F}\\ \|f-g\|_{P,1}\leq\delta}} n^{-1/2}|S_k(f) - S_k(g)| \Big\|_p \leq 8\varepsilon + \mathcal{N}_m^{2/p} \sup_{\substack{f,g\in\mathcal{F}\\ \|f-g\|_{P,1}\leq 2\delta}} \Big\| \max_{1\leq k\leq n} n^{-1/2}|S_k(f) - S_k(g)| \Big\|_p \,.$$

Since by (3.33),

$$\sup_{\substack{f,g\in\mathcal{F}\\ \|f-g\|_{P,1}\leq 2\delta}} \Big\| \max_{1\leq k\leq n} n^{-1/2}|S_k(f) - S_k(g)| \Big\|_p \leq C((2\delta)^{1/r} + n^{1/p-1/2}) \,,$$

it follows that

$$\Big\| \max_{1\leq k\leq n} \sup_{\substack{f,g\in\mathcal{F}\\ \|f-g\|_{P,1}\leq\delta}} n^{-1/2}|S_k(f) - S_k(g)| \Big\|_p \leq 8\varepsilon + C\mathcal{N}_m^{2/p}((2\delta)^{1/r} + n^{1/p-1/2}) \,,$$

which proves (3.34).

Let us now prove (3.35). We apply (3.43) with $\varepsilon = 1$: for $n \geq \delta^{-1}N(1)$, we infer from (3.43) that there exists $f_{[n\delta],m}$ in $\mathcal{F}_m$ for which

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

Hence

$$\Big\| \max_{1\leq k\leq [n\delta]} \sup_{f\in\mathcal{F}} n^{-1/2}|S_k(f)| \Big\|_p \leq \sqrt{\delta} + \Big\| \max_{1\leq k\leq [n\delta]} \sup_{f\in\mathcal{F}} n^{-1/2}|S_k(f_{[n\delta],m})| \Big\|_p \,. \tag{3.44}$$

Now, since $\mathcal{F}_m$ contains $2\mathcal{N}_m$ functions $(g_\ell)_{\ell\in\{1,\dots,2\mathcal{N}_m\}}$ (each g_ℓ being one of the functions f_i^m or $f_i^{m,-}$ in $\mathcal{F}_m$), it follows that

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

Let $K_m = \max_{f\in\mathcal{F}_m} \|f\|_{P,1}$. Applying (3.33), we infer that

$$\Big\| \max_{1\leq k\leq [n\delta]} \sup_{f\in\mathcal{F}} n^{-1/2}|S_k(f_{[n\delta],m})| \Big\|_p \leq 2C\mathcal{N}_m(K_m^{1/r}\sqrt{\delta} + n^{-(p-2)/2p}\delta^{1/p}) \,. \tag{3.45}$$

Since $m = m(1)$ is fixed, (3.35) follows from (3.44) and (3.45) and the fact that $p > 2$. □

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

$$\Big\|\sum_{i=1}^d x_i v_i\Big\| = \max_{i=1,\dots,d} |x_i|$$

and $d_0(\cdot,\cdot)$ be 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 by: for every $\delta > 0$,

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

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

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 $S^{-1}_{|E_u}$. For every $\rho_u \in (r_u, 1)$, there exists $K > 0$ such that, for every integer $n \ge 0$, we have

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

and

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

The following inequality can be viewed as an extension to continuous functions of a result for Hölder functions established in Le Borgne and Pène (2005) but with a σ-algebra satisfying $\mathcal{F}_0 \subseteq T^{-1}\mathcal{F}_0$ (this condition is not satisfied in the construction of $\mathcal{F}_0$ considered in Le Borgne and Pène (2005)). For the next result, we shall then use the construction of $\mathcal{F}_0$ given in Lind (1982); Le Borgne (1999) combined with some arguments developed in Le Borgne and Pène (2005).

Theorem 4.1. *Let $\rho_u \in (r_u, 1)$ and $\zeta \in (\rho_u^{1/(3(d+2)(d_e+d_s))}, 1)$. There exist $C > 0$, $N \ge 0$, $\xi \in (0,1)$, a sequence of measurable sets $(\mathcal{V}_n)_{n\ge0}$ 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 $n \ge N$, we have*

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

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

and

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

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

Remark 4.2. With the notations of Theorem 4.1, (4.5) and (4.6) imply that, for every $p \geq 1$ and every (ρ_u, ζ) as in Theorem 4.1, 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 4.1 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 Lind (1982); Le Borgne (1999) 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:

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

Next, for every integer n, we consider the σ-algebras $\mathcal{F}_n$ 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 4.3 (Lind (1982); Le Borgne (1999) applied to T^{-1}, see also Dedecker et al. (2013b))**.** *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 set $\bar{x} + \overline{F_{\bar{x}}}$, with $0 \in F_{\bar{x}} \subseteq E_u$ and such that $F_{\bar{x}}$ is a uniformly bounded convex set 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

$$\partial\mathcal{C}(\beta) := \{y \in \mathcal{C} \ : \ d_0(y, \partial\mathcal{C}) \leq \beta\} \ \text{ for any } \mathcal{C} \subseteq E_u .$$

Recall now an exponential decorrelation result for Lipschitz continuous functions.

Proposition 4.4 (Lind (1982) and also section 4.1 of Pène (2002))**.** *There exist $C_0 > 0$ and $\xi_0 \in (0, 1)$ 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 4.3. The following result is an adaptation of Proposition 1.3 of Le Borgne and Pène (2005).

Proposition 4.5. *Let $\zeta_1 \in (\xi_0^{1/((d+2)(d_e+d_s))}, 1)$ where ξ_0 is given in Proposition 4.4. 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 $r := \xi_0^{-1/(d+2)}$. We take $\varepsilon_n = \alpha^n$ with $\alpha \in (0,1)$ such that $\zeta_1 > \alpha > \xi_0^{1/((d+2)(d_e+d_s))} \geq r^{-1}$ and n such that $\alpha^n < r_0$. Let $U := \overline{T^{-n}\bar x + \mathcal{C} + B_s(\varepsilon_n) + B_e(\varepsilon_n)}$. We have $T^n(U) = \overline{\bar x + S^n\mathcal{C} + S^nB_s(\varepsilon_n) + S^nB_e(\varepsilon_n)}$. We have

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

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

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

Hence, due to (4.3), we have

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

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 $Kn^{d_e}\varepsilon_n \leq \zeta_1^n$),

$$\begin{aligned}
\left|\frac{1}{\bar\lambda(U)}\int_{\mathbb{T}^d} \mathbf{1}_{T^nU}\varphi\, d\bar\lambda - \frac{1}{m_u(S^n\mathcal{C})}\int_{S^n\mathcal{C}} \varphi(\bar x + \overline{h_u})\, dm_u(h_u)\right| &\leq \omega_{(s,e)}(\varphi, Kn^{d_e}\varepsilon_n) \\
&\leq \omega_{(s,e)}(\varphi, \zeta_1^n).
\end{aligned}$$

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 non-negative $(d+1)r^{n(d+1)}2^{-d}$-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

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

Second, we have

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

and let us prove that

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

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 $\bar x\in\partial U(r^{-n})$ where $\partial U(r^{-n}) := \{x\in U\ :\ d_1(x,\partial U)<r^{-n}\}$, 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

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

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

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

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

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

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

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

Moreover, applying Proposition 4.4 with $f=\chi_n*\varphi$ and $g=\chi_n*\mathbf{1}_U$ and using the following facts

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

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

$$\begin{aligned}\frac{1}{\bar\lambda(U)}\left|\int_{\mathbb{T}^d}\mathbf{1}_U\circ T^{-n}.\varphi\,d\bar\lambda\right| &\le \tilde C_0\|\varphi\|_\infty\frac{r^{-n}+(1+r^{n(d+1)})\xi_0^n}{\varepsilon_n^{d_e+d_s}m_u(\mathcal{C})}+\omega(\varphi,r^{-n})\\ &\le 3\tilde C_0\|\varphi\|_\infty\frac{\xi_0^{n/(d+2)}}{\varepsilon_n^{d_e+d_s}m_u(\mathcal{C})}+\omega(\varphi,\zeta_1^n),\end{aligned}$$

since $r^{-1}=r^{d+1}\xi_0=\xi_0^{1/(d+2)}$. We take then $\xi_1:=\xi_0^{1/(d+2)}\alpha^{-(d_e+d_s)}<1$, which concludes the proof. □

In the next result (which is an adaptation of Proposition 1.4 of Le Borgne and Pène (2005)), we prove that Proposition 4.5 holds true with the stable-neutral continuity modulus $\omega_{(s,e)}$ instead of ω.

Proposition 4.6. *Let $\zeta_1\in(\xi_0^{1/((d+2)(d_e+d_s))},1)$ where ξ_0 is given in Proposition 4.4. There exist $C_2>0$, $N_2\ge 1$ 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 $n\ge N_2$ and every bounded convex set $\mathcal{C}\subseteq E_u$ with diameter smaller than r_0 and satisfying $m_u(\partial\mathcal{C}(\beta))\le Q\gamma$, we have*

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

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

$$\begin{aligned}\omega(\psi_i, \delta) &\le m_u(B_u(r_0)) \left[\|\varphi\|_\infty Lip(g)\delta + \|g\|_\infty \omega_{(s,e)}(\varphi_i, \delta)\right] \\ &\le m_u(B_u(r_0)) \Big[\|\varphi\|_\infty Lip(g)\delta + \|g\|_\infty \|\varphi\|_\infty Lip(H_i)\delta \\ &\qquad + \|g\|_\infty \omega_{(s,e)}(\varphi, \delta)\|H_i\|_\infty\Big].\end{aligned}$$

Now, applying Proposition 4.5 to ψ_i, for every $n \ge N_1$, we have

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

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

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

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

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

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

Proof of Theorem 4.1: We start by proving the first point. By Proposition 4.3,

$$\mathbb{E}[\varphi|\mathcal{F}_n](\bar{x})-\varphi(\bar{x}) = \frac{1}{m_u(S^{-n}F_{T^n\bar{x}})}\int_{S^{-n}F_{T^n\bar{x}}}\big(\varphi(\bar{x}+\overline{h_u})-\varphi(\bar{x})\big)\,dm_u(h_u)\,. \quad (4.16)$$

Let $h_u \in S^{-n}F_{T^n\bar{x}}$ and $y \in F_{T^n\bar{x}}$ such that $h_u = S^{-n}(y)$. Take now $\beta_u \in (r_u, \rho_u)$. From (4.2) and the fact that $F_{T^n\bar{x}}$ is uniformly bounded, we derive that there exists a positive constant C such that $\|h_u\| \leq C\beta_u^n$. Therefore, starting from (4.16), by definition of $\omega_{(u)}(\varphi,\delta)$, we get

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

The first point of Theorem 4.1 then comes from the fact that there exists $N > 0$ such that for any $n \geq N$, $C\beta_u^n \leq \rho_u^n$.

We turn now to the proof of the second point. Let ζ_1, C_2, ξ_2 and N_2 as in Proposition 4.6 with $\zeta_1 < \zeta$. Let $\beta \in (\xi_2, 1)$ and $\mathcal{V}_n := \{m_u(F_\cdot) \geq \beta^n\}$. 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 4.3 and we apply Proposition 4.6 with $\mathcal{C} = F_{T^{-n}(\bar{x})}$ with the notation of Proposition 4.3.

It remains to prove the last point of the theorem. It comes from the fact (proved in Proposition II.1 of Le Borgne (1999)) that

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

□

5. Proof of Theorems 2.1 and 2.4

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 $a_n \leq Cb_n$, for all positive integers n.

Proof of Theorem 2.1: The proof is based on Proposition 3.1 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 Pisier (1975)). Moreover it has a Schauder basis (and even an unconditional basis).

Hence it suffices to check (3.2) of Proposition 3.1. According to Lemma 6.1 of Dedecker et al. (2011) (with $b_k = 1$), 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

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

for some positive constant K. 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

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

with $F_{f \circ T^n | \mathcal{F}_0}(t) := \mathbb{P}(f \circ T^n \leq t | \mathcal{F}_0)$. 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 4.1 that

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

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 4.1, 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. □

Proof of Theorem 2.4: Our aim is to apply the tightness criterion given in Proposition 3.13. 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 5.1. *Let $g_{s,t}(v) = \mathbf{1}_{v \le t} - \mathbf{1}_{v \le s}$, and let P be the image measure of $\bar\lambda$ by f. Under the assumptions of Theorem 2.4, 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 5.2. *Under the assumptions of Theorem 2.4, we have, for any $p \ge 1$,*

$$\|\mathbb{E}_0(g_{s,t}(X_k)) - \mathbb{E}(g_{s,t}(X_k))\|_p \ll k^{-a\alpha/(\alpha+p)}$$
$$\|g_{s,t}(X_0) - \mathbb{E}_k(g_{s,t}(X_0))\|_p \ll k^{-a\alpha/(\alpha+p)} ,$$

and, for any $p \ge 2$,

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

where the coefficient $A(g_{s,t}(X) - \mathbb{E}[g_{s,t}(X)], j)$ is defined in (3.12). The constants involved in the symbol $\ll$ do not depend on (s,t).

Let us continue the proof of Theorem 2.4 with the help of these lemmas. From Proposition 3.8, Lemma 5.1 and Lemma 5.2, we derive that, for $p > 2$,

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

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

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

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

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

We shall apply the tightness criterion given in Proposition 3.13. Since $\mathcal{N}_{P,1}(x, \mathcal{F}) \le Cx^{-\ell}$ for the class $\mathcal{F} = \{u \mapsto \mathbf{1}_{u \le t}, t \in \mathbb{R}^\ell\}$, we get

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

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

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

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 the conditions (3.34) and (3.35) of Proposition 3.13 hold for this choice of a, which proves the tightness of the empirical process (see van der Vaart and Wellner (1996), page 227).

To prove the weak convergence of the finite dimensional distribution, it suffices to show that for any $(\alpha_1,\dots,\alpha_m) \in \mathbb{R}^m$ and any $(s_1,\dots,s_m) \in (\mathbb{R}^\ell)^m$, the process

$$\Big\{n^{-1/2}\sum_{i=1}^m \alpha_i S_{[nt]}(s_i)\,,\, t\in[0,1]\Big\} \quad \text{converges in distribution in } D_{\mathbb{R}}([0,1]) \text{ to } W,$$

where W is a Wiener process such that

$$\mathrm{Cov}(W_{t_1}, W_{t_2}) = \min(t_1,t_2)\sum_{i=1}^m\sum_{j=1}^m \alpha_i\alpha_j\Lambda(s_i,s_j)\,.$$

Note that $\sum_{i=1}^m \alpha_i S_{[nt]}(s_i) = \sum_{k=1}^{[nt]} Y_k$ where $Y_k = \sum_{i=1}^m \alpha_i\big(\mathbf{1}_{X_k\leq s_i} - F(s_i)\big)$. Therefore, the above convergence in distribution will follow from Proposition 5 in Dedecker et al. (2007) if we can prove that

$$\sum_{k=1}^\infty \frac{\|\mathbb{E}_0(Y_k)\|_2}{\sqrt{k}} < \infty \quad\text{and}\quad \sum_{k=1}^\infty \frac{\|Y_0 - \mathbb{E}_k(Y_0)\|_2}{\sqrt{k}} < \infty\,. \tag{5.3}$$

By the triangle inequality, it suffices to prove that (5.3) holds with $\mathbf{1}_{X_k\leq s} - F(s)$ in place of Y_k. This follows from Lemma 5.2 as soon as $a > (\alpha+2)/2\alpha$. □

Proof of Lemma 5.1: 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

$$g_{s,t} = g_{s,t,\varepsilon} + H_{s,t,\varepsilon}, \quad \text{with} \quad H_{s,t,\varepsilon} = R_{t,\varepsilon} - R_{s,\varepsilon}\,. \tag{5.4}$$

On the other hand, we have

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

Using (5.4), we get

$$\begin{aligned}&\mathbb{E}((g_{s,t}(X_0) - \mathbb{E}(g_{s,t}(X_0)|\mathcal{F}_{[k/2]}))g_{s,t}(X_k)) \\ &\quad = \mathbb{E}((g_{s,t,\varepsilon}(X_0) - \mathbb{E}(g_{s,t,\varepsilon}(X_0)|\mathcal{F}_{[k/2]}))g_{s,t}(X_k)) \\ &\qquad + \mathbb{E}((H_{s,t,\varepsilon}(X_0) - \mathbb{E}(H_{s,t,\varepsilon}(X_0)|\mathcal{F}_{[k/2]}))g_{s,t}(X_k))\,. \end{aligned} \tag{5.5}$$

Applying (4.4) of Theorem 4.1, we infer that

$$|\mathbb{E}((g_{s,t,\varepsilon}(X_0) - \mathbb{E}(g_{s,t,\varepsilon}(X_0)|\mathcal{F}_{[k/2]}))g_{s,t}(X_k))| \le C\|g_{s,t}\|_{P,1}\varepsilon^{-1}\omega_{(u)}(f, \rho_u^{[k/2]}) . \tag{5.6}$$

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

$$|\mathbb{E}((H_{s,t,\varepsilon}(X_0) - \mathbb{E}(H_{s,t,\varepsilon}(X_0)|\mathcal{F}_{[k/2]}))g_{s,t}(X_k))| \le C\|g_{s,t}\|_{P,1}^{(\beta-1)/\beta}\varepsilon^{\alpha/\beta} . \tag{5.7}$$

Using (5.4) again, we also have

$$\begin{aligned}\mathrm{Cov}(\mathbb{E}(g_{s,t}(X_0)|\mathcal{F}_{[k/2]}), g_{s,t}(X_k)) &= \mathrm{Cov}(\mathbb{E}(g_{s,t}(X_0)|\mathcal{F}_{[k/2]}), g_{s,t,\varepsilon}(X_k)) \\ &\quad + \mathrm{Cov}(\mathbb{E}(g_{s,t}(X_0)|\mathcal{F}_{[k/2]}), H_{s,t,\varepsilon}(X_k)) . \end{aligned} \tag{5.8}$$

To handle the first term in the right-hand side, we set $g^{(0)}_{s,t,\varepsilon}(X_0) = g_{s,t,\varepsilon}(X_0) - \mathbb{E}(g_{s,t,\varepsilon}(X_0))$ and note first that

$$\begin{aligned}\mathrm{Cov}(\mathbb{E}(g_{s,t}(X_0)|\mathcal{F}_{[k/2]}), g_{s,t,\varepsilon}(X_k)) &= \mathbb{E}(\mathbb{E}(g_{s,t}(X_{-k})|\mathcal{F}_{[k/2]-k})g^{(0)}_{s,t,\varepsilon}(X_0)) \\ &= \mathbb{E}(g_{s,t}(X_{-k})\mathbb{E}(g^{(0)}_{s,t,\varepsilon}(X_0)|\mathcal{F}_{[k/2]-k})) .\end{aligned}$$

Therefore, considering the set $\mathcal{V}_n$ introduced in Theorem 4.1, it follows that

$$\begin{aligned}|\mathrm{Cov}(\mathbb{E}(g_{s,t}(X_0)|\mathcal{F}_{[k/2]}), g_{s,t,\varepsilon}(X_k))| &\le 2\|g_{s,t,\varepsilon}(X_0)\|_\infty \mathbb{E}(|g_{s,t}(X_{-k})|\mathbf{1}_{\mathcal{V}^c_{k-[k/2]}}) \\ &\quad + \mathbb{E}(|g_{s,t}(X_{-k})||\mathbb{E}(g_{s,t,\varepsilon}(X_0)|\mathcal{F}_{[k/2]-k}) - \mathbb{E}(g_{s,t,\varepsilon}(X_0))|\mathbf{1}_{\mathcal{V}_{k-[k/2]}}) .\end{aligned}$$

On one hand, applying (4.5) of Theorem 4.1 with $\varphi = g_{s,t,\varepsilon} \circ f$ and using the fact that, since $h_{u,\varepsilon}$ is Lipschitz with Lipschitz constant ε^{-1}, $\omega_{(s,e)}(g_{s,t,\varepsilon} \circ f, \zeta^{[k/2]}) \le 4\varepsilon^{-1}\omega_{(s,e)}(f, \zeta^{[k/2]})$, we infer that

$$\begin{aligned}&\mathbb{E}(|g_{s,t}(X_{-k})||\mathbb{E}(g_{s,t,\varepsilon}(X_0)|\mathcal{F}_{[k/2]-k}) - \mathbb{E}(g_{s,t,\varepsilon}(X_0))|\mathbf{1}_{\mathcal{V}_{k-[k/2]}}) \\ &\qquad \le C\|g_{s,t}\|_{P,1}(\xi^{[k/2]} + \varepsilon^{-1}\omega_{(s,e)}(f, \zeta^{[k/2]})) .\end{aligned}$$

On the other hand, since $\bar{\lambda}(\mathcal{V}^c_{k-[k/2]}) \le C\xi^{[k/2]}$, applying Hölder's inequality, we get

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

So, overall,

$$\begin{aligned}|\mathrm{Cov}(\mathbb{E}(g_{s,t}(X_0)|\mathcal{F}_{[k/2]}), g_{s,t,\varepsilon}(X_k))| &\le C\|g_{s,t}\|_{P,1}^{(\beta+\alpha-1)/(\beta+\alpha)}\xi^{[k/2]/(\beta+\alpha)} \\ &\quad + C\|g_{s,t}\|_{P,1}(\xi^{[k/2]} + \varepsilon^{-1}\omega_{(s,e)}(f, \zeta^{[k/2]})) .\end{aligned} \tag{5.9}$$

We handle now the second term in the right-hand side of (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

$$|\mathrm{Cov}(\mathbb{E}(g_{s,t}(X_0)|\mathcal{F}_{[k/2]}), H_{s,t,\varepsilon}(X_k))| \le C\|g_{s,t}\|_{P,1}^{(\beta-1)/\beta}\varepsilon^{\alpha/\beta} . \tag{5.10}$$

Therefore, starting from (5.8) and considering (5.9) and (5.10), it follows that

$$\begin{aligned}|\mathrm{Cov}(\mathbb{E}(g_{s,t}(X_0)|\mathcal{F}_{[k/2]}), g_{s,t}(X_k))| &\le C\|g_{s,t}\|_{P,1}^{(\beta+\alpha-1)/(\beta+\alpha)}\xi^{[k/2]/(\beta+\alpha)} \\ &\quad + C\|g_{s,t}\|_{P,1}(\xi^{[k/2]} + \varepsilon^{-1}\omega_{(s,e)}(f, \zeta^{[k/2]}))C\|g_{s,t}\|_{P,1}^{(\beta-1)/\beta}\varepsilon^{\alpha/\beta} .\end{aligned} \tag{5.11}$$

Gathering the bounds (5.5), (5.6), (5.7) and (5.11), it follows that

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

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

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

The result follows by summing in k. □

Proof of Lemma 5.2: Using the same notations as in the proof of Lemma 5.1, and using that the distribution functions of f_1 and f_2 are Hölder continuous of order α, we obtain

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

Recall that the set $\mathcal{V}_n$ introduced in Theorem 4.1 is such that $\bar\lambda(\mathcal{V}_n^c) \le C\xi^n$. Applying Theorem 4.1 (see (4.7)), we obtain

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

Consequently

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

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

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

proving the first inequality.

In the same way

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

Applying (4.4) of Theorem 4.1, we obtain

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

Since $\omega_{(u)}(f, \rho_u^k) \le Ck^{-a}$, 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

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

Using the decomposition (5.4), and the fact that the distribution functions of f_1 and f_2 are Hölder continuous of order α, we get

$$\|\mathbb{E}_0(g_{s,t}^{(0)}(X_i) g_{s,t}^{(0)}(X_{j+i}))\|_{p/2} \le \|\mathbb{E}_0(g_{s,t,\varepsilon}^{(0)}(X_i) g_{s,t,\varepsilon}^{(0)}(X_{j+i}))\|_{p/2} + C\varepsilon^{2\alpha/p} , \quad (5.12)$$

and

$$\|\mathbb{E}_0(g_{s,t}^{(0)}(X_j) g_{s,t}^{(0)}(X_{j+i})) - \mathbb{E}(g_{s,t}^{(0)}(X_j) g_{s,t}^{(0)}(X_{j+i}))\|_{p/2} \le \|\mathbb{E}_0(g_{s,t,\varepsilon}^{(0)}(X_j) g_{s,t,\varepsilon}^{(0)}(X_{j+i})) - \mathbb{E}(g_{s,t,\varepsilon}^{(0)}(X_j) g_{s,t,\varepsilon}^{(0)}(X_{j+i}))\|_{p/2} + C\varepsilon^{2\alpha/p} . \quad (5.13)$$

Writing

$$\begin{aligned}
&\|\mathbb{E}_0(g^{(0)}_{s,t,\varepsilon}(X_i)g^{(0)}_{s,t,\varepsilon}(X_{j+i}))\|_{p/2} \\
&\qquad \le \|\mathbb{E}_0((g_{s,t,\varepsilon}(X_i)-\mathbb{E}(g_{s,t,\varepsilon}(X_i)|\mathcal{F}_{i+[j/2]}))g^{(0)}_{s,t,\varepsilon}(X_{j+i}))\|_{p/2} \\
&\qquad\qquad + \|\mathbb{E}_0(\mathbb{E}(g_{s,t,\varepsilon}(X_i)|\mathcal{F}_{i+[j/2]})g^{(0)}_{s,t,\varepsilon}(X_{j+i}))\|_{p/2}\,, \qquad (5.14)
\end{aligned}$$

and arguing as in Lemma 5.1, we infer that

$$\|\mathbb{E}_0(g^{(0)}_{s,t,\varepsilon}(X_i)g^{(0)}_{s,t,\varepsilon}(X_{j+i}))\|_{p/2} \le C\Big(\frac{1}{\varepsilon j^a}+\xi^{[j/2]}\Big)\,. \qquad (5.15)$$

From (5.12) and (5.15), we obtain the bound

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

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

$$\sup_{i\ge 0}\|\mathbb{E}_0(g^{(0)}_{s,t}(X_i)g^{(0)}_{s,t}(X_{j+i}))\|_{p/2} \le Cj^{-2a\alpha/(2\alpha+p)}\,. \qquad (5.16)$$

Let $\varphi := g_{s,t,\varepsilon}\circ f - \bar\lambda(g_{s,t,\varepsilon}\circ f)$. Applying Theorem 4.1 (see (4.7)), for $i\le j$,

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

By (4.3), $\omega_{(s,e)}(\varphi.\varphi\circ T^i,\zeta^j) \le 2\|\varphi\|_\infty\omega_{(s,e)}(\varphi,K\zeta^j j^{d_e})) \le 4\omega_{(s,e)}(\varphi,L\zeta_0^j)$ where $\zeta_0\in(\zeta,1)$. Hence,

$$\begin{aligned}
&\|\mathbb{E}_0(g^{(0)}_{s,t,\varepsilon}(X_j)g^{(0)}_{s,t,\varepsilon}(X_{j+i})) - \mathbb{E}(g^{(0)}_{s,t,\varepsilon}(X_j)g^{(0)}_{s,t,\varepsilon}(X_{j+i}))\|_{p/2} \\
&\qquad \le C(\xi^{2j/p}+\omega_{(s,e)}(\varphi,L\zeta_0^j))\,.
\end{aligned} \qquad (5.17)$$

Since $\omega_{(s,e)}(\varphi,L\zeta_0^j) \le \varepsilon^{-1}\omega_{(s,e)}(f,L\zeta_0^j) \le C\varepsilon^{-1}j^{-a}$, we obtain from (5.13) and (5.17) that

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

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

$$\sup_{0\le i\le j}\|\mathbb{E}_0(g^{(0)}_{s,t}(X_j)g^{(0)}_{s,t}(X_{j+i})) - \mathbb{E}(g^{(0)}_{s,t}(X_j)g^{(0)}_{s,t}(X_{j+i}))\|_{p/2} \le Cj^{-2a\alpha/(2\alpha+p)}\,. \qquad (5.18)$$

The third inequality of Lemma 5.2 follows from (5.16), (5.18) and from the definition of the quantity $A(g_{s,t}(X)-\mathbb{E}(g_{s,t}(X)),j)$ given in Proposition 3.8. □

6. Additional results for partial sums

Let T be an 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 $\omega(f,\cdot)$.

The inequalities given in Theorem 4.1 have been used to prove the tightness of the sequential empirical process, but they can be used in many other situations. Let us give three examples of application to the behavior of the partial sums (1.1).

(1) **Moment bounds for partial sums**. Using Corollary 3.9 together with Theorem 4.1 (see also Remark 4.2), we infer that if

$$\sum_{n>0} \frac{\omega(f,\zeta^n)}{\sqrt{n}} < \infty\,, \tag{6.1}$$

where $\zeta \in (0,1)$ is defined in Theorem 4.1, then for any $p > 2$,

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

Clearly, the condition (6.1) is equivalent to the integral condition

$$\int_0^{1/2} \frac{\omega(f,t)}{t|\log t|^{1/2}} dt < \infty\,. \tag{6.2}$$

(2) **Weak invariance principle**. If the integral condition (6.2) holds then the series

$$\sigma^2(f) = \bar\lambda((f-\bar\lambda(f))^2) + 2\sum_{k>0} \bar\lambda((f-\bar\lambda(f))\cdot f\circ T^k) \tag{6.3}$$

converges absolutely, and the process

$$\Big\{ \frac{1}{\sqrt{n}} \sum_{k=1}^{[nt]} (f\circ T^k - \bar\lambda(f)), t\in[0,1] \Big\}$$

converges to a Wiener process with variance $\sigma^2(f)$ in the space $D([0,1])$ of càdlàg function equipped with the uniform metric. This follows from Theorem 4.1 together with Proposition 5 in Dedecker et al. (2007).

(3) **Rates of convergence in the strong invariance principle**. Let $p \in]2,4]$, and assume that $\omega(f,x) \le C|\log(x)|^{-a}$ 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 $(Z_i)_{i\ge1}$ of independent and identically distributed Gaussian random variables with mean zero and variance $\sigma^2(f)$ defined in (6.3) such that, for any $t > 2/p$,

$$\sup_{1\le k\le n} \Big| \sum_{i=1}^{k} (f\circ T^i - \bar\lambda(f)) - \sum_{i=1}^{k} Z_i \Big| = o\big(n^{1/p}(\log(n))^{(t+1)/2}\big) \quad \text{a. s.}$$

as n tends to infinity. In particular, we obtain the rate of convergence $n^{1/2-\epsilon}$ for some $\epsilon > 0$ as soon as $a > 1/2$, and the rate $n^{1/4}\log(n)$ as soon as $a \ge 3/2$. This follows from Theorem 4.1 together with Theorem 3.1 in Dedecker et al. (2013b).

Appendix

In this section, we prove Remark 2.5, 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 2.5). 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

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

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.

Acknowledgements

The authors would like to thank the two referees for carefully reading the manuscript and for numerous suggestions that improved the presentation of this paper.

References

D. W. K. Andrews and D. Pollard. An introduction to functional central limit theorems for dependent stochastic processes. *International Statistical Review / Revue Internationale de Statistique* **62** (1), 119–132 (1994). JSTOR 1403549.

J. Dedecker, F. Merlevède and M. Peligrad. Invariance principles for linear processes with application to isotonic regression. *Bernoulli* **17** (1), 88–113 (2011). MR2797983.

J. Dedecker, F. Merlevède and F. Pène. Rates in the strong invariance principle for ergodic automorphisms of the torus (2013a). To appear in Stoch. Dyn. arXiv: 1206.4336.

J. Dedecker, F. Merlevède and F. Pène. Rates of convergence in the strong invariance principle for non-adapted sequences; application to ergodic automorphisms of the torus. In Christian Houdré, David M. Mason, Jan Rosiński and Jon A. Wellner, editors, *High Dimensional Probability VI*, volume 66 of *Progress in Probability*, pages 113–138. Springer Basel (2013b). ISBN 978-3-0348-0489-9. DOI: 10.1007/978-3-0348-0490-5_9.

J. Dedecker, F. Merlevède and D. Volný. On the weak invariance principle for non-adapted sequences under projective criteria. *J. Theoret. Probab.* **20** (4), 971–1004 (2007). MR2359065.

J. Dedecker and C. Prieur. An empirical central limit theorem for dependent sequences. *Stochastic Process. Appl.* **117** (1), 121–142 (2007). MR2287106.

H. Dehling and O. Durieu. Empirical processes of multidimensional systems with multiple mixing properties. *Stochastic Process. Appl.* **121** (5), 1076–1096 (2011). MR2775107.

O. Durieu and P. Jouan. Empirical invariance principle for ergodic torus automorphisms; genericity. *Stoch. Dyn.* **8** (2), 173–195 (2008). MR2429199.

J. Kiefer. Skorohod embedding of multivariate RV's, and the sample DF. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **24** (1), 1–35 (1972). MR1554013.

S. Le Borgne. Limit theorems for non-hyperbolic automorphisms of the torus. *Israel J. Math.* **109**, 61–73 (1999). MR1679589.

S. Le Borgne and F. Pène. Vitesse dans le théorème limite central pour certains systèmes dynamiques quasi-hyperboliques. *Bull. Soc. Math. France* **133** (3), 395–417 (2005). MR2169824.

D. A. Lind. Dynamical properties of quasihyperbolic toral automorphisms. *Ergodic Theory Dynamical Systems* **2** (1), 49–68 (1982). MR684244.

F. Merlevède and M. Peligrad. Rosenthal-type inequalities for the maximum of partial sums of stationary processes and examples. *Ann. Probab.* **41** (2), 914–960 (2013). MR3077530.

M. Peligrad, S. Utev and W. B. Wu. A maximal $\mathbb{L}_p$-inequality for stationary sequences and its applications. *Proc. Amer. Math. Soc.* **135** (2), 541–550 (electronic) (2007). MR2255301.

F. Pène. Averaging method for differential equations perturbed by dynamical systems. *ESAIM Probab. Statist.* **6**, 33–88 (electronic) (2002). MR1905767.

I. Pinelis. Optimum bounds for the distributions of martingales in Banach spaces. *Ann. Probab.* **22** (4), 1679–1706 (1994). MR1331198.

G. Pisier. Martingales with values in uniformly convex spaces. *Israel J. Math.* **20** (3-4), 326–350 (1975). MR0394135.

A. van der Vaart and J. Wellner. *Weak convergence and empirical processes.* Springer Series in Statistics. Springer-Verlag, New York (1996). ISBN 0-387-94640-3. With applications to statistics. MR1385671.

W. A. Woyczyński. A central limit theorem for martingales in Banach spaces. *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.* **23** (8), 917–920 (1975). MR0385961.