Bernoulli **21**(1), 2015, 374–400
DOI: 10.3150/13-BEJ571

A compact LIL for martingales in 2-smooth Banach spaces with applications

CHRISTOPHE CUNY

Laboratoire MAS, Ecole Centrale de Paris, Grande Voie des Vignes, 92295 Chatenay-Malabry cedex, France. E-mail: christophe.cuny@ecp.fr

We prove the compact law of the iterated logarithm for stationary and ergodic differences of (reverse or not) martingales taking values in a separable 2-smooth Banach space (for instance a Hilbert space). Then, in the martingale case, the almost sure invariance principle is derived from a result of Berger. From those results, we deduce the almost sure invariance principle for stationary processes under the Hannan condition and the compact law of the iterated logarithm for stationary processes arising from non-invertible dynamical systems. Those results for stationary processes are new, even in the real valued case. We also obtain the Marcinkiewicz–Zygmund strong law of large numbers for stationary processes with values in some smooth Banach spaces. Applications to several situations are given.

Keywords: Banach valued martingales; compact law of the iterated logarithm; Hannan's condition; strong invariance principle

1. Introduction

Let $(\mathcal{X}, |\cdot|_{\mathcal{X}})$ be a separable Banach space and $\mathcal{X}^*$ be its topological dual. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $(X_n)_{n\geq 0}$ be a strictly stationary sequence of $\mathcal{X}$-valued random variables. We are interested in the $\mathbb{P}$-a.s. behaviour of $(S_n/\sqrt{2nL(L(n))})_{n\geq 1}$, where $S_n := X_0 + \cdots + X_{n-1}$ and $L := \max(1, \log)$.

Definition 1.1. *We say that $(X_n)_{n\geq 0}$ satisfies the bounded law of the iterated logarithm* (*bounded LIL or BLIL*) *if $(S_n/\sqrt{2nL(L(n))})_{n\geq 1}$ is $\mathbb{P}$-a.s. bounded.*

Definition 1.2. *We say that $(X_n)_{n\geq 0}$ satisfies the compact law of the iterated logarithm* (*compact LIL or CLIL*) *if $(S_n/\sqrt{2nL(L(n))})_{n\geq 1}$ is $\mathbb{P}$-a.s. relatively compact.*

When $(X_n)_{n\geq 0}$ is a sequence of independent random variables, the bounded and compact LILs are well understood, thanks to a characterization due to Ledoux and Talagrand [22]. When the compact LIL holds, the cluster set of $S_n/\sqrt{2nL(L(n))})_{n\geq 1}$ may be identified thanks to a result of Kuelbs [21]. When X_0 is *pregaussian* (see next section), we have an almost sure invariance principle as well.

For Banach spaces of type 2 (see next section for the definition), the result of Ledoux–Talagrand takes the following particularly simple form.

Theorem 1.1 (Ledoux and Talagrand [23], Corollary 8.8). *Let $(X_n)_{n\geq 0}$ be a sequence of i.i.d. random variables with values in a Banach space of type 2. Then, $(X_n)_{n\geq 0}$ satisfies the* *bounded LIL* (*resp. the compact LIL*) *if and only if* $\mathbb{E}((x^*(X_0))^2) < \infty$ *for every* $x^* \in \mathcal{X}^*$ (*resp.* $((x^*(X_0))^2)_{x^*\in\mathcal{X}^*,|x^*|_{\mathcal{X}^*}\le 1}$ *is uniformly integrable*), $\mathbb{E}(|X_0|^2_{\mathcal{X}}/L(L(|X_0|_{\mathcal{X}}))) < \infty$ *and* $\mathbb{E}(X_0) = 0$.

In particular, a sequence of i.i.d. variables $(X_n)_{n\ge 0}$ with values in a Banach space of type 2 satisfies the compact LIL (hence, the bounded LIL) as soon as:

$$\mathbb{E}\big(|X_0|^2_{\mathcal{X}}\big) < \infty \quad \text{and} \quad \mathbb{E}(X_0) = 0. \tag{1}$$

Now (see Remark 2.4), by a result of Pisier [30], if $\mathcal{X}$ is a Banach space for which any sequence of $\mathcal{X}$-valued i.i.d. variables, such that (1) holds, satisfies the bounded LIL, then, $\mathcal{X}$ must be of type p for any $1 < p < 2$.

We are interested here in the case where $(X_n)_{n\ge 0}$ is a general stationary sequence, including the case of martingale differences (and of reverse martingale differences). The analogue of the notion of Banach space of type 2 in the case of martingale differences is the notion of 2-smooth Banach space (see the next section for the definition). One could wonder whether Theorem 1.1 is true in this context, or, at least, whether (1) is sufficient for the bounded LIL or the compact LIL, when $(X_n)_{n\ge 0}$ is a stationary sequence of martingale differences.

As far as we know, the latter question remained unsolved. Let us however mention some results in that direction. Morrow and Philipp [27] (see also [28] for an improved version) obtained an almost sure invariance principle (see the next section for the definition), hence a compact LIL (with an ad hoc normalization), for sequences of non-necessarily stationary martingale differences taking values in a Hilbert space. Dehling, Denker and Philipp [16] proved a bounded LIL in the same context. When applied to stationary sequences of martingale differences, the above results require higher moments than 2.

In this paper, we prove that condition (1) is sufficient for the compact LIL when $(X_n)_{n\ge 0}$ is a stationary sequence of martingale differences with values in a 2-smooth Banach space. When the sequence is ergodic, the cluster set of $(S_n/\sqrt{2nL(L(n))})_{n\ge 1}$ is identified as well as $\limsup_n |S_n|_{\mathcal{X}}/\sqrt{nL(L(n))}$. Then, using a result of Berger [2], we obtain an almost sure invariance principle for $(S_n)_{n\ge 1}$. Those results (except for the invariance principle) extend to *reverse* martingale differences. We do not know whether our results could be extended beyond the scope of 2-smooth Banach spaces. However, the above mentioned result of Pisier shows some limitation.

To prove those results, we first obtain integrability properties of the "natural" maximal function arising in that context, hence generalizing a result of Pisier [30] for i.i.d. variables. This step is crucial not only to prove the results for martingales (and reverse martingales), but also in order to extend the results to general stationary processes under projective conditions, such as the Hannan condition, see Theorem 2.10 or the Maxwell–Woodroofe condition, see Cuny [4]. We note that the almost sure invariance principle for Hilbert-valued stationary processes under mixing conditions have been obtained by Merlevède [26] and Dedecker and Merlevède [13]. Their results have different range of applications.

We also investigate the Marcinkiewicz–Zygmund strong law of large numbers for stationary processes taking values in a smooth Banach space. The maximal function arising in that other context has been studied by Woyczyński [31], for stationary martingale differences. We investigate the case of stationary processes under projective conditions. The main argument used is the same as the one for the law of the iterated logarithm. The Marcinkiewicz–Zygmund strong laws in smooth Banach spaces have been also studied by Dedecker and Merlevède [12] for stationary processes satisfying mixing conditions.

In the next section, we set our notations and state our results for martingales and then, for stationary processes, including non-adapted processes, functionals of Markov chains or iterates of non-invertible dynamical systems. In Section 3, we give several examples to which our conditions apply. In Section 4, we prove our martingale results and in Section 5 we prove our results for stationary processes. Finally, we postpone some technical proofs or results to the Appendix.

2. Main results

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. We will consider Banach-valued random variables. We refer to the book by Ledoux and Talagrand [23] for the basic facts on the topic (definition, conditional expectation...).

Let $(\mathcal{X}, |\cdot|_{\mathcal{X}})$ be a separable *real* Banach space. We endow $\mathcal{X}$ with its Borel σ-algebra. Denote by $L^0(\mathcal{X})$ the space (of classes modulo $\mathbb{P}$) of measurable random variables on Ω taking values in $\mathcal{X}$. We define, for every $p \geq 1$, the usual Bochner spaces L^p and their weak versions, as follows

$$L^p(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{X}) = \{Z \in L^0(\mathcal{X}) \colon \mathbb{E}(|Z|_{\mathcal{X}}^p) < \infty\};$$

$$L^{p,\infty}(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{X}) = \Big\{Z \in L^0(\mathcal{X}) \colon \sup_{t>0} t\big(\mathbb{P}(|Z|_{\mathcal{X}} > t)\big)^{1/p} < \infty\Big\}.$$

For every $Z \in L^p(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{X})$, write $\|Z\|_{p,\mathcal{X}} := (\mathbb{E}(|Z|_{\mathcal{X}}^p))^{1/p}$ and for every $Z \in L^{p,\infty}(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{X})$, write $\|Z\|_{p,\infty,\mathcal{X}} := \sup_{t>0} t(\mathbb{P}(|Z|_{\mathcal{X}} > t))^{1/p}$.

For the sake of clarity, when they are understood, some of the references to Ω, $\mathcal{F}$ or $\mathbb{P}$ may be omitted. Also, in the case when $\mathcal{X} = \mathbb{R}$, we shall simply write $\|\cdot\|_p$ or $\|\cdot\|_{p,\infty}$. Recall that for every $p > 1$ there exists a norm on $L^{p,\infty}(\mathbb{P}, \mathcal{X})$ (see, for instance, [23], Chapter "Notation"), equivalent to the quasi-norm $\|\cdot\|_{p,\infty,\mathcal{X}}$, that makes $L^{p,\infty}(\mathbb{P}, \mathcal{X})$ a Banach space.

The Banach spaces we will consider are the so-called *smooth* Banach spaces.

Definition 2.1. *We say that $\mathcal{X}$ is r-*smooth, *for some $1 < r \leq 2$, if there exists $L \geq 1$, such that*

$$|x+y|_{\mathcal{X}}^r + |x-y|_{\mathcal{X}}^r \leq 2\big(|x|_{\mathcal{X}}^r + L^r|y|_{\mathcal{X}}^r\big) \qquad \forall x, y \in \mathcal{X}.$$

Definition 2.2. *We say that $(d_n)_{1\leq n\leq N} \subset L^1(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{X})$ is a sequence of martingale differences, if there exists non-decreasing σ-algebras $(\mathcal{G}_n)_{0\leq n\leq N}$ such that for every $1 \leq n \leq N$, d_n is $\mathcal{G}_n$-measurable and $\mathbb{E}(d_n|\mathcal{G}_{n-1}) = 0$ $\mathbb{P}$-a.s. If $(\mathcal{G}_n)_{1\leq n\leq N+1}$ is non-increasing and $\mathbb{E}(d_n|\mathcal{F}_{n+1}) = 0$ $\mathbb{P}$-a.s., we speak about differences of* reverse *martingales.*

It is known, see, for instance, Proposition 1 of Assouad [1] (and its corollary), that when $\mathcal{X}$ is r-smooth, there exists $D \geq 1$, such that for every martingale differences $(d_n)_{1\leq n\leq N}$, we have

$$\mathbb{E}\big(|d_1 + \cdots + d_N|_{\mathcal{X}}^r\big) \leq D^r \sum_{n=1}^{N} \mathbb{E}\big(|d_n|_{\mathcal{X}}^r\big). \qquad (2)$$

When needed, we will say that $\mathcal{X}$ is (r, D)-smooth, where D is a constant such that condition (2) is satisfied (notice that this definition is compatible with the definition page 1680 of [29], see Proposition 2.5 there). Clearly, D must be greater than 1.

Any L^p space, $p > 1$ (of $\mathbb{R}$-valued functions), associated with a σ-finite measure is r-smooth for $r = \min(2, p)$ (one may take $D^2 = p - 1$ if $p \geq 2$, see [29], Proposition 2.1, and $D^2 = 2$ if $1 \leq p < 2$ by [1]). Any Hilbert space is (2, 1)-smooth.

Definition 2.3. *We say that $\mathcal{X}$ is a Banach space of type r, $1 < r \leq 2$, if* (2) *holds for every finite set $(d_n)_{1\leq n\leq N}$ of independent variables. Hence*, 2*-smooth Banach spaces are particular examples of spaces of type* 2.

Our goal is to study the law of the iterated logarithm and the Marcinkiewicz–Zygmund strong law of large numbers for the partial sums of an $\mathcal{X}$-valued stationary process. We will start by studying the maximal functions associated with these limit theorems. Let us specify some notations.

Let θ be a measurable measure preserving transformation on Ω. To any $X \in L^0(\Omega, \mathcal{X})$, we associate a stationary process $(X \circ \theta^n)_{n\geq 0}$ (when θ is invertible, we extend that definition to $n \in \mathbb{Z}$). Then, for every $n \geq 1$, write $S_n(X) = \sum_{i=0}^{n-1} X \circ \theta^i$.

We shall assume that there exists a suitable filtration on Ω. In order to cover more situations, we shall consider filtrations that are either non-decreasing or non-increasing. In spirit, the first case arise when θ is invertible and the second one when θ is non-invertible.

In particular, we assume that we are in one of the following situations.

If $\mathcal{F}_0 \subset \mathcal{F}$ is a σ-algebra such that $\mathcal{F}_0 \subset \theta^{-1}(\mathcal{F}_0)$, we define a *non-decreasing* filtration $(\mathcal{F}_n)_{n\geq 0}$ by $\mathcal{F}_n := \theta^{-n}(\mathcal{F}_0)$. Define then $\mathbb{E}_n = \mathbb{E}(\cdot|\mathcal{F}_n)$.

If $\mathcal{F}^0$ is such that $\theta^{-1}(\mathcal{F}^0) \subset \mathcal{F}^0$ (for instance, take $\mathcal{F}^0 = \mathcal{F}$), we define a *non-increasing* filtration $(\mathcal{F}^n)_{n\geq 0}$, by $\mathcal{F}^n := \theta^{-n}(\mathcal{F}^0)$. Define then $\mathbb{E}^n = \mathbb{E}(\cdot|\mathcal{F}^n)$.

Let $1 \leq p \leq 2$. Let $X \in L^p(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{X})$. We consider the following maximal functions

$$\mathcal{M}_p(X) := \sup_{n\geq 1} \frac{|\sum_{k=0}^{n-1} X \circ \theta^k|_{\mathcal{X}}}{n^{1/p}}, \qquad \text{if } 1 \leq p < 2, \tag{3}$$

$$\mathcal{M}_2(X) := \sup_{n\geq 1} \frac{|\sum_{k=0}^{n-1} X \circ \theta^k|_{\mathcal{X}}}{\sqrt{nL(L(n))}}, \tag{4}$$

where $L := \max(\log, 1)$.

The maximal operator $\mathcal{M}_1$ is related to Birkhoff's ergodic theorem, which asserts that for every $X \in L^1(\Omega, \mathcal{X})$, $((\sum_{k=0}^{n-1} X \circ \theta^k)/n)_{n\geq 1}$ converges $\mathbb{P}$-a.s. (see Theorem 2.1, page 167 of [20] for the $\mathcal{X}$-valued case). For every $X \in L^1(\Omega, \mathcal{X})$, by Hopf's dominated ergodic theorem for real-valued stationary processes (see [20], Corollary 2.2, page 8), applied to $(|X|_{\mathcal{X}} \circ \theta^n)_{n\geq 0}$, we have

$$\left\|\mathcal{M}_1(X)\right\|_{1,\infty} \leq \|X\|_{1,\mathcal{X}}. \tag{5}$$

Now, once we know that (5) holds, by the Banach principle (see [20], Theorem 7.2, page 64, or Proposition C.1), in order to prove Birkhoff's ergodic theorem, it suffices to prove it on a set of X's dense in L^1 (e.g., the θ invariant elements and the coboundaries). We want to use that strategy to study the Marcinkiewicz–Zygmund strong law of large numbers and versions of the law of the iterated logarithm. Of course, one cannot expect to have a version of (5) for $\mathcal{M}_p$, when $1 < p \le 2$ without any further assumption on $(X \circ \theta^n)_{n\ge 0}$.

2.1. Results for stationary (reverse) martingale differences

In this subsection, we consider stationary sequences of (reverse) martingale differences.

Let $d \in L^p(\Omega, \mathcal{F}_1, \mathcal{X})$ be such that $\mathbb{E}_0(d) = 0$ $\mathbb{P}$-a.s. Then, by our assumptions on $\mathcal{F}_0$, $(d \circ \theta^n)_{n\ge 0}$ is a stationary sequence of martingale differences.

Let $d \in L^p(\Omega, \mathcal{F}^0, \mathcal{X})$ be such that $\mathbb{E}^1(d) = 0$ $\mathbb{P}$-a.s. Then, by our assumption on $\mathcal{F}^0$, $(d \circ \theta^n)_{n\ge 0}$ is a stationary sequence of reverse martingale differences, that is, for every $n \ge 0$, $d \circ \theta^n$ is $\mathcal{F}^n$-measurable and $\mathbb{E}(d \circ \theta^n | \mathcal{F}^{n+1}) = 0$ $\mathbb{P}$-a.s.

There is no loss of generality in assuming that our stationary sequences of (reverse) martingale differences are given that way.

Indeed, it is well known (see, e.g., Doob [17], page 456) that, given a stationary sequence $(\tilde{d}_n)_{n\ge 1}$ on a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$, there exist another probability space $(\Omega, \mathcal{F}, \mathbb{P})$, an invertible bi-measurable measure-preserving transformation θ on Ω and a random variable d on Ω such that the sequences $(\tilde{d}_n)_{n\ge 1}$ and $(d \circ \theta^n)_{n\ge 1}$ have the same law.

Moreover, it follows from the construction, that if $(\tilde{d}_n)_{n\ge 1}$ are martingale differences (respectively, reverse martingale differences), $(d \circ \theta^n)_{n\ge 1}$ are martingale differences (respectively, reverse martingale differences) either.

Hence, since all the results we are concerned with in that paper only rely on the distribution of the processes under consideration, we shall assume (without loss of generality) that our stationary sequences of martingale differences are given thanks to a measure-preserving transformation.

We start with a result of Woyczyński about the Marcinkiewicz–Zygmund strong law of large numbers.

Proposition 2.1 (Woyczyński [31]). *Let $1 < p < r \le 2$ and $D \ge 1$. Let $\mathcal{X}$ be a separable (r, D)-smooth Banach space. There exists $C_{p,r} > 0$ such that for every $d \in L^p(\Omega, \mathcal{F}_1, \mathcal{X})$ (resp. $d \in L^p(\Omega, \mathcal{F}^0, \mathcal{X})$), with $\mathbb{E}_0(d) = 0$ (resp. $\mathbb{E}^1(d) = 0$), we have*

$$\left\| \mathcal{M}_p(d) \right\|_{p,\infty} \le C_{p,r} D^{r/p} \|d\|_{p,\mathcal{X}}. \tag{6}$$

Moreover,

$$\left| S_n(d) \right|_{\mathcal{X}} / n^{1/p} \to 0 \qquad \mathbb{P}\text{-a.s.} \tag{7}$$

Remark 2.2. Actually, Woyczyński proved that $\mathcal{M}_p(d)$ is in any L^r, $r < p$ and worked with martingale differences (not differences of reverse martingales). But his argument applies to obtain the above proposition. We give the proof of (6) in the Appendix, for completeness. The proof of (7) is done in [31]. The argument is very similar to the scalar case. Actually by the Banach principle (see Proposition C.1), using (6), it is enough to show (7) in the scalar case, see for instance the proof of Theorem 2.3.

Next, we obtain a similar result for $\mathcal{M}_2$, from which we derive the compact LIL for stationary martingale differences (or reverse martingale differences).

Theorem 2.3. *Let $\mathcal{X}$ be a $(2, D)$-smooth separable Banach space, for some $D \geq 1$. For every $1 \leq p < 2$, there exists a constant $C_p \geq 1$, such that for every $d \in L^2(\Omega, \mathcal{F}_1, \mathcal{X})$ (resp. every $d \in L^2(\Omega, \mathcal{F}^0, \mathcal{X})$) with $\mathbb{E}_0(d) = 0$ (resp. $\mathbb{E}^1(d) = 0$), we have*

$$\left\| \mathcal{M}_2(d) \right\|_{p,\infty} \leq C_p D \|d\|_{2,\mathcal{X}}. \tag{8}$$

In particular, $(d \circ \theta^n)_{n\geq 0}$ satisfies the compact LIL. Moreover, if θ is ergodic,

$$\limsup_n \frac{|S_n(d)|_{\mathcal{X}}}{\sqrt{2nL(L(n))}} = \sup_{x^* \in \mathcal{X}^*, |x^*|_{\mathcal{X}^*} \leq 1} \left\| x^*(d) \right\|_2 \leq \|d\|_{2,\mathcal{X}} \qquad \mathbb{P}\text{-a.s.} \tag{9}$$

and the cluster set of $(\frac{S_n(d)}{\sqrt{2nL(L(n))}})_{n\geq 1}$ is $\mathbb{P}$-a.s. a fixed compact set whose description is given in Appendix D.

Remark 2.4. Of course, (8) is equivalent to the fact that, for every $1 \leq p < 2$, there exists $\tilde{C}_p$, such that $\|\mathcal{M}_2(d)\|_p \leq \tilde{C}_p D \|d\|_{2,\mathcal{X}}$. This bound has been obtained in [30], Théorème 1, for i.i.d. variables with values in a Banach space of type 2. Moreover, it follows from Remarque 2 and the proposition page 208 of [30], that if every sequence of i.i.d. variables in $L^2(\Omega, \mathcal{X})$ satisfy the bounded LIL, the space $\mathcal{X}$ must be of type p, for every $1 < p < 2$.

Now, we deduce an almost sure invariance principle (ASIP) from Theorem 2.3. We first give the notations to specify what we mean by an ASIP, in the Banach space setting.

Recall, that we denote by $\mathcal{X}^*$ the topological dual of $\mathcal{X}$. Let $X \in L^2(\Omega, \mathcal{X})$ such that $\mathbb{E}(X) = 0$. We define a bounded *symmetric* bilinear operator $\mathcal{K} = \mathcal{K}_X$ from $\mathcal{X}^* \times \mathcal{X}^*$ to $\mathbb{R}$, by

$$\mathcal{K}(x^*, y^*) = \mathbb{E}\big(x^*(X) y^*(X)\big) \qquad \forall x^*, y^* \in \mathcal{X}^*.$$

The operator $\mathcal{K}_X$ is called the *covariance operator* associated with X.

Definition 2.4. *We say that a random variable $W \in L^2(\Omega, \mathcal{X})$ is* Gaussian *if, for every $x^* \in \mathcal{X}^*$, $x^*(W)$ has a normal distribution. We say that a random variable $X \in L^2(\Omega, \mathcal{X})$ is* pregaussian, *if there exists a Gaussian variable $W \in L^2(\Omega, \mathcal{X})$ with the same covariance operator, that is, such that $\mathcal{K}_X = \mathcal{K}_W$.*

Definition 2.5. *We say that $(X_n)_{n\geq 0}$ satisfies the almost sure invariance principle (ASIP) if, without changing its distribution, one can redefine the sequence $(X_n)_{n\geq 0}$ on a new probability space on which there exists a sequence $(W_n)_{n\geq 0}$ of centered i.i.d. Gaussian variables, such that*

$$\left| X_0 + \cdots + X_{n-1} - (W_0 + \cdots + W_{n-1}) \right|_{\mathcal{X}} = o\Big(\sqrt{nL\big(L(n)\big)}\Big) \qquad \mathbb{P}\text{-a.s.}$$

We shall say that $(X_n)_{n\geq 0}$ satisfies the ASIP of covariance $\mathcal{K}$, when $\mathcal{K} = \mathcal{K}_{W_0}$ is identified.

We now recall an important result of Berger on the ASIP for martingale differences.

Proposition 2.5 (Berger [2], Theorem 3.2). *Let $\mathcal{X}$ be a separable Banach space. Assume that θ is ergodic. Let $d \in L^2(\Omega, \mathcal{F}_1, \mathcal{X})$, with $\mathbb{E}_0(d) = 0$. Assume that d is pregaussian and that $(d \circ \theta^n)_{n\geq 0}$ satisfies the CLIL. Then, for every $Y \in L^2(\Omega, \mathcal{X})$, such that $|S_n(Y)|_{\mathcal{X}} = \mathrm{o}(\sqrt{nL(L(n))})$ $\mathbb{P}$-a.s., $((d+Y)\circ\theta^n)_{n\geq 0}$ satisfies the ASIP of covariance $\mathcal{K}_d$.*

Actually, Berger proved his result in the particular case where $Y = Z - Z\circ\theta$ for some $Z \in L^2(\Omega, \mathcal{X})$, but the proof applies in the slightly more general situation above.

By [23], Proposition 9.24, on any Banach space $\mathcal{X}$ of type 2 (in particular, on any 2-smooth Banach space), every $X \in L^2(\Omega, \mathcal{X})$ is pregaussian. Hence, Berger's result applies as soon as the CLIL is satisfied and we deduce the following corollary.

Corollary 2.6. *Let $\mathcal{X}$ be a 2-smooth separable Banach space. Assume that θ is ergodic. For every $d \in L^2(\Omega, \mathcal{F}_1, \mathcal{X})$, with $\mathbb{E}_0(d) = 0$, $(d\circ\theta^n)_{n\geq 0}$ satisfies the ASIP of covariance $\mathcal{K}_d$.*

Remark 2.7. Assume that $\dim\mathcal{X} = 1$ and that θ is ergodic. It follows from Corollary 2.5 of [6] that for $d \in L^2(\Omega, \mathcal{F}^0, \mathcal{X})$ such that $\mathbb{E}^1(d) = 0$, $(d\circ\theta^n)_{n\geq 0}$ satisfies the ASIP. We do not know whether the ASIP holds when $\dim\mathcal{X} \geq 2$. The proof of Proposition 2.5 given in [2] does not seem to pass to *reverse* martingale differences.

2.2. Results for not necessarily adapted stationary processes

We assume all along this subsection that θ is invertible and bi-measurable, in which case we extend our filtration to $(\mathcal{F}_n)_{n\in\mathbb{Z}}$. Then, we write $\mathcal{F}_{-\infty} := \bigcap_{n\in\mathbb{Z}} \mathcal{F}_n$, $\mathcal{F}_\infty := \bigvee_{n\in\mathbb{Z}} \mathcal{F}_n$, and for every $n \in \overline{\mathbb{Z}}$, $\mathbb{E}_n(\cdot) = \mathbb{E}(\cdot|\mathcal{F}_n)$ and $P_n := \mathbb{E}_n - \mathbb{E}_{n-1}$. We say that a random variable $X \in L^1(\Omega, \mathcal{X})$ is *regular* if $\mathbb{E}_{-\infty}(X) = 0$ and $X - \mathbb{E}_\infty(X) = 0$.

Theorem 2.8. *Let $1 < p < r \leq 2$ and $D > 0$. Let $\mathcal{X}$ be a (r, D)-smooth separable Banach space and $X \in L^p(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{X})$ be a regular variable. Assume moreover that*

$$\|X\|_{H_p} := \sum_{n\in\mathbb{Z}} \|P_n X\|_{p,\mathcal{X}} < \infty. \tag{10}$$

Then, there exists (a universal) $C_{p,r} > 0$, such that

$$\left\|\mathcal{M}_p(X)\right\|_{p,\infty} \leq C_{p,r} D^{r/p} \|X\|_{H_p}. \tag{11}$$

Moreover,

$$\left|S_n(X)\right|_{\mathcal{X}}/n^{1/p} \to 0 \qquad \mathbb{P}\text{-a.s.} \tag{12}$$

Remark 2.9. Theorem 2.8 improves Corollary 1 of [32], where (12) has been proved under a stronger condition than (10). The proof in [32] is done for real-valued variables but work in the above Banach setting as well.

Now, we give a result under condition (13), which has been introduced by Hannan [19].

Theorem 2.10. *Let $\mathcal{X}$ be a $(2, D)$-smooth separable Banach space, for some $D \geq 1$. Let $X \in L^2(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{X})$ be a regular random variable. Assume moreover that*

$$\|X\|_{H_2} := \sum_{n\in\mathbb{Z}} \|P_n X\|_{2,\mathcal{X}} < \infty. \tag{13}$$

Then, for every $1 \leq p < 2$, there exists (a universal) $C_p > 0$, such that

$$\big\|\mathcal{M}_2(X)\big\|_{p,\infty} \leq C_p D \|X\|_{H_2}. \tag{14}$$

The series $d = \sum_{n\in\mathbb{Z}} P_1(X \circ \theta^n)$ converges in $L^2(\Omega, \mathcal{F}_1, \mathcal{X})$ and $\mathbb{E}_0(d) = 0$. Moreover, writing $M_n := \sum_{k=0}^{n-1} d \circ \theta^k$, we have

$$|S_n - M_n|_{\mathcal{X}} = \mathrm{o}\Big(\sqrt{nL\big(L(n)\big)}\Big) \qquad \mathbb{P}\text{-a.s.} \tag{15}$$

Remark 2.11. Theorem 2.10 improves Theorem 2 of Wu [32], Theorem 2.1 of Liu and Lin [24] (for $p = 2$) and Corollary 5.3 of Cuny [5]. In [5,24,32] the authors prove (15) under stronger conditions than (13) and the proof do not apply to infinite dimensional Banach spaces.

In particular, we deduce the following corollary from Theorem 2.10, Theorem 2.3 and Proposition 2.5.

Corollary 2.12. *Under the assumptions of Theorem 2.10, $(X \circ \theta^n)_{n\geq 0}$ satisfies the CLIL and the ASIP of covariance $\mathcal{K}_d$, where, for every $x^*, y^* \in \mathcal{X}^*$, $\mathcal{K}_d(x^*, y^*)$, $\mathcal{K}_d = \sum_{n\in\mathbb{Z}} \mathbb{E}(x^*(X_n) y^*(X))$. Moreover, since, $\|d\|_{2,\mathcal{X}} \leq \|X\|_{H_2}$,*

$$\limsup_n \frac{|S_n(X)|_{\mathcal{X}}}{\sqrt{2nL(L(n))}} \leq \|X\|_{H_2} \qquad \mathbb{P}\text{-a.s.}$$

In order to check (13) or (10), it may be easier to use the condition (16) below.

Lemma 2.13. *Let $1 < p \leq 2$. Let $\mathcal{H}$ be a separable real Hilbert space. Assume that*

$$\sum_{n\geq 1} \frac{\|\mathbb{E}_{-n}(X)\|_{p,\mathcal{H}}}{\sqrt{n}} < \infty \quad \text{and} \quad \sum_{n\geq 1} \frac{\|X - \mathbb{E}_n(X)\|_{p,\mathcal{H}}}{\sqrt{n}} < \infty. \tag{16}$$

Then X is regular and $\sum_{n\in\mathbb{Z}} \|P_n X\|_{p,\mathcal{H}} < \infty$.

2.3. Functionals of Markov chains

The situation considered in the previous paragraph includes the case of stationary (ergodic) Markov chains. Let Q be a transition probability on a measurable space $(\mathbb{S}, \mathcal{S})$ admitting an invariant probability m. Let $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\in\mathbb{Z}}, \mathbb{P}, (W_n)_{n\in\mathbb{Z}})$ be the canonical Markov chain associated with Q, that is, $\Omega = \mathcal{S}^{\mathbb{Z}}$, $\mathcal{F} = \mathcal{S}^{\otimes\mathbb{Z}}$, $(W_n)_{n\in\mathbb{Z}}$ the coordinates, $\mathcal{F}_n = \sigma\{\dots, W_{n-1}, W_n\}$, $\mathbb{P}\circ W_0^{-1} = m$ and $\mathbb{P}(W_{n+1} \in A|\mathcal{F}_n) = Q(W_n, A)$. Finally, denote by θ the shift on Ω.

Recall that Q induces an operator on $L^2(\mathbb{S}, m)$ that we still denote by Q. If $\mathcal{H}$ is a separable real Hilbert space, we denote by $\mathbf{Q}$ the analogous operator on $L^2(\mathbb{S}, m, \mathcal{H})$. In particular, for every $f \in L^2(\Omega, \mathcal{H})$ and every $h \in \mathcal{H}$, $\langle \mathbf{Q}f, h\rangle_{\mathcal{H}} = Q(\langle f, h\rangle_{\mathcal{H}})$.

Theorem 2.10 applies to that setting with $X = f(W_0)$, where $f \in L^2(\mathbb{S}, \mathcal{H})$. Using Lemma 2.13, it suffices to check (16). In that situation, the process is adapted, that is, X_0 is $\mathcal{F}_0$-measurable. Hence, the second part of condition (16) is automatically satisfied while the first part reads as follows

$$\sum_{n\ge 1} \frac{\|\mathbf{Q}^n f\|_{2,\mathcal{H}}}{\sqrt{n}} < \infty. \tag{17}$$

2.4. Results for non-invertible dynamical systems

Here, we assume that θ is non-invertible. Let us write $\mathcal{F}^n = \theta^{-n}(\mathcal{F})$, for every $n \ge 0$. Denote $\mathcal{F}^\infty = \bigcap_{n\ge 0} \mathcal{F}^n$.

In this case, there exists a Markov operator K, known as the Perron–Frobenius operator, defined by

$$\int_\Omega X(Y\circ\theta)\,d\mathbb{P} = \int_\Omega (KX)Y\,d\mathbb{P} \qquad \forall X, Y \in L^2(\Omega, \mathcal{F}, \mathbb{P}). \tag{18}$$

Then, we have for every $X \in L^1(\Omega, \mathcal{F}^0, \mathbb{P})$,

$$\mathbb{E}^n(X) = \big(K^n X\big)\circ\theta^n. \tag{19}$$

If $\mathcal{H}$ is a separable real Hilbert space, we extend K to $L^2(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{H})$, in a way similar to (18). We denote by $\mathbf{K}$ the obtained operator.

Theorem 2.14. *Let $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ be a non-invertible dynamical system. Let $X \in L^2(\Omega, \mathcal{H})$ be such that*

$$\sum_{n\ge 0} \frac{\|\mathbf{K}^n X\|_{2,\mathcal{H}}}{\sqrt{n}} < \infty. \tag{20}$$

Then, for every $1 < p < 2$, there exists $C_p > 0$ such that

$$\big\|\mathcal{M}_2(X)\big\|_{p,\mathcal{H}} \le C_p \sum_{n\ge 0} \frac{\|\mathbf{K}^n X\|_{2,\mathcal{H}}}{\sqrt{n}}.$$

Moreover, there exists $d \in L^2(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{H})$ with $\mathbb{E}^1(d) = 0$, such that, writing $M_n := \sum_{k=0}^{n-1} d\circ\theta^k$, we have

$$|S_n - M_n|_{\mathcal{H}} = \mathrm{o}\Big(\sqrt{nL\big(L(n)\big)}\Big) \qquad \mathbb{P}\text{-}a.s. \tag{21}$$

Remark 2.15. It follows from (21) that $(X \circ \theta^n)_{n\geq 0}$ satisfies the CLIL, but we do not know whether it satisfies the ASIP in general, except when $\mathcal{H}$ has dimension one (see Remark 2.7).

3. Applications, examples

Now, we give several applications of the previous results. We do not intend to give all possible examples where our conditions apply, but we try to provide examples illustrating the different situations we have considered.

For instance, our results on the Marcinkiewicz–Zygmund strong laws (and on the LIL) may be used (in the one-dimensional case) to obtain almost-sure invariance principles with rate as in [32] (see also [8] or [6]).

We start with a one-dimensional situation.

3.1. ϕ-mixing sequences

Let us recall the definition of the ϕ-mixing coefficients, introduced by Dedecker and Prieur [15]. Examples of ϕ-mixing sequences may be found there as well.

Definition 3.1. *For any integrable random variable X, let us write $X^{(0)} = X - \mathbb{E}(X)$. For any random variable Y with values in $\mathbb{R}$ and any σ-algebra $\mathcal{F}$, let*

$$\phi(\mathcal{F}, Y) = \sup_{x\in\mathbb{R}} \big\| \mathbb{E}\big((\mathbf{1}_{Y\leq x})^{(0)} | \mathcal{F}\big)^{(0)} \big\|_\infty.$$

For a sequence $\mathbf{Y} = (Y_i)_{i\in\mathbb{Z}}$, where $Y_i = Y \circ \theta^i$ and Y is an $\mathcal{F}_0$-measurable and real-valued random variable, let

$$\phi_{\mathbf{Y}}(n) = \sup_{i\geq n} \phi(\mathcal{F}_0, Y_i).$$

We need also the following technical definition.

Definition 3.2. *If μ is a probability measure on $\mathbb{R}$ and $p \in]1,\infty)$, $M \in (0,\infty)$, let $\mathrm{Mon}_p(M,\mu)$ denote the set of functions $f:\mathbb{R}\to\mathbb{R}$ which are monotone on some interval and null elsewhere and such that $\mu(|f|^p) \leq M^p$. Let $\mathrm{Mon}_p^c(M,\mu)$ be the closure in $\mathbb{L}^p(\mu)$ of the set of functions which can be written as $\sum_{\ell=1}^L a_\ell f_\ell$, where $\sum_{\ell=1}^L |a_\ell| \leq 1$ and $f_\ell \in \mathrm{Mon}_p(M,\mu)$.*

Theorem 3.1. *Let $X = f(Y) - \mathbb{E}(f(Y))$, where Y is an $\mathcal{F}_0$-measurable random variable. Let P_Y be the distribution of Y and $p \in]1,\infty]$. Assume that f belongs to $\mathrm{Mon}_p^c(M, P_Y)$ for some $M > 0$, if $2 \leq p < \infty$ and that f has bounded variation if $p = \infty$. Assume moreover that*

$$\sum_{k\geq 1} \frac{\phi_{\mathbf{Y}}^{(p-1)/p}(k)}{k^{1/2}} < \infty. \tag{22}$$

Then, if $1 < p < 2$, $(X \circ \theta^n)_{n\in\mathbb{Z}}$ *satisfies the conclusion of Theorem* 2.8 *and if* $p \geq 2$, $(X \circ \theta^n)_{n\in\mathbb{Z}}$ *satisfies the conclusion of Theorem* 2.10.

Remark 3.2. When $p = 2$, Dedecker, Gouëzel and Merlevède [9] proved that the condition $\sum_{k\geq 1} k^{1/\sqrt{3}-1/2}\phi_{\mathbf{Y}}^{1/2}(k) < \infty$ implies that $\sum_{n\geq 1} \mathbb{P}(\max_{1\leq k\leq 2^n} |S_k| > C2^{n/2}(L(n))^{1/2}) < \infty$ (which implies the bounded LIL).

Proof of Theorem 3.1. Assume first that $1 < p < \infty$. Since $f \in \mathrm{Mon}_p^c(M, P_{Y_0})$, there exists a sequence of functions

$$f_L = \sum_{k=1}^{L} a_{k,L} f_{k,L},$$

such that for every $L \geq 1$, $\sum_{k=1}^{L} |a_{k,L}| \leq 1$, for every $1 \leq k \leq L$, $f_{k,L}$ is monotonic on some interval and null elsewhere, and $\|f_{k,L}(Y_0)\|_p \leq M$ and such that $(f_L)_{L\geq 1}$ converges in $L^p(P_{Y_0})$ to f. Hence,

$$\begin{aligned}
&\big\|\mathbb{E}_0\big(f(Y_n)\big) - \mathbb{E}\big(f(Y_n)\big)\big\|_p \\
&\quad = \lim_{L\to\infty} \big\|\mathbb{E}_0\big(f_L(Y_n)\big) - \mathbb{E}\big(f_L(Y_n)\big)\big\|_p \\
&\quad \leq \liminf_{L\to\infty} \sum_{k=1}^{L} |a_{k,L}| \big\|\mathbb{E}_0\big(f_{k,L}(Y_n)\big) - \mathbb{E}\big(f_{k,L}(Y_n)\big)\big\|_p \leq C_p M \phi_{\mathbf{Y}}^{(p-1)/p}(n),
\end{aligned}$$

where we used Lemma 5.2 of [10] for the last estimate.

To conclude in that case, we notice first that we are in the adapted case, and that Theorem 2.8 (when $1 < p < 2$) and Theorem 2.10 (when $p \geq 2$) apply by Lemma 2.13.

Assume that $p = \infty$ and that f has bounded variation. Hence f is the difference of two monotonic functions, to which we apply Lemma 5.2 of [10] with $p = \infty$. Then, we conclude as above. □

3.2. $\mathcal{X}$-valued linear processes

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and θ be an ergodic invertible and bi-measurable transformation on Ω. Let $\mathcal{X}$ be a separable r-smooth Banach space, for some $1 < r \leq 2$. Let $\xi \in L^p(\Omega, \mathcal{F}_0, \mathbb{P}, \mathcal{X})$ for some $p > 1$. Assume that $\mathbb{E}(\xi|\mathcal{F}_{-1}) = 0$ and define $\xi_n = \xi \circ \theta^n$, $n \in \mathbb{Z}$.

Let $(A^{(k)})_{k\in\mathbb{Z}}$ be a (not necessarily stationary) sequence of random variables with values in $L^\infty(\Omega, \mathcal{F}_{k-1}, \mathbf{B}(\mathcal{X}))$, where $\mathbf{B}(\mathcal{X})$ stands for the Banach space of bounded (linear) operators on $\mathcal{X}$. For every $k, n \in \mathbb{Z}$, define $A_n^{(k)} = A^{(k)} \circ \theta^n$. Assume that

$$\sum_{k\in\mathbb{Z}} \big\|A^{(k)}\big\|_{\infty,\mathbf{B}(\mathcal{X})} < \infty. \tag{23}$$

Then, the process

$$X_n := \sum_{k\in\mathbb{Z}} A_n^{(k)}\xi_{n+k}, \qquad n \in \mathbb{Z}$$

is well defined in $L^p(\Omega, \mathcal{X})$ and is stationary.

Corollary 3.3. *Assume that* $1 < p < r \le 2$ *or* $p = r = 2$. *Let* (X_n) *be a linear process as above. Then*,

$$\sum_{n\in\mathbb{Z}} \|P_n X_0\|_{p,\mathcal{X}} < \infty. \tag{24}$$

Hence, *Theorem* 2.8 *applies when* $1 < p < 2$ *and Theorem* 2.10 *applies when* $p = 2$.

3.3. Functions of real-valued linear processes

Let $(\xi_n)_{n\in\mathbb{Z}}$ be a sequence of *independent* identically distributed *real* random variables in $L^2(\Omega, \mathcal{F}, \mathbb{P})$. Let $(a_n)_{n\in\mathbb{Z}}$ be in ℓ^1. We consider a linear process defined by

$$Y_n := \sum_{k\in\mathbb{Z}} a_k \xi_{n-k} \qquad \forall n \in \mathbb{Z}.$$

For every $n \in \mathbb{Z}$, write $\mathcal{F}_n = \sigma\{\dots, \xi_{n-1}, \xi_n\}$.

We denote by Λ the class of non-decreasing continuous and bounded functions on $[0, +\infty[$, such that $\varphi(0) = 0$, and satisfying one of the following

$$\varphi^2 \text{ is concave;}$$
$$\varphi(x) = C\min\bigl(1, x^\alpha\bigr) \qquad \forall x \ge 0, \text{ for some } 0 < \alpha \le 1,\ C > 0.$$

Let $r \ge 1$. Let f be a real valued function such that

$$\bigl|f(x) - f(y)\bigr| \le \varphi\bigl(|x-y|\bigr)\bigl(1 + |x|^r + |y|^r\bigr) \qquad \forall x, y \in \mathbb{R}. \tag{25}$$

Our functions are unbounded and their continuity is locally controlled by φ.

We want to study the process $(X_n)_{n\in\mathbb{Z}}$ given by

$$X_n := f(Y_n) - \mathbb{E}\bigl(f(Y_n)\bigr) \qquad \forall n \in \mathbb{Z}.$$

Corollary 3.4. *Let* $\varphi \in \Lambda$ *and* $r \ge 1$. *Let* $\xi_0 \in L^{2r}(\Omega, \mathcal{F}, \mathbb{P})$ *and* f *satisfy* (25). *Let* $(a_n)_{n\in\mathbb{Z}} \in \ell^1$. *Consider the process* $(X_n)_{n\ge 0}$ *above. If*

$$\sum_{n\ge 1} \varphi\bigl(|a_n|\bigr) < \infty \quad \textit{or} \quad \sum_{n\ge 1} \frac{\varphi(\sum_{k\ge n}|a_k|)}{\sqrt{n}} < \infty,$$

then $(X_n)_{n\ge 0}$ *satisfies the conclusion of Theorem* 2.10.

We give the proof in the Appendix.

Remark 3.5. Notice that condition (3.1) of [24] implies (25) with $\varphi(x) = \min(1, x)$. Hence, Corollary 3.4 improves Corollary 3.1 of [24] when $p = 2$.

3.4. A non-adapted example

We now consider an example of a non-adapted process for which new ASIP with rates have been obtained very recently, see Dedecker, Merlevède and Pène [14] and the references therein.

Let $d \geq 2$ and θ be an ergodic automorphism of the d-dimensional torus $\Omega = \Omega_d = \mathbb{R}^d/\mathbb{Z}^d$. Denote by $\mathcal{F}$ the Borel σ-algebra of Ω and take $\mathbb{P}$ to be the Lebesgue measure on Ω.

For every $\mathbf{k} = (k_1, \dots, k_d) \in \mathbb{Z}^d$, write $|\mathbf{k}| := \max_{1 \leq i \leq d} |k_i|$. If $\mathcal{H}$ is a Hilbert space and if $f \in L^2(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{H})$, we denote by $(c_{\mathbf{k}})_{\mathbf{k}\in\mathbb{Z}^d} = (c_{\mathbf{k},\mathcal{H}})_{\mathbf{k}\in\mathbb{Z}^d}$ its Fourier coefficients, that is, $c_{\mathbf{k},\mathcal{H}} = \int_{[0,1]^d} f(x)\mathrm{e}^{-2\mathrm{i}\pi\langle x,\mathbf{k}\rangle_d}\mathbb{P}(\mathrm{d}x)$, for every $\mathbf{k} \in \mathbb{Z}^d$, where $\langle \cdot, \cdot \rangle_d$ stands for the inner product on $\mathbb{R}^d$.

Corollary 3.6. *Let $\mathcal{H}$ be a Hilbert space and $f \in L^2(\Omega, \mathcal{H})$. Assume that there exists $\beta > 2$ and $C > 0$ such that*

$$\sum_{|\mathbf{k}|\geq m} |c_{\mathbf{k}}|^2_{\mathcal{H}} \leq \frac{C}{L(m)(L(L(m)))^\beta} \qquad \forall m \geq 1.$$

Then, $(f \circ \theta^n)_{n\geq 0}$ satisfies the ASIP with covariance operator given by $\mathcal{K}(x, y) := \sum_{m\in\mathbb{Z}} \mathbb{E}(\langle x, f\rangle_{\mathcal{H}}\langle y, f \circ \theta^n\rangle_{\mathcal{H}})$, for every $x, y \in \mathcal{H}$.

Remark 3.7. Dedecker, Merlevède and Pène [14], Theorem 2.1, obtained the ASIP when $\mathcal{H} = \mathbf{R}^m$ and their condition requires $\beta > 4$. When $m = 1$, rates in the ASIP are also provided in [14].

Proof of Corollary 3.6. It follows from the proof of Propositions 4.2 and 4.3 of [14] (notice that the proofs work in the Hilbert space setting) that there exists a filtration $(\mathcal{F}_n)_{n\in\mathbb{Z}}$ (defined at the beginning of paragraph 3 of [14]) such that $\mathcal{F}_n = \theta^{-n}(\mathcal{F}_0)$ and

$$\|\mathbb{E}_{-n}(f)\|_{2,\mathcal{H}} = \mathrm{O}\left(\frac{1}{\sqrt{nL(n)^\beta}}\right) \quad \text{and} \quad \|\mathbb{E}_n(f) - f\|_{2,\mathcal{H}} = \mathrm{O}\left(\frac{1}{\sqrt{nL(n)^\beta}}\right).$$

Then, the result follows from Lemma 2.13. □

3.5. Cramer–von Mises statistics

We use our previous notations, see Section 2.2.

Let $Y \in L^0(\Omega, \mathcal{F}_0, \mathbb{P})$. For every $n \in \mathbb{Z}$, let $Y_n := Y \circ \theta^n$ and $X_n := t \mapsto \mathbf{1}_{Y_n\leq t} - F(t)$, where $F(t) = \mathbb{P}(Y \leq t)$.

Let $1 < r \le 2$. For every σ-finite Borel measure μ on $\mathbb{R}$, we may see $(X_n)_{n\in\mathbb{Z}}$ as a process with values in the r-smooth Banach space $L^r(\mathbb{R},\mu)$, as soon as

$$\int_0^{\infty} \bigl(1 - F(t)\bigr)^r \mu(\mathrm{d}t) + \int_{-\infty}^{0} F(t)^r \mu(\mathrm{d}t) < \infty, \tag{26}$$

which is satisfied whenever μ is finite.

Define F_μ by $F_\mu(x) = -\mu([x,0[)$ if $x \le 0$ and $F_\mu(x) = \mu([0,x[)$ if $x \ge 0$. Let $1 < p \le 2$. Then, under (26), $X_0 \in L^p(\Omega, L^r(\mu))$ if and only if

$$\mathbb{E}\bigl(\bigl|F_\mu(Y_0)\bigr|^{p/r}\bigr) < \infty. \tag{27}$$

We want to understand the asymptotic behaviour of the process $F_n = S_n(X)/n$ (with values in $L^2(\mathbb{R},\mu)$), and more particularly of $D_n(\mu) := \|F_n\|_{2,\mu}$. When $\mu = P_Y = \mathbb{P}\circ Y^{-1}$, $D_n(\mu)^2$ is known as the Cramer–von Mises statistics.

It follows from Lemma 2.13, that if $(X_n)_{n\in\mathbb{Z}}$ satisfies

$$\sum_{n\ge 1} \frac{(\mathbb{E}(\|\mathbb{E}_{-n}(X_0)\|_{2,\mu}^p))^{1/p}}{n^{1/2}} < \infty, \tag{28}$$

for some $1 < p \le 2$, then $(X_n)_{n\in\mathbb{Z}}$ satisfies Theorem 2.8 if $1 < p < 2$ and Theorem 2.10 if $p = 2$. Hence, we have the following corollary.

Corollary 3.8. *Let $1 < p < r \le 2$ or $p = r = 2$. With the above notations, assume that (26), (27) and (28) be satisfied. Then,*

$$\lim_n n^{1-1/p} D_n(\mu) = 0 \qquad \mathbb{P}\text{-}a.s. \qquad \text{if } 1 < p < 2;$$

$$\limsup_n \frac{n^{1/2}}{(2L(L(n)))^{1/2}} D_n(\mu) = \Lambda_\mu \qquad \mathbb{P}\text{-}a.s. \qquad \text{if } p = 2,$$

where $\Lambda_\mu^2 := \sup_{\|f\|_{2,\mu,\mathbb{R}}\le 1} \int_{\mathbb{R}^2} f(s)f(t)C(s,t)\mu(\mathrm{d}s)\mu(\mathrm{d}t)$ and $C(s,t) := \sum_{n\in\mathbb{Z}}(\mathbb{P}(Y_0 \le s, Y_n \le t) - F(s)F(t))$.

Proof. Apply Lemma 2.13, Theorem 2.8 and Theorem 2.10. The expression of Λ_μ^2 follows, for instance, from Proposition 1 of Merlevède [26]. □

In the context of ϕ-mixing sequences, when μ is finite, Corollary 3.8 applies as soon as $\sum_{n\ge 1} \frac{\phi_{\mathbf{Y}}(n)^{1/2}}{n^{1/2}} < \infty$.

Other examples where (28) is satisfied may be found in [11].

4. Proof of the results for Banach-valued martingales

Proof of Theorem 2.3. Let us prove (8). We start with the case $d \in L^2(\Omega, \mathcal{F}_1, \mathbb{P})$ and $\mathbb{E}_0(d) = 0$.

When $d \in L^2(\Omega, \mathcal{F}^0, \mathbb{P})$ and $\mathbb{E}^1(d) = 0$, the proof is the same, with the obvious changes, noticing that for every $n \geq 1$, $(S_n(d) - S_{n-k}(d))_{0\leq k\leq n}$ is a $(\mathcal{F}^{n-k})_{0\leq k\leq n}$-martingale and that $\max_{1\leq k\leq n} |S_k(d)|_{\mathcal{X}} \leq 2\max_{1\leq k\leq n} |S_n(d) - S_{n-k}(d)|_{\mathcal{X}}$.

Clearly, by homogeneity, it suffices to prove the result when $\|d\|_{2,\mathcal{X}} = 1$. Let $\lambda > 0$ and $1 \leq p < 2$. Let us prove that there exists $C_p \geq 1$, independent of λ such that

$$\lambda^p \mathbb{P}\big(M^* > \lambda\big) \leq D^p C_p^p, \tag{29}$$

where

$$M^* = M^*(d) := \sup_{s\geq 0} \frac{\max_{1\leq k\leq 2^s} |S_k(d)|_{\mathcal{X}}}{2^{s/2}(L(s))^{1/2}}.$$

Since $\mathcal{M}_2(d) \leq CM^*$, this will imply the desired result. Notice that (29) holds trivially when $0 < \lambda < D$. Assume then that $\lambda \geq D$.

Let $S \geq 1$ be an integer, fixed for the moment. For simplicity, we write $S_n := S_n(d)$.

We have, using Doob's maximal inequality for the submartingale $(|S_n|_{\mathcal{X}})_{n\geq 1}$, and (2)

$$\begin{aligned}\mathbb{P}\bigg(\sup_{1\leq s\leq S} \frac{\max_{1\leq k\leq 2^s} |S_k|_X}{2^{s/2}(L(s))^{1/2}} > \lambda\bigg) &\leq \frac{1}{\lambda^2}\sum_{s=1}^{S} \frac{\mathbb{E}(\max_{1\leq k\leq 2^s} |S_k|^2_{\mathcal{X}})}{2^s L(s)} \\ &\leq \frac{2}{\lambda^2}\sum_{s=1}^{S} \frac{\mathbb{E}(|S_{2^s}|^2_{\mathcal{X}})}{2^s L(s)} \leq \frac{2D^2 S}{\lambda^2}.\end{aligned} \tag{30}$$

We make use of truncations. Let $\alpha > 0$ be fixed for the moment. Let us write $d_n := d \circ \theta^{n-1}$, $n \geq 1$. For every $s \geq 1$, $k \geq 1$ define

$$e_k^{(s)} := d_k \mathbf{1}_{\{|d_k|_{\mathcal{X}} \leq \alpha\lambda 2^{s/2}/(L(s))^{1/2}\}}; \qquad d_k^{(s)} := e_k^{(s)} - \mathbb{E}\big(e_k^{(s)}|\mathcal{F}_{k-1}\big); \qquad \tilde{d}_k^{(s)} := d_k - d_k^{(s)},$$

$$S_k^{(s)} := \sum_{i=1}^{k} d_i^{(s)}; \qquad \tilde{S}_k^{(s)} := S_k - S_k^{(s)},$$

$$T_s := 4\sum_{i=1}^{2^s} \mathbb{E}\big(|d_i|^2_{\mathcal{X}}|\mathcal{F}_{i-1}\big); \qquad T_s^{(s)} := \sum_{i=1}^{2^s} \mathbb{E}\big(|d_i^{(s)}|^2_{\mathcal{X}}|\mathcal{F}_{i-1}\big).$$

Notice that, for every $s \geq 1$,

$$T_s^{(s)} \leq T_s. \tag{31}$$

Let $\beta > 0$ be fixed for the moment. Define the events

$$A_s := \left\{\frac{\max_{1\leq k\leq 2^s} |S_k|_{\mathcal{X}}}{2^{s/2}(L(s))^{1/2}} > \lambda\right\}; \qquad B_s := \left\{\frac{\max_{1\leq k\leq 2^s} |S_k^{(s)}|}{2^{s/2}(L(s))^{1/2}} > \lambda/2\right\},$$

$$C_s := \left\{\frac{\max_{1\leq k\leq 2^s} |\tilde{S}_k^{(s)}|_{\mathcal{X}}}{2^{s/2}(L(s))^{1/2}} > \lambda/2\right\}; \qquad D_s := \left\{\frac{T_s}{2^s} > \beta\lambda^2\right\}; \qquad E_s := B_s \cap \left\{\frac{T_s^{(s)}}{2^s} \leq \beta\lambda^2\right\}.$$

Using (31), we see that $B_s \cap D_s^c \subset E_s$. In particular, we have

$$A_s \subset B_s \cup C_s; \qquad B_s \subset D_s \cup E_s.$$

Hence,

$$\left\{\sup_{s\geq S} \frac{\max_{1\leq k\leq 2^s} |S_k|_{\mathcal{X}}}{2^{s/2}(L(s))^{1/2}} > \lambda\right\} = \bigcup_{s\geq S} A_s \subset \Big(\bigcup_{s\geq S} C_s\Big) \cup \Big(\bigcup_{s\geq S} D_s\Big) \cup \Big(\bigcup_{s\geq S} E_s\Big).$$

Now, $\bigcup_{s\geq S} D_s = \{\sup_{s\geq S} \frac{T_s}{2^s} > \beta\lambda^2\}$, hence by Hopf maximal inequality (5), using that $\mathbb{E}(|d_1|_{\mathcal{X}}^2) = 1$,

$$\mathbb{P}\Big(\bigcup_{s\geq S} D_s\Big) \leq \mathbb{P}\Big(\bigcup_{s\geq 1} D_s\Big) \leq \frac{4}{\beta\lambda^2}. \tag{32}$$

We also easily see that, interverting $\sum$ and $\mathbb{E}$ in (33),

$$\begin{aligned}\mathbb{P}\Big(\bigcup_{s\geq S} C_s\Big) &\leq \frac{2}{\lambda}\sum_{s\geq 0} \frac{\mathbb{E}(\max_{1\leq k\leq 2^s} |\tilde{S}_k^{(s)}|_{\mathcal{X}})}{2^{s/2}(L(s))^{1/2}} \\ &\leq \frac{4}{\lambda}\sum_{s\geq 1} \frac{2^{s/2}}{(L(s))^{1/2}} \mathbb{E}\big(|d_1|_{\mathcal{X}} \mathbf{1}_{\{|d_1|_{\mathcal{X}} \geq \alpha\lambda 2^{s/2}/(L(s))^{1/2}\}}\big) \leq \frac{4C}{\alpha\lambda^2},\end{aligned} \tag{33}$$

where we also used that there exists $C > 0$ such that for every $u > 0$,

$$\sum_{s\leq u} 2^{s/2}/\big(L(s)\big)^{1/2} \leq \sum_{s\leq\sqrt{u}} 2^{s/2} + \frac{1}{(L(\sqrt{u}))^{1/2}} \sum_{\sqrt{u}<s\leq u} 2^{s/2} \leq C2^{u/2}/L(u)^{1/2}.$$

It remains to deal with $\bigcup_{s\geq S} E_s$. We need the following lemma from Dedecker, Gouëzel and Merlevède [9], Proposition A.1 (see also Merlevède [26], Lemma 1), whose proof follows from Pinelis [29], Theorem 3.4. The proof in [9] is done in the scalar case (and in [26] in the Hilbert case) but it easily extends to 2-smooth Banach spaces, since Theorem 3.4 in [29] is proved in that setting. A related inequality in the scalar case is stated in Freedman [18], Theorem 1.6.

Lemma 4.1. *Let $\mathcal{X}$ be a $(2, D)$-smooth Banach space. Let $c > 0$. Let $(\mathcal{F}_j)_{j\geq 0}$ be a non-decreasing filtration and $(d_j)_{j\geq 1}$ a sequence of random variables adapted to $(\mathcal{F}_j)_{j\geq 0}$, such that for every $j \geq 1$, $|d_j|_{\mathcal{X}} \leq c$ a.s. and $\mathbb{E}(d_j|\mathcal{F}_{j-1}) = 0$ a.s. Then, for all $x, y > 0$ and all integer $n \geq 1$, we have*

$$\mathbb{P}\left(\max_{1\leq k\leq n}\left|\sum_{i=1}^k d_i\right|_{\mathcal{X}} > x; \sum_{i=1}^n \mathbb{E}\big(|d_i|_{\mathcal{X}}^2|\mathcal{F}_{i-1}\big) \leq y/D^2\right) \leq 2\exp\left(-\frac{y}{c^2}h\left(\frac{xc}{y}\right)\right), \tag{34}$$

where $h(u) = (1+u)\log(1+u) - u$.

Let $s \geq S$. Let us apply the lemma to the sequence of martingale differences $(d_i^{(s)})$ (in this case, we may take $c = 2\alpha\lambda 2^{s/2}/(L(s))^{1/2}$), with $x = \lambda 2^{s/2-1}(L(s))^{1/2}$, $y = \beta D^2\lambda^2 2^s$ and $n = 2^s$. We obtain, taking $\alpha = D^2\beta$,

$$\mathbb{P}(E_s) \leq 2\exp\left(-\frac{D^2\beta L(s)}{4\alpha^2}h\left(\frac{\alpha}{D^2\beta}\right)\right) = 2\exp\left(-\frac{L(s)h(1)}{4D^2\beta}\right) = \frac{2}{s^{h(1)/4D^2\beta}}.$$

Hence, if $h(1)/(4D^2\beta) > 1$, we see that

$$\sum_{s\geq S}\mathbb{P}(E_s) \leq \frac{2}{(h(1)/4D^2\beta - 1)S^{h(1)/4D^2\beta-1}}.$$

Take $\beta = \frac{(2-p)h(1)}{8D^2}$ and $S = [\lambda^{2-p}]$. Then, $h(1)/4D^2\beta - 1 = 2/(2-p) - 1 = p/(2-p)$ and

$$\sum_{s\geq S}\mathbb{P}(E_s) \leq \frac{C}{(2-p)\lambda^p}. \tag{35}$$

Recall that we assume that $\lambda \geq D$, in particular $\frac{1}{\lambda^2} \leq \frac{D^{p-2}}{\lambda^p}$. Combining (30), (32), (33) and (35), we infer that, there exists $C > 0$, such that

$$\lambda^p\mathbb{P}\big(M^* > \lambda\big) \leq \frac{CD^p}{2-p},$$

which ends the proof of (8).

Let us prove that $(d \circ \theta^n)_{n\in\mathbb{N}}$ satisfies the CLIL. We shall use the Banach principle, see Proposition C.1. By definition of the Bochner spaces, there exists $(d^{(m)})_{m\geq 1}$, converging in $L^2(\Omega, \mathcal{X})$ to d, such that for every $m \geq 1$, there exist $k_m \geq 1$, $\alpha_1, \ldots, \alpha_{k_m} \in \mathcal{X}$ and $A_1, \ldots, A_{k_m} \in \mathcal{F}_1$ such that

$$d^{(m)} = \sum_{i=1}^{k_m}\alpha_i\mathbf{1}_{A_i}.$$

Write $\tilde{d}^{(m)} := d^{(m)} - \mathbb{E}_0(d^{(m)})$. Then, $(\tilde{d}^{(m)})_{m\geq 1}$ converges in $L^2(\Omega, \mathcal{X})$ to d. Hence, by the Banach principle, it suffices to prove that $(\tilde{d}^{(m)} \circ \theta^n)_{n\in\mathbb{N}}$ satisfies the CLIL for every $m \geq 1$. But, by construction, $(\tilde{d}^{(m)} \circ \theta^n)_{n\in\mathbb{N}}$ is a stationary sequence of martingale differences taking values in a *finite* dimensional Banach space (i.e., $\mathrm{Vect}\{\alpha_i\colon 1 \leq i \leq k_m\}$), in which case the compact LIL and the bounded LIL are equivalent. But the bounded LIL in that case follows from (8), hence the result.

It remains to prove (9). By the bounded LIL the variable $\limsup_n \frac{|S_n(d)|_{\mathcal{X}}}{\sqrt{2nL(L(n))}}$ is well-defined $\mathbb{P}$-a.s. and must be θ-invariant. By ergodicity, there exists $S \geq 0$, such that $\limsup_n \frac{|S_n(d)|_{\mathcal{X}}}{\sqrt{2nL(L(n))}} = S$ $\mathbb{P}$-a.s. Let $M := \sup_{|x^*|_{\mathcal{X}^*}\leq 1}\|x^*(d)\|_2$. Let us prove that $S = M$. Let $\varepsilon > 0$. There exists $x_\varepsilon^* \in \mathcal{X}^*$, with $|x_\varepsilon^*|_{\mathcal{X}^*} \leq 1$, such that $\|x_\varepsilon^*(d)\|_2 \geq M - \varepsilon$. Since, $|S_n(d)|_{\mathcal{X}} \geq |x_\varepsilon^*(S_n(d))|$, it follows from the LIL for real-valued martingales (with stationary ergodic increments) that

$$S \geq M - \varepsilon.$$

Letting $\varepsilon \to 0$, we see that $S \geq M$. Let us prove the converse inequality.

Let $x^* \in \mathcal{X}^*$. By the LIL for real-valued, stationary and ergodic martingale differences, $\limsup_n S_n(x^*(d))/\sqrt{2nL(L(n))} = \|x^*(d)\|_2$ $\mathbb{P}$-a.s. Hence, by the compact LIL and Proposition D.1, there exists a compact set $K \in \mathcal{X}$, such that for $\mathbb{P}$-a.e. $\omega \in \Omega$, the cluster set of $\{S_n(d)(\omega)/\sqrt{2nL(L(n))}, n \geq 1\}$ is K. Let $x \in K$ be such that $|x|_{\mathcal{X}} = S$, and let $x^* \in \mathcal{X}^*$ be such that $|x^*|_{\mathcal{X}^*} = 1$ and $x^*(x) = |x|_{\mathcal{X}}$. For $\mathbb{P}$-a.e. $\omega \in \Omega$, there exists $(n_k = n_k(\omega))_{k\geq 1}$ such that $S_{n_k}(d)(\omega)\sqrt{2n_k L(L(n_k))} \xrightarrow[k\to\infty]{|\cdot|_{\mathcal{X}}} x$. In particular

$$x^*\big(S_{n_k}(d)(\omega)\sqrt{2n_k L\big(L(n_k)\big)}\big) \xrightarrow[k\to\infty]{|\cdot|_{\mathcal{X}}} x^*(x) = S \leq \limsup_n S_n\big(x^*(d)\big)(\omega)\sqrt{2nL\big(L(n)\big)}.$$

But, by the real LIL, for $\mathbb{P}$-a.e. $\omega \in \Omega$,

$$\limsup_n S_n\big(x^*(d)\big)(\omega)\sqrt{2nL\big(L(n)\big)} \leq \big\|x^*(d)\big\|_2 \leq M,$$

which ends the proof. □

5. Proof of the results for stationary processes

5.1. Proof of Theorem 2.10

Recall that we assume here θ to be invertible. Let $\mathcal{X}$ be a 2-smooth Banach space.

Define

$$H_2 := \Big\{Z \in L^2(\Omega, \mathcal{X}) \colon \mathbb{E}_{-\infty}(Z) = 0, \mathbb{E}_{\infty}(Z) = Z, \sum_{n\in\mathbb{Z}} \|P_n Z\|_{2,\mathcal{X}} < \infty\Big\}. \tag{36}$$

It is not difficult to see that, setting $\|Z\|_{H_2} := \sum_{n\in\mathbb{Z}} \|P_n Z\|_{2,\mathcal{X}}$, $(H_2, \|\cdot\|_{H_2})$ is a Banach space.

By our regularity conditions, we have, $Z = \sum_{k\in\mathbb{Z}} P_k Z$ in $L^2(\Omega, \mathcal{X})$ and $\mathbb{P}$-a.s. Hence, writing $S_n = S_n(Z) = \sum_{i=0}^{n-1} Z \circ \theta^i$, we have

$$S_n = \sum_{k\in\mathbb{Z}} \sum_{i=0}^{n-1} (P_k Z) \circ \theta^i.$$

This splitting of S_n into a series of martingales with (stationary) increments has been used already in [32] and [7] in a similar context. This idea seems to appear first (explicitly) in a paper by McLeish [25]. We deduce that

$$\mathcal{M}_2(Z) \leq \sum_{k\in\mathbb{Z}} \mathcal{M}_2\big(P_k(Z)\big).$$

But, for every $k \in \mathbb{Z}$, $((P_k Z) \circ \theta^i)_{i \geq 1}$ is a stationary sequence of martingale differences. Hence, by Theorem 2.3, for every $1 \leq p < 2$, there exists C_p, such that

$$\left\| \mathcal{M}_2(Z) \right\|_{p,\infty} \leq C_p D\left(\sum_{k \in \mathbb{Z}} \|P_k Z\|_{2,\mathcal{X}} \right). \tag{37}$$

We define a continuous operator $\mathcal{D}$ on H_2 with values in $\{d \in L^2(\Omega, \mathcal{F}_1)\colon \mathbb{E}(d_1|\mathcal{F}_0) = 0\}$, by setting, for every $Z \in H_2$, $\mathcal{D}Z := \sum_{n \in \mathbb{Z}} P_1(Z \circ \theta^n)$. Write $d = \mathcal{D}Z$. Let $M_n := \sum_{i=0}^{n-1} d \circ \theta^i$. We want to prove that

$$|S_n - M_n|_{\mathcal{X}} = \mathrm{o}\big(\sqrt{nL\big(L(n)\big)}\big) \qquad \mathbb{P}\text{-a.s.} \tag{38}$$

Since $\mathcal{M}_2(Z - d) \leq \mathcal{M}_2(Z) + \mathcal{M}_2(d)$, using (37), Theorem 2.3 and the Banach principle (see the Appendix), we see that the set $\{Z \in H_2\colon \text{(38) holds}\}$ is closed in H_2, and, by linearity, that set is a vector space.

Let $Z \in H_2$. Clearly, $Z = \sum_{k \in \mathbb{Z}} P_k Z$ in H_2. Hence it suffices to prove (38) for $P_k Z$, for every $k \in \mathbb{Z}$. Now, $\mathcal{D}(P_k Z) = (P_k Z) \circ \theta^{1-k}$. Let $k \leq 0$. We have

$$\begin{aligned} S_n(P_k Z) - M_n(P_k Z) &= \sum_{\ell=0}^{n-1} \big((P_k Z) \circ \theta^\ell - (P_k Z) \circ \theta^{\ell+1-k}\big) \\ &= \sum_{\ell=0}^{-k} (P_k Z) \circ \theta^\ell - \left(\sum_{\ell=0}^{-k} (P_k Z) \circ \theta^\ell \right) \circ \theta^n = \mathrm{o}(\sqrt{n}) \qquad \mathbb{P}\text{-a.s.,} \end{aligned}$$

where we used that for any $X \in L^2(\Omega, \mathcal{X})$, $\sum_{n \geq 1} \mathbb{P}(|X \circ \theta^n|_{\mathcal{X}} > \varepsilon\sqrt{n})$, for every $\varepsilon > 0$, which implies that $X \circ \theta^n = \mathrm{o}(\sqrt{n})$ $\mathbb{P}$-a.s., by the Borel–Cantelli lemma. The case $k \geq 1$ may be handled similarly.

5.2. Proof of Theorem 2.8

As in the proof of Theorem 2.10, we define a Banach space

$$H_p := \left\{ Z \in L^p(\Omega, \mathcal{X})\colon \mathbb{E}_{-\infty}(Z) = 0, \mathbb{E}_\infty(Z) = Z, \|Z\|_{H_p} := \sum_{n \in \mathbb{Z}} \|P_n Z\|_{p,\mathcal{X}} < \infty \right\}.$$

We see that

$$\|\mathcal{M}_p Z\|_{p,\infty} \leq C_{p,r} D^{1/p} \|Z\|_{H_p},$$

where $C_{r,p}$ is the constant appearing in Proposition 2.1, and that the operator $\mathcal{D}$ may be extended in a continuous operator from H_p to $\{d \in L^p(\Omega, \mathcal{F}_1, \mathcal{X})\colon \mathbb{E}_0(d) = 0\}$. Then, the proof follows the one of Theorem 2.10. We first see that $|S_n - M_n|_{\mathcal{X}} = \mathrm{o}(n^{1/p})$ $\mathbb{P}$-a.s. and then we use that the Marcinkiewicz–Zygmund strong law of large number is known for r-smooth valued stationary martingale differences, see, for example, [31].

5.3. Proof of Corollary 2.12

We only have to prove that $\mathcal{K}_d$ is given as in the corollary. By (13), we have $\sum_{n\in\mathbb{Z}} \|P_1 X_n\|_{2,\mathcal{X}} < \infty$. Hence, for every $f, g \in \mathcal{X}^*$, we have, with absolute convergence of all the series,

$$\begin{aligned}\mathcal{K}_d(f,g) &= \sum_{m,n\in\mathbb{Z}} \mathbb{E}\big(P_1\big(f(X_n)\big)P_1\big(g(X_m)\big)\big) = \sum_{m,n\in\mathbb{Z}} \mathbb{E}\big(f(X_0)P_{1-n}\big(g(X_{m-n})\big)\big)\\ &= \sum_{m,n\in\mathbb{Z}} \mathbb{E}\big(f(X_0)P_{-n}\big(g(X_m)\big)\big) = \sum_{m\in\mathbb{Z}} \mathbb{E}\big(f(X_0)g(X_m)\big).\end{aligned}$$

6. Proof of Lemma 2.13

Since the sequences $(\|\mathbb{E}_{-n}(X)\|_{p,\mathcal{H}})$ and $(\|X - \mathbb{E}_n(X)\|_{p,\mathcal{H}})$ are non-increasing, (16) is equivalent to

$$\sum_{n\geq 0} 2^{n/2}\big\|\mathbb{E}_{-2^n}(X)\big\|_{p,\mathcal{H}} < \infty \quad \text{and} \quad \sum_{n\geq 0} 2^{n/2}\big\|X - \mathbb{E}_{2^n}(X)\big\|_{p,\mathcal{H}} < \infty.$$

In particular, X is regular.

Assume $p = 2$. For every $n \geq 0$, using Cauchy–Schwarz and that $\mathbb{E}(\langle P_{-k}X, P_{-\ell}X\rangle_{\mathcal{H}}) = 0$ for every $k \neq \ell$, we have

$$\left(\sum_{k=2^n}^{2^{n+1}-1} \|P_{-k}X\|_{2,\mathcal{H}}\right)^2 \leq 2^n \sum_{k\geq 2^n} \mathbb{E}\big(|P_{-k}X|^2_{\mathcal{H}}\big) \leq 2^n \mathbb{E}\big(\big|\mathbb{E}_{-2^n}(X)\big|^2_{\mathcal{H}}\big),$$

and

$$\left(\sum_{k=2^n}^{2^{n+1}-1} \|P_k X\|_{2,\mathcal{H}}\right)^2 \leq 2^n \sum_{k\geq 2^n} \mathbb{E}\big(|P_k X|^2_{\mathcal{H}}\big) \leq 2^n \mathbb{E}\big(\big|X - \mathbb{E}_{2^n}(X)\big|^2_{\mathcal{H}}\big).$$

Assume $1 < p < 2$. By Hölder's inequality twice we have, with $1/p + 1/q = 1$,

$$\begin{aligned}\left(\sum_{k=2^n}^{2^{n+1}-1} \|P_{-k}X\|_{p,\mathcal{H}}\right)^p &\leq 2^{np/q}\mathbb{E}\left(\sum_{k=2^n}^{2^{n+1}-1} |P_{-k}X|^p_{\mathcal{H}}\right)\\ &\leq 2^{np/2}\mathbb{E}\left(\left(\sum_{k\geq 2^n} |P_{-k}X|^2_{\mathcal{H}}\right)^{p/2}\right) \leq C 2^{np/2}\big\|\mathbb{E}_{-2^n}(X)\big\|^p_{p,\mathcal{H}},\end{aligned}$$

and

$$\left(\sum_{k=2^n}^{2^{n+1}-1} \|P_k X\|_{p,\mathcal{H}}\right)^2 \le 2^{np/q}\mathbb{E}\left(\sum_{k=2^n}^{2^{n+1}-1} |P_k X|_{\mathcal{H}}^p\right)$$

$$\le 2^{np/2}\mathbb{E}\left(\left(\sum_{k\ge 2^n} |P_k X|_{\mathcal{H}}^2\right)^{2/p}\right) \le C2^{np/2}\big\|X - \mathbb{E}_{2^n}(X)\big\|_{p,\mathcal{H}}^p,$$

where we used Burkholder's inequality in Hilbert spaces, see [3]. Then, we conclude as above.

6.1. Proof of Theorem 2.14

For every $n \ge 0$ define $P^{(n)} := \mathbb{E}^n - \mathbb{E}^{n+1}$. It suffices to prove the theorem under the weaker condition

$$\mathbb{E}^\infty(X) = 0 \quad \text{and} \quad \sum_{n\ge 0} \big\| P^{(n)}(X)\big\|_{2,\mathcal{H}} < \infty.$$

The fact that (20) implies the above condition may be proved as Lemma 2.13, using (19).

Then, the proof may be done exactly as the proof of Theorem 2.10 except that we make use of reverse martingales.

Appendix A: Proof of Proposition 2.1

We start with the case $d \in L^p(\Omega, \mathcal{F}_1, \mathbb{P})$ and $\mathbb{E}_0(d) = 0$. Define

$$M^* = M^*(d) := \sup_{s\ge 0} \frac{\max_{1\le n\le 2^s} |S_n(d)|_{\mathcal{X}}}{2^{s/p}}.$$

Let $s \ge 0$. For every $2^s \le n \le 2^{s+1} - 1$, we have

$$\frac{|S_n(d)|_{\mathcal{X}}}{n^{1/p}} \le \frac{\max_{1\le n\le 2^s} |S_n(d)|_{\mathcal{X}}}{2^{s/p}} \le M^*.$$

Hence, it suffices to prove the result for M^* instead of $\mathcal{M}_p(d)$. Let $\lambda > 0$. We proceed by truncation. For every $s \ge 0$, $k \ge 1$ define

$$e_k^{(s)} := d_k \mathbf{1}_{\{|d_k|_{\mathcal{X}} \le \lambda 2^{s/p}\}}; \qquad d_k^{(s)} := e_k^{(s)} - \mathbb{E}\big(e_k^{(s)}|\mathcal{F}_{k-1}\big);$$
$$\tilde{e}_k^{(s)} := d_k - e_k^{(s)}; \qquad \tilde{d}_k^{(s)} := d_k - d_k^{(s)};$$
$$M_k^{(s)} := \sum_{i=1}^k d_i^{(s)}; \qquad \tilde{M}_k^{(s)} := M_k - M_k^{(s)}.$$

Let $\lambda > 0$. Then,

$$\begin{aligned}
&\mathbb{P}\big(M^* > \lambda\big) \\
&\le \sum_{s\ge 0}\mathbb{P}\bigg(\frac{\max_{1\le n\le 2^s}|\tilde{M}_n^{(s)}|_{\mathcal{X}}}{2^{s/p}} > \lambda/2\bigg) + \sum_{s\ge 0}\mathbb{P}\bigg(\frac{\max_{1\le n\le 2^s}|M_n^{(s)}|_{\mathcal{X}}}{2^{s/p}} > \lambda/2\bigg) \\
&\le \frac{4}{\lambda}\sum_{s\ge 0}2^{(1-1/p)s}\mathbb{E}\big(\big|\tilde{e}_1^{(s)}\big|_{\mathcal{X}}\big) + \frac{2^r}{\lambda^r}\sum_{s\ge 0}\frac{\mathbb{E}(\max_{1\le n\le 2^s}|M_n^{(s)}|_{\mathcal{X}}^r)}{2^{rs/p}}.
\end{aligned}$$

Now, by Fubini and stationarity,

$$\sum_{s\ge 0}2^{(1-1/p)s}\mathbb{E}\big(\big|\tilde{e}_1^{(s)}\big|_{\mathcal{X}}\big) \le \frac{C\mathbb{E}(|d_1|_{\mathcal{X}}^p)}{\lambda^p}.$$

To deal with the second term, we use Doob's maximal inequality in L^r, for the submartingale $(|M_n|_{\mathcal{X}})_{n\ge 1}$, and (2). We obtain

$$\begin{aligned}
\sum_{s\ge 0}\frac{\mathbb{E}(\max_{1\le n\le 2^s}|M_n^{(s)}|_{\mathcal{X}}^r)}{2^{rs/p}} &\le \sum_{s\ge 0}\frac{C_r}{2^{rs/p}\lambda^r}\mathbb{E}\big(\big|M_{2^s}^{(s)}\big|_{\mathcal{X}}^r\big) \\
&\le D^r C_r\sum_{s\ge 0}2^{(1-r/p)s}\mathbb{E}\big(\big|d_1^{(s)}\big|_{\mathcal{X}}^r\big) \le \frac{D^r C_{r,p}\mathbb{E}(|d_1|_{\mathcal{X}}^p)}{\lambda^{p-r}},
\end{aligned} \tag{39}$$

which proves the proposition, in that case. When $d \in L^2(\Omega, \mathcal{F}^0, \mathbb{P})$ and $\mathbb{E}^1(d) = 0$, the proof is the same, with the obvious changes, noticing that for every $n \ge 1$, $(S_n(d) - S_{n-k}(d))_{0\le k\le n}$ is a $(\mathcal{F}^{n-k})_{0\le k\le n}$-martingale and that $\max_{1\le k\le n}|S_k(d)|_{\mathcal{X}} \le 2\max_{1\le k\le n}|S_n(d) - S_{n-k}(d)|_{\mathcal{X}}$.

Appendix B: Proof of Corollary 3.4

Notice that, by (25), for every $x, h, h' \in \mathbb{R}$, we have

$$\big|f(x+h) - f\big(x+h'\big)\big| \le 2^r\varphi\big(\big|h-h'\big|\big)\big(1+|x|^r\big) + 2^{r-1}K\big(|h|^r + \big|h'\big|^r\big). \tag{40}$$

Recall that for every concave ψ with $\psi(0) = 0$, $x \to \psi(x)/x$ is non-increasing on $]0, +\infty[$ and ψ is sub-additive.

We want to apply Theorem 2.10 and Lemma 2.13. We shall evaluate $\|P_0(X_n)\|_2$, $\|\mathbb{E}_0(X_n)\|_2$ and $\|X_n - \mathbb{E}_n(X_n)\|$.

Enlarging our probability space if necessary, we assume that there exists (ξ_n') an independent copy of (ξ_n).

Then,

$$P_0X_n = \mathbb{E}_0\big(f(A_n + h_n) - f\big(A_n + h_n'\big)\big),$$

where $A_n := \sum_{k>-n} a_{-k}\xi'_{n+k} + \sum_{k>n} a_k\xi_{n-k}$, $h_n := a_n\xi_0$ and $h'_n := a_n\xi'_0$.

In particular, we have, by independence and using (40),

$$\mathbb{E}\big((P_0X_n)^2\big) \le C_r\big(\mathbb{E}\big(\varphi^2\big(|a_n|\big(|\xi_0| + \big|\xi'_0\big|\big)\big)\big)\mathbb{E}\big(|A_n|^{2r}\big) + |a_n|^{2r}\mathbb{E}\big(|\xi_0|^{2r}\big)\big).$$

We notice now that for every $\varphi \in \Lambda$, there exists $C > 0$ such that, for every $n \ge 1$

$$\mathbb{E}\big(\varphi^2\big(|a_n|\big(|\xi_0| + \big|\xi'_0\big|\big)\big)\big) \le C\varphi^2\big(|a_n|\big). \tag{41}$$

This follows from Jensen's inequality and the sub-additivity of φ^2 (using that $\xi_0 \in L^1(\Omega, \mathcal{F}, \mathbb{P})$) when φ^2 is sub-additive, and it is obvious when $\varphi(x) = \min(1, x^\alpha)$ (using that $\xi_0 \in L^{2\alpha}(\Omega, \mathcal{F}, \mathbb{P})$).

Clearly, $\mathbb{E}(|A_n|^{2r}) \le (\sum_{k\in\mathbb{Z}} |a_k|\|\xi_0\|_{2r})^{2r}$.

Since $x \to \varphi^2(x)/x$ is non-increasing, when φ^2 is concave, we see that whenever $\varphi \in \Lambda$, $|a_n|^{2r} \le C\varphi^2(|a_n|)$.

This finishes the proof of Corollary 3.4 under the assumption on $P_0(X_n)$.

We shall now evaluate $\|\mathbb{E}_0(X_n)\|_2$, the case of $\|X_n - \mathbb{E}_n(X_n)\|_2$ may be treated similarly. We have

$$\mathbb{E}_0(X_n) = \mathbb{E}_0\big(f(B_n + k_n) - f\big(B_n - k'_n\big)\big),$$

where $B_n := \sum_{k>-n} a_{-k}\xi_{n+k}$, $k_n = \sum_{k\ge n} a_k\xi_{n-k}$ and $k'_n = \sum_{k\ge n} a_k\xi'_{n-k}$. Hence, using (40),

$$\|E_0(X_n)\|_2^2 \le C_r\big(\mathbb{E}\big(\varphi^2\big(|k_n| + |k'_n|\big)\mathbb{E}\big(|A_n|^{2r}\big) + 2\|k_n\|_{2r}^{2r}\big)\big).$$

When φ^2 is concave, by Jensen's inequality,

$$\mathbb{E}\big(\varphi^2\big(|k_n| + \big|k'_n\big|\big)\big) \le \varphi^2\Bigg(2\mathbb{E}\big(|\xi_0|\big)\sum_{k\ge n}|a_k|\Bigg) \le \big(1 + 2\mathbb{E}\big(|\xi_0|\big)\big)\varphi^2\Bigg(\sum_{k\ge n}|a_k|\Bigg).$$

When $\varphi(x) = \min(1, x^\alpha)$, assuming that $1/2 \le \alpha \le 1$ (otherwise we are in the previous case), we have

$$\mathbb{E}\big(\varphi^2\big(|k_n| + \big|k'_n\big|\big)\big) \le \Bigg(\sum_{k\ge n}|a_k|\|\xi_0\|_{2\alpha}\Bigg)^{2\alpha} \le C\varphi^2\Bigg(\sum_{k\ge n}|a_k|\Bigg).$$

Clearly, $\mathbb{E}(|B_n|^{2r}) \le (\sum_{k\in\mathbb{Z}} |a_k|\|\xi_0\|_{2r})^{2r}$.

Finally, we have

$$\|k_n\|_{2r}^{2r} \le \|\xi_0\|_{2r}^{2r}\Bigg(\sum_{k\ge n}|a_k|\Bigg)^{2r}.$$

Since $x \to \varphi^2(x)/x$ is non-decreasing, when φ^2 is concave, we see that whenever $\varphi \in \Lambda$,

$$\|k_n\|_{2r}^{2r} \le C\varphi^2\Bigg(\sum_{k\ge n}|a_k|\Bigg).$$

Appendix C: The Banach principle

The following is an extension of the Banach principle as stated in Theorem 7.2, page 64 of [20].

Proposition C.1. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\mathcal{X}, \mathbf{B}$ be Banach spaces. Let $\mathcal{C}$ be a vector space of measurable functions from Ω to $\mathcal{X}$. Let $(T_n)_{n\geq 1}$ be a sequence of linear maps from $\mathbf{B}$ to $\mathcal{C}$. Assume that there exists a positive decreasing function L on $]0,+\infty[$, with $\lim_{\lambda\to\infty} L(\lambda)=0$, such that*

$$\mathbb{P}\Big(\sup_{n\geq 1}|T_n x|_{\mathcal{X}} > \lambda |x|_{\mathbf{B}}\Big) \leq L(\lambda) \qquad \forall \lambda > 0, x\in \mathbf{B}. \tag{42}$$

Then the set $\{x\in\mathbf{B}\colon (T_n x)_{n\geq 1}$ is $\mathbb{P}$-a.s. relatively compact in $\mathcal{X}\}$ and the set $\{x\in\mathbf{B}\colon |T_n x|_{\mathcal{X}}\to 0$ $\mathbb{P}$-a.s.$\}$ are closed in $\mathbf{B}$.

Proof. We prove that the first set is closed, the proof for the second one being similar, but easier. Let $x\in\mathbf{B}$ and $(x_m)_{m\geq 1}\subset \mathbf{B}$ be such that $|x_m - x|_{\mathbf{B}} \underset{m\to\infty}{\longrightarrow} 0$ and such that for every $m\geq 1$, $(T_n x_m)_{n\geq 1}$ is $\mathbb{P}$-a.s. relatively compact in $\mathcal{X}$. We want to prove that $(T_n x)_{n\geq 1}$ is $\mathbb{P}$-a.s. relatively compact.

By (42), for every integers $m, p\geq 1$ (assume that $x\neq x_m$ otherwise there is nothing to do)

$$\mathbb{P}\Big(\sup_{n\geq 1}\big|T_n(x-x_m)\big|_{\mathcal{X}} > 1/p\Big) \leq L\bigg(\frac{1}{p|x-x_m|_{\mathbf{B}}}\bigg) \qquad \forall\lambda>0, x\in\mathbf{B}.$$

Since $\lim_{\lambda\to\infty} L(\lambda)=0$, there exists a subsequence $(m_k)_{k\geq 1}$ and a set $\Omega_0\in\mathcal{F}$ with $\mathbb{P}(\Omega_0)=1$, such that for every $\omega\in\Omega_0$,

$$\sup_{n\geq 1}\big|T_n(x-x_{m_k})\big|_{\mathcal{X}}(\omega) \underset{k\to\infty}{\longrightarrow} 0.$$

There exists $\Omega_1\in\mathcal{F}$, with $\mathbb{P}(\Omega_1)=1$, such that, for every $\omega\in\Omega_1$ and every $k\geq 1$, $((T_n x_{m_k})(\omega))_{n\geq 1}$ is relatively compact in $\mathcal{X}$.

Let $\omega\in\Omega_0\cap\Omega_1$ be fixed. Let φ_0 be an increasing function from $\mathbb{N}$ to $\mathbb{N}$. We want to prove that $(T_{\varphi_0(n)}x(\omega))_{n\geq 1}$ admits a convergent subsequence.

For every $k\geq 1$, $((T_{\varphi_0(n)}x_{m_k})(\omega))_{n\geq 1}$ admits a Cauchy subsequence. We construct by induction some increasing functions $(\varphi_k)_{k\geq 1}$ such that, for every $k\geq 1$, setting $\psi_k := \varphi_0\circ\varphi_1\circ\cdots\circ\varphi_k$, we have for every $p\geq n\geq 1$,

$$\big|T_{\psi_k(n)}x_{m_k}(\omega) - T_{\psi_k(p)}x_{m_k}(\omega)\big|_{\mathcal{X}} \leq 1/n.$$

Then, $(T_{\psi_n(n)}x(\omega))$ is Cauchy. Indeed, for every $N\geq 1$, and every $p>n\geq N$, we have

$$\begin{aligned}
&\big|T_{\psi_n(n)}x(\omega) - T_{\psi_p(p)}x(\omega)\big|_{\mathcal{X}} \\
&\quad\leq \big|T_{\psi_n(n)}x_{m_n}(\omega) - T_{(\psi_n\circ\varphi_{n+1}\circ\cdots\circ\varphi_p)(p)}x_{m_n}(\omega)\big|_{\mathcal{X}} + 2\sup_{r\geq 1}\big|T_r(x_{m_n}-x)\big|_{\mathcal{X}} \underset{N\to\infty}{\longrightarrow} 0,
\end{aligned}$$

and the result follows. $\square$

Appendix D: Identification of the cluster set

Denote by $\mathcal{X}^*$ the topological dual of $\mathcal{X}$. Let $X \in L^2(\Omega, \mathcal{X})$ such that $\mathbb{E}(X) = 0$. Following Kuelbs [21] (we refer to [21] for more details on the construction below), we define a bounded linear operator $\mathcal{S}$ from $\mathcal{X}^*$ to $\mathcal{X}$ and a bounded *symmetric* bilinear operator $\mathcal{K}$ from $\mathcal{X}^* \times \mathcal{X}^*$ to $\mathbb{R}$, by

$$\mathcal{S}(x^*) = \mathbb{E}\big(x^*(X)X\big) \qquad \forall x^* \in \mathcal{X}^*,$$
$$\mathcal{K}(x^*, y^*) = \mathbb{E}\big(x^*(X)y^*(X)\big) = y^*\big(\mathcal{S}(x^*)\big) = x^*\big(\mathcal{S}(y^*)\big) \qquad \forall x^*, y^* \in \mathcal{X}^*.$$

Let $\mathcal{H}_X$ be the closure of the range of $\mathcal{S}$ with respect to the following inner product:

$$\langle \mathcal{S}x^*, \mathcal{S}y^* \rangle_{\mathcal{H}_X} = \mathcal{K}(x^*, y^*).$$

Notice that the definition of $\langle \cdot, \cdot \rangle_{\mathcal{H}_X}$ does not depend on the chosen representatives (i.e., if $x^* \in \operatorname{Ker} \mathcal{S}$, $\langle \mathcal{S}x^*, \mathcal{S}y^* \rangle_{\mathcal{H}_X} = 0$ for every $y^* \in \mathcal{X}^*$) and that this inner product is really *positive definite*.

Finally, denote by $K = K_X$, the unit ball of $(\mathcal{H}_X, \|\cdot\|_{\mathcal{H}_X})$, K is compact by (iv), Lemma 2.1 of [21]. We recall an important result of Kuelbs, see [21], Theorem 3.1, II, where we denote by $C(\{x_n\})$ the cluster set of a sequence $(x_n) \subset \mathcal{X}$.

Proposition D.1 (Kuelbs [21]). *Let $X \in L^2(\Omega, \mathcal{X})$. Assume that $(X_n)_{n\ge 0}$ satisfies the CLIL and that,*

$$\limsup_n \frac{S_n(x^*(X))}{\sqrt{2nL(L(n))}} = \|x^*(X)\|_2 \qquad \mathbb{P}\text{-}a.s. \qquad \forall x^* \in \mathcal{X}^*. \tag{43}$$

Then,

$$C\left(\left\{\frac{S_n(X)}{\sqrt{2nL(L(n))}}\right\}\right) = K \qquad \mathbb{P}\text{-}a.s., \tag{44}$$

and

$$\limsup_n \frac{|S_n(X)|_{\mathcal{X}}}{\sqrt{2nL(L(n))}} = \sup_{x^* \in \mathcal{X}^*, |x^*|_{\mathcal{X}^*} \le 1} \|x^*(X)\|_2 \le \|X\|_{2,\mathcal{X}} \qquad \mathbb{P}\text{-}a.s. \tag{45}$$

Acknowledgements

The author is very grateful to the referees. One of them provided a very detailed and accurate report that surely helped to improve the presentation of the paper. The author is also grateful to Florence Merlevède for reading a first version of the paper and for her comments.

References

[1] Assouad, P. (1975). Espaces p-lisses et q-convexes, inégalités de Burkholder. In *Séminaire Maurey-Schwartz* 1974–1975: *Espaces L^p, Applications Radonifiantes et Géométrie des Espaces de Banach*, *Exp. No. XV* 8. Centre Math., École Polytech., Paris. MR0407963

[2] Berger, E. (1990). An almost sure invariance principle for stationary ergodic sequences of Banach space valued random variables. *Probab. Theory Related Fields* **84** 161–201. MR1030726

[3] Burkholder, D.L. (1988). Sharp inequalities for martingales and stochastic integrals. *Astérisque* **157–158** 75–94. Colloque Paul Lévy sur les Processus Stochastiques (Palaiseau, 1987). MR0976214

[4] Cuny, C. Limit theorems under the Maxwell–Woodroofe condition in the Banach-valued case. Preprint.

[5] Cuny, C. (2011). Pointwise ergodic theorems with rate with applications to limit theorems for stationary processes. *Stoch. Dyn.* **11** 135–155. MR2771346

[6] Cuny, C. and Merlevède, F. (2014). Strong invariance principles with rate for "reverse" martingales and applications. *J. Theoret. Prob.* To appear, available at arXiv:1209.3677.

[7] Cuny, C. and Volný, D. (2013). A quenched invariance principle for stationary processes. *ALEA Lat. Am. J. Probab. Math. Stat.* **10** 107–115. MR3083921

[8] Dedecker, J., Doukhan, P. and Merlevède, F. (2012). Rates of convergence in the strong invariance principle under projective criteria. *Electron. J. Probab.* **17** 31. MR2900457

[9] Dedecker, J., Gouëzel, S. and Merlevède, F. (2010). Some almost sure results for unbounded functions of intermittent maps and their associated Markov chains. *Ann. Inst. Henri Poincaré Probab. Stat.* **46** 796–821. MR2682267

[10] Dedecker, J., Gouëzel, S. and Merlevède, F. (2012). The almost sure invariance principle for unbounded functions of expanding maps. *ALEA Lat. Am. J. Probab. Math. Stat.* **9** 141–163. MR2904480

[11] Dedecker, J. and Merlevède, F. (2007). The empirical distribution function for dependent variables: Asymptotic and nonasymptotic results in $\mathbb{L}^p$. *ESAIM Probab. Stat.* **11** 102–114. MR2299650

[12] Dedecker, J. and Merlevède, F. (2007). Convergence rates in the law of large numbers for Banach-valued dependent variables. *Teor. Veroyatn. Primen.* **52** 562–587. MR2743029

[13] Dedecker, J. and Merlevède, F. (2010). On the almost sure invariance principle for stationary sequences of Hilbert-valued random variables. In *Dependence in Probability, Analysis and Number Theory* 157–175. Heber City, UT: Kendrick Press. MR2731073

[14] Dedecker, J., Merlevède, F. and Pène, F. (2014). Rates in the strong invariance principle for ergodic automorphisms of the torus. *Stoch. Dyn.* To appear, available at arXiv:1206.4336.

[15] Dedecker, J. and Prieur, C. (2005). New dependence coefficients. Examples and applications to statistics. *Probab. Theory Related Fields* **132** 203–236. MR2199291

[16] Dehling, H., Denker, M. and Philipp, W. (1986). A bounded law of the iterated logarithm for Hilbert space valued martingales and its application to U-statistics. *Probab. Theory Related Fields* **72** 111–131. MR0835162

[17] Doob, J.L. (1990). *Stochastic Processes*. *Wiley Classics Library*. New York: Wiley. Reprint of the 1953 original, A Wiley-Interscience Publication. MR1038526

[18] Freedman, D.A. (1975). On tail probabilities for martingales. *Ann. Probab.* **3** 100–118. MR0380971

[19] Hannan, E.J. (1979). The central limit theorem for time series regression. *Stochastic Process. Appl.* **9** 281–289. MR0562049

[20] Krengel, U. (1985). *Ergodic Theorems*. *de Gruyter Studies in Mathematics* **6**. Berlin: de Gruyter. With a supplement by Antoine Brunel. MR0797411

[21] Kuelbs, J. (1976). A strong convergence theorem for Banach space valued random variables. *Ann. Probab.* **4** 744–771. MR0420771

[22] Ledoux, M. and Talagrand, M. (1988). Characterization of the law of the iterated logarithm in Banach spaces. *Ann*. *Probab*. **16** 1242–1264. MR0942766
[23] Ledoux, M. and Talagrand, M. (1991). *Probability in Banach Spaces*: *Isoperimetry and Processes*. *Ergebnisse der Mathematik und Ihrer Grenzgebiete* (3) [*Results in Mathematics and Related Areas* (3)] **23**. Berlin: Springer. MR1102015
[24] Liu, W. and Lin, Z. (2009). Strong approximation for a class of stationary processes. *Stochastic Process*. *Appl*. **119** 249–280. MR2485027
[25] McLeish, D.L. (1975). A maximal inequality and dependent strong laws. *Ann*. *Probab*. **3** 829–839. MR0400382
[26] Merlevède, F. (2008). On a maximal inequality for strongly mixing random variables in Hilbert spaces. Application to the compact law of the iterated logarithm. *Ann*. *I.S.U.P.* **52** 47–60. MR2435040
[27] Morrow, G. and Philipp, W. (1982). An almost sure invariance principle for Hilbert space valued martingales. *Trans*. *Amer. Math*. *Soc*. **273** 231–251. MR0664040
[28] Philipp, W. (1986). A note on the almost sure approximation of weakly dependent random variables. *Monatsh*. *Math*. **102** 227–236. MR0863219
[29] Pinelis, I. (1994). Optimum bounds for the distributions of martingales in Banach spaces. *Ann*. *Probab*. **22** 1679–1706. MR1331198
[30] Pisier, G. (1976). Sur la loi du logarithme itéré dans les espaces de Banach. In *Probability in Banach Spaces* (*Proc*. *First Internat*. *Conf*., *Oberwolfach*, 1975) 203–210. *Lecture Notes in Math*. **526**. Berlin: Springer. MR0501237
[31] Woyczyński, W.A. (1982). Asymptotic behavior of martingales in Banach spaces. II. In *Martingale Theory in Harmonic Analysis and Banach Spaces* (*Cleveland*, *Ohio*, 1981). *Lecture Notes in Math*. **939** 216–225. Berlin: Springer. MR0668549
[32] Wu, W.B. (2007). Strong invariance principles for dependent random variables. *Ann*. *Probab*. **35** 2294–2320. MR2353389

Received October 2012 and revised June 2013