LIMIT THEOREMS UNDER THE MAXWELL-WOODROOFE CONDITION IN BANACH SPACES

CHRISTOPHE CUNY

ABSTRACT. We prove that, for (adapted) stationary processes, the so-called Maxwell-Woodroofe condition is sufficient for the law of the iterated logarithm and that it is optimal in some sense. We obtain a similar conclusion concerning the Marcinkiewicz-Zygmund strong law of large numbers. Those results actually hold in the context of Banach valued stationary processes, including the case of \( L^r \)-valued random variables, with \( 1 \leq r < \infty \). In this setting we also prove the weak invariance principle, under a version of the Maxwell-Woodroofe condition, generalizing a result of Peligrad and Utev [38]. Our results extend to non-adapted processes as well, and, partly to stationary processes arising from dynamical systems. The proofs make use of a new maximal inequality and of approximation by martingales, for which some of our results are also new.

MSC 2010 subject classification: 60F17, 60F25, 60B12; Secondary: 37A50

1. Introduction

Let \( (\Omega, \mathcal{F}, \mathbb{P}) \) be a probability space, \( \theta \) be an invertible bi-measurable measure preserving transformation on \( \Omega \) and \( \mathcal{F}_0 \subset \mathcal{F} \) a \( \sigma \)-algebra such that \( \mathcal{F}_0 \subset \theta^{-1}(\mathcal{F}_0) \). Define a non-decreasing filtration by \( \mathcal{F}_n = \theta^{-n}(\mathcal{F}_0) \), for every \( n \in \mathbb{Z} \) and denote \( \mathbb{E}_n := \mathbb{E}(\cdot|\mathcal{F}_n) \). For every \( X \in L^1(\Omega, \mathcal{F}, \mathbb{P}) \) write \( S_n(X) = X + \ldots + X \circ \theta^{n-1} \).

In 2000, Maxwell and Woodroofe [33] proved the CLT for \( (X \circ \theta^n)_{n \geq 0} \) under the condition

\[
\sum_{n \geq 1} \frac{\|\mathbb{E}_0(S_n)\|_2}{n^{3/2}} < \infty .
\]

Actually, Maxwell and Woodroofe worked in a Markov chain setting, but in our context their condition reads as above.

This was a considerable improvement of the martingale-coboundary condition of Gordin and Lifšic [24] which in our setting is equivalent to the boundedness of \( (\|\mathbb{E}_0(S_n(X))\|_2)_{n \geq 1} \).

Moreover, condition (1) proved to be useful in applications. It is directly checkable for linear processes with innovations that are martingale differences, see e.g. Zhao and Woodroofe (Proposition 5 and its proof). It leads to the optimal sufficient condition for the CLT in the case of \( \rho \)-mixing processes, see Merlevède, Peligrad and Utev [36] pages 14-15. It is implied by the condition \( \sum_n (\log n)^{1+\epsilon} \frac{\|\mathbb{E}_0(S_n)\|_2^2}{n^2} < \infty \), which can be checked in the case of Markov chains with normal Markov operator, see Cuny [7]. Finally, it is implied by the following condition which is easier to check in applications (see e.g. [8] sections 3.1 and 3.3)

\[
\sum_{n \geq 1} \frac{\|\mathbb{E}_0(X \circ \theta^n)\|_2}{n^{1/2}} < \infty .
\]

For more situations where the conditions (1) and (2) can be checked we refer to [36] and the references therein.

Key words and phrases. Banach valued processes, compact law of the iterated logarithm, invariance principles, Maxwell-Woodroofe’s condition, Marcinkiewicz-Zygmund strong law.

The author is thankful to Jérôme Dedecker for providing him with a copy of [48].
Because of those potential applications several authors tried to have a better understanding of the condition \((1)\) and its connection with probabilistic results such as maximal inequalities, the weak invariance principle, the law of the iterated logarithm (LIL) and others.

A key step toward that better understanding was the paper \([33,(2005)]\) by Peligrad and Utev \([38]\) who proved a new maximal inequality and applied it to deduce the weak invariance principle (WIP) under \((1)\). Moreover, they proved that \((1)\) is, in some sense, optimal for the CLT.

Later, Peligrad, Utev and Wu \([39]\) and Wu and Zhao \([50]\) proved \(L^p\)-versions of that maximal inequality, in the cases \(p \geq 2\) and \(1 < p \leq 2\) respectively and obtained new results under \(L^p\)-versions of \((1)\).

Further extensions of those maximal inequalities have been obtained recently by Merlevède and Peligrad \([35]\).

On another hand, the quenched CLT (a strengthening of the CLT), the quenched invariance principle and the law of the iterated logarithm (LIL) have been obtained, under various strengthening of \((1)\), by Derriennic and Lin \([20]\), Zhao and Woodroofe \([51]\), Cuny and Lin \([9]\) and Cuny \([7]\).

Very recently, Cuny and Merlevède \([11]\) investigated the martingale approximation method under \(L^p\)-versions of \((1)\) and, using a new maximal inequality inspired by \([35]\), they proved the quenched invariance principle under \((1)\). In view of all the above mentioned results, it seems very probable that the LIL (and its invariance principle) be true under \((1)\) with some kind of optimality.

In this paper we answer positively to that question and show that the example of Peligrad and Utev \([38]\) ensures the optimality. Actually our results hold in a Banach space setting, including any (separable) \(L^r\) spaces of \(\sigma\)-finite measure space. More precisely, we prove the almost sure invariance principle (ASIP) in 2-smooth Banach spaces or in \(L^r\) spaces with \(1 \leq r < 2\). We also obtain the WIP for dependent variables taking values in a 2-smooth Banach space or a Banach space of cotype 2. Finally, we investigate the Marcinkiewicz-Zygmund strong law of large numbers (MZ-SLLN). In the course we extend the maximal inequalities of \([39]\) and \([50]\) to Banach space valued variables, with improved constants.

The main motivation for considering Banach-valued variables (especially the \(L^r\) case, with \(1 \leq r < \infty\)) is the fact that there are applications in statistics, in the study of the empirical process, see section 8.2. Let us mention some papers in this vein: del Barrio, Giné and Matrán \([2]\), Berkes, Horváth, Shao and Steinebach \([3]\), Dedecker and Merlevède \([16],[14]\) and \([15]\) or Dédé \([12]\).

To give a flavour of our results we shall state here a theorem in \(L^r\), \(1 \leq r < 2\).

Let \((S,\mathcal{S},\mu)\) be a \(\sigma\)-finite measure space such that \(L^1(S,\mathcal{S},\mu)\) is separable (for instance assume that \(S\) be countably generated). Let \(X(s)\) be a random variable on \((\Omega,\mathcal{F}_0,\mathbb{P})\) with values in \(L^r(S,\mathcal{S},\mu)\), for some \(1 \leq r < \infty\). We shall often consider \(X\) as a (class of a) measurable function on \((\Omega \times S,\mathcal{F}_0 \otimes S,\mathbb{P} \otimes \mu)\), without mentionning it.

For every integer \(n \geq 0\), write \(X_n = X \circ \theta^n\). For every \(t \in [0,1]\) and every integer \(n \in \mathbb{N}\), write \(S_{n,t}(X) := \sum_{k=0}^{\lfloor nt \rfloor - 1} X_k + (nt - \lfloor nt \rfloor)X_{\lfloor nt \rfloor}\) and \(T_{n,t} := S_{n,t}/\sqrt{n}\).

For the sake of readability, we state the next theorem under a condition in the spirit of \([2]\) rather than \((1)\). With this formulation, the ASIP has already been obtained by the author \([8]\), when \(r \geq 2\), and the CLT has been obtained by Dédé \([12]\) when \(r = 1\).

We denote by \(\| \cdot \|_2\) the \(L^2\)-norm on \((\Omega,\mathbb{P})\).
Theorem 1.1. Assume that \( \theta \) is ergodic. Let \( X \in L^2(\Omega, F_0, \mathbb{P}, L^r(S)) \) (1 \( \leq r < \infty \)) be such that \( N_r(X) < \infty \), where

\[
N_r(X) = \sum_{n\geq 1} \left( \int_S \left| \mathbb{E}_0(X_n(s)) \right|^r \mu(ds) \right)^{1/r} n^{1/2} \quad \text{if } 1 \leq r < 2,
\]

\[
N_r(X) = \sum_{n\geq 1} \left( \int_S \left| \mathbb{E}_0(X_n(s)) \right|^r \mu(ds) \right)^{1/r} \| \right\|_2 \quad \text{if } r \geq 2.
\]

Then, the process \(((T_n,t)_{0\leq t \leq 1})_{n\geq 1}\) converges in law in \( C([0,1], L^r(S,\mu)) \) (to an \( L^r(S,\mu) \)-valued brownian motion); \( (S_n(X)/\sqrt{nL(L(n))})_{n\geq 1} \) is \( \mathbb{P} \)-a.s. relatively compact in \( L^r(S,\mu) \).

Moreover, there exists a universal constant \( C > 0 \), such that

\[
\limsup_{n \to +\infty} \left( \int_S \left| \sum_{k=0}^{n-1} X_k(s) \right|^r \mu(ds) \right)^{1/r} \sqrt{2nL(L(n))} \leq CN_r \quad \mathbb{P}\text{-a.s.}
\]

The exact value of the (\( \mathbb{P} \)-a.s. contant) limsup above may be derived from the proof. The above WIP is new, apart from the above mentionned result of Dédé.

Under the assumptions of the theorem an ASIP holds as well, see Theorem 5.2 and Theorem 5.3.

Our method of proof follows a classical line. To prove the weak invariance principle, we first prove tightness of the underlying process and then prove convergence in law of the finite-dimensional distributions. To prove the almost sure invariance principle (in particular the functional law of the iterated logarithm) we first prove a compact law of the iterated logarithm (CLIL) and then invoke an important result of Berger, see Theorem B.3. The tightness and the CLIL are obtained thanks to suitable maximal inequalities. Our proofs make also use of martingale approximation arguments, in particular we first prove all results for martingale differences.

The paper is organised as follows. In section 2, we recall some definitions and lemmas, about probability in Banach spaces, that are necessary for the understanding of the statement and/or the proofs of the results. In section 3, we state all the results (some of them are new) for martingale with stationnary (and ergodic) increments that are needed in the sequel. In section 4, we state maximal inequalities under projective conditions. In section 5, we state our limit theorems under projective conditions. In section 6, we explain how to extend our results to non-adapted processes or to processes arising from non-invertible dynamical systems. In section 7, we prove the optimality of our conditions, and in section 8, we provide several examples including the case of the empirical process. All the results are proved in the appendix.

2. Generalities on probability on Banach spaces

Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space. We will consider Banach-valued random variables. We refer to the book by Diestel and Uhl [21] for the basic facts on the topic (definition, conditional expectation...). We shall also use results or notations from Ledoux and Talagrand [30]. In all the paper, we shall be concerned only with separable Banach spaces, in which case the definitions of a random variable of [21] and [30] coincide.

In all the paper, \((\mathcal{X}, | \cdot |_\mathcal{X})\) will be a real separable Banach space. Denote by \( L^0(\mathcal{X}) \) the space of (classes modulo \( \mathbb{P} \) of) functions from \( \Omega \) to \( \mathcal{X} \) that are limits \( \mathbb{P} \)-a.s. of simple (or step) functions. We define, for every \( p \geq 1 \), the usual Bochner spaces \( L^p \) and their weak versions, as follows.
Lemma 2.1. For the next (easy) two results, whose proofs are given in the appendix.

Let \( p > 4 \) \( \forall \end{align*} \)

\[ X \]

\[ L_p^p(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{X}) = \{ Z \in L^0(\mathcal{X}) : \mathbb{E}(|Z|_p^p) < \infty \} ; \]

\[ L_{p,\infty}(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{X}) = \{ Z \in L^0(\mathcal{X}) : \sup_{t>0} t(\mathbb{P}(|Z|_t > t))^{1/p} < \infty \}. \]

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

For the sake of clarity, when they are understood, some of the references to \( \Omega, \mathcal{F} \) 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 \( L_p^p(\mathcal{X}, \mathcal{Y}) \) (see the proof of Lemma [2]), equivalent to the quasi-norm \( \|\cdot\|_{p,\infty,\mathcal{Y}} \), that makes \( L_{p,\infty}(\mathcal{X}, \mathcal{Y}) \) a Banach space.

We will state our results in the context of Banach spaces that are \( r \)-smooth or of cotype 2. Let us recall the definitions of those 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

\[ |x + y|_\mathcal{X}^r + |x - y|_\mathcal{X}^r \leq 2(|x|_\mathcal{X}^r + L'r|y|_\mathcal{X}^r) \quad \forall x, y \in \mathcal{X}. \]

We shall speak about \( (r, L) \)-smooth spaces to emphasize the constant \( L \) such that (4) is satisfied.

**Remark.** A Banach space is said to be \( r \)-convex whenever (4) holds in the reverse direction.

**Definition 2.2.** We say that \( (d_n)_{1 \leq n \leq N} \subseteq L^1(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{X}) \) is a sequence of martingale differences, if there exist non-decreasing \( \sigma \)-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 \( \mathbb{E}(d_n|\mathcal{G}_{n-1}) = 0 \ P \)-a.s.

The notion of smooth Banach spaces is very useful due to the following results, see for instance Proposition 1 of Assouad [1] (and its corollary) or see the proof of Corollary [2.2] below.

Assume that \( \mathcal{X} \) is \( (r, L) \)-smooth. Then, for every martingale differences \( (d_n)_{1 \leq n \leq N} \), we have

\[ \mathbb{E}(|d_1 + \cdots + d_N|_\mathcal{X}^r) \leq 2L'r \sum_{n=1}^{N} \mathbb{E}(|d_n|_\mathcal{X}^r). \]

Notice that a Banach space which is \( (r, L) \)-smooth is \( (r', L') \)-smooth for any \( 1 \leq r' \leq r \). Any Hilbert space is \( (2, 1) \)-smooth.

Any \( L^p \) space, \( p > 1 \), (of \( \mathbb{R} \)-valued functions) associated with a \( \sigma \)-finite measure is \( (r, L) \)-smooth with \( r = \min(2, p) \) and \( L = \max(1, \sqrt{p-1}) \) (see [40] Proposition 2.1 for the case \( p \geq 2 \) and Clarkson [6] for the case \( 1 < p \leq 2 \)).

Let us recall the following way to generate more smooth spaces. We did not find references for the next (easy) two results, whose proofs are given in the appendix.

**Lemma 2.1.** Let \( \mathcal{X} \) be a \( (r, L) \)-smooth Banach space. Then, for every \( p \geq r \), \( L^p(\Omega, \mathcal{X}) \) is \( (r, L') \)-smooth with \( L' = L^r + (\max(1, \sqrt{p-1}))^r \).

From this, one derives easily the following extension of [5].

**Corollary 2.2.** Let \( \mathcal{X} \) be a \( (r, L) \)-smooth separable Banach space and \( p \geq r \). Let \( X, Y \in L^p(\Omega, \mathcal{X}) \) be such that \( \mathbb{E}(Y|X) = 0 \). Then, writing \( \tilde{L}' := L' + (\max(1, \sqrt{p-1}))^r \),

\[ \|X + Y\|_{p,\mathcal{X}}^{r'} \leq \|X\|_{p,\mathcal{X}}^{r'} + 2\tilde{L}'\|Y\|_{p,\mathcal{X}}^{r'}. \]
In particular, if $X_1, \ldots, X_n \in L^p(\Omega, \mathcal{X})$ with $\mathbb{E}(X_k | X_1 + \ldots + X_{k-1}) = 0$, for every $2 \leq k \leq n$, then
\begin{equation}
\|X_1 + \ldots + X_n\|_{p,\mathcal{X}} \leq 2L^r(\|X_1\|_{p,\mathcal{X}} + \ldots + \|X_n\|_{p,\mathcal{X}}).
\end{equation}

\textbf{Remark.} When $p = r$, we recover (5). When $r = 2$ and $\mathcal{X} = \mathbb{R}$, (6) has been obtained by Rio \cite{22} with the constant $p - 1$ instead of $2(p - 1 + L^2)$. Rio proved that his constant is optimal.

We shall also need the concept of Banach spaces of type 2 and of cotype 2. These concepts are relevant in the study of the central limit theorem in Banach spaces, in particular in their relation with the notion of pregaussian variables that we shall introduce later.

\textbf{Definition 2.3.} We say that a separable Banach space $\mathcal{X}$ is of type 2 (respectively of cotype 2) if there exists $L > 0$ such that for every independent random variables $a_1, \ldots, a_n \in L^r(\Omega, \mathcal{X})$, (5) holds with $r = 2$ (respectively, such that (5) holds in the reverse direction with $r = 2$).

Of course, any 2-smooth Banach space is of type 2.

Now, we explain what we mean by an invariance principle in a Banach space.

Let us denote by $\mathcal{X}^*$ the topological dual of $\mathcal{X}$. Let $X \in L^0(\Omega, \mathcal{X})$ be such that for every $x^* \in \mathcal{X}^*$, $\mathbb{E}(x^*(X))^2 < \infty$ and $\mathbb{E}(x^*(X)) = 0$. We define a bounded symmetric bilinear operator $K = K_X$ from $\mathcal{X}^* \times \mathcal{X}^*$ to $\mathbb{R}$, by
\[ K(x^*, y^*) = \mathbb{E}(x^*(X)y^*(X)) \quad \forall x^*, y^* \in \mathcal{X}^*. \]
The operator $K_X$ is called the covariance operator associated with $X$.

\textbf{Definition 2.4.} We say that a random variable $W \in L^0(\Omega, \mathcal{X})$ is gaussian if, for every $x^* \in \mathcal{X}^*$, $x^*(W)$ has a normal distribution. We say that a random variable $X \in L^0(\Omega, \mathcal{X})$, such that for every $x^* \in \mathcal{X}^*$, $\mathbb{E}(x^*(X))^2 < \infty$ and $\mathbb{E}(x^*(X)) = 0$, is pregaussian, if there exists a gaussian variable $W \in L^0(\Omega, \mathcal{X})$ with the same covariance operator, i.e. such that $K_X = K_W$. As in \cite{30}, when $X$ is pregaussian, we shall denote (abusively) by $G(X)$ a gaussian variable with same covariance operator than $X$.

\textbf{Definition 2.5.} We say that a process $(W_t)_{0 \leq t \leq 1} \in L^0(\Omega, C([0, 1], \mathcal{X}))$ is a Brownian motion with covariance operator $K$ if it is a gaussian process such that for every $x^*, y^* \in \mathcal{X}^*$ and every $0 \leq s, t \leq 1$, $\operatorname{cov}(x^*(W_s), y^*(W_t)) = \min(s, t)K(x^*, y^*)$.

\textbf{Definition 2.6.} 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 iid gaussian variables, such that
\[ |X_0 + \cdots + X_{n-1} - (W_0 + \cdots + W_{n-1})|_{\mathcal{X}} = o(\sqrt{nL(L(n))}) \quad \mathbb{P}\text{-a.s.} \]

We say that $(X_n)_{n \geq 0}$ satisfies the weak invariance principle (WIP) of covariance operator $K$ if $(T_{n,t})_{0 \leq t \leq 1}$ converges weakly in $C([0, 1], \mathcal{X})$ to a brownian motion of covariance operator $K$, where for every $t \in [0, 1]$ and every $n \geq 1$, $T_{n,t} = S_{n,t}/\sqrt{n}$ and $S_{n,t} = X_0 + \cdots + X_{[nt]-1} + (nt - [nt])X_{[nt]}$.

It is known that in order to have a central limit theorem (or a WIP) for a sequence of iid $\mathcal{X}$-valued random variables it is necessary that the variables be pregaussian. Hence, to prove invariance principles for stationary sequences, we shall consider only pregaussian variables. Let us state some useful facts about the connection between the notions of type 2 or cotype 2 and the property of being pregaussian.

Recall, see \cite{20} Lemma 3.8], that a gaussian variables admits moments of any order.
**Lemma 2.3** (Ledoux-Talagrand, [30] Lemma 9.23). Let \( X \in L^0(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{X}) \) be a pregaussian random variable. Let \( Y \in L^0(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{X}) \) be such that for every \( x^* \in \mathcal{X}^* \), \( \mathbb{E}(x^*(Y)^2) \leq \mathbb{E}(x^*(X)^2) \) and \( \mathbb{E}(x^*(Y)) = 0 \). Then, \( Y \) is pregaussian and, for every \( p > 0 \), we have \( \mathbb{E}((G(Y))_{\mathcal{X}}^p) \leq 2\mathbb{E}((G(X))_{\mathcal{X}}^p) \).

This lemma, in conjunction with the next one, will be very useful for proving results concerning pregaussian random variable in presence of cotype 2.

**Lemma 2.4** (Ledoux-Talagrand, [30] Proposition 9.25). Let \( \mathcal{X} \) be a real separable Banach space of cotype 2. Let \( X \in L^2(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{X}) \) be a pregaussian random variable. Then, \( X \in L^2(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{X}) \) is pregaussian if and only if \( \mathbb{E}((G(X))_{\mathcal{X}}^2) \leq C \mathbb{E}((G(X))_{\mathcal{X}}^2) \), for some \( C > 0 \) depending only on the constant in the definition of the cotype of \( \mathcal{X} \).

In presence of type 2 any square integrable variable is pregaussian.

**Lemma 2.5** (Ledoux-Talagrand, [30] Proposition 9.24). Let \( \mathcal{X} \) be a real separable Banach space of type 2. Let \( X \in L^2(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{X}) \) be a centered random variable. Then, \( X \) is pregaussian and \( \mathbb{E}((|G(X)|_{\mathcal{X}}^2)) \leq C \mathbb{E}((|G(X)|_{\mathcal{X}}^2)) \), for some \( C > 0 \) depending only on the constant in the definition of the type of \( \mathcal{X} \).

**Definition 2.7.** Let \( G(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{X}) = G(\mathcal{X}) \) be the set of pregaussian random variables that are in \( L^2(\Omega, \mathcal{X}) \). For every \( X \in G(\mathcal{X}) \), denote \( ||X||_{G(\mathcal{X})} := ||X||_{2,\mathcal{X}} + ||G(X)||_{2,\mathcal{X}} \).

**Lemma 2.6.** Let \( \mathcal{X} \) be a real separable Banach space. Then, for every pregaussian variables \( X, Y \), the variable \( X + Y \) is pregaussian and \( ||G(X + Y)||_{2,\mathcal{X}} \leq ||G(X)||_{2,\mathcal{X}} + ||G(Y)||_{2,\mathcal{X}} \). In particular, \( (G(\mathcal{X}), || \cdot ||_{G(\mathcal{X})}) \) is a normed vector space. Actually, it is a Banach space.

The proof is given in the appendix. The following result is an obvious consequence of Lemma 2.3, hence the proof is omitted.

**Lemma 2.7.** Let \( \mathcal{X} \) be a real separable Banach space. Let \( (\Omega, \mathcal{F}, \mathbb{P}) \) be a probability space, \( \theta \) be an invertible bi-measurable transformation on \( \Omega \). Let \( \mathcal{F}' \) be a sub-\( \sigma \)-algebra of \( \mathcal{F} \). Let \( X \in L^1(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{X}) \) be pregaussian. Then \( \mathbb{E}(X|\mathcal{F}') \) is pregaussian and for every \( n \geq 0 \), \( X \circ \theta^n \) is pregaussian. Moreover, \( ||\mathbb{E}(X|\mathcal{F}')||_{G(\mathcal{X})} \leq \sqrt{2} ||X||_{G(\mathcal{X})} \) and \( ||X \circ \theta^n||_{G(\mathcal{X})} \leq \sqrt{2} ||X||_{G(\mathcal{X})} \).

**Lemma 2.8.** Let \( \mathcal{X} \) be a real separable Banach space. Let \( (\mathcal{H}_n)_{n \geq 1} \) be a non-decreasing filtration and let \( \mathcal{H}_\infty := \bigvee_{n \geq 1} \mathcal{H}_n \). For every \( X \in G(\mathcal{X}) \), \( ||\mathbb{E}(X|\mathcal{H}_n) - \mathbb{E}(X|\mathcal{H}_\infty)||_{G(\mathcal{X})} \xrightarrow{n \to \infty} 0 \).

The proof is given in the appendix.

From the above lemmas, we see that it will be very convenient to work in \( G(\mathcal{X}) \) in order to obtain invariance principles for a sequence \( (X \circ \theta^n)_{n \geq 0} \) under conditions involving terms of the type \( \mathbb{E}_0(X \circ \theta^n) \). In order to have tractable conditions it is necessary to be able to compute \( ||X||_{G(\mathcal{X})} \). When \( \mathcal{X} \) is of type 2, by Lemma 2.3, \( || \cdot ||_{G(\mathcal{X})} \) is equivalent to \( || \cdot ||_{2,\mathcal{X}} \).

Let \( X = L^p(S, \mu) \) (1 \( p \leq 2 \)), for some \( \sigma \)-finite measure, (recall that, then, \( \mathcal{X} \) is of cotype 2). In this case, the following characterization of pregaussian variables is part of the folklore. It is due to Vakhania [46] when \( \mu \) is discrete (see [30] p. 262 for a proof). It seems to be essentially due to Rajput [41] for a general \( \sigma \)-finite measure \( \mu \). We provide more details in the appendix.

**Lemma 2.9.** Let \( X = L^p(S, \mu) \) (1 \( p \leq 2 \)), for some \( \sigma \)-finite measure. Then, \( X(s) \in L^2(\Omega, \mathbb{P}, \mathcal{X}) \) is pregaussian if and only if \( X \) is pregaussian and \( \int_S (\mathbb{E}|X(s)|^2)^{p/2} \mu(ds) < \infty \). Moreover, there exists \( C_p > 0 \), depending only on \( p \), such that

\[
(7) \quad \frac{||G(X)||_2}{C_p} \leq \left( \int_S (\mathbb{E}|X(s)|^2)^{p/2} \mu(ds) \right)^{1/p} \leq C_p ||G(X)||_2 \quad \forall X \in G(L^p(\mu)) .
\]

Hence, \( G(L^p(\mu)) \) may be identified with \( \{X \in L^p(S, L^2(\Omega, \mathbb{R})) : \mathbb{E}(X) = 0 \} \).

**Remark.** The above identification makes use of the natural embedding of \( L^p(S, L^2(\Omega, \mathbb{R})) \) into \( L^2(\Omega, \mathbb{P}, L^p(S, \mu)) \) (when 1 \( p \leq 2 \)), see Lemma 2.2.
3. Some limit theorems for Banach-valued martingales

In this section, we give maximal inequalities and limit theorems (WIP, ASIP and MZ-SLLN) for martingales with stationary differences \((d_n)_{n \geq 0}\). As mentioned in [8], there is no loss of generality in assuming that \(d_n = d \circ \theta^n\), where \(\theta\) is an invertible bi-measurable measure preserving transformation. Hence we shall use the notations of the introduction.

Let us mention that all the results of this section, except the ASIP in Proposition 3.5, hold for stationary differences of reverse martingales. Recall that \((d_n)_{n \geq 1} \subset L^1(\Omega, \mathcal{X})\) is a sequence of differences of reverse martingale if \(\mathbb{E}(d_n|\sigma\{d_k : k \geq n+1\}) = 0\).

Nevertheless, for stationary sequences of reverse martingales we know that the ASIP (as stated in Proposition 3.5) holds in the particular case where \(\mathcal{X} = \mathbb{R}\), see Cuny and Merlevède [11 Corollary 2.5].

Part of the results stated here are new. We shall discuss their novelty in the sequel.

3.1. Maximal inequalities. As mentioned, we use the notations from the introduction. Notice that we do not assume \(\theta\) to be ergodic here.

The following lemma is part of the folklore. Let us notice however that in [8] we obtain the right order of magnitude (\(\sqrt{p}\)) when \(p \to \infty\), see Rio [22, Remark 2.1].

**Lemma 3.1.** (i) Let \(p > 1\) and \(r := \min(2,p)\). Let \(\mathcal{X}\) be a real separable \((r, L)\)-smooth Banach space, for some \(L \geq 1\). Then, there exists \(C_{p, L} > 0\) such that, for every \(d \in L^p(\Omega, \mathcal{F}_0, \mathcal{X})\) with \(\mathbb{E}_{-1}(d) = 0\), we have

\[
\| \max_{1 \leq k \leq n} |S_k(d)|_{\mathcal{X}} \|_p \leq C_{p, L} n^{1/r} \|d\|_{p, L},
\]

where \(C_{p, L} = 2^{1/p} L^p/(p-1)\) if \(1 < p \leq 2\) and \(C_{p, L} = \sqrt{\frac{p}{p-1}}(p-1 + L^2)^{1/2}\) if \(p \geq 2\).

(ii) Let \(\mathcal{X}\) be a real separable Banach space of cotype \(2\). There exists \(C > 0\) such that for every pregaussian random variable \(d \in L^2(\Omega, \mathcal{X})\)

\[
\| \max_{1 \leq k \leq n} |S_k(d)|_{\mathcal{X}} \|_2 \leq C n^{1/2} \|G(d)\|_{2, \mathcal{X}},
\]

where (see Definition [2.3]) \(G(d)\) stands for a gaussian variable with same covariance operator than \(d\).

**Remarks.** The constant \(C\) in item (ii) depends only on the constant in the definition of the cotype of \(\mathcal{X}\).

The estimates [8] (with \(p = r = 2\)) and [9] are useful in order to prove tightness results related to the WIP. We notice (using Lemmas 2.4 and 2.5) that, whenever \(\mathcal{X}\) is \(2\)-smooth or has cotype \(2\), if \(d \in \mathcal{G}(\mathcal{X})\), \(\| \max_{1 \leq k \leq n} |S_k(d)|_{\mathcal{X}} \|_2 \leq C n^{1/2} \|d\|_{\mathcal{G}(\mathcal{X})}\).

We now state some maximal inequalities that are related to the ASIP or the MZ-SLLN.

For every \(X \in L^p(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{X})\), we consider the following maximal functions

\[
\mathcal{M}_p(X, \theta, \mathcal{X}) := \sup_{n \geq 1} \left| \sum_{k=0}^{n-1} X \circ \theta^k \right|_{\mathcal{X}}, \quad \text{if } 1 \leq p < 2,
\]

\[
\mathcal{M}_2(X, \theta, \mathcal{X}) := \sup_{n \geq 1} \left| \sum_{k=0}^{n-1} X \circ \theta^k \right|_{\mathcal{X}} / \sqrt{n L(L(n))},
\]

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

We shall omit the dependence in the parameters \(\theta\) and/or \(\mathcal{X}\) when they are understood.

**Proposition 3.2.** Let \(1 < p < 2\) and \(1 \leq r < 2\) with \(r \neq p\). Let \(\mathcal{X}\) be a Banach space. Assume that \(\mathcal{X}\) is \((r, L)\)-smooth, if \(p < r\) or that \(\mathcal{X} = L^r(S, \mathcal{S}, \mu)\) with \(\mu\) \(\sigma\)-finite, if \(1 \leq r < p\). Then, there exists \(C_{p, r} > 0\) such that, for every \(d \in L^p(\Omega, \mathcal{F}_0, \mathcal{X})\) with \(\mathbb{E}_{-1}(d) = 0\), we have

\[
\| \mathcal{M}_p(d) \|_{p, \infty} \leq C_{p, r} K_r(d),
\]
where \( K_r(d) = L^{r/p} \|d\|_{p,X} \) if \( X \) is \((r,L)\)-smooth and \( K_r(d) = \left( \int_S (\|d(s)\|_p)^r \, d\mu(s) \right)^{1/r} \).

Moreover,
\[
|S_n(d)|_{X}/n^{1/p} \xrightarrow{n \to \infty} 0 \quad \mathbb{P}\text{-a.s.}
\]

**Remark.** Only the case \( 1 \leq r < p \) is new here. The case \( p < r \leq 2 \) is just Proposition 2.1 of [8].

**Proposition 3.3.** Let \( X \) be a Banach space. Assume either that \( X \) is a \((2, L)\)-smooth Banach space or \( X = L^r(S, \mathcal{S}, \mu) \) with \( \mu \) \(\sigma\)-finite and \( 1 \leq r \leq 2 \). Then, for every \( 1 < p < 2 \) (resp. for every \( r < p < 2 \)) there exists \( C_p > 0 \) (resp. \( C_{p,r} > 0 \)) such that
\[
\|M_2(d)\|_{p,\infty} \leq C \|d\|_{G(X)},
\]
where \( C = LC_p \) (resp. \( C = C_{p,r} \)).

**Remark.** Again, only the case \( X = L^r(S, \mathcal{S}, \mu) \), \( 1 \leq r < 2 \) is new here. The proposition is proved in [8] when \( X \) is \((2, L)\)-smooth.

### 3.2. WIP and ASIP for martingales in Banach spaces

We shall assume here that \( \theta \) is ergodic.

The CLT for martingales with stationary and ergodic increments in 2-smooth Banach spaces (admitting a Schauder basis) has been obtained by Woyczynski [45]. Rosinski [43] considered the case of general arrays of martingale increments a la Brown (in the \( r \)-smooth case). As far as we know, the only CLT for martingales taking values in a Banach space of cotype 2 has been obtained by Dédé [12] in the special case where \( X = L^1(S, \mathcal{S}, \mu) \), with \( \mu \) \(\sigma\)-finite.

Hence, the CLT in the next proposition is only partly new, while the WIP seems to be new.

**Proposition 3.4.** Let \( X \) be a real separable Banach space that is either 2-smooth or of cotype 2. Let \( d \in \mathcal{G}(\mathcal{F}_0, X) \) such that \( \mathbb{E}_{-1}(d) = 0 \). Then \((d \circ \theta^n)_{n \geq 0}\) satisfies the WIP of covariance \( K_d \).

**Proposition 3.5.** Let \( X \) be either a 2-smooth Banach space or \( X = L^p(S, \mathcal{S}, \mu) \), for some \( 1 \leq p \leq 2 \) and \( \sigma\)-finite \( \mu \). Let \( d \in \mathcal{G}(X) \) such that \( \mathbb{E}_{-1}(d) = 0 \). Then, \((d \circ \theta^n)_{n \geq 0}\) satisfies the CLIL, hence, in particular, the ASIP. Moreover
\[
\limsup_{n \to \infty} \frac{|S_n(d)|_{X}}{\sqrt{nL(L(n))}} = \sup_{x^* \in X^*, \|x^*\|_{X^*} \leq 1} \|x^*(d)\|_2 \quad \mathbb{P}\text{-a.s.}
\]

**Remarks.**
1. The ergodicity of \( \theta \) is not necessary for the CLIL. As already mentioned the CLIL also holds for stationary differences of reverse martingales.
2. Only the case \( X = L^r(S, \mathcal{S}, \mu) \), \( 1 \leq r < 2 \) is new here. The case where \( X \) is 2-smooth has been obtained in [8].

### 4. Maximal inequalities under projective conditions

In all of this section we do not require \( \theta \) to be ergodic.

Before going further, let us introduce the generalized version of the Maxwell-Woodroofe condition that we shall need in the sequel. Their relevance will be clear from the next results.

Let \( X \in L^p(\Omega, \mathcal{X}) \) for some \( 1 < p \leq 2 \). We define \( \|X\|_{MW_p} \) as follows.
\[
\|X\|_{MW_p} := \sum_{n \geq 0} \frac{\|\mathbb{E}_0(S_{2n}(X))\|_{p,X}}{2^{n/p}},
\]
\[
\|X\|_{MW_2} := \sum_{n \geq 0} \frac{\|\mathbb{E}_0(S_{2n}(X))\|_{G(X)}}{2^{n/2}}.
\]
To have a better understanding of $\| \cdot \|_{MW_2}$ recall that if $X$ is of type 2 (in particular if $X$ is $2$-smooth), then $\| \cdot \|_{C(X)} \leq C \| \cdot \|_{2,X}$ and that if $X = L^r(S,S,\mu)$ with $1 \leq r \leq 2$ and $\mu$ $\sigma$-finite, we have \( \| \cdot \|_{MW_2} \).

In view of applications, let us mention the following easy fact based on the observation that $\| E_0(S_n) \| \leq \| E_0(X) \| + \ldots + \| E_0(X \circ \theta^n) \|$. There exists $C_p > 0$ such that

$$
\| X \|_{MW_p} \leq C_p \sum_{n \geq 1} \frac{\| E_0(X \circ \theta^n) \|_{p,X}}{n^{1/p}} ,
$$

$$
\| X \|_{MW_2} := C_2 \sum_{n \geq 1} \frac{\| E_0(X \circ \theta^n) \|_{C(X)}}{n^{1/2}} .
$$

We first give an almost-sure maximal inequality, whose proof is based on the dyadic chaining in its simplest form, taking into account our filtration. Then we derive several other maximal inequalities that will be needed later, and that have interest in their own.

There are two important points concerning the following proposition. Firstly, it involves the terms $(E_{-2k}(S_{2k}))_{k \geq 0}$ which appear in the Maxwell-Woodroofe condition (notice that $\| E_{-2k}(S_{2k}) \|_{p,X} \leq \| E_0(S_{2k}) \|_p$). Secondly, for every $k \geq 0$, the sequence $(d_k \circ \theta^{2k+1})_{\ell \geq 0}$ defined below is a stationary sequence of martingale differences.

**Proposition 4.1.** Let $X \in L^1(\Omega,F_0,\mathbb{P},X)$. For every $k \geq 0$, write $u_k := \| E_{-2k}(S_{2k}) \|_X$ and $d_k := E_{-2k}(S_{2k}) + E_{-2k}(S_{2k}) \circ \theta^{2k} - E_{-2k+1}(S_{2k+1}).$ Then, for every integer $d \geq 0$, we have (with the convention $\sum_{k=0}^{-1} = 0$)

$$
\max_{1 \leq i \leq 2^d} |S_i|_X \leq \max_{1 \leq i \leq 2^d} \left( X - E_{-1}(X) \right) \circ \theta^\ell \bigg|_X + \sum_{k=0}^{-1} \max_{1 \leq i \leq 2^{d-k-1}} \left| \sum_{\ell=0}^{d-1} d_k \circ \theta^{2k+1} \right|_X + u_d + \sum_{k=0}^{-1} \max_{0 \leq \ell \leq 2^{d-k-1}} u_k \circ \theta^{2k+1}. 
$$

In particular, there exists $C > 0$ such that for every $p \geq 1$,

$$
\mathcal{M}_p(X, \theta) \leq C \left( \sum_{k \geq 0} \frac{u_k}{2^{k/p}} + \sum_{k \geq 0} \frac{(\mathcal{M}_1(u^p_k, \theta^{2k+1})^{1/p}}{2^{k/p}} \right) + \mathcal{M}_p(X - E_{-1}(X), \theta) + \sum_{k \geq 0} \frac{\mathcal{M}_p(d_k, \theta^{2k+1})}{2^{k/p}} .
$$

**Remark.** The statement and the proof of the proposition are inspired by the works of Peligrad, Utev and Wu \[39\] and of Wu and Zhao \[50\].

We derive the following version of the maximal inequalities \[39\] Theorem 1 and \[50\] Theorem 3. Notice that our constants are better when $p \to 1$ or $p \to +\infty$.

**Corollary 4.2.** Let $p > 1$ and $r := \min(2,p)$. Let $X$ be a $(r,L)$-smooth Banach space for some $L \geq 1$. There exists $C_{p,L} > 0$ such that or every $X \in L^p(\Omega,F_0,\mathbb{P},X)$ and every integer $d \geq 0$, we have

$$
\| \max_{1 \leq i \leq 2^d} |S_i|_X \|_p \leq C_{p,L} 2^{d/r} \left( \| X \|_{p,X} + \sum_{k=0}^{d} 2^{-k/r} \| E_{-2k}(S_{2k}) \|_{p,X} \right) ,
$$

where $C_{p,L} = \frac{5pL2^{1/p}}{p-1}$ when $1 < p \leq 2$ and $C_p = \frac{5\sqrt{2}p}{p-1} (p-1 + L^2)^{1/2}$, when $p > 2$. 
Corollary 4.3. Let \( X \) be a Banach space that is \( 2 \)-smooth or of cotype 2. There exists \( C > 0 \) such that for every \( X \in \mathcal{G}(\mathcal{X}) \) and every integer \( d \geq 0 \), we have

\[
\| \max_{1 \leq k \leq d} |S_k|X \|_2 \leq C 2^{d/2} \left( \|X\|_{\mathcal{G}(\mathcal{X})} + \sum_{k=0}^{d} 2^{-k/2} \|\mathbb{E}_{-2k}(S_{2k})\|_{\mathcal{G}(\mathcal{X})} \right).
\]

In particular, if \( \|X\|_{ MW_2} < \infty \), then

\[
(18) \quad \sup_{n \geq 1} \frac{\|\max_{1 \leq k \leq n} |S_k(X)|X\|_2}{\sqrt{n}} \leq C \|X\|_{ MW_2}.
\]

Using Proposition 4.1 combined with Lemma 3.2 and 3.3 we easily derive the following maximal inequalities.

Proposition 4.4 (MZ-SLLN). Let \( 1 < p < r \leq 2 \) and \( L \geq 1 \). Let \( \mathcal{X} \) be a separable \((r,L)\) smooth Banach space. Let \( X \in L^p(\Omega, F_0, \mathbb{P}) \) such that \( \|X\|_{ MW_p} < \infty \). There exists a constant \( C_{p,r} > 0 \), depending only on \( p \) and \( r \) such that

\[
(19) \quad \|M_p(X)\|_{p,\infty,\mathcal{X}} \leq C_{p,r} L^{r/p} \|X\|_{ MW_p}.
\]

Moreover,

\[
(20) \quad |S_n|X = o(n^{1/p}) \quad \mathbb{P}\text{-a.s.}
\]

Remark. The conclusion (20) has been obtained in [11], when \( \mathcal{X} = \mathcal{H} \) is a Hilbert space, under the stronger condition \( \sum_{n \geq 1} n \|\mathbb{E}_0(S_{2n})\|_{p,\mathcal{H}} 2^{-n/2} < \infty \). However, a martingale approximation is also proved there. The full conclusion of the proposition has been proved in Theorem 2.8 of [8] under the condition \( \sum_{n \geq 0} \|\mathbb{E}_0(X \circ \theta^n) - \mathbb{E}_{-1}(X \circ \theta^n)\|_{p,\mathcal{X}} < \infty \).

Proposition 4.5 (MZ-SLLN). Let \( 1 \leq r < p < 2 \). Let \( \mathcal{X} = L^r(S, \mathcal{S}, \mu) \) with \( \mu \) \( \sigma \)-finite. Let \( X \in L^p(\Omega, F_0, \mathbb{P}, \mathcal{X}) \) be such that \( \|X\|_{ MW_p} < \infty \), where

\[
(21) \quad \|X\|_{ MW_p} := \sum_{n \geq 0} \left( \int_{S} \|\mathbb{E}_0(S_{2n})(s)\|_{p,\mu}^r ds \right)^{1/r}.
\]

Then, (19) holds with \( \|X\|_{ MW_p} \) in place of \( \|X\|_{ MW_p} \) and (20) holds as well.

Proposition 4.6. Let \( \mathcal{X} \) be either a \((2,L)\) smooth Banach space or \( \mathcal{X} = L^r(S, \mathcal{S}, \mu) \), with \( 1 \leq r \leq 2 \). Let \( X \in L^2(\Omega, F_0, \mathbb{P}, \mathcal{X}) \) be such that \( \|X\|_{ MW_2} < \infty \). For every \( 1 < p < 2 \), there exists a constant \( C_p > 0 \), such that

\[
(22) \quad \|M_2(X)\|_{p,\infty,\mathcal{X}} \leq C_p \|X\|_{ MW_2}.
\]

Remarks. The constant \( C_p \) depends on \( p \) and \( L \) if \( \mathcal{X} \) is \((2,L)\) smooth and on \( p \) and \( r \) if \( \mathcal{X} = L^r(S, \mathcal{S}, \mu) \). Define \( \|X\|_{ H_2} := \sum_{n \geq 0} \|\mathbb{E}_0(X \circ \theta^n) - \mathbb{E}_{-1}(X \circ \theta^n)\|_{2,\mathcal{X}} < \infty \). Then, if \( \|X\|_{ H_2} < \infty \), (22) holds with \( \|X\|_{ H_2} \) in place of \( \|X\|_{ MW_2} \). This follows from Theorem 2.10 of [8] when \( \mathcal{X} \) is \( 2 \)-smooth. The proof when \( \mathcal{X} = L^r(S, \mathcal{S}, \mu) \) may be done as the proof of Theorem 2.10 of [8], using (13).

5. WIP and ASIP under projective conditions

We first obtain martingale approximation results in Banach spaces of cotype 2.

Proposition 5.1. Let \( \mathcal{X} \) be a Banach space of cotype 2. Let \( X \in \mathcal{G}(\mathcal{X}, F_0) \) be such that \( \|X\|_{ MW_2} < \infty \). Then there exists \( d \in \mathcal{G}(\mathcal{X}, F_0) \) with \( \mathbb{E}_{-1}(d) = 0 \) such that

\[
(23) \quad \| \max_{1 \leq k \leq n} |S_k(X) - S_k(d)|X\|_2 = o(\sqrt{n}).
\]

In particular, \((X \circ \theta^n)_{n \geq 0}\) satisfies the WIP of covariance operator \( K_d \) and \( K_d(x^*, y^*) = \lim_n \text{cov}(S_n(x^*(X), S_n(y^*(X)))/n \) for every \( x^*, y^* \in \mathcal{X}^* \).
Remark. The martingale approximation (23) has been proved in [11], see Remark 2.4, in the case where \( \mathcal{X} \) is a Hilbert space (with an explicit expression for \( d \)). When \( \mathcal{X} = \mathbb{R} \) the martingale approximation (23) is due to Gordin and Peligrad [25] and the WIP to Peligrad and Utev [38].

Theorem 5.2. Let \( \mathcal{X} \) be either a Hilbert space or \( \mathcal{X} = L^r(S, \mathcal{S}, \mu) \), with \( 1 \leq r \leq 2 \) and \( \mu \) \( \sigma \)-finite. Let \( X \in \mathbb{G}(\mathcal{X}, \mathcal{F}_0) \) be such that \( \|X\|_{\text{MW}_2} < \infty \). Then there exists \( d \in \mathbb{G}(\mathcal{X}, \mathcal{F}_0) \) with \( \mathbb{E}_{-1}(d) = 0 \) such that

\[
|S_n(X) - S_n(d)|_{\mathcal{X}} = o(\sqrt{nL(L(n))}) \quad \mathbb{P}\text{-a.s.}
\]

In particular, \((X \circ \theta^n)_{n \geq 0}\) satisfies the ASIP of covariance operator \( \mathcal{K}_d \) and \( \mathcal{K}_d(x^*, y^*) = \lim_n \text{cov}(S_n(x^*(X)), S_n(y^*(X)))/n \) for every \( x^*, y^* \in \mathcal{X}^* \). Moreover,

\[
\limsup_n \frac{|S_n|_{\mathcal{X}}}{\sqrt{2nL(L(n))}} = \sup_{x^* \in \mathcal{X}^*, |x^*|_{\mathcal{X}^*} \leq 1} \|x^*(d)\|_2 \leq 10\sqrt{2}\|X\|_{\text{MW}_2} \quad \mathbb{P}\text{-a.s.}
\]

Remark. This result is new even when \( \mathcal{X} = \mathbb{R} \). In view of the previous proposition, one can wonder whether the theorem holds true for Banach spaces of cotype 2.

Theorem 5.3. Let \( \mathcal{X} \) be 2-smooth Banach space. Let \( X \in \mathbb{G}(\mathcal{X}, \mathcal{F}_0) \) be such that \( \|X\|_{\text{MW}_2} < \infty \). Then, \((X \circ \theta^n)_{n \geq 0}\) satisfies the WIP and the ASP of covariance operator \( \mathcal{K} \) given by \( \mathcal{K}(x^*, y^*) = \lim_n \text{cov}(S_n(x^*(X)), S_n(y^*(X)))/n \), for every \( x^*, y^* \in \mathcal{X}^* \). Moreover,

\[
\limsup_n \frac{|S_n|_{\mathcal{X}}}{\sqrt{2nL(L(n))}} = \sup_{x^* \in \mathcal{X}^*, |x^*|_{\mathcal{X}^*} \leq 1} (\mathcal{K}(x^*, x^*))^{1/2} \leq 10\sqrt{2}\|X\|_{\text{MW}_2} \quad \mathbb{P}\text{-a.s.}
\]

Remark. Let \( \mathcal{X} \) be either as in Theorem 5.2 or as in Theorem 5.3. Assume that \( \|X\|_{\text{H}_2} := \sum_{n \geq 0} \|E_0(X \circ \theta^n) - E_{-1}(X \circ \theta^n)\|_{\mathbb{G}(\mathcal{X})} < \infty \). Then, \((X \circ \theta^n)_{n \geq 0}\) satisfies the WIP and the ASP of covariance operator \( \mathcal{K} \) given by \( \mathcal{K}(x^*, y^*) = \lim_n \text{cov}(S_n(x^*(X)), S_n(y^*(X)))/n \), for every \( x^*, y^* \in \mathcal{X}^* \). Moreover, (25) holds with \( \|X\|_{\text{H}_2} \) in the right-hand side instead of \( 10\sqrt{2}\|X\|_{\text{MW}_2} \). This is proved in Theorem 2.10 (see also Corollary 2.12) of [8] when \( \mathcal{X} \) is 2-smooth and may be proved similarly when \( \mathcal{X} = L^r(S, \mathcal{S}, \mu) \) using the remark after Proposition 4.6.

6. Extensions

The general situation considered previously includes the situation where \( X_0 = f(W_0) \), with \((W_n)_{n \in \mathbb{Z}}\) a stationary Markov chain (with general state space). Then \( \theta \) is nothing but the associated shift. It is not hard to translate all the results in terms of the associated Markov operator, see e.g. section 2.3 of [8] for more details.

Now, we shall explain how to extend our results to non-adapted stationary processes or stationary processes arising from non-invertible dynamical systems.

6.1. Non-adapted stationary processes. There are situations where the observable \( X \) is not measurable with respect to the \( \sigma \)-algebra \( \mathcal{F}_0 \), in which case the above results do not apply. However, it is well-known (see Volný [47]) that a suitable version of the Maxwell-Woodroofe condition guarantees the WIP in the real-valued non-adapted case. More generally, methods of approximation by a martingale may be extended somehow easily from the adapted to the non-adapted setting.

All the results of sections 4 and 5 extend to the non-adapted case under the ”natural” strenghtening of the conditions imposed. In this context, we should introduce the following
conditions.

\[ \|X\|_{MW_p} := \sum_{n \geq 0} \left( \|\mathbb{E}(S_{2^n}(X))\|_{p,\mathcal{F}} + \|S_{2^n}(X) - \mathbb{E}_{2^n-1}(S_{2^n}(X))\|_{p,\mathcal{F}} \right) < \infty, \text{ for } 1 < p < 2, \]

\[ \|X\|_{MW_2} := \sum_{n \geq 0} \left( \|\mathbb{E}(S_{2^n}(X))\|_{G(\mathcal{F})} + \|S_{2^n}(X) - \mathbb{E}_{2^n-1}(S_{2^n}(X))\|_{G(\mathcal{F})} \right) < \infty. \]

Notice, that when \( X \) is \( \mathcal{F}_0 \)-measurable the above definitions of \( \| \cdot \|_{MW_p} \) coincide with the previous ones, hence it is legitimate to use the same notation.

We start with a simple observation. Let \( X \in L^1(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{X}) \), then

\[ X = \mathbb{E}_0(X) + X - \mathbb{E}_0(X), \]

where \( \mathbb{E}_0(X) \) is \( \mathcal{F}_0 \) measurable and \( \mathbb{E}_0(X - \mathbb{E}_0(X)) = 0 \).

We can apply the results of section 4 to \( X \), hence, it remains to prove maximal inequalities for \( X - \mathbb{E}_0(X) \). The main new ingredient in section 4 was Proposition 4.1. We shall give its non-adapted analogue.

**Proposition 6.1.** Let \( X \in L^1(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{X}) \) be such that \( \mathbb{E}_0(X) = 0 \). For every \( k \geq 0 \), write \( v_k := |S_{2k} - \mathbb{E}_{2k+1-1}(S_{2k})|_{\mathcal{X}} \) and \( e_k := \mathbb{E}_{2k+1-1}(S_{2k+1}) - (\mathbb{E}_{2k+1-1}(S_{2k}) + \mathbb{E}_{2k+1-1}(S_{2k}) \circ \theta^{2k}) \).

For every integer \( d \geq 0 \), we have (with the convention \( \sum_{k=0}^{\infty} = 0 \))

\[ \max_{1 \leq i \leq 2^d} |S_i|_{\mathcal{X}} \leq \max_{1 \leq i \leq 2^d} \left( \sum_{\ell=0}^{i-1} E(X) \circ \theta^{\ell} \right)_{\mathcal{X}} + \sum_{k=0}^{d-1} \max_{1 \leq i \leq 2^{d-k}-1} \left( \sum_{\ell=0}^{i-1} e_k \circ \theta^{2k+1+\ell} \right)_{\mathcal{X}} + v_d + \sum_{k=0}^{d-1} \max_{0 \leq \ell \leq 2^{d-k}-1} v_k \circ \theta^{2k+1+\ell}. \]

Then, it is not hard to see that all the results of section 4 (with the ad hoc changes) may be proved as in the adapted case, using (27) and (28).

We shall now explain how to prove the non-adapted versions of the results of section 5.

For every \( X \in L^1(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{X}) \) with \( \mathbb{E}_0(X) = 0 \), define \( RX = X \circ \theta^{-1} - \mathbb{E}_0(X \circ \theta^{-1}) \). Then, for every \( n \geq 1 \), \( R^n X = X \circ \theta^{-n} - \mathbb{E}_0(X \circ \theta^{-n}) \). Hence the operator \( R \) is power-bounded on every \( L^p(\Omega, \mathcal{X}) \), that is, there exists \( C > 0 \) such that, for every \( X \in L^p(\Omega, \mathcal{X}) \), and every \( n \geq 1 \), \( \|R^n X\|_{p,\mathcal{X}} \leq C \|X\|_{p,\mathcal{X}} \). Moreover, by Lemma 2.7, \( R \) is also power-bounded on \( G(\mathcal{X}) \).

Finally, let us notice that, for every \( n \geq 1 \), \( S_n(X) - \mathbb{E}_{n-1}(S_n(X)) = (X + \ldots + R^{n-1}X) \circ \theta^{n-1} \).

Then, the non-adapted versions of Proposition 5.1, Theorem 5.2 and Theorem 5.3 may be proved as in the adapted case, making use of the operator \( R \) instead of \( Q \). We leave the task of checking the details to the reader.

For the sake of applications, let us notice that

\[ \|X\|_{MW_p} \leq C_p \sum_{n \geq 1} \left( \|\mathbb{E}(O \circ \theta^n)\|_{p,\mathcal{F}} + \|X \circ \theta^n - \mathbb{E}_n(X \circ \theta^n)\|_{p,\mathcal{F}} \right), \]

\[ \|X\|_{MW_2} \leq C \sum_{n \geq 1} \left( \|\mathbb{E}(O \circ \theta^n)\|_{G(\mathcal{F})} + \|X \circ \theta^n - \mathbb{E}_n(X \circ \theta^n)\|_{G(\mathcal{F})} \right). \]

Dedecker, Merlevède and Pène [17] obtained the WIP when \( \mathcal{X} = L'(S, S, \mu) \), \( r \geq 2 \) under the condition

\[ \sum_{n \geq 0} \|\mathbb{E}(O \circ \theta^n) - \mathbb{E}_{-1}(X \circ \theta^n)\|_{p,\mathcal{X}} < \infty. \]
Generally speaking, their condition is independent to the condition \( \|X\|_{MW_2} < \infty \). However, in their applications they need a condition in terms of \( \mathbb{E}_0(X \circ \theta^n) \) and \( X \circ \theta^n - \mathbb{E}_n(X \circ \theta^n) \), namely they need (with \( \mathcal{X} = L^r(\mu) \))

\[
\sum_{n \geq 1} \frac{\|\mathbb{E}_0(X \circ \theta^n)\|_{r,\mathcal{X}} + \|X \circ \theta^n - \mathbb{E}_n(X \circ \theta^n)\|_{r,\mathcal{X}}}{n^{1/r}},
\]

which is stronger than the sufficient condition that we obtain from (22), when \( r > 2 \).

### 6.2. Results for non-invertible dynamical systems.

Let \( (\mathcal{Y}, \Sigma, \nu) \) be a probability space and \( \tau \) be a measurable (non-invertible) measure preserving transformation on \( \mathcal{Y} \). Let us write \( \mathcal{G}_n = \theta^{-n}(\Sigma) \), for every \( n \geq 0 \).

In this case there exists a (unique) Markov operator \( K \), known as the Perron-Frobenius operator, defined by

\[
\int_{\mathcal{Y}} f(g \circ \tau) \, d\nu = \int_{\mathcal{Y}} (Kf) \, g \, d\nu \quad \forall f, g \in L^2(\mathcal{Y}, \Sigma, \nu).
\]

Then, we have for every \( f \in L^1(\mathcal{Y}, \Sigma, \nu) \),

\[
\mathbb{E}(f|\mathcal{G}_n) = (K^n f) \circ \tau^n.
\]

The operator \( K \) may be extended to a contraction of every \( L^p(\mathcal{Y}, \Sigma, \nu, \mathcal{X}) \), for every real separable Banach space \( \mathcal{X} \).

The main difference between this situation and the one considered in sections 4 and 5 is that, here we cannot in general approximate \( f + \ldots + f \circ \tau^{n-1} \) thanks to stationary martingale differences but, by mean of stationary differences of reverse martingales.

As mentionned there, all the results of section 3 except the ASIP (for \( \mathcal{X} \neq \mathbb{R} \)), hold for stationary differences of reverse martingales. In particular, we will have versions of the results of section 4 provided that we can state (and prove) an analogue to Proposition 4.1, what we do right now.

**Proposition 6.2.** Let \( f \in L^1(\mathcal{Y}, \Sigma, \nu, \mathcal{X}) \). For every \( k \geq 0 \), write \( S_k = f + \ldots + f \circ \tau^{k-1} \),

\[
w_k := \|\mathbb{E}(S_{2^k}|\mathcal{G}_{2^k+1-1})\|_{\mathcal{X}} \quad \text{and} \quad f_k := \mathbb{E}(S_{2^k}|\mathcal{G}_{2^k+1-1}) + \mathbb{E}(S_{2^k}|\mathcal{G}_{2^k+1-1}) \circ \tau^{2^k} - \mathbb{E}(S_{2^k+1}|\mathcal{G}_{2^k+2-1}).
\]

With the above notations, we have, for every integer \( d \geq 0 \):

\[
\max_{1 \leq i \leq 2^d} \|S_i\|_{\mathcal{X}} \leq w_d + \sum_{k=0}^{d-1} \max_{0 \leq i \leq 2^{d-1-k}-1} w_k \circ \tau^{2^{k+1}+1} + \max_{1 \leq i \leq 2^d} \sum_{0 \leq \ell \leq i-1} \|f - \mathbb{E}(f|\mathcal{G}_{1})\|_{\mathcal{X}} \circ \tau^\ell
\]

\[
+ \sum_{k=0}^{d-1} \max_{0 \leq i \leq 2^{d-k-1}-1} \sum_{\ell=0}^{i} f_k \circ \tau^{2^{k+1}+1} \|\theta_i\|_{\mathcal{X}}.
\]

We leave the proof to the reader. Let us notice that \( (f_k \circ \tau^{2^{k+1}+1})_{\ell \geq 0} \) is a stationary sequence of differences of reverse martingales and that \( \mathbb{E}(S_{2^k}|\mathcal{G}_{2^k+1-1}) = \left( \sum_{\ell=2^k}^{2^{k+1}-1} K^\ell f \right) \circ \tau^{2^{k+1}-1} \).

Let us denote \( \|\cdot\|_{p,\nu,\mathcal{X}} \) the norm in \( L^p(\mathcal{Y}, \nu, \mathcal{X}) \). Let us introduce the ”natural” conditions that should be used in this new setting.

\[
\|f\|_{MW_2^p} := \sum_{n \geq 0} \frac{\|f + \ldots + K^{2^n-1}f\|_{p,\nu,\mathcal{X}}}{2^{n/p}} < \infty \quad \text{if } 1 < p < 2,
\]

\[
\|f\|_{MW_2^2} := \sum_{n \geq 0} \frac{\|f + \ldots + K^{2^n-1}f\|_{\mathcal{G}(\mathcal{X},\nu)}}{2^{n/2}} < \infty.
\]

Then, all the results of section 4 hold with the ad hoc changes.

Concerning the results of section 5, we can say the following.
The conclusion of Proposition 5.1 holds under (32), except that $d$ should be a difference of reverse martingales, which does not bring any trouble for the WIP.

Assume that (32) holds.

If $\mathcal{X}$ is a Hilbert space or if $\mathcal{X} = L^s(S,S,\mu)$, $1 \leq r \leq 2$, then (24), with $d$ a reverse martingale difference, and the CLIL hold.

If $\mathcal{X}$ is a 2-smooth Banach space, then the WIP and CLIL hold. The ASIP holds as well, in the particular case where $\mathcal{X} = \mathbb{R}$.

7. **Optimality of the Conditions**

We prove that our sufficient conditions for the MZ-SLLN and for the CLIL are somehow optimal (in terms of conditions relying on $\mathbb{E}_0(S_n)$), in the real valued case.

The optimality of the condition $\|X\|_{MW_2} < \infty$ for the CLT has been obtained by Peligrad and Utev [38]. Actually, we shall use their example in the sequel.

**Proposition 7.1.** Let $(a_n)_{n \geq 0}$ be a sequence of positive numbers with $a_n \to 0$ as $n \to \infty$. Let $1 < p \leq 2$. There exist a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with a transformation $\theta$ and a filtration $(\mathcal{F}_n)_{n \in \mathbb{Z}}$, as in the Introduction, such that there exists $X \in L^p(\Omega, \mathcal{F}_0, \mathbb{P})$ for which

\[
\sum_{n \geq 1} a_n \frac{\|\mathbb{E}_0(S_n(\theta^n))\|_p}{n^{1+1/p}} < \infty,
\]

but

\[
\limsup_n \frac{|S_n(\theta^n)|}{b_n} = +\infty \quad \mathbb{P}\text{-a.s.},
\]

where $b_n = n^{1/p}$ if $1 < p < 2$ and $b_n = \sqrt{nL(L(n))}$ if $p = 2$.

**Remark.** It would be interesting to know whether the condition $\sum_{n \geq 1} \frac{\|\mathbb{E}_0(X\circ\theta^n)\|_p}{n^{1/p}} < \infty$ is also optimal.

**Proof.** We consider the Markov chain $(W_n)_{n \geq 0}$ with state space $\mathbb{N} := \{0, 1, \ldots\}$ and transition probability given by $p_{i+1} = 1$ and $p_{i+1} = 1$ for every $i \geq 1$, and $p_{i,j} = 0$ otherwise. The stationarity is guaranteed by the condition $\mathbb{E}(\tau) < \infty$ and then, the stationary distribution $\pi := (\pi_i)_{i \geq 0}$ is given by $\pi_0 = 1/\mathbb{E}(\tau)$ and $\pi_i = \pi_0 \sum_{j=i+1}^{\infty} p_{j,i}$.

Since our Markov chain is stationary, we may consider its two-sided version $(W_n)_{n \in \mathbb{Z}}$, taking for $(\Omega, \mathcal{F}, \mathbb{P})$ the canonical space, for $\theta$ the shift and for $\mathcal{F}_0, \sigma\{W_n : n \leq 0\}$. Then we are exactly in the situation considered in our paper.

Without loss of generality, we shall assume that $(a_n)_{n \geq 0}$ decreases to 0.

Let $(u_n)_{n \geq 1}$ be a sequence of positive integers such that (its existence is clear)

\[
(34) \quad u_1 = 1, \quad u_2 = 2, \quad u_k^2 + 1 < u_{k+1} \quad \forall k \geq 3, \quad a_t \leq k^{-2} \quad \forall t \geq u_k.
\]

Then, for every $n \geq 1$, define $p_{u_n} = c/u_n^2$ and $p_i = 0$ if $i \notin \{u_n, n \geq 1\}$, where $c$ ensures that we have a probability.

Notice that the sequence $(u_n)_{n \geq 1}$ has a super-exponential growth. In particular, $\mathbb{E}(\tau) < \infty$. Moreover $\mathbb{E}(\tau^p) = +\infty$.

As in [38], we shall take $X := 1_{\{W_0=0\}} - \pi_0$. Let $\nu := \min\{m \geq 1 : W_m = 0\}$. The fact that (33) holds may be proved exactly as in [38]. We shall give only some key steps, based on the computations in [38].

Define $\tilde{S}_n = X \circ \theta + \ldots + X \circ \theta^n$. Clearly it suffice $s$ to prove the proposition for $\tilde{S}_n$ in place of $S_n$. It follows from the proof of Proposition 3.1 of [38] that (notice that our notations are different)

\[
\|\mathbb{E}_0(\tilde{S}_n)\|_p \leq \|\nu \land n\|_p + \max_{1 \leq i \leq n} |\mathbb{E}(\tilde{S}_i|W_0 = 0)| := I_n + J_n.
\]
Then, one can prove that for every $u_k < n \leq u_{k+1}$,
\[ I_n^p \leq C(u_k + n^{p+1}/u_{k+1}^p + n^p/u_{k+1}^p), \]
and then, that $\sum_{n \geq 1} u_n a_n b_n^{1/p} < \infty$.

Using the notation of \cite{8} (page 812), one can prove that $\mathbb{E}(M_n) \leq C(n^{1/2p} + n/u_{k+1}^p + u_n \min(1, n/u_{k+1}^p)$, and then that
\[ \sum_{n \geq 1} a_n J_n / n^{3/2} \leq C \sum_{n \geq 1} n^{-1-1/2p} + \sum_{k \geq 1} 1/(k^2 u_{k+1}^{p-2}) + \sum_{k \geq 1} 1/k^2 < \infty, \]
where we used that $p + 1/p > 2$ for $p > 1$.

Let us prove that $\limsup_n |S_n| / b_n = +\infty$ $\mathbb{P}$-a.s.

Let $T_0 := 0$ and, for $k \geq 1$, $T_k := \min\{t > T_{k-1} : W_t = 0\}$. Define then, $\tau_k := T_k - T_{k-1}$. Then, $(\tau_k)_{k \geq 1}$ is iid, distributed like $\tau$ and $S_{T_k} = \sum_{i=1}^k (1 - \pi_0 \tau_i)$.

It is enough to prove that $\limsup_k |S_{T_k}| / b_k = +\infty$ $\mathbb{P}$-a.s.

Since $\mathbb{E}(\tau) < \infty$, by the strong law of large numbers, $T_n/n \longrightarrow E(\tau)$ $\mathbb{P}$-a.s., hence it is enough to prove that $\limsup_k |S_{T_k}| / b_k = +\infty$ $\mathbb{P}$-a.s. In particular, it is enough to prove that
\[ \limsup_k \left( \sum_{i=0}^k (1 - \pi_0 \tau_i) / b_k \right) = +\infty \quad \mathbb{P}\text{-a.s.} \tag{35} \]

Recall that $\mathbb{E}(\tau^p) = +\infty$.

By the Borel-Cantelli Lemma, $\limsup_n \tau_n / n^{1/p} = +\infty$ $\mathbb{P}$-a.s. Hence, if $1 < p < 2$, (35) must hold (recall that $\mathbb{E}(\tau) < \infty$).

When $p = 2$, we conclude thanks to Strassen’s converse to the law of the iterated logarithm, see for instance \cite{30} page 203-204. \hfill \Box

8. Examples

8.1. Some direct examples. We shall give here examples where the conditions $MW_p$ are used rather than their reinforcements. We start with the case of linear processes.

Let $1 < p \leq 2$. Let $d \in L^p(\Omega, \mathcal{F}_0, \mathbb{P}, \mathbb{R})$, with $\mathbb{E}_{-1}(d) = 0$. Let $(a_n)_{n \geq 0} \in L^p(\mathbb{R})$. Define a random variable in $L^p(\Omega, \mathcal{F}_0, \mathbb{P})$ by
\[ X := \sum_{n \geq 0} a_n d \circ \theta^{-n}. \]

Then, $(X \circ \theta^n)_{n \geq 0}$ is a stationary linear process.

**Proposition 8.1.** Let $(X \circ \theta^n)_{n \geq 0}$ be the above linear process. Assume further that
\[ \sum_{n \geq 1} \left( \sum_{k \geq 0} |a_k + \ldots + a_{k+n}|^p \right)^{1/p} / n^{1+1/p} < \infty. \]

Then $\|X\|_{MW_p} < \infty$.

**Proof.** Use Corollary 2.2 and more particularly \cite{8}. \hfill \Box

We now consider the case of $\rho$-mixing processes for which it is known that the Maxwell-Woodroofe condition is well-adapted.

Let $(X_n)_{n \in \mathbb{Z}}$ be a stationary $\mathcal{H}$-valued sequence. Define
\[ \rho(n) = \rho(\mathcal{F}_n^0, \mathcal{F}_n^\infty) \quad \text{and} \quad \psi(n) = \psi(\mathcal{F}_n^0, \mathcal{F}_n^\infty). \tag{36} \]
Sharipov obtained the conclusion of the corollary under the condition $\varepsilon > 0$. When $H(38)$ holds, by (37).

\[\text{Proof.}\]

In a 2-smooth Banach space. However, he assumes weaker moment conditions and the variables are allowed to take values in the min(2, $r$)-smooth Banach space.

Then, $\|X\|_{MW2} < \infty$.

**Remarks** The condition $\rho(2^n) = O(1/n^{1+\varepsilon})$ has been proven to be sufficient in [44] (for any $\varepsilon > 0$), when $H = \mathbb{R}$. The sufficiency of (37) has been obtained very recently by Lin and Zhao [32], when $H = \mathbb{R}$.

Sharipov [45] obtained the conclusion of the corollary under the condition $\sum_n \psi(n) < \infty$. However, he assumes weaker moment conditions and the variables are allowed to take values in a 2-smooth Banach space.

**Proof.** It suffices to prove that

\[\sum_n \frac{\|E_0(S_{2^n}(X_0))\|_{2,H}}{2^{n/2}} < \infty,\]

Let $(e_i)_{i \geq 0}$ be an orthonormal basis of $H$, and write $Y_0^{(i)} := \langle X_0, e_i \rangle_H$. We have

\[\|E_0(S_{2^n}(X_0))\|_{2,H} = \sum_{i \geq 0} E \left[ \left( E_0(S_{2^n}(Y_0^{(i)})) \right)^2 \right].\]

Now, it follows from the computations page 15 of [36] combined with Lemma 3.4 of [?] that

\[E \left[ \left( E_0(S_{2^n}(Y_0^{(i)})) \right)^2 \right] \leq CE((Y_0^{(i)})^2)\left( \sum_{k=0}^{n} 2^{k/2} \rho(2^k) \right)^2.\]

Using that $\sum_{i \geq 0} (Y_0^{(i)})^2 = |X_0|_{H}^2$, we see that (38) is satisfied as soon as

\[\sum_n \frac{1}{2^{n/2}} \sum_{k=0}^{n} 2^{k/2} \rho(2^k) < \infty,\]

which holds, by (37). \qed

8.2. **Applications to the empirical process.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $\theta$ be an invertible bi-measurable measure preserving transformation on $\Omega$ and $\mathcal{F}_0 \subset \mathcal{F}$ a $\sigma$-algebra such that $\mathcal{F}_0 \subset \theta^{-1}(\mathcal{F}_0)$. Define a non-decreasing filtration by $\mathcal{F}_n = \theta^{-n}(\mathcal{F}_0)$, for every $n \in \mathbb{Z}$ and denote $\mathbb{E}_n := \mathbb{E}(\cdot | \mathcal{F}_n)$.

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 $X_n := t \mapsto 1_{Y_n \leq t} - F(t)$, where $F(t) = \mathbb{P}(Y \leq t)$.

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

\[\int_0^\infty (1 - F(t))^r \mu(dt) + \int_{-\infty}^0 F(t)^r \mu(dt) < \infty,\]

which is satisfied whenever $\mu$ is finite.
Define $F_\mu$ by $F_\mu(x) = -\mu([x, 0])$ if $x < 0$ and $F_\mu(x) = \mu([0, x])$ if $x \geq 0$. Let $1 < p \leq 2$. Then, under (39), $X_0 \in L^p(\Omega, L^r(\mu))$ if and only if

$$E(\|F_\mu(Y_0)\|^{p/r}) < \infty.$$  

We want to understand the asymptotic behaviour of the process $F_n = S_n(X)/n$ (with values in $L^p(\Omega, \mathcal{F}_0, \mathbb{P}, L^r(\mathbb{R}, \mu))$, and more particularly of $D_{n,r}(\mu) := \|F_n\|_{r,\mu}$.

Notice that when $\mu$ is the Lebesgue measure $\lambda$ and $r = 1$, $D_{n,1}(\lambda)$ represents the Wasserstein distance between the empirical distribution and the true distribution.

Let us introduce some dependence coefficients. For every $Y \in L^1(\Omega, \mathcal{F}, \mathbb{P})$, every $1 \leq r \leq \infty$ and every $1 \leq p \leq \infty$, define

$$\tilde{\tau}_{\mu,r,p}(\mathcal{F}_0, Y_n) := \left\| \left( \int_{\mathbb{R}} \mathbb{P}(Y_n \leq t | \mathcal{F}_0) - F(t) \right)^{1/r} \mu(dt) \right\|_p \quad \text{if } r \geq 1$$

$$\tilde{\tau}_{\mu,r,p}(\mathcal{F}_0, Y_n) := \left( \int_{\mathbb{R}} \mathbb{P}(Y_n \leq t | \mathcal{F}_0) - F(t) \right)^{1/r}_p \mu(dt) \quad \text{if } 1 \leq r \leq p$$

Notice that we use two ways to define $\tilde{\tau}_{\mu,p,p}(\mathcal{F}_0, Y_n)$ but both definitions coincide by Fubini’s theorem. When $r \geq p$, $\tilde{\tau}_{\mu,r,p}(\mathcal{F}_0, Y_n) = \tau_{\mu,r,p}(\mathcal{F}_0, Y_n)$, where $\tau_{\mu,r,p}(\mathcal{F}_0, Y_n)$ appears for instance in (40) (notice that our notations are slightly different).

Let us notice that both (39) and (40) are satisfied as soon as $\tau_{\mu,r,p}(\mathcal{F}_0, Y_n) < \infty$.

**Theorem 8.3.** Let $Y \in L^1(\Omega, \mathcal{F}_0, \mathbb{P})$ and $(S, S, \mu)$ be a $\sigma$-finite measure space. Let $1 \leq p \leq 2$ and $1 \leq r < \infty$. Assume that

$$\sum_{n \geq 0} \frac{\tilde{\tau}_{\mu,r,p}(\mathcal{F}_0, Y_n)}{n^{1/p}} < \infty.$$  

(i) If $p < 2$ and $r \neq p$, then

$$\limsup_n n^{-1/p} D_{n,r}(\mu) = 0 \quad \mathbb{P} \text{-a.s.}$$

(ii) If $p = 2$, then $(X_n)_{n \geq 1}$ satisfies the WIP and the ASIP. Moreover,

$$\limsup_n \frac{n^{1/2}}{\sqrt{2nL(L(n))}} D_{n,r}(\mu) = \Lambda_\mu \quad \mathbb{P} \text{-a.s.},$$

for some $\Lambda_\mu \geq 0$.

**Remark.** Actually, if $r'$ denotes the conjugate of $r$, we have $\Lambda_{\mu,r} = \sup \Gamma_{\mu,r}(f)$ where $\Gamma_{\mu,r}(f) = \lim_n \|F_n f(S_n(s) \mu(ds))\|_2 / \sqrt{n}$. Since Theorem 8.3 is a straightforward application of the results of section 5 we omit the proof.

Dedecker and Merlevède obtained limit theorems under mixing conditions, including conditions on the coefficients $\tilde{\tau}_{\mu,r,p}$, when $r \geq p$. In (15) they studied the MZ-SLLN (with a Baum-Katz type result), in (14) they studied the WIP and in (16) the ASIP. When $r > p$, $r \neq 2$, their results rely on a condition a la Gordin, hence yield to stronger conditions than ours. When $r = 2$, they use a very different approach and their results have different range of applicabilities. Their approach does not seem to apply when $1 < r < p$. In this case (1 $\leq r < p$), the only other result that we are aware of is the CLT of Dédé [12] that we already mentioned.

In order to apply Theorem 8.3 we shall further study the coefficients $\tilde{\tau}$, and estimate them thanks to other coefficients that are known to be computable in many situations (see e.g. Dedecker and Prieur [19]).
Let us define the coefficients \( \hat{\phi} \) and \( \tilde{\alpha} \), as defined in Dedeker and Prieur [19]. For every \( n \geq 1 \), define
\[
\hat{\phi}(n) := \sup_{t \in \mathbb{R}} \| E_0(\mathbf{1}_{\{Y_n \leq t\}}) - F(t) \|_\infty
\]
\[
\tilde{\alpha}(n) := \sup_{t \in \mathbb{R}} \| E_0(\mathbf{1}_{\{Y_n \leq t\}}) - F(t) \|_1 .
\]
Notice that clearly, \( \tau_{\mu,\infty,\infty}(\mathcal{F}_0, Y_n) \leq \hat{\phi}(n) \) and \( \tau_{\mu,\infty,1}(\mathcal{F}_0, Y_n) \leq \tilde{\alpha}(n) \). It is not hard to prove (using Jensen’s inequality and Fubini) that if \( q = \max(p, r) \), then
\[
\tau_{\mu,r,p}(\mathcal{F}_0, Y_n) \leq \tau_{\mu,q,q}(\mathcal{F}_0, Y_n) \leq (\tau_{\mu,1,1}(\mathcal{F}_0, Y_n))^{1/q} .
\]

**Lemma 8.4.** Assume that \( \mu \) is finite. Let \( 1 \leq p, r \leq 1 \) and define \( q := \max(p, r) \). For every \( n \geq 1 \), we have
\[
\tau_{\mu,r,p}(\mathcal{F}_0, Y_n) \leq \mu(\mathbb{R})^{1/r} \hat{\phi}(n) ,
\]
\[
\tau_{\mu,r,p}(\mathcal{F}_0, Y_n) \leq \mu(\mathbb{R})^{1/r} \tilde{\alpha}(n)^{1/q} .
\]

**Proof.** The first inequality is obvious. The second one follows easily from (41). \( \square \)

We shall focus now on the case \( 1 \leq r \leq p \).

**Lemma 8.5.** Let \( 1 \leq r \leq p \). For every \( n \geq 1 \), we have
\[
\tau_{\lambda,r,p}(\mathcal{F}_0, Y_n) \leq 2^{1/p} \left( \int_{0}^{\infty} (\mathbb{P}(|Y| \geq t))^{r/p} \mu(dt) \right)^{1/r} \hat{\phi}(n)^{(p-1)/p} .
\]
\[
\tau_{\lambda,r,p}(\mathcal{F}_0, Y_n) \leq 2^{1/p} \left( \int_{0}^{\infty} \min \left[ \tilde{\alpha}_n, (\mathbb{P}(|Y| \geq t)) \right] \right)^{r/p} \mu(dt) \right)^{1/r} .
\]

**Proof.** Notice that, for every \( t \in \mathbb{R} \),
\[
\| \mathbb{P}(Y \leq t| \mathcal{F}_0) - F(t) \|_p^n \leq 2^{\hat{\phi}(n)^{p-1}} (1 - F(t)) F(t) \leq 2^{\hat{\phi}(n)^{p-1}} \mathbb{P}(|Y| \geq t) .
\]
Hence, (42) follows.

Using that for every \( t \in \mathbb{R} \),
\[
\| \mathbb{P}(Y \leq t| \mathcal{F}_0) - F(t) \|_p^n \leq \| \mathbb{P}(Y \leq t| \mathcal{F}_0) - F(t) \|_{1} \leq \tilde{\alpha}(n) ,
\]
and
\[
\| \mathbb{P}(Y \leq t| \mathcal{F}_0) - F(t) \|_p^n \leq 2F(t)(1 - F(t)) \leq 2\mathbb{P}(|Y| \geq t) ,
\]
we see that
\[
\hat{\tau}_{\mu,r,p}(\mathcal{F}_0, Y_n) \leq 2^{1/p} \left( \int_{0}^{\infty} \min \left[ \tilde{\alpha}_n, (\mathbb{P}(|Y| \geq t)) \right] \right)^{r/p} \mu(dt) \right)^{1/r} .
\]

**Theorem 8.6.** Let \( 1 \leq r \leq p \), with \( r \neq p \) if \( p < 2 \). Assume either of the following items.

(i) \( \int_{0}^{\infty} (\mathbb{P}(|Y| > t))^{r/p} \mu(dt) < \infty \) and \( \sum_{n \geq 1} \hat{\phi}(n)^{(p-1)/p} = \infty \).

(ii) Assume that \( \mu \) is the Lebesgue measure and that \( \sum_{n \geq 1} \left( \int_{0}^{\hat{\phi}(n)^{p-1}} (1 - x/Q(x)) Q(x) dx \right)^{1/r} < \infty \),

where \( Q(x) := \inf \{ t \geq 0 : \mathbb{P}(|Y| > t) > x \} \).

Then, the conclusion of Theorem 8.3 holds.

8.3. **Proof of Theorem 8.6** The conclusion under (i) follows from Theorem 8.3 and Lemma 8.5. To prove item (ii), in view of Theorem 8.3 and Lemma 8.5, it suffices to prove that (notice that \( \mathbb{P}(|Y| \geq t) = \mathbb{P}(|Y| > t) \) for \( \lambda \)-a.e. \( t \in \mathbb{R} \))
\[
\sum_{n \geq 1} \frac{1}{n^{1/p}} \left( \int_{0}^{\infty} \min \left[ \tilde{\alpha}_n, (\mathbb{P}(|Y| > t)) \right] \right)^{r/p} \mu(dt) \right)^{1/r} < \infty .
\]

Now, \( \int_{0}^{\infty} \left( \min \left[ \tilde{\alpha}_n, (\mathbb{P}(|Y| > t)) \right] \right) dt \leq \hat{\alpha}(n)^{r/p} Q(\hat{\alpha}(n)) + \int_{Q(\hat{\alpha}(n))}^{\infty} (\mathbb{P}(|Y| > t))^{r/p-1} dt \)
Since $Q$ is non-increasing, we see that $(ii)$ implies that
\[ \sum_{n \geq 1} \frac{\hat{f}(n)^{1/p} (\hat{Q}(\hat{f}(n)))^{1/r}}{n^{1/p}} < \infty, \]
hence, it remains to deal with the second term in the right-hand side of (11).

We have
\[ \int_{Q(\hat{f}(n))}^{+\infty} (\mathbb{P}(Y > t))^{r/p-1} dt = \int_{Q(\hat{f}(n))}^{+\infty} \left( \int_{0}^{1} \frac{r}{p} x^{r/p-1} 1_{x \leq \mathbb{P}(Y > t)} dx \right) dt \]
\[ \leq \int_{0}^{\hat{f}(n)} \frac{r}{p} x^{r/p-1} \left( \int_{0}^{Q(x)} dt \right) dx = \int_{0}^{\hat{f}(n)} \frac{r}{p} x^{r/p-1} Q(x) dx, \]
and the proof is complete. \hfill \Box

Appendix A. Proof of the results of section 2

A.1. Proof of Lemma 2.1. Let $X, Y \in L^p(\Omega, \mathcal{X})$. According to definition 2.1, we have to prove that
\[ \left( \mathbb{E}( |X + Y|^p ) \right)^{r/p} + \left( \mathbb{E}( |X - Y|^p ) \right)^{r/p} \leq 2 \left( \mathbb{E}( |X|^p ) \right)^{r/p} + 2 \tilde{L}^r \left( \mathbb{E}( |Y|^p ) \right)^{r/p}, \]
with $\tilde{L}^r = L^r + \max(1, \sqrt{p-1})$. Define $x := (|X + Y| + |X - Y|)/2$ and $y := (|X + Y| - |X - Y|)/2$. We have, using that $L^p(\Omega, \mathbb{R})$ is $(r, \max(1, \sqrt{p-1}))$-smooth,
\[ \left( \mathbb{E}( |X + Y|^p ) \right)^{r/p} + \left( \mathbb{E}( |X - Y|^p ) \right)^{r/p} = \|x + y\|^p_r + \|x - y\|^p_r \]
\[ \leq 2 \left( \|x\|^p_{p,r} + \max(1, \sqrt{p-1}) \right) \|y\|^p_{p,r}. \]
Now, since $\mathcal{X}$ is $(r, L)$-smooth, $x \leq \left( (|X + Y| + |X - Y|)/2 \right)^{1/r} \leq \left( |X|_\mathcal{X} + L^r |Y|_\mathcal{X} \right)^{1/r} \leq \left( |X|_\mathcal{X} + L |Y|_\mathcal{X} \right)^{1/r}$ and
\[ \|x\|^p_r \leq \|X|_\mathcal{X} + L^r |Y|_\mathcal{X} \leq \|X\|^p_{p, \mathcal{X}} + L^r \|Y\|^p_{p, \mathcal{X}}. \]
To conclude, we combine (45) and (46) with the fact that $|y| \leq |Y|_\mathcal{X}$. \hfill \Box

A.2. Proof of Corollary 2.2. By Lemma 2.1, we have
\[ \|X + Y\|^p_{p, \mathcal{X}} + \|X - Y\|^p_{p, \mathcal{X}} \leq 2 \|X\|^p_{p, \mathcal{X}} + 2 \left( \max(1, \sqrt{p-1}) \right)^r \|Y\|^p_{p, \mathcal{X}}. \]
Using the conditional Jensen inequality and that $|X|_\mathcal{X} = |E(X - Y)|_\mathcal{X} \leq |E(X - Y)|_\mathcal{X}$, we obtain that $\|X\|^p_{p, \mathcal{X}} \leq \|X - Y\|^p_{p, \mathcal{X}}$ and the result follows. \hfill \Box

A.3. Proof of Lemma 2.6. Let $X, Y \in G(\mathcal{X})$. Consider the Banach space $C := \mathcal{X} \times \mathcal{X}$ with norm $\| (x, y) \|_C := (|x|_\mathcal{X}^2 + |y|_\mathcal{X}^2)^{1/2}$. Let us prove that $(X, Y) \in G(C)$. Let $G(X)$ and $G(Y)$ be independent gaussian variables with same covariance operator as $X$ and $Y$ respectively. Then, $(G(X), G(Y))$ is a gaussian variable taking values in $C$. Now, for every $x^*, y^* \in \mathcal{X}^*$, we have
\[ \mathbb{E}( (x^*(X) + y^*(Y))^2 ) \leq 2 \mathbb{E} \left[ (x^*(G(X))^2 + (y^*(G(Y))^2 \right] = 2 \mathbb{E}((x^*(G(X)) + y^*(G(Y)))^2). \]
Hence, by Lemma 2.3, $(X, Y) \in G(C)$. Let $(U, V)$ be a gaussian variable with values in $C$ with same covariance operator as $(X, Y)$. Clearly, $U + V$ is gaussian and has same covariance operator as $X + Y$. Hence, $X + Y$ is pregaussian and we may take $G(X + Y) = U + V$. Similarly, we may take $G(X) = U$ and $G(Y) = V$. Now,
\[ \|G(X + Y)\|_{2, \mathcal{X}} = \|U + V\|_{2, \mathcal{X}} \leq \|U\|_{2, \mathcal{X}} + \|V\|_{2, \mathcal{X}} = \|X\|_{2, \mathcal{X}} + \|Y\|_{2, \mathcal{X}}. \]
Hence, $\| \cdot \|_{G(\mathcal{X})}$ is a norm on $G(\mathcal{X})$.

Let us prove that $(G(\mathcal{X})$ is a Banach space.
Let \((X_n)_{n \geq 1}\) be Cauchy in \((\mathcal{G}(X)), \| \cdot \|_{\mathcal{G}(X)}\). Hence, \((X_n)_{n \geq 1}\) is Cauchy in \(L^2(\Omega, \mathcal{X})\), so it converges, say to \(X\) in \(L^2(\Omega, \mathcal{X})\). We just have to prove that \(X\) is pregaussian and that \((X_n)_{n \geq 1}\) admits a subsequence converging to \(X\) for \(\| \cdot \|_{\mathcal{G}(X)}\). By assumption, there exists a subsequence \((X_{n_k})_{k \geq 1}\) such that \(\|X_{n_k} - X_{n_k+1}\|_{\mathcal{G}(X)} \leq 2^{-k}\). Then \(X = -X_{n_1} + \sum_{k \geq 1} X_{n_k} - X_{n_{k+1}}\) with convergence in \(L^2(\Omega, \mathcal{X})\).

Extending our probability space, if necessary, we may assume that there exists a sequence \((G_k)_{k \geq 0}\) of independent Gaussian variables taking values in \(\mathcal{X}\), such that \(G_0 = G(X_{n_1})\) and for every \(k \geq 1\), \(G_k = G(X_{n_k+1} - X_{n_k})\). Then, \(G := \sum_{k \geq 0} 2^{k/2} G_k\) defines a Gaussian variable. Moreover, for every \(x^* \in \mathcal{X}^*\), we have, using Cauchy-Schwarz,

\[
\mathbb{E}(x^*(X))^2 = \mathbb{E}\left( (x^*(-X_{n_1}) + \sum_{k \geq 1} x^*(X_{n_k+1} - X_{n_k}))^2 \right) \\
\leq 2\left( \mathbb{E}(x^*(-X_{n_1}))^2 + \sum_{k \geq 1} 2^k \mathbb{E}(x^*(X_{n_k+1} - X_{n_k}))^2 \right) = 2\mathbb{E}\left( (x^*(\sum_{k \geq 0} 2^{k/2} G_k))^2 \right).
\]

It follows from Lemma 2.3 that \(X\) is pregaussian. By a similar argument, using the second half of Lemma 2.3, we see that \(\mathbb{E}(\|X - X_{n_m}\|^2) \to 0\) as \(m \to +\infty\), and the proof is finished. \(\square\)

A.4. Proof of Lemma 2.8. Let \(x^* \in \mathcal{X}^*\). Clearly, by Lemma 2.7, we may assume that \(X\) is \(\mathcal{H}_\infty\)-measurable. Denote \(X_n := \mathbb{E}(X|\mathcal{H}_n)\). Then \((X_n)_{n \geq 1}\) is a martingale converging in \(L^2(\Omega, \mathcal{X})\) to \(X\) (see for instance Proposition V.2.6. of Neveu [37]). It suffices to prove that \(\|G(X - X_n)\|_{2,\mathcal{X}}\) converges to 0. Using Lemma 2.4, see also the remark thereafter, we have

\[
\mathbb{E}(x^*(X - X))^2 \leq 2(\mathbb{E}(x^*(X))^2 + \mathbb{E}(x^*(X_n))^2) \leq 6\mathbb{E}(x^*(G(X))^2).
\]

Since \(X_n - X\) is (clearly) pregaussian, we infer that

\[
\mathbb{E}(x^*(G(X_n - X)))^2 \leq \mathbb{E}(x^*(G(X))^2).
\]

Then, it follows from the discussion pages 73-74 of [30], that \(G(X_n - X)\) is tight, hence converges in probability to 0, since for every \(x^* \in \mathcal{X}^*\), \((x^*(G(X_n - X)))_{n \geq 1}\) converges in probability to 0 (recall that \(\|x^*(G(X_n - X))\|_2 = \|x^*(X_n - X)\|_2 \to 0\) as \(n \to \infty\).

Let \(\varepsilon > 0\). There exists \(n_\varepsilon \geq 1\) such that \(\mathbb{P}(\|G(X_{n_\varepsilon} - X)|_{\mathcal{X}} > \varepsilon) < 1/2\). In particular, the median of the Gaussian variable \(G(\bar{X}_{n_\varepsilon} - X)\) is smaller than \(\varepsilon\), and it follows from the last assertion of Lemma 3.2 of [30], that there exists a universal \(C > 0\) such that \(\|G(\bar{X}_{n_\varepsilon} - X)\|_{2,\mathcal{X}} \leq C\varepsilon^2\), and the proof is finished. \(\square\)

A.5. Proof of Lemma 2.9. Let \(X(s) \in L^2(\Omega, \mathbb{P}, L^p(\mathbb{S}, \mu))\) be pregaussian. Hence there exists a Gaussian variable \(W\) on \((\Omega, \mathcal{F}, \mathbb{P})\) with values in \(L^p(\mathbb{S}, \mu)\) with same covariance operator than \(X\). By Theorem 3.1 of Rajput [41], we may see \(W\) as a Gaussian process \((W(s))_{s \in \mathbb{S}}\) whose paths are \(\mathbb{P}\)-a.s. in \(L^p(\mathbb{S}, \mu)\). Then,

\[
\|G(X)\|_{2, L^p(\mu)} = \|W\|_{2, L^p(\mu)} \geq C_p \|W\|_{p, L^p(\mu)} = C_p \left( \int_{\mathbb{S}} \mathbb{E}(|W(s)|^p) \mu(ds) \right)^{1/p} \\
= \tilde{C}_p \left( \int_{\mathbb{S}} (\mathbb{E}(|W(s)|^2))^{p/2} \mu(ds) \right)^{1/p} = \tilde{C}_p \left( \int_{\mathbb{S}} (\mathbb{E}(|X(s)|^2))^{p/2} \mu(ds) \right)^{1/p}.
\]

the reverse inequality may be proved similarly.

The fact that a centered \(X\) such that \(\int_{\mathbb{S}} (\mathbb{E}(|X(s)|^2))^{p/2} \mu(ds) < \infty\) is pregaussian follows from Lemma 5.1 of [41]. \(\square\)
APPENDIX B. PROOF OF THE MARTINGALE RESULTS

B.1. Proof of Lemma 3.1. For any Banach space $X$, by Doob’s maximal inequality for the submartingale $(|S_n|X)_n \geq 0$, $\|x\|_p \leq \frac{1}{p} \|S_n(d)\|_{p,X}$. Then, item (i) follows from (5) when $1 < p \leq 2$ and from (6) when $p > 2$.

Let us prove item (ii). By orthogonality of real-valued martingale differences, we see that $S_n(d)/\sqrt{n}$ is pregaussian and that $G(S_n(d)/\sqrt{n}) = G(d)$. Then, we conclude thanks to Lemma 2.4.

B.2. Proof of Proposition 3.2. The case where $X$ is $r$-smooth is just Proposition 2.1 of [8]. Assume that $X = L^p(S)$. It suffices to prove the result when $d \in L^p(S, L^p(\Omega, F_0))$, otherwise $K_r(d) = +\infty$. There exists a sequence of step functions $(d_n)_{n \geq 1}$ converging in $L^p(S, L^p(\Omega, F_0))$ to $d$. We may write $d_n(s, \omega) = \sum_{k=1}^{m_n} f_k, n(\omega) 1_{A_k, n}(s)$, where $A_k, n \in S$ and $f_k, n \in L^p(\Omega, \mathbb{F})$. Let $d_n := \sum_{k=1}^{m_n} (f_k, n - E_1(f_k, n)) 1_{A_k, n}$. Then $(d_n)_{n \geq 1}$ converges to $d$ in $L^p(S, L^p(\Omega, F_0))$ as well (hence also in $L^p(\Omega, L^p(S))$, by Lemma 2.2) and for every $s \in S$, $d_n(s, \cdot)$ is a real-valued martingale difference in $L^p(\Omega, F_0, \mathbb{P})$. Hence, applying Proposition 3.2 to the $(2, 1)$-smooth Banach space $\mathbb{R}$, we obtain that there exists $C_p > 0$ such that for every $s \in S$,

\begin{align}
\|M_p(d_n(s, \cdot))\|_{p, \infty} \leq C_p \|d_n(s, \cdot)\|_p.
\end{align}

Notice that $M_p(d_n, L^r(S)) \leq \left( \int_S (\|M_p\| d\mu(s))^{r} \right)^{1/r}$. Writing $\varphi(s, \cdot) = M_p(d_n(s, \cdot), \mathbb{R})$, it follows from lemma 2.2 that

\begin{align}
\|M_p(d_n, L^r(S))\|_{p, \infty} \leq C_{p,r} \left( \int_S \|\varphi(s, \cdot)\|^{r}_{p, \infty} \right)^{1/r}.
\end{align}

Then, we infer from (47) that

\begin{align}
\|M_p(d_n, L^r(S))\|_{p, \infty} \leq C_{p,r} \left( \int_S \|\varphi(s, \cdot)\|_{p, \infty} \right)^{1/r}.
\end{align}

The desired result then follows by letting $n \to \infty$ (approximate first $M_p$ by a supremum over a finite set of integers and use the monotone convergence theorem).

To prove the a.s. convergence, we first notice that, by (19) and the Banach principle, see Lemma 2.1 the set

\begin{align}
\{d \in L^p(\Omega, F_0, X) : \mathbb{E}_1(d) = 0 \text{ and } |S_n(d)| \leq o(n^{1/p}) \text{ a.s.}\}
\end{align}

is closed in $L^p(\Omega, X)$. Hence it suffices to prove the result for stationary martingale differences taking values in a finite dimensional Banach space, and the result is well-known in that case.

□

B.3. Proof of Proposition 3.3. This is just Proposition 3.3 of [8] when $X$ is $(2, L)$-smooth. When $X = L^p(S, S, \mu)$, the proof may be done as the proof of Proposition 3.2.

B.4. Proof of Proposition 3.4. As usual we shall first prove a tightness result. Let us recall the definition of tightness required here.

Let $X \in L^0(\Omega, F_0, \mathbb{P}, X)$. Recall that $S_{n,t} = S_{n,t}(X) := S_{[nt]} + (nt - [nt])X_{[nt]}$ and $T_{n,t} := \frac{S_{n,t}}{\sqrt{n}}$. We consider $((T_{n,t})_{0 \leq t \leq 1})_{n \geq 0}$ as a process taking values in $C([0, 1], X)$, the Banach space of continuous functions from $[0, 1]$ to $X$.

Definition B.1. We say that $((T_{n,t})_{0 \leq t \leq 1})_{n \geq 0}$ is tight if for every $\varepsilon > 0$, there exists a compact set $\kappa$ of $C([0, 1], X)$ such that

\begin{align}
\mathbb{P}((T_{n,t})_{0 \leq t \leq 1} \in \kappa) \geq 1 - \varepsilon \quad \forall n \geq 0.
\end{align}
Let $\mathcal{X}$ be either 2-smooth or of cotype 2. Let $d \in G(\mathcal{X})$ with $\mathbb{E}_{-1}(d) = 0$. Let us prove the tightness of $((T_{n,t}(d))_{0 \leq t \leq 1})_{n \geq 1}$ in $C([0,1], \mathcal{X})$.

We first recall the following tightness criteria that may be easily deduced from Theorem 11.5.4 of Dudley \[22\].

**Lemma B.1.** Let $(\Gamma, \delta)$ be a separable complete metric space endowed with its Borel $\sigma$-algebra. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(Z_n)_{n \geq 1}$ be a sequence of random variables on $\Omega$ taking values in $\Gamma$. Assume that, for every $\varepsilon > 0$, there exist $n_0 \geq 1$ and random variables $(Z'_n)_{n \geq n_0}$ such that

- (i) $(Z'_n)_{n \geq n_0}$ is tight;
- (ii) $\sup_{n \geq n_0} \mathbb{E}(\delta(Z_n, Z'_n)) < \varepsilon$.

Then $(Z_n)_{n \geq 1}$ is tight.

Since $\mathcal{X}$ is separable, $\sigma(d)$ (the $\sigma$-algebra generated by $d$) is countably generated and there exists an increasing filtration $(G_m)_{m \geq 1}$ such that $G_m$ is finite for every $m \geq 1$ and $\sigma(d) = \bigvee_{m \geq 1} G_m$. For every $m \geq 1$, let $d_m := \mathbb{E}(d|G_m)$. Since $G_m$ is finite, there exists $A_{1,m}, \ldots, A_{k_m,m} \in G_m$ and $x_{1,m}, \ldots, x_{k_m,m} \in \mathcal{X}$ such that $d_m = \sum_{1 \leq k \leq k_m} x_k 1_{A_{k,m}}$. By Lemma 2.28 $(d_m)_{m \geq 1}$ converges in $G(\mathcal{X})$ to $d$. Hence, writing $d := d_m - \mathbb{E}_{-1}(d_m)$ and using Lemma 2.27 $(d_m)_{m \geq 1}$ converges in $G(\mathcal{X})$ to $d$.

By the WIP for real-valued martingales with stationary and ergodic increments, for every $m \geq 1$, $(T_{n,t}(d_m))_{0 \leq t \leq 1}$, $n \geq 0$, is tight in $C([0,1], \mathcal{X})$.

Now, by Lemma 3.1 (using (38) when $\mathcal{X}$ is $(2, L)$-smooth and (9) when $\mathcal{X}$ has cotype 2),

$$\| \sup_{0 \leq t \leq 1} |T_{n,t}(d_m) - T_{n,t}(d)|\|_2 \leq \frac{3}{\sqrt{n}} \max_{1 \leq k \leq n} \| S_k(d_m) - S_k(d) \|_2 \leq C \| d_m - d \|_{G(\mathcal{X})} \underset{m \to \infty}{\longrightarrow} 0,$$

and the tightness of $((T_{n,t}(d))_{0 \leq t \leq 1})_{n \geq 0}$ in $C([0,1], \mathcal{X})$ follows from Lemma 3.1.

Let us write $T_{n,t}(d) = T_{n,t}$. The second step consists in proving the convergence of the finite-dimensional laws. That is, it remains to prove that, for any $0 = t_0 < \ldots < t_m = 1$,

$$((T_{n,t_i} - T_{n,t_{i-1}})_{1 \leq i \leq m})_{n \geq 1}$$

converges in law to $(W_t - W_{t_{i-1}})_{1 \leq i \leq m}$, where $(W_t)_{0 \leq t \leq 1}$ is a brownian motion with covariance operator $K_d$. Using tightness again (and the Cramer-Wold device), it suffices to prove that for any $0 = t_0 < \ldots < t_m = 1$ and any $x_1^*, \ldots, x_m^* \in \mathcal{X}^*$,

$$\sum_{i=1}^m x_i^*(T_{n,t_i} - T_{n,t_{i-1}})$$

converges in law to $\sum_{i=1}^m x_i^*(W_{t_i} - W_{t_{i-1}})$ as $n \to \infty$.

Hence, we are back to prove a CLT for an array of martingale differences. Let us recall the following CLT of McLeish, as stated in Theorem 3.2 page 58 of Hall and Heyde \[27\].

**Proposition B.2.** Let $(X_{n,j})_{1 \leq j \leq k_n}$ be (real valued) martingale differences for every $n \geq 1$. Assume that there exists $\sigma \geq 0$ such that

- (i) $\max_{1 \leq j \leq k_n} |X_{n,j}| \overset{p}{\to} 0$;
- (ii) $\sum_{1 \leq j \leq k_n} X_{n,j}^2 \overset{p}{\to} \sigma^2$;
- (iii) $\sup_{n \geq 1} \mathbb{E}(\max_{1 \leq j \leq k_n} X_{n,j}^2) < \infty$.

Then $(\sum_{1 \leq j \leq k_n} X_{n,j})_{n \geq 1}$ converges in law to a normal law $N(0, \sigma^2)$.

Take $k_n := n$ and for every $1 \leq i \leq m$ and every $[nt_{i-1}] \leq j \leq [nt_i] - 1$, take $X_{n,j} := x_j^*(d) \circ \theta^j/\sqrt{n}$.

Then, setting $Z := \max_{1 \leq j \leq m} |X_{n,j}|$ (which belongs to $L^2(\Omega)$), we have $\max_{1 \leq j \leq k_n} |X_{n,j}| \leq \max_{1 \leq j \leq k} Z \circ \theta^j/\sqrt{n}$ which implies (i), by the Borel-Cantelli lemma, and (iii) by standard arguments. Now, by the ergodic theorem we have

$$\frac{1}{n} \sum_{j=\lfloor nt_{i-1} \rfloor}^{\lfloor nt_i \rfloor} (x_j^*(d))^2 \circ \theta^j \overset{n \to \infty}{\longrightarrow} (t_i - t_{i-1}) \mathbb{E}((x_j^*(d))^2) \quad \mathbb{P}\text{-a.s.},$$

hence in probability. Hence the proof is complete.

$\square$
B.5. Proof of Proposition 3.5. Let us prove the CLIL. Notice that \( \mathcal{G}_0(\mathcal{X}) := \{d \in \mathcal{G}(\mathcal{X}, \mathcal{F}_0) : \mathbb{E}_-^{-1}(d) = 0\} \) is a closed subspace of \( \mathcal{G}(\mathcal{X}) \). By (13) and Proposition E.1, the set of \( d \in \mathcal{G}_0(\mathcal{X}) \), such that \( (d \circ \theta_n)_{n \geq 0} \) satisfies the CLIL is closed in \( \mathcal{G}_0(\mathcal{X}) \). Then, the CLIL follows by approximating any \( d \in \mathcal{G}_0(\mathcal{X}) \) by a martingale difference with values in a finite dimensional Banach space as in the proof of Proposition 3.3.

Then, (13) follows from a result of Kuelbs (see e.g. Proposition D. of [8]) combined with the LIL for real valued stationary (and ergodic) martingale differences.

To prove the ASIP, we just apply the following version of Theorem 3.2 of Berger [4] whose proof may be done similarly.

**Theorem B.3.** Let \( \mathcal{X} \) be a real separable Banach space. Assume that \( \theta \) is ergodic. Let \( X \in L^0(\Omega, \mathcal{F}_0, \mathbb{P}, \mathcal{X}) \) be such that \( \mathbb{E}(x^*(X)^2) < \infty \), for every \( x^* \in \mathcal{X}^* \). Assume that \((X \circ \theta_n)_{n \geq 0}\) satisfies the CLIL and that for every \( x^* \in \mathcal{X}^* \), there exists \( Z = Z_{x^*} \in L^2(\Omega, \mathcal{F}_0, \mathbb{R}) \) with \( \mathbb{E}_-^{-1}(Z) = 0 \) such that

\[
S_n(x^*(X)) - S_n(Z) = o(\sqrt{nL(n)}) \quad \mathbb{P}\text{-a.s.} \tag{48}
\]

\[
\|S_n(x^*(X)) - S_n(Z)\|_2 = o(\sqrt{n}) \tag{49}
\]

Then, for every \( x^*, y^* \in \mathcal{X}^* \), \( \mathbb{K}(x^*, y^*) := \lim_{n \to \infty} \frac{\mathbb{E}(x^*(S_n(Y)) y^*(S_n(X)))}{n} \) exists. Assume moreover that \( \mathbb{K} \) is the covariance operator of a gaussian variable. Then, \((X \circ \theta_n)_{n \geq 0}\) satisfies the ASIP.

**Appendix C. Proof of the maximal inequalities**

C.1. Proof of Proposition 4.1. We make the proof by induction. For \( d = 0 \) we have

\[
S_1 = X - \mathbb{E}_-^{-1}(X) + \mathbb{E}_-^{-1}(X) = (X - \mathbb{E}_-^{-1}(X)) + \mathbb{E}_-^{-1}(S_1)
\]

and the result follows in that case.

Assume that we already proved the result for some \( d \geq 0 \). For every \( 1 \leq i \leq 2^{d+1} \), we have

\[
S_i = \sum_{\ell=0}^{i-1} (X - \mathbb{E}_-^{-1}(X)) \circ \theta^\ell + \sum_{\ell=0}^{i-1} (\mathbb{E}_-^{-1}(X)) \circ \theta^\ell,
\]

and for every \( 1 \leq j \leq 2^d \) (with \( \sum_{\ell=0}^{j-1} = 0 \)),

\[
\sum_{\ell=0}^{2j-1} (\mathbb{E}_-^{-1}(X)) \circ \theta^\ell = \sum_{\ell=0}^{j-1} (\mathbb{E}_-^{-1}(X) + \mathbb{E}_-^{-1}(X) \circ \theta) \circ \theta^{2\ell};
\]

\[
\sum_{\ell=0}^{2j-2} (\mathbb{E}_-^{-1}(X)) \circ \theta^\ell = (\mathbb{E}_-^{-1}(X)) \circ \theta^{2j-2} + \sum_{\ell=0}^{j-2} (\mathbb{E}_-^{-1}(X) + \mathbb{E}_-^{-1}(X) \circ \theta) \circ \theta^{2\ell}.
\]

Hence,

\[
\max_{1 \leq i \leq 2^{d+1}} |S_i|_{\mathcal{X}} \leq \max_{1 \leq i \leq 2^{d+1}} \left| \sum_{\ell=0}^{i-1} (X - \mathbb{E}_-^{-1}(X)) \circ \theta^\ell \right|_{\mathcal{X}} + \max_{1 \leq j \leq 2^d} \left| \mathbb{E}_-^{-1}(X) \circ \theta^{2j-2} \right|_{\mathcal{X}}.
\]

(50)

We shall apply our induction hypothesis to the following situation: \( \bar{X} := \mathbb{E}_-^{-1}(X) + \mathbb{E}_-^{-1}(X) \circ \theta \), the transformation \( \bar{\theta} := \theta^2 \) and the filtration given by \( \mathcal{F}_n := \bar{\theta}^{-n}(\mathcal{F}) = \mathcal{F}_{2n} \) for every \( n \in \mathbb{Z} \).

We shall also use the notation \( \bar{\mathcal{E}}_n(\cdot) := \mathbb{E}(\cdot|\mathcal{F}_n) \) and \( \bar{S}_n = \sum_{\ell=0}^{n-1} \bar{X} \circ \bar{\theta}^\ell \).

Notice then that we have
\[ \tilde{S}_n = \sum_{\ell=0}^{n-1} (E_{-1}(X) + E_{-1}(X) \circ \theta) \circ \theta^{2\ell}, \quad \tilde{E}_{-2k}(\tilde{S}_{2k}) = E_{-2k+1}(S_{2k+1}) \]
and
\[ \tilde{X} - \tilde{E}_{-1}(\tilde{X}) = E_{-1}(S_1) + E_{-1}(S_1) \circ \theta - E_{-2}(S_2). \]

Hence, by our induction hypothesis and using the change of index \( k \to k + 1 \), we infer that

\[
(51) \quad \max_{1 \leq \ell \leq 2^d} |\tilde{S}|_\mathcal{X} \leq |E_{-2d+1}(S_{2d+1})|_\mathcal{X} + \sum_{k=1}^{(d+1)-1} \max_{0 \leq \ell \leq 2^{(d+1)-1-k-1}} |E_{-2k}(S_{2k})|_\mathcal{X} \circ \theta^{2k+1} \ell

+ \sum_{k=1}^{(d+1)-1} \max_{1 \leq \ell \leq 2^{(d+1)-k-1}} \left| \sum_{\ell=0}^{i-1} \left[ E_{-2k}(S_{2k}) + E_{-2k}(S_{2k}) \circ \theta^{2k} - E_{-2k+1}(S_{2k+1}) \right] \circ \theta^{2k+1} \ell \right|_\mathcal{X}.
\]

Then, the result follows by combining (50) and (51). \( \square \)

C.2. Proof of Corollary 4.2 We shall use Proposition 4.1 We first notice that

\[
0 \leq \ell \leq 2^{d-1-k-1} \quad |E_{-2k}(S_{2k})|_\mathcal{X} \circ \theta^{2k+1} \ell \leq \left( \sum_{0 \leq \ell \leq 2^{d-1-k-1}} |E_{-2k}(S_{2k})|_\mathcal{X}^p \circ \theta^{2k+1} \ell \right)^{1/p}.
\]

Hence, using that \( \theta \) preserves \( \mathbb{P} \), we infer that

\[
\left\| \max_{0 \leq \ell \leq 2^{d-1-k-1}} \left| E_{-2k}(S_{2k})|_\mathcal{X} \circ \theta^{2k+1} \ell \right|_p \right\| \leq 2^{(d-1-k)/p} \|E_{-2k}(S_{2k})\|_p \mathcal{X}.
\]

Applying (3) to (the martingale difference) \( d = X - E_{-1}(X) \) we see that

\[
\left\| \max_{1 \leq \ell \leq 2^d} \left( \sum_{\ell=0}^{i-1} (X - E_{-1}(X)) \circ \theta^\ell \right)_\mathcal{X} \right\|_p \leq C_{p,L}(\|X\|_p \mathcal{X} + \|E_{-1}(X)\|_p \mathcal{X}).
\]

Similarly, we may apply (3) with \( d_k = E_{-2k}(S_{2k}) + E_{-2k}(S_{2k}) \circ \theta^{2k} - E_{-2k+1}(S_{2k+1}) \) (and \( \theta^{2k+1} \) instead of \( \theta \)). To conclude we just notice that \( \|X - E_{-1}(X)\|_p \mathcal{X} \leq 2\|X\|_p \mathcal{X} \) and that \( \|E_{-2k}(S_{2k}) + E_{-2k}(S_{2k}) \circ \theta^{2k} - E_{-2k+1}(S_{2k+1}) \|_p \mathcal{X} \) and that \( \|E_{-2k}(S_{2k}) + E_{-2k}(S_{2k}) \circ \theta^{2k} - E_{-2k+1}(S_{2k+1}) \|_p \mathcal{X} \leq 4\|E_{-2k}(S_{2k})\|_p \mathcal{X} \).

C.3. Proof of Corollary 4.3 We proceed as in the proof of Corollary 4.2 except that we apply (3) in the 2-smooth case and (9) in the case of cotype 2. \( \square \)

Appendix D. Proof of the Limit Theorems under Projective Conditions

Before doing the proof, let us give general facts about \( \| \cdot \|_{MW_p} \), that will be used in the sequel.

Let \( 1 < p \leq 2 \). Define \( MW_{p} := \{X \in L^p(\Omega, \mathcal{F}_0, \mathbb{P}, \mathcal{X}) : \|X\|_{MW_p} < \infty \} \). Then, \( (MW_p, \| \cdot \|_{MW_p}) \) is a Banach space.

For every \( X \in L^1(\Omega, \mathcal{F}_0, \mathbb{P}, \mathcal{X}) \) define \( QX = E_0(X \circ \theta) \). Notice that \( Q^n(X) = E_0(X \circ \theta^n) \). Then, clearly \( Q \) is a contraction of \( L^1(\Omega, \mathcal{F}_0, \mathcal{X}) \) and, by Lemma 2.7 \( Q \) is power bounded on \( \mathbb{G}(\mathcal{X}) \), i.e., for every \( X \in \mathbb{G}(\mathcal{X}) \), \( \sup_{n \geq 1} \|Q^nX\|_{\mathbb{G}(\mathcal{X})} \leq C\|X\|_{\mathbb{G}(\mathcal{X})} \), for some universal \( C > 0 \). Now, we see that

\[
\|X\|_{MW_p} = \sum_{n \geq 0} \frac{\|\sum_{k=0}^{n-1} Q^kX\|_{p, \mathcal{X}}}{2^n/p}, \text{ if } 1 < p < 2,
\]

\[
\|X\|_{MW_2} = \sum_{n \geq 0} \frac{\|\sum_{k=0}^{n-1} Q^kX\|_{\mathbb{G}(\mathcal{X})}}{2^n/2}.
\]

Hence, in any case, \( Q \) is power bounded on \( MW_p \).
Writing $V_n := I + \cdots + Q^{n-1}$ and using that $\|V_n V_k X\|_{p,\mathcal{X}} \leq C \min(k\|V_n\|_{p,\mathcal{X}}, n\|V_k X\|_{p,\mathcal{X}})$, we see that, for every $X \in MW_p$,

$$
\frac{\|V_n^2 X\|_{MW_p}}{2^n} \leq C_p \left( \frac{\|V_n\|_{p,\mathcal{X}}}{2^{n/p}} + \sum_{k \geq n+1} \frac{\|V_{2^k} X\|_{p,\mathcal{X}}}{2^{k/p}} \right) \to 0. 
$$

Now, for every $n \geq 1$, taking $m$ such that $2^m \leq n < 2^{m+1}$, we have $\|V_n X\|_{MW_p} \leq C \sum_{k=0}^{m} \|V_{2^k}\|_{MW_p} = o(n)$.

In particular, we see that $Q$ is mean ergodic on $MW_p$ and has no non trivial fixed point (see e.g. Theorem 1.3 p. 73 of [28]), i.e.,

$$
MW_p = (I - Q)MW_p. 
$$

D.1. Proof of Proposition 5.1 and Theorem 5.2. In both results, $\mathcal{X}$ is a Banach space of cotype 2. Let $X \in (I - Q)MW_2$. Let $Y \in MW_2$ be the unique (notice that $Q$ has no fixed point on $MW_2$) solution to $X = (I - Q)Y$. Then, one may define

$$
D(X) := Y - E_{-1}(Y) = Y - QY \circ \theta^{-1}. 
$$

Notice that $X = D(X) + QY - QY \circ \theta^{-1}$ and that $D(X)$ is a martingale difference. In particular

$$
\|G(S_n(D(X)))\|_{2,\mathcal{X}} = \sqrt{n}\|G(D(X))\|_{2,\mathcal{X}}. 
$$

Recall that, since $\mathcal{X}$ has cotype 2, there exists $C > 0$, such that for every $Z \in G(\mathcal{X})$,

$$
\|G(Z)\|_{2,\mathcal{X}}/C \leq \|Z\|_{G(\mathcal{X})} \leq C\|G(Z)\|_{2,\mathcal{X}}. 
$$

Now, it follows from the proof of Proposition 4.1 (combined with 5.1) applied to the martingales with stationary increments that appear in the proof) that there exists $D > 0$ such that for every $d \geq 0$,

$$
\|G(S_{2^d}(X))\|_{2,\mathcal{X}} \leq D 2^{d/2} \left( \|G(X)\|_{2,\mathcal{X}} + \sum_{k=0}^{d} 2^{-k} \|G(E_0(S_{2^k}(X)))\|_{2,\mathcal{X}} \right). 
$$

Notice that $\|S_{2^d}(QY - QY \circ \theta^{-1})\|_{G(\mathcal{X})} \leq \|QY \circ \theta^{-1}\|_{G(\mathcal{X})} + \|QY \circ \theta^{2^d-1}\|_{G(\mathcal{X})} = o(2^{d/2})$ and that $\|G(S_{2^d}(D(X)))\|_{2,\mathcal{X}} \leq \|G(S_{2^d}(X))\|_{2,\mathcal{X}} + \|G(S_{2^d}(QY - QY \circ \theta^{-1}))\|_{2,\mathcal{X}}$.

Combining this with (56), (55), and (55) and letting $d \to \infty$, we infer that

$$
\|D(X)\|_{G(\mathcal{X})} \leq C\|X\|_{MW_2}. 
$$

Hence, we may extend our linear operator $D$ continuously to $(I - Q)MW_2^{MW_2} = MW_2$. Notice that $D$ takes values in $G_0(\mathcal{X}) = \{ Z \in G(\mathcal{X}), F_0 : E_{-1}(Z) = 0 \}$.

Let us prove Proposition 5.1. By Corollary 4.3 and (9), there exists $C > 0$ such that

$$
\| \max_{1 \leq k \leq n} |S_k(X) - S_k(D(X))|_{\mathcal{X}} \|_2 \leq C \sqrt{n}\|X\|_{MW_2}. 
$$

By linearity of $D$ (and of $X \mapsto S_k(X)$) it then suffices to prove (23) for a set of $X$‘s that is dense in $MW_2$, in particular for $X \in (I - Q)MW_2$. But if $X = (I - Q)Y$ with $Y \in MW_2$, we have, for every $K > 0$

$$
\| \max_{1 \leq k \leq n} |S_k(X) - S_k(D(X))|_{\mathcal{X}} \|_2 \leq \| \max_{1 \leq k \leq n} |Q_k(QY - QY \circ \theta^{-1})|_{\mathcal{X}} \|_2 
\leq \|QY\|_{2,\mathcal{X}} + \| \max_{1 \leq k \leq n} |QY \circ \theta^{k-1}|_{\mathcal{X}} \|_2 \leq \|QY\|_{2,\mathcal{X}} + K + n\|QY\|_{1,\{QY|_{\mathcal{X}} \geq K\}} \to 0, 
$$

Hence

$$
\limsup_{n \to \infty} \max_{1 \leq k \leq n} |S_k(X) - S_k(D(X))|_{\mathcal{X}} \|_2 \leq \|QY\|_{1,\{QY|_{\mathcal{X}} \geq K\}} \to 0 \quad \text{as} \quad K \to \infty, 
$$

and (23) holds. Then, the proof of the WIP follows from Lemma 4.1 and Proposition 3.3.
Let us prove Theorem 5.2. By Proposition 3.3 and (22), for every $1 < p < 2$, there exists $C_p > 0$ such that
\[ \|M_2(X - D(X))\|_{p, \infty} \leq C_p \|X\|_{MW_2}. \]

Hence, by the Banach principle, see Lemma E.1, it suffices to prove (24) for $X = (I - Q)Y$, with $Y \in MW_2$. But in this case the result is obvious, since $|QY|_{X} \in L^2(\Omega)$ and, by the Borel-Cantelli lemma, $|QY|_{X} \circ \theta^{n-1} = o(\sqrt{n}) \, P$-a.s. By (24) and Proposition 5.1 (X \circ \theta^n)_{n \geq 0} satisfies the CLIL. Then, the ASIP follows from Proposition 4.3 using that $D(X)$ is pregaussian.

It remains to prove (25). The first equality follows from (24) and (14). Let us prove that, with $d = D(X)$, $\sup_{x^* \in X^*, |x^*| \leq 1} \|x^*(d)\|_2 \leq 10\sqrt{2} \|X\|_{MW_2}$. We first notice that $x^*(d) = D(x^*(X))$ (with the obvious “new” meaning of the operator $D$). Proceeding as above one can prove that for every $m \geq 0$,
\[ \|x^*(d)\|_2 = 2^{m/2} \|S_{2m}(d)\|_2/2^{m/2} \leq \|S_{2m}(X)\|_2/2^{m/2} + \|S_{2m}(d) - S_{2m}(X)\|_2/2^{m/2}. \]

Applying Proposition 5.1 (noticing that $\|x^*(X)\|_{MW_2} \leq \|X\|_{MW_2}$) and Corollary 4.2 to $x^*(X)$, we derive that $\|x^*(d)\|_{MW_2} \leq 10\sqrt{2} \|X\|_{MW_2}$ and the proof is complete.

D.2. Proof of Theorem 5.3. Let us prove the WIP. As above we shall first prove tightness. Let $X \in MW_2$. Let $\varepsilon > 0$. By (53), there exists $Y \in MW_2$ such that $\|X - (I - Q)Y\|_{MW_2} \leq \varepsilon$.

Then, by (4),
\[ \max_{1 \leq k \leq n} \|S_k(X) - S_k((I - Q)Y)\|_{X} \|_2 \leq C\varepsilon. \]

Now, as in the proof of Proposition 5.1 for every $K > 0$ we have
\[ \max_{1 \leq k \leq n} \|S_k((I - Q)Y) - S_k(Y - E_{-1}(Y))\|_{X} \|_2 \leq \|QY\|_{2, X} + K + n\|QY\|_{\chi 1_{\{|QY|_{X} \geq K\}}} 2. \]

Chose $K$ such that $\|QY\|_{2, X} < \varepsilon$ and then chose $n_0 \geq (\|QY\|_{2, X} + K)^2/\varepsilon^2$.

Then, $\max_{n \geq 1} \|T_{n,t}(X) - T_{n,t}(Y - E_{-1}(Y))\|_{X} \leq (C + 6)\varepsilon$. Now, $Y - E_{-1}(Y)$ is a martingale difference, hence, by Proposition 3.3, $\{(T_{n,t}(Y - E_{-1}(Y))_{0 \leq n \leq 1})n \geq 0$ is tight in $C([0, 1], \chi)$. Then, the tightness of $\{(T_{n,t}(X)_{0 \leq n \leq 1})n \geq 0$ follows from Proposition 5.1.

The proof of the finite-dimensional laws may be done exactly as the proof of the martingale case, hence is omitted. The fact that the covariance operator is given as stated follows from the fact that for any $x^* \in X^*$, $x^*(X)$ satisfies the assumption of Proposition 6.1.

Let us prove the ASIP. We shall use Theorem 5.3. In particular, we have to prove that $(X \circ \theta^n)_{n \geq 0}$ satisfies the CLIL.

By (22) and Lemma E.1, the set $\{X \in MW_2 : (X \circ \theta^n)_{n \geq 0}$ satisfies the CLIL} is closed in $MW_2$. Hence, it suffices to prove the CLIL for $X = (I - Q)Y$, with $Y \in MW_2$. But then, $X = Y - E_{-1}(Y) + QY \circ \theta^{-1} - QY$ and $(I - QY) \circ \theta^n)_{n \geq 0}$ satisfies the CLIL by Proposition 3.5, while $|S_n(QY \circ \theta^{-1} - QY)|_{X} = o(\sqrt{n}) \, P$-a.s., by the Borel-Cantelli lemma. Hence the CLIL is proved.

Now, let $x^* \in X^*$. Clearly, $x^*(X)$ satisfies the assumption of Theorem 5.2 taking for $\chi$ the Hilbert space $\mathbb{R}$. In particular, there exists $Z \in L^2(\Omega, F_0, \mathbb{R})$ with $E_{-1}(Z) = 0$ such that
\[ S_n(x^*(X)) - S_n(Z) = o(\sqrt{nL(L(n)))} \, P$-a.s. \]
\[ \|S_n(x^*(X)) - S_n(Z)\|_2 = o(\sqrt{n}). \]

The fact that $K(x^*, y^*) := \lim_{n \to \infty} \frac{\cos(x^*(S_n(X)))y^*(S_n(X))}{n}$ is the covariance operator of a gaussian variable, follows from the WIP.
To prove the inequality in (26), by a result of Kuelbs (see e.g. Proposition D.1 in [8]), we have to prove that for every \( x^* \in X^* \), we have

\[
\limsup_n \frac{S_n(x^*(X))}{\sqrt{2nL(L(n))}} = (K(x^*, x^*))^{1/2} \quad \mathbb{P}\text{-a.s.}
\]

But this follows from Theorem 5.2 applied to \( x^*(X) \). Then, the inequality in (26) may be proved as the inequality in (25). \( \square \)

**APPENDIX E. TECHNICAL RESULT**

We recall here the Banach principle that we need (see Proposition C.1 of [8]).

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

\[
\mathbb{P}(\sup_{n \geq 1} |T_n| \lambda > \lambda |x|_B) \leq L(\lambda) \quad \forall \lambda > 0, x \in B.
\]

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

We give here a technical result concerning \(L^p\) spaces of \(L^r\)-valued variables.

**Lemma E.2.** Let \(1 \leq r < p < \infty\). Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and \((S, S, \nu)\) be a \(\sigma\)-finite measure spaces. There is a continuous embedding from \(L^r(S, L^p(\Omega))\) (resp. \(L^r(S, L^p(\Omega))\)) into \(L^p(\Omega, L^r(S))\) (resp. \(L^p(\Omega, L^r(S))\)).

**Proof.** We first recall some useful fact about weak \(L^p\)-spaces (see Exercise 1.1.11 p. 13 of Grafakos [29]). For every \(p > 1\) and every \(0 < t < p\), let

\[
N_{p,t}(\|X\|_X) := \sup_{\mathbb{P}(A) > 0} \frac{1}{\mathbb{P}(A)^{1/p-1/t} \left(\mathbb{E}(\|X\|_X 1_A)^{1/t}\right)}.
\]

Then, there exists \(C_{p,t}\) such that

\[
\|X\|_{p,\infty, X} / C_{p,t} \leq N_{p,t}(\|X\|_X) \leq C_{p,t}\|X\|_{p,\infty, X},
\]

and for \(t = 1, N_{s,1}\) is a norm.

Let \(f(s, \omega) = \sum_{i=1}^n f_i(\omega) 1_{A_i(s)}\) be a step function of \(L^r(S, L^p(\Omega))\), i.e. \(A_i \in S\) and \(f_i \in L^p(\Omega)\). We may consider \(f\) as an element of \(L^0(S \times \Omega, S \otimes F)\) or as an element of \(L^0(\Omega, F, L^0(S, S))\).

Take \(X = L^r(\mu)\) and \(t = r\). We have, using Fubini,

\[
\mathbb{E}(\|f\|_{L^r(\mu)}^r 1_A) = \int_S \mathbb{E}(\|f(s, \cdot)\|_{L^r}^r 1_A) d\mu(s) \leq \mathbb{E}(A)^{1/p-1/r} \int_S N_{p,r}(\|f(s, \cdot)\|_{L^r}^r 1_A) d\mu(s).
\]

Hence,

\[
\|f\|_{p,\infty, L^r(S)} \leq C_{p,r}^2 \left(\int_S \|f(s, \cdot)\|_{p,\infty}^r d\mu(s)\right)^{1/r}.
\]

Hence, the identity map sends step functions of \(L^r(S, L^p(\Omega))\) to elements of \(L^p(\Omega, L^r(S))\) in a continuous way. In particular, it can be extended continuously in an injective map to the whole \(L^r(S, L^p(\Omega))\). \( \square \)
References

[1] P. Assouad, *Espaces p-lisses et q-concaves, inégalités de Burkholder*, (French) Séminaire Maurey-Schwartz 1974–1975 : Espaces Lp, applications radonifiantes et géométrie des espaces de Banach, Exp. No. XV, 8 pp. Centre Math., École Polytech., Paris, 1975.

[2] E. del Barrio, E. Giné and C. Matrán, *Central limit theorems for the Wasserstein distance between the empirical and the true distributions*, Ann. Probab. 27 (1999), no. 2, 1009-1071.

[3] I. Berkes, L. Horváth, Q.-M. Shao, and J. Steinebach, *Strong laws for Lp-norms of empirical and related processes*, Endre Csáki 65. Period. Math. Hungar. 41 (2000), no. 1-2, 35-69.

[4] E. Berger, *An almost sure invariance principle for stationary ergodic sequences of Banach space valued random variables*, Probab. Theory Related Fields 84 (1990), no. 2, 161-201.

[5] D. L. Burkholder, *Sharp inequalities for martingales and stochastic integrals*. Colloque Paul Lévy sur les Processus Stochastiques (Palaiseau, 1987). Astérisque No. 157-158 (1988), 75-94.

[6] J. A. Clarkson, *Uniformly convex spaces*, Trans. Amer. Math. Soc. 40 (1936), no. 3, 396-414.

[7] C. Cuny, *Pointwise ergodic theorems with rate with applications to limit theorems for stationary processes*, Stoch. Dyn. 11 (2011), no. 1, 135-155

[8] C. Cuny, *A compact LIL for martingales in 2-smooth Banach spaces with applications*, accepted for publications in Bernoulli.

[9] C. Cuny and M. Lin, *Pointwise ergodic theorems with rate and application to the CLT for Markov chains*, Ann. Inst. Henri Poincar Probab. Stat. 45 (2009), no. 3, 710-733.

[10] C. Cuny and F. Merlevède, *On martingale approximations and the quenched weak invariance principle*, accepted for publicaiton in Ann. Probab. [arXiv:1202.2964]

[11] C. Cuny and F. Merlevède, *Strong invariance principles with rate for “reverse” martingales and applications*, accepted for publication in J. Theoret. Probab. [arXiv:1209.3677]

[12] S. Déét, *An empirical central limit theorem in L1 for stationary sequences*. Stochastic Process. Appl. 119 (2009), no. 10, 3494-3515

[13] J. Dedoeker, P. Doukhan and F. Merlevède, *Rates of convergence in the strong invariance principle under projective criteria*, Electron. J. Probab. 17 (2012), no. 16, 31 pp.

[14] J. Dedoeker and F. Merlevède, *The empirical distribution function for dependent variables: asymptotic and nonasymptotic results in Lp*, ESAIM Probab. and Stat. 11, (2007) 102-114.

[15] J. Dedoeker and F. Merlevède, *Convergence rates in the law of large numbers for Banach-valued dependent variables*, Teor. Veroyatn. Primen. 52 (2007), no. 3, 562-587; translation in Theory Probab. Appl. 52 (2008), no. 3, 416–438.

[16] J. Dedoeker and F. Merlevède, *On the almost sure invariance principle for stationary sequences of Hilbert-valued random variables*, Dependence in probability, analysis and number theory, 157-175, Kendrick Press, Heber City, UT, 2010.

[17] J. Dedoeker, F. Merlevède and F. Pène, *Rates in the strong invariance principle for ergodic automorphisms of the torus*, [arXiv:1206.3336]

[18] J. Dedoeker, F. Merlevède and F. Pène, *Empirical central limit theorems for ergodic automorphisms of the torus*, accepted for publication in ALEA, Lat. Am. J. Probab. Math. Stat.

[19] J. Dedoeker and C. Prieur, *New dependence coefficients. Examples and applications to statistics*, Probab. Theory Related Fields 132 (2005), no. 2, 203-236.

[20] Y. Derriennic and M. Lin, *The central limit theorem for Markov chains with normal transition operators, started at a point*, Probab. Theory Related Fields 119 (2001), no. 4, 508-528.

[21] J. Diestel and J. J. Uhl, *Vector measures*. With a foreword by B. J. Pettis, Mathematical Surveys, No. 15. American Mathematical Society, Providence, R.I., 1977. xiii

[22] R. M. Dudley, *Real analysis and probability*. Revised reprint of the 1989 original. Cambridge Studies in Advanced Mathematics, 74. Cambridge University Press, Cambridge, 2002.

[23] O. Durieu, *Independence of four projective criteria for the weak invariance principle*, ALEA Lat. Am. J. Probab. Math. Stat. 5 (2009), 21-26.

[24] M. I. Gordin and B. A. Lifšic, *Central limit theorem for stationary Markov processes*, (Russian) Dokl. Akad. Nauk SSSR 239 (1978), no. 4, 766-767.

[25] M. Gordin and M. Peligrad, *On the functional central limit theorem via martingale approximation*, Bernoulli 17 (2011), no. 1, 424-440.

[26] L. Grafakos, *Classical Fourier analysis*, Second edition. Graduate Texts in Mathematics, 249. Springer, New York, 2008.

[27] P. Hall and C. C. Heyde, *Martingale limit theory and its application*. Probability and Mathematical Statistics, Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, 1980.

[28] U. Krengel, *Ergodic theorems*, de Gruyter Studies in Mathematics, 6. Walter de Gruyter & Co., Berlin, (1985).

[29] J. Kuelbs, *A strong convergence theorem for Banach space valued random variables*, Ann. Probab. 4 (1976), no. 5, 744-771.
[30] M. Ledoux and M. Talagrand, Probability in Banach spaces. Isoperimetry and processes. Springer-Verlag, Berlin, 1991. xii+480 pp.

[31] Z. Lin and W. Liu, Strong approximation for a class of stationary processes, Stochastic Process. Appl. 119 (2009), no. 1, 249-280.

[32] Z. Lin and Y. Zhao, Strong approximation for ρ-mixing sequences, Sci. China Math. 55 (2012), no. 10, 2159-2182.

[33] M. Maxwell and M. Woodroofe, Central limit theorems for additive functionals of Markov chains, Ann. Probab. 28 (2000), no. 2, 713-724.

[34] F. Merlevède, 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 (2008), no. 1-2, 47-60.

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

[36] F. Merlevède, M. Peligrad and S. Utev, Recent advances in invariance principles for stationary sequences, Probab. Surv. 3 (2006), 1-36.

[37] J. Neveu, Discrete-parameter martingales. Translated from the French by T. P. Speed. Revised edition. North-Holland Mathematical Library, Vol. 10. North-Holland Publishing Co., Amsterdam-Oxford; American Elsevier Publishing Co., Inc., New York, 1975.

[38] M. Peligrad and S. Utev, A new maximal inequality and invariance principle for stationary sequences, Ann. Probab. 33 (2005), no. 2, 798-815.

[39] I. Pinelis, Optimum bounds for the distributions of martingales in Banach spaces, Ann. Probab. 22 (1994), no. 4, 1679-1706.

[40] B. S. Rajput, Gaussian measures on Lp spaces, 1?p < ∞, J. Multivariate Anal. 2 (1972), 382-403.

[41] E. Rio, inequalities for sums of dependent random variables under projective conditions, J. Theoret. Probab. 22 (2009), no. 1, 146-163.

[42] J. Rosiński, Central limit theorems for dependent random vectors in Banach spaces, Martingale theory in harmonic analysis and Banach spaces (Cleveland, Ohio, 1981), pp. 157-180, Lecture Notes in Math., 939, Springer, Berlin-New York, 1982.

[43] Q. M., Shao, Almost sure invariance principles for mixing sequences of random variables, Stochastic Process. Appl. 48 (1993), no. 2, 319-334.

[44] O. Sharipov, Laws of the iterated logarithm and an almost sure invariance principle for mixing B-valued random variables and autoregressive processes, Lith. Math. J. 49 (2009), no. 2, 203-215.

[45] N. Vakhania, Sur une propriété des répartitions normales de probabilités dans les espaces lp(1?p < ∞) et H, (French) C. R. Acad. Sci. Paris 260 1965 1334-1336.

[46] W. B. Wu, Strong invariance principles for dependent random variables. Ann. Probab. 35 (2007), 2294-2320.

[47] W. B. Wu, and Z. Zhao, Moderate deviations for stationary processes, Statist. Sinica 18 (2008), no. 2, 769-782.

[48] O. Zhao and M. Woodroofe, Law of the iterated logarithm for stationary processes, Ann. Probab. 36 (2008), no. 1, 127-142.

Laboratoire MAS, Ecole Centrale de Paris, Grande Voie des Vignes, 92295 Chatenay-Malabry Cedex, FRANCE.

E-mail address: christophe.cuny@ecp.fr