The Annals of Probability
2017, Vol. 45, No. 3, 1578–1611
DOI: 10.1214/16-AOP1095

INVARIANCE PRINCIPLES UNDER THE MAXWELL–WOODROOFE CONDITION IN BANACH SPACES

BY CHRISTOPHE CUNY

Centralesupelec

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. That result actually holds in the context of Banach valued stationary processes, including the case of L^p-valued random variables, with $1 \le p < \infty$. In this setting, we also prove the weak invariance principle, hence generalizing a result of Peligrad and Utev [*Ann. Probab.* **33** (2005) 798–815]. The proofs make use of a new maximal inequality and of approximation by martingales, for which some of our results are also new.

1. Introduction. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, θ be an invertible bi-measurable measure preserving transformation on Ω and $\mathcal{F}_0 \subset \mathcal{F}$ a σ-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 $S_n(X) = X + \cdots + X \circ \theta^{n-1}$.

In 2000, Maxwell and Woodroofe [34] proved the CLT for $(X \circ \theta^n)_{n\ge 0}$ under the condition

$$\sum_{n\ge 1} \frac{\|\mathbb{E}_0(S_n)\|_2}{n^{3/2}} < \infty. \tag{1}$$

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 [25] which in our setting is equivalent to the boundedness of $(\|\mathbb{E}_0(S_n(X))\|_2)_{n\ge 1}$.

Moreover, the condition (1) proved to be useful in applications. It is directly checkable for linear processes with innovations that are martingale differences; see, for example, Zhao and Woodroofe [53] (Proposition 5 and its proof). It leads to the optimal sufficient condition for the CLT in the case of ρ-mixing processes; see Merlevède, Peligrad and Utev [36], pages 14–15. It is implied by the condition $\sum_n (\log n)^{1+\varepsilon} \frac{\|\mathbb{E}_0(S_n)\|_2^2}{n^2} < \infty$, which can be checked in the case of Markov chains with normal Markov operator; see Cuny [6]. Finally, it is implied by the follow-

Received March 2015; revised January 2016.
MSC2010 subject classifications. Primary 60F17, 60F25, 60B12; secondary 37A50.
Key words and phrases. Banach valued processes, compact law of the iterated logarithm, invariance principles, Maxwell–Woodroofe's condition.

ing condition which is easier to check in applications (see, e.g., [7], Sections 3.1 and 3.3)

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

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

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 [39] by Peligrad and Utev 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 [40] and Wu and Zhao [52] 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 [21], Rassoul-Agha and Seppäläinen [44], Zhao and Woodroofe [53], Wu and Woodroofe [51], Cuny and Lin [9] and Cuny [6].

Very recently, Cuny and Merlevède [10] 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 those results, one may expect that (1) be a (sharp) sufficient condition for the LIL, as well as for its invariance principle.

In this paper, we provide a positive answer to that question (the example of Peligrad and Utev [39] ensures the sharpness). Actually our results hold in a Banach space setting, including any (separable) L^p spaces of a σ-finite measure space. More precisely, we prove the almost sure invariance principle (ASIP) in 2-smooth Banach spaces or in L^p spaces with $1\leq p<2$. We also obtain the WIP for dependent variables taking values in a 2-smooth Banach space or in a Banach space of cotype 2.

The main motivation for considering Banach-valued variables (especially the L^p cases, with $1\leq p<\infty$) is the fact that there are applications in statistics, in the study of the empirical process; see Section 6.2. Let us mention some papers in this vein: del Barrio, Giné and Matrán [20], Berkes, Horváth, Shao and Steinebach [4], Dedecker and Merlevède [15] and [14] or Dédé [12]. Let us mention also the very recent preprint of Dedecker and Merlevède [16].

To give a flavour of our results, we shall state here a theorem in L^p, $p \geq 1$.

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

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

For the sake of clarity, 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 [7], when $p = 2$; the WIP follows from Theorem 3.1 of Dedecker–Merlevède–Pène [18] (see also Theorem 2.1 of Dedecker–Merlevède–Pène [17], when $p = 2$; and the CLT has been obtained by Dédé [12] when $p = 1$.

We denote by $\|\cdot\|_2$ the L^2-norm on $(\Omega, \mathbb{P})$.

THEOREM 1.1. *Assume that θ is ergodic. Let $X \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}, L^p(S))$ $(1 \leq p < \infty)$ be such that $N_p(X) < \infty$, where*

$$N_p(X) = \sum_{n \geq 1} \frac{(\int_S \|\mathbb{E}_0(X_{n-1}(s))\|_2^p \mu(ds))^{1/p}}{n^{1/2}} \qquad \text{if } 1 \leq p < 2, \tag{3}$$

$$N_p(X) = \sum_{n \geq 1} \frac{\|(\int_S |\mathbb{E}_0(X_{n-1}(s))|^p \mu(ds))^{1/p}\|_2}{n^{1/2}} \qquad \text{if } p \geq 2. \tag{4}$$

Then the process $((T_{n,t})_{0 \leq t \leq 1})_{n \geq 1}$ converges in law in $C([0, 1], L^p(S, \mu))$ [to an $L^p(S, \mu)$-valued Brownian motion]; $(S_n(X)/\sqrt{nL(L(n))})n \geq 1$ is $\mathbb{P}$-a.s. relatively compact in $L^p(S, \mu)$. Moreover, there exists a universal constant $C > 0$, such that

$$\limsup_{n \to +\infty} \frac{(\int_S |\sum_{k=0}^{n-1} X_k(s)|^p \mu(ds))^{1/p}}{\sqrt{2nL(L(n))}} \leq C N_p \qquad \mathbb{P}\text{-a.s.}$$

The exact value of the ($\mathbb{P}$-a.s. constant) limsup above may be derived from the proof.

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

Notice that if $(X_n)_{n \geq 0}$ is a sequence of martingale differences [i.e., $\mathbb{E}_{n-1}(X_n) = 0$ for every $n \geq 1$] in $L^2(\Omega, L^p(S))$ if $p \geq 2$ or in $L^p(S, L^2(\Omega))$ if $1 \leq p \leq 2$ then the condition $N_p(X) < \infty$ automatically holds. In this case, the WIP and the ASIP are new when $1 \leq p < 2$; see Section 3 for references when $p \geq 2$.

When $p > 2$, neither the ASIP nor the WIP, can be obtained under condition (4), by the method of [7] or [18]. Indeed, when $p > 2$, the only sufficient condition (for the WIP or the ASIP) relying on $(\mathbb{E}_0(X_n))_{n \geq 1}$ that may be derived from the results of [18] or [7] is the following (see the proof of Theorem 2.1 of [17], page 758): $\sum_{n\geq 1} \frac{\|(\int_S |\mathbb{E}_0(X_{n-1}(s))|^p \mu(ds))^{1/p}\|_p}{n^{1/p}} < \infty$.

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 stationary (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 provide several examples including the case of the empirical process. All the results of Sections 2–5 are proved in the Appendix. The fact that our examples satisfy the required conditions is checked in the Section 6 itself.

Let us mention that versions of our results may be obtained (with slight modifications) for nonadapted stationary processes or stationary processes arising in noninvertible dynamical systems.

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 [22] for the basic facts on the topic (definition, conditional expectation, etc.). We shall also use results or notation from Ledoux and Talagrand [31]. In all the paper, we shall be concerned only with separable Banach spaces, in which case the definitions of a random variable of [22] and [31] coïncide. Other relevant references on the topic are the books by Vakhania, Tarieladze and Chobanyan [49] and by Araujo and Giné [1].

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 Ω 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:

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

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

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

For the sake of clarity, when they are understood, some of the references to Ω, $\mathcal{F}$ or $\mathbb{P}$ may be omitted. Also, in the case when $\mathcal{X} = \mathbb{R}$, we shall simply write $\|\cdot\|_p$ or $\|\cdot\|_{p,\infty}$. Recall that for every $p > 1$ there exists a norm on $L^{p,\infty}(\mathbb{P}, \mathcal{X})$ (see the proof of Lemma E.2), equivalent to the quasi-norm $\|\cdot\|_{p,\infty,\mathcal{X}}$, that makes $L^{p,\infty}(\mathbb{P}, \mathcal{X})$ a Banach space.

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

DEFINITION 2.1. We say that $\mathcal{X}$ is 2-*smooth*, if there exists $L \geq 1$, such that

$$|x+y|_{\mathcal{X}}^2 + |x-y|_{\mathcal{X}}^2 \leq 2\big(|x|_{\mathcal{X}}^2 + L^2|y|_{\mathcal{X}}^2\big) \qquad \forall x, y \in \mathcal{X}. \tag{5}$$

We shall speak about (2, L)-smooth spaces to emphasize the constant L such that (5) is satisfied.

REMARK. A Banach space is said to be 2-convex whenever (5) holds in the reverse direction.

DEFINITION 2.2. We say that $(d_n)_{1\leq n\leq N} \subset L^1(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{X})$ is a sequence of martingale differences, if there exist nondecreasing σ-algebras $(\mathcal{G}_n)_{0\leq n\leq N}$ such that for every $1 \leq n \leq N$, d_n is $\mathcal{G}_n$-measurable and $\mathbb{E}(d_n|\mathcal{G}_{n-1}) = 0$ $\mathbb{P}$-a.s.

The notion of 2-smooth Banach spaces is very useful due to the inequality (6) below; see, for instance, Proposition 1 of Assouad [2] (and its corollary).

Assume that $\mathcal{X}$ is $(2, L)$-smooth. Then, for every martingale differences $(d_n)_{1\leq n\leq N}$, we have

$$\mathbb{E}\big(|d_1 + \cdots + d_N|_{\mathcal{X}}^2\big) \leq 2L^2 \sum_{n=1}^N \mathbb{E}\big(|d_n|_{\mathcal{X}}^2\big). \tag{6}$$

Any Hilbert space is (2, 1)-smooth.

Any L^p space, $p \geq 2$, (of $\mathbb{R}$-valued functions) associated with a σ-finite measure is (2, $\sqrt{p-1}$)-smooth (see [41], Proposition 2.1).

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 relationship with the notion of pre-Gaussian variables that we shall introduce later.

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 $d_1, \ldots, d_N \in L^2(\Omega, \mathcal{X})$, with $\mathbb{E}(d_1) = \cdots = \mathbb{E}(d_N) = 0$, (6) holds (resp., such that (6) holds in the reverse direction).

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 $x^* \in X^*$, $\mathbb{E}(x^*(X)^2) < \infty$ and $\mathbb{E}(x^*(X)) = 0$. We define a bounded *symmetric* bilinear operator $\mathcal{K} = \mathcal{K}_X$ from $\mathcal{X}^* \times \mathcal{X}^*$ to $\mathbb{R}$, by

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

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

DEFINITION 2.4. We say that a random variable $W \in L^0(\Omega, \mathcal{X})$ 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 $x^* \in X^*$, $\mathbb{E}(x^*(X)^2) < \infty$ and $\mathbb{E}(x^*(X)) = 0$, is *pre-Gaussian*, if there exists a Gaussian variable $W \in L^0(\Omega, \mathcal{X})$ with the same covariance operator, that is, such that $\mathcal{K}_X = \mathcal{K}_W$. As in [31], when X is pre-Gaussian, we shall denote (abusively) by $G(X)$ a Gaussian variable having the same covariance operator as X.

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

DEFINITION 2.6. We say that $(X_n)_{n\ge 0}$ satisfies the almost sure invariance principle (ASIP) if, without changing its distribution, one can redefine the sequence $(X_n)_{n\ge 0}$ on a new probability space on which there exists a sequence $(W_n)_{n\ge 0}$ of centered i.i.d. Gaussian variables, such that

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

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

DEFINITION 2.7. We say that $(X_n)_{n\ge 0}$ satisfies the compact law of the iterated logarithm (CLIL) if the sequence $((X_0 + \cdots + X_{n-1})/\sqrt{nL(L(n))})_{n\ge 1}$ is $\mathbb{P}$-almost surely relatively compact in $\mathcal{X}$. We say that $(X_n)_{n\ge 0}$ satisfies the bounded law of the iterated logarithm (BLIL) if the sequence $((X_0 + \cdots + X_{n-1})/\sqrt{nL(L(n))})_{n\ge 1}$ is $\mathbb{P}$-almost surely bounded in $\mathcal{X}$.

It has been well known that if $(X_n)_{n\ge 0}$ satisfies the ASIP, it satisfies the CLIL, also. However, we have not found a proper reference where this is explicitly mentioned, hence we shall provide some arguments. Let $(W_n)_{n\in\mathbb{N}}$ be i.i.d. Gaussian variables taking values in $\mathcal{X}$. Then, combining the theorem on page 107 of [32] and Lemma 3 of [30] (alternatively, combining Theorem 8.6 and Lemma 3.1 of [31]), it follows that $(W_n)_{n\in\mathbb{N}}$ satisfies the CLIL. Then the fact that if $(X_n)_{n\geq 0}$ satisfies the ASIP, it satisfies the CLIL also, and readily follows from a standard approximation argument.

It is known (see the discussion on page 274 of [31]) that in order to have a central limit theorem (or a WIP) for a sequence of i.i.d. $\mathcal{X}$-valued random variables it is necessary that the variables be pre-Gaussian. Hence, to prove invariance principles for stationary sequences, we shall consider only pre-Gaussian variables.

DEFINITION 2.8. Let $\mathbb{G}(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{X}) = \mathbb{G}(\mathcal{X})$ be the set of pre-Gaussian random variables that are in $L^2(\Omega, \mathcal{X})$. For every $X \in \mathbb{G}(\mathcal{X})$, denote $\|X\|_{\mathbb{G}(\mathcal{X})} := \|X\|_{2,\mathcal{X}} + \|G(X)\|_{2,\mathcal{X}}$.

LEMMA 2.1. *Let $\mathcal{X}$ be a real separable Banach space. Then, for every pre-Gaussian variables X, Y, the variable $X + Y$ is pre-Gaussian and $\|G(X+Y)\|_{2,\mathcal{X}} \leq \|G(X)\|_{2,\mathcal{X}} + \|G(Y)\|_{2,\mathcal{X}}$. In particular, $(\mathbb{G}(\mathcal{X}), \|\cdot\|_{\mathbb{G}(\mathcal{X})})$, 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 8.23 of [31], hence its proof is omitted.

LEMMA 2.2. *Let $\mathcal{X}$ be a real separable Banach space. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, θ be an invertible bi-measurable transformation on Ω. Let $\mathcal{F}'$ be a sub-σ-algebra of $\mathcal{F}$. Let $X \in L^1(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{X})$ be pre-Gaussian. Then $\mathbb{E}(X|\mathcal{F}')$ is pre-Gaussian and for every $n \geq 0$, $X \circ \theta^n$ is pre-Gaussian. Moreover, $\|\mathbb{E}(X|\mathcal{F}')\|_{\mathbb{G}(\mathcal{X})} \leq \sqrt{2}\|X\|_{\mathbb{G}(\mathcal{X})}$ and $\|X \circ \theta^n\|_{\mathbb{G}(\mathcal{X})} \leq \sqrt{2}\|X\|_{\mathbb{G}(\mathcal{X})}$.*

LEMMA 2.3. *Let $\mathcal{X}$ be real separable Banach space. Let $(\mathcal{H}_n)_{n\geq 1}$ be a nondecreasing filtration and let $\mathcal{H}_\infty := \bigvee_{n\geq 1} \mathcal{H}_n$. For every $X \in \mathbb{G}(\mathcal{X})$, $\|\mathbb{E}(X|\mathcal{H}_n) - \mathbb{E}(X|\mathcal{H}_\infty)\|_{\mathbb{G}(\mathcal{X})} \underset{n\to\infty}{\longrightarrow} 0$.*

The proof is given in the Appendix.

From the above lemmas, we see that it will be very convenient to work in $\mathbb{G}(\mathcal{X})$ in order to obtain invariance principles for a sequence $(X \circ \theta^n)_{n\geq 0}$ under conditions involving terms of the type $(\mathbb{E}_0(X \circ \theta^n))_{n\geq 0}$.

Of course, in order to have tractable conditions it is necessary to be able to compute $\|X\|_{\mathbb{G}(\mathcal{X})}$.

Let $\mathcal{X} = L^p(S, \mu)$ $(1 \leq p \leq 2)$, for some σ-finite measure (recall that, then, $\mathcal{X}$ is of cotype 2). In this case, the following characterization of pre-Gaussian variables is part of the folklore. It is due to Vakhania [48] when μ is discrete (see [31], page 262, for a proof). It seems to be essentially due to Rajput [43] for a general σ-finite measure μ. We provide more details in the Appendix.

LEMMA 2.4. *Let $\mathcal{X} = L^p(S,\mu)$ $(1 \le p \le 2)$, for some σ-finite measure. Then $X(s) \in L^2(\Omega,\mathbb{P},\mathcal{X})$ is pre-Gaussian if and only if (X is centered 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*

$$(7)\qquad \begin{aligned}\|G(X)\|_2/C_p &\le \left(\int_S \left(\mathbb{E}|X(s)|^2\right)^{p/2}\mu(ds)\right)^{1/p}\\ &\le C_p\|G(X)\|_2 \qquad \forall X \in \mathbb{G}\left(L^p(\mu)\right).\end{aligned}$$

Hence, $\mathbb{G}(L^p(\mu))$ 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 $1 < p \le 2$); see Lemma E.2.

On another hand, when $\mathcal{X}$ is of type 2, in particular when $\mathcal{X}$ is 2-smooth, by Proposition 9.24 of [31], $\|\cdot\|_{\mathbb{G}(\mathcal{X})}$ is equivalent to $\|\cdot\|_{2,\mathcal{X}}$.

Hence, we infer that when $\mathcal{X} = L^p(S,\mu)$, for some $1 \le p < \infty$, there exists $C_p > 0$ such that for every $X \in \mathbb{G}(\mathcal{X})$,

$$(8)\qquad \begin{aligned}\|X\|_{\mathbb{G}(\mathcal{X})}/C_p &\le \left(\int_S \left(\mathbb{E}|X(s)|^2\right)^{p/2}\mu(ds)\right)^{1/p}\\ &\le C_p\|X\|_{\mathbb{G}(\mathcal{X})} \qquad \text{if } 1 \le p \le 2,\end{aligned}$$

$$(9)\qquad \begin{aligned}\|X\|_{\mathbb{G}(\mathcal{X})}/C_p &\le \left[\mathbb{E}\left(\int_S |X(s)|^p\mu(ds)\right)^{2/p}\right]^{1/2}\\ &\le C_p\|X\|_{\mathbb{G}(\mathcal{X})} \qquad \text{if } p \ge 2.\end{aligned}$$

Let us conclude that section with some results concerning the necessity of geometric conditions for the WIP, the ASIP or the BLIL, in the case of i.i.d. sequences. Those results motivate some of our restrictions in the next sections.

PROPOSITION 2.5. *Let $\mathcal{X}$ be a separable Banach space. Assume that every i.i.d. $\mathcal{X}$-valued $(X_n)_{n\ge 0}$ in $L^2(\mathcal{X})$, satisfies the WIP (resp., the ASIP, resp., the BLIL). Then $\mathcal{X}$ is of type 2 (resp., of type 2, resp., of type p for every $1 \le p < 2$).*

In the case of the WIP, the proposition follows from Theorem 10.5 of [31] (there is even a converse result there). In the case of the BLIL, the result follows from Pisier [42] (see his Remark 2 and the proposition, page 208). We have no reference for the case of the ASIP, so we provide a proof in the Appendix.

PROPOSITION 2.6. *Let $\mathcal{X}$ be a separable Banach space. Assume that every i.i.d. $\mathcal{X}$-valued and pre-Gaussian $(X_n)_{n\ge 0}$ satisfies the WIP (resp., the BLIL). Then $\mathcal{X}$ is of cotype 2.*

In the case of the WIP, the proposition follows from Theorem 10.7 of [31] (there is even a converse result there). We have no reference for the case of the BLIL, so we will provide a proof in the Appendix.

3. The martingale case. In this section, we give maximal inequalities and invariance principles for martingales with stationary differences $(d_n)_{n\geq 0}$. As mentioned in [7], there is no loss of generality in assuming that $d_n = d\circ\theta^n$, where θ is an invertible bi-measurable measure preserving transformation. Hence, we shall use the notation in the Introduction.

Let us mention that all the results of this section, except the ASIP in Proposition 3.3, 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 3.3) 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.

As mentioned, we use the notation from the Introduction.

We first state a maximal inequality that is related to the ASIP.

For every $X\in L^2(\Omega,\mathcal{F},\mathbb{P},\mathcal{X})$, we consider the following maximal function:

$$\mathcal{M}_2(X,\theta,\mathcal{X}) := \sup_{n\geq 1}\frac{|\sum_{k=0}^{n-1} X\circ\theta^k|_{\mathcal{X}}}{\sqrt{nL(L(n))}}, \tag{10}$$

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

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

PROPOSITION 3.1. *Let $\mathcal{X}$ be a Banach space. Assume either that $\mathcal{X}$ is a $(2, L)$-smooth Banach space or $\mathcal{X} = L^p(S,\mathcal{S},\mu)$ with μ σ-finite and $1\leq p\leq 2$). Then, for every $1<r<2$ there exists $C_r>0$ such that*

$$\|\mathcal{M}_2(d)\|_{r,\infty} \leq LC_r\|d\|_{\mathbb{G}(\mathcal{X})}. \tag{11}$$

REMARKS. Only the case $\mathcal{X} = L^p(S,\mathcal{S},\mu)$, $1\leq p<2$ is new here. The proposition is proved in [7] when $\mathcal{X}$ is $(2, L)$-smooth. We do not require θ to be ergodic. One may wonder whether the proposition holds when $\mathcal{X}$ is of cotype 2, or at least 2-convex, which is an open question.

For martingales with stationary and ergodic increments in 2-smooth Banach spaces (admitting a Schauder basis), the CLT has been obtained by Woyczyński [50], and the WIP by Dedecker–Merlevède–Pène [17] (see the proof of their Proposition 6). Rosiński [45] considered the case of general arrays of martingale increments a la Brown (in the p-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 $\mathcal{X} = L^1(S, \mathcal{S}, \mu)$, with μ σ-finite.

Hence, the CLT in the next proposition is only partly new, while the WIP seems to be new. Recall that $\mathbb{G}(\mathcal{F}_0, \mathcal{X})$ has been defined in Definition 2.8.

PROPOSITION 3.2. *Assume that θ is ergodic. Let $\mathcal{X}$ be a real separable Banach space that is either 2-smooth or of cotype 2. Let $d \in \mathbb{G}(\mathcal{F}_0, \mathcal{X})$ such that $\mathbb{E}_{-1}(d) = 0$. Then $(d \circ \theta^n)_{n\geq 0}$ satisfies the WIP of covariance $\mathcal{K}_d$, and there exists $C > 0$, such that*

$$\Big\| \max_{1\leq k\leq n} \big|S_k(d)\big|_{\mathcal{X}} \Big\|_2 \leq C n^{1/2} \|d\|_{\mathbb{G}(\mathcal{X})}. \tag{12}$$

REMARK. The constant C depends only on $\mathcal{X}$.

PROPOSITION 3.3. *Assume that θ is ergodic. Let $\mathcal{X}$ be either a 2-smooth Banach space or $\mathcal{X} = L^p(S, \mathcal{S}, \mu)$, for some $1 \leq p \leq 2$ and σ-finite μ. Let $d \in \mathbb{G}(\mathcal{F}_0, \mathcal{X})$ such that $\mathbb{E}_{-1}(d) = 0$. Then $(d \circ \theta^n)_{n\geq 0}$ satisfies the ASIP. Moreover,*

$$\limsup_n \frac{|S_n(d)|_{\mathcal{X}}}{\sqrt{nL(L(n))}} = \sup_{x^*\in\mathcal{X}^*, |x^*|_{\mathcal{X}^*}\leq 1} \big\|x^*(d)\big\|_2 \qquad \mathbb{P}\text{-a.s.} \tag{13}$$

REMARKS. 1. Since $(d \circ \theta^n)_{n\geq 0}$ satisfies the ASIP, it satisfies the CLIL, also. However, it follows from the proof that the ergodicity of θ is not necessary for the CLIL. As already mentioned, the CLIL also holds for stationary differences of reverse martingales.

2. Only the case $\mathcal{X} = L^p(S, \mathcal{S}, \mu)$, $1 \leq p < 2$ is new here. The case where $\mathcal{X}$ is 2-smooth has been obtained in [7]. As in Proposition 3.1, one may wonder whether Proposition 3.3 holds if $\mathcal{X}$ is of cotype 2 or at least 2-convex.

4. Maximal inequalities under projective conditions. In all of this section, we do not require θ to be ergodic.

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

Let $X \in L^2(\Omega, \mathcal{X})$. Define

$$\|X\|_{\mathrm{MW}_2} := \sum_{n\geq 0} \frac{\|\mathbb{E}_0(S_{2^n}(X))\|_{\mathbb{G}(\mathcal{X})}}{2^{n/2}}. \tag{14}$$

To have a better understanding of $\|\cdot\|_{\mathrm{MW}_2}$, recall that if $\mathcal{X}$ is of type 2 (in particular if $\mathcal{X}$ is 2-smooth), then $\|\cdot\|_{\mathbb{G}(\mathcal{X})} \leq C\|\cdot\|_{2,\mathcal{X}}$ and that if $\mathcal{X} = L^r(S, \mathcal{S}, \mu)$ with $1 \leq r \leq 2$ and μ σ-finite, we have (8).

In view of applications, let us mention the following easy fact based on the observation that $\|\mathbb{E}_0(S_n)\| \le \|\mathbb{E}_0(X)\| + \cdots + \|\mathbb{E}_0(X \circ \theta^{n-1})\|$. There exists $C > 0$ such that

$$\|X\|_{\mathrm{MW}_2} \le C \sum_{n\ge 1} \frac{\|\mathbb{E}_0(X \circ \theta^n)\|_{\mathbb{G}(\mathcal{X})}}{n^{1/2}}.$$

In particular, when $\mathcal{X} = L^p(S, \mu)$, for some $1 \le p < \infty$, using (8) and (9), we see that $\|X\|_{\mathrm{MW}_2} < \infty$ when $N_p(X) < \infty$, where $N_p(X)$ is defined by (3) if $1 \le p \le 2$ and by (4) if $p \ge 2$.

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. First, it involves the terms $(\mathbb{E}_{-2^k}(S_{2^k}))_{k\ge 0}$ which appear in the Maxwell–Woodroofe condition [notice that, by Lemma 2.2, $\|\mathbb{E}_{-2^k}(S_{2^k})\|_{\mathbb{G}(\mathcal{X})} \le \sqrt{2}\|\mathbb{E}_0(S_{2^k})\|_{\mathbb{G}(\mathcal{X})}$]. Second, for every $k \ge 0$, the sequence $(d_k \circ \theta^{2^{k+1}\ell})_{\ell\ge 0}$ defined below is a stationary sequence of martingale differences. The proposition makes use of the following maximal function. For every $X \in L^1(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{X})$, define

$$\mathcal{M}_1(X, \theta, \mathcal{X}) := \sup_{n\ge 1} \frac{|\sum_{k=0}^{n-1} X \circ \theta^k|_{\mathcal{X}}}{n}. \tag{15}$$

Recall that, by Hopf's dominated ergodic theorem (see Corollary 2.2, page 6 of [28]), applied to the real variable $|X|_{\mathcal{X}}$, we have

$$\big\|\mathcal{M}_1(X, \theta, \mathcal{X})\big\|_{1,\infty} \le \|X\|_{1,\mathcal{X}}.$$

PROPOSITION 4.1. *Let $X \in L^1(\Omega, \mathcal{F}_0, \mathbb{P}, \mathcal{X})$. For every $k \ge 0$, write $u_k := |\mathbb{E}_{-2^k}(S_{2^k})|_{\mathcal{X}}$ and $d_k := \mathbb{E}_{-2^k}(S_{2^k}) + \mathbb{E}_{-2^k}(S_{2^k}) \circ \theta^{2^k} - \mathbb{E}_{-2^{k+1}}(S_{2^{k+1}})$. Then, for every integer $d \ge 0$, we have $\mathbb{P}$-almost surely (with the convention $\sum_{k=0}^{-1} = 0$)*

$$\begin{aligned}\max_{1\le i\le 2^d} |S_i|_{\mathcal{X}} \le{}& \max_{1\le i\le 2^d} \left|\sum_{\ell=0}^{i-1} \big(X - \mathbb{E}_{-1}(X)\big) \circ \theta^\ell\right|_{\mathcal{X}} \\ &+ \sum_{k=0}^{d-1} \max_{1\le i\le 2^{d-k-1}} \left|\sum_{\ell=0}^{i-1} d_k \circ \theta^{2^{k+1}\ell}\right|_{\mathcal{X}} \\ &+ u_d + \sum_{k=0}^{d-1} \max_{0\le \ell\le 2^{d-1-k}-1} u_k \circ \theta^{2^{k+1}\ell}.\end{aligned}$$

In particular, there exists $C > 0$, *such that*

$$\begin{aligned}\mathcal{M}_2(X,\theta) \le C\Big(\sum_{k\ge 0}\frac{u_k}{2^{k/2}} + \sum_{k\ge 0}\frac{(\mathcal{M}_1(u_k^2,\theta^{2^{k+1}}))^{1/2}}{2^{k/2}} \\ + \mathcal{M}_2(X-\mathbb{E}_{-1}(X),\theta) + \sum_{k\ge 0}\frac{\mathcal{M}_2(d_k,\theta^{2^{k+1}})}{2^{k/2}}\Big).\end{aligned} \tag{16}$$

REMARK. That proposition is inspired by the works of Peligrad, Utev and Wu [40] and of Wu and Zhao [52].

COROLLARY 4.2. *Let $\mathcal{X}$ be Banach space that is either 2-smooth or of cotype 2. There exists $C > 0$ such that for every $X \in \mathbb{G}(\mathcal{X})$ and every integer $d \ge 0$, we have*

$$\Big\|\max_{1\le i\le 2^d} |S_i|_{\mathcal{X}}\Big\|_2 \le C2^{d/2}\bigg(\|X\|_{\mathbb{G}(\mathcal{X})} + \sum_{k=0}^{d} 2^{-k/2}\|\mathbb{E}_{-2^k}(S_{2^k})\|_{\mathbb{G}(\mathcal{X})}\bigg).$$

In particular, if $\|X\|_{\mathrm{MW}_2} < \infty$, *then*

$$\sup_{n\ge 1}\frac{\|\max_{1\le k\le n}|S_k(X)|_{\mathcal{X}}\|_2}{\sqrt{n}} \le C\|X\|_{\mathrm{MW}_2}. \tag{17}$$

PROPOSITION 4.3. *Let $\mathcal{X}$ be either a $(2,L)$-smooth Banach space or $\mathcal{X} = L^p(S,\mathcal{S},\mu)$, with $1 \le p \le 2$. Let $X \in L^2(\Omega,\mathcal{F}_0,\mathbb{P},\mathcal{X})$ be such that $\|X\|_{\mathrm{MW}_2} < \infty$. For every $1 < r < 2$, there exists a constant $C_r > 0$, such that*

$$\big\|\mathcal{M}_2(X)\big\|_{r,\infty,\mathcal{X}} \le C_r\|X\|_{\mathrm{MW}_2}. \tag{18}$$

REMARKS. The constant C_r depends on r and L if $\mathcal{X}$ is $(2,L)$-smooth and on r and p if $\mathcal{X} = L^p(S,\mathcal{S},\mu)$. Define $\|X\|_{H_2} := \sum_{n\ge 0}\|\mathbb{E}_0(X\circ\theta^n) - \mathbb{E}_{-1}(X\circ\theta^n)\|_{2,\mathcal{X}} < \infty$. Then, if $\|X\|_{H_2} < \infty$, (18) holds with $\|X\|_{H_2}$ in place of $\|X\|_{\mathrm{MW}_2}$. This follows from Theorem 2.10 of [7] when $\mathcal{X}$ is 2-smooth and, when $\mathcal{X} = L^r(S,\mathcal{S},\mu)$; the proof may be done exactly as the proof of Theorem 2.10 of [7], using (11).

5. WIP and ASIP under projective conditions. In all of this section we DO require θ to be ergodic.

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 $X \in \mathbb{G}(\mathcal{X},\mathcal{F}_0)$ be such that $\|X\|_{\mathrm{MW}_2} < \infty$. Then there exists $d \in \mathbb{G}(\mathcal{X},\mathcal{F}_0)$ with $\mathbb{E}_{-1}(d) = 0$ such that*

$$\Big\|\max_{1\le k\le n}\big|S_k(X) - S_k(d)\big|_{\mathcal{X}}\Big\|_2 = o(\sqrt{n}). \tag{19}$$

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

REMARK. The martingale approximation (19) 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 (19) is due to Gordin and Peligrad [24] and the WIP to Peligrad and Utev [39].

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

$$|S_n(X) - S_n(d)|_{\mathcal{X}} = o\big(\sqrt{nL(L(n))}\big) \qquad \mathbb{P}\text{-a.s.} \tag{20}$$

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

$$\begin{aligned}\limsup_n \frac{|S_n|_{\mathcal{X}}}{\sqrt{2nL(L(n))}} &= \sup_{x^*\in\mathcal{X}^*, |x^*|_{\mathcal{X}^*}\leq 1} \|x^*(d)\|_2 \\ &\leq 10\sqrt{2}\|X\|_{\mathrm{MW}_2} \qquad \mathbb{P}\text{-a.s.}\end{aligned} \tag{21}$$

REMARK. This result is new even when $\mathcal{X} = \mathbb{R}$. In view of the previous proposition, one may wonder whether the theorem holds true for Banach spaces of cotype 2 or, at least, for 2-convex Banach spaces.

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

$$\begin{aligned}\limsup_n \frac{|S_n|_{\mathcal{X}}}{\sqrt{2nL(L(n))}} &= \sup_{x^*\in\mathcal{X}^*, |x^*|_{\mathcal{X}^*}\leq 1} \big(\mathcal{K}(x^*, x^*)\big)^{1/2} \\ &\leq 10\sqrt{2}\|X\|_{\mathrm{MW}_2} \qquad \mathbb{P}\text{-a.s.}\end{aligned} \tag{22}$$

REMARK. Let $\mathcal{X}$ be either as in Theorem 5.2 or as in Theorem 5.3. Assume that $\|X\|_{H_2} := \sum_{n\geq 0} \|\mathbb{E}_0(X \circ \theta^n) - \mathbb{E}_{-1}(X \circ \theta^n)\|_{\mathbb{G}(\mathcal{X})} < \infty$. Then $(X \circ \theta^n)_{n\geq 0}$ satisfies the WIP and the ASIP of covariance operator $\mathcal{K}$ given by $\mathcal{K}(x^*, y^*) = \lim_n \operatorname{cov}(S_n(x^*(X)), S_n(y^*(X)))/n$, for every $x^*, y^* \in \mathcal{X}^*$. Moreover, (21) holds with $\|X\|_{H_2}$ in the right-hand side instead of $10\sqrt{2}\|X\|_{\mathrm{MW}_2}$. This is proved in Theorem 2.10 (see also Corollary 2.12) of [7] 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 4.3.

Peligrad and Utev [39] proved that the condition $\|X\|_{\mathrm{MW}_2} < \infty$ is optimal (in the sense below) for the CLT. Actually, their example gives also the optimality of the condition $\|X\|_{\mathrm{MW}_2} < \infty$ for the LIL; see [8] for a proof.

PROPOSITION 5.4. *Let $(a_n)_{n\geq 0}$ be a sequence of positive numbers with $a_n \to 0$ as $n \to \infty$. There exist a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with a transformation θ and a filtration $(\mathcal{F}_n)_{n\in\mathbb{Z}}$, as in the Introduction, such that there exists $X \in L^2(\Omega, \mathcal{F}_0, \mathbb{P})$ for which*

$$\sum_{n\geq 1} a_n \frac{\|\mathbb{E}_0(S_n(X))\|_2}{n^{3/2}} < \infty, \tag{23}$$

but $(S_n/\sqrt{n})$ is not stochastically bounded and

$$\limsup_n \frac{|S_n(X)|}{\sqrt{nL(L(n))}} = +\infty \qquad \mathbb{P}\text{-a.s.}$$

REMARK. It would be interesting to know whether the condition $\sum_{n\geq 1} \frac{\|\mathbb{E}_0(X\circ\theta^n)\|_2}{n^{1/2}} < \infty$ is also optimal. The optimality of the latter condition for the CLT has been recently investigated by Dedecker [13]. His arguments do not seem to apply for the LIL.

6. Examples.

6.1. *A direct example.* We now consider the case of ρ-mixing processes for which it is known that the Maxwell–Woodroofe condition is well-adapted; see, for instance, pages 14–15 in [36] or the proof of Lemma 1 (page 548) in [40].

Let $(X_n)_{n\in\mathbb{Z}}$ be a stationary $\mathcal{H}$-valued sequence. Define

$$\rho(n) = \rho(\mathcal{F}_{-\infty}^0, \mathcal{F}_n^\infty) \quad \text{and} \quad \psi(n) = \psi(\mathcal{F}_{-\infty}^0, \mathcal{F}_n^\infty), \tag{24}$$

where $\mathcal{F}_i^j = \sigma(X_i, \dots, X_j)$ and

$$\rho(\mathcal{A}, \mathcal{B}) = \sup\left\{\frac{\mathrm{Cov}(X, Y)}{\|X\|_2\|Y\|_2} : X \in L^2(\mathcal{A}), Y \in L^2(\mathcal{B})\right\};$$

$$\psi(\mathcal{A}, \mathcal{B}) = \sup\left\{\frac{|\mathbb{P}(A\cap B) - \mathbb{P}(A)\mathbb{P}(B)|}{\mathbb{P}(A)\mathbb{P}(B)} : A \in \mathcal{A}, B \in (\mathcal{B})\right\}.$$

It is well known that $\rho(n) \leq \psi(n)$; see, for instance, Proposition 3.11, page 76 of [5].

We have the following.

COROLLARY 6.1. *Assume that*

$$\sum_{n\geq 1} \rho(2^n) < \infty. \tag{25}$$

Then $\|X\|_{\mathrm{MW}_2} < \infty$.

REMARKS. The condition $\rho(2^n) = O(1/n^{1+\varepsilon})$ has been proven to be sufficient in [46] (for any $\varepsilon > 0$), when $\mathcal{H} = \mathbb{R}$. The sufficiency of (25) has been obtained very recently by Lin and Zhao [33], when $\mathcal{H} = \mathbb{R}$.

Sharipov [47] 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 OF COROLLARY 6.1. It suffices to prove that

$$\sum_n \frac{\|\mathbb{E}_0(S_{2^n}(X_0))\|_{2,\mathcal{H}}}{2^{n/2}} < \infty. \tag{26}$$

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

$$\|\mathbb{E}_0(S_{2^n}(X_0))\|_{2,\mathcal{H}}^2 = \sum_{i\geq 0} \mathbb{E}[(\mathbb{E}_0(S_{2^n}(Y_0^{(i)})))^2].$$

Now, it follows from the computations, page 15 of [36] combined with Lemma 3.4 of [38] that

$$\mathbb{E}[(\mathbb{E}_0(S_{2^n}(Y_0^{(i)})))^2] \leq C\mathbb{E}((Y_0^{(i)})^2)\left(\sum_{k=0}^{n} 2^{k/2}\rho(2^k)\right)^2.$$

Using that $\sum_{i\geq 0}(Y_0^{(i)})^2 = |X_0|_{\mathcal{H}}^2$, we see that (26) 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 (25). □

6.2. *Applications to the empirical process*. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, θ be an invertible bi-measurable measure preserving transformation on Ω and $\mathcal{F}_0 \subset \mathcal{F}$ a σ-algebra such that $\mathcal{F}_0 \subset \theta^{-1}(\mathcal{F}_0)$. Define a nondecreasing 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 $X_n := t \mapsto \mathbf{1}_{Y_n \leq t} - F(t)$, where $F(t) = \mathbb{P}(Y \leq t)$.

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

$$\int_0^\infty (1 - F(t))^p \mu(dt) + \int_{-\infty}^0 F(t)^p \mu(dt) < \infty, \tag{27}$$

which is satisfied whenever μ is finite.

Define F_μ by $F_\mu(x) = -\mu([x, 0[)$ if $x \leq 0$ and $F_\mu(x) = \mu([0, x[)$ if $x \geq 0$. Then, under (27), $X_0 \in L^2(\Omega, L^p(\mu))$ if and only if

$$\mathbb{E}(|F_\mu(Y_0)|^{2/p}) < \infty. \tag{28}$$

We want to understand the asymptotic behaviour of the process $F_n = S_n(X)/n$ [with values in $L^2(\Omega, \mathcal{F}_0, \mathbb{P}, L^p(\mathbb{R}, \mu))$], and more particularly of $D_{n,p}(\mu) := \|F_n\|_{p,\mu}$.

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

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

$$\check{\tau}_{\mu,p}(\mathcal{F}_0, Y_n) := \left\| \left(\int_{\mathbb{R}} |\mathbb{P}(Y_n \le t|\mathcal{F}_0) - F(t)|^p \mu(dt) \right)^{1/p} \right\|_2 \qquad \text{if } p \ge 2,$$

$$\check{\tau}_{\mu,r}(\mathcal{F}_0, Y_n) := \left(\int_{\mathbb{R}} \|\mathbb{P}(Y_n \le t|\mathcal{F}_0) - F(t)\|_2^p \mu(dt) \right)^{1/p} \qquad \text{if } 1 \le p < 2.$$

When $p \ge 2$, $\check{\tau}_{\mu,p}(\mathcal{F}_0, Y_n) = \tau_{\mu,p}(\mathcal{F}_0, Y_n)$, where $\tau_{\mu,p}(\mathcal{F}_0, Y_n)$ appears for instance in [14] (notice that our notation are slightly different).

Let us notice that both (27) and (28) are satisfied as soon as $\tau_{\mu,p}(\mathcal{F}_0, Y_0) < \infty$.

THEOREM 6.2. *Let $Y \in L^0(\Omega, \mathcal{F}_0, \mathbb{P})$ and $(S, \mathcal{S}, \mu)$ be a σ-finite measure space. Let $1 \le p < \infty$. Assume that*

$$\sum_{n \ge 0} \frac{\check{\tau}_{\mu,p}(\mathcal{F}_0, Y_n)}{n^{1/2}} < \infty.$$

Then $(X_n)_{n\ge1}$ satisfies the WIP and the ASIP. In particular, $(n^{1/2} D_{n,p})$ converges in law to an L^p-valued Gaussian variable, with covariance operator given by $\mathcal{K}_\mu(f, g)$ and

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

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

Let p' be the conjugate of p. We have

$$\mathcal{K}_\mu(f, g) = \lim_{n\to+\infty} \mathbb{E}\left(\int_S f(s) S_n(s) \mu(ds) \int_S g(t) S_n(t) \mu(dt) \right) \Big/ n \qquad \forall f, g \in L^{p'}(S),$$

and $\Lambda_{\mu,p} = \sup_{\|f\|_{p',\mu} \le 1} \Gamma_{\mu,p}(f)$ where $\Gamma_{\mu,p}(f) = \lim_n \| \int_S f(s) S_n(s) \mu(ds)\|_2 / \sqrt{n}$.

REMARK. Actually, if p' denotes the conjugate of p, we have $\Lambda_{\mu,p} = \sup_{\|f\|_{p',\mu}\le1} \Gamma_{\mu,p}(f)$ where $\Gamma_{\mu,p}(f) = \lim_n \| \int_S f(s) S_n(s) \mu(ds)\|_2/\sqrt{n}$. Since Theorem 6.2 is a straightforward application of the results of Section 5, we omit the proof.

In a series of paper, Dedecker and Merlevède obtained the WIP or the ASIP under conditions on the coefficients $\check{\tau}_{\mu,p}$, when $p \geq 2$. In [14], they studied the WIP and in [15] the ASIP. When $p > 2$, their results rely on a condition a la Gordin, hence yield to stronger conditions than ours. When $p = 2$, they use a very different approach and their results have different range of applicability.

When $p = 1$, Dédé [12] obtained the CLT under the same condition as above.

In order to apply Theorem 6.2, we shall further study the coefficients $\check{\tau}$, 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 $\tilde{\phi}$ and $\tilde{\alpha}$, as defined in Dedecker and Prieur [19]. For every $n \geq 1$, define

$$\tilde{\phi}(n) := \sup_{t \in \mathbb{R}} \|\mathbb{P}(Y_n \leq t | \mathcal{F}_0) - F(t)\|_\infty,$$

$$\tilde{\alpha}(n) := \sup_{t \in \mathbb{R}} \|\mathbb{P}(Y_n \leq t | \mathcal{F}_0) - F(t)\|_1.$$

LEMMA 6.3. *Assume that μ is finite. Let $p \geq 1$ and define $q := \max(2, p)$. For every $n \geq 1$, we have*

$$\tau_{\mu,p}(\mathcal{F}_0, Y_n) \leq \mu(\mathbb{R})^{1/p} \tilde{\phi}(n),$$

$$\tau_{\mu,p}(\mathcal{F}_0, Y_n) \leq \mu(\mathbb{R})^{1/p} \tilde{\alpha}(n)^{1/q}.$$

PROOF. The first inequality is obvious. The second one follows from the fact that for every $s \geq 1$, $\|\mathbb{P}(Y_n \leq t|\mathcal{F}_0) - F(t)\|_s \leq \|\mathbb{P}(Y_n \leq t|\mathcal{F}_0) - F(t)\|_1^{1/s}$. □

LEMMA 6.4. *Let $1 \leq p \leq 2$. For every $n \geq 1$, we have*

$$\check{\tau}_{\mu,p}(\mathcal{F}_0, Y_n) \leq \sqrt{2}\left(\int_0^\infty \bigl(F(t)(1 - F(t))\bigr)\right)^{p/2} \mu(dt)^{1/p} \tilde{\phi}(n)^{1/2}, \tag{29}$$

$$\check{\tau}_{\mu,p}(\mathcal{F}_0, Y_n) \leq \sqrt{2}\left(\int_0^{+\infty} (\min[\tilde{\alpha}_n, F(t)(1 - F(t))])^{p/2} \mu(dt)\right)^{1/p}. \tag{30}$$

PROOF. Notice that, for every $t \in \mathbb{R}$,

$$\|\mathbb{P}(Y_n \leq t|\mathcal{F}_0) - F(t)\|_2^2 \leq 2\tilde{\phi}(n)\bigl(1 - F(t)\bigr)F(t).$$

Hence, (29) follows.

Using that for every $t \in \mathbb{R}$,

$$\|\mathbb{P}(Y_n \leq t|\mathcal{F}_0) - F(t)\|_2^2 \leq \|\mathbb{P}(Y_n \leq t|\mathcal{F}_0) - F(t)\|_1 \leq \tilde{\alpha}(n), \quad \text{and}$$

$$\|\mathbb{P}(Y_n \leq t|\mathcal{F}_0) - F(t)\|_2^2 \leq 2F(t)\bigl(1 - F(t)\bigr),$$

we see that (30) holds. □

THEOREM 6.5. *Let $1 \le p \le 2$. Assume either of the following items*:

(i) $\int_0^\infty (F(t)(1-F(t)))^{p/2}\mu(dt) < \infty$ *and* $\sum_{n\ge 1} n^{-1/2}\tilde{\phi}(n)^{1/2} < \infty$.

(ii) $\mu = \lambda$ *the Lebesgue measure and* $\sum_{n\ge 1} n^{-1/2}(\int_0^{\tilde{\alpha}(n)} x^{p/2-1}Q(x)\,dx)^{1/p} < \infty$, *where* $Q(x) := \inf\{t \ge 0 : \mathbb{P}(|Y| > t) \le x\}$.

Then the conclusion of Theorem 6.2 *holds*.

REMARK. A better sufficient condition, in terms of $(\tilde{\alpha}(n))$ for the WIP has been obtained by Dedecker and Merlevède [16] when $p = 1$; see their Sections 4.4 and 5.

6.3. *Proof of Theorem* 6.5. The conclusion under (i) follows from Theorem 6.2 and Lemma 6.4. To prove item (ii), in view of Theorem 6.2 and Lemma 6.4, it suffices to prove that [notice that $F(t)(1-F(t)) \le \mathbb{P}(|Y| \ge |t|) = \mathbb{P}(|Y| > |t|)$ for λ-a.e. $t \in \mathbb{R}$]

$$\sum_{n\ge 1} \frac{1}{n^{1/2}}\left(\int_0^{+\infty} (\min[\tilde{\alpha}_n, (\mathbb{P}(|Y| > t))])^{p/2}\lambda(dt)\right)^{1/p} < \infty.$$

Now,

$$\begin{aligned} &\int_0^{+\infty} \min[\tilde{\alpha}_n, (\mathbb{P}(|Y| > t))]\,dt \\ &\qquad \le \tilde{\alpha}(n)^{p/2} Q(\tilde{\alpha}(n)) + \int_{Q(\tilde{\alpha}(n))}^{+\infty} (\mathbb{P}(|Y| > t))^{p/2-1}\,dt. \end{aligned} \tag{31}$$

Since Q is nonincreasing, we see that (ii) implies that

$$\sum_{n\ge 1} \frac{\tilde{\alpha}(n)^{1/2}(Q(\tilde{\alpha}(n)))^{1/p}}{n^{1/2}} < \infty,$$

hence, it remains to deal with the second term in the right-hand side of (31).

We have

$$\begin{aligned} &\int_{Q(\tilde{\alpha}(n))}^{+\infty} (\mathbb{P}(|Y| > t))^{p/2-1}\,dt \\ &\quad = \int_{Q(\tilde{\alpha}(n))}^{+\infty}\left(\int_0^1 \frac{p}{2}x^{p/2-1}\mathbf{1}_{\{x \le \mathbb{P}(|Y|>t)\}}\,dx\right)dt \\ &\quad \le \int_0^{\tilde{\alpha}(n)} \frac{p}{2}x^{p/2-1}\left(\int_0^{Q(x)} dt\right)dx \\ &\quad = \int_0^{\tilde{\alpha}(n)} \frac{p}{2}x^{p/2-1}Q(x)\,dx, \end{aligned}$$

and the proof is complete. □

APPENDIX A: PROOF OF THE RESULTS OF SECTION 2

A.1. Proof of Lemma 2.1. Let $X, Y \in \mathbb{G}(\mathcal{X})$. Consider the Banach space $\mathcal{C} := \mathcal{X} \times \mathcal{X}$ with norm $|(x, y)|_{\mathcal{C}} := (|x|_{\mathcal{X}}^2 + |y|_{\mathcal{X}}^2)^{1/2}$. Let us prove that $(X, Y) \in \mathbb{G}(\mathcal{C})$. 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 $\mathcal{C}$. Now, for every $x^*, y^* \in \mathcal{X}^*$, we have

$$\begin{aligned}\mathbb{E}\big(\big(x^*(X) + y^*(Y)\big)^2\big) &\le 2\mathbb{E}\big[\big(x^*\big(G(X)\big)\big)^2 + \big(y^*\big(G(Y)\big)\big)^2\big] \\ &= 2\mathbb{E}\big(\big(x^*\big(G(X)\big) + y^*\big(G(Y)\big)\big)^2\big).\end{aligned}$$

Hence, by Lemma 9.23 of [31], $(X, Y) \in \mathbb{G}(\mathcal{C})$. Let (U, V) be a Gaussian variable with values in $\mathcal{C}$ with same covariance operator as (X, Y). Clearly, $U + V$ is Gaussian and has same covariance operator as $X + Y$. Hence, $X + Y$ is pre-Gaussian and we may take $G(X + Y) = U + V$. Similarly, we may take $G(X) = U$ and $G(Y) = V$. Now,

$$\begin{aligned}\big\|G(X + Y)\big\|_{2,\mathcal{X}} &= \|U + V\|_{2,\mathcal{X}} \\ &\le \|U\|_{2,\mathcal{X}} + \|V\|_{2,\mathcal{X}} = \|X\|_{2,\mathcal{X}} + \|Y\|_{2,\mathcal{X}}.\end{aligned}$$

Hence, $\|\cdot\|_{\mathbb{G}(\mathcal{X})}$ is a norm on $\mathbb{G}(\mathcal{X})$.

Let us prove that $\mathbb{G}(\mathcal{X})$ is a Banach space.

Let $(X_n)_{n\ge1}$ be Cauchy in $(\mathbb{G}(\mathcal{X}), \|\cdot\|_{\mathbb{G}(\mathcal{X})})$. Hence, $(X_n)_{n\ge1}$ 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 pre-Gaussian and that $(X_n)_{n\ge1}$ admits a subsequence converging to X for $\|\cdot\|_{\mathbb{G}(\mathcal{X})}$. By assumption, there exists a subsequence $(X_{n_k})_{k\ge1}$ such that $\|X_{n_k} - X_{n_{k+1}}\|_{\mathbb{G}(\mathcal{X})} \le 2^{-k}$. Then $X = -X_{n_1} + \sum_{k\ge1} X_{n_k} - X_{n_{k+1}}$ with convergence in $L^2(\Omega, \mathcal{X})$.

Extending our probability space, if necessary, we may assume that there exists a sequence $(G_k)_{k\ge0}$ of independent Gaussian variables taking values in $\mathcal{X}$, such that $G_0 = G(X_{n_1})$ and for every $k \ge 1$, $G_k = G(X_{n_{k+1}} - X_{n_k})$. Then $G := \sum_{k\ge0} 2^{k/2} G_k$ defines a Gaussian variable. Moreover, for every $x^* \in \mathcal{X}^*$, we have, using Cauchy–Schwarz,

$$\begin{aligned}\mathbb{E}\big(x^*(X)^2\big) &= \mathbb{E}\bigg[\bigg(x^*(-X_{n_1}) + \sum_{k\ge1} x^*(X_{n_{k+1}} - X_{n_k})\bigg)^2\bigg] \\ &\le 2\mathbb{E}\big(\big(x^*(-X_{n_1})\big)^2\big) + \sum_{k\ge1} 2^k \mathbb{E}\big(\big(x^*(X_{n_{k+1}} - X_{n_k})^2\big)\big) \\ &= 2\mathbb{E}\bigg[\bigg(x^*\bigg(\sum_{k\ge0} 2^{k/2} G_k\bigg)\bigg)^2\bigg].\end{aligned}$$

It follows from Lemma 9.23 of [31] that X is pre-Gaussian. By a similar argument, using the second half of Lemma 9.23 of [31], we see that $\mathbb{E}(|G(X - X_{n_m})|^2) \to 0$ as $m \to +\infty$, and the proof is complete. □

A.2. Proof of Lemma 2.3. Let $x^* \in \mathcal{X}^*$. Clearly, 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 2.2, we have

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

Since $X_n - X$ is (clearly) pre-Gaussian, we infer that

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

Then it follows from the discussion pages 73–74 of [31] that $(G(X_n - X))_{n\geq 1}$ 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 $\|x^*(G(X_n - X))\|_2 = \|x^*(X_n - X)\|_2 \underset{n\to\infty}{\longrightarrow} 0$].

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 $G(\tilde{X}_{n_\varepsilon} - X)$ is smaller than ε, and it follows from the last assertion of Lemma 3.2 of [31], that there exists a universal $C > 0$ such that $\|G(\tilde{X}_{n_\varepsilon} - X)\|_{2,\mathcal{X}} \leq C\varepsilon^2$, and the proof is complete. □

A.3. Proof of Lemma 2.4. Let $X(s) \in L^2(\Omega, \mathbb{P}, L^p(S, \mu))$ be pre-Gaussian. Hence, there exists a Gaussian variable W on $(\Omega, \mathcal{F}, \mathbb{P})$ with values in $L^p(S, \mu)$ with same covariance operator than X. By Theorem 3.1 of Rajput [43], we may see W as a Gaussian process $(W(s))_{s\in S}$ whose paths are $\mathbb{P}$-a.s. in $L^p(S, \mu)$. Then

$$\begin{aligned}
\infty > \big\|G(X)\big\|_{2,L^p(\mu)} &= \|W\|_{2,L^p(\mu)} \geq C_p\|W\|_{p,L^p(\mu)} \\
&= C_p\bigg(\int_S \mathbb{E}\big(\big|W(s)\big|^p\big)\mu(ds)\bigg)^{1/p} \\
&= \tilde{C}_p\bigg(\int_S \big(\mathbb{E}\big(\big|W(s)\big|^2\big)\big)^{p/2}\mu(ds)\bigg)^{1/p} \\
&= \tilde{C}_p\bigg(\int_S \big(\mathbb{E}\big(\big|X(s)\big|^2\big)\big)^{p/2}\mu(ds)\bigg)^{1/p},
\end{aligned}$$

the reverse inequality may be proved similarly.

The fact that a centered X such that $\int_S (\mathbb{E}(|X(s)|^2))^{p/2}\mu(ds) < \infty$ is pre-Gaussian follows from Lemma 5.1 of [43]. □

A.4. Proof of Proposition 2.5: The ASIP case. Let $(X_n)_{n\geq 0}$ be i.i.d. variables in $L^2(\mathcal{X})$. By assumption, they satisfy the ASIP. Hence, there exists i.i.d. Gaussian variables $(W_n)_{n\geq 0}$, such that

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

Let $x^* \in \mathcal{X}^*$. By the law of the iterated logarithm (in the real case), $\mathbb{E}((x^*(X_0))^2) = \mathbb{E}((x^*(W_0))^2)$. In particular, X_0 is pre-Gaussian. Then we conclude thanks to Proposition 9.24 of [31]. □

A.5. Proof of Proposition 2.6: The BLIL case. Let $(X_n)_{n\geq 0}$ be i.i.d. pre-Gaussian variables taking values in $\mathcal{X}$. By assumption, they satisfy the BLIL. Let $1 \leq p < 2$. It follows that $|X_n|_{\mathcal{X}}/n^{1/p} \underset{n\to+\infty}{\longrightarrow} 0$ $\mathbb{P}$-a.s. Hence, by the Borel–Cantelli lemma, $X_0 \in L^p(\mathcal{X})$. Then the result follows from the proof of Proposition 9.25 of [31]. □

APPENDIX B: PROOF OF THE MARTINGALE RESULTS

B.1. Proof of Proposition 3.1. This is just Proposition 3.3 of [7] when $\mathcal{X}$ is 2-smooth. Assume that $\mathcal{X} = L^p(S)$, $p \geq 1$. It suffices to prove the result when $d \in L^p(S, L^2(\Omega, \mathcal{F}_0))$, otherwise $K_p(d) = +\infty$. There exists a sequence of step functions $(d_n)_{n\geq 1}$ converging in $L^p(S, L^2(\Omega, \mathcal{F}_0))$ to d. We may write $d_n(s,\omega) = \sum_{k=1}^{m_n} f_{k,n}(\omega)\mathbf{1}_{A_{k,n}}(s)$, where $A_{k,n} \in \mathcal{S}$ and $f_{k,n} \in L^2(\Omega, \mathbb{P})$. Let $\tilde{d}_n := \sum_{k=1}^{m_n}(f_{k,n} - \mathbb{E}_{-1}(f_{k,n}))\mathbf{1}_{A_{k,n}}$. Then $(\tilde{d}_n)_{n\geq 1}$ converges to d in $L^p(S, L^2(\Omega, \mathcal{F}_0))$ as well [hence also in $L^2(\Omega, L^r(S))$, by Lemma E.2] and for every $s \in S$, $\tilde{d}_n(s,\cdot)$ is a real-valued martingale difference in $L^2(\Omega, \mathcal{F}_0, \mathbb{P})$. Hence, applying Proposition 3.1 to the (2, 1)-smooth Banach space $\mathbb{R}$, we obtain that there exists $C_p > 0$ such that for every $s \in S$,

$$\big\|\mathcal{M}_2\big(\tilde{d}_n(s,\cdot)\big)\big\|_{p,\infty} \leq C_p \big\|\tilde{d}_n(s,\cdot)\big\|_2. \tag{32}$$

Notice that $\mathcal{M}_2(\tilde{d}_n, L^p(S)) \leq (\int_S (\mathcal{M}_2(\tilde{d}_n(s,\cdot),\mathbb{R}))^p \,d\mu(s))^{1/p}$. Writing $\varphi(s,\cdot) = \mathcal{M}_2(\tilde{d}_n(s,\cdot),\mathbb{R})$, it follows from Lemma E.2 that

$$\big\|\mathcal{M}_2\big(\tilde{d}_n, L^p(S)\big)\big\|_{2,\infty} \leq C_p\bigg(\int_S \|\varphi(s,\cdot)\|_{r,\infty}^p \,d\mu(s)\bigg)^{1/p}.$$

Then we infer from (32) that

$$\big\|\mathcal{M}_2\big(\tilde{d}_n, L^p(S)\big)\big\|_{r,\infty} \leq C_p\bigg(\int_S \|\tilde{d}_n(s)\|_2^p \,d\mu(s)\bigg)^{1/p}.$$

The desired result then follows by letting $n \to \infty$ (approximate first $\mathcal{M}_2$ by a supremum over a finite set of integers and use the monoton convergence theorem). □

B.2. Proof of Proposition 3.2. We shall first prove (12) which will allow us to derive the required tightness for the WIP. By Doob's maximal inequality for submartingales, we have

$$\Big\|\max_{1\leq k\leq n}\big|S_k(d)\big|_{\mathcal{X}}\Big\|_2^2 \leq 2\big\||S_n(d)|_{\mathcal{X}}\big\|_2^2.$$

When $\mathcal{X}$ is 2-smooth, (12) then follows from (6) and the fact that, on Type 2 Banach spaces, the norms $\|\cdot\|_{\mathbb{G}(\mathcal{X})}$ and $\|\cdot\|_{2,\mathcal{X}}$ are equivalent, by Proposition 9.24 of [31].

Assume now that $\mathcal{X}$ has cotype 2. Since d is pre-Gaussian, so is $S_n(d)$. Moreover, by orthogonality of real-valued martingale increments, we see that $G(S_n(d)/\sqrt{n}) = G(d)$. Since $\mathcal{X}$ has cotype 2, by Proposition 9.25 of [31],

$$\|S_n(d)\|_{2,\mathcal{X}} \le C\|G(S_n(d))\|_{2,\mathcal{X}} = C\sqrt{n}\|G(d)\|_{2,\mathcal{X}} \le C\sqrt{n}\|d\|_{\mathbb{G}(\mathcal{X})},$$

and (12) follows.

Let us prove the WIP. Let us recall the definition of tightness required here.

Let $X \in L^0(\Omega, \mathcal{F}_0, \mathbb{P}, \mathcal{X})$. 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 $((T_{n,t})_{0\le t\le 1})_{n\ge 0}$ as a process taking values in $C([0,1], \mathcal{X})$, the Banach space of continuous functions from [0, 1] to $\mathcal{X}$.

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

$$\mathbb{P}\big((T_{n,t})_{0\le t\le 1} \in \kappa\big) \ge 1 - \varepsilon \qquad \forall n \ge 0.$$

Let $\mathcal{X}$ be either 2-smooth or of cotype 2. Let $d \in \mathbb{G}(\mathcal{X})$ with $\mathbb{E}_{-1}(d) = 0$. Let us prove the tightness of $((T_{n,t}(d))_{0\le t\le 1})_{n\ge 1}$ 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 [23].

LEMMA B.1. *Let (Γ, δ) be a separable complete metric space endowed with its Borel σ-algebra. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(Z_n)_{n\ge 1}$ be a sequence of random variables on Ω taking values in Γ. Assume that, for every $\varepsilon > 0$, there exist $n_0 \ge 1$ and random variables $(Z_n^\varepsilon)_{n\ge n_0}$ such that*:

(i) $(Z_n^\varepsilon)_{n\ge n_0}$ *is tight*;
(ii) $\sup_{n\ge n_0} \mathbb{E}(\delta(Z_n, Z_n^\varepsilon)) < \varepsilon$.

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

Since $\mathcal{X}$ is separable, $\sigma(d)$ (the σ-algebra generated by d) is countably generated and there exists an increasing filtration $(\mathcal{G}_m)_{m\ge 1}$ such that $\mathcal{G}_m$ is finite for every $m \ge 1$ and $\sigma(d) = \bigvee_{m\ge 1} \mathcal{G}_m$. For every $m \ge 1$, let $d_m := \mathbb{E}(d|\mathcal{G}_m)$. Since $\mathcal{G}_m$ is finite, there exists $A_{1,m}, \dots, A_{k_m,m} \in \mathcal{G}_m$ and $x_{1,m}, \dots, x_{k_m,m} \in \mathcal{X}$ such that $d_m = \sum_{1\le k\le k_m} x_k \mathbf{1}_{A_{k,m}}$. By Lemma 2.3, $(d_m)_{m\ge 1}$ converges in $\mathbb{G}(\mathcal{X})$ to d. Hence, writing $\tilde{d}_m := d_m - \mathbb{E}_{-1}(d_m)$ and using Lemma 2.2, $(\tilde{d}_m)_{m\ge 1}$ converges in $\mathbb{G}(\mathcal{X})$ to d.

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

Now, by (12),

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

and the tightness of $((T_{n,t}(d))_{0\le t\le 1})_{n\ge 0}$ in $C([0,1],\mathcal{X})$ follows from Lemma B.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 $0=t_0<\cdots<t_m=1$, $((T_{n,t_i}-T_{n,t_{i-1}})_{1\le i\le m})_{n\ge 1}$ converges in law to $(W_{t_i}-W_{t_{i-1}})_{1\le i\le m}$, where $(W_t)_{0\le t\le 1}$ is a Brownian motion with covariance operator $\mathcal{K}_d$. Using tightness again (and the Cramer–Wold device), it suffices to prove that for any $0=t_0<\cdots<t_m=1$ and any $x_1^*,\dots,x_m^*\in\mathcal{X}^*$, $\sum_{i=1}^m x_i^*(T_{n,t_i}-T_{n,t_{i-1}})$ 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 $(X_{n,j})_{1\le j\le k_n}$ be (real valued) martingale differences for every $n\ge 1$. Assume that there exists $\sigma\ge 0$ such that*:

(i) $\max_{1\le j\le k_n}|X_{n,j}|\xrightarrow{\mathbb{P}}0$;

(ii) $\sum_{1\le j\le k_n}X_{n,j}^2\xrightarrow{\mathbb{P}}\sigma^2$;

(iii) $\sup_{n\ge 1}\mathbb{E}(\max_{1\le j\le k_n}X_{n,j}^2)<\infty$.

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

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

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

$$\frac{1}{n}\sum_{j=[nt_{i-1}]}^{[nt_i]}\big(x_i^*(d)\big)^2\circ\theta^j\underset{n\to\infty}{\longrightarrow}(t_i-t_{i-1})\mathbb{E}\big(x_i^*(d)^2\big)\qquad\mathbb{P}\text{-a.s.},$$

hence in probability. Hence, the proof is complete. □

B.3. Proof of Proposition 3.3. Let us prove the CLIL. Notice that $\mathbb{G}_0(\mathcal{X}):=\{d\in\mathbb{G}(\mathcal{X},\mathcal{F}_0):\mathbb{E}_{-1}(d)=0\}$ is a closed subspace of $\mathbb{G}(\mathcal{X})$. By (11) and Proposition E.1, the set of $d\in\mathbb{G}_0(\mathcal{X})$, such that $(d\circ\theta^n)_{n\ge 0}$ satisfies the CLIL is closed in $\mathbb{G}_0(\mathcal{X})$. Then the CLIL follows by approximating any $d\in\mathbb{G}_0(\mathcal{X})$ by a martingale difference with values in a finite dimensional Banach space as in the proof of Proposition 3.2.

Then (13) follows from a result of Kuelbs [29] (see, e.g., Proposition D of [7]) 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 [3] whose proof may be done similarly.

THEOREM B.3. *Let $\mathcal{X}$ be a real separable Banach space. Assume that θ 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*

$$S_n(x^*(X)) - S_n(Z) = o\big(\sqrt{nL(L(n))}\big) \qquad \mathbb{P}\text{-a.s.}, \tag{33}$$

$$\|S_n(x^*(X)) - S_n(Z)\|_2 = o(\sqrt{n}). \tag{34}$$

Then, for every $x^*, y^* \in \mathcal{X}^*$, $\mathcal{K}(x^*, y^*) := \lim_{n\to\infty} \frac{\operatorname{cov}(x^*(S_n(X)), y^*(S_n(X)))}{n}$ exists. Assume moreover that $\mathcal{K}$ 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

$$S_1 = X - \mathbb{E}_{-1}(X) + \mathbb{E}_{-1}(X) = \big(X - \mathbb{E}_{-1}(X)\big) + \mathbb{E}_{-1}(S_1)$$

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

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

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

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

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

Hence,

$$\begin{aligned}
\max_{1\leq i\leq 2^{d+1}} |S_i|_{\mathcal{X}} \leq{}& \max_{1\leq i\leq 2^{d+1}} \left|\sum_{\ell=0}^{i-1}\big(X - \mathbb{E}_{-1}(X)\big) \circ \theta^\ell\right|_{\mathcal{X}} \\
&+ \max_{1\leq j\leq 2^d} \big|\mathbb{E}_{-1}(X)\big|_{\mathcal{X}} \circ \theta^{2j-2} \\
&+ \max_{1\leq j\leq 2^d} \left|\sum_{\ell=1}^{j}\big(\mathbb{E}_{-1}(X) + \mathbb{E}_{-1}(X) \circ \theta\big) \circ \theta^{2\ell}\right|_{\mathcal{X}}.
\end{aligned} \tag{35}$$

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

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

Notice then that we have

$$\tilde{S}_n = \sum_{\ell=0}^{n-1} \big(\mathbb{E}_{-1}(X) + \mathbb{E}_{-1}(X) \circ \theta\big) \circ \theta^{2\ell},$$

$$\tilde{\mathbb{E}}_{-2^k}(\tilde{S}_{2^k}) = \mathbb{E}_{-2^{k+1}}(S_{2^{k+1}}) \quad \text{and}$$

$$\tilde{X} - \tilde{\mathbb{E}}_{-1}(\tilde{X}) = \mathbb{E}_{-1}(S_1) + \mathbb{E}_{-1}(S_1) \circ \theta - \mathbb{E}_{-2}(S_2).$$

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

$$\begin{aligned} &\max_{1 \le i \le 2^d} |\tilde{S}_i|_{\mathcal{X}} \\ &\quad \le \big|\mathbb{E}_{-2^{d+1}}(S_{2^{d+1}})\big|_{\mathcal{X}} + \sum_{k=1}^{(d+1)-1} \max_{0 \le \ell \le 2^{(d+1)-1-k}-1} \big|\mathbb{E}_{-2^k}(S_{2^k})\big|_{\mathcal{X}} \circ \theta^{2^{k+1}\ell} \\ &\qquad + \sum_{k=1}^{(d+1)-1} \max_{1 \le i \le 2^{(d+1)-k-1}} \Bigg|\sum_{\ell=0}^{i-1} \big[\mathbb{E}_{-2^k}(S_{2^k}) + \mathbb{E}_{-2^k}(S_{2^k}) \circ \theta^{2^k} \\ &\qquad - \mathbb{E}_{-2^{k+1}}(S_{2^{k+1}})\big] \circ \theta^{2^{k+1}\ell}\Bigg|_{\mathcal{X}}. \end{aligned} \tag{36}$$

Then the result follows by combining (35) and (36). $\square$

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

$$\max_{0 \le \ell \le 2^{d-1-k}-1} \big|\mathbb{E}_{-2^k}(S_{2^k})\big|_{\mathcal{X}} \circ \theta^{2^{k+1}\ell} \le \bigg(\sum_{0 \le \ell \le 2^{d-1-k}-1} \big|\mathbb{E}_{-2^k}(S_{2^k})\big|_{\mathcal{X}}^2 \circ \theta^{2^{k+1}\ell}\bigg)^{1/2}.$$

Hence, using that θ preserves $\mathbb{P}$, we infer that

$$\Big\|\max_{0 \le \ell \le 2^{d-1-k}-1} \big|\mathbb{E}_{-2^k}(S_{2^k})\big|_{\mathcal{X}} \circ \theta^{2^{k+1}\ell}\Big\|_2 \le 2^{(d-1-k)/2} \|\mathbb{E}_{-2^k}(S_{2^k})\|_{2,\mathcal{X}}.$$

Applying (12) to (the martingale difference) $d = X - \mathbb{E}_{-1}(X)$, we see that

$$\Bigg\|\max_{1 \le i \le 2^d} \bigg|\sum_{\ell=0}^{i-1} \big(X - \mathbb{E}_{-1}(X)\big) \circ \theta^{\ell}\bigg|_{\mathcal{X}}\Bigg\|_{\mathbb{G}(\mathcal{X})} \le C\big(\|X\|_{\mathbb{G}(\mathcal{X})} + \|\mathbb{E}_{-1}(X)\|_{\mathbb{G}(\mathcal{X})}\big).$$

Similarly, we may apply (12) with $d_k = \mathbb{E}_{-2^k}(S_{2^k}) + \mathbb{E}_{-2^k}(S_{2^k}) \circ \theta^{2^k} - \mathbb{E}_{-2^{k+1}}(S_{2^{k+1}})$ (and $\theta^{2^{k+1}}$ instead of θ). To conclude, we just notice that, by

Lemma 2.2, $\|X - \mathbb{E}_{-1}(X)\|_{\mathbb{G}(\mathcal{X})} \le (1+\sqrt{2})\|X\|_{\mathbb{G}(\mathcal{X})}$ and that $\|[\mathbb{E}_{-2^k}(S_{2^k}) + \mathbb{E}_{-2^k}(S_{2^k}) \circ \theta^{2^k} - \mathbb{E}_{-2^{k+1}}(S_{2^{k+1}})] \circ \theta^{2^{k+1}\ell}\|_{\mathbb{G}(\mathcal{X})} \le (1+\sqrt{2})^2\|\mathbb{E}_{-2^k}(S_{2^k})\|_{\mathbb{G}(\mathcal{X})}$. □

C.3. Proof of Proposition 4.3. By Hopf's maximal inequality, for every $X \in L^1(\Omega, \mathbb{R})$, and every measure preserving θ

$$\|\mathcal{M}_1(X,\theta)\|_{1,\infty} \le \|X\|_1.$$

Then the proposition follows from (16) combined (11). □

APPENDIX D: PROOF OF THE LIMIT THEOREMS UNDER PROJECTIVE CONDITIONS

Before doing the proof, let us give general facts about $\|\cdot\|_{\mathrm{MW}_2}$, that will be used in the sequel.

Define $\mathrm{MW}_2 := \{X \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}, \mathcal{X}) : \|X\|_{\mathrm{MW}_2} < \infty\}$. Then $(\mathrm{MW}_2, \|\cdot\|_{\mathrm{MW}_2})$ is a Banach space.

For every $X \in L^1(\Omega, \mathcal{F}_0, \mathbb{P}, \mathcal{X})$, define $QX = \mathbb{E}_0(X \circ \theta)$. Notice that $Q^n(X) = \mathbb{E}_0(X \circ \theta^n)$. Then clearly Q is a contraction of $L^2(\Omega, \mathcal{F}_0, \mathcal{X})$ and, by Lemma 2.2, Q is power bounded on $\mathbb{G}(\mathcal{X})$, that is, for every $X \in \mathbb{G}(\mathcal{X})$, $\sup_{n\ge 1}\|Q^n X\|_{\mathbb{G}(\mathcal{X})} \le C\|X\|_{\mathbb{G}(\mathcal{X})}$, for some universal $C > 0$.

Now, we see that

$$\|X\|_{\mathrm{MW}_2} = \sum_{n\ge 0} \frac{\|\sum_{k=0}^{2^n-1} Q^k X\|_{\mathbb{G}(\mathcal{X})}}{2^{n/2}}.$$

Hence, Q is power bounded on MW_2.

Writing $V_n := I + \cdots + Q^{n-1}$ and using that $\|V_n V_k X\|_{\mathbb{G}(\mathcal{X})} \le C\min(k\|V_n\|_{2,\mathcal{X}}, n\|V_k X\|_{\mathbb{G}(\mathcal{X})})$, we see that, for every $X \in \mathrm{MW}_2$,

$$\frac{\|V_{2^n} X\|_{\mathrm{MW}_2}}{2^n} \le C\left(\frac{\|V_{2^n}\|_{\mathbb{G}(\mathcal{X})}}{2^{n/2}} + \sum_{k\ge n+1} \frac{\|V_{2^k} X\|_{\mathbb{G}(\mathcal{X})}}{2^{k/2}}\right) \underset{n\to+\infty}{\longrightarrow} 0. \tag{37}$$

In particular, for every $m \ge 1$, taking n such that $2^n \le m < 2^{n+1}$, we have $\|V_m X\|_{\mathrm{MW}_2} \le C\sum_{k=0}^n \|V_{2^k}\|_{\mathrm{MW}_2} = o(2^n) = o(m)$.

In particular, we see that Q is mean ergodic on MW_2 and has no nontrivial fixed point (see, e.g., Theorem 1.3, page 73 of [28]), that is,

$$\mathrm{MW}_2 = \overline{(I-Q)\mathrm{MW}_2}^{\mathrm{MW}_2}. \tag{38}$$

D.1. Proof of Proposition 5.1 and Theorem 5.2. In both results, $\mathcal{X}$ is a Banach space of cotype 2. Let $X \in (I-Q)\mathrm{MW}_2$. Let $Y \in \mathrm{MW}_2$ be the unique (notice that Q has no fixed point on MW_2) solution to $X = (I-Q)Y$. Then one may define

$$\mathcal{D}(X) := Y - \mathbb{E}_{-1}(Y) = Y - QY \circ \theta^{-1}.$$

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

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

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

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

Now, it follows from the proof of Proposition 4.1 [combined with (39) applied to the martingales with stationary increments that appear in the proof] that there exists $D > 0$ such that for every $d \ge 0$,

(41) $$\|G(S_{2^d}(X))\|_{2,\mathcal{X}} \le D2^{d/2}\left(\|G(X)\|_{2,\mathcal{X}} + \sum_{k=0}^{d} 2^{-k}\|G(\mathbb{E}_0(S_{2^k}(X)))\|_{2,\mathcal{X}}\right).$$

Notice that

$$\|S_{2^d}(QY - QY \circ \theta^{-1})\|_{\mathbb{G}(\mathcal{X})} \le \|QY \circ \theta^{-1}\|_{\mathbb{G}(\mathcal{X})} + \|QY \circ \theta^{2^d-1}\|_{\mathbb{G}(\mathcal{X})} = o(2^{d/2})$$

and that

$$\|G(S_{2^d}(\mathcal{D}(X)))\|_{2,\mathcal{X}} \le \|G(S_{2^d}(X))\|_{2,\mathcal{X}} + \|G(S_{2^d}(QY - QY \circ \theta^{-1}))\|_{2,\mathcal{X}}.$$

Combining this with (41), (40) and (39) and letting $d \to \infty$, we infer that

$$\|\mathcal{D}(X)\|_{\mathbb{G}(\mathcal{X})} \le C\|X\|_{\mathrm{MW}_2}.$$

Hence, we may extend our linear operator $\mathcal{D}$ continuously to $\overline{(I - Q)\mathrm{MW}_2}^{\mathrm{MW}_2} = \mathrm{MW}_2$. Notice that $\mathcal{D}$ takes values in $\mathbb{G}_0(\mathcal{X}) = \{Z \in \mathbb{G}(\mathcal{X}, \mathcal{F}_0) : \mathbb{E}_{-1}(Z) = 0\}$.

Let us prove Proposition 5.1. By Corollary 4.2 and (12), there exists $C > 0$ such that

$$\left\|\max_{1\le k\le n}|S_k(X) - S_k(\mathcal{D}(X))|_{\mathcal{X}}\right\|_2 \le C\sqrt{n}\|X\|_{\mathrm{MW}_2}.$$

By linearity of $\mathcal{D}$ [and of $X \mapsto S_k(X)$] it then suffices to prove (19) for a set of X's that is dense in MW_2, in particular for $X \in (I - Q)\mathrm{MW}_2$. But if $X = (I - Q)Y$ with $Y \in \mathrm{MW}_2$, we have, for every $K > 0$

$$\begin{aligned}
&\left\|\max_{1\le k\le n}|S_k(X) - S_k(\mathcal{D}(X))|_{\mathcal{X}}\right\|_2 \\
&\quad \le \left\|\max_{1\le k\le n}|S_k(QY - QY \circ \theta^{-1})|_{\mathcal{X}}\right\|_2 \\
&\quad \le \|QY\|_{2,\mathcal{X}} + \left\|\max_{1\le k\le n}|QY \circ \theta^{k-1}|_{\mathcal{X}}\right\|_2 \\
&\quad \le \|QY\|_{2,\mathcal{X}} + K + n\||QY|_{\mathcal{X}}\mathbf{1}_{\{|QY|_{\mathcal{X}}\ge K\}}\|_2.
\end{aligned}$$

Hence,

$$\limsup_{n\to\infty}\Big\|\max_{1\le k\le n}\big|S_k(X)-S_k\big(\mathcal{D}(X)\big)\big|_{\mathcal{X}}\Big\|_2\le\big\||QY|_{\mathcal{X}}\mathbf{1}_{\{|QY|_{\mathcal{X}}\ge K\}}\big\|_2\underset{K\to\infty}{\longrightarrow}0,$$

and (19) holds. Then the proof of the WIP follows from Lemma B.1 and Proposition 3.2.

Let us prove Theorem 5.2. By Proposition 3.1 and (18), for every $1<p<2$, there exists $C_p>0$ such that

$$\big\|\mathcal{M}_2\big(X-\mathcal{D}(X)\big)\big\|_{p,\infty}\le C_p\|X\|_{\mathrm{MW}_2}.$$

Hence, by the Banach principle (see Lemma E.1), it suffices to prove (20) for $X=(I-Q)Y$, with $Y\in\mathrm{MW}_2$. But in this case the result is obvious, since $|QY|_{\mathcal{X}}\in L^2(\Omega)$ and, by the Borel–Cantelli lemma, $|QY|_{\mathcal{X}}\circ\theta^{n-1}=o(\sqrt{n})$ $\mathbb{P}$-a.s. By (20) and Proposition 3.3, $(X\circ\theta^n)_{n\ge 0}$ satisfies the CLIL. Then the ASIP follows from Proposition B.3, using that $\mathcal{D}(X)$ is pre-Gaussian.

It remains to prove (21). The first equality follows from (20) and (13). Let us prove that, with $d=\mathcal{D}(X)$, $\sup_{x^*\in\mathcal{X}^*,|x^*|_{\mathcal{X}^*}\le 1}\|x^*(d)\|_2\le 10\sqrt{2}\|X\|_{\mathrm{MW}_2}$. We first notice that $x^*(d)=\mathcal{D}(x^*(X))$ (with the obvious "new" meaning of the operator $\mathcal{D}$). Proceeding as above, one can prove that for every $m\ge 0$,

$$\begin{aligned}\big\|x^*(d)\big\|_2&=2^{m/2}\big\|S_{2^m}(d)\big\|_2/2^{m/2}\\&\le\big\|S_{2^m}(X)\big\|_2/2^{m/2}+\big\|S_{2^m}(d)-S_{2^m}(X)\big\|_2/2^{m/2}.\end{aligned}$$

Applying Proposition 5.1 [noticing that $\|x^*(X)\|_{\mathrm{MW}_2}\le\|X\|_{\mathrm{MW}_2}$] and Corollary 4.2 to $x^*(X)$, we derive that $\|x^*(d)\|_{\mathrm{MW}_2}\le 10\sqrt{2}\|X\|_{\mathrm{MW}_2}$ 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 $X\in\mathrm{MW}_2$. Let $\varepsilon>0$. By (38), there exists $Y\in\mathrm{MW}_2$ such that $\|X-(I-Q)Y\|_{\mathrm{MW}_2}\le\varepsilon$.

Then, by Corollary 4.2,

$$\Big\|\max_{1\le k\le n}\big|S_k(X)-S_k\big((I-Q)Y\big)\big|_{\mathcal{X}}\Big\|_2\le C\varepsilon\sqrt{n}.$$

Now, as in the proof of Proposition 5.1, for every $K>0$ we have

$$\begin{aligned}&\Big\|\max_{1\le k\le n}\big|S_k\big((I-Q)Y\big)-S_k\big(Y-\mathbb{E}_{-1}(Y)\big)\big|_{\mathcal{X}}\Big\|_2\\&\qquad\le\|QY\|_{2,\mathcal{X}}+K+n\big\||QY|_{\mathcal{X}}\mathbf{1}_{\{|QY|_{\mathcal{X}}\ge K\}}\big\|_2.\end{aligned}$$

Choose K such that $\||QY|_{\mathcal{X}}\mathbf{1}_{\{|QY|_{\mathcal{X}}\ge K\}}\|_2\le\varepsilon$ and then choose $n_0\ge(\|QY\|_{2,\mathcal{X}}+K)^2/\varepsilon^2$.

Then $\sup_{0\le t\le 1}|T_{n,t}(X)-T_{n,t}(Y-\mathbb{E}_{-1}(Y))|_{\mathcal{X}}\le C\varepsilon$. Now, $Y-\mathbb{E}_{-1}(Y)$ is a martingale difference, hence, by Proposition 3.2, $(T_{n,t}(Y-\mathbb{E}_{-1}(Y))_{0\le t\le 1})_{n\ge 0}$ is tight in $C([0,1],\mathcal{X})$. Then the tightness of $(T_{n,t}(X)_{0\le t\le 1})_{n\ge 0}$ follows from Proposition B.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 $x^*\in\mathcal{X}^*$, $x^*(X)$ satisfies the assumption of Proposition 5.1.

Let us prove the ASIP. We shall use Theorem B.3. In particular, we have to prove that $(X\circ\theta^n)_{n\ge 0}$ satisfies the CLIL.

By (18) and Lemma E.1, the set $\{X\in \mathrm{MW}_2 : (X\circ\theta^n)_{n\ge 0}$ satisfies the CLIL$\}$ is closed in MW_2. Hence, it suffices to prove the CLIL for $X=(I-Q)Y$, with $Y\in\mathrm{MW}_2$. But then $X=Y-\mathbb{E}_{-1}(Y)+QY\circ\theta^{-1}-QY$ and $((Y-QY)\circ\theta^n)_{n\ge 0}$ satisfies the CLIL by Proposition 3.3, while $|S_n(QY\circ\theta^{-1}-QY)|_{\mathcal{X}}=o(\sqrt{n})$ $\mathbb{P}$-a.s., by the Borel–Cantelli lemma. Hence, the CLIL is proved.

Now, let $x^*\in\mathcal{X}^*$. Clearly, $x^*(X)$ satisfies the assumption of Theorem 5.2, taking for $\mathcal{X}$ the Hilbert space $\mathbb{R}$. In particular, there exists $Z\in L^2(\Omega,\mathcal{F}_0,\mathbb{R})$ with $\mathbb{E}_{-1}(Z)=0$ such that

$$S_n\big(x^*(X)\big)-S_n(Z)=o\Big(\sqrt{nL\big(L(n)\big)}\Big)\qquad \mathbb{P}\text{-a.s.,}\tag{42}$$

$$\big\|S_n\big(x^*(X)\big)-S_n(Z)\big\|_2=o(\sqrt{n}).\tag{43}$$

The fact that $\mathcal{K}(x^*,y^*):=\lim_{n\to\infty}\frac{\mathrm{cov}(x^*(S_n(X)),y^*(S_n(X)))}{n}$ is the covariance operator of a Gaussian variable, follows from the WIP.

To prove the equality in (22), by a result of Kuelbs (see, e.g., Proposition D.1 in [7]), we have to prove that for every $x^*\in\mathcal{X}^*$, we have

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

But this follows from Theorem 5.2 applied to $x^*(X)$. Then the inequality in (22) may be proved as the inequality in (21). □

D.3. Proof of Proposition 5.4. We first recall the construction of Peligrad and Utev [39].

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

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 θ the shift and for $\mathcal{F}_0$, $\sigma\{W_n:n\le 0\}$. Then we are exactly in the situation considered in our paper.

Let $(a_n)_{n\in\mathbb{N}}$ be a sequence of positive numbers with $a_n\to 0$ as $n\to\infty$. It is proved in [39] that there exists a choice of $(p_n)_{n\ge 0}$, such that $\mathbb{E}(\tau)<\infty$, $\mathbb{E}(\tau^2)=+\infty$ and such that (23) holds with $X:=\mathbf{1}_{\{W_0=0\}}-\pi_0$.

Define $b_n := \sqrt{n \log\log n}$. 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 i.i.d., distributed like τ and $S_{T_k} = \sum_{i=1}^{k}(1-\pi_0\tau_i)$.

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

Since $\mathbb{E}(\tau) < \infty$, by the strong law of large numbers, $T_n/n \underset{n\to\infty}{\longrightarrow} \mathbb{E}(\tau)$ $\mathbb{P}$-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

$$\limsup_k \left|\sum_{i=0}^{k}(1-\pi_0\tau_i)\right| \Big/ b_k = +\infty \qquad \mathbb{P}\text{-a.s.} \tag{44}$$

But (44) follows from Strassen's converse to the law of the iterated logarithm (see, for instance, [31], pages 203–204), since $\mathbb{E}(\tau^2) = +\infty$. $\square$

APPENDIX E: TECHNICAL RESULTS

We recall here the Banach principle that we need (see Proposition C.1 of [7]).

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

$$\mathbb{P}\Big(\sup_{n\geq 1} |T_n x|_{\mathcal{X}} > \lambda |x|_{\mathbf{B}}\Big) \leq L(\lambda) \qquad \forall \lambda > 0, x \in \mathbf{B}. \tag{45}$$

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

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

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

PROOF. We first recall some useful fact about weak L^r-spaces (see Exercise 1.1.11, page 13 of Grafakos [26]). For every $r > 1$ and every $0 < t < r$, let

$$N_{r,t}(|X|_{\mathcal{X}}) := \sup_{\mathbb{P}(A)>0} \frac{1}{\mathbb{P}(A)^{1/r-1/t}} \big(\mathbb{E}(|X|_{\mathcal{X}}^t \mathbf{1}_A)\big)^{1/t}.$$

Then there exists $C_{r,t}$ such that

$$\|X\|_{r,\infty,\mathcal{X}}/C_{r,t} \leq N_{r,t}(|X|_{\mathcal{X}}) \leq C_{r,t}\|X\|_{r,\infty,\mathcal{X}},$$

and for $t = 1$, $N_{s,1}$ is a norm.

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

Take $\mathcal{X}=L^p(\mu)$ and $t=p$. We have, using Fubini,

$$\begin{aligned}\mathbb{E}\big(\|f\|_{L^p(\mu)}^p\mathbf{1}_A\big)&=\int_S\mathbb{E}\big(|f(s,\cdot)|^p\mathbf{1}_A\big)\,d\mu(s)\\&\le\mathbb{P}(A)^{1/r-1/p}\int_S N_{r,p}\big(|f|(s,\cdot)\big)\,d\mu(s).\end{aligned}$$

Hence,

$$\|f\|_{r,\infty,L^p(S)}\le C_{r,p}^2\bigg(\int_S\|f(s,\cdot)\|_{r,\infty}^p\,d\mu(s)\bigg)^{1/p}.$$

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

Acknowledgments. I am thankful to Jérôme Dedecker for providing me with a copy of [50]. I wish to thank two anonymous referees for valuable remarks that yielded an improved presentation of the paper.

REFERENCES

[1] ARAUJO, A. and GINÉ, E. (1980). *The Central Limit Theorem for Real and Banach Valued Random Variables*. Wiley, New York. MR0576407

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

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

[4] BERKES, I., HORVÁTH, L., SHAO, Q.-M. and STEINEBACH, J. (2000). Strong laws for L_p-norms of empirical and related processes. *Period*. *Math*. *Hungar.* **41** 35–69. MR1812797

[5] BRADLEY, R. C. (2007). *Introduction to Strong Mixing Conditions*, *Vol*. 1. Kendrick Press, Heber City, UT. MR2325294

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

[7] CUNY, C. (2015). A compact LIL for martingales in 2-smooth Banach spaces with applications. *Bernoulli* **21** 374–400. MR3322323

[8] CUNY, C. (2016). Limit theorems under the Maxwell–Woodroofe condition in Banach spaces. Available at arXiv:1403.0772.

[9] CUNY, C. and LIN, M. (2009). Pointwise ergodic theorems with rate and application to the CLT for Markov chains. *Ann*. *Inst*. *Henri Poincaré Probab*. *Stat*. **45** 710–733. MR2548500

[10] CUNY, C. and MERLEVÈDE, F. (2014). On martingale approximations and the quenched weak invariance principle. *Ann*. *Probab*. **42** 760–793. MR3178473

[11] CUNY, C. and MERLEVÈDE, F. (2015). Strong invariance principles with rate for “reverse” martingale differences and applications. *J. Theoret. Probab.* **28** 137–183. MR3320963

[12] DEDE, S. (2009). An empirical central limit theorem in $\mathbf{L}^1$ for stationary sequences. *Stochastic Process. Appl.* **119** 3494–3515. MR2568284

[13] DEDECKER, J. (2015). On the optimality of McLeish’s conditions for the central limit theorem. *C. R. Math. Acad. Sci. Paris* **353** 557–561. MR3348992

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

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

[16] DEDECKER, J. and MERLEVÈDE, F. (2017). Behavior of the Wasserstein distance between the empirical and the marginal distributions of stationary α-dependent sequences. **23** 2083–2127. MR3624887

[17] DEDECKER, J., MERLEVÈDE, F. and PÈNE, F. (2013). Empirical central limit theorems for ergodic automorphisms of the torus. *ALEA Lat. Am. J. Probab. Math. Stat.* **10** 731–766. MR3125745

[18] DEDECKER, J., MERLEVÈDE, F. and PÈNE, F. (2014). Rates in the strong invariance principle for ergodic automorphisms of the torus. *Stoch. Dyn.* **14** 1350021. MR3190216

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

[20] DEL BARRIO, E., GINÉ, E. and MATRÁN, C. (1999). Central limit theorems for the Wasserstein distance between the empirical and the true distributions. *Ann. Probab.* **27** 1009–1071. MR1698999

[21] DERRIENNIC, Y. and LIN, M. (2001). The central limit theorem for Markov chains with normal transition operators, started at a point. *Probab. Theory Related Fields* **119** 508–528. MR1826405

[22] DIESTEL, J. and UHL, J. J. JR. (1977). *Vector Measures. Mathematical Surveys* **15**. Amer. Math. Soc., Providence, RI. With a foreword by B. J. Pettis. MR0453964

[23] DUDLEY, R. M. (2002). *Real Analysis and Probability. Cambridge Studies in Advanced Mathematics* **74**. Cambridge Univ. Press, Cambridge. Revised reprint of the 1989 original. MR1932358

[24] GORDIN, M. and PELIGRAD, M. (2011). On the functional central limit theorem via martingale approximation. *Bernoulli* **17** 424–440. MR2797997

[25] GORDIN, M. I. and LIFŠIC, B. A. (1978). Central limit theorem for stationary Markov processes. *Dokl. Akad. Nauk SSSR* **239** 766–767. MR0501277

[26] GRAFAKOS, L. (2008). *Classical Fourier Analysis*, 2nd ed. *Graduate Texts in Mathematics* **249**. Springer, New York. MR2445437

[27] HALL, P. and HEYDE, C. C. (1980). *Martingale Limit Theory and Its Application*. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York. MR0624435

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

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

[30] KUELBS, J. and LEPAGE, R. (1973). The law of the iterated logarithm for Brownian motion in a Banach space. *Trans. Amer. Math. Soc.* **185** 253–265. MR0370725

[31] LEDOUX, M. and TALAGRAND, M. (1991). *Probability in Banach Spaces: Isoperimetry and Processes. Ergebnisse der Mathematik und Ihrer Grenzgebiete* (3) [*Results in Mathematics and Related Areas* (3)] **23**. Springer, Berlin. MR1102015

[32] LePage, R. D. (1972/1973). Loglog law for Gaussian processes. *Z. Wahrsch. Verw. Gebiete* **25** 103–108. MR0331456

[33] Lin, Z. and Zhao, Y. (2012). Strong approximation for ρ-mixing sequences. *Sci. China Math.* **55** 2159–2182. MR2972636

[34] Maxwell, M. and Woodroofe, M. (2000). Central limit theorems for additive functionals of Markov chains. *Ann. Probab.* **28** 713–724. MR1782272

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

[36] Merlevède, F., Peligrad, M. and Utev, S. (2006). Recent advances in invariance principles for stationary sequences. *Probab. Surv.* **3** 1–36. MR2206313

[37] Neveu, J. (1975). *Discrete-Parameter Martingales*, revised ed. *North-Holland Mathematical Library* **10**. North-Holland, Amsterdam; American Elsevier Publishing Co., Inc., New York. Translated from the French by T. P. Speed. MR0402915

[38] Peligrad, M. (1982). Invariance principles for mixing sequences of random variables. *Ann. Probab.* **10** 968–981. MR0672297

[39] Peligrad, M. and Utev, S. (2005). A new maximal inequality and invariance principle for stationary sequences. *Ann. Probab.* **33** 798–815. MR2123210

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

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

[42] Pisier, G. (1976). Sur la loi du logarithme itéré dans les espaces de Banach. In *Probability in Banach Spaces* (*Proc. First Internat. Conf., Oberwolfach*, 1975). *Lecture Notes in Math.* **526** 203–210. Springer, Berlin. MR0501237

[43] Rajput, B. S. (1972). Gaussian measures on L_p spaces, $1 \le p < \infty$. *J. Multivariate Anal.* **2** 382–403. MR0345157

[44] Rassoul-Agha, F. and Seppäläinen, T. (2008). An almost sure invariance principle for additive functionals of Markov chains. *Statist. Probab. Lett.* **78** 854–860. MR2398359

[45] Rosiński, J. (1982). Central limit theorems for dependent random vectors in Banach spaces. In *Martingale Theory in Harmonic Analysis and Banach Spaces* (*Cleveland, Ohio*, 1981). *Lecture Notes in Math.* **939** 157–180. Springer, Berlin. MR0668545

[46] Shao, Q. M. (1993). Almost sure invariance principles for mixing sequences of random variables. *Stochastic Process. Appl.* **48** 319–334. MR1244549

[47] Sharipov, O. S. (2009). Laws of the iterated logarithm and an almost sure invariance principle for mixing B-valued random variables and autoregressive processes. *Lith. Math. J.* **49** 203–215. MR2532035

[48] Vakhania, N. (1965). Sur une propriété des répartitions normales de probabilités dans les espaces l_p $(1 \le p < \infty)$ et H. *C. R. Acad. Sci. Paris* **260** 1334–1336. MR0174075

[49] Vakhania, N. N., Tarieladze, V. I. and Chobanyan, S. A. (1987). *Probability Distributions on Banach Spaces*. *Mathematics and Its Applications* (*Soviet Series*) **14**. D. Reidel Publishing Co., Dordrecht. Translated from the Russian and with a preface by Wojbor A. Woyczynski. MR1435288

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

[51] Wu, W. B. and Woodroofe, M. (2004). Martingale approximations for sums of stationary processes. *Ann. Probab.* **32** 1674–1690. MR2060314

[52] Wu, W. B. and Zhao, Z. (2008). Moderate deviations for stationary processes. *Statist. Sinica* **18** 769–782. MR2411619

[53] ZHAO, O. and WOODROOFE, M. (2008). Law of the iterated logarithm for stationary processes. *Ann. Probab.* **36** 127–142. MR2370600

LABORATOIRE MICS
CENTRALESUPELEC
GRANDE VOIE DES VIGNES
92295 CHATENAY-MALABRY CEDEX
FRANCE
E-MAIL: christophe.cuny@ecp.fr