Convergence rates in the law of large numbers for Banach valued dependent variables.

Jérôme Dedecker*and Florence Merlevède†
Université Paris VI

March 13, 2006

Abstract

We extend Marcinkievicz-Zygmund strong laws of large numbers for martingales to weakly dependent random variables with values in smooth Banach spaces. The conditions are expressed in terms of conditional expectations. In the case of Hilbert spaces, we show that our conditions are weaker than optimal ones for strongly mixing sequences (which were previously known for real-valued variables only). As a consequence, we give rates of convergence for Cramér-von Mises statistics and for the empirical estimator of the covariance operator of an Hilbert-valued autoregressive process.

Mathematics Subject Classifications (2000): 60 F 15, 60 E 15.
Key words: smooth Banach spaces, Hilbert spaces, Marcinkievicz-Zygmund strong laws of large numbers, almost sure convergence, martingales, weak dependence, Cramér-von Mises statistics.
Short Title: Dependent strong laws in Banach spaces.

---

*Jérôme Dedecker, Université Paris VI, Laboratoire de Statistique Théorique et Appliquée, Boîte 158, Plateau A, 8 ème étage, 175 rue du Chevaleret, 75013 Paris, France. Email: dedecker@ccr.jussieu.fr

†Florence Merlevède, Laboratoire de Probabilités et Modèles Aléatoires, Université Paris VI, et C.N.R.S UMR 7599, Boîte 188, 175 rue du Chevaleret, 75013 Paris, FRANCE. Email: merleve@ccr.jussieu.fr.

1 Introduction

The problem of rates of convergence in the strong law of large numbers for i.i.d. random variables with moments of order $1 \leq p < 2$ has been solved by Baum and Katz (1965) for real-valued variables, and by de Acosta (1981) for variables with values in Banach spaces of type p. To our knowledge the first extensions of this type of results to martingale differences sequences were derived by Woyczyński (1981) who considered random variables with values in smooth separable Banach spaces (see Definition 2 and Theorem 1, Section 2). Starting from Woyczyński's result and using a coboundary decomposition due to Gordin (1969), Theorem 1 can be extended to stationary sequences under mixingale-type conditions (see Theorem 3, Section 2.2). At the boundary (the case $p = 2$), this criterion is the same as that given by Gordin (1969) for the central limit theorem.

For strongly mixing real-valued random variables, this approach does not lead to optimal results, as quoted in the introduction of Section 3. Concerning the convergence rates in the strong law of large numbers for weakly dependent sequences of real-valued random variables, we mention the papers by Lai (1977), Hipp (1982), Berbee (1987), Peligrad (1985, 1989), Shao (1993) and Rio (1995). In the two latter ones, the optimality of the results is discussed.

In Theorem 4 of Section 3, we establish convergence rates in the strong law for partial sums of Hilbert-valued random variables under a projective criterion, which seems to be new even in the real case. This criterion can be viewed as a mixingale-type condition, since it is verified by martingale differences sequences and leads to the optimal results for strongly mixing processes. At the boundary (the case $p = 2$) our condition is the same as that obtained in Corollary $2(\beta)$ of Dedecker and Merlevède (2003) for the central limit theorem in Hilbert spaces. This generalization is important since it covers a much broader line of examples than mixing processes (see the examples in McLeish (1975) and Dedecker and Doukhan (2003)). The key of the proof of Theorem 4 is a new maximal inequality (see Proposition 1, Section 3) in which the dependence coefficients involved are expressed in terms of conditional expectations. This maximal inequality is more precise than a related one stated in Lemma 3.3 of Merlevède (2003). As in Merlevède's paper, the proof combines a martingale approximation of blocks as done in Shao (1993) and the quantile method as developed in Rio (1995, 2000).

In Proposition 2 of Section 4, we apply Theorem 4 to describe the asymptotic behaviour of Cramér-von Mises statistics obtained from dependent variables. In this context, it is natural to consider the empirical distribution function as a random variable with values in $\mathbb{L}^2(\mu)$ for an appropriate measure μ on the real line. As a by-product, we obtain the asymptotic behaviour of the supremum of the empirical process over generalized Sobolev balls (the class $W_1(\mu)$ defined in Lemma 1).

In Section 5, we study the empirical covariance operator C_n of a stationary Hilbert-valued autoregressive process. Such processes may be useful when considering estimation and forecasting problems for several classes of continuous time processes (see Bosq (2000) for more details). In Proposition 3 we give the rates of convergence of C_n to the covariance operator C_X when the innovations have moments of order $2 < p \leq 4$.

2 The case of smooth separable Banach spaces

In order to develop our result, we need some definitions.

Definition 1. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $(\mathbb{B}, \|\cdot\|_{\mathbb{B}})$ be a separable Banach space. For any real $p \geq 1$, denote by $\mathbb{L}^p_{\mathbb{B}}$ the space of $\mathbb{B}$-valued random variables such that $\|X\|^p_{\mathbb{L}^p_{\mathbb{B}}} = \mathbb{E}(\|X\|^p_{\mathbb{B}})$ is finite. Let $(\mathcal{M}_i)_{i>0}$ be an increasing sequence of σ-algebras of $\mathcal{A}$. We say that $(X_i)_{i>0}$ is a sequence of $\mathbb{B}$-valued martingale differences (with respect to the filtration $(\mathcal{M}_i)_{i>0}$) if:

1. For any positive i, X_i is $\mathcal{M}_i$-measurable and belongs to $\mathbb{L}^1_{\mathbb{B}}$.

2. For any $i > 1$, $\mathbb{E}(X_i|\mathcal{M}_{i-1}) = 0$ almost surely.

Definition 2. Following Pisier (1975), we say that a Banach space $(\mathbb{B}, \|\cdot\|_{\mathbb{B}})$ is r-smooth $(1 < r \leq 2)$ if there exists an equivalent norm $\|\cdot\|$ such that

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

Clearly, if $\mathbb{B}$ is r-smooth, then it is r'-smooth for any $r' \leq r$. A Banach space is said to be super-reflexive if it is r-smooth for some $1 < r \leq 2$. From Assouad (1975), we know that if $\mathbb{B}$ is r-smooth and separable, then there exists a constant D such that, for any sequence of $\mathbb{B}$-valued martingale differences $(X_i)_{i\geq 1}$,

$$\mathbb{E}(\|X_1 + \cdots + X_n\|^r_{\mathbb{B}}) \leq D \sum_{i=1}^n \mathbb{E}(\|X_i\|^r_{\mathbb{B}})\,. \tag{2.1}$$

From (2.1), we see that r-smooth Banach spaces play the same role for martingales as space of type r do for sums of independent variables. When the constant D needs to be specified, we shall say that $\mathbb{B}$ is (r, D)-smooth. Note that, for any measure space $(T, \mathcal{A}, \nu)$, $\mathbb{L}^p(T, \mathcal{A}, \nu)$ is $p \wedge 2$-smooth for any $p > 1$, and that any separable Hilbert space is $(2, 1)$-smooth.

Definition 3. For any non-increasing cadlag function f from $\mathbb{R}^+$ to $\mathbb{R}^+$, define the generalized inverse $f^{-1}(u) = \inf\{t \geq 0 : f(t) \leq u\}$. For any nonnegative random variable X, define the upper tail function $L_X(t) = \mathbb{P}(X > t)$ and the quantile function $Q_X = L_X^{-1}$. Let $(X_i)_{i>0}$ be a sequence of $\mathbb{B}$-valued random variables. Following Woyczińsky (1981), we write $(X_i) \prec X$ is there exists a positive random variable X such that $Q_X \geq \sup_{k \geq 1} Q_{\|X_k\|_{\mathbb{B}}}$.

2.1 The martingale case

In (1981), W. A. Woyczińsky proved the following theorem.

Theorem 1. *Let $\mathbb{B}$ be a separable Banach space and $(X_i)_{i>0}$ be a sequence of $\mathbb{B}$-valued martingale differences. Assume that $(X_i) \prec X$ for some positive random variable X and define $S_n = X_1 + \cdots + X_n$.*

1. *Let $1 < p < 2$. If X belongs to $\mathbb{L}^p_{\mathbb{R}}$ and $\mathbb{B}$ is r-smooth for some $r > p$, then $n^{-1/p} S_n$ tends to 0 almost surely.*
2. *If $\mathbb{E}(X \ln^+(X)) < \infty$ and $\mathbb{B}$ is super-reflexive, then $n^{-1} S_n$ tends to 0 almost surely.*

The next theorem is slightly more precise than Woyczińsky's result, as we shall see in Remarks 1 and 2. The proof of this result is given in the appendix.

Theorem 2. *Let $\mathbb{B}$, $(X_i)_{i>0}$, X and S_n be defined as in Theorem 1.*

1. *Let $1 < p < 2$. If X belongs to $\mathbb{L}^p_{\mathbb{R}}$ and $\mathbb{B}$ is r-smooth for some $r > p$, then, for any $1 \leq 1/\alpha \leq p$,*

$$\sum_{n=1}^{\infty} n^{\alpha p - 2} \mathbb{P}\Big(\max_{1 \leq k \leq n} \|S_k\|_{\mathbb{B}} \geq \varepsilon\, n^{\alpha} \Big) < \infty \,. \tag{2.2}$$

2. *If $\mathbb{E}(X \ln^+(X)) < \infty$ and $\mathbb{B}$ is super-reflexive, then*

$$\sum_{n=1}^{\infty} \frac{1}{n} \mathbb{P}\Big(\max_{1 \leq k \leq n} \|S_k\|_{\mathbb{B}} \geq \varepsilon\, n \Big) < \infty \,. \tag{2.3}$$

Remark 1. The sequence $\max_{1\leq k\leq n}\|S_k\|_{\mathbb{B}}$ being non decreasing, Property (2.2) with $\alpha p = 1$ is equivalent to

$$\sum_{N=1}^{\infty}\mathbb{P}\Big(\max_{1\leq k\leq 2^N}\|S_k\|_{\mathbb{B}}\geq \varepsilon\, 2^{N/p}\Big)<\infty\,. \tag{2.4}$$

We infer from (2.4) that $n^{-1/p}S_n$ tends to 0 almost surely, so that Theorem 2 contains Theorem 1.

Remark 2. Property (2.2) describes speed of convergence in the strong law. Indeed by Lemma 4 in Lai (1977), it implies in case $\alpha p > 1$ that

$$\sum_{n=1}^{\infty}n^{\alpha p-2}\mathbb{P}\Big(\sup_{k\geq n}k^{-\alpha}\|S_k\|_{\mathbb{B}}\geq \varepsilon\Big)<\infty\,. \tag{2.5}$$

Since the probabilities in (2.5) are non-increasing in n, it follows that

$$\mathbb{P}\Big(\sup_{k\geq n}k^{-\alpha}\|S_k\|_{\mathbb{B}}\geq \varepsilon\Big)=o\Big(\frac{1}{n^{\alpha p-1}}\Big)\,.$$

Remark 3. The moment condition $\mathbb{E}(X\ln^+(X))<\infty$ in Item 2 of Theorems 1 and 2 cannot be removed. More precisely, Elton (1981) proved the following result: if Y is any real-valued integrable and centered random variable such that $\mathbb{E}(|Y|\ln^+(|Y|))=\infty$, then there exists a sequence $(X_i)_{i>0}$ of martingale differences with the same marginal distribution as Y and such that $n^{-1}S_n$ diverges almost surely.

Remark 4. To be complete about strong laws of large numbers, let us quote the following remarkable result, which is a particular case of Theorem 1 in Woyczyński (1975b). For a separable Banach space $\mathbb{B}$, the following two conditions are equivalent:

1. $\mathbb{B}$ is r-smooth.
2. For every sequence $(X_n)_{n>0}$ of martingale-differences with values in $\mathbb{B}$ such that $\sum n^{-r}\mathbb{E}(\|X_n\|_{\mathbb{B}}^r)<\infty$, we have that $n^{-1}S_n$ tends to 0 almost surely and in $\mathbb{L}^r_{\mathbb{B}}$.

2.2 Application to stationary sequences

Let $(\Omega,\mathcal{A},\mathbb{P})$ be a probability space and $T:\Omega\mapsto\Omega$ be a bijective bimeasurable transformation preserving the probability $\mathbb{P}$. An element A of $\mathcal{A}$ is said to be invariant if $T(A)=A$. We denote by $\mathcal{I}$ the σ-algebra of all invariant sets. The probability $\mathbb{P}$ is ergodic if each element of $\mathcal{I}$ has measure 0 or 1.

Theorem 3. *Let $\mathcal{M}_0$ be a σ-algebra of $\mathcal{A}$ satisfying $\mathcal{M}_0 \subseteq T^{-1}(\mathcal{M}_0)$ and define the nondecreasing filtration $(\mathcal{M}_i)_{i\in\mathbb{Z}}$ by $\mathcal{M}_i = T^{-i}(\mathcal{M}_0)$. Let $\mathbb{B}$ be a separable Banach space and X_0 be a random variable in $\mathbb{L}^1_{\mathbb{B}}$ with mean zero. Define the sequence $(X_i)_{i\in\mathbb{Z}}$ by $X_i = X_0 \circ T^i$, and $S_n = X_1 + \cdots + X_n$. For $1 < p < 2$, consider the condition*

$$G(p): \quad \sum_{n=0}^{\infty} \mathbb{E}(X_n|\mathcal{M}_0) \quad \text{and} \quad \sum_{n=0}^{\infty} (X_{-n} - \mathbb{E}(X_{-n}|\mathcal{M}_0)) \quad \text{converge in } \mathbb{L}^p_{\mathbb{B}}\,.$$

If $G(p)$ holds and $\mathbb{B}$ is r-smooth for some $r > p$, then Property (2.2) holds for any $1 \le 1/\alpha \le p$.

Remark 5. For the strong law of large numbers (case $p = 1$), no additional condition is needed. It follows from Mourier's ergodic theorem (1953) that if $\mathbb{E}(\|X_0\|_{\mathbb{B}}) < \infty$, then

$$\frac{1}{n}\sum_{k=1}^{n} X_k \quad \text{converges almost surely to } \mathbb{E}(X_0|\mathcal{I}).$$

In fact, Mourier's result holds in any separable Banach space $\mathbb{B}$. Note that, if $\mathbb{P}$ is ergodic and $\mathbb{B} = \mathbb{R}$, we cannot obtain Property (2.3) without additional assumptions (see the example page 117 in Baum and Katz (1965)).

Remark 6. For the central limit theorem (case $p = 2$), Woyczyński (1975a) proved that: if $\mathbb{B}$ is 2-smooth and has a Schauder basis, if the X_i of Theorem 3 are martingale-differences with respect to the filtration $\mathcal{M}_i$, and if $\mathbb{P}$ is ergodic, then

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i \quad \text{converges in distribution to a Gaussian measure on } \mathbb{B}\,. \tag{2.6}$$

In fact (2.6) still holds if we replace the martingale assumption by Condition $G(2)$ (this comes from the coboundary decomposition (2.7) below). For real-valued variables, this result is due to Gordin (1969).

2.3 Proof of Theorem 3.

It is based on a coboundary decomposition due to Gordin (1969). More precisely, according to Theorem 4.3 in Lesigne and Volnỳ (2001), $G(p)$ holds if and only if

$$X_0 = M_0 + Z_0 - Z_0 \circ T\,, \tag{2.7}$$

where both M_0 and Z_0 belong to $\mathbb{L}^p_{\mathbb{B}}$, M_0 is $\mathcal{M}_0$ measurable and $\mathbb{E}(M_0|\mathcal{M}_{-1}) = 0$ a.s.. Note that the result in Lesigne and Volnỳ (2001) is given for real valued variables. For Banach valued variables, the proof is unchanged.

Let $M_i = M_0 \circ T^i$ and $Z_i = Z_0 \circ T^i$. Note that $Q_n = M_1 + \cdots + M_n$ is a martingale adapted to the filtration $\mathcal{M}_n$. Clearly $S_k = Q_k + Z_1 - Z_{k+1}$ and

$$\max_{1\le k\le n} \|S_k\|_{\mathbb{B}} \le \|Z_1\|_{\mathbb{B}} + \max_{1\le k\le n} \|Q_k\|_{\mathbb{B}} + \max_{2\le k\le n+1} \|Z_k\|_{\mathbb{B}} .$$

According to Theorem 2, it suffices to prove that $\sum_{n=1}^{\infty} n^{\alpha p-2}\mathbb{P}(\max_{2\le k\le n+1} \|Z_k\|_{\mathbb{B}} \ge \varepsilon\, n^{\alpha})$ is finite. By stationarity of Z_k, we infer that

$$\sum_{n=1}^{\infty} n^{\alpha p-2}\mathbb{P}\Big(\max_{2\le k\le n+1} \|Z_k\|_{\mathbb{B}} \ge \varepsilon\, n^{\alpha}\Big) \le \sum_{n=1}^{\infty} n^{\alpha p-1}\mathbb{P}\left(\|Z_0\|_{\mathbb{B}} \ge \varepsilon\, n^{\alpha}\right) .$$

Applying Fubini and using that Z_0 belongs to $\mathbb{L}^p_{\mathbb{B}}$, we infer that the right hand term is finite, which completes the proof.

3 The case of separable Hilbert spaces

Condition $G(p)$ is expressed in terms of conditional expectations, so that it seems to be a reasonable extension of the martingale case. However, it does not lead to the optimal condition for strongly mixing sequences of real-valued variables, as we shall see in the sequel.

In this section, we obtain a sufficient condition for Hilbert-valued variables, which contains both the martingale case and the case of strongly mixing sequences. In order to develop our results, we need more definitions.

Definition 4. For any nonnegative integrable random variable X with quantile function Q_X, let H_X be the function $x \to \int_0^x Q_X(u)du$. Note that, on the set $[0, \mathbb{P}(X > 0)]$, H_X is an absolutely continuous and increasing function with values in $[0, \mathbb{E}(X)]$. Denote by G_X the inverse of H_X.

Definition 5. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $(\mathbb{B}, \|\cdot\|_{\mathbb{B}})$ be a separable Banach space. For any σ-algebra $\mathcal{M}$ of $\mathcal{A}$ and any random variable Y in $\mathbb{L}^1(\mathbb{B})$, we consider the coefficient $\gamma(\mathcal{M}, Y)$ of weak dependence

$$\gamma(\mathcal{M}, Y) = \|\mathbb{E}(Y|\mathcal{M}) - \mathbb{E}(Y)\|_{\mathbb{L}^1_{\mathbb{B}}} . \tag{3.1}$$

Denote by $\mathbb{P}_Y$ the distribution of Y and by $\mathbb{P}_{Y|\mathcal{M}}$ a regular distribution of Y given $\mathcal{M}$. The strong mixing coefficient between $\mathcal{M}$ and $\sigma(Y)$ introduced by Rosenblatt (1956) may be defined as follows (see for instance Bradley (2002), Proposition 3.22)

$$\alpha(\mathcal{M}, \sigma(Y)) = \sup_{A \in \mathcal{B}(\mathbb{R})} \|\mathbb{P}_{Y|\mathcal{M}}(A) - \mathbb{P}_Y(A)\|_1 \tag{3.2}$$

(note that, with this definition, $\alpha(\mathcal{M}, \sigma(Y))$ is two times the usual one). Let $(X_i)_{i>0}$ be a sequence of $\mathbb{B}$-valued random variables and let $(\mathcal{M}_i)_{i\geq 0}$ be a sequence of σ-algebras of $\mathcal{A}$. The sequences of coefficients γ_i and α_i are then defined by

$$\gamma_i = \sup_{k\geq 0} \gamma(\mathcal{M}_k, X_{i+k}) \quad \text{and} \quad \alpha_i = \sup_{k\geq 0} \alpha(\mathcal{M}_k, \sigma(X_{i+k})) \tag{3.3}$$

Let $(X_i)_{i\in\mathbb{Z}}$ and $(\mathcal{M}_i)_{i\in\mathbb{Z}}$ be defined as in Theorem 3, and assume furthermore that X_0 is $\mathcal{M}_0$-measurable and in $\mathbb{L}^p_{\mathbb{H}}$ for some p in $]1, 2[$. Let $X = \|X_0\|_{\mathbb{H}}$. We can prove (see Subsection 6.2) that Condition $G(p)$ holds as soon as

$$\sum_{i\geq 0} (i+1)^{p-1} \int_0^{\gamma_i} Q_X^{p-1} \circ G_X(u) du < \infty\,. \tag{3.4}$$

According to the inequality $2G_X(\gamma_k/18) \leq \alpha_k$, proved in Dedecker and Merlevède (2003), page 250, we infer that (3.4) holds as soon as

$$\sum_{i\geq 0} (i+1)^{p-1} \int_0^{\alpha_i} Q_X^p(u) du < \infty\,. \tag{3.5}$$

From Rio (1995), we know that, for real-valued variables, Property (2.2) holds for $\alpha p = 1$ and p in $]1, 2[$ as soon as

$$\sum_{i\geq 0} (i+1)^{p-2} \int_0^{\alpha_i} Q_X^p(u) du < \infty\,, \tag{3.6}$$

which is clearly less restrictive than (3.5). Comparing (3.4) and (3.6), a natural question arises: does Property (2.2) still holds for $\mathbb{H}$-valued random variables, under the condition

$$DM(p, \gamma, X): \quad \sum_{i\geq 0} (i+1)^{p-2} \int_0^{\gamma_i} Q_X^{p-1} \circ G_X(u) du < \infty\,.$$

which is implied by either (3.4) or (3.6)? We shall see in Theorem 4 below that the answer is positive, even in the non stationary case. Note that, in the stationary case, the conditions $G(p)$ and $DM(p, \gamma, X)$ cannot be compared. However, it may happen that $G(p)$ does not hold for any p in $]1, 2[$, while $DM(p, \gamma, X)$ holds for any p in $]1, 2[$ (see again Subsection 6.2).

Theorem 4. *Let $(X_k)_{k>0}$ be a sequence of random variables in $\mathbb{L}^1_{\mathbb{H}}$, and let $\mathcal{M}_k = \sigma(X_i, i \leq k)$. Let $S_n = \sum_{i=1}^n (X_i - \mathbb{E}(X_i))$ and define the coefficients $(\gamma_i)_{i\geq 0}$ as in (3.3). Let X be a positive random variable such that $(X_i) \prec X$.*

1. *If $\mathbb{E}(X^p) < \infty$ and $DM(p, \gamma, X)$ holds for some p in $]1, 2[$, then (2.2) holds for any $1 \leq 1/\alpha \leq p$.*
2. *If $\mathbb{E}(X \ln^+(X)) < \infty$ and $\sum_{i\geq 1} \gamma_i / i < \infty$, then (2.3) holds.*

Remark 7. For a stationary sequence $(X_i)_{i\in\mathbb{Z}}$ of centered $\mathbb{H}$-valued random variables, Condition $DM(2, \gamma, \|X_0\|_{\mathbb{H}})$ implies that $n^{-1/2} S_n$ converges weakly to a mixture of Gaussian distributions in $\mathbb{H}$. This result is proved in Corollary 2 of Dedecker and Merlevède (2003).

From Lemma 2 in Dedecker and Doukhan (2003), we obtain sufficient conditions for $DM(p, \gamma, X)$ to hold.

Corollary 1. *Let $1 < p \leq 2$. Any of the following conditions implies $DM(p, \gamma, X)$.*

1. *$\mathbb{P}(X > x) \leq (c/x)^r$ for some $r > p$, and $\sum_{i\geq 0} (i+1)^{p-2} (\gamma_i)^{(r-p)/(r-1)} < \infty$.*
2. *$\|X\|_r < \infty$ for some $r > p$, and $\sum_{i\geq 1} i^{(pr-2r+1)/(r-p)} \gamma_i < \infty$.*
3. *$\mathbb{E}(X^p (\ln(1+X))^{p-1}) < \infty$ and $\gamma_i = O(a^i)$ for some $a < 1$.*

If $\mathbb{H} = \mathbb{R}$, the condition (3.6) may be weakened as shown in Rio (2000, Corollary 3.1). For any real valued random variable Y, let $F_Y(t) = \mathbb{P}_Y(]-\infty, t])$ and $F_{Y|\mathcal{M}}(t) = \mathbb{P}_{Y|\mathcal{M}}(]-\infty, t])$. Define

$$\alpha(\mathcal{M}, Y) = \sup_{t\in\mathbb{R}} \|F_{Y|\mathcal{M}}(t) - F_Y(t)\|_1 \text{ and } \widetilde{\alpha}_i = \sup_{k>0} \alpha(\mathcal{M}_k, X_{i+k}) \,. \tag{3.7}$$

From Dedecker and Prieur (2005, Proposition 2 Item 2), we know that, for real-valued variables, $G_X(\gamma_k/2) \leq \widetilde{\alpha}_k$, so that $DM(p, \gamma, X)$ holds as soon as

$$\sum_{i\geq 0} (i+1)^{p-2} \int_0^{\widetilde{\alpha}_i} Q_X^p(u) du < \infty \,. \tag{3.8}$$

Consequently, for $1 < p < 2$, Condition $DM(p, \gamma, X)$ is weaker than Rio's criterion (1995, 2000). For $p = 1$, Rio does not assume that $\mathbb{E}(X \ln^+(X))$ is finite. In that case, the difference between our result and Rio's is the same as the difference between independent variables and martingale differences for the strong law of large numbers. Note also that Theorem 4 sharpens Theorem 1 in Shao (1993) in the special case where $p \in]1,2[$. However the maximal inequality stated in Proposition 1 below does not allow to obtain sharp results in the case where $1 \leq 1/\alpha < 2 \leq p < \infty$, which is also considered by Shao.

To understand the difference between α_i and $\widetilde{\alpha}_i$, let us give the following example: if $(\epsilon_i)_{i>0}$ is i.i.d. with marginal $\mathcal{B}(1/2)$, then, for the stationary solution $(X_i)_{i>0}$ of the equation $2X_n = X_{n-1} + \epsilon_n$, we have $\alpha_i = 1/2$ and $\widetilde{\alpha}_i \leq 2^{-i}$.

Now to better understand the difference between $\widetilde{\alpha}_i$ and γ_i, note that one can build a sequence of martingale differences (i.e. $\gamma_i = 0$ for $i > 0$) such that $\widetilde{\alpha}_i$ does not converge to 0. For the one-sided linear processes, γ_i can be easily computed: let $(\epsilon_i)_{i\in\mathbb{Z}}$ be a stationary sequence of centered r.v.'s in $\mathbb{L}^1_{\mathbb{H}}$, $(a_i)_{i\geq 0}$ a sequence of linear operators from $\mathbb{H}$ to $\mathbb{H}$ such that $\sum_{i=0}^{\infty} \|a_i\| < \infty$, and $X_n = \sum_{i=0}^{\infty} a_i(\epsilon_{n-i})$. For $i \geq 0$, let $\gamma_i^{\epsilon} = \|\mathbb{E}(\epsilon_i|\sigma(\epsilon_j, j \leq 0))\|_{\mathbb{L}^1_{\mathbb{H}}}$ and $\gamma_i^X = \|\mathbb{E}(X_i|\sigma(X_j, j \leq 0))\|_{\mathbb{L}^1_{\mathbb{H}}}$. We have the following upper bound

$$\gamma_n^X \leq \sum_{i=0}^{n-1} \|a_i\|\gamma_{n-i}^{\epsilon} + \|\epsilon_0\|_{\mathbb{L}^1_{\mathbb{H}}} \sum_{i=n}^{\infty} \|a_i\| \,.$$

Now, according to the inequality $G_X(\gamma_k/2) \leq \widetilde{\alpha}_k$ and to the examples given in Rio (1995), we can see that Condition $DM(p,\gamma,X)$ is essentially optimal. For instance, Corollary 2 below follows easily from Theorem 2 in Rio (1995) (apply Rio's result with $a = r(p-1)/(r-p)$ and $b = r/(r-p)$).

Corollary 2. *For any $1 < p < 2$ and any $r > p$, there exists a strictly stationary real-valued Markov chain $(X_i)_{i\in\mathbb{Z}}$ such that $\mathbb{E}(X_0) = 0$ and*

1. *For any nonnegative real x, $\mathbb{P}(|X_0| > x) = \min(1, x^{-r})$.*
2. *The sequence $(\gamma_i)_{i\geq 0}$ satisfies $\sup_{i>0} i^{(p-1)} \ln(i)(\gamma_i)^{(r-p)/(r-1)} < \infty$.*
3. $\limsup_{n\to\infty} \dfrac{|S_n|}{n^{1/p}} = +\infty$ *almost surely.*

We conclude this section by giving a maximal inequality (Proposition 1) which is the key result to prove Theorem 4. The proof of the proposition 1 will be done in Appendix.

Definition 6. For any non-increasing sequence $(\delta_i)_{i\geq 0}$ of nonnegative numbers, define $\delta^{-1}(u) = \sum_{i\geq 0} \mathbb{I}_{u<\delta_i} = \inf\{k \in \mathbb{N} : \delta_k \leq u\}$. For any nonincreasing cadlag function f, define the generalized inverse by $f^{-1}(u) = \inf\{t : f(t) \leq u\}$. Note that δ^{-1} is the generalized inverse of the cadlag function $x \to \delta_{[x]}$, $[.]$ denoting the integer part.

Proposition 1. *Let $(X_k)_{k>0}$ be a sequence of random variables in $\mathbb{L}^1_{\mathbb{H}}$, and let $\mathcal{M}_k = \sigma(X_i, i \leq k)$. Let $S_n = \sum_{i=1}^n (X_i - \mathbb{E}(X_i))$ and define the coefficients $(\gamma_i)_{i\geq 0}$ as in (3.3). Let X be a positive random variable such that $(X_i) \prec X$. Let $R_X = ((\gamma/2)^{-1} \circ G_X^{-1} \wedge n) Q_X$ and $S_X = R_X^{-1}$. For any $x > 0$ and $r \geq 1$,*

$$\mathbb{P}\Big(\max_{1\leq k\leq n} \|S_k\|_{\mathbb{H}} \geq 5x\Big) \leq \frac{14n}{x} \int_0^{S_X(x/r)} Q_X(u)du + \frac{4n}{x^2} \int_{S_X(x/r)}^1 R_X(u) Q_X(u) du\,. \tag{3.9}$$

Remark 8. Note that for $p \geq 1$, $\mathbb{E}(|Z|^p) = p\,5^p \int_0^\infty x^{p-1} \mathbb{P}(|Z| \geq 5x)dx$. With the same notations as in Proposition 1, we obtain from Inequality (3.9) that, for any fixed real number p in $]1, 2[$,

$$\mathbb{E}\Big(\max_{1\leq k\leq n} \|S_k\|^p_{\mathbb{H}}\Big) \leq C_p\, n \int_0^1 R_X^{p-1}(u) Q_X(u) du\,,$$

where $C_p = 5^p 2p\,(12 - 5p)(p-1)^{-1}(2-p)^{-1}$. By the definition of R_X, it follows that

$$\mathbb{E}\Big(\max_{1\leq k\leq n} \|S_k\|^p_{\mathbb{H}}\Big) \leq C_p\, n \sum_{i=0}^{n-1} (i+1)^{p-2} \int_0^{\gamma_i/2} Q_X^{p-1} \circ G_X(u) du\,.$$

3.1 Proof of Theorem 4.

Proof of Item 1. For the sake of brevity, write L, Q, R, S and G for L_X, Q_X, R_X, S_X and G_X respectively. Applying Inequality (3.9) with $x = x_n = (\varepsilon n^\alpha)/5$ and $r = 1$, we obtain that, for any $\varepsilon \in]0, 1]$,

$$n^{\alpha p-2} \mathbb{P}\Big(\max_{1\leq k\leq n} \|S_k\|_{\mathbb{H}} \geq \varepsilon n^\alpha\Big) \leq \frac{70}{\varepsilon} n^{\alpha(p-1)-1} \int_0^{S(x_n)} Q(u)du + \frac{100}{\varepsilon^2} n^{\alpha(p-2)-1} \int_{S(x_n)}^1 R(u)Q(u)du\,. \tag{3.10}$$

Since R is right-continuous and non-increasing,

$$u < S(x_n) \iff R(u) > (\varepsilon n^\alpha)/5 \iff n < \Big(\frac{5R(u)}{\varepsilon}\Big)^{1/\alpha}\,. \tag{3.11}$$

Ending the proof as in Rio (2000), page 59, we infer that there exists a finite constant C depending only on α and on ε, such that for all $\varepsilon \in]0, 1]$,

$$\begin{aligned}\sum_{n\geq 1} n^{\alpha p-2} \mathbb{P}\Big(\max_{1\leq k\leq n} \|S_k\|_{\mathbb{H}} \geq \varepsilon n^\alpha\Big) &\leq C \int_0^1 R^{p-1}(u) Q(u) du \\ &\leq C \int_0^1 ((\gamma/2)^{-1} \circ G^{-1}(u))^{p-1} Q^p(u) du\,.\end{aligned}$$

Setting $v = H(u)$, the right hand side is finite as soon as

$$\int_0^1 ((\gamma/2)^{-1}(u))^{p-1} Q^{p-1} \circ G(u) du < \infty ,$$

which is equivalent to $DM(p, \gamma, X)$ (see for instance Rio (2000), Appendix C).

Proof of Item 2. We apply Inequality (3.9) with $x = x_n = (\varepsilon n)/5$ and $r = 1$:

$$n^{-1}\mathbb{P}\Big(\max_{1\le k\le n} \|S_k\|_{\mathbb{H}} \ge \varepsilon n^{\alpha}\Big) \le I_1(n) + I_2(n) + I_3(n) \tag{3.12}$$

where

$$\begin{aligned} I_1(n) &= \frac{70}{n\varepsilon}\int_0^1 Q(u)\mathbb{1}_{u<L(x_n)} du \\ I_2(n) &= \frac{70}{n\varepsilon}\int_0^{S(x_n)} Q(u)\mathbb{1}_{u\ge L(x_n)} du \\ I_3(n) &= \frac{100}{\varepsilon^2 n^2}\int_{S(x_n)}^1 R(u)Q(u) du \end{aligned}$$

Note first that, for any positive real A we have,

$$Q_{X\mathbb{1}(X>A)}(u) = Q(u)\mathbb{1}_{u<L(A)} . \tag{3.13}$$

Applying first (3.13), we obtain that $(\varepsilon/70)\sum_{n\ge1} I_1(n) = \sum_{n\ge1} n^{-1}\mathbb{E}(X\mathbb{1}_{X>n\varepsilon/5})$. Then applying Fubini, it follows that $\sum_{n\ge1} I_1(n) < \infty$ as soon as $\mathbb{E}(X \ln^+(X)) < \infty$. On the other hand, using the fact that $L(x_n) \le u < S(x_n)$ if and only if $Q(u) \le x_n < R(u)$, we get

$$\frac{\varepsilon}{70}\sum_{n\ge1} I_2(n) \le \int_0^1 Q(u)\Big(\sum_{1\vee 5Q(u)/\varepsilon\le n<5R(u)/\varepsilon} \frac{1}{n}\Big) du .$$

Next, using the elementary inequalities $\ln(K+1) - \ln 2 \le \sum_{n=2}^{K} n^{-1} \le \ln(K)$, we easily infer that

$$\frac{\varepsilon}{70}\sum_{n\ge1} I_2(n) \le \int_0^1 Q(u) du + \int_0^1 Q(u)\ln\big(1 + (\gamma/2)^{-1}\circ G^{-1}(u)\big) du .$$

Then, setting $v = H(u)$, it follows that

$$\frac{\varepsilon}{70}\sum_{n\ge1} I_2(n) \le \mathbb{E}\|X\|_{\mathbb{H}} + \int_0^{\mathbb{E}\|X\|_{\mathbb{H}}} \ln(1+\gamma^{-1}(u)) du ,$$

and the right hand side is finite if and only if $\sum_{i\geq 1}\gamma_i/i<\infty$.

Finally, using again (3.11), we obtain that

$$\sum_{n\geq 1} n^{-2}\mathbb{1}(u\geq S(x_n)) = \sum_{n\geq 1\vee 5R(u)/\varepsilon} n^{-2} \leq 2\Big(\frac{5R(u)}{\varepsilon}\vee 1\Big)^{-1}.$$

Consequently $\varepsilon^2\sum_{n\geq 1} I_3(n)\leq K\,\mathbb{E}(\|X\|_{\mathbb{H}})$ for some K, which completes the proof. ■

4 Cramér-von Mises statistics

Let $(\Omega,\mathcal{A},\mathbb{P})$, T and $\mathcal{I}$ be as in Section 2.2. Let X_0 be a real-valued random variable and $X_i=X_0\circ T^i$. Let F be the distribution function of X_0 and define

$$F_n(t)=\frac{1}{n}\sum_{i=1}^n \mathbb{1}_{X_i\leq t}\,.$$

Let μ be a σ-finite measure on $\mathbb{R}$ and suppose that F satisfies

$$\int_{\mathbb{R}_-}(F(t))^2\mu(dt)+\int_{\mathbb{R}_+}(1-F(t))^2\mu(dt)<\infty. \tag{4.1}$$

Under this assumption, the process $\{t\to F_n(t)-F(t), t\in\mathbb{R}\}$ may be viewed as a random variable with values in the Hilbert space $\mathbb{L}^2(\mu)$. Define then

$$D_n(\mu)=\Big(\int (F_n(t)-F(t))^2\mu(dt)\Big)^{1/2}.$$

When $\mu=dF$, $D_n(\mu)$ is known as the Cramér-von Mises statistics, and is commonly used for testing goodness of fit. It is also interesting to write $D_n(\mu)$ as the supremum of the empirical process over a particular class of functions. For this task, we need the following lemma, whose proof will be done in the appendix.

Lemma 1. *For any two distibutions functions F and G satisfying (4.1), define*

$$D(F,G,\mu)=\Big(\int (F(t)-G(t))^2\mu(dt)\Big)^{1/2}.$$

Let $W_1(\mu)$ be the set of functions

$$\Big\{f\,:\,f(t)=f(0)+\Big(\int_{[0,t[} g(x)\mu(dx)\Big)\mathbb{1}_{t>0}-\Big(\int_{[t,0[} g(x)\mu(dx)\Big)\mathbb{1}_{t\leq 0},\ \int (g(x))^2\mu(dx)\leq 1\Big\}.$$

With these notations, we have that

$$D(F,G,\mu)=\sup_{f\in W_1(\mu)}\Big|\int f dF-\int f dG\Big|.$$

According to Lemma 1, since $D_n(\mu) = D(F_n, F, \mu)$, we get that

$$D_n(\mu) = \sup_{f \in W_1(\mu)} \left| \frac{1}{n} \sum_{i=1}^{n} \big(f(X_i) - \mathbb{E}(f(X_i)) \big) \right| .$$

In particular, if μ is the Lebesgue measure on the real line, $W_1(\mu)$ contains the unit ball of the Sobolev space of order 1.

We now define the dependence coefficients which naturally appear in this context. Let X be a real valued random variable and $\mathcal{M}$ be a σ-algebra of $\mathcal{A}$. Keeping the same notations as in Definition 5, define the coefficient $\tau_\mu(\mathcal{M}, X)$ by

$$\tau_\mu(\mathcal{M}, X) = \Big\| \Big(\int (F_{X|\mathcal{M}}(t) - F_X(t))^2 \mu(dt) \Big)^{1/2} \Big\|_1$$

The sequence $(\tau_\mu(i))_{i \geq 0}$ of $(X_i)_{i \in \mathbb{Z}}$ is then defined by

$$\tau_\mu(i) = \tau_\mu(\mathcal{M}_0, X_i) \quad \text{for} \quad \mathcal{M}_0 = \sigma(X_i, i \leq 0) .$$

With the help of this coefficient, we can describe the asymptotic behavior of $D_n(\mu)$.

Proposition 2. *Assume that the distribution function F of X_0 satisfies (4.1). Define the function F_μ by: $F_\mu(x) = \mu([0, x[)$ if $x \geq 0$ and $F_\mu(x) = -\mu([x, 0[)$ if $x \leq 0$. Define also $Y_\mu = \sqrt{|F_\mu(X_0)|}$.*

1. *If $F_{X_0|\mathcal{I}} = F$ and Y_μ is integrable, then $D_n(\mu)$ converges to 0 almost surely.*
2. *If $DM(p, \tau_\mu, Y_\mu)$ holds for some p in $]1, 2[$, then $n^{(p-1)/p} D_n(\mu)$ converges to 0 almost surely.*
3. *If $DM(2, \tau_\mu, Y_\mu)$ holds, then $\sqrt{n} D_n(\mu)$ converges in distribution to $\sqrt{\int G^2(x)\mu(dx)}$, where G is a mixture of gaussian processes.*

Proof of Proposition 2. Define the variable $Z_i = \{t \to \mathbb{1}_{X_i \leq t} - F(t), t \in \mathbb{R}\}$ which belongs to $\mathbb{H} = \mathbb{L}^2(\mu)$ as soon as (4.1) holds. Clearly

$$(4.2) \quad \|Z_i\|_{\mathbb{H}} \leq \Big(\int_{]-\infty,0[} (\mathbb{1}_{X_i \leq t})^2 \mu(dt) + \int_{[0,\infty[} (1 - \mathbb{1}_{X_i \leq t})^2 \mu(dt) \Big)^{1/2} + \Big(\int_{]-\infty,0[} (F(t))^2 \mu(dt) + \int_{[0,\infty[} (1 - F(t))^2 \mu(dt) \Big)^{1/2} ,$$

so that $\|Z_i\|_{\mathbb{H}} \leq \sqrt{|F_\mu(X_i)|} + \mathbb{E}(\sqrt{|F_\mu(X_i)|})$ and $\mathbb{E}(\|Z_i\|_{\mathbb{H}}) \leq 2\mathbb{E}(Y_\mu)$. Item 1 follows from (4.2) and Mourier's ergodic theorem, as quoted in Remark 5. Now, by definition of γ_k, we have that $\gamma_k = \mathbb{E}(\|\mathbb{E}(Z_k|\mathcal{M}_0)\|_{\mathbb{H}}) = \tau_\mu(k)$. On the other hand, we infer from (4.2) that

$$Q_{\|Z_i\|_{\mathbb{H}}} \leq Q_{Y_\mu + \mathbb{E}(Y_\mu)} \leq Q_{Y_\mu} + \mathbb{E}(Y_\mu)\,. \tag{4.3}$$

Since $\mathbb{E}\|Y_\mu\|_{\mathbb{H}} \leq \int_0^1 Q_{Y_\mu}(u)du$ and since Q_{Y_μ} is non-increasing, we get for all $x \in [0,1]$,

$$\int_0^x Q_{\|Z_0\|_{\mathbb{H}}}(u)du \leq \int_0^x Q_{Y_\mu}(u)du + x\int_0^1 Q_{Y_\mu}(u)du \leq 2\int_0^x Q_{Y_\mu}(u)du\,. \tag{4.4}$$

Now for two increasing continuous functions f and g, we have that $f \leq g$ if and only if $f^{-1} \geq g^{-1}$. In addition $[2g(x)]^{-1} = g^{-1}(x/2)$ and consequently $G_{\|Z_0\|_{\mathbb{H}}}(u) \geq G_{Y_\mu}(u/2)$. From (4.3) and the last inequality, we infer that, for any $1 < p \leq 2$,

$$\begin{aligned} \int_0^{\gamma_k} Q^{p-1}_{\|Z_0\|_{\mathbb{H}}} \circ G_{\|Z_0\|_{\mathbb{H}}}(u)du &\leq \int_0^{\gamma_k} Q^{p-1}_{Y_\mu} \circ G_{\|Z_0\|_{\mathbb{H}}}(u)du + \int_0^{\gamma_k} (\mathbb{E}(Y_\mu))^{p-1}du \\ &\leq 2\int_0^{\gamma_k/2} Q^{p-1}_{Y_\mu} \circ G_{Y_\mu}(u)du + \gamma_k(\mathbb{E}(Y_\mu))^{p-1}\,. \end{aligned} \tag{4.5}$$

Since $\gamma_k = \tau_\mu(k)$, we infer from (4.5) that $DM(p, \gamma, \|Z_0\|_{\mathbb{H}})$ holds as soon as $DM(p, \tau_\mu, Y_\mu)$ does. Hence Item 2 follows from Theorem 4 (together with Remark 1) and Item 3 from Corollary 2 in Dedecker and Merlevède (2003) (the covariance structure of G conditionally to $\mathcal{I}$ is given in Example 2 of the same paper).

4.1 Examples

The coefficients $\tau_\mu(\mathcal{M}, X)$ may be compared to $\alpha(\mathcal{M}, \sigma(X))$, $\alpha(\mathcal{M}, X)$ and to other dependence coefficients introduced in Dedecker and Prieur (2005). Define $\tau(\mathcal{M}, X)$ and $\beta(\mathcal{M}, X)$ by

$$\begin{aligned} \tau(\mathcal{M}, X) &= \int \|F_{X|\mathcal{M}}(t) - F_X(t)\|_1 dt \\ \beta(\mathcal{M}, X) &= \|\sup_{t\in\mathbb{R}} |F_{X|\mathcal{M}}(t) - F_X(t)|\,\|_1 \end{aligned}$$

Lemma 2. *For any real random variable X and any σ-algebra $\mathcal{M}$, we have*

1. *Let $Y_\mu = \sqrt{|F_\mu(X)|}$. We have the bound $\tau_\mu(\mathcal{M}, X) \leq 36\int_0^{\alpha(\mathcal{M},\sigma(X))/2} Q_{Y_\mu}(u)du$.*
2. *If μ is a probability measure, $\tau_\mu(\mathcal{M}, X) \leq \sqrt{\alpha(\mathcal{M}, X)}$ and $\tau_\mu(\mathcal{M}, X) \leq \beta(\mathcal{M}, X)$.*

3. If F_μ is K-Lipschitz, then $\tau_\mu(\mathcal{M}, X) \leq \sqrt{K\tau(\mathcal{M}, X)}$.

Recall that $\tau(\mathcal{M}, X)$ has the following property: if X^* is any random variable independant of $\mathcal{M}$ and distributed as X, then $\tau(\mathcal{M}, X)) \leq \|X - X^*\|_1$. Further, if Ω is rich enough, one can choose X^* such that $\tau(\mathcal{M}, X)) = \|X - X^*\|_1$. Due to this property, the coefficient τ is easy to compute in many situations. Among the large variety of examples given in Dedecker and Prieur (2005), let us choose two important cases: if $X_n = \sum_{j\geq 0} a_j \xi_{n-j}$, where $(\xi_i)_{i\in\mathbb{Z}}$ is i.i.d., then $\tau(\mathcal{M}_0, X_n) \leq 2\|\xi_0\|_1(\sum_{j\geq n} |a_j|)$ and also $\tau(\mathcal{M}_0, X_n) \leq (2\mathrm{Var}(\xi_0)\sum_{j\geq n} a_j^2)^{1/2}$. If $(X_n)_{n\geq 0}$ is the stationary solution of the equation $X_n = f(X_{n-1}) + \xi_n$, where f is κ-Lipshitz for some $\kappa < 1$ and $(\xi_i)_{i\in\mathbb{Z}}$ is i.i.d, then $\tau(\mathcal{M}_0, X_n) \leq 2\|X_0\|_1 \kappa^n$. Note also that the coefficient $\beta(\mathcal{M}, X)$ is well adapted to dynamical systems (see Dedecker and Prieur (2005), Section 4.4).

Proof of Lemma 2. Define the variable $Z = \{t \to \mathbb{1}_{X\leq t} - F(t), t \in \mathbb{R}\}$ which belongs to $\mathbb{H} = \mathbb{L}^2(\mu)$ as soon as (4.1) holds. Using the definition of $\tau(\mathcal{M}, X)$ and an inequality of Dedecker and Merlevède (2003) (stated at the end of the proof of their Corollary 2), we infer that

$$\tau_\mu(\mathcal{M}, X) = \mathbb{E}(\|\mathbb{E}(Z|\mathcal{M})\|_{\mathbb{H}}) \leq 18 \int_0^{\alpha(\mathcal{M},\sigma(X))/2} Q_{\|Z\|_{\mathbb{H}}}(u)du\,. \tag{4.6}$$

Now, it follows from (4.3) that $Q_{\|Z\|_{\mathbb{H}}} \leq Q_{Y_\mu} + \mathbb{E}(Y_\mu)$. Since Q_{Y_μ} is nonincreasing, Item 1 follows from (4.4) and (4.6).

If μ is a probability measure, then

$$\tau_\mu(\mathcal{M}, X) \leq \Big(\int \|F_{X|\mathcal{M}}(t) - F_X(t)\|_1 \mu(dt)\Big)^{1/2} \leq \sqrt{\alpha(\mathcal{M}, X)}\,.$$

The second inequality of Item 2 follows from the bound

$$\tau_\mu(\mathcal{M}, X) \leq \Big\|\Big(\int \Big(\sup_{t\in\mathbb{R}} |F_{X|\mathcal{M}}(t) - F_X(t)|\Big)^2 \mu(dt)\Big)^{1/2}\Big\|_1 = \beta(\mathcal{M}, X)\,.$$

We now prove Item 3. Since F_μ is K-Lipschitz, μ is absolutely continuous with respect to the Lebesgue measure, with a density bounded by K. Consequently

$$\tau_\mu(\mathcal{M}, X) \leq \Big(K\int \|F_{X|\mathcal{M}}(t) - F_X(t)\|_1 dt\Big)^{1/2} = \sqrt{K\tau(\mathcal{M}, X)}\,.$$

5 Estimation of the covariance operator of an Hilbertian autoregressive process

Let $\mathbb{H}$ be a separable Hilbert space with norm $\|\cdot\|_{\mathbb{H}}$ and inner product $\langle\cdot,\cdot\rangle_{\mathbb{H}}$. Denote by $L(\mathbb{H})$ the class of bounded linear operators from $\mathbb{H}$ to $\mathbb{H}$ and by $\|\cdot\|_{L(\mathbb{H})}$ the usual norm on $L(\mathbb{H})$. A strictly stationary $\mathbb{H}$-valued process $X = (X_n)_{n\in\mathbb{Z}}$ is an autoregressive process of order 1 (AR$\mathbb{H}$(1)) if it satisfies

$$X_n - \mu = \rho(X_{n-1} - \mu) + \xi_n,\ n \in \mathbb{Z}\,, \tag{5.1}$$

where $\mu \in \mathbb{H}$, $\rho \in L(\mathbb{H})$ and $(\xi_n)_{n\in\mathbb{Z}}$ is a strictly stationary sequence of centered random variables in $\mathbb{L}^1_{\mathbb{H}}$. The existence of such a process is ensured by the mild condition (see Bosq (2000))

$$\text{there exists an integer } j_0 \geq 1 \text{ such that } \|\rho^{j_0}\|_{L(\mathbb{H})} < 1. \tag{5.2}$$

Indeed if (5.2) holds then (5.1) has an unique stationary solution given by

$$X_n = \mu + \sum_{j\geq 0} \rho^j(\xi_{n-j}),\ n \in \mathbb{Z}\,,$$

where the series is convergent almost surely and in $\mathbb{L}^1_{\mathbb{H}}$. These types of processes with values in functional spaces facilitate the study of estimation and forecasting problems for several classes of continuous time processes. For more details we refer to Bosq (2000).

Consider an AR$\mathbb{H}$(1) $X = (X_n)_{n\in\mathbb{Z}}$ generated by a sequence $\xi = (\xi_n)_{n\in\mathbb{Z}}$ of i.i.d. random variables in $\mathbb{L}^2_{\mathbb{H}}$. Define the covariance operator C_{X_0} of X_0 by

$$C_{X_0}(x) = \mathbb{E}(\langle X_0, x\rangle_{\mathbb{H}} X_0),\ x \in \mathbb{H}\,.$$

The natural estimator of C_{X_0} is the empirical covariance operator, defined as

$$C_n(x) = \frac{1}{n}\sum_{i=1}^{n} \langle X_i, x\rangle_{\mathbb{H}} X_i,\ x \in \mathbb{H}\,.$$

We wish to study the behaviour of the random variable $C_n - C_{X_0}$ with values in a certain functional space. Such results are useful for the estimation of the eigenelements of C_{X_0} which allows to derive results on the estimation of ρ (see Chapter 8 in Bosq (2000)). The main tool for studying C_n is an autoregressive representation. Set

$$Z_i(x) = \langle X_i, x\rangle_{\mathbb{H}} X_i - C_{X_0}(x),\ x \in \mathbb{H},\ i \in \mathbb{Z}\,. \tag{5.3}$$

Let $\mathcal{S}$ be the space of Hilbert-Schmidt operators on $\mathbb{H}$, that is the space of bounded linear operators s from $\mathbb{H}$ to $\mathbb{H}$ such that the quantity $\|s\|_{\mathcal{S}} = \big(\sum_{1\leq i\leq\infty} \|s(g_i)\|_{\mathbb{H}}^2\big)^{1/2}$ is finite for any orthonormal basis $(g_i)_{i\geq 1}$ of $\mathbb{H}$. This is a separable Hilbert space with respect to the scalar product

$$\langle s_1, s_2\rangle_{\mathcal{S}} = \sum_{1\leq i,j\leq\infty} \langle s_1(g_i), h_j\rangle_{\mathbb{H}} \langle s_2(g_i), h_j\rangle_{\mathbb{H}}\,,$$

where $(g_i)_{i\geq 1}$ and $(h_j)_{j\geq 1}$ are two arbitrary orthonormal bases of $\mathbb{H}$. According to Lemma 4.1 in Bosq (2000), the $\mathcal{S}$-valued process $Z = (Z_n)_{n\in\mathbb{Z}}$ admits the following AR$\mathcal{S}(1)$ representation

$$Z_i = R(Z_{i-1}) + E_i,\ i\in\mathbb{Z}\,, \tag{5.4}$$

where $R\in L(\mathcal{S})$ is defined by $R(s) = \rho s \rho^*$ for all $s\in\mathcal{S}$ and satisfies $\|R^h\|_{L(\mathcal{S})} \leq \|\rho^h\|^2_{L(\mathbb{H})}$, and $E = (E_n)_{n\in\mathbb{Z}}$ is a sequence of $\mathcal{S}$-valued martingale differences with respect to the filtration $\sigma(\xi_k, k\leq n)$. Note that E_n belongs to $\mathbb{L}^p_{\mathcal{S}}$ as soon as ξ belongs to $\mathbb{L}^{2p}_{\mathbb{H}}$. Concerning the behaviour of the random variable $\|C_n - C_{X_0}\|_{\mathcal{S}}$, the following results hold:

Proposition 3. *Let ρ be an operator of $L(\mathbb{H})$ satisfying (5.2). Assume that $(\xi_n)_{n\in\mathbb{Z}}$ is an i.i.d. sequence of random variables in $\mathbb{L}^2_{\mathbb{H}}$. Let $(X_n)_{n\in\mathbb{Z}}$ be the strictly stationary solution of (5.1).*

1. *The sequence $\|C_n - C_{X_0}\|_{\mathcal{S}}$ converges almost surely to 0.*
2. *If $\mathbb{E}\|\xi_0\|^{2p}_{\mathbb{H}} < \infty$ for some p in $]1,2[$, then $n^{(p-1)/p}\|C_n - C_{X_0}\|_{\mathcal{S}}$ converges almost surely to 0.*
3. *If $\mathbb{E}\|\xi_0\|^4_{\mathbb{H}} < \infty$, then $\sqrt{n}\|C_n - C_{X_0}\|_{\mathcal{S}}$ converges in distribution to $\|G\|_{\mathcal{S}}$, where G is an $\mathcal{S}$-valued centered Gaussian random variable.*

Remark 9. Item 3 has been proved by Bosq (2000), Corollary 4.6 and can be proved also by noting that Condition $G(2)$ is satisfied (see Remark 6). The covariance operator of the $\mathcal{S}$-valued Gaussian random variable G is given by Bosq.

Proof of Proposition 3. The fact that $(\xi_n)_{n\in\mathbb{Z}}$ is an i.i.d. sequence ensures the ergodicity of the sequence $(Z_n)_{n\in\mathbb{Z}}$ defined by (5.3). Hence, Item 1 follows from Mourier's ergodic theorem (1953) as quoted in Remark 5. To prove Item 2, note that E_0 belongs to $\mathbb{L}^p_{\mathcal{S}}$ and that

$$\|\mathbb{E}(Z_n|\mathcal{M}_0)\|_{\mathbb{L}^p_{\mathcal{S}}} \leq \|E_0\|_{\mathbb{L}^p_{\mathcal{S}}} \sum_{j\geq n} \|R^j\|_{L(\mathcal{S})}\,.$$

Since (5.2) is satisfied and $\|R^j\|_{L(\mathcal{S})} \leq \|\rho^j\|^2_{L(\mathbb{H})}$, we infer that condition $G(p)$ of Theorem 3 holds. This completes the proof.

6 Appendix

6.1 Proof of Theorem 2.

Proof of Item 1. Define the four variables

$$X_i' = X_i \mathbb{1}\left(\|X_i\|_{\mathbb{B}} \leq n^\alpha\right) \quad \text{and} \quad X_i'' = X_i \mathbb{1}\left(\|X_i\|_{\mathbb{B}} > n^\alpha\right)$$
$$Y_i' = X_i' - \mathbb{E}(X_i'|\mathcal{M}_{i-1}) \quad \text{and} \quad Y_i'' = X_i'' - \mathbb{E}(X_i''|\mathcal{M}_{i-1})\,.$$

Since $(X_i)_{i\geq 1}$ is a sequence of martingale differences, it follows that $X_i = Y_i' + Y_i''$. Hence, for every positive ε,

(6.1)
$$\mathbb{P}\Big(\max_{1\leq k\leq n} \|S_k\|_{\mathbb{B}} \geq 2\varepsilon\, n^\alpha\Big) \leq \mathbb{P}\Big(\max_{1\leq k\leq n}\Big\|\sum_{j=1}^k Y_j'\Big\|_{\mathbb{B}} \geq \varepsilon\, n^\alpha\Big) + \mathbb{P}\Big(\max_{1\leq k\leq n}\Big\|\sum_{j=1}^k Y_j''\Big\|_{\mathbb{B}} \geq \varepsilon\, n^\alpha\Big)\,.$$

Applying Markov's inequality, we obtain that

(6.2)
$$\mathbb{P}\Big(\max_{1\leq k\leq n}\Big\|\sum_{j=1}^k Y_j''\Big\|_{\mathbb{B}} \geq \varepsilon\, n^\alpha\Big) \leq \frac{2}{\varepsilon\, n^\alpha}\sum_{k=1}^n \mathbb{E}(\|X_k''\|_{\mathbb{B}}) = \frac{2}{\varepsilon\, n^\alpha}\sum_{k=1}^n \mathbb{E}(\|X_k\|_{\mathbb{B}} \mathbb{1}\left(\|X_k\|_{\mathbb{B}} > n^\alpha\right))\,.$$

Note that for any $k \geq 1$,

$$\mathbb{E}\|X_k\|_{\mathbb{B}} \mathbb{1}\left(\|X_k\|_{\mathbb{B}} > n^\alpha\right) = \int_0^1 Q_{\|X_k\|_{\mathbb{B}}\mathbb{1}(\|X_k\|_{\mathbb{B}}>n^\alpha)}(u)du \leq \int_0^1 Q_{X\mathbb{1}(X>n^\alpha)}(u)du\,. \tag{6.3}$$

For simplicity, let $Q = Q_X$ and $L = L_X$. From (6.2), (6.3) and (3.13), we infer that

$$\mathbb{P}\Big(\max_{1\leq k\leq n}\Big\|\sum_{j=1}^k Y_j''\Big\|_{\mathbb{B}} \geq \varepsilon\, n^\alpha\Big) \;\leq\; \frac{2\, n^{1-\alpha}}{\varepsilon}\int_0^1 Q(u)\mathbb{1}(u < L(n^\alpha))\,. \tag{6.4}$$

Since $u < L(n^\alpha)$ if and only if $n < Q^{1/\alpha}(u)$, we infer that there exists a finite constant C depending only on p and α such that

$$\begin{aligned}\sum_{n\geq 1} n^{\alpha p-2}\mathbb{P}\Big(\max_{1\leq k\leq n}\Big\|\sum_{j=1}^k Y_j''\Big\|_{\mathbb{B}} \geq \varepsilon\, n^\alpha\Big) \;&\leq\; \frac{2}{\varepsilon}\int_0^1 Q(u)\Big(\sum_{n=1}^{[Q^{1/\alpha}(u)]} n^{\alpha(p-1)-1}\Big)du \\ &\leq\; C\int_0^1 Q^p(u)du = C\,\mathbb{E}(X^p)\,. \end{aligned}\tag{6.5}$$

It remains to control the first term on right hand in (6.1). Applying Doob's inequality to the submartingale $\|\sum_{j=1}^k Y'_j\|_{\mathbb{B}}$ we have that

$$\mathbb{P}\Big(\max_{1\le k\le n}\Big\|\sum_{j=1}^k Y'_j\Big\|_{\mathbb{B}} \ge \varepsilon\, n^{\alpha}\Big) \le \frac{1}{\varepsilon^r n^{r\alpha}}\mathbb{E}\Big(\Big\|\sum_{j=1}^n Y'_j\Big\|^r_{\mathbb{B}}\Big)\,.$$

Since $\mathbb{B}$ is (r,D)-smooth, we obtain, applying Inequality (2.1)

$$\mathbb{P}\Big(\max_{1\le k\le n}\Big\|\sum_{j=1}^k Y'_j\Big\|_{\mathbb{B}} \ge \varepsilon\, n^{\alpha}\Big) \le \frac{D}{\varepsilon^r n^{r\alpha}}\sum_{j=1}^n \mathbb{E}(\|Y'_j\|^r_{\mathbb{B}}) \le \frac{2^r D}{\varepsilon^r n^{r\alpha}}\sum_{j=1}^n \mathbb{E}(\|X'_j\|^r_{\mathbb{B}})\,.$$

Bearing in mind the definition of X'_j we infer that

$$\mathbb{P}\Big(\max_{1\le k\le n}\Big\|\sum_{j=1}^k Y'_j\Big\|_{\mathbb{B}} \ge \varepsilon\, n^{\alpha}\Big) \le \frac{2^r D}{\varepsilon^r n^{r\alpha}}\sum_{i=1}^n \int_0^1 Q^r_{\|X_i\|_{\mathbb{B}}\mathbb{1}(\|X_i\|_{\mathbb{B}}\le n^{\alpha})}(u)du\,. \tag{6.6}$$

Note that for any $A>0$ and any $k\ge 1$, $\|X_k\|_{\mathbb{B}}\mathbb{1}(\|X_k\|_{\mathbb{B}}\le A)\le \|X_k\|_{\mathbb{B}}\wedge A$. Now, for any $u\in[0,1]$,

$$Q_{\|X_k\|_{\mathbb{B}}\mathbb{1}(\|X_k\|_{\mathbb{B}}\le A)}(u) \le Q_{\|X_k\|_{\mathbb{B}}\wedge A}(u) \le Q_{X\wedge A}(u) \le Q(u)\vee A\,. \tag{6.7}$$

From (6.6), (6.7) and the fact that $A<Q(u)$ if and only if $u<L(A)$, we infer that

$$\sum_{n\ge 1} n^{\alpha p-2}\mathbb{P}\Big(\max_{1\le k\le n}\Big\|\sum_{j=1}^k Y'_j\Big\|_{\mathbb{B}} \ge \varepsilon\, n^{\alpha}\Big) \le \frac{2^r D}{\varepsilon^r}(A_1(n)+A_2(n))\,, \tag{6.8}$$

where

$$A_1(n) = \sum_{n\ge 1} n^{\alpha(p+1-r)-1}\int_0^1 Q^{r-1}(u)\mathbb{1}(u<L(n^{\alpha}))du$$
$$A_2(n) = \sum_{n\ge 1} n^{\alpha(p-r)-1}\int_0^1 Q^{r}(u)\mathbb{1}(L(n^{\alpha})\le u\le 1)du\,.$$

Now, since $u\ge L(n^{\alpha})$ if and only if $n\ge Q^{1/\alpha}(u)$, there exists two finite constants K_1 and K_2 depending only on α, p and r, such that

$$A_1(n) \le \int_0^1 Q^{r-1}(u)\Big(\sum_{n=1}^{[Q^{1/\alpha}(u)]} n^{\alpha(p+1-r)-1}\Big)du \le K_1\int_0^1 Q^p(u)du \tag{6.9}$$

$$A_2(n) \le \int_0^1 Q^{r}(u)\Big(\sum_{n\ge Q^{1/\alpha}(u)} n^{\alpha(p-r)-1}\Big)du \le K_2\int_0^1 Q^p(u)du\,, \tag{6.10}$$

Finally, $A_1(n)$ and $A_2(n)$ are finite as soon as $\mathbb{E}(X^p)$ is finite, which completes the proof.

Proof of Item 2. Since $\mathbb{B}$ is super-reflexive, it is r-smooth for some $1 < r \leq 2$. Without loss of generality, assume that $r < 2$. According to Inequality (6.3), for all $k \geq 1$,

$$\mathbb{E}\|X_k\|_{\mathbb{B}} \mathbb{1}\left(\|X_k\|_{\mathbb{B}} > n\right) \leq \mathbb{E}(X \mathbb{1}(X > n))\,.$$

Consequently, by applying Fubini, we infer from (6.2) that if $\mathbb{E}(X \ln^+(X)) < \infty$, then

$$\sum_{n\geq 1} \frac{1}{n} \mathbb{P}\Big(\max_{1\leq k\leq n} \Big\| \sum_{j=1}^{k} Y_j'' \Big\|_{\mathbb{B}} \geq \varepsilon n \Big) < \infty\,.$$

It remains to prove that $\sum_{n\geq 1} n^{-1} \mathbb{P}\big(\max_{1\leq k\leq n} \| \sum_{j=1}^{k} Y_j' \|_{\mathbb{B}} \geq \varepsilon n \big) < \infty$. Starting from (6.8) with $\alpha = p = 1$ and applying Inequalities (6.9) and (6.10), we infer that $A_1(n)$ and $A_2(n)$ are finite as soon as $\mathbb{E}(X)$ is finite, which completes the proof.

6.2 Comparison of the conditions $G(p)$, (3.4) and $DM(p, \gamma, X)$.

In this section, we consider a stationary sequence $(X_i)_{i\in\mathbb{Z}}$ with values in $\mathbb{L}^p_{\mathbb{H}}$ with $p \in]1, 2]$, such that X_0 is centered and $\mathcal{M}_0$-measurable. Let $X = \|X_0\|_{\mathbb{H}}$.

Proof of the implication : (3.4) $\Rightarrow$ $G(p)$**.** Notice first that $G(p)$ can be rewritten as

$$\lim_{m\to\infty} \sup_{n>m} \Big\| \mathbb{E}\Big(\sum_{i=m+1}^{n} X_i \Big| \mathcal{M}_0 \Big) \Big\|_{\mathbb{L}^p_{\mathbb{H}}} = 0\,. \tag{6.11}$$

Since $p > 1$, using first the duality between $\mathbb{L}^p_{\mathbb{H}}$ and $\mathbb{L}^q_{\mathbb{H}}$ for $q = (p-1)/p$, followed by the inequality (3.29) in Dedecker and Merlevède (2003), we successively have

$$\begin{aligned} \Big\| \mathbb{E}\Big(\sum_{i=m+1}^{n} X_i \Big| \mathcal{M}_0 \Big) \Big\|_{\mathbb{L}^p_{\mathbb{H}}} &= \sup_{Y\in\mathcal{M}_0, \|Y\|_{\mathbb{L}^q_{\mathbb{H}}}\leq 1} \mathbb{E}\Big\langle Y, \sum_{i=m+1}^{n} X_i \Big\rangle_{\mathbb{H}} \\ &\leq \sup_{\|Y\|_{\mathbb{L}^q_{\mathbb{H}}}\leq 1} \sum_{i=m+1}^{n} \int_0^1 Q_Y(u) Q_X(u) \mathbb{1}_{u\leq G_X(\gamma_i)} du\,. \end{aligned}$$

Then Hölder's inequality yields

$$\Big\| \mathbb{E}\Big(\sum_{i=m+1}^{n} X_i \Big| \mathcal{M}_0 \Big) \Big\|_{\mathbb{L}^p_{\mathbb{H}}} \leq \sup_{\|Y\|_{\mathbb{L}^q_{\mathbb{H}}}\leq 1} \Big(\int_0^1 Q_Y^q(u) \Big)^{1/q} \Big(\int_0^1 Q_X^p(u) \Big(\sum_{i=m+1}^{n} \mathbb{1}_{u\leq G_X(\gamma_i)} \Big)^p du \Big)^{1/p}\,.$$

Since

$$\Big(\sum_{i=m+1}^{\infty}\mathbb{I}_{u\le G_X(\gamma_i)}\Big)^p \le \sum_{j=m+1}^{\infty}(j+1)^p\,\mathbb{I}_{G_X(\gamma_{j+1})\le u\le G_X(\gamma_j)}$$

and since, for $1<p\le 2$, $(j+1)^p \le 2\sum_{\ell=0}^{j}(\ell+1)^{p-1}$, we get:

$$\Big(\sum_{i=m+1}^{\infty}\mathbb{I}_{u\le G_X(\gamma_i)}\Big)^p \le 2\sum_{j=m+1}^{\infty}\Big(\sum_{\ell=0}^{j}(\ell+1)^{p-1}\Big)\mathbb{I}_{G_X(\gamma_{j+1})\le u\le G_X(\gamma_j)}\,.$$

Then,

$$\begin{aligned}\Big\|\mathbb{E}\Big(\sum_{i=m+1}^{n}X_i\Big|\mathcal{M}_0\Big)\Big\|_{\mathbb{L}^p_{\mathbb{H}}} &\le 2^{1/p}\left(\sum_{i=0}^{\infty}(i+1)^{p-1}\int_0^{G_X(\gamma_i)\wedge G_X(\gamma_{m+1})}Q_X^p(u)du\right)^{1/p}\\ &\le 2^{1/p}\left(\sum_{i=0}^{\infty}(i+1)^{p-1}\int_0^{\gamma_i\wedge\gamma_{m+1}}Q_X^{p-1}\circ G_X(u)du\right)^{1/p},\end{aligned}$$

which implies that (6.11) holds (and hence $G(p)$) as soon as (3.4) does.

Some examples showing that $G(p)$ and $DM(p,\gamma,X)$ cannot be compared. Consider the simple case $X_n=\sum_{i\ge 0}a_i\epsilon_{n-i}$, where $(\epsilon_i)_{i\in\mathbb{Z}}$ is a sequence of iid real-valued random variables with distribution $\mathcal{N}(0,1)$, and $(a_i)_{i\ge0}$ is a sequence of real numbers such that $\sum_{i\ge0}a_i^2<\infty$. Let $\mathcal{M}_i$ be the natural σ-algebra, that is $\mathcal{M}_i=\sigma(\epsilon_j, j\le i)$. Since $\mathbb{E}(X_n|\mathcal{M}_0)$ is gaussian with mean 0 and variance $\sum_{i\ge n}a_i^2$, one has that

$$\gamma_n=\sqrt{\frac{2}{\pi}\sum_{i\ge n}a_i^2}\,.$$

On an other hand, since $(\mathbb{E}(X_n|\mathcal{M}_0))_{n\ge0}$ is a Gaussian process, it follows that for any p in $]1,2[$, the condition $G(p)$ is equivalent to $G(2)$. Then according to the condition (5.37) in Hall and Heyde (1980), we derive that for any p in $]1,2[$, the condition $G(p)$ holds if and only if

$$\sum_{k=0}^{n}a_k \quad \text{converges, and} \quad \sum_{i\ge0}\Big(\sum_{k=i}^{\infty}a_k\Big)^2<\infty\,.$$

Consequently, if $a_i=(-1)^i(\sqrt{i+1}\ln(i+1))^{-1}$, then the condition $G(p)$ holds for any p in $]1,2[$, while $\sum_{i>0}\gamma_i/i=\infty$, so that $DM(p,\gamma,X)$ cannot holds for any p in $]1,2[$. On the contrary, if $a_i=(i+1)^{-3/2}$, then $G(p)$ does not hold for any p in $]1,2[$, while $DM(p,\gamma,X)$ holds for any p in $]1,2[$ (to see this, apply Corollary 1 item 2. with $r>1/(2-p)$).

6.3 Proof of Proposition 1.

We need the following preliminary result.

Proposition 4. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $(X_k)_{1\le k\le n}$ be n random variables with values in a (r, D)-smooth separable Banach space $\mathbb{B}$ such that $\mathbb{P}(\|X_k\|_{\mathbb{B}} \le T) = 1$ for any $1 \le k \le n$. For $i \le n$, define the σ-algebras $\mathcal{M}_i$ by: $\mathcal{M}_i = \{\emptyset, \Omega\}$ if $i \le 0$ and $\mathcal{M}_i = \sigma(X_j, 1 \le j \le i)$ if $1 \le i \le n$. Let $S_0 = 0_{\mathbb{B}}$ and $S_k = \sum_{i=1}^k (X_i - \mathbb{E}(X_i))$. Define the random variables U_i by: $U_{[n/q]+1} = S_n - S_{q[n/q]}$ and $U_i = S_{iq} - S_{iq-q}$ for $1 \le i \le [n/q]$. For any $x \ge Tq$, the following inequality holds*

$$\mathbb{P}\Big(\max_{1\le k\le n} \|S_k\|_{\mathbb{B}} \ge 4x\Big) \le \frac{D}{x^r} \sum_{i=1}^{[n/q]+1} \mathbb{E}(\|U_i - \mathbb{E}(U_i|\mathcal{M}_{(i-2)q})\|_{\mathbb{B}}^r) + \frac{1}{x} \sum_{i=3}^{[n/q]+1} \mathbb{E}(\|\mathbb{E}(U_i|\mathcal{M}_{(i-2)q})\|_{\mathbb{B}}). \tag{6.12}$$

Proof of Proposition 4. Any integer j being distant from at most $[q/2]$ of an element of $q\mathbb{N}$, we have that

$$\max_{1\le k\le n} \|S_k\|_{\mathbb{B}} \le 2[q/2]T + \max_{1\le j\le [n/q]+1} \Big\| \sum_{i=1}^j U_i \Big\|_{\mathbb{B}}.$$

Hence Inequality (6.12) follows from the bound

$$\mathbb{P}\Big(\max_{1\le j\le [n/q]+1} \Big\| \sum_{i=1}^j U_i \Big\|_{\mathbb{B}} \ge 3x\Big) \le \frac{D}{x^r} \sum_{i=1}^{[n/q]+1} \mathbb{E}(\|U_i - \mathbb{E}(U_i|\mathcal{M}_{(i-2)q})\|_{\mathbb{B}}^r) + \frac{1}{x} \sum_{i=3}^{[n/q]+1} \mathbb{E}(\|\mathbb{E}(U_i|\mathcal{M}_{(i-2)q})\|_{\mathbb{B}}). \tag{6.13}$$

Consider the σ-algebras $\mathcal{F}_i^U = \mathcal{M}_{iq}$ and define the variables $\tilde{U}_i$ as follows: $\tilde{U}_1 = U_1$, and $\tilde{U}_{2i-1} = U_{2i-1} - \mathbb{E}(U_{2i-1}|\mathcal{F}^U_{2(i-1)-1})$ for $i > 1$, $\tilde{U}_2 = U_2$ and $\tilde{U}_{2i} = U_{2i} - \mathbb{E}(U_{2i}|\mathcal{F}^U_{2(i-1)})$ for $i > 1$. Substituting $\tilde{U}_i$ to U_i, we obtain the inequality

$$\max_{1\le j\le [n/q]+1} \Big\| \sum_{i=1}^j U_i \Big\|_{\mathbb{B}} \le \max_{2\le 2j\le [n/q]+1} \Big\| \sum_{i=1}^j \tilde{U}_{2i} \Big\|_{\mathbb{B}} + \max_{1\le 2j-1\le [n/q]+1} \Big\| \sum_{i=1}^j \tilde{U}_{2i-1} \Big\|_{\mathbb{B}} + \sum_{i=1}^{[n/q]+1} \|U_i - \tilde{U}_i\|_{\mathbb{B}}. \tag{6.14}$$

Note that $(\tilde{U}_{2i})_{i\geq 1}$ (resp. $(\tilde{U}_{2i-1})_{i\geq 1}$) is a $\mathbb{B}$-valued martingale difference sequence with respect to the filtration $(\mathcal{F}^U_{2i})_{i\geq 1}$ (resp. $(\mathcal{F}^U_{2i-1})_{i\geq 1}$). Applying Doob's inequality, to the submartingale $\| \sum_{i=1}^j \tilde{U}_{2i}\|_{\mathbb{B}}$ and next Inequality (2.1), we obtain that
(6.15)

$$\mathbb{P}\Big(\max_{2\leq 2j\leq [n/q]+1}\Big\|\sum_{i=1}^{j}\tilde{U}_{2i}\Big\|_{\mathbb{B}}\geq x\Big)\leq \frac{1}{x^r}\mathbb{E}\Big(\Big\|\sum_{i=1}^{([n/q]+1)/2}\tilde{U}_{2i}\Big\|^r_{\mathbb{B}}\Big)\leq \frac{D}{x^r}\sum_{i=1}^{([n/q]+1)/2}\mathbb{E}(\|\tilde{U}_{2i}\|^r_{\mathbb{B}})\,.$$

In the same way

$$\mathbb{P}\Big(\max_{1\leq 2j-1\leq [n/q]+1}\Big\|\sum_{i=1}^{j}\tilde{U}_{2i-1}\Big\|_{\mathbb{B}}\geq x\Big)\leq \frac{D}{x^r}\sum_{i=1}^{[n/q]/2+1}\mathbb{E}(\|\tilde{U}_{2i-1}\|^r_{\mathbb{B}})\,. \tag{6.16}$$

On the other hand, we have

$$\mathbb{P}\Big(\sum_{i=1}^{[n/q]+1}\|U_i-\tilde{U}_i\|_{\mathbb{B}}\geq x\Big)\leq \frac{1}{x}\sum_{i=3}^{[n/q]+1}\mathbb{E}\|\mathbb{E}(U_i|\mathcal{M}_{(i-2)q})\|_{\mathbb{B}}\,. \tag{6.17}$$

Inequality (6.13) follows from (6.14), (6.15), (6.16) and (6.17). ■

End of the proof of Proposition 1. For every $v\in[0,1]$, we introduce the variables

$$X'_i = X_i \mathbb{1}\,(\|X_i\|_{\mathbb{H}}\leq Q(v)) \ \text{ and } X''_i = X_i\mathbb{1}\,(\|X_i\|_{\mathbb{H}}> Q(v))\,.$$

Let $S'_n=\sum_{i=1}^n(X'_i-\mathbb{E}X'_i)$ and $S''_n=\sum_{i=1}^n(X''_i-\mathbb{E}X''_i)$ and write

$$\mathbb{P}\Big(\max_{1\leq k\leq n}\|S_k\|_{\mathbb{H}}\geq 5\,x\Big)\ \leq\ \mathbb{P}\Big(\max_{1\leq k\leq n}\|S'_k\|_{\mathbb{H}}\geq 4\,x\Big)+\mathbb{P}\Big(\max_{1\leq k\leq n}\|S''_k\|_{\mathbb{H}}\geq x\Big)\,. \tag{6.18}$$

Arguing as for (6.4) and using the fact that

$$u<L(Q(v)) \iff Q(v)<Q(u) \iff u<v\,,$$

we obtain that

$$\mathbb{P}\Big(\max_{1\leq k\leq n}\|S''_k\|_{\mathbb{H}}\geq x\Big)\ \leq\ \frac{2n}{x}\int_0^v Q(u)du\,. \tag{6.19}$$

To control the first term in decomposition (6.18), we apply Inequality (6.12) with $r=2$ and $D=1$. Since in Hilbert spaces, $\mathbb{E}(\|U'_i-\mathbb{E}(U'_i|\mathcal{M}_{(i-2)q})\|^2_{\mathbb{H}})\leq \mathbb{E}\|U'_i\|^2_{\mathbb{H}}$, we get that for any integer q smaller than n and any $x\geq qQ(v)$,

$$\mathbb{P}\Big(\max_{1\leq k\leq n}\|S'_k\|_{\mathbb{H}}\geq 4x\Big)\leq \frac{1}{x^2}\sum_{i=1}^{[n/q]+1}\mathbb{E}\|U'_i\|^2_{\mathbb{H}}+\frac{1}{x}\sum_{i=3}^{[n/q]+1}\mathbb{E}\|\mathbb{E}(U'_i|\mathcal{M}_{(i-2)q})\|_{\mathbb{H}}\,, \tag{6.20}$$

where $U'_i = S'_{iq} - S'_{iq-q}$ for $1 \le i \le [n/q]$, and $U'_{[n/q]+1} = S'_n - S'_{q[n/q]}$. Define the random variables

$$Y_k = X_k - \mathbb{E}X_k \, , \; Y'_k = X'_k - \mathbb{E}X'_k \text{ and } Y''_k = X''_k - \mathbb{E}X''_k \, .$$

For all $1 \le i \le [n/q]$, we have that

$$\mathbb{E}\|U'_i\|^2_{\mathbb{H}} = \sum_{j=(i-1)q+1}^{iq} \sum_{\ell=(i-1)q+1}^{iq} \mathbb{E}\langle Y'_\ell, Y'_j\rangle_{\mathbb{H}} = \sum_{j=(i-1)q+1}^{iq} \sum_{\ell=(i-1)q+1}^{iq} \mathbb{E}\langle X'_\ell, Y'_j\rangle_{\mathbb{H}} \, .$$

Let $1 \le i \le [n/q]$. Clearly

$$\mathbb{E}\|U'_i\|^2_{\mathbb{H}} \le 2 \sum_{j=(i-1)q+1}^{iq} \sum_{\ell=(i-1)q+1}^{j} \mathbb{E}\big(\|X'_\ell\|_{\mathbb{H}} \|\mathbb{E}(Y'_j|\mathcal{M}_\ell)\|_{\mathbb{H}}\big) \le A_1 + A_2 \, , \tag{6.21}$$

where

$$\begin{aligned} A_1 &= \sum_{j=(i-1)q+1}^{iq} \sum_{\ell=(i-1)q+1}^{j} 2\mathbb{E}\big(\|X'_\ell\|_{\mathbb{H}} \|\mathbb{E}(Y_j|\mathcal{M}_\ell)\|_{\mathbb{H}}\big) \\ A_2 &= \sum_{j=(i-1)q+1}^{iq} \sum_{\ell=(i-1)q+1}^{j} 2\mathbb{E}\big(\|X'_\ell\|_{\mathbb{H}} \|\mathbb{E}(Y''_j|\mathcal{M}_\ell)\|_{\mathbb{H}}\big) \end{aligned}$$

Since $\|X'_k\|_{\mathbb{H}} \le Q(v)$, it follows that

$$\begin{aligned} \frac{A_2}{2Q(v)} &\le \sum_{j=(i-1)q+1}^{iq} \sum_{\ell=(i-1)q+1}^{j} \mathbb{E}\|\mathbb{E}(Y''_j|\mathcal{M}_\ell)\|_{\mathbb{H}} \le \sum_{j=(i-1)q+1}^{iq} (j-(i-1)q)\mathbb{E}\|Y''_j\|_{\mathbb{H}} \\ &\le 2 \sum_{j=(i-1)q+1}^{iq} (j-(i-1)q)\mathbb{E}\|X''_j\|_{\mathbb{H}} \, . \end{aligned}$$

Next arguing as in (6.3) and using (3.13), we easily obtain that for all $j \ge 1$,

$$\mathbb{E}\|X''_j\|_{\mathbb{H}} \le \int_0^v Q(u)du \, . \tag{6.22}$$

It follows that

$$A_2 \le 2q(q+1)Q(v) \int_0^v Q(u)du \, . \tag{6.23}$$

On the other hand, applying Inequality (3.29) in Dedecker and Merlevède (2003), we infer that

$$A_1 \le 2 \sum_{j=(i-1)q+1}^{iq} \sum_{\ell=(i-1)q+1}^{j} \int_0^{\gamma_{j-\ell}} Q_{\|X'_\ell\|_{\mathbb{H}}} \circ G_{\|Y_j\|_{\mathbb{H}}}(u)du \, . \tag{6.24}$$

Note that, for all $j \geq 1$, $Q_{\|Y_j\|_{\mathbb{H}}} \leq Q_{\|X_j\|_{\mathbb{H}}} + \mathbb{E}\|X_j\|_{\mathbb{H}}$. Arguing as in (4.4), we infer that $G_{\|Y_j\|_{\mathbb{H}}}(u) \geq G(u/2)$. From this inequality and (6.24), we obtain the bound

$$A_1 \leq 4 \sum_{j=(i-1)q+1}^{iq} \sum_{\ell=(i-1)q+1}^{j} \int_0^{\frac{\gamma_{j-\ell}}{2}} Q_{\|X'_\ell\|_{\mathbb{H}}} \circ G(u)du\,.$$

Since H is absolutely continuous and monotonic, we can make the change-of-variables $v = H(u)$ (see Theorem 6.26 in Rudin (1986) and the example given page 156), which yields

$$A_1 \leq 4 \sum_{j=(i-1)q+1}^{iq} \sum_{\ell=(i-1)q+1}^{j} \int_0^{G(\gamma_{j-\ell}/2)} Q_{\|X'_\ell\|_{\mathbb{H}}}(u)Q(u)du\,.$$

Applying (6.7), we infer that for any $\ell \geq 1$ and any $u \in [0,1]$,

$$Q_{\|X'_\ell\|_{\mathbb{H}}}(u) \leq Q(v \vee u)\,. \tag{6.25}$$

Using (6.25), we get that

$$A_1 \quad \leq \quad 4\sum_{k=0}^{q-1}(q-k)\int_0^{G(\gamma_k/2)} Q(v \vee u)Q(u)du\,. \tag{6.26}$$

Starting from (6.21) and collecting (6.23) and (6.26), we infer that for all $1 \leq i \leq [n/q]$,

$$\mathbb{E}\|U'_i\|^2_{\mathbb{H}} \leq 4q\Big(2qQ(v)\int_0^v Q(u)du + \int_0^1 \big((\gamma/2)^{-1} \circ G^{-1}\big)(u)Q^2(u)\mathbb{1}(v \leq u \leq 1)du\Big)\,. \tag{6.27}$$

On the other hand, for all $3 \leq i \leq [n/q]$,

$$\mathbb{E}\|\mathbb{E}(U'_i|\mathcal{M}_{(i-2)q})\|_{\mathbb{H}} \leq \mathbb{E}\|\mathbb{E}(U_i|\mathcal{M}_{(i-2)q})\|_{\mathbb{H}} + 2\sum_{k=(i-1)q+1}^{iq} \mathbb{E}\|X''_k\|_{\mathbb{H}}\,. \tag{6.28}$$

Applying Inequality (6.22), we obtain that

$$\sum_{k=(i-1)q+1}^{iq} \mathbb{E}\|X''_k\|_{\mathbb{H}} \leq q\int_0^v Q(u)du\,.$$

According to the definition of the coefficients γ_i, we have that

$$\mathbb{E}\|\mathbb{E}(U_i|\mathcal{M}_{(i-2)q})\|_{\mathbb{H}} \leq q\gamma_q = 2qH \circ G(\gamma_q/2) = 2q\int_0^{G(\gamma_q/2)} Q(u)du\,.$$

Starting from (6.28), we infer that for all $3 \le i \le [n/q]$,

$$\mathbb{E}\|\mathbb{E}(U_i'|\mathcal{M}_{(i-2)q})\|_{\mathbb{H}} \le 2q\Big(\int_0^{G(\gamma_q/2)} Q(u)du + \int_0^v Q(u)du\Big)\,. \tag{6.29}$$

The terms involving the quantity $U'_{[n/q]+1}$ are treated similarly: we obtain the same bound as (6.27) and (6.29) but with $n - q[n/q]$ instead of q. Starting from Inequality (6.18) and collecting (6.19), (6.20), (6.27) and (6.29), we obtain that for any $v \in [0,1]$, any positive integer q smaller than n and any $x \ge qQ(v)$,

$$\begin{aligned}\mathbb{P}\Big(\max_{1\le k\le n}\|S_k\|_{\mathbb{H}} \ge 5\,x\Big) \;\le\; & \frac{12n}{x}\int_0^v Q(u)du + \frac{2n}{x}\int_0^{G(\gamma_q/2)} Q(u)du \\ & + \frac{4n}{x^2}\int_v^1 \big((\gamma/2)^{-1}\circ G^{-1}\big)(u)Q^2(u)du\,.\end{aligned} \tag{6.30}$$

Now choose $v = S(x/r)$ and $q = ((\gamma/2)^{-1}\circ G^{-1}(v))\wedge n$. This choice implies that $G(\gamma_q/2) \le v$ and that

$$qQ(v) = R(v) = R(S(x/r)) \le x/r \le x\,.$$

Applying Inequality (6.30), we obtain the desired result. ■

6.4 Proof of Lemma 1.

Take $f \in W_1(\mu)$. We first check that under (4.1), $|f|$ is integrable with respect to dF. Without loss of generality, assume that $f(0) = 0$. Clearly

$$\int |f|dF \le \int_{\mathbb{R}^+}\Big(\int_{[0,t[}|g(x)|\mu(dx)\Big)dF(t) + \int_{\mathbb{R}^-}\Big(\int_{[t,0[}|g(x)|\mu(dx)\Big)dF(t)\,.$$

Applying Fubini, we obtain that

$$\int |f|dF \le \int_{\mathbb{R}^+}|g(x)|(1-F(x))\mu(dx) + \int_{\mathbb{R}^-}|g(x)|F(x)\mu(dx)\,.$$

Since g belongs to $\mathbb{L}^2(\mu)$, the right hand side is finite as soon as (4.1) holds. In the same way, we have both

$$\begin{aligned}\int f dF &= \int_{\mathbb{R}^+} g(x)(1-F(x))\mu(dx) - \int_{\mathbb{R}^-} g(x)F(x)\mu(dx)\\ \int f dG &= \int_{\mathbb{R}^+} g(x)(1-G(x))\mu(dx) - \int_{\mathbb{R}^-} g(x)G(x)\mu(dx).\end{aligned}$$

Consequently

$$\int f dF - \int f dG = \int g(x)(G(x) - F(x))\mu(dx)\,.$$

The result follows by noting that

$$D(F, G, \mu) = \sup_{\|g\|_{\mathbb{L}^2(\mu)} \leq 1} \Big| \int_{\mathbb{R}} g(x)(F(x) - G(x)) d\mu(x) \Big|\,. \quad \blacksquare$$

Acknowledgment. The authors are indebted to the referee for carefully reading the manuscript and for helpful comments.

References

[1] de Acosta, A. (1981). Inequalities for B-valued random vectors with applications to the strong law of large numbers, *Ann. Probab.* **9** 157-161.

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

[3] Baum, L.B. and Katz, M. (1965). Convergence rates in the law of large numbers, *Trans. Amer. Math. Soc.* **120** 108-123.

[4] Berbee, H.C.P. (1987). Convergence rates in the strong law for bounded mixing sequences, *Probab. Theory Related Fields* **74** 255-270.

[5] Bosq, D. (2000). Linear Processes in Function Spaces. Theory and Applications. Lecture Notes in Statistics **149** Springer.

[6] Bradley, R.C. (2002). *Introduction to Strong Mixing Conditions, Volume 1.* Technical Report, Department of Mathematics, Indiana University, Bloomington. Custom Publishing of I.U., Bloomington, March 2002.

[7] Dedecker, J. and Doukhan, P. (2003) A new covariance inequality and applications, *Stoch. Processes Appl.* **106** 63-80.

[8] Dedecker, J. and Merlevède, F. (2003). The conditional central limit theorem in Hilbert spaces, *Stoch. Processes Appl.* **108** 229-262.

[9] Dedecker, J. and Prieur, C. (2005). New dependence coefficients. Examples and applications to statistics, *Probab. Theory Relat. Fields* **132** 203-236.

[10] Elton, J. (1981). A law of large numbers for identically distributed martingale differences, *Ann. Probab.* **9** 405-412.

[11] Gordin, M. I. (1969). The central limit theorem for stationary processes, *Dokl. Akad. Nauk SSSR* **188** 739-741.

[12] Hall, P. and Heyde, C. C. (1980). *Martingale Limit Theory and its Applications.* Academic Press, New York-London.

[13] Hipp, C. (1979). Convergence rates of the strong law for stationary mixing sequences, *Z. Wahrsch. Verw. Gebiete* **49** 49-62.

[14] Lai, T.L. (1977). Convergence rates and r-quick versions of the strong law for stationary mixing sequences, *Ann. Probab.* **5** 693-706.

[15] Lesigne, E. and Volný, D. (2001). Large deviations for martingales, *Stoch. Processes Appl.* **96** 143-159.

[16] McLeish, D. L. (1975). A maximal inequality and dependent strong laws. *Ann. Probab.* **3** 829-839.

[17] Merlevède, F. (2003). On the central limit theorem and its weak invariance principle for strongly mixing sequences with values in a Hilbert space via martingale approximation, *J. Theor. Probab.* **16** 625-653.

[18] Peligrad, M. (1985). Convergence rates of the strong law for stationary mixing sequences. *Z. Wahrsch. Verw. Gebiete* **70** 307-314.

[19] Peligrad, M. (1989). The r-quick version of the strong law for stationary φ-mixing sequences. *Almost everywhere convergence (Columbus, OH, 1988)*, 335–348, Academic Press, Boston, MA.

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

[21] Rio, E. (1995). A maximal inequality and dependent Marcinkiewicz-Zygmund strong laws. *Ann. Probab.*, **23** 918-937.

[22] Rio, E. (2000). *Théorie asymptotique des processus aléatoires faiblement dépendants. Mathématiques et applications de la SMAI.* **31** Springer.

[23] Rosenblatt, M. (1956). A central limit theorem and a strong mixing condition. *Proc. Nat. Acad. Sci.*, **42** 43-47.

[24] Rudin, W. (1987). *Real and complex analysis.* Third edition. McGraw-Hill Book Co., New York.

[25] Shao, Q-M. (1993). Complete convergence for α-mixing sequences. *Statist. Probab. Letters*, **16** 279-287.

[26] Woyczyński, W.A. (1975a). A central limit theorem for martingales in Banach spaces. *Bull. Acad. Polon. Sci. Sr. Sci. Math. Astronom. Phys.*, **23**, no. 8, 917-920.

[27] Woyczyński, W.A. (1975b). Laws of large numbers for vector-valued martingales. *Bull. Acad. Polon. Sci. Sr. Sci. Math. Astronom. Phys.*, **23**, no. 11, 1199-1201.

[28] Woyczyński, W.A. (1981). Asymptotic behavior of martingales in Banach spaces. II. *Martingale theory in harmonic analysis and Banach spaces (Cleveland, Ohio, 1981)*, pp. 216-225, Lecture Notes in Math., 939, Springer, Berlin-New York, 1982.