Inequalities for partial sums of Hilbert valued dependent sequences and applications.

Jérôme Dedecker* and Florence Merlevède†
Université Paris VI

Abstract

By using coupling arguments, we prove a Fuk-Nagaev inequality for the deviation of the maximum of partial sums of Hilbert valued random variables. The upper bound is expressed in terms of some dependence coefficients which naturally appear when using such coupling arguments. These coefficients may be computed in many situations. In particular, we show that they are well adapted to functions of mixing sequences, iterated random functions, and a class of expanding maps of $[0,1]^k$. We apply our maximal inequality to obtain almost sure convergence results for the partial sums, such as complete convergence or compact law of the iterated logarithm. An application to Cramér-von Mises statistics is given.

Mathematics Subject Classifications (2000): 60 B 12, 60 F 15, 60 F 17.
Key words: Almost sure convergence, almost sure invariance principle, weak dependence, maximal inequalities, coupling, Hilbert space.
Short Title: Maximal inequalities in Hilbert 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, 75 013 Paris, FRANCE. Email: merleve@ccr.jussieu.fr.

1 Introduction

Let $(X_k)_{k>0}$ be a sequence of centered random variables with values in a separable Hilbert space $(\mathbb{H}, \|\cdot\|_{\mathbb{H}})$, and let $S_n = X_1 + \cdots + X_n$. In this paper we are interested in the almost sure behavior of the sequence S_n. To be more precise we are interested by sufficient conditions under which, for $1/2 < \alpha \leq 1$ and $1/\alpha \leq p < \infty$,

$$\sum_{n=1}^{\infty} n^{\alpha p-2} \mathbb{P}\Big(\max_{1 \leq k \leq n} \|S_k\|_{\mathbb{H}} \geq \varepsilon\, n^{\alpha} \Big) < \infty, \text{ for all } \varepsilon > 0, \tag{1.1}$$

or sufficient conditions under which

$$\text{the sequence } \frac{S_n}{\sqrt{n \ln \ln n}} \text{ is almost surely relatively compact.} \tag{1.2}$$

To address these problems, one of the main tools is to establish suitable upper bounds for the quantity

$$\mathbb{P}\left(\max_{1 \leq k \leq n} \|S_k\|_{\mathbb{H}} \geq x \right). \tag{1.3}$$

For instance, applying Proposition 3.5 in Dedecker and Merlevède (2004), we derive sharp sufficient conditions for (1.1) in the case where $p \in [1, 2[$. In this note, the dependence conditions are expressed in terms of conditional expectations, and are weaker than Rio's conditions (1995) for strongly mixing sequences. Note that the maximal inequality stated in our Proposition 3.5 is obtained via martingale approximations, and cannot be applied to prove (1.1) for $p \geq 2$.

A suitable tool to prove (1.1) for $p \geq 2$ as well as the compact law (1.2), is a Fuk-Nagaev inequality for (1.3). A way to prove it, is to use approximations by independent random variables instead of martingale approximations, as done in Rio (2000, Theorem 6.2) for real valued random variables. Since the works of Berbee (1979) and Rüschendorf (1985), we know that the price to pay for such approximations with respect to a distance d on $\mathbb{H}$ is exactly the value of some dependence coefficients having the *coupling property* for d (*i.e.* the property (2.3) of Lemma 1). To control (1.3), the appropriate distance is $d(x, y) = \|x - y\|_{\mathbb{H}}$, and the appropriate measure of dependence is the coefficient τ_d defined in Definition 1. Following the proofs of Theorem 6.2 in Rio (2000) and Theorem 2 in Dedecker and Prieur (2004), and using a delicate truncation argument for $\mathbb{H}$-valued random variables, we obtain the Fuk-Nagaev type inequality given in (4.28). From this inequality, we derive sharp sufficient conditions for (1.1) and for the compact law (1.2) in

Theorems 2 and 3 respectively. The optimality of these conditions is discussed in Sections 5 and 6, and an application to Cramér-von Mises statistics is given in Section 7.

One of the main interests of the coefficients considered in this paper, is that they can be computed in many situations. In particular, we show in Section 3 that they are well adapted to functions of mixing sequences, iterated random functions, and a class of iterates of expanding maps of $[0, 1]^k$. Note that it is well known that the processes defined as iterates of maps cannot be mixing in the sense of Rosenblatt (1956).

2 Definitions and Properties

Definition 1. Let $\mathcal{X}$ be a Polish space and let d be a distance on $\mathcal{X}$ (the space $\mathcal{X}$ need not be Polish with respect to d). Let $\Lambda_1(\mathcal{X}, d)$ be the set of 1-Lipschitz functions from $\mathcal{X}$ to $\mathbb{R}$ with respect to d. Assume that the distance d satisfies

$$d(x, y) = \sup_{f \in \Lambda_1(\mathcal{X},d)} |f(x) - f(y)| . \tag{2.1}$$

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. Let $p \geq 1$. We say that a random variable X with values in $\mathcal{X}$ belongs to $\mathbb{L}^p(\mathcal{X}, d)$ if the variable $d(X, s)$ belongs to $\mathbb{L}^p(\mathbb{R})$ for some (and therefore any) s in $\mathcal{X}$. For any random variable X in $\mathbb{L}^p(\mathcal{X}, d)$ and any σ-algebra $\mathcal{M}$ of $\mathcal{A}$, let $\mathbb{P}_{X|\mathcal{M}}$ be a conditional distribution of X given $\mathcal{M}$ and let $\mathbb{P}_X$ be the distribution of X. We consider the coefficient $\tau_{d,p}(\mathcal{M}, X)$ of weak dependence (introduced for $p = 1$ by Dedecker and Prieur (2005)) which is defined by

$$\tau_{d,p}(\mathcal{M}, X) = \Big\| \sup_{f \in \Lambda_1(\mathcal{X},d)} \Big| \int f(x) \mathbb{P}_{X|\mathcal{M}}(dx) - \int f(x) \mathbb{P}_X(dx) \Big| \Big\|_p , \tag{2.2}$$

When $p = 1$, we write $\tau_d(\mathcal{M}, X)$ in place of $\tau_{d,1}(\mathcal{M}, X)$ and when $p = \infty$, we write $\varphi_d(\mathcal{M}, X)$ in place of $\tau_{d,\infty}(\mathcal{M}, X)$.

The coefficient $\tau_{d,p}(\mathcal{M}, X)$ has an interpretation in terms of coupling. The following result can be deduced from Proposition 4 in Rüschendorf (1985) (see also Dedecker *et al.* (2005), Theorem 2).

Lemma 1. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, X a random variable with values in some Polish space $\mathcal{X}$, and $\mathcal{M}$ a σ-algebra of $\mathcal{A}$. Let d be a distance on $\mathcal{X}$ satisfying (2.1). Assume that there exists a random variable U uniformly distributed over $[0, 1]$, independent of $\sigma(X) \vee \mathcal{M}$. Then there exists a random variable X^* measurable with* *respect to $\sigma(U) \vee \sigma(X) \vee \mathcal{M}$, distributed as X and independent of $\mathcal{M}$, such that for any $p \geq 1$,*

$$(2.3) \qquad \tau_d(\mathcal{M}, X) = \mathbb{E}(d(X, X^*)) \ \textit{ and } \ \tau_{d,p}(\mathcal{M}, X) = \|\mathbb{E}(d(X, X^*)|\mathcal{M})\|_p \,.$$

The coefficients $\tau_{d,p}(\mathcal{M}, X)$ can be compared to other dependence coefficients. Let $\mathcal{B}(\mathcal{X})$ be the class of Borel sets of $\mathcal{X}$, and define

$$(2.4) \qquad \beta_p(\mathcal{M}, \sigma(X)) \;=\; \Big\| \sup_{A \in \mathcal{B}(\mathcal{X})} |\mathbb{P}_{X|\mathcal{M}}(A) - \mathbb{P}_X(A)| \Big\|_p = \frac{1}{2} \| \|\mathbb{P}_{X|\mathcal{M}} - \mathbb{P}_X\|_v \|_p \,,$$

where $\| \cdot \|_v$ is the variation norm.

When $p = 1$, we write $\beta(\mathcal{M}, \sigma(X))$ in place of $\beta_1(\mathcal{M}, \sigma(X))$ and when $p = \infty$, we write $\phi(\mathcal{M}, \sigma(X))$ in place of $\beta_\infty(\mathcal{M}, \sigma(X))$. The coefficient $\beta(\mathcal{M}, \sigma(X))$ has been introduced by Rozanov and Volkonskii (1959), and $\phi(\mathcal{M}, \sigma(X))$ has been introduced by Ibragimov (1962). Note that $\beta_p(\mathcal{M}, \sigma(X)) = \tau_{d_0,p}(\mathcal{M}, X)$, where d_0 is the discrete metric $d_0(x, y) = \mathbb{1}_{x \neq y}$. In particular, Berbee's coupling lemma (1979) follows from Lemma 1.

Definition 2. 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 Y, define the upper tail function $L_Y(t) = \mathbb{P}(Y > t)$ and the quantile function $Q_Y = L_Y^{-1}$. On the set $[0, \mathbb{P}(Y > 0)]$, the function $H_Y : x \to \int_0^x Q_Y(u)du$ is an absolutely continuous and increasing function with values in $[0, \mathbb{E}(Y)]$. Denote by G_Y the inverse of H_Y.

We can now compare $\tau_{d,p}(\mathcal{M}, X)$ and $\beta_p(\mathcal{M}, \sigma(X))$.

Lemma 2. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, X a random variable with values in some Polish space $\mathcal{X}$, and $\mathcal{M}$ a σ-algebra of $\mathcal{A}$. Let d be a distance on $\mathcal{X}$ satisfying (2.1). For any X^* distributed as X and independent of $\mathcal{M}$, the following inequality hold:*

$$(2.5) \qquad \tau_{d,p}(\mathcal{M}, X) \;\leq\; \mathbb{E}^{1/p}\big(d^p(X, X^*)\mathbb{1}_{X \neq X^*}\mathbb{E}^{p-1}(\mathbb{1}_{X \neq X^*}|\mathcal{M})\big) \,, \ \textit{for } p \in [1, \infty[\,.$$

$$(2.6) \qquad \varphi_d(\mathcal{M}, X) \;\leq\; \|d(X, X^*)\|_\infty \phi(\mathcal{M}, \sigma(X)) \,.$$

Consequently, for any $s \in \mathcal{X}$ and $p \in [1, \infty[$,

$$(2.7) \qquad \tau_{d,p}(\mathcal{M}, X) \;\leq\; 2\Big(\int_0^{(\beta_p(\mathcal{M}, \sigma(X)))^p} Q^p_{d(X,s)}(u)du \Big)^{1/p} \,, \textit{ and}$$

$$(2.8) \qquad \tau_{d,p}(\mathcal{M}, X) \;\leq\; 2\big(\phi(\mathcal{M}, \sigma(X))\big)^{\frac{p-1}{p}} \Big(\int_0^{\beta(\mathcal{M}, \sigma(X))} Q^p_{d(X,s)}(u)du \Big)^{1/p} \,.$$

Remark 1. Let us introduce the usual strong mixing coefficient of Rosenblatt (1956):

$$\alpha(\mathcal{M}, \sigma(X)) = \sup_{A \in \mathcal{B}(\mathcal{X})} \|\mathbb{P}_{X|\mathcal{M}}(A) - \mathbb{P}_X(A)\|_1 . \tag{2.9}$$

With this definition and if $\mathcal{X} = \mathbb{R}$ and $d(x, y) = |x - y|$, then

$$\tau_d(\mathcal{M}, X) \le 2 \int_0^{\alpha(\mathcal{M}, \sigma(X))} Q_{|X|}(u) du .$$

This inequality has been proved in Dedecker and Prieur (2004). From an example given by Dehling (1983), we know that this inequality cannot be extended to separable Hilbert spaces.

Proof of Lemma 2. According to the definition of $\tau_{d,p}(\mathcal{M}, X)$, we clearly have that for any X^* distributed as X and independent of $\mathcal{M}$,

$$\tau_{d,p}(\mathcal{M}, X) \le \|\mathbb{E}(d(X, X^*)|\mathcal{M})\|_p = \|\mathbb{E}(d(X, X^*)\mathbb{1}_{X \neq X^*}|\mathcal{M})\|_p . \tag{2.10}$$

Applying Hölder's inequality conditionally to $\mathcal{M}$, we derive that

$$\tau_{d,p}(\mathcal{M}, X) \le \Big(\mathbb{E}\Big(\mathbb{E}\big(d^p(X, X^*)|\mathcal{M}\big)\mathbb{E}^{p-1}\big(\mathbb{1}_{X \neq X^*}|\mathcal{M})\big)\Big)\Big)^{1/p} ,$$

and (2.5) follows.

To prove (2.6), we first notice that for any X^* distributed as X and independent of $\mathcal{M}$,

$$\varphi_d(\mathcal{M}, X) \le \|\mathbb{E}(d(X, X^*)\mathbb{1}_{X \neq X^*}|\mathcal{M})\|_\infty \le \|d(X, X^*)\|_\infty \|\mathbb{E}(\mathbb{1}_{X \neq X^*}|\mathcal{M})\|_\infty .$$

According to Lemma 1 with the discrete metric $d_0(x, y) = \mathbb{1}_{x \neq y}$, we can choose X^* such that $\|\mathbb{E}(\mathbb{1}_{X \neq X^*}|\mathcal{M})\|_\infty = \phi(\mathcal{M}, \sigma(X))$.

To prove (2.7), we first give the following elementary result:

Lemma 3. *For any positive random variables U and V with $V \in [0, 1]$ a.s., we have that*

$$\mathbb{E}(UV) \le \int_0^{\mathbb{E}(V)} Q_U(u) du . \tag{2.11}$$

On the other hand, from (2.5) we have for any $s \in \mathcal{X}$ that

$$\tau_{d,p}(\mathcal{M}, X) \le \|d(X, s)\mathbb{1}_{X \neq X^*}\mathbb{E}^{\frac{p-1}{p}}(\mathbb{1}_{X \neq X^*}|\mathcal{M})\|_p + \|d(X^*, s)\mathbb{1}_{X \neq X^*}\mathbb{E}^{\frac{p-1}{p}}(\mathbb{1}_{X \neq X^*}|\mathcal{M})\|_p . \tag{2.12}$$

Applying Lemma 3 with $U = d^p(X,s)$ (or $U = d^p(X^*,s)$) and $V = \mathbb{1}_{X\neq X^*}\mathbb{E}^{p-1}(\mathbb{1}_{X\neq X^*}|\mathcal{M})$, and noticing that $\mathbb{E}(V) = (\beta_p(\mathcal{M},\sigma(X)))^p$, (2.7) follows. Now from (2.12), we derive that

$$\tau_{d,p}(\mathcal{M},X) \leq \big(\phi(\mathcal{M},\sigma(X))\big)^{\frac{p-1}{p}}\Big(\|d(X,s)\mathbb{1}_{X\neq X^*}\|_p + \|d(X^*,s)\mathbb{1}_{X\neq X^*}\|_p\Big)\,.$$

Applying Lemma 3 with $U = d^p(X,s)$ (or $U = d^p(X^*,s)$) and $V = \mathbb{1}_{X\neq X^*}$, and noticing that $\mathbb{E}(V) = \beta(\mathcal{M},\sigma(X))$, (2.8) follows. Finally, to prove (2.11), note that for any random variable T uniformly distributed on $[0,1]$ and independent of U and V, we have: $\mathbb{E}(UV) = \mathbb{E}(U\mathbb{1}_{T\leq V})$. Hence using Fréchet's inequality (1957), we get that

$$\mathbb{E}(UV) \leq \int_0^1 Q_U(u)Q_{\mathbb{1}_{T\leq V}}(u)du = \int_0^{\mathbb{E}(V)} Q_U(u)du\,.$$

3 Examples

Definition 3. Let $\mathcal{X}$ be some Polish space, and let d be a distance on $\mathcal{X}$ (the space $\mathcal{X}$ need not be Polish with respect to d). On $\mathcal{X}^k$ we put the distance $d_{(k)}$ defined by $d_{(k)}(x,y) = d(x_1,y_1)+\cdots+d(x_k,y_k)$. Let $(\Omega,\mathcal{A},\mathbb{P})$ be a probability space, $(X_i)_{i\geq 1}$ a sequence of $\mathcal{X}$-valued random variables and $(\mathcal{M}_i)_{i\geq 1}$ a sequence of σ-algebras of $\mathcal{A}$. For any positive integer k, define

$$(3.13)\qquad \tau_{d,p,k}(i) = \max_{1\leq \ell\leq k}\frac{1}{\ell}\sup\Big\{\tau_{d_{(\ell)},p}(\mathcal{M}_q,(X_{j_1},\cdots,X_{j_\ell})),\, q+i\leq j_1<\cdots<j_\ell\Big\}$$

and $\tau_{d,p,\infty}(i) = \sup_{k\geq 0}\tau_{d,p,k}(i)$. When $p=1$, we set $\tau_{d,1,k}(i) = \tau_{d,k}(i)$, and when $p=\infty$, we set $\tau_{d,\infty,k}(i) = \varphi_{d,k}(i)$.

In the same way, the coefficients $\beta_p(i)$ are defined by

$$\beta_{p,k}(i) \;=\; \sup\{\beta_p(\mathcal{M}_q,\sigma(X_{j_1},\ldots,X_{j_k})), q+i\leq j_1<\cdots<j_k\} \text{ and } \beta_{p,\infty}(i) = \sup_{k>0}\beta_{p,k}(i).$$

When $p=1$, we set $\beta_{1,k}(i) = \beta_k(i)$, and when $p=\infty$, we set $\beta_{\infty,k}(i) = \phi_k(i)$.

Starting from Inequalities (2.12) and (2.6), and arguing as in Lemma 2, we can prove the following result.

Lemma 4. *Let $(\Omega,\mathcal{A},\mathbb{P})$ be a probability space, let $(X_i)_{i>0}$ be a sequence of random variables with values in a Polish space $\mathcal{X}$, and let $\mathcal{M}_i$ be a sequence of σ-algebra of $\mathcal{A}$. Let d be a distance on $\mathcal{X}$ satisfying (2.1). For any $s\in\mathcal{X}$, the following inequalities hold*

$$\begin{aligned}\tau_{d,p,k}(i) &\leq 2\Big(\sup_{j>0}\int_0^{\beta^p_{p,k}(i)} Q^p_{d(X_j,s)}(u)du\Big)^{1/p}, \quad \textit{for any } p\in[1,\infty[,\\ \varphi_{d,k}(i) &\leq 2\phi_k(i)\sup_{j>0}\|d(X_j,s)\|_\infty\,.\end{aligned}$$

Remark 2. Let $\Lambda_1(\mathcal{X}^\ell, \mathcal{X}, d)$ be the set of functions f from $\mathcal{X}^\ell$ to $\mathcal{X}$ such that, for any x, y in $\mathcal{X}^\ell$, $d(f(x), f(y)) \leq d_{(\ell)}(x, y)$. We can define the following weaker version of $\tau_{d,k}(i)$:

$$(3.14)\quad \tilde{\tau}_{d,k}(i) = \max_{1\leq \ell \leq k} \frac{1}{\ell} \sup\Big\{ \sup_{f \in \Lambda_1(\mathcal{X}^\ell, \mathcal{X}, d)} \tau_d(\mathcal{M}_p, f(X_{j_1}, \cdots, X_{j_\ell})),\ p+i \leq j_1 < \cdots < j_\ell \Big\}.$$

Define also the sequence of strong mixing coefficients $(\alpha_\infty(i))_{i>0}$ by:

$$\alpha_\infty(i) = \sup_{k>0} \alpha(\mathcal{M}_k, \sigma(X_j, j \geq k+i))\,. \tag{3.15}$$

In the case $\mathcal{X} = \mathbb{R}$ and $d(x, y) = |x - y|$, we can obtain the bound

$$\tilde{\tau}_{d,\infty}(i) \leq 2 \sup_{j>0} \int_0^{\alpha_\infty(i)} Q_{|X_j|}(u) du\,. \tag{3.16}$$

3.1 Example 1: causal functions of stationary sequences

Let $(\xi_i)_{i\in\mathbb{Z}}$ be a stationary sequence of random variables with values in a measurable space $\mathcal{S}$. Assume that there exists a function H defined on a subset of $\mathcal{S}^{\mathbb{N}}$, with values in a Polish space $\mathcal{X}$ and such that $H(\xi_0, \xi_{-1}, \xi_{-2}, \ldots,)$ is defined almost surely. The stationary sequence $(X_n)_{n\in\mathbb{Z}}$ defined by $X_n = H(\xi_n, \xi_{n-1}, \xi_{n-2}, \ldots)$ is called a causal function of $(\xi_i)_{i\in\mathbb{Z}}$.

Assume that there exists a stationary sequence $(\xi'_i)_{i\in\mathbb{Z}}$ distributed as $(\xi_i)_{i\in\mathbb{Z}}$ and independent of $(\xi_i)_{i\leq 0}$. Define $X^*_n = H(\xi'_n, \xi'_{n-1}, \xi'_{n-2}, \ldots)$. Clearly X^*_n is independent of $\mathcal{M}_0 = \sigma(X_i, i \leq 0)$ and distributed as X_n. For any $p \geq 1$ (p may be infinite), let $(\delta_p(i))_{i>0}$ be a non increasing sequence such that

$$\|\mathbb{E}(d(X_i, X^*_i)|\mathcal{M}_0)\|_p \leq \delta_p(i)\,, \tag{3.17}$$

where d is a distance on $\mathcal{X}$ satisfying (2.1) (the space $\mathcal{X}$ need not be Polish with respect to d). Let $\mathcal{M}_i = \sigma(X_j, j \leq i)$. From Lemma 1, we infer that the coefficients $\tau_{d,p,\infty}$ of the sequence $(X_n)_{n\geq 0}$ satisfy

$$\tau_{d,p,\infty}(i) \leq \delta_p(i)\,. \tag{3.18}$$

For instance, according to Theorem 4.4.7 in Berbee (1979), if Ω is rich enough, there exists $(\xi'_i)_{i\in\mathbb{Z}}$ distributed as $(\xi_i)_{i\in\mathbb{Z}}$ and independent of $(\xi_i)_{i\leq 0}$ such that

$$\mathbb{P}(\xi_i \neq \xi'_i \text{ for some } i \geq k|\mathcal{F}_0) = \frac{1}{2}\|\mathbb{P}_{\tilde{\xi}_k|\mathcal{F}_0} - \mathbb{P}_{\tilde{\xi}_k}\|_v\,, \tag{3.19}$$

where $\widetilde{\xi}_k = (\xi_k, \xi_{k+1}, \ldots)$, $\mathcal{F}_0 = \sigma(\xi_i, i \leq 0)$. In particular if the sequence $(\xi_i)_{i\in\mathbb{Z}}$ is iid, it suffices to take $\xi'_i = \xi_i$ for $i > 0$ and $\xi'_i = \xi''_i$ for $i \leq 0$, where $(\xi''_i)_{i\in\mathbb{Z}}$ is an independent copy of $(\xi_i)_{i\in\mathbb{Z}}$.

Application: causal linear processes in Banach spaces. Let $(\mathbb{B}_1, \|\cdot\|_{\mathbb{B}_1})$ and $(\mathbb{B}_2, \|\cdot\|_{\mathbb{B}_2})$ be two Banach spaces. For any linear application A from $\mathbb{B}_1$ to $\mathbb{B}_2$, let $\|A\| = \sup\{\|A(b)\|_{\mathbb{B}_2}, \|b\|_{\mathbb{B}_1} \leq 1\}$. Let $(A_i)_{i\geq 0}$ be a sequence of linear operators from $\mathbb{B}_1$ to $\mathbb{B}_2$ such that $\sum_{i\geq 0}\|A_i\| < \infty$, and let $(\xi_i)_{i\in\mathbb{Z}}$ be a stationary sequence of $\mathbb{B}_1$-valued random variables. Define the random variables $X_n = \sum_{j\geq 0} A_j(\xi_{n-j})$ with values in $\mathbb{B}_2$. For any $p \geq 1$, observe that

$$\|\mathbb{E}(\|X_i - X_i^*\|_{\mathbb{B}_2}|\mathcal{M}_0)\|_p \leq \sum_{j\geq 0}\|A_j\|\|\mathbb{E}(\|\xi_{i-j} - \xi'_{i-j}\|_{\mathbb{B}_1}|\mathcal{M}_0)\|_p$$

and consequently

$$\|\mathbb{E}(\|X_i - X_i^*\|_{\mathbb{B}_2}|\mathcal{M}_0)\|_p \leq \|\,\|\xi_0 - \xi'_0\|_{\mathbb{B}_1}\|_p \sum_{j\geq i}\|A_j\| + \sum_{j=0}^{i-1}\|A_j\|\|\mathbb{E}(\|\xi_{i-j} - \xi'_{i-j}\|_{\mathbb{B}_1}|\mathcal{M}_0)\|_p\,.$$

Taking $(\xi_i)_{i\in\mathbb{Z}}$ as in (3.19) and arguing as in Lemma 2, we obtain the following inequalities: for $p \in [1, \infty[$ and and $\mathcal{G}_i = \sigma(\xi_k, k \geq i)$,

$$\begin{aligned}&\|\mathbb{E}(\|X_i - X_i^*\|_{\mathbb{B}_2}|\mathcal{M}_0)\|_p \\ &\qquad\leq \|\,\|\xi_0 - \xi'_0\|_{\mathbb{B}_1}\|_p \sum_{j\geq i}\|A_j\| + \sum_{j=0}^{i-1}\|A_j\|\Big(2^p\int_0^{(\beta_p(\mathcal{F}_0,\mathcal{G}_{i-j}))^p} Q^p_{\|\xi_0\|_{\mathbb{B}_1}}(u)\Big)^{1/p} du\,,\end{aligned}$$

and

$$\|\mathbb{E}(\|X_i - X_i^*\|_{\mathbb{B}_2}|\mathcal{M}_0)\|_\infty \leq \|\,\|\xi_0 - \xi'_0\|_{\mathbb{B}_1}\|_\infty\Big(\sum_{j\geq i}\|A_j\| + \sum_{j=0}^{i-1}\|A_j\|\phi(\mathcal{F}_0, \mathcal{G}_{i-j})\Big)\,.$$

Hence we can take $\delta_p(i)$ such that

$$\begin{aligned}\delta_p(i) &\geq \|\,\|\xi_0 - \xi'_0\|_{\mathbb{B}_1}\|_p \sum_{j\geq i}\|A_j\| + \sum_{j=0}^{i-1}\|A_j\|\Big(2^p\int_0^{(\beta_p(\mathcal{F}_0,\mathcal{G}_{i-j}))^p} Q^p_{\|\xi_0\|_{\mathbb{B}_1}}(u)\Big)^{1/p} du \\ \text{and } \delta_\infty(i) &\geq \|\,\|\xi_0 - \xi'_0\|_{\mathbb{B}_1}\|_\infty\Big(\sum_{j\geq i}\|A_j\| + \sum_{j=0}^{i-1}\|A_j\|\phi(\mathcal{F}_0, \mathcal{G}_{i-j})\Big)\,.\end{aligned}$$

If the sequence $(\xi_i)_{i\in\mathbb{Z}}$ is iid, we can take $\delta_p(i) = \|\,\|\xi_0 - \xi'_0\|_{\mathbb{B}_1}\|_p \sum_{j\geq i}\|A_j\|$. If furthermore, $\mathbb{B}_2$ is of type 2, then we can take $\delta_2(i) = (C\mathbb{E}\|\xi_0 - \xi'_0\|^2_{\mathbb{B}_2}\sum_{j\geq i}\|A_j\|^2)^{1/2}$ where C

is the type constant. For instance, if $\mathbb{B}_1 = \mathbb{B}_2 = \mathbb{R}$, $A_i = 2^{-i-1}$ and $\xi_0 \sim \mathcal{B}(1/2)$, then we can take $\delta_2(i) = 2^{-i}/\sqrt{6}$ and $\delta_\infty(i) = 2^{-i}$. Since X_0 is uniformly distributed over $[0,1]$, we have $\varphi_{d,\infty}(i) \leq 2^{-i}$ for $d(x-y) = \|x-y\|_{\mathbb{B}_2}$. Recall that this sequence is not strongly mixing.

3.2 Example 2 : iterated random functions

Let $(X_n)_{n\geq 0}$ be a $\mathcal{X}$-valued stationary Markov chain, such that $X_n = F(X_{n-1}, \xi_n)$ for some measurable function F and some iid sequence $(\xi_i)_{i>0}$ independent of X_0. Let X_0^* be a random variable distributed as X_0 and independent of $(X_0, (\xi_i)_{i>0})$. Define $X_n^* = F(X_{n-1}^*, \xi_n)$. The sequence $(X_n^*)_{n\geq 0}$ is distributed as $(X_n)_{n\geq 0}$ and independent of X_0. Let $\mathcal{M}_i = \sigma(X_j, 0 \leq j \leq i)$. As in Example 1, let $(\delta_p(i))_{i>0}$ be a non increasing sequence satisfying (3.17). The coefficients $\tau_{d,p,\infty}$ of the sequence $(X_n)_{n\geq 0}$ satisfy the bound (3.18) of Example 1.

Let μ be the distribution of X_0 and let $(X_n^x)_{n\geq 0}$ be the chain starting from $X_0^x = x$. With these notations, we can take $\delta_p(i)$ such that

$$\delta_p(i) \geq \|d(X_i, X_i^*)\|_p = \Big(\iint \| d(X_i^x, X_i^y)\|_p^p \, \mu(dx)\mu(dy)\Big)^{1/p} .$$

For instance, if there exists a non increasing sequence $(a_p(i))_{i\geq 0}$ of positive numbers such that

$$\| d(X_i^x, X_i^y)\|_p \leq a_p(i) d(x,y) \, ,$$

then we can take $\delta_p(i) = a_p(i)\| d(X_0, X_0^*)\|_p$. For example, in the usual case where

$$\|d(F(x,\xi_0), F(y,\xi_0))\|_p \leq \kappa d(x,y) \, ,$$

for some $\kappa < 1$, we can take $a_p(i) = \kappa^i$ and $\delta_p(i) = \kappa^i \| d(X_0, X_0^*)\|_p$.

An important example is $X_n = f(X_{n-1}) + \xi_n$ for some function f which is κ-lipschitz with respect to d. If X_0 belongs to $\mathbb{L}^p(\mathcal{X}, d)$, then we can take $\delta_p(i) = \kappa^i \| d(X_0, X_0^*)\|_p$.

3.3 Example 3 : multidimensional expanding maps

Let $\mathcal{X}$ be a Polish space and let d be a distance on $\mathcal{X}$ satisfying (2.1). Let T be a map from $\mathcal{X}$ to $\mathcal{X}$ and define $X_i = T^i$. If μ is invariant by T, the sequence $(X_i)_{i\geq 0}$ of random variables from $(\mathcal{X}, \mu)$ to $\mathcal{X}$ is strictly stationary.

Notations. For any finite measure ν on $\mathcal{X}$, let $\mathbb{L}^1(\nu)$ be the space of all functions f from $\mathcal{X}$ to $\mathbb{R}$, such that: $\|f\|_{\mathbb{L}^1(\nu)} = \int_{\mathcal{X}} |f(x)|\nu(dx) < \infty$ and let $B^1(\nu) = \{f : \|f\|_{\mathbb{L}^1(\nu)} \leq 1\}$. Let $\mathbb{L}^\infty(\nu)$ be the space of all functions f from $\mathcal{X}$ to $\mathbb{R}$, such that:

$$\|f\|_{\mathbb{L}^\infty(\nu)} = \inf_{M>0} \{\nu(|f| > M) = 0\} < \infty \,.$$

We also use the notations $\nu(h) = \int_{\mathcal{X}} h(x)\nu(dx)$. If h is a Lipshitz function with respect to d, we note

$$Lip(h) = \sup_{(x,y)\in\mathcal{X}^2} \frac{|h(x)-h(y)|}{d(x,y)} \,.$$

Covariance inequalities. In many interesting cases, one can prove that, for any Lipshitz function h and any f in $\mathbb{L}^1(\mu)$,

$$|\mathrm{Cov}(f(X_n), h(X_0))| \leq a_n \|f\|_{\mathbb{L}^1(\mu)} Lip(h)\,, \tag{3.20}$$

for some non increasing sequence a_n tending to zero as n tends to infinity.
Since

$$\varphi_{d,1}(\sigma(X_n), X_0) = \sup_{h\in\Lambda_1(\mathcal{X},d)} \sup_{f\in B^1(\mu)} |\mathrm{Cov}(f(X_n), h(X_0))|\,,$$

we derive that if (3.20) holds then

$$\varphi_{d,1}(\sigma(X_n), X_0) \leq a_n \,.$$

If $\mathcal{X} = [0,1]$ and $d(x,y) = |x-y|$, then the inequality (3.20) is satisfied with $a_n = C\rho^n$, where $\rho \in]0,1[$, for the class of uniformly expanding maps given in Collet *et al.* (2002).

We now give an example where the inequality (3.20) holds for $\mathcal{X}$ a compact of $\mathbb{R}^k$ and d the Euclidean distance on $\mathbb{R}^k$. Let T be a map from $\mathcal{X}$ to $\mathcal{X}$ satisfying the assumptions (PE1-PE5) in Saussol (2000). Let $|h|_\alpha$ be defined as in page 232 in Saussol and $\|h\|_\alpha = |h|_\alpha + \lambda(|h|)$ where λ is the Lebesgue measure on $\mathcal{X}$. Let $\widetilde{P}$ be the Perron-Frobenius of T defined on $\mathbb{L}^1(\mu)$, that is for any bounded function f and any $h \in \mathbb{L}^1(\mu)$,

$$\int_{\mathcal{X}} (f \circ T^n) h d\mu = \int_{\mathcal{X}} f \widetilde{P}(h) d\mu \,.$$

By Theorem 5.1 in Saussol (2000), we know that there exists a finite number of invariant probabilities by T which are absolutely continuous with respect to λ. Let μ be such a probability, and assume furthermore that μ is mixing in the ergodic-theoric sense. Denote by $h_\star = d\mu/d\lambda$. By Theorem 6.2 in Saussol (2000), there exist a positive constant C and $\rho \in]0,1[$ such that for any $\alpha \in]0,1]$ and any h such that $\|h\|_\alpha < \infty$,

$$\|(\widetilde{P}^n(h) - \mu(h))h_\star\|_\alpha \leq C\|h\|_\alpha \rho^n \,. \tag{3.21}$$

Following the proof of Theorem 6.1 in Saussol, we have that

$$\begin{aligned} |\mathrm{Cov}(f(X_n), h(X_0))| &= \Big| \int_{\mathcal{X}} f(\widetilde{P}^n(h) - \mu(h))h_\star d\lambda \Big| \\ &\leq \|f\|_{\mathbb{L}^1(\mu)} \|\widetilde{P}^n(h) - \mu(h)\|_{\mathbb{L}^\infty(\mu)} \\ &\leq \frac{\|f\|_{\mathbb{L}^1(\mu)}}{\|h_\star\|_{\mathbb{L}^\infty(\lambda)}} \|(\widetilde{P}^n(h) - \mu(h))h_\star\|_{\mathbb{L}^\infty(\lambda)} \,. \end{aligned}$$

Then using Inequality (6) of Proposition 3.4 in Saussol, we infer that there exists a positive constant K such that

$$|\mathrm{Cov}(f(X_n), h(X_0))| \leq K \frac{\|f\|_{\mathbb{L}^1(\mu)}}{\|h_\star\|_{\mathbb{L}^\infty(\lambda)}} \|(\widetilde{P}^n(h) - \mu(h))h_\star\|_\alpha \,.$$

Hence using (3.21), we obtain that

$$|\mathrm{Cov}(f(X_n), h(X_0))| \leq KC \frac{\|f\|_{\mathbb{L}^1(\mu)}}{\|h_\star\|_{\mathbb{L}^\infty(\lambda)}} \|h\|_\alpha \rho^n \,. \tag{3.22}$$

Now notice that for any $s \in \mathcal{X}$,

$$|\mathrm{Cov}(f(X_n), h(X_0))| = |\mathrm{Cov}(f(X_n), h(X_0) - h(s))| \,.$$

From Inequality (3.22) applied with $\alpha = 1$ and the fact that $|h|_1 \leq Lip(h)$, we then derive that

$$\begin{aligned} |\mathrm{Cov}(f(X_n), h(X_0))| &\leq KC \frac{\|f\|_{\mathbb{L}^1(\mu)}}{\|h_\star\|_{\mathbb{L}^\infty(\lambda)}} \big(|h|_1 + \lambda(|h - h(s)|)\big)\rho^n \\ &\leq \frac{KC(1 + \lambda(\mathcal{X}))}{\|h_\star\|_{\mathbb{L}^\infty(\lambda)}} \rho^n \|f\|_{\mathbb{L}^1(\mu)} Lip(h) \,. \end{aligned}$$

Hence (3.20) follows with $a_n = \rho^n KC(1 + \lambda(\mathcal{X}))/\|h_\star\|_{\mathbb{L}^\infty(\lambda)}$ and consequently

$$\varphi_{d,1}(\sigma(X_n), X_0) \leq A\rho^n \ \text{ with } A = \frac{KC(1 + \lambda(\mathcal{X}))}{\|h_\star\|_{\mathbb{L}^\infty(\lambda)}} \,. \tag{3.23}$$

Remark 3. Arguing as in Dedecker and Prieur (2005) page 230, there exists a stationary Markov chain $(Y_i)_{i\geq 0}$ with transition kernel $\widetilde{P}$ such that for all $n \geq 0$, $(Y_0, \ldots, Y_n)$ has the same distribution as $(X_n, \ldots, X_0)$. Hence, for the Markov chain $(Y_i)_{i\geq 0}$ and the σ-algebras $\mathcal{M}_i = \sigma(Y_j, j \leq i)$ we obtain from (3.23) that $\varphi_{d,1}(i) \leq A\rho^i$. From Saussol's paper, it seems difficult to prove that $\varphi_{d,\infty}(i)$ decreases exponentially fast. However for $\mathcal{X} = [0, 1]$ and a map T satisfying the assumptions of the paper of Collet *et al.* (2002), we can prove that for the Markov chain $(Y_i)_{i\geq 0}$: $\varphi_{d,\infty}(i) \leq B\rho^i$ (see Dedecker and Prieur (2005) page 230 for more details).

Remark 4. In Buzzi and Maume-Deschamps (2002), others examples of multidimensional expanding maps for which (3.23) holds are given. They also give sufficient conditions to obtain an arithmetic decay of the coefficients (see their main theorem in Section 0.2).

4 A Fuk-Nagaev Inequality

¿From now, we shall only consider separable Banach spaces $(B, \|\cdot\|_{\mathbb{B}})$. When $d(x, y) = \|x - y\|_{\mathbb{B}}$, we shall write $\tau(\mathcal{M}, X)$ instead of $\tau_d(\mathcal{M}, X)$ and $\tau_k(i)$ instead of $\tau_{d,k}(i)$.

Definition 4. Following Pisier (1975), we say that a Banach space $(\mathbb{B}, \|\cdot\|_{\mathbb{B}})$ is 2-smooth if there exists an equivalent norm $\|\cdot\|$ such that

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

¿From Assouad (1975), we know that if $\mathbb{B}$ is 2-smooth and separable, then there exists a constant D such that, for any sequence of $\mathbb{B}$-valued martingale differences $(X_i)_{i\geq 1}$,

$$\mathbb{E}(\|X_1 + \cdots + X_n\|_{\mathbb{B}}^2) \leq D\sum_{i=1}^{n}\mathbb{E}(\|X_i\|_{\mathbb{B}}^2)\,. \tag{4.24}$$

¿From (4.24), we see that 2-smooth Banach spaces play the same role for martingales as space of type 2 do for sums of independent variables. When the constant D needs to be specify, we shall say that $\mathbb{B}$ is $(2, D)$-smooth. Note that, for any measure space $(T, \mathcal{A}, \nu)$, $\mathbb{L}^p(T, \mathcal{A}, \nu)$ is $(2, \sqrt{p-1})$-smooth for any $p \geq 2$, and that any separable Hilbert space is $(2, 1)$-smooth.

Definition 5. Given a separable Banach space $(\mathbb{B}, \|\cdot\|_{\mathbb{B}})$, let $(X_i)_{i>0}$ be a sequence of $\mathbb{B}$-valued random variables. Following Woyczińsky (1981), we write $(X_i) \prec X$ if there exists a nonnegative random variable X such that $Q_X \geq \sup_{k\geq 1} Q_{\|X_k\|_{\mathbb{B}}}$.

Theorem 1. *Let $(\mathbb{B}, \|\cdot\|_{\mathbb{B}})$ be a $(2,D)$-smooth separable Banach space. Let $(X_k)_{k>0}$ be a sequence of centered random variables with values in $\mathbb{B}$. Let $S_n = X_1 + \cdots + X_n$ and $\mathcal{M}_i = \sigma(X_k, 1 \le k \le i)$. Let $\tau_\infty^{-1}(u) = \sum_{i\ge 0} \mathbb{1}_{u<\tau_\infty(i)}$. Let X be a nonnegative random variable such that $(X_i) \prec X$. Let $R_X = (\tau_\infty^{-1} \circ G_X^{-1} \wedge n) Q_X$ and $S_X = R_X^{-1}$. Then, for any $x > 0$ and $r \ge 1$ and every quantity s_n^2 such that*

$$s_n^2 \ge \max_{1\le q\le n} \sum_{i=1}^{[n/q]} \mathbb{E}\|S_{iq} - S_{(i-1)q}\|_{\mathbb{B}}^2 \,,$$

one has that

$$\mathbb{P}\Big(\max_{1\le k\le n} \|S_k\|_{\mathbb{B}} \ge 4x\Big) \;\le\; 4\left(1 + \frac{x^2}{D^2\, r s_n^2}\right)^{-r/8} + \frac{9n}{x}\int_0^{S_X(x/r)} Q_X(u)du\,. \tag{4.25}$$

Remark 5. In the proof of Theorem 1, note that the inequality (4.41) can be weakened in $\tau(\mathcal{F}_{i-2}, U_i) \le q\widetilde{\tau}_q(q+1)$, where $\widetilde{\tau}_q(i)$ is defined from (3.14). Consequently, Inequality (4.25) remains valid with the sequence $\widetilde{\tau}_\infty = (\widetilde{\tau}_\infty(i))_{i\ge 0}$ replacing $\tau_\infty = (\tau_\infty(i))_{i\ge 0}$.

Corollary 1. *Let $(\mathbb{B}, \|\cdot\|_{\mathbb{B}})$ be a $(2,D)$-smooth separable Banach space. Let $(X_k)_{k>0}$ be a sequence of centered random variables with values in $\mathbb{B}$. Let X be a nonnegative random variable such that $(X_i) \prec X$. Then with the notations of Theorem 1, for any $p \ge 2$,*

$$\mathbb{E}\Big(\max_{1\le k\le n} \|S_k\|_{\mathbb{H}}^p\Big) \le a_p s_n^p + b_p\, n \int_0^1 R_X^{p-1}(u) Q_X(u) du\,, \tag{4.26}$$

where

$$a_p = p\, 4^{p+2} D^p (4p+1)^{p/2} \text{ and } b_p = 9\frac{p}{p-1} 4^p (4p+1)^{p-1}\,.$$

By the definition of R_X, it follows that

$$\mathbb{E}\Big(\max_{1\le k\le n} \|S_k\|_{\mathbb{H}}^p\Big) \le a_p s_n^p + (p-1) b_p\, n \sum_{i=0}^{n-1} (i+1)^{p-2} \int_0^{\tau_\infty(i)} Q_X^{p-1} \circ G_X(u) du\,. \tag{4.27}$$

Proof of Corollary 1 Inequality (4.26) follows from the fact that for $p \ge 1$, $\mathbb{E}(|Z|^p) = p\, 4^p \int_0^\infty x^{p-1} \mathbb{P}(|Z| \ge 4x) dx$, and the application of Inequality (4.25) with $r = 4p+1$. Inequality (4.27) comes from (4.26) by making the change of variables $v = H(u)$ and by using the fact that $(i+1)^{p-1} - i^{p-1} \le c_p (i+1)^{p-2}$ where $c_p = 1 \vee (p-1)$.

Remark 6. In the case where $\mathbb{B} = \mathbb{H}$ is a separable Hilbert space the constants involved in the right-hand side of Inequality (4.25) can be sharpened. First we can take $D = 1$.

In addition, with the notations of the proof of Theorem 1, $\mathbb{E}\|\tilde{U}_i\|_{\mathbb{B}}^2 \le \mathbb{E}\|U_i'\|_{\mathbb{H}}^2 \le \mathbb{E}\|U_i\|_{\mathbb{H}}^2$. Consequently, we can apply Lemma 6 of the appendix with $y_n = s_n^2$ and $c = 4qM$ to obtain a better inequality than (4.38). The upper bound (4.25) becomes

$$\mathbb{P}\Big(\max_{1\le k\le n} \|S_k\|_{\mathbb{H}} \ge 4x\Big) \le 4\Big(1+\frac{4x^2}{rs_n^2}\Big)^{-r/8} + \frac{9n}{x}\int_0^{S_X(x/r)} Q_X(u)du\,, \tag{4.28}$$

and we can take

$$s_n^2 = \sum_{i=1}^n \sum_{j=1}^n |\mathbb{E} < X_i, X_j >_{\mathbb{H}} |\,. \tag{4.29}$$

The constant a_p involved in Inequalities (4.26) and (4.27) becomes in this case $a_p = p\,2^{p+4}(4p+1)^{p/2}$ (the constant b_p is unchanged).

Proof of Theorem 1. For the sake of brevity, write Q, R, S and G for Q_X, R_X, S_X and G_X respectively. Let q be a positive integer and $M > 0$. Define the random variables $U_i = S_{iq} - S_{iq-q}$ for $1 \le i \le [n/q]$. Let

$$U_i' = U_i \mathbb{1}_{\|U_i\|_{\mathbb{B}} \le 2qM} \quad \text{and} \quad U_i'' = U_i \mathbb{1}_{\|U_i\|_{\mathbb{B}} > 2qM}\,. \tag{4.30}$$

With these notations, it is clear that $U_i = U_i' + U_i''$. Define $\varphi_M(x) = (|x| - M)_+$. We first show that

$$\max_{1\le k\le n} \|S_k\|_{\mathbb{B}} \le \max_{1\le j\le [n/q]} \Big\|\sum_{i=1}^j U_i'\Big\|_{\mathbb{B}} + qM + 2\sum_{k=1}^n \varphi_M(\|X_k\|_{\mathbb{B}})\,. \tag{4.31}$$

To prove (4.31), it suffices to note that, if the maximum of $\|S_k\|_{\mathbb{B}}$ is obtained in k_0, then for $j_0 = [k_0/q]$,

$$\max_{1\le k\le n} \|S_k\|_{\mathbb{B}} \le \Big\|\sum_{i=1}^{j_0} U_i'\Big\|_{\mathbb{B}} + \sum_{i=1}^{j_0} \|U_i''\|_{\mathbb{B}} + \sum_{k=qj_0+1}^{k_0} \|X_k\|_{\mathbb{B}}\,. \tag{4.32}$$

Now, by definition of φ_M,

$$\sum_{k=qj_0+1}^{k_0} \|X_k\|_{\mathbb{B}} \le (k_0 - qj_0)M + \sum_{k=qj_0+1}^{k_0} \varphi_M(\|X_k\|_{\mathbb{B}})\,. \tag{4.33}$$

On the other hand, since $|x|\mathbb{1}(|x| > 2A) \le 2(|x| - A)_+$, we have $\|U_i''\|_{\mathbb{B}} \le 2\varphi_{qM}(\|U_i\|_{\mathbb{B}})$. By convexity of the function φ_{qM}, we derive that

$$\sum_{i=1}^{j_0} \|U_i''\|_{\mathbb{B}} \le 2\sum_{k=1}^{qj_0} \varphi_M(\|X_k\|_{\mathbb{B}})\,. \tag{4.34}$$

Starting from (4.32) and using (4.33) and (4.34), we obtain (4.31). Now, since the random variables are centered, we have that $\mathbb{E}(U_i') = -\mathbb{E}(U_i'')$ and hence

$$\max_{1\le j\le [n/q]} \Big\| \sum_{i=1}^{j} U_i' \Big\|_{\mathbb{B}} \le \max_{1\le j\le [n/q]} \Big\| \sum_{i=1}^{j} (U_i' - \mathbb{E}(U_i')) \Big\|_{\mathbb{B}} + \sum_{i=1}^{[n/q]} \mathbb{E}\|U_i''\|_{\mathbb{B}} \,.$$

Using (4.34), it follows that

$$\max_{1\le j\le [n/q]} \Big\| \sum_{i=1}^{j} U_i' \Big\|_{\mathbb{B}} \le \max_{1\le j\le [n/q]} \Big\| \sum_{i=1}^{j} (U_i' - \mathbb{E}(U_i')) \Big\|_{\mathbb{B}} + 2\sum_{k=1}^{n} \mathbb{E}(\varphi_M(\|X_k\|_{\mathbb{B}})) \,. \tag{4.35}$$

To control the quantity: $\max_{1\le j\le [n/q]} \| \sum_{i=1}^{j} (U_i' - \mathbb{E}(U_i'))\|_{\mathbb{B}}$, we proceed as follows. Let $(\delta_i)_{1\le i\le [n/q]}$ be a sequence of independent random variables with uniform distribution over $[0,1]$, independent of $(U_i)_{1\le i\le [n/q]}$. Let $U_1^* = U_1'$ and $U_2^* = U_2'$, and apply Lemma 1. For any $3 \le i \le [n/q]$, there exists a measurable function F_i such that $U_i^* = F_i(U_1', \cdots, U_{i-2}', U_i', \delta_i)$ satisfies the conditions of Lemma 1 with $\mathcal{M} = \sigma(U_\ell, \ell \le i-2)$. The sequence $(U_i^*)_{1\le i\le [n/q]}$ has the following properties:

1. for any $1 \le i \le [n/q]$, the random variable U_i^* has the same distribution as U_i'.
2. The random variables $(U_{2i}^*)_{2\le 2i\le [n/q]}$ are independent as well as the random variables $(U_{2i-1}^*)_{1\le 2i-1\le [n/q]}$.
3. for any $3 \le i \le [n/q]$, $\mathbb{E}\|U_i' - U_i^*\|_{\mathbb{B}} = \tau\big(\sigma(U_\ell, \ell \le i-2), U_i'\big)$.

Substituting the variables U_i^* to U_i', we obtain the inequality

$$\begin{aligned}\max_{1\le j\le [n/q]} \Big\| \sum_{i=1}^{j} (U_i' - \mathbb{E}(U_i')) \Big\|_{\mathbb{B}} &\le \max_{2\le 2j\le [n/q]} \Big\| \sum_{i=1}^{j} \tilde{U}_{2i} \Big\|_{\mathbb{B}} \\ &\quad + \max_{1\le 2j-1\le [n/q]} \Big\| \sum_{i=1}^{j} \tilde{U}_{2i-1} \Big\|_{\mathbb{B}} + \sum_{i=3}^{[n/q]} \|U_i' - U_i^*\|_{\mathbb{B}} \,,\end{aligned} \tag{4.36}$$

where $\tilde{U}_i = U_i^* - \mathbb{E}(U_i^*)$ for any $1 \le i \le [n/q]$. Combining (4.31), (4.35) and (4.36), we obtain the following upper bound

$$\begin{aligned}\max_{1\le k\le n} \|S_k\|_{\mathbb{B}} &\le qM + \max_{2\le 2j\le [n/q]} \Big\| \sum_{i=1}^{j} \tilde{U}_{2i} \Big\|_{\mathbb{B}} + \max_{1\le 2j-1\le [n/q]} \Big\| \sum_{i=1}^{j} \tilde{U}_{2i-1} \Big\|_{\mathbb{B}} \\ &\quad + \sum_{i=3}^{[n/q]} \|U_i' - U_i^*\|_{\mathbb{B}} + 2\sum_{k=1}^{n} \Big(\varphi_M(\|X_k\|_{\mathbb{B}}) + \mathbb{E}(\varphi_M(\|X_k\|_{\mathbb{B}})) \Big) \,.\end{aligned} \tag{4.37}$$

Since $\|U_i'\|_{\mathbb{B}} \leq 2qM$ almost surely, it follows that $\|\tilde{U}_i\|_{\mathbb{B}} \leq 4qM$ almost surely. Since $\mathbb{E}\|\tilde{U}_i\|_{\mathbb{B}}^2 \leq 4\mathbb{E}\|U_i'\|_{\mathbb{B}}^2 \leq 4\mathbb{E}\|U_i\|_{\mathbb{B}}^2$, we obtain from Lemma 6 of the appendix, with $y_n = 4s_n^2$ and $c = 4qM$,

$$(4.38) \qquad \mathbb{P}\Big(\max_{2\leq 2j\leq [n/q]}\Big\|\sum_{i=1}^{j}\tilde{U}_{2i}\Big\|_{\mathbb{B}} \geq x\Big) \leq 2\exp\left(-\frac{x}{8qM}\ln\left(1+\frac{xqM}{D^2s_n^2}\right)\right).$$

Obviously, the same bound holds for $\max_{1\leq 2j-1\leq [n/q]}\|\sum_{i=1}^{j}\tilde{U}_{2i-1}\|_{\mathbb{B}}$.

Now from Markov's inequality, we get that

$$(4.39) \quad \mathbb{P}\Big(\sum_{i=3}^{[n/q]+1}\|U_i' - U_i^*\|_{\mathbb{B}} + 2\sum_{k=1}^{n}\Big(\varphi_M(\|X_k\|_{\mathbb{B}}) + \mathbb{E}(\varphi_M(\|X_k\|_{\mathbb{B}}))\Big) \geq x\Big) \\ \leq \frac{1}{x}\Big(\sum_{i=3}^{[n/q]+1}\mathbb{E}\|U_i' - U_i^*\|_{\mathbb{B}} + 4\sum_{k=1}^{n}\mathbb{E}(\varphi_M(\|X_k\|_{\mathbb{B}}))\Big).$$

Let $\mathbb{P}_{U_i',U_i|\mathcal{M}}$ be a conditional distribution of (U_i', U_i) given $\mathcal{M}$, and define the σ-algebras $\mathcal{F}_i = \sigma(U_i, 1\leq j\leq i)$. For any $3\leq i\leq [n/q]$, we have for $d(x,y) = \|x-y\|_{\mathbb{B}}$,

$$(4.40) \quad \|U_i' - U_i^*\|_{\mathbb{L}^1_{\mathbb{B}}} = \Big\|\sup_{\Lambda_1(\mathbb{B},d)}\Big|\iint f(x)\mathbb{P}_{U_i',U_i|\mathcal{F}_{i-2}}(dx,dy) - \iint f(x)\mathbb{P}_{U_i',U_i}(dx,dy)\Big|\Big\|_1,$$

and consequently $\|U_i' - U_i^*\|_{\mathbb{L}^1_{\mathbb{B}}} \leq A_1 + A_2$, where

$$A_1 = \Big\|\sup_{f\in\Lambda_1(\mathbb{B},d)}\Big|\int f(y)\mathbb{P}_{U_i|\mathcal{F}_{i-2}}(dy) - \int f(y)\mathbb{P}_{U_i}(dy)\Big|\Big\|_1$$
$$A_2 = \Big\|\sup_{f\in\Lambda_1(\mathbb{B},d)}\Big|\iint (f(x)-f(y))(\mathbb{P}_{U_i',U_i|\mathcal{F}_{i-2}}(dx,dy) - \mathbb{P}_{U_i',U_i}(dx,dy))\Big|\Big\|_1.$$

Clearly $A_2 \leq 2\mathbb{E}\|U_i' - U_i\|_{\mathbb{B}} = 2\mathbb{E}\|U_i''\|_{\mathbb{B}}$, and $A_1 = \tau(\mathcal{F}_{i-2}, U_i)$. By definition of $\tau_q(k)$, we have, for any $3\leq i\leq [n/q]$,

$$(4.41) \qquad \tau(\mathcal{F}_{i-2}, U_i) \leq q\tau_q(q+1).$$

Hence, it follows from (4.34) that

$$(4.42) \qquad \sum_{i=3}^{[n/q]}\mathbb{E}\|U_i' - U_i^*\|_{\mathbb{B}} \leq n\tau_q(q+1) + 4\sum_{k=1}^{n}\mathbb{E}(\varphi_M(\|X_k\|_{\mathbb{B}}).$$

Applying (4.42), we infer from (4.39) that

$$(4.43)\quad \mathbb{P}\Big(\sum_{i=3}^{[n/q]+1}\|U_i'-U_i^*\|_{\mathbb{B}}+2\sum_{k=1}^{n}\big(\varphi_M(\|X_k\|_{\mathbb{B}})+\mathbb{E}(\varphi_M(\|X_k\|_{\mathbb{B}}))\big)\geq x\Big)$$
$$\leq \frac{1}{x}\Big(n\tau_q(q+1)+8\sum_{k=1}^{n}\mathbb{E}(\varphi_M(\|X_k\|_{\mathbb{B}}))\Big)\,.$$

Consequently, from (4.37), (4.38) and (4.43),

$$(4.44)\quad \mathbb{P}(\max_{1\leq k\leq n}\|S_k\|_{\mathbb{B}}\geq 4x)\leq 4\exp\Big(-\frac{x}{8qM}\ln\Big(1+\frac{xqM}{s_n^2D^2}\Big)\Big)$$
$$+\frac{1}{x}\Big(n\tau_q(q+1)+8\sum_{k=1}^{n}\mathbb{E}(\varphi_M(\|X_k\|_{\mathbb{B}}))\Big)\,.$$

Choose $v=S(x/r)$, $q=(\tau_\infty^{-1}\circ G^{-1}(v))\wedge n$ and $M=Q(v)$. We have that

$$qM=R(v)=R(S(x/r))\leq x/r\leq x\,.$$

Note that $Q_{\varphi_M(\|X_k\|_{\mathbb{B}})}=(Q_{\|X_k\|_{\mathbb{B}}}-M)_+$. Consequently

$$\sum_{k=1}^{n}\mathbb{E}(\varphi_M(\|X_k\|_{\mathbb{B}}))=\sum_{k=1}^{n}\int_0^1(Q_{\|X_k\|_{\mathbb{B}}}(u)-Q_X(v))_+du\leq n\int_0^v(Q(u)-Q(v))du\,.$$

In addition the choice of q implies that $\tau_q(q+1)\leq\tau_\infty(q)\leq\int_0^v Q(u)du$, and Theorem 1 follows easily from (4.44). ■

5 Application to the complete convergence

In this section, we are interested in the complete convergence (or the convergence rate of the strong law of large numbers) for a τ-dependent sequence with values in a Hilbert space.

Let $(X_k)_{k>0}$ be a sequence of random variables with values in a separable Hilbert space $(\mathbb{H},\|\cdot\|_{\mathbb{H}})$ and let $\mathcal{M}_k=\sigma(X_i,i\leq k)$. Let $S_n=\sum_{i=1}^n(X_i-\mathbb{E}(X_i))$. Here we are interested by conditions under which (1.1) holds for $1/2<\alpha\leq 1$ and $1/\alpha\leq p<\infty$. Note that Property (1.1) with $\alpha p=1$ is equivalent to

$$(5.1)\qquad \sum_{N=1}^{\infty}\mathbb{P}\Big(\max_{1\leq k\leq 2^N}\|S_k\|_{\mathbb{B}}\geq\varepsilon\,2^{N/p}\Big)<\infty\,.$$

Using the monotonicity of the sequence $\max_{1\leq k\leq n}\|S_k\|_{\mathbb{B}}$, we infer from (5.1) that $n^{-1/p}S_n$ tends to 0 almost surely.

In addition Property (1.1) describes speed of convergence in the strong law. Indeed by Lemma 4 in Lai (1977), it implies in case $\alpha p > 1$ that

$$\sum_{n=1}^{\infty} n^{\alpha p-2}\mathbb{P}\Big(\sup_{k\geq n} k^{-\alpha}\|S_k\|_{\mathbb{B}} \geq \varepsilon\Big) < \infty\,. \tag{5.2}$$

Since the probabilities in (5.2) are non-increasing in n, it follows that

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

Definition 6. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. Let X be a random variable with values in $\mathbb{H}$ and let $\mathcal{M}$ be a σ-algebra of $\mathcal{A}$. For any X in $\mathbb{L}^1(\mathbb{H})$, define

$$\gamma(\mathcal{M}, X) = \mathbb{E}(\|\mathbb{E}(X|\mathcal{M}) - \mathbb{E}(X)\|_{\mathbb{H}})\,, \tag{5.3}$$

The coefficients $\gamma(i)$ of the sequence $(X_k)_{k>0}$ are then defined by

$$\gamma(i) = \sup_{k>0} \gamma(\mathcal{M}_k, X_{k+i})\,.$$

In the special case where $p \in]1,2[$, Dedecker and Merlevède (2004, Theorem 3.3) sharpened Theorem 1 in Shao (1993), and proved that (1.1) holds as soon as $(X_i) \prec X$ (see Definition 5) and

$$DM(p,\gamma,X):\quad \sum_{i\geq 0}(i+1)^{p-2}\int_0^{\gamma(i)} Q_X^{p-1}\circ G_X(u)du < \infty\,. \tag{5.4}$$

They also proved that if $\mathbb{E}(X\ln^+(X)) < \infty$ and $\sum_{i\geq 1}\gamma(i)/i < +\infty$ then Property (1.1) holds true with $\alpha = p = 1$. The key of the proof of Theorem 4 in Dedecker and Merlevède is a new maximal inequality in which the dependence coefficients involved are expressed in terms of conditional expectations. However the maximal inequality stated in this paper 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 (1993). Here we shall see that an application of Inequality (4.28) can cover the case where $1 \leq 1/\alpha < 2 \leq p < \infty$.

Theorem 2. *Let $1/2 < \alpha \leq 1$ and p such that $p > 1$ and $1/\alpha \leq p < \infty$. Let $(X_k)_{k>0}$ be a sequence of random variables with values in a separable Hilbert space $(\mathbb{H}, \|\cdot\|_{\mathbb{H}})$ 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 let X be a nonnegative random variable such that $(X_i) \prec X$.*

1. *If $p \in]1,2[$ and $DM(p,\tau_1,X)$ holds then (1.1) is satisfied.*

2. *If $p \geq 2$ and $DM(p,\tau_\infty,X)$ holds then (1.1) is satisfied.*

Remark 7. Since the proof of Theorem 2 comes from an application of Theorem 1, it follows from Remark 5 that Item 2 of Theorem 2 remains valid under the condition $DM(p,\widetilde{\tau}_\infty,X)$.

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

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

1. *$\mathbb{P}(X > x) \leq (c/x)^r$ for some $r > p$, and $\sum_{i\geq 0}(i+1)^{p-2}(\tau_k(i))^{(r-p)/(r-1)} < \infty$.*

2. *$\|X\|_r < \infty$ for some $r > p$, and $\sum_{i\geq 1} i^{(pr-2r+1)/(r-p)}\tau_k(i) < \infty$.*

3. *$\mathbb{E}(X^p(\ln(1+X))^{p-1}) < \infty$ and $\tau_k(i) = O(a^i)$ for some $a < 1$.*

Using Lemma 4, we obtain the following corollary for β-mixing sequences.

Corollary 3. *The following condition implies $DM(p,\tau_k,X)$:*

$$\sum_{i\geq 0}(i+1)^{p-2}\int_0^{\beta_k(i)} Q_X^p(u)du < \infty\,. \tag{5.5}$$

Hence if $\sup_{n\geq 1}\|X_n\|_{\mathbb{H}} < C$ almost surely, then (1.1) holds under the condition $\sum_{i\geq 0}(i+1)^{p-2}\beta_\infty(i) < \infty$, which was first obtained by Berbee (1987) for real valued random variables.

Remark 8. In the case $\mathbb{H} = \mathbb{R}$ and $d(x,y) = |x-y|$, by taking into account Remark 7 and Inequality (3.16), we can derive the following result: Let $1/2 < \alpha \leq 1$ and p such that $p > 1$ and $1/\alpha \leq p < \infty$. Define the coefficients $(\alpha_\infty(i))_{i\geq 0}$ by (3.15). If

$$\sum_{i\geq 0}(i+1)^{p-2}\int_0^{\alpha_\infty(i)} Q_X^p(u)du < \infty\,, \tag{5.6}$$

then (1.1) holds. The condition (5.6) was first used by Rio (1995a) to prove (1.1), but only in the case where $\alpha p = 1$ and $p \in]1,2[$. Let us compare this result with Theorem

1 in Shao (1993): Shao proved that, if $\mathbb{E}(X_n) = 0$ and there exists $r > p$ such that $\sup_{n\geq 1} \|X_n\|_r < \infty$, and if

$$\alpha_\infty(n) = O\left(n^{-\frac{r(p-1)}{r-p}} (\log n)^{-\beta}\right) \text{ with } \beta \geq rp/(r-p)\,, \tag{5.7}$$

then (1.1) holds. However, in his concluding remarks, Shao (1993) made the conjecture that the condition on β may be weakened in $\beta > r/(r-p)$. He also gives an example showing that we cannot take $\beta \leq r/(r-p)$. As quoted in Rio (1995a), page 923, if $\sup_{n\geq 1} \|X_n\|_r < \infty$, then there exists X such that $(X_n) \prec X$ with $\mathbb{P}(X > x) = O(x^{-r})$, so that (1.1) holds as soon as

$$\sum_{i\geq 0} (i+1)^{p-2} \alpha_\infty^{(r-p)/r}(i) < \infty\,.$$

Clearly, this proves Shao's conjecture. Now if the condition $\sup_{n\geq 1} \|X_n\|_r < \infty$ is strengthened in: there exists X such that $(X_n) \prec X$ and $\mathbb{E}(X^r) < \infty$, then according to Relation (C.8) in Rio (2000), (1.1) holds as soon as

$$\sum_{i\geq 0} (i+1)^{(pr-2r+p)/(r-p)} \alpha_\infty(i) < \infty\,,$$

which is true by only assuming $\beta > 1$ in (5.7).

Proof of Theorem 2. Item 1 follows from (5.4) and Item 1 of the following lemma.

Lemma 5. *The following upper bounds hold:*

1. $\gamma(\mathcal{M}, X) \leq \tau(\mathcal{M}, X)$.
2. $\gamma(\mathcal{M}, X) \leq 18 \int_0^{\alpha(\mathcal{M},\sigma(X))} Q_{\|X\|_{\mathbb{H}}}(u)du$.

Proof of Lemma 5. For Item 1, notice that

$$\begin{aligned}\gamma(\mathcal{M}, X) &= \mathbb{E}\Big(\Big\|\int x\, (\mathbb{P}_{X|\mathcal{M}} - \mathbb{P}_X)(dx)\Big\|_{\mathbb{H}}\Big) \\ &= \mathbb{E}\Big(\sup_{y\in\mathbb{H},\, \|y\|_{\mathbb{H}}\leq 1} \int < y, x >_{\mathbb{H}} (\mathbb{P}_{X|\mathcal{M}} - \mathbb{P}_X)(dx)\Big)\,.\end{aligned}$$

Since $x \to < y, x >_{\mathbb{H}}$ belongs to $\Lambda_1(\mathbb{H}, \|\cdot\|_{\mathbb{H}})$, the result follows from the definition of $\tau(\mathcal{M}, X)$. The item 2 is proved in Dedecker and Merlevède (2003), page 250. ■

We now turn to the proof of Item 2 of Theorem 2. Note first that, if s_n^2 is defined as in (4.29), we infer from Inequality (3.33) in Dedecker and Merlevède (2003) that $s_n^2 \leq nK$ for some positive constant K depending on the distribution of X. For the sake of brevity, write Q, R, S and G for Q_X, R_X, S_X and G_X respectively. Applying Inequality (4.28) with $x = x_n = (\varepsilon n^\alpha)/4$, we obtain that, for any $r \geq 1$ and $\varepsilon > 0$,

$$n^{\alpha p-2}\mathbb{P}\Big(\max_{1\leq k\leq n} \|S_k\|_{\mathbb{H}} \geq \varepsilon n^\alpha\Big) \leq I_1(n) + I_2(n),$$

where

$$I_1(n) = 4n^{\alpha p-2}\left(1+\frac{\varepsilon^2 n^{2\alpha-1}}{4rK}\right)^{-r/8} \quad \text{and} \quad I_2(n) = \frac{36 n^{\alpha(p-1)-1}}{\varepsilon}\int_0^{S(x_n/r)} Q(u)du\,.$$

Choose $r > \max(1, 8(\alpha p-1)/(2\alpha-1))$, so that $\sum_{n\geq 1} I_1(n) < \infty$. Now since R is right-continuous and non-increasing,

$$u < S(x_n/r) \iff R(u) > (\varepsilon n^\alpha)/4r \iff n < \Big(\frac{4rR(u)}{\varepsilon}\Big)^{1/\alpha}. \tag{5.8}$$

Applying Fubini, it follows that there exists a finite constant C depending only on α, p and ε, such that

$$\sum_{n\geq 1} I_2(n) \leq C\int_0^1 R^{p-1}(u)Q(u)du \leq C\int_0^1 (\tau_\infty^{-1}\circ G^{-1}(u))^{p-1}Q^p(u)du\,.$$

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

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

which is equivalent to $DM(p, \tau_\infty, X)$ (see for instance Rio (2000), Appendix C). This completes the proof.

6 Application to strong invariances principles

Let us first recall a bounded law of the iterated logarithm for $\mathbb{H}$-valued stationary and ergodic martingale difference sequences $(d_i)_{i\in\mathbb{Z}}$, which can be deduced from Theorem 1 in Morrow and Philipp (1982) (see also the remarks page 112 in Dehling *et al.* (1986)). If $\mathbb{E}(\|d_0\|_{\mathbb{H}}^{2+\varepsilon}) < \infty$ for some $\varepsilon > 0$, then

$$\limsup_{n\to\infty}\frac{\|\sum_{i=1}^n d_i\|_{\mathbb{H}}}{\sqrt{2n\ln\ln n}} = \sqrt{\sup_{\|y\|_{\mathbb{H}}\leq 1}\mathrm{Var}(< y, d_0 >_{\mathbb{H}})} \quad \text{almost surely}\,.$$

Starting from this result, we can derive a bounded law of the iterated logarithm for stationary and ergodic sequences $(X_i)_{i\in\mathbb{Z}}$ of $\mathbb{H}$-valued random variables, under a projective criterion in the style of Gordin (1969). Let $\mathcal{M}_k = \sigma(X_i, i \le k)$. If

$$\sum_{k=1}^{n} \mathbb{E}(X_k|\mathcal{M}_0) \quad \text{converges in } \mathbb{L}^{2+\varepsilon}_{\mathbb{H}}, \tag{6.9}$$

then $S_n = Z_1 - Z_{n+1} + \sum_{i=1}^n d_i$ where $(Z_i)_{i\in\mathbb{Z}}$ is a stationary sequence in $\mathbb{L}^{2+\varepsilon}_{\mathbb{H}}$, and $(d_i)_{i\in\mathbb{Z}}$ is a stationary and ergodic martingale difference sequence in $\mathbb{L}^{2+\varepsilon}_{\mathbb{H}}$. Hence, if (6.9) holds for some $\varepsilon > 0$, then

$$\limsup_{n\to\infty} \frac{\|S_n\|_{\mathbb{H}}}{\sqrt{2n \ln\ln n}} = \sqrt{\sup_{\|y\|_{\mathbb{H}}\le 1} \mathrm{Var}(< y, d_0 >_{\mathbb{H}})} \quad \text{almost surely}\,.$$

A sufficient condition for (6.9) in terms of the coefficients $\gamma(i)$ of Definition 6, is

$$\sum_{i\ge 0} (i+1)^{1+\varepsilon} \int_0^{\gamma(i)} Q^{1+\varepsilon}_{\|X_0\|_{\mathbb{H}}} \circ G_{\|X_0\|_{\mathbb{H}}}(u) < \infty\,. \tag{6.10}$$

By using Item 2 of Lemma 5, it follows easily that (6.10) is satisfied for some $\varepsilon > 0$ as soon as

$$\mathbb{E}(\|X_0\|^{2+\delta}_{\mathbb{H}}) < +\infty \quad \text{and} \quad \alpha_\infty(n) = O(n^{-(2+\epsilon)(1+2/\delta)})\,, \tag{6.11}$$

for some $\delta > 0$ and $\epsilon > 0$. However, from Theorem 1 in Dehling and Philipp (1982), we know that the condition (6.11) can be improved to

$$\mathbb{E}(\|X_0\|^{2+\delta}_{\mathbb{H}}) < +\infty \quad \text{and} \quad \alpha_\infty(n) = O(n^{-(1+\epsilon)(1+2/\delta)})\,, \tag{6.12}$$

at least in the case where $0 < \delta \le 1$.

In fact we shall see in the next theorem that, if we consider the stronger coefficients $\tau_\infty(i)$ instead of $\gamma(i)$, the condition (6.10) can be improved to

$$DM(2, \tau_\infty, \|X_0\|_{\mathbb{H}}) : \quad \sum_{i\ge 0} \int_0^{\tau_\infty(i)} Q_{\|X_0\|_{\mathbb{H}}} \circ G_{\|X_0\|_{\mathbb{H}}}(u)du < \infty\,.$$

As in Theorem 1 in Dehling and Philipp (1982), this condition implies also an almost sure invariance principle. In the real case, this result was proved by Dedecker and Prieur (2004, Theorem 3), and is known to be essentially optimal according to Proposition 3 in Doukhan *et al.* (1994).

Notice that according to Corollary 3, $DM(2, \tau_\infty, \|X_0\|_{\mathbb{H}})$ holds as soon as

$$\sum_{i\geq 0} \int_0^{\beta_\infty(i)} Q^2_{\|X_0\|_{\mathbb{H}}}(u)du < \infty\,, \tag{6.13}$$

which is true under (6.12) with $\beta_\infty(n)$ replacing $\alpha_\infty(n)$ and no restriction on δ (except $\delta > 0$). If $\mathbb{H} = \mathbb{R}$, Rio (1995b) proved that the strong invariance principle holds under the condition (6.13) with $\alpha_\infty(n)$ replacing $\beta_\infty(n)$. It is an open question to know if Rio's result can be extended to separable Hilbert spaces.

Definition 7. A nonnegative self adjoint operator Λ on a separable Hilbert space $\mathbb{H}$ will be called an $\mathcal{S}(\mathbb{H})$-operator if it has finite trace, i.e. for some (and therefore every) orthonormal basis $(e_l)_{l\geq 1}$ of $\mathbb{H}$, $\sum_{l\geq 1} < \Lambda e_l, e_l >_{\mathbb{H}} < \infty$.

Theorem 3. *Let $(X_n)_{n\in\mathbb{Z}}$ be a strictly stationary sequence of random variables with values in a separable Hilbert space $(\mathbb{H}, \|\cdot\|_{\mathbb{H}})$, such that $\mathbb{E}\|X_0\|^2_{\mathbb{H}}$ is finite and $\mathbb{E}(X_0) = 0$. Let $\mathcal{M}_k = \sigma(X_i, i \leq k)$ and $S_n = X_1 + \cdots + X_n$.*

1. *If the sequence is ergodic and $DM(2, \tau_1, \|X_0\|_{\mathbb{H}})$ holds, then $n^{-1/2}S_n$ converges in distribution to $\mathcal{N}(0, \Lambda)$, where the operator $\Lambda \in \mathcal{S}(\mathbb{H})$ is defined by*

$$\Lambda(x, y) = \mathbb{E}(< X_0, x >< X_0, y >) + \sum_{k=1}^{\infty} \mathbb{E}(< X_0, x >< X_k, y >) + \sum_{k=1}^{\infty} \mathbb{E}(< X_0, y >< X_k, x >)\,.$$

2. *If $DM(2, \tau_\infty, \|X_0\|_{\mathbb{H}})$ holds, then there exists a sequence $(Y_n)_{n\in\mathbb{N}}$ of independent $\mathcal{N}(0, \Lambda)$-distributed random variables (possibly degenerate) such that*

$$\Big\| \sum_{i=1}^{n} (X_i - Y_i) \Big\|_{\mathbb{H}} = o\left(\sqrt{n \ln \ln n}\right) \quad a.s.$$

Remark 9. Using Remark 5, the conclusion of Theorem 3.2 remains valid under the condition $DM(2, \widetilde{\tau}_\infty, \|X_0\|_{\mathbb{H}})$, which was first obtained by Dedecker and Prieur (2004) in the real case.

Proof of Theorem 3. Item 1 follows from Corollary 2(β) in Dedecker and Merlevède (2003) and Lemma 5. We now turn to the proof of Item 2.

Definition 8. If $Lx = \max(1, \ln x)$, define the set

$$\Psi = \left\{ \psi : \mathbb{N} \to \mathbb{N}\,,\ \psi \text{ increasing}\,,\ \frac{\psi(n)}{n} \xrightarrow[n\to+\infty]{} +\infty\,,\ \psi(n) = o\,(n\sqrt{LLn}\,) \right\}.$$

If ψ is some function of Ψ, let $M_1 = 0$ and $M_n = \sum_{k=1}^{n-1}(\psi(k) + k)$ for $n \geq 2$. For $n \geq 1$, define the random variables

$$U_n = \sum_{i=M_n+1}^{M_n+\psi(n)} X_i\,, \quad V_n = \sum_{i=M_{n+1}+1-n}^{M_{n+1}} X_i\,, \quad \text{and} \quad Z_n = \sum_{i=M_n+1}^{M_{n+1}} \|X_i\|_{\mathbb{H}}\,.$$

Define the truncated random variables $\overline{U}_n = U_n \mathbb{1}_{\|U_n\|_{\mathbb{H}} \leq n/\sqrt{LLn}}$.

Item 2 of Theorem 3 is a consequence of the following Proposition

Proposition 1. *Let $(X_n)_{n\in\mathbb{Z}}$ be a strictly stationary sequence of random variables with values in a separable Hilbert space $(\mathbb{H}, \|\cdot\|_{\mathbb{H}})$, such that $\mathbb{E}\|X_0\|_{\mathbb{H}}^2$ is finite and $\mathbb{E}(X_0) = 0$. Assume that $DM(2, \tau_\infty, \|X_0\|_{\mathbb{H}})$ holds. There exist a function $\psi \in \Psi$ and a sequence $(W_n)_{n\in\mathbb{N}}$ of independent $\mathcal{N}(0, \psi(n)\Lambda)$-distributed random variables (possibly degenerate) such that*

(a) $\left\| \sum_{i=1}^{n} (W_i - \overline{U}_i) \right\|_{\mathbb{H}} = o\left(\sqrt{M_n LLn}\right)$ *a.s.*

(b) $\sum_{n=1}^{\infty} \frac{\mathbb{E}(\|U_n - \overline{U}_n\|_{\mathbb{H}})}{n\sqrt{LLn}} < \infty$

(c) $Z_n = o\left(n\sqrt{LLn}\right)$ *a.s.*

Proof of Proposition 1. It is adapted from the proof of Proposition 2 in Rio (1995b).

Proof of (b). Note first that $\mathbb{E}\|U_n - \overline{U}_n\|_{\mathbb{H}} \leq 2\mathbb{E}(\|U_n\|_{\mathbb{H}} - n/2\sqrt{LLn})_+$, so that

$$\mathbb{E}\|U_n - \overline{U}_n\|_{\mathbb{H}} \leq 2\int_{\frac{n}{2\sqrt{LLn}}}^{+\infty} \mathbb{P}(\|U_n\|_{\mathbb{H}} > t)dt. \tag{6.14}$$

In the following we write Q instead of $Q_{\|X_0\|_{\mathbb{H}}}$. Since U_n is distributed as $S_{\psi(n)}$, we obtain from (4.28) with s_n^2 as in (4.29),

$$\mathbb{P}(\|U_n\|_{\mathbb{H}} > t) \leq 4\left(1 + \frac{t^2}{4rs_{\psi(n)}^2}\right)^{-\frac{r}{8}} + \frac{36\,\psi(n)}{t}\int_0^{S(t/4r)} Q(u)du. \tag{6.15}$$

We finish the proof as for the proof of (5.6) in Dedecker and Prieur (2004).

Proof of (c). Let $T_n = \sum_{i=M_n+1}^{M_{n+1}} \left(\|X_i\|_{\mathbb{H}} - \mathbb{E}\|X_i\|_{\mathbb{H}} \right)$. We easily see that

$$Z_n = (\psi(n) + n)\,\mathbb{E}\|X_1\|_{\mathbb{H}} + T_n. \tag{6.16}$$

By definition of Ψ, we have $\psi(n) = o\left(n\sqrt{LLn}\right)$. Here note that

$$T_n \le \frac{n}{\sqrt{LLn}} + \left(T_n - \frac{n}{\sqrt{LLn}}\right)_+ . \tag{6.17}$$

Since $\|\cdot\|_{\mathbb{H}}$ is 1-lipschitz, the coefficients $\tau(i)$ of the sequence $(\|X_i\|_{\mathbb{H}} - \mathbb{E}\|X_i\|_{\mathbb{H}})_{i>0}$ are smaller than those of the sequence $(X_i)_{i>0}$. Hence, using similar arguments as for the proof of (b), we obtain that

$$\sum_{n\ge 1} \frac{\mathbb{E}\left(T_n - \frac{n}{\sqrt{LLn}}\right)_+}{n\sqrt{LLn}} < +\infty, \text{ so that } \sum_{n\ge 1} \frac{\left(T_n - \frac{n}{\sqrt{LLn}}\right)_+}{n\sqrt{LLn}} < +\infty \text{ a.s.}$$

Consequently $(T_n - n(LLn)^{-1/2})_+ = o(n\sqrt{LLn})$ almost surely, and the result follows from (6.16) and (6.17).

Proof of (a). In the following, $(\delta_n)_{n\ge1}$ and $(\eta_n)_{n\ge1}$ denote independent sequences of independent random variables with uniform distribution over $[0,1]$, independent of $(X_n)_{n\ge1}$. Using Lemma 1 and arguing as in the proof of Theorem 1, we get the existence of a sequence $(\overline{U}_n^*)_{n\ge1}$ of independent random variables with the same distribution as the random variables $\overline{U}_n$ such that $\overline{U}_n^*$ is a measurable function of $\left(\overline{U}_l, \delta_l\right)_{l\le n}$ and

$$\mathbb{E}\|\overline{U}_n - \overline{U}_n^*\|_{\mathbb{H}} = \tau(\sigma(U_i, i \le n-1), \overline{U}_n)\,.$$

Arguing as in (4.40) of the proof of Theorem 1, we obtain that

$$\begin{aligned} \mathbb{E}\|\overline{U}_n - \overline{U}_n^*\|_{\mathbb{H}} &\le \tau(\sigma(U_i, i\le n-1), U_n) + 2\mathbb{E}\|U_n - \overline{U}_n\|_{\mathbb{H}} \\ &\le \psi(n)\tau(n) + 2\mathbb{E}\|U_n - \overline{U}_n\|_{\mathbb{H}}\,. \end{aligned} \tag{6.18}$$

Since $DM(2, \tau_\infty, \|X_0\|_{\mathbb{H}})$ holds, we obtain from (6.18) and (b) that

$$\sum_{n\ge1} \frac{\mathbb{E}\|\overline{U}_n - \overline{U}_n^*\|_{\mathbb{H}}}{\sqrt{M_n\,LLn}} < +\infty \quad \text{so that} \quad \sum_{n\ge1} \frac{\|\overline{U}_n - \overline{U}_n^*\|_{\mathbb{H}}}{\sqrt{M_n\,LLn}} < +\infty \text{ a.s.}$$

Applying Kronecker's lemma, we obtain that

$$(6.19)\qquad \sum_{i=1}^{n} \|\overline{U}_i - \overline{U}_i^*\|_{\mathbb{H}} = o\left(\sqrt{M_n\, LLn}\right) \text{ a.s.}$$

¿From Corollary 2(β) in Dedecker and Merlevède (2003), we know that $\psi(n)^{-1}\|U_n\|_{\mathbb{H}}^2$ is uniformly integrable and that $\psi(n)^{-1/2}U_n$ converges in distribution to a centered Gaussian mesure P_Λ with covariance operator Λ belonging to $\mathcal{S}(\mathbb{H})$. Using the uniform integrability of $\psi(n)^{-1}\|U_n\|_{\mathbb{H}}^2$ and the fact that $\overline{U}_n^*$ has the same distribution as $\overline{U}_n$, it follows that $\psi(n)^{-1/2}\overline{U}_n^*$ converges in distribution to $\mathcal{N}(0, \Lambda)$. Consequently, if $\rho(P, Q)$ is the Prohorov distance between P and Q, we have that $\rho(P_{\psi(n)^{-1/2}\overline{U}_n^*}, P_\Lambda)$ tends to 0 as n tends to infinity. From Theorem 2 in Dudley (1968), it follows that there exists a sequence μ_n of probability on $(\mathbb{H}\times\mathbb{H}, \mathcal{B}(\mathbb{H})\otimes\mathcal{B}(\mathbb{H}))$ with marginals $P_{\psi(n)^{-1/2}\overline{U}_n^*}$ and P_Λ, such that $\mu_n(\|x-y\|_{\mathbb{H}} > \epsilon)$ converges to 0 for any positive ϵ. By Skorohod's lemma (1976), one can construct some sequence $(W_n)_{n\geq 1}$ of $\sigma(\overline{U}_n^*, \eta_n)$-measurable random variables with respective distribution $\mathcal{N}(0, \psi(n)\Lambda)$ such that, for any $\epsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}\left(\|\overline{U}_n^* - W_n\|_{\mathbb{H}} > \epsilon\sqrt{\psi(n)}\right) = 0\,.$$

Using the uniform integrability of $\psi(n)^{-1}\|\overline{U}_n^*\|_{\mathbb{H}}^2$ and of $\psi(n)^{-1}\|W_n\|_{\mathbb{H}}^2$, it follows that

$$(6.20)\qquad \mathbb{E}\left(\|\overline{U}_n^* - W_n\|_H^2\right) = o\left(\psi(n)\right) \quad \text{as } n \to +\infty\,.$$

Let $\overline{W}_n = W_n \mathbb{1}_{\|W_n\|_{\mathbb{H}}\leq n/\sqrt{LLn}}$. We have

$$\mathbb{E}\|W_n - \overline{W}_n\|_{\mathbb{H}} \leq 2\mathbb{E}(\|W_n\|_{\mathbb{H}} - n/2\sqrt{LLn})_+ = 2\int_{n/2\sqrt{LLn}}^{+\infty} \mathbb{P}(\|W_n\|_{\mathbb{H}} > t)dt\,.$$

Now according to Remark 4 in Pinelis and Sakhanenko (1985), if G is a centered Gaussian random variable with values in $\mathbb{H}$ then

$$(6.21)\qquad \mathbb{P}(\|G\|_{\mathbb{H}} > t) \leq 2\exp\big(-\frac{t^2}{2\mathbb{E}\|G\|_{\mathbb{H}}^2}\big)\,.$$

It follows that

$$\mathbb{E}\|W_n - \overline{W}_n\|_{\mathbb{H}} \leq 8\frac{\mathbb{E}\|W_n\|_{\mathbb{H}}^2}{n}\sqrt{LLn}\exp\big(-\frac{n^2}{8(LLn)\mathbb{E}\|W_n\|_{\mathbb{H}}^2}\big)\,.$$

Then there exist constants C_1 and C_2 depending on ψ and Λ such that

$$(6.22)\qquad \mathbb{E}\|W_n - \overline{W}_n\|_{\mathbb{H}} \leq C_1(LLn)\exp\left(\frac{-C_2 n}{(LLn)^{3/2}}\right)\,,$$

so that $\sum_{n>0} \mathbb{E}\|W_n - \overline{W}_n\|_{\mathbb{H}}$ is finite. By Kronecker's lemma, it follows that

$$\sum_{i=1}^{n} \|W_i - \overline{W}_i\|_{\mathbb{H}} = o(\sqrt{M_n LLn}) \quad \text{a.s.} \tag{6.23}$$

In view of (6.19) and (6.23), it only remains to prove that

$$\Big\| \sum_{i=1}^{n} (\overline{W}_i - \overline{U}_i^*) \Big\|_{\mathbb{H}} = o(\sqrt{M_n LLn}) \quad \text{a.s.}$$

Since $\overline{U}_i^*$ is distributed as $\overline{U}_i$ and since the random variables are centered, $\mathbb{E}(\overline{U}_i^*) = \mathbb{E}(\overline{U}_i - U_i)$. Consequently, Proposition 1(b) yields $\sum_{i=1}^n \|\mathbb{E}(\overline{U}_i^*)\|_{\mathbb{H}} = o(\sqrt{M_n LLn})$. In the same way, $\mathbb{E}(\overline{W}_i) = \mathbb{E}(W_i - \overline{W}_i)$. Then according to (6.22), $\sum_{i=1}^n \|\mathbb{E}(\overline{W}_i)\|_{\mathbb{H}} = o(\sqrt{M_n LLn})$. Hence, it remains to prove that

$$\Big\| \sum_{i=1}^{n} \overline{W}_i - \mathbb{E}(\overline{W}_i) - \overline{U}_i^* + \mathbb{E}(\overline{U}_i^*) \Big\|_{\mathbb{H}} = o(\sqrt{M_n LLn}) \quad \text{a.s.} \tag{6.24}$$

Proof of (6.24). Notice first that

$$\begin{aligned} \mathbb{E}\|\overline{W}_n - \mathbb{E}(\overline{W}_n) - \overline{U}_n^* + \mathbb{E}(\overline{U}_n^*)\|_{\mathbb{H}}^2 &\leq \mathbb{E}\|\overline{W}_n - \overline{U}_n^*\|_{\mathbb{H}}^2 \\ &\leq 2\mathbb{E}\|\overline{W}_n - W_n\|_{\mathbb{H}}^2 + 2\mathbb{E}\|W_n - \overline{U}_n^*\|_{\mathbb{H}}^2 . \end{aligned}$$

Since $\psi(n)^{-1}\|W_n\|_{\mathbb{H}}^2$ is uniformly integrable, $\mathbb{E}\|W_n - \overline{W}_n\|_{\mathbb{H}}^2 = o(\psi(n))$, which combined with (6.20) implies that

$$\mathbb{E}\|\overline{W}_n - \mathbb{E}(\overline{W}_n) - \overline{U}_n^* + \mathbb{E}(\overline{U}_n^*)\|_{\mathbb{H}}^2 \leq \epsilon_n \psi(n)$$

for some sequence ϵ_n of positive reals decreasing to 0 as n tends to infinity. Since the random variables $(\overline{W}_i - \mathbb{E}(\overline{W}_i) - \overline{U}_i^* + \mathbb{E}(\overline{U}_i^*))_{1\leq i \leq n}$ are independent and almost surely bounded by $4n/\sqrt{LLn}$, we obtain from the last inequality of Lemma 6 that for all $x > 0$,

$$\mathbb{P}\left(\sup_{1\leq j\leq n} \Big\| \sum_{i=1}^{j} \Big(\overline{W}_i - \mathbb{E}(\overline{W}_i) - \overline{U}_i^* + \mathbb{E}(\overline{U}_i^*) \Big) \Big\|_{\mathbb{H}} \geq x \right) \\ \leq 2\exp\left(\frac{-x^2}{4\sum_{i=1}^n \epsilon_i \psi(i)} \right) \vee 2\exp\left(\frac{-3x\sqrt{LLn}}{16n} \right).$$

Taking

$$x_n = \max\Big(32n\sqrt{LLn}/3, \Big(8LLn \sum_{i=1}^{n} \epsilon_i \psi(i) \Big)^{1/2} \Big),$$

we obtain that

$$\sum_{n>0} \frac{1}{n} \mathbb{P}\left(\sup_{1\le j\le n} \Big\| \sum_{i=1}^{j} \Big(\overline{W}_i - \mathbb{E}(\overline{W}_i) - \overline{U}_i^* + \mathbb{E}(\overline{U}_i^*) \Big) \Big\|_{\mathbb{H}} \ge x_n \right) < \infty ,$$

which implies (6.24), since $x_n = o(\sqrt{M_n LLn})$.

Proof of Theorem 3. By Skohorod's lemma (1976), there exists a sequence $(Y_i)_{i\ge1}$ of independent $\mathcal{N}(0,\Lambda)$-distributed random variables satisfying $W_n = \sum_{i=M_n+1}^{M_n+\psi(n)} Y_i$ for all positive n. Define the random variable $V_n' = \sum_{i=M_{n+1}+1-n}^{M_{n+1}} Y_i$.

Define $n(k) = \sup\{n \ge 0 : M_n \le k\}$, and note that by definition of M_n we have $n(k) = o(\sqrt{k})$. Applying Proposition 1(c) we see that

$$\Big\| \sum_{i=1}^{k} X_i - \sum_{i=1}^{n(k)} (U_i + V_i) \Big\|_{\mathbb{H}} \le Z_{n(k)} = o\left(\sqrt{k\, LLk}\right) \text{ a.s.} \tag{6.25}$$

Using the same arguments as in the proof of (a) and (b) of Proposition 1, one can prove that there exists a sequence $(W_n')_{n>0}$ of independent $\mathcal{N}(0,n\Lambda)$-distributed random variables such that

$$\Big\| \sum_{i=1}^{n} (V_i - W_i') \Big\|_{\mathbb{H}} = o(\sqrt{M_n LLn}) \text{ a.s.}$$

Since $\| \sum_{i=1}^{n} W_i' \|_{\mathbb{H}} = O(n\sqrt{LLn})$ almost surely, by the bounded law of the iterated logarithm for Gaussian random variables with values in $\mathbb{H}$, we infer that

$$\Big\| \sum_{i=1}^{n(k)} V_i \Big\|_{\mathbb{H}} = o\left(\sqrt{k\, LLk}\right) \text{ a.s.} \quad \text{and also} \quad \Big\| \sum_{i=1}^{n(k)} V_i' \Big\|_{\mathbb{H}} = o\left(\sqrt{k\, LLk}\right) \text{ a.s.} \tag{6.26}$$

Gathering (6.25), (6.26) and Proposition 1(a) and (b), we obtain that

$$\Big\| \sum_{i=1}^{k} X_i - \sum_{i=1}^{n(k)} (W_i + V_i') \Big\|_{\mathbb{H}} = o\left(\sqrt{k\, LLk}\right) \text{ a.s.} \tag{6.27}$$

Clearly $\sum_{i=1}^{k} Y_i - \sum_{i=1}^{n(k)} (W_i + V_i')$ is normally distributed with covariance $(M_{n(k)+1} - k)\Lambda$. Since $n(k) = o(\sqrt{k})$ we have that $M_{n(k)+1} - k \le \psi(n(k)) + n(k) = o(\sqrt{kLLk})$ by definition of ψ. Applying again Inequality (6.21), we infer that there exists a positive constant C depending on Λ and ψ such that, for any $\epsilon > 0$,

$$\mathbb{P}\Big(\Big\| \sum_{i=1}^{k} Y_i - \sum_{i=1}^{n(k)} (W_i + V_i') \Big\|_{\mathbb{H}} > \epsilon\sqrt{kLLk} \Big) \le 2\exp(-C\epsilon^2\sqrt{kLLk}) .$$

Applying Borel-Cantelli, we infer that

$$\Big\| \sum_{i=1}^{k} Y_i - \sum_{i=1}^{n(k)} (W_i + V_i') \Big\|_{\mathbb{H}} = o\Big(\sqrt{k\, LLk}\Big) \text{ a.s.} \tag{6.28}$$

Theorem 3 follows from (6.27) and (6.28).

7 Cramér-von Mises statistics

Let $(X_i)_{1\le i\le n}$ be a strictly stationary sequence of real-valued random variables with common distribution function F. Let F_n be the empirical distribution function $F_n(t) = n^{-1}\sum_{i=1}^{n} \mathbb{I}_{X_i\le t}$. Let μ be a σ-finite measure on $\mathbb{R}$. Suppose that F satisfies

$$\int_{\mathbb{R}_-} (F(t))^2 \mu(dt) + \int_{\mathbb{R}_+} (1 - F(t))^2 \mu(dt) < \infty\,. \tag{7.1}$$

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 space $\mathbb{L}^2(\mu)$. Let $\|.\|_{\mathbb{L}^2(\mu)}$ be the $\mathbb{L}^2$-norm with respect to μ, and define

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

When $\mu = dF$, $D_n^2(\mu)$ is known as the Cramér-von Mises statistics, and is commonly used for testing goodness of fit. It is interesting to write $D_n(\mu)$ as the supremum of the empirical process over a particular class of functions. Indeed,

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

where $W_1(\mu)$ is the set of functions

$$\Big\{ f \,:\, f(t) = f(0) + \Big(\int_{[0,t[} f'(x)\mu(dx)\Big) \mathbb{I}_{t>0} - \Big(\int_{[t,0[} f'(x)\mu(dx)\Big) \mathbb{I}_{t\le 0},\ \|f'\|_{\mathbb{L}^2(\mu)} \le 1 \Big\}.$$

Notice that if λ is the Lebesgue measure on the real line, $W_1(\lambda)$ contains the unit ball of the Sobolev space of order 1 with respect to $\mathbb{L}^2(\lambda)$.

We now define the dependence coefficients which naturally appear in this context. Define first the function F_μ by

$$F_\mu(x) = \mu([0, x[) \text{ if } x \ge 0 \text{ and } F_\mu(x) = -\mu([x, 0[) \text{ if } x \le 0\,. \tag{7.2}$$

Let now d_μ be the distance defined by: for all x and y in $\mathbb{R}$

$$d_\mu(x,y) = \sqrt{|F_\mu(x) - F_\mu(y)|}\,. \tag{7.3}$$

For the sequence $(X_i)_{i\in\mathbb{Z}}$ and the σ-algebras $\mathcal{M}_i = \sigma(X_k, k \le i)$, we define the sequence $(\tau_{d_\mu,\infty}(i))_{i\ge 0}$ as in Definition 3.

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 (7.1). Define the function F_μ by (7.2). Define also $Y_\mu = \sqrt{|F_\mu(X_0)|}$.*

1. *If the sequence is ergodic and Y_μ is integrable, then $D_n(\mu)$ converges to 0 almost surely.*

2. *If $DM(p, \tau_{d_\mu,1}, Y_\mu)$ holds for some $p \in]1,2[$, then for α such that $1/2 < \alpha \le 1$ and $\alpha \ge 1/p$, we have*

$$\sum_{n=1}^{\infty} n^{\alpha p-2}\mathbb{P}\Big(\max_{1\le k\le n} kD_k(\mu) \ge \varepsilon\, n^{\alpha}\Big) < \infty\,. \tag{7.4}$$

   In particular, $n^{(p-1)/p}D_n(\mu)$ converges to 0 almost surely.

3. *If $DM(p, \tau_{d_\mu,\infty}, Y_\mu)$ holds for some $p \ge 2$, then for α such that $1/2 < \alpha \le 1$ and $\alpha \ge 1/p$, the condition (7.4) is satisfied.*

4. *If the sequence is ergodic and $DM(2, \tau_{d_\mu,1}, Y_\mu)$ holds, then $\sqrt{n}D_n(\mu)$ converges in distribution to $\sqrt{\int G^2(x)\mu(dx)}$, where G is a gaussian process in $\mathbb{L}^2(\mu)$ with covariance function defined by*

$$\text{for } (f,g) \text{ in } \mathbb{L}^2(\mu)\times\mathbb{L}^2(\mu), \quad \Lambda(f,g) = \iint f(s)g(t)C(s,t)\mu(dt)\mu(ds),$$

   where $C(s,t) = F(t\wedge s) - F(t)F(s) + 2\sum_{k\ge 1}(\mathbb{P}(Y_0 \le t, Y_k \le s) - F(t)F(s))$.

5. *If $DM(2, \tau_{d_\mu,\infty}, Y_\mu)$ holds then*

$$\limsup_{n\to\infty} \sqrt{\frac{n}{2LLn}}D_n(\mu) = \sqrt{\rho(\Lambda)} \ \text{ almost surely.} \tag{7.5}$$

   where $\rho(\Lambda)$ is the spectral radius of Λ, that is $\rho(\Lambda) = \sup_{\|y\|_{\mathbb{H}}\le 1} < y, \Lambda(y) >_{\mathbb{H}}$.

We now give two sufficient conditions for $DM(p,\tau_{d_\mu,k},Y_\mu)$ to hold.

Corollary 4. *The condition $DM(p,\tau_{d_\mu,k},Y_\mu)$ holds if one of the two following conditions does:*

1. $\displaystyle\sum_{i\geq 1} i^{p-2}\int_0^{\beta_k(i)} Q^p_{Y_\mu}(u)du<\infty.$
2. *F_μ is α-Hölder, that is $|F_\mu(x)-F_\mu(y)|\leq C|x-y|^\alpha$ for $\alpha\in]0,1]$ and $C>0$, and*
$$\sum_{i\geq 1} i^{p-2}\int_0^{(\tau_k(i))^{\alpha/2}} Q^{p-1}_{Y_\mu}\circ G_{Y_\mu}(u)du<\infty\,,$$
where $\tau_k(i)=\tau_{d,k}(i)$ with $d(x,y)=|x-y|$.

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 (7.1) holds. Clearly $D_n(\mu)=n^{-1}\|\sum_{i=1}^n Z_i\|_{\mathbb{H}}$ and

$$(7.6)\quad \|Z_i\|_{\mathbb{H}}\leq \Big(\int_{]-\infty,0[}(\mathbb{1}_{X_i\leq t})^2\mu(dt)+\int_{[0,\infty[}(1-\mathbb{1}_{X_i\leq t})^2\mu(dt)\Big)^{1/2} + \Big(\int_{]-\infty,0[}(F(t))^2\mu(dt)+\int_{[0,\infty[}(1-F(t))^2\mu(dt)\Big)^{1/2},$$

so that $\|Z_i\|_{\mathbb{H}}\leq\sqrt{|F_\mu(X_i)|}+\mathbb{E}(\sqrt{|F_\mu(X_i)|})$ and $\mathbb{E}(\|Z_i\|_{\mathbb{H}})\leq 2\mathbb{E}(Y_\mu)$. Hence Item 1 follows from Mourier's ergodic theorem (1953). Now let $(\tau_k^Z(i))$ be the sequence of coefficients associated to the Hilbert valued random sequence (Z_i) and to the distance $\|\cdot\|_{\mathbb{H}}$. Let $(x_1,\ldots,x_\ell)$ and $(y_1,\ldots,y_\ell)$ be two elements of $\mathbb{R}^\ell$ and define the functions f_i and g_i of $\mathbb{L}^2(\mu)$ by
$$f_i(t)=\mathbb{1}_{x_i\leq t}-F(t)\text{ and }g_i(t)=\mathbb{1}_{y_i\leq t}-F(t)\,.$$
Since for any f in $\Lambda_1(\mathbb{H}^\ell,\|\cdot\|_{\mathbb{H}^\ell})$,
$$|f(f_1,\ldots,f_\ell)-f(g_1,\ldots,g_\ell)|\leq\sum_{i=1}^\ell\sqrt{|F_\mu(x_i)-F_\mu(y_i)|}\,,$$
we clearly get that $\tau^Z(i)\leq\tau_{d_\mu,\infty}(i)$. On the other hand, we infer from (7.6) that
$$(7.7)\qquad Q_{\|Z_i\|_{\mathbb{H}}}\leq Q_{Y_\mu+\mathbb{E}(Y_\mu)}\leq Q_{Y_\mu}+\mathbb{E}(Y_\mu)\,.$$
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 $x\in[0,1]$,
$$(7.8)\qquad \int_0^x Q_{\|Z_0\|_{\mathbb{H}}}(u)du\leq\int_0^x Q_{Y_\mu}(u)du+x\int_0^1 Q_{Y_\mu}(u)du\leq 2\int_0^x Q_{Y_\mu}(u)du\,.$$

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 (7.7) and the last inequality, we infer that

$$\begin{aligned} \int_0^{\tau_k^Z(i)} Q_{\|Z_0\|_{\mathbb{H}}}^{p-1} \circ G_{\|Z_0\|_{\mathbb{H}}}(u)du &\leq 2^{p-1}\Big(\int_0^{\tau_{d_\mu,k}(i)} Q_{Y_\mu} \circ G_{\|Z_0\|_{\mathbb{H}}}(u)du + \int_0^{\tau_{d_\mu,k}(i)} (\mathbb{E}(Y_\mu))^{p-1}du\Big) \\ &\leq 2^p \int_0^{\tau_{d_\mu,k}(i)/2} Q_{Y_\mu} \circ G_{Y_\mu}(u)du + 2^{p-1}\tau_{d_\mu,k}(i)(\mathbb{E}(Y_\mu))^{p-1} . \end{aligned} \tag{7.9}$$

It follows that $DM(p, \tau_k^Z, \|Z\|_{\mathbb{H}})$ holds as soon as $DM(p, \tau_{d_\mu,k}, Y_\mu)$ does. Items 2 and 3 follow by applying Theorem 2 to the sequence $\{Z_i\}_{i\in\mathbb{Z}}$. Item 4 follows from Item 1 of Theorem 3. Now by applying Item 2 of Theorem 3 to the sequence $\{Z_i\}_{i\in\mathbb{Z}}$, we deduce that $\sum_{k=1}^n Z_k$ satisfies the Strassen form of the law of the iterated logarithm with covariance structure Λ, as described in Section 8.2 in Ledoux and Talagrand (1991). Hence Item 5 follows from the limit (8.22) in Ledoux and Talagrand. ■

Proof of Corollary 4. Note first that

$$\int_0^{\tau_{d_\mu,k}(i)} Q_{Y_\mu}^{p-1} \circ G_{Y_\mu}(u)du \leq 2 \int_0^{\tau_{d_\mu,k}(i)/2} Q_{Y_\mu}^{p-1} \circ G_{Y_\mu}(u)du . \tag{7.10}$$

Applying Lemma 4, we get that

$$\tau_{d_\mu,k}(i) \leq 2 \int_0^{\beta_k(i)} Q_{d_\mu(X_0,0)}(u)du = 2 \int_0^{\beta_k(i)} Q_{Y_\mu}(x)dx.$$

Hence $G_{Y_\mu}(\tau_{d_\mu,k}(i)/2) \leq \beta_k(i)$. Using the change-of-variables $v = G_{Y_\mu}(u)$ in (7.10), Item 1 follows. To prove Item 2, notice first that for all $(X_{j_1}^*, \ldots, X_{j_\ell}^*)$ independent of $\mathcal{M}_0$ and distributed as $(X_{j_1}, \ldots, X_{j_\ell})$, we have that

$$\frac{1}{\ell}\tau_{d_\mu}(\mathcal{M}_0, (X_{j_1}, \ldots, X_{j_\ell})) \leq \frac{1}{\ell} \sum_{i=1}^{\ell} \mathbb{E}\sqrt{|F_\mu(X_{j_i}) - F_\mu(X_{j_i}^*)|} .$$

Since F_μ is α-Hölder, we get

$$\frac{1}{\ell}\tau_{d_\mu}(\mathcal{M}_0, (X_{j_1}, \ldots, X_{j_\ell})) \leq \frac{\sqrt{C}}{\ell} \sum_{i=1}^{\ell} \mathbb{E}(|X_{j_i} - X_{j_i}^*|^{\alpha/2}) .$$

Applying Lemma 1, we can choose the ℓ-tuple $(X_{j_1}^*, \ldots, X_{j_\ell}^*)$ such that

$$\sum_{i=1}^{\ell} \mathbb{E}(|X_{j_i} - X_{j_i}^*|) = \tau(\mathcal{M}_0, (X_{j_1}, \ldots, X_{j_\ell})) .$$

Hence, using Jensen's inequality, it follows that

$$\frac{1}{\ell}\tau_{d_\mu}(\mathcal{M}_0,(X_{j_1},\dots,X_{j_\ell}))\le\sqrt{C}\Big(\frac{1}{\ell}\tau(\mathcal{M}_0,(X_{j_1},\dots,X_{j_\ell})\Big)^{\alpha/2}.$$

Consequently, $\tau_{d_\mu,k}(i)\le\sqrt{C}(\tau_k(i))^{\alpha/2}$, and the result follows.

8 Appendix

In this section, we recall a result given in Pinelis (1994, Theorem 3.4).

Lemma 6. *Let $(\mathbb{B},\|\cdot\|_{\mathbb{B}})$ be a $(2,D)$-smooth separable Banach space. Let $\{d_j,\mathcal{F}_j\}_{j\ge1}$ be a sequence of $\mathbb{B}$-valued martingale differences such that*

$$\|\,\|d_j\|_{\mathbb{B}}\|_\infty\le c \quad \textit{and} \quad \Big\|\sum_{j=1}^n\mathbb{E}(\|d_j\|^2_{\mathbb{B}}|\mathcal{F}_{j-1})\Big\|_\infty\le y_n\,.$$

Set $M_j=\sum_{i=1}^j d_i$. Then for all $x>0$,

$$\mathbb{P}\left(\sup_{1\le j\le n}\|M_j\|_{\mathbb{B}}\ge x\right)\le 2\exp\left(-\frac{y_nD^2}{c^2}\,h\left(\frac{xc}{y_nD^2}\right)\right) \tag{8.1}$$

where $h(u)=(1+u)\ln(1+u)-u$. Consequently, we have the bounds

$$\begin{aligned}\mathbb{P}\left(\sup_{1\le j\le n}\|M_j\|_{\mathbb{B}}\ge x\right)&\le 2\exp\left(-\frac{x}{2c}\ln\left(1+\frac{xc}{y_nD^2}\right)\right),\ \textit{and}\\ \mathbb{P}\left(\sup_{1\le j\le n}\|M_j\|_{\mathbb{B}}\ge x\right)&\le 2\exp\left(\frac{-x^2}{2y_nD^2+2cx/3}\right)\\ &\le 2\max\left(\exp\left(\frac{-x^2}{4y_nD^2}\right),\exp\left(\frac{-3x}{4c}\right)\right).\end{aligned}$$

References

[1] 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.*

[2] Berbee, H.C.P. (1979). Random walks with stationary increments and renewal theory. *Math. Cent. Tracts.* Amsterdam.

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

[4] Buzzi, J. and Maume-Deschamps, V. (2002). Decay of correlations for piecewise invertible maps in higher dimension. *Israel J. Math.* **131** 203-220.

[5] Collet, P., Martinez, S. and Schmitt, B. (2002). Exponential inequalities for dynamical measures of expanding maps of the interval. *Probab. Theory. Relat. Fields* **123** 301-322.

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

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

[8] Dedecker, J. and Merlevède, F. (2004). Vitesses de convergence dans la loi forte des grands nombres pour des variables dépendantes. *C. R. Math. Acad. Sci. Paris* **339** 883-886.

[9] Dedecker, J. and Prieur, C. (2004). Coupling for τ-dependent sequences and applications. *J. Theoret. Probab.* **17**, 861–885.

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

[11] Dedecker, J., Prieur, C. and Raynaud De Fitte, P. (2005). Parametrized Kantorovich-Rubinštein theorem and application to the coupling of random varaibles. *To appear.*

[12] Dehling, H. (1983). A note on a theorem of Berkes and Philipp, *Z. Wahrsch. Verw. Gebiete* **62** 39-42.

[13] Dehling, H. and Philipp, W. (1982). Almost sure invariance principles for weakly dependent vector-valued random variables. *Ann. Probab.* **10** 689-701.

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

[15] Doukhan, P., Massart, P. and Rio, E. (1994). The functional central limit theorem for strongly mixing processes. *Ann. Inst. H. Poincaré Probab. Statist.* **30** 63-82.

[16] Dudley, R. M. (1968) Distances of probability measures and random variables. *Ann. Math. Statist.* **39** 1563-1572.

[17] Fréchet, M. (1957). Sur la distance de deux lois de probabilité, *C.R. Acad. Sci. Paris*, **244** 689-692.

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

[19] Ibragimov, I. A. (1962). Some limit theorems for stationary processes. *Theory Probab. Appl.* **7** 349-382.

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

[21] Ledoux, M. and Talagrand, M. (1991). *Probability in Banach spaces.* Springer-Verlag, Berlin.

[22] Morrow, G. and Philipp, W. (1982). An almost sure invariance principle for Hilbert space valued martingales. *Trans. Amer. Math. Soc.* **273** 231-251.

[23] Mourier, E. (1953). Eléments aléatoires dans un espace de Banach. *Ann. Inst. H. Poincaré* **13**. 161-244.

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

[25] Pinelis, I. F. and Sakhanenko, A. I. (1985). Remarks on inequalities for probabilities of large deviations. *Theory Probab. Appl.* **30**, 143-148.

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

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

[28] Rio, E. (1995b). The functional law of the iterated logarithm for stationary strongly mixing sequences. *Ann. Probab.* **23** 1188–1203.

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

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

[31] Rozanov, Y. A. and Volkonskii, V. A. (1959). Some limit theorems for random functions I. *Theory Probab. Appl.* **4** 178-197.

[32] Rüschendorf, L. (1985). The Wasserstein Distance and Approximation Theorems. *Z. Wahrsch. Verw. Gebiete.*, **70** 117-129.

[33] Saussol, B. (2000). Absolutely continuous invariant measures for multidimensional expanding maps. *Isr. J. Math.*, **116**, 223-248.

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

[35] Skorohod, A. V. (1976). On a representation of random variables. *Theor. Probability Appl.* **21** 628-632

[36] 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.