On a maximal inequality for strongly mixing random variables in Hilbert spaces. Application to the compact law of the iterated logarithm.

Florence Merlevède*
Université Paris VI

May 23, 2008

Abstract

In this paper, we state a maximal inequality for the partial sums of strongly mixing sequences of Hilbert space valued random variables. This inequality allows to derive the almost sure compactness of the partial sums divided by the normalizing sequence $(n \log\log n)^{1/2}$. As a consequence, we derive the compact law of the iterated logarithm under the same condition than the one required in the real case, which is known to be essentially optimal. An application to Cramér-von Mises statistics is given.

Mathematics Subject Classifications (2000): 60 F 15.
Key words: strong mixing sequences, Hilbert space, maximal inequalities, compact law of the iterated logarithm.
Short Title: Limit theorems in Hilbert spaces.

*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: florence.merlevede@upmc.fr

1 Introduction

Having suitable upper bounds for the deviation of the maximum of partial sums of random variables is a way to derive almost sure convergence results for the partial sums. For real valued random variables, by using Bernstein blocking technique and coupling arguments allowing to approximate the blocks by independent ones having the same law, Rio (1995, 2000) derived some maximal inequalities which are closer to Fuk-Nagaev inequalities than to classical exponential inequalities. Indeed, the upper bound depends on the strong mixing coefficient α of the underlying sequence and for arithmetical rates of convergence of the strong mixing coefficients, this upper bound also decreases arithmetically. For random variables taking values in a real separable Hilbert space $(\mathbb{H}, \|\cdot\|_{\mathbb{H}})$, Dedecker and Merlevède (2006) also used blocking techniques and coupling arguments to get a similar upper bound than the one obtained by Rio but with an other coefficient of dependence replacing the strong mixing one. Indeed, since the works of Berbee (1979) and Rüschendorf (1985), we know that the price to pay for replacing the initial dependent sequence by an independent one having the same marginals, is exactly the value of some dependence coefficients τ having the *coupling property* for $\|\cdot\|_{\mathbb{H}}$ (see for instance Lemma 1 in Dedecker and Merlevède (2006)). In the real case, this coupling coefficient τ can be controlled by the strong mixing one (see Remarks 2 and 3 in Dedecker and Merlevède (2006)). Unfortunately if we deal with infinite dimensional spaces such a control cannot hold. As a matter of fact, Dehling (1983) constructed examples of strongly mixing sequences of ℓ^2-valued random variables X_k which cannot be approximated by independent random variables Y_k in any useful way. By this we mean convergence of $X_k - Y_k$ to zero in probability. Consequently, in order to get a Fuk-Nagaev type inequality for the maximum of the partial sums of strongly mixing random variables with values in an infinite dimensional space, some other ideas than approximations by independent random variables are needed. In this paper, rather than using coupling arguments, we use a martingale approximation of blocks. Obviously, the price to pay for such a technique is the fact that Fuk-Nagaev type inequalities for martingales are valid but only with an appropriate control of the conditional variance. Hence compared to the method using coupling arguments, an additional term, namely the probability of deviation of the conditional variance, has to be controlled. However as we shall see in the section 2.2, the maximal inequality gives a suitable upper bound to derive the almost sure relative compactness of the partial sums normalized by $\sqrt{n \log\log n}$ which is one of the essential tool to derive the compact law of the iterated logarithm. More precisely, we show that the compact law of the iterated logarithm holds for strongly mixing sequences of Hilbert space valued random variables satisfying the same condition than the one required by Rio (1995) in the real case, which is known to be essentially optimal. In this section, we also provide an application to Cramér-von Mises statistics. All the proofs are postponed to

Section 3.

2 Results

In this paper we deal with strongly mixing sequences of random variables defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, and taking their values in a real separable Hilbert space $\mathbb{H}$ with norm $\|\cdot\|_{\mathbb{H}}$ generated by an inner product $< \cdot, \cdot >_{\mathbb{H}}$. Let us first recall the definition of the strong mixing coefficient α introduced by Rosenblatt (1956): For any two σ algebras $\mathcal{A}$ and $\mathcal{B}$, we define the α-mixing coefficient by

$$\alpha(\mathcal{A}, \mathcal{B}) = \sup_{A \in \mathcal{A}, B \in \mathcal{B}} |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|.$$

Let $(X_k, k \in \mathbb{Z})$ be a sequence of random variables defined on $(\Omega, \mathcal{A}, \mathbb{P})$ with values in $\mathbb{H}$. This sequence will be called strongly mixing if

$$\alpha(n) := \sup_{k \in \mathbf{Z}} \alpha\left(\mathcal{M}_k, \mathcal{G}_{k+n}\right) \to 0 \text{ as } n \to \infty\,, \tag{2.1}$$

where $\mathcal{M}_j := \sigma(X_i, i \le j)$ and $\mathcal{G}_j := \sigma(X_i, i \ge j)$ for $j \in \mathbb{Z}$.

2.1 A maximal inequality for Hilbert valued random variables

Before stating the maximal inequality proved in this paper, we shall introduce some notations.

Definition 1. For any nonnegative integrable random variable X, define the "upper tail" quantile function Q_X by $Q_X(u) = \inf\{t \ge 0 : \mathbb{P}(X > t) \le u\}$. Given a real separable Hilbert space $(\mathbb{H}, \|\cdot\|_{\mathbb{H}})$, let $(X_i)_{i \in \mathbf{Z}}$ be a sequence of $\mathbb{H}$-valued random variables. We write $(X_i) \prec X$ if there exists a nonnegative random variable X such that $Q_X \ge \sup_{k \in \mathbf{Z}} Q_{\|X_k\|_{\mathbb{H}}}$.

Theorem 1. *Let $(\mathbb{H}, \|\cdot\|_{\mathbb{H}})$ be a real separable Hilbert space. Let $\{X_k\}_{k \in \mathbb{Z}}$ be a sequence of centered random variables with values in $\mathbb{H}$. Define the coefficients $\alpha(i)$ as in (2.1) and let $\alpha^{-1}(u) = \sum_{i \ge 0} \mathbb{I}_{u < \alpha(i)}$. Let X be a nonnegative random variable such that $(X_i) \prec X$. Let $R_X(u) = (\alpha^{-1}(u) \wedge n) Q_X(u)$ and $H_X(u) = R_X^{-1}(u)$. Then, for any $x > 0$ and $r \ge 1$ and every quantity s_n^2 such that*

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

one has that

$$\mathbb{P}(\sup_{1 \le k \le n} \|S_k\|_{\mathbb{H}} \ge 4x) \quad \le \quad 4 \exp\left(-\frac{r^2 s_n^2}{8x^2} h\left(\frac{2x^2}{r s_n^2}\right)\right) + n\Big\{\frac{22}{x} + \frac{32x}{r s_n^2}\Big\} \int_0^{H_X(x/r)} Q_X(u)du\,,$$

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

Remark 1. Since $h(u) \geq u \ln(1+u)/2$, under the notation and assumptions of the above theorem, we get that for any $x > 0$ and $r \geq 1$,

$$\mathbb{P}(\sup_{1\leq k\leq n} \|S_k\|_{\mathbb{H}} \geq 4x) \quad \leq \quad 4\left(1+\frac{2x^2}{rs_n^2}\right)^{-r/8} + n\{\frac{22}{x}+\frac{32x}{rs_n^2}\}\int_0^{H_X(x/r)} Q_X(u)du\,, \qquad (2.2)$$

2.2 The compact law of the iterated logarithm in Hilbert space

In this section, we write Lu to denote $\log u$ for $u \geq e$ and 1 otherwise. The function $L(Lu)$ is written LLu. Also for a sequence $(x_n) \in \mathbb{H}$ and $K \subset \mathbb{H}$, we denote $d(x_n, K) = \inf\{\|x_n - y\|_{\mathbb{H}};\, y \in K\}$ and $C(x_n)$ the set of the cluster points of (x_n) (that is all possible limit points of the sequence (x_n)).

Definition 2. 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 $\mathbb{H}$, $\sum_{l\geq 1} < \Lambda e_l, e_l >_{\mathbb{H}} < \infty$.

The following theorem shows that the compact law of the iterated logarithm holds for strongly mixing sequences of Hilbert space valued random variables satisfying the same condition than the one required by Rio (1995) in the real case, and which is known to be essentially optimal according to Proposition 3 in Doukhan, Massart and Rio (1994).

Theorem 2. *Let $(\mathbb{H}, \|\cdot\|_{\mathbb{H}})$ be a real separable Hilbert space. Let $(X_k)_{k\in\mathbb{Z}}$ is a strictly stationary sequence of centered random variables with values in $\mathbb{H}$ such that $\mathbb{E}(\|X_0\|_{\mathbb{H}}^2) < \infty$. Define the coefficients $\alpha(k)$ as in (2.1). Assume that the following condition is satisfied*

$$\sum_{k\geq 1} \int_0^{\alpha(k)} Q^2_{\|X_0\|_{\mathbb{H}}}(u)du < \infty\,. \qquad (2.3)$$

Then the operator Λ defined for any x and y in $\mathbb{H}$ by

$$\begin{aligned}\Lambda(x,y) \quad &= \quad \mathbb{E}(< X_0, x >_{\mathbb{H}} < X_0, y >_{\mathbb{H}}) + \sum_{k=1}^{\infty} \mathbb{E}(< X_0, x >_{\mathbb{H}} < X_k, y >_{\mathbb{H}}) \qquad (2.4)\\ &\quad + \sum_{k=1}^{\infty} \mathbb{E}(< X_0, y >_{\mathbb{H}} < X_k, x >_{\mathbb{H}})\end{aligned}$$

is in $\mathcal{S}(\mathbb{H})$. In addition if K denotes the unit ball of the reproducing kernel Hilbert space associated with Λ, then with probability one,

$$\lim_{n\to\infty} d\left(\frac{S_n}{\sqrt{2nLLn}}\,, K\right) = 0 \text{ and } C\left(\frac{S_n}{\sqrt{2nLLn}}\right) = K\,.$$

Remark 2. Notice that Dedecker and Merlevède (2006, Theorem 3) obtained the conclusion of the above theorem with the absolute regular coefficient β replacing the strong mixing one in the condition (2.3) (see their condition (6.5)).

Remark 3. In view of applications, we mention that the couple of conditions

$$\mathbb{E}(\|X_0\|_{\mathbb{H}}^{2+\delta}) < \infty, \text{ and } \sum_{k=1}^{\infty} k^{2/\delta}\alpha(k) < \infty, \text{ for some } \delta > 0$$

are sufficient for the validity of (2.3). This couple of conditions has to be compared with the conditions required in Theorem 1 of Dehling and Philipp (1982). As a consequence of their result, they derived the conclusion of Theorem 2 under the conditions that for $0 < \delta \leq 1$ and $0 < \varepsilon \leq 1$

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

Theorem 2 then shows that for arithmetic decrease of the strong mixing coefficients, the compact law of the iterated logarithm holds without the limitation $0 < \delta \leq 1$. However, it still an open question if the almost sure invariance principle holds under (2.3).

Example 1. *Cramér-von Mises statistics.* Let $(X_i)_{i\in\mathbb{Z}}$ be a strictly stationary sequence of real-valued random variables with common distribution function F, and let $\alpha_X(i)$ be the sequence of strong mixing coefficients defined by (2.1) associated to the sequence $(X_i)_{i\in\mathbb{Z}}$. Let F_n be the empirical distribution function $F_n(t) = n^{-1}\sum_{i=1}^{n} \mathbb{1}_{X_i \leq t}$. Let μ be a σ-finite measure on $\mathbb{R}$. Suppose that F satisfies

$$\int_{\mathbb{R}_-} (F(t))^2 \mu(dt) + \int_{\mathbb{R}_+} (1 - F(t))^2 \mu(dt) < \infty\,. \tag{2.5}$$

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 $\mathbb{L}^2(\mu)$. 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. We refer to the paper of Dedecker and Merlevède (2006, page 201) where $D_n(\mu)$ is rewritten as the supremum of the empirical process over a particular class of functions. We now define

$$F_\mu(x) = \mu([0, x[) \text{ if } x \geq 0\ ,\ F_\mu(x) = -\mu([x, 0[) \text{ if } x \leq 0 \text{ and } Y_\mu = \sqrt{|F_\mu(X_0)|}\,.$$

Let $Z_i = \{t \to \mathbb{1}_{X_i \leq t} - F(t), t \in \mathbb{R}\}$ which belongs to $\mathbb{L}^2(\mu)$ as soon as (2.5) holds. Notice that the strong mixing coefficient $\alpha_Z(k)$ associated with the sequence $(Z_i)_{i\in\mathbb{Z}}$ is bounded by $\alpha_X(k)$.

Also, we have that $\|Z_i\|_{\mathbb{L}^2(\mu)} \le \sqrt{|F_\mu(X_i)|} + \mathbb{E}\sqrt{|F_\mu(X_i)|}$. Hence applying Theorem 2 to the variables Z_i, we derive the following proposition:

Proposition 1. *If (2.5) holds and if*

$$\sum_{k\ge 1} \int_0^{\alpha_X(k)} Q^2_{Y_\mu}(u)du < \infty \,, \tag{2.6}$$

then

$$\limsup_{n\to\infty} \sqrt{\frac{n}{2LLn}} D_n(\mu) = \sqrt{\rho(\Lambda)} \ \text{ almost surely,} \tag{2.7}$$

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

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

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

We would like to mention that in case where F_μ is Lipschitz, then (2.7) has been obtained by Dedecker and Merlevède (2006) but with $\sqrt{\alpha_X(k)}$ replacing $\alpha_X(k)$ in the condition (2.6), which is clearly more restrictive.

3 Proofs

3.1 Proof of Theorem 1

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

$$U_i' = U_i \mathbb{1}_{\|U_i\|_{\mathbb{H}}\le 2qM} \quad \text{and} \quad U_i'' = U_i \mathbb{1}_{\|U_i\|_{\mathbb{H}}> 2qM} \,. \tag{3.1}$$

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

$$\sup_{1\le k\le n} \|S_k\|_{\mathbb{H}} \le \sup_{1\le j\le [n/q]} \|\sum_{i=1}^{j} U_i'\|_{\mathbb{H}} + qM + 2\sum_{k=1}^{n} \varphi_M(\|X_k\|_{\mathbb{H}}) \,. \tag{3.2}$$

To prove (3.2), it suffices to notice that, if the maximum of $\|S_k\|_{\mathbb{H}}$ is obtained in k_0, then for $j_0 = [k_0/q]$,

$$\sup_{1\le k\le n} \|S_k\|_{\mathbb{H}} \le \|\sum_{i=1}^{j_0} U_i'\|_{\mathbb{H}} + \sum_{i=1}^{j_0} \|U_i''\|_{\mathbb{H}} + \sum_{k=qj_0+1}^{k_0} \|X_k\|_{\mathbb{H}} \,. \tag{3.3}$$

Notice now that

$$\sum_{k=qj_0+1}^{k_0}\|X_k\|_{\mathbb{H}} \le (k_0-qj_0)M + \sum_{k=qj_0+1}^{k_0}\varphi_M(\|X_k\|_{\mathbb{H}})\,. \tag{3.4}$$

On the other hand, using the fact that $|x|\mathbb{1}(|x|>2A)\le 2(|x|-A)_+$, we get that

$$\sum_{i=1}^{j_0}\|U_i''\|_{\mathbb{H}} \le 2\sum_{i=1}^{j_0}\varphi_{qM}(\|U_i\|_{\mathbb{H}})\,.$$

Then by convexity of the function φ_{qM}, we derive that

$$\sum_{i=1}^{j_0}\|U_i''\|_{\mathbb{H}} \le 2\sum_{k=1}^{qj_0}\varphi_M(\|X_k\|_{\mathbb{H}})\,. \tag{3.5}$$

Starting from (3.3) and using (3.4) and (3.5), we get (3.2).

Setting now for all $i\ge 1$, $\mathcal{F}_i^U=\sigma(X_j, j\le iq)$, we define a sequence $(\tilde U_i)_{i\ge 1}$ as follows: for all $i\ge 1$, $\tilde U_{2i-1}=U'_{2i-1}-\mathbb{E}(U'_{2i-1}|\mathcal{F}^U_{2(i-1)-1})$ and $\tilde U_{2i}=U'_{2i}-\mathbb{E}(U'_{2i}|\mathcal{F}^U_{2(i-1)})$. Notice that $(\tilde U_i)_{i\ge1}$. is a sequence of martingale differences with respect to $(\mathcal{F}_i^U)$. Substituting the variables $\tilde U_i$ to the initial variables, in the inequality (3.2), we derive the following upper bound

$$\max_{1\le k\le n}\|S_k\|_{\mathbb{H}} \le qM + \max_{2\le 2j\le [n/q]}\Big\|\sum_{i=1}^{j}\tilde U_{2i}\Big\|_{\mathbb{H}} + \max_{1\le 2j-1\le [n/q]}\Big\|\sum_{i=1}^{j}\tilde U_{2i-1}\Big\|_{\mathbb{H}} + \sum_{i=1}^{[n/q]}\|U_i'-\tilde U_i\|_{\mathbb{H}} + 2\sum_{k=1}^{n}\varphi_M(\|X_k\|_{\mathbb{H}})\,. \tag{3.6}$$

Since $\|U_i'\|_{\mathbb{H}}\le 2qM$ almost surely, it follows that $\|\tilde U_i\|_{\mathbb{H}}\le 4qM$ almost surely. Then applying Lemma 1 of the appendix with $y=2s_n^2$, we derive that

$$\mathbb{P}\Big(\max_{2\le 2j\le [n/q]}\Big\|\sum_{i=1}^{j}\tilde U_{2i}\Big\|_{\mathbb{H}}\ge x\Big) \le 2\exp\Big(-\frac{s_n^2}{8(qM)^2}h\Big(\frac{2xqM}{s_n^2}\Big)\Big) + \mathbb{P}\Big(\sum_{i=1}^{[[n/q]/2]}\mathbb{E}(\|\tilde U_{2i}\|^2_{\mathbb{H}}|\mathcal{F}^U_{2(i-1)})\ge 2s_n^2\Big)\,. \tag{3.7}$$

Now notice that

$$\mathbb{E}(\|\tilde U_{2i}\|^2_{\mathbb{H}}|\mathcal{F}^U_{2(i-1)}) = \mathbb{E}(\|U'_{2i}\|^2_{\mathbb{H}}|\mathcal{F}^U_{2(i-1)}) - \|\mathbb{E}(U'_{2i}|\mathcal{F}^U_{2(i-1)})\|^2_{\mathbb{H}} \le \mathbb{E}(\|U'_{2i}\|^2_{\mathbb{H}}|\mathcal{F}^U_{2(i-1)})\,.$$

Then it follows that

$$\mathbb{P}\Big(\sum_{i=1}^{[[n/q]/2]}\mathbb{E}(\|\tilde U_{2i}\|^2_{\mathbb{H}}|\mathcal{F}^U_{2(i-1)})\ge 2s_n^2\Big) \le \mathbb{P}\Big(\sum_{i=1}^{[[n/q]/2]}\mathbb{E}(\|U'_{2i}\|^2_{\mathbb{H}}|\mathcal{F}^U_{2(i-1)})\ge 2s_n^2\Big)\,. \tag{3.8}$$

Since $\sum_{i=1}^{[[n/q]/2]} \mathbb{E}(\|U'_{2i}\|_{\mathbb{H}}^2) \leq \sum_{i=1}^{[[n/q]/2]} \mathbb{E}(\|U_{2i}\|_{\mathbb{H}}^2) \leq s_n^2$, we clearly get from Markov's inequality that

$$\mathbb{P}\Big(\sum_{i=1}^{[[n/q]/2]} \mathbb{E}(\|U'_{2i}\|_{\mathbb{H}}^2 | \mathcal{F}^U_{2(i-1)}) \geq 2s_n^2 \Big) \leq \frac{1}{s_n^2} \sum_{i=1}^{[[n/q]/2]} \mathbb{E}|\mathbb{E}(\|U'_{2i}\|_{\mathbb{H}}^2 | \mathcal{F}^U_{2(i-1)}) - \mathbb{E}\|U'_{2i}\|_{\mathbb{H}}^2| \,. \tag{3.9}$$

Then starting from (3.7) and using (3.8) and (3.9), we get that

$$\begin{aligned} \mathbb{P}\Big(\max_{2\leq 2j \leq [n/q]} \Big\| \sum_{i=1}^{j} \tilde{U}_{2i} \Big\|_{\mathbb{H}} \geq x \Big) &\leq 2\exp\left(-\frac{s_n^2}{8(qM)^2} h\left(\frac{2xqM}{s_n^2}\right)\right) \\ &\quad + \frac{1}{s_n^2} \sum_{i=1}^{[[n/q]/2]} \mathbb{E}|\mathbb{E}(\|U'_{2i}\|_{\mathbb{H}}^2 | \mathcal{F}^U_{2(i-1)}) - \mathbb{E}\|U'_{2i}\|_{\mathbb{H}}^2| \,. \end{aligned} \tag{3.10}$$

Obviously similar computations allow to treat the quantity $\max_{1\leq 2j-1 \leq [n/q]} \Big\| \sum_{i=1}^{j} \tilde{U}_{2i-1} \Big\|_{\mathbb{H}}$. Hence we get that

$$\begin{aligned} \mathbb{P}\Big(\max_{2\leq 2j \leq [n/q]} \Big\| \sum_{i=1}^{j} \tilde{U}_{2i} \Big\|_{\mathbb{H}} + \max_{1\leq 2j-1 \leq [n/q]} \Big\| \sum_{i=1}^{j} \tilde{U}_{2i-1} \Big\|_{\mathbb{H}} \geq 2x \Big) &\leq 4\exp\left(-\frac{s_n^2}{8(qM)^2} h\left(\frac{2xqM}{s_n^2}\right)\right) \\ &+ \frac{1}{s_n^2} \sum_{i=1}^{[n/q]} \mathbb{E}|\mathbb{E}(\|U'_{i}\|_{\mathbb{H}}^2 | \mathcal{M}_{(i-2)q}) - \mathbb{E}\|U'_{i}\|_{\mathbb{H}}^2| \,, \end{aligned}$$

where we recall that $\mathcal{M}_k = \sigma(X_j, j \leq k)$. Setting $A_i := \|U'_i\|_{\mathbb{H}}^2 - \mathbb{E}\|U'_i\|_{\mathbb{H}}^2$, we have that

$$\begin{aligned} &\mathbb{E}|\mathbb{E}(\|U'_{i}\|_{\mathbb{H}}^2 | \mathcal{M}_{(i-2)q}) - \mathbb{E}(\|U'_{i}\|_{\mathbb{H}}^2| \\ &\quad = \mathbb{E}\Big\{ \mathbb{E}(A_i|\mathcal{M}_{(i-2)q}) \Big(\mathbb{1}(\mathbb{E}(A_i|\mathcal{M}_{(i-2)q}) \geq 0) - \mathbb{1}(\mathbb{E}(A_i|\mathcal{M}_{(i-2)q}) < 0) \Big) \Big\} \\ &\quad = \mathbb{E}\Big\{ A_i \Big(\mathbb{1}(\mathbb{E}(A_i|\mathcal{M}_{(i-2)q})) \geq 0) - \mathbb{1}(\mathbb{E}(A_i|\mathcal{M}_{(i-2)q})) < 0) \Big) \Big\} \,. \end{aligned}$$

Hence using Rio's covariance inequality (1993) and the fact that $\|A_i\|_\infty \leq 2 \times (2qM)^2$, we derive that

$$\begin{aligned} \mathbb{E}|\mathbb{E}(\|U'_{i}\|_{\mathbb{H}}^2 | \mathcal{M}_{(i-2)q}) - \mathbb{E}(\|U'_{i}\|_{\mathbb{H}}^2| &\leq 4\int_0^{\alpha(q)} Q_{|A_i|}(u)du \\ &\leq 32(qM)^2 \alpha(q) \,. \end{aligned} \tag{3.11}$$

It follows that

$$\mathbb{P}\Big(\max_{2\leq 2j \leq [n/q]} \Big\| \sum_{i=1}^{j} \tilde{U}_{2i} \Big\|_{\mathbb{H}} + \max_{1\leq 2j-1 \leq [n/q]} \Big\| \sum_{i=1}^{j} \tilde{U}_{2i-1} \Big\|_{\mathbb{H}} \geq 2x \Big) \tag{3.12}$$

$$\leq 4\exp\left(-\frac{s_n^2}{8(qM)^2} h\left(\frac{2xqM}{s_n^2}\right)\right) + \frac{32nqM}{s_n^2} M\alpha(q) \,. \tag{3.13}$$

Now by using Markov's inequality, we get that

$$\mathbb{P}\Big(\sum_{i=3}^{[n/q]}\|U_i'-\tilde{U}_i\|_{\mathbb{H}}+2\sum_{k=1}^{n}\varphi_M(\|X_k\|_{\mathbb{H}})\geq x\Big)$$
$$\leq x^{-1}\Big(\sum_{i=1}^{[n/q]}\mathbb{E}\|\mathbb{E}(U_i'|\mathcal{M}_{(i-2)q})\|_{\mathbb{H}}+2\sum_{k=1}^{n}\mathbb{E}\varphi_M(\|X_k\|_{\mathbb{H}})\Big)$$

Since for every $i\geq 1$, $U_i'=U_i-U_i''$, we get that

$$\mathbb{E}\|\mathbb{E}(U_i'|\mathcal{M}_{(i-2)q})\|_{\mathbb{H}}\leq\mathbb{E}\|\mathbb{E}(U_i|\mathcal{M}_{(i-2)q})\|_{\mathbb{H}}+\mathbb{E}\|U_i''\|_{\mathbb{H}}\,,$$

which implies using Inequality (3.5) that

$$\sum_{i=1}^{[n/q]}\mathbb{E}\|\mathbb{E}(U_i'|\mathcal{M}_{(i-2)q})\|_{\mathbb{H}}\leq\sum_{i=1}^{[n/q]}\mathbb{E}\|\mathbb{E}(U_i|\mathcal{M}_{(i-2)q})\|_{\mathbb{H}}+2\sum_{k=1}^{n}\mathbb{E}\varphi_M(\|X_k\|_{\mathbb{H}})\,,$$

Now

$$\mathbb{E}\|\mathbb{E}(U_i|\mathcal{M}_{(i-2)q})\|_{\mathbb{H}}\leq\sum_{j=(i-1)q+1}^{iq}\mathbb{E}\|\mathbb{E}(X_j|\mathcal{M}_{(i-2)q})\|_{\mathbb{H}}$$

Next we write that

$$\begin{aligned}\mathbb{E}\|\mathbb{E}(X_j|\mathcal{M}_{(i-2)q})\|_{\mathbb{H}} &= \mathbb{E}<\mathbb{E}(X_j|\mathcal{M}_{(i-2)q}),\frac{\mathbb{E}(X_j|\mathcal{M}_{(i-2)q})}{\|\mathbb{E}(X_j|\mathcal{M}_{(i-2)q})\|_{\mathbb{H}}}>_{\mathbb{H}}\\ &= \mathbb{E}<X_j,\frac{\mathbb{E}(X_j|\mathcal{M}_{(i-2)q})}{\|\mathbb{E}(X_j|\mathcal{M}_{(i-2)q})\|_{\mathbb{H}}}>_{\mathbb{H}}\ . \qquad (3.14)\end{aligned}$$

(Interpret 0/0 to be 0.)
Inequality (3.14) together with the covariance inequality stated in Lemma 2 in Merlevède, Peligrad and Utev (1997), then entails that for all $i\geq 1$,

$$\begin{aligned}\mathbb{E}\|\mathbb{E}(U_i|\mathcal{M}_{(i-2)q})\|_{\mathbb{H}} &\leq 18\sum_{j=(i-1)q+1}^{iq}\int_0^{\alpha(j-(i-2)q)}Q_X(u)du\\ &\leq 18q\int_0^{\alpha(q)}Q_X(u)du\,. \qquad (3.15)\end{aligned}$$

It follows that

$$\begin{aligned}\mathbb{P}\Big(\sum_{i=3}^{[n/q]}\|U_i'-\tilde{U}_i\|_{\mathbb{H}}+2\sum_{k=1}^{n}\varphi_M(\|X_k\|_{\mathbb{H}})\geq x\Big) &\leq \frac{18n}{x}\int_0^{\alpha(q)}Q_X(u)du \qquad (3.16)\\ &\quad+\frac{4}{x}\sum_{k=1}^{n}\mathbb{E}\varphi_M(\|X_k\|_{\mathbb{H}})\,.\end{aligned}$$

Then starting from (3.6), if q and M are chosen in such a way that $qM \le x$, we derive from (3.12) and (3.16) that

$$\mathbb{P}\Big(\max_{1\le k\le n} \|S_k\|_{\mathbb{H}} \ge 4x\Big) \le 4\exp\left(-\frac{s_n^2}{8(qM)^2}h\left(\frac{2xqM}{s_n^2}\right)\right) + \frac{32nqM}{s_n^2}M\alpha(q) + \frac{18n}{x}\int_0^{\alpha(q)} Q_X(u)du + \frac{4}{x}\sum_{k=1}^n \mathbb{E}\varphi_M(\|X_k\|_{\mathbb{H}})\,. \quad (3.17)$$

Now choose $v = H_X(x/r)$, $q = \alpha^{-1}(v)$ and $M = Q_X(v)$. Then we have that

$$qM = R_X(v) = R_X(H_X(x/r)) \le x/r \le x\,.$$

Note also that $Q_{\varphi_M(\|X_k\|_{\mathbb{H}})} = \sup(Q_{\|X_k\|_{\mathbb{H}}} - M, 0)$. Then we get that

$$\sum_{k=1}^n \mathbb{E}\varphi_M(\|X_k\|_{\mathbb{H}}) = \sum_{k=1}^n \int_0^1 (Q_{\|X_k\|_{\mathbb{H}}}(u) - Q_X(v))_+ du \le n\int_0^v (Q_X(u) - Q_X(v))du\,. \quad (3.18)$$

In addition, the choice of q implies that $\alpha(q) \le v$ and $M\alpha(q) \le vQ_X(v)$. Hence the inequality is obtained by taking into account (3.18) in (3.17) with the fact that $qM \le x/r$. ■

3.2 Proof of Theorem 2

The theorem will follow from Item II of Theorem 3.1 in Kuelbs (1976) if we can prove that

1) there exists a mean zero gaussian measure ν with covariance function Λ,

2) $\mathbb{P}\Big(w : \limsup \frac{<S_n,x>_{\mathbb{H}}}{\sqrt{2nLLn}} = \sup_{y\in K} <x,y>_{\mathbb{H}}\Big) = 1$ for any $x \in \mathbb{H}$

and that

3) the sequence $\{S_n/\sqrt{nLLn}\,,\, n\ge 1\}$ is almost surely relatively compact in $\mathbb{H}$.

In what follows we prove these three points.

1) The central limit theorem in $\mathbb{H}$. According to Theorem 4 in Merlevède, Peligrad and Utev (1997), the condition (2.3) implies that the sequence $n^{-1/2}S_n$ converges in distribution to $\mathcal{N}(0,\Lambda)$ where the operator $\Lambda \in \mathcal{S}(\mathbb{H})$ is defined by (2.4).

2) The LIL for the finite-dimensional laws. Since for any $x \in \mathbb{H}$, $| <X_0, x>_{\mathbb{H}} | \le \|x\|_{\mathbb{H}}\|X_0\|_{\mathbb{H}}$, the condition (1.3) in Rio (1995) is satisfied under (2.3). Hence for any $x \in \mathbb{H}$, Theorem 2 in Rio (1995) entails that

$$\lim n^{-1}\mathrm{Var}(<S_n, x>_{\mathbb{H}}) = \sigma_x^2 \in [0,\infty[\quad (3.19)$$

and

$$\mathbb{P}\Big(w : \limsup \frac{< S_n, x >_{\mathbb{H}}}{\sqrt{2nLLn}} = \sigma_x\Big) = 1\,. \tag{3.20}$$

Now from the central limit theorem in $\mathbb{H}$, we get that necessarily

$$\sigma_x = \sqrt{\Lambda(x,x)} = \sup_{y \in K} < x, y >_{\mathbb{H}}\ . \tag{3.21}$$

3) The sequence $\{S_n/\sqrt{nLLn}\,,\, n \geq 1\}$ *is almost surely relatively compact in* $\mathbb{H}$. To prove this step, we first prove that under the conditions of Theorem 2, one has that

$$\limsup_{n\to\infty} \frac{\|S_n\|_{\mathbb{H}}}{\sqrt{2nLLn}} \leq 8\sqrt{V} \text{ a.s.}\,, \tag{3.22}$$

where $V = \mathbb{E}\|X_0\|_{\mathbb{H}}^2 + 2\sum_{i\geq 1} |\mathbb{E} < X_0, X_i >_{\mathbb{H}}|$.

Notice that if V equals to zero then necessarily $\mathbb{E}\|X_0\|_{\mathbb{H}}^2 = 0$ implying that for all i, $X_i = 0_{\mathbb{H}}$ almost surely (since the variables are centered). Hence in this case (3.22) is trivial. Hence, we assume now that $V > 0$.

To prove (3.22), it suffices to show that

$$\sum_{n\geq 1} n^{-1}\mathbb{P}\left(\sup_{1\leq k\leq n} \|S_k\|_{\mathbb{H}} \geq 8\sqrt{2VnLLn}\right) < \infty\,, \tag{3.23}$$

and after to apply Borel-Cantelli Lemma (see for instance Stout (1974), Chapter 5). With this aim, we proceed as in Rio (2000, page 89).

Notice first that by the covariance inequality stated in Lemma 2 in Merlevède, Peligrad and Utev (1997), we get for any $i \geq 0$, that $|\mathbb{E} < X_0, X_i >_{\mathbb{H}}| \leq 18 \int_0^{\alpha(i)} Q^2_{\|X_0\|_{\mathbb{H}}}(u)du$. Hence, if we set $s_n^2 = \sum_{i=1}^n \sum_{j=1}^n |\mathbb{E} < X_i, X_j >_{\mathbb{H}}|$, under Condition (2.3) and by stationarity, we get that

$$\lim_{n\to\infty} n^{-1}s_n^2 = \mathbb{E}\|X_0\|_{\mathbb{H}}^2 + 2\sum_{i\geq 1} |\mathbb{E} < X_0, X_i >_{\mathbb{H}}| = V > 0\,. \tag{3.24}$$

To show (3.23), we shall apply Theorem 1 with $r = 8LLn$ and $x_n = 2\sqrt{2VnLLn}$. Set $y_n = x_n/r = \frac{\sqrt{2V}}{4}\sqrt{n}(LLn)^{-1/2}$ and notice that $Vn = x_n^2/r$. Applying Inequality (2.2) and setting $Q(u) = Q_{\|X_0\|_{\mathbb{H}}}(u)$, $R(u) = R_{\|X_0\|_{\mathbb{H}}}(u)$ and $H(u) = H_{\|X_0\|_{\mathbb{H}}}(u)$, one gets that

$$\begin{aligned}\sum_{n\geq 1} n^{-1}\mathbb{P}(&\sup_{1\leq k\leq n} \|S_k\|_{\mathbb{H}} \geq 8\sqrt{2VnLLn}) \\ &\leq 4\sum_{n\geq 1} n^{-1}\big(1 + 2V\frac{n}{s_n^2}\big)^{-LLn} + \sum_{n\geq 1}\frac{32}{x_n}\big(1 + V\frac{n}{s_n^2}\big)\int_0^{H(y_n)} Q(u)du\,. \end{aligned} \tag{3.25}$$

By taking into account (3.24), the first series in the right-hand side is obviously convergent. To study the second one, we use (3.24) and we notice that $u < H(y_n) \iff y_n < R(u)$, then

$$\sum_{n\geq 1} \frac{1}{x_n} \int_0^1 Q(u)\mathbb{I}(u < H(y_n))du = V^{-1} \int_0^1 Q(u)\Big(\sum_{n\geq 1} \frac{y_n}{n} \mathbb{I}(y_n < R(u))\Big)du \,.$$

Since the general term of the above series is equivalent to $(8nVLLn)^{-1/2}$, there exists a constant $C > 0$ such that: $\sum_{n\geq 1} \frac{y_n}{s_n^2} \mathbb{I}(y_n < R(u)) \leq CR(u)$. It follows that

$$\int_0^1 Q(u)\Big(\sum_{n\geq 1} \frac{y_n}{s_n^2} \mathbb{I}(y_n < R(u))\Big)du \leq C\int_0^1 R(u)Q(u)du = C\sum_{k\geq 1} \int_0^{\alpha(k)} Q^2(u)du \,.$$

which ends the proof of (3.23), and then of (3.22).

We turn now to prove that the sequence $\{S_n/\sqrt{nLLn}\,,\, n \geq 1\}$ is almost surely relatively compact in $\mathbb{H}$. With this aim, we argue as page 698 in Dehling and Philipp (1982, proof of their theorem 1), with the help of (3.22). Let $\{e_i, i \geq 1\}$ be a complete orthonormal basis for $\mathbb{H}$. We write for each $k \in \mathbf{Z}$

$$X_k = \sum_{i\geq 1} < X_k, e_i >_{\mathbb{H}} e_i \text{ and } P_N(X_k) = \sum_{i=1}^{N} < X_k, e_i >_{\mathbb{H}} e_i \,.$$

Applying (3.22) to the sequence $\{X_k - P_N(X_k), k \in \mathbf{Z}\}$, we get that with probability one

$$\limsup_{n\to\infty} \frac{\| \sum_{k=1}^n (X_k - P_N(X_k))\|_{\mathbb{H}}}{\sqrt{2nLLn}} \leq 8\sqrt{V_N}\,,$$

where $V_N = \mathbb{E}\|X_0 - P_N(X_0)\|_{\mathbb{H}}^2 + 2\sum_{i\geq 1} |\mathbb{E} < X_0 - P_N(X_0), X_i - P_N(X_i) >_{\mathbb{H}} |$. But by using Lemma 2 in Merlevède, Peligrad and Utev (1997), we have that

$$\sum_{i\geq 0} |\mathbb{E} < X_0 - P_N(X_0), X_i - P_N(X_i) >_{\mathbb{H}} | \leq 18\sum_{i\geq 0} \int_0^{\alpha(i)} Q^2_{\|X_0 - P_N(X_0)\|_{\mathbb{H}}}(u)du \,.$$

Noticing that for any $u \in [0, 1]$, $Q^2_{\|X_0 - P_N(X_0)\|_{\mathbb{H}}}(u) \leq Q^2_{\|X_0\|_{\mathbb{H}}}(u)$ and that $\lim_{N\to\infty} Q^2_{\|X_0 - P_N(X_0)\|_{\mathbb{H}}} = 0$, we get by using (2.3) and the Lebesgue dominated Theorem that for each $\rho > 0$ there is an $N_0(\rho)$ such that for all $N \geq N_0(\rho)$,

$$V_N \leq \rho \,.$$

Hence with probability one,

$$\limsup_{n\to\infty} \frac{\| \sum_{k=1}^n (X_k - P_N(X_k))\|_{\mathbb{H}}}{\sqrt{2nLLn}} \leq 8\sqrt{\rho}\,. \tag{3.26}$$

Now applying again (3.22), we get that the sequence $\Big\{\frac{\sum_{k=1}^n P_{N_0(\rho)}(X_k)}{\sqrt{2nLLn}}, n \geq 1\Big\}$ is with probability one relatively compact. This fact combining with (3.26) establishes the almost sure relative compactness of $\{S_n/\sqrt{nLLn}\,,\, n \geq 1\}$.

4 Appendix

In this section, we state a modified version of a result given in Pinelis (1994, Theorem 3.4).

Lemma 1. *Let* $(\mathbb{H}, \|\cdot\|_{\mathbb{H}})$ *be a real separable separable Hilbert space. Let* $\{d_j, \mathcal{F}_j\}_{j\geq 1}$ *be a sequence of* $\mathbb{H}$*-valued martingale differences with* $\|d_j\|_{\mathbb{H}} \leq c$*. Set* $M_j = \sum_{i=1}^{j} d_i$*. Then for all* $x, y > 0$,

$$\mathbb{P}\left(\sup_{1\leq j\leq n} \|M_j\|_{\mathbb{H}} \geq x, \sum_{j=1}^{n} \mathbb{E}(\|d_j\|^2_{\mathbb{H}}|\mathcal{F}_{j-1}) \leq y\right) \leq 2\exp\left(-\frac{y}{c^2}\, h\Big(\frac{xc}{y}\Big)\right),$$

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

Proof of Lemma 1. First, given $\lambda > 0$, we set for all $j \geq 1$,

$$e_j = \mathbb{E}(e^{\lambda\|d_j\|_{\mathbb{H}}} - 1 - \lambda\|d_j\|_{\mathbb{H}}|\mathcal{F}_{j-1}),\ G_0 = 1 \text{ and } G_j = \cosh(\lambda\|M_j\|_{\mathbb{H}})/\Pi_{i=1}^{j}(1+e_i)\,.$$

Now using the fact that for all $j \in [1, n]$, $\Pi_{i=1}^{j}(1+e_i) \leq \Pi_{i=1}^{n}(1+e_i)$, we derive that

$$\mathbb{P}\left(\sup_{1\leq j\leq n} \|M_j\|_{\mathbb{H}} \geq x, \sum_{i=1}^{n} \mathbb{E}(\|d_j\|^2_{\mathbb{H}}|\mathcal{F}_{j-1}) \leq y\right)$$
$$\leq \mathbb{P}\left(\sup_{1\leq j\leq n} G_j \geq \cosh(\lambda x)/\Pi_{i=1}^{n}(1+e_i), \sum_{i=1}^{n} \mathbb{E}(\|d_j\|^2_{\mathbb{H}}|\mathcal{F}_{j-1}) \leq y\right) \tag{4.1}$$

Now because the function $g(u) := u^{-2}(e^u - 1 - u)$ for $u \neq 0$, $g(0) := \frac{1}{2}$ is increasing in $u \in \mathbb{R}$, and since $\|d_j\|_{\mathbb{H}} \leq c$, we easily infer that

$$\Pi_{i=1}^{n}(1+e_i) \leq \exp\left(\frac{e^{\lambda c} - 1 - \lambda c}{c^2} \sum_{i=1}^{n} \mathbb{E}(\|d_j\|^2_{\mathbb{H}}|\mathcal{F}_{j-1})\right). \tag{4.2}$$

Starting from (4.1) and using (4.2), we get that

$$\mathbb{P}\left(\sup_{1\leq j\leq n} \|M_j\|_{\mathbb{H}} \geq x, \sum_{i=1}^{n} \mathbb{E}(\|d_j\|^2_{\mathbb{H}}|\mathcal{F}_{j-1}) \leq y\right)$$
$$\leq \mathbb{P}\left(\sup_{1\leq j\leq n} G_j \geq \cosh(\lambda x)\exp(-y \times \frac{e^{\lambda c} - 1 - \lambda c}{c^2})\right). \tag{4.3}$$

But since $\{G_i, \mathcal{F}_i\}$ is a positive supermartingale, it follows from (4.3) and the maximum inequality (see Stout (1974), page 299) that

$$\mathbb{P}\left(\sup_{1\leq j\leq n} \|M_j\|_{\mathbb{H}} \geq x, \sum_{i=1}^{n} \mathbb{E}(\|d_j\|^2_{\mathbb{H}}|\mathcal{F}_{j-1}) \leq y\right) \leq \frac{1}{\cosh(\lambda x)} \exp(y \times \frac{e^{\lambda c} - 1 - \lambda c}{c^2}).$$

Now the result follows from the fact that $\cosh u > e^u/2$ and the minimization in λ.

Acknowledgments. I am indebted to J. Dedecker for the interesting discussions allowing some improvements in this paper.

References

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

[2] Dedecker, J. and Merlevède, F. (2006). Inequalities for partial sums of Hilbert valued dependent sequences and applications, *Mathematical methods of Statistics* **15** 176-206.

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

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

[5] 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.

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

[7] Merlevède, F., Peligrad, M. and Utev, S. (1997). Sharp conditions for the CLT of linear processes in a Hilbert space. *J. of Theor. Probab.*, **10** 681-693.

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

[9] Rio, E. (1993). Covariance inequalities for strongly mixing processes. *Ann. Inst. H. Poincaré Probab. Statist.*, **29** 587-597.

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

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

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

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

[14] Stout, W. F. (1974). Almost sure convergence. *Academic Press, New-York.*