The Annals of Probability
1982, Vol. 10, No. 3, 689–701

ALMOST SURE INVARIANCE PRINCIPLES FOR WEAKLY DEPENDENT VECTOR-VALUED RANDOM VARIABLES

BY HEROLD DEHLING[1] AND WALTER PHILIPP[2]

University of Illinois, Urbana

We obtain the almost sure approximation of the partial sums of random variables with values in a separable Hilbert space and satisfying a strong mixing condition by a suitable Brownian motion. This is achieved by a modification of the proof of a similar result by Kuelbs and Philipp (1980) on ϕ-mixing Banach space valued random variables. As by-products we get almost sure invariance principles for sums of absolutely regular sequences of random variables with values in a Banach space and necessary and sufficient conditions for the almost sure invariance principle for sums of independent, identically distributed random variables.

1. Introduction. Let $\{x_\nu, \nu \geq 1\}$ be a sequence of random variables with values in a real separable Banach space. Let $\mathscr{M}_a^b$ denote the σ-field generated by the random variables $x_a, x_{a+1}, \cdots, x_b$. The sequence $\{x_\nu, \nu \geq 1\}$ is said to satisfy a strong mixing condition if there exists a sequence of real numbers $\rho(n) \downarrow 0$ such that

$$|P(A \cap B) - P(A)P(B)| \leq \rho(n) \tag{1.1}$$

for all $A \in \mathscr{M}_1^k$ and $B \in \mathscr{M}_{k+n}^\infty$ and all $k, n \geq 1$. A sequence $\{x_\nu, \nu \geq 1\}$ is called absolutely regular if for some $\beta(n) \downarrow 0$

$$E\{\sup_{B \in \mathscr{M}_{k+n}^\infty} |P(B | \mathscr{M}_1^k) - P(B)|\} \leq \beta(n) \tag{1.2}$$

for all $k, n \geq 1$. (The supremum in (1.2) is measurable since in a Polish space it is sufficient to extend the supremum only over countably many sets B. This observation is also useful for the following two remarks.) Finally, the sequence $\{x_\nu, \nu \geq 1\}$ is called ϕ-mixing if for some $\phi(n) \downarrow 0$

$$|P(A \cap B) - P(A)P(B)| \leq \phi(n)P(A) \tag{1.3}$$

for all $A \in \mathscr{M}_1^k$ and $B \in \mathscr{M}_{k+n}^\infty$ and $k, n \geq 1$. Since (1.3) is equivalent with

$$|P(B | \mathscr{M}_1^k) - P(B)| \leq \phi(n) \tag{1.3'}$$

with probability 1 for all $B \in \mathscr{M}_{k+n}^\infty$ and all $k, n \geq 1$ we note that every ϕ-mixing sequence is absolutely regular. We also note that (1.1) is equivalent with

$$E|P(B | \mathscr{M}_1^k) - P(B)| \leq \rho'(n) \tag{1.1'}$$

for all $B \in \mathscr{M}_{k+n}^\infty$ and all $k, n \geq 1$; here $\rho(n) \leq \rho'(n) \leq 2\rho(n)$. In other words a strongly mixing sequence is ϕ-mixing in L^1. Consequently every absolutely regular sequence is strongly mixing.

On the other hand, by comparing the spectral densities of weakly dependent stationary Gaussian sequences it is clear that there are examples of strongly mixing sequences which are not absolutely regular, and of absolutely regular sequences which are not ϕ-mixing. (See Ibragimov and Rozanov (1978), Chapters 4 and 5.)

Received April 1980; revised July 1981.
[1] Supported by the Studienstiftung des Deutschen Volkes.
[2] Supported in part by a grant from the National Science Foundation.
AMS 1980 *subject classification*. Primary, 60B12.
Key words and phrases. Almost sure invariances principles, mixing and absolutely regular sequences of random variables, Hilbert space, Banach space, Brownian motion.

In a recent paper [7] Kuelbs and Philipp established several almost sure invariance principles for sums of ϕ-mixing random variables with values in a separable Banach space. However, in the strongly mixing case only results for $\mathbb{R}^d$-valued random variables were obtained. In order to facilitate subsequent "research" the argument in [7] was set up in such a way that once the proper generalization was obtained for Theorem 2 of [1] Berkes and Philipp to strongly mixing sequences of random variables with values in a Polish space, the generalization of most of the results of [7] would become rather trivial.

Unfortunately, this hoped-for generalization of Theorem 2 in [1] cannot hold. As a matter of fact, H. Dehling (1982) recently 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 prove almost sure invariance principles for sums of strongly mixing random variables with values in an infinite-dimensional space, some new ideas will be needed.

As indicated in the abstract, the purpose of this paper is threefold. Our foremost goal is to prove the following result.

THEOREM 1. *Let $\{x_\nu, \nu \geq 1\}$ be a strictly stationary sequence of random variables with values in a real separable Hilbert space H, centered at expectations and with finite* $(2+\delta)$th *moments where* $0 < \delta \leq 1$. *Suppose that the sequence satisfies a strong mixing condition* (1.1) *with a mixing rate*

$$\rho(n) \ll n^{-(1+\varepsilon)(1+2/\delta)} \tag{1.4}$$

for some $0 < \varepsilon \leq 1$. *Then the two series defining the covariance function T of the sequence $\{x_\nu, \nu \geq 1\}$, defined as*

$$(Tx, y) = E\{(x, x_1)(y, x_1)\} + \textstyle\sum_{\nu\geq 2} E\{(x, x_1)(y, x_\nu)\} + \sum_{\nu\geq 2} E\{(x, x_\nu)(y, x_1)\}$$

converge absolutely for all $x, y \in H$. Moreover, without changing its distribution we can redefine the sequence $\{x_\nu, \nu \geq 1\}$ on a new probability space on which there exists a Brownian motion $\{X(t), t \geq 0\}$ with covariance structure given by T such that with probability 1

$$\| \textstyle\sum_{\nu\leq t} x_\nu - X(t) \| = o((t \log\log t)^{1/2}).$$

NOTE. Proposition 4.2 in [7] implies the central limit theorem under the hypotheses of Theorem 1.

The proofs of most of the results in [7] were based on Theorem 6 of [7]. However, for the proof of Theorem 1, instead of Theorem 6 of [7], we have to use the following stronger variant which, at the same time, is also much simpler. Let B be a separable Banach space.

THEOREM 2. *Let $\{x_\nu, \nu \geq 1\}$ be a weak sense stationary sequence of B-valued random variables centered at expectations and with* $(2+\delta)$th *moments with* $0 < \delta \leq 1$ *uniformly bounded by* 1. *Suppose that $\{x_\nu, \nu \geq 1\}$ satisfies a strong mixing condition* (1.1) *with mixing rate given by* (1.4). *Then the series defining the covariance function*

$$T(f, g) = E\{f(x_1)g(x_1)\} + \textstyle\sum_{\nu\geq 2} E\{f(x_1)g(x_\nu)\} + \sum_{\nu\geq 2} E\{f(x_\nu)g(x_1)\}, \quad f, g \in B^*$$

converge absolutely. Moreover, suppose that

$$E \| \textstyle\sum_{\nu=a+1}^{a+n} x_\nu \|^2 \leq \sigma^2 n \tag{1.5}$$

for some σ and all $a \geq 0$, $n \geq 1$. Let $S_n = \sum_{\nu\leq n} x_\nu$. Then the following two conditions are equivalent.

(i) *There exists a mean zero Gaussian measure μ with covariance function T and $\{(n \log\log n)^{-1/2} S_n, n \geq 1\}$ is with probability* 1 *relatively compact.*

(ii) *Without changing its distribution, we can redefine the sequence* $\{x_\nu, \nu \geq 1\}$ *on a new probability space on which there exists a Brownian motion* $\{X(t), t \geq 0\}$ *with covariance function* T *such that with probability* 1

$$\| \textstyle\sum_{\nu \leq t} x_\nu - X(t) \| = o((t \log \log t)^{1/2}). \tag{1.6}$$

Theorem 2 is proved in Section 4 by a modification of the proof of Theorem 6 in [7]. Using the same kind of ideas, we modify in Section 6 the proof of Corollary 3 in [9] to get necessary and sufficient conditions for the almost sure invariance principles for independent identically distributed random variables.

THEOREM 3. *Let* $\{x_\nu, \nu \geq 1\}$ *be a sequence of independent identically distributed B-valued random variables centered at expectations. Then the following two statements are equivalent.*

(i) x_1 *is pregaussian and* $\{(n \log \log n)^{-1/2} S_n, n \geq 1\}$ *is with probability* 1 *relatively compact.*

(ii) *Without changing its distribution we can redefine* $\{x_\nu, \nu \geq 1\}$ *on a new probability space on which there exists a Brownian motion (which necessarily has the same covariance function as* x_1*) such that with probability* 1, *relation* (1.6) *holds.*

In view of Theorem 1.1 of Pisier (1975), the following corollary is immediate.

COROLLARY. *Let* $\{x_\nu, \nu \geq 1\}$ *be a sequence of independent, identically distributed random variables. Then* $\{x_\nu, \nu \geq 1\}$ *satisfies an almost sure invariance principle* (1.6) *if and only if it satisfies the compact law of the iterated logarithm and* x_1 *is pregaussian.*

The interesting feature in Theorem 3 and its corollary is that explicitly no assumptions on the finiteness of moments of order higher than one have to be made. On the other hand, the compact law of the iterated logarithm, the relative compactness of $\{(n \log \log n)^{-1/2} S_n, n \geq 1\}$ and the condition that x_ι be pregaussian all implicitly contain assumptions on these moments.

Our third goal, finally, is to generalize most of the results of [7] to absolutely regular equences of random variables. We list these results as well as Theorem 5 below, on which their proof rests, only for the sake of completeness and easy reference. The proofs for which no new ideas are needed will be given in Sections 8 and 7 respectively.

THEOREM 4. *Let* $\{x_\nu, \nu \geq 1\}$ *be a weak sense stationary sequence of B-valued random variables with uniformly bounded* $(2 + \delta)$th *moments. We assume* $0 < \delta \leq 1/5$. *Suppose that the sequence satisfies the absolute regularity condition* (1.2) *with rate* $\beta(n)$ *bounded by*

$$\beta(n) \ll n^{-(1+\varepsilon)(1+2/\delta)}$$

for some $\varepsilon > 0$. *Then Theorems* 1 *and* 2 *and Corollaries* 1 *and* 2 *of* [7] *all remain valid.*

Note that for small δ the rate of decay for $\beta(n)$ can now be even somewhat slower than in [7].

THEOREM 5. *Let* $\{B_k, m_k, k \geq 1\}$ *be a sequence of Polish spaces. Let* $\mathscr{B}_k$ *denote the Borel field of* B_k, *let* $\{X_k, k \geq 1\}$ *be a sequence of random variables with values in* B_k *and let* $\mathscr{F}_k$ *be a non-decreasing sequence of σ-fields such that* X_k *is* $\mathscr{F}_k$*-measurable. Suppose that for some sequence* $\{\beta_k, k \geq 1\}$ *of non-negative numbers*

$$E \sup_{A \in \mathscr{B}_k} | P(X_k \in A \mid \mathscr{F}_{k-1}) - P(X_k \in A) | \leq \beta_k \tag{1.7}$$

for all $k \geq 1$. *Denote by* F_k *the distribution of* X_k *and let* $\{G_k, k \geq 1\}$ *be a sequence of* *distributions on B_k such that*

$$F_k(A) \le G_k(A^{\rho_k}) + \sigma_k \quad \textit{for all} \quad A \in \mathscr{B}_k. \tag{1.8}$$

Here ρ_k and σ_k are non-negative numbers and $A^\varepsilon = \bigcup_{x\in A}\{y : m_k(x, y) < \varepsilon\}$. Then without changing its distribution we can redefine the sequence $\{X_k, k \ge 1\}$ on a richer probability space on which there exists a sequence $\{Y_k, k \ge 1\}$ of independent random variables Y_k with distribution G_k such that for all $k \ge 1$

$$P\{m_k(X_k, Y_k) \ge 2(\beta_k^{1/2} + \rho_k)\} \le 2(\beta_k^{1/2} + \sigma_k). \tag{1.9}$$

In general, if (1.7) *is replaced by*

$$E^{1/p}\{\cdots\}^p \le \beta_k, \quad 1 \le p \le \infty \tag{1.7*}$$

then (1.9) *is to be replaced by*

$$P\{m_k(X_k, Y_k) \ge 2(\beta_k^{p/(p+1)} + \rho_k)\} \le 2(\beta_k^{p/(p+1)} + \sigma_k). \tag{1.9*}$$

REMARK. For $p = \infty$ we obtain Theorem 3 in [9] as a special case.

Finally, a comment on the proofs of Theorems 1 and 4, which might be useful in connection with proofs of almost sure invariance principles in similar circumstances. As is evident from Sections 5 and 8, in view of Theorem 2 the proofs of Theorems 1 and 4 can in effect be reduced to the proof of a bounded law of the iterated logarithm. For sequences of Hilbert-space valued random variables satisfying the hypotheses of Theorem 1, such a bounded law of the iterated logarithm will be established in Section 3.

2. Lemmas on mixing random variables.

LEMMA 2.1. *Let $\{\xi_\nu, \nu \ge 1\}$ be a sequence of random variables with values in a separable Banach space satisfying a strong mixing condition* (1.1) *with mixing rate* (1.4). *Suppose that their $(2 + \delta)$th moments are uniformly bounded by M, where $0 < \delta \le 1$ and that* (1.5) *holds. Then for all $a \ge 0$ and all $0 \le \alpha \le \varepsilon\delta/8$*

$$E\,\Big\| \sum_{\nu=a+1}^{a+n} \xi_\nu \Big\|^{2+\alpha} \ll n^{1+\alpha/2}(\sigma^{2+\alpha} + M)$$

where the constant implied by $\ll$ only depends on ε, δ and the constant implied by $\ll$ in (1.4).

The proof as given in Sotres and Malay Ghosh (1977) still works for B-valued random variables. The bound for α can be obtained by a careful analysis of their proof and of Serfling's (1968) paper on which their proof rests.

LEMMA 2.2. *Let $\mathscr{F}$ and $\mathscr{G}$ be two σ-fields. Define*

$$\rho(\mathscr{F}, \mathscr{G}) = \sup |P(A \cap B) - P(A)P(B)|$$

the supremum being extended over all $A \in \mathscr{F}$ and $B \in \mathscr{G}$. Let ξ and η be random variables with values in a separable Hilbert space H measurable $\mathscr{F}$ and $\mathscr{G}$ respectively. If ξ and η are essentially bounded then

$$|E(\xi, \eta) - (E\xi, E\eta)| \le 10\rho(\mathscr{F}, \mathscr{G})\,\|\xi\|_\infty \|\eta\|_\infty. \tag{2.1}$$

Here $\|\cdot\|_\infty$ denotes the essential supremum with respect to H. Moreover, let $r, s, t > 1$ with $r^{-1} + s^{-1} + t^{-1} = 1$. If ξ and η have finite rth and sth moments respectively then

$$|E(\xi, \eta) - (E\xi, E\eta)| \le 15\rho^{1/t}(\mathscr{F}, \mathscr{G})\,\|\xi\|_r \|\eta\|_s. \tag{2.2}$$

PROOF. For real-valued random variables, (2.1) with 10 replaced by 4 is due to Volkonskii and Rozanov (1959), (2.2) is due to Davydov (1970). The proof as given in Deo (1973) shows that (2.1) implies (2.2). It still works for H-valued random variables. Hence it remains to show (2.1).

The proof of (2.1) combines the fact that (2.1) holds for real-valued random variables and the following theorem which is due to Grothendieck. For its proof see Lindenstrauss, Tzafriri (1977), page 68.

THEOREM. *Let* $(\alpha_{ij}, 1 \le i, j \le n)$ *be a matrix of real numbers such that*

$$|\textstyle\sum_{1\le i,j\le n} \alpha_{ij} s_i t_j| \le 1 \tag{2.3}$$

for all real numbers s_i, t_j *with* $|s_i| \le 1$, $|t_j| \le 1$ $(1 \le i, j \le n)$. *Then for all vectors* x_i, y_j $(1 \le i, j \le n)$ *in a Hilbert space,*

$$|\textstyle\sum_{1\le i,j\le n} \alpha_{ij}(x_i, y_j)| \le K_G \max_{1\le i\le n} \|x_i\| \cdot \max_{1\le j\le n} \|y_j\|. \tag{2.4}$$

Here $K_G \le \frac{1}{2}(\exp(\frac{1}{2}\pi) - \exp(-\frac{1}{2}\pi)) < 2.5$.

For the proof of (2.1) we assume without loss of generality that $\xi = \sum_{i\le n} x_i 1_{A_i}$ and $\eta = \sum_{j\le n} y_j 1_{B_j}$ where $\{A_i, 1 \le i \le n\}$ and $\{B_j, 1 \le j \le n\}$ are partitions of the sample space and where $A_i \in \mathscr{F}$ and $B_j \in \mathscr{G}$ $(1 \le i, j \le n)$. Put $\rho = \rho(\mathscr{F}, \mathscr{G})$ and

$$\alpha_{ij} = \frac{1}{4\rho}(P(A_i \cap B_j) - P(A_i)P(B_j)). \tag{2.5}$$

To verify (2.3) let s_i, t_j be real numbers with $|s_i| \le 1$ and $|t_j| \le 1$ and define the random variables $X = \sum_{i\le n} s_i 1_{A_i}$ and $Y = \sum_{j\le n} t_j 1_{B_j}$. Then

$$|\textstyle\sum \alpha_{ij} s_i t_j| = \frac{1}{4\rho}|\sum s_i t_j(P(A_i \cap B_j) - P(A_i)P(B_j))| = \frac{1}{4\rho}|EXY - EXEY| \le 1$$

since, as was observed above, (2.1) with 10 replaced by 4 holds for real-valued random variables. Hence by (2.4) and (2.5)

$$\begin{aligned} |E(\xi, \eta) - (E\xi, E\eta)| &= |\textstyle\sum_{1\le i,j\le n} (x_i, y_j)(P(A_i \cap B_j) - P(A_i)P(B_j))| \\ &\le 4\rho K_G \max\|x_i\| \cdot \max\|y_j\| < 10\rho\|\xi\|_\infty \|\eta\|_\infty. \quad \square \end{aligned}$$

LEMMA 2.3. *Let* $\{x_\nu, \nu \ge 1\}$ *be a sequence of H-valued random variables, centered at expectations and with* $(2 + \delta)$th *moments uniformly bounded by M (say) and where* $0 < \delta \le 1$. *Suppose that* $\{x_\nu, \nu \ge 1\}$ *satisfies a strong mixing condition with mixing rate* (1.4). *The* (1.5) *holds with*

$$\sigma^2 \ll \varepsilon^{-1} M^{2/(2+\delta)}. \tag{2.6}$$

Here the constant implied by $\ll$ *only depends on the constant implied by* $\ll$ *in* (1.4).

For a proof see e.g. the proof of Lemma 2.3 in [7].

3. A bounded law of the iterated logarithm. In this section we prove the following theorem. Without loss of generality we assume that in (2.6), $\sigma = 1$.

THEOREM 6. *Let* $\{x_\nu, \nu \ge 1\}$ *be a sequence of H-valued random variables, centered at expectations, with* $(2 + \delta)$th *moments uniformly bounded where* $0 < \delta \le 1$. *Suppose that* $\{x_\nu, \nu \ge 1\}$ *satisfy a strong mixing condition with rate* (1.4). *Then with probability* 1

$$\limsup_{N\to\infty}(N \log\log N)^{-1/2} \|\textstyle\sum_{\nu\le N} x_\nu\| \le 200(\varepsilon\delta)^{-3/2}.$$

For the proof of Theorem 6 we define inductively blocks H_j, I_j, $j = 1, 2, \cdots$ of consecutive integers leaving no gaps between the blocks. Each block H_j and I_j consists of $[j^\beta]$ integers, $j = 1, 2, \cdots$ where

$$\beta = \frac{16}{\varepsilon\delta}. \tag{3.1}$$

The order is $H_1, I_1, H_2, I_2, \cdots$. We write

$$t_n = \textstyle\sum_{j\le n} \operatorname{card}(H_j \cup I_j). \tag{3.2}$$

Then

$$n^{\beta+1} \ll t_n \ll n^{\beta+1}. \tag{3.3}$$

LEMMA 3.1. *There is a constant* $\lambda > 0$ *such that as* $n \to \infty$

$$\max_{t_n < N \le t_{n+1}} \| \textstyle\sum_{\nu=t_n+1}^{N} x_\nu \| \ll t_n^{1/2-\lambda}, \quad \text{a.s.}$$

In view of Lemmas 2.1 and 2.3, we observe that the proof of Lemma 3.1 is the same as the proof of Proposition 2.2 in [7]. Note that the discussion preceding (2.25) in [7] becomes now redundant.

Next we define random variables y_k and z_k by

$$y_k = \textstyle\sum_{\nu\in H_k} x_\nu \quad \text{and} \quad z_k = \sum_{\nu\in I_k} x_\nu, \quad k = 1, 2, \cdots. \tag{3.4}$$

Because of Lemma 3.1 it is enough to show that with probability 1

$$\limsup_{n\to\infty} (t_n \log\log t_n)^{-1/2} \| \textstyle\sum_{j\le n} y_j \| \le 100(\varepsilon\delta)^{-3/2}. \tag{3.5}$$

We truncate y_j by setting

$$Y_j = y_j 1\{\| y_j \| \le j^{\beta+1}\} \tag{3.6}$$

and observe that by Lemma 2.3 and Chebyschev's inequality we have

$$P(Y_j \ne y_j) = P(\| y_j \| > j^{\beta+1}) \ll j^{-2}.$$

Hence by the Borel Cantelli Lemma we have $Y_j = y_j$ with probability 1 for all sufficiently large j. Consequently, and in view of (3.5) for the proof of Theorem 6, it is enough to prove the following proposition.

PROPOSITION 1. *As* $n \to \infty$ *we have with probability* 1

$$\limsup_{n\to\infty} (t_n \log\log t_n)^{-1/2} \| \textstyle\sum_{j\le n} Y_j \| \le 100(\varepsilon\delta)^{-3/2}.$$

In the proof of Proposition 1, we make heavy use of ideas of Goodman, Kuelbs and Zinn (1981). Writing

$$W_n = \textstyle\sum_{j\le n} Y_j \tag{3.7}$$

we obtain

$$\| W_n \|^2 = \textstyle\sum_{j\le n} \| Y_j \|^2 + 2 \sum_{j\le n} (Y_j, W_{j-1}). \tag{3.8}$$

In Lemma 3.2 we prove that the first sum is small. The estimate of the second sum depends on the observation that it is close to a real-valued martingale to which, when properly truncated, the standard exponential bounds apply. We now carry out this program in detail.

LEMMA 3.2. *As* $n \to \infty$ *we have with probability* 1

$$\textstyle\sum_{j\le n} \| Y_j \|^2 \ll n^{\beta+1}.$$

PROOF. Define $U_j = \| Y_j \|^2 - E\| Y_j \|^2$. Since by (3.6) and Lemma 2.3

(3.9) $$E\| Y_j \|^2 \ll j^{\beta}$$

we get

$$\sum_{j\le n} E\| Y_j \|^2 \ll n^{\beta+1}.$$

Hence it suffices to show that with probability 1

(3.10) $$\sum_{j\le n} U_j \ll n^{\beta+1}.$$

We first apply Theorem 1 of Berkes and Philipp (1979) to the sequence $\{j^{-\beta}U_j, j\ge 1\}$ and distributions $G_j = \mathscr{L}(j^{-\beta}U_j)$, $j\ge 1$ and the σ-fields $\mathscr{F}_j = \mathscr{M}_1^{t_j}$. By (3.17) below we have for all u

$$E\,|\,E\{\exp(iuU_j)\,|\,\mathscr{F}_{j-1}\} - E\{\exp(iuU_j)\}\,| \le 20\alpha(j^{\beta}).$$

Using Lemma 2.1 and Markov's inequality we get

$$G_j\{u: |u| \ge \tfrac14 T_j\} \ll T_j^{-1-\alpha/2}.$$

We choose $T_j = 10^8 \vee j^2$. Then by Theorem 1 in [1] we can find a sequence of independent random variables V_j having the same distribution as $j^{-\beta}U_j$ such that

$$P\{|\,j^{-\beta}U_j - V_j\,| \ge \alpha_j\} \le \alpha_j$$

where

$$\alpha_j \ll T_j^{-1}\log T_j + \alpha^{1/2}(j^{\beta})T_j + T_j^{-1-\alpha/2} \ll j^{-3/2}.$$

Hence

(3.11) $$\sum_{j\ge1} |\,j^{-\beta}U_j - V_j\,| < \infty, \quad \text{a.s.}$$

Since $E\,|\,V_j\,|^{1+\alpha/2} = j^{-(1+\alpha/2)}E\,|\,U_j\,|^{1+\alpha/2} \ll 1$ by Lemma 2.1, we obtain

(3.12) $$\sum_{j\ge1} j^{-1-\alpha/2}E\,|\,V_j\,|^{1+\alpha/2} < \infty.$$

Hence by a standard stability result (see e.g. Corollary 2.8.5 of Stout (1974), page 67) we get

$$\sum_{j\ge1} j^{-1}V_j < \infty, \quad \text{a.s.}$$

Hence by (3.11) and Kronecker's lemma

$$\sum_{j\le n} U_j = o(n^{\beta+1}), \quad \text{a.s.}$$

This yields (3.10). □

We now put

(3.13) $$Z_j = (Y_j, W_{j-1})1\{|(Y_j, W_{j-1})| \le j^{\beta+1}\}$$

and

(3.14) $$M_n = \sum_{j\le n}(Z_j - E(Z_j\,|\,\mathscr{F}_{j-1}))$$

where as before $\mathscr{F}_j = \mathscr{M}_1^{t_j}$. Then $\{M_n, \mathscr{F}_n, n\ge 1\}$ is a martingale. Moreover, we have the following lemma.

LEMMA 3.3. *As $n\to\infty$ we have with probability* 1

$$M_n - \sum_{j\le n}(Y_j, W_{j-1}) \ll n^{\beta+1}.$$

PROOF. By Kronecker's lemma, (3.13) and (3.14) it is enough to show that

$$\sum_{j\ge 1} j^{-\beta-1}E\{|(Y_j, W_{j-1})|1\{|(Y_j, W_{j-1})| > j^{\beta+1}\}\} < \infty \tag{3.15}$$

and

$$\sum_{j\ge 1} j^{-\beta-1}\, E\,|E\{(Y_j, W_{j-1})\,|\,\mathscr{F}_{j-1}\}\,| < \infty. \tag{3.16}$$

Applying Lemma 2.1 and (2.2) and (3.9) we see that a typical term of the series in (3.15) is bounded by

$$\begin{aligned} &j^{-(\beta+1)(2+\alpha)}E\{\|Y_j\|^{2+\alpha}\|W_{j-1}\|^{2+\alpha}\}\\ &\quad\ll j^{-(\beta+1)(2+\alpha}(E\|Y_j\|^{2+\alpha}E\|W_{j-1}\|^{2+\alpha} + \|Y_j\|_\infty^{2+\alpha}\|W_{j-1}\|_\infty^{2+\alpha}\rho(j^\beta))\\ &\quad\ll j^{-(\beta+1)(2+\alpha)}(j^{(2\beta+1)(1+\alpha/2)} + j^{(2\beta+3)(2+\alpha)}j^{-3\beta}) \ll j^{-1-\alpha/2}. \end{aligned}$$

Next we observe that by Lemma 2.2 and (1.4) we have for any bounded random variable ξ with mean zero and measurable with respect to $\mathscr{F}_j$

$$\begin{aligned} E\,\|E(\xi\,|\,\mathscr{F}_{j-1})\| &= E\left(E(\xi\,|\,\mathscr{F}_{j-1}), \frac{E(\xi\,|\,\mathscr{F}_{j-1})}{\|\cdot\|}\right)\\ &= E\left(\xi, \frac{E(\xi\,|\,\mathscr{F}_{j-1})}{\|\cdot\|}\right) \le 10\,\|\xi\|_\infty\cdot\rho(j^\beta). \end{aligned} \tag{3.17}$$

Hence by (3.6) and Lemma 2.2

$$\begin{aligned} E\,\|E(Y_j\,|\,\mathscr{F}_{j-1})\| &\le E\,\|E(Y_j\,|\,\mathscr{F}_{j-1}) - EY_j\| + \|EY_j\|\\ &\ll j^{(\beta+1)}j^{-\beta(1+\varepsilon)(1+2/\delta)} + E\{\|y_j\|1\{\|y_j\| \ge j^{\beta+1}\}\}\\ &\ll j^{-(\beta+1)(\alpha+1)}\,E\,\|y_j\|^{2+\alpha}\\ &\ll j^{-2-\alpha}. \end{aligned}$$

Hence by (3.6) we see that a typical term of the series in (3.16) is bounded by

$$j^{-\beta-1}E(\|W_{j-1}\|\cdot\|E(Y_j|\mathscr{F}_{j-1})\|) \ll j^{-\beta-\alpha-3}\|W_{j-1}\|_\infty \ll j^{-1-\alpha}. \quad \square$$

LEMMA 3.4. *Put*

$$s_n^2 = \sum_{j<n}\max(\|W_j\|^2E(\|Y_{j+1}\|^2\,|\,\mathscr{F}_j), 4(\beta+1)j^{2\beta+1}\log\log j)$$

and

$$r_n^2 = \sum_{j<n}\max(2\,\|W_j\|^2 j^\beta, 4(\beta+1)j^{2\beta+1}\log\log j). \tag{3.18}$$

Then with probability 1

$$\limsup_{n\to\infty} s_n/r_n \le 1.$$

PROOF. Using (3.6), (3.17), (1.4) and Lemma 2.3 we obtain

$$\begin{aligned} P\{E(\|Y_j\|^2\,|\,\mathscr{F}_{j-1}) \ge 2j^\beta\} &\ll P\{|E(\|Y_j\|^2\,|\,\mathscr{F}_{j-1}) - E\,\|Y_j\|^2\,| \ge \tfrac{1}{2}j^\beta\}\\ &\ll j^{-\beta}\|Y_j\|_\infty^2\rho(j^\beta) \ll j^{-2}. \end{aligned}$$

The result follows now from the Borel Cantelli Lemma.

LEMMA 3.5. *With probability* 1

$$\limsup_{n\to\infty}(s_n^2\log\log s_n)^{-1/2}M_n \le 2.$$

PROOF. Recall that by (3.13) and (3.14) M_n is a martingale with

$$M_n - M_{n-1} \le 2n^{\beta+1}.$$

Moreover,

$$E\{(M_n - M_{n-1})^2 \mid \mathscr{F}_{n-1}\} < E\{(Y_n, W_{n-1})^2 \mid \mathscr{F}_{n-1}\} \le \| W_{n-1} \|^2 E(\| Y_n \|^2 \mid \mathscr{F}_{n-1}).$$

Now the proof of the lemma can be completed in the same way as the proof of (4.32) in Goodman, Kuelbs and Zinn (1981) subject to the following minor changes.

Let λ and c be positive constants with $\lambda c \le 1$. Then $\{V_n, n \le n_0\}$ is a supermartingale. Here n_0 satisfies $n_0^{\beta+1} \le \frac{1}{2} c$. We follow the argument in [5] until (4.43) and observe that

$$s_{n+1}^2 \ge \sum_{j \le n} 4(\beta + 1) j^{2\beta+1} \log \log j.$$

Hence for n sufficiently large

$$s_n^2 \ge 2\theta^{-1} n^{2\beta+2} \log \log n$$

and thus

$$\tau_k^{\beta+1} \le \tfrac{1}{2} \theta^{k+1} (\log \log \theta^{2k})^{-1/2}.$$

The remainder of the proof is the same as in [5]. □

The following lemma is also a minor modification of Lemma 4.4 in [5], and so is its proof.

LEMMA 3.6. *Let $\{c_n\}$ be a sequence of positive numbers such that for some non-negative constants ρ and τ and for all sufficiently large n*

$$c_n^2 \le \rho n^{\tau} \sum_{k<n} c_k \log \log \left(\sum_{k<n} c_k\right).$$

Then

$$\limsup_{n\to\infty} \frac{c_n}{n^{\tau+1} \log \log n} \le \rho.$$

We finally can finish the proof of Proposition 1. Using (3.8), (3.18), Lemmas 3.2, 3.3, 3.5, and 3.4 we conclude that with probability 1

$$\limsup_{n\to\infty} (r_n^2 \log \log r_n)^{-1/2} \| W_n \|^2 \le 4. \tag{3.19}$$

Next we observe that by (3.18) we have for $\theta > 1$

$$r_n^2 \ge 2\theta^{-1} n^{2\beta+2} \log \log n$$

for sufficiently large n and hence

$$r_n^2 \log \log r_n \ge 2\theta^{-1} n^{2\beta+2} (\log \log n)^2 = (2\theta(\beta + 1)^2)^{-1} b_n^2$$

where $b_n^2 = 4(\beta + 1)^2 n^{2\beta+2} (\log \log n)^2$. Consequently we have for $\theta > 1$ and all sufficiently large n

$$\begin{aligned}(\| W_n \|^2 \vee b_n)^2 &\le 2\theta(\beta + 1)^2 r_n^2 \log \log r_n \\ &\le 2\theta^2(\beta + 1) n^{\beta} \sum_{j<n} (\| W_j \|^2 \vee b_j) \cdot \log \log\left(\sum_{j<n} \| W_j \|^2 \vee b_j\right)\end{aligned}$$

by (3.19) and (3.18). Proposition 1 follows now immediately from Lemma 3.6 and (3.3). This concludes the proof of Theorem 6.

4. Proof of Theorem 2. The proof that in Theorem 2 (i) implies (ii) is a minor modification of the proof of Theorem 6 in [7]. We first apply Theorem 3.1 of [8] Kuelbs (1976). In view of the classical law of the iterated logarithm for mixing sequences of random variables, which is easily obtained from Theorem 4 in [7], we conclude that Condition (3.1) in [8] is satisfied since the lim sup is with probability 1 $T^{1/2}(f, f) = \sup_{x \in K} f(x)$. Here K is the unit ball of the Hilbert space H_μ as described in Lemma 2.1 of [8]. Hence in view of (i), Condition (4.3) in [7] is satisfied.

Let Π_N be the maps associated with μ, as described in Lemma 2.1 of [8]. Then $\{\Pi_N x_j, j \geq 1\}$ is a weak sense stationary sequence of random variables centered at expectations with $(2 + \delta)$th moments uniformly bounded by $\| \Pi_N \|^{2+\delta}$. Hence by Proposition 2.1 in [7] for any fixed $a \geq 0$ the sequence $\{\Pi_N x_{j+a}, j \geq 1\}$ satisfies the central limit theorem with limiting Gaussian measure $\mu \circ \Pi_N^{-1}$. Thus μ has the properties required in Theorem 6 of [7].

Hence $\{x_\nu, \nu \geq 1\}$ satisfies two out of the three basic hypotheses of Theorem 6 in [7]. With regard to the third one, it might, perhaps, come as a surprise that Condition (4.2) of [7] can simply be replaced by Condition (1.5) of the present paper. Indeed, in the proof of Theorem 6 in [7] the only use of Condition (4.2) was made in the proof of Proposition 4.1 of [7]. Consequently the only thing that remains to show is Proposition 4.1 of [7] under the hypothesis (1.5) of the present paper. We shall do that by modifying the proof of Proposition 2.2 of [7] instead. We define $F(r, s)$ by (2.25) in [7], replacing η_ν by x_ν. We follow the argument in [7] until (2.28), replacing (2.28) by

$$G_k(m, \ell) = \{F(m2^{\ell+1}, 2^\ell) \geq (t_k/\log^3 t_k)^{1/2}\}.$$

Then Lemma 2.9 in [7] remains valid. Indeed, using Lemma 2.1 we have

$$P(G_k(m, \ell)) \ll \exp(-(1 + \tfrac{1}{2}\alpha)k^{\alpha/10})k^{3\alpha/20} \cdot 2^{\ell(1+\alpha/2)}$$

and thus

$$P(G_k) \ll \exp(-(1 + \tfrac{1}{2}\alpha)k^{\alpha/10})k^{3\alpha/20} \textstyle\sum_{\ell \leq n_k} 2^{\ell(1+\alpha/2)} \cdot 2^{n_k-\ell}$$

$$\ll k^{3\alpha/20}k^{-(1+\alpha/2)(1-\alpha/10)} \ll k^{-1-\alpha/20}.$$

Hence applying the Borel Cantelli Lemma we conclude that with probability 1

$$\max_{t_k<N\leq t_{k+1}} F(0, N - t_k) \ll n_k(t_k/\log^3 t_k)^{1/2} \ll (t_k/\log t_k)^{1/2}.$$

This completes the proof of Proposition 4.1 of [7] under the hypotheses of Theorem 2. There are no more changes necessary. This concludes the proof that (i) implies (ii).

The proof that (ii) implies the existence of a Gaussian measure μ with covariance function T is a minor modification of the corresponding proof in Section 5 in [7]. The only change necessary is the application of the law of the iterated logarithm for mixing random variables, an immediate consequence of Theorem 4 in [7], instead of the classical law of the iterated logarithm.

The compact law of the iterated logarithm for Brownian motion and (1.6) imply that with probability 1 $\{(n \log \log n)^{1/2} S_n, n \geq 1\}$ is relatively compact.

5. Proof of Theorem 1. The proof of Theorem 1 is the same as the proof of Theorem 1 of [7] except for two minor modifications. First we apply Theorem 2 instead of Theorem 6 of [7]. Second, in order to establish relative compactness of $\{S_n/a_n, n \geq 1\}$, we argue as follows. Let $\{e_i, i \geq 1\}$ be a complete orthonormal basis for H. We write

$$x_1 = \textstyle\sum_{i\geq 1} (x_1, e_i)e_i$$

and

$$P_N x_1 = \textstyle\sum_{i\leq N}(x_1, e_i)e_i.$$

Then for each $\rho > 0$ there is an $N_0(\rho)$ such that

$$E \| x_1 - P_N x_1 \|^{2+\delta} \leq \rho^{2+\delta} \tag{5.1}$$

for all $N \geq N_0(\rho)$. We set $\Lambda_\rho = P_{N_0(\rho)}$ and observe that by Theorem 6 and stationarity the sequence $\{a_n^{-1} \sum_{\nu\leq n} \Lambda_\rho(x_\nu), n \geq 1\}$ is with probability 1 relatively compact. We apply now Theorem 6 to the sequence $\{x_\nu - \Lambda_\rho(x_\nu), \nu \geq 1\}$ and obtain that for some constant $C(\varepsilon, \delta)$

$$\limsup_{n\to\infty} a_n^{-1} \| \textstyle\sum_{\nu\leq n} x_\nu - \Lambda_\rho(x_\nu) \| \leq \rho C(\varepsilon, \delta)$$

with probability 1. This establishes the relative compactness of $\{S_n/a_n, n \geq 1\}$. Finally we observe that in view of (5.1) and Lemma 2.3, Proposition 4.2 of [7] applies. □

We observe that Theorem 1 remains valid for weakly stationary sequences $\{x_\nu, \nu \geq 1\}$ satisfying (5.1) uniformly for all random variables x_ν. We also observe that Corollary 3 of Kuelbs and Philipp (1980) remains valid for H-valued random variables satisfying an absolute regularity condition with mixing rate (1.4). The required changes in the proof of Corollary 3 are routine.

6. Proof of Theorem 3. We first prove that (i) implies (ii). Let

$$a_n = (2n \log\log\ n)^{1/2}, \quad S_n = \textstyle\sum_{j \leq n} x_j$$

and let K denote the unit ball of $H_{L(x_1)}$ as defined in Lemma 2.1 in [5]. Then by Theorem 1.1 of Pisier (1975), K is a compact set and

$$\lim_{n\to\infty} \| S_n/a_n - K \| = 0, \quad \text{a.s.} \tag{6.1}$$

Here $\| x - K \| = \inf\{ \| x - y \| : y \in K\}$. Let Π_N be the maps as defined in Lemma 2.1 of [5] and let I be the identity map on B. Since the map $I - \Pi_N$ is continuous, we have by (6.1)

$$\| a_n^{-1}(S_n - \Pi_N S_n) - (I - \Pi_N)K \| \to 0, \quad \text{a.s.}$$

Since K is compact and since $\cap_{N\geq 1}(I - \Pi_N)K = \{0\}$ we obtain for each $\eta > 0$

$$(I - \Pi_N)K \subseteq \{x \in B : | x | < \eta\}$$

for all sufficiently large N. Hence the conclusion of Lemma 3.1 in [9] remains valid.

Since $\Pi_N x_1$ is a $\mathbb{R}^N$-valued pregaussian random variable we conclude that $\Pi_N x_1$ has finite second moment. We pick up the proof of Theorem 1 in [9] at Lemma 3.2 and as we go along, reinterpreting $\{X(t), t \geq 0\}$ as a Brownian motion determined by the Gaussian measure μ having the same covariance function as x_1, we obtain (ii).

Conversely, suppose that (ii) holds. Since $\{X(t), t \geq 0\}$ satisfies the compact law of the iterated logarithm we conclude that with probability 1 the sequence $\{S_n/a_n, n \geq 1\}$ is relatively compact. Hence with probability 1 the sequence $\{f(S_n)/a_n, n \geq 1\}$ is relatively compact for each $f \in B^*$. The Kolmogorov zero-one law therefore implies that for some constant $c < \infty$

$$\limsup_{n\to\infty} f(S_n)/a_n = c, \quad \text{a.s.} \tag{6.2}$$

Thus by Strassen's converse of the law of the iterated logarithm (see Stout, 1975, page 297), $E(f^2(x_1)) < \infty$. Hence the constant c in (6.2) is $c = (E(f^2(x_1)))^{1/2} = T^{1/2}(f, f)$. We follow the remainder of the proof of the corresponding statement of Theorem 3 in [7], Section 5 and conclude that x_1 is pregaussian.

The relative compactness of $\{S_n/a_n, n \geq 1\}$ follows as in the proof of Theorem 2.

7. Proof of Theorem 5. The proof of Theorem 5 is a minor modification of the proofs of Theorem 2 in [1] and of Theorem 3 in [9] with some of the ingredients from the proof of Theorem 1 in [1] added. As usual, we can assume without loss of generality that the σ-fields $\mathscr{F}_k$ are atomless.

We first prove Theorem 5 in the case that the random variables X_k are all discrete and that the distributions G_k are equal F_k. We follow the proof of Theorem 2 in [1], page 33 replacing the last line on page 33 by

$$E \sup_{A\in\mathscr{B}_k} | P(X_k \in A \mid \mathscr{G}_{k-1}) - P(X_k \in A) | \leq \beta_k \tag{7.1}$$

where $\mathscr{G}_{k-1}$ is the σ-field generated by $Y_1, \cdots, Y_{k-1}$. This follows from (1.7) and Lemma 2.6 in [1] since $\mathscr{G}_{k-1} \subset \mathscr{F}_{k-1}$. If (1.7*) holds then (7.1) is to be replaced by a corresponding relation. Let $\varepsilon_k > 0$. Then by (7.1)

$$\sup_{A \in B_k} | P(X_k \in A \mid \mathscr{G}_{k-1}) - P(X_k \in A) | \le \varepsilon_k$$

except on a set A_k with $P(A_k) \le \eta_k = \beta_k/\varepsilon_k$. For each atom $D \in \mathscr{G}_{k-1}$ we either have $D \subset A_k$ or $D \subset A_k^c$. Hence

$$\sup_{A \in B_k} | P_1(A) - P_2(A) | \le \varepsilon'_k$$

where

$$\varepsilon'_k \begin{cases} = 1 & \text{if } D \subset A_k \\ = \varepsilon_k & \text{if } D \subset A_k^c. \end{cases}$$

Hence the Prohorov distance of P_1 and P_2, as defined in (2.1.2) of [1], does not exceed ε'_k. We follow now the proof of Theorem 1 in [1], page 40 starting from (2.3.10), but choosing $\varepsilon_k = \beta_k^{1/2}$ at the end. We thus obtain a sequence $\{Y_k, k \ge 1\}$ of independent random variables having the same distribution as the (discrete random variables) X_k such that

$$P(| X_k - Y_k | \ge 2\beta_k^{1/2}) \le 2\beta_k^{1/2}. \tag{7.2}$$

If (1.7*) holds, then the exponent ½ is to be replaced by $p/(p + 1)$.

For the proof of the general case of Theorem 5, we copy the proof of Theorem 3 in [9] except that we replace relation (2.1.7) of [1] by (7.2) substituting $\beta^{1/2}$ for ϕ as we go along.

8. Proof of Theorem 4. In view of relations (1.7) through (1.9) of [7], the proofs of the results mentioned in Theorem 4 reduce to a proof of a bounded law of the iterated logarithm. This follows at once from Theorem 2. Indeed (1.7) through (1.9) of [7] immediately imply condition (1.5), and the existence of a mean zero Gaussian measure with covariance function T follows from Proposition 4.2 in [7]. Finally the proof of Lemma 4.6 of [7] shows the relative compactness of $\{a_n^{-1}S_n, n \ge 1\}$ provided that we have a bounded law of the iterated logarithm.

But as the proof of Theorem 5 in [7] shows, for this purpose we only need an exponential bound such as the one given in Proposition 3.1 of [7]. In the proof of Proposition 3.1 of [7] there were only two places where the ϕ-mixing condition was used decisively. The first one occurred when Lemma 3.1 of [7] was applied. We replace it by Lemma 2.1 which holds in the present set-up since condition (1.5) holds as already observed in the previous paragraph. That δ is to be replaced by α does not matter (as long as $\alpha > 0$). The second use of the ϕ-mixing condition was made when Theorem 2 of [1] was applied in (3.9) of [7]. We replace it by Theorem 5. The slight worsening of the mixing rate (β_k instead of ϕ_k) is compensated by the restriction of δ to $0 < \delta \le 1/5$.

This concludes the proof of the exponential bound and thus the proof of Theorem 4.

Acknowledgment. We are grateful to V. Goodman, J. Kuelbs and J. Zinn for showing us their unpublished manuscript.

REFERENCES

[1] BERKES, ISTVÁN and PHILIPP, WALTER (1979). Approximation theorems for independent and weakly dependent random vectors. *Ann. Probability* **7** 29–54.

[2] DAVYDOV, YU. A. (1970). The invariance principle for stationary processes. *Theory Probability Appl.* **15** 487–498.

[3] DEO, CHANDRAKANT M. (1973). A note on empirical processes of strong mixing sequences. *Ann. Probability* **1** 870–875.

[4] DEHLING, HEROLD (1982), A note on a theorem of Berkes and Philipp *Z. Wahrsch. verw. Gebiete*, to appear.

[5] GOODMAN, V., KUELBS, J. and ZINN, J. (1981). Some results on the LIL in Banach space with applications to the weighted empirical process. *Ann. Probability* **9** 713–752.

[6] IBRAGIMOV, I. A. and ROZANOV, YU. A. (1978). *Gaussian Random Processes.* Springer, New York.

[7] KUELBS, J. and PHILIPP, WALTER (1980). Almost sure invariance principles for partial sums of mixing B-valued random variables. Ann. Probability **8** 1003–1036.

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

[9] PHILIPP, WALTER (1979). Almost sure invariance principles for sums of B-valued random variables. Probability in Banach spaces. *Lecture Notes in Mathematics* **709** 171–193. Springer, New York.
[10] PISIER, G. (1975). Le théorème de la limite centrale et la loi du logarithme itéré dans les espaces de Banach. Seminaire Maurey-Schwartz 1975–76, Exposes 3 et 4 Ecole Polytechnique, Paris.
[11] LINDENSTRAUSS, JORAM and TZAFRIRI, LIOR (1977). Classical Banach spaces I. *Ergebnisse Math. Grenzgebiete.* Springer, Berlin.
[12] SERFLING, R. J. (1968). Contributions to central limit theory for dependent variables. *Ann. Math. Statist.* **39** 1158–1175.
[13] SOTRES, DAVID A. and GHOSH, MALAY (1977). Strong convergence of linear rank statistics for mixing processes. *Sankhya* Ser. B, **39** 1–11.
[14] STOUT, W. F. (1974). *Almost Sure Convergence.* Academic, New York.
[15] VOLKONSKII, V. A. and ROZANOV, YU. A. (1959). Some limit theorems for random functions I. *Theory Probability Appl.* **4** 178–197.

INSTITUT FÜR MATHEMATISCHE STOCHASTIK
UNIVERSITÄT GÖTTINGEN
LOTZESTR. 13
D-3400 GÖTTINGEN, WEST GERMANY

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ILLINOIS
URBANA, ILLINOIS 61801