APPROXIMATION THEOREMS FOR STRONGLY MIXING RANDOM VARIABLES

Richard C. Bradley

1. Introduction. As a technique for proving limit theorems for dependent random variables, the direct approximation of dependent r.v.'s by independent ones has been gaining popularity since it was introduced in articles by Berkes and Philipp ([2], [3]). The idea is to carry out such an approximation in a manner that permits limit theorems for the independent r.v.'s to carry over directly to the dependent ones. Our purpose is to derive some sharp "approximation theorems" for random sequences satisfying the "strong mixing" condition. Before we discuss some of the results in the literature that pertain to this problem, it will be convenient to define some terminology.

First, a "uniform-[0, 1]" random variable is simply a r.v. which is uniformly distributed on the interval [0, 1].

A "Borel space" is a measurable space $(\mathcal{S}, \mathcal{D})$ which is (bimeasurably) isomorphic to a Borel subset of the real number line **R**. (No metric is needed in this definition.) When we refer to an "$\mathcal{S}$-valued" r.v. X on a probability space $(\Omega, \mathcal{F}, P)$, the σ-algebra $\mathcal{D}$ that accompanies $\mathcal{S}$ will usually be suppressed, but it is implicitly understood that $\forall D \in \mathcal{D}$ the set of sample points $\{\omega : X(\omega) \in D\}$ is an element of $\mathcal{F}$. The σ-field of such events $\{X \in D\}$, $D \in \mathcal{D}$, is denoted by $\mathcal{B}(X)$. The Euclidian spaces $\mathbf{R}^J$, $1 \leqslant J \leqslant \infty$, are always accompanied by the usual (J-dimensional) Borel σ-algebra, and are well known to be Borel spaces.

Let $(\Omega, \mathcal{F}, P)$ be a probability space. For any two σ-fields $\mathcal{A}$ and $\mathcal{B}$ $(\subset \mathcal{F})$ define the following measures of dependence

$$\alpha(\mathcal{A}, \mathcal{B}) = \sup |P(A \cap B) - P(A)P(B)| \quad A \in \mathcal{A},\ B \in \mathcal{B}$$

$$\rho(\mathcal{A}, \mathcal{B}) = \sup |\operatorname{Corr}(f, g)| \quad f \in \mathcal{L}^2(\mathcal{A}),\ g \in \mathcal{L}^2(\mathcal{B})$$

$$\beta(\mathcal{A}, \mathcal{B}) = \sup (1/2) \sum_{i=1}^{I} \sum_{j=1}^{J} |P(A_i \cap B_j) - P(A_i)P(B_j)|$$

where this latter sup is taken over all pairs of partitions $\{A_1, \ldots, A_I\}$ and $\{B_1, \ldots, B_J\}$ of Ω such that each $A_i \in \mathcal{A}$ and each $B_j \in \mathcal{B}$. In the definition of $\rho(\mathcal{A}, \mathcal{B})$ it is understood that $\operatorname{Corr}(f, g) \equiv 0$ if f or g is constant a.s. Obviously $\alpha(\mathcal{A}, \mathcal{B}) \leqslant \rho(\mathcal{A}, \mathcal{B}) \leqslant 1$ and $\alpha(\mathcal{A}, \mathcal{B}) \leqslant \beta(\mathcal{A}, \mathcal{B}) \leqslant 1$, but there is no general comparison either way between $\rho(\mathcal{A}, \mathcal{B})$ and $\beta(\mathcal{A}, \mathcal{B})$.

The following approximation theorem comes from Berbee's [1] book.

THEOREM A ([1, Corollary 4.2.5]). *Suppose X and Y are* r.v.*'s taking their values in Borel spaces $\mathcal{S}_1$ and $\mathcal{S}_2$ respectively, and suppose U is a uniform-*[0, 1]

Received March 17, 1982. Revision received May 24, 1982.
This work was partially supported by NSF grant MCS 81-01583.
Michigan Math. J. 30 (1983).

r.v. *independent of* (X, Y). *Then there exists an* $\mathcal{S}_2$*-valued* r.v. $Y^* = f(X, Y, U)$ *where* f *is a measurable function from* $\mathcal{S}_1 \times \mathcal{S}_2 \times [0,1]$ *into* $\mathcal{S}_2$, *such that*

(i) Y^* *is independent of* X,

(ii) *the probability distributions of* Y^* *and* Y *on* $\mathcal{S}_2$ *are identical, and*

(iii) $P(Y^* \neq Y) = \beta(\mathcal{B}(X), \mathcal{B}(Y))$.

This theorem is proved on pp. 91–95 of Berbee [1] and extends an earlier similar result of Schwarz [17]. (The publication of Schwarz' article was delayed.) We have simply restated [1, p. 94, Corollary 4.2.5] in a more explicit manner (consistent with Berbee's proof). Theorem A is sharp in the sense that one cannot have $P(Y^* \neq Y) < \beta(\mathcal{B}(X), \mathcal{B}(Y))$ and still retain (i) and (ii); see for example [1, p. 92, Proposition 4.2.2] (Schwarz' result). The use of a uniform-[0, 1] r.v. here (and in similar theorems to be given below) is purely a matter of taste; U could be replaced by any continuous r.v. Bryc [7] gives a further extension of Theorem A.

In the case where $\mathcal{S}_1$ and $\mathcal{S}_2$ are finite-dimensional Euclidian spaces one has, as a special case of Theorem 1 of Berkes and Philipp [3], a result like Theorem A but with (iii) replaced by an inequality of the form $P(d(Y^*, Y) > c) < c$ where d denotes Euclidian distance on $\mathcal{S}_2$ and c is a positive quantity that depends partly on $\alpha(\mathcal{B}(X), \mathcal{B}(Y))$. (See the remarks on p. 31 of [3].) Dehling [11] showed with a counterexample that a nontrivial general result of this kind does not exist for $\alpha(\mathcal{B}(X), \mathcal{B}(Y))$ or even $\rho(\mathcal{B}(X), \mathcal{B}(Y))$ when $\mathcal{S}_1$ and $\mathcal{S}_2$ are permitted to be arbitrary metric spaces; his counterexample is discussed in terms of $\alpha(\ ,\ )$ but carries over verbatim to $\rho(\ ,\ )$. We will seek fairly sharp approximation theorems for $\alpha(\ ,\ )$ and $\rho(\ ,\ )$ of a certain specific nature; our main tools will be Theorem A and the following statement:

THEOREM 1. *Suppose* $\mathcal{A}$ *and* $\mathcal{B}$ *are σ-fields, with* $\mathcal{B}$ *being purely atomic with exactly* N *atoms. Then* $\beta(\mathcal{A}, \mathcal{B}) \leq (8N)^{1/2}\alpha(\mathcal{A}, \mathcal{B})$.

Our study of approximation theorems will follow lines similar to work done by W. Bryc [6, 7, 8]. Under the hypothesis of Theorem 1 the inequality $\beta(\mathcal{A}, \mathcal{B}) \leq N\cdot\alpha(\mathcal{A}, \mathcal{B})$ is easy to establish, and one approach used by Bryc in [6] was to combine this inequality with Theorem A in order to derive a new approximation theorem under strong mixing, similar to [3, Theorem 1]. In deriving our new approximation theorems (Theorems 2 and 3 below) we will essentially follow Bryc's argument, except that Theorem A will be combined with Theorem 1 rather than with the inequality $\beta(\mathcal{A}, \mathcal{B}) \leq N\cdot\alpha(\mathcal{A}, \mathcal{B})$. Using a construction in [8], we will show later on that Theorem 1 is sharp up to a constant factor, even if $\alpha(\mathcal{A}, \mathcal{B})$ is replaced by $\rho(\mathcal{A}, \mathcal{B})$.

Our first approximation theorem is as follows:

THEOREM 2. *Suppose* X *and* Y *are* r.v.*'s taking their values in Borel spaces* $\mathcal{S}_1$ *and* $\mathcal{S}_2$ *respectively; and suppose* U *is a uniform*-[0, 1] r.v. *independent of* (X, Y). *Suppose* N *is a positive integer and* $\mathcal{H} = \{H_1, H_2, \ldots, H_N\}$ *is a measurable partition of* $\mathcal{S}_2$. *Then there exists an* $\mathcal{S}_2$*-valued* r.v. $Y^* = f(X, Y, U)$ *where* f *is a measurable function from* $\mathcal{S}_1 \times \mathcal{S}_2 \times [0,1]$ *into* $\mathcal{S}_2$, *such that*

(i) *Y^* is independent of X,*
(ii) *the probability distributions of Y^* and Y on $\mathcal{S}_2$ are identical, and*
(iii) *$P(Y^*$ and Y are* not *elements of the same $H_i \in \mathcal{H}) \leqslant (8N)^{1/2}\alpha(\mathcal{B}(X), \mathcal{B}(Y))$.*

In Theorem 2 the terminology "$\mathcal{H}$ is a measurable partition of $\mathcal{S}_2$" means of course that each H_i is an element of the σ-algebra that accompanies $\mathcal{S}_2$. Condition (ii) here is a matter of convenience; it will become clear from the proof that Y^* can have any preassigned distribution such that $P(Y^* \in H_i) = P(Y \in H_i)$ would hold for all i. If one wants to have $P(Y^* \in H_i) \neq P(Y \in H_i)$ for some i but retain (i), then another term may be needed on the right-hand side of (iii). The next statement is a corollary of Theorem 2.

THEOREM 3. *Suppose X and Y are* r.v.*'s taking their values on $\mathcal{S}_1$ and* **R**, *respectively, where $\mathcal{S}_1$ is a Borel space; suppose U is a uniform-*[0, 1] r.v. *independent of (X, Y); and suppose q and γ are positive numbers such that $q \leqslant \|Y\|_\gamma < \infty$. Then there exists a real-valued* r.v. *$Y^* = f(X, Y, U)$ where f is a measurable function from $\mathcal{S}_1 \times \mathbf{R} \times [0,1]$ into* **R**, *such that*
(i) *Y^* is independent of X,*
(ii) *the probability distributions of Y^* and Y are identical, and*
(iii) *$P(|Y^* - Y| \geqslant q) \leqslant 18(\|Y\|_\gamma/q)^{\gamma/(2\gamma+1)}[\alpha(\mathcal{B}(X), \mathcal{B}(Y))]^{2\gamma/(2\gamma+1)}$.*

In Theorem 3 and in the rest of this article, $\|Y\|_\gamma \equiv (E|Y|^\gamma)^{1/\gamma}$. The proof of Theorem 3 (to be given later) is simple and can be adapted to cases where Y takes its values on a finite-dimensional Euclidian space or—under rather stringent conditions on the distribution of Y—on more general metric spaces; the inequality in Theorem 3(iii) has to be revised accordingly. One can also derive Theorem 3 for the case $\gamma = \infty$ (i.e. if $\|Y\|_\infty \equiv \text{ess sup}|Y| < \infty$) and then apply that to truncations of r.v.'s to get bounds similar to those in [3, Theorem 1]; here the r.h.s. of Theorem 3(iii) is interpreted to be $18(\|Y\|_\infty/q)^{1/2}\alpha(\mathcal{B}(X), \mathcal{B}(Y))$.

To illustrate the use of Theorem 3 (although [3, Theorem 1] can be used for the same purpose) we will prove an "almost sure invariance principle" under the strong mixing condition with a logarithmic mixing rate. Suppose $(X_k, k = \ldots, -1, 0, 1, \ldots)$ is a strictly stationary sequence of real-valued r.v.'s. For each real number $t \geqslant 0$ define $S(t) \equiv \sum_{1 \leqslant k \leqslant t} X_k$ (with $S(t) \equiv 0$ if $0 \leqslant t < 1$). For each positive integer n define $S_n \equiv X_1 + X_2 + \cdots + X_n \equiv S(n)$ and $\alpha(n) \equiv \alpha(\mathcal{B}(X_k, k \leqslant 0), \mathcal{B}(X_k, k \geqslant n))$. In what follows, "log" will mean the natural logarithm.

THEOREM 4. *Suppose (X_k) is a strictly stationary sequence of real-valued* r.v.*'s with $EX_k = 0$, $EX_k^2 < \infty$, and* $\operatorname{Var} S_n \to \infty$ *as $n \to \infty$. Suppose $\delta > 0$ and $\lambda > 1 + 3/\delta$ are real numbers, $\alpha(n) = o((\log n)^{-\lambda})$ as $n \to \infty$, and*

$$\sup_n E|S_n|^{2+\delta}/(\operatorname{Var} S_n)^{(2+\delta)/2} < \infty. \tag{1.1}$$

Then there exists σ^2, $0 < \sigma^2 < \infty$, such that $\lim_{n\to\infty} n^{-1} \operatorname{Var} S_n = \sigma^2$. *Without changing its probability distribution the process $(S(t), t \geqslant 0)$ can be redefined on* *another probability space, together with a standard Wiener process* $(W(t), t \geq 0)$, *such that*

$$(1.2)\qquad P(|S(t)-W(\sigma^2 t)| = o(t^{1/2}(\log\log t)^{-1/2}) \text{ as } t \to \infty) = 1.$$

This result is motivated by two a.s. invariance principles that have been proved under ϕ-mixing with a logarithmic mixing rate: the first by Berkes and Philipp [3, Theorem 4] and the second by Dabrowski [9]. The result in [9] was obtained with a sharp treatment of the basic approach of Berkes and Philipp; the mixing rate was thereby considerably relaxed, at the expense of a larger error term. The quantity $\sigma^2 \equiv \lim n^{-1} \operatorname{Var} S_n$ played the same role in [9] as in Theorem 4 here. The proof of Theorem 4 will follow Dabrowski's argument rather closely, but will be given in detail because of several necessary minor modifications. In Theorem 4, in exchange for the use of the "strong mixing" condition $\alpha(n) \to 0$ (which is weaker than ϕ-mixing), a faster mixing rate than Dabrowski's is imposed (in our language, he used $\lambda > 1 + 2/\delta$ where $E|X_k|^{2+\delta} < \infty$), along with the rather stringent condition (1.1). If (1.1) were replaced by the weaker condition $E|X_k|^{2+\delta} < \infty$ used in both [3, Theorem 4] and [9], then Theorem 4 would fail badly; see for example Davydov's [10, pp. 320–324] counterexamples to the central limit theorem. This can be remedied either by imposing a polynomial mixing rate on $\alpha(n)$ as in Theorem 4 of Kuelbs and Philipp [15], or else by imposing, say, an additional condition on the maximal correlation coefficients $\rho(n) \equiv \rho(\mathcal{B}(X_k, k \leq 0), \mathcal{B}(X_k, k \geq n))$ to insure that (1.1) holds, as in the central limit theorems [12, Theorem 2.1] (where $\rho(n) \to 0$ is assumed) and [4, Theorem 5] (i.e. assuming $\delta \leq 1$ and a small positive limit for $\rho(n)$).

Section 2 will deal with the question of "sharpness" for Theorem 1; Section 3 will give the proofs of Theorems 1, 2, and 3; and Section 4 will give the proof of Theorem 4. The following notations will be used:

(i) I_F denotes the indicator function of a set F.
(ii) $f(t) \approx g(t)$ means $\lim_{t\to\infty} f(t)/g(t) = 1$.
(iii) $g(t) \ll h(t)$ will mean $g(t) = O(h(t))$ as $t \to \infty$, (as in [3, 9, 15]).
(iv) To avoid subscripts of subscripts, terms like a_b will often be written as $a(b)$.

2. An example for Theorem 1. It was mentioned above that Theorem 1 is sharp up to a constant factor, even if $\alpha(\mathcal{A}, \mathcal{B})$ is replaced by $\rho(\mathcal{A}, \mathcal{B})$. We will give an example to show this for even integers N; with trivial modifications the same example will work to show this for odd N.

Let N be an arbitrary positive even integer. We will construct a probability space with σ-fields $\mathcal{A}$ and $\mathcal{B}$ such that $\mathcal{B}$ is purely atomic with exactly N atoms and

$$(2.1)\qquad \beta(\mathcal{A}, \mathcal{B}) = 1/2$$

$$(2.2)\qquad \rho(\mathcal{A}, \mathcal{B}) \leq (2/N)^{1/2}$$

The same construction was used in [8] for a related purpose.

Let Ω_1 be the interval $[0, 1]$, let $\mathcal{F}_1$ be the σ-algebra of Borel subsets of Ω_1, and let P_1 be Lebesgue measure on $(\Omega_1, \mathcal{F}_1)$.

Let $\Omega_2 = \{1, 2, \ldots, N\}$, let $\mathfrak{F}_2$ be the σ-algebra of all subsets of Ω_2, and let P_2 be the "uniform" probability measure on $(\Omega_2, \mathfrak{F}_2)$ given by $P_2(\{j\}) = 1/N \ \forall j \in \Omega_2$.

Let $m = N/2$; let $h_1(x), h_2(x), \ldots, h_m(x)$ denote the first m Rademacher functions on $[0,1]$; and define the probability measure P on $(\Omega_1 \times \Omega_2, \mathfrak{F}_1 \times \mathfrak{F}_2)$ by

$$dP(x,j) = \begin{cases} (1+h_j(x))d(P_1 \times P_2)(x,j) & \text{if } x \in [0,1] \text{ and } 1 \leqslant j \leqslant m \\ (1-h_{j-m}(x))d(P_1 \times P_2)(x,j) & \text{if } x \in [0,1] \text{ and } m+1 \leqslant j \leqslant N. \end{cases}$$

We are taking the liberty of using the same letter P here as in the rest of the article; for this example the probability space is $(\Omega_1 \times \Omega_2, \mathfrak{F}_1 \times \mathfrak{F}_2, P)$. Note that the marginals of P are P_1 and P_2 respectively.

Define the σ-fields $\mathcal{A} \equiv \{F \times \Omega_2 : F \in \mathfrak{F}_1\}$ and $\mathcal{B} \equiv \{\Omega_1 \times F : F \in \mathfrak{F}_2\}$. Since P is absolutely continuous with respect to $P_1 \times P_2$ we have that

$$\beta(\mathcal{A}, \mathcal{B}) = (1/2) \int_{\Omega_1 \times \Omega_2} |dP/d(P_1 \times P_2) - 1| d(P_1 \times P_2)$$

$$= (1/2) \int_{\Omega_1 \times \Omega_2} 1 d(P_1 \times P_2) = 1/2$$

since $h_j(x) = \pm 1 \ \forall j, x$. That is, (2.1) holds.

Let f be an arbitrary Borel function on $[0,1]$ with $\int_0^1 f(x)\,dx = 0$ and $0 < \int_0^1 [f(x)]^2\,dx < \infty$. Let g be an arbitrary function on $\{1, 2, \ldots, N\}$. Now f and g can each be regarded naturally as a r.v. on $\Omega_1 \times \Omega_2$. To show (2.2) it suffices to show $|\mathrm{Corr}(f,g)| \leqslant (2/N)^{1/2}$.

Now $g = g_1 + g_2$ where the functions g_1 and g_2 satisfy $g_1(j+m) = -g_1(j)$ and $g_2(j+m) = g_2(j)$ for all $j = 1, \ldots, m$. With some simple calculations and the fact $Ef = 0$ one has

$$\mathrm{Var}\, g \geqslant \mathrm{Var}\, g_1 + 2(Eg_1g_2 - Eg_1Eg_2) = \mathrm{Var}\, g_1 + 2(0 - 0Eg_2)$$

$$\mathrm{Cov}(f,g) = Efg = Efg_1 + Efg_2 = Efg_1 + 0 = \mathrm{Cov}(f, g_1)$$

and hence $|\mathrm{Corr}(f,g)| \leqslant |\mathrm{Corr}(f, g_1)|$. To show (2.2) it suffices to show

$$|\mathrm{Corr}(f, g_1)| \leqslant (2/N)^{1/2}.$$

For each $j = 1, \ldots, m$ let $c_j \equiv g_1(j) = -g_1(j+m)$. Since $Ef = Eg_1 = 0$ and $\int_0^1 h_j(x)h_k(x)\,dx = 0$ for $j \neq k$ we have

$$|\mathrm{Cov}(f, g_1)| = |Efg_1| = \left| \sum_{j=1}^{m} Efg_1(I_{[0,1]\times\{j\}} + I_{[0,1]\times\{j+m\}}) \right|$$

$$= \left| \sum_{j=1}^{m} \left[\int_0^1 c_j f(x)[1 + h_j(x)](1/N)\,dx \right.\right.$$

$$\left.\left. + \int_0^1 (-c_j) f(x)[1 - h_j(x)](1/N)\,dx \right] \right|$$

$$= \left| (2/N) \int_0^1 \sum_{j=1}^{m} c_j h_j(x) f(x)\,dx \right| \leqslant$$

$$\leq (2/N)\left[\int_0^1 \left[\sum_{j=1}^m c_j h_j(x)\right]^2 dx\right]^{1/2} \left[\int_0^1 [f(x)]^2 dx\right]^{1/2}$$
$$= (2/N)\left[\int_0^1 \sum_{j=1}^m c_j^2 h_j^2(x)\, dx\right]^{1/2} (\operatorname{Var} f)^{1/2}$$
$$= (2/N)\left[\sum_{j=1}^m c_j^2\right]^{1/2} (\operatorname{Var} f)^{1/2}$$
$$= (2/N)[(N/2)\operatorname{Var} g_1]^{1/2}(\operatorname{Var} f)^{1/2}.$$

Hence $|\operatorname{Corr}(f, g_1)| \leq (2/N)^{1/2}$, and hence (2.2) holds.

3. Proofs of Theorems 1–3. For the proof of Theorem 1 we will need a couple of technical lemmas.

LEMMA 1. *Suppose* $Z_1, Z_2, Z_3, \ldots$ *are* i.i.d. r.v.*'s with*

$$P(Z_k = 1) = P(Z_k = -1) = 1/2.$$

Then for any positive integer N and any real numbers $a_1, a_2, \ldots, a_N$ one has $E|a_1 Z_1 + a_2 Z_2 + \cdots + a_N Z_N| \geq (2N)^{-1/2}(|a_1| + |a_2| + \cdots + |a_N|)$.

Proof. This is an immediate consequence of Szarek's bound [19, p. 198, Theorem 1] for the Khinchin inequality. We may take our probability space to be the unit interval [0, 1] with P being Lebesgue measure, and we may assume that for each k the r.v. Z_k is the kth Rademacher function, which Szarek calls $r_k(t)$. Then by [19, Theorem 1] and the Cauchy–Schwarz inequality we have

$$E\left|\sum_{k=1}^N a_k Z_k\right| = \int_0^1 \left|\sum_{k=1}^N a_k r_k(t)\right| dt \geq 2^{-1/2}\left(\sum_{k=1}^N a_k^2\right)^{1/2} \geq 2^{-1/2} N^{-1/2} \sum_{k=1}^N |a_k|$$

and Lemma 1 is verified. □

LEMMA 2. *Suppose $A = (a_{ij},\ 1 \leq i \leq M,\ 1 \leq j \leq N)$ is an $M \times N$ matrix (of real numbers). Then there exists a subset $S \subset \{1, 2, \ldots, M\}$ and a subset $T \subset \{1, 2, \ldots, N\}$ such that*

$$\left|\sum_{i \in S}\sum_{j \in T} a_{ij}\right| \geq (32 \cdot \min\{M, N\})^{-1/2} \cdot \sum_{i=1}^M \sum_{j=1}^N |a_{ij}|.$$

Proof. Without loss of generality we assume $M \geq N$.

Let $Z_1, Z_2, \ldots, Z_N$ be i.i.d. random variables (on some probability space) such that $P(Z_k = 1) = P(Z_k = -1) = \frac{1}{2}\ \forall k$.

Define the random variable $X \equiv \sum_{i=1}^M |\sum_{j=1}^N a_{ij} Z_j|$. Letting $a \equiv \sum_{i=1}^M \sum_{j=1}^N |a_{ij}|$ we have by Lemma 1, $EX = \sum_{i=1}^M E|\sum_{j=1}^N a_{ij} Z_j| \geq (2N)^{-1/2} a$.

Hence there exists an N-tuple $(z_1, \ldots, z_N)$ of fixed numbers $\in \{-1, +1\}$ such that $\sum_{i=1}^M |\sum_{j=1}^N a_{ij} z_j| \geq (2N)^{-1/2} a$.

Define the sets

$$S(1) \equiv \left\{ i : \sum_{j=1}^{N} a_{ij} z_j \geq 0 \right\} \qquad S(2) \equiv \left\{ i : \sum_{j=1}^{N} a_{ij} z_j < 0 \right\}.$$

Also define the number $b \equiv (2N)^{-1/2} a$. Then either

$$\sum_{i \in S(1)} \sum_{j=1}^{N} a_{ij} z_j \geq b/2 \quad \text{or} \quad -\sum_{i \in S(2)} \sum_{j=1}^{N} a_{ij} z_j \geq b/2.$$

In either case we have a subset $S \subset \{1, 2, \ldots, M\}$ such that $|\sum_{i \in S} \sum_{j=1}^{N} a_{ij} z_j| \geq b/2$.

Define the sets $T(1) \equiv \{j : z_j = +1\}$ and $T(2) \equiv \{j : z_j = -1\}$. Then either

$$\left| \sum_{i \in S} \sum_{j \in T(1)} a_{ij} \right| \geq b/4 \quad \text{or} \quad \left| \sum_{i \in S} \sum_{j \in T(2)} a_{ij} \cdot (-1) \right| \geq b/4.$$

In either case we have a subset $T \subset \{1, 2, \ldots, N\}$ such that $|\sum_{i \in S} \sum_{j \in T} a_{ij}| \geq b/4 = (32N)^{-1/2} a$, and Lemma 2 is proved. □

Proof of Theorem 1: Let $B_1, B_2, \ldots, B_N$ denote the atoms of $\mathcal{B}$. Let $\{A_1, A_2, \ldots, A_M\}$ be an arbitrary finite partition of our probability space Ω such that $A_i \in \mathcal{A}$ $\forall i$. It suffices to prove

$$\sum_{i=1}^{M} \sum_{j=1}^{N} |P(A_i \cap B_j) - P(A_i) P(B_j)| \leq (32N)^{1/2} \alpha(\mathcal{A}, \mathcal{B}). \tag{3.1}$$

For each pair (i, j) let $a_{ij} \equiv P(A_i \cap B_j) - P(A_i) P(B_j)$. Then by Lemma 2, for some subsets $S \subset \{1, \ldots, M\}$ and $T \subset \{1, \ldots, N\}$ we have

$$\left| \sum_{i \in S} \sum_{j \in T} a_{ij} \right| \geq (32N)^{-1/2} \sum_{i=1}^{M} \sum_{j=1}^{N} |a_{ij}|.$$

Define the events $A \equiv \bigcup_{i \in S} A_i$ and $B \equiv \bigcup_{j \in T} B_j$. Then $A \in \mathcal{A}$ and $B \in \mathcal{B}$ and hence

$$\alpha(\mathcal{A}, \mathcal{B}) \geq |P(A \cap B) - P(A) P(B)| = \left| \sum_{i \in S} \sum_{j \in T} a_{ij} \right| \geq (32N)^{-1/2} \sum_{i=1}^{M} \sum_{j=1}^{N} |a_{ij}|$$

and we have (3.1). This completes the proof of Theorem 1. □

Proof of Theorem 2. Without losing generality we assume that $\mathcal{S}_2 = \mathbf{R}$ (for otherwise we could simply employ a one-to-one bimeasurable mapping from $\mathcal{S}_2$ into $\mathbf{R}$) and that $P(Y \in H_i) > 0$ $\forall i$. For each i let y_i be an element of H_i. Define the r.v. Y_1 by $Y_1 \equiv \sum_{i=1}^{N} y_i I_{\{Y \in H(i)\}}$.

Let U_1 and U_2 be independent uniform-[0, 1] r.v.'s which are each a measurable function of U. (For example let the binary expansions of U_1 and U_2 consist of alternating digits in the binary expansion of U.)

Using Theorem A let Y_1^* be defined as a measurable function of (X, Y_1, U_1) such that Y_1^* is independent of X and has the same distribution as Y_1 and such that

$$\begin{aligned} P(Y_1^* \neq Y_1) &= \beta(\mathcal{B}(X), \mathcal{B}(Y_1)) \leq (8N)^{1/2} \alpha(\mathcal{B}(X), \mathcal{B}(Y_1)) \\ &\leq (8N)^{1/2} \alpha(\mathcal{B}(X), \mathcal{B}(Y)). \end{aligned} \tag{3.2}$$

The inequalities here come respectively from Theorem 1 and the fact $\mathcal{B}(Y_1) \subset \mathcal{B}(Y)$.

Now one can define a r.v. Y^* as a measurable function of (Y_1^*, U_2) such that Y^* has the same distribution as Y and $Y_1^* = \sum_{i=1}^N y_i I_{\{Y^* \in H(i)\}}$ a.s. (For example, for each i, within the event $\{Y_1^* = y_i\}$ simply define Y^* to be the U_2-th quantile of the conditional distribution of Y given $Y_1 = y_i$.) Since U_2 is independent of (X, Y_1^*) and Y_1^* is independent of X, we have that Y^* is independent of X. Also Theorem 2(iii) holds by (3.2), and the proof of Theorem 2 is complete. □

Proof of Theorem 3. We assume $\alpha(\mathcal{B}(X), \mathcal{B}(Y)) > 0$, for otherwise Theorem 3 would hold trivially with $Y^* \equiv Y$. Let

$$(3.3) \qquad \begin{aligned} &M \equiv \|Y\|_\gamma \qquad \alpha \equiv \alpha(\mathcal{B}(X), \mathcal{B}(Y)) \\ &x \equiv [(1/\alpha)(M/q)^\gamma \max\{\gamma, 1\}]^{2/(2\gamma+1)} \end{aligned}$$

Since $M/q \geqslant 1$ by hypothesis, we have $x \geqslant 1$. Let m be an integer that satisfies $x \leqslant m \leqslant 2x$.

For $-m \leqslant j \leqslant m$ let H_j denote the half-open interval $(jq - q/2, jq + q/2]$. Define the half-lines $H_{m+1} \equiv (mq + q/2, \infty)$ and $H_{-m-1} \equiv (-\infty, -mq - q/2]$. Then $\{H_j: -m-1 \leqslant j \leqslant m+1\}$ is a partition of $\mathbf{R}$.

Applying Theorem 2, let Y^* be defined as a measurable function of (X, Y, U) such that Y^* is independent of X and has the same distribution as Y and such that $P(Y^*$ and Y belong to different H_j's$) \leqslant [8(2m+3)]^{1/2}\alpha$. Then using (3.3), the definition of m, and Markov's inequality at the appropriate places, we get

$$\begin{aligned} P(|Y^* - Y| \geqslant q) &\leqslant [8(2m+3)]^{1/2}\alpha + P(|Y^*| \geqslant mq + q/2) + P(|Y| \geqslant mq + q/2) \\ &\leqslant (56x)^{1/2}\alpha + 2P(|Y| \geqslant xq) \leqslant (56x)^{1/2}\alpha + 2(M/q)^\gamma x^{-\gamma} \\ &\leqslant C_\gamma (M/q)^{\gamma/(2\gamma+1)} \alpha^{2\gamma/(2\gamma+1)} \end{aligned}$$

where $C_\gamma \equiv 2 + 8[\max\{\gamma, 1\}]^{1/(2\gamma+1)}$. Now $C_\gamma \leqslant \max\{10, 2 + 8\gamma^{1/\gamma}\}$ and as a simple exercise in calculus one has that $\forall x > 0$, $x^{1/x} \leqslant e^{1/e} \leqslant 3^{1/2} < 2$ and hence $C_\gamma \leqslant 18$. Theorem 3 is proved. □

4. Proof of Theorem 4. Throughout this proof we follow the arguments in [3, Theorem 4] and [9]; several sections of our proof are taken directly from [9].

We will first prove the existence of the quantity σ^2 by applying [5, p. 587, Theorem 1(i)]. If $I \geqslant 0$ and $n \geqslant 1$ are integers, then letting

$$W_1 \equiv \sum_{k=-I}^{0} X_k / (\operatorname{Var} S_{I+1})^{1/2} \quad \text{and} \quad W_2 \equiv \sum_{k=n}^{n+I} X_k / (\operatorname{Var} S_{I+1})^{1/2}$$

we have, by [13, p. 307, Theorem 17.2.2], that

$$(4.1) \qquad \begin{aligned} \left| \operatorname{Corr}\left(\sum_{k=-I}^{0} X_k, \sum_{k=n}^{n+I} X_k \right) \right| &= |\operatorname{Cov}(W_1, W_2)| \\ &\leqslant 10D[\alpha(n)]^{\delta/(2+\delta)} \ll (\log n)^{-\lambda\delta/(2+\delta)} \end{aligned}$$

where $D \equiv \sup_n E|S_n|^{2+\delta} / (\operatorname{Var} S_n)^{(2+\delta)/2}$. Hence by [5, Theorem 1(i)] there exists σ^2, $0 < \sigma^2 < \infty$, such that $\lim_{n \to \infty} n^{-1} \operatorname{Var} S_n = \sigma^2$ (since $\lambda\delta/(2+\delta) > 1$).

For the remainder of the proof of Theorem 4 we assume without loss of generality that

(4.2) $$\sigma^2 \equiv \lim_{n\to\infty} n^{-1} \operatorname{Var} S_n = 1$$

By [5, Theorem 1(ii)] (use any $\gamma \in (0,1)$ there) along with (4.1) and some routine calculations we have

(4.3) $$|1 - n^{-1} \operatorname{Var} S_n| \ll (\log n)^{1-\lambda\delta/(2+\delta)}.$$

We also have

(4.4) $$E|S_n|^{2+\delta} \leqslant E\left[\max_{1\leqslant m\leqslant n} |S_m|\right]^{2+\delta} \ll n^{(2+\delta)/2}$$

by (1.1), (4.2), and [18, p. 1231, Theorem B]. (In that Theorem B let $g(n) \equiv Cn$ where $E|S_n|^{2+\delta} \leqslant (Cn)^{(2+\delta)/2}$ using (1.1) and (4.2).)

We will refer to (4.2), (4.3), and (4.4) again when we need to. Now we start the task of constructing the Wiener process $(W(t))$.

We first need to define some parameters and sequences of numbers. First define the function

(4.5) $$a(x,y) \equiv [(2+\delta)/(5+2\delta)]\cdot[(1/2)(1-x) - 2\lambda y]$$

where δ and λ are as in the hypothesis of Theorem 4. Since $\lambda > 1 + 3/\delta$ we get $a(\delta/(2+\delta), \delta/(2+\delta)) < -1$ after some arithmetic. Let z, $0 < z < \delta/(2+\delta)$, be such that $a(x,y) < -1$ $\forall x$, $y \in [z, \delta/(2+\delta)]$. Choose numbers γ, η, τ, and ν such that

(4.6) $$z < \tau < \eta < \gamma < \delta/(2+\delta)$$
$$\nu + a(\gamma, \tau) < -1 \qquad 0 < \nu < 1/2.$$

Define the numbers g_l, h_l, and q_l, $l = 1, 2, \ldots$ as follows:

(4.7) $$g_l = [[\gamma l^{\gamma-1} \exp(l^\gamma)]]$$
$$h_l = [[\eta l^{\eta-1} \exp(l^\eta)]]$$
$$q_l = [(1/2)\gamma l^{\gamma-\nu\gamma-1} - \nu\gamma l^{-\nu\gamma-1}] \exp[(1/2) l^\gamma]$$

where in the definition of g_l and h_l, $[[x]] \equiv$ the greatest integer $\leqslant x$ if $x \geqslant 1$ and (to avoid some trivial technicalities) $[[x]] \equiv 1$ if $x < 1$. The numbers q_l need not be integers, but they are clearly positive. It is easily seen that as $l \to \infty$,

(4.8) $$\sum_{j=1}^{l} g_j \approx \exp(l^\gamma), \qquad \sum_{j=1}^{l} h_j \approx \exp(l^\eta), \qquad \sum_{j=1}^{l} q_j \approx l^{-\nu\gamma} \exp[(1/2) l^\gamma].$$

Now we define blocks G_j and H_j of consecutive positive integers, leaving no gaps between the blocks. The order is $G_1, H_1, G_2, H_2, G_3, H_3, \ldots$ (with $1 \in G_1$). The lengths of the blocks are defined by $\operatorname{card} G_k = g_k$ and $\operatorname{card} H_k = h_k$.

For each $j \geqslant 1$ define the r.v.'s

$$Y_j \equiv \sum_{k \in G(j)} X_k, \qquad Z_j \equiv \sum_{k \in H(j)} X_k$$

and define the positive number $y_j \equiv \operatorname{Var} Y_j$.

Next, enlarging the probability space if necessary, we introduce a sequence $(U_1, U_2, U_3, \dots)$ of independent uniform-[0, 1] r.v.'s, this random sequence being independent of our given sequence (X_k).

By (4.2) and (4.7) there is an integer $J \geqslant 2$ such that $\forall j \geqslant J$, $\| Y_j \|_{2+\delta} \geqslant \| Y_j \|_2 \geqslant q_j$. By Theorem 3 there is, for each $j \geqslant J$, a r.v. Y_j^* which is a measurable function of $(Y_1, \dots, Y_{j-1}, Y_j, U_j)$ such that Y_j^* is independent of $(Y_1, \dots, Y_{j-1})$, has the same distribution as Y_j, and satisfies

$$P(|Y_j^* - Y_j| \geqslant q_j) \leqslant 18(\| Y_j \|_{2+\delta}/q_j)^{(2+\delta)/(5+2\delta)}[\alpha(h_{j-1})]^{(4+2\delta)/(5+2\delta)}. \tag{4.9}$$

For $1 \leqslant j < J$ define Y_j^* as a function of U_j so that it has the same distribution as Y_j. (This is purely a technicality.) It is easily seen that for each $j \geqslant 2$, Y_j^* is independent of $(Y_1^*, \dots, Y_{j-1}^*)$. Hence $(Y_1^*, Y_2^*, Y_3^*, \dots)$ is a sequence of independent r.v.'s.

We will now approximate the sequence $(Y_j^*, j = 1, 2, \dots)$ by a Wiener process by applying Theorem 3.1 on pp. 122–123 of Jain, Jogdeo, and Stout [14]. In that theorem we will take $\alpha = 2$ (and hence $f_\alpha(t) \equiv t \cdot (\log\log t)^{-2}$ there) and the quantity V_n there is given by

$$V_n \equiv y_1 + \cdots + y_n \approx g_1 + \cdots + g_n \approx \exp(n^\gamma) \tag{4.10}$$

by (4.2) and (4.8). For each fixed $c > 0$ one has by (4.4), (4.7), (4.8), and (4.10),

$$\begin{aligned} E(|Y_j^*| \cdot I[Y_j^{*2} \geqslant cf_\alpha(V_j)]) &\leqslant E(|Y_j^*|^{2+\delta}/(cf_\alpha(V_j))^{(1+\delta)/2}) \\ &\ll g_j^{(2+\delta)/2}(f_\alpha(V_j))^{-(1+\delta)/2} \\ &\ll V_j^{1/2} \cdot j^{(\gamma-1)(2+\delta)/2}(\log j)^{1+\delta}, \end{aligned}$$

$$E(Y_j^{*2} \cdot I[Y_j^{*2} \geqslant cf_\alpha(V_j)]) \ll V_j \cdot j^{(\gamma-1)(2+\delta)/2}(\log j)^\delta \qquad \text{(similarly)},$$

$$\begin{aligned} E(Y_j^{*4} \cdot I[Y_j^{*2} \leqslant cf_\alpha(V_j)]) &\leqslant E((cf_\alpha(V_j))^{(2-\delta)/2}|Y_j^*|^{2+\delta}) \\ &\ll (f_\alpha(V_j))^{(2-\delta)/2} g_j^{(2+\delta)/2} \\ &\ll V_j^2 \cdot j^{(\gamma-1)(2+\delta)/2}(\log j)^{\delta-2}. \end{aligned}$$

Since $(\gamma-1)(2+\delta)/2 < -1$ by (4.6), all conditions in [14, Theorem 3.1] (for $\alpha = 2$) are satisfied for our r.v.'s Y_j^*. Hence by that theorem we can redefine the process $(Y_j^*, j \geqslant 1)$ on another probability space, together with a standard Wiener process $(W(t), t \geqslant 0)$ on the same space, such that

$$\left| \sum_{j=1}^n Y_j^* - W(V_n) \right| = o(V_n^{1/2}(\log\log V_n)^{-1/2}) \text{ a.s.} \quad \text{as} \quad n \to \infty. \tag{4.11}$$

Using [3, p. 53, Lemma A1] we redefine the three processes

$$(X_k, -\infty < k < \infty;\ Y_l, Z_l, l \geqslant 1;\ S(t), t \geqslant 0), \qquad (Y_l^*, l = 1, 2, \dots),$$

and $(W(t), t \geqslant 0)$ together on a new probability space so that the joint distribution of the first two processes remains unchanged and the joint distribution of the second and third processes remains unchanged.

Define $r(0) \equiv 0$ and for $l = 1, 2, \ldots$ define the integer $r(l) \equiv \sum_{j=1}^{l}(g_j + h_j)$. For each $t \geqslant 0$ let $l(t)$ denote the integer l such that $r(l) \leqslant t < r(l+1)$. To prove (1.2) in Theorem 4 it suffices to prove that each of statements A–E below is valid; here $V_n \equiv y_1 + \cdots + y_n$ as in (4.10) and (4.11):

A. $$|S(t) - S_{r(l(t))}| = o(t^{1/2}(\log\log t)^{-1/2}) \text{ a.s.}$$

B. $$\left| S_{r(l(t))} - \sum_{j=1}^{l(t)} Y_j \right| = o(t^{1/2}(\log\log t)^{-1/2}) \text{ a.s.}$$

C. $$\left| \sum_{j=1}^{l(t)} Y_j - \sum_{j=1}^{l(t)} Y_j^* \right| = o(t^{1/2}(\log\log t)^{-1/2}) \text{ a.s.}$$

D. $$\left| \sum_{j=1}^{l(t)} Y_j^* - W(V_{l(t)}) \right| = o(t^{1/2}(\log\log t)^{-1/2}) \text{ a.s.}$$

E. $$|W(V_{l(t)}) - W(t)| = o(t^{1/2}(\log\log t)^{-1/2}) \text{ a.s.}$$

In the proofs of these statements we will use the functions $b_\mu(t) \equiv t^{1/2}(\log t)^{-\mu}$ for $t > 1$ and $\mu > 0$, and will sometimes refer to these trivial facts:

$$r(l) \approx \sum_{j=1}^{l} g_j \approx \exp(l^\gamma) \approx V_l \tag{4.12}$$
$$r(l(t)) \leqslant t < r(l(t)+1) \qquad t \approx r(l(t)).$$

Proof of A. It suffices to prove $|S(t) - S_{r(l(t))}| \ll b_\mu(r(l(t)))$ a.s. for some (positive) μ sufficiently small.

For each $l = 1, 2, \ldots$ define the r.v. $T_l \equiv \max\{|S_k - S_{r(l)}| : r(l) \leqslant k \leqslant r(l+1)\}$. It suffices to prove $T_l \ll b_\mu(r(l))$ a.s. as $l \to \infty$.

By stationarity, (4.4), and (4.7) we have

$$ET_l^{2+\delta} \ll [r(l+1) - r(l)]^{(2+\delta)/2} = (g_{l+1} + h_{l+1})^{(2+\delta)/2} \approx g_l^{(2+\delta)/2}.$$

Using (4.7), (4.12), and Markov's inequality we get

$$P[T_l > b_\mu(r(l))] \leqslant ET_l^{2+\delta}/[b_\mu(r(l))]^{2+\delta} \ll l^{(\gamma-1)(2+\delta)/2 + \mu\gamma(2+\delta)}.$$

By (4.6), $(\gamma-1)(2+\delta)/2 + \mu\gamma(2+\delta) < -1$ for sufficiently small μ. By the Borel–Cantelli Lemma, $P(T_l > b_\mu(r(l)) \text{ i.o.}) = 0$ and we have verified Statement A.

Proof of B. It suffices to prove $|\sum_{j=1}^{l} Z_j| \ll b_\mu(r(l))$ a.s. for μ sufficiently small.

Let $\epsilon = (\gamma - \eta)/2$. Then by (4.4), (4.7), and Minkowski's inequality we have

$$\left\| \sum_{j=1}^{l} Z_j \right\|_{2+\delta} \leqslant \sum_{j=1}^{l} \|Z_j\|_{2+\delta} \ll \sum_{j=1}^{l} h_j^{1/2}$$
$$\ll \sum_{j=1}^{l} [(1/2)(\eta+\epsilon) j^{\eta+\epsilon-1} \exp((1/2) j^{\eta+\epsilon})] \approx \exp[(1/2) l^{\eta+\epsilon}]$$

and hence by Markov's inequality,

$$P\left[\left|\sum_{j=1}^{l} Z_j\right| > b_\mu(r(l))\right] \leqslant E\left|\sum_{j=1}^{l} Z_j\right|^{2+\delta} \Big/ [b_\mu(r(l))]^{2+\delta} \ll \exp(-l^\gamma).$$

By the Borel-Cantelli Lemma, $P(|\sum_{j=1}^{l} Z_j| > b_\mu(r(l))\text{ i.o.}) = 0$ and Statement B has been verified.

Proof of C. For each $j \geqslant 1$ define the event $E_j \equiv \{|Y_j^* - Y_j| > q_j\}$. By (4.8) and (4.12) we have $\sum_{j=1}^{l} q_j \ll b_\nu(r(l))$ as $l \to \infty$, so to prove C it suffices to prove $P(E_j\text{ i.o.}) = 0$.

Now $\| Y_j \|_{2+\delta} \ll g_j^{1/2}$ by (4.4) and $\exp(j^\tau) \ll h_{j-1}$ by (4.6) and (4.7), and hence

$$P(E_j) \ll [(g_j^{1/2}/q_j)^{(2+\delta)/(5+2\delta)}] \cdot [j^{\tau(-\lambda)(4+2\delta)/(5+2\delta)}] \ll j^{\nu + a(\gamma,\tau)}$$

by (4.9), (4.7), and (4.5). Now by (4.6) and the Borel-Cantelli Lemma, $P(E_j\text{ i.o.}) = 0$ and Statement C is proved.

Proof of D. Apply (4.10), (4.11), and (4.12).

Proof of E. We will first get a bound on $|t - V_{l(t)}|$. Let $\beta > 0$ satisfy $\beta < \min\{1-\gamma, \gamma(-1+\lambda\delta/(2+\delta))\}$ (both numbers are positive). Then by (4.3), (4.7), (4.8), (4.10), and (4.12),

$$\begin{aligned} |t - V_{l(t)}| &\leqslant |t - r(l(t))| + \left| r(l(t)) - \sum_{j=1}^{l(t)} g_j \right| + \left| \sum_{j=1}^{l(t)} (g_j - y_j) \right| \\ &\ll \left[g_{l(t)+1} + h_{l(t)+1} + \sum_{j=1}^{l(t)} h_j \right] + \left[\sum_{j=1}^{l(t)} g_j (\log g_j)^{1-\lambda\delta/(2+\delta)} \right] \\ &= o(V_{l(t)} (\log V_{l(t)})^{-\beta}). \end{aligned}$$

The rest is routine. For numbers $0 \leqslant a < b < \infty$ define the r.v. $R(a,b) \equiv \max\{|W(u) - W(v)| : a \leqslant u, v \leqslant b\}$. For all positive integers l such that $V_l > 3$ define the numbers $c_l \equiv V_l(1 - (\log V_l)^{-\beta})$ and $d_l \equiv V_l(1 + (\log V_l)^{-\beta})$. To prove Statement E it suffices to prove that with probability 1, $R(c_l, d_l) > b_{\beta/4}(V_l)$ for at most finitely many l. By the Borel-Cantelli Lemma it suffices to prove $P[R(c_l, d_l) > b_{\beta/4}(V_l)] \ll l^{-2}$. But by [14, pp. 121-122, Lemma 2.2] and (4.10),

$$\begin{aligned} P[R(c_l, d_l) > b_{\beta/4}(V_l)] &= P[R(0,1) > 2^{-1/2} (\log V_l)^{\beta/4}] \\ &\ll \exp[-(\log V_l)^{\beta/2}/4] \ll l^{-2} \end{aligned}$$

Statement E is proved, and this completes the proof of Theorem 1. □

ACKNOWLEDGMENTS. W. Philipp acquainted the author with the work of Dehling and Dabrowski, and suggested the use of [14, Theorem 3.1], which considerably strengthened Theorem 4 over an earlier version. G. M. Constantine pointed out that in Lemma 2 an additional assumption that the row- and column-sums of A be zero was superfluous. W. Bryc acquainted the author with his own research and also suggested the use of [19, Theorem 1], which made a longer proof of Lemma 1 unnecessary. The author also benefited from conversations with V. Goodman and L. Tran.

REFERENCES

1. H. C. P. Berbee, *Random walks with stationary increments and renewal theory,* Math. Centre Tracts, 112, Mathematisch Centrum, Amsterdam, 1979.
2. I. Berkes and W. Philipp, *An almost sure invariance principle for the empirical distribution function of mixing random variables,* Z. Wahrsch. Verw. Gebiete 41 (1977/78), 115-137.
3. ———, *Approximation theorems for independent and weakly dependent random vectors,* Ann. Probab. 7 (1979), 29-54.
4. R. C. Bradley, *Central limit theorems under weak dependence,* J. Multivariate Anal. 11 (1981), 1-16.
5. ———, *A sufficient condition for linear growth of variances in a stationary random sequence,* Proc. Amer. Math. Soc. 83 (1981), 586-589.
6. W. Bryc, *Sequences of weakly dependent random variables,* Ph.D. Dissertation. (In Polish), 1981.
7. ———, *On the approximation theorem of I. Berkes and W. Philipp,* Demonstratio Math., to appear.
8. ———, *On almost sure convergence and integrability of sums of weakly dependent random vectors,* Rev. Roumaine Math. Pures Appl., to appear.
9. A. R. Dabrowski, *A note on a theorem of Berkes and Philipp for dependent sequences,* Statistic and Probability Letters 1 (1982), 53-55.
10. Yu. A. Davydov, *Mixing conditions for Markov chains,* Theory Probab. Appl. 18 (1973), 312-328.
11. H. Dehling, *A note on a theorem of Berkes and Philipp,* Z. Wahrsch. Verw. Gebiete 62 (1983), 39-42.
12. I. A. Ibragimov, *A note on the central limit theorem for dependent random variables,* Theory Probab. Appl. 20 (1975), 135-141.
13. I. A. Ibragimov and Yu. V. Linnik, *Independent and stationary sequences of random variables,* Wolters-Noordhoff, Groningen, 1971.
14. N. C. Jain, K. Jogdeo and W. Stout, *Upper and lower functions for martingales and mixing processes,* Ann. Probability 3 (1975), 119-145.
15. J. Kuelbs and W. Philipp, *Almost sure invariance principles for partial sums of mixing B-valued random variables,* Ann. Probab. 8 (1980), 1003-1036.
16. W. Philipp and W. Stout, *Almost sure invariance principles for partial sums of weakly dependent random variables,* Mem. Amer. Math. Soc. 2 (1975), issue 2, no. 161.
17. G. Schwarz, *Finitely determined processes—an indiscrete approach,* J. Math. Anal. Appl. 76 (1980), 146-158.
18. R. J. Serfling, *Moment inequalities for the maximum cumulative sum,* Ann. Math. Statist. 41 (1970), 1227-1234.
19. S. Szarek, *On the best constants in the Khinchin inequality,* Studia Math. 58 (1976), 197-208.

Department of Mathematics
Indiana University
Bloomington, Indiana 47405