

MOMENT CONDITIONS FOR ALMOST SURE CONVERGENCE OF WEAKLY CORRELATED RANDOM VARIABLES

W. BRYC AND W. SMOLENSKI

(Communicated by Lawrence F. Gray)

ABSTRACT. For random sequences with unrestricted maximal correlation coefficient strictly less than 1, sufficient moment conditions for almost sure convergence of a series and for the strong law of large numbers are given.

NOTATION AND RESULTS

Suppose $(\xi_k)_{k \in \mathbb{N}}$ is a real random sequence on a probability space (Ω, \mathcal{M}, P) . For $S \subset \mathbb{N}$ define σ -fields $\mathcal{F}_S = \sigma\{\xi_k : k \in S\}$. Given σ -fields \mathcal{F}, \mathcal{G} in \mathcal{M} let

$$\rho(\mathcal{F}, \mathcal{G}) := \sup\{\text{corr}(V; W) : V \in L_2(\mathcal{F}), W \in L_2(\mathcal{G})\}.$$

Following Bradley [3] for $k \geq 0$ we define the following coefficients of dependence:

$$(1) \quad \tilde{\rho}(k) := \sup\{\rho(\mathcal{F}_S, \mathcal{F}_T)\},$$

where the supremum is taken over all finite subsets $S, T \subset \mathbb{N}$ such that $\text{dist}(S, T) \geq k$. Clearly, $0 \leq \tilde{\rho}(k+1) \leq \tilde{\rho}(k) \leq 1$, $k \geq 0$, and $\tilde{\rho}(0) = 1$.

Definition (1) resembles the definition of the so-called *maximal correlation coefficient*, which is defined by (1) with index sets restricted to subsets S of $[1, n]$ and subsets T of $[n+k, \infty)$, $n, k \in \mathbb{N}$. The maximal correlation coefficient is usually denoted in the literature by $\rho(k)$; Bradley [2, 3] denotes by $\rho^*(r)$ what in our notation is $\tilde{\rho}(r)$. Conditions such as $\tilde{\rho}(k) \rightarrow 0$ (as $k \rightarrow \infty$) have been used in the study of weak limit theorems for random fields, see Bradley [2, 3] and the references therein.

We would like to point out explicitly that the condition $\tilde{\rho}(k) \rightarrow 0$ follows from the following *hypercontractivity* condition.

(H) There exist $q(k) \rightarrow \infty$ (as $k \rightarrow \infty$) such that if $S, T \subset \mathbb{N}$ satisfy $\text{dist}(S, T) \geq k$ then the norm of conditional expectation $E\{\cdot | \mathcal{F}_S\}$ as a linear operator from $L_2(\mathcal{F}_T)$ to $L_{q(k)}(\mathcal{F}_S)$ is 1.

This can be easily seen from Deuschel and Stroock [6, Lemma 5.5.11]; their proof gives $\tilde{\rho}(k) \ll 1/\sqrt{q(k)-1}$ as $k \rightarrow \infty$.

Received by the editors June 27, 1991 and, in revised form, February 26, 1992.

1991 *Mathematics Subject Classification.* Primary 60F15; Secondary 60E15.

Key words and phrases. Maximal correlation, almost surely convergent series, strong law of large numbers.

The purpose of this note is to show that the condition $\lim_{k \rightarrow \infty} \tilde{\rho}(k) < 1$ suffices to get the criterion for almost sure convergence of a random series in terms of the moments of individual random variables. We also prove the strong law of large numbers under a somehow weaker mixing condition. For other almost sure convergence results using different mixing measures of dependence, see Peligrad [9], Stout [11], and the references therein.

Theorem 1. *Suppose $\tilde{\rho}(k) < 1$ for some $k < \infty$. Assume $E\{\xi_k\} = 0$, $E\{\xi_k^2\} = 1$ for all k , and suppose there is $\delta > 0$ such that $\sup_k E\{|\xi_k|^{2+\delta}\} < \infty$. If $\sum_{k=1}^{\infty} a_k^2 < \infty$, then the series $\sum_{k=1}^{\infty} a_k \xi_k$ converges almost surely.*

By Kronecker's lemma, Theorem 1 implies the law of large numbers. However, using Szablowski [12], a stronger result can be obtained, assuming both weaker mixing and weaker moment conditions. The weaker mixing condition is based on the following coefficient. For $k \geq 0$, let

$$\tilde{r}(k) = \sup\{\text{corr}(V; W)\},$$

where the supremum is taken over all finite subsets $S, T \subset \mathbb{N}$ such that $\text{dist}(S, T) \geq k$ and over all linear combinations V of variables $\{\xi_k\}_{k \in S}$ and all linear combinations W of variables $\{\xi_k\}_{k \in T}$.

Clearly, $\tilde{r}(k) \leq \tilde{\rho}(k)$; $\tilde{r}(1) = 0$ iff $\{\xi_k\}$ are uncorrelated; and $\tilde{r}(k)$ is a nonincreasing function of k . For an L_2 -stationary sequence, condition $\tilde{r}(k) \rightarrow 0$ as $k \rightarrow \infty$ implies that $\{\xi_k\}$ has a continuous spectral density; [$\tilde{r}(1) < 1$ and $\tilde{r}(k) \rightarrow 0$] is equivalent to $\{\xi_k\}$ having continuous and positive spectral density, see Bradley [2, Theorem 2]. Spectral density conditions equivalent to either the exponential or polynomial rate of convergence of $\tilde{r}(k)$ to 0 are given in Cheng [5, Theorem 1.1].

The following result resembles the well-known martingale law of large numbers (cf. Feller [7, Chapter VII.8]), improves upon a direct application of Kronecker's lemma to Theorem 1 and extends Szablowski [12, Corollary 9] to the "weakly orthogonal case".

Theorem 2. *If $\tilde{r}(k) < 1$ for some k , $E\{\xi_k\} = 0$, for all k , and*

$$(2) \quad \sum_k k^{-3/2} E\{\xi_k^2\} < \infty,$$

then $\frac{1}{n} \sum_{k=1}^n \xi_k \rightarrow 0$ almost surely.

Remark 1. For simplicity of notation, throughout this note we consider \mathbb{N} -indexed random sequences only. Theorem 1 holds true for the d -dimensional index set \mathbb{Z}^d , $d \geq 1$, too; the only change needed in the proof is to use Moricz [8, Theorem 7] instead of Lemma A below.

Remark 2. Under the assumptions of Theorem 1, one can show that $\sum_{k=1}^{\infty} a_k \xi_k$ converges in $L_{2+\delta}$. If $0 \leq \delta < 1$, this is established in the course of the proof below; some modifications are needed to cover the case when $\delta \geq 1$.

Remark 3. Even in the stationary case, it may happen that $\tilde{\rho}(1) < 1$ while $\lim_{k \rightarrow \infty} \tilde{\rho}(k) \neq 0$; take a gaussian sequence with the noncontinuous, positive, bounded, and bounded away from zero spectral density, see the proof of Theorem 2 in Bradley [2].

PROOFS

Throughout the proofs we use the same symbol C for various constants that may depend on the value of $\tilde{\rho}(1)$; $\|\cdot\|_p$ denotes the L_p -norm.

Let a_1, a_2, \dots be fixed. In what follows $X_k = a_k \xi_k$ and Y_k denotes arbitrary centered $\sigma(\xi_k)$ -measurable random variable, $k \geq 1$. Theorem 1 asserts that $\sum_{k=1}^\infty X_k$ converges almost surely.

Without loss of generality we may assume that (X_k) is such that $\tilde{\rho}(1) < 1$. Indeed, if $\tilde{\rho}(N) < 1$ for some $N > 1$, then $\sum_{k=1}^\infty X_k$ can be written as the sum of N terms of the form $\sum_{k=0}^\infty X_{Nk+j}$, $j = 1, 2, \dots, N$, and the coefficient $\tilde{\rho}(1)$ defined for $(X_{Nk+j})_{k \in \mathbb{N}}$ is dominated by $\tilde{\rho}(N)$.

Lemma A. *If for some $\delta > 0$, $C < \infty$, all $N \geq 1$, and all $k \leq N$,*

$$(3) \quad \mathbb{E} \left\{ \left| \sum_{j=k}^N X_j \right|^{2+\delta} \right\} \leq C \left(\sum_{j=k}^N a_j^2 \right)^{1+\delta/2},$$

then there is $K < \infty$ such that for all $N \geq 1$, $t > 0$,

$$P \left(\max_{k \leq N} \left| \sum_{j=1}^k X_j \right| \geq t \right) \leq \frac{K}{t^{2+\delta}} \left(\sum_{j=1}^N a_j^2 \right)^{1+\delta/2}.$$

This lemma is known, see Billingsley [1, Chapter 2, Theorem 12.2].

The following result is essentially contained in [4].

Lemma 1. *Fix $1 < p < \infty$ and σ -fields \mathcal{F} , \mathcal{F}_* in \mathcal{M} . If $\rho = \rho(\mathcal{F}, \mathcal{F}_*) < 1$ then there is $C = C(p, \rho)$ such that the following implication holds. If X (respectively, X_*) is a centered \mathcal{F} -measurable (respectively, \mathcal{F}_* -measurable) random variable, $\mathbb{E}\{|X|^p\} < \infty$ and $\mathbb{E}\{|X_*|^p\} < \infty$, then there is a random variable Z on the same probability space (Ω, \mathcal{M}, P) such that*

$$(4) \quad X = \mathbb{E}\{Z|\mathcal{F}\} \quad \text{and} \quad X_* = \mathbb{E}\{Z|\mathcal{F}_*\};$$

$$(5) \quad \|Z\|_p \leq C(\|X\|_p + \|X_*\|_p).$$

Remark 4. For the proof of Theorem 1, we shall need only values $1 < p \leq 2$. In this case our proof gives $C(p, \rho) = 2^{2/p}(1 - \rho^{2-2/p})^{-1}$.

Proof. Suppose that $1 < p \leq 2$. Let A be a linear operator defined by $A(\cdot) = \mathbb{E}\{\mathbb{E}\{\cdot|\mathcal{F}_*\}|\mathcal{F}\}$ and let $E(\cdot)$ denote the linear operator of taking the expected value $\mathbb{E}\{\cdot\}$. Clearly, $\|A - E\|_{L_2 \rightarrow L_2} \leq \rho$ and $\|A^k - E\|_{L_1 \rightarrow L_1} \leq 2$. Furthermore, it is easy to check that for $k \geq 1$ we have $(A - E)^k = A^k - E$. Therefore, for $k \geq 1$, the Riesz interpolation theorem (here we use the real version of the Riesz-Thörin theorem, see [10, §22]) implies

$$\|A^k - E\|_{L_p \rightarrow L_p} \leq (\|(A - E)^k\|_{L_2 \rightarrow L_2})^{2-2/p} (\|A^k - E\|_{L_1 \rightarrow L_1})^{2/p-1} \leq \rho^{2k(1-1/p)} 2^{2/p-1}.$$

In particular, for centered Y we have $\|A^k Y\|_p \leq 2^{2/p-1} \rho^{2k(1-1/p)} \|Y\|_p$. Since the estimate also holds true for $k = 0$ and, by symmetry, also for $A^*(Y) = \mathbb{E}\{\mathbb{E}\{Y|\mathcal{F}\}|\mathcal{F}_*\}$,

$$Z = \sum_{k=0}^\infty A^k (X - \mathbb{E}\{X_*|\mathcal{F}\}) + \sum_{k=0}^\infty (A^*)^k (X_* - \mathbb{E}\{X|\mathcal{F}_*\})$$

is well defined and (5) holds by the triangle inequality. By a simple computation (4) holds. This ends the proof in the case $1 < p \leq 2$. If $2 < p < \infty$ the only change needed in the proof is to use $\|A^k - E\|_{L_\infty \rightarrow L_\infty} \leq 2$ for the Riesz-Thörin interpolation theorem.

The following result is related to Lemma 1 and Lemma 5 in Bradley [2].

Lemma 2. *Let $\tilde{p} := \tilde{p}(1) < 1$, $q \geq 1$, be fixed. There is $C = C(q, \tilde{p}) < \infty$, which depends on \tilde{p} and q only such that if Y_j are centered $\sigma(\xi_j)$ -measurable with finite q th moments, $j \geq 1$, then for all $N \geq 1$ and all $k \leq N$,*

$$(6) \quad E \left\{ \left| \sum_{j=k}^N Y_j \right|^q \right\} \leq CE \left\{ \left(\sum_{j=k}^N Y_j^2 \right)^{q/2} \right\}.$$

Proof. Fix $k \leq N$. Let (ε_j) be i.i.d. random variables independent of (Y_j) with $P(\varepsilon_j = \pm 1) = 1/2$. We claim that there is $K = K(q, \tilde{p}) < \infty$ such that

$$(7) \quad E \left\{ \left| \sum_{j=k}^N Y_j \right|^q \right\} \leq KE \left\{ \left| \sum_{j=k}^N \varepsilon_j Y_j \right|^q \right\}.$$

Clearly, (6) follows from (7) by the Khinchine inequality.

To prove (7), following Bradley [2] we consider random variables S, S_* on the product probability space $2^{\{k, k+1, \dots, N\}} \times \Omega$ with the product measure $P_1(Q \times A) = P(A)/2^{N-k+1}$, $Q \subset \{k, k+1, \dots, N\}$; S and S_* are defined by

$$S(Q, \omega) = S_Q(\omega) = \sum_{j \in Q} Y_j(\omega), \quad S_*(Q, \omega) = S_{Q^*}(\omega) = \sum_{j \in Q^*} Y_j(\omega),$$

where $Q^* = \{k, k+1, \dots, N\} \setminus Q$.

Clearly, for each Q and every $\omega \in \Omega$ we have $S_Q + S_{Q^*} = \sum_{j=k}^N Y_j$. In particular,

$$(8) \quad \left\| \sum_{j=k}^N Y_j \right\|_q \leq \|S_Q\|_q + \|S_{Q^*}\|_q.$$

Moreover, by a well-known consequence of the Hahn-Banach theorem applied to each of the terms on the right-hand side of (8) separately, we get

$$(9) \quad \|S_Q\|_q + \|S_{Q^*}\|_q = E\{YS_Q\} + E\{Y_*S_{Q^*}\},$$

where Y (respectively, Y_*) is \mathcal{F}_Q -measurable (respectively, \mathcal{F}_{Q^*} -measurable) and

$$E\{|Y_*|^p\} = E\{|Y|^p\} = 1, \quad 1/p + 1/q = 1.$$

Since by assumption $E\{S_Q\} = 0$, we can replace Y and Y_* in (9) by centered variables

$$X = Y - E\{Y\}, \quad X_* = Y_* - E\{Y_*\}.$$

For each fixed Q , by Lemma 1 applied to σ -fields $\mathcal{F} = \mathcal{F}_Q, \mathcal{F}_* = \mathcal{F}_{Q^*}$, and random variables $X, -X_*$, there is Z such that

$$\|Z\|_p \leq 4C(p, \tilde{p}), \quad E\{Z|\mathcal{F}_Q\} = X, \quad E\{Z|\mathcal{F}_{Q^*}\} = -X_*.$$

Therefore

$$\begin{aligned} E\{YS_Q\} + E\{Y_*S_{Q^*}\} &= E\{XS_Q\} - E\{(-X_*)S_{Q^*}\} = E\{ZS_Q\} - E\{ZS_{Q^*}\} \\ &\leq \|Z\|_p \|S_Q - S_{Q^*}\|_q. \end{aligned}$$

This shows that there is $K = (4C(p, \tilde{\rho}))^q$ such that for all $Q \subset \{k, k + 1, \dots, N\}$ we have

$$E \left\{ \left| \sum_{j=k}^N Y_j \right|^q \right\} \leq KE\{|S_Q - S_{Q^*}|^q\}.$$

Since in distribution $S - S_* \simeq \sum_{j=k}^N \varepsilon_j Y_j$, averaging the last inequality over all subsets $Q \subset \{k, k + 1, \dots, N\}$ gives (7).

Lemma 3. *Let $2 \leq q \leq 4$ be fixed. If $\tilde{\rho} := \tilde{\rho}(1) < 1$ then there is $C = C(q, \tilde{\rho}) < \infty$, which depends on $\tilde{\rho}$ and q only and such that for all $N \geq 1$ and all $k \leq N$*

$$(10) \quad E \left\{ \left| \sum_{j=k}^N X_j \right|^q \right\} \leq C \left(\left(\sum_{j=k}^N E\{X_j^2\} \right)^{q/2} + \sum_{j=k}^N E\{|X_j|^q\} \right).$$

Proof. Fix $k \leq N$. Let $U_j = X_j^2 - E\{X_j^2\}$. By the triangle inequality

$$\left| \sum_{j=k}^N X_j^2 \right| \leq \left| \sum_{j=k}^N U_j \right| + \sum_{j=k}^N E\{X_j^2\},$$

hence (6) applied to $Y_j = X_j$ gives

$$(11) \quad E \left\{ \left| \sum_{j=k}^N X_j \right|^q \right\} \leq C' \left(E \left\{ \left| \sum_{j=k}^N U_j \right|^{q/2} \right\} + \left(\sum_{j=k}^N E\{X_j^2\} \right)^{q/2} \right).$$

By (6) applied to $Y_j = U_j$, we have

$$E \left\{ \left| \sum_{j=k}^N U_j \right|^{q/2} \right\} \leq C'' E \left\{ \left(\sum_{j=k}^N U_j^2 \right)^{q/4} \right\}.$$

Since $q \leq 4$, this and (11) imply

$$E \left\{ \left| \sum_{j=k}^N X_j \right|^q \right\} \leq CE \left\{ \sum_{j=k}^N |U_j|^{q/2} \right\} + C \left(\sum_{j=k}^N E\{X_j^2\} \right)^{q/2}.$$

The trivial inequality $\|U_j\|_{q/2} \leq 2(\|X_j\|_q)^2$ now ends the proof of (10).

Proof of Theorem 1. Without loss of generality we may assume $0 < \delta \leq 2$. By Lemma 3 with $q = 2 + \delta$, there is C such that for all $N \geq 1$ and all $k \leq N$

$$(12) \quad E \left\{ \left| \sum_{j=k}^N X_j \right|^{2+\delta} \right\} \leq C \left(\sum_{j=k}^N E\{X_j^2\} \right)^{1+\delta/2} + C \sum_{j=k}^N E\{|X_j|^{2+\delta}\}.$$

From (12) we easily get (3). Indeed, $E\{X_j^2\} = a_j^2$ and

$$\sum_{j=k}^N E\{|X_j|^{2+\delta}\} \leq \sup_j E\{|\xi_j|^{2+\delta}\} \sum_{j=k}^N |a_j|^{2+\delta} \leq C \left(\sum_{j=k}^N a_j^2 \right)^{1+\delta/2}.$$

By the definition of almost sure convergence, Theorem 1 now follows from Lemma A.

Proof of Theorem 2. The result is an immediate consequence of Theorem 4 in Szablowski [12] used together with Lemma 1 in Bradley [2]. Since in our case there is no need for using the full scope of Szablowski's theory, a short proof based on his ideas is given below.

Proof. As in the proof of Theorem 1, without loss of generality we may assume $\tilde{r} = \tilde{r}(1) < 1$. Denote $S_n = \sum_{k=1}^n \xi_k$, $\bar{X}_n = \frac{1}{n} S_n$. By Lemma 1 in Bradley [2], we have

$$(13) \quad E\{(S_n)^2\} \leq \frac{1+\tilde{r}}{1-\tilde{r}} \sum_{k=1}^n E\{(\xi_k)^2\}.$$

This implies that $\bar{X}_n \rightarrow 0$ as $n \rightarrow \infty$ in L_2 and in probability. Indeed, $E\{(\bar{X}_n)^2\} \leq C(\frac{1}{n})^2 \sum_{k=1}^n E\{(\xi_k)^2\} \rightarrow 0$ by (2). To prove the theorem, it suffices therefore to verify that $(\bar{X}_n)^2$ converges with probability 1. To this end, squaring the trivial recurrence $\bar{X}_{n+1} = n\bar{X}_n/(n+1) + \xi_{n+1}/(n+1)$ we get

$$(14) \quad (\bar{X}_{n+1})^2 - (\bar{X}_n)^2 \leq \left(\frac{1}{n+1}\right)^2 (\xi_{n+1})^2 + 2\left(\frac{1}{n+1}\right)^2 |S_n| |\xi_{n+1}| \\ = A_n \quad (\text{say}).$$

From (2) we get

$$\sum_n \left(\frac{1}{n+1}\right)^2 E\{(\xi_{n+1})^2\} < \infty,$$

hence

$$\sum_n \left(\frac{1}{n+1}\right)^2 (\xi_{n+1})^2$$

converges a.s. Also it is easy to see that by the trivial inequality $2|ab| \leq a^2 + b^2$ and (13)

$$\sum_n 2\left(\frac{1}{n+1}\right)^2 E\{|S_n| |\xi_{n+1}|\} \leq \sum_n n^{-3/2} E\{(\xi_n)^2\} + \frac{1+\tilde{r}}{1-\tilde{r}} \sum_n n^{-5/2} \sum_{k=1}^n E\{(\xi_k)^2\}.$$

Since

$$\sum_n n^{-5/2} \sum_{k=1}^n E\{(\xi_k)^2\} \leq C \sum_k k^{-3/2} E\{(\xi_k)^2\},$$

thus by (2) the series $\sum_n A_n$ converges a.s. From (14) we have

$$(\bar{X}_{n+m})^2 - (\bar{X}_n)^2 \leq \sum_{k=n}^{n+m} A_k.$$

Therefore $\limsup_{n \rightarrow \infty} (\bar{X}_n)^2 \leq \liminf_{n \rightarrow \infty} (\bar{X}_n)^2$ a.s., which concludes the proof.

ACKNOWLEDGMENT

The authors would like to thank R. C. Bradley for information about his research and for helpful discussion, in particular, for Remark 3. We would like also to thank the referee for his careful reading of the manuscript and M. Peligrad and P. Szablowski for helpful comments.

REFERENCES

1. P. Billingsley, *Convergence of probability measures*, Wiley, New York, 1968.
2. R. C. Bradley, *On the spectral density and asymptotic normality of weakly dependent random fields*, J. Theoret. Probab. **5** (1992), 355–374.
3. ———, *Equivalent mixing conditions for random fields*, Technical Report No. 336, Center for Stochastic Processes, Univ. of North Carolina, Chapel Hill, 1990.
4. W. Bryc, *Conditional expectation with respect to dependent σ -fields*, Proc. Seventh Conf. on Probability Theory; Brasov (M. Iosifescu, ed.), Editura Acad. Rep. Soc. Romania, Bucurestic, 1984, pp. 409–411.
5. R. Cheng, *On the rate of strong mixing in stationary Gaussian fields*, Studia Math. **101** (1992), 183–192.
6. J. D. Deuschel and D. W. Stroock, *Large deviations*, Pure Appl. Math., vol. 137, Academic Press, Boston, MA, 1989.
7. W. Feller, *An introduction to probability theory and its applications*, Wiley, New York, 1966.
8. F. Moricz, *Moment inequalities for the maximum of partial sums of random fields*, Acta Sci. Math. **39** (1977), 353–366.
9. M. Peligrad, *The r -quick version of the strong law for stationary ϕ -mixing sequences*, Proc. Internat. Conf. on Almost Everywhere Convergence in Probab. and Statist., Academic Press, New York, 1989, pp. 335–348.
10. M. Riesz, *sur les maxima des formes bilinéaires et sur les fonctionelles linéaires*, Acta Math. **49** (1926), 465–497.
11. W. F. Stout, *On convergence of ϕ -mixing sequences of random variables*, Z. Wahrsch. Verw. Gebiete **31** (1974), 69–70.
12. P. J. Szablowski, *Generalized laws of large numbers and auxiliary results concerning stochastic approximation with dependent disturbances. II*, Comm. Math. Appl. **13** (1987), 973–987.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CINCINNATI, CINCINNATI, OHIO 45221-0025
E-mail address: BRYC@UCBEH.SAN.UC.EDU

DEPARTMENT OF MATHEMATICS, THE UNIVERSITY OF KANSAS, LAWRENCE, KANSAS 66045-2142
E-mail address: SMOLENSK@KUHUB.CC.UKAN.EDU