

## Convergence Rates in the Strong Law for Bounded Mixing Sequences

Henry Berbee

Stichting Mathematisch Centrum, Kruislaan 413, 1098 SJ Amsterdam,  
The Netherlands

**Summary.** Speed of convergence is studied for a Marcinkiewicz-Zygmund strong law for partial sums of bounded dependent random variables under conditions on their mixing rate. Though  $\alpha$ -mixing is also considered, the most interesting result concerns absolutely regular sequences. The results are applied to renewal theory to show that some of the estimates obtained by other authors on coupling are best possible. Another application sharpens a result for averaging a function along a random walk.

### 1. Introduction

Suppose  $S_n$  are the partial sums of i.i.d. random variables with distribution  $F$ . Baum and Katz (1965) showed that if  $\alpha p \geq 1$ ,  $\alpha > \frac{1}{2}$

$$\sum_{n \geq 1} n^{\alpha p - 2} P(|S_n| \geq \varepsilon n^\alpha) < \infty \quad \text{for } \varepsilon > 0 \tag{1.1}$$

if and only if  $F$  has finite  $p^{\text{th}}$  moment and moreover in case  $\alpha \leq 1$  has vanishing mean.

We consider  $S_n = \eta_1 + \dots + \eta_n$ ,  $n \geq 1$ , the partial sums of random variables bounded by 1 with vanishing mean. We want to prove (1.1) under a sharp condition on the mixing rate. Let  $\mathcal{F}_n^- = \sigma(\eta_i, i \leq n)$  and  $\mathcal{F}_n^+ = \sigma(\eta_i, i \geq n)$ . Define for  $n \geq 0$

$$\alpha_n = \sup_{k \geq 1} \sup_{A \in \mathcal{F}_k^-, B \in \mathcal{F}_{n+k}^+} |P(AB) - P(A)P(B)|, \tag{1.2}$$

$$\beta_n = \sup_{k \geq 1} E \sup_{A \in \mathcal{F}_k^-} |P(A | \mathcal{F}_{n+k}^+) - P(A)|. \tag{1.3}$$

The process  $(\eta_i)_{i \geq 1}$  is called *strong mixing* if  $\alpha_n \downarrow 0$  and *absolutely regular* (or *weak Bernoulli*) if  $\beta_n \downarrow 0$ . Because  $\alpha_n \leq \beta_n$  the latter concept is stronger.

Our results deal mainly with absolutely regular processes. This leaves open a wide range of applications because e.g. an aperiodic, positive recurrent Markov chain is absolutely regular. Berbee (1984) studies  $\beta_n$ ,  $n \rightarrow \infty$ , for stationary processes.

The main result, given below is a convergence theorem for dependent random variables. It is interesting because it is a sharp result. The reader will notice that it is a rate of convergence result for the ergodic theorem. The theorem is valid also for non-stationary sequences.

**Theorem 1.1.** *Suppose  $(\eta_i)_{i \geq 1}$  is a sequence of random variables bounded by 1 with vanishing mean. Let  $\frac{1}{2} < \alpha \leq 1$  and  $\alpha p \geq 1$ . If for  $p \geq 1$*

$$\sum_{n \geq 1} n^{p-2} \beta_n < \infty \tag{1.4}$$

then

$$\sum_{n \geq 1} n^{\alpha p - 2} P(\max_{k \leq n} |S_k| \geq \varepsilon n^\alpha) < \infty \quad \text{for } \varepsilon > 0. \tag{1.5}$$

To judge the mixing rate (1.4) in this theorem we consider in Sect. 6 a special class of processes that is still rich enough to discriminate effectively between mixing rates. We consider a stationary renewal process on the positive integers, aperiodic with finite mean. Let  $\tilde{S}_n$  be the number of renewals in  $\{1, \dots, n\}$  and take  $S_n := \tilde{S}_n - E\tilde{S}_n$ . Then  $(\eta_i)$  is a stationary sequence of random variables bounded by 1 with vanishing mean. We show for this class of processes that (1.4) and (1.5) are equivalent. Let  $F$  be the distribution of the increments of the renewal process. We also show that (1.4) is equivalent to the finiteness of the  $p^{\text{th}}$  moment of  $F$  in case  $p > 1$ . It will appear that these results are closely related to the theorem in Baum and Katz (1965), quoted above.

For stationary, strong mixing sequences Lai (1977) obtained earlier results related to Theorem 1.1. The main interest of that paper is to prove under certain conditions the equivalence between (1.5) and finiteness of the  $p^{\text{th}}$  moment of  $\eta_i$ . If the results of Lai (1977) are applied to our case with bounded random variables, then to get (1.5) stronger conditions are imposed on the mixing rate  $\alpha_n$  than we impose on  $\beta_n$ .

Property (1.5) describes speed of convergence in the strong law. It implies in case  $\alpha p > 1$

$$\sum_{n \geq 1} n^{\alpha p - 2} P\left(\sup_{k \geq n} \frac{1}{k^\alpha} |S_k| \geq \varepsilon\right) < \infty \tag{1.6}$$

by Lemma 4 in Lai (1977). Because the probabilities above are non-increasing in  $n$  we may infer

$$P\left(\sup_{k \geq n} \frac{1}{k^\alpha} |S_k| > \varepsilon\right) = o\left(\frac{1}{n^{\alpha p - 1}}\right).$$

The proof of Theorem 1.1 bears some resemblance to ideas used by Baum and Katz (1965). In that paper the increments of  $S_n$  are truncated at height  $n^\alpha$  and then the Markov inequality is applied to get (1.1). We do not truncate but replace  $S_n$  by sums of independent random variables of size  $o(n^\alpha)$ . At this point we use the absolute regularity to estimate the error. Lemma 2.1 is used for this purpose. Then we apply the Markov inequality and finally we use bounds for  $E(S_n)^{2m}$  to obtain (1.5). These bounds are derived in Sect. 3 and are valid for strong mixing sequences. They are closely related to similar bounds in Yokoyama (1980).

Strong mixing is easier to formulate but technically less nice because the coupling argument of Sect. 2 is not available (in the same way). Therefore we

only discuss averaging, i.e. (1.6) with  $\alpha=1$ . We could also get results for  $\alpha < 1$  but they may not be sharp and we do not present these.

**Theorem 1.2.** *Suppose  $(\eta_i)_{i \geq 1}$  is a sequence of random variables bounded by 1 with vanishing mean. If for  $p=1$*

$$\sum_{n \geq 1} n^{p-2} \alpha_n < \infty \tag{1.7}$$

*then the strong law  $\frac{1}{n} S_n \rightarrow 0$  holds a.s. and if  $p > 1$  then (1.7) implies moreover*

$$\sum_{n \geq 1} n^{p-2} P \left( \sup_{k \geq n} \frac{1}{k} |S_k| > \varepsilon \right) < \infty. \tag{1.8}$$

Averaging as is done in (1.8) is a much easier procedure than comparison with  $n^\alpha$  with  $\frac{1}{2} < \alpha < 1$ . The proof of the last theorem is therefore much simpler because we do not use the trick of replacing  $S_n$  by sums of independent random variables of  $o(n)$  (the analogue of truncation). Etemadi (1981) gave an elementary proof of the strong law (so for averaging) that does not need Kolmogorov's inequality.

## 2. A Dependence Coefficient

The total variation of a signed measure  $\nu$  is defined as

$$\|\nu\| = \sup \nu(A) - \nu(A^c)$$

where the supremum is taken over all measurable sets  $A$ .

Suppose  $X$  and  $Y$  are real random variables (or vectors). Define their dependence as

$$\beta(X, Y) = \frac{1}{2} \|P_{X,Y} - P_X \times P_Y\|.$$

This coefficient vanishes if and only if  $X$  and  $Y$  are independent. We can rewrite

$$\beta(X, Y) = E \frac{1}{2} \|P_{Y|X} - P_Y\|$$

where  $P_{X|Y}$  is the conditional distribution of  $X$  given  $Y$ . Hence if  $Y'$  is  $Y$ -measurable

$$\beta(X, Y') \leq \beta(X, Y) \tag{2.1}$$

(see e.g. Schwarz (1980)). Also (1.3) is equivalent to

$$\beta_n = \sup_{k \geq 1} \beta((\eta_i, i \leq k), (\eta_i, i \geq n+k)).$$

In case  $(\eta_i)$  is stationary and is extended to be defined for all  $i \in \mathbb{Z}$  the supremum is increasing by (2.1) as  $k \rightarrow \infty$  with limit

$$\beta_n = \beta((\eta_i, i \leq 0), (\eta_i, i \geq n)). \tag{2.2}$$

We use the following technical lemma.

**Lemma 2.1.** Let  $X_1, \dots, X_n$  be real random variables on the same probability space. Define for  $1 \leq i < n$

$$\beta^{(i)} = \beta(X_i, (X_{i+1}, \dots, X_n)).$$

The probability space can be extended with random variables  $\tilde{X}_i$  distributed as  $X_i$  such that  $\tilde{X}_1, \dots, \tilde{X}_n$  are independent and

$$P(\tilde{X}_j \neq X_j \text{ for some } j) \leq \beta^{(1)} + \dots + \beta^{(n-1)}.$$

*Proof.* Write  $X = (X_1, \dots, X_n)$  and write  $\tilde{X} = (\tilde{X}_1, \dots, \tilde{X}_n)$  for the random vector to be constructed. The distributions of these random variables are known. It is well-known (see e.g. Schwarz 1980) that one can construct  $\tilde{X}$  as in the statement of the lemma such that

$$P(X \neq \tilde{X}) = \frac{1}{2} \|P_X - P_{\tilde{X}}\|.$$

To get (2.2) we have to estimate the right hand side.

If  $X, Y$  and  $\tilde{Y}$  are random variables, with  $Y$  and  $\tilde{Y}$  having values in the same space, then

$$\begin{aligned} \|P_{X,Y} - P_{X,\tilde{Y}}\| &\leq \|P_{X,Y} - P_X \times P_Y\| + \|P_X \times P_Y - P_X \times P_{\tilde{Y}}\| \\ &= 2\beta(X, Y) + \|P_Y - P_{\tilde{Y}}\| \end{aligned}$$

as one easily observes. Applying this rule successively one obtains

$$\begin{aligned} &\frac{1}{2} \|P_{X_1, \dots, X_n} - P_{X_1} \times \dots \times P_{X_n}\| \\ &\leq \beta^{(1)} + \|P_{X_2, \dots, X_n} - P_{X_2} \times \dots \times P_{X_n}\| \\ &\leq \dots \leq \beta^{(1)} + \dots + \beta^{(n-1)}. \quad \square \end{aligned}$$

To obtain Theorem 1.2 we want to apply Lemma 2.1 with

$$X_i = \sum_{j=k_{2i}+1}^{k_{2i+1}} \eta_j, \quad i=1, \dots, m_n,$$

where  $k_{i+1} - k_i = M_n$  for all  $i$ . So we apply the lemma on block sums with indices in different blocks separated by distance of at least  $M_n$ . We also want that  $m_n M_n$  is about  $n$ . To apply the lemma successfully we need that  $m_n \beta_{M_n}$  is sufficiently small. This is guaranteed by the following lemma.

**Lemma 2.2.** Suppose that  $\sum_{n \geq 1} n^{p-2} \beta_n < \infty$  for some given  $p \geq 1$ . Let  $\frac{1}{2} < \alpha \leq 1$ .

Then there exists  $M_n = o(n^\alpha)$  such that if  $m_n = \left\lfloor \frac{n}{M_n} \right\rfloor$

$$\sum_{n \geq 1} n^{\alpha p - 2} m_n \beta_{M_n} < \infty. \tag{2.3}$$

*Proof.* Observe that for each  $j \geq 1$

$$j^p \sum_{n \geq k} n^{p-2} \beta_n \rightarrow 0 \quad \text{as } k \rightarrow \infty.$$

Choose a sequence  $l = k_1 < k_2 < \dots$  such that

$$\sum_{j \geq 1} j^p \sum_{n \geq \lceil \frac{1}{j} k_j^2 \rceil} n^{p-2} \beta_n < \infty. \quad (2.4)$$

Define

$$M_n = \left\lfloor \frac{1}{j} n^2 \right\rfloor \quad \text{for } k_j \leq n < k_{j+1}, \quad j \geq 1.$$

Because  $M_n \leq \frac{n^2}{j}$  if  $n \geq k_j$ , we have  $M_n = o(n^2)$ . Observe that (2.3) is bounded by

$$\sum_{j \geq 1} \sum_{n \geq k_j} n^{2p-2} \left\lfloor \frac{n}{\left\lfloor \frac{1}{j} n^2 \right\rfloor} \right\rfloor \beta_{\left\lfloor \frac{1}{j} n^2 \right\rfloor}. \quad (2.5)$$

To prove that this expression is finite observe first that for  $m, j \geq 1$

$$\begin{aligned} \# \left\{ n: \left\lfloor \frac{1}{j} n^2 \right\rfloor = m \right\} &\leq ((m+1)j)^{\frac{1}{2}} - ((m-1)j)^{\frac{1}{2}} \\ &\leq j^{\frac{1}{2}} \int_{m-1}^{m+1} \frac{1}{x^{\frac{3}{2}}} dx \leq c_2 j^{\frac{1}{2}} m^{\frac{1}{2}-1} \end{aligned}$$

for some constant  $c_2$ . If  $\left\lfloor \frac{1}{j} n^2 \right\rfloor = m$  then  $(mj)^{\frac{1}{2}} \leq n \leq 4(mj)^{\frac{1}{2}}$ . Hence (2.5) is for a suitable constant  $c'_2$  bounded by

$$\begin{aligned} c'_2 \sum_{j \geq 1} \sum_{m \geq \lceil \frac{1}{j} k_j^2 \rceil} (mj)^{\frac{2p-2}{2}} \frac{(mj)^{\frac{1}{2}}}{m} \beta_m (c_2 j^{\frac{1}{2}} m^{\frac{1}{2}-1}) \\ = c_2 c'_2 \sum_{j \geq 1} j^p \sum_{m \geq \lceil \frac{1}{j} k_j^2 \rceil} m^{p-2} \beta_m < \infty \end{aligned}$$

by (2.4).  $\square$

### 3. Moment Bounds

In this section we assume that  $(\eta_i)_{i \geq 1}$  is a sequence of random variables bounded by 1 with vanishing mean such that for  $p \geq 1$

$$\sum_{n \geq 1} n^{p-2} \alpha_n < \infty \quad (3.1)$$

where  $(\alpha_n)$  is defined by (1.2). For  $1 \leq p < 2$  we use the estimate

**Lemma 3.1.**

$$E(S_n)^2 \leq 4n \sum_{k=0}^{n-1} \alpha_k, \quad n \geq 1. \quad (3.2)$$

*Proof.* By Ibragimov and Linnik (1971), Th. 17.2.1

$$|E(\eta_i \eta_j)| \leq 4\alpha(|j-i|).$$

After summing over  $1 \leq i, j \leq n$  we easily obtain the inequality.  $\square$

For  $p \geq 2$  we need bounds on higher moments of  $S_n$ . Yokoyama (1980) discusses this subject for stationary sequences. For bounded sequences he proves

$$E|S_n|^{2(p-1)} = O(n^{p-1}), \quad p > 2. \tag{3.3}$$

We need bounds for  $E(S_n)^{2m}$ , for integer  $m \geq 1$ . To be able to obtain Theorem 1.1 we need a bound that is more precise than (3.3) in the way it depends on the mixing coefficients, like in (3.2).

**Lemma 3.2.** *If  $p \geq 2$  then for each integer  $m \geq 1$  there is a finite constant  $c$  depending on the mixing coefficients (1.2) and  $m$  such that*

$$E(S_n)^{2m} \leq c \max \left( n^{m+(m-p)+}, n \sum_{k=0}^{n-1} (k+1)^{2m-2} \alpha_k \right), \quad n \geq 1. \tag{3.4}$$

Note that if  $2m \leq p$  then the second term at the right is  $O(n)$  by (3.1). For larger  $m \geq 1$  this term is  $O(n^{2m-p+1})$  by (3.1). So for large  $m$ , and in fact  $m \geq p-1$ , the second term is more important than the first term. Thus we obtain

**Corollary 3.3.** *If  $p \geq 2$  then*

$$\begin{aligned} E(S_n)^{2m} &= O(n^{2m-p+1}) && \text{if } m \geq p-1, \\ &= O(n^m) && \text{if } 1 \leq m \leq p-1, \end{aligned}$$

and the same estimates hold for  $E(S_{j+n} - S_j)^{2m}$ , uniformly in  $j$ .

*Proof of Lemma 3.2.* As in the corresponding part of Yokoyama (1980) our proof is similar to Sen (1974).

We denote by  $\sum_{n,j}$  the summation over all  $1 \leq i_1 \leq \dots \leq i_j \leq n$ . Let

$$b_k(n, j) = \sum_{n,j} |E(\eta_{k+1} \eta_{k+i_1} \dots \eta_{k+i_j})|,$$

$$b(n, j) = \sup_{k \geq 0} b_k(n, j).$$

Note that

$$\begin{aligned} E(S_n)^{2m} &\leq (2m)! \sum_{n, 2m} |E(\eta_{i_1} \dots \eta_{i_{2m}})| \\ &= (2m)! \sum_{i_1=1}^n b_{i_1-1}(n-i_1+1, 2m-1) \\ &\leq (2m)! nb(n, 2m-1). \end{aligned}$$

Define  $r_j = i_j - i_{j-1}$  (with  $i_0 = 0$ ) and let  $\sum_{n,j}^{(h)} 1 \leq h \leq j$ , be the components of  $\sum_{n,j}$  for which  $r_h = \max \{r_1, \dots, r_j\}$ .

Our aim is to find a bound on  $b$ . By Ibragimov (1962), Lemma 1.2

$$nb(n, 1) = n \sup_{k \geq 0} \sum_{n, 1} |E(\eta_{k+1} \eta_{k+i_1})| \leq 4n \sum_{k=0}^{n-1} \alpha_k. \tag{3.6}$$

We want to prove by induction on  $j \geq 2$  that there are finite constants  $c$  and  $c'$  depending on  $j$  and the mixing coefficients (1.2) such that

$$nb(n, j-1) \leq c \max \left( n^{j^*}, n \sum_{k=0}^{n-1} (k+1)^{j-2} \alpha_k \right) \leq c' n^{j^{**}} \tag{3.7}$$

where

$$j^* = \max \left( \left[ \frac{j}{2} \right], j-p \right)$$

$$j^{**} = \max \left( \left[ \frac{j}{2} \right], j-p+1 \right).$$

If this is proved then the assertion of the lemma follows by applying the first inequality of (3.7) with  $j=2m$  in (3.5).

Let us note first that the second inequality of (3.7) is obvious, following in the same way as in the proof of Corollary 3.3, i.e. by applying (3.1). So we are only concerned with the first inequality in the induction proof.

For  $j=2$  the first inequality of (3.7) follows from (3.6). Suppose now that (3.7) holds for all  $j \leq m$  with  $m \geq 2$  arbitrary. It suffices to prove the first inequality of (3.7) for  $j=m+1$ . Note

$$nb_k(n, m) = \sum_{h=1}^m \left[ n \sum_{n, m}^{(h)} |E(\eta_{k+1} \eta_{k+i_1} \dots \eta_{k+i_m})| \right].$$

We consider the term between  $[ \cdot ]$  separately. The total number of these terms is  $m$  and so does not depend on  $n$ . Again using Ibragimov and Linnik (1971), Th. 17.2.1 the term is bounded by

$$n \sum_{n, m}^{(h)} |E(\eta_{k+1} \eta_{k+i_1} \dots \eta_{k+i_{h-1}}) E(\eta_{k+i_h} \dots \eta_{k+i_m})| \tag{3.8}$$

$$+ n \sum_{n, m}^{(h)} 4\alpha(r_h). \tag{3.9}$$

Note that (3.9) is bounded by

$$n \sum_{r_h=0}^{n-1} (r_h+1)^{m-1} 4\alpha(r_h)$$

in accordance with the induction assertion.

We still have to consider (3.8). It vanishes for  $h=1, m$ . For other values of  $h$  it is bounded by

$$n \sum_{i=1}^n b_k(i, h-1) b_{i+k-1}(n-i+1, m-h)$$

with  $i = i_h$ . By the induction assumption (3.8) is bounded by

$$nc \sum_{i=1}^n i^{h^{**}-1} (n-i+1)^{(m-h+1)^{**}-1}$$

where  $c$  depends on  $m$  and (1.2). Because  $2 \leq h \leq m-1$  all powers are non-negative and so (3.8) is bounded by

$$\begin{aligned} n \cdot c \cdot n^{h^{**}-1} n^{(m-h+1)^{**}-1} \\ = c \cdot n^{h^{**}+(m-h+1)^{**}}. \end{aligned}$$

If we show

$$h^{**} + (m-h+1)^{**} \leq (m+1)^* \quad \text{for } 2 \leq h \leq m-1 \tag{3.10}$$

then it follows that (3.8) is bounded by  $cn^{(m+1)^*}$  and we completed the proof of the induction step.

To prove (3.10) it suffices by symmetry to consider only values  $h \leq \frac{1}{2}(m+1)$ . Let us say that an integer  $j \geq 2$  is *in the area of constant increase of  $(\cdot)^{**}$*  if  $j-p+1 \geq \left\lfloor \frac{j}{2} \right\rfloor$ , so for sufficiently large  $j$ . In this area  $(\cdot)^{**}$  increases with unit speed and outside this area  $(\cdot)^{**}$  increases at most with unit speed. Suppose first that  $m-h+1$  is in the area of constant increase. Replacing  $h$  by 2 in the left hand side of (3.10), and so replacing  $m-h+1$  by  $m-1$  would certainly not decrease this expression. Because

$$2^{**} + (m-1)^{**} = 1 + \max \left( \left\lfloor \frac{m-1}{2} \right\rfloor, m-p \right) = (m+1)^*$$

this proves (3.10). In case  $m-h+1$  is not in the area of constant increase, also  $h \leq \frac{1}{2}(m+1) \leq m-h+1$  is not in this area. In that case the left side of (3.10) equals

$$\left\lfloor \frac{h}{2} \right\rfloor + \left\lfloor \frac{m-h+1}{2} \right\rfloor \leq \left\lfloor \frac{m+1}{2} \right\rfloor \leq (m+1)^*.$$

Thus (3.10) holds for all  $h$ .  $\square$

#### 4. Proof of Theorem 1.1

*Proof.* We split up  $S_k$  in blocks to which we can apply Lemma 2.1. Define  $m_n$  and  $M_n$  as in Lemma 2.2. Write for  $k \leq n$

$$\begin{aligned} S_k = & \left[ \sum_{1 \leq j \leq m_n, j \text{ odd}} \sum_{(j-1)M_n < i \leq jM_n, i \leq k} \eta_i \right] \\ & + \left[ \sum_{1 \leq j \leq m_n, j \text{ even}} \sum_{(j-1)M_n < i \leq jM_n, i \leq k} \eta_i \right] \\ & + \left[ \sum_{m_n M_n < i \leq n, i \leq k} \eta_i \right]. \end{aligned}$$

Denote by  $S_k^{(1)}$ ,  $S_k^{(2)}$  and  $S_k^{(3)}$  the first, second and third term of the sum above. Observe

$$\max_{k \leq n} |S_k| \leq \max_{k \leq n} |S_k^{(1)}| + \max_{k \leq n} |S_k^{(2)}| + n - m_n M_n. \tag{4.1}$$

The last term is at most  $M_n$  and for  $\rho = 1, 2$

$$\max_{k \leq n} |S_k^{(\rho)}| \leq \max_{1 \leq j \leq m_n} |S_{jM_n}^{(\rho)}| + M_n. \tag{4.2}$$

By Lemma 2.2  $M_n = o(n^\alpha)$ . Choose  $N$  so large that  $M_n \leq \frac{1}{6} \varepsilon n^\alpha$  for  $n \geq N$ . Hence by (4.1) and (4.2)

$$P(\max_{k \leq n} |S_k| \geq \varepsilon n^\alpha) \leq \sum_{\rho=1,2} P(\max_{j \leq m_n} |S_{jM_n}^{(\rho)}| \geq \frac{1}{4} \varepsilon n^\alpha) \quad \text{for } n \geq N. \tag{4.3}$$

Consider, e.g.,  $\rho = 1$ . Let  $\sigma_j := \sum_{(j-1)M_n < i \leq jM_n} \eta_i$  and observe

$$S_{jM_n}^{(1)} = \sum_{1 \leq k \leq j, k \text{ odd}} \sigma_k.$$

We want to replace  $\sigma_j$ ,  $j$  odd, by independent random variables  $\tilde{\sigma}_j$ ,  $j$  odd. Observe that by (2.1)

$$\begin{aligned} &\beta(\sigma_j, (\sigma_k: k > j, k \text{ odd})) \\ &\leq \beta((\eta_i: i \leq jM_n), (\eta_i: i \geq (j+1)M_n)) = \beta_{M_n}. \end{aligned}$$

Hence by Lemma 2.1, there are independent random variables  $\tilde{\sigma}_j$ ,  $1 \leq j \leq m_n$ ,  $j$  odd, distributed as  $\sigma_j$  such that

$$P(\tilde{\sigma}_j \neq \sigma_j \text{ for some } j) \leq m_n \beta_{M_n}.$$

So with probability at least  $1 - m_n \beta_{M_n}$

$$\begin{aligned} S_{jM_n}^{(1)} &= \sum_{1 \leq k \leq j, k \text{ odd}} \tilde{\sigma}_k \\ &= \sum_{1 \leq k \leq [\frac{1}{2}(j+1)]} \tilde{\sigma}_{2k-1}, \end{aligned}$$

where the last sum consists of independent random variables. Hence

$$\begin{aligned} &P(\max_{1 \leq j \leq m_n} |S_{jM_n}^{(1)}| \geq \frac{1}{4} \varepsilon n^\alpha) \\ &\leq m_n \beta_{M_n} + P\left(\max_{1 \leq j \leq m_n} \left| \sum_{k=1}^{[\frac{1}{2}(j+1)]} \tilde{\sigma}_{2k-1} \right| \geq \frac{1}{4} \varepsilon n^\alpha\right) \\ &\leq m_n \beta_{M_n} + \frac{E\left(\sum_{k=1}^{[\frac{1}{2}(m_n+1)]} \tilde{\sigma}_{2k-1}\right)^{2d}}{(\frac{1}{4} \varepsilon n^\alpha)^{2d}} \end{aligned}$$

for any  $d \geq 1$  by the Markov inequality. A similar estimate holds for  $\rho = 2$ . So by Lemma 2.2 and (4.3) it is sufficient to show

$$\sum_{n \geq 1} n^{\alpha p - 2} \frac{E \left( \sum_{k=1}^{m'_n} \bar{\sigma}_k \right)^{2d}}{n^{2\alpha d}} < \infty \tag{4.4}$$

for  $m'_n \leq m_n$  while  $\bar{\sigma}_k$  consists of independent random variables distributed as the sums  $\sigma_k$ . We take here  $d$  to be integer such that

$$d = 1 \quad \text{if } 1 \leq p < 2$$

$$d \geq p \frac{p\alpha - 1}{2\alpha - 1} \quad \text{otherwise.}$$

The expectation in (4.4) may be written as

$$\sum_{1 \leq i_1, \dots, i_{2d} \leq m'_n} E \bar{\sigma}_{i_1} \dots \bar{\sigma}_{i_{2d}}.$$

For given  $(i_1, \dots, i_{2d})$  write, using independence

$$E \bar{\sigma}_{i_1} \dots \bar{\sigma}_{i_{2d}} = E \sigma_{j_1}^{k_1} \dots E \sigma_{j_r}^{k_r}.$$

Here  $k_1 \geq k_2 \geq \dots \geq k_r$  are the  $i$ -multiplicities, satisfying  $k_1 + \dots + k_r = 2d$ , and  $j_1, \dots, j_r$  are indices between 1 and  $m'_n$ . Write

$$\mu^{(k)} = \max_{j \geq 1} E \sigma_j^k.$$

The total number of  $(i_1, \dots, i_{2d})$  above with given multiplicity configuration  $k_1 \geq \dots \geq k_r$  is bounded by

$$c_k \binom{m'_n}{r} \leq c_k (m_n)^r$$

where  $c_k$  is a constant depending on the multiplicity configuration  $k_1 \geq \dots \geq k_r$  but not on  $n$ . To prove (4.4) it is sufficient to show for each of the finitely many multiplicity configurations  $k_1 \geq \dots \geq k_r$  that

$$\sum_{n \geq 1} n^{\alpha p - 2} \frac{(m_n)^r \mu^{(k_1)} \dots \mu^{(k_r)}}{n^{2\alpha d}} < \infty. \tag{4.5}$$

Because  $E \sigma_j = 0$  we only have to consider the case where all  $k_j \geq 2$ . In case  $1 \leq p < 2$  this leaves only one configuration to consider, because  $d = 1$ .

*Case 1.*  $r = 1$  and so the entire configuration consists of  $k_1 = 2d$ . We use the moment bounds of Sect. 3. First consider

$$\sum_{n \geq 1} n^{\alpha p - 2} \frac{m_n \left[ M_n \sum_{k=0}^{M_n - 1} (k + 1)^{2d - 2} \alpha_k \right]}{n^{2\alpha d}}$$

where  $\alpha_k$  is defined by (1.2). Because  $m_n M_n \leq n$  and because for sufficiently large  $n$  we have  $M_n \leq n^\alpha$  the sum above is finite if

$$\begin{aligned} & \sum_{n \geq 1} n^{\alpha p - 2} \frac{n \sum_{0 \leq k < n^\alpha} (k+1)^{2d-2} \alpha_k}{n^{2\alpha d}} \\ &= \sum_{k \geq 1} \sum_{n > k^{\frac{1}{\alpha}}} n^{\alpha p - 1 - 2\alpha d} (k+1)^{2d-2} \alpha_k < \infty. \end{aligned}$$

Because  $2d > p$  this sum is finite if

$$\sum_{k \geq 1} k^{\frac{1}{\alpha}(\alpha p - 2\alpha d)} (k+1)^{2d-2} \alpha_k \leq c \sum_{k \geq 1} k^{p-2} \alpha_k < \infty$$

with  $c$  a finite constant. Because  $\alpha_n \leq \beta_n$  this holds by our assumption (1.4). If  $1 \leq p < 2$  this proves (4.5) by Lemma 3.1 because  $d = 1$ . Hence the theorem is proved for  $1 \leq p < 2$ . We can assume now that  $p \geq 2$ . To prove (4.5) for  $r = 1$ , it is sufficient by Lemma 3.2 to show that also

$$\sum_{n \geq 1} n^{\alpha p - 2} \frac{m_n M_n^{d+(d-p)+}}{n^{2\alpha d}} < \infty.$$

Because  $m_n M_n \leq n$  and  $M_n = o(n^\alpha)$  it is sufficient to show

$$\sum_{n \geq 1} n^{\alpha p - 2} \frac{n \cdot n^{\alpha(d-1+(d-p)+)}}{n^{2\alpha d}} = \sum_{n \geq 1} n^\beta < \infty.$$

Because  $p \geq 2$  we have  $d \geq p$  and so  $\beta = -\alpha - 1$ . Hence  $\beta < -1$  and (4.5) follows for  $r = 1$ .

Case 2.  $r \geq 2$  and  $k_1 \geq 2p$ . If  $k_1$  is even then by Corollary 3.3

$$\mu^{(k_1)} = O(M_n^{k_1 - p + 1}).$$

If  $k_1$  is odd then also by this corollary

$$\begin{aligned} E|\bar{\sigma}_j|^{k_1} &\leq M_n \mu^{(k_1-1)} = M_n \cdot O(M_n^{k_1-1-p+1}) \\ &= O(M_n^{k_1-p+1}). \end{aligned}$$

For all other  $k_i \geq 2$  by this corollary

$$\begin{aligned} \mu^{(k_i)} &\leq M_n^{k_i-2} \mu^{(2)} = M_n^{k_i-2} \cdot O(M_n) \\ &= O(M_n^{k_i-1}), \end{aligned} \tag{4.6}$$

because  $p \geq 2$ . To prove (4.5) it is sufficient to show

$$\sum_{n \geq 1} n^{\alpha p - 2} \frac{(m_n)^r M_n^{k_1 - p + 1} \cdot M_n^{k_2 - 1 + \dots + k_r - 1}}{n^{2\alpha d}} < \infty.$$

Because  $m_n M_n \leq n$  and we can use  $M_n \leq n^\alpha$  for large  $n$  it is sufficient to prove

$$\sum_{n \geq 1} n^{\alpha p - 2} \frac{n^r n^{\alpha(k_1 + \dots + k_r - (p-2) - 2r)}}{n^{2\alpha d}} = \sum_{n \geq 1} n^\beta < \infty.$$

Because  $2d = k_1 + \dots + k_r$ ,

$$\begin{aligned} \beta &= \alpha p - 2 - r(2\alpha - 1) - \alpha(p - 2) \\ &\leq -2\alpha < -1 \end{aligned}$$

because  $r \geq 2$  and also  $\alpha > \frac{1}{2}$ . Thus (4.5) holds in Case 2.

Case 3.  $r \geq 2$  and all  $k_i < 2p$ . Again using (4.6) as above we obtain now that it is sufficient to show  $\sum_{n \geq 1} n^\beta < \infty$  with

$$\beta = \alpha p - 2 - r(2\alpha - 1).$$

Now  $2d = k_1 + \dots + k_r < 2p \cdot r$  so  $r > \frac{d}{p}$ . By our choice of  $d$  it follows that indeed  $\beta < -1$ . This completes the proof of Theorem 1.1.  $\square$

In the case of unbounded random variables  $(\eta_n)$  we could also use this proving technique but it is not clear how to get a nice split-up with random variables of  $o(n^\alpha)$ . This makes it not probable that our results would be precise and we do not discuss this here further.

Absolute regularity was needed in the application of Lemma 2.1, for a split-up in random variables of size  $o(n^\alpha)$ . In case  $\alpha = 1$  we can however get the corresponding results for  $\alpha$ -mixing. Using the Markov inequality and the lemmas of Sect. 3 gives (1.1) for all  $p \geq 1$ . If  $\alpha = 1$  and  $p > 2$  we can also get

$$\sum_{n \geq 1} n^r P(\max_{k \leq n} |S_k| \geq n^\alpha \varepsilon) < \infty, \quad \varepsilon > 0, \tag{4.7}$$

for all  $r < \alpha p - 2$  by using the Markov inequality, Corollary 3.3 and Serfling (1970), Theorem B. Generally for  $\frac{1}{2} < \alpha \leq 1$  and  $p > 2$  the last method gives (4.7) only for all

$$r < (2\alpha - 1)(p - 1) - 1. \tag{4.8}$$

An application of Lai (1977), Theorem 2 for stationary bounded  $(\eta_i)$  gives for  $\frac{1}{2} < \alpha \leq 1$  and  $p > 2$  that (4.7) holds for

$$r < \alpha(p - 1) - 2. \tag{4.9}$$

It is interesting to see that (4.8) is sometimes better and sometimes worse than (4.9). As we might have expected the upper bounds in (4.8) and (4.9) are both dominated by  $\alpha p - 2$ . We also mention here that Hipp (1979) has stronger bounds than ours but his results for the  $\alpha$ -mixing case are in conflict with Theorem 6.1. For  $\phi$ -mixing his results were studied later by Peligrad (1985).

### 5. Strong Mixing and Averaging: $\alpha = 1$

*Proof of Theorem 1.2.* Apply Lemma 3.1 (3.2) with  $m \geq p$  and the Markov inequality to get after a simple calculation that

$$\sum_{k \geq 1} \gamma^{k(p-1)} P(|S_{\gamma^k}| \geq \varepsilon \gamma^k) < \infty \quad \text{for any } \varepsilon > 0$$

and  $\gamma > 1$  arbitrary. Because for  $\gamma^k \leq j < \gamma^{k+1}$  holds  $|S_j - S_{\gamma^k}| \leq (\gamma - 1)\gamma^k$  it follows that

$$\sum_{k \geq 1} \gamma^{k(p-1)} P \left( \sup_{\gamma^k \leq j < \gamma^{k+1}} \frac{1}{j} |S_j| \geq \varepsilon' \right) < \infty \quad \text{for any } \varepsilon' > \gamma - 1 > 0. \quad (5.1)$$

For  $p = 1$  this implies the strong law by Borel-Cantelli. For  $p > 1$  note that (1.8) is dominated by

$$\begin{aligned} & \sum_{k \geq 1} \sum_{\gamma^{k-1} \leq n < \gamma^k} n^{p-2} P \left( \sup_{j \geq \gamma^{k-1}} \frac{1}{j} |S_j| > \varepsilon \right) \\ & \leq c \sum_{k \geq 1} \gamma^{k(p-1)} \sum_{m \geq k-1} P \left( \sup_{\gamma^m \leq j < \gamma^{m+1}} \frac{1}{j} |S_j| > \varepsilon \right). \end{aligned}$$

Exchange summation and observe that for suitable choice of  $\gamma$  we obtain from (5.1) that the last expression is finite.  $\square$

### 6. Sharpness and Renewal Theory for Independent Random Variables

We show that both theorems of Sect. 1 are sharp for renewal sequences. Let  $F$  be an aperiodic probability distribution on  $\mathbb{N} = \{1, 2, \dots\}$  and let  $X_1, X_2, \dots$  have distribution  $F$ . Assume

$$P(X_0 = i) = \frac{1}{\mu} P(X_1 > i), \quad i \geq 0,$$

where  $\mu$  is the mean of  $F$  and let  $X_0, X_1, X_2, \dots$  be independent. Then

$$\begin{aligned} \tilde{\eta}_n &= 1 && \text{if } n = X_0 + X_1 + \dots + X_k \text{ for some } k \\ &= 0 && \text{otherwise} \end{aligned}$$

defines a ‘‘renewal’’ sequence  $(\tilde{\eta}_n)_{n \geq 0}$ . It is stationary and its distribution is uniquely determined by  $F$ . Let  $\eta_n = \tilde{\eta}_n - \frac{1}{\mu}$ ,  $n \in \mathbb{Z}$ . We study the strong law (1.5) for  $\eta_n$ . Note

$$\tilde{S}_n \geq a \quad \text{if and only if } X_0 + X_1 + \dots + X_{[a]} \leq n. \quad (6.1)$$

We write  $P^*(\cdot)$  for  $P(\cdot | X_0 = 0)$ . Using the notations of Sect. 1 also, we have

**Theorem 6.1.** *The following assertions are equivalent for  $p > 1$ :*

- (i)  *$F$  has finite  $p^{\text{th}}$  moment*
- (ii)  $\sum_{n \geq 1} n^{p-2} \beta_n < \infty$
- (iii)  $\sum_{n \geq 1} n^{p-2} \alpha_n < \infty$
- (iv)  $\sum_{n \geq 1} n^{\alpha p - 2} P^*(|S_n| \geq n^\alpha \varepsilon) < \infty$  for any given  $\alpha, p$  with  $\frac{1}{2} < \alpha \leq 1$  and  $\alpha p > 1$
- (v)  $\sum_{n \geq 1} n^{\alpha p - 2} P(\max_{k \leq n} |S_k| \geq n^\alpha \varepsilon) < \infty$  for all  $\alpha, p$  with  $\frac{1}{2} < \alpha \leq 1$  and  $\alpha p > 1$ .

*Proof.* (i)  $\xrightarrow{L}$  (ii)  $\Rightarrow$  (iii)  $\xrightarrow{T'}$  (iv) for  $\alpha=1$  and  $p>1 \xrightarrow{K}$  (i)  $\xrightarrow{L}$  (ii)  $\xrightarrow{T}$  (v)  $\Rightarrow$  (iv). Here  $L$  is Lindvall (1979) which we explain below; (ii)  $\Rightarrow$  (iii) is trivial because  $\alpha_n \leq \beta_n$ ;  $T'$  is Theorem 1.2;  $K$  is Baum and Katz (1965);  $T$  is Theorem 1.1 and (v)  $\Rightarrow$  (iv) is trivial because

$$P(|S_n| \geq \varepsilon n^\alpha) \geq P^*(|S_n| \geq \varepsilon n)P(X_0=0).$$

So from both (v), (iv) we have

$$\sum n^{\alpha p-2} P^*(|S_n| \geq \varepsilon n^\alpha) < \infty.$$

To get (i) from (iv) one uses  $K$  and the following considerations

$$\begin{aligned} \{S_n \geq \varepsilon n^\alpha\} &= \{\tilde{S}_n \geq \mu^{-1}n + \varepsilon n^\alpha\} \\ &= \{X_0 + \dots + X_{[\mu^{-1}n + \varepsilon n^\alpha]} \leq n\} \end{aligned}$$

by (6.1). So also by using that  $P^*(X_0=0)=1$ , we have

$$\begin{aligned} \infty &> \sum_{n \geq 1} n^{\alpha p-2} P^*(|S_n| \geq \varepsilon n^\alpha) \\ &= \sum_{n \geq 1} n^{\alpha p-2} [P^*(X_1 + \dots + X_{[\mu^{-1}n + \varepsilon n^\alpha]} \leq n) \\ &\quad + P^*(X_1 + \dots + X_{[\mu^{-1}n - \varepsilon n^\alpha]} > n)] \\ &\geq C_\varepsilon \sum_{n \geq 1} n^{\alpha p-2} P^*(|X_1 + \dots + X_n - \mu n| \geq \gamma n^\alpha), \end{aligned}$$

where  $\gamma = \gamma(\varepsilon) \downarrow 0$  as  $\varepsilon \downarrow 0$ . This implies by  $K$  that (i) holds.

Let us now explain (i)  $\xrightarrow{L}$  (ii). The stationary process  $(\tilde{\eta}_n)_{n \geq 0}$  can be extended to a stationary process  $(\tilde{\eta}_n)_{n = -\infty}^\infty$  and it is well known that this process has an imbedded Markov structure in the sense that  $Z_n = \inf \{k \geq n: \tilde{\eta}_k = 1\}$  is a Markov sequence and so  $(\tilde{\eta}_n)_{n \leq 0}$  is independent of  $(\tilde{\eta}_n)_{n \geq 0}$  given  $X_0 \equiv Z_0$ .

As in Lindvall (1979) one constructs also a second renewal sequence  $(\tilde{\eta}''_n)$  coupled with  $(\tilde{\eta}_n)$  such that  $\tilde{\eta}''_n = \tilde{\eta}_n$  for  $n \geq T$  where  $T$  is a random time. We say that  $X''_0$  is distributed as  $X_0$  and also that  $\tilde{\eta}$  and  $X''_0$  is independent. By our remark in the paragraph above we may even get that  $\tilde{\eta}$  and  $(X''_0, (\tilde{\eta}''_n)_{n \leq 0})$  are independent, while  $\tilde{\eta}_\cdot$  and  $\tilde{\eta}''_\cdot$  have the same distribution.

Now note that by (2.2) and by well-known coupling inequalities

$$\begin{aligned} \beta &= \frac{1}{2} \|P_{\tilde{\eta}_\cdot, \tilde{\eta}''_\cdot} - P_{\tilde{\eta}_\cdot, \tilde{\eta}_\cdot}\| \\ &\leq P(\tilde{\eta}''_{n+} \neq \tilde{\eta}_{n+}) \leq P(T \geq n). \end{aligned}$$

Here  $\tilde{\eta}_{n+} = (\tilde{\eta}_k)_{k \geq n}$  and similarly  $\tilde{\eta}''_{n+} = (\tilde{\eta}''_k)_{k \geq n}$ . Use of Lindvall (1979), Proposition 1 gives the assertion (i)  $\Rightarrow$  (ii).

The result (ii)  $\Rightarrow$  (i) is known in case  $p \geq 2$  by Davies and Grübel (1983) where a similar stronger result is given for  $p \geq 2$ . In this and in the related paper Davies and Grübel (1981) these authors use analytic techniques for function algebras. A nice aspect of such results is that they form a converse to the Pitman-Lindvall results (see Lindvall 1977).

Also a very interesting aspect of these results is that speed of convergence in the ergodic theorem is coupled to the mixing rate of a stationary sequence

i.e. (ii) or (iii)  $\Leftrightarrow$  (iv) or (v). Note that here  $\eta_n$  is a functional of the Markov chain

$$Y_n = n - \max \{j: \eta_j = 1 \text{ for } j < n\}.$$

The functional has the property that the event  $\{\eta_0 = 1\}$  disconnects past and future in the sense that  $(Y_n)_{n \leq 0}$  and  $(Y_n)_{n \geq 0}$  are independent given this event. The question arises how much (iii)  $\Leftrightarrow$  (v) can be generalized for subclasses of the set of strongly mixing sequences.

### 7. Averaging a Function Along a Random Walk

Consider a random walk  $(T_n)_{n \geq 0}$  started at the origin with strictly positive increments, having an aperiodic distribution  $F$ . Let  $F$  have finite mean  $\mu$ . Suppose there is given a bounded function  $f$  on the integers.

**Theorem 7.1.** *If  $F$  has finite  $p^{\text{th}}$  moment,  $\frac{1}{2} < \alpha \leq 1$ ,  $\alpha p > 1$ , then*

$$\sum_{N \geq 0} N^{\alpha p - 2} P \left( \sup_{n \geq N} \frac{1}{n^\alpha} \left| \sum_{k=1}^n f(T_k) - \frac{1}{\mu} \sum_{0 \leq k < n\mu} f(k) \right| \geq \varepsilon \right) < \infty.$$

*Proof.* Section 2 of Bingham and Goldie (1981) shows that this is implied by Theorem 6.1 (iv) and (i).  $\square$

Meilijson (1973) and Stam (1968) discuss related results. Investigation of distributions  $F$  with mass on the negative axis would go too far here, but Bingham and Goldie (1982) contain a useful idea for this case.

*Acknowledgement.* The last application stimulated the research of this paper and therefore I thank N. Bingham for a visit to Westfield College in connection with these investigations. I also thank a referee for indicating improvements to the paper.

### References

1. Baum, L.B., Katz, M.: Convergence rates in the law of large numbers. *Trans. Amer. Math. Soc.* **120**, 108-123 (1965)
2. Berbee, H.C.P.: Random walks with stationary increments and renewal theory. Amsterdam: Mathematical Centre tract 112 (1979)
3. Berbee, H.C.P.: Periodicity and absolute regularity. To appear in *Israel J. Math.* (1984)
4. Bingham, N.H., Goldie, C.M.: Probabilistic and deterministic averaging. *Trans. Amer. Math. Soc.* **269**, 453-480 (1982)
5. Davies, P.L., Grübel, R.: Spaces of summable sequences in renewal theory and the theory of Markov chains. *Math. Nachr.* **104**, 119-128 (1981)
6. Davies, P.L., Grübel, R.: A smoothness property of renewal sequences. *Z. Angew. Math. Mech.* **63**, 455-456 (1983)
7. Etemadi, N.: An elementary proof of the strong law of large numbers. *Z. Wahrscheinlichkeitstheor. Verw. Geb.* **55**, 119-122 (1981)
8. Hipp, C.: Convergence rates of the strong law for stationary mixing sequences. *Z. Wahrscheinlichkeitstheor. Verw. Geb.* **49**, 49-62 (1979)
9. Ibragimov, I.A., Linnik, Yu.V.: Independent and stationary sequences of random variables. Wolters-Noordhoff, Groningen (1971)

10. Lai, T.L.: Convergence rates and  $r$ -quick versions of the strong law for stationary mixing sequences. *Ann. Probab.* **5**, 693–706 (1977)
11. Lindvall, T.: On coupling of discrete renewal processes. *Z. Wahrscheinlichkeitstheor. Verw. Geb.* **48**, 57–70 (1979)
12. Mailijson, I.: The average of the values of a function at random points. *Israel J. Math.* **15**, 193–203 (1973)
13. Peligrad, M.: Convergence rates of the strong law for stationary mixing sequences. *Z. Wahrscheinlichkeitstheor. Verw. Geb.* **70**, 307–314 (1985)
14. Schwarz, G.: Finitely determined processes – an indiscrete approach. *J. Math. Anal. Appl.* **76**, 146–158 (1980)
15. Sen, P.K.: Weak convergence of multidimensional empirical processes for stationary  $\phi$ -mixing processes. *Ann. Probab.* **2**, 147–154 (1974)
16. Serfling, R.J.: Moment inequalities for the maximum cumulative sum. *Ann. Math. Statist.* **41**, 1227–1234 (1970)
17. Stam, A.J.: Laws of large numbers for functions of random walks with positive drift. *Compos. Math.* **19**, 299–333 (1968)
18. Yokoyama, R.: Moment bounds for stationary mixing sequences. *Z. Wahrscheinlichkeitstheor. Verw. Geb.* **52**, 45–57 (1980)

Received December 3, 1984; in revised form November 8, 1985