DOI 10.1007/s10986-021-09528-7
Lithuanian Mathematical Journal, Vol. 61, No. 4, October, 2021, pp. 471–482

The strong law of large numbers for sums of randomly chosen random variables

Agnieszka M. Gdula and Andrzej Krajka

Maria-Curie Skłodowska University, ul. Akademicka 9, 20-033 Lublin, Poland
(e-mail: gdula.agnieszka@gmail.com; akrajka@gmail.com)

Received January 30, 2020; revised October 16, 2020

Abstract. Let $\{X_n,\, n \geqslant 1\}$ be a sequence of independent or identically distributed dependent random variables, and let $\{A_n,\, n \geqslant 1\}$ be a sequence of random subsets of natural numbers independent of $\{X_n,\, n \geqslant 1\}$. In this paper, we describe the strong law of large numbers (SLLN) of the form $\sum_{i\in A_n}(X_i - \mathbf{E}\sum_{i\in A_n} X_i)/b_n \to 0$ a.s. as $n \to \infty$ for some sequence of nondecreasing positive numbers $\{b_n,\, n \geqslant 1\}$. There often arises an assumption that $\{A_n,\, n \geqslant 1\}$ are almost surely increasing: $A_n \subset A_{n+1}$, a.s. $n \geqslant 1$.

MSC: 60F15

Keywords: strong law of large numbers, randomly indexed sums, random sets

1 Introduction

Let $\{X_n,\, n \geqslant 1\}$ be a sequence of independent random variables defined on a probability space $(\Omega, \mathfrak{F}, \mathbf{P})$ such that $\mathbf{E}|X_n| < \infty$, $n \geqslant 1$. A random set with values in $2^{\mathbb{N}}$ is a map $A : \Omega \to 2^{\mathbb{N}}$ such that $A^{-1}(B) = \{\omega \in \Omega\colon\ A(\omega) = B\} \in \mathfrak{F}$ for any $B \subset \mathbb{N}$ (cf. [7, p. 35, Def. 3.1 and the remark on p. 72]). Obviously, if A and B are random sets, then $A \cup B$, $A \cap B$, and $A \setminus B$ are also random sets. Furthermore, $|A|$ is a nonnegative random variable, where $|A|$ is the cardinality of the set A. The basis of the theory of random sets can be found in a classic book by Matheron [5] or, more currently, in Molchanov's book [6] or Nguyen's book [7]. For an arbitrary (random or nonrandom) subset A of natural numbers, we put

$$S(A) = \sum_{i\in A} X_i - \mathbf{E}\sum_{i\in A} X_i.$$

Let $\{A_n,\, n \geqslant 1\}$ (we always put $A_0 = \emptyset$) be a sequence of arbitrarily dependent subsets of positive integer numbers $\mathbb{N}$ that are almost surely bounded, that is, there exists a sequence of positive reals $\{\overline{\alpha}_n,\, n \geqslant 1\}$, possibly, divergent to infinity, such that

$$\mathbf{P}\bigl[\sup\{k\colon\ k \in A_n\} \leqslant \overline{\alpha}_n\bigr] = 1. \tag{1.1}$$

Throughout the paper, we assume that $\{A_n,\, n \geqslant 1\}$ and $\{X_n,\, n \geqslant 1\}$ are independent, that is, for every sequence $\{B_n,\, n \geqslant 1\}$ of subsets of $2^{\mathbb{N}}$ and every sequence $\{C_n,\, n \geqslant 1\}$ of measurable Borel sets on $\mathbb{R}$, for every $n \in \mathbb{N}$, we have

$$\mathbf{P}[A_k \in B_k,\ X_k \in C_k,\ 1 \leqslant k \leqslant n] = \mathbf{P}[A_k \in B_k,\ 1 \leqslant k \leqslant n]\,\mathbf{P}[X_k \in C_k,\ 1 \leqslant k \leqslant n].$$

The general aim of this paper is to establish the strong law of large numbers for the sums $\{S(A_n),\ n \geqslant 1\}$. Randomly indexed sums have not been considered yet. Generally, this problem seems very difficult.

In our results, there often (although not always) arises the assumption

$$A_n \subset A_{n+1} \text{ a.s.}, \quad n \geqslant 1. \tag{1.2}$$

Therefore the investigation of the behavior $\{S(A_n)/b_n,\ n \geqslant 1\}$ for some sequence of divergent to infinity positive reals $\{b_n,\ n \geqslant 1\}$ under (1.2) is equivalent to investigating $\{\sum_{i=1}^{n} S_{(i)}/b_n,\ n \geqslant 1\}$, where

$$S_{(n)} = \sum_{i \in A_n \setminus A_{n-1}} X_i - \mathbf{E} \sum_{i \in A_n \setminus A_{n-1}} X_i, \quad n \geqslant 1,$$

is the sequence of dependent random variables.

The most recent result for a sequence of dependent mixing random variables is due to Hu and Weber [4] (see also [2] and [3]). Their result, formulated in our terms, may be stated as follows.

Theorem 1. (*See* [4, *Thm.* 1.1 *and Cor.* 1.2].) *Let $\{X_n,\ n \geqslant 1\}$ be a sequence of random variables, and let $\{A_n,\ n \geqslant 1\}$ be a sequence of random subsets of $\mathbb{N}$ such that $A_n \subset A_{n+1}$ a.s. $n \geqslant 1$. Let $\{b_n,\ n \geqslant 1\}$ be an increasing sequence of positive constants. Assume that there exists a constant K such that, for all $n \geqslant 1$,*

$$\frac{n}{b_n} \leqslant K. \tag{1.3}$$

Suppose that

$$\sum_{n=1}^{\infty} \frac{\operatorname{Var}(\sum_{i \in A_n \setminus A_{n-1}} X_i) \log^2 n}{b_n^2} < \infty \tag{1.4}$$

and

$$\sum_{k=1}^{\infty} \sup_{n \geqslant 1} \left| \operatorname{Cov}\Biggl(\sum_{i \in A_{n+1} \setminus A_n} X_i, \sum_{i \in A_{n+k+1} \setminus A_{n+k}} X_i \Biggr) \right| \frac{\log^2 k}{k} < \infty. \tag{1.5}$$

Then

$$\lim_{n \to \infty} \frac{S(A_n)}{b_n} = 0 \quad \textit{a.s.}$$

The general aim of our paper is to obtain a new SLLN for $\{S(A_n)/b_n,\ n \geqslant 1\}$ with the following improvements:

- We assume nothing about the mixing structure type (1.5). In Example 1 and Remark 2 in Section 4, we construct the sequence $\{X_n,\ n \geqslant 1\}$ such that $\{S_{(n)},\ n \geqslant 1\}$ are dependent and not satisfying (1.5) but such that for this sequence, our result holds.
- We remove assumption (1.3).
- Our results essentially weaken assumption (1.4).
- We consider both the Marcinkiewicz–Zygmund and Kolmogorov SLLNs.
- The technique of proof of our results is essentially different from that presented in papers [3, 4] and [2]. We develop the technique described in [1] and [8].

We postpone a discussion and comparison of our results with Theorem 1 to Section 4 devoted to remarks, examples, and conclusions. Section 2 contains the main results, which are proved in Section 3.

Throughout the paper, C denotes the generic constants, and we always assume that $\sum_{i \in \emptyset} a_i = 0$.

2 Main results

Let $\{X_n,\, n \geqslant 1\}$ be a sequence of random variables such that $\mathbf{E}|X_n| < \infty$, $n \geqslant 1$. For an arbitrary random subset A of natural numbers, we denote

$$\begin{aligned} V(A) &= \sum_{i\in A} X_i - \sum_{i\in A} \mathbf{E}X_i = \sum_{i\in A}(X_i - \mathbf{E}X_i), \\ Z(A) &= \sum_{i\in A} \mathbf{E}X_i - \mathbf{E}\sum_{i\in A}\mathbf{E}X_i = \sum_{i=1}^{\infty}\bigl(\mathbf{I}[i\in A] - \mathbf{P}[i\in A]\bigr)\mathbf{E}X_i, \\ S(A) &= \sum_{i\in A} X_i - \mathbf{E}\sum_{i\in A} X_i. \end{aligned} \tag{2.1}$$

Note that if A is independent of $\{X_n,\, n \geqslant 1\}$, then

$$S(A) = V(A) + Z(A). \tag{2.2}$$

Theorem 2. *Let $\{X_n,\, n \geqslant 1\}$ be a sequence of independent random variables such that $\mathbf{E}|X_n| < \infty$, $n \geqslant 1$, and let $\{A_n,\, n \geqslant 0\}$ be a sequence of random subsets of $\mathbb{N}$ ($A_o = \emptyset$) independent of $\{X_n,\, n \geqslant 1\}$ and satisfying* (1.1). *Let $\{b_n,\, n \geqslant 1\}$ be a nondecreasing unbounded sequence of positive reals. Introduce the following conditions*:

(a) *for $q > 1$,*

$$\sum_{n=1}^{\infty}\mathbf{E}\sum_{j\in A_n\setminus A_{n-1}}\mathbf{E}|X_j - \mathbf{E}X_j|^{2q}\sum_{k=n+1}^{\infty}\frac{|A_k|^{q-1} - |A_{k-1}|^{q-1}}{b_k^{2q}} < \infty; \tag{2.3}$$

(b) *for $q \geqslant 1$,*

$$\sum_{n=1}^{\infty}\mathbf{E}\sum_{j\in A_n\setminus A_{n-1}}\mathbf{E}|X_j - \mathbf{E}X_j|^{2q}\frac{|A_n|^{q-1}}{b_n^{2q}} < \infty; \tag{2.4}$$

(c) $$A_n \subset A_{n+1} \ \textit{a.s.}, \quad n \geqslant 1;$$

(d) *for every $\epsilon > 0$,*

$$\sum_{n=1}^{\infty}\Bigl(\mathrm{e}^{-b_n\epsilon + \mathbf{E}\sum_{i\in A_n} X_i}\,\mathbf{E}\prod_{i\in A_n}\mathrm{e}^{-\mathbf{E}X_i}\Bigr) < \infty, \tag{2.5}$$

$$\sum_{n=1}^{\infty}\Bigl(\mathrm{e}^{-b_n\epsilon - \mathbf{E}\sum_{i\in A_n} X_i}\,\mathbf{E}\prod_{i\in A_n}\mathrm{e}^{\mathbf{E}X_i}\Bigr) < \infty. \tag{2.6}$$

If for some $q \geqslant 1$, $\mathbf{E}|X_n|^{2q} < \infty$, $n \geqslant 1$, and (a)–(c) *are satisfied, then*

$$\lim_{n\to\infty}\frac{V(A_n)}{b_n} = 0 \quad \textit{a.s.} \tag{2.7}$$

If (d) *is satisfied, then*

$$\lim_{n\to\infty}\frac{Z(A_n)}{b_n} = 0 \quad \textit{a.s.} \tag{2.8}$$

If for some $q \geqslant 1$, $\mathbf{E}|X_n|^{2q} < \infty$, $n \geqslant 1$, *and* (a)–(d) *are satisfied, then*

$$\lim_{n\to\infty} \frac{S(A_n)}{b_n} = 0 \quad \textit{a.s.} \tag{2.9}$$

Corollary 1. *Let* $\{X, X_n,\, n \geqslant 1\}$ *be a sequence of independent identically distributed random variables such that* $\mathbf{E}|X| < \infty$, *and let* $\{A_n,\, n \geqslant 0\}$ *be a sequence of random subsets of* $\mathbb{N}$ ($A_o = \emptyset$) *independent of* $\{X_n,\, n \geqslant 1\}$ *and satisfying* (1.1). *Let* $\{b_n,\, n \geqslant 1\}$ *be a nondecreasing unbounded sequence of positive reals. Introduce the following conditions*:

(i)
$$A_n \subset A_{n+1} \ \textit{a.s.}, \quad n \geqslant 1;$$

(ii) *for* $q > 1$,

$$\sum_{n=1}^{\infty} \mathbf{E}\left(\big(|A_n| - |A_{n-1}|\big) \sum_{k=n+1}^{\infty} \frac{|A_k|^{q-1} - |A_{k-1}|^{q-1}}{b_k^{2q}} \right) < \infty;$$

(iii) *for* $q \geqslant 1$,

$$\sum_{n=1}^{\infty} \mathbf{E} \frac{(|A_n| - |A_{n-1}|)|A_n|^{q-1}}{b_n^{2q}} < \infty;$$

(iv) *for every* $\epsilon > 0$,

$$\sum_{n=1}^{\infty} \mathrm{e}^{-b_n\epsilon} \mathbf{E}\big(\mathrm{e}^{||A_n| - E|A_n||}\big)^{\mathbf{E}X} < \infty.$$

If $\mathbf{E}|X|^{2q} < \infty$, $q \geqslant 1$, *and* (i)–(iii) *are satisfied, then*

$$\lim_{n\to\infty} \frac{V(A_n)}{b_n} = 0 \quad \textit{a.s.}$$

If (iv) *is satisfied, then*

$$\lim_{n\to\infty} \frac{Z(A_n)}{b_n} = 0 \quad \textit{a.s.}$$

If $\mathbf{E}|X|^{2q} < \infty$, $q \geqslant 1$, *and* (i)–(iv) *are satisfied, then*

$$\lim_{n\to\infty} \frac{S(A_n)}{b_n} = 0 \quad \textit{a.s.}$$

It is worth noting that condition (iv) of the corollary is satisfied when

$$\lim_{n\to\infty} \frac{\log \mathbf{E}(\mathrm{e}^{||A_n| - \mathbf{E}|A_n||})^{\mathbf{E}X}}{b_n} = 0$$

and, for every $\epsilon > 0$,

$$\sum_{n=1}^{\infty} \mathrm{e}^{-\epsilon b_n} < \infty.$$

Let us now consider the case where $\{X_n,\, n \geqslant 1\}$ are arbitrary dependent but identically distributed.

Theorem 3. *Let $\{X, X_n,\, n \geqslant 1\}$ be a sequence of arbitrary dependent identically distributed random variables such that $\mathbf{E}|X| < \infty$, and let $\{A_n,\, n \geqslant 0\}$ be a sequence of a random subsets of $\mathbb{N}$ ($A_o = \emptyset$) independent of $\{X, X_n,\, n \geqslant 1\}$ such that*

$$A_{n-1} \subset A_n \ \textit{a.s.}, \quad n \geqslant 1.$$

Suppose that $b_0 = 0$ and $\{b_n, n \geqslant 1\}$ is a nondecreasing divergent to infinity sequence of positive constants and

$$\sum_{k=1}^{\infty} b_k P[b_{k-1} < |X - \mathbf{E}X| \leqslant b_k] \sum_{n=1}^{\infty} \frac{\mathbf{E}|(A_n \setminus A_{n-1}) \cap [k, \infty)|}{b_n} < \infty, \tag{2.10}$$

$$\liminf_{k\to\infty} \frac{b_k}{k} \sum_{n=1}^{\infty} \frac{\mathbf{E}|(A_n \setminus A_{n-1}) \cap [k, \infty)|}{b_n} > 0. \tag{2.11}$$

Assume additionally that, for every $\epsilon > 0$, we have

$$\sum_{n=1}^{\infty} \mathrm{e}^{-b_n \epsilon} \bigl(\mathrm{e}^{\mathbf{E}X\mathbf{E}|A_n|} \mathbf{E}\bigl(\mathrm{e}^{-\mathbf{E}X}\bigr)^{|A_n|} + \mathrm{e}^{-\mathbf{E}X\mathbf{E}|A_n|} \mathbf{E}\bigl(\mathrm{e}^{\mathbf{E}X}\bigr)^{|A_n|}\bigr) < \infty. \tag{2.12}$$

Then

$$\lim_{n\to\infty} \frac{S(A_n)}{b_n} = 0 \quad \textit{a.s.}$$

It is easy to check that in the case of $A_n = \{1, 2, \dots, n\}$, $n \geqslant 1$, from Theorem 3 we obtain Theorem 2.1 of [8].

3 Proofs

In this section, for an arbitrary set $A \subset \mathbb{N}$ and for an arbitrary sequence of random variables $\{X_n,\, n \geqslant 1\}$, we will use notation (2.1). Because $V(A_n)$ may be written as the sum of random variables, the classical Hájek–Rényi inequality holds.

Lemma 1. (*See Hájek–Rényi-type maximal inequality,* [1, *Thm.* 1.1].) *Let $\{A_n,\, n \geqslant 0\}$, $A_0 = \emptyset$, be an a.s. increasing sequence of random subsets of $\mathbb{N}$ satisfying* (1.1) *and independent of the sequence of random variables $\{X_n,\, n \geqslant 1\}$. Let $\beta_1, \beta_2, \dots$ be a nondecreasing sequence of positive numbers. Let $\alpha_1, \alpha_2, \dots$ be nonnegative numbers. Let r be a fixed positive number. Assume that for each m with $1 \leqslant m \leqslant n$,*

$$\mathbf{E}\Bigl[\max_{1\leqslant l\leqslant m} \bigl|V(A_l)\bigr|\Bigr]^r \leqslant \sum_{l=1}^{m} \alpha_l.$$

Then

$$\mathbf{E}\biggl[\max_{1\leqslant l\leqslant n} \biggl|\frac{V(A_l)}{\beta_l}\biggr|\biggr]^r < 4 \sum_{l=1}^{n} \frac{\alpha_l}{\beta_l^r}.$$

Lemma 2. (*See* [1, *Thm.* 2.1].) *Let $\{A_n,\, n \geqslant 0\}$ be a sequence of a.s. increasing random subsets of $\mathbb{N}$ satisfying* (1.1):

$$A_n : \Omega \to 2^{\mathbb{N}}, \quad A_n \subset A_{n+1} \ \textit{a.s.}, \quad n \geqslant 1.$$

Let $\{b_n,\, n \geqslant 1\}$ be a nondecreasing unbounded sequence of positive numbers, and let $\{\alpha_n,\, n \geqslant 1\}$ be a sequence of nonnegative numbers. Let r be a fixed positive number. Assume that for each $n \geqslant 1$,

$$\mathbf{E}\Big[\max_{1\leqslant l\leqslant n}\big|V(A_l)\big|\Big]^r \leqslant \sum_{l=1}^{n} \alpha_l.$$

If $\sum_{l=1}^{\infty} \alpha_l/b_l^r < \infty$, then

$$\lim_{n\to\infty} \frac{V(A_n)}{b_n} = 0 \quad \textit{a.s.}$$

Proof of Theorem 2. We first consider convergence (2.7). The proof essentially runs similarly as that of Corollary 3.1 in [1]. We only remark that in our case the Doob inequality is

$$\begin{aligned}
\mathbf{E}\Big[\max_{1\leqslant k\leqslant n}\big|V(A_k)\big|\Big]^{2q} &= \sum{}' \mathbf{E}\Big[\max_{1\leqslant k\leqslant n}\big|V(B_k)\big|\Big]^{2q} \mathbf{P}[A_k = B_k,\ 1 \leqslant k \leqslant n] \\
&\leqslant \left(\frac{2q}{2q-1}\right)^{2q} \sum{}' \mathbf{E}\big[\big|V(B_n)\big|\big]^{2q} \mathbf{P}[A_k = B_k,\ 1 \leqslant k \leqslant n] \\
&\leqslant \left(\frac{2q}{2q-1}\right)^{2q} \mathbf{E}\big|V(A_n)\big|^{2q},
\end{aligned}$$

where $\sum{}'$ is the sum taken over all possible sets $\{B_k,\, 1 \leqslant k \leqslant n\}$ such that $B_i \subset B_{i+1}$, $1 \leqslant i \leqslant n-1$, and $B_n \subset \{1, 2, \ldots, \overline{\alpha}_n\}$ (cf. (1.1)), and the Burkholder inequality is

$$\mathbf{E}\big[\big|V(A_n)\big|\big]^{2q} \leqslant c_q \mathbf{E}\bigg(\sum_{j\in A_n} (X_j - \mathbf{E}X_j)^2\bigg)^q.$$

Applying the Hölder inequality for $q > 1$, we get

$$\mathbf{E}\big[\big|V(A_n)\big|\big]^{2q} \leqslant c_q \mathbf{E}\bigg(\sum_{j\in A_n} |X_j - \mathbf{E}X_j|^{2q} |A_n|^{q-1}\bigg).$$

Now putting, for $n \geqslant 1$,

$$\alpha_n = \begin{cases} c(q)(\mathbf{E}\sum_{j\in A_n\setminus A_{n-1}} \mathbf{E}|X_j - \mathbf{E}X_j|^{2q}|A_n|^{q-1} & \\ \quad +\mathbf{E}\sum_{j\in A_{n-1}} \mathbf{E}|X_j - \mathbf{E}X_j|^{2q}(|A_n|^{q-1} - |A_{n-1}|^{q-1})) & \text{if } q > 1, \\ c(q)\mathbf{E}\sum_{j\in A_n\setminus A_{n-1}} \mathrm{Var}(X_j) & \text{if } q = 1, \end{cases}$$

where $c(q) = c_q(2q/(2q-1))^{2q}$, we get

$$\mathbf{E}\Big[\max_{1\leqslant k\leqslant n}\big|V(A_k)\big|\Big]^{2q} \leqslant \sum_{j=1}^{n} \alpha_j, \quad n \geqslant 1.$$

Thus (a), (b), and Lemma 2 end the proof of (2.7).

To prove (2.8), it suffices to show that for all $\varepsilon > 0$,

$$\sum_{n=1}^{\infty} \mathbf{P}\bigg[\bigg|\frac{Z(A_n)}{b_n}\bigg| > \varepsilon\bigg] < \infty.$$

Notice that by the Chebyshev exponential inequality we have

$$\sum_{n=1}^{\infty} \mathbf{P}\big[Z(A_n) > \varepsilon b_n\big] \leqslant \sum_{n=1}^{\infty} \frac{\mathbf{E}\mathrm{e}^{Z(A_n)}}{\mathrm{e}^{\varepsilon b_n}} = \sum_{n=1}^{\infty} \mathrm{e}^{-\varepsilon b_n} \mathbf{E}\, \mathrm{e}^{\sum_{i=1}^{\infty}(\mathbf{I}[i\in A_n]-p_{n,i})\mathbf{E}X_i},$$

where $p_{n,i} = \mathbf{P}[i \in A_n]$, $i, n \geqslant 1$. Next, taking into account that

$$\sum_{i=1}^{\infty} p_{n,i}\mathbf{E}X_i = \mathbf{E}\sum_{i\in A_n} X_i, \quad n \geqslant 1,$$

and by the property of the indicator function we have

$$\begin{aligned} \mathbf{E}\, \mathrm{e}^{\sum_{i=1}^{\infty}(\mathbf{I}[i\in A_n]-p_{n,i})\mathbf{E}X_i} &= \mathrm{e}^{-\sum_{i=1}^{\infty} p_{n,i}\mathbf{E}X_i}\, \mathbf{E}\, \mathrm{e}^{\sum_{i=1}^{\infty} \mathbf{I}[i\in A_n]X_i} \\ &= \mathrm{e}^{-\mathbf{E}\sum_{i\in A_n} X_i}\, \mathbf{E} \prod_{i\in A_n} \mathrm{e}^{\mathbf{E}X_i}. \end{aligned}$$

Thus from (2.6) we obtain

$$\sum_{n=1}^{\infty} \mathbf{P}\big[Z(A_n) > \varepsilon b_n\big] < \infty.$$

Similarly, we have

$$\mathbf{E}\, \mathrm{e}^{-\sum_{i=1}^{\infty}(\mathbf{I}[i\in A_n]-p_{n,i})\mathbf{E}X_i} = \mathrm{e}^{\mathbf{E}\sum_{i\in A_n} X_i}\, \mathbf{E} \prod_{i\in A_n} \mathrm{e}^{-\mathbf{E}X_i},$$

and from (2.5) we have

$$\sum_{n=1}^{\infty} \mathbf{P}\big[Z(A_n) < -\varepsilon b_n\big] < \infty. \qquad \square$$

Proof of Theorem 3. Let us first prove that

$$\lim_{n\to\infty} \frac{V(A_n)}{b_n} = 0 \quad \text{a.s.} \tag{3.1}$$

Because

$$\begin{aligned} V(A_n) = &\sum_{i\in A_n} (X_i - \mathbf{E}X_i)\mathbf{I}\big[|X_i - \mathbf{E}X_i| > b_i\big] \\ &+ \sum_{k=1}^{n} \sum_{i\in A_k\setminus A_{k-1}} (X_i - \mathbf{E}X_i)\mathbf{I}\big[|X_i - \mathbf{E}X_i| \leqslant b_i\big], \end{aligned}$$

to obtain (3.1), it suffices to prove that

$$\sum_{i=1}^{\infty} |X_i - \mathbf{E}X_i|\mathbf{I}\big[|X_i - \mathbf{E}X_i| > b_i\big] < \infty \quad \text{a.s.} \tag{3.2}$$

and (by the Kronecker lemma)

$$\sum_{n=1}^{\infty} \frac{1}{b_n} \sum_{i\in A_n\setminus A_{n-1}} |X_i - \mathbf{E}X_i|\mathbf{I}[|X_i - \mathbf{E}X_i| \leqslant b_i] < \infty \quad \text{a.s.} \tag{3.3}$$

Proof of (3.2). From (2.10) and (2.11) we may conclude that

$$\sum_{k=1}^{\infty} \mathbf{P}\big[|X-\mathbf{E}X|>b_k\big] = \sum_{k=1}^{\infty} k\mathbf{P}\big[b_{k-1}<|X-\mathbf{E}X|\leqslant b_k\big] - \mathbf{P}\big[|X-\mathbf{E}X|\neq 0\big] < \infty.$$

Thus by the Borel–Cantelli lemma there exists the positive integer-valued random variable Y such that

$$\mathbf{P}\Big[\bigcup_{n\geqslant Y}\big[|X_n-\mathbf{E}X_n|>b_n\big]\Big]=0,$$

and therefore

$$\sum_{i=1}^{\infty}|X_i-\mathbf{E}X_i|\mathbf{I}\big[|X_i-\mathbf{E}X_i|>b_i\big] = \sum_{i\leqslant Y}|X_i-\mathbf{E}X_i|\mathbf{I}\big[|X_i-\mathbf{E}X_i|>b_i\big] < \infty \quad \text{a.s.},$$

which ends the proof of (3.2).

Proof of (3.3). From (2.10) we have

$$\begin{aligned}
&\sum_{n=1}^{\infty}\frac{1}{b_n}\mathbf{E}\sum_{i\in A_n\setminus A_{n-1}}|X_i-\mathbf{E}X_i|\mathbf{I}\big[|X_i-\mathbf{E}X_i|\leqslant b_i\big]\\
&=\sum_{n=1}^{\infty}\frac{1}{b_n}\mathbf{E}\sum_{i\in A_n\setminus A_{n-1}}\sum_{k=1}^{i}|X_i-\mathbf{E}X_i|\mathbf{I}\big[b_{k-1}<|X_i-\mathbf{E}X_i|\leqslant b_k\big]\\
&\leqslant\sum_{n=1}^{\infty}\frac{1}{b_n}\mathbf{E}\sum_{i\in A_n\setminus A_{n-1}}\sum_{k=1}^{i}b_k\mathbf{P}\big[b_{k-1}<|X-\mathbf{E}X|\leqslant b_k\big]\\
&=\sum_{n=1}^{\infty}\frac{1}{b_n}\sum_{k=1}^{\infty}b_k\mathbf{P}\big[b_{k-1}<|X-\mathbf{E}X|\leqslant b_k\big]\mathbf{E}\sum_{i\in A_n\setminus A_{n-1},\, i\geqslant k}1\\
&=\sum_{n=1}^{\infty}\frac{1}{b_n}\sum_{k=1}^{\infty}b_k\mathbf{P}\big[b_{k-1}<|X-\mathbf{E}X|\leqslant b_k\big]\mathbf{E}\big|(A_n\setminus A_{n-1})\cap[k,\infty)\big|\\
&=\sum_{k=1}^{\infty}b_k\mathbf{P}\big[b_{k-1}<|X-\mathbf{E}X|\leqslant b_k\big]\sum_{n=1}^{\infty}\frac{\mathbf{E}|(A_n\setminus A_{n-1})\cap[k,\infty)|}{b_n}\\
&<\infty,
\end{aligned}$$

and since L^1 bounded series are almost surely bounded, we also get (3.3).

To prove Theorem 3, we need to show that

$$\lim_{n\to\infty}\frac{Z(A_n)}{b_n}=0 \quad \text{a.s.},$$

but because complete convergence implies almost sure convergence, it suffices to prove that for every $\epsilon>0$,

$$\sum_{n=1}^{\infty}\mathbf{P}\big[\big|Z(A_n)\big|>\epsilon b_n\big]<\infty. \tag{3.4}$$

By the Chebyshev exponential inequality we have

$$\sum_{n=1}^{\infty}\mathbf{P}\big[Z(A_n)>\varepsilon b_n\big]\leqslant\sum_{n=1}^{\infty}\mathrm{e}^{-\varepsilon b_n}\mathbf{E}\mathrm{e}^{\mathbf{E}X[\sum_{i=1}^{\infty}(\mathbf{I}[i\in A_n]-p_{n,i})]},$$
$$\sum_{n=1}^{\infty}\mathbf{P}\big[Z(A_n)<-\varepsilon b_n\big]\leqslant\sum_{n=1}^{\infty}\mathrm{e}^{-\varepsilon b_n}\mathbf{E}\mathrm{e}^{\mathbf{E}X[\sum_{i=1}^{\infty}(p_{n,i}-\mathbf{I}[i\in A_n])]},$$

where $p_{n,i}=P[i\in A_n]$, $n\geqslant 1$, $1\leqslant i\leqslant\overline{\alpha}_n$. Furthermore, because $|A_n|=\sum_{i=1}^{\infty}\mathbf{I}[i\in A_n]$, $\mathbf{E}|A_n|=\sum_{i=1}^{\infty}p_{n,i}$, and (3.4) follows from (2.12). $\square$

4 Remarks and examples

Remark 1. Let us remark that the assumption $\mathbf{E}|X|<\infty$ in Theorem 3 is superfluous. Replacing $X-\mathbf{E}X$ by X in the proof, we obtain the SLLN of the following form:

$$\lim_{n\to\infty}\frac{\sum_{i\in A_n}X_i}{b_n}=0.$$

However, in this paper, we emphasize the influence of decomposition (2.2) on SLLN, and therefore we assume the existence of the first moment of X.

Let us consider the increments of terms $V(A_n)$, $Z(A_n)$, $S(A_n)$, that is, the terms $V_{(n)}=V(A_{n+1}\setminus A_n)$, $Z_{(n)}=Z(A_{n+1}\setminus A_n)$, $S_{(n)}=S(A_{n+1}\setminus A_n)$, $n\geqslant 1$, respectively.

Remark 2. Let $\{X_n,\, n\geqslant 1\}$ be a sequence of independent random variables, and let $\{A_n,\, n\geqslant 1\}$ be a sequence of almost surely increasing $2^{\mathbb{N}}$-valued random sets independent of $\{X_n,\, n\geqslant 1\}$. We have:

(a) The increments of $V(A_n)$ are uncorrelated.
(b) The increments of $Z(A_n)$ and $S(A_n)$ may be correlated.
(c) The increments of $V(A_n)$, $Z(A_n)$, and $S(A_n)$ may be dependent.

Proof of Remark 2*(a).* Because $\{A_n,\, n\geqslant 1\}$ and $\{X_n,\, n\geqslant 1\}$ are independent, and $\{X_n,\, n\geqslant 1\}$ are independent random variables, we have

$$\begin{aligned}\mathrm{Cov}(V_{(n)},V_{(n+k)})&=\mathrm{Cov}\biggl(\sum_{j\in A_{n+1}\setminus A_n}(X_j-\mathbf{E}X_j),\sum_{j\in A_{n+k+1}\setminus A_{n+k}}(X_j-\mathbf{E}X_j)\biggr)\\&=\sum{}'\sum_{j\in B\setminus A}\sum_{k\in D\setminus C}\mathbf{E}(X_j-\mathbf{E}X_j)(X_k-\mathbf{E}X_k)\\&\quad\times\mathbf{P}[A_n=B_1,\ A_{n+1}=B_2,\ A_{n+k}=B_3,\ A_{n+k+1}=B_4]=0,\end{aligned}$$

where the summation in $\sum'$ is taken over all possible sets $\{B_k, 1\leqslant k\leqslant 4\}$ such that $B_i\subset B_{i+1}$, $i=1,2,3$, and $B_4\subset\{1,2,\dots,\overline{\alpha}_{n+k+1}\}$ (cf. (1.1)). $\square$

For points (b) and (c) of Remark 2, we construct the following example.

Example 1. Let $\{X_n,\, n\geqslant 1\}$ be a sequence of independent Gaussian $N(\mu_n,\sigma_n)$ random variables defined on the probability space $([0,1],\mathcal{B},\lambda)$ ($\mathcal{B}$ denotes the family of Borel subsets of $[0,1]$, and λ is the Lebesgue measure on $[0,1]$). On the same probability space $([0,1],\mathcal{B},\lambda)$, we will define the sequence of random sets

$\{A_n,\, n \geqslant 1\}$ and expand the definitions of $\{X_n, A_n,\, n \geqslant 1\}$ on the product space $([0,1]^2, \mathcal{B}\times\mathcal{B}, \lambda\times\lambda)$ such that $\{X_n,\, n \geqslant 1\}$ and $\{A_n,\, n \geqslant 1\}$ will be independent ($X_n(\omega_1,\omega_2) = X_n(\omega_1)$, $A_n(\omega_1,\omega_2) = A_n(\omega_2)$). For some $0 < p < 1$, we define

$$I_n = \begin{cases} [0,p) & \text{if } n \text{ is even,} \\ [p,1] & \text{if } n \text{ is odd,} \end{cases}$$

and let

$$A_n = \begin{cases} \{1,2,\dots,n,\, n+1\}, & \omega \in [0,1] \times I_n, \\ \{1,2,\dots,n,\, n+2\}, & \omega \in [0,1] \times ([0,1] \setminus I_n). \end{cases} \tag{4.1}$$

Obviously $A_n \subset A_{n+1}$ a.s. $n \geqslant 1$. In the case where n is even, $n+1$ is odd, and $I_n = [0,p)$, $I_{n+1} = [p,1]$. Therefore for $\omega \in [0,1] \times [0,p)$, $A_n(\omega)$ and $A_{n+1}(\omega)$ are defined by the first and second formulas in (4.1), respectively. Thus

$$\begin{aligned} A_{n+1}(\omega) \setminus A_n(\omega) &= \{1,2,\dots,n,\, n+1,\, n+3\} \setminus \{1,2,\dots,n,\, n+1\} \\ &= \{n+3\}, \end{aligned}$$

whereas for $\omega \in [0,1] \times [p,1]$ we take $A_{n+1}(\omega)$ defined by first formula and $A_n(\omega)$ by the second formula in (4.1). Thus

$$\begin{aligned} A_{n+1}(\omega) \setminus A_n(\omega) &= \{1,2,\dots,n,\, n+1,\, n+2\} \setminus \{1,2,\dots,n,\, n+2\} \\ &= \{n+1\}. \end{aligned}$$

Proceeding similarly for the case of odd n, we establish

$$A_{n+1}(\omega) \setminus A_n(\omega) = \begin{cases} \{n+1\}, & \omega \in [0,1] \times ([0,1] \setminus I_n), \\ \{n+3\}, & \omega \in [0,1] \times I_n. \end{cases}$$

Proof of items (b) *and* (c) *of Remark* 2. In Example 1, from the independency of $\{X_n,\, n \geqslant 1\}$ and $\{A_n,\, n \geqslant 1\}$ we have

$$\begin{aligned} \operatorname{Cov}(S_{(n)}, S_{(n+k)}) &= \operatorname{Cov}\biggl(\sum_{j \in A_{n+1} \setminus A_n} X_j, \sum_{j \in A_{n+k+1} \setminus A_{n+k}} X_j\biggr) \\ &= (-1)^{I[k \text{ is odd}]} p(1-p)(\mu_{n+3} - \mu_{n+1})(\mu_{n+k+3} - \mu_{n+k+1}), \end{aligned}$$

where $I[B]$ is the indicator (or characteristic function) of event B. Now if $\mu_n = n\delta$ for some $\delta > 0$, then we have

$$\operatorname{Cov}(S_{(n)}, S_{(n+k)}) = (-1)^{I[k \text{ is odd}]} 4p(1-p)\delta^2 \neq 0.$$

By similar computations we get

$$\operatorname{Cov}(Z_{(n)}, Z_{(n+k)}) = (-1)^{I[k \text{ is odd}]} 4p(1-p)\delta^2 \neq 0,$$

which ends the proof of (b).

For (c), let us assume that k is even and $\mu_n = 0$. Then

$$\begin{aligned} \phi_{V_{(n)}, V_{(n+k)}}(t,s) &= \mathbf{P}[I_n] \mathrm{e}^{-t^2\sigma_{n+3}^2/2 - s^2\sigma_{n+k+3}^2/2} + \bigl(1 - \mathbf{P}[I_n]\bigr) \mathrm{e}^{-t^2\sigma_{n+1}^2/2 - s^2\sigma_{n+k+1}^2/2}, \\ \phi_{V_{(n)}}(t) &= \mathbf{P}[I_n] \mathrm{e}^{-t^2\sigma_{n+3}^2/2} + \bigl(1 - \mathbf{P}[I_n]\bigr) \mathrm{e}^{-t^2\sigma_{n+1}^2/2}, \\ \phi_{V_{(n+k)}}(s) &= \mathbf{P}[I_n] \mathrm{e}^{-s^2\sigma_{n+k+3}^2/2} + \bigl(1 - \mathbf{P}[I_n]\bigr) \mathrm{e}^{-s^2\sigma_{n+k+1}^2/2}, \end{aligned}$$

where $\phi_X(t)$ denotes the characteristic function of X. Now we put $\sigma_n^2 = \log(\mathrm{e}n/(n+1))$, $t = s = \sqrt{2}$, $p = 0.5$ (note that $\sigma_n \to 1$, as $n \to \infty$). Then we have

$$\begin{aligned}
&\phi_{V_{(n)},V_{(n+k)}}(\sqrt{2},\sqrt{2}) - \phi_{V_{(n)}}(\sqrt{2})\phi_{V_{(n+k)}}(\sqrt{2}) \\
&\quad = \frac{1}{2\mathrm{e}^2}\left(\frac{(n+4)(n+k+4)}{(n+3)(n+k+3)} + \frac{(n+2)(n+k+2)}{(n+1)(n+k+1)}\right) \\
&\qquad - \frac{1}{4\mathrm{e}^2}\left(\frac{n+4}{n+3} + \frac{n+2}{n+1}\right)\left(\frac{n+k+4}{n+k+3} + \frac{n+k+2}{n+k+1}\right) \\
&\quad = \frac{1}{4\mathrm{e}^2}\left(\frac{1}{n+k+1} - \frac{1}{n+k+3}\right)\left(\frac{1}{n+1} - \frac{1}{n+3}\right) \\
&\quad = \frac{1}{\mathrm{e}^2(n+1)(n+3)(n+k+1)(n+k+3)} \neq 0.
\end{aligned}$$

Obviously, correlated random variables are dependent, and thus statement (c) follows from (b). □

Remark 3. Defining $\{X_n, A_n,\, n \geqslant 1\}$ as in Example 1 with $\mu_n = n\delta$, $\sigma_n = 1$, $n \geqslant 1$, $\delta > 0$, Theorem 1 fails for every choice of the sequence $\{b_n,\, n \geqslant 1\}$, whereas if for some $q \geqslant 1$,

$$\sum_{n=1}^{\infty}\sum_{k=n}^{\infty}\frac{(k+1)^{q-1} - k^{q-1}}{b_k^{2q}} < \infty, \tag{4.2}$$

$$\sum_{n=1}^{\infty}\frac{n^{q-1}}{b_n^{2q}} < \infty, \tag{4.3}$$

$$\sum_{n=1}^{\infty}\mathrm{e}^{-\epsilon b_n} < \infty, \tag{4.4}$$

then (2.7)–(2.9) hold.

Proof. We note that when $X_j \sim N(j\delta, 1)$, $X_j - \mathbf{E}X_j \sim N(0,1)$ and $\mathbf{E}|X_j - \mathbf{E}X_j|^{2q} = \kappa_q$, say. Then we have

$$\sum_{k=1}^{\infty}\sup_{n\geqslant 1}\left|\operatorname{Cov}\left(\sum_{i\in A_{n+1}\setminus A_n} X_i, \sum_{i\in A_{n+k+1}\setminus A_{n+k}} X_i\right)\right|\frac{\log^2 k}{k} = 4p(1-p)\delta^2\sum_{n=1}^{\infty}\frac{\log^2 n}{n} = \infty,$$

so that (1.5) fails. Now we show that there exists a sequence $\{b_n,\, n \geqslant 1\}$ such that the assumptions of Theorem 2 hold. Because

$$\mathbf{E}\sum_{i\in A_n} X_i = \frac{1}{2}\big((n+1)(n+2)+2\big)\delta - \mathbf{P}[I_n]\delta,$$

$$\mathbf{E}\prod_{i\in A_n}\mathrm{e}^{\pm\mathbf{E}X_i} = \mathrm{e}^{\pm(n+1)(n+2)\delta/2}\big(\mathbf{P}[I_n] + \big(1-\mathbf{P}[I_n]\big)\mathrm{e}^{\pm\delta}\big),$$

(2.5) and (2.6) hold if and only if $\sum_{n=1}^{\infty}\mathrm{e}^{-\epsilon b_n} < \infty$. Conditions (2.3) and (2.4) are reduced to those (4.2) and (4.3), respectively. For example, we may put $b_n = n$, $q = 2$, and then (4.2)–(4.4) are fulfilled. □

References

1. I. Fazekas and O. Klesov, A general approach to the strong law of large numbers, *Theory Probab. Appl.*, **45**(3):436–449, 2001.

2. T.-C. Hu, A. Rosalsky, and A. Volodin, On convergence properties of sums of dependent random variables under second moment and covariance restrictions, *Stat. Probab. Lett.*, **78**(14):1999–2005, 2008.

3. T.-C. Hu and R.L. Taylor, On the strong law for arrays and for the bootstrap mean and variance, *Int. J. Math. Math. Sci.*, **20**(2):375–382, 1997.

4. T.-C. Hu and N.C. Weber, A note on strong convergence of sums of dependent random variables, *J. Probab. Stat.*, **2009**:873274, 2009.

5. G. Matheron, *Random Sets and Integral Geometry*, Wiley Ser. Probab. Math. Stat., Wiley, New York, 1975.

6. I. Molchanov, *Theory of Random Sets*, Probab. Appl., Springer, London, 2005.

7. H.T. Nguyen, *An Introduction to Random Sets*, Chapman & Hall/CRC, Boca Raton, FL, 2006.

8. A. Rosalsky and G. Stoica, On the strong law of large numbers for identically distributed random variables irrespective of their joint distributions, *Stat. Probab. Lett.*, **80**(17–18):1265–1270, 2010.