

Quantitative Strong Laws of Large Numbers*

Morenikeji Neri[†] 

Abstract

We give novel computationally effective limit theorems for the convergence of the Cesaro-means of certain sequences of random variables. These results are intimately related to various Strong Laws of Large Numbers and, in that way, allow for the extraction of quantitative versions of many of these results. In particular, we produce optimal polynomial bounds in the case of pairwise independent random variables with uniformly bounded variance, improving on known results; furthermore, we obtain a new Baum-Katz type result for this class of random variables. Lastly, we are able to provide a fully quantitative version of a recent result of Chen and Sung that encompasses many limit theorems in the Strong Laws of Large Numbers literature.

Keywords: Laws of Large Numbers; large deviations; limit theorems; proof mining.

MSC2020 subject classifications: 60F10; 60F15.

Submitted to EJP on July 30, 2024, final version accepted on January 19, 2025.

1 Introduction

Throughout this article, fix a probability space $(\Omega, \mathbf{F}, \mathbb{P})$ which all the random variables we work with will act on.

The classical Strong Law of Large Numbers is the following fundamental result due to Kolmogorov:

Theorem 1.1 (The classical Strong Law of Large Numbers, c.f. Theorem 6.6.1 of [20]). *Suppose X_1, X_2, \dots are independent, identically distributed (iid) real-valued random variables with $\mathbb{E}(|X_1|) < \infty$. Then,*

$$\frac{1}{n} \sum_{i=1}^n X_i \rightarrow \mathbb{E}(X_1) \tag{1.1}$$

almost surely, that is, with probability 1.

*The author was partially supported by the EPSRC Centre for Doctoral Training in Digital Entertainment (EP/L016540/1).

[†]University of Bath, United Kingdom. E-mail: mn728@bath.ac.uk <https://kejineri.github.io>

For ease write $S_n := \sum_{i=1}^n X_i$ and assume $\mathbb{E}(X_1) = 0$. By multiple applications of the continuity of the probability measure, one can show (following the notation in [47]) that the conclusion of the Strong Law of Large Numbers is equivalent to the sequence of real numbers

$$P_{n,\varepsilon}^* := \mathbb{P} \left(\sup_{m \geq n} \left| \frac{S_m}{m} \right| > \varepsilon \right)$$

converging to 0 as $n \rightarrow \infty$ ¹, for any $\varepsilon > 0$, and so the computational content of the Strong Law of Large Numbers can be given an explicit meaning by studying the convergence speed of this sequence of real numbers. That is, for each $\varepsilon > 0$ can we find a function $r_\varepsilon : (0, \infty) \rightarrow (0, \infty)$ such that

$$\forall \lambda > 0 \forall n \geq r_\varepsilon(\lambda), P_{n,\varepsilon}^* \leq \lambda?$$

Such a function is known as a rate of convergence, more generally a rate of convergence for a sequence of real numbers $\{x_n\}$ converging to some real number x is any function $r : (0, \infty) \rightarrow (0, \infty)$ satisfying

$$\forall \varepsilon > 0 \forall n \geq r(\varepsilon), (|x_n - x| \leq \varepsilon).$$

A strictly decreasing rate of convergence, r , that has an inverse can easily be converted to an asymptotic upper bound as follows,

$$\forall n \in \mathbb{N}, (|x_n - x| \leq r^{-1}(n)).$$

Furthermore, it is also clear that a function bounding a rate of convergence will also be a rate of convergence.

A particular instance of a rate of convergence that will occur a lot in this article is that for the partial sums of a convergent series, $\sum_{i=1}^\infty x_i < \infty$, of nonnegative real numbers $\{x_i\}$. In this case, a rate of convergence for the partial sums, $\sum_{i=1}^n x_i$, to their limit, $\sum_{i=1}^\infty x_i$, will be a function $r : (0, \infty) \rightarrow (0, \infty)$ satisfying, for all $\varepsilon > 0$ and $n \geq r(\varepsilon)$,

$$\sum_{i=n+1}^\infty x_i \leq \varepsilon.$$

It has been of interest to study sufficient conditions for $\frac{S_n}{n} \rightarrow 0$ almost surely when we weaken the iid condition. In [17], Etemadi demonstrates that $\frac{S_n}{n} \rightarrow 0$ almost surely if we only assume the random variables are pairwise iid. Furthermore, Etemadi's proof is rather elementary compared to Kolmogorov's original proof of this result for iid random variables.

To obtain $\frac{S_n}{n} \rightarrow 0$, almost surely, dropping the assumption that the random variables are identically distributed requires that other assumptions are placed on the random variables instead. If we keep the independence condition, one can obtain another classical result of Kolmogorov:

Theorem 1.2 (Kolmogorov's Strong Law of Large Numbers). *Suppose $\{X_n\}$ is a sequence of independent random variables, each with expected value 0 and*

$$\sum_{n=1}^\infty \frac{\text{Var}(X_n)}{n^2} < \infty. \tag{1.2}$$

Then $\frac{S_n}{n} \rightarrow 0$ almost surely.

¹The proof of equivalence is similar to how one proves Egorov's theorem and is well known in the literature.

In [13] Csörgő et al. demonstrate that Kolmogorov’s Strong Law of Large Numbers does not hold if we weaken the independence condition, in Theorem 1.2, to even pairwise independence (see [13, Theorem 3]). They instead obtain the following:

Theorem 1.3 (cf. Theorem 1 of [13]). *Suppose $\{X_n\}$ is a sequence of pairwise independent random variables, each with expected value 0, satisfying (1.2) and*

$$\frac{1}{n} \sum_{k=1}^n \mathbb{E}(|X_k|) = O(1). \tag{1.3}$$

Then $\frac{S_n}{n} \rightarrow 0$ almost surely.

This article will study the quantitative content of the Strong Law of Large Numbers when the iid assumption is weakened by calculating explicit rates of convergences for $P_{n,\varepsilon}^*$. We will do this by proving an abstract technical theorem (in Section 2) whose quantitative content captures the key combinatorial idea in the proof of Theorem 1.3. Our abstract theorem will allow us to produce our first main contribution:

Theorem 1.4. *Suppose $\{X_n\}$ is a sequence of pairwise independent random variables satisfying, $\mathbb{E}(X_n) = 0$, $\mathbb{E}(|X_n|) \leq \tau$ and $\text{Var}(X_n) \leq \sigma^2$, for all $n \in \mathbb{N}$ and some $\tau, \sigma > 0$. There exists a universal constant $\kappa \leq 1536$ such that for all $0 < \varepsilon \leq \tau$,*

$$P_{n,\varepsilon}^* \leq \frac{\kappa \sigma^2 \tau}{n \varepsilon^3}.$$

The above is an improvement of the best known asymptotic upper bound, given by Luzia [36], who showed (with notation as in Theorem 1.4) that for all $\beta > 1$ and $0 < \varepsilon \leq \tau$, there exists $N(\beta, \varepsilon, \tau)$ such that, for all $n \geq N(\beta, \varepsilon, \tau)$,

$$P_{n,\varepsilon}^* \leq \frac{\sigma^2}{n \varepsilon^2} (C_\beta + D_\beta \log(n)^{\beta-1}),$$

for some $C_\beta, D_\beta > 0$ depending only on β . Thus, if we fix ε, σ , and τ the above tells us that for each $\beta > 1$,

$$P_{n,\varepsilon}^* = O\left(\frac{\log(n)^{\beta-1}}{n}\right)$$

as $n \rightarrow \infty$, for such a class of random variables, whereas the bound in Theorem 1.4 yields:

$$P_{n,\varepsilon}^* = O\left(\frac{1}{n}\right).$$

Observe we can take $\tau = \sigma$ by Jensen’s inequality, so Theorem 1.4 in particular yields

$$P_{n,\varepsilon}^* \leq \frac{\kappa \sigma^3}{n \varepsilon^3}.$$

The above bound bears a resemblance to the bound one obtains in the case where the random variables are assumed to be independent, namely, Hájek and Rényi [21] state that if $\{X_n\}$ is a sequence of independent random variables each with expected value 0 then, for all $\varepsilon > 0$ and $n < m$

$$\mathbb{P}\left(\max_{n \leq k \leq m} \left|\frac{S_k}{k}\right| > \varepsilon\right) \leq \frac{1}{\varepsilon^2} \left(\frac{1}{n^2} \sum_{k=1}^n \text{Var}(X_k) + \sum_{k=n+1}^m \frac{\text{Var}(X_k)}{k^2}\right).$$

From this, one easily obtains that if $\{X_n\}$ is also assumed to have a common bound on their variance σ^2 ,

$$P_{n,\varepsilon}^* \leq \frac{2\sigma^2}{n \varepsilon^2}. \tag{1.4}$$

However, the Hájek and Rényi inequality generalises Kolmogorov’s inequality, which is known to fail for pairwise independent random variables; therefore, more work is needed to obtain a bound in this case.

In addition, through a simple construction, we demonstrate that $O(1/n)$ is the best polynomial bound for $P_{n,\varepsilon}^*$ in the case that $\{X_n\}$ is iid with finite variance. That is, for each $\delta > 0$, we construct a sequence of random variables with finite variance, satisfying

$$\mathbb{P}\left(\sup_{m \geq n} \left| \frac{S_m}{m} \right| > \varepsilon\right) \geq \frac{\omega}{n^{1+\delta}}$$

for some $\omega > 0$ depending only on δ , for all $0 < \varepsilon \leq 1$. All of these results are in Section 3.

The following result is mainly attributed to Baum and Katz [6] as well as Chow [11]:

Theorem 1.5. *Let $\{X_n\}$ be a sequence of iid random variables satisfying $\mathbb{E}(X_1) = 0$ and let $r \geq -1$. Then for all $\varepsilon > 0$, the following are equivalent:*

- (i) $\mathbb{E}(|X_1|^{r+2}) < \infty$,
- (ii) $\sum_{n=1}^{\infty} n^r \mathbb{P}\left(\left|\frac{1}{n}S_n\right| > \varepsilon\right) < \infty$,
- (iii) $\sum_{n=1}^{\infty} n^r \mathbb{P}\left(\sup_{m \geq n} \left|\frac{1}{m}S_m\right| > \varepsilon\right) < \infty$,
- (iv) $\sum_{n=1}^{\infty} n^r \mathbb{P}\left(\max_{1 \leq m \leq n} |S_m| > n\varepsilon\right) < \infty$.

To prove this result, independence is crucial. Work has been done to extend this result to the case where the random variables are pairwise independent. It is clear that (iii) and (iv) both imply (ii) in the non-independent case. However, as noted in [5], it is possible that (iv) is strictly stronger than (ii) in the non-independent case, and work has been done in establishing the convergence of (iv) in the case where the random variables are pairwise iid. Many authors have established the following theorem, but the result goes back to Rio [45].

Theorem 1.6. *Suppose $\{X_n\}$ are pairwise independent, identically distributed random variables with $\mathbb{E}(X_1) = 0$. For all $-1 \leq r < 0$: $\mathbb{E}(|X_1|^{2+r}) < \infty$ iff*

$$\sum_{n=1}^{\infty} n^r \mathbb{P}\left(\max_{1 \leq m \leq n} |S_m| > n\varepsilon\right) < \infty$$

for all $\varepsilon > 0$.

There does not appear to be any results in the literature for the convergence of the sum (iii), assuming the random variables are pairwise independent. However, a simple application of Theorem 1.4 gives the following:

Corollary 1.7. *Suppose $\{X_n\}$ are pairwise independent, identically distributed random variables with $\mathbb{E}(X_1) = 0$ and $\text{Var}(X_1) < \infty$. Then, for all $\varepsilon > 0$ and $r < 0$:*

$$\sum_{n=1}^{\infty} n^r P_{n,\varepsilon}^* < \infty.$$

Proof. This result simply follows from the fact that $P_{n,\varepsilon}^* = O\left(\frac{1}{n}\right)$. □

Furthermore, it appears to be open whether it is the case that condition (iii) in Theorem 1.5 holds in the case $r = 0$ and if the random variables are only assumed to be pairwise independent, which is the case for iid random variables, by Theorem 1.5.

In [17], Etemaidi’s novel insight in demonstrating that $\frac{S_n}{n} \rightarrow 0$ almost surely for pairwise iid random variables was that one could first assume that the random variables were nonnegative, in which case one can take advantage of the monotonicity of the

partial sums. The general case is then obtained by using the decomposition of a random variable into its positive and negative parts (that is, writing a random variable, X , as $X = X^+ - X^-$ where $X^+ = \max\{X, 0\}$ and $X^- = \max\{-X, 0\}$). Due to this insight, there has been a lot of interest in studying when $\frac{S_n}{n} \rightarrow 0$, almost surely, for nonnegative random variables that are not assumed to be iid, as e.g. in Petrov [41], which was later generalised by Korchevsky et al. [32] and further generalised again by Korchevsky in [30]. In addition, Chandra et al. [9] generalise Theorem 1.3, with Chen and Sung [10] later producing a result which unifies [9] and [30], as well as generalising results from [8, 26, 46, 23]. The proofs of all the results Chen and Sung generalise are adaptations of the proof of Theorem 1.3, and they established the following general sufficient condition, which encompasses all the results mentioned:

Theorem 1.8 (cf. Theorem 2.1 of [10]). *Let $\{X_n\}$ be a sequence of nonnegative random variables with finite p -th moment (for some fixed $p \geq 1$) and respective expected values $\{\mu_n\}$. Let $S_n := \sum_{k=1}^n X_k$ and $z_n := \sum_{i=1}^n \mu_i$. Suppose that*

$$\frac{z_n}{n} = O(1)$$

and that there exists a sequence of nonnegative real numbers $\{\gamma_n\}$ satisfying

- $\mathbb{E}(|S_n - z_n|^p) \leq \sum_{k=1}^n \gamma_k$,
- $\sum_{n=1}^{\infty} \frac{\gamma_n}{n^p} < \infty$.

Then

$$\frac{S_n}{n} - \frac{z_n}{n} \rightarrow 0$$

almost surely.

Building on our first main contribution, which was an improvement of the bound in [36], our second main contribution will be a fully quantitative version of Theorem 1.8.

Theorem 1.9. *Let $\{X_n\}$ be a sequence of nonnegative random variables with finite p -th moment (for some fixed $p \geq 1$) and respective expected values $\{\mu_n\}$. Let $S_n := \sum_{k=1}^n X_k$ and $z_n := \sum_{i=1}^n \mu_i$. Suppose there exists a sequence of nonnegative real numbers $\{\gamma_n\}$ satisfying*

$$\mathbb{E}(|S_n - z_n|^p) \leq \sum_{k=1}^n \gamma_k$$

and

$$\sum_{m=1}^{\infty} \frac{\gamma_m}{m^p} \leq \Gamma$$

for some $\Gamma \geq 1$, with the partial sums of the above series converging to their limit with a strictly decreasing rate of convergence Ψ . Furthermore, assume for all $n \in \mathbb{N}$,

$$\frac{z_n}{n} \leq W,$$

for some $W \geq 1$. Then for all $0 < \varepsilon \leq 1, \lambda > 0$ and all

$$n \geq A_p \left(\frac{W\Gamma}{\lambda\varepsilon^{p+1}} \right)^{\frac{1}{p}} \Psi \left(\frac{B_p \lambda \varepsilon^{p+1}}{W} \right),$$

it holds that

$$\mathbb{P} \left(\sup_{m \geq n} \left| \frac{S_m}{m} - \frac{z_m}{m} \right| > \varepsilon \right) \leq \lambda.$$

Here, A_p and B_p are constants that only depend on p .

From Theorem 1.9, one can obtain quantitative versions of many of the “Strong Law of Large Numbers”-type results discussed above. For example, we can easily obtain a quantitative version of Theorem 1.3:

Theorem 1.10. *Suppose $\{X_n\}$ is a sequence of pairwise independent random variables, each with expected value 0 and finite variance. Let $S_n := \sum_{k=1}^n X_k$ and $z_n := \sum_{i=1}^n \mathbb{E}(|X_i|)$. Further, assume*

$$\sum_{n=1}^{\infty} \frac{\text{Var}(X_n)}{n^2} \leq \Gamma$$

for some $\Gamma \geq 1$ and that the partial sums of the above series converge to their limit with a strictly decreasing rate of convergence Ψ . Furthermore, assume for all $n \in \mathbb{N}$,

$$\frac{z_n}{n} \leq W,$$

for some $W \geq 1$. For all $0 < \varepsilon \leq 1, \lambda > 0$ and all

$$n \geq A \left(\frac{W\Gamma}{\lambda\varepsilon^3} \right)^{\frac{1}{2}} \Psi \left(\frac{B\lambda\varepsilon^3}{W} \right),$$

it holds that

$$P_{n,\varepsilon}^* = \mathbb{P} \left(\sup_{m \geq n} \left| \frac{S_m}{m} \right| > \varepsilon \right) \leq \lambda.$$

Here, A and B are universal constants.

All of these results are in Section 4.

1.1 Related work

The study of large deviations in the Strong Law of Large Numbers starts with Cramér’s 1938 article [12], where he determined large deviation probabilities for the sums of iid random variables up to asymptotic equivalence. Furthermore, in this work, he introduced the moment generating function condition (the moment generating function of the random variables is finite on an interval), which has become a standard assumption in this area.

The subsequent notable work in this direction was in 1960 by Bahadur and Ranga Rao [4], where they built on Cramér’s work to calculate large deviation probabilities for the weak law of large numbers up to asymptotic equivalence (again assuming the moment generating function condition from Cramér).

Then, in 1975, Siegmund [47] (see also [18]) was able to determine $P_{n,\varepsilon}^*$ up to asymptotic equivalence, again assuming the moment generating function condition. Thus, [47] provides the first quantitative interpretation of the Strong Law of Large Numbers. Furthermore, Siegmund’s bounds heavily depend on the distribution of the random variables, so although their results are much stronger than those in this article, they assume a lot more about the sequence of random variables.

Not much work has been done to study the large deviations without strong conditions, such as the moment-generating function condition. This may be because, for weaker conditions, one cannot hope to calculate these probabilities up to asymptotic equivalence. The best we can hope for are bounds on the large deviation probabilities. As discussed already, in 2018, Luzia [36] obtained distribution independent bounds under milder assumptions on the random variables, which are improved in this article.

Work has been done to study the large deviation probabilities for sequences of random variables that are not necessarily identically distributed. In 1943, Feller was able to generalise Cramér’s 1938 article to random variables that are not necessarily

identically distributed; however, his assumptions were too restrictive (he assumed the random variables only took values in finite intervals) that the result was not a complete generalisation of Cramér’s. Petrov [40], in 1954, was able to provide a full generalisation of Cramér’s result and has been able to strengthen this result (by relaxing the moment generating function condition) a further two times, with the most recent in 2006 [42] jointly with Robinson.

We also note that Pointwise Ergodic Theorem can be used to show that the Strong Law of Large Numbers also holds for stationary sequences of random variables and obtaining rates for $P_{n,\varepsilon}^*$, in this case, has been of great interest. For example, Gaposhkin [19] provides an asymptotic upper bound for $P_{n,\varepsilon}^*$ (which they demonstrate is optimal) for second-order stationary sequences of random variables with finite variance, with more recent work being done by Kachurovskii on this topic, see [25, 24].

Baum-Katz type rates have been obtained in the Strong Law of Large Numbers for nonnegative random variables where both the independence and identical distribution conditions are weakened. In 2018, Korchevsky [31] obtained a Baum-Katz type rate for the Chen-Sung Strong Law of Large Numbers [10] under stronger assumptions. This result generalised the work of Kuczmaszewska [35], in 2016, who was able to obtain rates for a Strong Law of Large Numbers result of Korchevsky in [30], under stronger assumptions. No Baum-Katz type rates have been found for the full results in [30] and [10].

Lastly, Baum-Katz type results can be used to obtain results concerning large deviation probabilities. For example, if $\{X_n\}$ are iid random variables with, $\mathbb{E}(X_1) = 0$ and $\text{Var}(X_1) < \infty$ then condition (iii) of Theorem 1.5 with $r = 0$ implies,

$$P_{n,\varepsilon}^* = o\left(\frac{1}{n}\right).$$

The above is a stronger result than what one gets in (1.4) through the Hájek and Rényi inequality (although one must assume more, namely that the random variables are identically distributed). Unlike the bound in Theorem 1.4, this result is ineffective in the sense that it does not explicitly tell you the constant C such that $P_{n,\varepsilon}^* \leq \frac{C}{n}$, in addition, one cannot determine, a priori, that such a constant is independent of the distribution of the random variables. Furthermore, the bound in Theorem 1.4 only requires the assumption that the random variables are pairwise independent.

2 A general theorem

In this section, we shall state and prove the general quantitative theorem we alluded to in Section 1. This theorem will be a quantitative version of (a generalisation of) a critical step in proving [13, Theorem 1], which is a result that has been modified many times to obtain various “Strong Law of Large Numbers”-type results as discussed in the introduction.

For this, we now first introduce the following definitions that are mostly as presented in [13]: Let $\{X_n\}$ be a sequence of nonnegative random variables with respective expected values $\{\mu_n\}$. Let $S_n := \sum_{k=1}^n X_k$, $z_n := \sum_{i=1}^n \mu_i$ and suppose there exists $W > 0$ such that

$$\frac{z_n}{n} \leq W$$

for all $n \in \mathbb{N}$. Further, we make use of the following definitions:

- For each $\delta > 0$, let $L_\delta := \lfloor \frac{W}{\delta} \rfloor$.
- For each $\delta > 0$, $\alpha > 1$ and natural numbers m and $0 \leq s \leq L_\delta$, let

$$C_{\alpha,s,\delta,m} := \left\{ \alpha^m \leq n < \alpha^{m+1} \mid \frac{z_n}{n} \in [s\delta, (s+1)\delta) \right\}.$$

- Let $k_s^-(m) := \min C_{\alpha,s,\delta,m}$ and $k_s^+(m) := \max C_{\alpha,s,\delta,m}$ if $C_{\alpha,s,\delta,m}$ is nonempty.
- Let $k_s^-(m) = k_s^+(m) := \lfloor \alpha^m \rfloor$ if $C_{\alpha,s,\delta,m}$ is empty.

One should note that $k_s^+(m)$ and $k_s^-(m)$ depend on δ, α but (following the convention of [13]) we hid this dependence to make the notation less cumbersome. We shall also adopt the convention (used in [13]) that $k_s^\pm(m)$ being used in a relationship (an equation, an inequality, a limit, etc.) is short-hand for that relationship holding for both $k_s^+(m)$ and $k_s^-(m)$.

Our general theorem is now the following:

Theorem 2.1. For all $\varepsilon, \delta > 0, \alpha > 1$ and $0 \leq s \leq L_\delta$, if

$$\sum_{n=1}^{\infty} \mathbb{P} \left(\left| \frac{S_{k_s^\pm(n)}}{k_s^\pm(n)} - \frac{z_{k_s^\pm(n)}}{k_s^\pm(n)} \right| > \varepsilon \right) < \infty, \tag{2.1}$$

then ²

$$\frac{S_n}{n} - \frac{z_n}{n} \rightarrow 0$$

almost surely.

Proof. The Borel-Cantelli Lemma and (2.1) implies that for all $\varepsilon, \delta > 0, \alpha > 1$ and all $0 \leq s \leq L_\delta$:

$$\frac{1}{k_s^\pm(n)} S_{k_s^\pm(n)} - \frac{1}{k_s^\pm(n)} z_{k_s^\pm(n)} \rightarrow 0 \tag{2.2}$$

almost surely. For all $m \in \mathbb{N}$, we can take a natural number $0 \leq s \leq L_\delta$ such that

$$\frac{1}{m} z_m \in [s\delta, (s+1)\delta) \tag{2.3}$$

since $z_n/n \leq W$ and $L_\delta = \lfloor \frac{W}{\delta} \rfloor$. Thus, if we take $p \in \mathbb{N}$ such that $\alpha^p \leq m < \alpha^{p+1}$, then $m \in C_{\alpha,s,\delta,p}$ by definition, so $C_{\alpha,s,\delta,p}$ is non-empty. Therefore, $k_s^-(p) \leq m \leq k_s^+(p)$ and, since $k_s^\pm(p) \in C_{\alpha,s,\delta,p}$, we have

$$\frac{1}{k_s^\pm(p)} z_{k_s^\pm(p)} \in [s\delta, (s+1)\delta)$$

which implies, by (2.3), that

$$\left| \frac{1}{m} z_m - \frac{1}{k_s^\pm(p)} z_{k_s^\pm(p)} \right| \leq \delta. \tag{2.4}$$

Now we have the following chain of inequalities,

$$\begin{aligned} & -\delta - \left(1 - \frac{1}{\alpha}\right) W + \frac{1}{\alpha} \frac{1}{k_s^-(p)} \left(S_{k_s^-(p)} - z_{k_s^-(p)} \right) \\ & \leq -\delta - \left(1 - \frac{1}{\alpha}\right) \frac{1}{k_s^-(p)} z_{k_s^-(p)} + \frac{1}{\alpha} \frac{1}{k_s^-(p)} \left(S_{k_s^-(p)} - z_{k_s^-(p)} \right) \\ & \leq \frac{1}{m} S_{k_s^-(p)} - \frac{1}{m} z_m \\ & \leq \frac{1}{m} (S_m - z_m) \\ & \leq \frac{1}{m} S_{k_s^+(p)} - \frac{1}{k_s^+(p)} z_{k_s^+(p)} + \delta \\ & \leq \frac{\alpha}{k_s^+(p)} \left(S_{k_s^+(p)} - z_{k_s^+(p)} \right) + (\alpha - 1)W + \delta. \end{aligned} \tag{2.5}$$

²Recall that the use of the \pm notation means we are actually assuming the convergence of two sums in the premise.

Here, the first inequality follows since

$$\frac{1}{k_s^-(p)} z_{k_s^-(p)} \leq W.$$

The second inequality follows from expanding brackets, using (2.4) and the fact that $m \leq \alpha k_s^-(p)$ (since $m \in C_{\alpha,s,\delta,p}$, so by definition, $m < \alpha^{p+1}$ and $k_s^-(p) \in C_{\alpha,s,\delta,p}$, and so $\alpha^p \leq k_s^-(p)$). The third inequality follows from the fact that $\{S_n\}$ is monotone (since $\{X_n\}$ is nonnegative) and $k_s^-(p) \leq m$. The remaining inequalities are justified using similar reasoning to the above (see also [13]).

Thus, by (2.2) and the fact that $p \rightarrow \infty$ as $m \rightarrow \infty$, we have

$$-\delta - \left(1 - \frac{1}{\alpha}\right) W \leq \liminf_{n \rightarrow \infty} \frac{1}{m} (S_m - z_m) \leq \limsup_{n \rightarrow \infty} \frac{1}{m} (S_m - z_m) \leq (\alpha - 1)W + \delta$$

almost surely. So, taking $\delta \rightarrow 0$ and $\alpha \rightarrow 1$ gives our result. \square

Remark 2.2. $\{X_n\}$ (not assumed to be nonnegative) is said to converge completely to 0 if

$$\sum_{n=1}^{\infty} \mathbb{P}(|X_n| > \varepsilon) < \infty$$

for all $\varepsilon > 0$. Hsu and Robbins first introduced this notion of convergence in [22], where they demonstrated that if $\{X_n\}$ were iid random variables with finite variance (again, not assumed to be nonnegative), then

$$\frac{S_n}{n} - \mathbb{E}(X_1)$$

converges to 0 completely. Furthermore, complete convergence implies almost sure convergence by the Borel-Cantelli Lemma, so Theorem 2.1 says that if specifically chosen sub-sequences of

$$\frac{S_n}{n} - \frac{z_n}{n}$$

converge completely to 0, then

$$\frac{S_n}{n} - \frac{z_n}{n}$$

converges to 0 almost surely.

Remark 2.3. To prove [13, Theorem 1], it is shown that

$$\sum_{n=1}^{\infty} \mathbb{E} \left(\left(\frac{S_{k_s^\pm(n)}}{k_s^\pm(n)} - \frac{z_{k_s^\pm(n)}}{k_s^\pm(n)} \right)^2 \right) < \infty. \tag{2.6}$$

(2.1) in Theorem 2.1 follows from this by Chebyshev's inequality, so the result in [13] follows by our theorem. Therefore, Theorem 2.1 generalises the key step in proving [13, Theorem 1].

We now give a quantitative version of Theorem 2.1:

Theorem 2.4. Suppose for each $\varepsilon, \delta > 0, \alpha > 1$ and $0 \leq s \leq L_\delta$:

$$\sum_{n=1}^{\infty} \mathbb{P} \left(\left| \frac{S_{k_s^\pm(n)}}{k_s^\pm(n)} - \frac{z_{k_s^\pm(n)}}{k_s^\pm(n)} \right| > \varepsilon \right) < \infty. \tag{2.7}$$

Furthermore, suppose that the partial sums of both sums converge to their respective limits with a rate of convergence $\Lambda_{\varepsilon,\delta,\alpha} : \mathbb{R} \rightarrow \mathbb{R}$, independent of s .³

³We can always obtain a rate independent of s by taking the maximum value of all such rates that depend on s , as s can only take the value of finitely many natural numbers. Furthermore, if both sums (the plus one and the minus one) have different rates, we can obtain one that works for both by taking the maximum of the two rates.

More explicitly, for each $\varepsilon, \delta > 0$, $\alpha > 1$, $0 \leq s \leq L_\delta$, $\lambda > 0$ and $p \geq \Lambda_{\varepsilon, \delta, \alpha}$, we have,

$$\sum_{n=p+1}^{\infty} \mathbb{P} \left(\left| \frac{S_{k_s^-(n)}}{k_s^-(n)} - \frac{z_{k_s^-(n)}}{k_s^-(n)} \right| > \varepsilon \right) \leq \lambda \text{ and } \sum_{n=p+1}^{\infty} \mathbb{P} \left(\left| \frac{S_{k_s^+(n)}}{k_s^+(n)} - \frac{z_{k_s^+(n)}}{k_s^+(n)} \right| > \varepsilon \right) \leq \lambda.$$

Then, for all $\varepsilon > 0$,

$$\mathbb{P} \left(\sup_{m \geq n} \left| \frac{S_m}{m} - \frac{z_m}{m} \right| > \varepsilon \right) \rightarrow 0$$

with a rate of convergence given by

$$\Phi_{\varepsilon, \Lambda}(\lambda) := \alpha^{\Pi_\varepsilon(\lambda)},$$

where

$$\Pi_\varepsilon(\lambda) := \Lambda_{\frac{\varepsilon}{3\alpha}, \frac{\varepsilon}{3}, \alpha} \left(\frac{\lambda}{2} \right) + 1 \text{ and } \alpha := 1 + \frac{\varepsilon}{3W}.$$

Proof. First we observe that, for all $\delta, \lambda, \varepsilon > 0$, $\alpha > 1$, natural numbers $0 \leq s \leq L_\delta$ and $p \geq \Lambda_{\varepsilon, \delta, \alpha}(\lambda) + 1$:

$$\begin{aligned} \mathbb{P} \left(\sup_{q \geq p} \left| \frac{S_{k_s^\pm(q)}}{k_s^\pm(q)} - \frac{z_{k_s^\pm(q)}}{k_s^\pm(q)} \right| > \varepsilon \right) &= \mathbb{P} \left(\bigcup_{q=p}^{\infty} \left(\left| \frac{S_{k_s^\pm(q)}}{k_s^\pm(q)} - \frac{z_{k_s^\pm(q)}}{k_s^\pm(q)} \right| > \varepsilon \right) \right) \\ &\leq \sum_{q=p}^{\infty} \mathbb{P} \left(\left| \frac{S_{k_s^\pm(q)}}{k_s^\pm(q)} - \frac{z_{k_s^\pm(q)}}{k_s^\pm(q)} \right| > \varepsilon \right) \leq \lambda. \end{aligned}$$

Here, the last inequality follows from the fact that $p - 1 \geq \Lambda_{\varepsilon, \delta, \alpha}(\lambda)$ (and that Λ is a rate of convergence).

Now, fix $\varepsilon, \lambda > 0$ and

$$n \geq \Phi_{\varepsilon, \Lambda}(\lambda) = \alpha^{\Lambda_{\frac{\varepsilon}{3\alpha}, \frac{\varepsilon}{3}, \alpha}(\frac{\lambda}{2}) + 1}.$$

We must show,

$$\mathbb{P} \left(\sup_{m \geq n} \left| \frac{S_m}{m} - \frac{z_m}{m} \right| > \varepsilon \right) \leq \lambda.$$

Set $\delta = \frac{\varepsilon}{3}$ and observe that having $\alpha = 1 + \frac{\varepsilon}{3W}$ ensures that

$$-\frac{\varepsilon}{3} \leq -\left(1 - \frac{1}{\alpha}\right)W \text{ and } (\alpha - 1)W = \frac{\varepsilon}{3}. \tag{2.8}$$

Take $p \in \mathbb{N}$ such that $\alpha^p \leq n < \alpha^{p+1}$. Then we have

$$\alpha^{\Lambda_{\frac{\varepsilon}{3\alpha}, \frac{\varepsilon}{3}, \alpha}(\frac{\lambda}{2}) + 1} \leq n < \alpha^{p+1}$$

which implies

$$p \geq \Lambda_{\frac{\varepsilon}{3\alpha}, \frac{\varepsilon}{3}, \alpha} \left(\frac{\lambda}{2} \right) + 1.$$

Thus, by the very first step of the proof, we have, for each $0 \leq r \leq L_\delta$,

$$\mathbb{P} \left(\sup_{q \geq p} \left| \frac{1}{k_r^\pm(q)} S_{k_r^\pm(q)} - \frac{1}{k_r^\pm(q)} z_{k_r^\pm(q)} \right| > \frac{\varepsilon}{3\alpha} \right) \leq \frac{\lambda}{2}. \tag{2.9}$$

Thus, it suffices to show that there exists $0 \leq r \leq L_\delta$ such that

$$\sup_{m \geq n} \left| \frac{S_m}{m} - \frac{z_m}{m} \right| > \varepsilon$$

implies that

$$\sup_{q \geq p} \left| \frac{1}{k_r^-(q)} S_{k_r^-(q)} - \frac{1}{k_r^-(q)} z_{k_r^-(q)} \right| > \frac{\varepsilon}{3\alpha}$$

or that

$$\sup_{q \geq p} \left| \frac{1}{k_r^+(q)} S_{k_r^+(q)} - \frac{1}{k_r^+(q)} z_{k_r^+(q)} \right| > \frac{\varepsilon}{3\alpha},$$

as then we would have, for such an r ,

$$\begin{aligned} \mathbb{P} \left(\sup_{m \geq n} \left| \frac{S_m}{m} - \frac{z_m}{m} \right| > \varepsilon \right) &\leq \mathbb{P} \left(\sup_{q \geq p} \left| \frac{1}{k_r^-(q)} S_{k_r^-(q)} - \frac{1}{k_r^-(q)} z_{k_r^-(q)} \right| > \frac{\varepsilon}{3\alpha} \right) \\ &\quad + \mathbb{P} \left(\sup_{q \geq p} \left| \frac{1}{k_r^+(q)} S_{k_r^+(q)} - \frac{1}{k_r^+(q)} z_{k_r^+(q)} \right| > \frac{\varepsilon}{3\alpha} \right) \leq \lambda, \end{aligned}$$

which is what we are required to show (with the final inequality following from (2.9)).

Suppose, for contradiction, that the above was not the case. Then, for all $0 \leq r \leq L_\delta$, we have

$$\sup_{m \geq n} \left| \frac{1}{m} S_m - \frac{1}{m} z_m \right| > \varepsilon$$

and

$$\left| \frac{1}{k_r^\pm(q)} S_{k_r^\pm(q)} - \frac{1}{k_r^\pm(q)} z_{k_r^\pm(q)} \right| \leq \frac{\varepsilon}{3\alpha} \tag{2.10}$$

for all $q \geq p$. Take $m \geq n$ such that

$$\left| \frac{1}{m} S_m - \frac{1}{m} z_m \right| > \varepsilon. \tag{2.11}$$

We now use arguments similar to the proof of Theorem 2.1. We can find $0 \leq r \leq L_\delta$ such that

$$\frac{1}{m} z_m \in [r\delta, (r+1)\delta],$$

so, taking $q \in \mathbb{N}$ such that $\alpha^q \leq m < \alpha^{q+1}$, ensures that $m \in C_{\alpha,r,\delta,q}$. Furthermore, as $m \geq n$, we have $q \geq p$.

Now, since $k_r^\pm(q) \in C_{\alpha,r,\delta,q}$, we have

$$\frac{1}{k_r^\pm(q)} z_{k_r^\pm(q)} \in [r\delta, (r+1)\delta]$$

which implies

$$\left| \frac{1}{m} z_m - \frac{1}{k_r^\pm(q)} z_{k_r^\pm(q)} \right| \leq \delta.$$

Now, following the exact same reasoning as (2.5), we have

$$\begin{aligned} -\delta - \left(1 - \frac{1}{\alpha}\right) W + \frac{1}{\alpha} \frac{1}{k_r^-(q)} \left(S_{k_r^-(q)} - z_{k_r^-(q)} \right) \\ \leq \frac{1}{m} (S_m - z_m) \\ \leq \frac{\alpha}{k_r^+(q)} \left(S_{k_r^+(q)} - z_{k_r^+(q)} \right) + (\alpha - 1)W + \delta. \end{aligned}$$

So, (2.8) implies that (recalling that $\delta = \varepsilon/3$),

$$\begin{aligned} -\frac{2\varepsilon}{3} + \frac{1}{\alpha} \frac{1}{k_r^-(q)} \left(S_{k_r^-(q)} - z_{k_r^-(q)} \right) &\leq \frac{1}{m} (S_m - z_m) \\ &\leq \frac{\alpha}{k_r^+(q)} \left(S_{k_r^+(q)} - z_{k_r^+(q)} \right) + \frac{2\varepsilon}{3}. \end{aligned} \tag{2.12}$$

Now, the above and (2.10) (and the fact that $\alpha > 1$) implies $|1/m(S_m - z_m)| \leq \varepsilon$, which contradicts (2.11). \square

Remark 2.5. Applying Theorem 2.4 to obtain explicit rates of convergences for results related to the Strong Laws of Large Numbers requires that we must find an explicit rate of convergence for the (2.1). It is a well known result in computability theory that the Monotone Convergence Theorem is computationally ineffective, in the sense that a computable bound on a series does not necessarily yield a computable rate of convergence (a monotone bounded sequence can converge arbitrarily slowly, see [15]). Since this step still features in the proof of Theorem 2.1, we require a rate of convergence for (2.1) in Theorem 2.4 to provide a construction for a rate of convergence for the conclusion. Furthermore, when the strategy captured by Theorem 2.1 is applied to obtain results related to the Strong Law of Large Numbers, one typically just bounds (2.1), thus more work will be required than what is offered by the proof of these results as we will have to explicitly calculate rates of convergences. This will be the case when we apply Theorem 2.4 to prove Theorem 1.9 in Section 4.

3 Application I: pairwise independent with finite variance

In this section, we shall prove Theorem 1.4. First, we calculate a rate under the assumption that the random variables are nonnegative.

Fix a sequence of nonnegative, pairwise independent random variables $\{Y_n\}$ with, $\mathbb{E}(Y_n) \leq \mu \neq 0$, $\text{Var}(Y_n) \leq \sigma_Y^2$ for all $n \in \mathbb{N}$ and some $\mu, \sigma_Y > 0$. Set $S_n^Y := \sum_{i=1}^n Y_i$ and $z_n^Y := \sum_{i=1}^n \mathbb{E}(Y_n)$.

Lemma 3.1. For all $\varepsilon, \delta > 0$, $\alpha > 1$ and $0 \leq s \leq L_\delta$,

$$R_{\varepsilon, \alpha}(\lambda) = \log_\alpha \left(\frac{2\sigma_Y^2}{\lambda\varepsilon^2(\alpha - 1)} \right) - 1$$

is a rate of convergence for the partial sums of

$$\sum_{n=1}^{\infty} \mathbb{P} \left(\left| \frac{S_{k_s^\pm}^Y(n)}{k_s^\pm(n)} - \frac{z_{k_s^\pm}^Y(n)}{k_s^\pm(n)} \right| > \varepsilon \right) \tag{3.1}$$

to their respective limits.

Proof. Fix $\lambda, \varepsilon, \delta > 0$ and $\alpha > 1$ as well as $0 \leq s \leq L_\delta$, $Q \geq R_{\varepsilon, \alpha}(\lambda)$. Then:

$$\begin{aligned} \sum_{n=Q+1}^{\infty} \mathbb{P} \left(\left| \frac{1}{k_s^\pm(n)} S_{k_s^\pm}^Y(n) - \frac{1}{k_s^\pm(n)} z_{k_s^\pm}^Y(n) \right| > \varepsilon \right) &\leq \frac{1}{\varepsilon^2} \sum_{n=Q+1}^{\infty} \frac{\text{Var}(S_{k_s^\pm}^Y(n))}{k_s^\pm(n)^2} \\ &\leq \frac{\sigma_Y^2}{\varepsilon^2} \sum_{n=Q+1}^{\infty} \frac{1}{k_s^\pm(n)} \\ &\leq \frac{2\sigma_Y^2}{\varepsilon^2} \sum_{n=Q+1}^{\infty} \alpha^{-n} \\ &\leq \frac{2\sigma_Y^2 \alpha^{-(Q+1)}}{\varepsilon^2(\alpha - 1)} \leq \lambda. \end{aligned}$$

We get the first inequality from Chebyshev’s inequality, the second inequality by pairwise independence, the third inequality by using $k_s^\pm(n) \geq \lfloor \alpha^n \rfloor > \alpha^n/2$, the fourth inequality by using the sum of an infinite geometric sequence and the last inequality from the assumption that $Q \geq R_{\varepsilon, \alpha}(\lambda)$. \square

We can now apply Theorem 2.4 with the rate we obtained above, observing that R is independent of s (and δ), to easily obtain the following:

Lemma 3.2. For all $\varepsilon, \lambda > 0$ and all

$$n \geq \Delta_{\varepsilon, \mu, \sigma_Y}(\lambda) := \Phi_{\varepsilon, R}(\lambda) := \frac{36\alpha^2 \sigma_Y^2}{\lambda \varepsilon^2 (\alpha - 1)},$$

it holds that

$$\mathbb{P} \left(\sup_{m \geq n} \left| \frac{S_m^Y}{m} - \frac{z_{k_s^\pm(n)}}{k_s^\pm(n)} \right| > \varepsilon \right) \leq \lambda.$$

Here, $\alpha := 1 + \frac{\varepsilon}{3\mu}$, R is defined as in the previous lemma and Φ is defined as in Theorem 2.4.

Proof. This follows immediately from Theorem 2.4. Note we may take W to be μ . □

We can now obtain a rate where the random variables are not assumed to be nonnegative.

Proposition 3.3. Let $\{X_n\}$ be a sequence of pairwise independent random variables with $\mathbb{E}(X_n) = 0$, $\text{Var}(X_n) \leq \sigma^2$ and $\mathbb{E}(|X_n|) \leq \tau$ for all $n \in \mathbb{N}$ and some $\tau, \sigma > 0$. Furthermore, let $S_n := \sum_{i=1}^n X_i$. Then for all $\varepsilon, \lambda > 0$ and all $n \geq \Delta_{\frac{\varepsilon}{2}, \frac{\tau}{2}, \sigma}(\frac{\lambda}{2})$:

$$P_{n, \varepsilon}^* = \mathbb{P} \left(\sup_{m \geq n} \left| \frac{S_m}{m} \right| > \varepsilon \right) \leq \lambda.$$

Here,

$$\Delta_{\frac{\varepsilon}{2}, \frac{\tau}{2}, \sigma} \left(\frac{\lambda}{2} \right) := \frac{288\alpha^2 \sigma^2}{\lambda \varepsilon^2 (\alpha - 1)}$$

as before with $\alpha := 1 + \frac{\varepsilon}{3\tau}$. Thus, $P_{n, \varepsilon}^*$ converges to 0 with a rate of convergence given above.

Proof. We have, for all $n \in \mathbb{N}$, $\sigma^2 \geq \text{Var}(X_n) = \mathbb{E}(X_n^2) \geq \text{Var}(X_n^+) + \text{Var}(X_n^-)$ which implies $\sigma^2 \geq \text{Var}(X_n^\pm)$. Furthermore, we have $\mathbb{E}(X_n^\pm) \leq \frac{\tau}{2}$ since $\mathbb{E}(X_n) = 0 = \mathbb{E}(X_n^+) - \mathbb{E}(X_n^-)$. Thus, if we take $\{Y_n\} = \{X_n^\pm\}$, we can set $\sigma_Y := \sigma$ and $\mu := \tau/2$. Furthermore we can set

$$z_n := \sum_{i=1}^n \mathbb{E}(X_i^+) = \sum_{i=1}^n \mathbb{E}(X_i^-).$$

Thus, from the previous lemma:

$$\mathbb{P} \left(\sup_{m \geq n} \left| \frac{1}{m} S_m^\pm - \frac{1}{m} z_m \right| > \frac{\varepsilon}{2} \right) \leq \frac{\lambda}{2}$$

for all $\varepsilon, \lambda > 0$ and $n \geq \Delta_{\frac{\varepsilon}{2}, \mu, \sigma_Y}(\frac{\lambda}{2})$. Thus, if $n \geq \Delta_{\frac{\varepsilon}{2}, \frac{\tau}{2}, \sigma}(\frac{\lambda}{2}) = \Delta_{\frac{\varepsilon}{2}, \mu, \sigma_Y}(\frac{\lambda}{2})$, then

$$\begin{aligned} \mathbb{P} \left(\sup_{m \geq n} \left| \frac{1}{m} S_m \right| > \varepsilon \right) &= \mathbb{P} \left(\sup_{m \geq n} \left| \left(\frac{1}{m} S_m^+ - \frac{1}{m} z_m \right) - \left(\frac{1}{m} S_m^- - \frac{1}{m} z_m \right) \right| > \varepsilon \right) \\ &\leq \mathbb{P} \left(\sup_{m \geq n} \left| \frac{1}{m} S_m^+ - \frac{1}{m} z_m \right| > \frac{\varepsilon}{2} \right) + \mathbb{P} \left(\sup_{m \geq n} \left| \frac{1}{m} S_m^- - \frac{1}{m} z_m \right| > \frac{\varepsilon}{2} \right) \\ &\leq \lambda. \end{aligned}$$

Here $S_n^\pm := \sum_{i=1}^n X_n^\pm$. □

This, in particular, allows us rather immediately to deduce Theorem 1.4:

Proof of Theorem 1.4. Using the above proposition, we have

$$P_{n,\varepsilon}^* \leq \frac{288\alpha^2\sigma^2}{n\varepsilon^2(\alpha-1)}$$

for all $n \in \mathbb{N}$. So Theorem 1.4 follows by noting that if $\varepsilon \leq \tau$, we will have $\alpha \leq 4/3$. \square

Remark 3.4. In the case $\varepsilon > \tau$, observe that $\alpha < 4\varepsilon/3\tau$ and so we can deduce

$$P_{n,\varepsilon}^* \leq \frac{1536\sigma^2}{n\varepsilon\tau}.$$

We shall now discuss the optimality of the bound we obtained.

Example 3.5. For $\delta > 0$, let $\{X_n\}$ be a sequence of integer-valued iid random variables such that

$$\mathbb{P}(X_1 = n) = \frac{c}{n^{3+\delta}} \text{ for } c = \left(\sum_{n=1}^{\infty} \frac{1}{n^{3+\delta}} \right)^{-1}$$

for all $n \in \mathbb{Z}^+$ (and probability 0 for all other integers). Then $\text{Var}(X_1) < \infty$ and for all $1 \geq \varepsilon > 0$ and any $n \in \mathbb{N}$:

$$\mathbb{P} \left(\sup_{m \geq n} \left| \frac{S_m}{m} - \mu \right| > \varepsilon \right) \geq \frac{\omega}{n^{1+\delta}}$$

where μ is the mean of X_1 , given by

$$\mu = \sum_{n=1}^{\infty} \frac{c}{n^{2+\delta}},$$

and

$$\omega = \frac{c}{2 \times 3^{2+\delta}(2+\delta)}.$$

Proof. These random variables clearly have finite variance. For any $1 \geq \varepsilon > 0$, we have

$$\mathbb{P} \left(\sup_{m \geq n} \left| \frac{S_m}{m} - \mu \right| > \varepsilon \right) \geq \mathbb{P} \left(\sup_{m \geq n} \left| \frac{S_m}{m} - \mu \right| \geq 1 \right).$$

Observe that

$$\mu = \frac{\sum_{n=1}^{\infty} \frac{1}{n^{2+\delta}}}{\sum_{n=1}^{\infty} \frac{1}{n^{3+\delta}}} < \frac{\zeta(2)}{\zeta(4)} = \frac{15}{\pi^2} < 2.$$

Thus, we get

$$\begin{aligned} \mathbb{P} \left(\sup_{m \geq n} \left| \frac{S_m}{m} - \mu \right| \geq 1 \right) &\geq \mathbb{P} \left(\sup_{m \geq n} \frac{S_m}{m} \geq 3 \right) \\ &\geq \mathbb{P} \left(\frac{S_n}{n} \geq 3 \right) \\ &\geq \mathbb{P}(X_1 \geq 3n \cup \dots \cup X_n \geq 3n) \\ &= 1 - \mathbb{P}(X_1 < 3n \cap \dots \cap X_n < 3n) \\ &= 1 - (\mathbb{P}(X_1 < 3n))^n = 1 - (1 - \mathbb{P}(X_1 \geq 3n))^n. \end{aligned}$$

We now have

$$\mathbb{P}(X_1 \geq 3n) = c \sum_{k=3n}^{\infty} \frac{1}{k^{3+\delta}} \geq c \int_{3n}^{\infty} \frac{1}{x^{3+\delta}} dx = \frac{c}{(3n)^{2+\delta}(2+\delta)} = \frac{\omega}{n^{2+\delta}},$$

where $w = \frac{c}{3^{2+\delta}(2+\delta)}$. This implies,

$$\mathbb{P} \left(\sup_{m \geq n} \left| \frac{S_m}{m} - \mu \right| > \varepsilon \right) \geq 1 - \left(1 - \frac{w}{n^{2+\delta}} \right)^n \geq 1 - \frac{1}{1 + \frac{w}{n^{1+\delta}}} \geq \frac{w}{2n^{1+\delta}},$$

where we used the inequality $(1+x)^n \leq \frac{1}{1-nx}$ and $w < 1$. This yields the result. \square

Therefore, for every $\delta > 0$, by translation, we can obtain a sequence of iid random variables with expected values equal to 0 and finite variance, such that

$$\mathbb{P} \left(\sup_{m \geq n} \left| \frac{1}{m} S_m \right| > \varepsilon \right) \geq \frac{\omega}{n^{1+\delta}}.$$

This example demonstrates that $O\left(\frac{1}{n}\right)$ is an optimal general power of n bound for $P_{n,\varepsilon}^*$ in the case of finite variance. It however does not rule out the possibility that $P_{n,\varepsilon}^* = O\left(\frac{1}{n \log(n)}\right)$, for example.

4 Application II: the Chen-Sung Strong Law of Large Numbers

Throughout this section, let $\{X_n\}$ be a sequence of random variables with finite p -th moment (for some fixed $p \geq 1$) and respective means $\{\mu_n\}$. Let $S_n := \sum_{k=1}^n X_k$ and $z_n := \sum_{i=1}^n \mu_i$.

To use Theorem 2.4 to obtain a quantitative version of Theorem 1.8, we must find a rate of convergence for (2.1). To do this, we need some lemmas. The first is a technical lemma that resembles the Hájek and Rényi inequality.

Lemma 4.1. *Suppose $\{\gamma_n\}$ is a sequence of nonnegative real numbers satisfying*

$$\mathbb{E}(|S_n - z_n|^p) \leq \sum_{k=1}^n \gamma_k.$$

For all $\varepsilon, \delta > 0$, $\alpha > 1$ and $0 \leq s \leq L_\delta$:

$$\begin{aligned} \sum_{n=Q+1}^{\infty} \mathbb{P} \left(\left| \frac{S_{k_s^\pm(n)}}{k_s^\pm(n)} - \frac{z_{k_s^\pm(n)}}{k_s^\pm(n)} \right| > \varepsilon \right) \\ \leq \frac{2^p \alpha^{2p}}{\varepsilon^p [\alpha^{Q+2}]^p (\alpha^p - 1)} \sum_{m=1}^{\lfloor \alpha^{Q+2} \rfloor} \gamma_m + \frac{2^p \alpha^{2p}}{\varepsilon^p (\alpha^p - 1)} \sum_{m=\lfloor \alpha^{Q+1} \rfloor + 1}^{\infty} \frac{\gamma_m}{m^p}. \end{aligned} \tag{4.1}$$

Proof. Fix $M \in \mathbb{N}$. By the generalised Chebyshev's inequality, we have

$$\begin{aligned} \sum_{n=Q+1}^M \mathbb{P} \left(\left| \frac{S_{k_s^\pm(n)}}{k_s^\pm(n)} - \frac{z_{k_s^\pm(n)}}{k_s^\pm(n)} \right| > \varepsilon \right) &\leq \frac{1}{\varepsilon^p} \sum_{n=Q+1}^M \frac{\mathbb{E} \left(\left| S_{k_s^\pm(n)} - z_{k_s^\pm(n)} \right|^p \right)}{k_s^\pm(n)^p} \\ &\leq \frac{1}{\varepsilon^p} \sum_{n=Q+1}^M \frac{1}{k_s^\pm(n)^p} \sum_{m=1}^{k_s^\pm(n)} \gamma_m. \end{aligned}$$

Now, splitting the inner sum into two parts, observing that if $n > Q + 1$ then $k_s^\pm(n) > k_s^\pm(Q + 1)$, we have that the above sum is equal to

$$\frac{1}{\varepsilon^p} \sum_{n=Q+1}^M \frac{1}{k_s^\pm(n)^p} \sum_{m=1}^{k_s^\pm(Q+1)} \gamma_m + \frac{1}{\varepsilon^p} \sum_{n=Q+1}^M \frac{1}{k_s^\pm(n)^p} \sum_{m=k_s^\pm(Q+1)+1}^{k_s^\pm(n)} \gamma_m. \tag{4.2}$$

Now, interchanging summations in the first term and using $k_s^\pm(n) \geq \lfloor \alpha^n \rfloor > \alpha^n/2$, we can bound the first term from above by

$$\begin{aligned} \frac{2^p}{\varepsilon^p} \sum_{m=1}^{k_s^\pm(Q+1)} \gamma_m \sum_{n=Q+1}^M \alpha^{-pn} &\leq \frac{2^p \alpha^p}{\varepsilon^p \alpha^{p(Q+1)} (\alpha^p - 1)} \sum_{m=1}^{k_s^\pm(Q+1)} \gamma_m \\ &\leq \frac{2^p \alpha^{2p}}{\varepsilon^p \alpha^{p(Q+2)} (\alpha^p - 1)} \sum_{m=1}^{k_s^\pm(Q+1)} \gamma_m \\ &\leq \frac{2^p \alpha^{2p}}{\varepsilon^p \lfloor \alpha^{Q+2} \rfloor^p (\alpha^p - 1)} \sum_{m=1}^{\lfloor \alpha^{Q+2} \rfloor} \gamma_m. \end{aligned}$$

We bound the first line by an infinite geometric series to get the second line, and we use $\alpha^{Q+2} > k_s^\pm(Q+1)$ to get from the penultimate line to the last line.

We now bound the second term in (4.2) from above. Again, interchanging summations and using similar manipulations as used to obtain the bound for the first term, we get that the second is bounded above by

$$\begin{aligned} \frac{1}{\varepsilon^p} \sum_{m=k_s^\pm(Q+1)+1}^{k_s^\pm(M)} \gamma_m \sum_{\{M \geq n \geq Q+1: k_s^\pm(n) \geq m\}} \frac{1}{k_s^\pm(n)^p} \\ \leq \frac{2^p}{\varepsilon^p} \sum_{m=k_s^\pm(Q+1)+1}^{k_s^\pm(M)} \gamma_m \sum_{\{M \geq n \geq Q: \alpha^{n+1} \geq m\}} \alpha^{-pn}. \end{aligned}$$

The inner sum is bounded by an infinite geometric series with the first term $\leq m^{-p} \alpha^p$. Thus, the above is again bounded above by

$$\frac{2^p \alpha^{2p}}{\varepsilon^p (\alpha^p - 1)} \sum_{m=k_s^\pm(Q+1)+1}^{k_s^\pm(M)} \frac{\gamma_m}{m^p} \leq \frac{2^p \alpha^{2p}}{\varepsilon^p (\alpha^p - 1)} \sum_{m=\lfloor \alpha^{Q+1} \rfloor + 1}^{\lfloor \alpha^{M+1} \rfloor} \frac{\gamma_m}{m^p}.$$

Taking $M \rightarrow \infty$ gives the required result. □

To find a rate of convergence for (2.1), we must find one for the two terms on the right-hand side of (4.1). A rate for the second term can easily be calculated given one for $\sum_{m=1}^\infty \frac{\gamma_m}{m^p}$. To obtain a rate for the second term, we need a quantitative version of what is known as Kronecker’s Lemma. One can turn to [20, Theorem A.6.2] for a proof of the non-quantitative result.

Lemma 4.2 (Quantitative Kronecker’s lemma). *Let x_1, x_2, \dots be a sequence of nonnegative real numbers such that $\sum_{i=1}^\infty x_i < \infty$ and let $0 < a_1 \leq a_2 \leq \dots$ be such that $a_n \rightarrow \infty$. Quantitatively, suppose $\sum_{i=1}^\infty x_i < S$ for some $S > 0$ and that $s_n := \sum_{i=1}^n x_i$ converges to $\sum_{i=1}^\infty x_i$ with rate of convergence ϕ . Further, suppose that there is a function $f : \mathbb{R} \rightarrow \mathbb{N}$ such that $a_{f(\omega)} \geq \omega$ for all $\omega > 0$. Then*

$$\frac{1}{a_n} \sum_{i=1}^n a_i x_i \rightarrow 0$$

as $n \rightarrow \infty$ with rate of convergence

$$K_{\phi, f, \{a_n\}, S}(\varepsilon) = \max \left\{ \phi \left(\frac{\varepsilon}{4} \right), f \left(\frac{4a_{\phi(\frac{\varepsilon}{4})} S}{\varepsilon} \right) \right\}.$$

Proof. Take $\varepsilon > 0$ and $n \geq K_{\phi, f, \{a_n\}, S}(\varepsilon)$ and let $M = \phi(\frac{\varepsilon}{4})$. From the definition of ϕ , we have for all $i \geq M$ that $|s_M - s_i| \leq \frac{\varepsilon}{4}$. Summation by parts gives

$$\left| \frac{1}{a_n} \sum_{i=1}^n a_i x_i \right| = \left| s_n - \frac{1}{a_n} \sum_{i=1}^{n-1} (a_{i+1} - a_i) s_i \right|$$

which in turn is equal to

$$\begin{aligned} & \left| s_n - \frac{1}{a_n} \sum_{i=1}^{M-1} (a_{i+1} - a_i) s_i - \frac{1}{a_n} \sum_{i=M}^{n-1} (a_{i+1} - a_i) s_M - \frac{1}{a_n} \sum_{i=M}^{n-1} (a_{i+1} - a_i) (s_i - s_M) \right| \\ & \leq \left| s_n - (1 - \frac{a_M}{a_n}) s_M \right| + \left| \frac{1}{a_n} \sum_{i=1}^{M-1} (a_{i+1} - a_i) s_i \right| + \left| \frac{1}{a_n} \sum_{i=M}^{n-1} (a_{i+1} - a_i) (s_i - s_M) \right| \\ & \leq |s_n - s_M| + \left| \frac{a_M s_M}{a_n} \right| + \left| \frac{1}{a_n} \sum_{i=1}^{M-1} (a_{i+1} - a_i) s_i \right| + \left| \frac{1}{a_n} \sum_{i=M}^{n-1} (a_{i+1} - a_i) (s_i - s_M) \right| \\ & \leq \frac{\varepsilon}{4} + \left| \frac{a_M S}{a_n} \right| + \left| \frac{S}{a_n} \sum_{i=1}^{M-1} (a_{i+1} - a_i) \right| + \left| \frac{1}{a_n} \frac{\varepsilon}{4} \sum_{i=M}^{n-1} (a_{i+1} - a_i) \right| \\ & \leq \frac{\varepsilon}{4} + \left| \frac{a_M S}{a_n} \right| + \left| \frac{S a_M}{a_n} \right| + \frac{\varepsilon}{4} \leq \varepsilon. \quad \square \end{aligned}$$

We can now calculate a rate of convergence for (2.1)

Lemma 4.3. *Suppose $\{X_n\}$ and $\{\gamma_n\}$ are as in Theorem 1.8. Suppose $\sum_{m=1}^{\infty} \frac{\gamma_n}{m^p} \leq \Gamma$ for some $\Gamma > 0$ and that the partial sums converge to their limit with a strictly decreasing rate of convergence Ψ . For all $\varepsilon, \delta > 0, \alpha > 1$ and $0 \leq s \leq L_\delta$, the function $\chi_{\varepsilon, \alpha, \Psi}$ is a rate of convergence for the partial sums of*

$$\sum_{m=1}^{\infty} \mathbb{P} \left(\left| \frac{S_{k_s^\pm(m)}}{k_s^\pm(m)} - \frac{z_{k_s^\pm(m)}}{k_s^\pm(m)} \right| > \varepsilon \right)$$

to their respective limits, where

$$\chi_{\varepsilon, \alpha, \Psi}(\lambda) = \max \left\{ \log_\alpha \left(2\Psi \left(\frac{\lambda \varepsilon^p (\alpha^p - 1)}{2^{p+1} \alpha^{2p}} \right) \right), \log_\alpha \left(2K_{\psi, f_p, \{n^p\}, R} \left(\frac{\lambda}{2} \right) \right) \right\}$$

with

$$\psi(\lambda) = \Psi \left(\frac{\lambda \varepsilon^p (\alpha^p - 1)}{2^p \alpha^{2p}} \right), \quad f_p(\omega) = \left\lceil \omega^{\frac{1}{p}} \right\rceil, \quad R = \frac{2^p \Gamma \alpha^{2p}}{\varepsilon^p (\alpha^p - 1)}.$$

Proof. Let $\lambda, \varepsilon, \delta > 0, \alpha > 1$ and $0 \leq s \leq L_\delta$ as well as $n \geq \chi_{\varepsilon, \alpha, \Psi}(\lambda)$ be given. We have, by Lemma 4.1, that

$$\begin{aligned} & \sum_{m=n+1}^{\infty} \mathbb{P} \left(\left| \frac{S_{k_s^\pm(m)}}{k_s^\pm(m)} - \frac{z_{k_s^\pm(m)}}{k_s^\pm(m)} \right| > \varepsilon \right) \\ & \leq \frac{2^p \alpha^{2p}}{\varepsilon^p [\alpha^{n+2}]^p (\alpha^p - 1)} \sum_{m=1}^{\lfloor \alpha^{n+2} \rfloor} \gamma_m + \frac{2^p \alpha^{2p}}{\varepsilon^p (\alpha^p - 1)} \sum_{m=\lfloor \alpha^{n+1} \rfloor + 1}^{\infty} \frac{\gamma_m}{m^p}. \end{aligned}$$

Now, $n \geq \chi_{\varepsilon, \alpha, \Psi}(\lambda)$ implies

$$n \geq \log_\alpha \left(2\Psi \left(\frac{\lambda \varepsilon^p (\alpha^p - 1)}{2^{p+1} \alpha^{2p}} \right) \right)$$

and from this, we deduce

$$\lfloor \alpha^{n+1} \rfloor \geq \alpha^{n+1}/2 \geq \Psi \left(\frac{\lambda \varepsilon^p (\alpha^p - 1)}{2^{p+1} \alpha^{2p}} \right),$$

which in turn implies

$$\frac{2^p \alpha^{2p}}{\varepsilon^p (\alpha^p - 1)} \sum_{m=\lfloor \alpha^{n+1} \rfloor + 1}^{\infty} \frac{\gamma_m}{m^p} \leq \frac{\lambda}{2}.$$

Now, observe that the partial sums of

$$\frac{2^p \alpha^{2p}}{\varepsilon^p (\alpha^p - 1)} \sum_{m=1}^{\infty} \frac{\gamma_m}{m^p}$$

converge, to their limit, with rate

$$\psi(\lambda) = \Psi \left(\frac{\lambda \varepsilon^p (\alpha^p - 1)}{2^p \alpha^{2p}} \right).$$

Therefore, by Lemma 4.2, we have that

$$\frac{2^p \alpha^{2p}}{\varepsilon^p n^p (\alpha^p - 1)} \sum_{m=1}^n \gamma_m$$

converges to 0 with rate $K_{\psi, f_p, \{n^p\}, R}$. Now, $n \geq \chi_{\varepsilon, \alpha, \Psi}(\lambda)$ further implies

$$n \geq \log_{\alpha} \left(2K_{\psi, f_p, \{n^p\}, R} \left(\frac{\lambda}{2} \right) \right)$$

and, arguing as above, we get

$$\lfloor \alpha^{n+2} \rfloor \geq K_{\psi, f_p, \{n^p\}, R} \left(\frac{\lambda}{2} \right).$$

This allows us to conclude

$$\frac{2^p \alpha^{2p}}{\varepsilon^p \lfloor \alpha^{n+2} \rfloor^p (\alpha^p - 1)} \sum_{m=1}^{\lfloor \alpha^{n+2} \rfloor} \gamma_m \leq \frac{\lambda}{2}$$

and we are done. □

This, in particular, allows us rather immediately to deduce Theorem 1.9 and Theorem 1.10.

Proof of Theorem 1.9. In the context of the assumptions of Theorem 1.9, using the assumption that Ψ is strictly decreasing, the rate χ in the previous Lemma 4.3 simplifies to

$$\log_{\alpha} \left(\max \left\{ 2\Psi \left(\frac{\lambda \varepsilon^p (\alpha^p - 1)}{2^{p+3} \alpha^{2p}} \right), 2 \left[\frac{2\alpha^2}{\varepsilon} \Psi \left(\frac{\lambda \varepsilon^p (\alpha^p - 1)}{2^{p+3} \alpha^{2p}} \right) \left(\frac{8\Gamma}{\lambda(\alpha^p - 1)} \right)^{\frac{1}{p}} \right] \right\} \right).$$

We can now apply Theorem 2.4 with the rate we obtained above, observing that χ is independent of s (and δ) to deduce Theorem 1.9, noting that we can take $\alpha = 1 + \frac{\varepsilon}{3W}$ and so the assumption that $\varepsilon \leq 1 \leq W$ implies $\alpha < 4/3$ and $\alpha^p - 1 \geq \alpha - 1$. Furthermore, the assumptions that $\Gamma, W \geq 1$ and $0 < \varepsilon \leq 1$, as well as the strictly decreasing condition on Ψ , allows us to conclude that the second argument in the max function above is bigger than the first argument. □

Proof of Theorem 1.10. We write $X_n = X_n^+ - X_n^-$ and apply Theorem 1.9 to each sequence $\{X_n^{\pm}\}$, with $p = 2$ and $\gamma_n = \text{Var}(X_n^{\pm}) \leq \text{Var}(X_n)$. We then obtain the result for $\{X_n\}$ by arguing as in Proposition 3.3. □

5 Comments and questions

There are still a lot of questions that can be asked in the area of large deviations in the Strong Law of Large Numbers. We conclude with some potential avenues of further study:

Firstly, as mentioned in Section 1.1, Cramér's original 1938 results have been generalised in [42] to sequences of random variables not assumed to be identically distributed. However, the non-identically distributed generalisation of large deviation probabilities related to the weak law of large numbers [4] has not been studied. So, we can ask if one can obtain such a generalisation. Furthermore, one can ask if such a generalisation exists for the Strong Law of Large Numbers.

Baum-Katz type rates are typically given by showing some series converges through the bounding of the series. As already mentioned in this article, it is a well known result in computability theory that given a bound on a sum, there is no general computable process to extract a computable rate of convergence of that sum to its limit. Thus, one avenue of study is to try to extract rates of convergence for these Baum-Katz type results. Doing so will give us a more descriptive picture of how these large deviation probabilities behave. This problem appears to have already been considered in passing by Erdős in [14]. In the case $r = 0$ of Theorem 1.5, Erdős provides an elementary proof that condition (i) implies condition (ii) (this was first demonstrated by Hsu and Robbins [22] by techniques involving Fourier analysis) as well as the converse implication. Erdős's approach to demonstrating that (i) implies (ii) was to split the sum (ii) into three parts. For two parts, Erdős calculates explicit rates of convergence that are independent of the distribution of the random variables; however, for the last part, Erdős bounds the sum, and it is unclear how one obtains a rate from this bound (that is independent of the distribution of the random variables). Thus, getting a rate from Erdős's proof is not straightforward. One could try to obtain a rate by analysing Hsu and Robbins's proof.

Uniform rates of convergence (uniform in that they do not depend on the distribution of the random variables) have been found for the Central Limit Theorem:

Theorem 5.1 (Berry [7] and Esseen [16]). *Let $\{X_n\}$ be iid random variables satisfying $\mathbb{E}(X_1) = 0$, $\text{Var}(X_1) = \sigma^2 > 0$, $\mathbb{E}(|X_1|^3) = \rho < \infty$. Let*

$$S_n = \frac{\sum_{i=1}^n X_i}{\sqrt{\sum_{i=1}^n \sigma_i^2}}$$

F_n be the cumulative distribution function of S_n and Φ the cumulative distribution function of the standard normal distribution. Then for all $n \in \mathbb{N}$ and $x \in \mathbb{R}$

$$|F_n(x) - \Phi(x)| \leq \frac{C\rho}{\sqrt{n}\sigma^3}$$

For some $C > 0$.

Suppose we do not assume the random variables have a finite third moment but do have finite variance. In that case, we can still deduce that $F_n(x) \rightarrow \Phi(x)$ by the Central Limit Theorem, but we do not get such uniform rates of convergence (Berry initially attempted to do this and could not [7]). A similar phenomenon appears to occur in the Strong Law of Large Numbers, as the result holds if we assume the random variables have a finite first moment. It is also not clear if one can obtain rates that are independent of the distribution by only assuming a finite first moment. However, taking one moment higher (so finite variance) allows one to obtain rates independent of the distribution of the random variables. One can investigate whether this occurs in other limit theorems. A potential case study could investigate whether a rate of convergence

for the Hsu-Robbins-Erdős sum [14, 22] (condition (ii) in Theorem 1.5 in the case $r = 0$) can be obtained if we assume finite third moment instead of just finite variance.

This article can be seen as a contribution to the proof mining program, which is a research area that aims to use tools and ideas from logic to extract quantitative data from proofs that appear nonconstructive. Although the program has its origins in Kreisel’s “unwinding” program of the 1950s [33, 34], its emergence as a mature subfield of applied logic was due to the work of Kohlenbach and his collaborators (see [27] for a comprehensive overview of the program and the recent survey papers [29, 28] for applications). This program has enjoyed a lot of success in analysis but has also expanded into various other areas of mathematics, where we in particular mention Tauberian theory [43, 44], differential algebra [48] and probability theory [1, 2, 3], with the latter references being the only publications of applications of the techniques of proof mining to probability theory, up until recently. Progress has been made both in the development of logical tools to study nonconstructive proofs [37] and in the extraction of computational content from proofs in probability theory [38, 39], with this article adding to the growing body of work in the latter regard.

This logical background was crucial for obtaining the results of this paper⁴; however, we did not assume any familiarity with the concepts or techniques of proof mining, and, even further, none of the results or proofs presented here make any explicit use of such methods. Such a style of presentation is common practice in the proof mining literature.

References

- [1] R. Arthan and P. Oliva, *On the Borel-Cantelli lemmas, the Erdős-Rényi Theorem, and the Kochen-Stone Theorem*, Journal of Logic and Analysis **13** (2021), no. 6, 23pp. MR4359804
- [2] J. Avigad, E. Dean, and J. Rute, *A metastable dominated convergence theorem*, Journal of Logic and Analysis **4** (2012), no. 3, 1–19. MR2889823
- [3] J. Avigad, P. Gerhardy, and H. Towsner, *Local stability of ergodic averages*, Transactions of the American Mathematical Society **362** (2010), no. 1, 261–288. MR2550151
- [4] R.R. Bahadur and R.R. Ranga, *On deviations of the sample mean*, The Annals of Mathematical Statistics **31** (1960), no. 4, 1015–1027. MR0117775
- [5] P. Bai, P. Chen, and S.H. Sung, *On complete convergence and the strong law of large numbers for pairwise independent random variables*, Acta Mathematica Hungarica **142** (2014), 502–518. MR3165498
- [6] L.E. Baum and M.L. Katz, *Convergence rates in the law of large numbers*, Transactions of the American Mathematical Society **120** (1965), no. 1, 108–123. MR0198524
- [7] A.C. Berry, *The accuracy of the Gaussian approximation to the sum of independent variates*, Transactions of the American Mathematical Society **49** (1941), no. 1, 122–136. MR0003498
- [8] T. Birkel, *A note on the strong law of large numbers for positively dependent random variables*, Statistics & Probability Letters **7** (1988), no. 1, 17–20. MR0996845
- [9] T. Chandra and A. Goswami, *Cesaro uniform integrability and the strong law of large numbers*, Sankhyā: The Indian Journal of Statistics, Series A (1992), 215–231. MR1192096
- [10] P. Chen and S.H. Sung, *A strong law of large numbers for nonnegative random variables and applications*, Statistics & Probability Letters **118** (2016), 80–86. MR3531485
- [11] Y.S. Chow, *Delayed sums and Borel summability of independent, identically distributed random variables*, Bulletin of the Institute of Mathematics, Academia Sinica **1** (1973), no. 2, 207–220. MR0343357
- [12] H. Cramér, *On a new limit theorem of probability theory*, Actualités scientifiques et industrielles **736** (1938), 5–23.

⁴The insights from proof mining allowed us to obtain quantitative forms of lemmas we relied on to prove our main quantitative results and provide methods of extracting the computational content of these lemmas.

- [13] S. Csörgő, K. Tandori, and V. Totik, *On the strong law of large numbers for pairwise independent random variables*, Acta Mathematica Hungarica **42** (1983), 319–330. MR0722846
- [14] P. Erdős, *On a Theorem of Hsu and Robbins*, The Annals of Mathematical Statistics **20** (1949), no. 2, 286 – 291. MR0030714
- [15] E. Specker, *Nicht konstruktiv beweisbare Sätze der Analysis*, Journal of Symbolic Logic **14** (1949), 145–158. MR0031447
- [16] C.-G. Esseen, *On the Liapunov limit error in the theory of probability*, Arkiv för matematik, astronomi och fysik **28** (1942), 1–19. MR0011909
- [17] N. Etemadi, *An elementary proof of the strong law of large numbers*, Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete **55** (1981), no. 1, 119–122. MR0606010
- [18] J.A. Fill, *Convergence Rates Related to the Strong Law of Large Numbers*, The Annals of Probability **11** (1983), no. 1, 123 – 142. MR0682804
- [19] V. Gaposhkin, *Some Examples of the Problem of ε -Deviations for Stationary Sequences*, Theory of Probability & Its Applications **46** (2002), no. 2, 341–346. MR1968692
- [20] A. Gut, *Probability: a graduate course*, Springer, 2013. MR2977961
- [21] J. Hájek and A. Rényi, *Generalization of an inequality of Kolmogorov*, Acta Mathematica Hungarica **6** (1955), no. 3-4, 281–283. MR0076207
- [22] P. Hsu and H. Robbins, *Complete convergence and the law of large numbers*, Proceedings of the National Academy of Sciences of the United States of America **33** (1947), no. 2, 25–31. MR0019852
- [23] H. Jabbari, *On almost sure convergence for weighted sums of pairwise negatively quadrant dependent random variables*, Statistical Papers **54** (2013), 765–772. MR3072899
- [24] A.G. Kachurovskii and I.V. Podvigin, *Measuring the rate of convergence in the Birkhoff ergodic theorem*, Mathematical Notes **106** (2019), 52–62. MR3981324
- [25] A.G. Kachurovskii, I.V. Podvigin, and A.A. Svishchev, *The maximum pointwise rate of convergence in Birkhoff's ergodic theorem*, Journal of Mathematical Sciences **255** (2021), no. 2, 119–123. MR4252946
- [26] T.-S. Kim and J.I. Baek, *The strong laws of large numbers for weighted sums of pairwise quadrant dependent random variables*, Journal of the Korean Mathematical Society **36** (1999), no. 1, 37–49. MR1669137
- [27] U. Kohlenbach, *Applied Proof Theory: Proof Interpretations and their Use in Mathematics*, Springer Monographs in Mathematics, Springer, 2008. MR2445721
- [28] U. Kohlenbach, *Recent progress in proof mining in nonlinear analysis*, Journal of Applied Logics - IfCoLog Journal **4** (2017), no. 10, 3361–3410. MR4542239
- [29] U. Kohlenbach, *Proof-theoretic methods in nonlinear analysis*, Proceedings of the International Congress of Mathematicians 2018, vol. 2, World Scientific, 2019, pp. 61–82. MR3966757
- [30] V.M. Korchevsky, *A generalization of the Petrov strong law of large numbers*, Statistics & Probability Letters **104** (2015), 102–108. MR3360711
- [31] V.M. Korchevsky, *On the Rate of Convergence in the Strong Law of Large Numbers for Nonnegative Random Variables*, Journal of Mathematical Sciences **229** (2018), no. 6, 719–727. MR3602409
- [32] V.M. Korchevsky and V.V. Petrov, *On the strong law of large numbers for sequences of dependent random variables*, Vestnik St. Petersburg University: Mathematics **43** (2010), 143–147. MR2741955
- [33] G. Kreisel, *On the Interpretation of Non-Finitist Proofs, Part I*, Journal of Symbolic Logic **16** (1951), 241–267. MR0049135
- [34] G. Kreisel, *On the interpretation of Non-Finitist Proofs, Part II: Interpretation of Number Theory*, Journal of Symbolic Logic **17** (1952), 43–58. MR0051193
- [35] A. Kuczmaszewska, *Convergence rate in the Petrov SLLN for dependent random variables*, Acta Mathematica Hungarica **148** (2016), 56–72. MR3439282

- [36] N. Luzia, *A simple proof of the strong law of large numbers with rates*, Bulletin of the Australian Mathematical Society **97** (2018), no. 3, 513–517. MR3802466
- [37] M. Neri and N. Pischke, *Proof mining and probability theory*, Preprint, available at <https://arxiv.org/abs/2403.00659>, 2024.
- [38] M. Neri and T. Powell, *On quantitative convergence for stochastic processes: Crossings, fluctuations and martingales*, Preprint, available at <https://arxiv.org/abs/2406.19979>, 2024.
- [39] M. Neri and T. Powell, *A quantitative Robbins-Siegmund theorem*, Preprint, available at <https://arxiv.org/abs/2410.15986>, 2024.
- [40] V.V. Petrov, *Generalization of Cramér’s limit theorem*, Uspekhi Matematicheskikh Nauk **9** (1954), no. 4, 195–202. MR0065058
- [41] V.V. Petrov, *On the Strong Law of Large Numbers for Nonnegative Random Variables*, Theory of Probability & Its Applications **53** (2009), no. 2, 346–349. MR3691828
- [42] V.V. Petrov and J. Robinson, *Large deviations for sums of independent non identically distributed random variables*, Communications in Statistics—Theory and Methods **37** (2008), no. 18, 2984–2990. MR2467746
- [43] T. Powell, *A note on the finitization of Abelian and Tauberian theorems*, Mathematical Logic Quarterly **66** (2020), no. 3, 300–310. MR4174107
- [44] T. Powell, *A finitization of Littlewood’s Tauberian theorem and an application in Tauberian remainder theory*, Annals of Pure and Applied Logic **174** (2023), no. 4, 103231. MR4535406
- [45] E. Rio, *Convergence speeds in the strong law for dependent sequences*, Proceedings of the Academy of Sciences. Series 1, Mathématique **320** (1995), no. 4, 469–474. MR1320123
- [46] H.R. Nili Sani, H.A. Azarnoosh, and A. Bozorgnia, *The strong law of large numbers for pairwise negatively dependent random variables*, Iranian Journal of Science **28** (2004), no. 2, 211–217. MR2136407
- [47] D. Siegmund, *Large deviation probabilities in the strong law of large numbers*, Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete **31** (1975), no. 2, 107–113. MR0365672
- [48] W. Simmons and H. Towsner, *Proof mining and effective bounds in differential polynomial rings*, Advances in Mathematics **343** (2019), 567–623. MR3883215

Acknowledgments. This article was written as part of the author’s PhD studies under the supervision of Thomas Powell, and I would like to thank him for his support and invaluable guidance. The author would also like to thank Nicholas Pischke for his advice on the presentation and formatting of the results in this article, as well as Cécile Mailler and Nathan Creighton for their helpful comments.