Necessary and Sufficient Conditions for the Lacunary/Hereditary Laws of Large Numbers *

István Berkes † Ioannis Karatzas ‡ Walter Schachermayer §

February 23, 2026

Abstract

The celebrated theorem of Komlós [31] asserts that $\mathbb{L}^1-$boundedness is sufficient for a given sequence of functions to contain a subsequence along which (in a "lacunary" manner), and along whose every further subsequence ("hereditarily"), a strong law of large numbers holds. We identify here slightly weaker, Egorov-type conditions, as not only sufficient in this context, but necessary as well. Necessary and sufficient conditions are developed also for the lacunary/hereditary version of the weak law of large numbers for general sequences, as well as for the weak law of large numbers in the context of exchangeable sequences, both long-open questions.

AMS 2020 Subject Classification: Primary 60A10, 60F15; Secondary 60G57, 60G09, 60G42.

Keywords: Subsequences, hereditary convergence, stable convergence, exchangeability, Komlós theorem, laws of large numbers

1 Introduction

The lacunary/hereditary strong law of large numbers established in the seminal paper of Komlós [31], singles out the boundedness-in-$\mathbb{L}^1$ condition

$$\sup_{n\in\mathbb{N}} \mathbb{E}^{\mathbb{P}}\big(|f_n|\big) < \infty\,, \tag{1.1}$$

pertaining to a given sequence of measurable functions $f_1, f_2, \cdots$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, as sufficient for the existence of a measurable $f_* : \Omega \to \mathbb{R}$ and of a subsequence $f_{k_1}, f_{k_2}, \cdots$ satisfying the convergence in Cesàro mean

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f_{k_n} = f_* \tag{1.2}$$

$\mathbb{P}-$a.e.; as well as hereditarily, that is, along all further subsequences. The limiting function f_* plays here the rôle of "randomized mean", whose $\mathbb{P}-$integrability the condition (1.1) also guarantees.

* The authors are grateful to Tomoyuki Ichiba and Nicola Doninelli for pointing out crucial results in this field.

† Institute of Statistics, Graz University of Technology, Kopernikusgasse 24, 8010 Graz, Austria (*berkes@tugraz.at*) and Alfred Rényi Institute of Mathematics, Reáltanoda utca 13-15, 1053 Budapest, Hungary (*berkes.istvan@renyi.hu*).

‡ Departments of Mathematics and Statistics, Columbia University, New York, NY 10027 (*ik1@columbia.edu*). Support from the NSF under Grant DMS-25-06199, and from a Lenfest Award at Columbia, is gratefully acknowledged.

§ Faculty of Mathematics, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria (*walter.schachermayer@univie.ac.at*). Support from the Austrian Science Fund (FWF) under grants P-35519 and P-35197 is gratefully acknowledged.

This is a striking, and very useful, result; its proof was a major early success of martingale methods in classical probability limit theory. It extends the sufficiency statements, not only of the KOLMOGOROV ([29], [30]) strong law of large numbers, but also of the DUNFORD-PETTIS ([20], [21]) theorem in functional analysis. It has found numerous applications in that field (e.g., [18], [26], [32], [38]), as well as in stochastic analysis (e.g., [19], [6], [7], [25]). It inspired the development of sufficient conditions for the lacunary/hereditary versions of the law of the iterated logarithm and of the central limit theorem, leading to a heuristic "principle of subsequences" ([13], [24], [14]).

Now it is well-known that, when the $\mathsf{f}_1, \mathsf{f}_2, \cdots$ are independent and identically distributed (I.I.D.), the integrability condition $\mathbb{E}\big(\big|\mathsf{f}_1\big|\big) < \infty$ is not only sufficient for the validity of the classical strong law of large numbers, but also necessary (e.g. [33], p. 251). Likewise, uniform integrability in the DUNFORD-PETTIS theorem is not just sufficient for relative compactness in the space $\mathbb{L}^1$ equipped with the $\sigma(\mathbb{L}^1, \mathbb{L}^\infty)$ topology, but necessary as well (Theorem T23, page 20 in [34]).

We identify here a condition, slightly weaker than (1.1), as not only sufficient for the validity of the lacunary/hereditary Strong Law of Large Numbers (1.2) with some measurable $\mathsf{f}_* : \Omega \to \mathbb{R}$, but also necessary. Developed in Theorem 2.1, this condition posits the existence of sets $\mathsf{A}_1 \subseteq \mathsf{A}_2 \subseteq \cdots \subseteq \mathsf{A}_\mathsf{j} \subseteq \cdots$ in $\mathcal{F}$ with $\lim_{\mathsf{j}\to\infty} \mathbb{P}(\mathsf{A}_\mathsf{j}) = 1$, and of a subsequence $\mathsf{f}_{\mathsf{k}_1}, \mathsf{f}_{\mathsf{k}_2}, \cdots$, such that

$$\Big(\mathsf{f}_{\mathsf{k}_\mathsf{n}}\, \mathbf{1}_{\mathsf{A}_\mathsf{j}}\Big)_{\mathsf{n}\in\mathbb{N}} \quad \text{is bounded in } \mathbb{L}^1(\mathbb{P}), \text{ for each } \mathsf{j} \in \mathbb{N}. \tag{1.3}$$

This is equivalent to the existence of a probability measure $\mathbb{Q} \sim \mathbb{P}$, under which a subsequence $\mathsf{f}_{\mathsf{k}_1}, \mathsf{f}_{\mathsf{k}_2}, \cdots$ is bounded in $\mathbb{L}^1(\mathbb{Q})$; whereas, the slightly stronger condition (2.4) of Proposition 2.3 is necessary and sufficient for the CESÀRO limit f_* to be $\mathbb{P}-$integrable.

Similarly, Theorem 3.1 develops necessary and sufficient conditions for the lacunary/hereditary version of the KOLMOGOROV-FELLER ([28], [23]) *Weak Law of Large Numbers* in the context of general sequences, a pivotal result of the paper. Its proof turns out to be considerably more involved than the one for the Strong Law of Large Numbers.

The arguments rest on three pillars: (i) the notion and properties of stable convergence (cf. [35]); (ii) ALDOUS's profound subsequence theory in [2], using exchangeability; and (iii) *conditions necessary and sufficient for the Weak Law of Large Numbers in the context of exchangeable sequences.*

These latter conditions settle another long-open question. It is quite striking that, the necessary and sufficient conditions for the exchangeable case, are *not* the corresponding classical conditions for the I.I.D. case conditioned on the tail σ-algebra of the sequence.

2 Sufficiency and Necessity in the Lacunary/Hereditary SLLN

Here is a reformulation of KOMLÓS's result [31] in terms of conditions both sufficient and necessary. It is proved in section 5; a result of independent interest, crucial for the proof, is in section 6.

Theorem 2.1. Sufficiency/Necessity in the Lacunary/Hereditary SLLN for General Sequences: *On a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$, *consider real-valued, measurable functions* $\mathsf{f}_1, \mathsf{f}_2, \cdots$.

(i) *Suppose that for some subsequence* $\mathsf{f}_{\mathsf{k}_1}, \mathsf{f}_{\mathsf{k}_2}, \cdots$ *and sets* $\mathsf{A}_1 \subseteq \mathsf{A}_2 \subseteq \cdots \subseteq \mathsf{A}_\mathsf{j} \subseteq \cdots$ *in* $\mathcal{F}$ *with* $\lim_{\mathsf{j}\to\infty} \mathbb{P}(\mathsf{A}_\mathsf{j}) = 1$, *the condition* (1.3) *holds.*

There exist then a measurable function $\mathsf{f}_* : \Omega \to \mathbb{R}$ *and a* (*further, relabelled*) *subsequence along which, and along whose every subsequence, the* CESÀRO*-mean convergence in* (1.2) *holds* $\mathbb{P}-$*a.e.*

(ii) *Conversely, suppose that along some subsequence* $f_{k_1}, f_{k_2}, \cdots$ *and each of its subsequences, the* Cesàro*-mean convergence* (1.2) *holds* $\mathbb{P}-$*a.e. for some measurable* $f_* : \Omega \to \mathbb{R}$ *. There exist then sets* $A_1 \subseteq A_2 \subseteq \cdots \subseteq A_j \subseteq \cdots$ *in* $\mathcal{F}$ *with* $\lim_{j\to\infty} \mathbb{P}(A_j) = 1$ *, such that* (1.3) *is satisfied.*

(iii) *The validity of* (1.3) *for some subsequence* $f_{k_1}, f_{k_2}, \cdots$ *and for sets* $A_1 \subseteq A_2 \subseteq \cdots \subseteq A_j \subseteq \cdots$ *in* $\mathcal{F}$ *with* $\lim_{j\to\infty} \mathbb{P}(A_j) = 1$ *, is equivalent to the existence of a probability measure* $\mathbb{Q} \sim \mathbb{P}$ *on* $(\Omega, \mathcal{F})$ *with the property that some subsequence* $f_{k_1}, f_{k_2}, \cdots$ *is bounded in* $\mathbb{L}^1(\mathbb{Q})$ *.*

Remark 2.2. It makes good sense, that probability measures equivalent to $\mathbb{P}$ should enter the picture here, as they do in part (iii) above: the property in question, almost-everywhere convergence to a real-valued limit, is invariant under equivalent changes of probability measure.

In the special context of a necessary condition for the Komlós theorem with nonnegative functions, this feature was pointed out and established by von Weizsäcker [37].

2.1 Stable Convergence

We shall endow the space $\mathbb{L}^0 \equiv \mathbb{L}^0(\mathbb{P})$ of real-valued, measurable functions on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with the topology induced by the metric $\mathbb{L}^0 \times \mathbb{L}^0 \ni \big(f, g\big) \longmapsto \mathbb{E}^{\mathbb{P}}\big(1 \wedge |f - g|\big) \in [0, 1]$. Convergence under this metric is equivalent to convergence in $\mathbb{P}-$measure ("in probability"). The space $\mathbb{L}^0$ and its topology depend only on the equivalence class, modulo null sets, to which $\mathbb{P}$ belongs: they are the same for any probability measure $\mathbb{Q} \sim \mathbb{P}$ on $(\Omega, \mathcal{F})$.

We call a set $\mathbb{G} \subset \mathbb{L}^0$ *bounded in probability*, if $\lim_{M\to\infty} \sup_{g\in\mathbb{G}} \mathbb{P}\big(|g| > M\big) = 0$ holds.

Let us consider now real-valued, measurable $f_1, f_2, \cdots$ on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, along with their tail $\sigma-$algebra

$$\mathcal{T} := \bigcap_{n\in\mathbb{N}} \mathcal{T}_n \,, \qquad \mathcal{T}_n := \boldsymbol{\sigma}\big(f_n, f_{n+1}, \cdots\big) \,. \tag{2.1}$$

Here is a basic result, which goes back to Rényi [35]: *Every bounded in probability sequence* $f_1, f_2, \cdots$ *contains a subsequence* $f_{k_1}, f_{k_2}, \cdots$ *which is "determining"*, i.e., satisfies the *stable convergence* (or extended Helly-Bray property, proved in [5]; see also [4], and Theorem 2.2 in [10])

$$\lim_{n\to\infty} \mathbb{P}\big(f_{k_n} \leqslant x, B\big) = \int_B \mathbf{H}(x, \omega)\, \mathbb{P}(\mathrm{d}\,\omega) \,, \qquad \forall \;\; B \in \mathcal{F} \tag{2.2}$$

at each point x in a countable, dense set $\mathbf{D} \subset \mathbb{R}$. Here $\mathbb{R} \times \Omega \ni (x, \omega) \longmapsto \mathbf{H}(x, \omega) \in [0, 1]$ is a limiting random probability distribution function; for $\mathbb{P}-$a.e. $\omega \in \Omega$, it induces on $\mathcal{B}(\mathbb{R})$ the probability measure $\boldsymbol{\mu}_\omega \equiv \boldsymbol{\mu}(\omega)$, where $\omega \mapsto \boldsymbol{\mu}_\omega$ is measurable with respect to the "tail $\sigma-$algebra" of (2.1).

2.2 Integrability of the Limit in Theorem 2.1

The following result addresses the $\mathbb{P}-$integrability of the Cesàro limit ("randomized mean") f_* in Theorem 2.1. Its first claim can be argued as in § 6.1.4 of [8]; the second as in section 5 here.

Proposition 2.3. *On a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *consider real-valued, measurable functions* $f_1, f_2, \cdots$.
(i) *Suppose that the sequence* $f_1, f_2, \cdots$ *is bounded in* $\mathbb{L}^0$ *with "determining" subsequence* $f_{k_1}, f_{k_2}, \cdots$ *, whose limiting random distribution* $\omega \mapsto \boldsymbol{\mu}_\omega$ *, as in and below* (2.2)*, satisfies*

$$\mathbb{E}^{\mathbb{P}} \int_{\mathbb{R}} \big|x\big|\, \boldsymbol{\mu}(\mathrm{d}x, \cdot) \,<\, \infty \,. \tag{2.3}$$

Then the Cesàro *convergence* (1.2) *holds* $\mathbb{P}-$*a.e. for the* $\mathbb{P}-$*integrable function* $\mathsf{f}_* = \int_{\mathbb{R}} \mathsf{x}\, \boldsymbol{\mu}(\mathrm{d}\mathsf{x}, \cdot) : \Omega \to \mathbb{R}$; *and there exist a* (*further, relabelled*) *subsequence* $\mathsf{f}_{\mathsf{k}_1}, \mathsf{f}_{\mathsf{k}_2}, \cdots$, *as well as a sequence of sets* $\mathsf{A}_1 \subseteq \mathsf{A}_2 \subseteq \cdots \subseteq \mathsf{A}_{\mathsf{n}} \subseteq \cdots$ *in* $\mathcal{F}$ *with* $\lim_{\mathsf{n}\to\infty} \mathbb{P}(\mathsf{A}_{\mathsf{n}}) = 1$, *such that the sequence*

$$\left(\mathsf{f}_{\mathsf{k}_{\mathsf{n}}}\, \mathbf{1}_{\mathsf{A}_{\mathsf{n}}}\right)_{\mathsf{n}\in\mathbb{N}} \quad \textit{is bounded in } \mathbb{L}^1(\mathbb{P})\,. \tag{2.4}$$

(ii) *Conversely, suppose there exist a subsequence* $\mathsf{f}_{\mathsf{k}_1}, \mathsf{f}_{\mathsf{k}_2}, \cdots$, *and a sequence of sets* $\mathsf{A}_1 \subseteq \mathsf{A}_2 \subseteq \cdots \subseteq \mathsf{A}_{\mathsf{n}} \subseteq \cdots$ *in* $\mathcal{F}$, *satisfying* $\lim_{\mathsf{n}\to\infty} \mathbb{P}(\mathsf{A}_{\mathsf{n}}) = 1$ *as well as the property* (2.4).

Then the Cesàro*-mean convergence* (1.2) *holds* $\mathbb{P}-$*a.e. along a* (*relabelled, "determining"*) *subsequence, with a limiting function* f_* *which is* $\mathbb{P}-$*integrable*: $\mathbb{E}^{\mathbb{P}}\big(|\mathsf{f}_*|\big) < \infty$.

The condition (2.4) is slightly stronger than (1.3), though still weaker than the $\mathbb{L}^1(\mathbb{P})-$boundedness $\sup_{\mathsf{n}\in\mathbb{N}} \mathbb{E}^{\mathbb{P}}\big(|\mathsf{f}_{\mathsf{k}_{\mathsf{n}}}|\big) < \infty$ associated with the Komlós theorem as in (1.1); it is the exact analogue, in the present context, of the Egorov-type condition in Theorem 2.2 of [8].

3 Sufficiency and Necessity in the Lacunary/Hereditary WLLN

We move now from the Strong to the Weak Law of Large Numbers (WLLN), and formulate in Theorem 3.1 right below conditions both sufficient and necessary for the validity of its Lacunary/Hereditary version in the context of general sequences. The proof of this result is presented in section 10; it relies very crucially on the novel sufficient/necessary conditions in Theorem 4.2 pertaining to the Weak Law of Large Numbers for Exchangeable Sequences, a heretofore open question.

3.1 General Sequences

Theorem 3.1. Sufficiency/Necessity in the Lacunary/Hereditary WLLN for General Sequences: *Consider a sequence of measurable functions* $\mathsf{f}_1, \mathsf{f}_2, \cdots$ *on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$, *and recall the tail* $\sigma-$*algebra of* (2.1). *Then the following conditions are equivalent:*

(i) Lacunary-Hereditary WLLN: *There exist real-valued,* $\mathcal{T}-$*measurable "correctors"* $\mathsf{D}_1, \mathsf{D}_2, \cdots$ *with* $\mathbb{P}\big(|\mathsf{D}_{\mathsf{N}}| \leqslant \mathsf{N}, \forall\, \mathsf{N} \in \mathbb{N}\big) = 1$ *and a subsequence* $\mathsf{f}_{\mathsf{k}_1}, \mathsf{f}_{\mathsf{k}_2}, \cdots$ *along which, and along whose every subsequence,*

$$\lim_{\mathsf{N}\to\infty} \left(\frac{1}{\mathsf{N}} \sum_{\mathsf{n}=1}^{\mathsf{N}} \mathsf{f}_{\mathsf{k}_{\mathsf{n}}} - \mathsf{D}_{\mathsf{N}} \right) = 0 \quad \textit{holds in } \mathbb{P}-\textit{probability.} \tag{3.1}$$

(i)′ Lacunary-Hereditary Conditional WLLN: *There exist real-valued* $\mathcal{T}-$*measurable "correctors"* $\mathsf{D}_1, \mathsf{D}_2, \cdots$ *with* $\mathbb{P}\big(|\mathsf{D}_{\mathsf{N}}| \leqslant \mathsf{N}, \forall\, \mathsf{N} \in \mathbb{N}\big) = 1$ *and a subsequence* $\mathsf{f}_{\mathsf{k}_1}, \mathsf{f}_{\mathsf{k}_2}, \cdots$ *along which, and along whose every subsequence, we have for every* $\varepsilon > 0$ *the convergence*

$$\lim_{\mathsf{N}\to\infty} \mathbb{P}\left(\left| \frac{1}{\mathsf{N}} \sum_{\mathsf{n}=1}^{\mathsf{N}} \mathsf{f}_{\mathsf{k}_{\mathsf{n}}} - \mathsf{D}_{\mathsf{N}} \right| > \varepsilon \,\middle|\, \mathcal{T} \right) = 0 \quad \textit{in } \mathbb{P}-\textit{probability.} \tag{3.2}$$

(ii) *The sequence* $\mathsf{f}_1, \mathsf{f}_2, \cdots$ *contains a* determining *subsequence* $\mathsf{f}_{\mathsf{k}_1}, \mathsf{f}_{\mathsf{k}_2}, \cdots$ *whose limit random probability distribution* $\boldsymbol{\mu}$, *as in and below* (2.2), *satisfies in* $\mathbb{P}-$*probability*

$$\lim_{\mathsf{N}\to\infty} \Big(\mathsf{N}\, \boldsymbol{\mu}\big(\mathbb{R} \setminus [-\mathsf{N}, \mathsf{N}]\big) \Big) = 0\,, \qquad \lim_{\mathsf{N}\to\infty} \frac{1}{\mathsf{N}} \int_{[-\mathsf{N},\mathsf{N}]} \mathsf{x}^2\, \boldsymbol{\mu}(\mathrm{d}\mathsf{x}, \cdot\,) = 0\,. \tag{3.3}$$

Each corrector D_{N} *in* **(i)**, **(i)′** *can be taken as the weak*$-\mathbb{L}^2$ *limit of* $\big(\mathbb{E}\big(\mathsf{f}_{\mathsf{k}_{\mathsf{n}}}\, \mathbf{1}_{\{|\mathsf{f}_{\mathsf{k}_{\mathsf{n}}}| \leqslant \mathsf{N}\}} \,\big|\, \mathcal{T}\big)\big)_{\mathsf{n}\in\mathbb{N}}$; *and* $\lim_{\mathsf{N}\to\infty} \big(\mathsf{D}_{\mathsf{N}}/\sqrt{\mathsf{N}}\,\big) = 0$ *holds in* $\mathbb{P}-$*probability.*

Remark 3.2. A Sufficient Condition: It was shown in [27] that the KOMLÓS-type condition $\lim_{N\to\infty}\big(N\cdot\sup_{n\in\mathbb{N}}\mathbb{P}\big(|f_n|>N\big)\big)=0$, is *sufficient* for the validity of the Lacunary-Hereditary Weak Law of Large Numbers (3.1). As we shall see in Proposition 4.3 ($\mathfrak{A}$), this sufficient condition is *not* necessary for this WLLN to hold.

4 The WLLN for Exchangeable Sequences

The proof of Theorem 3.1 relies crucially on conditions not just sufficient, but also *necessary*, for the validity of the Weak Law of Large Numbers (WLLN) in the context of exchangeable sequences — a question that has been open for a long time (cf. [36] and its citations).

We develop such conditions now. This is a considerable task. It encompasses Theorem 4.1 for the classical WLLN in the I.I.D. case; and Theorem 4.2 for the exchangeable case, a central result of the present paper.

4.1 The Classical Case

On a probability space $(\Omega,\mathcal{F},\mathbb{P})$, let us consider real-valued, I.I.D. $h_1,h_2,\cdots$, and adopt the notation

$$\delta_t(\varepsilon):=\mathbb{E}\big(h_1\cdot\mathbf{1}_{\{|h_1|\leqslant\varepsilon t\}}\big)\,,\quad \delta_t:=\delta_t(1) \tag{4.1}$$

$$\pi_N(\varepsilon):=\mathbb{P}\bigg(\Big|\frac{1}{N}\sum_{n=1}^{N}h_n-\delta_N\Big|>\varepsilon\bigg),\quad \pi_N:=\pi_N(1) \tag{4.2}$$

$$\tau_t(\varepsilon):=t\cdot\mathbb{P}\big(|h_1|>\varepsilon\,t\big)\,,\quad \tau_t:=\tau_t(1) \tag{4.3}$$

$$\sigma_t(\varepsilon):=\frac{1}{t}\cdot\mathbb{E}\big(h_1^2\cdot\mathbf{1}_{\{|h_1|\leqslant\varepsilon t\}}\big)\,,\quad \sigma_t:=\sigma_t(1) \tag{4.4}$$

$$\nu_t(\varepsilon):=\frac{1}{t}\cdot\mathrm{Var}\big(h_1\cdot\mathbf{1}_{\{|h_1|\leqslant\varepsilon t\}}\big)\;=\;\sigma_t(\varepsilon)-\frac{1}{t}\,\delta_t^2(\varepsilon)\,,\quad \nu_t:=\nu_t(1) \tag{4.5}$$

$$\rho_t(\varepsilon):=\frac{1}{t}\cdot\mathbb{E}\Big(\big(|h_1|\wedge\varepsilon t\big)^2\Big)=\varepsilon^2\cdot\tau_t(\varepsilon)+\sigma_t(\varepsilon)\,,\quad \rho_t:=\rho_t(1) \tag{4.6}$$

for $\varepsilon>0$, $N\in\mathbb{N}$, $t\in(0,\infty)$; all these are real numbers. We have then the celebrated KOLMOGOROV–FELLER ([28], [23]) Weak Law of Large Numbers (cf. [30]; [16], § 5.2; [22], § 2.2.3). The equivalence of **(i)**, **(ii)** below is in FELLER [23], pp. 235-236; the other equivalent conditions seem to be new.

Theorem 4.1. Necessity and Sufficiency in the Classical WLLN for I.I.D. Sequences: *For real-valued I.I.D. functions* $h_1,h_2,\cdots$, *the following are equivalent:*

(i) WLLN: *There exist "correctors"* $d_1,d_2,\cdots$ *in* $\mathbb{R}$, *so that* $\lim_{N\to\infty}\Big(\frac{1}{N}\sum_{n=1}^{N}h_n-d_N\Big)=0$ *holds in* $\mathbb{P}$*—probability; i.e., for each* $\varepsilon>0$ *we have*

$$\pi_N^{\star}(\varepsilon):=\mathbb{P}\bigg(\Big|\frac{1}{N}\sum_{n=1}^{N}h_n-d_N\Big|>\varepsilon\bigg)\;\longrightarrow\;0\,,\quad \textit{as}\;\;N\to\infty\,. \tag{4.7}$$

(ii) *With the notation of* (4.3), *for every* $\varepsilon>0$ *we have*

$$\lim_{M\to\infty}\tau_M(\varepsilon)=0\,. \tag{4.8}$$

(ii)′ *With the notation of* (4.4), *for every* $\varepsilon>0$ *we have*

$$\lim_{M\to\infty}\sigma_M(\varepsilon)=0\,. \tag{4.9}$$

(ii)″ *With the notation of* (4.5)*, for every* $\varepsilon > 0$ *we have*

$$\lim_{M\to\infty} \nu_M(\varepsilon) = 0\,. \tag{4.10}$$

(ii)‴ *With the notation of* (4.6)*, for every* $\varepsilon > 0$ *we have*

$$\lim_{M\to\infty} \rho_M(\varepsilon) = 0\,. \tag{4.11}$$

(iii) *With the notation of* (4.3)*, for every* $\varepsilon > 0$ *we have*

$$\lim_{M\to\infty} \frac{1}{M}\int_0^M \tau_t(\varepsilon)\,dt \;=\; 0\,. \tag{4.12}$$

(iii)′ *With the notation of* (4.4)*, for every* $\varepsilon > 0$ *we have*

$$\lim_{M\to\infty} M\int_M^\infty \frac{\sigma_t(\varepsilon)}{t^2}\,dt \;=\; 0\,. \tag{4.13}$$

(iii)″ *With the notation of* (4.5)*, for every* $\varepsilon > 0$ *we have*

$$\lim_{M\to\infty} M\int_M^\infty \frac{\nu_t(\varepsilon)}{t^2}\,dt \;=\; 0\,. \tag{4.14}$$

When any (*therefore, all*) *of* (4.8)–(4.14) *hold, then so does* (4.7) *with* $d_N \equiv \delta_N$ *as in* (4.1)*, and* $\pi^\star_N(\varepsilon) \equiv \pi_N(\varepsilon)$ *as in* (4.2)*.*

4.2 The Exchangeable Case

Let us consider now exchangeable, real-valued $g_1, g_2, \cdots$ on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with

$$\mathcal{T}_* := \bigcap_{n\in\mathbb{N}} \sigma\big(g_n, g_{n+1}, \cdots\big) \tag{4.15}$$

their tail $\sigma-$algebra (which is trivial when the $g_1, g_2, \cdots$ are independent, by the KOLMOGOROV zero-one law). Conditioned on this $\sigma-$algebra, the $g_1, g_2, \cdots$ become I.I.D., by the DE FINETTI theorem (e.g. [15], p. 222). For $\varepsilon > 0$, $N \in \mathbb{N}$, $t \in (0,\infty)$ we introduce the conditional versions of the quantities in (4.1)–(4.5), namely:

$$\Delta_t(\varepsilon) := \mathbb{E}\big(g_1 \cdot \mathbf{1}_{\{|g_1|\leqslant \varepsilon t\}} \,\big|\, \mathcal{T}_*\big)\,, \quad \Delta_t := \Delta_t(1) \tag{4.16}$$

$$\mathbf{\Pi}_N(\varepsilon) := \mathbb{P}\bigg(\,\bigg|\,\frac{1}{N}\sum_{n=1}^N g_n - \Delta_N\bigg| > \varepsilon\,\bigg|\,\mathcal{T}_*\bigg)\,, \quad \mathbf{\Pi}_N := \mathbf{\Pi}_N(1) \tag{4.17}$$

$$p_N(\varepsilon) := \mathbb{E}\big[\,\mathbf{\Pi}_N(\varepsilon)\,\big] = \mathbb{P}\bigg(\,\bigg|\,\frac{1}{N}\sum_{n=1}^N g_n - \Delta_N\bigg| > \varepsilon\bigg), \quad p_N := p_N(1) \tag{4.18}$$

$$T_t(\varepsilon) := t\cdot\mathbb{P}\big(\,|g_1| > \varepsilon\, t\,\big|\,\mathcal{T}_*\big)\,, \quad T_t := T_t(1) \tag{4.19}$$

$$\Sigma_t(\varepsilon) := \frac{1}{t}\cdot\mathbb{E}\big(\,g_1^2\cdot\mathbf{1}_{\{|g_1|\leqslant\varepsilon t\}}\,\big|\,\mathcal{T}_*\big)\,, \quad \Sigma_t := \Sigma_t(1) \tag{4.20}$$

$$V_t(\varepsilon) := \frac{1}{t}\cdot\mathrm{Var}\big(\,g_1\cdot\mathbf{1}_{\{|g_1|\leqslant\varepsilon t\}}\,\big|\,\mathcal{T}_*\big)\,, \quad V_t := V_t(1) \tag{4.21}$$

$$R_t(\varepsilon) := \frac{1}{t}\cdot\mathbb{E}\Big[\,\Big(\,|g_1|\wedge\varepsilon t\Big)^2\,\Big|\,\mathcal{T}_*\Big] = \varepsilon^2\cdot T_t(\varepsilon) + \Sigma_t(\varepsilon)\,, \quad R_t := R_t(1)\,. \tag{4.22}$$

It has been an open question for quite some time (cf. [36] and its references), whether conditions both necessary and sufficient exist for the Weak Law of Large Numbers in the context of exchangeable sequences, just as they exist for sequences of independent functions with common distribution.

The following result settles this question in the affirmative.

Theorem 4.2. Necessity and Sufficiency in the WLLN for Exchangeable Sequences: *For real-valued, exchangeable functions* $\mathsf{g}_1, \mathsf{g}_2, \cdots$, *the following are equivalent:*

(i) WLLN: *There exist real-valued,* $\mathcal{T}_*$*−measurable "correctors"* $\mathsf{D}_1, \mathsf{D}_2, \cdots$ *such that, for each* $\varepsilon > 0$,

$$\mathsf{p}_{\mathsf{N}}^{\dagger}(\varepsilon) := \mathbb{P}\Big(\Big| \frac{1}{\mathsf{N}} \sum_{\mathsf{n}=1}^{\mathsf{N}} \mathsf{g}_{\mathsf{n}} - \mathsf{D}_{\mathsf{N}} \Big| > \varepsilon \Big) \longrightarrow 0\,, \quad \textit{as} \;\; \mathsf{N} \to \infty\,. \tag{4.23}$$

(i)′ Conditional WLLN: *There exist real-valued,* $\mathcal{T}_*$*−measurable "correctors"* $\mathsf{D}_1, \mathsf{D}_2, \cdots$ *such that, for each* $\varepsilon > 0$,

$$\mathbf{\Pi}_{\mathsf{N}}^{\dagger}(\varepsilon) := \mathbb{P}\Big(\Big| \frac{1}{\mathsf{N}} \sum_{\mathsf{n}=1}^{\mathsf{N}} \mathsf{g}_{\mathsf{n}} - \mathsf{D}_{\mathsf{N}} \Big| > \varepsilon \,\Big|\, \mathcal{T}_* \Big) \longrightarrow 0 \quad \textit{as} \;\; \mathsf{N} \to \infty\,, \qquad \textit{in } \mathbb{P}\textit{−probability}\,. \tag{4.24}$$

(ii) *With the notation of* (4.19), (4.20), *for every* $\varepsilon > 0$ *we have* both

$$\lim_{\mathsf{t}\to\infty} \mathsf{T}_{\mathsf{t}}(\varepsilon) = 0 \quad \text{and} \quad \lim_{\mathsf{t}\to\infty} \mathsf{\Sigma}_{\mathsf{t}}(\varepsilon) = 0\,, \quad \textit{in } \mathbb{P}\textit{−probability}\,. \tag{4.25}$$

(ii)′ *With the notation of* (4.19), (4.21), *for every* $\varepsilon > 0$ *we have* both

$$\lim_{\mathsf{t}\to\infty} \mathsf{T}_{\mathsf{t}}(\varepsilon) = 0 \quad \text{and} \quad \lim_{\mathsf{t}\to\infty} \mathsf{V}_{\mathsf{t}}(\varepsilon) = 0\,, \qquad \textit{in } \mathbb{P}\textit{−probability}\,. \tag{4.26}$$

(ii)″ *With the notation of* (4.22), *for every* $\varepsilon > 0$ *we have*

$$\lim_{\mathsf{t}\to\infty} \mathsf{R}_{\mathsf{t}}(\varepsilon) = 0\,, \quad \textit{in } \mathbb{P}\textit{−probability}\,. \tag{4.27}$$

(iii) *With the notation of* (4.19), *for every* $\varepsilon > 0$ *we have*

$$\lim_{\mathsf{M}\to\infty} \frac{1}{\mathsf{M}} \int_0^{\mathsf{M}} \mathsf{T}_{\mathsf{t}}(\varepsilon)\, \mathrm{d}\mathsf{t} \;=\; 0\,, \quad \textit{in } \mathbb{P}\textit{−probability}\,. \tag{4.28}$$

(iii)′ *With the notation of* (4.20), *for every* $\varepsilon > 0$ *we have*

$$\lim_{\mathsf{M}\to\infty} \mathsf{M} \int_{\mathsf{M}}^{\infty} \frac{\mathsf{\Sigma}_{\mathsf{t}}(\varepsilon)}{\mathsf{t}^2}\, \mathrm{d}\mathsf{t} \;=\; 0\,, \quad \textit{in } \mathbb{P}\textit{−probability}\,. \tag{4.29}$$

(iii)″ *With the notation of* (4.21), *for every* $\varepsilon > 0$ *we have*

$$\lim_{\mathsf{M}\to\infty} \mathsf{M} \int_{\mathsf{M}}^{\infty} \frac{\mathsf{V}_{\mathsf{t}}(\varepsilon)}{\mathsf{t}^2}\, \mathrm{d}\mathsf{t} \;=\; 0\,, \quad \textit{in } \mathbb{P}\textit{−probability}\,. \tag{4.30}$$

When any (therefore, all) of the conditions (4.25)–(4.30) *hold, then so do* **(i)**, **(i)′** *with* $\mathsf{D}_{\mathsf{N}} \equiv \Delta_{\mathsf{N}}$ *as in* (4.16); *and then* $\mathsf{p}_{\mathsf{N}}^{\dagger}(\varepsilon) \equiv \mathsf{p}_{\mathsf{N}}(\varepsilon)$ *holds in* (4.23), (4.18), *as does* $\mathbf{\Pi}_{\mathsf{N}}^{\dagger}(\varepsilon) \equiv \mathbf{\Pi}_{\mathsf{N}}(\varepsilon)$ *in* (4.17), (4.24). *The correctors of* (4.23), (4.24) *satisfy* $\lim_{\mathsf{N}\to\infty} \big(\mathsf{D}_{\mathsf{N}}/\sqrt{\mathsf{N}}\,\big) = 0$ *in* $\mathbb{P}$*−probability.*

4.3 Some Counterexamples

Before presenting the proofs of (the essentially classical) Theorem 4.1 and of (the novel) Theorem 4.2, we discuss one subtle difference: *"There is no counterpart in Theorem 4.2 to the conditions* **(ii)**, **(ii)′**, **(ii)″** *of Theorem 4.1."* Rather, these conditions *have to be combined,* as they are in **(ii)**, **(ii)′**, **(ii)″** of Theorem 4.2, in order for their conditional versions to yield equivalent conditions in that context.

Here is some intuition for this phenomenon. First, (4.6) gives the identity $\rho_t = \sigma_t + \tau_t$. As we shall show in Lemma 7.2 below, for I.I.D. $h_1, h_2, \cdots$ *it is indeed true* that $\big(\sigma_t\big)_{t>0}$ remains bounded (respectively, tends to zero) if, and only if, $\big(\tau_t\big)_{t>0}$ does so.

But there is a subtle caveat: *it may very well happen, for an exchangeable sequence* $g_1, g_2, \cdots$ *and for some sequence of natural numbers* $N_1, N_2, \cdots$ *increasing to infinity, that we have*

$$\lim_{k\to\infty} T_{N_k} = 0 \qquad \text{while} \qquad \lim_{k\to\infty} \Sigma_{N_k} = \infty\,,$$

as well as

$$\lim_{k\to\infty} T_{N_k-1} = \infty \qquad \text{while} \qquad \lim_{k\to\infty} \Sigma_{N_k-1} = 0\,.$$

Indeed, it suffices to consider g_1 with $\mathbb{P}(g_1 = N_k) = k/N_k$ for $k \in \mathbb{N}$, and zero otherwise; if $\big(N_k\big)_{k\in\mathbb{N}} \subset \mathbb{N}$ tends to infinity sufficiently fast, g_1 is well-defined and both statements above hold.

Thus, the conjunction of the two requirements in (4.25), Theorem 4.2 (ii)′ is necessary, as shown by examples $(\mathfrak{B})$, $(\mathfrak{C})$ of Proposition 4.3. This will be proved in section 9, based on the above idea.

Proposition 4.3. Counterexamples: *The following hold.*

$(\mathfrak{A})$ *There exist exchangeable functions* $g_1, g_2, \cdots$ *with (common) symmetric distribution, for which*

$(\mathfrak{a})$ *the WLLN holds, namely,* $\lim_{N\to\infty} \mathbb{P}\Big(\Big|\sum_{n=1}^{N} g_n\Big| > \varepsilon \cdot N\Big) = 0$ *for every* $\varepsilon > 0$;

$(\mathfrak{a})'$ *the conditional (in the a.e. sense, on the tail* $\sigma-$*algebra* $\mathcal{T}_*$*) WLLN fails to hold on* $(\Omega, \mathcal{T}_*, \mathbb{P})$; *namely, we have* $\overline{\lim}_{N\to\infty} \mathbb{P}\Big(\Big|\sum_{n=1}^{N} g_n\Big| > C \cdot N \,\Big|\, \mathcal{T}_*\Big) > 0$, $\mathbb{P}-$*a.e., for every* $C > 0$;

$(\mathfrak{b})$ $\overline{\lim}_{N\to\infty} \Sigma_N = 1$, $\mathbb{P}-$*a.e.*

$(\mathfrak{c})$ $\overline{\lim}_{N\to\infty} T_N = 1$, $\mathbb{P}-$*a.e.*

$(\mathfrak{B})$ *There exist exchangeable functions* $g_1, g_2, \cdots$ *with (common) symmetric distribution, for which*

$(\mathfrak{a})$ *the WLLN fails, namely,* $\overline{\lim}_{N\to\infty} \mathbb{P}\Big(\Big|\sum_{n=1}^{N} g_n\Big| > C \cdot N\Big) = 1$ *for every* $C > 0$;

$(\mathfrak{b})$ $\lim_{N\to\infty} \Sigma_N = 0$, *in* $\mathbb{P}-$*probability; and*

$(\mathfrak{c})$ $\overline{\lim}_{N\to\infty} \mathbb{P}\big(T_N \geqslant C\big) = 1$, $\forall\ C > 0$.

$(\mathfrak{C})$ *There exist exchangeable functions* $g_1, g_2, \cdots$ *with (common) symmetric distribution, for which*

$(\mathfrak{a})$ *the WLLN fails, namely,* $\overline{\lim}_{N\to\infty} \mathbb{P}\Big(\Big|\sum_{n=1}^{N} g_n\Big| > C \cdot N\Big) = 1$ *for every* $C > 0$;

$(\mathfrak{b})$ $\lim_{N\to\infty} T_N = 0$, *in* $\mathbb{P}-$*probability; and*

$(\mathfrak{c})$ $\overline{\lim}_{N\to\infty} \mathbb{P}\big(\Sigma_N \geqslant C\big) = 1$, $\forall\ C > 0$.

4.3.1 Consequences in the General Setting

Let us consider an *arbitrary* sequence $f_1, f_2, \cdots$ of measurable functions on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and introduce, by analogy with the notation in subsection 4.2, the $\mathcal{T}-$measurable quantities

$$T_t^{\ell h}(\varepsilon) := t \cdot \sup_{n\in\mathbb{N}} \mathbb{P}\big(\,|f_n| > t\,\big|\,\mathcal{T}\big), \qquad \Sigma_t^{\ell h}(\varepsilon) := \frac{1}{t} \cdot \sup_{n\in\mathbb{N}} \mathbb{E}\Big(f_n^2 \cdot \mathbf{1}_{\{|f_n|\leqslant \varepsilon t\}} \,\Big|\, \mathcal{T}\Big), \tag{4.31}$$

$$\frac{1}{2}\left(\varepsilon^2\cdot \mathsf{T}_t^{\ell h}(\varepsilon)+\Sigma_t^{\ell h}(\varepsilon)\right)\leqslant \mathsf{R}_t^{\ell h}(\varepsilon):=\frac{1}{t}\cdot\sup_{n\in\mathbb{N}}\mathbb{E}\Big(\big(\,|f_n|\wedge\varepsilon t\big)^2\,\Big|\,\mathcal{T}\Big)\leqslant \varepsilon^2\cdot \mathsf{T}_t^{\ell h}(\varepsilon)+\Sigma_t^{\ell h}(\varepsilon)\,, \tag{4.32}$$

$$\mathsf{V}_t^{\ell h}(\varepsilon):=\frac{1}{t}\cdot\sup_{n\in\mathbb{N}}\mathrm{Var}\Big(\,|f_n|\cdot\mathbf{1}_{\{|f_n|\leqslant\varepsilon t\}}\,\Big|\,\mathcal{T}\Big)\,; \tag{4.33}$$

here the superscript "ℓh" denotes "lacunary/hereditary", and $\mathcal{T}$ is the tail-σ-algebra in (2.1).

• We **posit** now the following conditions, analogues of those in Theorem 4.2 (conditions **(i)**, **(i)′** below are identical to their "namesakes" in Theorem 3.1):

(i) *Lacunary/Hereditary WLLN:* There exist $\mathcal{T}-$measurable "correctors" $D_1, D_2, \cdots$ with $\mathbb{P}\big(|D_N|\leqslant N,\,\forall\, N\in\mathbb{N}\big)=1$ and a subsequence $f_{k_1}, f_{k_2}, \cdots$ along which, and along whose every subsequence,

$$\lim_{N\to\infty}\left(\frac{1}{N}\sum_{n=1}^{N} f_{k_n}-D_N\right)=0\quad\text{holds in }\mathbb{P}-\text{probability.} \tag{4.34}$$

(i)′ *Lacunary/Hereditary Conditional WLLN:* There exist $\mathcal{T}-$measurable "correctors" $D_1, D_2, \cdots$ with $\mathbb{P}\big(|D_N|\leqslant N,\,\forall\, N\in\mathbb{N}\big)=1$ and a subsequence $f_{k_1}, f_{k_2}, \cdots$ along which, and along whose every subsequence, we have for every $\varepsilon>0$,

$$\lim_{N\to\infty}\mathbb{P}\left(\left|\frac{1}{N}\sum_{n=1}^{N} f_{k_n}-D_N\right|>\varepsilon\,\middle|\,\mathcal{T}\right)=0\quad\text{in }\mathbb{P}-\text{probability.} \tag{4.35}$$

(ii) With the notation of (4.31), for every $\varepsilon>0$ we have both

$$\lim_{t\to\infty}\mathsf{T}_t^{\ell h}(\varepsilon)=0\quad\text{and}\quad\lim_{t\to\infty}\Sigma_t^{\ell h}(\varepsilon)=0\,,\quad\text{in }\mathbb{P}-\text{probability}\,. \tag{4.36}$$

(ii)′ With the notation of (4.31), (4.32), for every $\varepsilon>0$ we have both

$$\lim_{t\to\infty}\mathsf{T}_t^{\ell h}(\varepsilon)=0\quad\text{and}\quad\lim_{t\to\infty}\mathsf{V}_t^{\ell h}(\varepsilon)=0\,,\quad\text{in }\mathbb{P}-\text{probability.} \tag{4.37}$$

(ii)″ With the notation of (4.32), for every $\varepsilon>0$ we have

$$\lim_{t\to\infty}\mathsf{R}_t^{\ell h}(\varepsilon)=0\,,\quad\text{in }\mathbb{P}-\text{probability.} \tag{4.38}$$

(iii) With the notation of (4.31), for every $\varepsilon>0$ we have

$$\lim_{M\to\infty}\frac{1}{M}\int_0^M \mathsf{T}_t^{\ell h}(\varepsilon)\,dt\;=\;0\,,\quad\text{in }\mathbb{P}-\text{probability.} \tag{4.39}$$

(iii)′ With the notation of (4.31), for every $\varepsilon>0$ we have

$$\lim_{M\to\infty}M\int_M^\infty\frac{\Sigma_t^{\ell h}(\varepsilon)}{t^2}\,dt\;=\;0\,,\quad\text{in }\mathbb{P}-\text{probability.} \tag{4.40}$$

(iii)″ With the notation of (4.33), for every $\varepsilon>0$ we have

$$\lim_{M\to\infty}M\int_M^\infty\frac{\mathsf{V}_t^{\ell h}(\varepsilon)}{t^2}\,dt\;=\;0\,,\quad\text{in }\mathbb{P}-\text{probability.} \tag{4.41}$$

It is very tempting at this point, to conjecture that all these conditions above are equivalent. *The counterexamples of Proposition 4.3 demonstrate, however, that this conjecture is false.* While **(i)** $\Leftrightarrow$ **(i)′** still holds in the above list, Proposition 4.3 ($\mathfrak{A}$) shows that **(i)** does not imply any of the conditions (4.36)–(4.41) of the above list.

In the reverse direction, however, *each of* (4.36)–(4.41) *is strong enough to imply both* **(i)**, **(i)′**. Indeed, after passing to a determining subsequence $f_{k_1}, f_{k_2}, \cdots$ of $f_1, f_2, \cdots$ with associated exchangeable $g_1, g_2, \cdots$ as in (5.5) below, the assumptions (4.36)–(4.41) on the $f_1, f_2, \cdots$ imply the corresponding properties for the sequence $g_1, g_2, \cdots$. Hence, by Theorem 4.2, the $g_1, g_2, \cdots$ satisfy the WLLN, thus also condition **(ii)** of Theorem 3.1. This shows the following.

Corollary 4.4. Sufficient Conditions: *Any one of the conditions* (4.36)–(4.41) *implies both the Lacunary/Hereditary WLLN* (4.34)*, and the Lacunary/Hereditary Conditional WLLN* (4.35)*.*

5 The Proof of Theorem 2.1

We start with a straightforward observation.

Lemma 5.1. *Suppose that the real-valued, measurable* $f_1, f_2, \cdots$ *contain a subsequence* $f_{k_1}, f_{k_2}, \cdots$ *along which, and along whose every further subsequence,* (3.1) *holds for some real-valued* $D_1, D_2, \cdots$ *measurable with respect to the tail* $\sigma-$*algebra* $\mathcal{T}$ *of* (2.1)*.*

Then we may assume (*passing to a further subsequence*) *that* $f_{k_1}, f_{k_2}, \cdots$ *is bounded in probability.*

Proof: Assume the contrary; i.e., that $f_{k_1}, f_{k_2}, \cdots$ is *not* bounded in probability. Then

$$\lim_{M\to\infty} \sup_{n\in\mathbb{N}} \mathbb{P}\left(\left|f_{k_n}\right| > M\right) > \alpha \tag{5.1}$$

holds for some $\alpha > 0$; and we construct, inductively, a subsequence $f_{k_{n_1}}, f_{k_{n_2}}, \cdots$ with the property

$$\mathbb{P}\left(\left| \frac{1}{L} \sum_{\ell=1}^{L} f_{k_{n_\ell}} - D_L \right| > 1 \right) > \frac{\alpha}{2}\,, \qquad \forall\ L \in \mathbb{N}\,. \tag{5.2}$$

This will provide the desired contradiction to the hereditary validity of (4.34).

Indeed, suppose that the terms $f_{k_{n_1}}, \cdots, f_{k_{n_L}}$ have been chosen to satisfy (5.2); selecting then $M > 2$ in (5.1) large enough, we determine the index $k_{n_{L+1}} > k_{n_L}$ of the next term, so that

$$\mathbb{P}\left(\left|\frac{1}{L+1}\sum_{\ell=1}^{L+1} f_{k_{n_\ell}} - D_{L+1}\right| > \frac{M}{2}\right) > \mathbb{P}\left(\frac{\left|f_{k_{n_{L+1}}}\right|}{L+1} > M\right) - \mathbb{P}\left(\left|\frac{1}{L+1}\sum_{\ell=1}^{L} f_{k_{n_\ell}} - D_{L+1}\right| > \frac{M}{2}\right) > \frac{\alpha}{2}$$

holds. Here, the last displayed probability is less than $\alpha/2$ for $M > 2$ sufficiently large, because of (4.34); whereas, after fixing such M, we can choose $k_{n_{L+1}} > k_{n_L}$ so that the next-to-last displayed probability is bigger than α, on the strength of (5.1).

Recalling $M > 2$, we see that the inductive step $L \longmapsto L+1$ for (5.2) has been established. □

5.1 Proof of Theorem 2.1 (i)

This is a consequence of the Komlós [31] result, and of the fact that the Cesàro limit f_* can be defined on the increasing sequence of sets $A_1, A_2, \cdots$ consistently, via diagonalization, and in the spirit of the "Chacon Biting Lemma" in [12].

More specifically, from the KOMLÓS theorem there exist for each $j \in \mathbb{N}$ a measurable $f_*^{(j)} : \Omega \to \mathbb{R}$ and a subsequence $f_{k_1}^{(j)}, f_{k_2}^{(j)}, \cdots$ of $f_{k_1}, f_{k_2}, \cdots$, so that $f_*^{(j)} = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f_{k_n}^{(j)} \mathbf{1}_{A_j}$ holds $\mathbb{P}-$a.e. Clearly, the $f_{k_1}^{(j+1)}, f_{k_2}^{(j+1)}, \cdots$ can be selected as a subsequence of $f_{k_1}^{(j)}, f_{k_2}^{(j)}, \cdots$; and $f_*^{(j+1)} = f_*^{(j)}$ holds $\mathbb{P}-$a.e. on A_j. We define now, consistently,

$$f_*(\omega) := f_*^{(j)}(\omega)\,, \quad \omega \in A_j\,; \qquad f_*(\omega) := 0\,, \quad \omega \in \Omega \setminus A$$

with $A := \bigcup_{j\in\mathbb{N}} A_j$ a set of full measure $\mathbb{P}(A) = 1$, and notice that for each fixed $j \in \mathbb{N}$ we have

$$f_*(\omega) = f_*^{(j)}(\omega) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f_{k_n}^{(j)}(\omega) = \lim_{N\to\infty} \frac{1}{N} \left(\sum_{n=1}^{j} f_{k_n}^{(j)}(\omega) + \sum_{n=j+1}^{N} f_{k_n}^{(j)}(\omega) \right)$$

$$= \lim_{N\to\infty} \frac{1}{N} \sum_{n=j+1}^{N} f_{k_n}^{(n)}(\omega) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f_{k_n}^{(n)}(\omega)\,, \qquad \omega \in A_j\,. \tag{5.3}$$

The fourth equality recalls that $\left(f_{k_n}^{(n)}\right)_{n>j}$ is a subsequence of $\left(f_{k_n}^{(j)}\right)_{n\in\mathbb{N}}$, and the hereditary nature of the CESÀRO-mean convergence in the KOMLÓS theorem. The first and last terms of (5.3) do not involve j, thus agree on the set of full measure $A = \bigcup_{j\in\mathbb{N}} A_j$. This is the CESÀRO-mean convergence along the "diagonal" subsequence $\left(f_{k_n}^{(n)}\right)_{n\in\mathbb{N}}$; its hereditary character is clear. □

5.2 Proof of Theorem 2.1 (ii)

STEP 1: Determining Subsequence. On the strength of the boundedness-in-probability established in Lemma 5.1, we may assume (passing to a further subsequence, if needed) the subsequence $f_{k_1}, f_{k_2}, \cdots$ to be "determining" in the sense, and with the notation, of (2.2), (2.1).

STEP 2: Approximation by an Exchangeable Sequence. Following the trail blazed by ALDOUS [2]-[4] (cf. DACUNHA-CASTELLE [17]; [9]), we summon now an *exchangeable* sequence of functions $g_1, g_2, \cdots$, conditionally independent and with the random probability measure $\omega \mapsto \mu_\omega$ (cf. (2.2) and the paragraph below it) as their common conditional distribution given the tail-σ-algebra

$$\mathcal{T}_* := \bigcap_{n\in\mathbb{N}} \sigma\big(g_n, g_{n+1}, \cdots\big) \subseteq \mathcal{T}\,; \tag{5.4}$$

and also such that, passing to a further subsequence of the determining $f_{k_1}, f_{k_2}, \cdots$ if necessary,

$$\big(f_{k_{n+1}}, \cdots, f_{k_{n+M}}\big) \text{ converges in distribution as } n \to \infty \text{ to } \big(g_1, \cdots, g_M\big) \tag{5.5}$$

for all $M \in \mathbb{N}$ and $M-$tuples $(k_{n+1}, \cdots, k_{n+M})$ with $k_{n+1} < \cdots < k_{n+M}$, $\lim_{n\to\infty} k_{n+1} = \infty$.

It follows now from Proposition 6.1(ii), that the assumed hereditary $\mathbb{P}-$a.e. validity of (1.2), for some $f_* \in \mathbb{L}^0$, implies the $\mathbb{P}-$a.e. finiteness

$$\mathbb{P}\big(\kappa < \infty\big) = 1 \quad \text{for the function} \quad \Omega \ni \omega \longmapsto \kappa(\omega) := \int_{\mathbb{R}} |x|\, \mu_\omega(dx) \in [0, \infty]\,. \tag{5.6}$$

STEP 3: The condition (1.3). On the strength of (5.6), we find for any given $\varepsilon \in (0,1)$ some $K \in (0,\infty)$ such that the set $B := \{\kappa \leqslant K\} \in \mathcal{T}$ has measure $\mathbb{P}(B) > 1-\varepsilon$. The sequence $f_{k_1}, f_{k_2}, \cdots$ is determining, so $f_{k_1}\mathbf{1}_B$, $f_{k_2}\mathbf{1}_B, \cdots$ converges in distribution to a random variable ξ with distribution ν and first absolute moment bounded by K (this holds even conditionally on $\mathcal{T}$).

Then, under very mild conditions on the underlying space $(\Omega, \mathcal{F})$, which we can certainly assume here, there exist a (relabelled) subsequence $f_{k_1}, f_{k_2}, \cdots$, as well as a sequence of random variables $\xi_1, \xi_2, \cdots$ with common distribution $\boldsymbol{\nu}$, such that $\lim_{n\to\infty} \big(f_{k_n} \mathbf{1}_B - \xi_n\big) = 0$ holds $\mathbb{P}-$a.e. (in the spirit of the SKOROKHOD, a.k.a. STRASSEN-DUDLEY, construction in [11], p. 70; a more direct reference is Lemma 4, p. 41 of [1]). But now EGOROV's theorem gives a subset $A \in \mathcal{T}$ of B with $\mathbb{P}(A) > 1 - 2\varepsilon$ and a (further, relabelled) subsequence with ξ bounded on A, $\sup_{\omega\in A} \big|f_{k_n}(\omega) - \xi_n(\omega)\big| \to 0$ as $n \to \infty$, and $\sup_{n\in\mathbb{N}} \big\| f_{k_n} \mathbf{1}_A \big\|_{\mathbb{L}^1(\mathbb{P})} < K + 1$.

Iterating this procedure, we come up with real numbers $0 < K_1 < K_2 < \cdots < K_j \uparrow \infty$ and EGOROV-type sets $A_1 \subseteq A_2 \subseteq \cdots \subseteq A_j \subseteq \cdots$ in $\mathcal{T}$, such that:
(i) for each $j \in \mathbb{N}$, we have the bounds $\mathbb{P}\big(A_j\big) \geqslant 1 - 2^{-j}$, $\kappa(\omega) \leqslant K_j$ for $\mathbb{P}-$a.e. $\omega \in A_j$; and
(ii) the sequence $\big(f_{k_n} \cdot \mathbf{1}_{A_j}\big)_{n\in\mathbb{N}}$ is bounded in $\mathbb{L}^1(\mathbb{P})$, as claimed in (1.3). □

Remark 5.2. The inclusion $\mathcal{T}_* \subseteq \mathcal{T}$ in (5.4) follows from the study of "almost exchangeable" sequences in [10], most notably its Proposition 2.3. *This inclusion can, very easily, be strict.*

Indeed, consider the sequence of uniformly bounded functions $f_n := (1/n) \cdot \mathbf{1}_{[0,1/2]}$, $n \in \mathbb{N}$ on $([0,1], \mathcal{B}, \boldsymbol{\lambda})$. This sequence is determining, and the associated exchangeable functions $g_1 \equiv 0$, $g_2 \equiv 0, \cdots$ have trivial tail $\sigma-$algebra $\mathcal{T}_*$; yet the tail $\sigma-$algebra $\mathcal{T}$ of the $f_1, f_2, \cdots$ contains the interval $(1/2, 1]$. All results below continue to hold, if conditioning on $\mathcal{T}_*$ is replaced by conditioning on $\mathcal{T}$.

5.3 Proof of Theorem 2.1 (iii)

Assuming the validity of (1.3) for some subsequence $f_{k_1}, f_{k_2}, \cdots$ and sets $A_1 \subseteq A_2 \subseteq \cdots \subseteq A_j \subseteq \cdots$ in $\mathcal{F}$ with $\lim_{j\to\infty} \mathbb{P}(A_j) = 1$, we define a new probability measure $\mathbb{Q} \sim \mathbb{P}$ via

$$\frac{d\mathbb{Q}}{d\mathbb{P}}(\omega) := \Big(\sum_{j\in\mathbb{N}} \alpha_j \mathbf{1}_{A_j}(\omega) \Big) \Big/ \sum_{j\in\mathbb{N}} \alpha_j \mathbb{P}(A_j), \tag{5.7}$$

where the numbers $\alpha_j \in (0,1)$ satisfy $\alpha_j \cdot \sup_{n\in\mathbb{N}} \big\| \big(1 + |f_{k_n}|\big) \mathbf{1}_{A_j} \big\|_{\mathbb{L}^1(\mathbb{P})} \leqslant (1/j^2)$, $\forall\, j \in \mathbb{N}$; thus, $\sup_{n\in\mathbb{N}} \big\|f_{k_n}\big\|_{\mathbb{L}^1(\mathbb{Q})} < \infty$ follows. Whereas, if $\sup_{n\in\mathbb{N}} \big\|f_{k_n}\big\|_{\mathbb{L}^1(\mathbb{Q})} < \infty$ holds for some probability measure $\mathbb{Q} \sim \mathbb{P}$, then so does (1.3) for the sets $A_j := \big\{\big(d\mathbb{P}/d\mathbb{Q}\big) \leqslant j\big\}$, $j \in \mathbb{N}$. □

6 Linking the Lacunary/Hereditary Case with Exchangeability

We state now and prove a result of considerable independent interest, crucial for proving Theorems 2.1, 3.1. It comes in two versions, one for the Weak, the other for the Strong, Law of Large Numbers.

Proposition 6.1. *On a given probability space $(\Omega, \mathcal{F}, \mathbb{P})$, consider a sequence of real-valued, measurable functions $f_1, f_2, \cdots$ containing a subsequence $f_{k_1}, f_{k_2}, \cdots$ with the stable convergence (2.2).*
(i) *Suppose that $\lim_{N\to\infty} \big(\frac{1}{N}\sum_{n=1}^{N} f_{k_n} - D_N\big) = 0$ holds in $\mathbb{P}-$probability for some $\mathcal{T}-$measurable correctors $D_1, D_2, \cdots$, along this subsequence $f_{k_1}, f_{k_2}, \cdots$ and all its subsequences. Then, the approximating exchangeable functions $g_1, g_2, \cdots$ of (5.5) also satisfy*

$$\lim_{N\to\infty} \Big(\frac{1}{N} \sum_{n=1}^{N} g_n - D_N \Big) = 0 \quad \text{in } \mathbb{P}-\text{probability} \tag{6.1}$$

with the same $\mathcal{T}-$measurable $D_1, D_2, \cdots$, as well as with the $\mathcal{T}_*-$measurable $\Delta_1, \Delta_2, \cdots$ of (4.16).

(ii) *In the same setting, suppose that the subsequence* $f_{k_1}, f_{k_2}, \cdots$, *and all its subsequences, satisfies* $\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f_{k_n} = 0$, $\mathbb{P}-$*a.e. Then the exchangeable* $g_1, g_2, \cdots$ *of* (5.5) *also have this property:*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} g_n \,=\, 0\,, \qquad \mathbb{P}-a.e. \tag{6.2}$$

6.1 Proof of Proposition 6.1(i)

To alleviate notation somewhat, let us write $f_1, f_2, \cdots$ for the sequence $f_{k_1}, f_{k_2}, \cdots$. We shall argue by contradiction: If (6.1) fails, there exist $\alpha \in (0,1)$ and a sequence $N_1, N_2, \cdots$ in $\mathbb{N}$ with

$$\mathbb{P}\left(\left| \sum_{n=1}^{N_j} g_n - N_j \cdot D_{N_j} \right| > \alpha N_j \right) \,\geqslant\, \alpha\,, \qquad \forall\ j \in \mathbb{N}\,. \tag{6.3}$$

Passing inductively to a (relabelled) subsequence of this $N_1, N_2, \cdots$ we may suppose that, for each $\ell \in \mathbb{N}$ and already defined $N_1, N_2, \cdots, N_\ell$, the next term $N_{\ell+1}$ is chosen big enough, so that, with $M_\ell := N_1 + \cdots + N_\ell$,

$$\mathbb{P}\left(\left|F\right| > \frac{\alpha}{4} N_{\ell+1} \right) \,<\, \frac{\alpha}{4}$$

holds for any sum F of at most $M_\ell -$many terms from the sequences $f_1, f_2, \cdots$ or $g_1, g_2, \cdots$.

Next, we choose a sequence $K_1, K_2, \cdots$ in $\mathbb{N}$ which converges to infinity sufficiently fast, and is such that for every $j \in \mathbb{N}$ we have $K_j + N_j \,<\, K_{j+1}$ as well as

$$\boldsymbol{\rho}_j \left(\text{distribution}\big(f_{K_j+1}, \cdots, f_{K_j+N_j}\big),\, \text{distribution}\big(g_1, \cdots, g_{N_j}\big) \right) \,<\, \varepsilon_j\,. \tag{6.4}$$

Here $\boldsymbol{\rho}_j$ denotes the Prokhorov distance ([11], p. 72) between probability measures on the Borel sets of $\mathbb{R}^{N_j}$, and $\varepsilon_j \in (0,1)$ is small enough so we have also

$$\mathbb{P}\left(\left| \sum_{n=1}^{N_j} f_{K_j+n} - \sum_{n=1}^{N_j} g_n \right| > \frac{\alpha}{4} \right) \,<\, \frac{\alpha}{4}\,, \qquad \forall\ j \in \mathbb{N}\,. \tag{6.5}$$

We denote now by $f_{k_1}, f_{k_2}, \cdots$ the subsequence of $f_1, f_2, \cdots$ obtained by gluing together consecutive blocks $\mathcal{I}_j := \big\{K_j + 1, \cdots, K_j + N_j\big\}$, $j \in \mathbb{N}$ of indices and, for each $L \in \mathbb{N}$, estimate

$$\mathbb{P}\left(\left| \sum_{n=1}^{M_{L+1}} f_{k_n} - M_{L+1} \cdot D_{M_{L+1}} \right| > \frac{\alpha}{4} M_{L+1} \right) \geqslant \mathbb{P}\left(\left| \sum_{n=M_L+1}^{M_{L+1}} f_{k_n} - M_{L+1} \cdot D_{M_{L+1}} \right| > \frac{2\,\alpha}{4} M_{L+1} \right) - \frac{\alpha}{4}$$

$$\geqslant \mathbb{P}\left(\left| \sum_{n=M_L+1}^{M_{L+1}} g_n - M_{L+1} \cdot D_{M_{L+1}} \right| > \frac{3\,\alpha}{4} M_{L+1} \right) - \frac{2\,\alpha}{4}$$

$$\geqslant \mathbb{P}\left(\left| \sum_{n=1}^{M_{L+1}} g_n - M_{L+1} \cdot D_{M_{L+1}} \right| > \alpha M_{L+1} \right) - \frac{3\,\alpha}{4} \,\geqslant\, \frac{\alpha}{4}\,;$$

here, the second inequality holds in light of (5.5), (6.4)–(6.5), and the last in light of (6.3). But the resulting subsequence of $f_{k_1}, f_{k_2}, \cdots$ fails then to satisfy (3.1), contrary to the premise of Proposition 6.1(i) and leading to the desired contradiction.

The Weak Law of Large Numbers (6.1) is thus established. □

6.2 Proof of Proposition 6.1(ii)

We argue again by contradiction, assuming that the approximating exchangeable functions $g_1, g_2, \cdots$ of (5.5) fail to converge in CESÀRO mean to $f_* \equiv 0$, $\mathbb{P}-$a.e.; there exists then $\alpha \in (0,1)$ with

$$\mathbb{P}(B) \geqslant \alpha\,, \quad \text{for} \quad B := \bigcap_{M\in\mathbb{N}} \bigcup_{N\geqslant M,\, N\in\mathbb{N}} B_N^{(g)}\,, \quad B_N^{(g)} := \left\{ \left| \sum_{n=1}^{N} g_n \right| > \alpha N \right\}. \tag{6.6}$$

We proceed as in § 6.1: consider sequences of natural numbers $N_1, N_2, \cdots$ and $K_1, K_2, \cdots$ with $K_j + N_j < K_{j+1}$, so that the terms over the disjoint blocks of indices $\mathcal{I}_j := \{K_j+1, \cdots, K_j+N_j\}$, $j \in \mathbb{N}$ are "glued together" in a (relabelled) subsequence $f_{k_1}, f_{k_2}, \cdots$.

There is now the following additional feature: we choose also for each $j \in \mathbb{N}$ a (large) number $C_j \ll N_j$ and ensure that, in each inductive step $L \mapsto L+1$, the following holds for these choices: Suppose K_L, C_L, N_L have been chosen, and let $M_L := N_1 + \cdots + N_L$; choose C_{L+1} first, so that

$$\mathbb{P}\left(\left| \sum_{n=1}^{M_L} f_{k_n} \right| > \frac{\alpha}{4}\, C_{L+1} \right) < \frac{\alpha}{4}\,, \qquad \mathbb{P}\left(\left| \sum_{n=1}^{M_L} g_n \right| > \frac{\alpha}{4}\, C_{L+1} \right) < \frac{\alpha}{4}$$

hold, then choose $N_{L+1} > C_{L+1}$ so that, in the notation of (6.6) for the sets defined there, we have

$$\mathbb{P}\left(\bigcup_{N=M_L+C_{L+1}}^{N_{L+1}} B_N^{(g)} \right) > \alpha\,.$$

Finally, we choose K_L large enough, so that the distributions of $(f_{K_j+1}, \cdots, f_{K_j+N_j})$ and $(g_1, \cdots, g_{N_j})$ are close enough, in such a manner that the probability of the set

$$\bigcup_{N=M_L+C_{L+1}}^{N_{L+1}} B_N^{(g)} \quad \text{is close to that of} \quad \bigcup_{N=M_L+C_{L+1}}^{N_{L+1}} B_N^{(f)}\,, \qquad \text{where} \quad B_N^{(f)} := \left\{ \left| \sum_{n=1}^{N} f_{K_j+n} \right| > \alpha N \right\}.$$

We follow then the chain of inequalities at the end of § 6.1, and obtain the desired contradiction.

The Strong Law of Large Numbers (6.2) is thus established. □

Remark 6.2. In the context of Proposition 6.1(ii), the exchangeable functions $g_1, g_2, \cdots$ of (5.5), conditionally independent and with common distribution $\mu_\omega \equiv \mu(\omega)$ given $\sigma(\mu)$, converge in CESÀRO mean to zero, *for* $\mathbb{P}-$*a.e.* $\omega \in \Omega$ (i.e., (6.2) holds). From the converse to the Strong Law of Large Numbers (LOÈVE [33], p. 251), this common conditional distribution satisfies then

$$\mathbb{E}\big[\, |g_1| \,\big|\, \mathcal{I}_* \,\big](\omega) = \int_{\mathbb{R}} |x|\ \mu_\omega(dx) < \infty \qquad \text{for } \mathbb{P}-\text{a.e. } \omega \in \Omega\,. \tag{6.7}$$

(The corresponding statement for the Weak Law of Large Numbers, i.e., in the context of Proposition 6.1(i), is developed in subsection 10.1.)

Conversely, (6.7) implies the CESÀRO-convergence (6.2) of the $g_1, g_2, \cdots$ to zero, $\mathbb{P}-$a.e.; as well as this same convergence for the "determining" subsequence $f_{k_1}, f_{k_2}, \cdots$ and for all its subsequences. Compare with the general formulation of the hereditary SLLN in Theorem 3 in ALDOUS [2].

7 The Proof of Theorem 4.1

We start with a couple of preliminary results pertaining to the setting of Theorem 4.1.

Lemma 7.1. *For any given* $\eta > 0$ *there exists an integer* $N_\eta \in \mathbb{N}$ *such that, in the notation of* (4.1), (4.4)*, we have*

$$\delta_N^2 \leqslant \eta\, N \big(1 + \sigma_N\big)\,, \qquad \forall\ \ N \geqslant N_\eta\,. \tag{7.1}$$

Proof: We select $p < \eta/2$, and $M > 0$ big enough so that $\mathbb{P}\big(|h_1| > M\big) < p$. Then we have

$$\mathbb{E}\big(h_1^2 \cdot \mathbf{1}_{\{|h_1| \leqslant M\}}\big) \leqslant M^2 \leqslant \big(\eta/2\big)\, N \qquad \text{as long as}\ \ N > M\,,\ \eta\, N > 2\, M^2\,.$$

In particular, with $h^{(M)} := h_1 \cdot \mathbf{1}_{\{|h_1| \leqslant M\}}$, we have $\mathbb{E}\big[\,\big(h^{(M)}\big)^2\,\big] \leqslant \big(\eta/2\big)\, N$.

We note now that the difference $h^{(N)} - h^{(M)} = h_1 \cdot \mathbf{1}_{\{M < |h_1| \leqslant N\}}$ is supported on a subset of $\{\,|h_1| > M\,\}$, whose $\mathbb{P}-$measure does not exceed p. Therefore, by Cauchy-Schwarz we have

$$\begin{aligned} \big|\,\mathbb{E}\big(h^{(N)} - h^{(M)}\big)\,\big|^2 \leqslant \Big(\,\mathbb{E}\big(\,\big|h^{(N)} - h^{(M)}\big|\,\big)\Big)^2 &\leqslant \mathbb{E}\Big(\Big(\big|h_1\big| \cdot \mathbf{1}_{\{M < |h_1| \leqslant N\}}\Big)\Big)^2 \\ &\leqslant \mathbb{E}\Big(h_1^2 \cdot \mathbf{1}_{\{|h_1| \leqslant N\}}\Big) \cdot \mathbb{P}\big(|h_1| > M\big) \ \leqslant\ N\, \sigma_N \cdot p\,. \end{aligned}$$

Therefore, $\delta_N = \mathbb{E}\big[\,h^{(N)}\,\big] = \mathbb{E}\big[\,\big(h^{(N)} - h^{(M)}\big) + h^{(M)}\,\big]$ satisfies

$$\delta_N^2 \ \leqslant\ 2 \cdot \mathbb{E}\big[\,\big(h^{(N)} - h^{(M)}\big)^2\,\big] + 2 \cdot \mathbb{E}\big[\,\big(h^{(M)}\big)^2\,\big] \ \leqslant\ 2\, p\, N\, \sigma_N + \eta\, N \ <\ \eta\, N\,\big(1 + \sigma_N\big)\,. \qquad \square$$

Lemma 7.2. *In the notation of* (4.3), (4.4)*, we have*

$$\lim_{M \to \infty} \tau_M = 0 \iff \lim_{M \to \infty} \sigma_M = 0\,. \tag{7.2}$$

Proof: We recall the quantities of (4.3), (4.4) as $\tau(t) = t \cdot \mathbb{P}\big(|h_1| > t\big)$, $\sigma(t) = \frac{1}{t} \cdot \mathbb{E}\big(h_1^2 \cdot \mathbf{1}_{\{|h_1| \leqslant t\}}\big)$ for $t \in (0, \infty)$; setting $\tau(0) := \tau(0+) = 0$, $\sigma(0) := \sigma(0+) = 0$, we obtain functions defined and right-continuous on $[0, \infty)$, as well as bounded on compact intervals. Then integration by parts gives

$$\tau(M) = \frac{2}{M} \int_0^M \tau(t)\, dt - \sigma(M)\,, \qquad M > 0 \tag{7.3}$$

as in (7.7) of p. 235 in Feller, Vol. II [23] (after correcting for a typo there), which leads to the implication $\lim_{M \to \infty} \tau(M) = 0 \implies \lim_{M \to \infty} \sigma(M) = 0$.

Now let us look at (7.3) as an integral equation for the function $\tau(\cdot)$, in terms of some "given" function $\sigma(\cdot)$. We assume for a moment that both $\tau(\cdot)$ and $\sigma(\cdot)$ are continuous and continuously differentiable, and obtain from (7.3) by differentiation $\tau(t) - t\,\tau'(t) = \sigma(t) + t\,\sigma'(t)$, thus also

$$\left(\frac{\tau(t)}{t}\right)' = -\frac{1}{t^2}\,\big(t\,\sigma(t)\big)'. \qquad \text{Noting} \quad \lim_{t \downarrow 0}\left(\frac{\tau(t)}{t}\right) = 1\,, \quad \lim_{t \uparrow \infty}\left(\frac{\tau(t)}{t}\right) = 0\,, \quad \lim_{t \downarrow 0}\left(\frac{\sigma(t)}{t}\right) = 0$$

from (7.3), and integrating by parts, we obtain the "solution" of the above "integral equation" as

$$\tau(M) \ =\ 2\, M \int_M^\infty \frac{\sigma(t)}{t^2}\, dt - \sigma(M)\,, \qquad M > 0\,. \tag{7.4}$$

Having thus arrived at the expression (7.4), it is easy to argue that it follows from (7.3) in general (i.e., without assuming continuous differentiability) via repeated application of Tonelli's theorem. The implication $\lim_{M \to \infty} \sigma(M) = 0 \implies \lim_{M \to \infty} \tau(M) = 0$ is now clear from (7.4). $\square$

7.1 Proof of Theorem 4.1

With these two results in place, and noting that it suffices to deal with $\varepsilon = 1$ in Theorem 4.1, we start by observing that the equivalence **(ii)** $\Leftrightarrow$ **(ii)′** is the subject of Lemma 7.2.

The implication **(i)** $\Rightarrow$ **(ii)** is from FELLER, Vol. II [23], p. 236; and **(ii)** $\Rightarrow$ **(i)** is proved in FELLER, Vol. II ([23], p. 235) where it is pointed out that, with $\mathsf{d}_N \equiv \delta_N = \mathbb{E}\big(h_1 \cdot \mathbf{1}_{\{|h_1| \leqslant N\}}\big)$, the bound

$$\pi_N(\eta) \,=\, \mathbb{P}\bigg(\Big| \frac{1}{N} \sum_{n=1}^{N} h_n - \delta_N \Big| > \eta \bigg) \,\leqslant\, \tau_N + \frac{1}{\eta^2}\, \sigma_N \,\longrightarrow\, 0\,, \qquad \text{as} \;\; N \to \infty\,, \tag{7.5}$$

holds for $\eta > 0$. We note here $\lim_{N\to\infty} \tau_N = 0 \;\Rightarrow\; \lim_{N\to\infty} \sigma_N = 0$ from Lemma 7.2.

We turn to the equivalence **(ii)′** $\Leftrightarrow$ **(ii)″**. From Lemma 7.1, for any given $\eta \in (0,1)$ we have

$$\sigma_M = \nu_M + \frac{1}{M}\,\delta_M^2 \geqslant \nu_M = \sigma_M - \frac{1}{M}\,\delta_M^2 \geqslant \sigma_M - \eta\big(1+\sigma_M\big) = \big(1-\eta\big)\sigma_M - \eta \tag{7.6}$$

for every $M \geqslant N_\eta$. The equivalence $\lim_{M\to\infty} \nu_M = 0 \iff \lim_{M\to\infty} \sigma_M = 0$ follows from (7.6), as does the equivalence **(iii)′** $\Leftrightarrow$ **(iii)″**. Now, we have shown that the conditions **(ii)**, **(ii)′** are equivalent, so each of them implies also $\lim_{M\to\infty} \rho_M = \lim_{M\to\infty} \big(\tau_M + \sigma_M\big) = 0$, namely, **(ii)‴**; and is in turn implied by it, because all quantities in question are positive.

But **(ii)‴** is also equivalent to **(iii)** on account of the identity $\rho(M) = (2/M)\int_0^M \tau(t)\,dt$ from (7.3); and to the condition **(iii)′** on account of $\rho(M) = 2M\int_0^M (\sigma(t)/t^2)\,dt$ from (7.4). □

8 The Proof of Theorem 4.2

The conditions **(i)**, **(i)′** *are equivalent.* Indeed, for all $\varepsilon > 0$, $N \in \mathbb{N}$, we have $\pi_N^*(\varepsilon) = \mathbb{E}\big[\,\mathbf{\Pi}_N^*(\varepsilon)\,\big]$ and $0 \leqslant \mathbf{\Pi}_N^*(\varepsilon) \leqslant 1$, $\mathbb{P}$–a.e. Now, convergence in probability implies $\mathbb{L}^1$–convergence under uniform boundedness (e.g., Theorem 4.5.4 in CHUNG [16]), so **(i)** $\Rightarrow$ **(i)′** follows. On the other hand, the reverse implication **(i)′** $\Rightarrow$ **(i)** is straightforward, since $\mathbb{L}^1-$ convergence implies convergence in probability.

For the remaining equivalences, we invoke Proposition 8.1 and Corollary 8.2 right below.

Proposition 8.1. *In the context of Theorem 4.1 and with the notation of* (4.2)–(4.5)*, there is a universal constant* $C \in (0,\infty)$ *such that, for all* $N \in \mathbb{N}$ *sufficiently large, we have*

$$\pi_N \,\leqslant\, C \cdot \big[\,\big(\tau_N \wedge 1\big) + \big(\sigma_N \wedge 1\big)\,\big] \tag{8.1}$$

$$\big(\tau_N(4) \wedge 1\big) + \big(\sigma_N \wedge 1\big) \,\leqslant\, C \cdot \pi_N\,, \qquad \big(\tau_N(4) \wedge 1\big) + \big(\nu_N \wedge 1\big) \,\leqslant\, C \cdot \pi_N\,. \tag{8.2}$$

Corollary 8.2. *In the context of Theorem 4.1 and with the notation of* (4.2)–(4.5)*, the following are equivalent:*

(a) $\lim_{M\to\infty} \pi_M(\varepsilon) = 0$, *for all* $\varepsilon > 0$,

(b) $\lim_{M\to\infty} \tau_M(\varepsilon) = 0$ and $\lim_{M\to\infty} \sigma_M(\varepsilon) = 0$, *for all* $\varepsilon > 0$,

(c) $\lim_{M\to\infty} \tau_M(\varepsilon) = 0$ and $\lim_{M\to\infty} \nu_M(\varepsilon) = 0$, *for all* $\varepsilon > 0$.

The Corollary follows directly from Proposition 8.1, and completes the proof of Theorem 4.2. □

8.1 Proof of Proposition 8.1

The inequality (8.1) restates the bound (7.5) (cf. [23], p. 235); in fact, the (formally stronger) bound $\pi_{\mathsf{N}} \leqslant \mathsf{C} \cdot \big[\big(\tau_{\mathsf{N}} \wedge 1\big) + \big(\nu_{\mathsf{N}} \wedge 1\big) \big]$ also holds for $\mathsf{N} \in \mathbb{N}$ sufficiently large, by the double inequality (7.6). This latter shows also that the first inequality in (8.2) follows from the second.

It remains, therefore, to prove the second inequality in (8.2). This, in turn, follows from Lemmata 8.3, 8.4 below; these analyze the symmetric case first, and provide the desired bound

$$\big(\tau_{\mathsf{N}}(4) \wedge 1\big) + \big(\nu_{\mathsf{N}} \wedge 1\big) \leqslant 4 \Big[\big(\tau_{\mathsf{N}}^{\mathrm{sym}}(2) \wedge 1\big) + \big(\nu_{\mathsf{N}}^{\mathrm{sym}}(2) \wedge 1\big) \Big] \leqslant 4\mathsf{C}^{\mathrm{sym}}\, \pi_{\mathsf{N}}^{\mathrm{sym}}(2) \leqslant 8\mathsf{C}^{\mathrm{sym}}\, \pi_{\mathsf{N}}$$

for all $\mathsf{N} \in \mathbb{N}$ sufficiently large. This settles the general case, and completes the proof. □

Lemma 8.3. Bounds for the Symmetric Case: *Let $\mathsf{h}_1, \mathsf{h}_2, \cdots$ be I.I.D. with symmetric distribution. Recalling (4.2)–(4.5) we have, for all $\mathsf{N} \in \mathbb{N}$ sufficiently large and some universal constant $\mathsf{C}^{\mathrm{sym}} > 0$,*

$$\tau_{\mathsf{N}} \wedge 1 \leqslant 6 \cdot \pi_{\mathsf{N}}\,, \tag{8.3}$$

$$\big(\tau_{\mathsf{N}} \wedge 1\big) + \big(\sigma_{\mathsf{N}} \wedge 1\big) \leqslant \mathsf{C}^{\mathrm{sym}} \cdot \pi_{\mathsf{N}}\,. \tag{8.4}$$

Proof: With $\mathsf{A}_{\mathsf{N}} := \bigcup_{\mathsf{n}=1}^{\mathsf{N}} \big\{ \big|\mathsf{h}_{\mathsf{n}}\big| > \mathsf{N} \big\}$ we have $\mathbb{P}(\mathsf{A}_{\mathsf{N}}) = 1 - \big(1 - (\tau_{\mathsf{N}}/\mathsf{N})\big)^{\mathsf{N}} \geqslant 1 - \mathrm{e}^{-\tau_{\mathsf{N}}}$, thus also $3\,\mathbb{P}(\mathsf{A}_{\mathsf{N}}) \geqslant \tau_{\mathsf{N}} \wedge 1$; whereas, by the symmetry assumption and Lemma 2 on p. 149 of Feller [23], we obtain the claim of (8.3) as

$$\pi_{\mathsf{N}} = \mathbb{P}\Big(\Big| \sum_{\mathsf{n}=1}^{\mathsf{N}} \mathsf{h}_{\mathsf{n}} \Big| > \mathsf{N} \Big) \geqslant \frac{1}{2}\, \mathbb{P}(\mathsf{A}_{\mathsf{N}}) \geqslant \frac{1}{6}\, \big(\tau_{\mathsf{N}} \wedge 1\big)\,.$$

Next, we need to estimate σ_{N} in terms of π_{N}; and for this, we may assume $\mathbb{P}(\mathsf{A}_{\mathsf{N}}) \leqslant 1/2$, as otherwise (8.4) holds for the constant $\mathsf{C}^{\mathrm{sym}} = 12$. We introduce the probability measure

$$\mathbb{P}^{(\mathsf{N})} := \frac{1}{1 - \mathbb{P}(\mathsf{A}_{\mathsf{N}})} \cdot \mathbb{P}\,\Big|_{\Omega \setminus \mathsf{A}_{\mathsf{N}}}\,, \tag{8.5}$$

the normalized restriction of $\mathbb{P}$ to $\Omega \setminus \mathsf{A}_{\mathsf{N}} = \bigcap_{\mathsf{n}=1}^{\mathsf{N}} \big\{ \big|\mathsf{h}_{\mathsf{n}}\big| \leqslant \mathsf{N} \big\}$; and note that, under this measure, the functions $\mathsf{h}_{\mathsf{n}}^{(\mathsf{N})} := \mathsf{h}_{\mathsf{n}} \cdot \mathbf{1}_{\{|\mathsf{h}_{\mathsf{n}}| \leqslant \mathsf{N}\}}$, $\mathsf{n} = 1, \cdots, \mathsf{N}$ are independent, with common distribution supported on $\big[-\mathsf{N}, \mathsf{N} \big]$. Then

$$\pi_{\mathsf{N}}^{(\mathsf{N})} := \mathbb{P}^{(\mathsf{N})}\Big(\Big| \sum_{\mathsf{n}=1}^{\mathsf{N}} \mathsf{h}_{\mathsf{n}} \Big| > \mathsf{N} \Big) \geqslant \mathsf{c} \cdot \big(\sigma_{\mathsf{N}} \wedge 1\big) \tag{8.6}$$

holds for some real constant $\mathsf{c} > 0$, as shown in subsection 8.2 below. But from (8.6), (8.5) we deduce the inequality right below, which establishes (8.4) in conjunction with (8.3):

$$\pi_{\mathsf{N}} = \mathbb{P}\Big(\Big| \sum_{\mathsf{n}=1}^{\mathsf{N}} \mathsf{h}_{\mathsf{n}} \Big| > \mathsf{N} \Big) \geqslant \frac{\pi_{\mathsf{N}}^{(\mathsf{N})}}{2} \geqslant \frac{\mathsf{c}}{2}\, \big(\sigma_{\mathsf{N}} \wedge 1\big)\,. \qquad \square$$

We pass now to the general case: with $\mathsf{h}_1^{\pm}, \mathsf{h}_2^{\pm}, \cdots$ independent copies of $\mathsf{h}_1, \mathsf{h}_2, \cdots$, we consider the sequence of "symmetrized" versions $\mathsf{h}_{\mathsf{n}}^{\mathrm{sym}} := \mathsf{h}_{\mathsf{n}}^{+} - \mathsf{h}_{\mathsf{n}}^{-}$, $\mathsf{n} \in \mathbb{N}$, which are I.I.D. with symmetric distribution. We denote then by $\tau_{\mathsf{t}}^{\mathrm{sym}}(\varepsilon)$, $\sigma_{\mathsf{t}}^{\mathrm{sym}}(\varepsilon)$, $\pi_{\mathsf{t}}^{\mathrm{sym}}(\varepsilon), \cdots$ the so-symmetrized versions of the quantities in (4.3), (4.4), (4.2), $\cdots$ pertaining to $\mathsf{h}_1^{\mathrm{sym}}$.

Lemma 8.4. Bounds for the General Case: *Let* $\mathsf{h}_1, \mathsf{h}_2, \cdots$ *be I.I.D. With the notation of* (4.2)–(4.5) *we have, for all* $\mathsf{N} \in \mathbb{N}$ *sufficiently large,*

$$\nu_{\mathsf{N}} \leqslant 2 \cdot \nu_{2\mathsf{N}}^{\mathfrak{sym}} = 2 \cdot \sigma_{2\mathsf{N}}^{\mathfrak{sym}} \tag{8.7}$$

$$\tau_{\mathsf{N}} \leqslant 4 \cdot \tau_{\mathsf{N}}^{\mathfrak{sym}}(1/2) \tag{8.8}$$

$$\pi_{\mathsf{N}} \geqslant (1/2) \cdot \pi_{\mathsf{N}}^{\mathfrak{sym}}(2)\,. \tag{8.9}$$

Proof: With δ_{N} as in (4.1), integers $\mathsf{N} \geqslant \mathsf{N}_0$ such that $\mathbb{P}\big(|\mathsf{h}_1| > \mathsf{N}_0\big) < 1/2$ (we note that this N_0 depends only on the distribution of h_1), and setting $\mathsf{q}_{\mathsf{N}} := 1/\mathbb{P}\big(|\mathsf{h}_1| \leqslant \mathsf{N}\big) < 2$, we obtain (8.7) from

$$\begin{aligned}
\mathsf{N}\,\nu_{\mathsf{N}} &= \mathbb{E}\Big[\Big(\mathsf{h}_1^+ - \delta_{\mathsf{N}}\Big)^2 \mathbf{1}_{\{|\mathsf{h}_1^+| \leqslant \mathsf{N}\}}\Big] = \mathsf{q}_{\mathsf{N}} \cdot \mathbb{E}\Big[\Big(\mathsf{h}_1^+ - \delta_{\mathsf{N}}\Big)^2 \mathbf{1}_{\{|\mathsf{h}_1^+| \leqslant \mathsf{N},\, |\mathsf{h}_1^-| \leqslant \mathsf{N}\}}\Big] \\
&= \frac{\mathsf{q}_{\mathsf{N}}}{2} \cdot \mathbb{E}\Big[\Big(\big(\mathsf{h}_1^+ - \delta_{\mathsf{N}}\big) - \big(\mathsf{h}_1^- - \delta_{\mathsf{N}}\big)\Big)^2 \mathbf{1}_{\{|\mathsf{h}_1^+| \leqslant \mathsf{N},\, |\mathsf{h}_1^-| \leqslant \mathsf{N}\}}\Big] \\
&\leqslant \mathbb{E}\Big[\Big(\mathsf{h}_1^+ - \mathsf{h}_1^-\Big)^2 \mathbf{1}_{\{|\mathsf{h}_1^+ - \mathsf{h}_1^-| \leqslant 2\,\mathsf{N}\}}\Big] = \mathbb{E}\Big[\Big(\mathsf{h}_1^{\mathfrak{sym}}\Big)^2 \mathbf{1}_{\{|\mathsf{h}_1^{\mathfrak{sym}}| \leqslant 2\,\mathsf{N}\}}\Big] = 2\,\mathsf{N}\,\nu_{2\mathsf{N}}^{\mathfrak{sym}}\,.
\end{aligned}$$

On the other hand, the technique of symmetrization (cf. Lemma 1, p. 149 in FELLER [23]) gives also

$$\tau^{\mathfrak{sym}} = \mathsf{N} \cdot \mathbb{P}\big(|\mathsf{h}_1^+ - \mathsf{h}_1^-| > \mathsf{N}\big) > \frac{\mathsf{N}}{2} \cdot \mathbb{P}\big(|\mathsf{h}_1^+| > 2\,\mathsf{N}\big) = \frac{\tau_{\mathsf{N}}(2)}{4}\,, \qquad \forall\ \mathsf{N} \geqslant \mathsf{N}_0$$

with $\mathbb{P}\big(|\mathsf{h}_1| > \mathsf{N}_0\big) < 1/2$, leading to (8.8). Finally, with $\mathsf{S}_{\mathsf{N}}^{\pm} := \sum_{\mathfrak{n}=1}^{\mathsf{N}} \mathsf{h}_{\mathfrak{n}}^{\pm}$, $\mathsf{S}_{\mathsf{N}}^{\mathfrak{sym}} := \sum_{\mathfrak{n}=1}^{\mathsf{N}} \mathsf{h}_{\mathfrak{n}}^{\mathfrak{sym}}$, the inequalities $\big|(\mathsf{S}_{\mathsf{N}}^+/\mathsf{N}) - \mathsf{D}_{\mathsf{N}}\big| \leqslant 1$, $\big|(\mathsf{S}_{\mathsf{N}}^-/\mathsf{N}) - \mathsf{D}_{\mathsf{N}}\big| \leqslant 1$, lead to $\big|\mathsf{S}_{\mathsf{N}}^{\mathfrak{sym}}/\mathsf{N}\big| \leqslant 2$, and this in turn to the inequality $2 \cdot \mathbb{P}\Big(\big|(\mathsf{S}_{\mathsf{N}}^+/\mathsf{N}) - \mathsf{D}_{\mathsf{N}}\big| > 1\Big) \geqslant \mathbb{P}\big(\,|\mathsf{S}_{\mathsf{N}}^{\mathfrak{sym}}| > 2\,\mathsf{N}\big)$, i.e., (8.9). □

8.2 Proof of the Lower Bound (8.6)

The inequality (8.6) will follow from the following result.

Lemma 8.5. *Suppose* $\mathsf{h}_1, \cdots, \mathsf{h}_{\mathsf{N}}$ *are independent functions with common, symmetric distribution supported on the interval* $\big[-\mathsf{N}, \mathsf{N}\big]$, *and with* $\sigma_{\mathsf{N}} = \nu_{\mathsf{N}} = \mathbb{E}\big(\mathsf{h}_1^2\big)/\mathsf{N} > 0$ *in the notation of* (4.4), (4.5). *There is then a uniform constant* $\mathsf{c}^{\mathfrak{sym}} > 0$ *such that*

$$\pi_{\mathsf{N}} = \mathbb{P}\Big(\Big|\sum_{\mathfrak{n}=1}^{\mathsf{N}} \mathsf{h}_{\mathfrak{n}}\Big| > \mathsf{N}\Big) \geqslant \mathsf{c}^{\mathfrak{sym}} \cdot \sigma_{\mathsf{N}} = \mathsf{c}^{\mathfrak{sym}} \cdot \nu_{\mathsf{N}}\,. \tag{8.10}$$

Proof: We note that $\mathsf{S}_{\mathfrak{n}} := \mathsf{h}_1 + \cdots + \mathsf{h}_{\mathfrak{n}}$, $\mathfrak{n} = 1, \cdots, \mathsf{N}$ is a martingale (of its own filtration), and set $\tau_1 := \min\big\{\mathfrak{n} : |\mathsf{S}_{\mathfrak{n}}| \geqslant \mathsf{N}\big\}$ with the understanding $\min \emptyset = \infty$. With $\alpha := \mathsf{c}^{\mathfrak{sym}} \cdot \sigma_{\mathsf{N}}$ the right-hand side in (8.10) and $\beta := \mathbb{P}\big(\tau_1 < \infty\big)$, we have to show $\beta > 2\,\alpha$; because then symmetry gives (8.10), via $\pi_{\mathsf{N}} = \mathbb{P}\big(|\mathsf{S}_{\mathsf{N}}| \geqslant \mathsf{N}\big) \geqslant \mathbb{P}\big(|\mathsf{S}_{\tau_1 \wedge \mathsf{N}}| \geqslant \mathsf{N}\big)\,/\,2 = \beta\,/\,2 > \alpha$.

We define now inductively the stopping times $\tau_{\mathsf{k}+1} := \min\big\{\mathfrak{n} > \tau_{\mathsf{k}} : |\mathsf{S}_{\mathfrak{n}} - \mathsf{S}_{\tau_{\mathsf{k}}}| \geqslant \mathsf{N}\big\}$, $\mathsf{k} \in \mathbb{N}$ and note $\mathbb{P}\big(\tau_2 < \infty \,\big|\, \tau_1 < \infty\big) \leqslant \beta$; we rely on the homogeneity of the increments of $\big(\mathsf{S}_{\mathfrak{n}}\,,\, \mathfrak{n} = 1, \cdots, \mathsf{N}\big)$ and the fact that, after time τ_1, this martingale has less time to reach the barrier N than when starting at $\tau_0 = 0$. We obtain thus $\mathbb{P}\big(\tau_2 < \infty\big) \leqslant \beta^2$; as well as $\mathbb{P}\big(\tau_{\mathsf{k}} < \infty\big) \leqslant \beta^{\mathsf{k}}$, $\mathsf{k} \in \mathbb{N}_0$ by an obvious induction, and setting $\tau_0 \equiv 0$.

Now, $|\mathsf{S}_{\tau_1 \wedge \mathsf{N}}| = |\mathsf{S}_{\mathsf{N}}| \leqslant \mathsf{N}$ holds on $\big\{\tau_1 = \infty\big\}$; the assumption $\mathbb{P}\big(|\mathsf{h}_1| \leqslant \mathsf{N}\big) = 1$ implies $|\mathsf{S}_{\tau_{\mathsf{k}} \wedge \mathsf{N}}| = |\mathsf{S}_{\mathsf{N}}| \leqslant \big(2\,\mathsf{k} - 1\big)\,\mathsf{N}$ on $\big\{\tau_{\mathsf{k}-1} < \infty\,, \tau_{\mathsf{k}} = \infty\big\}$; and $|\mathsf{S}_{\mathsf{N}} - \mathsf{S}_{\tau_{\mathsf{k}-1}}| \leqslant \mathsf{N}$ holds on $\big\{\tau_{\mathsf{k}} = \infty\big\}$. Writing $\mathsf{S}_{\mathsf{N}} = \sum_{\mathsf{k} \in \mathbb{N}_0} \mathsf{S}_{\mathsf{N}} \cdot \mathbf{1}_{\{\tau_{\mathsf{k}-1} < \infty\,,\, \tau_{\mathsf{k}} = \infty\}}$, we obtain the inequality $\beta > 2\,\alpha$ from $\sigma_{\mathsf{N}} = \mathsf{N}^{-2}\,\mathbb{E}(\mathsf{S}_{\mathsf{N}}^2) \leqslant \big(1 - \beta\big) \sum_{\mathsf{k} \in \mathbb{N}_0} \beta^{\mathsf{k}-1} \big(2\,\mathsf{k} - 1\big)^2 < \beta\,/\,\big(2\,\mathsf{c}^{\mathfrak{sym}}\big)$ for a suitable constant $\mathsf{c}^{\mathfrak{sym}} > 0$, as long as β is bounded away from 1. □

9 The Proof of Proposition 4.3

The idea behind the proof, is the familiar theme of "gliding humps". In its simplest form, this is reflected in the following arch-example of functions $\mathsf{f}_1, \mathsf{f}_2, \cdots$ in $\mathbb{L}^0\big([0,1)\big)$ given by

$$\mathsf{f}_{2^m+j} := \mathbf{1}_{\mathcal{I}_{m,j}} \quad \text{where} \quad \mathcal{I}_{m,j} := \big[\,(j-1)\,2^{-m}, j\,2^{-m}\,\big)\,; \qquad m \in \mathbb{N}_0\,,\ \ j = 1, \cdots, 2^m\,.$$

These converge to zero in probability, but not a.e. By defining instead $\mathsf{f}^{\dagger}_{2^m+j} := m \cdot \mathbf{1}_{\mathcal{I}_{m,j}}$, we obtain a sequence converging to zero in probability, yet with $\overline{\lim}_{n\to\infty}\big|\mathsf{f}^{\dagger}_n(\omega)\big| = \infty$ for each $\omega \in [0,1)$.

9.1 Proof of Proposition 4.3 ($\mathfrak{A}$)

Denoting by $\mathsf{h}^\alpha, \mathsf{h}^\alpha_1, \mathsf{h}^\alpha_2, \cdots$ an I.I.D. sequence with $\mathbb{P}(\mathsf{h}^\alpha = \pm \mathsf{N}_\alpha) = \alpha/2$, $\mathbb{P}(\mathsf{h}^\alpha = 0) = 1-\alpha$, and $\mathsf{N}_\alpha = 1/\alpha$ a large natural number, $\mathbb{P}\Big(\Big|\sum_{n=1}^{\mathsf{N}_\alpha} \mathsf{h}^\alpha_n\Big| \geqslant \mathsf{N}_\alpha\Big) > \mathsf{C}$ holds for some universal constant $\mathsf{C} \in (0,1)$. We may also find, for given $\varepsilon > 0$, a constant $\rho = \rho(\varepsilon) \in (0,1)$ such that, for each $\alpha > 0$, we have

$$\mathbb{P}\Big(\Big|\sum_{n=1}^{N} \mathsf{h}^\alpha_n\Big| > \varepsilon \cdot N\Big) < \varepsilon\,, \qquad \text{for all} \quad N \notin \big[\,\rho\,\mathsf{N}_\alpha\,, (1/\rho)\,\mathsf{N}_\alpha\,\big]\,. \tag{9.1}$$

More generally, for a sequence $\big(\mathsf{N}_{\alpha_k}\big)_{k\in\mathbb{N}} \subset \mathbb{N}$ increasing to infinity so fast that $\sum_{k\in\mathbb{N}}(1/\mathsf{N}_{\alpha_k}) < 1$, and for the sequence $\boldsymbol{\alpha} := \big(\alpha_k\big)_{k\in\mathbb{N}}$ with $\alpha_k := 1/\mathsf{N}_{\alpha_k}$, consider I.I.D. random variables

$$\mathsf{h}^{(\boldsymbol{\alpha})}, \mathsf{h}^{(\boldsymbol{\alpha})}_1, \mathsf{h}^{(\boldsymbol{\alpha})}_2, \cdots \text{ with } \ \mathbb{P}\big(\mathsf{h}^{(\boldsymbol{\alpha})} = \pm \mathsf{N}_{\alpha_k}\big) = \alpha_k/2 \ \text{ for } k \in \mathbb{N}\,, \quad \mathbb{P}\big(\mathsf{h}^{(\boldsymbol{\alpha})} = 0\big) = 1 - \sum_{k\in\mathbb{N}} \alpha_k\,. \tag{9.2}$$

We may choose a sequence $\big(\mathsf{N}_{\alpha_k},\, k \in \mathbb{N}\big)$, increasing sufficiently fast to infinity, so that the intervals

$$\mathcal{L}_k := \Big\{N \in \mathbb{N} :\ \mathbb{P}\Big(\Big|\sum_{n=1}^{N} \mathsf{h}^{\alpha_k}_n\Big| > N\,2^{-k}\Big) < 2^{-k}\Big\} \tag{9.3}$$

are disjoint. Now, for each $x \in [0,1)$, we let $j^{(m)}(x)$ be the number in $\{1, \cdots, 2^m\}$ with the property $\big(j^{(m)}(x) - 1\big)\,2^{-m} \leqslant x < j^{(m)}(x)\,2^{-m}$. Setting $k_m(x) := 2^m + j^{(m)}(x)$, we consider the sequence $\big(\alpha_{k_m(x)}\big)_{m\in\mathbb{N}}$, and let $\mathsf{h}^x, \mathsf{h}^x_1, \mathsf{h}^x_2, \cdots$ be the I.I.D. sequence attached to this subsequence as in (9.2).

Fixing $x \in [0,1]$ we thus have that the quantities $\mathsf{N}_{\alpha_{k_m(x)}}\,\mathbb{P}\big(|\mathsf{h}^x_1| \geqslant \mathsf{N}_{\alpha_{k_m(x)}}\big)$ as well as the quantities $\mathsf{N}^{-1}_{\alpha_{k_m(x)}}\,\mathbb{E}\Big(\big(\mathsf{h}^x_1\big)^2 \mathbf{1}_{\{|\mathsf{h}^x_1| \leqslant \mathsf{N}_{\alpha_{k_m(x)}}\}}\Big)$ remain bounded away from zero as $m \to \infty$. On the other hand, when we replace in these two expressions $\mathsf{N}_{\alpha_{k_m(x)}}$ by $N \in \mathbb{N}$, they tend to zero as $N \to \infty$ except for the case where N is in one of the intervals $\mathcal{L}_{k_m(x)}$ defined in and below (9.3). $\square$

9.2 Proof of Proposition 4.3 ($\mathfrak{B}$)

As in the proof of part ($\mathfrak{A}$), we construct inductively an increasing sequence $\big(\mathsf{N}_k\big) = \big(\mathsf{N}_{2^m+j}\big) \subset \mathbb{N}$. Fixing $m \in \mathbb{N}$ and $k = 2^m + j$ we define also random variables h^k with values in $\{\pm \mathsf{N}_k, 0\}$, with $\mathbb{P}\big(\mathsf{h}^k = \pm \mathsf{N}_k\big) = \mathsf{p}_m/2$ and m sufficiently small, to wit, $\mathsf{p}_m < 2^{-m}$.

The inductive step $m-1 \longmapsto m$ of this construction is described in the following Lemma, where we assume $\mathsf{N}^{(m-1)}_{2^{m-1}+1}$ has already been defined.

Lemma 9.1. *For $m \in \mathbb{N}$, there exist integers $N_{2^{m-1}+1}^{(m-1)} < N_1^{(m)} < \cdots < N_{2^m}^{(m)} < N_{2^m+1}^{(m)}$ and numbers $p_m \in (0, 2^{-m})$, such that, for h_j^m taking values in $\{\pm N_j^{(m)}, 0\}$ with $\mathbb{P}\big(h_j^m = \pm N_j^{(m)}\big) = p_m/2$, and setting*

$$\tau_N^{(m,j)} := N \cdot \mathbb{P}\left(\left|h_j^m\right| \geqslant N\right), \qquad \sigma_N^{(m,j)} := \mathbb{E}\left(\left(h_j^m\right)^2 \mathbf{1}_{\{|h_j^m| \leqslant N\}}\right) \Big/ N, \tag{9.4}$$

we have

(i) $\tau_{N_1^{(m)}}^{(m,j)} = m$ *for* $j = 1, \cdots, 2^m$,

(ii) $\sigma_N^{(m,j)} < 2^{-m}$ *for* $j = 1, \cdots, 2^m$, $N \notin \left[\, N_j^{(m)},\ N_{j+1}^{(m)} \,\right[$.

Proof: We choose $N_1^{(m)} > N_{2^{m-1}+1}^{(m-1)}$ big enough, so that $p_m = m \,/\, N_1^{(m)} < 2^{-m}$; then (i) follows for $j = 1$. We note also that $\sigma_N^{(m,1)} = \left(N_1^{(m)}\right)^2 p_m \,/\, N$ tends to zero as $N \to \infty$, so we can find $N_2^{(m)}$ with $\sigma_{N_2^{(m)}}^{(m,1)}$ smaller than 2^{-m}. This last property holds also for all $N \geqslant N_2^{(m)}$, while we have $\sigma_N^{(m,1)} = 0$ for $N < N_1^{(m)}$; thus, the claim (i) is established for $j = 1$.

We define h_2^m to take values in $\left\{ \pm N_2^{(m)}, 0 \right\}$ with $\mathbb{P}\big(h_2^m = \pm N_2\big) = p_m/2$, and observe that (i) also holds for $j = 2$, i.e., $\tau_{N_1^{(m)}}^{(m,2)} = m$. Next, we choose $N_3^{(m)}$ big enough, so that (ii) holds for $j = 2$.

Continuing in an obvious manner, we select $N_1^{(m)} < \cdots < N_{2^m}^{(m)}$; as well as $N_{2^m+1}^{(m)}$, whose only role is to make sure that in the next inductive step $m \longmapsto m+1$ of this construction, the number $N_1^{(m+1)}$ is big enough so $\sigma_N^{(m,2^m)} < 2^{-m}$ holds for $N \geqslant N_1^{(m+1)}$.

This completes the inductive step, and establishes the Lemma. □

As in subsection 9.1, we arrange now these $\left(\left(N_j^{(m)}\right)_{j=1}^{2^m}\right)_{m\in\mathbb{N}_0}$ as $\left(N_k\right)_{k=2}^{\infty} = \left(\left(N_{j+2^m}\right)_{j=1}^{2^m}\right)_{m\in\mathbb{N}_0}$; associate to each $x \in [0,1)$ the integer $J^{(m)}(x)$ as in subsection 9.1 above; and let $\left(h_n^x\right)_{n\in\mathbb{N}}$ be I.I.D. taking the values 0, $\pm N_{J^{(1)}(x)}^{(1)}$, $\pm N_{J^{(2)}(x)}^{(2)}, \cdots$ on disjoint sets, with $\mathbb{P}\big(h^x = \pm N_{J^{(m)}(x)}^{(m)}\big) = p_m/2$, $m \in \mathbb{N}$. This latter is possible, because $p_m < 2^{-m}$. Once again, we consider exchangeable $g_1^X, g_2^X, \cdots$ with the property that, for X uniformly distributed on [0,1), the conditional distribution of $g_1^X, g_2^X, \cdots$, given $\{X = x\}$, is that of the sequence $h_1^x, h_2^x, \cdots$.

For each $m \in \mathbb{N}$ we consider the first element $\widehat{N}_M := N_{2^m+1} = N_1^{(m)}$ of the above inductive construction. Lemma 9.1 (i) shows that $\widehat{N}_M \cdot \mathbb{P}\big(\left|g_m^X\right| \geqslant \widehat{N}_M \,\big|\, X = x\big) \geqslant \widehat{N}_M \cdot \mathbb{P}\big(\left|h_m^x\right| \geqslant \widehat{N}_M\big) = m$ holds for each $x \in [0,1)$. On the other hand, for every integer N between $N_1^{(m)}$ and $N_1^{(m+1)}$, there is at most one $\widehat{j} \in \left\{1, \cdots, 2^m\right\}$ so that, for $x \in \left[\left(\widehat{j}-1\right)2^{-m}, \widehat{j}\,2^{-m}\right)$, $\sigma_N^{(m,\widehat{j})}$ of (9.4) is big; while, for $j \neq \widehat{j}$, we have $\sigma_N^{(m,j)} < 2^{-m}$. This leads to (ii), and completes the proof of part ($\mathfrak{B}$). □

9.3 Proof of Proposition 4.3 ($\mathfrak{C}$)

We carry out a construction similar to that in the proof of part ($\mathfrak{B}$), defining inductively an increasing sequence $\left(N_k\right) = \left(N_{2^m+j}\right)$ of natural numbers. Fixing again $m \in \mathbb{N}$ we shall construct this time $\left(N_k\right)_{k=2}^{\infty} := \left(\left(N_{2^m+j}\right)_{j=1}^{2^m}\right)_{m\in\mathbb{N}} \equiv \left(\left(N_j^{(m)}\right)_{j=1}^{2^m}\right)_{m\in\mathbb{N}}$ via *backwards induction:*

We choose first a (very large) number $N_{2^m}^{(m)}$, then inductively a decreasing string of natural numbers $N_{2^m}^{(m)} > N_{2^m-1}^{(m)} > \cdots > N_1^{(m)}$, all the while making sure that $N_1^{(m)}$ is bigger than a threshold $N_{2^{m-1}}^{(m-1)}$ given from the previous inductive step $(m-1)$.

Lemma 9.2. *Consider natural numbers* m *and* $N_{2^{m-1}}^{(m-1)}$, *as well as a number* $\delta > 0$ *and a rational* $\varepsilon > 0$. *There exist then natural numbers* $N_{2^{m-1}}^{(m-1)} < N_1^{(m)} < \cdots < N_{2^m}^{(m)}$, *and positive numbers* $p_1^{(m)}, \cdots, p_{2^m}^{(m)}$ *with* $\sum_{j=1}^{2^m} p_j^{(m)} < \delta$, *such that*

$$p_j^{(m)} \left(N_j^{(m)}\right)^2 = m \cdot N_{2^m}^{(m)}, \qquad j = 1, \cdots, 2^m, \tag{9.5}$$

$$p_j^{(m)}\, N_{j-1}^{(m)} = \varepsilon, \qquad j = 2, \cdots, 2^m. \tag{9.6}$$

Proof: The above inductive relationships (9.5), (9.6) hold if, and only if, the pairs $\left(N_j^{(m)}, p_j^{(m)}\right)$, $j = 1, \cdots 2^m$ satisfy the following explicit relations:

$$N_j^{(m)} = N_{2^m}^{(m)} \Big(\frac{\varepsilon}{m}\Big)^{2^{2^m - j} - 1}, \quad p_j^{(m)} = \frac{m}{N_{2^m}^{(m)}} \Big(\frac{m}{\varepsilon}\Big)^{2^{2^m}}, \qquad j = 1, \cdots, 2^m.$$

As $\varepsilon > 0$ is assumed rational, we can choose a sufficiently large multiple $N_{2^m}^{(m)}$ of $\left(m/\varepsilon\right)^{2^{2^m}}$, so that $N_1^{(m)}$ is a natural number bigger than the given $N_{2^{m-1}}^{(m-1)}$ and such that $\sum_{j=1}^{2^m} p_j^{(m)} < \delta$ holds. □

As in the proof of part $(\mathfrak{B})$, the Lemma that follows describes the inductive step $m - 1 \longmapsto m$, where we assume that $N_{2^{m-1}}^{(m-1)}$ has already been defined.

Lemma 9.3. *For* $m \in \mathbb{N}$ *and* $(\varepsilon, \delta) \in (0, \infty)^2$, *there exist integers* $N_{2^{m-1}}^{(m-1)} < N_1^{(m)} < \cdots < N_{2^m}^{(m)}$, *as well as positive numbers* $p_1^{(m)}, \cdots, p_{2^m}^{(m)}$ *with* $\sum_{j=1}^{2^m} p_j^{(m)} < \delta$, *and random variables* h_j^m *with values in* $\left\{ \pm N_j^{(m)}, 0 \right\}$ *and* $\mathbb{P}\big(h_j^m = \pm N_j^{(m)}\big) = p_j^{(m)}/2$, *so that*

$$\tau_N^{(m,j)} := N \cdot \mathbb{P}\left(\left|h_j^m\right| > N\right), \qquad \sigma_N^{(m,j)} := \mathbb{E}\left(\left(h_j^m\right)^2 \mathbf{1}_{\{|h_j^m| \leqslant N\}}\right) \Big/ N \tag{9.7}$$

satisfy

(i) $\tau_N^{(m,j)} = 0$ *for* $j = 1, \cdots, 2^m$, $N \geqslant N_j^{(m)}$,

(i)′ $\tau_N^{(m,j)} \leqslant \varepsilon^\ell \cdot \tau_N^{(m,j+\ell)}$ *for* $j = 1, \cdots, 2^m$, $\ell = 1, \cdots, 2^m - j$, $N \leqslant N_j^{(m)}$,

(ii) $\sigma_N^{(m,j)} = \left(N_{2^m}^{(m)}\right)^{-1} p_j^{(m)} \left(N_j^{(m)}\right)^2 = m$ *for* $j = 1, \cdots, 2^m$, $N = N_{2^m}^{(m)}$.

Proof: Part (i) is obvious; (ii) follows from (9.6); and (iii) from applying (9.5) inductively. □

Again as in subsection 9.1, we arrange the $\left(\left(N_j^{(m)}\right)_{j=1}^{2^m}\right)_{m \in \mathbb{N}_0}$ as $\left(N_k\right)_{k=2}^{\infty} = \left(\left(N_{j+2^m}\right)_{j=1}^{2^m}\right)_{m \in \mathbb{N}_0}$; associate to each $x \in [0,1)$ an integer $J^{(m)}(x)$; and let $\left(h_n^x\right)_{n \in \mathbb{N}}$ be I.I.D. taking the values $0, \pm N_{J^{(1)}(x)}^{(1)}, \pm N_{J^{(2)}(x)}^{(2)}, \cdots$ on disjoint sets, with $\mathbb{P}\big(h^x = \pm N_{J^{(m)}(x)}^{(m)}\big) = p_m/2$, $m \in \mathbb{N}$. We summon also exchangeable $g_1^X, g_2^X, \cdots$ with the property that, for X uniformly distributed on [0,1), the conditional distribution of $g_1^X, g_2^X, \cdots$, given $\{X = x\}$, is that of the sequence $h_1^x, h_2^x, \cdots$.

For each $m \in \mathbb{N}$ we consider now the last element $M(m) := N_{2^m}^{(m)}$ of the above string $N_1^{(m)} < \cdots < N_{2^m}^{(m)}$. It follows from Lemma 9.3 (ii) that property (c) of Proposition 4.3 $(\mathfrak{C})$ holds as $\mathbb{P}\big(\Sigma_{M(m)} \geqslant m\big) = 1$; this implies also property (a) there, i.e., the failure of the WLLN.

We still have to show property (b), i.e., $\lim_{N \to \infty} T_N = 0$, in $\mathbb{P}-$probability. With $m \in \mathbb{N}$ and $N \in \left(N_{2^{m-1}}^{(m-1)}, N_{2^m}^{(m)}\right]$ there is at most one interval of the form $\left(N_j^{(m)}, N_{j+1}^{(m)}\right]$ with $j = 1, \cdots, 2^m - 1$ such that $N \in \left(N_j^{(m)}, N_{j+1}^{(m)}\right]$. For such N and $\ell \neq j$, we have then $\tau_N\left(g_1^X\right) < (1/m) = \varepsilon$ on the event $\left\{(j-1)\, 2^{-m} < X \leqslant j\, 2^{-m}\right\}$. This proves $\lim_{N \to \infty} T_N = 0$, in $\mathbb{P}-$probability. □

10 The Proof of Theorem 3.1

The equivalence of the conditions **(i)**, **(i)′** is shown exactly as in the first paragraph of section 8, for the proof of Theorem 4.2. We need to prove the equivalence of conditions **(i)** and **(ii)**.

10.1 The Implication (i) ⇒ (ii)

Suppose that there exists a sequence of "correctors" $D_1, D_2, \cdots$, measurable with respect to the tail $\sigma-$algebra $\mathcal{T}$ of (2.1), such that the Weak Law of Large Numbers (3.1) holds for some subsequence $f_{k_1}, f_{k_2}, \cdots$ and all its subsequences. Then, from Lemma 5.1, this $f_{k_1}, f_{k_2}, \cdots$ is bounded in $\mathbb{L}^0$; so we may assume it is also determining, with associated exchangeable approximating sequence $g_1, g_2, \cdots$ as in (5.5). On the strength of Proposition 6.1(i), the exchangeable $g_1, g_2, \cdots$ satisfy this Weak Law of Large Numbers as well; in other words, (6.1) holds.

Hence, we may apply Theorem 4.2 to obtain all the properties (4.25)–(4.30) listed there, for the "statistics" (4.19)–(4.21) of the (random) conditional distribution $\boldsymbol{\mu}$ of g_1, given the tail $\sigma-$algebra $\mathcal{T}_*$ of (4.15). In particular, the conditions of (4.25) with $\varepsilon = 1$ lead now to those in (3.3).

10.2 The Implication (ii) ⇒ (i)

We assume that the determining subsequence $f_{k_1}, f_{k_2}, \cdots$, with limit random probability distribution $\boldsymbol{\mu}$ as in and below (2.2), satisfies the conditions in (3.3); in particular, boundedness in $\mathbb{L}^0$. We select a sequence of natural numbers J_N, $N \in \mathbb{N}$ increasing to infinity so slowly, that

$$\lim_{N\to\infty} \left(\frac{1}{N} \sum_{n=J_N}^{N} f_{k_n} - D_N \right) = 0 \quad \text{in } \mathbb{P}-\text{probability} \tag{10.1}$$

implies the desired (3.1); and set out to prove (10.1).

We summon now Theorem 2 of [9] to obtain, for each $N \in \mathbb{N}$ and after passing to an appropriate subsequence, a partition $\big(A_i^{(N)}\big)_{i=0}^{r_N}$ of the space Ω with $\mathbb{P}\big(A_0^{(N)}\big) \leqslant 2^{-N}$ and such that, for each remaining set $A = A_i^{(N)}$, $i = 1, \cdots, r_N$, there exists a sequence $h_j^{(A)}$, $j \in \mathbb{N}$ of I.I.D. random variables with common distribution μ^A and the property

$$\mathbb{P}^A\Big(\big| f_j - h_j^{(A)} \big| > 2^{-N} \Big) \leqslant 2^{-N}, \qquad j = J_N, J_{N+1}, \cdots . \tag{10.2}$$

Here, we have renamed as $\big(f_j,\ j \in \mathbb{N}\big)$ the determining subsequence $f_{k_1}, f_{k_2}, \cdots$; and denoted by $\mathbb{P}^A$ the conditional probability measure $\mathbb{P}^A(\cdot) = \mathbb{P}(\cdot \cap A)\,/\,\mathbb{P}(A)$, by μ^A the limiting distribution of the $f_1, f_2, \cdots$ on the set A. The inequality (7.5) (also on page 235 of Feller, Vol. II [23]) gives now

$$\mathbb{P}^A\left(\Big| \frac{1}{N} \sum_{j=J_N}^{N} h_j^{(A)} - \delta_N^{(A)} \Big| > \varepsilon \right) \leqslant N \cdot \mu^A\big(\mathbb{R} \setminus [-N, N]\big) + \frac{1}{N\varepsilon^2} \int_{[-N,N]} x^2\, \mu^A(dx) \tag{10.3}$$

for the correctors

$$\delta_N^{(A)} := \mathbb{E}^{\mathbb{P}^A}\Big(h_j^{(A)} \cdot \mathbf{1}_{\{|h_j^{(A)}| \leqslant N\}} \Big), \quad N \in \mathbb{N}. \tag{10.4}$$

Whereas, on account of (10.3), (10.2), we deduce for each atom $A = A_i^{(N)}$, $i = 1, \cdots, r_N$ the bound

$$\mathbb{P}^A\left(\Big| \frac{1}{N} \sum_{j=J_N}^{N} f_j - \delta_N^{(A)} \Big| > \varepsilon + 2^{-N} \right) \leqslant N \cdot \mu^A\big(\mathbb{R} \setminus [-N, N]\big) + \frac{1}{N\varepsilon^2} \int_{[-N,N]} x^2\, d\mu^A(x) + 2^{-N}. \tag{10.5}$$

We denote now by $\mathcal{G}_N$ the σ–algebra generated by the sets $A_i^{(N)}$, $i = 1, \cdots, r_N$, and observe that the construction in [9] can be made to guarantee the filtration structure $\mathcal{G}_1 \subseteq \mathcal{G}_2 \subseteq \cdots$; in this manner, (10.5) becomes

$$\mathbb{P}\Big(\Big| \frac{1}{N} \sum_{j=J_N}^{N} f_j - D_N \Big| > \varepsilon + 2^{-N} \,\Big|\, \mathcal{G}_N \Big) \tag{10.6}$$

$$\leqslant \mathbb{E}\Big(N\,\mu\big(\mathbb{R} \setminus [-N, N], \cdot\big) \,\Big|\, \mathcal{G}_N \Big) + \mathbb{E}\Big(\frac{1}{N\varepsilon^2} \int_{[-N,N]} x^2\, \mu(dx, \cdot) \,\Big|\, \mathcal{G}_N \Big) + 2^{-N}$$

with "randomized correctors"

$$D_N := \sum_{i=1}^{r_N} \delta_N^{(A_i^{(N)})} \cdot \mathbf{1}_{A_i^{(N)}} \tag{10.7}$$

in the notation of (10.4). Whereas, passing in (10.6) to a suitable subsequence of $f_1, f_2, \cdots$ and relabelling, this inequality remains valid when conditioning with respect to $\mathcal{G}_N$ in (10.6) is replaced by conditioning with respect to $\mathcal{G}_{s_N}$, where $(s_N)_{N\in\mathbb{N}}$ grows as rapidly as desired.

Finally, we argue that

$$\mathbb{P}\Big(\Big| \frac{1}{N} \sum_{j=J_N}^{N} f_j - D_N \Big| > \varepsilon + 2^{-N} \,\Big|\, \mathcal{G}_{s_N} \Big) \leqslant N\,\mu\big(\mathbb{R} \setminus [-N, N], \cdot\big) + \frac{1}{N\varepsilon^2} \int_{[-N,N]} x^2\, \mu(dx, \cdot) + R_N$$

holds, where $\lim_{N\to\infty} R_N = 0$ in probability. Indeed, recalling the assumption (ii) and (10.4), (10.7), invoking the martingale convergence theorem, and always assuming that $(s_N)_{N\in\mathbb{N}}$ grows sufficiently rapidly, we deduce that the left-hand side of the above display, as well as the difference $D_N - \int_{[-N,N]} x\,\mu(dx, \cdot)$, converge in probability to zero as $N \to \infty$.

Integration leads now to the desired conclusion (10.1). □

References

[1] Aistleitner, Ch., Berkes, I. and Tichy, R. (2024) Lacunary Sequences in Analysis, Probability and Number Theory. In "Diophantine Problems: Determinism, Randomness, and Applications". *Panoramas et Synthèses* **62**, 1-60. Societé Mathématique de France, Paris.

[2] Aldous, D. (1977) Limit theorems for subsequences of arbitrarily-dependent sequences of random variables. *Zeitschrift für Wahrscheinlichkeitstheorie & Verwandte Gebiete* **40**, 59-82.

[3] Aldous, D. (1977) Subsequences of sequences of random variables. *Bulletin of the American Mathematical Society* **83**, 121-123.

[4] Aldous, D. and Eagleson, G.K. (1978) On mixing and stability of limit theorems. *Annals of Probability* **6**, 325-331.

[5] Baxter, J.R. and Chacon, R.V. (1977) Compactness of stopping times. *Zeitschrift für Wahrscheinlichkeitstheorie & Verwandte Gebiete* **40**, 169-182.

[6] Beiglböck, M., Schachermayer, W. and Veliyev, B. (2011) A short proof of the Doob-Meyer theorem. *Stochastic Processes & Their Applications* **122**, 1204-1209.

[7] Beiglböck, M., Schachermayer, W. and Veliyev, B. (2011) A direct proof of the Bichteler-Dellacherie theorem. *Annals of Probability* **39**, 2424-2440.

[8] Berkes, I., Karatzas, I. and Schachermayer, W. (2025) A hereditary Hsu-Robbins-Erdős law of large numbers. *Available on arXiv:2503.19484 . Submitted for publication.*

[9] Berkes, I. and Péter, E. (1986) Exchangeable random variables and the subsequence principle. *Probability Theory & Related Fields* **73**, 395-413.

[10] Berkes, I. and Rosenthal, H.P. (1985) Almost exchangeable sequences of random variables. *Probability Theory & Related Fields* **70**, 473-507.

[11] Billingsley, P. (1968) *Convergence of Probability Measures.* J. Wiley & Sons, New York.

[12] Brooks, J.K. & Chacon, R.V. (1980) Continuity and compactness of measures. *Advances in Mathematics* **37**, 16 -26.

[13] Chatterji, S.D. (1970) A general strong law. *Inventiones Math.* **9**, 235-245.

[14] Chatterji, S.D. (1974) A subsequence principle in probability theory: the law of the iterated logarithm. *Inventiones Math.* **25**, 241-251.

[15] Chow, Y.S. and Teicher, H. (1978) *Probability Theory: Independence, Interchangeability, Martingales.* Springer Verlag, New York

[16] Chung, K.L. (1974) *A Course in Probability Theory.* Second Edition. Probability and Mathematical Statistics: A Series of Monographs and Textbooks, Vol. 21. Academic Press, New York.

[17] Dacunha-Castelle, D. (1975) Indiscernibility and exchangeability in $\mathbb{L}^{\mathrm{p}}-$spaces. *Proceedings of the Seminar on Random Series, Convex Sets, and Geometry of Banach Spaces,* Aarhus Universitet, Matematisk Institut Various Publ. **25**, 50-56.

[18] Delbaen, F. and Schachermayer, W. (1994) A general version of the fundamental theorem of asset pricing. *Mathematische Annalen* **300**, 463-520.

[19] Delbaen, F. and Schachermayer, W. (1999) A compactness principle for bounded sequences of martingales, with applications. In "Proceedings of the Seminar on Stochastic Analysis, Random Fields and Applications" (R. Dalang, M. Dozzi, F. Russo, editors). *Progress in Probability* **45**, 137-173. Birkhäuser, Basel and Boston.

[20] Dunford, N. and Pettis, B.J. (1940) Linear operations on summable functions. *Transactions of the American Mathematical Society* **47**, 323-392.

[21] Dunford, N. and Schwartz, J.T. (1988) *Linear Operators. Part I: General Theory.* Reprint of the 1958 Original Edition in "Wiley Classics Library". John Wiley & Sons, New York.

[22] Durrett, R. (2010) *Probability: Theory and Examples.* Cambridge University Press.

[23] Feller, W. (1971) *An Introduction to Probability Theory and Its Applications: Volume II.* Second Edition. John Wiley & Sons, New York.

[24] Gaposhkin, V.F. (1972) Convergence and limit theorems for sequences of random variables. *Theory of Probability & Its Applications* **17**, 379-400.

[25] Jakubowski, A. (2005) An almost sure approximation for the predictable process in the Doob-Meyer decomposition theorem. In "Séminaire de Probabilités XXXVIII". *Lecture Notes in Mathematics* **1857**, 158-164. Springer-Verlag, New York.

[26] Karatzas, I. and Kardaras, C. (2021) *Portfolio Theory & Arbitrage.* Graduate Studies in Mathematics, Volume **214**. American Mathematical Society, Providence, RI.

[27] Karatzas, I. and Schachermayer, W. (2023) A weak law of large numbers for dependent random variables. *Theory of Probability & Its Applications* **48**, 501-509.

[28] Kolmogorov, A.N. (1928/29) Über die Summen durch den Zufall bestimmter unabhängiger Grössen. *Mathematische Annalen* **99**, 309-319; **102**, 484-488.

[29] Kolmogorov, A.N. (1930) Sur la loi forte des grandes nombres. *Comptes Rendus de l' Académie des Sciences, Paris* **191**, 910-912.

[30] Kolmogorov, A.N. (1933) Grundbegriffe der Wahrscheinlichkeitsrechnung. *Ergebnisse der Mathematik* **2** (3), Springer-Verlag, Berlin. English Translation published as "Foundations of Probability Theory" by Chelsea Publishing Co., New York, 1950.

[31] Komlós, J. (1967) A generalization of a problem of Steinhaus. *Acta Mathematica Academiae Scientiarum Hungaricae* **18**, 217-229.

[32] Larsen, K. and Žitković, G. (2007) Stability of utility maximization in incomplete markets. *Stochastic Processes & Their Applications* **117**, 1642-1662.

[33] Loève, M. (1977) *Probability Theory, Volume I.* Graduate Texts in Mathematics, Volume **45**. Springer Verlag, New York.

[34] Meyer, P.A. (1966) *Probability & Potentials.* Blaisdell Publishing Company, Waltham, Mass.

[35] Rényi, A. (1965) On stable sequences of events. *Sankhyã, Ser. A* **16**, 310-318.

[36] Stoica, G. (2010) An extension of the weak law of large numbers for exchangeable sequences. *Acta Applicandae Mathematicae* **109**, 759-763.

[37] von Weizsäcker, H. (2004) Can one drop $\mathbb{L}^1$—boundedness in Komlós's subsequence theorem? *American Mathematical Monthly* **111**, 900-903.

[38] Žitković, G. (2010) Convex compactness and its applications. *Mathematical & Financial Economics* **3**, 1-12.