The functional law of the iterated logarithm for the empirical process based on sample means

Einmahl, J.H.J.; Rosalsky, A.

Published in:
Journal of Theoretical Probability

Publication date:
2001

Journal of Theoretical Probability, Vol. 14, No. 2, 2001

The Functional Law of the Iterated Logarithm for the Empirical Process Based on Sample Means

John H. J. Einmahl[1] and Andrew Rosalsky[2]

Received October 27, 1999; revised September 29, 2000

Consider a double array $\{X_{i,j}; i \geqslant 1, j \geqslant 1\}$ of i.i.d. random variables with mean μ and variance σ^2 $(0 < \sigma^2 < \infty)$ and set $Z_{i,n} = n^{-1/2} \sum_{j=1}^{n} (X_{i,j} - \mu)/\sigma$. Let $\hat{\Phi}_{N,n}$ denote the empirical distribution function of $Z_{1,n},..., Z_{N,n}$ and let Φ be the standard normal distribution function. The main result establishes a functional law of the iterated logarithm for $\sqrt{N}(\hat{\Phi}_{N,n} - \Phi)$, where $n = n(N) \to \infty$ as $N \to \infty$. For the proof, some lemmas are derived which may be of independent interest. Some corollaries of the main result are also presented.

KEY WORDS: Empirical process based on sample means; functional law of the iterated logarithm; double array; relative compactness; central limit theorem; Berry–Esseen inequality.

1. INTRODUCTION

Consider a sequence of independent, identically distributed (i.i.d.) random variables $\{X_n; n \geqslant 1\}$ with $EX_1 = \mu$ and $0 < \text{Var } X_1 = \sigma^2 < \infty$. The sequence of *sample means* is defined as usual by $\bar{X}_n = n^{-1} \sum_{j=1}^{n} X_j$, $n \geqslant 1$. Let

$$\Phi_n(x) = P\left(\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \leqslant x\right), \qquad x \in \mathbb{R}, \quad n \geqslant 1$$

Now Paul Lévy's version of the central limit theorem (CLT) asserts that

$$\lim_{n \to \infty} \Phi_n(x) = \Phi(x), \qquad x \in \mathbb{R} \tag{1.1}$$

[1] EURANDOM and Department of Mathematics and Computing Science, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands. E-mail: j.h.j.einmahl@tue.nl

[2] Department of Statistics, University of Florida, Box 118545, Gainesville, Florida 32611-8545, U.S.A. E-mail: rosalsky@stat.ufl.edu

0894-9840/01/0400-0577$19.50/0

where Φ is the distribution function of the standard normal distribution:

$$\Phi(x)=\int_{-\infty}^{x} \phi(t)\,dt \qquad \text{where} \quad \phi(x)=\frac{1}{\sqrt{2\pi}}\,e^{-x^2/2}, \qquad x\in\mathbb{R}$$

Since Φ is continuous, the convergence in (1.1) is uniform in x by Polya's theorem.

The presentation of the CLT provided in elementary textbooks often contains illustrations purporting to demonstrate that empirical histograms of sample means approach a normal density. This was first studied rigorously in Freedman.[6] For integer valued i.i.d. random variables with period 1 and a finite third moment, Freedman[6] showed that the empirical histogram of sample means (suitably centered and scaled) converges uniformly *in probability* to ϕ if $N/(\sqrt{n}\log n)\to\infty$ as $n\to\infty$ where N is the number of independent samples each of size n.

The CLT is, however, a statement about the limiting behavior of distribution functions and not about density functions. It is thus more natural to compare the empirical distribution function based on sample means (from independent samples) with Φ than it is to compare an empirical histogram based on sample means with ϕ. This observation is the main motivation for the present study. In order to be more explicit let us specify our setup and introduce more notation. Let $\{X_{i,j};\ i\geqslant 1,\ j\geqslant 1\}$ be a double array of i.i.d. nondegenerate random variables with $EX_{1,1}^2<\infty$. Using statistical terminology, for each $N\geqslant 1$ and $n\geqslant 1$, the family $\{\{X_{i,j};\ 1\leqslant j\leqslant n\};\ 1\leqslant i\leqslant N\}$ can be interpreted as consisting of N independent *random samples*, each of *size* n, from a population with some mean μ and variance σ^2. Without loss of generality we suppose in the sequel that $\mu=0$ and $\sigma^2=1$. Set

$$Z_{i,n}=\frac{1}{\sqrt{n}}\sum_{j=1}^{n} X_{i,j}, \qquad n\geqslant 1, \quad i\geqslant 1$$

and (as before)

$$\Phi_n(x)=P(Z_{1,n}\leqslant x), \qquad x\in\mathbb{R}, \quad n\geqslant 1$$

For $N\geqslant 1$, the empirical distribution function of the $\{Z_{i,n};\ 1\leqslant i\leqslant N\}$ is

$$\hat{\Phi}_{N,n}(x)=\frac{1}{N}\sum_{i=1}^{N} 1_{(-\infty,x]}(Z_{i,n})$$

We are interested in the difference $\hat{\Phi}_{N,n}-\Phi$. It follows from

$$\hat{\Phi}_{N,n}-\Phi=(\hat{\Phi}_{N,n}-\Phi_n)+(\Phi_n-\Phi) \tag{1.2}$$

the Glivenko–Cantelli theorem (applied to the first term on the right-hand-side of (1.2)) and the CLT (applied to the second term on the right-hand-side of (1.2)) that

$$\lim_{n\to\infty} \lim_{N\to\infty} \sup_{x\in\mathbb{R}} |\hat{\Phi}_{N,n}(x) - \Phi(x)| = 0 \qquad \text{a.s.}$$

Let us take $n = n(N)$ in the sequel. Rosalsky[12] proved the following result: Let $\{n(N);\ N \geqslant 1\}$ be a sequence of positive integers, let $\{b_N;\ N \geqslant 1\}$ be a sequence of positive constants, and suppose $0 \leqslant \delta \leqslant 1$ satisfies $E\,|X_{1,1}|^{2+\delta} < \infty$ where if $\delta = 0$ then $\lim_{N\to\infty} n(N) = \infty$ and $b_N = O(1)$ and if $\delta > 0$, then $b_N^2 = o((n(N))^\delta)$. If

$$\sum_{N=1}^{\infty} \exp\left\{-\frac{\lambda N}{b_N^2}\right\} < \infty \qquad \text{for all} \quad \lambda > 0 \tag{1.3}$$

then

$$\lim_{N\to\infty} b_N \sup_{x\in\mathbb{R}} |\hat{\Phi}_{N,n(N)}(x) - \Phi(x)| = 0 \qquad \text{a.s.} \tag{1.4}$$

If $b_N \equiv 1$, then the convergence of the series of (1.3) for all $\lambda > 0$ is automatic and so $EX_{1,1}^2 < \infty$ and $\lim_{N\to\infty} n(N) = \infty$ ensure that

$$\lim_{N\to\infty} \sup_{x\in\mathbb{R}} |\hat{\Phi}_{N,n(N)}(x) - \Phi(x)| = 0 \qquad \text{a.s.} \tag{1.5}$$

The functional CLT for $\sqrt{N}(\hat{\Phi}_{N,n(N)} - \Phi)$ follows also readily along these lines, under natural conditions; clearly the limiting process is $B \circ \Phi$, with B a standard Brownian bridge (cf. Section 3 of Rosalsky[12]).

It should be noted that (1.4) is in the spirit of a *strong law of large numbers* for empirical distribution functions based on sample means. This serves as a point of departure and the challenging and natural problem that then arises is to obtain a nondegenerate version of (1.4), i.e., to obtain the rate of convergence in (1.5). Note that it is not at all obvious that the *law of the iterated logarithm* (LIL) holds here, since in principle fluctuations infinitely larger than $\sqrt{\log\log N}$ may occur, due to the fact that the $Z_{i,n(N)}$ change with N. However, under suitable conditions, we will prove that the *functional* LIL for the process $\sqrt{N}(\hat{\Phi}_{N,n(N)} - \Phi)$ (see Theorem 1) holds, thereby combining two fundamental results of probability theory, the CLT and the LIL, in one result.

The functional LIL for the classical empirical process based on the first N of a sequence of i.i.d. random variables has been established in Finkelstein,[5] motivated by the celebrated Strassen[16] functional LIL for the partial-sum process. A functional LIL, in particular, determines the set of almost sure subsequential limit functions. It is the strong analogue of the functional CLT, but typically the proof is much more involved because of the "blocking argument" needed to show that between the terms of an appropriately chosen subsequence (along which the LIL is proved first) the empirical processes do not oscillate too much. In the present situation the blocking argument leads to detailed insight into the fine structure of the successive sample means, which are the building blocks of the empirical distribution function. Once a functional LIL is established, the LIL for continuous functionals of the empirical process (divided by $\sqrt{2 \log \log N}$) follows immediately. In particular Theorem 1 below yields the following LIL as a corollary (cf. Corollary 2(ii)):

$$\limsup_{N \to \infty} \sup_{x \in \mathbb{R}} \frac{\sqrt{N}\, |\hat{\Phi}_{N, n(N)}(x) - \Phi(x)|}{\sqrt{2 \log \log N}} = \frac{1}{2} \qquad \text{a.s.} \tag{1.6}$$

specifying the rate (and also the constant) in (1.5).

It is interesting to note that the convergence criteria in (1.6) and in Theorem 1 wherein there are *iterated* logarithms is markedly different from the convergence criteria of the $Z_{N,N}$ alone obtained in Hu and Weber[8] wherein there is a *single* logarithm. That result asserts that if $EX_{1,1}^4 < \infty$, then

$$\limsup_{N \to \infty} \frac{\pm Z_{N,N}}{\sqrt{2 \log N}} = 1 \qquad \text{a.s.} \tag{1.7}$$

This is not surprising in view of (1.1) and the well-known fact that for a sequence $\{Z_N; N \geqslant 1\}$ of independent standard normal random variables

$$\limsup_{N \to \infty} \frac{\pm Z_N}{\sqrt{2 \log N}} = 1 \qquad \text{a.s.}$$

The result (1.7) was improved in Li *et al.*,[9] where it was shown that

$$E\left(\frac{|X_{1,1}|^{2(1+\alpha^{-1})}}{(\log(e \vee |X_{1,1}|))^{1+\alpha^{-1}}}\right) < \infty, \qquad \alpha > 0$$

is necessary and sufficient for

$$\limsup_{N \to \infty} \frac{\pm Z_{N, \lfloor N^\alpha \rfloor}}{\sqrt{2 \log N}} = 1 \qquad \text{a.s.}$$

This had been obtained in Li *et al.*[10] and in Qi[11] for the case $\alpha = 1$.

This paper is organized as follows. In Section 2 the main result, the aforementioned functional LIL, is presented along with some corollaries. Some lemmas needed in the proof of the main result will be established in Section 3. Finally, in Section 4, the proof of the main result is given.

2. THE MAIN RESULT

With the preliminaries accounted for, the main result may be stated. For convenience, let us introduce the notation $\gamma_N = \sqrt{N}(\hat{\Phi}_{N,n(N)} - \Phi)$, $N \geqslant 1$, where $\hat{\Phi}_{N,n(N)}$, Φ, and $\{X_{i,j};\ i \geqslant 1,\ j \geqslant 1\}$ are as in Section 1. Theorem 1 asserts under suitable conditions that $\gamma_N/(2 \log \log N)^{1/2}$ is a.s. relatively compact in $B(\mathbb{R})$ (with respect to the sup-norm) and identifies the set of its a.s. limit functions. As in Shorack and Wellner,[15] p. 69, we define $Y_m \rightsquigarrow L$ a.s. $(m \to \infty)$, to mean that Y_m is a.s. relatively compact with limit set L.

Theorem 1. Assume that $E|X_{1,1}|^\beta < \infty$ for some $\beta > 2$ and let $\alpha \geqslant ((\beta \wedge 3) - 2)^{-1}$. Then as $N \to \infty$, with $n = n(N) = \lfloor N^\alpha \rfloor$,

$$\frac{\gamma_N}{\sqrt{2 \log \log N}} \rightsquigarrow \mathscr{F}_\Phi \qquad \text{a.s.} \tag{2.1}$$

on $B(\mathbb{R})$, where $\mathscr{F}_\Phi = \{f \circ \Phi;\ f \in \mathscr{F}\}$ with

$$\mathscr{F} = \{f;\ f \text{ is an absolutely continuous function on } [0, 1] \text{ with } f(0) = f(1) = 0 \text{ and } \textstyle\int_0^1 (f'(t))^2\, dt \leqslant 1\}$$

Remarks. (i) Theorem 1 should be compared with the functional LIL of Finkelstein[5] for the empirical distribution functions formed from a sequence of i.i.d. random variables. Observe that although these results look very similar, our setup is completely different from the i.i.d. case: in no way does $1_{(-\infty, x]}(Z_{i,n(N)})$ (which is a summand in $\hat{\Phi}_{N,n(N)}$) relate to $1_{(-\infty, x]}(Z_i)$, $\{Z_i;\ i \geqslant 1\}$ i.i.d., not even for normal $X_{i,j}$'s and normal Z_i's. The Z_i's do not depend on N, whereas the $Z_{i,n(N)}$ oscillate according to their own strong limiting behavior.

(ii) Observe that if $E|X_{1,1}|^3 < \infty$, then the sample size n can be taken to be equal to the number of samples N.

(iii) Recent and somewhat related papers are Deheuvels and Steinebach[1] and Deheuvels and Teicher,[2] where the strong limiting behavior of *maxima* of sample means has been studied in great detail for the same double array setup. Also see Li *et al.*[9] where a strong invariance principle and a Strassen-type functional limit law for the array $\{X_{N,j}; 1 \leqslant j \leqslant \lfloor N^\alpha \rfloor, N \geqslant 1\}$, $\alpha > 0$, of i.i.d. random variables are obtained.

Several corollaries of Theorem 1 will now be presented.

Corollary 1. When $X_{1,1}$ is normally distributed, the result in (2.1), with $n(N) = \lfloor N^\alpha \rfloor$, holds for all $\alpha > 0$.

Proof. The restriction $\alpha \geqslant ((\beta \wedge 3) - 2)^{-1}$ is only needed for the proof of Theorem 1 at places where the Berry–Esseen inequality is applied. Since the $X_{i,j}$ themselves are already normal now, this restriction is not needed, meaning that the theorem holds for any $\alpha > 0$. □

The following corollary is a standard application of Theorem 1.

Corollary 2. Let $h: B(\mathbb{R}) \to \mathbb{R}$ be continuous on $\mathscr{F}_\Phi$ and suppose that the hypotheses of Theorem 1 (or Corollary 1) are satisfied. Then, as $N \to \infty$,

$$h\left(\frac{\gamma_N}{\sqrt{2 \log \log N}}\right) \rightsquigarrow h(\mathscr{F}_\Phi) \qquad \text{a.s.}$$

In particular:

(i) For all fixed $x \in \mathbb{R}$,

$$\frac{\gamma_N(x)}{\sqrt{2 \log \log N}} \rightsquigarrow [-(\Phi(x)(1-\Phi(x)))^{1/2}, (\Phi(x)(1-\Phi(x)))^{1/2}] \qquad \text{a.s.}$$

(ii)

$$\sup_{x \in \mathbb{R}} \frac{|\gamma_N(x)|}{\sqrt{2 \log \log N}} \rightsquigarrow \left[0, \frac{1}{2}\right] \qquad \text{a.s.}$$

(iii) For either choice of sign

$$\sup_{x \in \mathbb{R}} \frac{\pm \gamma_N(x)}{\sqrt{2 \log \log N}} \rightsquigarrow \left[0, \frac{1}{2}\right] \qquad \text{a.s.}$$

(iv)

$$\frac{\int_{-\infty}^{\infty} \gamma_N^2(x)\, d\Phi(x)}{2 \log \log N} \rightsquigarrow \left[0, \frac{1}{\pi^2}\right] \qquad \text{a.s.}$$

The final corollary is a functional LIL for the "uniformized" quantile process based on sample means. It follows readily from Theorem 1 in conjunction with Vervaat's[18] lemma.

Corollary 3. Under the hypotheses of Theorem 1 (or Corollary 1), as $N \to \infty$,

$$\frac{\sqrt{N}(\Phi \circ \hat{\Phi}_{N,n(N)}^{-1} - I)}{\sqrt{2 \log \log N}} \rightsquigarrow \mathscr{F} \qquad \text{a.s.}$$

on $B([0, 1])$, where $\hat{\Phi}_{N,n(N)}^{-1}$ denotes the left-continuous quantile function corresponding to $\hat{\Phi}_{N,n(N)}$ and I denotes the identity function on $[0, 1]$.

3. LEMMAS

The first lemma extends part of Lemma 1.1.1 in van Zuijlen[17] (which gives bounds for moments of binomial random variables) to the trinomial case.

Lemma 1. Let $T_n = \sum_{i=1}^{n} X_i$ where $X_1, \ldots, X_n$ are i.i.d. random variables with

$$P(X_1 = 1) = p, \qquad P(X_1 = -1) = q, \qquad P(X_1 = 0) = 1 - p - q$$

where $p \geqslant \frac{1}{n}$, $q \geqslant \frac{1}{n}$, and $p + q \leqslant 1$. Then for every $v > 1$, there exists $K(v) < \infty$ such that for all n $(\geqslant 2)$

$$\begin{aligned} E\,|T_n - n(p-q)|^{v} &\leqslant \tfrac{1}{3} K(v)(n(p(1-p) + q(1-q) + 2pq))^{v/2} \\ &\leqslant K(v)(n(p \vee q))^{v/2} \end{aligned}$$

The proof of the first inequality is essentially the same as that in van Zuijlen[17] and will hence be omitted; the second inequality is trivial.

The second lemma is a modification of Theorem B in Serfling,[13] tailor made for our situation. Let S_N, $N \geqslant 1$, be the partial sums of some sequence of random variables (which are not necessarily independent or identically distributed) and set $S_{N,b} = S_{N+b} - S_N$, $N, b \geqslant 1$; also write $M_{N,b} = \max_{1 \leqslant i \leqslant b} |S_{N,i}|$. Set as in Section 4, $N_k = \lceil e^{k^{\eta}} \rceil$, $0 < \eta < 1$.

Lemma 2. Let $v \geqslant 7$ be an integer. Assume that for all $k \geqslant k_0$ (some k_0) and all $N_k \leqslant N < N + b \leqslant N_{k+1}$,

$$E\left(\frac{|S_{N,b}|}{N_k^{1/3}}\right)^{v} \leqslant C_1 b^{v/6} \tag{3.1}$$

for some $C_1 = C_1(\nu) < \infty$. Then there exists $C = C(\nu) < \infty$ such that for all $k \geqslant k_0$, $N_k \leqslant N \leqslant N + b \leqslant N_{k+1}$,

$$E\left(\frac{M_{N,b}}{N_k^{1/3}}\right)^\nu \leqslant Cb^{\nu/6} \tag{3.2}$$

Proof. We will give a proof by induction on b. First consider all $b \leqslant b_0$ (for any fixed b_0). Then trivially by (3.1)

$$E\left(\frac{M_{N,b}}{N_k^{1/3}}\right)^\nu \leqslant \frac{1}{N_k^{\nu/3}} E(|S_{N,1}|^\nu + \cdots + |S_{N,b}|^\nu) \leqslant \frac{1}{N_k^{\nu/3}} bC_1 b^{\nu/6} \leqslant Cb^{\nu/6}$$

Now let $B > b_0$ and suppose (3.2) holds for all $b < B$. We will show that (3.2) holds for B. Set $K = \lfloor \frac{1}{2}(B+1) \rfloor$ and note that $2K \geqslant B$ and hence $B - K \leqslant K$. When $b \leqslant K$ we have $|S_{N,b}|^\nu \leqslant M_{N,b}^\nu$. When $K \leqslant b \leqslant B$ we obtain

$$\begin{aligned}|S_{N,b}|^\nu &\leqslant (|S_{N,K}| + M_{N+K,B-K})^\nu \\ &= |S_{N,K}|^\nu + M_{N+K,B-K}^\nu + \sum_{j=1}^{\nu-1} \binom{\nu}{j} |S_{N,K}|^j M_{N+K,B-K}^{\nu-j}\end{aligned}$$

Hence

$$M_{N,B}^\nu \leqslant M_{N,K}^\nu + M_{N+K,B-K}^\nu + \sum_{j=1}^{\nu-1} \binom{\nu}{j} |S_{N,k}|^j M_{N+K,B-K}^{\nu-j} \tag{3.3}$$

By Hölder's inequality

$$E\,|S_{N,K}|^j M_{N+K,B-K}^{\nu-j} \leqslant (E\,|S_{N,K}|^\nu)^{j/\nu} (EM_{N+K,B-K})^{(\nu-j)/\nu}$$

So from (3.3)

$$\begin{aligned}E\left(\frac{M_{N,B}}{N_k^{1/3}}\right)^\nu &\leqslant 2CK^{\nu/6} + \sum_{j=1}^{\nu-1} \binom{\nu}{j} C_1^{j/\nu} K^{j/6} C^{(\nu-j)/\nu} K^{(\nu-j)/6} \\ &= CK^{\nu/6} \left\{2 + \sum_{j=1}^{\nu-1} \binom{\nu}{j} \left(\frac{C_1}{C}\right)^{j/\nu}\right\}\end{aligned}$$

Now using $\nu \geqslant 7$ and taking C and b_0 above large enough, it is clear that this last expression is bounded from above by $CB^{\nu/6}$. □

The next lemma is a kind of a.s. analogue of the well-known Cramér–Wold device for establishing convergence in distribution of random vectors, and may be of independent interest.

Lemma 3. Let $\{X_n; n \geqslant 1\}$ be a sequence of random vectors in $\mathbb{R}^m$, $m \geqslant 1$. Suppose that

$$\limsup_{n \to \infty} a'X_n = \sqrt{a'\Sigma a} \quad \text{a.s.} \qquad \text{for all} \quad a \in \mathbb{R}^m$$

where Σ is a symmetric, nonnegative definite matrix with determinant $|\Sigma| \neq 0$. Then

$$X_n \rightsquigarrow L \qquad \text{a.s.}$$

where

$$\{x \in \mathbb{R}^m; x'\Sigma^{-1}x = 1\} \subseteq L \subseteq \{x \in \mathbb{R}^m; x'\Sigma^{-1}x \leqslant 1\}$$

Proof. Denote the entries of Σ with σ_{ij}, $i, j = 1,..., m$. Set $K = \{x \in \mathbb{R}^m; x'\Sigma^{-1}x = 1\}$ and $M = \{x \in \mathbb{R}^m; x'\Sigma^{-1}x \leqslant 1\}$. Also observe that it easily follows that

$$P(\limsup_{n \to \infty} a'X_n = \sqrt{a'\Sigma a} \text{ for all } a \in \mathbb{R}^m) = 1 \tag{3.4}$$

We need to prove that almost surely the following hold:

(i) $\{X_n\}$ is relatively compact,

(ii) all the subsequential limits of $\{X_n\}$ are contained in M, and

(iii) every $b \in K$ is a subsequential limit of $\{X_n\}$.

For (i) we take a to be the vector with the jth entry being 1 or -1, $j = 1,..., m$, and all the other entries 0. Both cases give $\sqrt{\sigma_{jj}}$ as the a.s. value for $\limsup_{n \to \infty} a'X_n$. Hence for $\varepsilon > 0$ and for almost every ω there exists an $n_\varepsilon(\omega)$ such that for $n \geqslant n_\varepsilon(\omega)$ we have that $X_n \in \times_{j=1}^m [-\sqrt{\sigma_{jj}} - \varepsilon, \sqrt{\sigma_{jj}} + \varepsilon]$. This settles the relative compactness of $\{X_n\}$.

For (ii), let us assume that ω is such that

$$\limsup_{n \to \infty} a'X_n = \sqrt{a'\Sigma a} \qquad \text{for all} \quad a \in \mathbb{R}^m \tag{3.5}$$

and that there exists a subsequential limit, b say, of $\{X_n\}$ that is not contained in M, i.e., $b'\Sigma^{-1}b > 1$. Now take $a = \Sigma^{-1}b$. So $a'\Sigma a = b'\Sigma^{-1}b > 1$ and hence

$$\limsup_{n \to \infty} a'X_n \geqslant a'b = a'\Sigma a > \sqrt{a'\Sigma a}$$

This contradicts (3.5).

Finally, for (iii) set $\|x\|_{\Sigma^{-1}} = \sqrt{x'\Sigma^{-1}x}$, $x \in \mathbb{R}^m$. Because of (ii) we have

$$\limsup_{n\to\infty} \|X_n\|_{\Sigma^{-1}} \leqslant 1 \qquad \text{a.s.} \tag{3.6}$$

By (3.6) and (3.4), with $a = \Sigma^{-1}b$, we have a.s. for some subsequence $\{n_k\}$

$$\begin{aligned}
&\limsup_{k\to\infty} \|X_{n_k} - b\|^2_{\Sigma^{-1}} \\
&\quad = \limsup_{k\to\infty} \{ \|X_{n_k}\|^2_{\Sigma^{-1}} + \|b\|^2_{\Sigma^{-1}} - 2b'\Sigma^{-1}X_{n_k} \} \\
&\quad \leqslant 1 + 1 - 2\sqrt{b'\Sigma^{-1}b} = 0
\end{aligned}$$

and hence also, since all norms on $\mathbb{R}^m$ are equivalent,

$$\lim_{k\to\infty} \|X_{n_k} - b\| = 0$$

with $\|x\| = \sqrt{x'x}$, the Euclidean norm of $x \in \mathbb{R}^m$. □

Observe from the proof that Lemma 3 is essentially an analytic lemma; it remains true for deterministic vectors (and without the "a.s.").

Next we state Bennett's inequality, see Shorack and Wellner,[15] p. 851.

Lemma 4. Let $\{X_i;\ 1 \leqslant i \leqslant n\}$ be i.i.d. random variables with $EX_i = 0$, $|X_i| \leqslant b$ and Var $X_i = \sigma^2$. Then

$$P\left(\frac{|\sum_{i=1}^n X_i|}{\sqrt{n}} \geqslant \lambda\right) \leqslant 2\exp\left(\frac{-\lambda^2}{2\sigma^2}\,\psi\left(\frac{\lambda b}{\sqrt{n}\sigma^2}\right)\right), \qquad \lambda > 0$$

where $\psi(\lambda) = 2\lambda^{-2}((1+\lambda)\log(1+\lambda) - \lambda)$, $\lambda > 0$.

The last lemma provides a probability inequality for the oscillation modulus ω_n of the uniform empirical process; see Einmahl and Ruymgaart[4] or Einmahl,[3] Chapter 5.

Lemma 5. Let $0 < \delta \leqslant \frac{1}{2}$ and $0 < a < \frac{1}{4}$. Then there exists a constant $C_\delta < \infty$ such that

$$P(\omega_n(a) \geqslant \lambda) \leqslant C_\delta\,\frac{1}{a}\exp\left(\frac{-(1-\delta)\,\lambda^2}{2a}\,\psi\left(\frac{\lambda}{\sqrt{na}}\right)\right), \qquad \lambda > 0$$

4. PROOFS

Let us first remark that

$$\sup_{x \in \mathbb{R}} |\Phi_{\lfloor N^\alpha \rfloor}(x) - \Phi(x)| = O(N^{-(1/2)\alpha((\beta \wedge 3)-2)})$$

by the Berry–Esseen inequality; see, e.g., Shorack and Wellner,[15] p. 848. Hence

$$\sqrt{\frac{N}{2 \log \log N}} \sup_{x \in \mathbb{R}} |\Phi_{\lfloor N^\alpha \rfloor}(x) - \Phi(x)| \to 0$$

if $\alpha \geqslant ((\beta \wedge 3)-2)^{-1}$. This means that we can and will redefine γ_N by *replacing* Φ *by* $\Phi_{\lfloor N^\alpha \rfloor}$ *in its definition* without affecting our results. Also in the sequel we often simply write n for $\lfloor N^\alpha \rfloor$.

We need some more notation. Let $m \geqslant 2$ (m typically large) and set

$$Y_{N,j} = \frac{\gamma_N\left(\Phi^{-1}\left(\frac{j}{m}\right)\right)}{\sqrt{2 \log \log N}}, \qquad j = 1,..., m-1$$

Let $\{\varepsilon_i; i \geqslant 1\}$ be a sequence of independent Rademacher random variables, i.e., $P(\varepsilon_i = 1) = P(\varepsilon_i = -1) = \frac{1}{2}$. Assume that the $\{\varepsilon_i\}$ are independent of the $\{X_{i,j}\}$. Define

$$Y_{N,m} = \frac{\sum_{i=1}^N \varepsilon_i}{\sqrt{2N \log \log N}}$$

and write

$$Y_N = \begin{pmatrix} Y_{N,1} \\ \vdots \\ Y_{N,m-1} \end{pmatrix}, \qquad \tilde{Y}_N = \begin{pmatrix} Y_{N,1} \\ \vdots \\ Y_{N,m} \end{pmatrix}.$$

Define the $m \times m$-matrix

$$\tilde{\Sigma} = \begin{pmatrix} \Sigma & 0 \\ 0 & 1 \end{pmatrix}$$

with Σ being the symmetric $(m-1) \times (m-1)$-matrix having entries

$$\sigma_{ij} = \frac{i}{m}\left(1 - \frac{j}{m}\right), \qquad i \leqslant j,$$

$$\sigma_{ij} = \sigma_{ji}, \qquad i > j$$

The proof of Theorem 1 relies heavily on the following result.

Proposition 1. For any vector $a \in \mathbb{R}^m$

$$\limsup_{N \to \infty} a' \tilde{Y}_N = \sqrt{a' \tilde{\Sigma} a} \qquad \text{a.s.} \tag{4.1}$$

Observe that this result specializes to

$$\limsup_{N \to \infty} \pm Y_{N,j} = \sqrt{\sigma_{jj}} = \frac{1}{m} \sqrt{j(m-j)} \qquad \text{a.s.}$$

for $j = 1, ..., m-1$;

$$\limsup_{N \to \infty} \pm Y_{N,m} = 1 \qquad \text{a.s.}$$

by taking the jth component of a to be ± 1 and by taking all the other components to be 0.

Proof of Proposition 1. First we consider the upper bound part of (4.1); cf. the proof of Theorem 4.1 in Serfling.[14] For the upper bound it suffices to show that for any $\varepsilon > 0$

$$\limsup_{N \to \infty} a' \tilde{Y}_N \leqslant (1+\varepsilon) \sqrt{a' \tilde{\Sigma} a} \qquad \text{a.s.} \tag{4.2}$$

Set $N_k = \lceil e^{k^\eta} \rceil$, $0 < \eta < 1$; η to be specified later on. First we prove that

$$\limsup_{k \to \infty} a' \tilde{Y}_{N_k} \leqslant (1+\varepsilon) \sqrt{a' \tilde{\Sigma} a} \qquad \text{a.s.}$$

Write $c_{j,m} = \Phi^{-1}(j/m)$. Note that for $j < m$

$$\begin{aligned} Y_{N,j} &= \frac{\sqrt{N}(\hat{\Phi}_{N,n}(c_{j,m}) - \Phi_n(c_{j,m}))}{\sqrt{2 \log \log N}} \\ &= \frac{1}{\sqrt{2N \log \log N}} \sum_{i=1}^{N} \{ 1_{(-\infty, c_{j,m}]}(Z_{i,n}) - \Phi_n(c_{j,m}) \} \end{aligned}$$

We will now apply Bennett's inequality (Lemma 4) taking the X_i there to be

$$\sum_{j=1}^{m} a_j \{ 1_{(-\infty, c_{j,m}]}(Z_{i,n}) - \Phi_n(c_{j,m}) \} + a_m \varepsilon_i$$

where a_j is the jth component of a. Hence the b in Lemma 4 can be taken to be $\sum_{j=1}^{m} |a_j|$ and we have that

$$\sigma^2 = \sigma^2(X_i) \to a'\tilde{\Sigma}a \quad (N \to \infty)$$

So, for large k

$$\begin{aligned} p_k &= P\left(\frac{|\sum_{i=1}^{N_k} X_i|}{\sqrt{N_k}} \geqslant (1+\varepsilon)\sqrt{a'\tilde{\Sigma}a}\sqrt{2\log\log N_k}\right) \\ &\leqslant 2\exp\left(\frac{-(1+\varepsilon)^2\, a'\tilde{\Sigma}a \cdot 2\log\log N_k}{2\sigma^2}\right. \\ &\quad \left.\times \psi\left(\frac{(1+\varepsilon)\sqrt{a'\tilde{\Sigma}a}\sqrt{2\log\log N_k}\, \sum_{j=1}^{m}|a_j|}{\sqrt{N_k}\sigma^2}\right)\right) \\ &\leqslant 2\exp(-(1+\varepsilon)\,\eta\log k) \end{aligned}$$

where it is used that $\lim_{\lambda \downarrow 0} \psi(\lambda) = 1$. Hence $\sum_{k=1}^{\infty} p_k < \infty$, provided we choose $\eta > 1/(1+\varepsilon)$.

So to establish (4.2) it remains to show that

$$\lim_{k\to\infty} \max_{N_k < N < N_{k+1}} \frac{|a'(\sqrt{2N\log\log N}\,\tilde{Y}_N - \sqrt{2N_k\log\log N_k}\,\tilde{Y}_{N_k})|}{\sqrt{N_k\log\log N_k}} = 0 \qquad \text{a.s.} \tag{4.3}$$

For $x \in \mathbb{R}$ set $S_N = S_N(x) = \sum_{i=1}^{n} \{1_{(-\infty, x]}(Z_{i,n}) - \Phi_n(x)\}$. Then it is clear that for (4.3) it suffices to show that for any $x \in \mathbb{R}$

$$\lim_{k\to\infty} \max_{N_k < N < N_{k+1}} \frac{|S_N - S_{N_k}|}{\sqrt{N_k\log\log N_k}} = 0 \qquad \text{a.s.} \tag{4.4}$$

Note that a similar statement dealing with the mth component of $\tilde{Y}_N$ is needed also. However, that holds because of the proof of the ordinary LIL, so we do not need to consider it here.

In order to deal with (4.4) fix a large k and let N and b be such that $N_k \leqslant N < N + b \leqslant N_{k+1}$. Write

$$
\begin{aligned}
S_{N,b} &= S_{N+b} - S_N \\
&= \left\{ \sum_{i=N+1}^{N+b} 1_{(-\infty,x]}(Z_{i,\lfloor (N+b)^\alpha \rfloor}) - b\Phi_{\lfloor (N+b)^\alpha \rfloor}(x) \right\} \\
&\quad + \left\{ \sum_{i=1}^{N} 1_{(-\infty,x]}(Z_{i,\lfloor (N+b)^\alpha \rfloor}) - \sum_{i=1}^{N} 1_{(-\infty,x]}(Z_{i,\lfloor N^\alpha \rfloor}) \right. \\
&\quad \left. - N(\Phi_{\lfloor (N+b)^\alpha \rfloor}(x) - \Phi_{\lfloor N^\alpha \rfloor}(x)) \right\} \\
&= S_{1,N,b} + S_{2,N,b}
\end{aligned} \tag{4.5}
$$

We will show that for $\nu \geqslant 7$ and some $C_\ell = C_\ell(\nu) < \infty$

$$
E\,|S_{\ell,N,b}|^\nu \leqslant C_\ell N_k^{\nu/3} b^{\nu/6}, \qquad \ell = 1, 2 \tag{4.6}
$$

and hence, since $E\,|X+Y|^\nu \leqslant 2^{\nu-1}(E\,|X|^\nu + E\,|Y|^\nu)$ for $\nu \geqslant 1$,

$$
E\,|S_{N,b}|^\nu \leqslant C N_k^{\nu/3} b^{\nu/6} \tag{4.7}
$$

for some $C = C(\nu) < \infty$. Showing (4.6) for $\ell = 1$ is easy: applying bounds for the absolute moments of binomial random variables (see, e.g., van Zuijlen,[17] Lemma 1.1.1) yields for k large enough

$$
E\,|S_{1,N,b}|^\nu \leqslant C_1 b^{\nu/2} \leqslant C_1 N_k^{\nu/3} b^{\nu/6}
$$

Now we consider (4.6) for $\ell = 2$. Set $Z' = Z_{1,\lfloor (N+b)^\alpha \rfloor}$ and $Z'' = Z_{1,\lfloor N^\alpha \rfloor}$. Then

$$
\begin{aligned}
&P(Z' \leqslant x, Z'' \geqslant x) \\
&\quad = P\left(Z' \leqslant x, Z'' \geqslant x, Z'' - Z' \leqslant \left(\frac{b}{N_k} \right)^{1/3} \right) \\
&\quad\quad + P\left(Z' \leqslant x, Z'' \geqslant x, Z'' - Z' > \left(\frac{b}{N_k} \right)^{1/3} \right) \\
&\quad = P\left(x - \left(\frac{b}{N_k} \right)^{1/3} \leqslant Z' \leqslant x \right) + P\left(Z'' - Z' > \left(\frac{b}{N_k} \right)^{1/3} \right)
\end{aligned} \tag{4.8}
$$

By the Berry–Esseen inequality we have for large k

$$P\left(x-\left(\frac{b}{N_k}\right)^{1/3} \leqslant Z' \leqslant x\right) \leqslant \frac{1}{\sqrt{2\pi}}\left(\frac{b}{N_k}\right)^{1/3} + O(\lfloor (N+b)^{\alpha} \rfloor^{(-(\beta \wedge 3)-2)/2})$$
$$\leqslant \left(\frac{b}{N_k}\right)^{1/3}$$

and by Chebyshev's inequality for large k

$$P\left(Z''-Z' > \left(\frac{b}{N_k}\right)^{1/3}\right) \leqslant \frac{2(1-(\lfloor N^{\alpha}\rfloor/\lfloor (N+b)^{\alpha}\rfloor)^{1/2})}{(b/N_k)^{2/3}} = O\left(\left(\frac{b}{N_k}\right)^{1/3}\right)$$

Hence for some c_2, the left-hand-side of (4.8) is bounded from above by $c_2(b/N_k)^{1/3}$. Similarly we can bound $P(Z'' \leqslant x, Z' \geqslant x)$ by the same quantity. Hence by Lemma 1

$$E\,|S_{2,N,b}|^{\nu} \leqslant K(\nu)\, N^{\nu/2}\left(c_2\left(\frac{b}{N_k}\right)^{1/3}\right)^{\nu/2} \leqslant C_2 N_k^{\nu/3} b^{\nu/6}$$

This is (4.6) for $\ell = 2$. Hence we have (4.7).

Now applying Lemma 2 we obtain

$$E \max_{N_k < N < N_{k+1}} |S_{N_k, N-N_k}|^{\nu} \leqslant C N_k^{\nu/3}(N_{k+1}-N_k)^{\nu/6}$$

So for any $\varepsilon > 0$,

$$\begin{aligned} q_k &= P(\max_{N_k < N < N_{k+1}} |S_N - S_{N_k}| > \varepsilon\sqrt{N_k \log\log N_k}) \\ &\leqslant \frac{C N_k^{\nu/3}(N_{k+1}-N_k)^{\nu/6}}{\varepsilon^{\nu} N_k^{\nu/2}(\log\log N_k)^{\nu/2}} \\ &= \frac{C}{\varepsilon^{\nu}(\log\log N_k)^{\nu/2}}\left(\frac{N_{k+1}-N_k}{N_k}\right)^{\nu/6} \\ &= O\left(\frac{1}{(\log k)^{\nu/2}\, k^{(1-\eta)\nu/6}}\right) \end{aligned}$$

Hence for ν such that $\nu \geqslant 6/(1-\eta) \vee 7$ we see that $\sum_{k=1}^{\infty} q_k < \infty$. This yields (4.4) and hence (4.2).

Next we turn to the easier lower bound in (4.1). We will actually show that for any $0 < \varepsilon < 1$

$$\limsup_{k \to \infty} a' \tilde{Y}_{N_k} \geqslant (1-\varepsilon)\sqrt{a'\tilde{\Sigma} a} \qquad \text{a.s.} \tag{4.9}$$

with again $N_k = \lceil e^{k^\eta} \rceil$, but now with $\eta > 1$. Note that $\eta > 1$ implies $\lim_{k\to\infty} N_k/N_{k-1} = \infty$. Hence, it is easy to show that for any $x \in \mathbb{R}$, with $n_k = \lfloor N_k^\alpha \rfloor$,

$$\lim_{k\to\infty} \frac{N_k^{1/2}}{(\log\log N_k)^{1/2}} |\hat{\Phi}_{N_k, n_k}(x) - \hat{\Phi}^{(N_k)}_{(N_{k-1})}(x)| = 0 \qquad \text{a.s.} \tag{4.10}$$

with $\hat{\Phi}^{(N_k)}_{(N_{k-1})}(x) = (1/(N_k - N_{k-1})) \sum_{i=N_{k-1}+1}^{N_k} 1_{(-\infty, x]}(Z_{i,n_k})$. In fact (4.10) follows by writing

$$\begin{aligned}&\hat{\Phi}_{N_k, n_k}(x) - \hat{\Phi}^{(N_k)}_{(N_{k-1})}(x)\\ &= \frac{N_{k-1}}{N_k}\left\{\frac{1}{N_{k-1}} \sum_{i=1}^{N_{k-1}} 1_{(-\infty,x]}(Z_{i,n_k}) - \frac{1}{N_k - N_{k-1}} \sum_{i=N_{k-1}+1}^{N_k} 1_{(-\infty,x]}(Z_{i,n_k})\right\}\end{aligned} \tag{4.11}$$

and by applying the Hoeffding[7] exponential inequality to the two probabilities that have to be considered by separately treating the two terms on the right-hand-side of (4.11), after centering. This is straight forward, so we omit details. Also note that a statement similar to (4.10) can be easily proven for the ε_i's of $Y_{N_k, m}$.

Hence to prove (4.9), it suffices to show

$$\limsup_{k\to\infty} a' \bar{Y}_{N_k} \geqslant (1-\varepsilon)\sqrt{a'\tilde{\Sigma} a} \qquad \text{a.s.} \tag{4.12}$$

where $\bar{Y}_{N_k}$ is defined in the same way as $\tilde{Y}_{N_k}$, but with $\hat{\Phi}_{N_k, n_k}(c_{j,m})$ replaced by $\hat{\Phi}^{(N_k)}_{(N_{k-1})}(c_{j,m})$, $j = 1,..., m-1$, and with $Y_{N_k, m}$ replaced by

$$\left(\frac{N_k}{N_k - N_{k-1}}\right) \frac{1}{\sqrt{2N_k \log\log N_k}} \sum_{i=N_{k-1}+1}^{N_k} \varepsilon_i$$

But, in contrast to the $\tilde{Y}_{N_k}$, the $\bar{Y}_{N_k}$ are independent, so we can apply the Borel–Cantelli lemma directly. Kolmogorov's exponential inequality (see, e.g., Shorack and Wellner,[15] p. 855) yields for large k, using similar arguments as in the upper bound proof,

$$\begin{aligned} r_k &= P(a' \bar{Y}_{N_k} \geqslant (1-\varepsilon)\sqrt{a'\tilde{\Sigma} a})\\ &\geqslant \exp\left(\frac{-(1-\varepsilon)\, a'\tilde{\Sigma} a 2 \log\log N_k \cdot [(N_k - N_{k-1})/N_k]}{2\sigma^2}\right)\\ &\geqslant \exp\left(-\left(1 - \frac{\varepsilon}{2}\right)\eta \log k\right)\end{aligned}$$

Hence $\sum_{k=1}^{\infty} r_k = \infty$ by choosing $\eta \leqslant 1/(1-\frac{\varepsilon}{2})$. This proves (4.12) which in conjunction with (4.10) yields (4.9). □

Proof of Theorem 1. Instead of (2.1) we will actually prove the equivalent statement

$$\frac{\gamma_N \circ \Phi^{-1}}{\sqrt{2 \log \log N}} \rightsquigarrow \mathscr{F} \qquad \text{a.s.} \tag{4.13}$$

on $B([0, 1])$ with the sup-norm.

By Lemma 3 and Proposition 1 we obtain that

$$\tilde{Y}_N \rightsquigarrow L \qquad \text{a.s.}$$

with $\{y \in \mathbb{R}^m; y'\tilde{\Sigma}^{-1}y = 1\} = \{(x, x_m) \in \mathbb{R}^m; x'\Sigma^{-1}x + x_m^2 = 1\} \subseteq L \subseteq \{y \in \mathbb{R}^m; y'\tilde{\Sigma}^{-1}y \leqslant 1\} = \{(x, x_m) \in \mathbb{R}^m; x'\Sigma^{-1}x + x_m^2 \leqslant 1\}$, where we have used

$$\tilde{\Sigma}^{-1} = \begin{pmatrix} \Sigma^{-1} & 0 \\ 0 & 1 \end{pmatrix}$$

Now project $\mathbb{R}^m$ on $\mathbb{R}^{m-1}$, eliminating the last coordinate (which deals with the auxiliary ε_i's). It is clear that the projection of L is equal to $\{x \in \mathbb{R}^{m-1}; x'\Sigma^{-1}x \leqslant 1\}$. Hence

$$Y_N \rightsquigarrow \{x \in \mathbb{R}^{m-1}; x'\Sigma^{-1}x \leqslant 1\} \qquad \text{a.s.} \tag{4.14}$$

By direct calculations one can show that

$$\Sigma^{-1} = \begin{pmatrix} 2m & -m & 0 & \cdot & \cdot & \cdot & 0 \\ -m & 2m & -m & \cdot & \cdot & \cdot & \cdot \\ 0 & -m & 2m & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & -m \\ 0 & \cdot & \cdot & \cdot & 0 & -m & 2m \end{pmatrix}$$

This yields that the set of continuous piecewise linear functions f on $[0, 1]$ with "support points" j/m, $j = 1,..., m-1$, (and being 0 at 0 and 1) satisfying $x'\Sigma^{-1}x \leqslant 1$, with $x_j = f(j/m)$, is the same as $\mathscr{F} \cap \{$continuous piecewise linear functions on j/m, $j = 0, 1,..., m\}$, see Finkelstein.[5]

Since $m \geqslant 2$ is arbitrary (large), we have now proven a result on an arbitrarily fine grid on $[0, 1]$. To prove the theorem, we have to extend this result to the whole interval $[0, 1]$. In order to achieve this, it remains to show that the oscillations between two adjacent grid points vanish as $m \to \infty$, see Finkelstein[5] or Shorack and Wellner,[15] pp. 76–79.

To be more precise, we will show that for all $j = 0, 1, ..., m-1$,

$$\limsup_{N \to \infty} \sup_{\Phi^{-1}(j/m) < x < \Phi^{-1}((j+1)/m)} \frac{|\gamma_N(x) - \gamma_N(\Phi^{-1}(j/m))|}{\sqrt{\log \log N}} \leqslant \frac{3}{m^{1/2}} \qquad \text{a.s.} \tag{4.15}$$

(Actually at the cost of a harder proof we could replace $3/m^{1/2}$ by $(2/m)^{1/2}$, cf. Finkelstein,[5] but (4.15) as it stands is sufficient for our needs, since $\lim_{m \to \infty} 3/m^{1/2} = 0$.) Let j be fixed. We deal with the small interval $[j/m, (j+1)/m]$, by splitting it up into smaller intervals of length $1/\log N$.

First we show

$$\lim_{N \to \infty} V_N = 0 \qquad \text{a.s.} \tag{4.16}$$

where

$$V_N = \sup_{\substack{\Phi^{-1}(j/m) \leqslant x < y \leqslant \Phi^{-1}((j+1)/m) \\ \Phi(y) - \Phi(x) \leqslant 1/\log N}} \frac{|\gamma_N(x) - \gamma_N(y)|}{\sqrt{\log \log N}}$$

From Lemma 5, with $\delta = \frac{1}{2}$, it follows that for any $\varepsilon > 0$ and all large N

$$\begin{aligned} P(V_N \geqslant \varepsilon) &\leqslant C(\log N) \exp\left(\frac{-\varepsilon^2(\log \log N) \log N}{5} \psi\left(\frac{\varepsilon \sqrt{\log \log N} \log N}{\frac{1}{2}\sqrt{N}}\right)\right) \\ &\leqslant C(\log N) \exp\left(\frac{-\varepsilon^2(\log \log N) \log N}{6}\right) \\ &= \frac{C \log N}{N^{(1/6)\varepsilon^2 \log \log N}} \end{aligned}$$

which is summable in N thereby yielding (4.16). Note that the interval length $1/\log N$ is just small enough to prove (4.16) without appealing to a subsequence.

Next we will prove

$$\limsup_{N \to \infty} W_N \leqslant \frac{3}{m^{1/2}} \qquad \text{a.s.} \tag{4.17}$$

with

$$W_N = \max_{\ell = 1,\dots,\lfloor(\log N)/m\rfloor} \frac{|\gamma_N(\Phi^{-1}(j/m + \ell/\log N)) - \gamma_N(\Phi^{-1}(j/m))|}{\sqrt{\log\log N}}$$

Obviously (4.17) and (4.16) give (4.15). To prove (4.17), we again take the subsequence $\lceil e^{k^\eta} \rceil$, but since the upper bound $3/m^{1/2}$ is not sharp we can now take $\eta = \frac{1}{4}$. We again apply Lemma 5 for large k:

$$\begin{aligned} &P\left(W_{N_k} \geqslant \frac{3}{m^{1/2}}\right) \\ &\quad \leqslant C_\delta m \exp\left(\frac{-(1-2\delta)\, 9 \log\log N_k}{2m \cdot (1/m)}\, \psi\left(\frac{3\sqrt{\log\log N_k}}{\sqrt{m}\,\sqrt{N_k} \cdot (1/2m)}\right)\right) \\ &\quad \leqslant C_\delta m \exp\left(\frac{-(1-3\delta)\, 9 \cdot (1/4) \log k}{2}\right) \end{aligned} \tag{4.18}$$

which is clearly summable in k for δ small enough, i.e., we have shown (4.17) along the subsequence N_k.

So finally let us prove that

$$\lim_{k \to \infty} U_k = 0 \qquad \text{a.s.} \tag{4.19}$$

with

$$U_k = \max_{N_k < N < N_{k+1}} \max_{\ell = 0, 1,\dots,\lfloor(\log N)/m\rfloor} \frac{\left|\sqrt{\frac{N}{N_k}}\, \gamma_N\left(\Phi^{-1}\left(\frac{j}{m} + \frac{\ell}{\log N_k}\right)\right) - \gamma_{N_k}\left(\Phi^{-1}\left(\frac{j}{m} + \frac{\ell}{\log N_k}\right)\right)\right|}{\sqrt{\log\log N_k}}$$

(The change from $\log N$ to $\log N_k$ in γ_N is negligible, cf. (4.16).) Note that (4.19) is similar to (4.4), but with the difference that there is now also the maximum over ℓ. On the other hand the subsequence N_k is tending to infinity much slower now, which means that we can simply take both maxima out of the probabilities involved. We again split U_k into two terms, as in (4.5), and call them $U_{k,1}$ and $U_{k,2}$, respectively. Then, by Lemma 4, we have for $\varepsilon > 0$ and large k

$$
\begin{aligned}
P(|U_{k,1}| \geqslant \varepsilon) \\
&\leqslant (N_{k+1} - N_k)(\log N_k) \\
&\quad \cdot \max_{N_k < N < N_{k+1}} \sup_{x \in \mathbb{R}} P\left(\left| \sum_{i=N_k+1}^{N} 1_{(-\infty, x]}(Z_{i, \lfloor N^\alpha \rfloor}) - (N - N_k)\, \Phi_{\lfloor N^\alpha \rfloor}(x) \right| \right. \\
&\qquad \left. \geqslant \varepsilon \sqrt{\frac{N_k \log\log N_k}{N_{k+1} - N_k}} \right) \\
&\leqslant 2(N_{k+1} - N_k)(\log N_k) \exp\left(\frac{-\varepsilon^2 N_k \log\log N_k}{N_{k+1} - N_k} \right) \\
&= O\left(\frac{1}{k^{3/4}} e^{k^{1/4}} (\log k)\, e^{-k^{3/4}} \right) \qquad (4.20)
\end{aligned}
$$

which is summable in k.

Similarly we have for large k, using the bound $c_2(b/N_k)^{1/3}$ for the left-hand-side of (4.8)

$$
\begin{aligned}
P(|U_{k,2}| \geqslant \varepsilon) &\leqslant (N_{k+1} - N_k)(\log N_k) \\
&\quad \cdot \max_{N_k < N < N_{k+1}} \sup_{x \in \mathbb{R}} P\left(\frac{1}{\sqrt{N_k}} \left| \sum_{i=1}^{N_k} 1_{(-\infty, x]}(Z_{i, \lfloor N^\alpha \rfloor}) \right. \right. \\
&\quad - \sum_{i=1}^{N_k} 1_{(-\infty, x]}(Z_{i, \lfloor N_k^\alpha \rfloor}) \\
&\quad \left. \left. - N_k(\Phi_{\lfloor N^\alpha \rfloor}(x) - \Phi_{\lfloor N_k^\alpha \rfloor}(x)) \right| \geqslant \varepsilon \sqrt{\log\log N_k} \right) \\
&\leqslant 2(N_{k+1} - N_k)(\log N_k) \exp(-\varepsilon^2 c_3 (\log\log N_k)\, k^{1/4}) \\
&= O\left(\frac{1}{k^{3/4}} e^{k^{1/4}} (\log k)\, e^{-c_4 k^{1/4} \log k} \right) \qquad (4.21)
\end{aligned}
$$

for some constants c_3 and c_4. The last expression in (4.21) in clearly summable in k. Since $\varepsilon > 0$ is arbitrary, this in conjunction with (4.20) proves (4.19).

Recall that (4.19) and (4.18) yield (4.17), and (4.17) and (4.16), in turn, establish (4.15). Finally combining (4.14) and (4.15) completes the proof of the theorem. □

ACKNOWLEDGMENTS

We thank the referee for carefully reading the manuscript and for some useful suggestions. We are grateful to Frits H. Ruymgaart for bringing the two of us together to work on this project and for helpful comments. The research of John Einmahl was supported in part by a Fulbright Grant, European Union HCM Grant ERB CHRX-CT 940693, and a grant from the Netherlands Organization for Scientific Research (NWO). These grants enabled him to visit the Departments of Statistics at the University of Florida and Florida State University. He thanks both departments for their kind hospitality during the writing of this article.

REFERENCES

1. Deheuvels, P., and Steinebach, J. (1996). Strong laws for extreme values of sequences of partial sum processes. *Studia Sci. Math. Hungar.* **31**, 121–144.
2. Deheuvels, P., and Teicher, H. (1996). A new law of the iterated logarithm for arrays. *Studia Sci. Math. Hungar.* **31**, 145–185.
3. Einmahl, J. H. J. (1987). *Multivariate Empirical Processes*, CWI Tract, Vol. 32, Amsterdam.
4. Einmahl, J. H. J., and Ruymgaart, F. H. (1987). The almost sure behavior of the oscillation modulus of the multivariate empirical process. *Statist. Probab. Lett.* **6**, 87–96.
5. Finkelstein, H. (1971). The law of the iterated logarithm for empirical distributions. *Ann. Math. Statist.* **42**, 607–615.
6. Freedman, D. A. (1977). A central limit theorem for empirical histograms. *Z. Wahrsch. Verw. Gebiete* **41**, 1–11.
7. Hoeffding, W. (1963). Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58**, 13–30.
8. Hu, T. C., and Weber, N. C. (1992). On the rate of convergence in the strong law of large numbers for arrays. *Bull. Austral. Math. Soc.* **45**, 279–282.
9. Li, D., Huang, M. L., and Rosalsky, A. (2000). Strong invariance principles for arrays. *Bull. Inst. Math. Acad. Sinica* **28**, 167–181.
10. Li, D., Rao, M. B., and Tomkins, R. J. (1995). A strong law for *B*-valued arrays. *Proc. Amer. Math. Soc.* **123**, 3205–3212.
11. Qi, Y. C. (1994). On strong convergence of arrays. *Bull. Austral. Math. Soc.* **50**, 219–223.
12. Rosalsky, A. (1983). Some central limit, Glivenko–Cantelli type theorems for empirical distribution functions of sample means. *Calcutta Statist. Assoc. Bull.* **32**, 9–21.
13. Serfling, R. J. (1970). Moment inequalities for the maximum cumulative sum. *Ann. Math. Statist.* **41**, 1227–1234.
14. Serfling, R. J. (1970). Convergence properties of S_n under moment restrictions. *Ann. Math. Statist.* **41**, 1235–1248.
15. Shorack, G. R., and Wellner, J. A. (1986). *Empirical Processes with Applications to Statistics*, John Wiley, New York.
16. Strassen, V. (1964). An invariance principle for the law of the iterated logarithm. *Z. Wahrsch. Verw. Gebiete* **3**, 211–226.
17. van Zuijlen, M. C. A. (1977). *Empirical Distributions and Rank Statistics*, Mathematical Centre Tracts, Vol. 79, CWI, Amsterdam.
18. Vervaat, W. (1972). Functional central limit theorems for processes with positive drift and their inverses. *Z. Wahrsch. Verw. Gebiete* **23**, 245–253.