The Annals of Probability
1990, Vol. 18, No. 4, 1693–1722

NONSTANDARD FUNCTIONAL LAWS OF THE ITERATED LOGARITHM FOR TAIL EMPIRICAL AND QUANTILE PROCESSES

BY PAUL DEHEUVELS AND DAVID M. MASON[1]

Université Paris VI and University of Delaware

Let $\{\alpha_n(t),\ 0 \le t \le 1\}$ and $\{\beta_n(s),\ 0 \le s \le 1\}$ denote the uniform empirical and quantile processes. We show that, for suitable sequences $A(n, \kappa_n)$ and $B(n, l_n)$, the tail empirical process $\{A(n, \kappa_n)\alpha_n(n^{-1}\kappa_n t),\ 0 \le t \le 1\}$ and the tail quantile process $\{B(n, l_n)\beta_n(n^{-1}l_n s),\ 0 \le s \le 1\}$ are almost surely relatively compact in appropriate topological spaces, where $0 \le \kappa_n \le n$ and $0 \le l_n \le n$ are sequences such that κ_n and l_n are $O(\log\log n)$ as $n \to \infty$. The limit sets of functions are defined through integral conditions and differ from the usual Strassen set obtained when κ_n and l_n are $\infty(\log\log n)$ as $n \to \infty$. Our results enable us to describe the strong limiting behavior of classical statistics based on the top extreme order statistics of a sample or on the empirical distribution function considered in the tails.

1. Introduction. Let $U_1, U_2, \dots$ be a sequence of independent random variables with a uniform distribution on $(0,1)$. Denote by $U_n(t) = n^{-1}\#\{U_i \le t\colon 1 \le i \le n\}$ for $-\infty < t < \infty$ the right-continuous empirical distribution function [df], and by $V_n(s) = \inf\{t \ge 0\colon U_n(t) \ge s\}$ for $0 < s \le 1$, with $V_n(0) = 0$, the left-continuous empirical quantile function [qf] based on the first n of these random variables. Let $\alpha_n(t) = n^{1/2}(U_n(t) - t)$ for $0 \le t \le 1$ be the *uniform empirical process* and let $\beta_n(s) = n^{1/2}(V_n(s) - s)$ for $0 \le s \le 1$ be the *uniform quantile process*.

Let $0 < \kappa_n \le n$, $n = 1, 2, \dots,$ be a sequence of real numbers. The aim of this paper is to obtain functional strong limit laws for $\{A(n, \kappa_n)\alpha_n(n^{-1}\kappa_n t),\ 0 \le t \le 1\}$ and $\{B(n, \kappa_n)\beta_n(n^{-1}\kappa_n s),\ 0 \le s \le 1\}$, where $A(n, \kappa_n)$ and $B(n, \kappa_n)$ are appropriate norming sequences.

The best known result of this kind has been obtained in 1971 by Finkelstein for $\kappa_n = n$ [see, e.g., Shorack and Wellner (1986), page 513] and is stated in Theorem A below. Denote by $(B(0,1), U)$ the set $B(0,1)$ of all bounded functions on $[0,1]$, endowed with the topology U of uniform convergence on $[0,1]$ (see Section 2.1 in the sequel).

THEOREM A. *The sequences* $\{(2\log\log n)^{-1/2}\alpha_n\}$ *and* $\{(2\log\log n)^{-1/2}\beta_n\}$ *are almost surely relatively compact in* $(B(0,1), U)$ *with set of limit points equal to* $\mathbb{S}_{0,1}$, *where* $\mathbb{S}_{0,1}$ *consists of all absolutely continuous functions* f *on*

Received January 1989; revised August 1989.
[1]Research partially supported by NSF Grant DMS-88-03209.
AMS 1980 *subject classifications*. Primary 60F15, 60F05, 62G30; secondary 60F17.
Key words and phrases. Functional laws of the iterated logarithm, empirical and quantile processes, order statistics, extreme values, large deviations, strong laws.

[0, 1] *such that*

$$f(0) = f(1) = 0 \quad \textit{and} \quad \int_0^1 (\dot{f}(u))^2 \, du \le 1, \tag{1.1}$$

where $\dot{f}$ *denotes the Lebesgue derivative of* f.

A version of Theorem A for the *tail empirical and quantile processes* has recently been obtained by Mason (1988) for α_n and by Einmahl and Mason (1988) for β_n. Their results are stated in Theorem B.

THEOREM B. *Assume that* $\{\kappa_n, n \ge 1\}$ *satisfies the conditions*

$$0 < \kappa_n \le n, \quad \kappa_n \uparrow, \quad n^{-1}\kappa_n \downarrow 0 \quad \textit{and} \quad \kappa_n/\log\log n \to \infty \quad \textit{as } n \to \infty. \tag{1.2}$$

Then the sequences of tail processes $\{(2n^{-1}\kappa_n \log\log n)^{-1/2}\alpha_n(n^{-1}\kappa_n t), 0 \le t \le 1\}$ *and* $\{(2n^{-1}\kappa_n \log\log n)^{-1/2}\beta_n(n^{-1}\kappa_n s), 0 \le s \le 1\}$ *are almost surely relatively compact in* $(B(0,1), U)$ *with set of limit points equal to* $\mathbb{S}_0$, *where* $\mathbb{S}_0$ *consists of all absolutely continuous functions* f *on* [0, 1] *such that*

$$f(0) = 0 \quad \textit{and} \quad \int_0^1 (\dot{f}(u))^2 \, du \le 1, \tag{1.3}$$

where $\dot{f}$ *denotes the Lebesgue derivative of* f.

Notice that $\mathbb{S}_0$ is the so-called *Strassen set*, named after the famous functional law of the iterated logarithm for partial sums due to Strassen (1964). A simple analysis shows that the functions $f \in \mathbb{S}_0$ which maximize $|f(1)|$ are $f(u) = u$ and $f(u) = -u$ for $0 \le u \le 1$. Therefore we see that, whenever the conclusion of Theorem B holds, we have

$$\limsup_{n\to\infty} \pm (2\kappa_n \log\log n)^{-1/2}\big(nU_n(n^{-1}\kappa_n) - \kappa_n\big) = 1 \quad \text{a.s.} \tag{1.4}$$

and

$$\limsup_{n\to\infty} \pm (2\kappa_n \log\log n)^{-1/2}\big(nV_n(n^{-1}\kappa_n) - \kappa_n\big) = 1 \quad \text{a.s.} \tag{1.5}$$

By a result due to Kiefer (1972), the limits in (1.4) and (1.5) are no longer true when $\kappa_n = O(\log\log n)$, so Theorem B cannot be extended to this case. His result is stated in Theorem C below. Introduce the functions

$$\begin{aligned} h(x) &= x\log x - x + 1 \quad \text{for } 0 \le x < \infty, \\ h(x) &= \infty \quad \text{for } x < 0, \end{aligned} \tag{1.6}$$

with the convention that $0 \log 0 = 0$, and

$$\begin{aligned} l(x) &= x - 1 - \log x \quad \text{for } 0 < x < \infty, \\ l(x) &= \infty \quad \text{for } x \le 0. \end{aligned} \tag{1.7}$$

Notice that the function $\exp(-h)$ is the Chernoff function of the Poisson random variable with mean 1 and that $\exp(-l)$ is the Chernoff function of the exponential random variable with mean 1 [see, e.g., Shorack and Wellner (1986), pages 416 and 856].

For any $0 < c < \infty$, denote by $0 \le \delta_c^- < 1 < \delta_c^+ < \infty$ the roots of the equation (in δ) $h(\delta) = 1/c$, with the convention that $\delta_c^- = 0$ for $0 < c < 1$. Likewise, for any $0 < c < \infty$, denote by $0 < \gamma_c^- < 1 < \gamma_c^+ < \infty$ the roots of the equation (in γ) $l(\gamma) = 1/c$.

THEOREM C. *Assume that* $\{\kappa_n,\ n \ge 1\}$ *satisfies the conditions*

$$0 < \kappa_n \le n \quad \text{and} \quad \kappa_n/\log\log n \to c \in (0,\infty) \quad \text{as } n \to \infty. \tag{1.8}$$

Then almost surely

$$\begin{aligned} \limsup_{n\to\infty} \pm \frac{nU_n(n^{-1}\kappa_n)}{\log\log n} &= \pm c\delta_c^{\pm}, \\ \limsup_{n\to\infty} \pm \frac{nV_n(n^{-1}\kappa_n)}{\log\log n} &= \pm c\gamma_c^{\pm}. \end{aligned} \tag{1.9}$$

In view of (1.4), (1.5) and Theorem B, Theorem C leads to the conjecture that the sequences $\{A(n,\kappa_n)\alpha_n(n^{-1}\kappa_n t),\ 0 \le t \le 1\}$ and $\{B(n,\kappa_n)\beta_n(n^{-1}\kappa_n s),\ 0 \le s \le 1\}$ might be relatively compact in some topological space of functions defined on $[0,1]$ when $\kappa_n/\log\log n \to c \in (0,\infty)$ and for the choices of $A(n,\kappa_n)$ and $B(n,\kappa_n)$ given by

$$A(n,\kappa_n) = B(n,\kappa_n) = n^{1/2}/\log\log n. \tag{1.10}$$

In Section 2, we will prove that this conjecture is true, and give explicit descriptions of the limit sets. Our main tools will be the functional large deviation theorem of Varadhan (1966) and its extension obtained recently by Lynch and Sethuraman (1987). Our results can be considered as *nonstandard laws of the iterated logarithm* [LIL] for tail processes, since the limiting sets of functions differ from the Strassen set $\mathbb{S}_0$.

In order to complete our description of functional LILs for tail processes, we must consider the remaining case where $\kappa_n/\log\log n \to 0$ as $n \to \infty$. Here the analog of Theorem C is as follows.

THEOREM D. (a) *Assume that* $\{\kappa_n,\ n \ge 1\}$ *satisfies the conditions*

$$0 < \kappa_n \le n \quad \text{and} \quad \kappa_n/\log\log n \to 0 \quad \text{as } n \to \infty. \tag{1.11}$$

Then we have

$$P\big(nU_n(n^{-1}\kappa_n) = 0 \text{ i.o.}\big) = 1 \tag{1.12}$$

and

$$\limsup_{n\to\infty} \frac{nV_n(n^{-1}\kappa_n)}{\log\log n} = 1 \quad a.s. \tag{1.13}$$

(b) *Assume, in addition to* (1.11), *that* $\{\kappa_n,\ n \geq 1\}$ *satisfies*:

(1.14) *There exists a sequence* $\{\chi_n,\ n \geq 1\}$ *such that* $\chi_n \downarrow 0$ *and* $(\chi_n \log\log n)/\log(\kappa_n^{-1}\log\log n) \to 1$ *as* $n \to \infty$.

Then we have

$$(1.15)\qquad \limsup_{n\to\infty} \frac{nU_n(n^{-1}\kappa_n)}{\log\log n}\log\left(\frac{\log\log n}{\kappa_n}\right) = 1 \quad a.s.$$

(c) *Assume, in addition to* (1.11), *that* $\{\kappa_n,\ n \geq 1\}$ *is such that*

$$(1.16)\qquad k_n = \lceil \kappa_n \rceil \uparrow,$$

where $\lceil u \rceil \geq u > \lceil u \rceil - 1$ *denotes the smallest integer greater than or equal to* u. *Then we have*

$$(1.17)\quad P\left(nV_n(n^{-1}\kappa_n) \leq k_n \exp\left(-(1+\varepsilon)k_n^{-1}\log\log n\right)\ i.o.\right) = 0 \text{ or } 1,$$

according as $\varepsilon > 0$ *or* $\varepsilon < 0$.

REMARK 1.1. Theorem D is essentially due to Kiefer (1972). However, Kiefer (1972) makes use of the additional regularity condition that $n^{-1}\kappa_n \downarrow 0$ for (1.15) and (1.17). Deheuvels and Mason (1988) have proved that this last condition is superfluous for (1.17). We will adapt their result to show that this extra condition is also not needed for (1.15) (see Proposition 3.1 below).

REMARK 1.2. Denote by $0 \leq U_{1,n} \leq \cdots \leq U_{n,n} \leq 1$ the order statistics of $U_1, \dots, U_n$. Notice that $V_n(n^{-1}\kappa_n) = U_{k_n,n}$, where $k_n = \lceil \kappa_n \rceil$. Moreover, (1.17) does not hold with k_n replaced by κ_n if $\kappa_n = O(1)$ as $n \to \infty$. On the other hand, this replacement is permissible when $\kappa_n \uparrow \infty$ as $n \to \infty$ and (1.11) holds.

Motivated by (1.13) and (1.15), we will show in Section 3 that, whenever $\kappa_n/\log\log n \to 0$ as $n \to \infty$ (under additional regularity conditions to be stated later on), the sequences $\{A(n,\kappa_n)\alpha_n(n^{-1}\kappa_n t),\ 0 \leq t \leq 1\}$ and $\{B(n,\kappa_n)\beta_n(n^{-1}\kappa_n s),\ 0 \leq s \leq 1\}$ are relatively compact in appropriate topological spaces, with $A(n,\kappa_n)$ and $B(n,\kappa_n)$ given by

$$(1.18)\quad A(n,\kappa_n) = \frac{n^{1/2}}{\log\log n}\log\left(\frac{\log\log n}{\kappa_n}\right) \quad\text{and}\quad B(n,\kappa_n) = \frac{n^{1/2}}{\log\log n}.$$

Here also, a complete description of the limit sets is given. In Section 4, we describe what happens in the extreme case when κ_n is a fixed positive integer.

It turns out that these nonstandard functional LILs, in combination with Theorem B, yield strong laws for practically all known statistics based either on the top (or on the bottom) $k_n = \lceil \kappa_n \rceil$ order statistics of a sample when $n^{-1}\kappa_n \to 0$, or on the empirical distribution function considered in the tails. For a general sequence $X_1, X_2, \dots,$ of independent random variables with common df F this is accomplished via a simple quantile transform. A thorough exposition of such applications is provided in Deheuvels and Mason (1990). In

Section 5, we give a simple example in order to briefly indicate our general approach.

2. The intermediate case. This section is devoted to the study of functional LILs for the *intermediate case when* $\kappa_n/\log\log n \to c \in (0, \infty)$ as $n \to \infty$. We start with some technical results and conventions which are needed to present our results.

2.1. *Preliminary results and notation.* In the following, we consider the restriction to $[0,1]$ of functions f defined on open neighborhoods of this interval. In order to keep track of the behavior of these functions on the left of 0 and on the right of 1, it will be useful to define $f(0-) = \limsup_{\varepsilon\downarrow 0} f(-\varepsilon)$ and $f(1+) = \liminf_{\varepsilon\downarrow 0} f(1+\varepsilon)$, and to use throughout the convention that the functions we consider are defined on $[0-, 1+] = [0,1] \cup \{0-\} \cup \{1+\}$. Let $B(0,1)$ denote the functions induced on $[0-, 1+]$ by the bounded functions on $(-\infty, \infty)$. Likewise let $D_{\mathrm{LC}}(0,1)$, $D(0,1)$ and $C(0,1)$ be the functions induced on $[0-, 1+]$ by the left-continuous functions with right-hand limits, the right-continuous functions with left-hand limits and the continuous functions on $(-\infty, \infty)$, respectively. Denote by $\mathrm{AC}(0,1)$ the space of all functions induced on $[0-, 1+]$ by df's of Radon measures on $[0,1]$ which are absolutely continuous with respect to Lebesgue measure. In other words, $f \in \mathrm{AC}(0,1)$ is of the form

$$(2.1)\qquad \begin{aligned} f(x) &= \int_0^x \dot{f}(t)\,dt \quad \text{for } 0 < x \le 1, \\ f(0-) &= f(0) = 0 \quad \text{and} \quad f(1+) = f(1). \end{aligned}$$

Let $M^+(0,1)$ be the set of all nonnegative bounded measures on $[0,1]$. For any $\mu \in M^+(0,1)$, set

$$(2.2)\qquad \begin{aligned} f_\mu^+(x) &= \mu([0,x)) \quad \text{for } 0 < x \le 1, \\ f_\mu^+(0-) &= f_\mu^+(0) = 0 \quad \text{and} \quad f_\mu^+(1+) = \mu([0,1]) \end{aligned}$$

and

$$(2.3)\qquad \begin{aligned} f_\mu^-(x) &= \mu([0,x]) \quad \text{for } 0 \le x \le 1, \\ f_\mu^-(0-) &= 0 \quad \text{and} \quad f_\mu^-(1+) = f_\mu^-(1) = \mu([0,1]). \end{aligned}$$

Set $I(0,1)$ [resp. $I_{\mathrm{RC}}(0,1)$] to be the space formed by f_μ^+ (resp. f_μ^-) when μ varies in $M^+(0,1)$. Notice that $\{f_\mu^+(x),\ 0 \le x \le 1\}$ does not completely determine μ if the value of $f^+(1+)$ is unknown.

For any $\mu \in M^+(0,1)$, let $\mu = \mu_{\mathrm{AC}} + \mu_S$ be the corresponding Lebesgue decomposition of μ, where $\mu_{\mathrm{AC}} \ll dx$ and $\mu_S \perp dx$ (here dx is Lebesgue measure). For any $f = f_\mu^+ \in I(0,1)$ [resp. $f = f_\mu^- \in I_{\mathrm{RC}}(0,1)$], set $\dot{f} = \dot{f}_{\mu_{\mathrm{AC}}}^{\pm}$ and $f_S = f_{\mu_S}^+$ (resp. $f_S = f_{\mu_S}^-$) so that $0 \le \dot{f} < \infty$ a.e., and

$$(2.4)\qquad f(x) = \int_0^x \dot{f}(t)\,dt + f_S(x) \quad \text{for } 0- \le x \le 1+\,.$$

Let AC $I(0,1) = I(0,1) \cap \mathrm{AC}(0,1) = \{f \in I(0,1)\colon f_S \equiv 0\}$.

We endow the above spaces of functions on $[0-, 1+]$ by the topologies induced by the following topologies for sets of functions defined on $(-\infty, \infty)$ or on an interval containing $(-\varepsilon, 1+\varepsilon)$ for some $\varepsilon > 0$. Denote by U the topology of uniform convergence, by J_1 [following the notation originally introduced by Skorohod (1956); see, e.g., Billingsley (1968), page 111] the classical Skorohod topology, and by W the topology induced on $I(0,1)$ or $I_{\mathrm{RC}}(0,1)$ by the weak (abbreviated from weak*) convergence of measures on $M^+(0,1)$. A space $\mathbb{F}$ endowed with a topology $\mathscr{T}$ will be denoted by $(\mathbb{F}, \mathscr{T})$.

Note here that the definition of $(B(0,1), U)$ so obtained is slightly different from that given in the Introduction, and that the statements of Theorems A and B are valid with the new definition [here and in the sequel, we define $U_n(u) = V_n(u) = 0$ for $u < 0$ and $U_n(u) = V_n(u) = 1$ for $u > 1$].

In the sequel, we will consider possibly infinite valued functions $\{\Phi(x), -\infty < x < \infty\}$ such that the following assumptions hold:

(C.1) Φ is convex and nonnegative;
(C.2) $\Phi < \infty$ on some nondegenerate interval;
(C.3) $0 < C'_\Phi = \lim_{x\to\infty}(\Phi(x)/x) \le \infty$ and $0 < C''_\Phi = \lim_{x\to-\infty}(\Phi(x)/|x|) \le \infty$.

Note for further use that if Φ satisfies (C.1)–(C.3), then the same holds for the function $v\Phi(\cdot/v)$ for any fixed $v > 0$. Moreover, the constants in (C.3) are independent of v.

For any Φ satisfying (C.1)–(C.3), $f \in I(0,1)$ having the decomposition in (2.4), and $0 < v < \infty$, set

$$J_{\Phi,v}(f) = v\int_0^1 \Phi(\dot f(u)/v)\,du + C'_\Phi f_S(1+). \tag{2.5}$$

Obviously, $J_{\Phi,v}(f)$ does not depend upon the representative of $\dot f$, and defines a mapping of $I(0,1)$ into $[0,\infty]$. In the particular case when $C'_\Phi = \infty$, we see that $J_{\Phi,v}(f) < \infty$ if and only if $f \in$ AC $I(0,1)$, in which case, we have

$$J_{\Phi,v}(f) = v\int_0^1 \Phi(\dot f(u)/v)\,du. \tag{2.6}$$

One could give an analogous definition of $J_{\Phi,v}(f)$ for f nonincreasing, by replacing C'_Φ by C''_Φ in (2.5). The main point of these extensions is that they coincide with that given in (2.6) in the case where $C'_\Phi = C''_\Phi = \infty$. If this last condition holds, it can be seen that $J_{\Phi,v}$ can be extended as a mapping of $C(0,1)$ into $[0,\infty]$ by defining $J_{\Phi,v}(f)$ as in (2.6) when $f - f(0) \in \mathrm{AC}(0,1)$, and by letting

$$J_{\Phi,v}(f) = \infty \quad \text{for } f - f(0) \in C(0,1) - \mathrm{AC}(0,1). \tag{2.7}$$

The following important result, due to Varadhan (1966), pages 262, 263 and 272, shows the interest of the definitions above.

LEMMA 2.1. *Let Φ satisfy* (C.1)–(C.3) *with $C'_\Phi = C''_\Phi = \infty$, and let $0 < v < \infty$ be fixed. Then $J_{\Phi,v}$ defined by* (2.6) *and* (2.7) *is a lower-semicontinuous* *mapping of* $(C(0,1), U)$ *into* $[0, \infty]$. *Moreover, for any* $c < \infty$, *the set* $\{f \in C(0,1)\colon f(0) = 0$ *and* $J_{\Phi, v}(f) \le c\}$ *is a compact subset of* $(\mathrm{AC}(0,1), U)$.

It turns out that the assumption that $C'_\Phi = C''_\Phi = \infty$ is crucial for the validity of Lemma 2.1 as shown by Lynch and Sethuraman (1987). They also proved the following remarkable new version of Lemma 2.1.

LEMMA 2.2. *Let* Φ *satisfy* (C.1)–(C.3) *and let* $0 < v < \infty$ *be fixed. Then* $J_{\Phi, v}$ *defined by* (2.5) *is a lower-semicontinuous mapping of* $(I(0,1), W)$ *into* $[0, \infty]$. *Moreover, for any* $c < \infty$, *the set* $\{f \in I(0,1)\colon J_{\Phi, v}(f) \le c\}$ *is a compact subset of* $(I(0,1), W)$.

Recall by (2.2) that any $f \in I(0,1)$ satisfies $f(0) = 0$, so that the statement of Lemma 2.2 is similar to that of Lemma 2.1, with the replacement of the uniform topology by the weak topology.

In applying Lemmas 2.1 and 2.2, we shall consider the following three examples.

EXAMPLE 2.1. Let $\Phi(x) = x^2$. We see that for any $0 < v < \infty$, $v\Phi(\cdot/v) = \Phi(\cdot)$, and that $C'_\Phi = C''_\Phi = \infty$, so that Lemma 2.1 may be applied. An application of this lemma shows that the Strassen set $\mathbb{S}_0 = \{f \in C(0,1)\colon f(0) = 0$ and $J_{\Phi,1}(f) \le 1\}$ used in Theorem B is a compact subset of $(C(0,1), U)$. The same holds naturally for the *Finkelstein set* $\mathbb{S}_{0,1} = \{f \in \mathbb{S}_0\colon f(1) = 0\}$. Moreover, both $\mathbb{S}_0$ and $\mathbb{S}_{0,1}$ are subsets of $\mathrm{AC}(0,1)$.

Notice that any compact subset of $(C(0,1), U)$ is a compact subset of $(D(0,1), J_1)$ and of $(B(0,1), U)$. This explains the formulation given in Theorem A, noting that Finkelstein's (1971) theorem is usually formulated in $(D(0,1), J_1)$ or $(D_{\mathrm{LC}}(0,1), J_1)$.

EXAMPLE 2.2. Let $\Phi(x) = h(x)$, where $h(x)$ is as in (1.6). Here again, $C'_\Phi = C''_\Phi = \infty$, so that we may apply Lemma 2.1 which shows that, for any $0 < v < \infty$, the set

$$\Delta_v = \{f \in I(0,1)\colon J_{h,v}(f) \le 1\} \tag{2.8}$$

is a compact subset of $(C(0,1), U)$ and is included in AC $I(0,1)$. Here we have made use of the fact that the set $I(0,1) \cap C(0,1)$ is closed in $(C(0,1), U)$.

EXAMPLE 2.3. Let $\Phi(x) = l(x)$, where $l(x)$ is as in (1.7). We now have $C'_\Phi = 1$ and $C''_\Phi = \infty$, so that we may only apply Lemma 2.2 to show that, for any $0 < w < \infty$, the set

$$\Gamma_w = \{g \in I(0,1)\colon J_{l,w}(g) \le 1\} \tag{2.9}$$

is relatively compact in $(I(0,1), W)$.

In this particular case, we may use (2.5) to show that if $g \in I(0,1)$ is decomposed as f in (2.4), we have

$$J_{l,w}(g) = g(1+) - w - w\int_0^1 \log(\dot{g}(u)/w)\,du, \tag{2.10}$$

which has the advantage of not using explicitly the singular component g_S of g.

A similar simplified expression may be given for $J_{h,v}(f)$ for $f \in \mathrm{AC}\, I(0,1)$, i.e.,

$$J_{h,v}(f) = v - f(1) + \int_0^1 \dot{f}(u)\log(\dot{f}(u)/v)\,du. \tag{2.11}$$

We will conclude this section by showing that the sets Δ_v [defined in (2.8)] and Γ_w [defined in (2.9)] are closely related via simple transformations. Toward this end, define for any $v > 0$ and $w > 0$ the sets

$$\Delta_{v,w} = \{f \in \Delta_v: f(1) = w\} \quad \text{and} \quad \Gamma_{w,v} = \{g \in \Gamma_w: g(1+) = v\}. \tag{2.12}$$

LEMMA 2.3. *For any $0 < v < \infty$, the set $\Delta_{v,w}$ is nonvoid if and only if*

$$v\delta_v^- \le w \le v\delta_v^+, \tag{2.13}$$

where $0 \le \delta_v^- \le 1 \le \delta_v^+$ are as in Theorem C. *Likewise, for any $0 < w < \infty$, the set $\Gamma_{w,v}$ is nonvoid if and only if*

$$w\gamma_w^- \le v \le w\gamma_w^+, \tag{2.14}$$

where $0 < \gamma_w^- \le 1 \le \gamma_w^+$ are as in Theorem C.

PROOF. Consider the case where $\dot{f} = w \ge 0$ is constant. In this case, $f(1) = w$, and $f \in \Delta_v$ if and only if $vh(w/v) \le 1$, which in turn is equivalent to $v\delta_v^- \le w \le v\delta_v^+$. By a similar argument used for the function $l(\cdot)$, we obtain the "if" part of both statements (2.13) and (2.14). The "only if" part is a simple consequence of Jensen's inequality. For instance, since $h(\cdot)$ is convex, we have

$$\begin{aligned} vh\left(\int_0^1 \dot{f}(u)v^{-1}\,du\right) &= vh(f(1)/v) \le v\int_0^1 h(\dot{f}(u)v^{-1})\,du \\ &= J_{h,v}(f) \le 1. \end{aligned} \tag{2.15}$$

A similar argument for $l(\cdot)$ completes the proof of the lemma. $\square$

For any $f \in I_{\mathrm{RC}}(0,1)$ such that $w = f(1) > 0$ [recall that $f(0-) = 0$], define the corresponding *rescaled inverse* $\overleftarrow{f} \in I(0,1)$ by

$$\begin{aligned} &\overleftarrow{f}(s) = \inf\{t: 0 \le t \le 1,\ f(t) \ge sf(1)\} \quad \text{for } 0 \le s \le 1, \\ &\overleftarrow{f}(1+) = 1. \end{aligned} \tag{2.16}$$

Likewise, for any $g \in I(0,1)$ such that $v = g(1+) > 0$ [recall that $g(0) = 0$], define the corresponding *rescaled inverse* $\overrightarrow{g} \in I_{\mathrm{RC}}(0,1)$ by

$$\text{(2.17)}\qquad \begin{aligned} &\overrightarrow{g}(t) = \sup\{s\colon 0 \le s \le 1,\ g(s) \le tg(1+)\} \quad \text{for } 0 \le t \le 1,\\ &\overrightarrow{g}(0-) = 0. \end{aligned}$$

Recall that the set of points of discontinuity of a monotone function is at most countable. Hence, for any $f \in I_{\mathrm{RC}}(0,1)$, there exists a set $D \subset [0,1]$ which is dense in $[0,1]$ and such that, for any $t \in D$,

$$\text{(2.18)}\qquad f(1)\big(\overleftarrow{f}\,\big)^{\rightarrow}(t) = f(t).$$

A similar statement holds for $(\overrightarrow{g})^{\leftarrow}$ for $g \in I(0,1)$.

LEMMA 2.4. *For any $v > 0$ and $w > 0$, $\Delta_{v,w}$ is nonvoid if and only if $\Gamma_{w,v}$ is nonvoid. Moreover, if either of the two sets is nonvoid, then the application $f \to v\overleftarrow{f}$ (resp. $g \to w\overrightarrow{g}$) defines a one-to-one mapping of $\Delta_{v,w}$ onto $\Gamma_{w,v}$ (resp. of $\Gamma_{w,v}$ onto $\Delta_{v,w}$).*

PROOF. The fact that $\Delta_{v,w} \ne \varnothing \Leftrightarrow \Gamma_{w,v} \ne \varnothing$ is a direct consequence of Lemma 2.3 and of the following simple identities (extended by continuity for $x = 0$):

$$\text{(2.19)}\qquad \begin{aligned} h(x) &= xl(1/x) \quad \text{for } 0 \le x < \infty,\\ l(y) &= yh(1/y) \quad \text{for } 0 < y < \infty. \end{aligned}$$

By (2.18), we see that $f \to T_{v,w}(f) := v\overleftarrow{f}$ defines a one-to-one mapping of the set $I_{\mathrm{RC},w} = \{f \in I_{\mathrm{RC}}(0,1)\colon f(1) = w\}$ onto the set $I_{\mathrm{LC},v} = \{g \in I(0,1)\colon g(1+) = v\}$. Therefore all we need is to prove that, for $f \in I_{\mathrm{RC},w}$ and $g = T_{v,w}(f)$, $J_{h,v}(f) \le 1 \Leftrightarrow J_{l,w}(g) \le 1$.

In order to clarify the mechanism of our proof, we consider the particular case when $f \in \mathrm{AC}\,I(0,1)$ and $\dot{f} > 0$ on $[0,1]$. We obtain directly the reciprocal relations

$$\text{(2.20)}\qquad \dot{g}(s) = vw/\dot{f}(g(s)/v), \quad s = f(t)/w \quad \text{and} \quad ds = \big(\dot{f}(t)/w\big)\,dt,$$

$$\text{(2.21)}\qquad \dot{f}(t) = vw/\dot{g}(f(t)/w), \quad t = g(s)/v \quad \text{and} \quad dt = (\dot{g}(s)/v)\,ds,$$

so that, by (2.19), (2.20) and (2.21), a simple change of variables yields

$$\text{(2.22)}\qquad J_{h,v}(f) = v\int_0^1 h\big(\dot{f}(t)/v\big)\,dt = w\int_0^1 l(\dot{g}(s)/w)\,ds = J_{l,w}(g).$$

In the general case where $f \in I_{\mathrm{RC}}(0,1)$, we still have the equality $J_{h,v}(f) = J_{l,w}(g)$ [even though $J_{l,w}(g)$ is then given by the general expression in (2.5) which does not always reduce to (2.6)]. The proof of this statement is nothing else but a repetition of the argument leading to (2.22) in a discretized version, used jointly with Theorem 3.2 of Lynch and Sethuraman (1987). By this theorem, if $\mathscr{P} = \{t_0 = 0 < t_1 < \cdots < t_k = 1\}$ denotes a partition of $[0, 1]$, and

$$I_{\mathscr{P}}(f) = \sum_{i=1}^{k} vh\left(\frac{f(t_i) - f(t_{i-1})}{v(t_i - t_{i-1})}\right)(t_i - t_{i-1}), \tag{2.23}$$

then $d(\mathscr{P}) = \max_{1 \le i \le k}(t_i - t_{i-1}) \to 0$ implies that $I_{\mathscr{P}}(f) \to J_{h,v}(f)$. Likewise, if $\mathscr{Q} = \{s_0 = 0 < s_1 < \cdots < s_k = 1\}$, and

$$J_{\mathscr{Q}}(g) = \sum_{i=1}^{k} wl\left(\frac{g(s_i) - g(s_{i-1})}{w(s_i - s_{i-1})}\right)(s_i - s_{i-1}), \tag{2.24}$$

then $d(\mathscr{Q}) \to 0$ implies that $J_{\mathscr{Q}}(g) \to J_{l,w}(g)$.

In view of (2.19) we see that if we set $t_i = \overleftarrow{f}(s_i) = g(s_i)/v$ and $t_{i-1} = \overleftarrow{f}(s_{i-1}) = g(s_{i-1})/v$, the equality

$$vh\left(\frac{f(t_i) - f(t_{i-1})}{v(t_i - t_{i-1})}\right)(t_i - t_{i-1}) = wl\left(\frac{g(s_i) - g(s_{i-1})}{w(s_i - s_{i-1})}\right)(s_i - s_{i-1}) \tag{2.25}$$

always holds if

$$\begin{gathered} t_{i-1} = \overleftarrow{f}(s_{i-1}) < t_i = \overleftarrow{f}(s_i), \\ f\left(\overleftarrow{f}(s_{i-1})\right) = s_{i-1}w, \\ f\left(\overleftarrow{f}(s_i)\right) = s_i w. \end{gathered} \tag{2.26}$$

Observe that if, for some $1 \le i \le k$, $s_{i-1} < s_i$ and $\overleftarrow{f}(s_{i-1}) = \overleftarrow{f}(s_i)$, then $\dot{g} = 0$ on (s_{i-1}, s_i) [which implies that $J_{l,w}(g) = \infty$] and f is discontinuous [which implies that $J_{h,v}(f) = \infty$]. Hence, in this case, $J_{l,w}(g) = J_{h,v}(f)$. If we exclude this possibility, then we may suppose that the inequality in (2.26) is satisfied for all possible choices of $0 \le s_{i-1} < s_i \le 1$. Since we now assume that f is continuous and that $\overleftarrow{f}$ is strictly increasing on $[0, 1]$, we also have $f(\overleftarrow{f}(s)) = s$ for all $s \in [0, 1]$, which implies that (2.26) and (2.25) hold, so that $J_{\mathscr{Q}}(g) = I_{\mathscr{P}}(f)$.

In order to use this last equality to prove that $J_{l,w}(g) = J_{h,v}(f)$, all we need is to choose $\mathscr{Q}$ in such a way that $d(\mathscr{Q}) \to 0$, jointly with $d(\mathscr{P}) \to 0$. Unfortunately, it is in general impossible to achieve $d(\mathscr{P}) \to 0$ because of possible discontinuities of g (or flat stretches in f). In particular, if we choose s_{i-1} and s_i in such a way that

$$\begin{gathered} f(t) < f(t_{i-1}) = f(t_i) < f(t') \quad \text{for} \\ t < t_{i-1} = \overleftarrow{f}(s_{i-1}) < \overleftarrow{f}(s_i) = t_i < t', \end{gathered} \tag{2.27}$$

the choice of $\mathscr{P}$ obtained through (2.26) implies that $d(\mathscr{P}) \ge t_i - t_{i-1}$ for i as in (2.27).

To overcome this difficulty, we add new points between t_{i-1} and t_i, i.e., by letting $t_{i-1} < t^*_{i,1} < \cdots < t^*_{1,k(i)} < t_i$ be added to $\mathscr{P}$ in order to obtain a new partition $\mathscr{P}^*$ of $[0, 1]$. Because of the fact that $h(0) = 1$, it is easily verified that $I_{\mathscr{P}}(f) = I_{\mathscr{P}^*}(f)$. Since we may now choose $\mathscr{Q}$ and $\mathscr{P}^*$ such that $d(\mathscr{Q}) \to 0$, $d(\mathscr{P}^*) \to 0$ and $J_{\mathscr{Q}}(g) = I_{\mathscr{P}^*}(f)$, an application of Theorem 3.2 in Lynch and Sethuraman (1987) shows that also in this case $J_{l,w}(g) = J_{h,v}(f)$. This completes the proof of Lemma 2.4, together with the relations

$$J_{h,v}(f) = J_{l,w}(v\overleftarrow{f}) \quad \text{and} \quad J_{h,v}(w\overrightarrow{g}) = J_{l,w}(g) \qquad (2.28)$$
$$\text{for } f \in \Delta_{v,w} \text{ and } g \in \Gamma_{w,v}. \qquad \square$$

We are now prepared to state the main theorems of this section.

2.2. *Theorems.*

THEOREM 2.1. *Assume that* $\{\kappa_n,\ n \geq 1\}$ *satisfies the conditions*

$$0 \leq \kappa_n \leq n \quad \textit{and} \quad \kappa_n/\log\log n \to v \in (0,\infty) \quad \textit{as } n \to \infty. \qquad (2.29)$$

Then the sequence of functions $\{(\log\log n)^{-1}(nU_n(n^{-1}\kappa_n t)),\ 0- \leq t \leq 1+\}$ *is almost surely relatively compact in* $(I_{\mathrm{RC}}(0,1), U)$, *with set of limit points equal to* Δ_v, *where*

$$\Delta_v = \Big\{ f \in \mathrm{AC}\, I(0,1) \colon f(0) = 0 \textit{ and } \qquad (2.30)$$
$$\int_0^1 \big(\dot f(t)\log(\dot f(t)/v) - \dot f(t) + v\big)\, dt \leq 1 \Big\}.$$

THEOREM 2.2. *Assume that* $\{l_n,\ n \geq 1\}$ *satisfies the conditions*

$$0 \leq l_n \leq n \quad \textit{and} \quad l_n/\log\log n \to w \in (0,\infty) \quad \textit{as } n \to \infty. \qquad (2.31)$$

Then the sequence of functions $\{(\log\log n)^{-1}(nV_n(n^{-1}l_n s)),\ 0- \leq s \leq 1+\}$ *is almost surely relatively compact in* $(I(0,1), W)$, *with set of limit points equal to* Γ_w, *where*

$$\Gamma_w = \Big\{ g \in I(0,1) \colon g(0-) = 0 \textit{ and } \qquad (2.32)$$
$$g(1+) - 1 - w\int_0^1 \log(\dot g(s)/w)\, ds \leq 1 \Big\}.$$

REMARK 2.1. It is obvious from Theorems 2.1 and 2.2 that we may reformulate the statements of these theorems in terms of

$$\Big\{ \big(2n^{-1}\kappa_n \log\log n\big)^{-1/2} \alpha_n\big(n^{-1}\kappa_n t\big),\ 0 \leq t \leq 1 \Big\}$$

and of $\{(2n^{-1}l_n \log\log n)^{-1/2}\beta_n(n^{-1}l_n s),\ 0 \leq s \leq 1\}$ by simple changes of scale. For these sequences, the limit sets become, respectively,

$$\Delta_v^* = \Big\{ f \in \mathrm{AC}(0,1) \colon f(0) = 0,\ \dot f \geq -\sqrt{v/2} \text{ a.e.}, \qquad (2.33)$$
$$\int_0^1 vh\big(1 + \sqrt{2/v}\, \dot f(s)\big)\, ds \leq 1 \Big\},$$

and

$$\Gamma_w^* = \left\{g\colon g + I\sqrt{w/2} \in I(0,1),\, g(0) = 0,\, J_{l,w}\left(I + \sqrt{2/w}\,g\right) \le 1\right\}, \tag{2.34}$$

where $I(s) := s$ is the identical mapping.

Observe that $h(1+u) = (1+u)\log(1+u) - u \le u^2/2$ for $u \ge -1$. Hence we always have the inequality

$$\int_0^1 vh\left(1 + \sqrt{2/v}\,\dot{f}(s)\right) ds \le \int_0^1 \left(\dot{f}(s)\right)^2 ds. \tag{2.35}$$

In view of (2.35), a comparison of (1.3) and (2.33) shows that

$$\left\{f \in \mathbb{S}_0\colon \dot{f} \ge -\sqrt{v/2}\right\} \subset \Delta_v^* \quad \text{for any } v > 0. \tag{2.36}$$

This, jointly with the fact that $h(1+u) \sim u^2/2$ as $u \to 0$, shows that the Strassen set $\mathbb{S}_0$ can be considered as the limit set of Δ_v^* as $v \to \infty$. A similar statement can be made for Γ_w^* as $w \to \infty$.

REMARK 2.2. It is remarkable that in the range covered by Theorems 2.1 and 2.2, the empirical and quantile processes exhibit a different behavior. In the first place, the limit functions corresponding to the empirical process are absolutely continuous. On the other hand, the limit functions corresponding to the quantile function may have discontinuities and are always strictly increasing on [0, 1].

PROOF OF THEOREM 2.2. The proof of Theorem 2.1 will be provided below in Section 2.3. In the following, we limit ourselves to the proof of Theorem 2.2, given the result of Theorem 2.1.

We start with the observation that both sequences

$$\left\{(\log\log n)^{-1}\left(nU_n\left(n^{-1}\kappa_n I\right)\right)\right\} \quad \text{and} \quad \left\{(\log\log n)^{-1}\left(nV_n\left(n^{-1}l_n I\right)\right)\right\}$$

are almost surely relatively compact in $(I_{\mathrm{RC}}(0,1), W)$ and $(I(0,1), W)$, respectively, if $\{\kappa_n,\ n \ge 1\}$ satisfies (2.29) and $\{l_n,\ n \ge 1\}$ satisfies (2.31). This is an obvious consequence of Helly's selection lemma, used jointly with (1.9).

Next we show that, if $\{l_n,\ n \ge 1\}$ satisfies (2.31), the sequence $g_n := (\log\log n)^{-1}(nV_n(n^{-1}l_n I))$ is relatively compact in $(I(0,1), W)$ with set of limit points included in Γ_w. For this, it is enough to show that from any sequence $\{1 \le n_1 < n_2 < \cdots\}$ one can extract a subsequence $\mathscr{S}$ such that $g_n \to g \in \Gamma_w$ along $\mathscr{S}$ (whenever $g_n \to g$ along $\mathscr{S}$ we shall say that g_n is $\mathscr{S}$-convergent to g). Consider therefore g such that g_n is $\mathscr{S}$-convergent to g [in $(I(0,1), W)$]. This is equivalent to having, along $\mathscr{S}$, $g_n(x) \to g(x)$ for all continuity points x of g, and $g_n(1+) \to g(1+)$. By (1.9), $v := g(1+) \in (0,\infty)$. Moreover, if we define a sequence $\{\kappa_n,\ n \ge 1\}$ by

$$\kappa_n = nV_n\left(n^{-1}l_n +\right) := \lim_{\zeta \downarrow 0} nV_n\left(n^{-1}l_n + \zeta\right), \tag{2.37}$$

we see that $\{\kappa_n,\ n \ge 1\}$ satisfies (2.29) along $\mathscr{S}$. Here a difficulty arises, coming from the fact that $\{\kappa_n\}$ is random. If we could directly apply Theorem

2.1 to this case we would obtain the existence of $\mathscr{S}' \subset \mathscr{S}$ such that $f_n := (\log\log n)^{-1}(nU_n(n^{-1}\kappa_n I))$ would $\mathscr{S}'$-converge in $(I_{\mathrm{RC}}(0,1), U)$ to some $f \in \Delta_{v,w}$. Since evidently $V_n = \overleftarrow{U_n}$ and $U_n = \overrightarrow{V_n}$, it would follow that if $g^* = w\overleftarrow{f}$, then g_n would be $\mathscr{S}'$-convergent to g^* in $(I(0,1), W)$. Since then by Lemma 2.4, $g^* \in \Gamma_{w,v} \subset \Gamma_w$, one would have $g = g^* \in \Gamma_w$ as sought.

To overcome the randomness of $\{\kappa_n\}$, we apply Theorem 2.1 to an auxiliary sequence $\kappa_n^* = (v + \varepsilon)\log\log n$, and use the continuity of the functions in $\Delta_{v+\varepsilon}$ (for some small $\varepsilon > 0$) to show that the conclusion of Theorem 2.1 remains valid for $\{\kappa_n\}$. We omit the details since a similar argument is used in Section 2.3 below for the proof of Theorem 2.1.

In order to complete the proof of Theorem 2.2, it remains to show that any $g \in \Gamma_w$ is the $\mathscr{S}$-limit of g_n for some suitable $\mathscr{S}$. The proof of this last statement can be made along the same lines as above, starting from $f = v\vec{g}$ for some arbitrary $v > 0$, then using Theorem 2.1 to exhibit a sequence $\mathscr{S}$ such that $f_n \to f$ along $\mathscr{S}$, then, using this fact to obtain that, along a subsequence $\mathscr{S}'$ of $\mathscr{S}$, $g_n \to g$ as desired. Here again, we make use of Lemma 2.4.

This completes the proof of Theorem 2.2. Note that the above arguments are valid throughout with probability 1. $\square$

2.3. *Proof of Theorem* 2.1. In this section, we give the proof of Theorem 2.1, together with some additional results of independent interest concerning the approximation of the empirical process in the tail by Poisson processes. A rough outline of our argument is as follows. In a first step, we show that we may replace our original process by a suitably chosen Poisson process with a negligible error. In a second step, we prove that the statement of Theorem 2.1 holds for this auxiliary Poisson process.

The following proposition provides us with the desired Poisson approximation.

PROPOSITION 2.1. *For any sequence* $\{t_n,\ n \geq 1\}$ *such that*

$$0 < t_n \leq n, \quad n^{-1}t_n \downarrow \quad \text{and} \quad \sum_n n^{-2}t_n^2 < \infty, \tag{2.38}$$

it is possible to define $\{U_n(\cdot),\ n \geq 1\}$ *on a probability space on which sits a standard homogeneous Poisson process* $N(\cdot)$ *on* $\mathbb{R}_+^2$ *such that*

$$\limsup_{n\to\infty}\ \sup_{0 \leq t \leq t_n} \left| nU_n(n^{-1}t) - S_n^*(n^{-1}t) \right| < \infty \quad a.s., \tag{2.39}$$

where $S_T^*(t) = N((0,T] \times (0,t])$ *with* $E(S_T^*(t)) = Tt$ *for* $T \geq 0$ *and* $t \geq 0$.

PROOF. Consider a standard homogeneous Poisson process $N(\cdot)$ on $\mathbb{R}_+^2$, set for $0 \leq s_n \leq 1$, $\nu_n = N((n-1, n] \times (0, s_n])$ for $n \geq 1$, and denote by τ_n the time of first arrival of $S_n^* - S_{n-1}^*$, i.e., $\tau_n = \inf\{t > 0\colon S_n^*(t) - S_{n-1}^*(t) \geq 1\}$.

Define a sequence $\{\xi_n,\ n \geq 1\}$ of independent random variables with distribution given by

$$P(\xi_n = 0) = 1 - P(\xi_n = 1) = (1 - s_n)e^{s_n} \quad \text{for } n = 1, 2, \dots . \tag{2.40}$$

Let $\{\theta_n,\ n \geq 1\}$ be a sequence of independent uniform $(0,1)$ random variables. Assume that $N(\cdot)$, $\{\xi_n,\ n \geq 1\}$ and $\{\theta_n,\ n \geq 1\}$ are independent. Further let

$$\begin{aligned} U_n = {} & \tau_n 1_{\{\nu_n = 1\}} + s_n \theta_n 1_{\{\nu_n \geq 2\}} + s_n \theta_n 1_{\{\nu_n = 0\}} 1_{\{\xi_n = 1\}} \\ & + (1 - \theta_n + \theta_n s_n) 1_{\{\nu_n = 0\}} 1_{\{\xi_n = 0\}}. \end{aligned} \tag{2.41}$$

It is easily verified that $\{U_n,\ n \geq 1\}$ so defined is an i.i.d. sequence of random variables with a uniform distribution on $(0,1)$. Moreover, if $U_n(\cdot)$ is the right-continuous empirical df of $U_1, \dots, U_n$, we have for $n = 1, 2, \dots,$

$$\begin{aligned} 1 - P(nU_n(t) - (n-1)U_{n-1}(t) = S_n^*(t) - S_{n-1}^*(t) \text{ for } 0 \leq t \leq s_n) \\ = s_n(1 - e^{-s_n}) \leq s_n^2, \end{aligned} \tag{2.42}$$

where $U_0 \equiv 0$ and $S_0 \equiv 0$.

Note that the construction given in (2.40) and (2.41) is well-known for the approximation of Bernoulli random variables by Poisson random variables [see, e.g., Serfling (1975), Section 5].

Now set $s_n = n^{-1}t_n$. By (2.42) and the Borel–Cantelli lemma, we see that (2.38) entails that almost surely for all n sufficiently large

$$nU_n(t) - (n-1)U_{n-1}(t) = S_n^*(t) - S_{n-1}^*(t) \quad \text{for all } 0 \leq t \leq n^{-1}t_n, \tag{2.43}$$

which since $n^{-1}t_n \downarrow$ yields (2.39). Notice that if the assumption $n^{-1}t_n \downarrow$ is not satisfied, then we can replace t_n in (2.39) by $n \min\{m^{-1}t_m\colon 1 \leq m \leq n\}$. $\square$

For the proof of Theorem 2.1, we will choose t_n in Proposition 2.1 as any sequence satisfying (2.38) and $t_n/\log\log n \to \infty$ as $n \to \infty$ (for instance $t_n = n^{1/4}$ will do). It is easily verified from the conclusion of Proposition 2.1 that we need only prove that Theorem 2.1 holds with $S_n(n^{-1}t)$ replacing $nU_n(n^{-1}t)$, where $S_n(t) := N((0,n] \times (0,t))$ denotes the left-continuous version of S_n^*. We will make use of the following lemmas.

LEMMA 2.5. *Let $\{\Pi(t),\ t \geq 0\}$ be a standard left-continuous Poisson process. Set*

$$\Theta(b,a) = \sup_{\substack{0 \leq t-s \leq a \\ 0 \leq s \leq t \leq 1}} |\Pi(tb) - \Pi(sb)| \quad \text{for } a > 0 \text{ and } b > 0. \tag{2.44}$$

Then, for all $b > 0$ and $x > 1$,

$$P(\Pi(b) > bx) \leq \exp(-bh(x)), \tag{2.45}$$

and for all $b > 0$, $0 < a \le \delta < 1$ *and* $x > 1$,

$$P(\Theta(b,a) > abx) \le \frac{20}{a\delta^3}\exp\bigl(-(1-\delta)^3 abh(x)\bigr), \tag{2.46}$$

where h *is as in* (1.6).

PROOF. (2.45) follows from Markov's inequality [see, e.g., Inequality 1, page 485 in Shorack and Wellner (1986)], while (2.46) is a version of Inequality 5, page 571 [see also pages 545–548 in Shorack and Wellner (1986). □

LEMMA 2.6. *For any subset* $\mathscr{A}$ *of* $I(0,1)$ *and* $v > 0$ *set*

$$J_{h,v}(\mathscr{A}) = \inf_{f\in\mathscr{A}} J_{h,v}(f), \tag{2.47}$$

where $J_{h,v}(f)$ *is as in* (2.11). *Let* $\{\Pi(t),\ t \ge 0\}$ *be a standard left-continuous Poisson process, and for* $T > 0$ *and* $v > 0$, *define* $Z_{v,T} \in I(0,1)$ *by*

$$Z_{v,T}(t) = T^{-1}\Pi(Tvt) \quad \text{for } 0 \le t \le 1,$$
$$Z_{v,T}(1+) = T^{-1}\lim_{\varepsilon\downarrow 0}\Pi(Tv+\varepsilon).$$

Then:

(a) *For each closed subset* $\mathscr{F} \subset I(0,1)$ *in* $(I(0,1), J_1)$,

$$\limsup_{T\to\infty} T^{-1}\log P(Z_{v,T}\in\mathscr{F}) \le -J_{h,v}(\mathscr{F}). \tag{2.48}$$

(b) *For each open subset* $\mathscr{G} \subset I(0,1)$ *in* $(I(0,1), J_1)$,

$$\liminf_{T\to\infty} T^{-1}\log P(Z_{v,T}\in\mathscr{G}) \ge -J_{h,v}(\mathscr{G}). \tag{2.49}$$

PROOF. This result is a consequence of the results of Lynch and Sethuraman (1987) (see their Example 1) and Varadhan (1966). □

We will now introduce some further notation. For any $v > 0$ and integer $n \ge 1$ define $f_{n,v} \in I(0,1)$ by

$$\begin{aligned} f_{n,v}(t) &= S_n\bigl(n^{-1}vl(n)t\bigr) \quad \text{for } 0 \le t \le 1,\\ f_{n,v}(1+) &= \lim_{\varepsilon\downarrow 0} S_n\bigl(n^{-1}vl(n)+\varepsilon\bigr), \end{aligned} \tag{2.50}$$

where here and in the sequel, $l(n) = \log\log(\max(n,3))$.

For any $\lambda > 1$, set

$$D_r(\lambda) = \max_{n_r < n \le n_{r+1}} \sup_{0\le t\le 1}\left| l(n_r)^{-1} f_{n_r,v}(t) - l(n)^{-1} f_{n,v}(t)\right|,$$

where $n_r := \lfloor \lambda^r \rfloor$, $r = 1, 2, \ldots,$ with $\lfloor x \rfloor \le x < \lfloor x \rfloor + 1$ denoting the integer part of x.

LEMMA 2.7. *There exists a bounded positive function ψ defined on $(1, \sqrt{2})$ such that $\psi(\lambda) \to 0$ as $\lambda \downarrow 1$, and*

$$\limsup_{r\to\infty} l(n_r)^{-1} D_r(\lambda) \le \psi(\lambda) \quad a.s.\ \text{for all } 1 < \lambda < \sqrt{2}. \tag{2.51}$$

PROOF. Choose any $1 < \lambda < \sqrt{2}$. We claim that

$$\limsup_{r\to\infty} l(n_r)^{-1} S_{n_r}\big(n_r^{-1} v l(n_r)\big) \le v\delta_v^+ \quad \text{a.s.,} \tag{2.52}$$

$$\limsup_{r\to\infty} l(n_r)^{-1} D_r'(\lambda) \le (\lambda - 1) v \delta_{v(\lambda-1)}^+ \quad \text{a.s.} \tag{2.53}$$

and

$$\limsup_{r\to\infty} l(n_r)^{-1} D_r''(\lambda) \le \tfrac{1}{8}(1-\lambda^{-2})\delta_{(v/8)(1-\lambda^{-2})}^+ \quad \text{a.s.,} \tag{2.54}$$

where δ_c^+ is as in Theorem C,

$$D_r'(\lambda) = S_{n_{r+1}}\big(n_r^{-1} v l(n_r)\big) - S_{n_r}\big(n_r^{-1} v l(n_r)\big)$$

and

$$D_r''(\lambda) = \sup_{0\le t\le 1} \left| S_{n_r}\big(n_r^{-1} v l(n_r) t\big) - S_{n_t}\big(n_{r+1}^{-1} v l(n_{r+1}) t\big)\right|.$$

First consider (2.52). Choose any $\varepsilon > 0$. Noting that $S_{n_r}(n_r^{-1} v l(n_r)) =_d \Pi(v l(n_r))$, where $\Pi(\cdot)$ is as in Lemma 2.5, we see that by (2.45)

$$P\Big(S_{n_r}\big(n_r^{-1} v l(n_r)\big) > v l(n_r)\delta_{v/(1+\varepsilon)}^+\Big) \le \exp(-(1+\varepsilon) l(n_r)). \tag{2.55}$$

Since the series formed by the right side of (2.55) is summable in r, the Borel–Cantelli lemma implies that

$$\limsup_{r\to\infty} l(n_r)^{-1} S_{n_r}\big(n_r^{-1} v l(n_r)\big) \le v\delta_{v/(1+\varepsilon)}^+ \quad \text{a.s.} \tag{2.56}$$

Letting $\varepsilon \downarrow 0$ in (2.56) completes the proof of (2.52). Assertion (2.53) is proven in much the same way.

Finally, consider (2.54). Notice that for all r sufficiently large, $D_r''(\lambda) \le D_r^*(\lambda)$, where

$$D_r^*(\lambda) = \sup_{\substack{0\le t-s\le 1-\lambda^{-2}\\ 0\le s\le t\le 1}} \left| S_{n_r}\big(n_r^{-1} v l(n_r) t\big) - S_{n_r}\big(n_r^{-1} v l(n_r) s\big)\right|,$$

which in turn is equal in distribution to $\Theta(v l(n_r), 1-\lambda^{-2})$, where $\Theta(\cdot,\cdot)$ is as in (2.44). Thus for all $\varepsilon > 0$,

$$\begin{aligned} &P\Big(D_r^*(\lambda) > \tfrac{1}{8}(1-\lambda^{-2})\delta_{(v/8)(1-\lambda^{-2})/(1+\varepsilon)}^+\Big) \\ &\quad = P\Big(\Theta\big(v l(n_r), 1-\lambda^{-2}\big) > \tfrac{1}{8}(1-\lambda^{-2})\delta_{(v/8)(1-\lambda^{-2})/(1+\varepsilon)}\Big), \end{aligned}$$

which by (2.46) taken with $\delta = 1/2$, $a = 1-\lambda^{-2}$ (note that $a \le \delta \Leftrightarrow \lambda \le \sqrt{2}$

for $\lambda > 1$) and $b = vl(n_r)$ is less than or equal to

$$160(1 - \lambda^{-2})^{-1} \exp(-(1 + \varepsilon) l(n_r)).$$

Assertion (2.54) now follows as above.

To complete the proof of (2.51), observe that

$$\begin{aligned} &l(n_r)^{-1} D_r(\lambda) \\ &\quad \le l(n_r)^{-1} \left\{ D_r'(\lambda) + D_r''(\lambda) + (1 - l(n_r)/l(n_{r+1})) S_{n_r}\left(n_r^{-1} v l(n_r)\right) \right\}. \end{aligned}$$

Therefore, by (2.52), (2.53) and (2.54),

$$\begin{aligned} \limsup_{r\to\infty} l(n_r)^{-1} D_r(\lambda) &\le (\lambda - 1) v \delta^+_{v(\lambda-1)} + \tfrac{1}{8}(1 - \lambda^{-2}) \delta^+_{(v/8)(1-\lambda^{-2})} \\ &=: \psi(\lambda) \quad \text{a.s.} \end{aligned} \tag{2.57}$$

Since $c\delta_c^+ \to 0$ as $c \downarrow 0$ [see, e.g., Shorack and Wellner (1986), page 433], (2.57) completes the proof of Lemma 2.7. Note that this lemma remains valid with S_n^* replacing S_n in (2.50). $\square$

Denote by $\|\cdot\|$ the supremum norm on $I(0,1)$. For any $\varepsilon > 0$ and $\mathscr{A} \subset I(0,1)$, denote by $\mathscr{A}^\varepsilon$ the subset of $I(0,1)$ such that, for every $q \in \mathscr{A}^\varepsilon$ there exists an $f \in \mathscr{A}$ with $\|f - q\| < \varepsilon$.

Lemma 2.8. *Let Δ_v be as in* (2.30). *Then, for any $v > 0$ and $\varepsilon > 0$,*

$$P\left(l(n)^{-1} f_{n,v} \notin \Delta_v^\varepsilon \ i.o.\right) = 0, \tag{2.58}$$

where $f_{n,v}$ is as in (2.50).

Proof. Choose by Lemma 2.7, $1 < \lambda < \sqrt{2}$ such that

$$\limsup_{n\to\infty} l(n_r)^{-1} D_r(\lambda) < \tfrac{1}{2}\varepsilon \quad \text{a.s.} \tag{2.59}$$

It is easy to see by (2.59) that

$$\left\{ l(n)^{-1} f_{n,v} \notin \Delta_v^\varepsilon \text{ i.o.} \right\} \subset \left\{ l(n_r)^{-1} f_{n_r,v} \notin \Delta_v^{\varepsilon/2} \text{ i.o.} \right\}. \tag{2.60}$$

Since the complement $\mathscr{F}$ of $\Delta_v^{\varepsilon/2}$ is a closed subset of $(I(0,1), J_1)$ and $l(n_r)^{-1} f_{n_r,v} =_d Z_{v,l(n_r)}$, where $Z_{v,T}$ is as in Lemma 2.6, we have by (2.48) that

$$\limsup_{r\to\infty} l(n_r)^{-1} \log P\left(l(n_r)^{-1} f_{n_r,v} \notin \Delta^{\varepsilon/2}\right) \le -J_{h,v}(\mathscr{F}).$$

It is straightforward to show using lower semicontinuity of $J_{h,v}$ and compactness of Δ_v for all $v > 0$ (see Lemma 2.1) that necessarily $J_{h,v}(\mathscr{F}) > 1$. Therefore, for some $\gamma > 1$ and all large r,

$$P\left(l(n_r)^{-1} f_{n_r,v} \notin \Delta_v^{\varepsilon/2}\right) \le \exp(-\gamma l(n_r)).$$

An application of the Borel–Cantelli lemma now shows that

$$P\Big(l(n_r)^{-1} f_{n_r,v} \notin \Delta_v^{\varepsilon/2} \text{ i.o. (in } r)\Big) = 0,$$

which by (2.60) gives (2.58). □

For any $f \in I(0,1)$ and $\varepsilon > 0$ let $B_\varepsilon(f) = \{q \in I(0,1): \|f - q\| < \varepsilon\}$.

LEMMA 2.9. *For all $f \in \Delta_v$, $v > 0$ and $\varepsilon > 0$,*

$$P\Big(l(n)^{-1} f_{n,v} \in B_\varepsilon(f) \text{ } i.o.\Big) = 1. \tag{2.61}$$

PROOF. Choose $\varepsilon > 0$ and $\lambda > 1$. It is routine to verify as in the proof of Lemma 2.7 that

$$\limsup_{r\to\infty} l(n_r)^{-1} S_{n_{r-1}}\big(n_r^{-1} v l(n_r)\big) \le (v/\lambda)\delta_{v/\lambda}^{+} \quad \text{a.s.} \tag{2.62}$$

Set for $r = 1, 2, \dots,$ and $t \ge 0$,

$$\begin{aligned} W_{n_r}(t) &= f_{n_r,v}(t) - S_{n_{r-1}}\big(n_r^{-1} v l(n_r) t\big) \\ &= S_{n_r}\big(n_r^{-1} v l(n_r) t\big) - S_{n_{r-1}}\big(n_r^{-1} v l(n_r) t\big), \\ W'_{n_r}(t) &= \{n_r/(n_r - n_{r-1})\} W_{n_r}(t) \end{aligned}$$

and

$$\rho_{n_r} = n_r^{-1}(n_r - n_{r-1}) l(n_r).$$

Since $l(n_r)^{-1} W'_{n_r} =_d Z_{\rho(n_r),v}$ we have by (2.49) and the fact that $B_{\varepsilon/3}(f)$ is open

$$\liminf_{r\to\infty} \rho(n_r)^{-1} P\Big(l(n_r)^{-1} W'_{n_r} \in B_{\varepsilon/3}(f)\Big) \ge -J_{h,v}\big(B_{\varepsilon/3}(f)\big). \tag{2.63}$$

Now there always exists a $q \in B_{\varepsilon/3}(f)$ such that $J_{h,v}(q) < 1$, i.e., choose

$$q(t) = (1 - \theta) f(t) + \theta t v \delta_{v'}^{+} \quad \text{for } 0 \le t \le 1,$$

where $v' > v$ and $\theta > 0$ are sufficiently small and use convexity of $J_{h,v}(\cdot)$ to show that $J_{h,v}(q) < 1$. By (2.63), it follows that for all r sufficiently large

$$P\Big(l(n_r)^{-1} W'_{n_r} \in B_{\varepsilon/3}(f)\Big) \ge \exp\big(-\gamma l(n_r)/(1 - \lambda^{-1})\big)$$

for some $0 < \gamma < 1$, which by making an initial choice of λ large enough is for some $0 < \gamma' < 1$ greater than or equal to

$$\exp(-\gamma' l(n_r)).$$

Thus, by using independence of the W'_{n_r}'s and the Borel–Cantelli lemma, we obtain for all large enough $\lambda > 1$,

$$P\Big(l(n_r)^{-1} W'_{n_r} \in B_{\varepsilon/3}(f) \text{ i.o.}\Big) = 1. \tag{2.64}$$

From (2.62) and (2.64) we have for any given $\lambda > 1$ sufficiently large, that almost surely, there exists a subsequence $\{r'\} \subset \{r\}$ such that for all r',

$$l(n_{r'})^{-1} W'_{n_{r'}} \in B_{\varepsilon/3}(f), \qquad l(n_{r'})^{-1} \|W'_{n_{r'}}\| < \tfrac{1}{3}\varepsilon + \|f\|, \tag{2.65}$$

$$l(n_{r'})^{-1} \|W_{n_{r'}} - W'_{n_{r'}}\| < \lambda^{-1}\left(\tfrac{1}{3}\varepsilon + \|f\|\right), \tag{2.66}$$

$$l(n_{r'})^{-1} \|f_{n_{r'},v} - W_{n_{r'}}\| < (v/\lambda)\delta^{+}_{v/\lambda} + \tfrac{1}{6}\varepsilon. \tag{2.67}$$

Noting that $(v/\lambda)\delta^{+}_{v/\lambda} \to 0$ as $\lambda \to \infty$ we see that by selecting λ so large that the right sides of (2.66) and (2.67) are each less than $\frac{1}{3}\varepsilon$, by (2.65) we have along $\{r'\}$,

$$l(n_{r'})^{-1} f_{n_{r'},v} \in B_{\varepsilon}(f).$$

This finishes the proof of (2.61). □

Combining Lemmas 2.8 and 2.9, we obtain the following proposition.

PROPOSITION 2.2. *Let $0 < v < \infty$ and let $f_{n,v}$ be as in* (2.50). *Then the sequence $\{f_{n,v}, n \geq 1\}$ is relatively compact in $(I(0,1), U)$ with set of limit points equal to Δ_v, where Δ_v is as in* (2.8).

PROOF OF THEOREM 2.1. As noted earlier, Propositions 2.1 and 2.2, jointly with the observation that $\|S_n - S_n^*\| \leq 1$ a.s., imply that Theorem 2.1 holds for all sequences $\{\kappa_n\}$ of the form $\kappa_n = vl(n)$ for $0 < v < \infty$. It remains to prove that the same result holds for all sequences satisfying

$$0 < \kappa_n < n \quad \text{and} \quad \kappa_n/\log\log n \to v \in (0,\infty). \tag{2.68}$$

Choose $\varepsilon > 0$. Using the fact that $\{l(n)^{-1} nU_n(n^{-1}v(1+\varepsilon)l(n)I)\}$ is relatively compact, and the observation (see, e.g., Example 2.2) that $\Delta_v \subset$ AC $I(0,1)$, we see that, w.p.1, any subsequence of $\{l(n)^{-1}nU_n(n^{-1}\kappa_n I)\}$ contains a further subsequence which converges in $(I_{\mathrm{RC}}(0,1), U)$ to a function of the form $f(t) = f_\varepsilon(t/(1+\varepsilon))$, where $f_\varepsilon \in \Delta_{v(1+\varepsilon)}$. Since evidently

$$J_{h,v}(f) = v(1+\varepsilon)\int_0^{1/(1+\varepsilon)} h\left(\frac{\dot{f}_\varepsilon(u)}{v(1+\varepsilon)}\right) du \leq J_{h,v(1+\varepsilon)}(f_\varepsilon) \leq 1, \tag{2.69}$$

we see that, under (2.68), $\{l(n)^{-1}nU_n(n^{-1}\kappa_n I)\}$ is relatively compact in $(I_{\mathrm{RC}}(0,1), U)$ with set of limit points included in Δ_v. Conversely, take $f \in \Delta_v$ and define f_ε by

$$f_\varepsilon(t) = \begin{cases} f(t(1+\varepsilon)) & \text{for } 0 \leq t \leq 1/(1+\varepsilon), \\ f(1) & \text{for } 1/(1+\varepsilon) \leq t \leq 1. \end{cases} \tag{2.70}$$

By (2.11) and (2.69), we see that $J_{h,v}(f) = J_{h,v(1+\varepsilon)}(f_\varepsilon)$. Thus, since f_ε is w.p.1 the limit of a subsequence of $\{l(n)^{-1}(nU_n(n^{-1}v(1+\varepsilon)l(n)I))\}$, f is the limit of a subsequence of $\{l(n)^{-1}(nU_n(n^{-1}\kappa_n I))\}$. The proof of Theorem 2.1 is now completed. □

3. The first nonstandard LIL in the extreme case. In this section, we consider nonstandard LILs obtained for the sequences $\{A(n,\kappa_n)\alpha_n(n^{-1}\kappa_n I)\}$ and $\{B(n,l_n)\beta_n(n^{-1}l_n I)\}$, where $A(n,\kappa_n)$ and $B(n,l_n)$ are as in (1.18), in the *extreme case* when $\kappa_n/\log\log n \to 0$ and $l_n/\log\log n \to 0$ as $n\to\infty$. A second type of nonstandard LIL will be presented in Section 4.

3.1. *Theorems.* We use throughout the notation introduced in Section 2.1.

THEOREM 3.1. *Assume that* $\{\kappa_n,\ n\ge 1\}$ *satisfies the conditions*

$$0<\kappa_n\le n,\qquad \kappa_n/\log\log n\to 0\quad \text{as } n\to\infty, \tag{3.1}$$

there exists a sequence χ_n *such that*

$$\chi_n\downarrow 0 \text{ and } \chi_n=(1+o(1))\log\left(\frac{\log\log n}{\kappa_n}\right)\bigg/\log\log n\quad \text{as } n\to\infty. \tag{3.2}$$

Then the sequence of functions

$$\left\{\frac{nU_n(n^{-1}\kappa_n t)}{\log\log n}\log\left(\frac{\log\log n}{\kappa_n}\right),0-\le t\le 1+\right\}$$

is almost surely relatively compact in $(I_{\mathrm{RC}}(0,1),W)$ *with set of limit points given by*

$$\Delta_0=\{f\in I_{\mathrm{RC}}(0,1)\colon f(1)\le 1\}. \tag{3.3}$$

THEOREM 3.2. *Assume that* $\{l_n,\ n\ge 1\}$ *satisfies the conditions*

$$0<l_n\le n,\qquad l_n/\log\log n\to 0\quad \text{and}\quad l_n\to\infty\quad \text{as } n\to\infty. \tag{3.4}$$

Then the sequence of functions

$$\left\{\frac{nV_n(n^{-1}l_n s)}{\log\log n},0-\le s\le 1+\right\}$$

is almost surely relatively compact in $(I(0,1),W)$ *with set of limit points given by*

$$\Gamma_0=\{g\in I(0,1)\colon g(1+)\le 1\}. \tag{3.5}$$

REMARK 3.1. In the range covered by Theorems 3.1 and 3.2, we can replace nU_n by $n^{1/2}\alpha_n$ and nV_n by $n^{1/2}\beta_n$.

REMARK 3.2. It is remarkable that the limit sets Δ_0 in (3.3) and Γ_0 in (3.5) are composed of right-continuous and left-continuous distribution functions for all nonnegative measures on $[0,1]$ having total mass less than or equal to 1. Thus both limit sets coincide up to the conventions of left or right continuity used for $U_n(\cdot)$ and $V_n(\cdot)$.

REMARK 3.3. Set $\kappa_n = v \log\log n$ for $0 < v < \infty$. By Theorem 2.1, the sequence

$$\left\{ \frac{nU_n(n^{-1}\kappa_n t)}{\log\log n} \log\left(\frac{\log\log n}{\kappa_n} \right), 0- \le t \le 1+ \right\}$$
$$= \left\{ \frac{nU_n(n^{-1}\kappa_n t)}{\log\log n} \log\left(\frac{1}{v} \right), 0- \le t \le 1+ \right\}$$

is relatively compact in $(I(0,1), U)$. The corresponding set of limit points is given by $\{ f \in I_{\mathrm{RC}}(0,1) \colon v \int_0^1 h(\dot{f}(x)/(v \log(1/v)))\, dx \le 1\}$. Since

$$\lim_{v \to 0} vh\left(z \Big/ \left(v \log\left(\frac{1}{v} \right) \right) \right) = z,$$

we see that the integral in the above limit set reduces for $v = 0$ to $\int_0^1 \dot{f}(x)\, dx = f(1)$. Hence the statement of Theorem 3.1 is in agreement with that of Theorem 2.1.

Likewise, if $l_n = w \log\log n$ for $0 < w < \infty$, by Theorem 2.2, the sequence

$$\left\{ \frac{nV_n(n^{-1} l_n s)}{\log\log n}, 0- \le s \le 1+ \right\}$$

is relatively compact in $(I(0,1), W)$ with limit set equal to

$$\Gamma_w = \left\{ g \in I(0,1) \colon g(1+) - w - w \int_0^1 \log(\dot{g}(u)/w)\, du \le 1 \right\}.$$

By letting $w \to 0$ in this expression, we obtain that this set reduces to $\{g \in I(0,1) \colon g(1+) \le 1\}$. Thus the statement of Theorem 3.2 is also in agreement with that of Theorem 2.2.

REMARK 3.4. If we have

$$\log\left(\frac{\log\log n}{\kappa_n} \right) \Big/ \log\log n = \nu > 0,$$

then [see Theorem 1 of Kiefer (1972) and Theorem E in the sequel]

$$\limsup_{n \to \infty} nU_n(n^{-1}\kappa_n) = \lfloor 1/\nu \rfloor \quad \text{a.s.}$$

A version of Theorem 3.1 valid in this case will be given in Section 4. We will also consider in Section 4 the case when l_n is constant in Theorem 3.2.

3.2. *Proofs*. We start with the following proposition which shows that Theorem 3 and (3.9) of Kiefer (1972) [see, e.g., Shorack and Wellner (1986), (12), page 234] hold under (3.1) and (3.2).

PROPOSITION 3.1. *Assume that* $\{\kappa_n, n \geq 1\}$ *satisfies* (3.1) *and* (3.2). *Then*

$$\limsup_{n\to\infty}\left(\frac{nU_n(n^{-1}\kappa_n)}{\log\log n}\log\left(\frac{\log\log n}{\kappa_n}\right)\right) = 1 \quad a.s. \tag{3.6}$$

PROOF. Let $U_{i,n} = V_n(i/n)$ for $1 \leq i \leq n$ be as in Remark 1.2. We have evidently the event identity $\{nU_n(n^{-1}\kappa_n) \geq k_n\} = \{nU_{k_n,n} \leq \kappa_n\}$. Let $\chi_n \downarrow 0$ be as in (3.2). It is straightforward that all we need for (3.6) is to prove that, for any $0 < \varepsilon < 1$,

$$P\left(nU_{k'_n,n} \leq \kappa_n \text{ i.o.}\right) = 0 \quad \text{for } k'_n = k'_{n,\varepsilon} = \lfloor (1+\varepsilon)/\chi_n \rfloor, \tag{3.7}$$

and

$$P\left(nU_{k''_n,n} \leq \kappa_n \text{ i.o.}\right) = 1 \quad \text{for } k''_n = k''_{n,\varepsilon} = \lfloor 1/(\chi_n(1+\varepsilon)) \rfloor. \tag{3.8}$$

Since (3.1) and (3.2) imply that, for $k_n = k'_n$ or k''_n and all n sufficiently large,

$$1 \leq k_n \leq n, \quad k_n \uparrow \infty \quad \text{and} \quad k_n/\log\log n \to 0 \quad \text{as } n \to \infty,$$

we may apply (1.17) to show that

$$P\left(nU_{k'_n,n} \leq k'_n \exp\left(-\left(\frac{1+\varepsilon}{1+\frac{1}{2}\varepsilon}\right)\frac{1}{k'_n}\log\log n\right) \text{ i.o.}\right) = 0, \tag{3.9}$$

and

$$P\left(nU_{k''_n,n} \leq k''_n \exp\left(-\left(\frac{1+\frac{1}{2}\varepsilon}{1+\varepsilon}\right)\frac{1}{k''_n}\log\log n\right) \text{ i.o.}\right) = 1. \tag{3.10}$$

A simple analysis now shows that we have ultimately in $n \to \infty$,

$$k''_n \exp\left(-\left(\frac{1+\frac{1}{2}\varepsilon}{1+\varepsilon}\right)\frac{1}{k''_n}\log\log n\right) \leq \kappa_n \leq k'_n \exp\left(-\left(\frac{1+\varepsilon}{1+\frac{1}{2}\varepsilon}\right)\frac{1}{k'_n}\log\log n\right).$$

Thus (3.9) and (3.10) imply (3.7) and (3.8), which completes the proof of the proposition. □

The following sequence of lemmas is directed to prove Theorem 3.1.

LEMMA 3.1. *Under* (3.1) *and* (3.2), *the sequence of functions*

$$\left\{\frac{nU_n(n^{-1}\kappa_n t)}{\log\log n}\log\left(\frac{\log\log n}{\kappa_n}\right), 0- \leq t \leq 1+\right\} \tag{3.11}$$

is almost surely relatively compact in $(I_{\mathrm{RC}}(0,1), W)$.

PROOF. It is an obvious consequence from Helly's selection lemma and Proposition 3.1. □

In the sequel, we shall use the following notation. Let $m_j = \lfloor \exp(j \log^2 j) \rfloor$ for $j = 1, 2, \dots,$

$$X_{j,N}(i) = S_{m_j}\left(\frac{i}{N} m_j^{-1} \kappa_{m_j}\right) - S_{m_{j-1}}\left(\frac{i}{N} m_j^{-1} \kappa_{m_j}\right) - S_{m_j}\left(\frac{i-1}{N} m_j^{-1} \kappa_{m_j}\right) + S_{m_{j-1}}\left(\frac{i-1}{N} m_j^{-1} \kappa_{m_j}\right), \tag{3.12}$$

for $1 \le i \le N$, and

$$x_j = (\log\log m_j) \Big/ \log\left(\frac{\log\log m_j}{\kappa_{m_j}}\right) \quad \text{for } j \ge j_0 \text{ sufficiently large.} \tag{3.13}$$

LEMMA 3.2. *Assume that* (3.1) *and* (3.2) *are satisfied. For any fixed integer* $N \ge 1$, $\varepsilon > 0$ *and coefficients* $\lambda_1, \dots, \lambda_N$ *such that*

$$\lambda_1 > 0, \dots, \lambda_N > 0 \quad \textit{and} \quad \sum_{i=1}^{N} \lambda_i < 1, \tag{3.14}$$

we have

$$P\left(\bigcap_{i=1}^{N} \{X_{j,N}(i) \in [\lambda_i x_j, (\lambda_i + \varepsilon) x_j)\} \ i.o. \ (in\ j) \right) = 1. \tag{3.15}$$

PROOF. Let $P_{j,N}(i) = P(X_{j,N}(i) \in [\lambda_i x_j, (\lambda_i + \varepsilon) x_j))$ and $Q_{j,N}(i) = P(X_{j,N}(i) \ge \lambda_i x_j)$. First we show that, as $j \to \infty$,

$$P_{j,N}(i) \sim Q_{j,N}(i) = \exp(-(1 + o(1)) \lambda_i \log\log m_j). \tag{3.16}$$

For this, with the notation of Lemma 2.5, observe that $Q_{j,N}(i) = P(\Pi(\Lambda_j) \ge \lambda_i x_j)$, where $\Lambda_j = m_j^{-1}(m_j - m_{j-1}) N^{-1} \kappa_{m_j} \sim N^{-1} \kappa_{m_j} = o(x_j)$ as $j \to \infty$. Recall the well-known inequalities for the Poisson distribution [see, e.g., Shorack and Wellner (1986), (8), page 485]

$$P(\Pi(\Lambda) = k) = p_k \le P(\Pi(\Lambda) \ge k) \le \frac{p_k}{1 - \Lambda/k + 1} \quad \text{for } k \ge \Lambda > 0. \tag{3.17}$$

By an application of Stirling's formula to $p_k = (\Lambda^k / k!) e^{-\Lambda}$, routine arguments show that for any fixed $\lambda > 0$,

$$P(\Pi(\Lambda_j) \ge \lambda x_j) = \exp(-(1 + o(1)) \lambda \log\log m_j) \quad \text{as } j \to \infty. \tag{3.18}$$

This, jointly with the observation that $P(\Pi(\Lambda_j) \ge (\lambda + \varepsilon) x_j) = o(P(\Pi(\Lambda_j) \ge \lambda x_j))$ as $j \to \infty$, proves (3.16).

Next, by (3.16) and using the independence of the $X_{j,N}(i)$'s, we have, as $j \to \infty$,

$$P\left(\bigcap_{i=1}^{N}\left\{X_{j,N}(i) \in \left[\lambda_i x_j, (\lambda_i + \varepsilon)x_j\right)\right\}\right) = \exp\left(-(1+o(1))\left(\sum_{i=1}^{N}\lambda_i\right)\log\log m_j\right). \tag{3.19}$$

Since the right-hand side of (3.19) is not summable in j, the Borel–Cantelli lemma implies (3.15) as sought. □

LEMMA 3.3. *Assume that* (3.1) *and* (3.2) *are satisfied. Then*

$$\lim_{j\to\infty}\left(S_{m_{j-1}}\left(m_j^{-1}\kappa_{m_j}\right)/x_j\right) = 0 \quad a.s. \tag{3.20}$$

PROOF. Choose any $\varepsilon > 0$. We have

$$P\left(S_{m_{j-1}}\left(m_j^{-1}\kappa_{m_j}\right) \ge \varepsilon x_j\right) = P\left(\Pi\left(\hat{\Lambda}_j\right) \ge \varepsilon x_j\right), \tag{3.21}$$

where $\hat{\Lambda}_j = m_{j-1}m_j^{-1}\kappa_{m_j}$. Now by using (3.17), Stirling's formula and $m_{j-1}/m_j \to 0$ as $j \to \infty$ one can easily show that for any $\lambda > 1$ one has for all j large enough

$$P\left(\Pi\left(\hat{\Lambda}_j\right) \ge \varepsilon x_j\right) \le \exp(-\lambda \log\log m_j). \tag{3.22}$$

Noting that the right side of (3.22) is summable in j, the Borel–Cantelli lemma yields (3.20). □

PROOF OF THEOREM 3.1. By Proposition 2.1 and the arguments used in the proof of Theorem 2.1, a joint use of Lemmas 3.2 and 3.3 shows that, for any integer $N \ge 1$, $\varepsilon > 0$ and $\lambda_1, \dots, \lambda_N$ satisfying (3.14), we have infinitely often

$$\sum_{i=1}^{k}(\lambda_i - \varepsilon) \le \frac{nU_n\left(n^{-1}\kappa_n(k/N)\right)}{\log\log n}\log\left(\frac{\log\log n}{\kappa_n}\right) \le \sum_{i=1}^{k}(\lambda_i + \varepsilon) \quad \text{for } k = 1, \dots, N.$$

This, jointly with Lemma 3.1 and Proposition 3.1, suffices for our needs. □

We now consider the proof of Theorem 3.2. Let (see Proposition 2.1)

$$T_j(t) = S_{m_j}^*(t) - S_{m_{j-1}}^*(t) \quad \text{for } t \ge 0 \text{ and } j \ge 2. \tag{3.23}$$

Define the corresponding left-continuous inverse by

$$R_j(s) = \inf\left\{t \ge 0 \colon T_j(t) \ge s\right\} \quad \text{for } s \ge 0. \tag{3.24}$$

LEMMA 3.4. *Assume that* $\{l_n, n \geq 1\}$ *is a sequence such that*

$$(3.25) \qquad 0 \leq l_n \leq n \quad \text{and} \quad l_n = O(\log\log n) \quad \text{as } n \to \infty.$$

Then, with probability 1, *there exists a* $j_1 < \infty$ *such that, for all* $j \geq j_1$,

$$(3.26) \qquad V_{m_j}(m_j^{-1}s) = R_j(s) \quad \text{for all } 0 \leq s \leq l_{m_j}.$$

PROOF. First, using Theorem C, Theorem E below and the notation of Remark 1.1, we see that almost surely for all n sufficiently large

$$(3.27) \qquad U_{l_{m_j}, m_j} \leq m_j^{-1} \log^2 m_j \leq m_{j-1} \log^{-2} m_{j-1} \leq U_{1, m_{j-1}}.$$

Next, we use (2.43) in the proof of Proposition 2.1 to show that almost surely for all large n,

$$(3.28) \qquad nU_n(n^{-1}t) - (n-1)U_{n-1}(n^{-1}t) = S_n^*(n^{-1}t) - S_{n-1}^*(n^{-1}t) \quad \text{for all } 0 \leq t \leq n^{1/4}.$$

Thus by (3.27) and (3.28) there exists w.p.1 a $j_1 < \infty$ such that for all $j \geq j_1$ and $0 \leq t \leq (\log\log m_j)^2$,

$$(3.29) \quad m_j U_{m_j}(m_j^{-1}t) = m_j U_{m_j}(m_j^{-1}t) - m_{j-1} U_{m_{j-1}}(m_j^{-1}t) = T_j(m_j^{-1}t),$$

which by (3.27) (again) implies (3.26). □

In view of (3.12), (3.13) and (3.26), let

$$(3.30) \qquad Y_{j,N}(i) = R_j\left(\frac{i}{N} l_{m_j}\right) - R_j\left(\frac{i-1}{N} l_{m_j}\right) \quad \text{for } i = 1, \dots, N,$$

and

$$(3.31) \qquad y_j = m_j^{-1} \log\log m_j \quad \text{for } j \geq j_0 \text{ sufficiently large.}$$

LEMMA 3.5. *Assume that* (3.4) *is satisfied. Then for any fixed integer* $N \geq 1$, *and coefficients* $\lambda_1, \dots, \lambda_N$ *such that* (3.14) *holds, we have*

$$(3.32) \qquad P\left(\bigcap_{i=1}^{N} \{Y_{j,N}(i) \in [\lambda_i y_j, (\lambda_i + \varepsilon) y_j)\} \; i.o. \; (in\, j)\right) = 1.$$

PROOF. The proof follows along the same lines as the proof of Lemma 3.2 with small changes. Notice that the $Y_{j,N}(i)$, $i = 1, \dots, N$, are independent with each following a $\Gamma(\rho_{i,j}, m_j - m_{j-1})$ distribution, where

$$|\rho_{i,j} - (1/N) l_{m_j}| \leq 2,$$

and $\Gamma(\rho, \mu)$ denotes a distribution with density $\mu^\rho \Gamma(\rho)^{-1} z^{\rho-1} e^{-\mu z}$ for $z > 0$. In order to use the arguments of the proof of Lemma 3.2, all we need is to show that, for any fixed $\lambda > 0$, we have uniformly over $1 \leq i \leq N$,

$$(3.33) \quad P(Y_{j,N}(i) > \lambda y_j) = \exp(-(1 + o(1))\lambda \log\log m_j) \quad \text{as } j \to \infty.$$

Since $T_j(\cdot)$ given in (3.23) takes integer values, so does $\rho_{i,j}$. Therefore we may use the well-known relation between gamma and Poisson distributions to obtain

$$P(Y_{j,N}(i) > \lambda y_j) = P\big(\Pi((m_j - m_{j-1})\lambda y_j) \le \rho_{i,j} - 1\big). \tag{3.34}$$

Next, we have [see, e.g., Shorack and Wellner (1986), (9), page 485]

$$P(\Pi(\Lambda) = k) = p_k \le P(\Pi(\Lambda) \le k) \le \frac{p_k}{1 - k/\Lambda} \quad \text{for } 0 \le k < \Lambda. \tag{3.35}$$

We now apply the inequality (3.35) with $k = \rho_{i,j} - 1$ and $\Lambda = (m_j - m_{j-1})y_j$ and Stirling's formula. Then routine computations using (3.4) yield (3.3). This suffices for the proof of Lemma 3.5. □

Proof of Theorem 3.2. The proof of Theorem 3.2 is identical to the proof of Theorem 3.1 with Proposition 2.1 replaced by Lemma 3.4 and Lemma 3.2 by Lemma 3.5. Therefore we omit the details. □

4. The second nonstandard LIL in the extreme case. Following Remark 1.2, we see that the process $\{V_n(n^{-1}l_n s),\ 0- \le s \le 1+\}$ is determined by $\{U_{i,n},\ 1 \le i \le k_n\}$, where $k_n = \lceil l_n \rceil$. In the case where $l_n = O(1)$ as $n \to \infty$, it is therefore simpler to work directly with $\{U_{i,n},\ 1 \le i \le k\}$ for a fixed $k \ge 1$ rather than with the functional versions considered up to now. In the following Section 4.1, we shall derive a LIL for the sequence $\{U_{i,n},\ 1 \le i \le k\}$. Finally, we complete our study in Section 4.2 by investigating the limiting asymptotic behavior of $\{U_n(n^{-1}\kappa_n s),\ 0- \le s \le 1+\}$ when $\kappa_n > 0$ is so small that $nU_n(n^{-1}\kappa_n) = O(1)$ almost surely as $n \to \infty$.

4.1. *Strong laws for a fixed number of extreme order statistics.* Using the notation introduced above, we shall prove the following theorem.

Theorem 4.1. *Let $k \ge 1$ be fixed. The sequence $\{(\log\log n)^{-1} nU_{i,n},\ 1 \le i \le k\}$ is almost surely relatively compact in $\mathbb{R}^k$ with set of limit points equal to*

$$L_k = \{(u_1, \dots, u_k) \colon 0 \le u_1 \le \cdots \le u_k \le 1\}. \tag{4.1}$$

Proof. In the first place, by (1.13), we have almost surely

$$0 \le \frac{nU_{1,n}}{\log\log n} \le \cdots \le \frac{nU_{k,n}}{\log\log n} \le 1 + o(1) \quad \text{as } n \to \infty. \tag{4.2}$$

Thus, if $\xi_n(i) = (\log\log n)^{-1} nU_{i,n}$ for $1 \le i \le n$, $\{\xi_n(i),\ 1 \le i \le k\}$ is almost surely relatively compact in $\mathbb{R}^k$ (endowed with the usual topology) with set of limit points included in L_k. Ina order to complete our proof by showing that the limit set is equal to L_k, we use Lemma 3.4 to reduce the argument to the proof of the following lemma.

LEMMA 4.1. *Under the notation of Lemma* 3.4, *for any fixed* $k \ge 1$, $V_i > 0$, $i = 1, \dots, k$ *with* $\sum_{i=1}^k V_i < 1$ *and* $\varepsilon > 0$,

$$(4.3)\quad P\left(\bigcap_{i=1}^{k}\left\{\frac{m_j(R_j(i) - R_j(i-1))}{\log\log m_j} \in [V_i, V_i + \varepsilon)\right\} i.o.\ (in\, j)\right) = 1.$$

PROOF. The proof of Lemma 4.1 boils down to showing that, for any fixed $\lambda > 0$,

$$(4.4)\quad \begin{aligned} &P\big(R_j(i) - R_j(i-1) > \lambda m_j^{-1}\log\log m_j\big)\\ &\qquad = \exp\big(-(1+o(1))\lambda\log\log m_j\big) \quad \text{as } j \to \infty. \end{aligned}$$

But since $R_j(i) - R_j(i-1)$ follows an exponential distribution with mean $m_j - m_{j-1}$, the proof of (4.4) is a direct consequence of the fact that $(m_j - m_{j-1})/m_j \to 1$ as $j \to \infty$. Thus (4.4) and (4.3) hold. This completes the proof of Lemma 4.1 and Theorem 4.1. □

4.2. *Extreme tails for the empirical process.* We complete our investigations by considering the limiting functional behavior of the empirical process $U_n(\cdot)$ considered in an interval $[0, n^{-1}\kappa_n]$, where κ_n is so small that $nU_n(n^{-1}\kappa_n) = O(1)$ a.s. as $n \to \infty$. This case is covered by Theorem 1 of Kiefer (1972), which we state in Theorem E below.

THEOREM E. *Assume that* $\{\kappa_n,\ n \ge 1\}$ *satisfies the conditions*

$$(4.5)\quad 0 < \kappa_n \le n \quad \text{and} \quad \kappa_n \downarrow.$$

Then, for any integer $k \ge 1$, $P(nU_n(n^{-1}\kappa_n) \ge k\ i.o.) = 0$ *or* 1, *according as*

$$(4.6)\quad \sum_n n^{-1}\kappa_n^k < \infty \text{ or } = \infty.$$

For any sequence $\{\kappa_n,\ n \ge 1\}$ satisfying (4.5), define $K = K(\{\kappa_n\})$ as the maximal value of k such that the series in (4.6) diverges ($K = \infty$ if this value does not exist). If $K < \infty$,

$$(4.7)\quad P\big(nU_n(n^{-1}\kappa_n) = K \text{ i.o.}\big) = 1 \quad\text{and}\quad P\big(nU_n(n^{-1}\kappa_n) = K + 1 \text{ i.o.}\big) = 0.$$

The following theorem describes the tail empirical process when $K < \infty$. By Remark 3.4, we see that this covers the case left open by Theorem 3.1 for small values of κ_n.

THEOREM 4.2. *Assume that* $\{\kappa_n,\ n \ge 1\}$ *satisfies the conditions, for some integer* $K \ge 0$,

$$(4.8)\quad 0 < \kappa_n \le n, \quad n^{-1}\kappa_n \downarrow, \quad \sum_n n^{-1}\kappa_n^K = \infty \quad \text{and} \quad \sum_n n^{-1}\kappa_n^{K+1} < \infty.$$

Then the sequence of functions $\{nU_n(n^{-1}\kappa_n t),\ 0- \le t \le 1+\}$ *is almost surely* *relatively compact in* $(I_{\mathrm{RC}}(0,1), W)$ *with set of limit points consisting of all functions of* $I_{\mathrm{RC}}(0,1)$ *taking integer values in the interval* $[0, K]$.

PROOF. By Theorem E, it is straightforward that $\{nU_n(n^{-1}\kappa_n I)\}$ is almost surely relatively compact, and that the limit set is included in the set described above. Therefore all we need is to show that both sets are equal. For this, we use the sequence $\nu_j = 2^j$, $j = 0, 1, \dots,$ and observe, as in Kiefer (1972), page 235, that the divergence of $\Sigma_n n^{-1}\kappa_n^K$ entails the divergence of $\Sigma_j \kappa_{\nu_j}^K$. Next, we use (2.43) in the proof of Proposition 2.1 to show that almost surely for all j sufficiently large and all $0 \le t \le 1$,

$$\begin{aligned} &\nu_j U_{\nu_j}\left(\nu_j^{-1}\kappa_{\nu_j} t\right) - \nu_{j-1} U_{\nu_{j-1}}\left(\nu_j^{-1}\kappa_{\nu_j} t\right) \\ &\quad = S_{\nu_j}^*\left(\nu_j^{-1}\kappa_{\nu_j} t\right) - S_{\nu_{j-1}}^*\left(\nu_j^{-1}\kappa_{\nu_j} t\right) =: T_j(t). \end{aligned} \tag{4.9}$$

Fix an integer $N > K$ and an integer sequence $1 \le i_1 < \cdots < i_K \le N$. Consider the event

$$H_j = \bigcap_{l=1}^{K} \left\{T_j(N^{-1}i_l) - T_j(N^{-1}(i_l - 1)) \ge 1\right\}. \tag{4.10}$$

We have

$$P(H_j) = \left(\left(1 - \exp\left(-\tfrac{1}{2}\kappa_{\nu_j}\right)\right)\right)^K.$$

Using the easily proved fact that $\Sigma_j \kappa_{\nu_j}^K = \infty \Leftrightarrow \Sigma_j P(H_j) = \infty$, the independence of the H_j's implies by the Borel–Cantelli lemma that $P(H_j \text{ i.o.}) = 1$ whenever $\Sigma_n n^{-1}\kappa_n^K = \infty$. Since we also know that $P(nU_n(n^{-1}\kappa_n) \ge K + 1 \text{ i.o.}) = 0$ under (4.8), we see that all functions f in $I_{\mathrm{RC}}(0,1)$ which are constant on $[(i-1)/N, i/N)$ for $i = 1, \dots, N$, and take integer values belong to the limit set if we have, in addition,

$$f(1) = K \quad \text{and} \quad f(i/N) \le f((i-1)/N) + 1 \quad \text{for } i = 1, \dots, N. \tag{4.11}$$

If we repeat the same argument for the functions $\{nU_n(2n^{-1}\kappa_n t), 0 \le t \le 1\}$, we see that we obtain also the functions f such that $0 \le f(1) \le K$. The proof of Theorem 4.2 may be now completed by an easy argument, based on the fact that the functions f introduced above are dense in $(I_{\mathrm{RC}}(0,1), W)$. □

5. Conclusion. The theorems presented in this article can be applied to obtain strong laws for statistics of the form $\Theta(f_n)$, where f_n is a tail empirical or quantile process and Θ a continuous functional on the space where f_n varies. If $\mathscr{L}$ denotes the almost sure limit set of the sequence $\{f_n\}$, then we may conclude that $\Theta(\mathscr{L})$ is the almost sure limit set of the sequence $\Theta(f_n)$. In particular, by compactness of $\mathscr{L}$, there exist $f_{\min}$ and $f_{\max}$ such that $f_{\min} \in \mathscr{L}$, $f_{\max} \in \mathscr{L}$ and almost surely

$$\begin{aligned} \Theta(f_{\min}) &= \inf_{f \in \mathscr{L}} \Theta(f) = \liminf_{n\to\infty} \Theta(f_n) \le \limsup_{n\to\infty} \Theta(f_n) \\ &= \sup_{f \in \mathscr{L}} \Theta(f) = \Theta(f_{\max}). \end{aligned}$$

Thus the problem of finding strong limiting bounds for $\Theta(f_n)$ reduces to that of finding the extrema of $\Theta(f)$ for $f \in \mathscr{L}$. This method also has the advantage that it gives additional insight on the behavior of f_n when the limiting bounds of $\Theta(f_n)$ are reached through the extremal functions $f_{\min}$ and $f_{\max}$. The explicit derivation of these extremal functions requires specific techniques which are beyond the scope of the present paper. We will content ourselves with the following simple example.

EXAMPLE. Let $\{\phi(s),\ 0 \le s \le 1\}$ be any bounded nonnegative continuous function on $[0,1]$. Define the statistic $W_n(\kappa_n)$ by (with the convention that $\Sigma_\varnothing := 0$)

$$W_n(\kappa_n) = \sum_{i:\, nU_{i,n} \le \kappa_n} \phi\left(\kappa_n^{-1} n U_{i,n}\right).$$

Since evidently

$$W_n(\kappa_n) = \int_0^1 \phi(t)\, dn\, U_n\left(n^{-1}\kappa_n t\right),$$

we see that under the assumptions (3.1) (i.e., $0 < \kappa_n \le n$ and $\kappa_n/\log\log n \to 0$) and (3.2) of Theorem 3.1, we have

$$\begin{aligned}(5.1)\qquad &\limsup_{n\to\infty} \frac{W_n(\kappa_n)}{\log\log n} \log\left(\frac{\log\log n}{\kappa_n}\right)\\ &\quad = \sup_{f\in\Delta_0} \int_0^1 \phi(s)\, df(s) = \sup_{0\le t\le 1} \phi(t) \quad \text{a.s.},\end{aligned}$$

whereas, under assumption (2.29) [i.e., $0 \le \kappa_n \le n$ and $\kappa_n/\log\log n \to v \in (0,\infty)$] of Theorem 3.2, we have

$$(5.2)\qquad \limsup_{n\to\infty} \frac{W_n(\kappa_n)}{\log\log n} = \sup_{f\in\Delta_v} \int_0^1 \phi(s)\, \dot{f}(s)\, ds \quad \text{a.s.}$$

Further applications of our theorems include nonstandard LILs for tail estimators [see, e.g., Deheuvels, Haeusler and Mason (1988)] and sums of extreme values [see, e.g., Deheuvels and Mason (1988)]. In Deheuvels and Mason (1990), we provide general techniques for computing the values of the constants that appear as the limit in expressions like (5.2).

We end with some concluding remarks. A referee has pointed out to us that in the proof of Theorem 2.1 we could have substituted Lemma 2.7 by an application of Theorem 2 of Vervaat (1987) which roughly says that whenever a functional LIL holds among all geometric subsequences it holds along the full sequence. Also the Associate Editor remarked that results related to those in Section 4 may be found in Wichura (1974) and Mori and Oodaira (1976). Indeed, the following result comparable to Theorem 4.1 can be inferred from Corollary 6.2 of Mori and Oodaira (1976). For each fixed integer $k \ge 1$ the sequence

$$(5.3)\qquad \left(-\frac{\log(nU_{1,n})}{\log\log n}, \dots, -\frac{\log(nU_{k,n})}{\log\log n}\right), \qquad n = k, k+1, \dots,$$

is almost surely relatively compact in $\mathbb{R}^k$ with set of limit points equal to

$$\left\{(x_1,\dots,x_k): x_1 \geq \cdots \geq x_k \geq 0 \text{ and } \sum_{j=1}^{k} x_j \leq 1\right\}. \tag{5.4}$$

Finally, Jon Wellner has kindly directed our attention to the fact that versions of Theorems 2.1 and 3.1 can be found in an unpublished dissertation of McBride (1974).

REFERENCES

BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

DEHEUVELS, P., HAEUSLER, E. and MASON, D. M. (1988). Almost sure convergence of the Hill estimator. *Math. Proc. Cambridge Philos. Soc.* **104** 371–384.

DEHEUVELS, P. and MASON, D. M. (1988). The asymptotic behavior of sums of exponential extreme values. *Bull. Sci. Math. (2)* **112** 211–233.

DEHEUVELS, P. and MASON, D. M. (1990). A tail empirical process approach to some non-standard laws of the iterated logarithm. *J. Theoret. Probab.* To appear.

EINMAHL, J. H. J. and MASON, D. M. (1988). Strong limit theorems for weighted quantile processes. *Ann. Probab.* **16** 1623–1643.

FINKELSTEIN, H. (1971). The law of the iterated logarithm for empirical distributions. *Ann. Math. Statist.* **42** 607–615.

KIEFER, J. (1972). Iterated logarithm analogues for sample quantiles when $p_n \downarrow 0$. *Proc. Sixth Berkeley Symp. Math. Statist. Probab.* **1** 227–244. Univ. California Press.

LYNCH, J. and SETHURAMAN, J. (1987). Large deviations for processes with independent increments. *Ann. Probab.* **15** 610–627.

MASON, D. M. (1988). A strong invariance theorem for the tail empirical process. *Ann. Inst. H. Poincaré Probab. Statist.* **24** 491–506.

MCBRIDE, J. (1974). Functional analogues of iterated logarithm type laws for empirical distribution functions whose arguments tend to 0 at an intermediate rate. Ph.D. dissertation, Univ. Chicago.

MORI, T. and OODAIRA, H. (1976). A functional law of the iterated logarithm for sample sequences. *Yokohama J. Math.* **24** 35–49.

SERFLING, R. J. (1975). A general Poisson approximation theorem. *Ann. Probab.* **3** 726–731.

SHORACK, G. R. and WELLNER, J. A. (1986). *Empirical Processes with Applications to Statistics*. Wiley, New York.

SKOROHOD, A. V. (1956). Limit theorems for stochastic processes. *Theory Probab. Appl.* **1** 261–290.

STRASSEN, V. (1964). An invariance principle for the law of the iterated logarithm. *Z. Wahrsch. Verw. Gebiete* **3** 211–226.

VARADHAN, S. R. S. (1966). Asymptotic probabilities and differential equations. *Comm. Pure Appl. Math.* **19** 261–286.

VERVAAT, W. (1987). Functional iterated logarithm laws for geometric subsequences and full sequences. Preprint.

WICHURA, M. J. (1974). On the functional form of the iterated logarithm for the partial maxima of independent identically distributed random variables. *Ann. Probab.* **2** 202–230.

L.S.T.A.
UNIVERSITÉ PARIS VI
4 PLACE JUSSIEU
75252 PARIS CEDEX 05
FRANCE

DEPARTMENT OF MATHEMATICAL SCIENCES
501 EWING HALL
UNIVERSITY OF DELAWARE
NEWARK, DELAWARE 19716