The Annals of Probability
1992, Vol. 20, No. 3, 1248–1287

FUNCTIONAL LAWS OF THE ITERATED LOGARITHM FOR THE INCREMENTS OF EMPIRICAL AND QUANTILE PROCESSES

BY PAUL DEHEUVELS[1] AND DAVID M. MASON[2]

Université Paris VI and University of Delaware

Let $\{\alpha_n(t), 0 \le t \le 1\}$ and $\{\beta_n(t), 0 \le t \le 1\}$ be the empirical and quantile processes generated by the first n observations from an i.i.d. sequence of uniformly distributed random variables on (0, 1). Let $0 < a_n < 1$ be a sequence of constants such that $a_n \to 0$ as $n \to \infty$. We investigate the strong limiting behavior as $n \to \infty$ of the increment functions

$$\{\alpha_n(t + a_n s) - \alpha_n(t), 0 \le s \le 1\}$$

and

$$\{\beta_n(t + a_n s) - \beta_n(t), 0 \le s \le 1\},$$

where $0 \le t \le 1 - a_n$. Under suitable regularity assumptions imposed upon a_n, we prove functional laws of the iterated logarithm for these increment functions and discuss statistical applications in the field of nonparametric estimation.

1. Introduction. Let $U_1, U_2, \dots$ be a sequence of independent and uniformly distributed on (0, 1) random variables. For each integer $n \ge 1$, denote by $U_n(t) = n^{-1}\#\{U_i \le t: 1 \le i \le n\}$, for $-\infty < t < \infty$, the right-continuous empirical distribution function, and denote by $V_n(t) = \inf\{u \ge 0: U_n(u) \ge t\}$, for $0 < t \le 1$, with $V_n(t) = 0$ for $t \le 0$ and $V_n(t) = V_n(1)$ for $t \ge 1$, the left-continuous empirical quantile function, based on the first n of these random variables. Let $\alpha_n(t) = n^{1/2}(U_n(t) - t)$, for $-\infty < t < \infty$, be the uniform empirical process and $\beta_n(t) = n^{1/2}(V_n(t) - t)$, for $-\infty < t < \infty$, be the uniform quantile process.

For any $0 < a < 1$ and integer $n \ge 1$ consider the increment functions

$$(1.1)\qquad \begin{aligned} \xi_n(a, t; s) &= \alpha_n(t + sa) - \alpha_n(t), \\ \zeta_n(a, t; s) &= \beta_n(t + sa) - \beta_n(t), \end{aligned}$$

for $0 \le s \le 1$ and $0 \le t \le 1$.

Received April 1990; revised January 1991.
[1]Invited IMS lecture, Baltimore, April 1990.
[2]Research partially supported by NSF Grant DMS-88-03209 and by a joint NSF-CNRS grant.
AMS 1980 *subject classifications*. Primary 60F15, 60F17; secondary 62G05.
Key words and phrases. Functional limit laws, laws of the iterated logarithm, empirical processes, quantile processes, order statistics, nonparametric estimation, density estimation, nearest neighbor estimates.

Much attention has been directed toward the investigation of the limiting behavior of the maximal oscillations of ξ_n and ζ_n. Let, namely, for $0 < a < 1$,

$$\begin{aligned}\Xi_n^{\pm}(a) &= \pm \sup_{0 \le t \le 1-a} \pm \xi_n(a, t; 1),\\ \Theta_n^{\pm}(a) &= \pm \sup_{0 \le t \le 1-a} \pm \zeta_n(a, t; 1).\end{aligned} \tag{1.2}$$

A sequence of constants $\{a_n,\ n \ge 1\}$ will be said to satisfy the Csörgő–Révész–Stute [CRS] conditions if the following hold:

(S.1) $0 < a_n < 1$ for $n \ge 1$, $a_n \downarrow 0$ and $na_n \uparrow \infty$ as $n \uparrow \infty$;

(S.2) $(\log(1/a_n))/\log\log n \to \infty$ as $n \to \infty$;

(S.3) $na_n/\log n \to \infty$ as $n \to \infty$.

Under the CRS conditions, Stute (1982a) [for $\Xi_n^{\pm}(a_n)$] and Mason (1984) [for $\Theta_n^{\pm}(a_n)$] obtained results which can be extended (see Remark 4.1 in the sequel) to show that

$$\begin{aligned}&\lim_{n\to\infty} \Xi_n^{\pm}(a_n)/(2a_n \log(1/a_n))^{1/2}\\ &\quad = \lim_{n\to\infty} \Theta_n^{\pm}(a_n)/(2a_n \log(1/a_n))^{1/2} = \pm 1 \quad \text{a.s.}\end{aligned} \tag{1.3}$$

In the boundary case when $na_n/\log n \to c \in (0,\infty)$, the limiting behavior of statistics related to $\Xi_n^{\pm}(a_n)$ has been studied by Mason, Shorack and Wellner (1983). Mason (1984) considers likewise statistics related to $\Theta_n^{\pm}(a_n)$, while Deheuvels and Devroye (1984) consider $\Theta_n^{\pm}(a_n)$ in the general case where $na_n = O(\log n)$ as $n \to \infty$. Whenever $na_n/\log n \to c \in (0,\infty)$, we have

$$\lim_{n\to\infty} \Xi_n^{\pm}(a_n)/(2a_n \log n)^{1/2} = (\delta_c^{\pm} - 1)(c/2)^{1/2} \quad \text{a.s.,} \tag{1.4}$$

and

$$\lim_{n\to\infty} \Theta_n^{\pm}(a_n)/(2a_n \log n)^{1/2} = (\gamma_c^{\pm} - 1)(c/2)^{1/2} \quad \text{a.s.,} \tag{1.5}$$

where $0 \le \delta_c^- < 1 < \delta_c^+ < \infty$ are the roots (in δ) of the equation $h(\delta) = 1/c$, with the convention that $\delta_c^- = 0$ for $0 < c < 1$, $0 < \gamma_c^- < 1 < \gamma_c^+ < \infty$ are the roots (in γ) of the equation $l(\gamma) = 1/c$ and the functions $h(\cdot)$ and $l(\cdot)$ are defined by

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

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

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

The results given in (1.3)–(1.5) bear a striking similarity to the limiting behavior of the tail empirical processes which we now describe. Introduce the conditions

(S.4) $$na_n/\log\log n \to \infty \quad \text{as } n \to \infty$$

and

(S.5) $$na_n/\log\log n \to c \in (0, \infty) \quad \text{as } n \to \infty.$$

Under (S.1) and (S.4), Kiefer (1972) proved that

(1.8) $$\begin{aligned}&\limsup_{n\to\infty} \pm \alpha_n(a_n)/(2a_n \log\log n)^{1/2}\\ &\quad = \limsup_{n\to\infty} \pm \beta_n(a_n)/(2a_n \log\log n)^{1/2} = 1 \quad \text{a.s.}\end{aligned}$$

Moreover, in the boundary case (S.5), Kiefer (1972) [see, e.g., Deheuvels (1986)] also proved that

(1.9) $$\limsup_{n\to\infty} \pm \alpha_n(a_n)/(2a_n \log\log n)^{1/2} = \pm(\delta_c^{\pm} - 1)(c/2)^{1/2} \quad \text{a.s.,}$$

and

(1.10) $$\limsup_{n\to\infty} \pm \beta_n(a_n)/(2a_n \log\log n)^{1/2} = \pm(\gamma_c^{\pm} - 1)(c/2)^{1/2} \quad \text{a.s.}$$

A new proof of (1.8) can be achieved by using the functional laws of the iterated logarithm [LIL] given for the tail empirical process $\xi_n(a_n, 0; \cdot)$ by Mason (1988) and for the tail empirical quantile process $\zeta_n(a_n, 0; \cdot)$ by Einmahl and Mason (1988). In the same way, (1.9) and (1.10) follow from the results of Deheuvels and Mason (1990) as an application of nonstandard functional laws of the iterated logarithm. In order to describe these laws, we need to introduce the following notation.

Denote by $I_{\mathrm{RC}}(0,1)$ [resp., $I(0,1)$] the set of all right-continuous (resp., left-continuous) distribution functions of nonnegative and bounded Radon measures with support in [0, 1]. For any $f \in I_{\mathrm{RC}}(0,1)$ [resp., $I(0,1)$] and $-\infty < x < \infty$, set $f(x \pm) = \lim_{\varepsilon \downarrow 0} f(x \pm \varepsilon)$. Further, for $f \in I_{\mathrm{RC}}(0,1)$ [resp., $I(0,1)$] set

(1.11) $$f(x) = \int_0^x \dot{f}(t)\, dt + f_S(x), \quad \text{for } -\infty < x < \infty,$$

where $f_S \in I_{\mathrm{RC}}(0,1)$ [resp., $I(0,1)$] is the distribution function of the singular component in the Lebesgue decomposition of df and $\dot{f}$ is the Lebesgue derivative (defined uniquely a.e.) of the absolutely continuous part of this decomposition.

We now state in Theorems A and B the nonstandard functional LIL's given in Deheuvels and Mason (1990).

THEOREM A. *Under* (S.5), *the sequence of functions* $\{nU_n(a_n s)/\log\log n, 0 \le s \le 1\}$ *is almost surely relatively compact in* $I_{\mathrm{RC}}(0,1)$ *endowed with the topology of uniform convergence, with set of limit points equal to* Δ_c, *where* Δ_c

consists of all absolutely continuous functions f in $I_{\mathrm{RC}}(0,1)$ *such that*

$$c\int_0^1 h\big(\dot{f}(s)/c\big)\,ds \le 1, \tag{1.12}$$

and $h(\cdot)$ *is as in* (1.6).

THEOREM B. *Under* (S.5), *the sequence of functions* $\{nV_n(a_n s)/\log\log n,\ 0 \le s \le 1+\}$ *is almost surely relatively compact in* $I(0,1)$ *endowed with the topology of weak convergence of the underlying measures, with set of limit points equal to* Γ_c, *where* Γ_c *consists of all functions g in* $I(0,1)$ *such that*

$$g_S(1+) + c\int_0^1 l(\dot{g}(s)/c)\,ds \le 1, \tag{1.13}$$

and $l(\cdot)$ *is as in* (1.7).

A description of how Theorems A and B may be used to obtain limiting laws such as (1.9) and (1.10) is given in Deheuvels and Mason (1991).

The similarity of (1.4) and (1.5) with (1.9) and (1.10) hints that versions of the functional LIL's given in Theorems A and B may exist for the increment functions ξ_n and ζ_n in the range corresponding to

$$na_n/\log n \to c \in (0,\infty) \quad \text{as } n \to \infty. \tag{S.6}$$

We will show in Section 2 that this is the case and describe the corresponding laws. Section 3 will be devoted to the case where the sequence $\{a_n,\ n \ge 1\}$ satisfies the CRS conditions. In this range, we shall obtain Strassen-type laws in the spirit of the well-known functional LIL due to Strassen (1964) and similar to the tail-process versions due to Mason (1988) and Einmahl and Mason (1988) under (S.1) and (S.4). The cases where $na_n/\log n \to 0$ or $(\log(1/a_n))/\log\log n \to c \in [0,\infty)$ as $n \to \infty$ will be considered elsewhere.

In Section 4, we will consider applications of our results. In particular, we shall prove laws of the iterated logarithm for nonparametric estimates of a density and of its derivatives which extend those obtained by Stute (1982b). Moreover, we obtain similar results for nearest-neighbor-type estimates.

2. Nonstandard functional LIL's for increments of empirical processes. Throughout this section, $\{a_n,\ n \ge 1\}$ will denote a sequence of constants satisfying assumption (S.6), that is,

$$\frac{na_n}{\log n} \to c \in (0,\infty) \quad \text{as } n \to \infty. \tag{2.1}$$

We assume implicitly that $n \ge n_0$, where $n_0 \ge 1$ is an integer such that $0 < a_n < 1$ for all $n \ge n_0$. Instead of $\xi_n(a_n, t; \cdot)$ and $\zeta_n(a_n, t; \cdot)$, we will consider, for $0 \le s \le 1+$ and $0 \le t \le 1 - a_n$,

$$\begin{aligned} \eta_n(a_n, t; s) &= \frac{n}{\log n}\big(U_n(t + sa_n) - U_n(t)\big), \\ v_n(a_n, t; s) &= \frac{n}{\log n}\big(V_n(t + sa_n) - V_n(t)\big), \end{aligned} \tag{2.2}$$

with the conventions that $U_n(u+) = U_n(u)$ and $V_n(u+) = \lim_{\varepsilon \downarrow 0} V_n(u+\varepsilon)$, for $0 \le u \le 1$. It will be convenient to set

$$\eta_n(a_n, t; s) = v_n(a_n, t; s) = 0, \quad \text{for } s < 0 \text{ and } 0 \le t \le 1,$$
$$\eta_n(a_n, t; s) = \eta_n(a_n, t; 1), \qquad v_n(a_n, t; s) = v_n(a_n, t; 1+),$$
$$\text{for } s > 1 \text{ and } 0 \le t \le 1 - a_n.$$

Moreover, for $1 - a_n < t \le 1$, let $\eta_n(a_n, t; s)$ and $v_n(a_n, t; s)$ be as in (2.2) for $0 \le s \le (1-t)/a_n+$, and set $\eta_n(a_n, t; s) = \eta_n(a_n, t; (1-t)/a_n) + a_n s - 1 + t$ and $v_n(a_n, t, s) = v_n(a_n, t; (1-t)/a_n+) + a_n s - 1 + t$ for $(1-t)/a_n < s \le 1+$. These conventions ensure that $\eta_n(a_n, t; \cdot) \in I_{\mathrm{RC}}(0,1)$ and $v_n(a_n, t; \cdot) \in I(0,1)$ for all $0 \le t \le 1$.

We will make use of the following notation. For any bounded function f on $[0,1]$, set $\|f\| = \sup_{0 \le s \le 1} |f(s)|$, and, for any $\varepsilon > 0$ and $A \subset I_{\mathrm{RC}}(0,1)$, denote by $\tilde{A}^\varepsilon$ the set of all functions $f \in I_{\mathrm{RC}}(0,1)$ such that there exists an $\tilde{f} = \tilde{f}_{f,\varepsilon} \in A$ with $\|f - \tilde{f}\| < \varepsilon$. Consider also the weak topology on $I(0,1)$ [see, e.g., Högnäs (1977)], which is conveniently defined via the Lévy metric given for $f \in I(0,1)$ and $g \in I(0,1)$ by

$$d_L(f, g) = \inf\{r \ge 0: f(s-r) - r \le g(s) \le f(s+r) + r, \text{all } s\}. \tag{2.3}$$

For any $\varepsilon > 0$ and $B \subset I(0,1)$, denote by $\hat{B}^\varepsilon$ the set of all functions $g \in I(0,1)$ such that there exists a $\hat{g} = \hat{g}_{g,\varepsilon} \in B$ with $d_L(g, \hat{g}) < \varepsilon$. We will endow at times $I_{\mathrm{RC}}(0,1)$ with the Skorohod J_1 topology as defined in Skorohod (1956) and Billingsley [(1968), pages 111–123]. This topology is metrizable and such that, whenever f is continuous, $f_n \to f$ in the Skorohod topology is equivalent to $\|f_n - f\| \to 0$. An obvious consequence of this fact and of the properties of Δ_c [see, e.g., Section 2 in Deheuvels and Mason (1990)] is that, for any $B \subset \Delta_c$ and $\varepsilon > 0$, $\tilde{B}^\varepsilon$ is a Skorohod neighborhood of B.

The main results of this section are stated in the following two theorems.

THEOREM 2.1. *Under* (S.6), *for any* $\varepsilon > 0$, *there exists almost surely a finite* N'_ε *such that, for all* $n \ge N'_\varepsilon$, *we have*

$$\{\eta_n(a_n, t; \cdot), 0 \le t \le 1 - a_n\} \subset \tilde{\Delta}_c^\varepsilon. \tag{2.4}$$

Moreover, for any $f \in \Delta_c$ *and* $\varepsilon > 0$, *there exists almost surely a finite* $N''_{\varepsilon,f}$ *such that, for all* $n \ge N''_{\varepsilon,f}$, *there exists a* $\tilde{t}$, $0 \le \tilde{t} = \tilde{t}_{n,\varepsilon,f} \le 1 - a_n$, *such that*

$$\|\eta_n(a_n, \tilde{t}; \cdot) - f\| < \varepsilon. \tag{2.5}$$

THEOREM 2.2. *Under* (S.6), *for any* $\varepsilon > 0$, *there exists almost surely a finite* M'_ε *such that, for all* $n \ge M'_\varepsilon$, *we have*

$$\{v_n(a_n, t; \cdot), 0 \le t \le 1 - a_n\} \subset \hat{\Gamma}_c^\varepsilon. \tag{2.6}$$

Moreover, for any $g \in \Gamma_c$ *and* $\varepsilon > 0$, *there exists almost surely a finite* $M''_{\varepsilon,g}$ *such that, for all* $n \ge M''_{\varepsilon,g}$, *there exists a* $\hat{t}$, $0 \le \hat{t} = \hat{t}_{n,\varepsilon,g} \le 1 - a_n$, *such that*

$$d_L(v_n(a_n, \hat{t}; \cdot), g) \le \varepsilon. \tag{2.7}$$

In the remainder of this section, we will prove Theorems 2.1 and 2.2. The following sequence of lemmas is directed toward the proof of Theorem 2.1.

Let $\{\Pi_n(t),\ t \ge 0\}$ denote a right-continuous Poisson process with parameter $n \ge 1$, that is, such that $E(\Pi_n(t)) = nt$, for $t \ge 0$. Define for $0 < a < 1$, $0 \le s \le 1$ and $0 \le t \le 1 - a$,

$$H_n(a, t; s) = \frac{1}{\log n}(\Pi_n(t + sa) - \Pi_n(t)). \tag{2.8}$$

LEMMA 2.1. *For any choice of* $\{t_1, \dots, t_m\} \subset \{ka_n\colon 0 \le k \le a_n^{-1} - 1\}$, *with* $0 < ma_n \le \frac{1}{2}$ *and Borel subsets* $A_1, \dots, A_m$ *of* $I_{\mathrm{RC}}(0,1)$, *endowed with the topology of uniform convergence, let*

$$E_1 = \{\eta_n(a_n, t_i; \cdot) \in A_i\colon i = 1, \dots, m\}$$

and

$$E_2 = \{H_n(a_n, t_i; \cdot) \in A_i\colon i = 1, \dots, m\}.$$

We have

$$P(E_1) \le 2P(E_2), \quad \textit{for all } n \ge 5. \tag{2.9}$$

PROOF. Set $R = \bigcup_{i=1}^m (t_i, t_i + a_n]$. Notice that the Lebesgue measure of R, written $|R|$, is equal to $ma_n \le \frac{1}{2}$. Set $\overline{R} = [0,1] - R$. Since $\{nU_n(s),\ 0 \le s \le 1\}$ has the same distribution as $\{\Pi_n(s),\ 0 \le s \le 1\}$, conditioned on $\Pi_n(1) = n$, we see that

$$\begin{aligned} P(E_1) = P(E_2 | \Pi_n(1) = n) &= \frac{P(E_2 \cap \{\Pi_n(1) = n\})}{P(\Pi_n(1) = n)} \\ &= \sum_{j=0}^{n} \frac{P\big(E_2 \cap \{\Pi_n(R) = j\} \cap \{\Pi_n(\overline{R}) = n - j\}\big)}{P(\Pi_n(1) = n)}, \end{aligned}$$

which, since the events $E_2 \cap \{\Pi_n(R) = j\}$ and $\{\Pi_n(\overline{R}) = n - j\}$ are independent, is equal to

$$\begin{aligned} &\sum_{j=0}^{n} P(E_2 \cap \{\Pi_n(R) = j\}) \frac{P\big(\Pi_n(\overline{R}) = n - j\big)}{P(\Pi_n(1) = n)} \\ &\quad \le \frac{P(E_2)}{P(\Pi_n(1) = n)} \max_{0 \le j \le n} P\big(\Pi_n(\overline{R}) = n - j\big) \\ &\quad \le \frac{P(E_2)}{P(\Pi_n(1) = n)} P\big(\Pi_n(\overline{R}) = [n|\overline{R}|]\big), \end{aligned}$$

where $[u] \le u < [u] + 1$ denotes the integer part of u. Here, we have used the facts that $\Pi_n(\overline{R})$ follows a Poisson distribution with parameter $n|\overline{R}|$, and that [see, e.g., Johnson and Kotz (1969), page 92] $P(\Pi_n(\Lambda) = j) \le P(\Pi_n(\Lambda) = [n|\Lambda|])$, for all j and $\Lambda \subset [0, \infty)$. Finally, using Stirling's formula [recall that

$n! = (n/e)^n(2\pi n)^{1/2}\exp(\theta_n/n)$, with $0 < \theta_n < 1/12$ for $n \geq 1$], we obtain that

$$\frac{P(\Pi_n(\overline{R}) = [n|\overline{R}|])}{P(\Pi_n(1) = n)} \leq \frac{n^{1/2}e^{1/12}}{\left(\frac{1}{2}n - 1\right)^{1/2}} \leq 2, \quad \text{for } n \geq 5,$$

which suffices for (2.9). □

REMARK 2.1. The proof of Lemma 2.1 was largely based on the proof of inequality 2.3 of Einmahl (1987). Note for further use that for (2.9) we need not require any specific assumption imposed on $0 < a_n < 1$.

LEMMA 2.2. *For any subset A of $I_{\mathrm{RC}}(0,1)$ and $c > 0$, set*

$$J_{h,c}(A) = \inf_{f \in A} J_{h,c}(f), \tag{2.10}$$

where for $f \in I_{\mathrm{RC}}(0,1)$ and $\dot{f}$ being as in (1.11),

$$J_{h,c}(f) = \begin{cases} c - f(1) + \int_0^1 \dot{f}(u)\log(\dot{f}(u)/c)\,du, & \text{for } f_s(1) = 0; \\ \infty, & \text{otherwise}. \end{cases} \tag{2.11}$$

Let $\{\Pi(t),\ t \geq 0\}$ be a standard right-continuous Poisson process. Then the following hold:

(i) *For any closed subset F of $I_{\mathrm{RC}}(0,1)$, endowed with the Skorohod topology, we have*

$$\limsup_{T \to \infty} T^{-1} \log P(T^{-1}\Pi(cT\,\cdot) \in F) \leq -J_{h,c}(F). \tag{2.12}$$

(ii) *For any open subset G of $I_{\mathrm{RC}}(0,1)$, endowed with the Skorohod topology, we have*

$$\liminf_{T \to \infty} T^{-1} \log P(T^{-1}\Pi(cT\,\cdot) \in G) \geq -J_{h,c}(G). \tag{2.13}$$

PROOF. This result is a consequence of the large deviation principle of Varadhan (1966) [see, e.g., Example 1 in Lynch and Sethuraman (1987)]. □

Note here that Theorems 4.1 and 4.2 of Lynch and Sethuraman (1987) cover the case where, in Lemma 2.2, F is closed (resp., G is open) when $I_{\mathrm{RC}}(0,1)$ is endowed with the weak topology. This, however, is not quite sufficient for our needs since a set may be closed (or open) in the Skorohod topology but not in the weak topology.

In the sequel, we will make use of the fact that whenever $\varepsilon > 0$ and $f \in \Delta_c$, $\tilde{N}_\varepsilon(f) = \{g \in I_{\mathrm{RC}}(0,1): \|g - f\| < \varepsilon\}$ contains an open Skorohod neighborhood of f, denoted by $\overline{\overline{N}}_\varepsilon(f)$. Likewise, for $\varepsilon > 0$, $\tilde{\Delta}_c^\varepsilon$ contains an open Skorohod neighborhood of Δ_c, denoted by $\overline{\overline{\Delta}}_c^\varepsilon$.

Our next lemma proves the second half of Theorem 2.1.

LEMMA 2.3. *Under* (S.6), *for every* $f \in \Delta_c$ *and* $\varepsilon > 0$, *there exists almost surely a finite* $N''_{\varepsilon,f}$ *such that, for all* $n \geq N''_{\varepsilon,f}$, *there exists a* $\tilde{t} = \tilde{t}_{n,\varepsilon,f} \in [0, 1 - a_n]$ *such that* (2.5) *holds.*

PROOF. For any $\varepsilon > 0$, set $\tilde{N}_\varepsilon(f) = \{g \in I_{\mathrm{RC}}(0,1): \|f - g\| < \varepsilon\} \supset \overline{\overline{N}}_\varepsilon(f)$. It is easily checked [see, e.g., the proof of Lemma 2.9 of Deheuvels and Mason (1990)] that, for $f \in \Delta_c$, $J_{h,c}(\overline{\overline{N}}_{\varepsilon/2}(f)) < 1$. Thus, for any fixed $0 \leq J_{h,c}(\overline{\overline{N}}_{\varepsilon/2}(f)) < \rho < 1$, by (2.13) and using the fact that $na_n/(c \log n) \to 1$ as $n \to \infty$, we obtain that, for all large n,

$$P\left(\frac{c}{na_n}\Pi(na_n\,\cdot) \in \tilde{N}_{\varepsilon/2}(f)\right) \geq P\left(\frac{c}{na_n}\Pi(na_n\,\cdot) \in \overline{\overline{N}}_{\varepsilon/2}(f)\right) \geq n^{-\rho}. \tag{2.14}$$

Next, observe that, for any $\lambda > 1$ such that $(\lambda - 1)(\varepsilon/2 + f(1)) < \varepsilon$, $\tilde{N}_{\lambda\varepsilon/2}(\lambda f) \subset \tilde{N}_\varepsilon(f)$. Since $na_n/(c \log n) \to 1$ as $n \to \infty$, we may choose $n_1 \geq 5$ so large that, for all $n \geq n_1$, (2.14) holds together with

$$\left(\frac{na_n}{c \log n} - 1\right)\left(\frac{\varepsilon}{2} + f(1)\right) < \varepsilon.$$

It follows that, for all $n \geq n_1$,

$$P\left(\frac{1}{\log n}\Pi(na_n\,\cdot) \in \tilde{N}_\varepsilon(f)\right) \geq n^{-\rho}. \tag{2.15}$$

We now apply Lemma 2.1 with $t_i = ia_n$, $i = 1, \dots, m_n := [1/(2a_n)]$ and $A_i = I_{\mathrm{RC}}(0,1) - \tilde{N}_\varepsilon(f)$, for $i = 1, \dots, m_n$. By (2.9) and (2.15), we have, for all $n \geq n_1$,

$$\begin{aligned} Q_n &:= P\big(\Pi_n(a_n, t_i; \cdot) \notin \tilde{N}_\varepsilon(f), i = 1, \dots, m_n\big) \\ &\leq 2P\big(H_n(a_n, t_i; \cdot) \notin \tilde{N}_\varepsilon(f), i = 1, \dots, m_n\big) \\ &= 2\left(1 - P\left(\frac{1}{\log n}\Pi(na_n\,\cdot) \in \tilde{N}_\varepsilon(f)\right)\right)^{m_n} \leq 2\exp(-n^{-\rho}m_n). \end{aligned}$$

Recalling that $m_n = (1 + o(1))n/(2c \log n)$ as $n \to \infty$, and that $\rho < 1$, we see that Q_n is ultimately less than $\exp(-n^\gamma)$ for any fixed $0 < \gamma < 1 - \rho$. From this last bound we get $\sum_{n=1}^\infty Q_n < \infty$. The conclusion of Lemma 2.3 now follows from the Borel–Cantelli lemma. □

REMARK 2.2. A close look at the arguments used in the proof of Lemma 2.3 shows that we need only assume that $(\log m_n)/\log n \to 1$ as $n \to \infty$. Thus, for any sequence I_n of subintervals of $[0, 1 - a_n]$ such that

$$(\log(n|I_n|/\log n))/\log n \to 1 \quad \text{as } n \to \infty,$$

there exists with probability 1, for all n sufficiently large, a $\tilde{t} = \tilde{t}_n \in I_n$ such that (2.5) holds. This condition is always satisfied when $I_n = (c_1, c_2)$ is a fixed nonvoid subinterval of $(0, 1)$.

For the proof of the first part of Theorem 2.1, we will make use of the following blocking argument. Introduce the sequence $\nu_k = [(1+\gamma)^k]$, $k = 1, 2, \dots,$ for some constant $\gamma > 0$, and consider, to start with, the case where exactly $a_n = (c \log n)/n$. Note that $\nu_{k+1}/\nu_k \to 1 + \gamma$ and $a_{\nu_{k+1}}/a_{\nu_k} \to 1/(1+\gamma)$ as $k \to \infty$. Since $n(U_n(t+sa) - U_n(t))$ is, for $s \geq 0$ and t fixed, a nondecreasing function of $n \geq 1$ and $a > 0$, it follows that, for all k sufficiently large, we have the inequalities

$$(1+\gamma)^{-2}\eta_{\nu_k}\big(a_{\nu_k}, t; (1+\gamma)^{-2}s\big) \leq \eta_n(a_n, t; s) \leq (1+\gamma)^2 \eta_{\nu_{k+1}}\big(a_{\nu_k}, t; s\big), \tag{2.16}$$

for all $\nu_k \leq n \leq \nu_{k+1}$, $s \geq 0$ and $0 \leq t \leq 1$.

We start by showing that a version of (2.4) is valid along the sequence $\{\nu_k\}$.

Lemma 2.4. *Under* (S.6), *for any* $\gamma > 0$ *and* $\varepsilon > 0$, (2.4) *holds with probability* 1 *ultimately in* k *along* $n = \nu_k$.

Proof. We start by the observation that, by Theorem 1(I) of Mason, Shorack and Wellner (1983), for any fixed $0 < \lambda < 1$, we have almost surely

$$\begin{aligned} &\limsup_{n\to\infty} \sup_{\substack{0 \leq t', t'' \leq 1 - a_n \\ |t' - t''| \leq \lambda a_n}} \|\eta_n(a_n, t'; \cdot) - \eta_n(a_n, t''; \cdot)\| \\ &\quad \leq 2 \limsup_{n\to\infty} \sup_{0 \leq t \leq 1 - \lambda a_n} \sup_{0 \leq h \leq \lambda a_n} \left\{ \frac{n}{\log n} |U_n(t+h) - U_n(t)| \right\} \\ &\quad = 2(\lambda c)\delta_{\lambda c}^+, \end{aligned} \tag{2.17}$$

where δ_C^+ is as in (1.4). Routine analysis shows that $C\delta_C^+ \sim -1/\log C \to 0$ as $C \to 0$. Thus, by choosing $\lambda > 0$ so small that $2(\lambda c)\delta_{\lambda c}^+ < \varepsilon/2$ there exists almost surely an $n_2 < \infty$ such that, for all $n \geq n_2$, if $M_n := [(1 - a_n)/(\lambda a_n)]$,

$$\big(\{\eta_n(a_n, j\lambda a_n; \cdot), j = 0, 1, \dots, M_n\} \subset \tilde{\Delta}_c^{\varepsilon/2}\big) \Rightarrow \big(\{\eta_n(a_n, t; \cdot), 0 \leq t \leq 1 - a_n\} \subset \tilde{\Delta}_c^{\varepsilon}\big). \tag{2.18}$$

Let $F_\varepsilon = I_{\mathrm{RC}}(0,1) - \tilde{\Delta}_c^{\varepsilon/2} \subset \overline{\overline{F}}_\varepsilon = I_{\mathrm{RC}}(0,1) - \overline{\overline{\Delta}}{}_c^{\varepsilon/2}$. Obviously, $\overline{\overline{F}}_\varepsilon$ is Skorohod-closed and such that $J_{h,c}(\overline{\overline{F}}_\varepsilon) > 1$ [see, e.g., the proof of Lemma 2.8 of Deheuvels and Mason (1990)]. By Lemma 2.1 applied with $m_n = 1$, it follows that, for all large n,

$$\begin{aligned} R_n &:= P\bigg(\bigcup_{0 \leq j \leq M_n} \{\eta_n(a_n, j\lambda a_n; \cdot) \in F_\varepsilon\}\bigg) \leq 4M_n P\big(H_n(a_n, 0; \cdot) \in \overline{\overline{F}}_\varepsilon\big) \\ &\leq \frac{8n}{\lambda c \log n} P\left(\frac{1}{\log n}\Pi(na_n \cdot) \in \overline{\overline{F}}_\varepsilon\right), \end{aligned}$$

which by an application of (2.12) and by using the same argument as in the proof of (2.15) is ultimately less than or equal to $n^{-\omega}$ for some $\omega > 0$. Since this evidently implies that $\sum_k R_{\nu_k} < \infty$, the Borel–Cantelli lemma in combination with (2.18) completes the proof of Lemma 2.4. □

LEMMA 2.5. *Under* (S.6), *for any* $\varepsilon > 0$, *there exists a* $\gamma > 0$ *such that almost surely*

$$(2.19)\quad \limsup_{k\to\infty} \sup_{0\le t\le 1-a_{\nu_k}} \left\| \frac{\log \nu_{k+1}}{\log \nu_k} \eta_{\nu_{k+1}}(a_{\nu_k}, t; \cdot) - \eta_{\nu_k}(a_{\nu_k}, t; \cdot) \right\| < \varepsilon.$$

PROOF. Set for $0 < a < 1$ and $n \ge 1$, $\omega_n(a) = \sup\{\|\xi_n(a, t; \cdot)\|: 0 \le t \le 1 - a\}$. By inequality 1 in Mason, Shorack and Wellner [(1983), page 86], for every $0 \le a \le \delta \le \frac{1}{2}$, $n \ge 1$ and $\lambda > 0$, we have

$$(2.20)\quad P\big(\omega_n(a) \ge \lambda\sqrt{a}\big) \le 20a^{-1}\delta^{-3}\exp\left(-(1-\delta)^4\frac{\lambda^2}{2}\psi\left(\frac{\lambda}{\sqrt{na}}\right)\right),$$

where $\psi(x) = 2x^{-2}h(1+x)$. Note for further use that $\psi(x)\uparrow 1$ as $x \downarrow 0$.

Let $N_k = \nu_{k+1} - \nu_k$ and observe that

$$\begin{aligned}
&\left\| \frac{\log \nu_{k+1}}{\log \nu_k} \eta_{\nu_{k+1}}(a_{\nu_k}, t; \cdot) - \eta_{\nu_k}(a_{\nu_k}, t; \cdot) \right\| \\
&\quad =_d \frac{\log N_k}{\log \nu_k} \|\eta_{N_k}(a_{\nu_k}, t; \cdot)\| \\
&\quad \le \frac{N_k}{\log \nu_k} a_{\nu_k} + \frac{N_k^{1/2}}{\log \nu_k} \omega_{N_k}(a_{\nu_k}) =: \Lambda_{1,k} + \Lambda_{2,k}.
\end{aligned}$$

Since $\Lambda_{1,k} = (1 + o(1))\gamma c$ as $k \to \infty$, a choice of $0 < \gamma < \varepsilon/(4c)$ ensures that, for all k sufficiently large, $\Lambda_{1,k} < \varepsilon/2$. Next, choose $a = a_{\nu_k}$, $n = N_k$, $\lambda = (\varepsilon/2)(\log \nu_k)/\sqrt{N_k a_{\nu_k}}$ and $\delta = \frac{1}{2}$ in (2.20). Obviously, we have as $k \to \infty$

$$\lambda^2 = (1 + o(1))\frac{\varepsilon^2}{4\gamma c}\log \nu_k \quad\text{and}\quad \frac{\lambda}{\sqrt{na}} = (1 + o(1))\frac{\varepsilon}{2\gamma c}.$$

Therefore, there exists a $\gamma_0 < \varepsilon/(4c)$ such that, for any $0 < \gamma < \gamma_0$, we have for all k sufficiently large the inequality [recall that $x\psi(x) \to \infty$ as $x \to \infty$]

$$(1-\delta)^4\frac{\lambda^2}{2}\psi\left(\frac{\lambda}{\sqrt{na}}\right) = \frac{\lambda^2}{32}\psi\left(\frac{\lambda}{\sqrt{na}}\right) \ge 3\log \nu_k.$$

This in turn implies by (2.20) that $P(\Lambda_{2,k} \ge \varepsilon/2)$ is ultimately less than ν_k^{-2}. Since $\sum_k \nu_k^{-2} < \infty$, the Borel–Cantelli lemma completes the proof of Lemma 2.5. □

PROOF OF THEOREM 2.1. We now collect the pieces of our puzzle. In a first step, we observe [see, e.g., Example 6 in Deheuvels and Mason (1991)] that, for any $0 < t < 1$,

$$\sup_{f \in \Delta_c} \sup_{\substack{0 \le s', s'' \le 1 \\ |s' - s''| \le t}} |f(s') - f(s'')| \le (ct)\delta_{ct}^+, \tag{2.21}$$

and

$$\sup_{f \in \Delta_c} \sup_{0 \le s \le 1} |f(s)| = c\delta_c^+. \tag{2.22}$$

It follows from (2.21) and (2.22) that for any $\theta > 0$,

$$\begin{aligned} &\sup_{f \in \tilde{\Delta}_c^\theta} \sup_{\substack{0 \le t' \le t'' \le 1 \\ t''/t' \le (1+\gamma)^2}} |f(t') - f(t'')| \\ &\quad \le c\big((1+\gamma)^2 - 1\big)\delta_{c((1+\gamma)^2-1)}^+ + 2\theta =: \theta_1(\gamma, \theta), \end{aligned} \tag{2.23}$$

and

$$\sup_{f \in \tilde{\Delta}_c^\theta} \sup_{(1+\gamma)^{-2} \le \Lambda \le 1} \|\Lambda f - f\| \le \big(1 - (1+\gamma)^{-2}\big)(\theta + c\delta_c^+) =: \theta_2(\gamma, \theta). \tag{2.24}$$

Thus, combining (2.23) and (2.24), we see that

$$\sup_{f \in \tilde{\Delta}_c^\theta} \|(1+\gamma)^{-2} f\big((1+\gamma)^{-2} \cdot\big) - f\| \le \theta_1(\gamma, \theta) + \theta_2(\gamma, \theta) := \theta_3(\gamma, \theta). \tag{2.25}$$

It is easily verified that $\theta_3(\gamma, \theta) \to 2\theta$ as $\gamma \downarrow 0$. Choose now $\theta = \varepsilon/8$, and select a $\gamma_1 > 0$ such that, for all $0 < \gamma \le \gamma_1$, $\theta_3(\gamma, \varepsilon/8) \le 3\varepsilon/8$. An application of Lemma 2.4 shows that, for any choice of $0 < \gamma \le \gamma_1$, the event $\{\eta_{\nu_k}(a_{\nu_k}, t; \cdot) \in \tilde{\Delta}_c^{\varepsilon/8}$ for $0 \le t \le 1 - a_{\nu_k}\}$ holds almost surely ultimately in k. This in turn implies that we have almost surely, for all k sufficiently large and all $0 \le t \le 1 - a_{\nu_k}$,

$$\Big\|(1+\gamma)^{-2}\eta_{\nu_k}\big(a_{\nu_k}, t; (1+\gamma)^{-2} \cdot\big) - \eta_{\nu_k}(a_{\nu_k}, t; \cdot)\Big\| \le 3\varepsilon/8. \tag{2.26}$$

Next, by Lemma 2.5, we see that there exists a $0 < \gamma_2 \le \gamma_1$ such that, for any choice of $0 < \gamma \le \gamma_2$, we have almost surely ultimately in k, for all $0 \le t \le 1 - a_{\nu_k}$, the event

$$\left\|\eta_{\nu_k}(a_{\nu_k}, t; \cdot) - \frac{\log \nu_{k+1}}{\log \nu_k}\eta_{\nu_{k+1}}(a_{\nu_k}, t; \cdot)\right\| \le \frac{\varepsilon}{8}, \tag{2.27}$$

which in turn implies that

$$\frac{\log \nu_{k+1}}{\log \nu_k}\eta_{\nu_{k+1}}(a_{\nu_k}, t; \cdot) \in \tilde{\Delta}_c^{\varepsilon/4}$$

and by (2.22) that

$$\begin{aligned} &\left\|\frac{\log \nu_{k+1}}{\log \nu_k}\eta_{\nu_{k+1}}(a_{\nu_k}, t; \cdot) - (1+\gamma)^2\eta_{\nu_{k+1}}(a_{\nu_k}, t; \cdot)\right\| \\ &\quad \le \left|(1+\gamma)^2\frac{\log \nu_k}{\log \nu_{k+1}} - 1\right|\left(\frac{\varepsilon}{4} + c\delta_c^+\right) =: \theta_4(\gamma, \varepsilon, k). \end{aligned} \tag{2.28}$$

Since $(\log \nu_k)/\log \nu_{k+1} \to 1$ as $k \to \infty$, we may select a $0 < \gamma \le \gamma_2$ such that for all large k, $\theta_4(\gamma, \varepsilon, k) \le \varepsilon/8$. Thus, by (2.16) and (2.26)–(2.28), we have almost surely for all k sufficiently large,

$$(2.29)\quad \left\| \eta_n(a_n, t; \cdot) - \eta_{\nu_k}(a_{\nu_k}, t; \cdot) \right\| \le 5\varepsilon/8 \quad \text{and} \quad \eta_{\nu_k}(a_{\nu_k}, t; \cdot) \in \tilde{\Delta}_c^{\varepsilon/8},$$

for all $\nu_k \le n \le \nu_{k+1}$ and $0 \le t \le 1 - a_{\nu_k}$.

By (2.29), it follows that the event $\{\eta_n(a_n, t; \cdot) \in \tilde{\Delta}_c^{3\varepsilon/4}\}$ holds almost surely for all $\nu_k \le n \le \nu_{k+1}$, $0 \le t \le 1 - a_{\nu_k}$ and k sufficiently large. To complete the Proof of Theorem 2.1, it suffices to consider the case where for $\nu_k \le n \le \nu_{k+1}$, $1 - a_{\nu_k} \le t \le 1 - a_n$. For this, notice that, for $\nu_k \le n \le \nu_{k+1}$,

$$(2.30)\quad \begin{aligned} &\sup_{1 - a_{\nu_k} \le t \le 1 - a_n} \left\| \eta_n(a_n, t; \cdot) - \eta_n(a_n, 1 - a_{\nu_k}; \cdot) \right\| \\ &\quad \le 2 \sup_{0 \le t \le 1 - (a_{\nu_k} - a_n)} \ \sup_{0 \le h \le a_{\nu_k} - a_n} \left\{ \frac{n}{\log n} |U_n(t + h) - U_n(t)| \right\}. \end{aligned}$$

Since for $\nu_k \le n \le \nu_{k+1}$,

$$a_{\nu_k} - a_n \le a_{\nu_k} - a_{\nu_{k+1}} = (1 + o(1)) \left(\frac{\gamma}{1 + \gamma} \right) a_{\nu_k} \quad \text{as } k \to \infty,$$

we can choose here $\gamma > 0$ such that, uniformly over $\nu_k \le n \le \nu_{k+1}$ for all large k, $a_{\nu_k} - a_n \le \lambda a_n$, where $\lambda > 0$ is a constant chosen in such a way that $2(\lambda c)\delta_{\lambda c}^+ \le \varepsilon/8$. By (2.17), it follows that the right-hand-side of (2.30) can be almost surely rendered less than $\varepsilon/4$ for all k sufficiently large. Since $\eta_n(a_n, 1 - a_{\nu_k}; \cdot) \in \tilde{\Delta}_c^{3\varepsilon/4}$ a.s. for all large n, it follows that $\eta_n(a_n, t; \cdot) \in \tilde{\Delta}_c^{\varepsilon}$ a.s. for $\nu_k \le n \le \nu_{k+1}$, $1 - a_{\nu_k} \le t \le 1 - a_n$ and k sufficiently large. Thus, we have (2.4), as sought.

We have just proved the first part of Theorem 2.1 in the case where $a_n = (c \log n)/n$. When a_n is a general sequence satisfying (S.6), we can readily use the fact, by (2.17), that, uniformly over $0 \le t \le 1$, we have almost surely

$$(2.31)\quad \left\| \eta_n(a_n, t; \cdot) - \eta_n\left(\frac{c \log n}{n}, t; \cdot \right) \right\| \to 0.$$

Thus, by (2.31), the proof of (2.4) can be reduced to the just-treated case where $a_n = (c \log n)/n$. This, in combination with Lemma 2.3, completes the proof of Theorem 2.1. □

PROOF OF THEOREM 2.2. The proof of Theorem 2.2, given Theorem 2.1, follows the lines of the proof of Theorem 2.2 in Deheuvels and Mason (1990), given Theorem 2.1 of the same paper. Following their notation, we define, for any $v > 0$ and $w \ge 0$,

$$(2.32)\quad \Delta_{v,w} = \{ f \in \Delta_v : f(1) = w \} \quad \text{and} \quad \Gamma_{w,v} = \{ g \in \Gamma_w : g(1+) = v \}.$$

Moreover, for any $f \in I_{RC}(0,1)$ such that $w = f(1) > 0$ and for any $g \in I(0,1)$ such that $v = g(1+) > 0$, we define $f \in I(0,1)$ and $g \in I_{RC}(0,1)$ by

$$(2.33)\quad \begin{aligned} f^{\leftarrow}(s) &= \inf\{t\colon 0 \le t \le 1,\ f(t) \ge sf(1)\}, \quad \text{for } 0 \le s \le 1,\\ f^{\leftarrow}(1+) &= f^{\leftarrow}(s) = 1, \quad \text{for } s > 1, \end{aligned}$$

and

$$(2.34)\quad \begin{aligned} g^{\rightarrow}(t) &= \sup\{s\colon 0 \le s \le 1, g(s) \le tg(1+)\}, \quad \text{for } 0 \le t \le 1,\\ g^{\rightarrow}(0-) &= 0 \quad \text{and} \quad g^{\rightarrow}(t) = 1, \quad \text{for } t > 1. \end{aligned}$$

The following facts will be instrumental for our needs.

FACT 1. For any $v > 0$ and $w > 0$, $\Delta_{v,w}$ is nonvoid if and only if $\Gamma_{w,v}$ is nonvoid. This condition is satisfied if and only if one of the following equivalent set of inequalities holds:

$$(2.35)\qquad v\delta_v^- \le w \le v\delta_v^+ \quad \text{or} \quad w\gamma_w^- \le v \le w\gamma_w^+.$$

Moreover, in this case, $f \to vf^{\leftarrow}$ (resp., $g \to wg^{\rightarrow}$) defines a one-to-one mapping of $\Delta_{v,w}$ onto $\Gamma_{w,v}$ (resp., of $\Gamma_{w,v}$ onto $\Delta_{v,w}$). These mappings have the property that, for any $f \in \Delta_{v,w}$, if $g = vf^{\leftarrow}$, then $f = wg^{\rightarrow}$ and conversely, if $f = wg^{\rightarrow}$, then $g = vf^{\leftarrow}$.

PROOF. See, for example, Lemmas 2.3 and 2.4 in Deheuvels and Mason (1990). □

FACT 2. For any $v > 0$ (resp., $w > 0$), the set Δ_v (resp., Γ_w) is a compact subset of $I_{RC}(0,1)$ [resp., $I(0,1)$] when endowed with the topology of weak convergence.

PROOF. See, for example, Definitions 2.1 and 2.2, Theorem 3.1 and Examples 1 and 2 in Lynch and Sethuraman (1987). □

We now complete the proof of Theorem 2.2 in the following steps. For notational convenience, we assume from now on that $na_n/\log n \to w \in (0,\infty)$ as $n \to \infty$.

STEP 1. We first observe that Theorem 1(I) of Mason (1984) [see, e.g., (1.5)] implies that, for any fixed $\varepsilon > 0$, the set of functions $\{v_n(a_n, t; \cdot), 0 \le t \le 1 - a_n\} \subset I(0,1)$ is almost surely ultimately included in the set $\{g \in I(0,1)\colon g(1+) \le w\gamma_w^+ + \varepsilon\}$. Since this set is weakly compact in $I(0,1)$, it follows that $\{v_n(a_n, t; \cdot), 0 \le t \le 1 - a_n\}$ is almost surely relatively compact [with respect to the weak topology on $I(0,1)$].

STEP 2. Consider a sequence of functions of the form $g_{N_k} = v_{N_k}(a_{N_k}, t_{N_k}; \cdot)$, $k \ge 1$, where $1 \le N_1 < N_2 < \cdots$, and $0 \le t_{N_K} \le 1 - a_{N_k}$ for $k \ge 1$. Assume that this sequence is weakly convergent in $I(0,1)$ to a function $g \in I(0,1)$ and set $v = g(1+)$. By (1.5), we have $0 < w\gamma_w^- \le v \le w\gamma_w^+ < \infty$. Let now $a'_{N_k} := V_{N_k}(t_{N_k} + a_{N_k} +) - V_{N_k}(t_{N_k})$ for $k \ge 1$. We have, almost surely,

$$N_k a'_{N_k}/\log N_k \to v \in (0,\infty) \quad \text{as } k \to \infty. \tag{2.36}$$

By (2.31) [notice that (2.31) holds whenever $na_n/\log n \to c$ a.s. as $n \to \infty$] taken with $c = v$, it follows from (2.30) that almost surely uniformly over $0 \le t \le 1 - a'_{N_k}$,

$$\left\| \eta_{N_k}(a'_{N_k}, t; \cdot) - \eta_{N_k}(vN_k^{-1}\log N_k, t; \cdot) \right\| \to 0 \quad \text{as } k \to \infty. \tag{2.37}$$

By Theorem 2.1, there exists a subsequence $\{n_k,\ k \ge 1\}$ of $\{N_k,\ k \ge 1\}$ such that, if $t'_{n_k} := V_{n_k}(t_{n_k})$, we have

$$\left\| \eta_{n_k}(vn_k^{-1}\log n_k, t'_{n_k}; \cdot) - f \right\| \to 0 \quad \text{a.s. as } k \to \infty, \tag{2.38}$$

where f is a suitable function in Δ_v. By (2.37) and (2.38), if we let $f_{n_k} := \eta_{n_k}(a'_{n_k}, t'_{n_k}; \cdot)$ we see that $\|f_{n_k} - f\| \to 0$ a.s. as $k \to \infty$. We will make use of this fact to show that $f \in \Delta_{v,w}$ and then, that $g \in \Gamma_w$.

For this, observe that

$$|U_n(V_n(t)) - t| \le 1/n, \quad \text{for all } 0 \le t \le 1. \tag{2.39}$$

Thus, by (2.39),

$$\begin{aligned} |f_{n_k}(1) - w| &= \left| \frac{n_k}{\log n_k}\left(U_{n_k}\left(V_{n_k}(t_{n_k} + a_{n_k} +)\right) - U_{n_k}\left(V_{n_k}(t_{n_k})\right)\right) - w \right| \\ &= \left| \frac{n_k a_{n_k}}{\log n_k} - w \right| + O\left(\frac{1}{\log n_k}\right) = o(1) \quad \text{as } k \to \infty, \end{aligned}$$

from where it follows that $f(1) = \lim_{k\to\infty} f_{n_k}(1) = w$ and $f \in \Delta_{v,w}$.

Recalling that $U_n(t) = \sup\{s\colon V_n(s) \ge t\}$, it follows from (2.39) that, uniformly over $0 \le t \le 1$, almost surely

$$\begin{aligned} wg_{n_k}^{\rightarrow}(t) &= \frac{w}{a_{n_k}}\left(U_{n_k}(t'_{n_k} + ta'_{n_k}) - U_{n_k}(t'_{n_k}) + U_{n_k}(t'_{n_k}) - t_{n_k}\right) \\ &= \frac{w\log n_k}{n_k a_{n_k}} f_{n_k}(t) + O\left(\frac{1}{\log n_k}\right) \to f(t) \quad \text{as } k \to \infty. \end{aligned} \tag{2.40}$$

Since $g_{n_k} \to g$ a.s. as $k \to \infty$, with respect to the weak topology in $I(0,1)$, we must have $g_{n_k}(1+) \to g(1+) = v$. It follows readily that $wg_{n_k}^{\rightarrow} \to wg^{\rightarrow}$, which by (2.40) implies that $wg^{\rightarrow} = f$. By Fact 1, this implies that $g = vf^{\leftarrow} \in \Gamma_{v,w} \subset \Gamma_w$.

STEP 3. Fix any $\varepsilon > 0$ and consider $\hat{\Gamma}_w^{\varepsilon}$. If we do not have ultimately in n the inclusion $\{v_n(a_n, t; \cdot),\ 0 \le t \le 1 - a_n\} \subset \hat{\Gamma}_w^{\varepsilon}$, then there exists a sequence $g_{N_k} = v_{N_k}(a_{N_k}, t_{N_k}; \cdot)$ with $N_k \to \infty$ and $0 \le t_{N_k} \le 1 - a_{N_k}$, such that $g_{N_k} \notin \hat{\Gamma}_w^{\varepsilon}$ for each $k \ge 1$. By Steps 1 and 2, we can almost surely extract a subsequence $g_{n_k} \to g \in \Gamma_w$ as $k \to \infty$. Since then we must have $d_L(g_{n_k}, g) \to 0$, we obtain a contradiction. This proves (2.6).

STEP 4. Choose any function $g \in \Gamma_w$ and let $v = g(1+)$. By Fact 1, we have $0 < w\gamma_w^- < v < w\gamma_w^+ < \infty$ and $f := wg^{\rightarrow} \in \Delta_{v,w} \subset \Delta_v$. Fix a $\theta > 0$, and define $f^{(\theta)}$ as follows:

$$f^{(\theta)}(t) = \begin{cases} v(1+\theta)t, & \text{for } 0 \le t \le \dfrac{\theta}{2(1+\theta)}, \\ f\left(\left(t - \dfrac{\theta}{2(1+\theta)}\right)(1+\theta)\right) + \dfrac{v\theta}{2}, & \\ \qquad\qquad \text{for } \dfrac{\theta}{2(1+\theta)} \le t \le 1 - \dfrac{\theta}{2(1+\theta)}, & \\ f(1) + \dfrac{v\theta}{2} + v(1+\theta)\left(t - 1 + \dfrac{\theta}{2(1+\theta)}\right), & \\ \qquad \text{for } 1 - \dfrac{\theta}{2(1+\theta)} \le t. & \end{cases}$$

It is readily verified that $f^{(\theta)} \in \Delta_{v(1+\theta)}$. Set $b_n' := v(1+\theta)((\log n)/n)$, and fix $0 < c_1 < c_2 < 1$ and $\varepsilon_1 > 0$. By Lemma 2.3 and Remark 2.2, there exists almost surely a sequence $c_1 \le s_n' \le c_2$ and a finite N_0 such that, for all $n \ge N_0$,

$$\|\eta_n(b_n', s_n'; \cdot) - f^{(\theta)}\| < \varepsilon_1. \tag{2.41}$$

Let $f_n^{(\theta)}(t) = \eta_n(b_n', s_n'; t)$, for $t \ge 0$, and set $g_n^{(\theta)}(s) = \inf\{t: f_n^{(\theta)}(t) \ge s\}$ and $g^{(\theta)}(s) = \inf\{t: f^{(\theta)}(t) \ge s\}$, for $s \ge 0$. It is readily verified that

$$g^{(\theta)}(s) = \begin{cases} \dfrac{s}{v(1+\theta)}, & \text{for } 0 \le s \le \dfrac{v\theta}{2}, \\ \dfrac{1}{v(1+\theta)} g\left(\dfrac{1}{w}\left(s - \dfrac{v\theta}{2}\right)\right) + \dfrac{\theta}{2(1+\theta)}, & \text{for } \dfrac{v\theta}{2} < s \le \dfrac{v\theta}{2} + w, \\ \dfrac{1}{1+\theta} + \dfrac{1}{v(1+\theta)}\left(s - w - \dfrac{v\theta}{2}\right), & \text{for } \dfrac{v\theta}{2} + w < s, \end{cases} \tag{2.42}$$

and, whenever (2.41) holds,

$$g_n^{(\theta)}(s) = \frac{1}{v(1+\theta)}\left(\frac{n}{\log n}\right)\left(V_n\left(U_n(s_n') + \frac{s\log n}{n}\right) - s_n'\right), \tag{2.43}$$
$$\text{for } 0 \le s \le w + v\theta - \varepsilon_1.$$

Observe that, for all $n \ge N_0$, by (2.41),

$$g^{(\theta)}(s - \varepsilon_1) \le g_n^{(\theta)}(s) \le g^{(\theta)}(s + \varepsilon_1), \quad \text{for } 0 \le s \le w + v\theta - \varepsilon_1, \tag{2.44}$$

where we use the convention that $g^{(\theta)}(s) = 0$ for $s < 0$. Assume, from now on, that $n \geq N_0$ and that $\varepsilon_1 > 0$ is such that $\varepsilon_1 \leq v\theta/4$. Let $u_0' = v\theta/2 - \varepsilon_1$. By setting $s = u_0'$ in (2.44) and by (2.42), we have

$$\frac{\theta}{2(1+\theta)} - \frac{2\varepsilon_1}{v(1+\theta)} < g_n^{(\theta)}(u_0') < \frac{\theta}{2(1+\theta)}. \tag{2.45}$$

Thus, combining (2.44) and (2.45), we obtain, for $0 \leq s \leq w + v\theta - \varepsilon_1$,

$$\begin{aligned} g^{(\theta)}(s - \varepsilon_1) - \frac{\theta}{2(1+\theta)} &< g_n^{(\theta)}(s) - g_n^{(\theta)}(u_0') \\ &< g^{(\theta)}(s + \varepsilon_1) - \frac{\theta}{2(1+\theta)} + \frac{2\varepsilon_1}{v(1+\theta)}. \end{aligned} \tag{2.46}$$

Set now

$$t_n = U_n(s_n') + \frac{\log n}{n} u_0'$$

and let

$$s = u_0' + \frac{na_n}{\log n} u$$

in (2.46). Since $c_1 \leq s_n' \leq c_2$, by the Glivenko–Cantelli theorem, for any fixed $0 < c_1' < c_1 < c_2 < c_2' < 1$, there exists a.s. an $N_1 \geq N_0$ such that $t_n \in (c_1', c_2')$ for $n \geq N_1$. Moreover, since $na_n/\log n \to w$ as $n \to \infty$, there exists an $N_2 \geq N_1$ such that, for all $0 \leq u \leq 1+$,

$$\begin{aligned} \frac{v\theta}{2} - \varepsilon_1 \leq u_0' + \frac{na_n}{\log n} u &= \frac{v\theta}{2} - \varepsilon_1 + \frac{na_n}{\log n} u \\ &\leq \frac{v\theta}{2} + w + \varepsilon_1 < w + v\theta - \varepsilon_1. \end{aligned} \tag{2.47}$$

Noting that

$$g_n^{(\theta)}\left(u_0' + \frac{na_n}{\log n} u\right) - g_n^{(\theta)}(u_0') = \frac{1}{v(1+\theta)} v_n(a_n, t_n; u), \quad \text{for } 0 \leq u \leq 1+\ ,$$

we see, by setting

$$s = u_0' + \frac{na_n}{\log n} u = \frac{v\theta}{2} - \varepsilon_1 + \frac{na_n}{\log n} u$$

in (2.46) and in view of (2.42) and (2.47), that, for all $0 \leq u \leq 1+$ and $n \geq N_2$,

$$\begin{aligned} \frac{1}{v(1+\theta)} g\left(\frac{na_n}{w \log n} u - \frac{2}{w}\varepsilon_1\right) &\leq \frac{1}{v(1+\theta)} v_n(t_n, a_n; u) \\ &\leq \frac{1}{v(1+\theta)} g\left(\frac{na_n}{w \log n} u\right) + \frac{2\varepsilon_1}{v(1+\theta)}, \end{aligned} \tag{2.48}$$

where we use the convention that $g(t) = 0$ for $t < 0$ and $g(t) = g(1+\ )$ for $t > 1$.

Next, observe that we may choose $N_3 \geq N_2$ such that, for all $n \geq N_3$ and $0 \leq u \leq 1$,

$$u - \frac{3}{w}\varepsilon_1 < \frac{na_n}{w \log n}u - \frac{2}{w}\varepsilon_1 < \frac{na_n}{w \log n}u < u + \frac{3}{w}\varepsilon_1. \tag{2.49}$$

Thus, by (2.48) and (2.49), we see that, for all $n \geq N_3$,

$$d_L(v_n(a_n, t_n; \cdot), g) \leq \varepsilon_1 \max\left(\frac{3}{w}, 2\right). \tag{2.50}$$

Since $\varepsilon_1 > 0$ may be chosen as small as desired, (2.50) implies (2.7) by choosing $\varepsilon_1 = \varepsilon/\max(3/w, 2)$ and $\theta > 4\varepsilon_1/v$. This completes the proof of Theorem 2.2. □

REMARK 2.3. In the proof of Theorem 2.2, we have shown that, in (2.7), we may always choose $\hat{t} = \hat{t}_{n,\varepsilon,g} \in (c_1', c_2')$, where $0 < c_1' < c_2' < 1$ are arbitrary but fixed points of $(0,1)$. With the notation of this proof and by using the classical Chung (1949) law of the iterated logarithm for empirical processes, we see that

$$|t_n - s_n'| = O\left(n^{-1/2}(\log\log n)^{1/2}\right) \quad \text{a.s. as } n \to \infty, \tag{2.51}$$

uniformly over $s_n' \in [0,1]$. Thus, by Remark 2.2, we obtain readily that we may restrict the choice of $\hat{t}_{n,\varepsilon,g}$ to an arbitrary sequence I_n' of subintervals of $(0,1)$ such that $(\log(n|I_n'|/\log n))/\log n \to 1$ as $n \to \infty$.

REMARK 2.4. By Fact 2, we see that, for any $\varepsilon > 0$, there exists a finite set $\{g_{1,\varepsilon}, \dots, g_{N,\varepsilon}\} \subset \Gamma_w$ such that, for any $g \in \Gamma_w$, there exists an i, $1 \leq i \leq N$, with $d_L(g, g_{i,\varepsilon}) < \varepsilon/2$. By applying Theorem 2.2 to each function $g_{i,\varepsilon}$, for $i = 1, \dots, N$, we see that, for $n \geq \max(M''_{\varepsilon/2, g_{1,\varepsilon}}, \dots, M''_{\varepsilon/2, g_{N,\varepsilon}})$, we have (2.7). Thus, in the statement of Theorem 2.2, we may take $M''_{\varepsilon,g}$ independent of $g \in \Gamma_w$. Likewise, in Theorem 2.1, we may take $N''_{\varepsilon,f}$ independent of $f \in \Delta_v$.

3. Strassen-type functional LIL's for the increments of empirical processes. In order to motivate the results of this section, we cite the functional LIL's obtained for the tail processes by Mason (1988) (for α_n) and by Einmahl and Mason (1988) (for β_n). Their results are stated in Theorem C. Denote by $B(0,1)$ the set of all bounded functions on $[0,1]$. For any $f \in B(0,1)$, set $\|f\| = \sup_{0 \leq s \leq 1}|f(s)|$ and, for any $\varepsilon > 0$ and $C \subset B(0,1)$, denote by C^ε the set of all functions $f \in B(0,1)$ such that there exists an $\bar{f} = \bar{f}_{f,\varepsilon} \in C$ with $\|f - \bar{f}\| < \varepsilon$.

THEOREM C. *Under* (S.1) *and* (S.4), *the sequences of functions*

$$\left\{\alpha_n(a_n s)/(2a_n \log\log n)^{1/2}, 0 \leq s \leq 1\right\}$$

and

$$\left\{\beta_n(a_n s)/(2a_n \log\log n)^{1/2}, 0 \leq s \leq 1\right\}$$

are both almost surely relatively compact in $B(0,1)$ *endowed with the topology of uniform convergence on* $[0,1]$. *In both cases, the set of limit points is equal to the Strassen set* $\mathbb{S}_0$ *which consists of all absolutely continuous functions* f *on* $[0,1]$ *such that*

$$f(0) = 0 \quad \text{and} \quad \int_0^1 \left(\dot{f}(s)\right)^2 ds \le 1, \tag{3.1}$$

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

We now state in Theorem 3.1 below the analogue of Theorem C for the increment processes ξ_n and ζ_n as defined in (1.1).

THEOREM 3.1. *Under the CRS conditions, for any* $\varepsilon > 0$, *there exists almost surely a finite* n'_ε *such that, for all* $n \ge n'_\varepsilon$, *we have*

$$\left\{\xi_n(a_n, t; \cdot)/(2a_n \log(1/a_n))^{1/2}, 0 \le t \le 1 - a_n\right\} \subset \mathbb{S}_0^\varepsilon, \tag{3.2}$$

and

$$\left\{\zeta_n(a_n, t; \cdot)/(2a_n \log(1/a_n))^{1/2}, 0 \le t \le 1 - a_n\right\} \subset \mathbb{S}_0^\varepsilon. \tag{3.3}$$

Moreover, for any $f \in \mathbb{S}_0$ *and* $\varepsilon > 0$, *there exists almost surely a finite* $n''_{\varepsilon, f}$ *such that, for all* $n \ge n''_{\varepsilon, f}$, *there exist a* t, $0 \le t = t_{n,\varepsilon,f} \le 1 - a_n$ *and a* $\bar{t}$, $0 \le \bar{t} = \bar{t}_{n,\varepsilon,f} \le 1 - a_n$, *such that*

$$\|\xi_n(a_n, t; \cdot)/(2a_n \log(1/a_n))^{1/2} - f\| < \varepsilon \tag{3.4}$$

and

$$\|\zeta_n(a_n, \bar{t}; \cdot)/(2a_n \log(1/a_n))^{1/2} - f\| < \varepsilon. \tag{3.5}$$

In the remainder of this section, we present the proof of Theorem 3.1. Throughout and unless otherwise specified, we assume that the CRS conditions [i.e., (S.1), (S.2) and (S.3)] are satisfied. First, we consider the case of ξ_n. Let $\{\Pi_n(t), t \ge 0\}$ be a right-continuous Poisson process with parameter $n \ge 1$ [see, e.g., (2.8)]. For $0 < a < 1$, let

$$L_n(a, t; s) = n^{-1/2}(\Pi_n(t + sa) - \Pi_n(t) - nsa), \quad \text{for } 0 \le s \le 1 \text{ and } t \ge 0. \tag{3.6}$$

LEMMA 3.1. *For any choice of* $\{t_1, \dots, t_m\} \subset \{ka_n\colon 0 \le k \le a_n^{-1} - 1\}$ *with* $0 < ma_n \le \frac{1}{2}$ *and Borel subsets* $B'_1, \dots, B'_m$ *of* $B(0,1)$ *endowed with the topology of uniform convergence on* $[0,1]$, *set* $A'_1 = \{\xi_n(a_n, t_i; \cdot) \in B'_i, i = 1, \dots, m\}$ *and* $A'_2 = \{L_n(a_n, t_i; \cdot) \in B'_i, i = 1, \dots, m\}$. *We have*

$$P(A'_1) \le 2P(A'_2), \quad \text{for all } n \ge 5. \tag{3.7}$$

PROOF. The proof of Lemma 3.1 is practically the same as the proof of Lemma 2.1 after a change of scale. Therefore, we omit details. □

Our next lemma gives a large deviation result which will be instrumental for our needs. For any $f \in B(0,1)$, set

$$(3.8)\quad J(f) = \begin{cases} \int_0^1 \left(\dot{f}(s)\right)^2 ds, & \text{if } f \text{ is absolutely continuous on } [0,1] \text{ with Lebesgue derivative } \dot{f}, \\ \infty, & \text{otherwise,} \end{cases}$$

and, for any $B \subset B(0,1)$, define

$$(3.9)\qquad J(B) = \inf_{f \in B} J(f).$$

LEMMA 3.2. *Let $\{W(t),\ t \geq 0\}$ be a standard Wiener process and, for any $\lambda > 0$, set $W_\lambda(t) = 2^{-1/2}\lambda^{-1}W(\lambda t)$, for $0 \leq t \leq 1$. Then the following hold:*

(i) *For each closed subset F of $B(0,1)$ endowed with the topology of uniform convergence on* $[0,1]$,

$$(3.10)\qquad \limsup_{\lambda \to \infty} \lambda^{-1} \log P(W_\lambda \in F) \leq -J(F).$$

(ii) *For each open subset G of $B(0,1)$ endowed with the topology of uniform convergence on* $[0,1]$,

$$(3.11)\qquad \liminf_{\lambda \to \infty} \lambda^{-1} \log P(W_\lambda \in G) \geq -J(G).$$

PROOF. See, for example, Ventsel (1976). □

The following lemma establishes the second part [i.e., (3.4)] of Theorem 3.1 for ξ_n.

LEMMA 3.3. *Under the CRS conditions, for every $f \in \mathbb{S}_0$ and $\varepsilon > 0$, there exists almost surely a finite $n''_{\varepsilon, f}$ such that, for all $n \geq n''_{\varepsilon, f}$, there exists a $t = t_{n,\varepsilon,f} \in [0, 1 - a_n]$ such that* (3.4) *holds.*

PROOF. Set $N_\varepsilon(f) = \{g \in B(0,1)\colon \|f - g\| < \varepsilon\}$. By Lemma 3.1 applied with $t_i = ia_n$, $i = 1, \dots, m_n := [1/(2a_n)]$ and $B_i' := B(0,1) - N_\varepsilon(f)$, for $i = 1, \dots, m_n$, we obtain

$$\begin{aligned} P_n &:= P\left(\bigcap_{i=1}^{m_n} \left\{ \xi_n(a_n, t_i; \cdot)/(2a_n \log(1/a_n))^{1/2} \notin N_\varepsilon(f) \right\} \right) \\ (3.12)\qquad &\leq 2P\left(\bigcap_{i=1}^{m_n} \left\{ L_n(a_n, t_i; \cdot)/(2a_n \log(1/a_n))^{1/2} \notin N_\varepsilon(f) \right\} \right) \\ &= 2\left(1 - P\left(L_n(a_n, 0; \cdot)/(2a_n \log(1/a_n))^{1/2} \in N_\varepsilon(f)\right)\right)^{m_n} \\ &:= 2(1 - P_{1,n})^{m_n}. \end{aligned}$$

In a second step, we evaluate $P_{1,n}$. For this, we make use of the approximation results of Komlós, Major and Tusnády (1975a, b), which enable us to construct on the same probability space a standard Poisson process $\{\Pi(t), t \geq 0\}$ and a standard Wiener process $\{W(t), t \geq 0\}$ such that, for universal constants $C_1 > 0$, $C_2 > 0$ and $C_3 > 0$,

$$P\Big(\sup_{0 \leq x \leq T} |\Pi(x) - x - W(x)| \geq C_1 \log T + z\Big) \leq C_2 \exp(-C_3 z), \tag{3.13}$$

for all $T > 0$ and $-\infty < z < \infty$.

By (3.6), (3.12) and (3.13), we see that

$$\begin{aligned} P_{1,n} &= P\left(\frac{(\Pi(na_n \cdot) - na_n \cdot)}{(2na_n \log(1/a_n))^{1/2}} \in N_\varepsilon(f)\right) \\ &\geq P\left(\frac{W(na_n \cdot)}{(2na_n \log(1/a_n))^{1/2}} \in N_{\varepsilon/2}(f)\right) \\ &\quad - P\left(\sup_{0 \leq s \leq 1} \frac{|\Pi(na_n s) - na_n s - W(na_n s)|}{(2na_n \log(1/a_n))^{1/2}} \geq \varepsilon/2\right) \\ &=: P_{2,n} - P_{3,n}^{\varepsilon/2}. \end{aligned} \tag{3.14}$$

Notice that

$$W(na_n \cdot)/(2na_n \log(1/a_n))^{1/2} =_d 2^{-1/2}(\log(1/a_n))^{-1} W((\log(1/a_n) \cdot),$$

so that

$$P_{2,n} = P\big(W_{\log(1/a_n)} \in N_{\varepsilon/2}(f)\big).$$

Since $f \in \mathbb{S}_0$, we have obviously $J(N_{\varepsilon/2}(f)) < 1$. Thus, by (3.11), it follows that for any $\rho \in (J(N_{\varepsilon/2}(f), 1)$, we have for all n sufficiently large

$$P_{2,n} \geq \exp(-\rho \log(1/a_n)) = a_n^\rho. \tag{3.15}$$

Next, observe that (S.1) implies that

$$(\log(na_n))/(2na_n \log(1/a_n))^{1/2} \to 0 \quad \text{as } n \to \infty.$$

This, when combined with (3.13), (3.14) and (S.3), implies that, for large n,

$$\begin{aligned} P_{3,n}^{\varepsilon/2} &\leq P\Bigg(\sup_{0 \leq x \leq na_n} |\Pi(x) - x - W(x)| \\ &\qquad \geq C_1 \log(na_n) + \frac{\varepsilon}{4}\left(2na_n \log\left(\frac{1}{a_n}\right)\right)^{1/2}\Bigg) \\ &\leq C_2 \exp\left(-C_3\left(\frac{\varepsilon}{4}\right)\left(2na_n \log\left(\frac{1}{a_n}\right)\right)^{1/2}\right) \\ &\leq \frac{1}{2} \exp\left(-\rho \log\left(\frac{1}{a_n}\right)\right) = \frac{1}{2} a_n^\rho. \end{aligned} \tag{3.16}$$

By (3.14), (3.15) and (3.16), we have ultimately as $n \to \infty$,

$$P_{1,n} \geq P_{2,n} - P_{3,n}^{\varepsilon/2} \geq \tfrac{1}{2} a_n^\rho,$$

which, when combined with (3.12), yields for all large n the inequalities

$$(3.17) \quad P_n \leq 2\exp\left(-\frac{1}{2} m_n a_n^\rho\right) \leq 2\exp\left(-\frac{1}{2}\left(\frac{1}{4a_n}\right) a_n^\rho\right) = 2\exp\left(-\frac{1}{8} a_n^{\rho-1}\right).$$

Since $\rho - 1 < 0$, by (S.2), the RHS of (3.17) is ultimately less than or equal to

$$2\exp\left(-\tfrac{1}{8}(\log n)^r\right)$$

for an arbitrary $r > 1$. It follows evidently that $\Sigma_n P_n < \infty$, which by the Borel–Cantelli lemma implies (3.4), as sought. □

REMARK 3.1. A close look at the arguments used in the proof of Lemma 3.3 shows that we need only assume that $(\log m_n)/\log(1/a_n) \to 1$ as $n \to \infty$. Thus, for any sequence I_n of subintervals of $[0, 1 - a_n]$ such that $(\log(|I_n|/a_n))/\log(1/a_n) \to 1$ as $n \to \infty$, there exists with probability 1, for all n sufficiently large, a $t = t_n \in I_n$ such that (3.4) holds. This condition is always satisfied when $I_n = (c_1, c_2)$ is a fixed nonvoid subinterval of $(0, 1)$. Here recall that $|I|$ denotes the length of the interval I.

We now turn to the proof of (3.2). Fix any $\gamma > 0$ and $\varepsilon > 0$. Let $\nu_k = [(1+\gamma)^k]$, $k = 1, 2, \dots$, $b_n = (2a_n \log(1/a_n))^{1/2}$, and consider the events

$$C_k(\varepsilon, \gamma) = \Big\{(n/\nu_{k+1})^{1/2} b_{\nu_{k+1}}^{-1} \xi_n(a_{\nu_{k+1}}, t; \cdot) \notin \mathbb{S}_0^\varepsilon$$
$$\text{for some } 0 \leq t \leq 1 - a_{\nu_{k+1}} \text{ and } \nu_k < n \leq \nu_{k+1}\Big\},$$

and

$$D_k(\varepsilon, \gamma) = \Big\{b_{\nu_{k+1}}^{-1} \xi_{\nu_{k+1}}(a_{\nu_{k+1}}, t; \cdot) \notin \mathbb{S}_0^\varepsilon \text{ for some } 0 \leq t \leq 1 - a_{\nu_{k+1}}\Big\}.$$

LEMMA 3.4. *For every $\varepsilon > 0$ and $\gamma > 0$, there exists a $K = K_{\varepsilon,\gamma}$ such that $k \geq K$ implies that*

$$(3.18) \quad P(C_k(\varepsilon, \gamma)) \leq 2P(D_k(\varepsilon/2, \gamma)).$$

PROOF. Let r_i, $i = 1, 2, \dots$, be a denumeration of the rationals in $[0, 1 - a_{\nu_{k+1}}]$, and introduce the events for $i = 1, 2, \dots$ and $\nu_k < n \leq \nu_{k+1}$,

$$E_{k,i,n}(\varepsilon) = \Big\{(n/\nu_{k+1})^{1/2} b_{\nu_{k+1}}^{-1} \xi_n(a_{\nu_{k+1}}, r_i; \cdot) \notin \mathbb{S}_0^\varepsilon\Big\},$$

$$E_{k,n}(\varepsilon) = \bigcup_{i \geq 1} E_{k,i,n}(\varepsilon)$$

and

$$F_{k,i,n}(\varepsilon) = \Big\{b_{\nu_{k+1}}^{-1} \big\| \xi_{\nu_{k+1}}(a_{\nu_{k+1}}, r_i; \cdot) - (n/\nu_{k+1})^{1/2} \xi_n(a_{\nu_{k+1}}, r_i; \cdot) \big\| < \varepsilon\Big\}.$$

Observe that for any $\varepsilon' > 0$ and $\varepsilon'' > 0$, the sequences of events $\{E_{k,i,n}(\varepsilon'), i \geq 1\}$ and $\{F_{k,i,n}(\varepsilon''), i \geq 1\}$ are independent. Denote by $\bar{A}$ the complement of the event A. We have

$$P(C_k(\varepsilon,\gamma)) = \sum_{q=\nu_k+1}^{\nu_{k+1}} \sum_{i=1}^{\infty} P\left(E_{k,i,q}(\varepsilon) \cap \bigcap_{j=1}^{i-1} \bar{E}_{k,j,q}(\varepsilon) \bigcap_{r=\nu_k+1}^{q-1} \bar{E}_{r,q}(\varepsilon)\right),$$

and hence

$$\begin{aligned}
&\inf_{\nu_k < n \leq \nu_{k+1}} \inf_{m \geq 1} P(F_{k,m,n}(\varepsilon/2)) P(C_k(\varepsilon,\gamma)) \\
&\quad \leq \sum_{q=\nu_k+1}^{\nu_{k+1}} \sum_{i=1}^{\infty} P\Bigg(E_{k,i,q}(\varepsilon) \cap F_{k,i,q}(\varepsilon/2) \\
&\qquad\qquad \cap \bigcap_{j=1}^{i-1} \bar{E}_{k,j,q}(\varepsilon) \bigcap_{r=\nu_k+1}^{q-1} \bar{E}_{r,q}(\varepsilon)\Bigg) \qquad (3.19)\\
&\quad \leq P\left(\bigcup_{i=1}^{\infty} E_{k,i,\nu_{k+1}}(\varepsilon/2)\right) = P(D_k(\varepsilon/2,\gamma)).
\end{aligned}$$

Next, we see that, for any $0 < a \leq \frac{1}{2}$, $\lambda > 0$ and $n \geq 1$, one has

$$P\left(\sup_{0 \leq u \leq a} n|U_n(u) - u| \geq \lambda\right) \leq P\left(\sup_{0 \leq u \leq a} n|U_n(u) - u|/(1-u) \geq \lambda\right),$$

which by the fact that $n(U_n(u) - u)/(1-u)$ is a martingale in u [see, e.g., Shorack and Wellner (1986), pages 271 and 870], is

$$\leq E\left(n^2|U_n(a) - a|^2\right)/(\lambda(1-a))^2 \leq 2na/\lambda^2.$$

From this last inequality, we obtain that, for all $\nu_k < n \leq \nu_{k+1}$, $m \geq 1$ and all k large enough,

$$\begin{aligned}
P(\bar{F}_{k,m,n}(\varepsilon/2)) &= P\left(\nu_{k+1}^{-1/2} \sup_{0 \leq u \leq a_{\nu_{k+1}}} (\nu_{k+1} - n)|U_{\nu_{k+1}-n}(u) - u| \geq (\varepsilon/2) b_{\nu_{k+1}}\right) \\
&\leq 8(\nu_{k+1} - \nu_k) a_{\nu_{k+1}} / \left(\varepsilon^2 \nu_{k+1} b_{\nu_{k+1}}^2\right) \\
&\leq 4/\left(\varepsilon^2 \log(1/a_{\nu_{k+1}})\right) \to 0 \quad \text{as } k \to \infty.
\end{aligned}$$

Using this bound in (3.19) yields (3.18). □

LEMMA 3.5. *We have*

$$\lim_{\gamma \downarrow 0}\left(\limsup_{k\to\infty} \max_{\nu_k < n \leq \nu_{k+1}} \sup_{0 \leq t \leq 1 - a_{\nu_{k+1}}} |(n/\nu_{k+1})^{1/2} b_{\nu_{k+1}}^{-1} - b_n^{-1}| \times \|\xi_n(a_{\nu_{k+1}}, t; \cdot)\|\right) = 0 \quad a.s. \qquad (3.20)$$

PROOF. The LHS of (3.20) is less than or equal to

$$\lim_{\gamma\downarrow 0}\limsup_{k\to\infty}\left(\max_{\nu_k<n\le\nu_{k+1}}|b_n b_{\nu_{k+1}}^{-1}-(\nu_{k+1}/n)^{1/2}|\right) \times \limsup_{n\to\infty} b_n^{-1}\omega_n(a_n), \tag{3.21}$$

where $\omega_n(a_n)$ is as in (2.20). By Theorem 0.2 of Stute (1982a), we have

$$\lim_{n\to\infty} b_n^{-1}\omega_n(a_n) = 1 \quad \text{a.s.} \tag{3.22}$$

Moreover, (S.1) implies that $na_n \uparrow \infty$ and $a_n \downarrow 0$ as $n\uparrow\infty$, so that $b_n = (2a_n \log(1/a_n))^{1/2}$ is ultimately nonincreasing and such that, for $\nu_k < n \le \nu_{k+1}$,

$$0 \le b_n/b_{\nu_{k+1}} - 1 \le b_{\nu_k}/b_{\nu_{k+1}} - 1 \le (1+o(1))(\nu_{k+1}/\nu_k) - 1 \to \gamma \quad \text{as } k\to\infty.$$

This, in combination with (3.21) and (3.22), readily yields (3.20). □

LEMMA 3.6. *We have*

$$\lim_{\gamma\downarrow 0}\left(\limsup_{k\to\infty}\max_{\nu_k<n\le\nu_{k+1}}\sup_{0\le t\le 1-a_{\nu_{k+1}}} b_n^{-1}\|\xi_n(a_{\nu_{k+1}},t;\cdot)-\xi_n(a_n,t;\cdot)\|\right) = 0 \quad a.s. \tag{3.23}$$

PROOF. From $a_n\downarrow 0$ we get that for $\nu_k < n \le \nu_{k+1}$,

$$a_n - a_{\nu_{k+1}} = a_n(1 - a_{\nu_{k+1}}/a_n) \le a_n(1-a_{\nu_{k+1}}/a_{\nu_k}).$$

From $na_n\uparrow$ we obtain that, for all k large enough,

$$a_{\nu_{k+1}}/a_{\nu_k} \ge \nu_k/\nu_{k+1} \ge 1-\gamma.$$

Therefore, for $\nu_k < n \le \nu_{k+1}$ and for all k large enough,

$$a_n - a_{\nu_{k+1}} \le a_n(1-a_{\nu_{k+1}}/a_{\nu_k}) \le \gamma a_n.$$

Hence the LHS of (3.23) is less than or equal to

$$\lim_{\gamma\downarrow 0}\left(\lim_{n\to\infty} b_n^{-1}\omega_n(\gamma a_n)\right). \tag{3.24}$$

By (3.22) and (3.24) it is easy to conclude (3.23). □

In view of Lemmas 3.4–3.6, we will show in the sequel that the proof of (3.2) boils down to showing that, for every $\varepsilon > 0$ and $\gamma > 0$,

$$\sum_{k=1}^{\infty} P(D_k(\varepsilon,\gamma)) < \infty. \tag{3.25}$$

Toward this end, we will make use of the following inequality.

LEMMA 3.7. *For all* $0 < \theta < 1$, $\varepsilon > 0$ *and* $n \geq 1$, *we have*

$$\begin{aligned}(3.26)\quad & P\big(b_n^{-1}\xi_n(a_n, t; \cdot) \notin \mathbb{S}_0^{\varepsilon} \text{ for some } 0 \leq t \leq 1 - a_n\big) \\ & \leq (\theta a_n)^{-1} P\big(b_n^{-1}\xi_n(a_n, 0; \cdot) \notin \mathbb{S}_0^{\varepsilon/2}\big) + P\big(b_n^{-1}\omega_n(\theta a_n) > \varepsilon/4\big)\end{aligned}$$

PROOF. For $\theta i a_n \leq t \leq \theta(i+1)a_n$ and $i = 0, 1, \dots, [(a_n^{-1} - 1)/\theta] - 1 =: \mu_n - 1$, or for $\mu_n a_n \leq t \leq 1 - a_n$ and $i = \mu_n$, we have uniformly over $0 \leq s \leq 1$,

$$\begin{aligned}&|\xi_n(a_n, t; s) - \xi_n(a_n, i\theta a_n; s)| \\ &\quad \leq |\alpha_n(t + sa_n) - \alpha_n(i\theta a_n + sa_n)| + |\alpha_n(t) - \alpha_n(i\theta a_n)| \\ &\quad \leq 2\omega_n(\theta a_n).\end{aligned}$$

The remainder of the proof is obvious. □

We are now prepared to show that (3.25) holds for all $\varepsilon > 0$ and $\gamma > 0$. We first use Lemma 3.7 to obtain that, for any $\theta > 0$ sufficiently small,

$$\begin{aligned}P(D_k(\varepsilon, \gamma)) &\leq \big(\theta a_{\nu_{k+1}}\big)^{-1} P\big(b_{\nu_{k+1}}^{-1}\xi_{\nu_{k+1}}(a_{\nu_{k+1}}, 0; \cdot) \notin \mathbb{S}_0^{\varepsilon/2}\big) \\ &\quad + P\big(b_{\nu_{k+1}}^{-1}\omega_{\nu_{k+1}}(\theta a_{\nu_{k+1}}) > \varepsilon/4\big) \\ &=: Q_{1,k} + Q_{2,k}.\end{aligned}$$

Next, we use Lemma 3.1 to obtain the inequality

$$P\big(b_{\nu_{k+1}}^{-1}\xi_{\nu_{k+1}}(a_{\nu_{k+1}}, 0; \cdot) \notin \mathbb{S}_0^{\varepsilon/2}\big) \leq 2P\big(b_{\nu_{k+1}}^{-1}L(a_{\nu_{k+1}}, 0; \cdot) \notin \mathbb{S}_0^{\varepsilon/2}\big) =: 2Q_{3,k}.$$

We now follow the arguments of the proof of Lemma 3.3 to obtain the following analogue of (3.14), with $P_{3,n}^{\varepsilon/2}$ being as in (3.16):

$$(3.27)\qquad Q_{3,k} \leq P\big(W_{\log(1/a_{\nu_{k+1}})} \notin \mathbb{S}_0^{\varepsilon/4}\big) + P_{3,\nu_{k+1}}^{\varepsilon/4} =: Q_{4,k} + Q_{5,k}.$$

Since $B(0,1) - \mathbb{S}_0^{\varepsilon/4}$ is closed (with respect to the topology of uniform convergence) and obviously satisfies $J(B(0,1) - \mathbb{S}_0^{\varepsilon/4}) > 1$, it follows from (3.10) that, for any $\rho \in (1, J(B(0,1) - \mathbb{S}_0^{\varepsilon/4}))$, we have, for all k sufficiently large,

$$(3.28)\qquad Q_{4,k} \leq \exp\big(-\rho \log(1/a_{\nu_{k+1}})\big) = a_{\nu_{k+1}}^{\rho}.$$

Moreover, the same arguments as used for (3.15) and (3.16) show that, for all k sufficiently large,

$$(3.29)\qquad Q_{5,k} \leq a_{\nu_{k+1}}^{\rho}.$$

Thus, combining (3.27), (3.28) and (3.29), we obtain, for all large k,

$$(3.30)\quad Q_{1,k} \leq 2\big(\theta a_{\nu_{k+1}}\big)^{-1} Q_{3,k} \leq \frac{4}{\theta} a_{\nu_{k+1}}^{\rho-1} = \frac{4}{\theta}\exp\left(-(\rho - 1)\log\left(\frac{1}{a_{\nu_{k+1}}}\right)\right).$$

Since (S.2) implies that $\log(1/a_{\nu_{k+1}})/\log k \to \infty$ as $k \to \infty$, we see from (3.30) that $Q_{1,k} \le 1/k^2$ for all large k, so that

$$\sum_{k=1}^{\infty} Q_{1,k} < \infty. \tag{3.31}$$

Next, we use (2.20) taken with $\delta = \frac{1}{2}$, $a = \theta a_{\nu_{k+1}}$, $n = \nu_{k+1}$ and $\lambda = \varepsilon\theta^{-1/2}(\log(1/a_{\nu_{k+1}}))^{1/2}$ to obtain

$$\begin{aligned} Q_{2,k} &\le 160\theta^{-1}a_{\nu_{k+1}}^{-1} \\ &\quad\times\exp\left(-\frac{\varepsilon^2}{32\theta}\left(\log\left(\frac{1}{a_{\nu_{k+1}}}\right)\right)\psi\left(\frac{\varepsilon}{\theta}\left(\frac{\nu_{k+1}a_{\nu_{k+1}}}{\log(1/a_{\nu_{k+1}})}\right)^{-1/2}\right)\right). \end{aligned} \tag{3.32}$$

Since (S.3) implies that $na_n/\log(1/a_n) \to \infty$, and $\psi(x)\uparrow 1$ as $x \downarrow 0$, the RHS of (3.32) is ultimately less than or equal to

$$160\theta^{-1}\exp\left(-\left(\frac{\varepsilon^2}{64\theta} - 1\right)\log\left(\frac{1}{a_{\nu_{k+1}}}\right)\right).$$

Thus, the same argument as used for (3.31) shows that, for any $0 < \theta < \varepsilon^2/64$, we have

$$\sum_{k=1}^{\infty} Q_{2,k} < \infty. \tag{3.33}$$

Combining (3.31) and (3.33), we see that (3.25) holds as sought. By the Borel–Cantelli lemma and Lemmas 3.4–3.6, it follows that for any $\varepsilon > 0$, whenever $\gamma > 0$ is sufficiently small, the event

$$\bigcup_{\nu_k < n \le \nu_{k+1}} \left\{b_n^{-1}\xi_n(a_n, t; \cdot), 0 \le t \le 1 - a_{\nu_{k+1}}\right\} \not\subset \mathbb{S}_0^{\varepsilon}$$

holds finitely often with probability 1. The following lemma shows that the same is true for the event

$$\bigcup_{\nu_k < n \le \nu_{k+1}} \left\{b_n^{-1}\xi_n(a_n, t; \cdot), 1 - a_{\nu_{k+1}} < t \le 1 - a_n\right\} \not\subset \mathbb{S}_0^{\varepsilon},$$

thus completing the proof of (3.2).

LEMMA 3.8. *We have*

$$\begin{aligned} \lim_{\gamma\downarrow 0}\Bigg(\limsup_{k\to\infty}\ &\sup_{\nu_k < n\le\nu_{k+1}}\ \sup_{1-a_{\nu_{k+1}}\le t\le 1-a_n} b_n^{-1} \\ &\times\|\xi_n(a_n, t; \cdot) - \xi_n(a_n, 1 - a_{\nu_{k+1}}; \cdot)\|\Bigg) = 0 \quad a.s. \end{aligned} \tag{3.34}$$

PROOF. By the same arguments as used in the proof of Lemma 3.6, we see that the RHS of (3.34) is less than or equal to

$$\lim_{\gamma\downarrow 0}\left(2\lim_{n\to\infty} b_n^{-1}\omega_n(\gamma a_n)\right),$$

which equals zero almost surely by (3.22) and (3.24). □

Having completed the proof of Theorem 3.1 for ξ_n, we turn our attention to ζ_n. We will show that the corresponding part of Theorem 3.1 holds by means of a Bahadur-type representation in the spirit of the well-known Bahadur representation for sample quantiles [Bahadur (1966), Kiefer (1967, 1970)]. The following sequence of lemmas is directed toward this aim.

LEMMA 3.9. *For any $n \geq 1$, $0 \leq t \leq 1 - a_n$ and $0 \leq s \leq 1$, we have*

$$|\zeta_n(a_n, t; s) - \{\alpha_n(V_n(t + a_n s)) - \alpha_n(V_n(t))\}| \leq 2n^{-1/2}. \tag{3.35}$$

PROOF. It follows from (1.6) in Shorack (1982), in combination with the triangle inequality. □

LEMMA 3.10. *We have*

$$\begin{aligned}\limsup_{n\to\infty} \sup_{0 \leq t \leq 1 - a_n} \sup_{0 \leq s \leq 1} & n^{1/4}(a_n \log(1/a_n))^{-1/4}(\log n)^{-1/2} \\ & \times |\alpha_n(V_n(t + a_n s)) - \alpha_n(V_n(t) + a_n s)| \leq 2^{3/4} \quad a.s.\end{aligned} \tag{3.36}$$

PROOF. Fix any $\varepsilon > 0$. By Theorem 1 of Mason (1984), we have almost surely for all n sufficiently large

$$|V_n(t + a_n s) - V_n(t) - a_n s| \leq \hat{a}_n := (1 + \varepsilon)^2(2n^{-1}a_n \log(1/a_n))^{1/2}, \tag{3.37}$$

for all $0 \leq t \leq 1 - a_n$ and $0 \leq s \leq 1$.

Consider $\hat{a}_n$. This sequence obviously satisfies (ultimately in n) the CRS conditions. By Theorem 0.2 of Stute (1982a) and the conventions of Section 1, it follows that

$$\begin{aligned}\limsup_{n\to\infty} \sup_{0 \leq V_n(u) \leq 1} \sup_{0 \leq v \leq \hat{a}_n} & |\alpha_n(V_n(u)) - \alpha_n(V_n(u) + v)| \\ & \times (2\hat{a}_n \log(1/\hat{a}_n))^{-1/2} \leq 1 \quad \text{a.s.}\end{aligned} \tag{3.38}$$

Next, we observe that the CRS conditions imposed upon $\{a_n, n \geq 1\}$ imply that

$$1/2 \leq \liminf_{n\to\infty}(\log(1/\hat{a}_n))/\log n \leq \limsup_{n\to\infty}(\log(1/\hat{a}_n))/\log n \leq 1. \tag{3.39}$$

By (3.37)–(3.39), the LHS of (3.36) is almost surely less than or equal to $(1 + \varepsilon)2^{3/4}$. Since $\varepsilon > 0$ is arbitrary, we have (3.36). □

Our next lemma gives the desired Bahadur-type representation of ζ_n in terms of ξ_n.

LEMMA 3.11. *We have*

$$\begin{aligned}\lim_{n\to\infty} \sup_{0 \leq t \leq 1 - a_n} & \|\zeta_n(a_n, t; \cdot) - \xi_n(a_n, V_n(t); \cdot)\| \\ & \times (2a_n \log(1/a_n))^{-1/2} = 0 \quad a.s.\end{aligned} \tag{3.40}$$

PROOF. By (3.35) and (3.36), it is readily verified that (3.40) follows from

$$\lim_{n\to\infty} \frac{\left(n^{-1/4}(a_n \log(1/a_n))^{1/4}(\log n)^{1/2}\right)}{(a_n \log(1/a_n))^{1/2}} = \lim_{n\to\infty}\left(\frac{\log^2 n}{na_n \log(1/a_n)}\right)^{1/4} = 0. \tag{3.41}$$

For the proof of (3.41), observe that we have either $a_n = \min(a_n, n^{-1/2})$ or $a_n = \max(a_n, n^{-1/2})$. In the first case, $\log(1/a_n) \ge \frac{1}{2}\log n$, so that, by (S.3),

$$(\log^2 n)/(na_n \log(1/a_n)) \le 2(\log n)/(na_n) \to 0 \quad \text{as } n \to \infty.$$

In the second case, by (S.1),

$$\begin{aligned}(\log^2 n)/(na_n \log(1/a_n)) &= O\left((\log^2 n)/(na_n)\right)\\ &= O\left((\log^2 n)/n^{1/2}\right) \to 0 \quad \text{as } n \to \infty.\end{aligned}$$

Thus, in both cases, (3.41) holds. □

LEMMA 3.12. *For any fixed $\lambda > 1$, we have almost surely*

$$\limsup_{n\to\infty} \sup_{1-\lambda a_n \le t \le 1-a_n} \|\zeta_n(a_n, t; \cdot) - \xi_n(a_n, 1-a_n; \cdot)\| \times (2a_n \log(1/a_n))^{-1/2} \le (2(\lambda-1))^{1/2}, \tag{3.42}$$

and, for all n sufficiently large,

$$V_n(1-\lambda a_n) < 1 - a_n. \tag{3.43}$$

PROOF. By (1.8), it is readily verified that

$$\lim_{n\to\infty} (V_n(1-\lambda a_n) - (1-a_n))/a_n = 1 - \lambda < 0 \quad \text{a.s.}, \tag{3.44}$$

which implies (3.43). Moreover, by (3.44) and (3.22),

$$\begin{aligned}\limsup_{n\to\infty} \sup_{1-\lambda a_n \le t \le 1-a_n} &\|\xi_n(a_n, V_n(t); \cdot) - \xi_n(a_n, 1-a_n; \cdot)\|\\ &\times (2a_n \log(1/a_n))^{-1/2}\\ &\le \lim_{n\to\infty} b_n^{-1}\omega_n(2(\lambda-1)a_n) = (2(\lambda-1))^{1/2} \quad \text{a.s.},\end{aligned} \tag{3.45}$$

which, when combined with (3.40), yields (3.42). □

We have now in hand all the ingredients needed for the proof of (3.3). Select by (3.2) an $n^{(1)}_{\varepsilon/2}$ such that for all $n \ge n^{(1)}_{\varepsilon/2}$,

$$\left\{\xi_n(a_n, u; \cdot)/(2a_n \log(1/a_n))^{1/2}, 0 \le u \le 1 - a_n\right\} \subset \mathbb{S}_0^{\varepsilon/2}. \tag{3.46}$$

Next, choose $\lambda = 1 + \frac{1}{16}\varepsilon^2$, and select by (3.42) and (3.43) an $n^{(2)}_{\varepsilon/2}$ such that, for all $n \geq n^{(2)}_{\varepsilon/2}$,

$$\sup_{1-\lambda a_n \leq t \leq 1-a_n} \|\zeta_n(a_n, t; \cdot) - \xi_n(a_n, 1 - a_n; \cdot)\|/(2a_n \log(1/a_n))^{1/2} \leq \varepsilon/2 \tag{3.47}$$

and

$$V_n(1 - \lambda a_n) < 1 - a_n. \tag{3.48}$$

Finally, by (3.40), select an $n^{(3)}_{\varepsilon/2}$ such that for all $n \geq n^{(3)}_{\varepsilon/2}$,

$$\sup_{0 \leq t \leq 1-\lambda a_n} \|\zeta_n(a_n, t; \cdot) - \xi_n(a_n, V_n(t); \cdot)\|/(2a_n \log(1/a_n))^{1/2} \leq \varepsilon/2. \tag{3.49}$$

By combining (3.46)–(3.49), we see that (3.3) holds whenever $n \geq \max(n^{(i)}_{\varepsilon/2}, 1 \leq i \leq 3)$. This completes the first half of the proof of Theorem 3.1 for ζ_n.

For the second half of this proof, we select an arbitrary $f \in \mathbb{S}_0$ and fix an $\varepsilon > 0$. By (3.4) and Remark 3.1, there exist almost surely an $n^{(4)}_{\varepsilon/4}$ and a sequence $\frac{1}{4} < s_n < \frac{3}{4}$ such that, for all $n \geq n^{(4)}_{\varepsilon/4}$,

$$\|\xi_n(a_n, s_n; \cdot)/(2a_n \log(1/a_n))^{1/2} - f\| < \varepsilon/4. \tag{3.50}$$

Set now $t_n = U_n(s_n)$. Observe that $t_n = i/n$, where $0 \leq i \leq n$ is an integer. Moreover, by the Glivenko–Cantelli theorem, we have almost surely for all large n, $\frac{1}{8} < t_n < \frac{7}{8}$, together with the inequality $V_n(t_n) \leq s_n < V_n(t_n + 1/n)$. This, in turn, implies [see, e.g., Devroye (1981, 1982a) and Deheuvels (1982)] that

$$\begin{aligned}&\limsup_{n\to\infty} |V_n(t_n) - s_n|(n/\log n)\\ &\quad \leq \lim_{n\to\infty} \max_{1\leq j\leq n} |V_n(j/n) - V_n((j-1)/n)|(n/\log n) = 1 \quad \text{a.s.}\end{aligned} \tag{3.51}$$

By (3.51) and the definition of t_n, it follows that

$$\begin{aligned}&\limsup_{n\to\infty} \|\xi_n(a_n, V_n(t_n); \cdot) - \xi_n(a_n, s_n; \cdot)\|/(2a_n \log(1/a_n))^{1/2}\\ &\quad \leq \limsup_{n\to\infty} \sup_{\substack{0\leq t', t''\leq 1\\ |t''-y'|\leq 2n^{-1}\log n}} \|\xi_n(a_n, t'; \cdot) - \xi_n(a_n, t''; \cdot)\|\\ &\qquad \times (2a_n \log(1/a_n))^{-1/2} = 0 \quad \text{a.s.,}\end{aligned} \tag{3.52}$$

where we have used (2.17) [note that (3.52) is a consequence of Theorem 1(I) of Mason, Shorack and Wellner (1983)].

Recalling that ultimately $\frac{1}{8} < t_n < \frac{7}{8}$, (3.49), (3.50) and (3.52) imply that whenever $n \geq n^{(4)}_{\varepsilon/4}$ is chosen so large that the LHS of (3.52) is less than $\varepsilon/4$ and $t_n < \lambda a_n$, we have

$$\|\zeta_n(a_n, t_n; \cdot)/(2a_n \log(1/a_n))^{1/2} - f\| < \varepsilon. \tag{3.53}$$

By (3.53), we have (3.5), which completes the proof of Theorem 3.1. □

REMARK 3.2. The same argument as in Remark 2.4 shows that $n''_{\varepsilon,f}$ in Theorem 3.1 may be chosen independently of $f \in \mathbb{S}$.

REMARK 3.3. It is readily verified from the arguments used for the proof of Theorem 3.1 that the conclusion of this theorem holds if there exists a sequence $\{\lambda_n,\ n \geq 1\}$ satisfying the CRS conditions and such that $\lambda_n/a_n \to 1$ as $n \to \infty$.

4. Applications.

4.1. *Oscillation moduli of the empirical and quantile processes.* Let $\Xi_n^{\pm}(a)$ and $\Theta_n^{\pm}(a)$ be as in (1.2), and define likewise

$$\overline{\overline{\Xi}}{}_n^{\pm}(a) = \pm \sup_{0 \leq t \leq 1-a} \sup_{0 \leq s \leq 1} \pm \xi_n(a,t;s) \tag{4.1.1}$$

and

$$\overline{\overline{\Theta}}{}_n^{\pm}(a) = \pm \sup_{0 \leq t \leq 1-a} \sup_{0 \leq s \leq 1} \pm \zeta_n(a,t;s). \tag{4.1.2}$$

Let $b_n = (2a_n \log(1/a_n))^{1/2}$. Theorem 3.1 shows that whenever $\{a_n,\ n \geq 1\}$ satisfies the CRS conditions, all four statistics $b_n^{-1}\Xi_n^{\pm}(a_n)$, $b_n^{-1}\overline{\overline{\Xi}}{}_n^{\pm}(a_n)$, $b_n^{-1}\Theta_n^{\pm}(a_n)$ and $b_n^{-1}\overline{\overline{\Theta}}{}_n^{\pm}(a_n)$ have almost sure limits when $n \to \infty$ equal to

$$\pm \sup_{f \in \mathbb{S}_0} \pm f(1) = \pm \sup_{f \in \mathbb{S}_0} \sup_{0 \leq s \leq 1} \pm f(s) = \pm 1. \tag{4.1.3}$$

The same holds obviously for $b_n^{-1}\max(\Xi_n^{+}(a_n), -\Xi_n^{-}(a_n))$, $b_n^{-1}\max(\Theta_n^{+}(a_n), -\Theta_n^{-}(a_n))$, $b_n^{-1}\max(\overline{\overline{\Xi}}{}_n^{+}(a_n), -\overline{\overline{\Xi}}{}_n^{-}(a_n))$ and $b_n^{-1}\max(\overline{\overline{\Theta}}{}_n^{+}(a_n), -\overline{\overline{\Theta}}{}_n^{-}(a_n))$ with "+" in (4.13).

Likewise, when $na_n/\log n \to c \in (0,\infty)$ as $n \to \infty$, we see from Theorems 2.1 and 2.2 that $b_n^{-1}\Xi_n^{\pm}(a_n)$ and $b_n^{-1}\overline{\overline{\Xi}}{}_n^{\pm}(a_n)$ have almost sure limits when $n \to \infty$ equal to

$$\begin{aligned} &\pm (2c)^{-1/2} \sup_{f \in \Delta_c} \pm (f(1) - c) \\ &= \pm (2c)^{-1/2} \sup_{f \in \Delta_c} \sup_{0 \leq s \leq 1} \pm (f(s) - cs) \\ &= (\delta_c^{\pm} - 1)(c/2)^{1/2}, \end{aligned} \tag{4.1.4}$$

while $b_n^{-1}\Theta_n^{\pm}(a_n)$ and $b_n^{-1}\overline{\overline{\Theta}}{}_n^{\pm}(a_n)$ have almost sure limits when $n \to \infty$ equal to

$$\begin{aligned} &\pm (2c)^{-1/2} \sup_{g \in \Gamma_c} \pm (g(1) - c) \\ &= \pm (2c)^{-1/2} \sup_{g \in \Gamma_c} \sup_{0 \leq s \leq 1} \pm (g(s) - cs) \\ &= (\gamma_c^{\pm} - 1)(c/2)^{1/2}. \end{aligned} \tag{4.1.5}$$

Similar results hold for

$$b_n^{-1}\max(\Xi_n^{+}(a_n), -\Xi_n^{-}(a_n)),\ b_n^{-1}\max\left(\overline{\overline{\Xi}}{}_n^{+}(a_n), -\overline{\overline{\Xi}}{}_n^{-}(a_n)\right),$$
$$b_n^{-1}\max(\Theta_n^{+}(a_n), -\Theta_n^{-}(a_n))$$

and $b_n^{-1}\max(\overline{\overline{\Theta}}_n^+(a_n), -\overline{\overline{\Theta}}_n^-(a_n))$ with "+" in (4.1.4) and (4.1.5), respectively. Here, we make use of the obvious inequalities

$$\delta_c^+ - 1 \geq 1 - \delta_c^- \quad \text{and} \quad \gamma_c^+ - 1 \geq 1 - \gamma_c^-, \quad \text{for } c > 0. \tag{4.1.6}$$

These results give new proofs for the theorems of Stute (1982a), Mason, Shorack and Wellner (1983) and Mason (1984). Moreover, they give an additional insight to the meaning of these theorems by showing in each case [see, e.g., Deheuvels and Mason (1991)] that the functions $\xi_n(a_n, t; \cdot)$ and $\zeta_n(a_n, t; \cdot)$ are nearly linear when the extremum in each of the eight statistics considered previously is reached.

REMARK 4.1. Notice that Stute (1982a) considers only the limiting behavior of $\max(\overline{\overline{\Xi}}_n^+, -\overline{\overline{\Xi}}_n^-)$. Likewise, Mason (1984) considers essentially $\max(\overline{\overline{\Theta}}_n^+, -\overline{\overline{\Theta}}_n^-)$. Finally, Mason, Shorack and Wellner (1983) consider only $\max(\overline{\Xi}_n^+, -\overline{\Xi}_n^-)$ and the Lipschitz-($\frac{1}{2}$) modulus of α_n. In that sense, the results obtained in this section for $\Xi_n^{\pm}$, $\overline{\overline{\Xi}}_n^{\pm}$, $\Theta_n^{\pm}$ and $\overline{\overline{\Theta}}_n^{\pm}$ are new. However, the methods used by the previously mentioned authors could have been applied without great difficulty to obtain (1.3)–(1.5) as cited in Section 1.

4.2. *Laws of the iterated logarithm for kernel density estimators.* Consider a sequence $X_1, X_2, \ldots$ of independent and identically distributed random variables on the real line having common distribution function $F(x) = P(X_1 \leq x)$ and density $f(x) = F'(x)$ assumed to be continuous and positive on a given bounded interval $[A, B]$ $(A < B)$. In the past years, much attention has been directed toward obtaining limiting properties of estimators of f such as the well-known Parzen–Rosenblatt kernel density estimator [Rosenblatt (1956) and Parzen (1962)] given by

$$\begin{aligned} f_n(x) &= (n\lambda_n)^{-1}\sum_{i=1}^{n} K((x - X_i)/\lambda_n) \\ &= \int_{-\infty}^{\infty} \lambda_n^{-1} K((x-t)/\lambda_n)\, dF_n(t), \end{aligned} \tag{4.2.1}$$

where $F_n(t) := n^{-1}\#\{X_i \leq t: 1 \leq i \leq n\}$, $\{\lambda_n, n \geq 1\}$ is a sequence of positive constants, and where the kernel $K(\cdot)$ is assumed to satisfy the conditions:

(K.1) $K(\cdot)$ is of bounded variation on $(-\infty, \infty)$.
(K.2) For some $0 < M < \infty$, $K(u) = 0$ for all $|u| \geq M$.
(K.3) $\int_{-\infty}^{\infty} K(u)\, du = 1$.

Conditions which ensure the consistency of such estimators with respect to various criteria (such as uniformity in the L^1 convergence) have been given by several authors, among whom we may cite Deheuvels (1974), Bertrand-Retali (1978), Silverman (1978), Devroye (1979) and Devroye and Wagner (1979). These well-known conditions [see, e.g., Devroye and Györfi (1985) and Devroye (1987) for additional references] are stated in Theorem D.

THEOREM D. *Under* (K.1)–(K.3), *for any given subinterval* $[C, D]$ *of* $[A, B]$ *with* $C < D$, *a necessary and sufficient condition to have* $\sup_{C \le x \le D} |f_n(x) - f(x)| \to 0$ *in probability as* $n \to \infty$ *for any density* f, *continuous on* $[A, B]$, *is that*

(S.7) $$\lambda_n \to 0 \textit{ together with } n\lambda_n/\log n \to \infty \quad \textit{as } n \to \infty.$$

Rates of strong uniform consistency of f_n to f have been obtained by Stute (1982b) who proved, under the CRS conditions, that whenever $A < C < D < B$, we have

(4.2.2) $$\lim_{n\to\infty} \sup_{C \le x \le D} \left| \frac{f_n(x) - E(f_n(x))}{\sqrt{f(x)}} \right| \left(\frac{n\lambda_n}{2\log(1/\lambda_n)} \right)^{1/2} = \left(\int_{-\infty}^{\infty} K^2(u)\, du \right)^{1/2} \quad \text{a.s.}$$

Our first result gives a new proof and an extension of (4.2.2).

THEOREM 4.1. *Under the CRS conditions imposed upon* $\{\lambda_n, n \ge 1\}$ *and* (K.1) *and* (K.2), *for any* $A < C < D < B$, *we have*

(4.2.3) $$\lim_{n\to\infty} \sup_{C \le x \le D} \pm \left(\frac{f_n(x) - E(f_n(x))}{\sqrt{f(x)}} \right) \left(\frac{n\lambda_n}{2\log(1/\lambda_n)} \right)^{1/2} = \left(\int_{-\infty}^{\infty} K^2(u)\, du \right)^{1/2} \quad a.s.$$

PROOF. Assume without loss of generality that $M = \frac{1}{2}$ in (K.2), and set $\tilde{K}(u) = K(\frac{1}{2} - u)$. Further, let

$$\tilde{f}_n(x) = (n\lambda_n)^{-1} \sum_{i=1}^{n} \tilde{K}\left(\frac{X_i - x}{\lambda_n} \right) = f_n\left(x + \frac{1}{2}\lambda_n \right).$$

Since f is continuous and $\lambda_n \to 0$, it is straightforward that all we need is to prove that, for any fixed $A < \tilde{C} < \tilde{D} < B$, we have

(4.2.4) $$\sup_{\tilde{C} \le x \le \tilde{D}} \pm \left(\frac{\tilde{f}_n(x) - E(\tilde{f}_n(x))}{\sqrt{f(x)}} \right) \left(\frac{n\lambda_n}{2\log(1/\lambda_n)} \right)^{1/2} \to \left(\int_0^1 \tilde{K}^2(u)\, du \right)^{1/2} \quad \text{a.s.}$$

In view of (4.2.4), observe that

(4.2.5) $$\begin{aligned} \tilde{f}_n(x) &= \int_0^1 \lambda_n^{-1} \tilde{K}(u)\, dF_n(x + \lambda_n u) \\ &= -\int_0^1 \lambda_n^{-1}(F_n(x + \lambda_n u) - F_n(x))\, d\tilde{K}(u). \end{aligned}$$

Without loss of generality, let $F_n(x) = U_n(F(x))$, where $U_n(\cdot)$ is as in Section 1. We obtain directly from (4.2.5) that

$$n^{1/2}\lambda_n\big(\tilde{f}_n(x) - E(\tilde{f}_n(x))\big) = -\int_0^1 \big(\alpha_n(F(x+\lambda_n u)) - \alpha_n(F(x))\big)\, d\tilde{K}(u). \tag{4.2.6}$$

Now set

$$a_n = \sup_{\tilde{C} \le x \le \tilde{D}} \{F(x+\lambda_n) - F(x)\} = (1+o(1))\lambda_n \sup_{\tilde{C}\le x\le \tilde{D}} f(x) \quad \text{as } \lambda_n \to 0.$$

Since $\{\lambda_n,\ n \ge 1\}$ satisfies the CRS conditions, by Theorem 3.1 and Remark 3.3, it follows that, for any $\varepsilon > 0$, there exists almost surely an $n'_\varepsilon < \infty$ such that, for all $n \ge n'_\varepsilon$

$$\Big\{(2a_n \log(1/a_n))^{-1/2}\big(\alpha_n(F(x) + a_n \cdot) - \alpha_n(F(x))\big) : \tilde{C} \le x \le \tilde{D}\Big\} \subset \mathbb{S}_0^{\varepsilon}. \tag{4.2.7}$$

For $\tilde{C} \le x \le \tilde{D}$ and $0 \le u \le 1$, set $v_n(u,x) = a_n^{-1}(F(x+\lambda_n u) - F(x))$. Obviously, $0 \le v_n(u,x) \le 1$. Moreover, by the continuity and positivity of f on $[A, B]$, for any $\rho > 1$, there exist a $\nu_\rho > 0$ and an $n_\rho < \infty$ such that $\tilde{D} - \tilde{C} \le \nu_\rho$ and $n \ge n_\rho$ imply that

$$\rho^{-1}u \le v_n(u,x) \le \rho u \quad \text{for all } 0 \le u \le 1 \text{ and } \tilde{C} \le x \le \tilde{D}.$$

By Schwarz's inequality, for any $0 \le s,\ t \le 1$ and $g \in \mathbb{S}_0$, we have $|g(s) - g(t)| \le \sqrt{|t-s|}$. It follows that for $n \le n_\rho$ and $\tilde{D} - \tilde{C} \le \nu_\rho$,

$$\sup_{g \in \mathbb{S}_0} \sup_{\tilde{C} \le x \le \tilde{D}} \|g(v_n(\cdot, x)) - g\| \le \sqrt{\rho - 1}. \tag{4.2.8}$$

Choose now $\rho = 1 + \varepsilon^2$ and assume that $\tilde{D} - \tilde{C} \le \nu_\rho$ and $n \ge \max(n'_\varepsilon, n_\rho)$. By (4.2.7) and (4.2.8), we have

$$\Big\{(2a_n \log(1/a_n))^{-1/2}\big(\alpha_n(F(x + \lambda_n \cdot)) - \alpha_n(F(x))\big) : \tilde{C} \le x \le \tilde{D}\Big\} \subset \tilde{\mathbb{S}}_0^{2\varepsilon}. \tag{4.2.9}$$

By (4.2.6) and (4.2.9), it follows that almost surely

$$\limsup_{n\to\infty}\Big\{n^{1/2}\lambda_n(2a_n \log(1/a_n))^{-1/2} \sup_{\tilde{C} < x < \tilde{D}} \pm \big(\tilde{f}_n(x) - E(\tilde{f}_n(x))\big)\Big\} \le 2\varepsilon\int_0^1 |d\tilde{K}(u)| + \sup_{g\in\mathbb{S}_0} \mp \int_0^1 g(u)\, d\tilde{K}(u), \tag{4.2.10}$$

where $\int_0^1 |d\tilde{K}(u)|$ [$< \infty$ by (K.1)] denotes the total variation of $\tilde{K}(\cdot)$. Next, by integrating by parts and by Schwarz's inequality, we have, for any $g \in \mathbb{S}_0$,

$$\begin{aligned}\left(\int_0^1 g(u)\, d\tilde{K}(u)\right)^2 &= \left(\int_0^1 \dot{g}(u)\tilde{K}(u)\, du\right)\\ &\le \left(\int_0^1 \dot{g}(u)^2\, du\right)\left(\int_0^1 \tilde{K}^2(u)\, du\right)\\ &\le \int_0^1 \tilde{K}^2(u)\, du.\end{aligned} \tag{4.2.11}$$

Next, we see that we may choose $\nu_\rho > 0$ so small that for $\tilde{D} - \tilde{C} \le \nu_\rho$ and $\tilde{C} \le x \le \tilde{D}$,

$$\left(\frac{n\lambda_n}{2\log(1/\lambda_n)}\right)^{1/2}\left(n^{1/2}\lambda_n\left(2a_n \log\left(\frac{1}{a_n}\right)\right)^{-1/2}\right)^{-1}(f(x))^{-1/2} \le 1+\varepsilon \tag{4.2.12}$$

for all large n. Thus, by (4.2.10)–(4.2.12), when applied repeatedly to each element of a subdivision of $[\tilde{C}, \tilde{D}]$ into subintervals of length less than ν_ρ, we obtain that the LHS of (4.2.4) is a.s. for all large n less than or equal to

$$2\varepsilon(1+\varepsilon)\int_0^1 |d\tilde{K}(u)| + (1+\varepsilon)\left(\int_0^1 \tilde{K}^2(u)\, du\right)^{1/2} + \varepsilon. \tag{4.2.13}$$

By taking $\varepsilon > 0$ arbitrarily small in (4.2.13), we obtain the upper half of (4.2.4). For the lower half, we use again Theorem 3.1 to show that, for any $g \in \mathbb{S}_0$ and $\varepsilon > 0$, there exists almost surely an $n''_{\varepsilon, g}$ such that for all $n \ge n''_{\varepsilon, g}$ there exists an $x_n \in (\tilde{C}, \tilde{D})$ (see, e.g., Remark 3.1) such that

$$\|(2a_n \log(1/a_n))^{-1/2}(\alpha_n(F(x_n) + a_n \cdot) - \alpha_n(F(x_n))) - g\| < \varepsilon. \tag{4.2.14}$$

By the same arguments as in (4.2.8) and (4.2.9), we obtain from (4.2.14) that, for all n sufficiently large,

$$\|(2a_n \log(1/a_n))^{-1/2}(\alpha_n(F(x_n + \lambda_n \cdot)) - \alpha_n(F(x_n))) - g\| < 2\varepsilon. \tag{4.2.15}$$

By (4.2.6) and (4.2.15), it follows that almost surely

$$\begin{aligned}&\liminf_{n\to\infty} \pm n^{1/2}\lambda_n(2a_n \log(1/a_n))^{-1/2}\left(\tilde{f}_n(x_n) - E\left(\tilde{f}_n(x_n)\right)\right)\\ &\quad\ge -2\varepsilon\int_0^1 d\tilde{K}(u) \mp \int_0^1 g(u)\, d\tilde{K}(u) \quad \text{a.s.}\end{aligned} \tag{4.2.16}$$

By (4.2.12) and (4.2.16), if we restrict x_n to vary in $(C_1, D_1) \subset (\tilde{C}, \tilde{D})$ with $D_1 - C_1 < \nu_{1+\varepsilon^2}$, we have, as in (4.2.12) and (4.2.13),

$$\begin{aligned}&\liminf_{n\to\infty} \pm \left(\frac{\tilde{f}_n(x_n) - E\left(\tilde{f}_n(x_n)\right)}{\sqrt{f(x_n)}}\right)\left(\frac{n\lambda_n}{2\log(1/\lambda_n)}\right)^{1/2}\\ &\quad\ge -2\varepsilon(1+\varepsilon)\int_0^1 |d\tilde{K}(u)| \mp (1+\varepsilon)\int_0^1 g(u)\, d\tilde{K}(u) - \varepsilon \quad \text{a.s.}\end{aligned} \tag{4.2.17}$$

Consider now the function $g(u) = \pm(\int_0^u \tilde{K}(s)\,ds)/(\int_0^1 \tilde{K}^2(s)\,ds)^{1/2}$, for $0 \le u \le 1$. It can be verified that $\int_0^1 g(u)\,d\tilde{K}(u) = \mp(\int_0^1 \tilde{K}^2(u)\,du)^{1/2}$ and that $g \in \mathbb{S}_0$. By making this choice for g in (4.2.17) and letting $\varepsilon > 0$ be arbitrarily small, we obtain the lower-half part of (4.2.4), which completes the proof of Theorem 4.1. □

REMARK 4.2. It is noteworthy that the conclusion of Theorem 4.1 holds without assuming (K.3). This assumption is only required to ensure that $\int_{-\infty}^{\infty} f_n(x)\,dx = 1$, that is, to obtain that f_n is a density. Moreover, no positivity assumption is required on $K(\cdot)$, so that we may apply Theorem 4.1 in the case of estimates of derivatives of f. Corollary 4.1 is obtained directly from this theorem by setting $K(\cdot) = H^{(p)}(\cdot)$, with $H^{(p)}$ and $f^{(p)}$ denoting the pth derivatives of H and f, respectively. Recently, Hall (1990) showed that any kernel $K(\cdot)$ with an unbounded support [i.e., not satisfying (K.2)] could be approximated by a kernel $K_\varepsilon(\cdot)$ in such a way that the formal replacement of K by K_ε in (4.2.1) modifies the limit in (4.2.2) by less than ε a.s. Therefore, we may use his argument to show that our results hold without assuming (K.2).

COROLLARY 4.1. *Let $H(\cdot)$ be p times differentiable on $(-\infty, \infty)$ and such that $H^{(p)}(\cdot)$ is of bounded variation on $(-\infty, \infty)$. Assume further that $K = H^{(p)}$ satisfies* (K.1) *and* (K.2) *and that $\int_{-\infty}^{\infty} H(u)\,du = 1$. Under the CRS conditions imposed upon $\{\lambda_n,\ n \ge 1\}$, if*

$$f_n^{(p)}(x) = n^{-1}\lambda_n^{-p-1} \sum_{i=1}^{n} H^{(p)}\left(\frac{x - X_i}{\lambda_n}\right), \tag{4.2.18}$$

then, for any $A < C < D < B$, we have

$$\begin{aligned}\lim_{n\to\infty}\ \sup_{C\le x\le D}\ \pm \left(\frac{f_n^{(p)}(x) - E(f_n^{(p)}(x))}{\sqrt{f(x)}}\right)\left(\frac{n\lambda_n^{2p+1}}{2\log(1/\lambda_n)}\right)^{1/2} \\ = \left(\int_{-\infty}^{\infty} H^{(p)}(u)^2\,du\right)^{1/2} \quad a.s.\end{aligned} \tag{4.2.19}$$

Interestingly, for (4.2.19), we need not assume the existence of $f^{(p)}$. Note here that

$$\begin{aligned}E(f_n^{(p)}(x)) &= \int_{-\infty}^{\infty} \lambda_n^{-p-1} H^{(p)}\left(\frac{x-t}{\lambda_n}\right) dF(t) \\ &= \int_{-\infty}^{\infty} \frac{d^p}{dx^p}\left(\frac{1}{\lambda_n} H\left(\frac{x-t}{\lambda_n}\right)\right) f(t)\,dt\end{aligned}$$

is always finite under the assumptions of Corollary 4.1. If f is p times continuously differentiable, integration by parts yields

$$E(f_n^{(p)}(x)) = \int_{-\infty}^{\infty}\left(\frac{1}{\lambda_n} H\left(\frac{x-t}{\lambda_n}\right)\right) f^{(p)}(t)\,dt \to f^{(p)}(x) \quad \text{as } \lambda_n \to 0.$$

Thus, under this last assumption, we see that a necessary and sufficient condition for the strong uniform consistency of $f_n^{(p)}(\cdot)$ to $f^{(p)}(\cdot)$ on $[C, D]$ (assuming that $\{\lambda_n, n \geq 1\}$ satisfies the CRS conditions) is for $p \geq 1$ that

$$\lambda_n \to 0 \text{ together with } n\lambda_n^{2p+1}/\log n \to \infty \quad \text{as } n \to \infty. \tag{4.2.20}$$

For $p = 0$, we may not conclude (4.2.19), since the CRS conditions impose that $n\lambda_n/\log n \to \infty$ as $n \to \infty$. However, in this case, we may use Theorem 2.1 to prove the following analogue of Theorem 4.1, covering the limiting behavior of $f_n(\cdot)$ when $n\lambda_n/\log n \to L \in (0, \infty)$. To introduce the statement of this theorem, observe (see, e.g., the proof of Theorem 4.1) that when $n\lambda_n/\log n \to \infty$, the RHS on (4.2.4) is governed by the constant

$$\begin{aligned} \sup_{g \in \mathbb{S}_0} \pm \int_0^1 g(t)\, d\tilde{K}(t) &= \sup_{g \in \mathbb{S}_0} \mp \int_0^1 \tilde{K}(t)\dot{g}(t)\, dt \\ &= \left(\int_0^1 \tilde{K}(t)^2\, dt \right)^{1/2}. \end{aligned} \tag{4.2.21}$$

It turns out that when $n\lambda_n/\log n \to L$, we need to evaluate the same expression as in (4.2.21) with the formal replacement of $\mathbb{S}_0$ by Δ_c for a suitable $c \in (0, \infty)$. By Theorems 3 and 4 of Deheuvels and Mason (1991), one must here consider the solutions $\lambda_c^- < 0 < \lambda_c^+$ of the equation in λ (assuming that $\tilde{K} \not\equiv 0$),

$$\int_{-\infty}^{\infty} h\big(\exp(\lambda \tilde{K}(t))\big)\, dt = 1/c, \tag{4.2.22}$$

where $h(\cdot)$ is as in (1.6). Here, $\lambda_c^+ \in (0, \infty)$ always exists, while $\lambda_c^- \in (0, \infty)$ is only defined when

$$c > c_0 := \lim_{\lambda \downarrow -\infty} \left\{ \int_{-\infty}^{\infty} h\big(\exp(\lambda \tilde{K}(t))\big)\, dt \right\}^{-1}. \tag{4.2.23}$$

Here, it may be seen that if M^+ (resp., M^-) denotes the Lebesgue measure of the set of points where $\tilde{K}(\cdot) > 0$ (resp., < 0), we have $c_0 = 0$ if $M^- > 0$, while $c_0 = 1/M^+$ when $M^- = 0$ (assuming that $\tilde{K} \not\equiv 0$).

Whenever $0 < c \leq c_0$, we will set $\lambda_c^- = -\infty$. By Theorems 3 and 4 of Deheuvels and Mason (1991), we have then, for $K \geq 0$,

$$\begin{aligned} \sup_{g \in \Delta_c} \pm \int_0^1 \tilde{K}(t)\dot{g}(t)\, dt &= \pm c \int_{-\infty}^{\infty} \tilde{K}(t) \exp\big(\lambda_c^{\pm} \tilde{K}(t)\big)\, dt \\ &=: \pm \Lambda^{\pm}(c, \tilde{K}), \end{aligned} \tag{4.2.24}$$

with the convention that $\Lambda^-(c, \tilde{K}) = 0$ when $\lambda_c^- = -\infty$.

We may now state the following theorem, which covers the case where $n\lambda_n/\log n \to L$.

THEOREM 4.2. *Assume that $n\lambda_n/\log n \to L \in (0, \infty)$ as $n \to \infty$. Then, under* (K.1)–(K.2) *and assuming that $K \not\equiv 0$ and $K \geq 0$, for any fixed $0 < \varepsilon < 1$*

and $A < C < D < B$, we have almost surely for all n sufficiently large and $C \le x \le D$,

$$\pm f_n(x) \le \pm \frac{1+\varepsilon}{L} \Lambda^{\pm}(Lf(x), K) + \varepsilon. \tag{4.2.25}$$

Moreover, there exists almost surely $C \le x_n^{\pm} \le D$ for all large n, such that

$$\pm f_n(x_n^{\pm}) \ge \pm \frac{1-\varepsilon}{L} \Lambda^{\pm}(Lf(x), K) - \varepsilon. \tag{4.2.26}$$

Proof. It will become obvious from the arguments used in the sequel that the restriction of positivity imposed upon K may easily be relaxed. However, this assumption simplifies the exposition and will be made from now on. We will follow the lines of the proof of Theorem 4.1 with small changes. In the first place, we reduce (4.2.25) and (4.2.26) to the corresponding statements with the formal replacements of f_n by $\tilde{f}_n$, of K by $\tilde{K}$ and of C, D by $\tilde{C}, \tilde{D}$. By letting $a_n = \sup_{\tilde{C} \le x \le \tilde{D}} \{F(x + \lambda_n) - F(x)\}$, we see that $na_n/\log n \to L \sup_{\tilde{C} \le x \le \tilde{D}} f(x)$, so that $\{a_n, n \ge 1\}$ satisfies (S.6). By Theorem 2.1, it follows that, for any $\varepsilon > 0$, there exists almost surely an $N'_\varepsilon < \infty$ such that, for all $n \ge N'_\varepsilon$,

$$\left\{ \frac{n}{\log n} (F_n(F(x) + a_n \cdot) - F_n(F(x))) : \tilde{C} \le x \le \tilde{D} \right\} \subset \Delta^{\varepsilon}_{L'}, \tag{4.2.27}$$

where $L' := L \sup_{\tilde{C} \le x \le \tilde{D}} f(x)$. By (2.21), for any $0 \le s, t \le 1$ and $g \in \Delta_{L'}$, we have $|g(s) - g(t)| \le L'|t - s|\delta_{L'|t-s|}$, so that we replace formally (4.2.8) by

$$\sup_{g \in \Delta_{L'}} \sup_{\tilde{C} \le x \le \tilde{D}} \|g(v_n(\cdot, x)) - g\| \le L'(\rho - 1)\delta_{L'(\rho-1)}. \tag{4.2.28}$$

By choosing $\rho > 1$ so small that the RHS of (4.2.28) is less than ε, we obtain the analogue of (4.2.9) which holds almost surely for all large n and $\tilde{D} - \tilde{C}$ sufficiently small:

$$\left\{ \frac{n}{\log n} (F_n(F(x + \lambda_n \cdot)) - F_n(F(x))) : \tilde{C} \le x \le \tilde{D} \right\} \subset \Delta^{2\varepsilon}_{L'}. \tag{4.2.29}$$

Given (4.2.29) and (4.2.5), the remainder of the proof is very similar to the proof of Theorem 4.1 and makes use of (4.2.24) and Remark 2.2. Therefore, we omit details. □

Remark 4.3. It is obvious from (4.2.24) that whenever (K.3) holds and $K \ge 0$, we have

$$\frac{1}{L} \Lambda^{\pm}(Lf(x), K) = f(x) \int_{-\infty}^{\infty} K(t) \exp\left(\lambda^{\pm}_{L f(x)} K(t)\right) dt, \tag{4.2.30}$$

which is greater than $f(x)$ in the "+" case and less than $f(x)$ in the "−" case. Thus we see from (4.2.26) and (4.2.29) that the condition $n\lambda_n/\log n \to \infty$ is necessary for strong uniform consistency of f_n to f [see, e.g., (4.2.20)].

Let $c \to \infty$ in (4.2.22). Since $h(1+u) \sim u^2/2$ and (uniformly over t) $\exp(\lambda K(t)) - 1 \sim \lambda K(t)$ as $\lambda \to 0$, the following expansion holds for (4.2.22):

$$\begin{aligned}\int_{-\infty}^{\infty} h(\exp(\lambda K(t)))\,dt &= (1+o(1))\frac{\lambda^2}{2}\int_{-\infty}^{\infty} K^2(t)\,dt \\ &= \frac{1}{c}, \quad \text{with } \lambda = \lambda_c^{\pm}, \quad \text{as } c \to \infty.\end{aligned} \tag{4.2.31}$$

By (4.2.31), we have as $c \to \infty$,

$$\lambda_c^{\pm} = \pm(1+o(1))\left(\frac{c}{2}\right)^{-1/2}\left(\int_{-\infty}^{\infty} K^2(t)\,dt\right)^{-1/2}, \tag{4.2.32}$$

so that, by (4.2.30) and (4.2.32), we have, under (K.3),

$$\begin{aligned}\frac{1}{L}\Lambda^{\pm}(Lf(x), K) - f(x) &= (1+o(1))\, f(x)\lambda_{Lf(x)}^{\pm}\int_{-\infty}^{\infty} K^2(t)\,dt \\ &= \pm(1+o(1))\left(\frac{2f(x)}{L}\right)^{1/2}\left(\int_{-\infty}^{\infty} K^2(t)\,dt\right)^{1/2} \\ &\qquad \text{as } L \to \infty.\end{aligned} \tag{4.2.33}$$

Consider now the case where $f(x)$ is constant on $[C, D]$ [so that $E(f_n(x)) = f(x)$]. It follows from (4.2.25), (4.2.26) and (4.2.33) that if $\lambda_n = L(\log n)/n$, then a.s.,

$$\begin{aligned}&\lim_{n\to\infty} \sup_{C \le x \le D} \pm\left(\frac{f_n(x) - E(f_n(x))}{\sqrt{f(x)}}\right)\left(\frac{n\lambda_n}{2\log(1/\lambda_n)}\right)^{1/2} \\ &\quad = \pm(1+o(1))\left(\int_{-\infty}^{\infty} K^2(t)\,dt\right)^{1/2} \quad \text{as } L \to \infty.\end{aligned}$$

This shows that the statements of Theorems 4.1 and 4.2 are in agreement.

4.3. *Laws of the iterated logarithm for nearest neighbor density estimators.* Assume that $X_1, X_2, \dots$ are as in Section 4.2, and let $0 < \lambda_n < 1$ be a real sequence. The nearest neighbor density estimator introduced by Fix and Hodges (1951) and later studied by Loftsgaarden and Quesenberry (1965) estimates the common density f of $X_1, X_2, \dots$ by

$$\hat{f}_n(x) = \lambda_n / \inf\left\{h > 0 : F_n\left(x + \frac{h}{2}\right) - F_n\left(x - \frac{h}{2}\right) \ge \lambda_n\right\}. \tag{4.3.1}$$

Denote by $M_n(x)$ the function of x defined for $n \ge 1$ by

$$M_n(x) = \lambda_n / \inf\left\{h > 0 : F\left(x + \frac{h}{2}\right) - F\left(x - \frac{h}{2}\right) \ge \lambda_n\right\}. \tag{4.3.2}$$

We obtain the following law of the iterated logarithm for $\hat{f}_n$.

THEOREM 4.3. *Under the CRS conditions imposed upon* $\{\lambda_n,\ n \geq 1\}$, *for any* $A < C < D < B$, *we have*

$$\lim_{n\to\infty} \sup_{C\leq x\leq D} \pm \left(\frac{\hat{f}_n(x) - M_n(x)}{f(x)}\right)\left(\frac{n\lambda_n}{2\log(1/\lambda_n)}\right)^{1/2} = \pm 1 \quad a.s. \tag{4.3.3}$$

PROOF. The proof is very similar to that just given for Theorem 4.1. Therefore, we omit details. □

REMARK 4.4. Further references to study of $\hat{f}_n$ are given in Devroye (1982b, 1987), and in Devroye and Györfi (1985). The result given in (4.3.3) is essentially due to Mack (1983), who proved that, under the CRS conditions imposed upon $\{\lambda_n,\ n \geq 1\}$,

$$\lim_{n\to\infty} \sup_{C\leq x\leq D} \left|\frac{f_n(x) - M_n(x)}{f(x)}\right|\left(\frac{n\lambda_n}{2\log(1/\lambda_n)}\right)^{1/2} = 1 \quad \text{a.s.} \tag{4.3.4}$$

Mack (1983) bases his proof on Theorem 2.15 of Stute (1982a) and makes use of the assumption [see his proof of (22), Mack (1983), page 191] that the sequence $b_n(x) = \frac{1}{2}M_n(x)n^{-1}k_n$ satisfies the conditions of Stute's Theorem 2.15. Since he also assumes that k_n is a sequence of positive integers such that $n^{-1}k_n \to 0$ as $n \to \infty$, this leads to a contradiction (there is no nonultimately constant sequence of positive integers k_n such that $k_n \uparrow 0$ together with $n^{-1}k_n \downarrow 0$ as $n \uparrow \infty$). It is however possible to prove by his methods that the result (4.3.4) holds when $k_n \sim n\lambda_n$ and λ_n is as in Theorem 4.3. These regularity conditions are not mentioned in the statements of his theorems, but are implicitly used in the proofs. Note further that (4.3.3) is more general than (4.3.4).

4.4. *Conclusion.* The application of our functional laws of the iterated logarithm just detailed comprises only a small subset of the kind of precise knowledge about the asymptotic a.s. behavior of function estimators that is now available through our results. Two further applications that immediately come to mind are the limiting behavior of regression and quantile function estimators. We will pursue these and related investigations elsewhere in the near future.

5. Acknowledgement. We thank the referee for a careful reading of our manuscript and for helpful suggestions.

REFERENCES

BAHADUR, R. (1966). A note on quantiles in large samples. *Ann. Math. Statist.* **37** 577–580.

BERTRAND-RETALI, M. (1978). Convergence uniforme d'un estimateur de la densité par la méthode du noyau. *Rev. Roumaine Math. Pures Appl.* **23** 361–385.

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

CHUNG, K. L. (1949). An estimate concerning the Kolmogorov limit distribution. *Trans. Amer. Math. Soc.* **67** 36–50.

DEHEUVELS, P. (1974). Conditions nécessaires et suffisantes de convergence presque sûre et uniforme presque sûre des estimateurs de la densité. *C. R. Acad. Sci. Paris Ser. A* **278** 1217–1220.

DEHEUVELS, P. (1982). Strong limiting bounds for maximal uniform spacings. *Ann. Probab.* **10** 1058–1065.

DEHEUVELS, P. (1986). Strong laws for the k-th order statistic when $k \le c \log_2 n$. *Probab. Theory Related Fields* **72** 133–154.

DEHEUVELS, P. and DEVROYE, L. (1984). Strong laws for the maximal k-spacing when $k \le c \log n$. *Z. Wahrsch. Verw. Gebiete* **66** 315–334.

DEHEUVELS, P. and MASON, D. M. (1990). Nonstandard functional laws of the iterated logarithm for tail empirical and quantile processes. *Ann. Probab.* **18** 1693–1722.

DEHEUVELS, P. and MASON, D. M. (1991). A tail empirical process approach to some non-standard laws of the iterated logarithm. *J. Theor. Probab.* **4** 53–85.

DEVROYE, L. (1979). On the pointwise and the integral covergence of recursive kernel estimates of probability densities. *Utilitas Math.* **15** 113–128.

DEVROYE, L. (1981). Laws of the iterated logarithm for order statistics of uniform spacings. *Ann. Probab.* **9** 860–867.

DEVROYE, L. (1982a). A log log law for maximal uniform spacings. *Ann. Probab.* **10** 863–868.

DEVROYE, L. (1982b). Necessary and sufficient conditions for the pointwise convergence of nearest neighbor regression function estimates. *Z. Wahrsch. Verw. Gebiete* **61** 467–481.

DEVROYE, L. (1987). *A Course in Density Estimation*. Birkhäuser, Boston.

DEVROYE, L. and GYÖRFI, L. (1985). *Density Estimation: The L_1 View*. Wiley, New York.

DEVROYE, L. and WAGNER, T. J. (1979). The L_1 convergence of density estimates. *Ann. Statist.* **7** 1136–1139.

EINMAHL, J. H. J. (1987). *Multivariate Empirical Processes. CWI Tract* **32** Cent. Wisk. Inform., Amsterdam.

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

FIX, E. and HODGES, J. L. (1951). Discriminatory analysis. Nonparametric discrimination: consistency properties. Technical Report 4, Project no. 21-49-004, USAF School of Aviation Medicine, Randoff Field, Texas.

HALL, P. (1990). On the law of the iterated logarithm for density estimators. *Statist. Probab. Lett.* **9** 237–240.

HÖGNÄS, G. (1977). Characterization of weak convergence of signed measures on [0, 1]. *Math. Scand.* **41** 175–184.

JOHNSON, N. J. and KOTZ, S. (1969). *Discrete Distributions*. Houghton-Mifflin, Boston.

KIEFER, J. (1967). On Bahadur's representation of sample quantiles. *Ann. Math. Statist.* **38** 1323–1342.

KIEFER, J. (1970). Deviations between the sample quantile process and the sample d.f. In *Nonparametric Techniques in Statistical Inference* (M. Puri, ed.). Cambridge Univ. Press.

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, Berkeley.

KOMLÓS, J., MAJOR, P. and TUSNÁDY, G. (1975a). Weak convergence and embedding. In *Colloq. Math. Soc. János Bolyai* (P. Révész, ed.). **11** 149–165. North-Holland, Amsterdam.

KOMLÓS, J., MAJOR, P. and TUSNÁDY, G. (1975b). An approximation of partial sums of independent rv's and the sample df. I. *Z. Wahrsch. Verw. Gebiete* **32** 111–131.

LOFTSGAARDEN, D. O. and QUESENBERRY, C. P. (1965). A nonparametric estimate of a multivariate density function. *Ann. Math. Statist.* **36** 1049–1051.

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

MACK, Y. P. (1983). Rate of strong uniform convergence of *k-nn* density estimates. *J. Statist. Plann. Inference* **8** 185–192.

MASON, D. M. (1984). A strong limit theorem for the oscillation modulus of the uniform empirical quantile process. *Stochastic Process. Appl.* **17** 127–136.

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

MASON, D. M., SHORACK, G. and WELLNER, J. A. (1983). Strong limit theorems for oscillation moduli of the empirical process. *Z. Wahrsch. Verw. Gebiete* **65** 83–97.

PARZEN, E. (1962). On the estimation of a probability density function and mode. *Ann. Math. Statist.* **33** 1065–1076.

ROSENBLATT, M. (1956). Remarks on some nonparametric estimates of a density function. *Ann. Math. Statist.* **27** 832–837.

SHORACK, G. R. (1982). Kiefer's theorem via the Hungarian construction. *Z. Wahrsch. Verw. Gebiete* **61** 369–373.

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

SILVERMAN, B. (1978). Weak and strong uniform consistency of the kernel estimate of a density and its derivatives. *Ann. Statist.* **6** 177–184. (Addendum: **8** 1175–1176.)

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.

STUTE, W. (1982a). The oscillation behavior of empirical processes. *Ann. Probab.* **10** 86–107.

STUTE, W. (1982b). A law of the iterated logarithm for kernel density estimators. *Ann. Probab.* **10** 414–422.

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

VENTSEL, A. D. (1976). Rough limit theorems on large deviations for Markov processes. I–II. *Theory Probab. Appl.* **21** 227–242, 499–512.

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