The Annals of Probability
2003, Vol. 31, No. 2, 926–947

STRONG APPROXIMATION OF MAXIMA BY EXTREMAL PROCESSES

BY I. FAHRNER AND U. STADTMÜLLER

Universität Ulm

Extending results by P. Deheuvels of 1981–1983 we give strong approximation results for sample maxima by simple transformations of extremal processes and discuss the quality of the approximations. The limiting process becomes stationary after a transformation of the argument.

1. Introduction. Let $\{X_n, n \in \mathbb{N}\}$ be a sequence of independent and identically distributed (i.i.d.) random variables on some probability space. We denote by $S_n = \sum_{k=1}^n X_k,\ n \in \mathbb{N}$, their partial sums and by $M_n = \max_{1 \le k \le n} X_k,\ n \in \mathbb{N}$, their partial maxima. Studying the almost sure behavior of partial sums or of functionals of the partial sums a powerful tool is given by so-called strong invariance principles. Under appropriate assumptions, the limiting process of partial sums is a Wiener process. It is well known by now that provided we have $E(X_1) = 0,\ E(X_1^2) = 1$ and $E(|X_1|^p) < \infty$ $(p \ge 2)$ there exists a version $\{\tilde{X}_n, n \in \mathbb{N}\}$ of the original sequence and a Wiener process W on some probability space such that

$$\begin{aligned} \tilde{S}_n - W(n) &= o_{\text{a.s.}}(\sqrt{n \log\log n}) \qquad && \text{if } p = 2,\\ \tilde{S}_n - W(n) &= o_{\text{a.s.}}(n^{1/p}) && \text{if } p > 2 \end{aligned}$$

holds. For further details see, for example, the book by Csörgő and Révész (1981). Examples for applications of strong invariance principles can be found in Csörgő and Révész (1981) or Stadtmüller (1986). For a concrete example, note that the law of the iterated logarithm for S_n follows from that for $W(n)$.

Here we are interested in the sequence $\{M_n, n \in \mathbb{N}\}$ instead of $\{S_n, n \in \mathbb{N}\}$ and seek for a similar result for the sequence of maxima.

Suppose there exist $a_n > 0$ and $b_n \in \mathbb{R}$ such that

$$\lim_{n\to\infty} P\big((M_n - b_n)/a_n \le x\big) = G(x) \qquad \text{for all continuity points of } G \tag{1}$$

for some nondegenerate distribution function $G(x)$. Then we can define a sequence $\{\tilde{Y}_n(t), n \in \mathbb{N}\}$ of stochastic processes in $D(0,\infty)$, the set of all right continuous real functions on $(0,\infty)$ with finite left limits existing everywhere, by

$$\tilde{Y}_n(t) := \begin{cases} (M_{[nt]} - b_n)/a_n, & t \ge 1/n,\\ (X_1 - b_n)/a_n, & 0 < t < 1/n, \end{cases} \tag{2}$$

where $[x]$ denotes the greatest integer less or equal to x.

Received April 2001; revised April 2002.
AMS 2000 subject classifications. Primary 60F17, 60G70; secondary 60F15.
Key words and phrases. Extremal process, strong approximation, speed of convergence, almost sure behavior of extremes.

For a distribution function F let $Y_F(t)$ denote the extremal-F process in $D(0,\infty)$ [see, e.g., Resnick (1987), Section 4.3] with finite-dimensional distributions

$$F_{t_1,\dots,t_k}(x_1,\dots,x_k) = F^{t_1}\Big(\min_{1\le i\le k} x_i\Big) F^{t_2-t_1}\Big(\min_{2\le i\le k} x_i\Big)\cdots F^{t_k-t_{k-1}}(x_k) \tag{3}$$

for $k\ge 1$, $0<t_1<t_2<\cdots<t_k$, $x_i\in\mathbb{R}$, $i=1,\dots,k$.

It is a well-known fact first given by Lamperti (1964) that (1) is equivalent to

$$\tilde{Y}_n(t) \Rightarrow Y_G(t) \qquad \text{in } D(0,\infty),$$

where $\Rightarrow$ denotes weak convergence.

In order to study the behavior of $\sum_{k=1}^n k^{-1} f((M_k-b_k)/a_k)$ in the context of a.s.-versions of weak limit theorems [see, e.g., Fahrner and Stadtmüller (1998), Fahrner (2000a, b, 2001)] we want to have an almost sure version of the weak invariance principle for maxima above.

Comparing with the strong invariance principle for sums we like to get a result which corresponds to a relation like

$$\frac{S_{[nt]}}{\sqrt{nt}} = \frac{W(nt+\overline{v}(nt))}{\sqrt{nt}} + \overline{r}(nt)$$

with $\overline{v}(nt) = [nt]-nt$ and $\overline{r}(nt) = o_{\text{a.s.}}(\sqrt{\log\log(nt)})$ as $nt\to\infty$, which follows from the result above.

Deheuvels (1981, 1982, 1983) proposed a method for deriving a strong approximation for the sequence $\{\tilde{Y}_n(t), n\in\mathbb{N}\}$. He shows that there exists a sequence $\{Y_G^{(n)}(t), n\in\mathbb{N}\}$ of versions of $Y_G(t)$ such that

$$\tilde{Y}_n(t) \le Y_G^{(n)}\big(t+u_n^1(t)\big) + v_n^1(t), \qquad \tilde{Y}_n(t) \ge Y_G^{(n)}\big(t-u_n^2(t)\big) - v_n^2(t)$$

with random variables $u_n^i(t)$ and $v_n^i(t)$ where $u_n^i(t)$ tend pointwise to zero as n tends to infinity with nonuniform rates. However, no rates for the convergence of $v_n^i(t)$ to zero are discussed.

In this paper we will consider the strong approximation of the sequence $\{Y_n(t), n\in\mathbb{N}\}$ of stochastic processes in $D(0,\infty)$ given by

$$Y_n(t) := \begin{cases} (M_{[nt]}-b_{[nt]})/a_{[nt]}, & t\ge 1/n,\\ (X_1-b_1)/a_1, & 0<t<1/n,\end{cases} \tag{4}$$

which also has a weak limit in $D(0,\infty)$ denoted by $Y_{0,G}(t)$. There are two advantages in considering $Y_n(t)$ instead of $\tilde{Y}_n(t)$: in applications one usually meets the maximum M_n together with the norming like in (4) rather than that in (2). Moreover, $Y_{0,G}(t)$ is a simple transformation of the stationary process $Y_{0,G}(e^t)$. We will give an approximation of roughly the form

$$Y_n(t) \le Y_{0,G}\big(nt+v(nt)\big) + R(nt), \qquad Y_n(t) \ge Y_{0,G}\big(nt-v(nt)\big) + R(nt).$$

Note that here we approximate $Y_n(t)$ by a *single* process $Y_{0,G}(t)$ and not a sequence of processes. There are two possibilities for stating the strong approximation results. Either we consider the whole probability space, then $v(nt)$ depends on ω or we neglect a small set and Egorov's theorem will yield a uniform bound for $v(nt)$ on the slightly smaller space. We will follow the latter path. In applications this reduction into a smaller space can often be eliminated. The random variables $R(nt)$ will tend to zero in probability as nt tends to infinity. In general this is the best possible outcome. But under some regularity conditions on the distribution of X_1 we can show almost sure convergence to zero for $R(nt)$ as nt tends to infinity.

We are dealing with two kinds of errors in this approximation. $Y_n(t)$ are pure jump processes where jumps can occur only at times $k/n,\ k \in \mathbb{N}$, whereas there is no restriction for the jumps of $Y_{0,G}(t)$. These different behaviors are brought together by the function $v(nt)$. Note that the distribution of the interarrival times of the jumps of an extremal-F process can be seen being invariant under the choice of the distribution function F, compare Resnick (1987), Proposition 4.8. Thus it is not surprising that $v(nt)$ will not depend on the underlying distribution of X_1 either. The second error term is the random variable $R(nt)$ which takes care of different jump heights. This quantity naturally depends on the distribution of X_1.

2. The strong approximation results. As before, let $\{X_n, n \in \mathbb{N}\}$ be a sequence of i.i.d. random variables with common distribution function F. If (1) holds then G is of extreme value type, that is, there exist $a > 0$ and $b \in \mathbb{R}$ such that $G(ax + b)$ equals one of the following three distribution functions:

$$\Lambda(x) = \exp(-e^{-x}),$$

$$\Phi_\alpha(x) = \exp(-x^{-\alpha}) I_{(0,\infty)}(x) \qquad \text{for some } \alpha > 0,$$

or

$$\Psi_\alpha(x) = \exp\bigl(-(-x)^\alpha\bigr) I_{(-\infty,0]}(x) + I_{(0,\infty)}(x) \qquad \text{for some } \alpha > 0,$$

and we say that F belongs to the domain of attraction of G and write $F \in \mathcal{D}(G)$. Defining $\overline{F}(x) := 1 - F(x),\ x_F := \sup\{x : F(x) < 1\}$ and $\mathcal{F}$ as the set of all distribution functions, we can characterize the domains of attraction as follows:

$$\mathcal{D}(\Lambda) = \Bigl\{F \in \mathcal{F} : \exists\, a(x) > 0 \text{ s.t. } \lim_{x \to x_F} \overline{F}\bigl(x + ta(x)\bigr)/\overline{F}(x) = e^{-t}\ \forall t \in \mathbb{R}\Bigr\},$$

$$\mathcal{D}(\Phi_\alpha) = \Bigl\{F \in \mathcal{F} : x_F = \infty \text{ and } \lim_{x \to \infty} \overline{F}(tx)/\overline{F}(x) = t^{-\alpha}\ \forall t > 0\Bigr\},$$

$$\mathcal{D}(\Psi_\alpha) = \Bigl\{F \in \mathcal{F} : x_F < \infty \text{ and } \lim_{x \to \infty} \overline{F}\bigl(x_F - 1/(tx)\bigr)/\overline{F}(x_F - 1/x) = t^{-\alpha}\ \forall t > 0\Bigr\}.$$

There is a canonical way of choosing the normalizing constants and we will denote those canonical constants by a_n^{can} (resp. b_n^{can}). These are given by

$$
\begin{aligned}
&a_n^{\text{can}} = a(V(n)), && b_n^{\text{can}} = V(n), && \text{when } G = \Lambda,\\
&a_n^{\text{can}} = V(n), && b_n^{\text{can}} = 0, && \text{when } G = \Phi_\alpha,\\
&a_n^{\text{can}} = x_F - V(n), && b_n^{\text{can}} = x_F, && \text{when } G = \Psi_\alpha,
\end{aligned}
\tag{5}
$$

where $V(t) := (1/\overline{F})^{\leftarrow}(t)$ and $H^{\leftarrow}(t) := \inf\{x : H(x) \geq t\}$ denotes the left continuous inverse of a nondecreasing function H. We will also need the right continuous inverse $H^{\rightarrow}(t) := \inf\{x : H(x) > t\}$. By the convergence of types theorem the normalizing constants can be replaced by a_n and b_n obeying $a_n/a_n^{\text{can}} \to a$ and $(b_n - b_n^{\text{can}})/a_n^{\text{can}} \to b$ with some $a > 0$ and $b \in \mathbb{R}$ leading to the limit distribution $G(ax + b)$ in (1). More information about convergence of normalized extremes can be found in the books of Haan (1970); Leadbetter, Lindgren and Rootzen (1983); Galambos (1987) or Resnick (1987).

We can show the following invariance principle:

THEOREM 1. *Let $\{X_n, n \in \mathbb{N}\}$ be a sequence of i.i.d. random variables and suppose that* (1) *holds. Define $Y_n(t)$ by* (4). *Then* (1) *is equivalent to*

$$Y_n \Rightarrow Y_{0,G} \qquad \text{in } D(0,\infty)$$

where

$$Y_{0,G}(t) := \begin{cases} Y_\Lambda(t) - \log t, & \text{when } G = \Lambda,\\ t^{-1/\alpha} Y_{\Phi_\alpha}(t), & \text{when } G = \Phi_\alpha,\\ t^{1/\alpha} Y_{\Psi_\alpha}(t), & \text{when } G = \Psi_\alpha. \end{cases}$$

Moreover, in any case $\{Y_{0,G}(e^t), t \in \mathbb{R}\}$ is a stationary Markov process.

Define $\log^* x = \log(\max\{e, x\})$ and $\log_2^* x = \log^* \log^* x$, $\log_n^* x = \log^*(\log_{n-1}^* x)$ for $x > 0$ and $n \geq 3$.

Then we have the following strong approximation result:

THEOREM 2. *Let $\{X_n, n \in \mathbb{N}\}$ be a sequence of i.i.d. random variables with common distribution function F on some probability space such that* (1) *holds.*

Then it is possible to reconstruct the sequence $\{X_n, n \in \mathbb{N}\}$ together with an extremal-G process $Y_G(u)$ in $D(0,\infty)$ and a family of random variables $\{R(u),\ u \geq 1\}$ on a possibly enlarged probability space (Ω, Σ, P) such that for every $\varepsilon > 0$ there exists a set $\Omega_\varepsilon \in \Sigma$ with $P(\Omega_\varepsilon) \geq 1 - \varepsilon$ and a constant $K_\varepsilon \geq 1$ independent of ω such that

$$Y_0\big(u, -v(u)\big) + R(u) \leq \frac{M_{[u]} - b_{[u]}}{a_{[u]}} \leq Y_0\big(u, v(u)\big) + R(u)$$

$$\forall u \geq K_\varepsilon^2,\ \forall \omega \in \Omega_\varepsilon,$$

holds with $v(u) := K_\varepsilon\sqrt{\log^* u \log_3^* u}$ *and*

$$Y_0(u,v) := \begin{cases} Y_\Lambda(u+v) - \log u, & \textit{when } G = \Lambda, \\ u^{-1/\alpha} Y_{\Phi_\alpha}(u+v), & \textit{when } G = \Phi_\alpha, \\ u^{1/\alpha} Y_{\Psi_\alpha}(u+v), & \textit{when } G = \Psi_\alpha. \end{cases}$$

Furthermore, $R(u)$ *tends to zero in probability as* u *tends to infinity and if* F *is continuous we have*

$$R(u) = (M_{[u]} - b_{[u]})/a_{[u]} - G^{\leftarrow}\big(F^u(M_{[u]})\big), \qquad u \ge 1.$$

The shift $v(u)$ is small compared to u and does not disturb the asymptotic behaviour of the process too much. Note that if F is continuous we have $P(G^{\leftarrow}(F^{[u]}(M_{[u]})) \le x) = G(x)$ for all real x, that is, $R([u])$ describes the error obtained by approximating $(M_{[u]} - b_{[u]})/a_{[u]}$ by a certain random variable which is distributed according to the limit distribution. This is why $R(u)$ appears on both sides of the inequality above. As already mentioned in the introduction, it is in general not true that $\lim_{u\to\infty} R(u) = 0$ a.s. See Section 5 for a counterexample. In order to give sufficient conditions for almost sure convergence of $R(u)$ we have to consider the cases Λ, Φ_α and Ψ_α seperately and study the corresponding von Mises conditions [cf. Resnick (1987), Section 1.4]. This will be done next.

We begin with the case $G = \Lambda$. Suppose that F has a negative second derivative in some interval (z_0, x_F). Define the function

$$a : (z_0, x_F) \to (0, \infty),\ x \mapsto a(x) := \frac{\overline{F}(x)}{F'(x)}. \tag{6}$$

If the von Mises condition

$$\lim_{x\to x_F} \frac{F''(x)(1-F(x))}{(F'(x))^2} = -1 \tag{7}$$

holds, then $F \in \mathcal{D}(\Lambda)$. Note that taking the derivative of $a(x)$ shows that (7) is equivalent to

$$\lim_{x\to x_F} a'(x) = 0. \tag{8}$$

On the other hand we have (7) iff F has the following representation:

$$\overline{F}(x) = c\exp\left(-\int_{z_0}^{x} \frac{dt}{a(t)}\right) \qquad \text{for all } x \in (z_0, x_F) \tag{9}$$

where $c > 0$ and $a(t) > 0$ is differentiable and satisfies (8). In this case $a(t)$ is given by (6); see, for example, Resnick [(1987), Section 1.4, Proposition 1.18c].

We need a condition on the speed of convergence in (7) [resp. (8)]. A similar condition was considered by de Haan and Hordijk (1972) in connection with large deviation results for extremes.

THEOREM 3. *Let $\{X_n, n \in \mathbb{N}\}$ be a sequence of i.i.d. random variables with common distribution function F on some probability space. Suppose F has a negative second derivative in some neighborhood of x_F and let $a(x) = \overline{F}(x)/F'(x)$. If*

$$\lim_{x \to x_F} a'(x)\big(\log\log\big(1/\overline{F}(x)\big)\big)^2 = 0, \tag{10}$$

it is possible to reconstruct the sequence $\{X_n, n \in \mathbb{N}\}$ together with an extremal-Λ process $Y_\Lambda(u)$ in $D(0,\infty)$ on a possibly enlarged probability space (Ω, Σ, P) such that for every $\varepsilon > 0$ there exists a set $\Omega_\varepsilon \in \Sigma$ with $P(\Omega_\varepsilon) \geq 1 - \varepsilon$, a constant $K_\varepsilon \geq 1$ and a function $r_\varepsilon(u) : [1,\infty) \to [0,\infty)$ such that

$$\begin{aligned} Y_\Lambda\big(u - v(u)\big) - \log u - r_\varepsilon(u) &\leq \frac{M_{[u]} - b_{[u]}^{\mathrm{can}}}{a_{[u]}^{\mathrm{can}}} \\ &\leq Y_\Lambda\big(u + v(u)\big) - \log u + r_\varepsilon(u) \end{aligned} \tag{11}$$

holds for all $u \geq K_\varepsilon^2$ and all $\omega \in \Omega_\varepsilon$, with $v(u) := K_\varepsilon\sqrt{\log^* u \log_3^* u}$ and $\lim_{u\to\infty} r_\varepsilon(u) = 0$.

Furthermore, if there exists an increasing function $r : (0,\infty) \to (0,\infty)$ such that

$$\begin{gathered} \lim_{x\to\infty} r(x) = \infty, \qquad \lim_{x\to\infty} r(x)\frac{\log x}{x} = 0, \\ \lim_{x \to x_F} a'(x)\big(\log\log\big(1/\overline{F}(x)\big)\big)^2 r\big(1/\overline{F}(x)\big) = 0 \end{gathered} \tag{12}$$

holds then $\lim_{u\to\infty} r_\varepsilon(u) r(u) = 0$.

If in addition for some positive sequence (a_k) and some real sequence (b_k)

$$\lim_{k\to\infty} \left(\left(\frac{a_k^{\mathrm{can}}}{a_k} - 1\right)\log\log k + (b_k^{\mathrm{can}} - b_k)/a_k\right) r(k) = 0 \tag{13}$$

holds, then the canonical constants in (11) *can be replaced by a_k (resp. b_k).*

The case $G = \Phi_\alpha$ can be attacked either directly or by a suitable embedding into the Λ-case. Both methods yield about the same results, however, with the second method less technical difficulties arise.

Let F have a second derivative on (z_0, ∞) for some real z_0. The von Mises condition

$$\lim_{x\to\infty} \frac{xF'(x)}{\overline{F}(x)} = \alpha > 0 \tag{14}$$

implies $F \in \mathcal{D}(\Phi_\alpha)$. Defining $\varepsilon(x) := xF'(x)/\overline{F}(x) - \alpha$, condition (14) can be rewritten as $\lim_{x\to\infty} \varepsilon(x) = 0$. Moreover, equation (14) is equivalent to the fact that F has representation

$$\overline{F}(x) = x^{-\alpha} c \exp\left(-\int_1^x \frac{\varepsilon(t)}{t}\,dt\right) \qquad \text{for all } x \geq z_0, \tag{15}$$

for some $z_0 \in \mathbb{R}$ where $c > 0$ and $\lim_{x\to\infty} \varepsilon(x) = 0$ [cf. Resnick (1987), Section 1.5]. Since F has a second derivative $\varepsilon(\cdot)$ is differentiable as well.

The transformed distribution function $F^{\#}(x) := F(e^x)$ is an element of $\mathcal{D}(\Lambda)$ but it does not necessarily satisfy the von Mises condition (7). But under the additional assumption $\lim_{x\to\infty} \varepsilon'(x)x = 0$ it does.

Again as in the Λ-case we need sufficiently fast convergence in (14).

THEOREM 4. *Let $\{X_n, n \in \mathbb{N}\}$ be a sequence of i.i.d. random variables with common distribution function F on some probability space. Suppose F has a second* (*not necessarily negative*) *derivative in some neighborhood of infinity and let $\varepsilon(x) = xF'(x)/\overline{F}(x) - \alpha$. If*

$$\lim_{x\to\infty} \varepsilon(x) \log\log x = 0 \quad \textit{and} \quad \lim_{x\to\infty} \varepsilon'(x)x(\log\log x)^2 = 0, \tag{16}$$

it is possible to reconstruct the sequence $\{X_n, n \in \mathbb{N}\}$ together with an extremal-Φ_α process $Y_{\Phi_\alpha}(u)$ in $D(0,\infty)$ on a possibly enlarged probability space (Ω, Σ, P) such that for every $\varepsilon > 0$ there exists a set $\Omega_\varepsilon \in \Sigma$ with $P(\Omega_\varepsilon) \geq 1 - \varepsilon$, a constant $K_\varepsilon \geq 1$ and a function $r_\varepsilon(u) : [1,\infty) \to [0,\infty)$ such that

$$u^{-1/\alpha} Y_{\Phi_\alpha}\big(u - v(u)\big) - r_\varepsilon(u) \leq \frac{M_{[u]} - b_{[u]}^{\mathrm{can}}}{a_{[u]}^{\mathrm{can}}} \leq u^{-1/\alpha} Y_{\Phi_\alpha}\big(u + v(u)\big) + r_\varepsilon(u) \tag{17}$$

holds for all $u \geq K_\varepsilon^2$ and all $\omega \in \Omega_\varepsilon$, with $v(u) := K_\varepsilon \sqrt{\log^* u \log_3^* u}$ and we have $\lim_{u\to\infty} r_\varepsilon(u)(\log u)^{-\beta} = 0$ *for all $\beta > 1/\alpha$.*

Furthermore, *if there exists an increasing function $r : (0,\infty) \to (0,\infty)$ such that*

$$\lim_{x\to\infty} r(x) = \infty, \qquad \lim_{x\to\infty} r(x)\frac{\log x}{x} = 0,$$

$$\lim_{x\to\infty} \varepsilon(x)(\log\log x) r\big(1/\overline{F}(x)\big) = 0, \tag{18}$$

$$\lim_{x\to\infty} \varepsilon'(x)x(\log\log x)^2 r\big(1/\overline{F}(x)\big) = 0 \tag{19}$$

holds then $\lim_{u\to\infty} r_\varepsilon(u)(\log u)^{-\beta} r(u) = 0$ *for all $\beta > 1/\alpha$.*

If in addition for some positive sequence (a_k) and some real sequence (b_k),

$$\lim_{k\to\infty} \left(\left(\frac{a_k^{\mathrm{can}}}{a_k} - 1\right)(\log k)^\beta - b_k/a_k\right) r(k) = 0 \qquad \textit{for some } \beta > 1/\alpha$$

holds, *then the canonical constants in* (17) *can be replaced by a_k* (*resp. b_k*).

The case $G = \Psi_\alpha$ is similar to the Φ_α-case. Use the embedding $F^{\#}(x) := F(x_F - e^{-x}) \in \mathcal{D}(\Lambda)$. Thus we can formulate:

THEOREM 5. *Let $\{X_n, n \in \mathbb{N}\}$ be a sequence of i.i.d. random variables with common distribution function F on some probability space. Suppose F has a second (not necessarily negative) derivative in some neighborhood of $x_F < \infty$ and let $\delta(x) = ((x_F - x)F'(x))/\overline{F}(x) - \alpha$. If*

$$\lim_{x\to\infty} \delta(x_F - 1/x)\log\log x = 0 \quad \textit{and}$$
$$\lim_{x\to\infty} \delta'(x_F - 1/x)x^{-1}(\log\log x)^2 = 0, \tag{20}$$

it is possible to reconstruct the sequence $\{X_n, n \in \mathbb{N}\}$ together with an extremal-Ψ_α process $Y_{\Psi_\alpha}(u)$ in $D(0,\infty)$ on a possibly enlarged probability space (Ω, Σ, P) such that for every $\varepsilon > 0$ there exists a set $\Omega_\varepsilon \in \Sigma$ with $P(\Omega_\varepsilon) \geq 1 - \varepsilon$, a constant $K_\varepsilon \geq 1$ and a function $r_\varepsilon(u) : [1,\infty) \to [0,\infty)$ such that

$$u^{1/\alpha} Y_{\Psi_\alpha}\big(u - v(u)\big) - r_\varepsilon(u) \leq \frac{M_{[u]} - b_{[u]}^{\mathrm{can}}}{a_{[u]}^{\mathrm{can}}} \leq u^{1/\alpha} Y_{\Psi_\alpha}\big(u + v(u)\big) + r_\varepsilon(u) \tag{21}$$

holds for all $u \geq K_\varepsilon^2$ and all $\omega \in \Omega_\varepsilon$, with $v(u) := K_\varepsilon\sqrt{\log^* u \log_3^* u}$ and we have $\lim_{u\to\infty} r_\varepsilon(u)(\log\log u)^{-\beta} = 0$ *for all $\beta > 1/\alpha$.*

Furthermore, if there exists an increasing function $r : (0,\infty) \to (0,\infty)$ such that

$$\lim_{x\to\infty} r(x) = \infty, \qquad \lim_{x\to\infty} r(x)\frac{\log x}{x} = 0,$$

$$\lim_{x\to\infty} \delta(x_F - 1/x)(\log\log x) r\big(1/\overline{F}(x_F - 1/x)\big) = 0, \tag{22}$$

$$\lim_{x\to\infty} \delta'(x_F - 1/x)x^{-1}(\log\log x)^2 r\big(1/\overline{F}(x_F - 1/x)\big) = 0 \tag{23}$$

holds then $\lim_{u\to\infty} r_\varepsilon(u)(\log\log u)^{-\beta} r(u) = 0$ *for all $\beta > 1/\alpha$.*

If in addition for some positive sequence (a_k) and some real sequence (b_k),

$$\lim_{k\to\infty}\left(\left(\frac{a_k^{\mathrm{can}}}{a_k} - 1\right)(\log\log k)^\beta + (x_F - b_k)/a_k\right) r(k) = 0 \qquad \textit{for some } \beta > 1/\alpha,$$

holds, then the canonical constants in (21) *can be replaced by a_k (resp. b_k).*

3. Proof of the invariance principle. By Lamperti's result (1) implies $\tilde{Y}_n(t) \Rightarrow Y_G(t)$ in $D(0,\infty)$, therefore [Resnick (1987), Proposition 4.18] $\tilde{Y}_n(t) \Rightarrow Y_G(t)$ in $D[a,b]$ for all $0 < a < b < \infty$. Note that $Y_G(t)$ is stochastically continuous for all $t > 0$.

In the case $G = \Lambda$ we have

$$\lim_{n\to\infty} \frac{a_n}{a_{[nt]}} = 1 \quad \text{and} \quad \lim_{n\to\infty} \frac{b_{[nt]} - b_n}{a_n} = \log t$$

uniformly in $t \in [a,b]$. This is true since $a \circ V$ [for V see below (5)] is slowly varying and V belongs to the function class Π [see, e.g., Bingham, Goldie and Teugels (1989)], arbitrary normalizing constants differ from the canonical choices just by the relations given below (5). Note that uniform convergence implies convergence in the metric of $D[a,b]$. Thus Slutzky's theorem yields

$$Y_n(t) = \tilde{Y}_n(t)\frac{a_n}{a_{[nt]}} - \frac{b_{[nt]} - b_n}{a_n}\frac{a_n}{a_{[nt]}} \Rightarrow Y_\Lambda(t) - \log t \qquad \text{in } D[a,b].$$

Since this is true for all $0 < a < b < \infty$, the claim follows in the case $G = \Lambda$. The other cases are similar.

The fact that $Y_{0,G}(e^t)$ is a stationary Markov process follows since $Y_G(t)$ is a Markov process [Resnick (1987), Proposition 4.7] and stationarity can be seen easily by writing down the one-dimensional and transition distributions of $Y_{0,G}(e^t)$. Recall the identities

$$\Lambda^b(a) = \Lambda(a - \log b), \qquad \Phi_\alpha^b(a) = \Phi_\alpha(b^{-1/\alpha}a) \quad \text{and} \quad \Psi_\alpha^b(a) = \Psi_\alpha(b^{1/\alpha}a).$$

These and Proposition 4.7 in Resnick (1987) yield, for example, in case $G = \Lambda$

$$P\big(Y_{0,\Lambda}(e^s) \le x\big) = P\big(Y_\Lambda(e^s) \le x + s\big) = \Lambda^{e^s}(x+s) = \Lambda(x)$$

and

$$\begin{aligned}
&P\big(Y_{0,\Lambda}(e^{t+s}) \le y \,\big|\, Y_{0,\Lambda}(e^s) = u\big)\\
&\quad = P\big(Y_\Lambda(e^{t+s} - e^s + e^s) \le y + t + s \,\big|\, Y_\Lambda(e^s) = u + s\big)\\
&\quad = \begin{cases} \Lambda^{(e^{t+s}-e^s)}(y+t+s), & \text{if } y+t+s \ge u+s,\\ 0, & \text{otherwise} \end{cases}\\
&\quad = \begin{cases} \Lambda^{(e^t-1)}(y+t), & \text{if } y+t \ge u,\\ 0, & \text{otherwise,} \end{cases}
\end{aligned}$$

similarly in the other cases. Hence we have the desired stationarity. $\square$

4. Proof of the strong approximation results. The method of proof is based on arguments by Deheuvels (1981, 1982), however, in comparison to the original result, we want to end with *one* and not a sequence of approximating extremal processes and more important we want to get uniform error bounds which need more care. In order to make the paper self-contained we repeat with minor changes parts of the proof which were already given by Deheuvels.

Assume that $X_1^E, X_2^E, \ldots$ are i.i.d. standard exponential [$X_1^E \stackrel{d}{=} \mathrm{Exp}(1)$] random variables on some probability space, set $M_0^E = 0$, $M_n^E := \max_{1 \le k \le n} X_k^E$ and define recursively the record times

$$L(1) := 1, \quad L(n+1) := \inf\{j > L(n) : X_j^E > X_{L(n)}^E\}, \qquad n \ge 1,$$

and the time when M_j^E exceeds the level t by

$$E(t) := \inf\{j \geq 0 : M_j^E > t\}.$$

As a shorthand write $z_n := X_{L(n)}^E = M_{L(n)}^E$, $n \geq 1$. It is well known that if $X_n^E \stackrel{d}{=} \mathrm{Exp}(1)$ then the sequence $(z_n)_{n=1}^{\infty}$ forms a homogeneous Poisson point process with unit intensity on $(0, \infty)$. Denote its counting process by $N(t)$; that is, $N(t) = \#\{k \in \mathbb{N} : z_k \leq t\}$, $t \geq 0$.

The following result was proved in Deheuvels (1982). For a proof of the result as stated here, see Fahrner (2000b).

THEOREM 6. *It is possible to complete* $\{z_k, 1 \leq k < \infty\}$ *to a homogeneous Poisson point process* $\{z_k, -\infty < k < \infty\}$ *with unit intensity on* $\mathbb{R}$ *and to reconstruct it together with the sequence* $\{X_n^E, n \in \mathbb{N}\}$ *and a sequence* $\{E_k, -\infty < k < \infty\}$ *of i.i.d. unit exponential random variables, independent of the Poisson process, on a possibly enlarged probability space* (Ω, Σ, P) *such that*

$$\limsup_{n\to\infty} \frac{1}{\sqrt{2n \log\log n}} \left| L(n+1) - \sum_{k=-\infty}^{n} E_k e^{z_k} \right| < \infty \qquad a.s. \tag{24}$$

Actually the limit superior was shown to be a constant. Let $\varepsilon > 0$ be given and define the process

$$Z(t) := \sum_{k=-\infty}^{N(t)} E_k e^{z_k}, \qquad 0 < t < \infty.$$

It is well known that $Z(t)$ has the same finite-dimensional distributions as $Y_{\Lambda}^{\rightarrow}(t)$; see Resnick [(1987), page 195f] or Fahrner [(2000b), Appendix A]. In the next lemma we approximate $E(t)$ by $Z(t)$.

LEMMA 1. *On the probability space* (Ω, Σ, P) *of Theorem* 6 *there exists a positive constant* C_2 *and an event* $A_2 \in \Sigma$ *with* $P(A_2) \geq 1 - \varepsilon/2$ *as well as processes* $\phi(t)$ *and* $\psi(t)$ *on* $(0, \infty)$ *such that*

$$Z(t) - \phi(t) < E(t) \leq Z(t) + \psi(t) \qquad \forall t > 0,\ \forall \omega \in A_2 \tag{25}$$

and:

(i) $\phi(t)$ *and* $\psi(t)$ *are constant except for jumps. Like* $E(t)$ *and* $Z(t)$ *these functions are constant on intervals of the form* (z_n, z_{n+1}),

(ii) $Z(t) - \phi(t)$ *and* $Z(t) + \psi(t)$ *are integers for every* $t > 0$ *and all* $\omega \in A_2$,

(iii) $0 \leq \phi(t), \psi(t) \leq C_2 \max\{1, \sqrt{t \log_2^* t}\}\ \forall t > 0,\ \forall \omega \in A_2$.

PROOF. Denote the lim sup in (24) by $\overline{s}$ and write (24) as

$$\left| L(n+1) - \sum_{k=-\infty}^{n} E_k e^{z_k} \right| \le \sqrt{2n \log\log n}\,(\overline{s} + B_n) \qquad \text{for } n \ge 3$$

where $\lim_{n\to\infty} B_n = 0$ a.s. Note that $E(t) = L(n+1)$ where $n+1 = \min\{k : X_{L(k)} \ge t\} = N(t) + 1$ a.s. $\forall t \in \mathbb{R}$. Thus

$$|E(t) - Z(t)| \le \sqrt{2N(t) \log\log N(t)}\,(\overline{s} + B_{N(t)}) \qquad \text{for } N(t) \ge 3. \tag{26}$$

We will use Egorov's theorem to get a nonrandom bound in (26) for large t. Since $\lim_{n\to\infty} B_n = 0$ a.s. there exists an event $A \in \Sigma$ with $P(A) \le \varepsilon/6$ and $\lim_{n\to\infty} B_n = 0$ uniformly on A^c. Furthermore, since $\lim_{t\to\infty} N(t)/t = 1$ a.s. there exists an event $\tilde{A} \in \Sigma$ with $P(\tilde{A}) \le \varepsilon/6$ and $\lim_{t\to\infty} N(t)/t = 1$ uniformly on $\tilde{A}^c$. Set $A_1 := (A \cup \tilde{A})^c$, then $P(A_1) \ge 1 - \varepsilon/3$ and there exists a t_0 independent of ω such that $3 \le N(t) \le 2t$ and $|B_{N(t)}| \le 1\ \forall t \ge t_0,\ \forall \omega \in A_1$; that is, there exists a C_1 such that

$$\begin{aligned} |E(t) - Z(t)| &\le \sqrt{4t \log\log(2t)}(\overline{s} + 1) \\ &\le C_1 \max\Bigl(1, \sqrt{t \log_2^* t}\Bigr) \qquad \forall t \ge t_0,\ \forall \omega \in A_1. \end{aligned}$$

Hence there exist functions $\phi(t)$ and $\psi(t)$ which satisfy the requirements of the lemma for $t \ge t_0$. Next we want to extend the functions for all $t > 0$, that is, for some event $A_2 \subset A_1$ we want to have

$$Z(t) - \phi(t) < E(t) \le Z(t) + \psi(t) \qquad \forall 0 < t \le t_0,\ \forall \omega \in A_2.$$

By definition we have $E(t) > 0$, therefore we may define $\phi(t) := Z(t) \le Z(t_0)$ for $t \le t_0$. It is also possible to find a function $\psi(t)$ with $E(t_0) \le \psi(t) \le E(t_0) + 1$ for $0 \le t \le t_0$ such that $Z(t) + \psi(t)$ is an integer for every t and since $Z(t) \ge 0$ and $E(t)$ is nondecreasing $Z(t) + \psi(t) \ge \psi(t) \ge E(t_0) \ge E(t)$. Let $\tilde{z}$ denote the time of the first jump after t_0. Defining $\phi(t) = \phi(t_0)$ and $\psi(t) = \psi(t_0)$ for $t \in (t_0, \tilde{z})$, we have (i) and (ii).

Now recall that $Z(t_0) \stackrel{d}{=} Y_\Lambda^{\rightarrow}(t_0)$. Thus by Proposition 4.8(iv) of Resnick (1987) $Z(t_0) \stackrel{d}{=} \operatorname{Exp}(-\log\Lambda(t_0))$. Furthermore $\mathbf{E}(E(t_0)) < \infty$ since

$$\sum_{n=0}^{\infty} P\bigl(E(t_0) > n\bigr) = \sum_{n=0}^{\infty} P(M_n^E \le t_0) = \sum_{n=0}^{\infty} (1 - e^{-t_0})^n < \infty.$$

Thus $Z(t_0)$, $E(t_0)$ are proper random variables, that is, there exists a $C_2 > 0$ such that

$$P\bigl(\max\{C_1, Z(t_0), E(t_0)\} > C_2\bigr) \le \varepsilon/6.$$

Setting $A_2 = A_1 \cap \{\max\{C_1, |Z(t_0)|, E(t_0)\} \le C_2\}$ completes the proof. $\square$

Inverting (25) we can approximate our maxima by an extremal-Exp(1) process:

LEMMA 2. *On the probability space* (Ω, Σ, P) *of Theorem* 6 *there exists an extremal*-Exp(1) *process* $Y_{\mathrm{Exp}(1)}(s)$ *in* $D(0,\infty)$, *an event* $A_3 \in \Sigma$ *with* $P(A_3) \geq 1 - 5\varepsilon/6$ *and a positive constant* C_4 *such that*

$$M^E_{[s-\tilde{v}(s)]} \leq Y_{\mathrm{Exp}(1)}(s) \leq M^E_{[s+\tilde{v}(s)]} \qquad \forall \omega \in A_3,\ \forall s > C_4^2 \tag{27}$$

where $\tilde{v}(s) := C_4\sqrt{\log^* s \log_3^* s}$.

PROOF. Note that by definition of E

$$E(t) > n \quad \Longrightarrow \quad M_n^E \leq t, \qquad E(t) \leq n \quad \Longrightarrow \quad M_n^E > t \qquad \forall t > 0,\ \forall n \in \mathbb{N}.$$

Let $E(t-)$ denote the left-hand limit of E at t. We will show

$$E(t-) \leq n \quad \Longrightarrow \quad M_n^E \geq t.$$

For this, take a sequence $\{t_k, k \in \mathbb{N}\}$ with $t_k < t$ and $t_k \uparrow t$, $k \to \infty$. By monotonicity of E we get $E(t_k) \leq E(t-) \leq n$, thus $M_n^E > t_k$ for all $k \in \mathbb{N}$ and therefore $M_n^E \geq t$.

Since $E(t-) \leq Z(t-) + \psi(t-)$ by (25), setting $t = Z^{\rightarrow}(x)$ and noting that we have $Z(Z^{\rightarrow}(x)-) \leq x$, we get $E(Z^{\rightarrow}(x)-) \leq [x + \psi(Z^{\rightarrow}(x))]$, thus $M^E_{[x+\psi(Z^{\rightarrow}(x))]} \geq Z^{\rightarrow}(x)$.

Since also $Z(Z^{\rightarrow}(x)) \geq x$ for $x > 0$, we get $E(Z^{\rightarrow}(x)) > x - \phi(Z^{\rightarrow}(x)) \geq [x - \phi(Z^{\rightarrow}(x))]$ and thus

$$M^E_{[x-\phi(Z^{\rightarrow}(x))]} \leq Z^{\rightarrow}(x) \leq M^E_{[x+\psi(Z^{\rightarrow}(x))]} \qquad \forall x > 0. \tag{28}$$

Define $S(s) := \Lambda^{\leftarrow}(1 - e^{-s}) = -\log(-\log(1 - e^{-s}))$. Using the inequalities $y \geq -\log(1-y) \geq 1/(1-y) - 1,\ 0 \leq y < 1$, we find putting $S(s) = s + h(s)$ that $-e^{-s}/(1 - e^{-s}) \leq h(s) \leq 0$ and $\{Z^{\rightarrow}(S(s)), s > 0\}$ is a version of an extremal-Exp(1) process [cf. Resnick (1987)] which we will denote by $Y_{\mathrm{Exp}(1)}(s)$. Note that this process is right continuous and thus an element of $D(0,\infty)$, and we have $(Y^{\rightarrow}_{\mathrm{Exp}(1)})^{\rightarrow}(s) = Y_{\mathrm{Exp}(1)}(s)$.

It is easy to see that $\lim_{s\to\infty} Y_{\mathrm{Exp}(1)}(s)/\log s = 1$ a.s. Hence, by Egorov's theorem there exists an event B with $P(B) \leq \varepsilon/6$ and a s_0 independent of $\omega \in B^c$ with $Y_{\mathrm{Exp}(1)}(s) \leq 2\log s\ \forall s \geq s_0,\ \forall \omega \in B^c$. Obviously there is a constant $C_3 > 0$ with $P(Y_{\mathrm{Exp}(1)}(s_0) > C_3) \leq \varepsilon/6$. Set $A_3 := A_2 \cap \{Y_{\mathrm{Exp}(1)}(s_0) \leq C_3\} \cap B^c$. Then $P(A_3^c) \leq 5\varepsilon/6$ and

$$Y_{\mathrm{Exp}(1)}(s) \leq 2C_3 \log^* s \qquad \forall s > 0,\ \forall \omega \in A_3.$$

By setting $x = S(s)$ in (28) the claim follows by observing

$$
\begin{aligned}
&x - \phi(Z^{\rightarrow}(x))\\
&\quad = s + h(s) - \phi\bigl(Y_{\mathrm{Exp}(1)}(s)\bigr)\\
&\quad \ge s - \Bigl(e^{-s}/(1-e^{-s}) + C_2 \max\Bigl\{1, \sqrt{Y_{\mathrm{Exp}(1)}(s) \log_2^*\bigl(Y_{\mathrm{Exp}(1)}(s)\bigr)}\Bigr\}\Bigr)\\
&\quad \ge s - C_4\sqrt{\log^* s \log_3^* s} \qquad \text{for some } C_4 > 0 \text{ and all } s > C_4^2.
\end{aligned}
$$

The proof of the upper bound is even simpler. □

REMARK. Note that the construction of $Y_{\mathrm{Exp}(1)}(s) = Z^{\rightarrow}(S(s))$ does *not* depend on the choice of ε.

PROOF OF THEOREM 2. Let $\{X_n, n \in \mathbb{N}\}$ be a sequence of random variables distributed according to F on some probability space as in the statement of Theorem 2. Assume that F is continuous except for jumps at $\{x_j : j \in J\}$ for some at most countable set J. Without loss of generality we can define a sequence $\{U_n', n \in \mathbb{N}\}$ of i.i.d. uniformly on (0, 1) distributed random variables independent of $\{X_n, n \in \mathbb{N}\}$ on this probability space. Define

$$
\begin{aligned}
U_n &= I(X_n \notin \{x_j : j \in J\}) F(X_n)\\
&\qquad + \sum_{j \in J} I(X_n = x_j)\bigl(F(x_j-) + \bigl(F(x_j) - F(x_j-)\bigr)U_n'\bigr)
\end{aligned}
$$

and observe that U_n are i.i.d. uniformly on (0, 1) distributed random variables and $X_n = F^{\leftarrow}(U_n)$ a.s.

Writing $F_E(x) := (1 - e^{-x}) I_{[0,\infty)}(x)$ the random variables $X_n^E := F_E^{\leftarrow}(U_n)$ have an unit exponential distribution. Let $M_n^U := \max_{1 \le k \le n} U_k$ and $M_n^E := \max_{1 \le k \le n} X_k^E = F_E^{\leftarrow}(M_n^U)$. Then by Lemma 2 applied to the sequence $\{X_n^E, n \in \mathbb{N}\}$ we get an enlarged probability space (Ω, Σ, P) and an extremal-Exp(1) process $Y_{\mathrm{Exp}(1)}(s)$ such that (27) holds. Define

$$
Y_G(s) := G^{\leftarrow}\bigl(F_E\bigl(Y_{\mathrm{Exp}(1)}(s)\bigr)\bigr). \tag{29}
$$

By writing down the finite-dimensional distributions, we see that $Y_G(s)$ is a version of an extremal-G process in $D(0, \infty)$. Let $T(u) = (M_{[u]} - b_{[u]})/a_{[u]}$ and

$$
R(u) := T(u) - G^{\leftarrow}\bigl(F_E^u\bigl(M_{[u]}^E\bigr)\bigr), \qquad u \ge 1.
$$

Note that if F is continuous $R(u) = T(u) - G^{\leftarrow}(F^u(M_{[u]}))$ as claimed in the statement of Theorem 2, since then $U_n = F(X_n)$ and $M_n^E = F_E^{\leftarrow}(F(M_n))$.

We have for $\omega \in A_3$ and for $s > C_4^2$

$$
\begin{aligned}
T\bigl(s - \tilde{v}(s)\bigr) &= G^{\leftarrow}\bigl(F_E^{s-\tilde{v}(s)}\bigl(M_{[s-\tilde{v}(s)]}^E\bigr)\bigr) + R\bigl(s - \tilde{v}(s)\bigr)\\
&\le G^{\leftarrow}\bigl(F_E^{s-\tilde{v}(s)}\bigl(Y_{\mathrm{Exp}(1)}(s)\bigr)\bigr) + R\bigl(s - \tilde{v}(s)\bigr) \qquad \text{by (27)}\\
&= Y_0\bigl(s - \tilde{v}(s), \tilde{v}(s)\bigr) + R\bigl(s - \tilde{v}(s)\bigr).
\end{aligned}
$$

Indeed, using the identities

$$\begin{aligned} &\Lambda^{\leftarrow}(a^b) = \Lambda^{\leftarrow}(a) - \log b, \qquad \Phi_\alpha(a^b) = b^{-1/\alpha}\Phi_\alpha^{\leftarrow}(a) \quad \text{and} \\ &\Psi_\alpha^{\leftarrow}(a^b) = b^{1/\alpha}\Psi_\alpha^{\leftarrow}(a) \end{aligned} \tag{30}$$

we get the a.s. equations

$$\begin{aligned} \Lambda^{\leftarrow}\big(F_E^{s-\tilde v(s)}\big(Y_{\mathrm{Exp}(1)}(s)\big)\big) &= Y_\Lambda(s) - \log\big(s - \tilde v(s)\big), \\ \Phi_\alpha^{\leftarrow}\big(F_E^{s-\tilde v(s)}\big(Y_{\mathrm{Exp}(1)}(s)\big)\big) &= \big(s - \tilde v(s)\big)^{-1/\alpha} Y_{\Phi_\alpha}(s), \\ \Psi_\alpha^{\leftarrow}\big(F_E^{s-\tilde v(s)}\big(Y_{\mathrm{Exp}(1)}(s)\big)\big) &= \big(s - \tilde v(s)\big)^{1/\alpha} Y_{\Psi_\alpha}(s), \end{aligned}$$

where the processes $Y_\Lambda(s)$, $Y_{\Phi_\alpha}(s)$ and $Y_{\Psi_\alpha}(s)$ are defined by (29). Let $g(s) := s - \tilde v(s)$. Eventually this is a strictly increasing differentiable function and without loss of generality we may therefore assume $C_4^2 \ge e^e$. Then

$$\sup_{s\ge C_4^2} \frac{\tilde v(s)}{s} \le C_4 \sup_{s \ge C_4^2} \frac{\log^* s}{s} = \frac{\log^*(C_4^2)}{C_4} =: C_5 < 1.$$

Thus $g(s) \ge (1 - C_5)s\ \forall s \ge C_4^2$ or equivalently $g^{\leftarrow}(u) \le u/(1 - C_5)\ \forall u \ge C_4^2$. For every $u > C_4^2$ there exists a unique $s > C_4^2$ with $u = g(s)$, namely $s = g^{\leftarrow}(u)$. Therefore

$$\begin{aligned} T(u) = T(g(s)) &\le Y_0\big(s - \tilde v(s), \tilde v(s)\big) + R\big(s - \tilde v(s)\big) \\ &= Y_0\big(u, \tilde v(g^{\leftarrow}(u))\big) + R(u) \qquad \forall u \ge C_4^2 \end{aligned}$$

and

$$\begin{aligned} \tilde v(g^{\leftarrow}(u)) \le \tilde v\left(\frac{u}{1 - C_5}\right) &\le C_4 \sqrt{\log^*\left(\frac{u}{1 - C_5}\right)\log_3^*\left(\frac{u}{1 - C_5}\right)} \\ &\le C_6\sqrt{\log^* u \log_3^* u} \qquad \forall u \ge C_6^2 \end{aligned}$$

for some $C_6 > 0$. The proof of the lower bound is similar. Thus we may set $\Omega_\varepsilon = A_3$ of Lemma 2 and $K_\varepsilon = \max\{C_4, C_6, e^{1/2}\}$.

Next we prove $R(u) \to 0$ in probability as $u \to \infty$. It suffices to show $R(k) \to 0$ in probability as $k \to \infty$, $k \in \mathbb{N}$, since using the identities (30) we get for $k \le u < k + 1$

$$R(k) \le R(u) \le \begin{cases} R(k) + \log(1 + 1/k), & \text{when } G = \Lambda, \\ R(k) + \big(1 - (1 + 1/k)^{-1/\alpha}\big)\Phi_\alpha^{\leftarrow}\big(F_E^k(M_k^E)\big), & \text{when } G = \Phi_\alpha, \\ R(k) + \big(1 - (1 + 1/k)^{1/\alpha}\big)\Psi_\alpha^{\leftarrow}\big(F_E^k(M_k^E)\big), & \text{when } G = \Psi_\alpha. \end{cases}$$

Note that $\Phi_\alpha^{\leftarrow}(F_E^k(M_k^E))$ and $\Psi_\alpha^{\leftarrow}(F_E^k(M_k^E))$ are both stochastically bounded and $|1 - (1 + 1/k)^{\pm 1/\alpha}| \sim 1/(\alpha k)$.

Define $V_k := (M_k^U)^k$. Then V_k is uniformly distributed on $(0, 1)$ and since $M_k = F^{\leftarrow}(M_k^U)$ a.s.

$$R(k) = \frac{M_k - b_k}{a_k} - G^{\leftarrow}\big(F_E^k(M_k^E)\big)$$
$$= \frac{F^{\leftarrow}(V_k^{1/k}) - b_k}{a_k} - G^{\leftarrow}(V_k) \qquad \text{a.s.}$$

By inversion of $F^k(a_k t + b_k) \to G(t)$ as $k \to \infty$ we get [Resnick (1987) Proposition 0.1] $(F^{\leftarrow}(t^{1/k}) - b_k)/a_k \to G^{\leftarrow}(t)$ as $k \to \infty$, and since F is increasing and G is continuous, the convergence is locally uniform in $(0, 1)$. Given $\delta > 0$ choose k_0 such that

$$\sup_{\delta/2 \le t \le 1-\delta/2} \big|\big(F^{\leftarrow}(t^{1/k}) - b_k\big)/a_k - G^{\leftarrow}(t)\big| \le \delta \qquad \forall\, k \ge k_0$$

whence

$$P\big(|R(k)| > \delta\big) \le P\big(|R(k)| > \delta;\, V_k \in [\delta/2, 1 - \delta/2]\big) + P\big(V_k \notin [\delta/2, 1 - \delta/2]\big)$$
$$\le 0 + \delta \qquad \forall\, k \ge k_0.$$

This completes the proof of Theorem 2. $\square$

PROOF OF THEOREM 3. Consider the construction of Theorem 2. Since

$$|R(u) - R(k)| \le 1/k \qquad \text{for } k \le u < k + 1,\ k \in \mathbb{N},$$

it suffices to show $R(k)r(k) \to 0$ a.s. as $k \to \infty$, $k \in \mathbb{N}$, where we set $r(k) \equiv 1$ for the first part of the theorem. If we have shown this, by Egorov's theorem, there exists a $\Omega_\varepsilon \subseteq A_3$ such that $P(\Omega_\varepsilon) \ge 1 - \varepsilon$ and $R(k)r(k) \to 0$ uniformly in $\omega \in \Omega_\varepsilon$, and we may take $r_\varepsilon(u) = \sup_{\omega \in \Omega(\varepsilon)} |R(u)|$. In the case of canonical constants we have

$$R(k) = (M_k - b_k)/a_k - \Lambda^{\leftarrow}\big(F^k(M_k)\big)$$
$$= \int_{b_k}^{M_k} \frac{dt}{a(b_k)} + \log\big(-\log\big(1 - \overline{F}(M_k)\big)\big) + \log k.$$

From $\overline{F}(b_k) = 1/k$ and (9) we get $-\log k = \log c - \int_{z_0}^{b_k} dt/a(t)$. Now $\overline{F}(M_k) = \exp(\log c - \int_{z_0}^{M_k} dt/a(t))$ and $-\log(-\log(1 - e^{-t})) = t + \mathcal{O}(e^{-t})$ as $t \to \infty$, thus

$$R(k) = \int_{b_k}^{M_k} \left(\frac{1}{a(b_k)} - \frac{1}{a(t)}\right) dt + \mathcal{O}\big(\overline{F}(M_k)\big). \tag{31}$$

Theorem 3.5.2 of Embrechts, Klüppelberg and Mikosch (1997) implies $\overline{F}(M_k) \le \log k/k$ a.s. for all sufficiently large k, hence substituting $u = (t - b_k)/(M_k - b_k)$

and writing $x_k := (M_k - b_k)/a(b_k)$ we get

$$R(k) = x_k \int_0^1 \left(1 - \frac{a(b_k)}{a(b_k + x_k a(b_k)u)}\right) du + \mathcal{O}(\log k/k).$$

Note that under condition (12) we have $x_k = \mathcal{O}(\log\log k)$ by Theorem 2 of de Haan and Hordijk (1972) (the restriction $x_F = \infty$ is removable) and following the proof of Theorem 3 of de Haan and Hordijk (1972) from their equation (25) to the end on page 1195 with $\phi(n) = \log\log n$ the first two parts of Theorem 3 are proven. The last part follows, since

$$\frac{M_k - b_k}{a_k} = \frac{M_k - b_k^{\mathrm{can}}}{a_k^{\mathrm{can}}} + \left(\frac{a_k^{\mathrm{can}}}{a_k} - 1\right)\frac{M_k - b_k^{\mathrm{can}}}{a_k^{\mathrm{can}}} + \frac{b_k^{\mathrm{can}} - b_k}{a_k}$$

and $(M_k - b_k^{\mathrm{can}})/a_k^{\mathrm{can}} = \mathcal{O}(\log\log k)$. $\square$

PROOF OF THEOREM 4. Define $X_k^\# := \log(\max\{X_k, 1\})$, $M_k^\# := \max_{1\le i\le k} X_i^\#$ and $F^\#(x) := F(e^x)$. Then $P(X_k^\# \le x) = F^\#(x)$ for $x > 0$ and $F^\# \in \mathcal{D}(\Lambda)$, since

$$\lim_{k\to\infty} P\big(\alpha\big(M_k^\# - \log V(k)\big) \le x\big) = \lim_{k\to\infty} P\big(M_k/V(k) \le e^{x/\alpha}\big)$$
$$= \Phi_\alpha\big(e^{x/\alpha}\big) = \Lambda(x) \qquad \forall x \in \mathbb{R}.$$

We will apply Theorem 3 to $\{X_k^\#, k \in \mathbb{N}\}$. Under (16) F has representation (15) with

$$F'(x) = \overline{F}(x)\frac{\alpha(x)}{x} \quad \text{and} \quad F''(x) = \frac{\overline{F}(x)}{x^2}\big(\alpha'(x)x - \alpha(x) - \alpha^2(x)\big)$$

where $\alpha(x) = \alpha + \varepsilon(x)$. From these identities it can be seen easily that $d^2/dx^2 F^\#(x) < 0$ for large x. Differentiating $a(x) := \overline{F^\#}(x)/F^{\#'}(x) = 1/\alpha(e^x)$ we see

$$a'(\log x) = -\frac{\alpha'(x)x}{\alpha^2(x)} \sim -\frac{\varepsilon'(x)x}{\alpha^2}.$$

Recall that $\overline{F}(x)$ is regularly varying, hence $\log\log(1/\overline{F^\#}(\log x)) \sim \log\log(1/\overline{F}(x)) \sim \log\log x$ as $x \to \infty$. Thus (16) implies (10) [resp. (19) implies (12)] of Theorem 3 for $F^\#$. The canonical constants for $F^\#$ are $b_k^{\mathrm{can}} = \log V(k)$ and $a_k^{\mathrm{can}} = 1/\alpha(V(k))$. Again since $V(k)$ is a regularly varying sequence we have $\log\log V(k) \sim \log\log k$ as $k \to \infty$ and under (18) we get (13) for $a_k = 1/\alpha$, $b_k = b_k^{\mathrm{can}}$. An Application of Theorem 3 yields an extremal-Λ process $Y_\Lambda(u)$ such that for given $\varepsilon > 0$ there is a set Ω_ε with $P(\Omega_\varepsilon) \ge 1 - \varepsilon$, a constant $K_\varepsilon \ge 1$ and a function $\overline{r}_\varepsilon(u)$, $u \ge 1$ such that

$$\alpha \log \frac{M_{[u]}}{V([u])} = \alpha\big(M_{[u]}^\# - \log V([u])\big) \le Y_\Lambda\big(u + v(u)\big) - \log u + \overline{r}_\varepsilon(u), \tag{32}$$

$$\alpha \log \frac{M_{[u]}}{V([u])} = \alpha\big(M_{[u]}^\# - \log V([u])\big) \ge Y_\Lambda\big(u - v(u)\big) - \log u - \overline{r}_\varepsilon(u) \tag{33}$$

for all $u \geq K_\varepsilon^2$ and all $\omega \in \Omega_\varepsilon$, where $\lim_{u\to\infty} \overline{r}_\varepsilon(u)r(u) = 0$. (Without loss of generality we may assume that $M_{[u]} > 1\ \forall\,\omega \in \Omega_\varepsilon$ and all $u \geq K_\varepsilon$. If not, increase K_ε and apply Egorov's theorem to get uniformity in ω on a slightly smaller set Ω'_ε.) Apply the function $e^{x/\alpha} = (\Phi_\alpha^{\leftarrow} \circ \Lambda)(x) = 1 + \mathcal{O}(x)$ for $x \to 0$ to the inequalities (32) and (33). Recall that $\Phi_\alpha^{\leftarrow} \circ \Lambda \circ Y_\Lambda(u)$ is an extremal-Φ_α process denoted by $Y_{\Phi_\alpha}(u)$. Therefore

$$\frac{M_{[u]}}{V([u])} \leq u^{-1/\alpha} Y_{\Phi_\alpha}\big(u + v(u)\big)\big(1 + \mathcal{O}(\overline{r}_\varepsilon(u))\big),$$
$$\frac{M_{[u]}}{V([u])} \geq u^{-1/\alpha} Y_{\Phi_\alpha}\big(u - v(u)\big)\big(1 + \mathcal{O}(\overline{r}_\varepsilon(u))\big).$$

To complete the proof of the theorem we need the following lemmas on the rate of growth of $Y_{\Phi_\alpha}(u)$ and M_n:

LEMMA 3. *The constructed process $Y_{\Phi_\alpha}(u)$ satisfies, for all $\beta > 1/\alpha$,*

$$\lim_{u\to\infty} \frac{u^{-1/\alpha} Y_{\Phi_\alpha}(u)}{(\log u)^\beta} = 0 \quad \textit{and} \quad \lim_{u\to\infty} (\log\log u)^\beta u^{-1/\alpha} Y_{\Phi_\alpha}(u) = \infty \qquad \textit{a.s.}$$

This is not true for $\beta = 1/\alpha$.

PROOF. For the first claim it suffices to show that $P(k^{-1/\alpha} Y_{\Phi_\alpha}(k) > \eta(\log k)^\beta \text{i.o.}) = 0$ for all $\eta > 0$. Now

$$\{Y_{\Phi_\alpha}(k), k \in \mathbb{N}\} \stackrel{d}{=} \left\{ \max_{1\leq j\leq k} Y_j, k \in \mathbb{N} \right\} \qquad \text{in } R^\infty,$$

where Y_j are i.i.d. Φ_α-distributed random variables Dwass (1964). Thus

$$P\big(k^{-1/\alpha} Y_{\Phi_\alpha}(k) > \eta(\log k)^\beta \text{i.o.}\big) = P\big(Y_k > \eta k^{1/\alpha}(\log k)^\beta \text{i.o.}\big) = 0$$

by the Borel–Cantelli lemma. For the second claim apply Theorem 3.5.2 of Embrechts, Klüppelberg and Mikosch (1997). □

It is an immediate consequence of Lemma 3 that for $\beta > 1/\alpha$,

$$\frac{u^{-1/\alpha} Y_{\Phi_\alpha}(u \pm v(u))}{(\log u)^\beta} \to 0 \qquad \text{a.s. as } u \to \infty$$

and the result follows for the case of canonical constants. For the last claim we need:

LEMMA 4. *If $F \in \mathcal{D}(\Phi_\alpha)$ then*

$$\lim_{k\to\infty} \frac{M_k}{V(k)} (\log k)^{-\beta} = 0 \qquad a.s.\ \forall\,\beta > 1/\alpha.$$

PROOF. Recall that $V(k)$ is regularly varying with index $1/\alpha$. Define $\gamma := (\alpha\beta - 1)/2 > 0$ then $1 + \gamma < \alpha\beta$ and using Potter bounds [Bingham, Goldie and Teugels (1989)] we see

$$\frac{V(k(\log k)^{1+\gamma})}{V(k)} \le (\log k)^{\beta}/c$$

for $k \ge k_0$ and some $0 < c < 1$. Thus for large k,

$$\begin{aligned} P\big(X_k > \eta V(k)(\log k)^{\beta}\big) &\le P\big(X_k > c\eta V\big(k(\log k)^{1+\gamma}\big)\big) \\ &= \frac{\overline{F}(c\eta V(k(\log k)^{1+\gamma}))}{\overline{F}(V(k(\log k)^{1+\gamma}))}\overline{F}\big(V\big(k(\log k)^{1+\gamma}\big)\big) \\ &\sim \frac{(c\eta)^{-1/\alpha}}{k(\log k)^{1+\gamma}}, \end{aligned}$$

and the lemma follows from the Borel–Cantelli lemma. □

The proof of Theorem 5 is similar to that of Theorem 4, using the embedding $X_k^{\#} := \log(\max\{1, (x_F - X_k)^{-1}\})$.

5. Remarks and applications. We first discuss that some additional conditions on the distribution function F are really needed to get almost sure convergence of the error term R.

1. Let $\overline{F}(x) = \exp(-\int_e^x \frac{dt}{a(t)})$ where $a(t) = t/\log\log t$. We will see that

$$\limsup_{k\to\infty} R(k) = \infty \qquad \text{a.s.}$$

Let $\overline{F}(b_k) = k^{-1}$ and $a_k = a(b_k)$ be the canonical choices. Due to the inequality $-\Lambda^{\leftarrow}(x^k) \ge \log(1-x) + \log k$ we get

$$\begin{aligned} R(k) &= (M_k - b_k)/a_k - \Lambda^{\leftarrow}\big(F^k(M_k)\big) \\ &\ge \frac{\log\log b_k}{b_k}(M_k - b_k) + \log\overline{F}(M_k) - \log\overline{F}(b_k) \\ &= \log\log b_k\left(\frac{M_k}{b_k} - 1\right) + \log\frac{\overline{F}(M_k)}{\overline{F}(b_k)}. \end{aligned}$$

Moreover,

$$\log\frac{\overline{F}(M_k)}{\overline{F}(b_k)} = -\int_{b_k}^{M_k}\frac{dt}{a(t)} = (\log b_k)(\log\log b_k - 1) - (\log M_k)(\log\log M_k - 1).$$

By Theorem 1 of de Haan and Hordijk (1972) $\limsup_{k\to\infty} M_k/b_k = e$ a.s., thus there exists for almost every ω a $k_0(\omega)$ such that

$$\frac{13}{5} \le \frac{M_{n_k}(\omega)}{b_{n_k}} \le e^{3/2}$$

for some subsequence $\{n_k, k \in \mathbb{N}\}$ of $\mathbb{N}$ and all $k \geq k_0(\omega)$. Since for large x the function $(\log x)(\log\log x - 1)$ is increasing, we get

$$(\log M_{n_k})(\log\log M_{n_k} - 1) \leq \log\bigl(e^{3/2} b_{n_k}\bigr)\bigl(\log\log\bigl(e^{3/2} b_{n_k}\bigr) - 1\bigr)$$
$$= (3/2 + \log b_{n_k})\bigl(\log(3/2 + \log b_{n_k}) - 1\bigr).$$

Writing x for $\log b_{n_k}$ yields

$$R(n_k) \geq (\log x)(13/5 - 1) + x(\log x - 1) - (3/2 + x)\bigl(\log(3/2 + x) - 1\bigr)$$
$$= 3/2 + \log\biggl(1 - \frac{3/2}{3/2 + x}\biggr)^{3/2+x} + \frac{1}{10}\log x \sim \frac{1}{10}\log x.$$

$\lim_{k\to\infty} b_{n_k} = \infty$ implies $R(n_k) \to \infty$ a.s. as $k \to \infty$.

2. In fact for the distribution discussed in 1, no extremal-Λ process $Y_\Lambda(u)$ can be constructed such that (11) holds, because (11) implies (cf. Application 1 below) that

$$\limsup_{n\to\infty} \frac{M_n - b_n^{\mathrm{can}}}{a_n^{\mathrm{can}} \log\log n} = 1$$

which is not true; compare de Haan and Hordjik [(1972), page 1193].

3. The most general $F \in \mathcal{D}(\Lambda)$ has representation

$$\overline{F}(x) = c(x)\exp\biggl(-\int_{z_0}^{x} \frac{dt}{a(t)}\biggr)$$

where $c(x) \to c > 0$ as $x \to x_F$ and $a(t)$ is absolutely continuous and satisfies (8). Thus weakening our assumptions on the smoothness of F we may need a nonconstant $c(x)$. In this case if $\lim_{x\to x_F} a'(x)(\log\log(1/\overline{F}(x))^2 = 0$ we still have $\lim_{k\to\infty} R(k) = 0$ a.s., but we cannot control the rate easily. Imitating the proof of Theorem 3, the analog of equation (31) reads

$$R(k) = \int_{b_k}^{M_k} \biggl(\frac{1}{a(b_k)} - \frac{1}{a(t)}\biggr)dt + \log\frac{c(b_k)}{c(M_k)} + \mathcal{O}\bigl(\overline{F}(M_k)\bigr). \tag{34}$$

Since $b_k \to x_F$ and $M_k \to x_F$ as $k \to \infty$, $\log(c(b_k)/c(M_k))$ tends to zero with an undetermined rate.

As an example consider $\overline{F}(x) = c(x)e^{-x}$ with a continuous function c which satisfies $\lim_{x\to\infty} c(x) = 1$. Here, $a(t) \equiv 1$ and $b_k = (1/\overline{F})^{\leftarrow}(k) = \log k + \log c(b_k)$, so

$$R(k) = M_k^F - b_k^F + \log\bigl(-\log\bigl(1 - \overline{F}(M_k^F)\bigr)\bigr) + \log k$$
$$\geq M_k^F + \log\overline{F}(M_k^F) + \log k - b_k^F = \log c(M_k^F) - \log c(b_k)$$

by using $\log(-\log(1 - x)) \geq \log x$.

Using $b_k = \log k + o(1)$ and Theorem 2 and Lemma 4 of de Haan and Hordijk (1972) we see $\limsup_{k\to\infty}(M_k - b_k)/\log\log k = 1$; that is, $M_k > b_k + 1/2\log\log k$ along a subsequence. In order to get a rate on $R(k)$ we would have to control the function c over intervals $[b_k, b_k + 1/2\log\log k]$.

EXAMPLES. Typical examples for $F \in \mathcal{D}(\Lambda)$ are the Weibull distributions given by $\overline{F}(x) = \exp(-cx^{-\tau})$, $x > 0$, for some $c, \tau > 0$. Here $a(x) = x^{1-\tau}/(c\tau)$, $x > 0$, and if $\tau \neq 1$, we may take $r(x) = \log x/(\log\log x)^3$. The case $\tau = 1$ corresponds to an exponential distribution with mean $1/c$, for which we may take any $r(x) = o(x/\log x)$, since $a'(x) \equiv 0$. If F is the standard normal distribution, then also $F \in \mathcal{D}(\Lambda)$ and Mill's ratio [Feller (1968), page 175] shows that again $r(x) = \log x/(\log\log x)^3$ is a valid choice.

A typical example for $F \in \mathcal{D}(\Phi_\alpha)$ is $\overline{F}(x) = c\,x^{-\alpha}(\log x)^\gamma$ $(x \geq x_0)$ for some real γ and some $c, x_0 > 0$. If $\gamma = 0$, then $\varepsilon(x) \equiv 0$ and we may take any $r(x) = o(x/\log x)$, if $\gamma \neq 0$ then $\varepsilon(x) = -\gamma/\log x$ and a possible choice for r is $r(x) = (\log x)^\eta$ for any $0 < \eta < 1$. Another example is $\overline{F}(x) = c\,x^{-\alpha}(1 + x^{-\gamma})$ for some $\gamma > 0$, for which $\varepsilon(x) = \gamma/(1 + x^\gamma)$ and we can take $r(x) = x^\eta$ with $0 < \eta < \min\{1, \gamma/\alpha\}$.

The uniform distribution on $(0, 1)$ is an element of $\mathcal{D}(\Psi_1)$ with $\delta(x) \equiv 0$ and thus $r(x) = o(x/\log x)$. A further example is $\overline{F}(x) = (x_F - x)^\alpha \times (-\log(x_F - x))^\gamma \in \mathcal{D}(\Psi_\alpha)$ with some $\alpha > 0$, $\gamma \neq 0$ and $x_F \in \mathbb{R}$. We have $\delta(x_F - 1/x) = \gamma/\log x$, therefore we may take $r(x) = (\log x)^\eta$ for any $0 < \eta < 1$.

APPLICATIONS.

1. As a first application of our invariance principle, recall that Resnick (1974) proved

$$\limsup_{t\to\infty} \frac{Y_\Lambda(t) - \sum_{j=1}^n \log_j^* t}{\log_{n+1}^* t} = 1, \qquad \liminf_{t\to\infty} \frac{Y_\Lambda(t) - \log^* t}{\log_3^* t} = -1 \qquad \text{a.s.}$$

for any $n \in \mathbb{N}$. Thus for an i.i.d. sequence $\{X_k, k \in \mathbb{N}\}$ of random variables with common distribution function F which satisfies (10), we get, for any $n \in \mathbb{N}$,

$$\limsup_{k\to\infty} \frac{(M_k - b_k^{\mathrm{can}})/a_k^{\mathrm{can}} - \sum_{j=2}^n \log_j^* k}{\log_{n+1}^* k} = 1,$$
$$\liminf_{k\to\infty} \frac{M_k - b_k^{\mathrm{can}}}{a_k^{\mathrm{can}} \log_3^* k} = -1 \qquad \text{a.s.}$$

by our strong approximation result and the Hewitt–Savage zero–one law. This is a refinement of Theorem 2 of de Haan and Hordijk (1972).

2. As mentioned in the Introduction our main motivation for developing a strong approximation result with error bounds is the study of logarithmic averages of sample maxima. More precisely, we want to have a strong invariance principle of the form

$$\sum_{k=1}^n \frac{1}{k} f\left(\frac{M_k - b_k}{a_k}\right) - m\log n = \sigma W(\log n) + o_{\mathrm{a.s.}}\big((\log n)^{1/2-\eta}\big)$$

where $m \in \mathbb{R}$, $\sigma, \eta > 0$, W is a Wiener process and f is some smooth function. Suppose that Theorem 3 holds with $r(x) = (\log x)^{1/2+2\eta}$, $\eta \in (0, 1/4)$. Then it can be shown that

$$\sum_{k=1}^{n} \frac{1}{k} f\left(\frac{M_k - b_k}{a_k}\right) = \int_1^n \frac{1}{u} f\big(Y_\Lambda(u) - \log u\big)\, du + o_{\text{a.s.}}\big((\log n)^{1/2-\eta}\big)$$
$$= \int_0^{\log n} f\big(Y_\Lambda(e^v) - v\big)\, dv + o_{\text{a.s.}}\big((\log n)^{1/2-\eta}\big)$$

and an invariance principle for the integral can be obtained with standard methods, since the process $\{Y_\Lambda(e^u) - u, u \geq 0\}$ is a stationary strong Markov process; see Fahrner (2001).

REFERENCES

BINGHAM, N. H., GOLDIE, C. M. and TEUGELS, J. L. (1987). *Regular Variation*. Cambridge Univ. Press.

CSÖRGŐ, M. and RÉVÉSZ, P. (1981). *Strong Approximations in Probability and Statistics*. Academic Press, New York.

DE HAAN, L. (1970). *On Regular Variation and Its Applications to the Weak Convergence of Sample Extremes. Math. Centre Tract* **32**. CWI, Amsterdam.

DE HAAN, L. and HORDIJK, A. (1972). The rate of growth of sample maxima. *Ann. Math. Statist.* **43** 1185–1196.

DEHEUVELS, P. (1981). The strong approximation of extremal processes. *Z. Wahrsch. Verw. Gebiete* **58** 1–6.

DEHEUVELS, P. (1982). Strong approximation in extreme value theory and applications. *Coll. Math. Soc. J. Bolyai* **36** 369–404.

DEHEUVELS, P. (1983). The strong approximation of extremal processes II. *Z. Wahrsch. Verw. Gebiete* **62** 7–15.

DWASS, M. (1964). Extremal processes. *Ann. Math. Statist.* **35** 1718–1725.

EMBRECHTS, P., KLÜPPELBERG, C. and MIKOSCH, T. (1997). *Modelling Extremal Events*. Springer, Berlin.

FAHRNER, I. (2000a). An extension of the almost sure max-limit theorem. *Statist. Probab. Lett.* **49** 93–103.

FAHRNER, I. (2000b). Almost sure versions of weak limit theorems. Thesis, Universität Ulm; Berichte aus der Mathematik, Shaker, Aachen.

FAHRNER, I. (2001). A strong invariance principle for the logarithmic average of sample maxima. *Stochastics Process Appl.* **93** 317–337.

FAHRNER, I. and STADTMÜLLER, U. (1998). On almost sure max-limit theorems. *Statist. Probab. Lett.* **37** 229–236.

FELLER, W. (1968). *An Introduction to Probability Theory and Its Applications* **1**, 3rd ed. Wiley, New York.

GALAMBOS, J. (1987). *The Asymptotic Theory of Extreme Order Statistics*, 2nd ed. Krieger, Melbourne, FL.

LAMPERTI, J. (1964). On extreme order statistics. *Ann. Math. Statist.* **35** 1726–1737.

LEADBETTER, M. R., LINDGREN, G. and ROOTZEN, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*. Springer, Berlin.

REISS, R. D. (1989). *Approximate Distributions of Order Statistics*. Springer, New York.
RESNICK, S. I. (1974). Inverses of extremal processes. *Adv. in Appl. Probab.* **6** 392–406.
RESNICK, S. I. (1987). *Extreme Values, Regular Variation, and Point Processes*. Springer, New York.
STADTMÜLLER, U. (1986). Asymptotic properties of nonparametric curve estimates. *Period. Math. Hungar.* **17** 83–108.

ABTEILUNG FÜR
ZAHLEN- U. WAHRSCHEINLICHKEITSTHEORIE
UNIVERSITÄT ULM
89069 ULM
GERMANY
E-MAIL: {fahrner, stamue}@mathematik.uni-ulm.de