Functionals of Clusters of Extremes

Segers, J.J.J.

Publication date:
2003

Citation for published version (APA):
Segers, J. J. J. (2003). *Functionals of Clusters of Extremes*. (CentER Discussion Paper; Vol. 2003-48). Econometrics.

No. 2003–48

FUNCTIONALS OF CLUSTERS OF EXTREMES

By Johan Segers

April 2003

ISSN 0924-7815

Applied Probability Trust (24 April 2003)

FUNCTIONALS OF CLUSTERS OF EXTREMES

JOHAN SEGERS,* *Tilburg University and Catholic University Leuven*

Abstract

For arbitrary stationary sequences of random variables satisfying a mild mixing condition, distributional approximations are established for functionals of clusters of exceedances over a high threshold. The approximations are in terms of the distribution of the process conditionally on the event that the first variable exceeds the threshold. This conditional distribution is shown to converge to a non-trivial limit if the finite-dimensional distributions of the process are in the domain of attraction of a multivariate extreme-value distribution. In this case, therefore, limit distributions are obtained for functionals of clusters of extremes, thereby generalizing results for higher-order stationary Markov chains by S. Yun (2000), *J. Appl. Probab.* **37**, 29–44.

Keywords: cluster functional; extremal index; extreme-value distribution; stationary sequence; tail sequence; threshold exceedance

AMS 2000 Subject Classification: Primary 60G70

Secondary 62G32

1. Introduction

Let $\{X_n\}_{n \geq 1}$ be a stationary sequence of random variables and let the thresholds $u_n \in \mathbb{R}$ be such that $\Pr(X_1 > u_n) > 0$, $n \geq 1$. The assumptions later on in the paper will imply $\lim_{n\to\infty} \Pr(X_1 > u_n) = 0$. For integer $r_n \geq 1$ the exceedances over u_n among the X_i, $i = 1, \ldots, r_n$, if any, are thought of as a single cluster. We will study statistics $c\{(X_i - u_n)_{i=1}^{r_n}\}$ that are functions of such a cluster as follows.

Definition 1. A measurable map $c : \mathbb{R} \cup \mathbb{R}^2 \cup \mathbb{R}^3 \cup \cdots \to \mathbb{R}$ is called a *cluster functional* if for all integers $1 \leq j \leq k \leq r$ and for all $(x_1, \ldots, x_r)$ such that $x_i \leq 0$ whenever $i = 1, \ldots, j-1$ or $i = k+1, \ldots, r$ we have $c(x_1, \ldots, x_r) = c(x_j, \ldots, x_k)$.

* Postal address: Tilburg University, P.O. Box 90153, NL-5000 LE Tilburg, the Netherlands (E-mail: *jsegers@uvt.nl*).

Typical examples are the number of threshold exceedances, the sum of all excesses, the number of up-crossings over a high level, the cluster maximum, etc. (Section 3).

We will study the asymptotic distribution of $c\{(X_i - u_n)_{i=1}^{r_n}\}$ conditionally on $\max_{i=1,\ldots,r_n} X_i > u_n$. For (higher-order) Markov chains this asymptotic distribution was described by Yun (2000) in terms of the process that arises as the weak limit of $(X_i - u_n)_{i\geq 1}$ conditionally on $X_1 > u_n$, the so-called tail chain. Well-known is the representation of this tail chain in terms of a certain random walk (Smith, 1992; Perfekt, 1994; Yun, 1998).

In Section 2, Yun's result is extended to arbitrary stationary sequences satisfying a weak mixing condition. Moreover, the distribution of $(X_i - u_n)_{i=1}^{m}$ conditionally on $X_1 > u_n$ is shown in Section 4 to converge for all $m = 1, 2, \ldots$ to a non-trivial limit in case the finite-dimensional distributions of the process are in the domain of attraction of a (multivariate) extreme-value distribution. Hence for such processes we can derive the weak limit behavior of functionals of clusters of extremes.

All proofs are deferred to Section 5. For integers $0 \leq j \leq k$ denote $M_{j,k} = \max\{X_{j+1}, \ldots, X_k\}$ (with $\max\emptyset = -\infty$) and $M_k = M_{0,k}$. Unless mentioned otherwise, asymptotic statements are for $n \to \infty$.

2. Approximate distribution of cluster statistics

Throughout, we work with a stationary process $\{X_n\}$ and thresholds $u_n \in \mathbb{R}$.

Theorem 1. *If the integers $1 \leq s_n \leq r_n$ are such that $s_n = o(r_n)$ and*

$$\Pr(M_{s_n,r_n} > u_n \mid X_1 > u_n) = o\left\{\frac{\Pr(M_{s_n} > u_n)}{s_n \Pr(X_1 > u_n)}\right\}, \tag{1}$$

then for every sequence of integers t_n such that $s_n \leq t_n \leq r_n$ and any uniformly bounded sequence $\{c_n\}$ of cluster functionals

$$\begin{aligned} &\mathrm{E}\left[c_n\{(X_i - u_n)_{i=1}^{r_n}\} \mid M_{r_n} > u_n\right] \\ &= \theta_n^{-1}\mathrm{E}\left[c_n\{(X_i - u_n)_{i=1}^{t_n}\} - c_n\{(X_i - u_n)_{i=2}^{t_n}\}\mathbf{1}(M_{1,t_n} > u_n) \,\middle|\, X_1 > u_n\right] \\ &\quad + o(1), \end{aligned} \tag{2}$$

with

$$\theta_n := \frac{\Pr(M_{r_n} > u_n)}{r_n \Pr(X_1 > u_n)} \sim \Pr(M_{1,t_n} \leq u_n \mid X_1 > u_n). \tag{3}$$

If θ_n converges to some $\theta \in [0,1]$, then under additional long-range dependence restrictions this θ is the *extremal index* of the process (Leadbetter, 1983; O'Brien, 1987).

Corollary 1. *Let $\{c_n\}$ be a sequence of (not necessarily bounded) cluster functionals. For integer $1 \le s_n \le t_n \le r_n$ as in Theorem 1 and for measurable $A_n \subset \mathbb{R}$ we have*

$$\begin{aligned}
&\Pr\left(c_n\{(X_i - u_n)_{i=1}^{r_n}\} \in A_n \mid M_{r_n} > u_n \right) \\
&= \theta_n^{-1} \big[\Pr\big(c_n\{(X_i - u_n)_{i=1}^{t_n}\} \in A_n \,\big|\, X_1 > u_n \big) \\
&\qquad - \Pr\big(c_n\{(X_i - u_n)_{i=2}^{t_n}\} \in A_n,\, M_{1,t_n} > u_n \,\big|\, X_1 > u_n \big) \big] + o(1). \qquad (4)
\end{aligned}$$

If also $c_n \ge 0$ and $c_n(x_1, \ldots, x_r) \ge c_n(x_2, \ldots, x_r)$ for all integer $n \ge 1$ and $r \ge 2$ and all $(x_1, \ldots, x_r) \in \mathbb{R}^r$, then

$$\begin{aligned}
&\Pr\left(c_n\{(X_i - u_n)_{i=1}^{r_n}\} > y_n \mid M_{r_n} > u_n \right) \\
&= \theta_n^{-1} \Pr\big(c_n\{(X_i - u_n)_{i=1}^{t_n}\} > y_n \ge c_n\{(X_i - u_n)_{i=2}^{t_n}\} \mid X_1 > u_n \big) + o(1) \qquad (5)
\end{aligned}$$

for arbitrary $0 \le y_n < \infty$.

Applications of Corollary 1 to particular cluster functionals are given in Section 3. In the remainder of the present section we mention some sufficient conditions for the restriction (1) on the long-range dependence in the process.

Corollary 2. *Let the integers r_n tend to infinity. If*

$$\lim_{m\to\infty} \limsup_{n\to\infty} \Pr(M_{m,r_n} > u_n \mid X_1 > u_n) = 0, \qquad (6)$$

then (1–3) hold for any sequence $s_n = 1, \ldots, r_n$ with $s_n \to \infty$, and $\liminf_{n\to\infty} \theta_n > 0$.

Observe that (6) is implied by

$$\lim_{m\to\infty} \limsup_{n\to\infty} \sum_{i=m+1}^{r_n} \Pr(X_i > u_n \mid X_1 > u_n) = 0.$$

Corollary 3. *For integers $1 \le s_n \le r_n$ define*

$$\Delta_n = \left| \Pr(X_1 > u_n,\, M_{s_n,r_n} > u_n) - \Pr(X_1 > u_n) \Pr(M_{s_n,r_n} > u_n) \right|.$$

If $r_n \Pr(X_1 > u_n) = o(1)$ and $\Delta_n = o\{s_n^{-1} \Pr(M_{s_n} > u_n)\}$, then (1) holds, and hence, if additionally $s_n = o(r_n)$, also the conclusions of Theorem 1.

The conditions of Corollary 3 are satisfied, for instance, if $\Delta_n = O(\rho^{s_n})$ for some $0 < \rho < 1$ and $\log(1/p_n) \ll s_n \ll r_n \ll 1/p_n$ with $p_n = \Pr(X_1 > u_n)$. Hence, Theorem 1 is readily applicable for geometrically mixing processes.

3. Examples of cluster statistics

All examples in this section are based on equation (5).

Example 1. If $c_n(x_1, \ldots, x_r) = \sum_{i=1}^r \phi_n(x_i)$ with $\phi_n : \mathbb{R} \to [0, \infty)$ measurable and $\phi_n(x) = 0$ for $x \leq 0$, then (5) becomes

$$\begin{aligned} &\Pr\left(\sum_{i=1}^{r_n} \phi_n(X_i - u_n) > y_n \,\middle|\, M_{r_n} > u_n \right) \\ &\quad = \theta_n^{-1} \Pr\left(\sum_{i=1}^{t_n} \phi_n(X_i - u_n) > y_n \geq \sum_{i=2}^{t_n} \phi_n(X_i - u_n) \,\middle|\, X_1 > u_n \right) + o(1), \end{aligned}$$

for arbitrary $y_n \geq 0$. Cluster statistics of this kind are considered in Rootzén, Leadbetter and de Haan (1998). In case $\phi_n = \mathbf{1}_{A_n}$ is the indicator function of a measurable set $A_n \subset (0, \infty)$, we obtain for arbitrary $k_n = 1, 2, \ldots$

$$\begin{aligned} &\Pr\left(\sum_{i=1}^{r_n} \mathbf{1}(X_i \in u_n + A_n) \geq k_n \,\middle|\, M_{r_n} > u_n \right) \\ &\quad = \theta_n^{-1} \Pr\left(X_1 \in u_n + A_n, \sum_{i=2}^{t_n} \mathbf{1}(X_i \in u_n + A_n) = k_n - 1 \,\middle|\, X_1 > u_n \right) + o(1). \end{aligned}$$

Example 2. For integer $0 \leq i < j$ and $1 \leq k \leq j - i$ denote the kth largest value of $X_{i+1}, \ldots, X_j$ by $M_{i,j}^{(k)}$ (in particular $M_{i,j}^{(1)} = M_{i,j}$). For convenience also set $M_{i,j}^{(k)} = -\infty$ if $k > j - i$, and denote $M_j^{(k)} = M_{0,j}^{(k)}$. Take $A_n = (x_n, \infty)$ with $0 \leq x_n < \infty$ in Example 1 to obtain for arbitrary $k_n = 1, 2, \ldots$

$$\begin{aligned} &\Pr\left(M_{r_n}^{(k_n)} > u_n + x_n \,\middle|\, M_{r_n} > u_n \right) \\ &\quad = \theta_n^{-1} \Pr\left(X_1 > u_n + x_n, \sum_{i=2}^{t_n} \mathbf{1}(X_i > u_n + x_n) = k_n - 1 \,\middle|\, X_1 > u_n \right) + o(1). \end{aligned}$$

The choice $k_n = 1$ leads to the asymptotic distribution of the cluster maximum:

$$\begin{aligned} &\Pr(M_{r_n} > u_n + x_n \mid M_{r_n} > u_n) \\ &\quad = \theta_n^{-1} \Pr(X_1 > u_n + x_n, M_{1,t_n} \leq u_n + x_n \mid X_1 > u_n) + o(1) \\ &\quad = \frac{\Pr(M_{1,r_n} \leq u_n + x_n \mid X_1 > u_n + x_n)}{\Pr(M_{1,r_n} \leq u_n \mid X_1 > u_n)} \Pr(X_1 > u_n + x_n \mid X_1 > u_n) + o(1), \end{aligned}$$

the latter equality in case $\Pr(X_1 > u_n + x_n) > 0$. In typical cases the ratio on the right-hand side of this equation tends to one, so that the cluster maximum has asymptotically the same distribution as a single threshold exceedance (Leadbetter, 1991).

Example 3. The choice $x_n = 0$ in Example 2 yields an approximation to the distribution of the cluster size: for arbitrary $k_n = 1, 2, \ldots$, we get

$$\begin{aligned}
&\Pr\left(\sum_{i=1}^{r_n} \mathbf{1}(X_i > u_n) \geq k_n \,\middle|\, M_{r_n} > u_n \right) \\
&= \theta_n^{-1} \Pr\left(\sum_{i=2}^{t_n} \mathbf{1}(X_i > u_n) = k_n - 1 \,\middle|\, X_1 > u_n \right) + o(1).
\end{aligned}$$

Observe that $k_n = 1$ yields (3) again, and that $\theta_n^{-1} = \mathrm{E}\left[\sum_{i=1}^{r_n} \mathbf{1}(X_i > u_n) \,|\, M_{r_n} > u_n\right]$ by definition.

Example 4. Generalizing Example 1, let $m = 2, 3, \ldots$ and let $\phi_n : \mathbb{R}^m \to [0, \infty)$ be a measurable function such that $\phi_n(x_1, \ldots, x_m) = 0$ if $\max_{i=1,\ldots,m} x_i \leq 0$. Define

$$\begin{aligned}
c_n(x_1, \ldots, x_r) &= \sum_{i=-m+2}^{r} \phi_n(x_i, \ldots, x_i + m - 1), \qquad (7) \\
&\quad \text{where } x_{-m+2} = \ldots = x_0 = 0 = x_{r+1} = \ldots = x_{r+m-1}.
\end{aligned}$$

Cluster functionals of this form are considered in Smith, Coles and Tawn (1997). Useful examples are generated through $\phi_n = \mathbf{1}_{A_n}$ for measurable $A_n \subset \mathbb{R}^m \setminus (-\infty, 0]^m$: (i) the number of up-crossings over the level $u_n + x_n$, with $0 \leq x_n < \infty$, is obtained by $m = 2$ and $A_n = (-\infty, x_n] \times (x_n, \infty)$; (ii) the number of local maxima exceeding $u_n + x_n$ is obtained by $m = 3$ and A_n the set of all (z_1, z_2, z_3) with $z_2 > \max(z_1, z_3, x_n)$; (iii) the number of times there are m consecutive exceedances over the level $u_n + x_n$, counting overlaps, is obtained by $A_n = (x_n, \infty)^m$; and so on.

Clearly, the c_n of (7) satisfies the conditions leading to (5). The latter can even be somewhat simplified because

$$\Pr(M_{m-1} > u_n \text{ or } M_{r-m+1,r} > u_n \mid M_{r_n > u_n}) \leq \frac{2(m-1)\Pr(X_1 > u_n)}{\Pr(M_{r_n} > u_n)} \to 0$$

and $\Pr(M_{t_n - m, t_n} > u_n \mid X_1 > u_n) = o(\theta_n)$ provided $s_n + m - 1 \leq t_n \leq r_n$ [use (1) and (19) below to prove these statements]. For instance, for the number of up-crossings over the level $v_n \geq u_n$ we have for $t_n = s_n + 1, \dots, r_n$, and $k_n = 1, 2, \dots$

$$\begin{aligned}
&\Pr\left(\sum_{i=1}^{r_n - 1} \mathbf{1}(X_i \leq v_n, X_{i+1} > v_n) \geq k_n \,\middle|\, M_{r_n} > u_n \right) \\
&= \theta_n^{-1} \Pr\left(X_1 > v_n,\, X_2 \leq v_n,\, \sum_{i=2}^{t_n - 1} \mathbf{1}(X_i \leq v_n, X_{i+1} > v_n) = k_n - 1 \,\middle|\, X_1 > u_n \right) \\
&\quad + o(1).
\end{aligned}$$

Example 5. The duration of a cluster corresponds to the functional

$$c(x_1, \dots, x_r) = \max\{j - i + 1 : 1 \leq i \leq j \leq r,\, x_i > 0,\, x_j > 0\}$$

(with $\max \emptyset = 0$). From (5) we obtain for arbitrary $k_n = 0, 1, 2, \dots$

$$\begin{aligned}
&\Pr(\exists 1 \leq i \leq j \leq r_n : X_i > u_n,\, X_j > u_n,\, j - i \geq k_n \mid M_{r_n} > u_n) \\
&= \theta_n^{-1} \Pr(\exists j = k_n + 1, \dots, t_n : M_{1, j-k_n} \leq u_n,\, X_j > u_n,\, M_{j, t_n} \leq u_n \mid X_1 > u_n) \\
&\quad + o(1).
\end{aligned}$$

4. Tail sequence of a stationary sequence

Let $\{X_n\}$ be a stationary sequence of random variables. For $m = 1, 2, \dots$ denote the distribution function of $(X_1, \dots, X_m)$ by F_m, and put $F = F_1$. We will be interested in the case that every F_m is in the domain of attraction of an m-dimensional extreme-value distribution. This assumption is, by work of Pickands (1975) and Marshall and Olkin (1983), equivalent to the existence of $\gamma \in \mathbb{R}$ and a positive function g defined in a left neighborhood of $x_+ = \sup\{x \in \mathbb{R} : F(x) < 1\}$ such that the following two conditions hold:

(i) We have

$$\lim_{u \uparrow x_+} \frac{1 - F(u + g(u)x)}{1 - F(u)} = \begin{cases} (1 + \gamma x)_+^{-1/\gamma} & \text{if } \gamma \neq 0 \\ \mathrm{e}^{-x} & \text{if } \gamma = 0 \end{cases} \tag{8}$$

locally uniformly in $\alpha < x < \infty$, where

$$\alpha = \begin{cases} -1/\gamma & \text{if } \gamma > 0 \\ -\infty & \text{if } \gamma \leq 0. \end{cases}$$

(ii) For every $m = 2, 3, \ldots$ and all $(x_1, \ldots, x_m) \in (\alpha, \infty)^m$

$$\lim_{u \uparrow x_+} \frac{1 - F_m(u + g(u)x_1, \ldots, u + g(u)x_m)}{1 - F(u)}$$

exists and is finite.

In this case, we have in fact for all $m = 1, 2, \ldots$ and all $(x_1, \ldots, x_m) \in (\alpha, \infty)^m$

$$\lim_{u \uparrow x_+} \frac{1 - F_m(u + g(u)x_1, \ldots, u + g(u)x_m)}{1 - F(u)} = -\log G_m(x_1, \ldots, x_m), \tag{9}$$

where the m-dimensional extreme-value distribution G_m is given by

$$G_m(x_1, \ldots, x_m) = \lim_{n \to \infty} F_m^n(u_n + g(u_n)x_1, \ldots, u_n + g(u_n)x_m) \tag{10}$$

for $(x_1, \ldots, x_m) \in \mathbb{R}^m$, with u_n being such that $1 - F(u_n) \sim 1/n$. Observe that α is the left end-point of G_1. De Haan (1984) established the following representation of the distributions G_m: there exist measurable functions $f_i : [0, 1] \to [0, \infty)$ $(i = 1, 2, \ldots)$, called *spectral functions*, such that $\int_0^1 f_i(u)\mathrm{d}u = 1$, $i = 1, 2, \ldots$, and

$$G_m(x_1, \ldots, x_m) = \exp\left\{ - \int_0^1 \max_{i=1,\ldots,m} \frac{f_i(u)}{z_i} \mathrm{d}u \right\} \tag{11}$$

for all $m = 1, 2, \ldots$ and $(x_1, \ldots, x_m) \in (\alpha, \infty)^m$, where

$$z_i = \begin{cases} (1 + \gamma x_i)_+^{1/\gamma} & \text{if } \gamma \neq 0 \\ \mathrm{e}^{x_i} & \text{if } \gamma = 0 \end{cases} \tag{12}$$

for $i = 1, \ldots, m$. [This shows in particular that the right-hand side of (9) is indeed finite.] The fact that the sequence $\{X_n\}$ is stationary leads to a certain property of the family of spectral functions which is explored in de Haan and Pickands (1986) but which will not be needed further on.

The interpretation of (8) is that the distribution of the scaled excess $(X_1 - u)/g(u)$ conditionally on $X_1 > u$ converges as $u \uparrow x_+$ to a generalized Pareto distribution. Equation (9) has a similar interpretation for the distribution of $\{(X_i - u)/g(u)\}_{i=1}^m$ conditionally on $X_1 > u$.

Theorem 2. *Let $\{X_n\}$ be a stationary sequence of random variables. If the marginal distribution function F satisfies (8), then (a) and (b) are equivalent:*

(a) For every $m = 1, 2, \ldots$ the distribution function F_m of $(X_1, \ldots, X_m)$ belongs to the domain of attraction of an extreme-value distribution.

(b) There exists a sequence of random variables $\{Y_m\}_{m\geq 1}$ such that for every $m = 1, 2, \ldots$ we have as $u \uparrow x_+$

$$\Pr\left\{ \left(\frac{X_1 - u}{g(u)} \vee \alpha, \ldots, \frac{X_m - u}{g(u)} \vee \alpha \right) \in \cdot \,\middle|\, X_1 > u \right\} \xrightarrow{w} \Pr\{(Y_1, \ldots, Y_m) \in \cdot\}. \quad (13)$$

In case (a–b) hold, we have for $m = 1, 2, \ldots$, $0 \leq x_1 < \infty$, and (in case $m \geq 2$) $\alpha < x_i < \infty$, $i = 2, \ldots, m$,

$$\begin{aligned}\Pr(Y_1 \leq x_1, \ldots, Y_m \leq x_m) &= \log \frac{G_m(x_1, x_2, \ldots, x_m)}{G_m(0, x_2, \ldots, x_m)} \\ &= \int_0^1 \left[f_1(u) - \max_{i=1,\ldots,m} \frac{f_i(u)}{z_i} \right]_+ \mathrm{d}u\end{aligned}$$

with G_m, f_i and z_i as in (9), (11) and (12) respectively.

We call the process $\{Y_m\}$ of Theorem 2(b) the *tail sequence* of $\{X_n\}$. In the special case $\gamma > 0$ we have $x_+ = \infty$, $g(u) \sim \gamma u$, and $\alpha = -1/\gamma$, so that (13) is equivalent to

$$\Pr\{ u^{-1}(X_1 \vee 0, \ldots, X_m \vee 0) \in \cdot \mid X_1 > u\} \xrightarrow{w} \Pr\{(Z_1, \ldots, Z_m) \in \cdot\}, \qquad u \to \infty,$$

where $Z_i = 1 + \gamma Y_i$ for $i \geq 1$. In case $\gamma < 0$ we have $x_+ < \infty$, $g(u) \sim (x_+ - u)|\gamma|$, and $\alpha = -\infty$, so that (13) is equivalent to

$$\Pr\left\{ \left(\frac{x_+ - u}{x_+ - X_1}, \ldots, \frac{x_+ - u}{x_+ - X_m} \right) \in \cdot \,\middle|\, X_1 > u \right\} \xrightarrow{w} \Pr\{(Z_1, \ldots, Z_m) \in \cdot\}, \qquad u \uparrow x_+,$$

where now $Z_i = 1/(1 + \gamma Y_i)$ for $i \geq 1$. In both cases $\Pr(Z_1 > z) = z^{-1/|\gamma|}$ for $1 \leq z < \infty$, while the Z_i for $i \geq 2$ are non-negative random variables, possibly with positive mass at 0. In case $\gamma = 0$ no simplification of (13) seems possible except that $\alpha = -\infty$ can be omitted. Finally, note that if $\gamma \leq 0$ the Y_i for $i \geq 2$ may attach positive mass at $-\infty$.

Example 6. Yun (1998), generalizing work of Smith (1992) and Perfekt (1994), studied extremes of k-order ($k = 1, 2, \ldots$) stationary Markov chains. Let $\{X_n\}$ be such a chain. If its marginal distribution satisfies (8) and if the joint distribution of $(X_1, \ldots, X_{k+1})$ is in the differentiable domain of attraction of a multivariate extreme-value distribution, then under certain side conditions there exists a unit exponential random variable T and a stationary $(k-1)$-order stationary Markov chain $\xi_1, \xi_2, \ldots$, independent of T, such that for every $m = 1, 2, \ldots$

$$\Pr\left\{ \left(\frac{X_1 - u}{g(u)}, \ldots, \frac{X_m - u}{g(u)} \right) \in \cdot \,\middle|\, X_1 > u \right\} \xrightarrow{w} \Pr\{(\phi(S_0 + T), \ldots, \phi(S_{m-1} + T)) \in \cdot\}$$

as $u \uparrow x_+$, where $\phi(x) = \int_0^x \mathrm{e}^{\gamma y} \mathrm{d}y$ for $x \in \mathbb{R}$, $S_0 = 0$, and $S_n = \sum_{i=1}^n \xi_i$ for $n \geq 1$.

Denote by $\mathbb{A}$ the set of all sequences $\{x_i\}_{i \geq 1}$ such that $x_i > 0$ for only a finite number of i. For $\{x_i\}_{i \geq 1} \in \mathbb{A}$ and a cluster functional c put $c\left(\{x_i\}_{i \geq 1}\right) := c(x_1, \ldots, x_m)$ with m such that $x_i \leq 0$ for all $i > m$.

Theorem 3. *Let $\{X_n\}$ be a stationary sequence with tail chain $\{Y_m\}$ as in (13). Let $u_n \in \mathbb{R}$ with $\Pr(X_1 > u_n) > 0$, let the integers r_n tend to infinity, and assume (6).*

(i) We have $\lim_{m \to \infty} Y_m = \alpha$ almost surely and

$$\Pr\left(M_{1,r_n} \leq u_n \mid X_1 > u_n\right) \to \theta := \Pr\left(\max_{i \geq 2} Y_i \leq 0\right) > 0. \tag{14}$$

(ii) Let c be a cluster functional and let $D \subset \mathbb{A}$ be the discontinuity set of c. If $\Pr\{(Y_1, Y_2, \ldots) \in D\} = 0 = \Pr\{(Y_2, Y_3, \ldots) \in D\}$, then

$$\begin{aligned} &\Pr\left\{ c\left(\left\{\frac{X_i - u_n}{g(u_n)} \vee \alpha\right\}_{i=1}^{r_n}\right) \in \cdot \,\middle|\, M_{r_n} > u_n \right\} \\ &\quad \xrightarrow{w} \theta^{-1}[\Pr\{c(Y_1, Y_2, \ldots) \in \cdot\,\} - \Pr\{c(Y_2, Y_3, \ldots) \in \cdot\,,\ \max_{i \geq 2} Y_i > 0\}]. \end{aligned} \tag{15}$$

In case $\gamma \leq 0$ and if some of the Y_i put positive mass at $\alpha = -\infty$, we implicitly assume that c is defined on $\bigcup_{r \geq 1} (\mathbb{R} \cup \{-\infty\})^r$. If c is such that $c(x_1, \ldots, x_r) = c(x_1 \vee 0, \ldots, x_r \vee 0)$ for all $r = 1, 2, \ldots$ and all $(x_1, \ldots, x_r)$ (as is the case for all examples of Section 3), then the α on the left-hand side of (15) can be omitted.

5. Proofs

5.1. Proof of Theorem 1

We split the proof in a series of lemmas, some of which seem of independent interest. Throughout this section, we consider integers $1 \leq s_n \leq t_n \leq r_n$. Additional assumptions are introduced when needed. Write

$$C_n(j, k) = c_n\{(X_i - u_n)_{i=j}^k\}, \qquad \text{integer } n \geq 1 \text{ and } 1 \leq j \leq k. \tag{16}$$

Without loss of generality we assume that $|c_n| \leq 1$. For a random variable X and an event A we write $\mathrm{E}[X; A] = \mathrm{E}[X \mathbf{1}_A]$; in particular $\mathrm{E}[X | A] = \mathrm{E}[X; A] / \Pr(A)$ provided $\Pr(A) > 0$.

Lemma 1. *(i) If (1), then*

$$\Pr(M_{t_n,r_n} > u_n \mid X_1 > u_n) = o\left\{ \frac{\Pr(M_{t_n} > u_n)}{t_n \Pr(X_1 > u_n)} \right\}. \tag{17}$$

(ii) If additionally $s_n = o(t_n)$*, then*

$$\Pr(M_{t_n,r_n} > u_n \mid M_{t_n} > u_n) = \Pr(M_{r_n - t_n} > u_n \mid M_{r_n - t_n, r_n} > u_n) \to 0. \tag{18}$$

Proof. *(i)* Since $\Pr(M_{t_n,r_n} > u_n \mid X_1 > u_n) \le \Pr(M_{s_n,r_n} > u_n \mid X_1 > u_n)$, we need only show that

$$\liminf_{n\to\infty} \frac{t_n^{-1} \Pr(M_{t_n} > u_n)}{s_n^{-1} \Pr(M_{s_n} > u_n)} > 0. \tag{19}$$

If $r_n \le 2s_n$ this is obviously true, so assume $r_n > 2s_n$. By stationarity

$$\begin{aligned} &\Pr(M_{2s_n,r_n} > u_n \mid M_{s_n} > u_n) \\ &\quad\le \frac{\sum_{i=1}^{s_n} \Pr(X_i > u_n,\, M_{2s_n,r_n} > u_n)}{\Pr(M_{s_n} > u_n)} \\ &\quad\le \frac{s_n \Pr(X_1 > u)}{\Pr(M_{s_n} > u)} \Pr(M_{s_n+1,r_n} > u_n \mid X_1 > u_n) \to 0. \end{aligned} \tag{20}$$

Denoting $k_n = \lfloor (1 + t_n/s_n)/2 \rfloor$, we have

$$\begin{aligned} \Pr(M_{t_n} > u) &\ge \Pr\left(\bigcup_{i=1}^{k_n} \{ M_{2(i-1)s_n,(2i-1)s_n} > u_n \} \right) \\ &\ge \sum_{i=1}^{k_n} \Pr(M_{2(i-1)s_n,(2i-1)s_n} > u_n,\, M_{2is_n,r_n} \le u_n) \\ &\ge k_n \Pr(M_{s_n} > u_n) \Pr(M_{2s_n,r_n} \le u_n \mid M_{s_n} > u). \end{aligned} \tag{21}$$

Now (20) and (21) imply (19) and therefore (17).

(ii) By stationarity

$$\begin{aligned} &\Pr(M_{t_n} > u_n,\, M_{t_n,r_n} > u_n) \\ &\quad= \Pr(M_{t_n} > u_n) + \Pr(M_{r_n - t_n} > u_n) - \Pr(M_{r_n} > u_n) \\ &\quad= \Pr(M_{r_n - t_n} > u_n,\, M_{r_n - t_n, r_n} > u_n). \end{aligned}$$

Denote $l_n = \lfloor t_n/s_n \rfloor$. We have

$$\begin{aligned} &\Pr(M_{t_n} > u_n,\, M_{t_n,r_n} > u_n) \\ &\quad\le \sum_{i=1}^{l_n} \Pr(M_{(i-1)s_n,is_n} > u_n,\, M_{t_n,r_n} > u_n) + \Pr(M_{l_n s_n, t_n} > u_n,\, M_{t_n,r_n} > u_n) \\ &\quad\le (l_n - 1) \Pr(M_{s_n} > u_n,\, M_{2s_n,r_n} > u_n) + 2 \Pr(M_{s_n} > u_n). \end{aligned}$$

Together with (21), this implies

$$\Pr(M_{t_n,r_n} > u_n \mid M_{t_n} > u_n) = O\{\Pr(M_{2s_n,r_n} > u_n \mid M_{s_n} > u_n)\} + O(s_n/t_n).$$

The first term converges to zero by (20), the second by the assumption $s_n = o(t_n)$.

Lemma 2. *If $t_n = o(r_n)$ and (18), then*

$$\Pr(M_{r_n} > u_n) \sim (r_n/t_n)\Pr(M_{t_n} > u_n) \tag{22}$$

$$\mathrm{E}[C_n(1,r_n) \mid M_{r_n} > u_n] = \mathrm{E}[C_n(1,t_n) \mid M_{t_n} > u_n] + o(1). \tag{23}$$

Proof. Denote $k_n = \lfloor r_n/t_n \rfloor$. Partition the event $\{M_{r_n} > u_n\}$ according to the first block of size t_n in which u_n is exceeded: by Definition 1

$$\begin{aligned}&\mathrm{E}[C_n(1,r_n); M_{r_n} > u_n]\\ &= \sum_{i=1}^{k_n} \mathrm{E}[C_n((i-1)t_n, r_n); M_{(i-1)t_n} \le u_n,\, M_{(i-1)t_n, it_n} > u_n]\\ &\quad + \mathrm{E}[C_n(k_n t_n, r_n); M_{k_n t_n} \le u_n,\, M_{k_n t_n, r_n} > u_n].\end{aligned}$$

Since $t_n = o(r_n)$ we have $k_n \sim r_n/t_n$, so that by (18) and stationarity

$$\begin{aligned}&\mathrm{E}[C_n(1,r_n); M_{r_n} > u_n]\\ &= k_n \mathrm{E}[C_n(1,t_n); M_{t_n} > u_n] + O\{\Pr(M_{t_n} > u_n)\}\\ &\quad + O\{(r_n/t_n)\Pr(M_{r_n - t_n} > u_n,\, M_{r_n - t_n, r_n} > u_n)\}\\ &\quad + O\{(r_n/t_n)\Pr(M_{t_n} > u_n,\, M_{t_n, r_n} > u_n)\}\\ &= (r_n/t_n)\mathrm{E}[C_n(1,t_n); M_{t_n} > u_n] + o\{(r_n/t_n)\Pr(M_{t_n} > u_n)\}.\end{aligned} \tag{24}$$

Choose $c_n = 1$ in (24) to get (22). For general c_n, (24) and (22) imply (23).

Lemma 3. *If $2t_n - 1 \le r_n$ and (18), then*

$$\begin{aligned}&\mathrm{E}[C_n(1,t_n) \mid M_{t_n} > u_n]\\ &= \frac{t_n \Pr(X_1 > u_n)}{\Pr(M_{t_n} > u_n)} \Big\{ \mathrm{E}[C_n(1,t_n) \mid X_1 > u_n]\\ &\qquad - \mathrm{E}[C_n(t_n, 2t_n - 1); M_{t_n - 1} > u_n \mid X_{t_n} > u_n] \Big\} + o(1).\end{aligned}$$

Proof. Since by stationarity $\mathrm{E}[C_n(1,t_n); X_1 > u_n] = \mathrm{E}[C_n(t_n, 2t_n-1); X_{t_n} > u_n]$, it is sufficient to show

$$\begin{aligned} &\mathrm{E}[C_n(1,t_n); M_{t_n} > u_n] \\ &\quad = t_n \mathrm{E}[C_n(t_n, 2t_n-1); M_{t_n-1} \le u_n,\, X_{t_n} > u_n] + o\{\Pr(M_{t_n} > u_n)\}. \end{aligned} \tag{25}$$

By stationarity and (18), we have

$$\begin{aligned} &\mathrm{E}[C_n(1,t_n); M_{t_n} > u_n] \\ &= \mathrm{E}[C_n(t_n, 2t_n-1); M_{t_n-1,2t_n-1} > u_n] \\ &= \mathrm{E}[C_n(t_n, 2t_n-1); M_{t_n-1} \le u_n,\, M_{t_n-1,2t_n-1} > u_n,\, M_{2t_n-1,3t_n-2} \le u_n] \\ &\quad + O\{\Pr(M_{t_n-1} > u_n,\, M_{t_n-1,2t_n-1} > u_n)\} \\ &\quad + O\{\Pr(M_{t_n-1,2t_n-1} > u_n,\, M_{2t_n-1,3t_n-2} > u_n)\} \\ &= \mathrm{E}[C_n(t_n, 2t_n-1); M_{t_n-1} \le u_n,\, M_{t_n-1,2t_n-1} > u_n,\, M_{2t_n-1,3t_n-2} \le u_n] \\ &\quad + o\{\Pr(M_{t_n} > u_n)\}. \end{aligned} \tag{26}$$

Now split the event $\{M_{t_n-1,2t_n-1} > u_n\}$ according to the smallest $i = t_n, \ldots, 2t_n - 1$ such that $X_i > u_n$:

$$\begin{aligned} &\mathrm{E}[C_n(t_n, 2t_n-1); M_{t_n-1} \le u_n,\, M_{t_n-1,2t_n-1} > u_n,\, M_{2t_n-1,3t_n-2} \le u_n] \\ &\quad = \sum_{i=t_n}^{2t_n-1} \mathrm{E}[C_n(t_n, 2t_n-1); M_{i-1} \le u_n,\, X_i > u_n,\, M_{2t_n-1,3t_n-2} \le u_n] \\ &\quad = \sum_{i=t_n}^{2t_n-1} \mathrm{E}[C_n(i, i+t_n-1); M_{i-1} \le u_n,\, X_i > u_n,\, M_{2t_n-1,3t_n-2} \le u_n], \end{aligned} \tag{27}$$

the last equality by Definition 1. Now for $i = t_n, \ldots, 2t_n - 1$,

$$\begin{aligned} &\big|\mathrm{E}[C_n(i, i+t_n-1); M_{i-t_n,i-1} \le u_n,\, X_i > u_n] \\ &\quad - \mathrm{E}[C_n(i, i+t_n-1); M_{i-1} \le u_n,\, X_i > u_n,\, M_{2t_n-1,3t_n-2} \le u_n]\big| \end{aligned} \tag{28}$$

is bounded by

$$\begin{aligned} &\Pr(M_{t_n-1} > u_n,\, M_{t_n-1,i-1} \le u_n,\, X_i > u_n) \\ &\quad + \Pr(M_{t_n-1,i-1} \le u_n,\, X_i > u_n,\, M_{2t_n-1,3t_n-2} > u_n), \end{aligned}$$

the sum of which over $i = t_n, \ldots, 2t_n - 1$ is

$$\Pr(M_{t_n-1} > u_n,\, M_{t_n-1,2t_n-1} > u_n) + \Pr(M_{t_n-1,2t_n-1} > u_n,\, M_{2t_n-1,3t_n-2} > u_n),$$

which is $o\{\Pr(M_{t_n} > u_n)\}$. Since for $i \geq t_n$

$$\begin{aligned} & \mathrm{E}[C_n(i, i+t_n-1); M_{i-t_n, i-1} \leq u_n,\, X_i > u_n] \\ = \;& \mathrm{E}[C_n(t_n, 2t_n-1); M_{t_n-1} \leq u_n,\, X_{t_n} > u_n] \end{aligned}$$

by stationarity, equation (25) follows from (26), (27), and the bound on (28). This proves the Lemma.

Lemma 4.

$$\begin{aligned} & \left|\mathrm{E}[C_n(t_n, 2t_n-1); M_{t_n-1} > u_n \mid X_{t_n} > u_n] - \mathrm{E}[C_n(2, t_n); M_{1,t_n} > u_n \mid X_1 > u_n]\right| \\ \leq \;& 2\Pr(M_{t_n, 2t_n-1} > u_n \mid X_1 > u_n). \end{aligned}$$

Proof. Decompose the event $\{M_{t_n-1} > u_n\}$ according to the largest $k = 1, \ldots, t_n - 1$ such that $X_k > u_n$: by Definition 1

$$\begin{aligned} & \mathrm{E}[C_n(t_n, 2t_n-1); M_{t_n-1} > u_n,\, X_{t_n} > u_n] \\ = & \sum_{k=1}^{t_n-1} \mathrm{E}[C_n(t_n, 2t_n-1); X_k > u_n,\, M_{k,t_n-1} \leq u_n,\, X_{t_n} > u_n] \\ = & \sum_{k=1}^{t_n-1} \mathrm{E}[C_n(k+1, 2t_n-1); X_k > u_n,\, M_{k,t_n-1} \leq u_n,\, X_{t_n} > u_n]. \end{aligned}$$

Stationarity implies

$$\begin{aligned} & \mathrm{E}[C_n(t_n, 2t_n-1); M_{t_n-1} > u_n,\, X_{t_n} > u_n] \\ = & \sum_{k=1}^{t_n-1} \mathrm{E}[C_n(2, 2t_n-k); X_1 > u_n,\, M_{1,t_n-k} \leq u_n,\, X_{t_n-k+1} > u_n] \\ = & \sum_{l=1}^{t_n-1} \mathrm{E}[C_n(2, t_n+l); X_1 > u_n,\, M_{1,l} \leq u_n,\, X_{l+1} > u_n], \end{aligned}$$

with $l = t_n - k$. On the other hand,

$$\begin{aligned} & \mathrm{E}[C_n(2, t_n); X_1 > u_n,\, M_{1,t_n} > u_n] \\ = & \sum_{l=1}^{t_n-1} \mathrm{E}[C_n(2, t_n); X_1 > u_n,\, M_{1,l} \leq u_n,\, X_{l+1} > u_n]. \end{aligned}$$

Since $C_n(2, t_n + l)$ and $C_n(2, t_n)$ are equal unless $M_{t_n, t_n+l} > u_n$, we have for $l = 1, \ldots, t_n - 1$

$$\begin{aligned} & \Big|\mathrm{E}[C_n(2, t_n+l); X_1 > u_n,\, M_{1,l} \leq u_n,\, X_{l+1} > u_n] \\ & - \;\mathrm{E}[C_n(2, t_n); X_1 > u_n,\, M_{1,l} \leq u_n,\, X_{l+1} > u_n]\Big| \end{aligned}$$

is bounded by $2\Pr(X_1 > u_n,\, M_{1,l} \le u_n,\, X_{l+1} > u_n,\, M_{t_n,t_n+l} > u_n)$. Therefore it remains to consider

$$\sum_{l=1}^{t_n-1} \Pr(X_1 > u_n,\, M_{1,l} \le u_n,\, X_{l+1} > u_n,\, M_{t_n,t_n+l} > u_n).$$

But the latter is clearly bounded by $\Pr(X_1 > u_n,\, M_{1,t_n} > u_n,\, M_{t_n,2t_n-1} > u_n)$.

Proof of Theorem 1. Choose t_n such that $s_n = o(t_n)$ and $t_n = o(r_n)$; take for instance $t_n = \lfloor (s_n r_n)^{1/2} \rfloor$. By Lemma 1(ii), we have (18). Therefore, by Lemma 2

$$\mathrm{E}[C_n(1, r_n) \mid M_{r_n} > u_n] = \mathrm{E}[C_n(1, t_n) \mid M_{t_n} > u_n] + o(1).$$

and thus, by Lemma 3,

$$\begin{aligned} &\mathrm{E}[C_n(1, r_n) \mid M_{r_n} > u_n] \\ &= \frac{t_n \Pr(X_1 > u_n)}{\Pr(M_{t_n} > u_n)} \Big\{ \mathrm{E}[C_n(1, t_n) \mid X_1 > u_n] \\ &\qquad - \mathrm{E}[C_n(t_n, 2t_n - 1); M_{t_n-1} > u_n \mid X_{t_n} > u_n] \Big\} + o(1). \end{aligned}$$

Using Lemma 1(i) and Lemma 4 we get

$$\begin{aligned} &\mathrm{E}[C_n(1, r_n) \mid M_{r_n} > u_n] \\ &= \frac{t_n \Pr(X_1 > u_n)}{\Pr(M_{t_n} > u_n)} \Big\{ \mathrm{E}[C_n(1, t_n) \mid X_1 > u_n] \\ &\qquad - \mathrm{E}[C_n(2, t_n - 1); M_{1,t_n} > u_n \mid X_1 > u_n] \Big\} + o(1). \end{aligned}$$

By Lemma 2 we have $\theta_n \sim \Pr(M_{t_n} > u_n)/\{t_n \Pr(X_1 > u_n)\}$, whence (2) and (3). These remain true for general $t_n = s_n, \ldots, r_n$ since each of

$$|\Pr(M_{1,s_n} \le u_n \mid X_1 > u_n) - \Pr(M_{1,t_n} \le u_n \mid X_1 > u_n)|,$$

$$|\mathrm{E}[C_n(1, s_n) \mid X_1 > u_n] - \mathrm{E}[C_n(1, t_n) \mid X_1 > u_n]|,$$

$$|\mathrm{E}[C_n(2, s_n); M_{1,s_n} > u_n \mid X_1 > u_n] - \mathrm{E}[C_n(2, t_n); M_{1,t_n} > u_n \mid X_1 > u_n]|,$$

are

$$O\{\Pr(M_{s_n,r_n} > u_n \mid X_1 > u_n)\} = o\left\{ \frac{\Pr(M_{s_n} > u_n)}{s_n \Pr(X_1 > u_n)} \right\} = o(\theta_n)$$

by (19).

5.2. Proofs of Corollaries 1–3

Proof of Corollary 1. Apply Theorem 1 to the cluster functionals $\mathbf{1}\{c_n(\cdot\,) \in A_n\}$.

Proof of Corollary 2. It is enough to show that (1) holds for arbitrary $s_n = 1, \ldots, r_n$ with $s_n \to \infty$. Choose a positive integer m. If n is large enough such that $s_n \geq m$, then $\Pr(M_{s_n,r_n} > u_n \mid X_1 > u_n) \leq \Pr(M_{m,r_n} > u_n \mid X_1 > u_n)$. Hence

$$\lim_{n\to\infty} \Pr(M_{s_n,r_n} > u_n \mid X_1 > u_n) \leq \lim_{m\to\infty} \limsup_{n\to\infty} \Pr(M_{m,r_n} > u_n \mid X_1 > u_n) = 0. \tag{29}$$

Moreover, denoting $k_n = \lfloor s_n/m \rfloor$, we have

$$\begin{aligned} \Pr(M_{s_n} > u_n) &\geq \sum_{i=1}^{k_n} \Pr(X_{(i-1)m+1} > u_n,\, M_{im,r_n} \leq u_n) \\ &\geq k_n \Pr(X_1 > u_n,\, M_{m,r_n} \leq u_n), \end{aligned}$$

and thus

$$\liminf_{n\to\infty} \frac{\Pr(M_{s_n} > u_n)}{s_n \Pr(X_1 > u_n)} \geq m^{-1} \left(1 - \limsup_{n\to\infty} \Pr(M_{m,r_n} > u_n \mid X_1 > u_n) \right).$$

Since $\limsup_{n\to\infty} \Pr(M_{m,r_n} > u_n \mid X_1 > u_n) < 1$ for large enough m, we obtain

$$\liminf_{n\to\infty} \frac{\Pr(M_{s_n} > u_n)}{s_n \Pr(X_1 > u_n)} > 0. \tag{30}$$

Equations (29–30) imply (1). Moreover, (30) for $s_n = r_n$ implies $\liminf \theta_n > 0$.

Proof of Corollary 3. Clearly

$$\Pr(M_{s_n,r_n} > u_n \mid X_1 > u_n) \leq \Pr(M_{r_n} > u_n) + \frac{\Delta_n}{\Pr(X_1 > u_n)},$$

Since $\{1, \ldots, r_n\} \subset \bigcup_{i=1}^{\lceil r_n/s_n \rceil} \{(i-1)s_n + 1, \ldots, is_n\}$ and $r_n \Pr(X_1 > u_n) = o(1)$,

$$\begin{aligned} \Pr(M_{r_n} > u_n) &\leq \lceil r_n/s_n \rceil \Pr(M_{s_n} > u_n) \\ &\leq 2(r_n/s_n) \Pr(M_{s_n} > u_n) = o\left(\frac{\Pr(M_{s_n} > u_n)}{s_n \Pr(X_1 > u_n)} \right). \end{aligned}$$

The assumption on Δ_n implies

$$\frac{\Delta_n}{\Pr(X_1 > u_n)} = o\left(\frac{\Pr(M_{s_n} > u_n)}{s_n \Pr(X_1 > u_n)} \right)$$

as well.

5.3. Proof of Theorem 2

As a preliminary to the proof, recall the following facts: the function g is unique up to asymptotic equivalence; for every $x \in \mathbb{R}$ such that $1+\gamma x > 0$ we have $u+g(u)x < x_+$ for u sufficiently close to x_+ and $u + g(u)x \to x_+$; and

$$\lim_{u \uparrow x_+} \frac{g(u+g(u)x)}{g(u)} = 1 + \gamma x \tag{31}$$

locally uniformly in $\{x \in \mathbb{R} : 1 + \gamma x > 0\}$. These properties easily follow from the fact that for fixed $u < x_+$ the function $x \mapsto [1 - F(u + g(u)x)]/[1 - F(u)]$ is non-increasing while the limit function on the right-hand side of (8) is decreasing and continuous.

(a) implies (b). Let $m \geq 1$ be integer. From (a) we obtain that for all $(x_1, \ldots, x_m) \in \mathbb{R}^m$ with $x_1 > 0$ and $x_i > \alpha$, $i = 2, \ldots, m$,

$$\begin{aligned} &\Pr\left\{ \frac{X_1 - u}{g(u)} \leq x_1, \ldots, \frac{X_m - u}{g(u)} \leq x_m \,\middle|\, X_1 > u \right\} \\ &= \frac{F_m(u + g(u)x_1, \ldots, u + g(u)x_m) - F_m(u, u + g(u)x_2, \ldots, u + g(u)x_m)}{1 - F(u)} \\ &\to -\log G_m(0, x_2, \ldots, x_m) - \{-\log G_m(x_1, x_2, \ldots, x_m)\} \end{aligned} \tag{32}$$

as $u \uparrow x_+$. Hence (32) is the distribution function of a random vector $(Y_1^{(m)}, \ldots, Y_m^{(m)})$ for which

$$\Pr\left\{ \left(\frac{X_1 - u}{g(u)} \vee \alpha, \ldots, \frac{X_m - u}{g(u)} \vee \alpha \right) \in \cdot \,\middle|\, X_1 > u \right\} \xrightarrow{w} \Pr\{(Y_1^{(m)}, \ldots, Y_m^{(m)}) \in \cdot\}$$

as $u \uparrow x_+$. For integer $1 \leq m \leq p$, the vectors $(Y_1^{(m)}, \ldots, Y_m^{(m)})$ and $(Y_1^{(p)}, \ldots, Y_m^{(p)})$ necessarily have the same distribution. By Kolmogorov's existence theorem, there exists a random sequence $\{Y_n\}$ such that $(Y_1, \ldots, Y_m)$ has the same distribution as $(Y_1^{(m)}, \ldots, Y_m^{(m)})$ for every m. For this sequence, (b) holds.

(b) implies (a). First we show that $\Pr(Y_j = x) = 0$ for $\alpha < x \leq \infty$ and $j = 1, 2, \ldots$. For $j = 1$ this follows from the fact that Y_1 has a generalized Pareto distribution. Let $j \geq 2$ be integer. Observe first that if $0 < x < \infty$ is a continuity point of Y_j, then by stationarity

$$\begin{aligned} \Pr(Y_j > x) &= \lim_{u \uparrow x_+} \Pr\left\{ \frac{X_j - u}{g(u)} > x \,\middle|\, X_1 > u \right\} \\ &= \lim_{u \uparrow x_+} \Pr\{X_1 > u,\, X_j > u + g(u)x\}/[1 - F(u)] \\ &\leq \lim_{u \uparrow x_+} \Pr\{X_j > u + g(u)x\}/[1 - F(u)] = \Pr(Y_1 > x). \end{aligned}$$

In particular, the right end-point of Y_j is bounded by the right end-point of Y_1. Now if $\alpha < x \leq \infty$, then either $1+\gamma x > 0$ or x is equal to or larger than the right end-point of Y_1. In the latter case, $\Pr(Y_j = x) = 0$, while in the former, we can find for arbitrarily small $\varepsilon > 0$ a number $y \in \mathbb{R}$ such that $\alpha < y < x < (y+\varepsilon)/(1-\gamma\varepsilon)$, and both y and $(y+\varepsilon)/(1-\gamma\varepsilon)$ are continuity points of Y_j. Denoting $v = v(u,\varepsilon) = u - g(u)\varepsilon$ and noting that $g(v)/g(u) \to 1-\gamma\varepsilon$ as $u \uparrow x_+$, we see that

$$\begin{aligned}
\Pr(Y_j > y) &= \lim_{u \uparrow x_+} \Pr\left\{ \frac{X_j - u}{g(u)} > y \,\middle|\, X_1 > u \right\} \\
&= \lim_{u \uparrow x_+} \frac{1-F(v)}{1-F(u)} \frac{\Pr\left\{ X_1 > v + g(v)\frac{g(u)}{g(v)}\varepsilon,\, X_j > v + g(v)\frac{g(u)}{g(v)}(y+\varepsilon) \right\}}{1-F(v)} \\
&= (1-\gamma\varepsilon)^{-1/\gamma} \Pr\left(Y_1 > \frac{\varepsilon}{1-\gamma\varepsilon},\, Y_j > \frac{y+\varepsilon}{1-\gamma\varepsilon} \right).
\end{aligned}$$

Hence

$$\Pr(Y_j = x) \leq \Pr(Y_j > y) - \Pr\left(Y_j > \frac{y+\varepsilon}{1-\gamma\varepsilon} \right) \leq (1-\gamma\varepsilon)^{-1/\gamma} - 1.$$

Let $\varepsilon \downarrow 0$ to obtain that $\Pr(Y_j = x) = 0$.

Now let $m = 1, 2, \ldots$ and $(x_1, \ldots, x_m) \in (\alpha, \infty)^m$; we have to show convergence of (9) to a finite limit as $u \uparrow x_+$. We proceed by induction on m. If $m = 1$, we can simply invoke the assumption (8). So let $m \geq 2$. Observe that

$$\begin{aligned}
&1 - F_m(u + g(u)x_1, \ldots, u + g(u)x_m) \\
&\quad = \Pr\left\{ \frac{X_1 - u}{g(u)} > x_1, \frac{X_2 - u}{g(u)} \leq x_2, \ldots, \frac{X_m - u}{g(u)} \leq x_m \right\} \\
&\qquad + [1 - F_{m-1}(u + g(u)x_2, \ldots, u + g(u)x_m)], \qquad (33)
\end{aligned}$$

By (33) and the induction hypothesis it is sufficient to show that

$$\Pr\left\{ \frac{X_1 - u}{g(u)} > x_1, \frac{X_2 - u}{g(u)} \leq x_2, \ldots, \frac{X_m - u}{g(u)} \leq x_m \right\} \Big/ [1 - F(u)] \qquad (34)$$

converges to a finite limit for all $(x_1, \ldots, x_m) \in (\alpha, \infty)^m$. On the one hand, if $1+\gamma x_1 \leq 0$, then x_1 must be equal to or larger than the right end-point of Y_1, and thus (34) is bounded by $[1-F(u+g(u)x_1)]/[1-F(u)]$, which converges to 0 as $u \uparrow x_+$. On the other hand, suppose $1+\gamma x_1 > 0$. Denote $v = u+g(u)x_1$ and observe that $g(v)/g(u) \to 1+\gamma x_1$

as $u \uparrow x_+$ by (31). We have $(x_i - x_1)/(1 + \gamma x_1) > \alpha$ for $i = 2, \ldots, m$, and thus

$$\begin{aligned} &\Pr\left\{ \frac{X_1 - u}{g(u)} > x_1, \frac{X_2 - u}{g(u)} \le x_2, \ldots, \frac{X_m - u}{g(u)} \le x_m \right\} \Big/ [1 - F(u)] \\ &= \frac{1 - F(v)}{1 - F(u)} \Pr\left\{ \frac{X_i - v}{g(v)} \le \frac{g(u)}{g(v)}(x_i - x_1),\, i = 2, \ldots, m \,\Big|\, X_1 > v \right\} \\ &\to (1 + \gamma x_1)^{-1/\gamma} \Pr\left(Y_i \le \frac{x_i - x_1}{1 + \gamma x_1},\, i = 2, \ldots, m \right), \qquad u \uparrow x_+. \end{aligned}$$

5.4. Proof of Theorem 3

(i) By Corollary 2

$$\theta_n := \frac{\Pr(M_{r_n} > u_n)}{r_n \Pr(X_1 > u_n)} = \Pr(M_{1,r_n} \le u_n \mid X_1 > u_n) + o(1)$$

and $\liminf_{n\to\infty} \theta_n > 0$; in particular $u_n \to x_+$. We show first that $Y_m \to \alpha$ almost surely. As $Y_m \ge \alpha$ almost surely, it suffices to show that

$$\lim_{m\to\infty} \Pr\left(\max_{i>m} Y_i > x \right) = 0 \tag{35}$$

for $x > \alpha$. First we treat the case $x = 0$. For arbitrary integers $1 \le m < l$

$$\begin{aligned} \Pr\left(\max_{m<i\le l} Y_i > 0 \right) &= \lim_{n\to\infty} \Pr(M_{m,l} > u_n \mid X_1 > u_n) \\ &\le \limsup_{n\to\infty} \Pr(M_{m,r_n} > u_n \mid X_1 > u_n). \end{aligned}$$

Let $l \to \infty$ to obtain $\Pr\left(\max_{i>m} Y_i > 0\right) \le \limsup_{n\to\infty} \Pr(M_{m,r_n} > u_n \mid X_1 > u_n)$. By (6) we get (35) for $x = 0$. Second, let $\alpha < x < 0$. Clearly $1 + \gamma x > 0$. For arbitrary integers $1 \le m < l$

$$\begin{aligned} &\Pr\left(\max_{m<i\le l} Y_i > x \right) \\ &= \lim_{u\uparrow x_+} \Pr\{M_{m,l} > u + g(u)x \mid X_1 > u\} \\ &\le \lim_{u\uparrow x_+} \Pr\{M_{m,l} > u + g(u)x \mid X_1 > u + g(u)x\} \frac{1 - F(u + g(u)x)}{1 - F(u)} \\ &= \Pr\left(\max_{m<i\le l} Y_i > 0 \right) (1 + \gamma x)^{-1/\gamma}. \end{aligned}$$

Since (35) is true for $x = 0$, we now see it is in fact true for arbitrary $x > \alpha$.

To show (14), write for $m = 1, \dots, r_n$

$$\begin{aligned}
&\left|\Pr(M_{1,r_n} \le u_n \mid X_1 > u_n) - \Pr\left(\max_{i\ge 2} Y_i \le 0\right)\right| \\
&\quad \le \Pr(M_{m,r_n} > u_n \mid X_1 > u_n) + \Pr\left(\max_{i>m} Y_i > 0\right) \\
&\qquad + \left|\Pr(M_{1,m} \le u_n \mid X_1 > u_n) - \Pr\left(\max_{i=2,\dots,m} Y_i \le 0\right)\right|.
\end{aligned}$$

Let first $n \to \infty$ and subsequently $m \to \infty$ to arrive at the result.

(ii) Assume first that c is a bounded cluster functional and $c(x_1, \dots, x_r) = 0$ whenever $\max_{i=1,\dots,r} x_i \le 0$. Let $\varepsilon_{n,m}$ be a generic array such that

$$\lim_{m\to\infty} \limsup_{n\to\infty} |\varepsilon_{n,m}| = 0.$$

By Theorem 1 and part (i)

$$\begin{aligned}
&\mathrm{E}\left[c\left\{\left(\frac{X_i - u_n}{g(u_n)} \vee \alpha\right)_{i=1}^{r_n}\right\} \,\middle|\, M_{r_n} > u_n\right] \\
&= \theta^{-1}\mathrm{E}\left[c\left\{\left(\frac{X_i - u_n}{g(u_n)} \vee \alpha\right)_{i=1}^{m}\right\} - c\left\{\left(\frac{X_i - u_n}{g(u_n)} \vee \alpha\right)_{i=2}^{m}\right\} \,\middle|\, X_1 > u_n\right] + \varepsilon_{n,m}.
\end{aligned}$$

Let D_m be the set of all discontinuity points of c in $[-\infty, \infty)^m$. Since

$$\begin{aligned}
0 &= \Pr\{(Y_1, Y_2, \dots) \in D, \max_{i>m} Y_i \le 0\} = \Pr\{(Y_1, \dots, Y_m) \in D_m, \max_{i>m} Y_i \le 0\} \\
&= \Pr\{(Y_1, \dots, Y_m) \in D_m\} - \Pr\{(Y_1, \dots, Y_m) \in D_m, \max_{i>m} Y_i > 0\}
\end{aligned}$$

we have $\Pr\{(Y_1, \dots, Y_m) \in D_m\} \le \Pr(\max_{i>m} Y_i > 0) \to 0$ as $m \to \infty$; similarly $\Pr\{(Y_2, \dots, Y_m) \in D_{m-1}\} \to 0$ as $m \to \infty$. Hence

$$\begin{aligned}
&\mathrm{E}\left[c\left\{\left(\frac{X_i - u_n}{g(u_n)} \vee \alpha\right)_{i=1}^{r_n}\right\} \,\middle|\, M_{r_n} > u_n\right] \\
&= \theta^{-1}\mathrm{E}\left[c(Y_1, \dots, Y_m) - c(Y_2, \dots, Y_m)\right] + \varepsilon_{n,m} \\
&= \theta^{-1}\mathrm{E}\left[c(Y_1, Y_2, \dots) - c(Y_2, Y_3, \dots)\right] + \varepsilon_{n,m}.
\end{aligned}$$

Since the first term on the right-hand side does not depend on m, we obtain

$$\begin{aligned}
&\mathrm{E}\left[c\left\{\left(\frac{X_i - u_n}{g(u_n)} \vee \alpha\right)_{i=1}^{r_n}\right\} \,\middle|\, M_{r_n} > u_n\right] \\
&= \theta^{-1}\mathrm{E}\left[c(Y_1, Y_2, \dots) - c(Y_2, Y_3, \dots)\right] + o(1). \qquad (36)
\end{aligned}$$

Now let c be an arbitrary cluster functional. Let $f : \mathbb{R} \to \mathbb{R}$ be bounded and continuous and define $\tilde{c}$ by $\tilde{c}(x_1, \dots, x_r) = f(c(x_1, \dots, x_r))\mathbf{1}(\max_{i=1,\dots,r} x_i > 0)$ for

$r = 1, 2, \ldots$ and $(x_1, \ldots, x_r) \in \mathbb{R}^r$. Clearly, $\tilde{c}$ is a bounded cluster functional and its discontinuity set is contained in $D \cup \bigcup_{i=1}^{\infty} \{(x_j)_{j \geq 1} \in \mathbb{A} : x_i = 0\}$. Since $\Pr(Y_i = 0) = 0$ for all $i = 1, 2, \ldots$, we can apply (36) on $\tilde{c}$ to find

$$\begin{aligned} &\mathrm{E}\left[f\left(c\left\{ \left(\frac{X_i - u_n}{g(u_n)} \vee \alpha \right)_{i=1}^{r_n} \right\} \right) \,\middle|\, M_{r_n} > u_n \right] \\ &= \theta^{-1} \mathrm{E}\left[f(c(Y_1, Y_2, \ldots)) - f(c(Y_2, Y_3, \ldots)) \mathbf{1}\left(\max_{i \geq 2} Y_i > 0 \right) \right] + o(1). \end{aligned}$$

Acknowledgements

The author is postdoctoral research fellow of the Fund for Scientific Research, Flanders. Warmly appreciated are numerous helpful discussions with Chris Ferro.

References

de Haan, L. (1984). A spectral representation for max-stable processes. *Ann. Probab.* **12**, 1194–1204.

de Haan, L. and Pickands, J. (1986). Stationary min-stable stochastic processes. *Probab. Th. Rel. Fields* **72**, 477–492.

Leadbetter, M.R. (1983). Extremes and local dependence of stationary sequences. *Z. Wahrscheinlichkeitsth.* **65**, 291–306.

Leadbetter, M.R. (1991). On a basis for 'Peaks over Threshold' modeling. *Statist. Probab. Letters* **12**, 357–362.

Marshall, A.W. and Olkin, I. (1983). Domains of attraction of multivariate extreme value distributions. *Ann. Probab.* **11**, 168–177.

O'Brien, G.L. (1987). Extreme values for stationary and Markov sequences. *Ann. Probab.* **15**, 281–291.

Perfekt, R. (1994). Extremal behaviour of stationary Markov chains with applications. *Ann. Appl. Probab.* **4**, 529–548.

Pickands, J. (1975). Statistical inference using extreme order statistics. *Ann. Statist.* **3**, 119–131.

Rootzén, H., Leadbetter, M.R., and de Haan, L. (1998). On the distribution of tail array sums for strongly mixing stationary sequences. *Ann. Appl. Probab.* **8**, 868–885.

Smith, R.L. (1992). The extremal index for a Markov chain. *J. Appl. Probab.* **29**, 37–45.

Smith, R.L., Tawn, J.A., and Coles, S.G. (1997). Markov chain models for threshold exceedances. *Biometrika* **84**, 249–268.

Yun, S. (1998). The extremal index of a higher-order stationary Markov chain. *Ann. Appl. Probab.* **8**, 408–437.

Yun, S. (2000). The distributions of cluster functionals of extreme events in a dth-order Markov chain. *J. Appl. Probab.* **37**, 29–44.