Functionals of Clusters of Extremes
Segers, J.J.J.

Publication date:
2003

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
FUNCTIONALS OF CLUSTERS OF EXTREMES

By Johan Segers

April 2003

ISSN 0924-7815
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 \( \max_{i=1,...,r_n} X_i > u_n \). 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

\[
\Pr(M_{s_n,r_n} > u_n \mid X_1 > u_n) = o\left(\frac{\Pr(M_{s_n} > u_n)}{s_n \Pr(X_1 > u_n)}\right),
\]

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

\[
E[c_n\{(X_i - u_n)_{i=1}^{r_n}\} \mid M_{r_n} > u_n]
\]

\[
= \theta_n^{-1}E\left[c_n\{(X_i - u_n)_{i=1}^{r_n}\} - c_n\{(X_i - u_n)_{i=2}^{r_n}\}I(M_{1,t_n} > u_n) \mid X_1 > u_n\right]
\]

\[
+ o(1),
\]

with

\[
\theta_n := \frac{\Pr(M_{r_n} > u_n)}{r_n \Pr(X_1 > u_n)} \sim \Pr(M_{1,t_n} \leq u_n \mid X_1 > u_n).
\]
If \( \theta_n \) converges to some \( \theta \in [0,1] \), then under additional long-range dependence restrictions this \( \theta \) 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 \( 1 \leq s_n \leq t_n \leq r_n \) as in Theorem 1 and for measurable \( A_n \subset \mathbb{R} \) we have

\[
\Pr \left( c_n \{ (X_i - u_n)_{i=1}^{r_n} \} \in A_n \mid M_{r_n} > u_n \right) = \theta_n^{-1} \Pr \left( c_n \{ (X_i - u_n)_{i=1}^{t_n} \} \in A_n \mid X_1 > u_n \right) - \Pr \left( c_n \{ (X_i - u_n)_{i=2}^{t_n} \} \in A_n, M_{1:t_n} > u_n \mid X_1 > u_n \right) + o(1). \tag{4}
\]

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

\[
\Pr \left( c_n \{ (X_i - u_n)_{i=1}^{r_n} \} > y_n \mid M_{r_n} > u_n \right) = \theta_n^{-1} \Pr \left( c_n \{ (X_i - u_n)_{i=1}^{t_n} \} > y_n \geq c_n \{ (X_i - u_n)_{i=2}^{t_n} \} \mid X_1 > u_n \right) + o(1) \tag{5}
\]

for arbitrary \( 0 \leq y_n < \infty \).

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

\[
\lim_{m \to \infty} \limsup_{n \to \infty} \Pr(M_{m:r_n} > u_n \mid X_1 > u_n) = 0, \tag{6}
\]

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 \( 1 \leq s_n \leq r_n \) define

\[
\Delta_n = |\Pr(X_1 > u_n, M_{s_n:r_n} > u_n) - \Pr(X_1 > u_n) \Pr(M_{s_n:r_n} > u_n)|.
\]

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 $\Delta_n = O(p^{\alpha_n})$ 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

\[
\Pr\left( \sum_{i=1}^{r} \phi_n(X_i - u_n) > y_n \left| M_{r_n} > u_n \right. \right) = \theta_n^{-1} \Pr\left( \sum_{i=1}^{r} \phi_n(X_i - u_n) > y_n \geq \sum_{i=2}^{r} \phi_n(X_i - u_n) \left| X_1 > u_n \right. \right) + o(1),
\]

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

\[
\Pr\left( \sum_{i=1}^{r} 1(X_i \in u_n + A_n) \geq k_n \left| M_{r_n} > u_n \right. \right) = \theta_n^{-1} \Pr\left(X_1 \in u_n + A_n, \sum_{i=2}^{r} 1(X_i \in u_n + A_n) = k_n - 1 \left| X_1 > u_n \right. \right) + o(1).
\]

Example 2. For integer $0 \leq i < j$ and $1 \leq k \leq j - i$ denote the $k$th 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 $k_n = 1, 2, \ldots$

\[
\Pr\left(M_{i,j}^{(k_n)} > u_n + x_n \left| M_{r_n} > u_n \right. \right) = \theta_n^{-1} \Pr\left(X_1 > u_n + x_n, \sum_{i=2}^{r} 1(X_i > u_n + x_n) = k_n - 1 \left| X_1 > u_n \right. \right) + o(1).
\]

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

\[
\Pr(M_{r_n} > u_n + x_n \left| M_{r_n} > u_n \right. \right) = \theta_n^{-1} \Pr(X_1 > u_n + x_n, M_{1,t_n} \leq u_n + x_n \left| X_1 > u_n \right. \right) + o(1)
\]

\[
= \frac{\Pr(M_{1,r_n} \leq u_n + x_n \left| X_1 > u_n + x_n \right. \right)}{\Pr(M_{1,r_n} \leq u_n \left| X_1 > u_n \right. \right)} \Pr(X_1 > u_n + x_n \left| X_1 > u_n \right. \right) + o(1),
\]
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

\[
\Pr \left( \sum_{i=1}^{r_n} 1(X_i > u_n) \geq k_n \bigg| M_{r_n} > u_n \right) = \theta_n^{-1} \Pr \left( \sum_{i=2}^{r_n} 1(X_i > u_n) = k_n - 1 \bigg| X_1 > u_n \right) + o(1).
\]

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

\[
c_n(x_1, \ldots, x_r) = \sum_{i=-m+2}^{r} \phi_n(x_i, \ldots, x_i + m - 1),
\]

where \( x_{-m+2} = \ldots = x_0 = 0 = x_{r+1} = \ldots = x_{r+m-1} \).

Cluster functionals of this form are considered in Smith, Coles and Tawn (1997). Useful examples are generated through \( \phi_n = 1_{A_n} \) 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

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

and \( \Pr(M_{t_n-m,t_n} > u_n \bigg| X_1 > u_n) = o(\theta_n) \) 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 \( t_n = s_n + 1, \ldots, r_n \), and \( k_n = 1, 2, \ldots \)

\[
\Pr \left( \sum_{i=1}^{r_n-1} 1(X_i \leq v_n, X_{i+1} > v_n) \geq k_n \mid M_{r_n} > u_n \right) = \theta_n^{-1} \Pr \left( X_1 > v_n, X_2 \leq v_n, \sum_{i=2}^{t_n-1} 1(X_i \leq v_n, X_{i+1} > v_n) = k_n - 1 \mid X_1 > u_n \right) + o(1).
\]

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

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

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

\[
\Pr(\exists 1 \leq i \leq r_n : X_i > u_n, X_j > u_n, j - i \geq k_n \mid M_{r_n} > u_n) = \theta_n^{-1} \Pr(\exists j = k_n + 1, \ldots, t_n : M_{1:j-k_n} \leq u_n, X_j > u_n, M_{j:t_n} \leq u_n \mid X_1 > u_n) + o(1).
\]

### 4. Tail sequence of a stationary sequence

Let \( \{X_n\} \) be a stationary sequence of random variables. For \( m = 1, 2, \ldots \) denote the distribution function of \((X_1, \ldots, X_m)\) 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

\[
\lim_{u \to x_+} \frac{1 - F(u + g(u)x)}{1 - F(u)} = \begin{cases} (1 + \gamma x)^{-1/\gamma} & \text{if } \gamma \neq 0 \\ e^{-x} & \text{if } \gamma = 0 \end{cases}
\]

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 \((x_1, \ldots, x_m) \in (\alpha, \infty)^m\)

\[
\lim_{u \to x^+} \frac{1 - F_m(u + g(u)x_1, \ldots, u + g(u)x_m)}{1 - F(u)}
\]

exists and is finite.

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

\[
\lim_{u \to x^+} \frac{1 - F_m(u + g(u)x_1, \ldots, u + g(u)x_m)}{1 - F(u)} = -\log G_m(x_1, \ldots, x_m),
\]

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

\[
G_m(x_1, \ldots, x_m) = \lim_{n \to \infty} F_{m}^n(u_n + g(u_n)x_1, \ldots, u_n + g(u_n)x_m) \quad (10)
\]

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 \( \alpha \) 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

\[
G_m(x_1, \ldots, x_m) = \exp \left\{ -\int_0^1 \max_{i=1,\ldots,m} \frac{f_i(u)}{z_i} \, du \right\}
\]

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

\[
z_i = \begin{cases} (1 + \gamma x_i)^{1/\gamma} & \text{if } \gamma \neq 0 \\ e^{x_i} & \text{if } \gamma = 0 \end{cases}
\]

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 $u \uparrow x^+$

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

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

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

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

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

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 $m = 1, 2, \ldots$

$$
\Pr \left\{ \left( \frac{X_1 - u}{g(u)}, \ldots, \frac{X_m - u}{g(u)} \right) \in \cdot \mid X_1 > u \right\} \xrightarrow{w} \Pr \{ (\phi(S_0 + T), \ldots, \phi(S_{m-1} + T)) \in \cdot \}
$$
as \( u \uparrow x_+ \), where \( \phi(x) = \int_0^x e^{uy} \, dy \) for \( x \in \mathbb{R} \), \( S_0 = 0 \), and \( S_n = \sum_{i=1}^n \xi_i \) for \( n \geq 1 \).

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

**Theorem 3.** Let \( \{X_n\} \) be a stationary sequence with tail chain \( \{Y_n\} \) 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
\[
\lim_{m \to \infty} Y_m = \alpha \quad \text{almost surely and}
\]
\[
\Pr (M_{1,r_n} \leq u_n \mid X_1 > u_n) \to \theta := \Pr \left( \max_{i \geq 2} Y_i \leq 0 \right) > 0.
\] (14)

(ii) Let \( c \) be a cluster functional and let \( D \subset \mathcal{A} \) 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
\[
\Pr \left\{ c \left( \frac{X_1 - u_n}{g(u_n)} \right)^{r_n}_{i=1} \in \cdot \mid M_{r_n} > u_n \right\}
\]
\[
\Rightarrow \quad \theta^{-1} [\Pr \{c(Y_1, Y_2, \ldots) \in \cdot\} - \Pr \{c(Y_2, Y_3, \ldots) \in \cdot, \max_{i \geq 2} Y_i > 0\}].
\] (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 \( c(x_1, \ldots, x_r) = c(x_1 \lor 0, \ldots, x_r \lor 0) \) 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 \( \alpha \) 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
\[
C_n(j, k) = c_n \{(X_i - u_n)_{i=j}^k\}, \quad \text{integer } n \geq 1 \text{ and } 1 \leq j \leq k.
\] (16)
Without loss of generality we assume that \( |c_n| \leq 1 \). For a random variable \( X \) and an event \( A \) we write \( E[X; A] = E[X 1_A] \); in particular \( E[X|A] = E[X; A]/\Pr(A) \) provided \( \Pr(A) > 0 \).
Lemma 1. (i) If (1), then
\[ Pr(M_{t_n, r_n} > u_n \mid X_1 > u_n) = o \left( \frac{Pr(M_{t_n} > u_n)}{t_n Pr(X_1 > u_n)} \right). \] (17)

(ii) If additionally \( s_n = o(t_n) \), then
\[ Pr(M_{t_n, r_n} > u_n \mid M_{t_n} > u_n) = Pr(M_{r_n - t_n} > u_n \mid M_{r_n - t_n, r_n} > u_n) \to 0. \] (18)

Proof. (i) Since \( Pr(M_{t_n, r_n} > u_n \mid X_1 > u_n) \leq Pr(M_{s_n, r_n} > u_n \mid X_1 > u_n) \), we need only show that
\[ \liminf_{n \to \infty} \frac{t_n^{-1} Pr(M_{t_n} > u_n)}{s_n^{-1} Pr(M_{s_n} > u_n)} > 0. \] (19)

If \( r_n \leq 2s_n \) this is obviously true, so assume \( r_n > 2s_n \). By stationarity
\[
\begin{align*}
Pr(M_{2s_n, r_n} > u_n \mid M_{s_n} > u_n) & \leq \frac{\sum_{i=1}^{s_n} Pr(X_i > u_n, M_{2s_n, r_n} > u_n)}{Pr(M_{s_n} > u_n)} \\
& \leq \frac{s_n Pr(X_1 > u_n)}{Pr(M_{s_n} > u_n)} Pr(M_{s_n+1, r_n} > u_n \mid X_1 > u_n) \to 0.
\end{align*}
\] (20)

Denoting \( k_n = \lfloor (1 + t_n/s_n)/2 \rfloor \), we have
\[
\begin{align*}
Pr(M_{t_n} > u) & \geq Pr \left( \bigcup_{i=1}^{k_n} \{ M_{2(i-1)s_n, (2i-1)s_n} > u_n \} \right) \\
& \geq \sum_{i=1}^{k_n} Pr(M_{2(i-1)s_n, (2i-1)s_n} > u_n, M_{2is_n, r_n} \leq u_n) \\
& \geq k_n Pr(M_{s_n} > u_n) Pr(M_{2s_n, r_n} \leq u_n \mid M_{s_n} > u).
\end{align*}
\] (21)

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

(ii) By stationarity
\[
\begin{align*}
Pr(M_{t_n} > u_n, M_{t_n, r_n} > u_n) & = Pr(M_{t_n} > u_n) + Pr(M_{r_n - t_n} > u_n) - Pr(M_{r_n} > u_n) \\
& = Pr(M_{r_n - t_n} > u_n, M_{r_n - t_n, r_n} > u_n).
\end{align*}
\]

Denote \( l_n = \lfloor t_n/s_n \rfloor \). We have
\[
\begin{align*}
Pr(M_{t_n} > u_n, M_{t_n, r_n} > u_n) & \leq \sum_{i=1}^{l_n} Pr(M_{(i-1)s_n, is_n} > u_n, M_{s_n, r_n} > u_n) + Pr(M_{s_n, s_n} > u_n, M_{t_n, r_n} > u_n) \\
& \leq (l_n - 1) Pr(M_{s_n} > u_n, M_{2s_n, r_n} > u_n) + 2 Pr(M_{s_n} > u_n).
\end{align*}
\]
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

\[
\Pr(M_{r_n} > u_n) \sim (r_n/t_n)\Pr(M_{t_n} > u_n) \quad (22)
\]

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

**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

\[
E[C_n(1, r_n); M_{r_n} > u_n] = \sum_{i=1}^{k_n} E[C_n((i-1)t_n, r_n); M_{(i-1)t_n} \leq u_n, M_{(i-1)t_n, t_n} > u_n] + E[C_n(k_n t_n, r_n); M_{k_n t_n} \leq u_n, M_{k_n t_n, r_n} > u_n].
\]

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

\[
E[C_n(1, r_n); M_{r_n} > u_n] = k_n E[C_n(1, t_n); M_{t_n} > u_n] + O\{\Pr(M_{t_n} > u_n)\}
\]

\[
+ O\{(r_n/t_n)\Pr(M_{r_n-t_n} > u_n, M_{r_n-t_n, r_n} > u_n)\}
\]

\[
+ O\{(r_n/t_n)\Pr(M_{t_n} > u_n, M_{t_n, r_n} > u_n)\}
\]

\[
= (r_n/t_n) E[C_n(1, t_n); M_{t_n} > u_n] + o((r_n/t_n)\Pr(M_{t_n} > u_n)). \quad (24)
\]

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

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

\[
E[C_n(1, t_n) \mid M_{t_n} > u_n] = \frac{t_n \Pr(X_1 > u_n)}{\Pr(M_{t_n} > u_n)} \left\{ E[C_n(1, t_n) \mid X_1 > u_n] - E[C_n(t_n, 2t_n - 1); M_{t_n-1} > u_n, X_{t_n} > u_n] \right\} + o(1).
\]
Proof. Since by stationarity \( \mathbb{E}[C_n(1, t_n); X_1 > u_n] = \mathbb{E}[C_n(t_n, 2t_n - 1); X_{t_n} > u_n] \), it is sufficient to show

\[
\mathbb{E}[C_n(1, t_n); M_{t_n} > u_n] = t_n \mathbb{E}[C_n(t_n, 2t_n - 1); M_{t_n-1} \leq u_n, X_{t_n} > u_n] + o(\Pr(M_{t_n} > u_n)).
\]

(25)

By stationarity and (18), we have

\[
\mathbb{E}[C_n(1, t_n); M_{t_n} > u_n]
\]

\[
= \mathbb{E}[C_n(t_n, 2t_n - 1); M_{t_n-1, 2t_n-1} > u_n]
\]

\[
= \mathbb{E}[C_n(t_n, 2t_n - 1); M_{t_n-1} \leq u_n, M_{t_n-1, 2t_n-1} > u_n, M_{2t_n-1, 3t_n-2} \leq u_n]
\]

\[
+ O\{\Pr(M_{t_n-1} > u_n, M_{t_n-1, 2t_n-1} > u_n)\}
\]

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

\[
= \mathbb{E}[C_n(t_n, 2t_n - 1); M_{t_n-1} \leq u_n, M_{t_n-1, 2t_n-1} > u_n, M_{2t_n-1, 3t_n-2} \leq u_n]
\]

\[
+ o\{\Pr(M_{t_n} > u_n)\}.
\]

(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 \( X_i > u_n \):

\[
\mathbb{E}[C_n(t_n, 2t_n - 1); M_{t_n-1} \leq u_n, M_{t_n-1, 2t_n-1} > u_n, M_{2t_n-1, 3t_n-2} \leq u_n]
\]

\[
= \sum_{i=t_n}^{2t_n-1} \mathbb{E}[C_n(t_n, 2t_n - 1); M_{i-1} \leq u_n, X_i > u_n, M_{2t_n-1, 3t_n-2} \leq u_n]
\]

\[
= \sum_{i=t_n}^{2t_n-1} \mathbb{E}[C_n(i, i + t_n - 1); M_{i-1} \leq u_n, X_i > u_n, M_{2t_n-1, 3t_n-2} \leq u_n],
\]

(27)

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

\[
|\mathbb{E}[C_n(i, i + t_n - 1); M_{i-1,i-1} \leq u_n, X_i > u_n] - \mathbb{E}[C_n(i, i + t_n - 1); M_{i-1} \leq u_n, X_i > u_n, M_{2t_n-1, 3t_n-2} \leq u_n]| \]

(28)

is bounded by

\[
\Pr(M_{i-1} > u_n, M_{i-1,i-1} \leq u_n, X_i > u_n) + \Pr(M_{i-1,i-1} \leq u_n, X_i > u_n, M_{2t_n-1, 3t_n-2} > u_n),
\]

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 \( i \geq t_n \)
\[
E[C_n(i, i + t_n - 1); M_{t_n-1,i} \leq u_n, X_i > u_n]
= E[C_n(t_n, 2t_n - 1); M_{t_n-1} \leq u_n, X_{t_n} > u_n]
\]
by stationarity, equation (25) follows from (26), (27), and the bound on (28). This
proves the Lemma.

**Lemma 4.**

\[
|E[C_n(t_n, 2t_n - 1); M_{t_n-1} > u_n | X_{t_n} > u_n] - E[C_n(2, t_n); M_{1,t_n} > u_n | X_1 > u_n]| \\
\leq 2 \Pr(M_{t_n,2t_n-1} > u_n | X_1 > u_n).
\]

**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
\[
E[C_n(t_n, 2t_n - 1); M_{t_n-1} > u_n, X_{t_n} > u_n] \\
= \sum_{k=1}^{t_n-1} E[C_n(t_n, 2t_n - 1); X_k > u_n, M_{k,t_n-1} \leq u_n, X_{t_n} > u_n] \\
= \sum_{k=1}^{t_n-1} E[C_n(k + 1, 2t_n - 1); X_k > u_n, M_{k,t_n-1} \leq u_n, X_{t_n} > u_n].
\]
Stationarity implies
\[
E[C_n(t_n, 2t_n - 1); M_{t_n-1} > u_n, X_{t_n} > u_n] \\
= \sum_{k=1}^{t_n-1} E[C_n(2, 2t_n - k); X_1 > u_n, M_{1,t_n-k} \leq u_n, X_{t_n-k+1} > u_n] \\
= \sum_{l=1}^{t_n-1} E[C_n(2, t_n + l); X_1 > u_n, M_{1,l} \leq u_n, X_{l+1} > u_n],
\]
with \( l = t_n - k \). On the other hand,
\[
E[C_n(2, t_n); X_1 > u_n, M_{1,t_n} > u_n] \\
= \sum_{l=1}^{t_n-1} E[C_n(2, t_n); X_1 > u_n, M_{1,l} \leq u_n, X_{l+1} > u_n].
\]
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 \( l = 1, \ldots, t_n-1 \)
\[
|E[C_n(2, t_n + l); X_1 > u_n, M_{1,l} \leq u_n, X_{l+1} > u_n] \\
- E[C_n(2, t_n); X_1 > u_n, M_{1,l} \leq u_n, X_{l+1} > u_n]|.
\]
is bounded by $2 \Pr(X_1 > u_n, M_{1,t} \leq u_n, X_{t+1} > u_n, M_{t_n,t_n + t} > u_n)$. Therefore it remains to consider

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

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

$$E[C_n|X_1 > u_n] = E[C_n|X_1 > u_n] + o(1).$$

and thus, by Lemma 3,

$$E[C_n|X_1 > u_n] = \frac{t_n \Pr(X_1 > u_n)}{\Pr(M_{t_n} > u_n)} \left\{ E[C_n|X_1 > u_n] - E[C_n(t_n, 2t_n - 1); M_{t_n - 1} > u_n \mid X_{t_n} > u_n] \right\} + o(1).$$

Using Lemma 1(i) and Lemma 4 we get

$$E[C_n|X_1 > u_n] = \frac{t_n \Pr(X_1 > u_n)}{\Pr(M_{t_n} > u_n)} \left\{ E[C_n(1, t_n) | X_1 > u_n] - E[C_n(2, t_n - 1); M_{t_n} > u_n \mid X_1 > u_n] \right\} + o(1).$$

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

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

$$|E[C_n(1, s_n) | X_1 > u_n] - E[C_n(1, t_n) | X_1 > u_n]|,$$

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

are $O\{Pr(M_{s_n,r_n} > u_n \mid X_1 > u_n)\} = o \left( \frac{Pr(M_{s_n,r_n} > u_n)}{s_n \Pr(X_1 > u_n)} \right) = o(\theta_n)$ by (19).
5.2. Proofs of Corollaries 1–3

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

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 $\Pr(M_{s_n, r_n} > u_n | X_1 > u_n) \leq \Pr(M_{m, r_n} > u_n | X_1 > u_n)$. Hence

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

(29)

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

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

and thus

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

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

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

(30)

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

Proof of Corollary 3. Clearly

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

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

$$\Pr(M_{r_n} > u_n) \leq \left[ r_n/s_n \right] \Pr(M_{s_n} > u_n) \leq 2(r_n/s_n) \Pr(M_{s_n} > u_n) = o\left( \frac{\Pr(M_{s_n} > u_n)}{s_n \Pr(X_1 > u_n)} \right).$$

The assumption on $\Delta_n$ 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

\[
\lim_{u \to x_+} \frac{g(u + g(u)x)}{g(u)} = 1 + \gamma x
\]

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 \( x_i > \alpha, i = 2, \ldots, m \),

\[
\Pr\left\{ \frac{X_1 - u}{g(u)} \leq x_1, \ldots, \frac{X_m - u}{g(u)} \leq x_m \left| X_1 > u \right. \right\} = F_m(u + g(u)x_1, \ldots, u + g(u)x_m) - F_m(u, u + g(u)x_2, \ldots, u + g(u)x_m) \leq 1 - F(u) \rightarrow -\log G_m(0, x_2, \ldots, x_m) - \{- \log G_m(x_1, x_2, \ldots, x_m)\}
\]

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

\[
\Pr\left\{ \frac{X_1 - u}{g(u)} \lor \alpha, \ldots, \frac{X_m - u}{g(u)} \lor \alpha \right\} \leq \Pr\left\{ (Y_1^{(m)}, \ldots, Y_m^{(m)}) \in \cdot \right\}
\]

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

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

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

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

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

(33)

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

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

(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

\[
\Pr\left\{ \frac{X_1 - u}{g(u)} > x_1, \frac{X_2 - u}{g(u)} \leq x_2, \ldots, \frac{X_m - u}{g(u)} \leq x_m \right\} \left/ \left[ 1 - F(u) \right] \right. = 1 - F(v)
\]

5.4. Proof of Theorem 3

(i) By Corollary 2

\[
\theta_n := \frac{\Pr(M_{m,n} > u_n)}{r_n \Pr(X_1 > u_n)} = \Pr(M_{1,r_n} \leq u_n \mid X_1 > u_n) + o(1)
\]

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 \( Y_m \geq \alpha \) almost surely, it suffices to show that

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

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

\[
\Pr\left( \max_{m < i \leq l} Y_i > 0 \right) = \lim_{n \to \infty} \Pr(M_{m,l} > u_n \mid X_1 > u_n) \leq \limsup_{n \to \infty} \Pr(M_{m,r_n} > u_n \mid X_1 > u_n).
\]

Let \( l \to \infty \) to obtain \( \Pr(\max_{i > m} Y_i > 0) \leq \limsup_{n \to \infty} \Pr(M_{m,r_n} > u_n \mid X_1 > u_n) \).

By (6) we get (35) for \( x = 0 \). Second, let \( \alpha < x < 0 \). Clearly \( 1 + \gamma x > 0 \). For arbitrary integers \( 1 \leq m < l \)

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

Since (35) is true for \( x = 0 \), we now see it is in fact true for arbitrary \( x > \alpha \).
To show (14), write for $m = 1, \ldots, r_n$

$$\left| \Pr(M_{1,r_n} \leq u_n \mid X_1 > u_n) - \Pr \left( \max_{i \geq 2} Y_i \leq 0 \right) \right|$$

$$\leq \Pr(M_{m,r_n} > u_n \mid X_1 > u_n) + \Pr \left( \max_{i > m} Y_i > 0 \right)$$

$$+ \left| \Pr(M_{1,m} \leq u_n \mid X_1 > u_n) - \Pr \left( \max_{i=2,\ldots,m} Y_i \leq 0 \right) \right|.$$

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 $c(x_1, \ldots, x_r) = 0$ whenever $\max_{i=1,\ldots,r} x_i \leq 0$. 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)

$$E \left[ c \left( \frac{X_i - u_n}{g(u_n)} \right)^{\alpha} \right]_{i=1}^{r_n} \left| M_{n,r_n} > u_n \right]$$

$$= \theta^{-1} E \left[ c \left( \frac{X_i - u_n}{g(u_n)} \right)^{m} \right] - c \left( \frac{X_i - u_n}{g(u_n)} \right)^{m} \left| X_1 > u_n \right] + \varepsilon_{n,m}.$$

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

$$0 = \Pr\{Y_1, Y_2, \ldots \in D, \max_{i > m} Y_i \leq 0\} = \Pr\{Y_1, \ldots, Y_m \in D_m, \max_{i > m} Y_i \leq 0\}$$

$$= \Pr\{Y_1, \ldots, Y_m \in D_m\} - \Pr\{Y_1, \ldots, Y_m \in D_m, \max_{i > m} Y_i > 0\}$$

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

$$E \left[ c \left( \frac{X_i - u_n}{g(u_n)} \right)^{\alpha} \right]_{i=1}^{r_n} \left| M_{n,r_n} > u_n \right]$$

$$= \theta^{-1} E[c(Y_1, \ldots, Y_m) - c(Y_2, \ldots, Y_m)] + \varepsilon_{n,m}$$

$$= \theta^{-1} E[c(Y_1, Y_2, \ldots) - c(Y_2, Y_3, \ldots)] + \varepsilon_{n,m}.$$

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

$$E \left[ c \left( \frac{X_i - u_n}{g(u_n)} \right)^{\alpha} \right]_{i=1}^{r_n} \left| M_{n,r_n} > u_n \right]$$

$$= \theta^{-1} E[c(Y_1, Y_2, \ldots) - c(Y_2, Y_3, \ldots)] + o(1). \quad (36)$$

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


