Characterization and Estimation of the Multivariate Extremal Index

R.L. Smith and I. Weissman

Third draft: December 1996

Abstract

The multivariate extremal index has been introduced by Nandagopalan as a measure of the clustering among the extreme values of a multivariate stationary process. In this paper, we derive some additional properties and use those to construct a statistical estimation scheme. Central to the discussion is a class of processes we call M_4 processes, which are characterized by means of a multivariate generalization of a characterization due to Deheuvels for univariate max-stable processes. The multivariate extremal index for M_4 processes is derived, and certain properties established. We then discuss estimation of the multivariate extremal index.

Keywords: Moving maxima processes, multivariate extreme value theory, stationary processes.

1 Introduction

Suppose $\{X_i,\ i = 1, 2, ...,\}$ is a stationary sequence in one dimension, with continuous marginal distribution function $F(x) = P\{X_i \leq x\}$. Let $M_n = \max\{X_1, ..., X_n\}$ and also let $\hat{M}_n = \max(\hat{X}_1, ..., \hat{X}_n)$ where $\{\hat{X}_i\}$ is the so-called *associated* sequence of i.i.d. random variables with the same marginal distribution function F. Suppose $\{u_n,\ n \geq 1\}$

is a sequence of thresholds such that $n\{1 - F(u_n)\} \to \tau$ as $n \to \infty$, where $0 < \tau < \infty$. Then it follows immediately that

$$P\{\hat{M}_n \leq u_n\} \to e^{-\tau}. \tag{1.1}$$

Under quite mild additional conditions, we also have that

$$P\{M_n \leq u_n\} \to e^{-\theta\tau}. \tag{1.2}$$

where θ is a parameter called the *extremal index* (Leadbetter 1983, Leadbetter, Lindgren and Rootzén 1983). Among its more elementary properties are that θ can take any value in the interval $[0, 1]$, and that if the limit (1.2) holds for some τ and corresponding $\{u_n\}$ sequence, then it holds for all; in other words, the value of θ does not depend on the particular τ or $\{u_n\}$ sequence chosen.

In this paper, we are interested in multivariate extensions of these results. Suppose $\{\boldsymbol{X}_i = (X_{i1}, \ldots, X_{iD}),\ i = 1, 2, \ldots,\}$ is a D-dimensional stationary stochastic process with marginal distribution functions $F(\boldsymbol{x}) = F(x_1, \ldots, x_D) = P\{X_{id} \leq x_d,\ d = 1, \ldots, D\}$ and $F_d(x) = P\{X_{id} \leq x\}$, $d = 1, \ldots, D$. Suppose $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_D)$ is a vector of nonnegative finite numbers, and suppose for each $d \in \{1, \ldots, D\}$, $\{u_{nd}, n \geq 1\}$ is a sequence of thresholds satisfying

$$n\{1 - F_d(u_{nd})\} \to \tau_d \tag{1.3}$$

as $n \to \infty$ for each d. Let $\boldsymbol{M}_n = (M_{n1}, \ldots, M_{nD})$ denote the vector of pointwise maxima, $M_{nd} = \max\{X_{id},\ 1 \leq i \leq n\}$. Also let $\{\hat{\boldsymbol{X}}_i\}$ denote the associated sequence of i.i.d. random vectors having the same D-dimensional distribution function F, and let $\hat{\boldsymbol{M}}_n = (\hat{M}_{n1}, \ldots, \hat{M}_{nD})$ denote the vector of pointwise maxima from $\{\hat{\boldsymbol{X}}_i\}$. We are interested in cases where the joint limits

$$\begin{aligned} H(\boldsymbol{\tau}) &= \lim_{n\to\infty} P\{M_{n1} \leq u_{n1}, \ldots, M_{nD} \leq u_{nD}\}\ , \\ \hat{H}(\boldsymbol{\tau}) &= \lim_{n\to\infty} P\{\hat{M}_{n1} \leq u_{n1}, \ldots, \hat{M}_{nD} \leq u_{nD}\}\ , \end{aligned} \tag{1.4}$$

both exist and are nonzero. In such cases, we may define a parameter $\theta(\boldsymbol{\tau})$ by the relation

$$H(\boldsymbol{\tau}) = \hat{H}(\boldsymbol{\tau})^{\theta(\boldsymbol{\tau})}. \tag{1.5}$$

With minor changes in notation, this is what Nandagopalan (1990, 1994) called the *multivariate extremal index.* Just as in one dimension, it is the key parameter relating the extreme-value properties of a stationary process to those of independent random vectors from the same D-dimensional marginal distribution. However, unlike the one-dimensional case, it is not a constant for the whole process, but instead depends on the vector $\boldsymbol{\tau}$. Some elementary properties include

(i) $0 \le \theta(\boldsymbol{\tau}) \le 1$ for all $\boldsymbol{\tau}$,

(ii) if $\tau_d > 0$ but $\tau_{d'} = 0$ for all $d' \neq d$, then $\theta(\boldsymbol{\tau}) = \theta_d$, the extremal index for the d^{th} component process; namely $\theta(0, \ldots, 0, \tau_d, 0, \ldots, 0) = \theta_d$

(iii) $\theta(c\boldsymbol{\tau}) = \theta(\boldsymbol{\tau})$ for all $c > 0$ (Theorem 1.1 of Nandagopalan, 1994).

However, these properties are not sufficient to characterize the function $\theta(\boldsymbol{\tau})$. We would like a more precise characterization, for two reasons. First, the number of examples for which the multivariate extremal index has been calculated is currently very small (Nandagopalan 1994, Weissman 1994) and it is important to be able to extend this class to cover a much broader range of processes. The second reason why we need a characterization is statistical: crude estimators of $\theta(\boldsymbol{\tau})$ are easy to construct, but would not correspond to the multivariate extremal index of any real stochastic process. The situation is analogous to estimating the dependence function of a multivariate extreme value distribution, for which it is necessary to impose some convexity conditions, cf. Pickands (1981), Smith (1985), Smith, Tawn and Yuen (1990). In the univariate case, estimators of the extremal index have been proposed by Leadbetter et al. (1989), Nandagopalan (1990), Hsing (1993) and Smith and Weissman (1994). Thus, a part of our purpose is to bring together two previously separate branches in the statistics of extreme values.

In Section 2 we introduce a particular class of stationary processes, *Multivariate Maxima of Moving Maxima* (henceforth M_4) *processes*, and prove a characterization result similar to Deheuvels (1983). In essence, our claim is that for a very wide class of processes, the multivariate extremal index may be approximated arbitrarily closely by one from the M_4 class. In Section 3 we study the M_4 class in more detail, derive the multivariate extremal index for this class, and establish some properties. The remaining sections are concerned with more practical aspects: some statistical theory in Section 4, then two examples in Section 5, along with some simulations to illustrate the methods in Section 6.

2 Multivariate Maxima of Moving Maxima

In this section we introduce the class of M_4 processes and argue that the problem can effectively be reduced to the study of these processes. There are two steps to this argument. First, under an assumption that the finite-dimensional distributions of the process are in the domain of max-attraction of a max-stable process, together with some mixing conditions, it is shown that the extremal index of a multivariate stationary process is the same as that of the limiting max-stable process. This reduces the problem to the study of max-stable processes. The second step is to adapt an argument originally given by Deheuvels for the characterization of one-dimensional max-stable processes. If we exclude processes containing a deterministic component, in a sense that will be explained later, then the process may be approximated arbitrarily closely by one in the M_4 class. In this sense, the multivariate extremal indices of M_4 processes form a rich subclass of those of general multivariate stationary sequences.

As a preliminary, observe that provided all the marginal distributions are continuous, the multivariate extremal index is invariant under pointwise strictly increasing continuous transformations. Therefore, there is no loss of generality in assuming that $F_1(x), ..., F_D(x)$ take any given (continuous) form. We choose to adopt the unit Fréchet form, i.e. $F_d(x) = e^{-1/x},\ 0 < x < \infty$, for $1 \leq d \leq D$. This assumption is made for the rest of the paper.

The following elementary result gives an alternative characterization of the multivariate extremal index:

Proposition 2.1 *Let $\{\boldsymbol{X}_i\}$ be a D-dimensional stationary sequence with unit Fréchet margins and multivariate extremal index $\theta(\boldsymbol{\tau})$. For each fixed $\boldsymbol{\tau} = (\tau_1, \ldots, \tau)$ $(\tau_d \geq 0,\ d = 1, \ldots, D)$ define the (univariate) stationary sequence $\{V_i(\boldsymbol{\tau}) = \max_d \tau_d X_{id} : i \geq 1\}$. Then $\theta(\boldsymbol{\tau})$ is the (univariate) extremal index of the sequence $\{V_i(\boldsymbol{\tau})\}$.*

Proof For each $u > 0$ and $i \geq 1$, $P\{V_i(\boldsymbol{\tau}) \leq u\} = P\{X_{id} \leq u/\tau_d\ \forall\ d\}$, thus for $M_n^V := \max\{V_1(\boldsymbol{\tau}), \ldots, V_n(\boldsymbol{\tau})\}$ we have

$$P\{M_n^V \leq n\} = P\{M_{n1} \leq n/\tau_1, \ldots, M_{nD} \leq n/\tau_D\}\ . \qquad (2.1)$$

Similarly, for the associated sequence $\{\hat{V}_i(\boldsymbol{\tau})\}$ we have

$$P\{\hat{M}_n^V \leq n\} = P\{\hat{M}_{n1} \leq n/\tau_1, \ldots, \hat{M}_{nD} \leq n/\tau_D\}\ . \qquad (2.2)$$

Since $u_{nd} := n/\tau_d$ satisfies (1.3), by assumption, both (2.1) and (2.2) converge, as $n \to \infty$, to $H(\boldsymbol{\tau})$ and $\hat{H}(\boldsymbol{\tau})$ respectively, and these limits determine $\theta(\boldsymbol{\tau})$ via (1.5). Hence, $\theta(\boldsymbol{\tau})$ is the extremal index of the sequence $\{V_i(\boldsymbol{\tau})\}$. □

Next, we define an M_4 process. Let $\{Z_{\ell,i},\ \ell \geq 1, -\infty < i < \infty\}$ denote an array of independent unit Fréchet random variables, and consider the process defined by

$$Y_{id} = \max_{\ell \geq 1} \max_{-\infty < k < \infty} a_{\ell k d} Z_{\ell, i-k} \tag{2.3}$$

for nonnegative constants $\{a_{\ell k d},\ \ell \geq 1, -\infty < k < \infty,\ 1 \leq d \leq D\}$, satisfying

$$\sum_{\ell=1}^{\infty} \sum_{k=-\infty}^{\infty} a_{\ell k d} = 1 \qquad \text{for } d = 1, \ldots, D\ . \tag{2.4}$$

For any set of indices $i = 1, 2, \ldots, n$ and positive constants $\{y_{id},\ 1 \leq i \leq n,\ 1 \leq d \leq D\}$ (where possibly $y_{id} = +\infty$ for some i and d), we have

$$\begin{aligned} & P\{Y_{id} \leq y_{id},\ 1 \leq i \leq n,\ 1 \leq d \leq D\} \\ = \ & P\left\{Z_{\ell,i-k} \leq \frac{y_{id}}{a_{\ell k d}} \ \text{ for } \ell \geq 1, -\infty < k < \infty,\ 1 \leq i \leq n,\ 1 \leq d \leq D\right\} \\ = \ & P\left\{Z_{\ell,m} \leq \min_{1-m \leq k \leq n-m} \min_{1 \leq d \leq D} \frac{y_{m+k,d}}{a_{\ell k d}}\ , \quad \ell \geq 1,\ -\infty < m < \infty\right\} \\ = \ & \exp\left[-\sum_{\ell=1}^{\infty} \sum_{m=-\infty}^{\infty} \max_{1-m \leq k \leq n-m} \max_{1 \leq d \leq D} \frac{a_{\ell k d}}{y_{m+k,d}}\right]\ . \end{aligned} \tag{2.5}$$

¿From (2.5), it can be seen that $P^n\{Y_{id} \leq n y_{id},\ 1 \leq i \leq n,\ 1 \leq d \leq D\}$ is independent of n, so the finite-dimensional distributions of $\{\boldsymbol{Y}_i\}$ are *max-stable* (also called multivariate extreme value). Resnick (1987) and Galambos (1987, Chapter 5) gave general surveys of multivariate extreme value distributions — in particular, there are a number of characterizations of multivariate extreme value distributions which rely on approximations similar to (2.5) (cf. Deheuvels, 1978, de Haan and Resnick, 1977, and especially Pickands, 1981). Following de Haan (1984), discrete-time processes whose finite-dimensional distributions are max-stable are known as *max-stable processes*, and there are a number of characterizations of these. We shall focus particularly on a characterization due to Deheuvels (1983) and its generalization to the multivariate case.

If the process (2.3) is restricted to a single value of ℓ and a single index d, then it is a moving maximum process. By allowing d to range over $1, ..., D$ and by also maximizing over independent processes indexed by ℓ, we obtain *maxima* of *multivariate* moving maxima, or M_4 for short.

Now we turn to the first of our two main tasks in this section, which is to establish conditions under which the problem of extremes in multivariate stationary processes may be reduced to the case where the process being studied is max-stable.

The principal assumption is that the finite-dimensional distributions of the original process $\{\boldsymbol{X}_i\}$ lie in the domain of max-attraction of a stationary process $\{\boldsymbol{Y}_i = (Y_{i1}, ..., Y_{iD}),\ i = 1, 2, ...,\}$. Specifically, if $r > 1$, $0 < y_{id} \le \infty$ for $i = 1, ..., r$, $d = 1, ..., D$, and $0 < u_{nid} \le \infty$ for $n \ge 1$, $i = 1, ..., r$, $d = 1, ..., D$ are constants such that $u_{nid}/n \to y_{id}$ as $n \to \infty$, then we assume

$$\lim_{n\to\infty} P^n \{X_{id} \le u_{nid},\ 1 \le i \le r,\ 1 \le d \le D\} = P\{Y_{id} \le y_{id},\ 1 \le i \le r,\ 1 \le d \le D\}. \tag{2.6}$$

If such a process $\{\boldsymbol{Y}_i\}$ exists, then it is necessarily max-stable.

The assumption that a max-stable limiting process exists is of course a new assumption and not a consequence of the fact that all the one-dimensional margins of $\{\boldsymbol{X}_i\}$ are max-stable. Nevertheless, it seems natural to make such an assumption in studying the extremal behavior of such sequences. Similar assumptions have been made in previous papers by Smith (1992), Perfekt (1994), Yun (1994) and the statistical treatment of Smith *et al.* (1997).

Fix $\boldsymbol{\tau} = \{\tau_1, ..., \tau_D\}$ where each $0 \le \tau_d < \infty$ for each d. Let $\{u_{nd},\ n \ge 1\}$ denote a sequence of thresholds such that $n\{1 - F_d(u_{nd})\} \to \tau_d$. In view of the unit Fréchet assumption, one choice is $u_{nd} = n/\tau_d$. Let $\boldsymbol{u}_n = (u_{n1}, ..., u_{nD})$. For $1 \le j \le k \le n$, let $\mathcal{B}_j^k(\boldsymbol{u}_n)$ denote the σ-field generated by the events $\{X_{id} \le u_{nd},\ j \le i \le k\}$, and for each integer t let

$$\alpha_{n,t} = \sup\left\{|P(A \cap B) - P(A)P(B)| :\ A \in \mathcal{B}_1^k(\boldsymbol{u}_n),\ B \in \mathcal{B}_{k+t}^n(\boldsymbol{u}_n)\right\} \tag{2.7}$$

where the supremum is taken not only over all events A and B in their respective σ-fields but also over k such that $1 \le k \le n - t$. As in Nandagopalan (1994), the mixing condition $\Delta(\boldsymbol{u}_n)$ is said to hold if there exists a sequence $\{t_n,\ n \ge 1\}$ such that

$$t_n \to \infty,\ t_n/n \to 0,\ \alpha_{n,t_n} \to 0 \text{ as } n \to \infty. \tag{2.8}$$

Assuming $\Delta(\boldsymbol{u}_n)$, we can find a sequence $\{k_n,\ n \ge 1\}$ such that

$$k_n \to \infty,\ k_n t_n/n \to 0,\ k_n \alpha_{n,t_n} \to 0 \text{ as } n \to \infty. \tag{2.9}$$

Also let $r_n = \lfloor n/k_n \rfloor$ (here $\lfloor\cdot\rfloor$ denotes integer part). The process from $i = 1$ to n is effectively broken up into k_n nearly independent sections of length r_n, and by making detailed assumptions about the exceedances of $\boldsymbol{u}_n$ within a block of length r_n, Nandagopalan (1994) was able to characterize the extremal index in terms of a limiting exceedance point process, so generalizing results which had been earlier obtained for univariate process by Hsing *et al.* (1988).

An alternative approach to the extremal index is due to O'Brien (1987), which leads to the following multivariate generalization:

Lemma 2.2 *Suppose (2.7)-(2.9) hold. Then*

$$\theta(\boldsymbol{\tau}) = \lim_{n\to\infty} P\left\{X_{id} \le u_{nd},\ 2 \le i \le r_n,\ 1 \le d \le D \,\middle|\, \max_d \left(\frac{X_{1d}}{u_{nd}}\right) > 1\right\}. \quad (2.10)$$

Alternatively, if we assume

$$\lim_{r\to\infty} \lim_{n\to\infty} \sum_{i=r}^{r_n} \sum_{d=1}^{D} P\left\{X_{id} > u_{nd} \,\middle|\, \max_d \left(\frac{X_{1d}}{u_{nd}}\right) > 1\right\} = 0, \quad (2.11)$$

then (2.10) is equivalent to

$$\theta(\boldsymbol{\tau}) = \lim_{r\to\infty} \lim_{n\to\infty} P\left\{X_{id} \le u_{nd},\ 2 \le i \le r,\ 1 \le d \le D \,\middle|\, \max_d \left(\frac{X_{1d}}{u_{nd}}\right) > 1\right\}. \quad (2.12)$$

Proof. Our assumption $\Delta(\boldsymbol{u}_n)$ is slightly stronger than the "asymptotic independence of maxima" for a one-dimensional sequence made by O'Brien, and it is readily verified that r_n as we have defined it satisfies the same properties as O'Brien's p_n. The result (2.10) then follows from O'Brien's Theorem 2.1 applied to the process $\{V_i(\boldsymbol{\tau}),\ i \ge 1\}$ of Prop. 2.1. The extension to (2.12) under the assumption (2.11) is immediate (cf. Section 2 of Smith, 1992). □

We now come to our main result:

Theorem 2.3 *Suppose $\Delta(\boldsymbol{u}_n)$ and (2.11) hold for $\{\boldsymbol{X}_i\}$, so that the multivariate extremal index $\theta^{\boldsymbol{X}}(\boldsymbol{\tau})$ is given by (2.12). Suppose also the same assumptions hold for $\{\boldsymbol{Y}_i\}$ (with the same t_n, k_n sequences). So the multivariate extremal index $\theta^{\boldsymbol{Y}}(\boldsymbol{\tau})$ is also given by (2.12) with Y_{id} replacing X_{id} everywhere. Then $\theta^{\boldsymbol{X}}(\boldsymbol{\tau}) = \theta^{\boldsymbol{Y}}(\boldsymbol{\tau})$.*

Remark. If the $\Delta(\boldsymbol{u}_n)$ condition holds for both $\boldsymbol{X}$ and $\boldsymbol{Y}$ but with respect to different $\{t_n\}$ sequences, say $\{t'_n\}$ for $\boldsymbol{X}$ and $\{t''_n\}$ for $\boldsymbol{Y}$, then we just define $t_n = \max(t'_n, t''_n)$ and corresponding k_n so as to satisfy (2.9) for each of the two α_{n,t_n} sequences. So there is no loss of generality in assuming that the $\{t_n\}$ and $\{k_n\}$ sequences are common to the two processes.

Proof. Since both $\theta^{\boldsymbol{X}}(\boldsymbol{\tau})$ and $\theta^{\boldsymbol{Y}}(\boldsymbol{\tau})$ are given by (2.12), it suffices to show that, for each r,

$$\lim_{n\to\infty} P\left\{X_{id} \le u_{nd},\ 2 \le i \le r,\ 1 \le d \le D \;\middle|\; \max_d \left(\frac{X_{1d}}{u_{nd}}\right) > 1\right\} \tag{2.13}$$

is unchanged when X_{id} is replaced by Y_{id} everywhere.

However we can write (2.13) as $\lim_{n\to\infty}\{(P_{n1} - P_{n2})/(1 - P_{n3})\}$, where

$$\begin{aligned} P_{n1} &= P\{X_{id} \le u_{nd},\ 2 \le i \le r,\ 1 \le d \le D\}, \\ P_{n2} &= P\{X_{id} \le u_{nd},\ 1 \le i \le r,\ 1 \le d \le D\}, \\ P_{n3} &= P\{X_{1d} \le u_{nd},\ 1 \le d \le D\}. \end{aligned}$$

By the assumption of convergence to a max-stable process, $P_{nj}^n \to \phi_j$ for $j = 1, 2, 3$, some $\phi_j > 0$. Then the limit in (2.13) is $(\log\phi_2 - \log\phi_1)/(-\log\phi_3)$.

Now do the same thing with Y_{id} replacing X_{id} everywhere. Let Q_{n1}, Q_{n2}, Q_{n3} denote the probabilities for the $\boldsymbol{Y}$ process corresponding to P_{n1}, P_{n2}, P_{n3} for the $\boldsymbol{X}$ process. By max-stability, $Q_{nj}^n \to \phi_j$ for $j = 1, 2, 3$, and so the limit of $\{(Q_{n1} - Q_{n2})/(1 - Q_{n3})\}$ is also $(\log\phi_2 - \log\phi_1)/(-\log\phi_3)$. This completes the proof. $\square$

As a result of Theorem 2.3, we may confine our subsequent attention to the study of stationary max-stable processes, where there is again no loss of generality in assuming unit Fréchet margins. We now turn to a detailed examination of these processes. When $D = 1$, Deheuvels (1983) obtained an interesting characterization of such processes in terms of moving maxima of the form (2.1). It should be noted that Deheuvels was studying min-stable processes and a moving minimum representation, but his results may be translated into our present form by taking reciprocals of all random variables involved. For the remainder of this section, we follow the same steps as in Deheuvels' paper to establish a similar result in the multivariate case.

As in Theorem 2 of Deheuvels (1983), for any n we may approximate the joint distribution of $\{\boldsymbol{Y}_i,\ 1 \le i \le n\}$ arbitrarily closely by a relation of the form

$$Y_{id} = \max_{-\infty<k<\infty} \alpha_{idk} Z_k, \quad 1 \le i \le n,\ 1 \le d \le D \tag{2.14}$$

where $\{\alpha_{idk}\}$ are nonnegative constants and $\{Z_k\}$ are independent unit Fréchet. The following two lemmas are the same as Deheuvels' Lemmas 1 and 2.

Lemma 2.4 *If the representation (2.14) exists, then without loss of generality we may assume that the sequence*

$$\left\{\frac{\alpha_{idk}}{\sum_{i=1}^{n}\sum_{d=1}^{D}\alpha_{idk}}\ ,\ 1 \le i \le n,\ 1 \le d \le D\right\} \tag{2.15}$$

is distinct for each k.

Lemma 2.5 *If there are two representations of form (2.14), one with constants $\{\alpha_{idk}\}$ and the other with constants $\{\beta_{idk}\}$, and if each of these representations satisfies the conclusion of Lemma 2.4, then there exists a bijective mapping $v : \mathbb{Z} \to \mathbb{Z}$ such that $\beta_{idk} = \alpha_{idv(k)}$ for all i, d, k.*

Proofs See Deheuvels (1983). □

Theorem 2.6 *Suppose the sequence $\{\boldsymbol{Y}_i = (Y_{i,d},\ 1 \le d \le D)\}$ defined by (2.14) is stationary on the index i. Suppose for each (i,d) the points $\{\alpha_{idk}, -\infty < k < \infty\}$ are distinct, and that there exists a sequence $1 < n_1 < n_2 < \cdots$, increasing to infinity, such that the values in (2.15) are distinct for each $n = n_k$. Then there exists a bijective mapping $v : \mathbb{Z} \to \mathbb{Z}$ such that, for any $i \ge 0$, $1 \le d \le D$, $-\infty < k < \infty$, $\alpha_{idk} = \alpha_{0dv^i(k)}$.*

Proof This follows Theorem 3 of Deheuvels (1983). For each n, the sequences $\{\boldsymbol{Y}_i,\ 1 \le i \le n\}$ and $\{\boldsymbol{Y}_i,\ 2 \le i \le n+1\}$ are identical in distribution by stationarity. By Lemma 2.5, there exists a bijective mapping $v_n : \mathbb{Z} \to \mathbb{Z}$ such that $\alpha_{i+1,d,k} = \alpha_{i,d,v_n(k)}$ for $1 \le i \le n$, $1 \le d \le D$ and all k. By the assumption that the $\{\alpha_{idk}, -\infty < k < \infty\}$ sequences are distinct for each i and d, it follows that v_n is unique. Moreover, $v_{n+1}(k) = v_n(k)$ so that, by letting $n \to \infty$ along the subsequence $\{n_\ell\}$, we get a limit v (independently of n). It follows that $\alpha_{i+1,d,k} = \alpha_{i,d,v(k)}$ for all i, d, k, and hence the result follows. □

As in Theorem 4 of Deheuvels (1983), it can easily be seen that the process (2.14), with $\alpha_{idk} = \alpha_{0dv^i(k)}$ for some bijection v, is stationary and max-stable so long as $\sum_k \alpha_{0dk} < \infty$ for each d. Define an equivalence relation $p \sim q$ for $p,\ q \in \mathbb{Z}$ if there exists $i \in \mathbb{Z}$ such that $p = v^i(q)$; this partitions $\mathbb{Z}$ into equivalence classes, some of which may be finite and others infinite. By reordering within each equivalence class, we obtain the following result.

Theorem 2.7 *If Y_{id} is given by (2.14) with $\alpha_{idk} = \alpha_{0dv^i(k)}$ for some bijection v, and if $\sum_k \alpha_{0dk} < \infty$ for each d, then there exists a decomposition $Y_{id} = \max(R_{id}, S_{id})$ where*

$$\begin{aligned} R_{id} &= \max_{\ell \in I} \max_{-\infty < k < \infty} a_{\ell k d} Z_{\ell, i-k}, \\ S_{id} &= \max_{\ell \in F} \max_{0 \leq k \leq N_\ell} b_{\ell k d} Z^*_{\ell, i-k}, \end{aligned} \tag{2.16}$$

where I and F are two subclasses of indices ℓ, all the $\{Z_{\ell,j}\}$ and $\{Z^*_{\ell,j}\}$ are mutually independent unit-Fréchet random variables, and $Z^*_{\ell,n} = Z^*_{\ell,n+N_\ell}$ for each ℓ, n.

Proof As in Theorem 5 of Deheuvels (1983). □

Although our final result is a direct multivariate generalization of Deheuvels (1983), our interpretation is a little different from that of Deheuvels. The process $\{S_{id}\}$ is a maximum over periodic sequences, and therefore in principle, a perfectly predictable process. It seems reasonable to assume that in most applications such components do not occur, and we therefore eliminate them from further study. With trivial changes of notation, and renormalization, the process $\{R_{id}\}$ reduces to the M_4 process introduced in (2.1). From now on, we take this as the main process of interest.

Deheuvels (1983, 1985) made a further simplification, reducing the process to a single sequence of moving maxima (equivalent to fixing $\ell = 1$ in (2.3)). This could be misleading, however: while such processes are certainly of interest as special cases, there is no reason to assume in general that the reduction to a single component is valid, and Deheuvels himself remarked that the set I in (2.16) cannot, in general, be reduced to a single element.

To summarize the two main conclusions of this section: under the mixing conditions $\Delta(\boldsymbol{u}_n)$ and (2.11), applied to both $\boldsymbol{X}$ and $\boldsymbol{Y}$, and the assumption (2.6) about the joint convergence of pointwise maxima over i.i.d. realizations, we deduce that the limiting joint distributions of maxima are taken from a multivariate max-stable process, where without loss of generality we assume unit Fréchet margins. Second, any such max-stable process may be approximated arbitrarily closely by one of the forms given in Theorem 2.6. Under the further assumption that the degenerate $\{S_{id}\}$ component is absent, we deduce the representation (2.3) as the main object of further study.

3 The Extremal Index of an M_4 Process

In the univariate case, the extremal index of a moving maximum process has been studied in detail in a recent paper of Weissman and Cohen (1995). They also considered an extension to moving maximum processes with random coefficients, and (in their Prop. 3.1) to a process consisting of pointwise maxima over independent processes. Thus for the special case of (2.3) with $\ell \equiv 1$, the extremal index is given by

$$\theta(\boldsymbol{\tau}) = \frac{\max_k \max_d a_{kd}\tau_d}{\sum_k \max_d a_{kd}\tau_d} \,. \tag{3.1}$$

For the double sequence (2.3), we have the following result.

Theorem 3.1 *The extremal index of $\{\boldsymbol{Y}_i\}$ defined by (2.3) is given by*

$$\theta(\boldsymbol{\tau}) = \frac{\sum_\ell \max_k \max_d a_{\ell kd}\tau_d}{\sum_\ell \sum_k \max_d a_{\ell kd}\tau_d} \,. \tag{3.2}$$

Proof. Let $\boldsymbol{u}_n = (n/\tau_1, \ldots, n/\tau_d)$, then it follows from (2.5) (applied to $i = n = 1$) that

$$F(\boldsymbol{u}_n) = P\{\boldsymbol{Y}_1 \le \boldsymbol{u}_n\} = \exp\left[-\frac{1}{n}\sum_\ell \sum_k \max_d a_{\ell kd}\tau_d\right] . \tag{3.3}$$

Similarly,

$$P\{\boldsymbol{M}_n \le \boldsymbol{u}_n\} = \exp\left[-\frac{1}{n}\sum_\ell \sum_m \max_{1-m\le k\le n-m} \max_d a_{\ell kd}\tau_d\right] , \tag{3.4}$$

where $\boldsymbol{M}_n$ now stands for the vector of componentwise maxima from $\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_n$. For fixed ℓ, let $b_{\ell k} = \max_d a_{\ell kd}\tau_d$ and note that by (2.4) $\sum_k b_{\ell k} < \infty$. Indeed we have $\sum_\ell \sum_k b_{\ell k} \le \sum_\ell \sum_k \sum_d a_{\ell kd}\tau_d = \sum_d \tau_d$. So $n^{-1}\sum_m \max_{1-m\le k\le n-m} b_{\ell k}$ is bounded for all n by $\sum_k b_{\ell k}$, which is summable in ℓ, while Lemma 3.2 below shows that $n^{-1}\sum_m \max_{1-m\le k\le n-m} b_{\ell k} \to \max_k b_{\ell k}$ for each ℓ. It then follows from the Dominated Convergence Theorem that

$$\lim_{n\to\infty} P\{\boldsymbol{M}_n \le \boldsymbol{u}_n\} = \exp\left[-\sum_\ell \max_k \max_d a_{\ell kd}\tau_d\right] . \tag{3.5}$$

By (3.3) and (3.5), the limit of $\log P\{\boldsymbol{M}_n \le \boldsymbol{u}_n\}/n \log F(\boldsymbol{u}_n)$ is the desired result (3.2).

□

Lemma 3.2 *Let $b_k \ge 0$ for each $k = 0, \pm 1, \pm 2, \ldots$, and suppose $\sum_k b_k < \infty$. Then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{m=-\infty}^{\infty} \max_{1-m\le k\le n-m} b_k = \max_k b_k. \tag{3.6}$$

Proof. Suppose b_k is maximized when $k = K$ (not necessarily unique). Break the sum in (3.6) into three smaller sums:

$$\frac{1}{n}\sum_{m<1-K}\max_{1-m\leq k\leq n-m} b_k + \frac{1}{n}\sum_{1-K\leq m\leq n-K}\max_{1-m\leq k\leq n-m} b_k + \frac{1}{n}\sum_{n-K<m}\max_{1-m\leq k\leq n-m} b_k. \quad (3.7)$$

The middle sum is exactly b_K. It therefore suffices to show that each of the two outer sums tends to 0. The treatment of the two sums is the same, so we just give the argument for the first sum.

Break the sum up into two further terms, one for which $m < 1 - K - \sqrt{n}$, the other for $1 - K - \sqrt{n} \leq m < 1 - K$. The second of these is at most $b_K/\sqrt{n}$, so tends to 0 as $n \to \infty$. The first term is bounded by

$$\frac{1}{n}\sum_{m<1-K-\sqrt{n}}\ \sum_{1-m\leq k\leq n-m} b_k. \quad (3.8)$$

However for any given k the number of times b_k appears is 0 if $k < K + \sqrt{n}$, otherwise at most n. Therefore the sum (3.8) is bounded by $\sum_{k>K+\sqrt{n}} b_k$, and this tends to 0 by summability of b_k. □

The numerator and denominator of (3.2) are each of the form of a dependence function for multivariate extremes (Pickands 1981), and is therefore a convex function of $\boldsymbol{\tau} \in \mathcal{S}_D$,

$$\mathcal{S}_D = \{(y_1, \ldots, y_D) : y_i \geq 0,\ \sum y_i = 1\}\ .$$

We may restrict attention to the simplex $\mathcal{S}_D$ in view of the relation $\theta(c\boldsymbol{\tau}) = \theta(\boldsymbol{\tau})$ for all $c > 0$, as is of course obvious from (3.2). Thus a general representation is of the form

$$\theta(\boldsymbol{\tau}) = \frac{f_1(\boldsymbol{\tau})}{f_2(\boldsymbol{\tau})}\quad , \quad (3.9)$$

where f_1 and f_2 are dependence functions on $\mathcal{S}_D$ with $f_1 \leq f_2$. It is not clear whether any further restrictions on f_1 and f_2 are needed to ensure that (3.9) is a valid expression for the multivariate extremal index, but we have been unable to find any.

Consider, for example, the case $D = 2$. Then we may identify $\boldsymbol{\tau} = (\tau_1, \tau_2) = (\omega, 1-\omega)$ for $\omega \in [0,1]$, and define $\theta(\omega) = f_1(\omega)/f_2(\omega)$, where

$$\begin{aligned} f_1(\omega) &= \sum_\ell \max_k \max\{a_{\ell k1}\omega,\ a_{\ell k2}(1-\omega)\}\ , \\ f_2(\omega) &= \sum_\ell \sum_k \max\{a_{\ell k1}\omega,\ a_{\ell k2}(1-\omega)\}\ . \end{aligned} \quad (3.10)$$

Both of these are special cases of the general form

$$g(\omega) = \sum_j \max\{c_{j1}\omega,\ c_{j2}(1-\omega)\} \tag{3.11}$$

with coefficients $\{c_{j1}\}$, $\{c_{j2}\}$. Suppose the indices are restricted to $1 \le j \le J$ and ordered so that $\{d_j = c_{j2}/(c_{j1}+c_{j2})\}$ are increasing with j. Then we easily see that

$$g'(\omega) = -\sum_{i>j} c_{i2} + \sum_{i\le j} c_{i1} \ \text{ in } \ d_j < \omega < d_{j+1}\ ,$$

and that $g'(\omega)$ has a positive jump $c_{j1}+c_{j2}$ at each d_j. Together with $g(0) = \sum_j c_{j2}$, $g(1) = \sum_j c_{j1}$, these properties establish a 1–1 correspondence between any continuous piecewise linear convex function on $[0,1]$ and the coefficients $\{c_{j1}\}$, $\{c_{j2}\}$. Moreover, an arbitrary continuous convex function may be approximated arbitrarily closely in this way.

Consider the implications for f_1 and f_2 given by (3.10). If $\{a_{\ell k d}\}$ are non-zero only for $k \ge 0$, and monotonically decreasing in k, then we have the following: if $0 \le \omega_1 < \omega_2 \le 1$ then

$$f_2'(\omega_2) - f_2'(\omega_1) = \sum_\ell \sum_k (a_{\ell k1} + a_{\ell k2})\ , \tag{3.12}$$

$$f_1'(\omega_2) - f_1'(\omega_1) = \sum_\ell (a_{\ell 01} + a_{\ell 02})\ , \tag{3.13}$$

where in (3.12), the summation is over all pairs (ℓ, k) such that $\omega_1 < a_{\ell k2}/(a_{\ell k1}+a_{\ell k2}) < \omega_2$, and in (3.13), the summation is over all ℓ such that $\omega_1 < a_{\ell 02}/(a_{\ell 01}+a_{\ell 02}) < \omega_2$. The former summation includes the latter and so it follows that

$$f_2'(\omega_2) - f_2'(\omega_1) \ge f_1'(\omega_2) - f_1'(\omega_1) \tag{3.14}$$

whenever $\omega_2 > \omega_1$ and all four derivatives in (3.14) are defined. It is clear from the construction (3.11) that the converse is also true: if f_1 and f_2 are convex functions satisfying (3.14), then we can approximate each by a representation (3.10) in which $\{a_{\ell k d}\}$ is decreasing in $k \ge 0$ for each ℓ and d. Note, in particular, that in this case $f_2 - f_1$ must be convex.

In general, however, there is no reason to assume that the $\{a_{\ell k d}\}$ are restricted to $k \ge 0$ or satisfy any monotonicity relationship, and by relaxing these restrictions we can easily construct examples for which $f_2 - f_1$ is not convex. Our conjecture is that arbitrary dependence functions may appear in (3.10), and more generally (3.9), subject only to $f_1 \le f_2$.

4 Estimation

Suppose now we have a set of data from a D-dimensional stationary process, and that we are interested in estimating the extremal behavior of the sequence. In particular, we have the problem of estimating the multivariate extremal index. The results of Sections 2 and 3 imply that we might try to do this by assuming that the high-level exceedances of the process are described by an M_4 process. We now consider ways in which we might estimate the coefficients $\{a_{\ell k d}\}$ in (2.5). As in previous sections, we assume that the marginal distributions of the process are unit Fréchet. In practice, a separate estimation of the marginal distribution, followed by appropriate pointwise transformation, would usually be required to achieve this, but we shall not consider that aspect of the problem here.

Suppose we fix some sequence of (univariate) thresholds $\{u_n\}$, and suppose we monitor all exceedances of u_n in any of the D components. Also let $\{r_n\}$ denote a sequence of integers such that $r_n \to \infty$, $r_n/u_n \to 0$. A key result to estimation is that, for sufficiently large n, all the exceedances of u_n within a block of length r_n are with high probability caused by a single very large value of $Z_{\ell,j}$. We make this precise as follows:

Theorem 4.1 *Suppose $Y_{i,d}$ is defined by (2.5) and $\{u_n\}$ and $\{r_n\}$ are as just described. Let $I(r,u)$ denote the set of all (i,d) pairs, $1 \le i \le r$, $1 \le d \le D$, such that $Y_{id} > u$. Then, given that $I(r_n,u_n)$ is non-empty, the probability that there exists a single (L,J) pair such that*

$$Y_{i,d} = a_{L,i+J,d}\ Z_{L,J} \text{ for all } (i,d) \in\ I(r_n,u_n) \tag{4.1}$$

tends to 1 as $n \to \infty$.

Proof. First note the following elementary fact which we state without proof: if for each $n \ge 1$, $\{E_{k,n},\ k \in \mathcal{K}\}$ is a countable set of independent events such that $\lim_{n\to\infty} \sum_k P\{E_{k,n}\} = 0$, and if $F_n = \cup_k E_{k,n}$, $G_n = \cup_k \cup_{k'\ne k} (E_{k,n} \cap E_{k',n})$, then $P\{G_n \mid F_n\} \to 0$. In words, if at least one of the events $\{E_{n,k},\ k \in \mathcal{K}\}$ occurs, then with probability tending to 1 as $n \to \infty$, it is the only one.

Now let us apply this to the case where $\mathcal{K} = \{(\ell,j),\ \ell \ge 1,\ -\infty < j < \infty\}$ and $E_{\ell,j,n}$ is the event

$$\max_d a_{\ell,i+j,d} Z_{\ell,j} > u_n, \quad \text{some } i \in \{1,...,r_n\}. \tag{4.2}$$

If we define $c_{\ell,k} = \max_d a_{\ell,k,d}$ then the probability of the event $E_{\ell,j,n}$ is

$$1 - \exp\left\{-\frac{1}{u_n}\max_{1\le i\le r_n} c_{\ell,i+j}\right\} \approx \frac{1}{u_n}\max_{1\le i\le r_n} c_{\ell,i+j}. \tag{4.3}$$

However it follows from the proof of Theorem 3.1 that

$$\frac{1}{r_n}\sum_\ell\sum_j \max_{1\le i\le r_n} c_{\ell,i+j} \to \sum_\ell \max_j c_{\ell,j} \le \sum_\ell\sum_j c_{\ell,j} \le D. \tag{4.4}$$

But $u_n/r_n \to \infty$ and so

$$\sum_\ell\sum_j P\{E_{\ell,j,n}\} = O\left(\frac{r_n}{u_n}\right) \to 0. \tag{4.5}$$

The result then follows from the elementary fact stated at the beginning of this proof.

□

This result gives us an intuitive feeling for how to estimate processes of this type. Our point of view is that the original observed process $\boldsymbol{X}$ is in the domain of max-attraction of a max-stable process $\boldsymbol{Y}$, and that the process $\boldsymbol{Y}$ may be closely approximated by an M_4 process. Suppose, first, the process is in fact a univariate simple moving maximum process, i.e. $Y_i = \max_k a_k Z_{i-k}$ for some sequence $\{a_k\}$. Suppose without loss of generality that the maximum of a_k is a_0. Then in each block of length r_n, with probability very close to 1 we will observe the following: either there are no exceedances of u_n, or the exceedances form a characteristic signature pattern with $Y_{j+k} = (a_k/a_0)Y_j$ for all sufficiently small indices k. In this way, it is in principle possible to read off any number of a_k/a_0 values from a single large excursion. We can then determine a_0 and hence each of the a_k from the condition $\sum_k a_k = 1$.

Of course, in practice we would not expect to see a single signature pattern near each high-level exceedance, and this explains intuitively why the simple moving maximum process does not appear realistic for practical applications. However, the M_3 process (i.e. maxima of moving maxima in one dimension) is much more realistic. In this case, there are a countable number of "signatures", indexed by ℓ, where the frequency of the ℓ'th signature is $\sum_k a_{k\ell}$. In practice we would not try to make the signatures identical to the observed excursion patterns, but would use some version of cluster analysis to classify the observed excursions into a finite number of clusters. Within each cluster, we would then pick out a characteristic shape of the cluster, and identify this with the $\{a_{\ell k}/a_{\ell 0}\}$ values for the ℓ'th cluster. Exactly how best to do this, and how many clusters to adopt, is a topic for future research.

The extension of this to the M_4 process is as follows. Fix a high threshold u_n, and divide the process into blocks of length r_n, where $r_n = o(u_n)$. An "excursion" is said to occur whenever $Y_{id} > u_n$ for at least one (i, d) pair within a block. According to Theorem 4.1, each excursion follows very closely one of the signature patterns

$$Y_{J+k,d} = a_{Lkd} Z_{L,J}, \;\; k = 0, \pm 1, \pm 2, ..., \;\; d = 1, ..., D \tag{4.6}$$

for some (L, J). By observing the ratio of each excursion to its maximum, we obtain the values of $\{a_{\ell kd} / \max_{k,d} a_{\ell kd}\}$ for each ℓ. Moreover, the relative frequency of the ℓ'th signature pattern is proportional to $\sum_k \max_d a_{\ell kd}$. By classifying the observed excursions into a finite number of clusters, and associating with each cluster a characteristic signature pattern, we obtain an approximation to the whole process as an M_4 process. We can then use the coefficients in this approximation to compute the multivariate extremal index via (3.2).

The approach that has just been outlined is "model-based": even though it need not be reduced to any finite-parameter model with a fixed number of parameters, it does rely on the model (2.3), which in most cases is only an approximation to the true model under study.

An alternative approach is to try to exploit directly the formula (3.9), in which $\theta(\boldsymbol{\tau})$ was expressed as a ratio of two multivariate extreme value dependence functions. A key point here is that, by Prop. 2.1, $\theta(\boldsymbol{\tau})$ is the univariate extremal index of a sequence $\{V_i(\boldsymbol{\tau})\}$, so for fixed $\boldsymbol{\tau}$ it can be estimated by univariate methods such as those proposed by Smith and Weissman (1994). This can be repeated on a fine grid of values $\boldsymbol{\tau}_j \in S_D$, thus obtaining a discrete approximation to the function $\theta(\boldsymbol{\tau})$.

For clarity, let us demonstrate these ideas by means of a bivariate process. So, suppose $\boldsymbol{X}_i = (X_{i1}, X_{i2})$, and define $V_i(\omega) = \max\{\omega X_{i1}, (1-\omega) X_{i2}\}$ for $0 \le \omega \le 1$ and $i = 1, \ldots, n$. Then, since $\boldsymbol{X}_i$ is max-stable,

$$P\{V_i(\omega) \le u\} = F\left(\frac{u}{\omega}\ ,\ \frac{u}{1-\omega}\right) = \exp\left[-\frac{1}{u}\ A(\omega)\right] \qquad (u > 0)\ ,$$

where $A(\omega)$ is the dependence function, introduced by Pickands (1981). The function $A(\omega)$ is convex on $[0, 1]$, $A(0) = A(1) = 1$, bounded above by 1 and below by $\max\{\omega, 1-\omega\}$.

For fixed ω, define now $M^V_{i,j} = \max\{V_{i+1}(\omega), \ldots, V_j(\omega)\}$ $(0 \le i < j \le n)$. For r and u large, since $\theta(\omega)$ is the extremal index of $\{V_i(\omega)\}$,

$$P\{M^V_{0,r} \le u\} \approx \exp\left[-\frac{r}{u}\ A(\omega)\theta(\omega)\right]\ .$$

In fact $M_{0,r}^V$ is itself asymptotically max-stable, so that $P\{M_{0,r}^V \le u\} = \exp\left[-\frac{r}{u}A_r(\omega)\right]$ where $A_r(\omega) \to A(\omega)\theta(\omega)$ as $r \to \infty$.

Let u_n be a threshold level and define $N_n(\omega) = \sum_{i=1}^n \mathbf{1}\{V_i(\omega) > u_n\}$, the number of exceedances over the level u_n. Fix $r = r_n$; let $k = [n/r]$ and let $Z_n(\omega) = \sum_{j=1}^k \mathbf{1}\{M_{(j-1)r,jr}^V > u_n\}$ be the number of blocks of size r which have at least one exceedance.

Since $EN_n(\omega) = n\left[1 - \exp\left\{-\frac{1}{u_n}A(\omega)\right\}\right] \sim \frac{n}{u_n}A(\omega)$ and $EZ_n(\omega) \sim \frac{n}{u_n}A(\omega)\theta(\omega)$, the ratio $\hat{\theta}(\omega) = Z_n(\omega)/N_n(\omega)$ is a reasonable estimator for $\theta(\omega)$. For this estimator to be consistent, one has to choose $r = r_n$ properly and $u_n \to \infty$ such that $n/u_n \to \infty$, so that $Z_n(\omega)$ and $N_n(\omega)$ will be quite large. This is the "blocks approach" to extremal index estimation considered by Smith and Weissman (1994), and they gave conditions to ensure its consistency as an estimator. However they also discussed an alternative *runs approach*, which they ultimately argued to be superior to the blocks approach. In the runs approach, the definition of $N_n(\omega)$ is the same, but the exceedances over the threshold u_n are grouped into clusters, where a new cluster is deemed to begin whenever there is a run of r_n consecutive values below the threshold. Then $Z_n(\omega)$ is the number of clusters in which there is at least one exceedance over the threshold u_n.

Using either the blocks or the runs approach, we can estimate $f_2(\omega) = A(\omega)$ by $\hat{f}_2(\omega) = u_n N_n(\omega)/n$, and $f_1(\omega) = A(\omega)\theta(\omega)$ by $\hat{f}_1(\omega) = u_n Z_n(\omega)/n$. The ratio of the two, $\hat{f}_1(\omega)/\hat{f}_2(\omega)$, is a crude estimator of $\theta(\omega)$. In Section 6 we shall use the notation $\hat{f}_{1B}, \hat{f}_{1R}$ for the blocks and runs, respectively.

For the problem of estimating a bivariate dependence function, Pickands (1981) suggested that the estimate should be modified to as to ensure that it was a convex function of ω. Pickands' proposal was to evaluate an estimator $\hat{A}(\omega)$ at a finite number of points $\{\omega_j\}$, and then to define a new estimator $\tilde{A}(\omega)$ as the greatest convex minorant of $\hat{A}(\omega)$. Smith, Tawn and Yuen (1990) proposed various more sophisticated methods of making $\hat{A}$ a convex function. We follow Pickands here. Thus, $\hat{f}_1(\omega)$ and $\hat{f}_2(\omega)$ are evaluated at a finite number of values of $\omega = \omega_j = j/m,\ j = 0, 1, ..., m$, and are then replaced by their greatest convex minorants $\tilde{f}_1(\omega)$ and $\tilde{f}_2(\omega)$. We then define $\tilde{\theta}(\omega) = \tilde{f}_1(\omega)/\tilde{f}_2(\omega)$.

Examples of this method will be presented in Section 6.

5 Examples

In this section, we show two bivariate examples. The first one is simple, with $\ell = 1$ The second example is an M_4 with $\ell = 1, 2$.

Example 5.1. Let $\{\xi_i : i \geq 1\}$ be a unit-Fréchet iid sequence and define $X_{i1} = \xi_i$ and $X_{i2} = \max(\xi_i, \xi_{i+1})/2$. Hence the random vector $\boldsymbol{X}_i = (X_{i1}, X_{i2})$ is max-stable with unit-Fréchet margins. Specifically for $u^{-1} = x_1^{-1} + x_2^{-1}$ and $\omega = u/x_1$ one has

$$\begin{aligned}
P\{X_{i1} \leq x_1, X_{i2} \leq x_2\} &= P\{\xi_i \leq \min(x_1, 2x_2),\ \xi_{i+1} \leq 2x_2\} \\
&= \exp\left[-\frac{1}{u}\{\max(\omega, (1-\omega)/2) + (1-\omega)/2\}\right] \\
&= \exp\left[-\frac{1}{u}A(\omega)\right] \\
&= P\{\max(\omega X_{i1}, (1-\omega)X_{i2}) \leq u\} \\
&= P\{V_i(\omega) \leq u\}\ , \quad 0 \leq \omega \leq 1,\ u > 0\ . \qquad (5.1)
\end{aligned}$$

The function $f_2(\omega) = A(\omega)$ on $[0,1]$ is convex, as discussed before. We note that

$$P\{M^V_{0,n} \leq u_n\} = P\{M^\xi_{0,n} \leq \min(u/\omega,\ 2u/(1-\omega))\}\ P\{\xi_{n+1} \leq 2u_n/(1-\omega)\}$$

$$= \exp\left[-\frac{n}{u_n}\cdot\max(\omega, (1-\omega)/2)\right]\cdot\exp[-(1-\omega)/(2u_n)]\ . \qquad (5.2)$$

If we let $u_n \to \infty$ as $n \to \infty$, the second factor of (5.2) tends to 1. So we can identify the function $f_1(\omega)$ as $\max(\omega, (1-\omega)/2)$ and hence the extremal index function $\theta(\omega)$ is given by

$$\theta(\omega) = \frac{f_1(\omega)}{f_2(\omega)} = \begin{cases} 1/2 & 0 \leq \omega \leq 1/3 \\ 2\omega/(1+\omega) & 1/3 \leq \omega \leq 1 \end{cases} \qquad (5.3)$$

(see Figure 1). It is easy to see that for our stationary sequence, clusters of exceedances (at least in one coordinate) are either of size 1 or of size 2. Suppose $(n/\tau_1, n/\tau_2)$ are the threshold levels, i.e., $\omega = \tau_1/(\tau_1+\tau_2)$ and $u_n = n/(\tau_1+\tau_2)$. If $\xi_i > n/\tau_1$ but $\xi_i < 2n/\tau_2$, then a cluster of size 1 is observed. If $\xi_i > \max(n/\tau_1, 2n/\tau_2)$ then the cluster is of size 2.

Let C denote the cluster size. Then if $\tau_1 < \tau_2/2$ (i.e., $0 \leq \omega \leq 1/3$) then $P\{C = 2|C \geq 1\} = 1$; if $\tau_1 > \tau_2/2$, then

$$P\{C = 1|C \geq 1\} = \lim_{n\to\infty} \frac{P\{\frac{n}{\tau_1} < \xi_i \leq \frac{2n}{\tau_2}\}}{P\{\frac{n}{\tau_1} < \xi_i\}} = \frac{\tau_1 - \tau_2/2}{\tau_1} = \frac{3\omega - 1}{2\omega}\ .$$

Hence, asymptotically

$$E(C|C \geq 1) = \begin{cases} 2 & \text{if } 0 \leq \omega \leq 1/3 \\ \frac{\omega+1}{2\omega} & \text{if } 1/3 \leq \omega \leq 1 \,, \end{cases} \tag{5.4}$$

and the reciprocal of (5.4) agrees with (5.3).

One realization of the process is shown in Figure 2. The peaks form one signature pattern.

Example 5.2. Let $\{Z_{\ell,i} : \ell = 1, 2; i \geq -19\}$ be two sequences of independent unit-Fréchet random variables. This example is a bivariate maxima of moving maxima as defined in (2.3).

For each $\ell = 1, 2$ and $d = 1, 2$, we define a sequence of coefficients $\{a_{\ell k d} : k = 0, \pm 1, \cdots, \pm 20\}$ as follows:

$$\begin{aligned} a_{1k1} &= 0.0882933 \times 0.7^{|k|}; & \textstyle\sum_k a_{1k1} &= 1/2. \\ a_{2k1} &= (1/6) \times 0.5^{|k|}; & \textstyle\sum_k a_{2k1} &= 1/2. \\ a_{1k2} &= 0.0396511 \times 0.9^{|k|}; & \textstyle\sum_k a_{1k2} &= 2/3. \\ a_{2k2} &= (1/9) \times 0.5^{|k|}; & \textstyle\sum_k a_{2k2} &= 1/3. \end{aligned}$$

The process $\{\boldsymbol{X}_i = (X_{i1}, X_{i2}) : i \geq 1\}$ is given by

$$\begin{aligned} X_{i1} &= \max\left(\max_k a_{1k1} Z_{1,i-k},\, \max_k a_{2k1} Z_{2,i-k}\right) \\ X_{i2} &= \max\left(\max_k a_{1k2} Z_{1,i-k},\, \max_k a_{2k2} Z_{2,i-k}\right). \end{aligned}$$

The extremal index function is $\theta(\omega) = f_1(\omega)/f_2(\omega), \omega \epsilon [0, 1]$, where f_1, f_2 are dependence functions given by (3.10)

$$\begin{aligned} f_1(\omega) &= \max_k \max\left(a_{1k1}\omega, a_{1k2}(1-\omega)\right) + \max_k \max\left(a_{2k1}\omega, a_{2k2}(1-\omega)\right) \\ f_2(\omega) &= \textstyle\sum_k \max\left(a_{1k1}\omega, a_{1k2}(1-\omega)\right) + \sum_k \max\left(a_{2k1}\omega, a_{2k2}(1-\omega)\right). \end{aligned}$$

(see Figure 3). One realization of $\{\boldsymbol{X}_i\}$ is shown in Figure 4 (first 500 observations, solid line for X_{i1}, dotted line for X_{i2}). The peaks do indeed form two different signature patterns. In particular, the peaks of $\{X_{i1}\}$ and $\{X_{i2}\}$ occur simultaneously.

6 Simulations

We have run many simulations for both models, but report the results for Example 5.2. In Figure 5 we show estimates of $\theta(\omega)$ evaluated at $\omega_j = j/m (j = 0, 1, \cdots, m)$ with $m = 20$, sample size $n = 10^5$, threshold level $u = 99.5$ (the 99th percentile of the unit-Fréchet distribution) and block and run size $r = 40$. The blocks-estimate $\hat{\theta}_B(\omega) = \hat{f}_{1B}(\omega)/\hat{f}_2(\omega)$ is represented by points and the corrected-for-convexity version $\tilde{\theta}_B = \tilde{f}_{1B}(\omega)/\tilde{f}_2(\omega)$ by a dashdot line. Similarly, the runs-estimate $\hat{\theta}_R(\omega) = \hat{f}_{1R}(\omega)/\hat{f}_2(\omega)$ is represented by points and the corrected-for-convexity version $\tilde{\theta}_R(\omega) = \tilde{f}_{1R}(\omega)\tilde{f}_2(\omega)$ by a dotted line. In this particular realization, the blocks-estimate is an over-estimate of θ and the runs-estimate is an under-estimate. A similar picture is seen in Figure 6, where the threshold level $u = 39.5$ is the 97.5th percentile; here, the runs-estimate is clearly inferior with respect to the blocks-estimate.

The main theme of this paper is not the performance of the blocks and runs estimators of the extremal index. However, we have run several simulations to check sensitivity to block (or run) size r and to threshold level u. A sample of size $n = 10^4$ was repeated 20 times. For each sample, the square root of mean square error (sqrmse) was computed, and the average over 20 samples is given in Table 1.

For $u = 39.5$ (97.5th percentile), the runs estimator performs best when $r = 10$ or 20, while the blocks estimator is at its best when $r = 40$ or 60 (with similar sqrmse). For $u = 99.5$, the performance is much poorer. A possible explanation is the lower number of exceedances. Indeed, further simulations show that for $n = 10^5$, $u = 99.5$ is a better choice than $u = 39.5$. Due to the enormous computing time required, we did not obtain Table 1 for $n = 10^5$.

7 Discussion

In this paper, we have provided conditions under which the extremal index of a multivariate stationary time series may be calculated in terms of an equivalent max-stable process (Theorem 2.3), and we have also provided a representation for a max-stable process as a limit of M_4 processes (Theorem 2.6). Our results on estimation are at the moment less well developed, but we have outlined two approaches, one based directly on the representation as an M_4 process, and the other based on estimating the numerator and denominator separately in (3.9). Examples have been provided of the latter approach. There are clearly many possibilities for further research developing alternative estimation procedures.

Acknowledgement. This research was supported by grant No. 9200008 from the United States-Israel Binational Science Foundation (BSF), Jerusalem, Israel. The authors would like to thank the BSF for the generous support.

References

Deheuvels, P. (1978). Caractérisation complète des lois extrèmes multivariées et de la convergence des types extrèmes. *Publ. Inst. Statist. Univ. Paris* **23**, 1–36.

Deheuvels, P. (1983). Point processes and multivariate extreme values. *J. Multivar. Anal.* **13**, 257–272.

Deheuvels, P. (1985). Point processes and multivariate extreme values II. In: *Multivariate Analysis* VI (ed.: P.R. Krishnaiah). North Holland, Amsterdam, 145–164.

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

de Haan, L. and Resnick, I. (1977). Limit theory for multivariate sample extremes. *Z.W.G.* **40**, 317–337.

Galambos, J. (1987). *The Asymptotic Theory of Extreme Order Statistics* (second edition). Krieger: Melbourne, Florida.

Hsing, T. (1989). Extreme value theory for multivariate stationary sequences. *J. Mult. Anal.* **29**, 274–291.

Hsing, T. (1993). Extremal index estimation for a weakly dependent stationary sequence. *The Annals of Statist.* **21**, 2043–2071.

Hsing, T., Hüsler, J. and Leadbetter, M.R. (1988). On the exceedance point process for a stationary sequence. *Probab. Theory Related Fields* **78**, 97–112.

Leadbetter, M.R. (1983). Extremes and local dependence in a stationary sequence. *Z.W.G.* **65**, 291–306.

Leadbetter, M.R., Lindgren, G. and Rootzén, H. (1983). *Extremes and Related Properties of Random Sequences and Processes.* Springer, Berlin.

Leadbetter, M.R., Weissman, I., de Haan, L. and Rootzén, H. (1989). On clustering of high values in statistically stationary series. In: *Proc. 4th Int. Meet. Statistical Climatology* (ed.: J. Sanson). Wellington: New Zealand Meteorological Service.

Nandagopalan, S. (1990). *Multivariate extremes and the estimation of the extremal index.* Ph.D. dissertation, Dept. of Statistics, University of North Carolina, Chapel Hill.

Nandagopalan, S. (1994). On the multivariate extremal index. *J. of Research, National Inst. of Standards and Technology* **99**, 543–550.

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

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

Pickands, J. III (1981). Multivariate extreme value distributions. *Proc. 43rd Session, ISI*, Buenos Aires, 859–878.

Resnick, S. (1987). *Extreme Values, Point Processes and Regular Variation.* Springer Verlag, New York.

Smith, R.L. (1985). Statistics of extreme values. *Proc. 45th Session, ISI*, Amsterdam, Book IV, Paper 26.1.

Smith, R.L. (1992). The extremal index for a Markov chain. *J. Applied Probability* **29**, 37-45.

Smith, R.L., Tawn, J.A. and Yuen, H.K. (1990). Statistics of multivariate extremes. *Inter. Statist. Rev.* **58**, 393–442.

Smith, R.L., Tawn, J.A. and Coles, S.G. (1997). Markov chain models for threshold exceedances. To appear, *Biometrika.*

Smith, R.L. and Weissman, I. (1994). Estimating the extremal index. *J. Royal Statist. Soc. B.* **56**, 515–528.

Weissman, I. (1994). On the extremal index of stationary sequences. Technical report, presented at the Conference on Multivariate Extreme Value Estimation with Applications to Economics and Finance, Erasmus University, Rotterdam.

Weissman, I. and Cohen, U. (1995). The extremal index and clustering of high values for derived stationary sequences. *J. of Appl. Probab.* **32**, 972–981.

Yun, S. (1994). *Extreme values in higher-order Markov chains.* [Check exact title] PhD Thesis, Department of Statistics, University of North Carolina, Chapel Hill.