12 September 2012

REPRESENTATIONS OF MAX-STABLE PROCESSES BASED ON SINGLE EXTREME EVENTS

SEBASTIAN ENGELKE,* *Georg-August-Universität Göttingen*
ALEXANDER MALINOWSKI,** *Georg-August-Universität Göttingen*
MARCO OESTING,*** *Universität Mannheim*
MARTIN SCHLATHER,**** *Universität Mannheim*

Abstract

This paper provides the basis for new methods of inference for max-stable processes ξ on general spaces that admit a certain incremental representation, which, in important cases, has a much simpler structure than the max-stable process itself. A corresponding peaks-over-threshold approach will incorporate all single events that are extreme in some sense and will therefore rely on a substantially larger amount of data in comparison to estimation procedures based on block maxima.
Conditioning a process η in the max-domain of attraction of ξ on being *extremal*, several convergence results for the increments of η are proved. In a similar way, the shape functions of mixed moving maxima (M3) processes can be extracted from suitably conditioned single events η. Connecting the two approaches, transformation formulae for processes that admit both an incremental and an M3 representation are identified.

Keywords: extreme value statistics; incremental representation; max-stable process; mixed moving maxima; peaks-over-threshold; weak convergence on function space

2010 Mathematics Subject Classification: Primary 60G70
Secondary 62G32; 62E20

1. Introduction

The joint extremal behavior at multiple locations of some random process $\{\eta(t) : t \in T\}$, T an arbitrary index set, can be captured via its limiting *max-stable process*, assuming the latter exists and is non-trivial everywhere. Then, for independent copies η_i of η, $i \in \mathbb{N}$, the functions $b_n : T \to \mathbb{R}$, $c_n : T \to (0, \infty)$

* Postal address: Institut für Mathematische Stochastik, Goldschmidtstr. 7, 37077 Göttingen, Germany, sengelk@uni-goettingen.de
** Postal address: Institut für Mathematische Stochastik, Goldschmidtstr. 7, 37077 Göttingen, Germany, malinows@math.uni-goettingen.de
*** Postal address: Institut für Mathematik, Universität Mannheim, A 5, 6, 68131 Mannheim, Germany, oesting@math.uni-mannheim.de
**** Postal address: Institut für Mathematik, Universität Mannheim, A 5, 6, 68131 Mannheim, Germany, schlather@math.uni-mannheim.de

can be chosen such that the convergence

$$\xi(t) = \lim_{n\to\infty} c_n(t)\Big(\max_{i=1}^{n} \eta_i(t) - b_n(t)\Big), \quad t \in T, \tag{1}$$

holds in the sense of finite-dimensional distributions. The process ξ is said to be *max-stable* and η is in its max-domain of attraction (MDA). The theory of max-stable processes is mainly concerned with the dependence structure while the marginals are usually assumed to be known. Even for finite-dimensional max-stable distributions, the space of possible dependence structures is uncountably infinite-dimensional and parametric models are required to find a balance between flexibility and analytical tractability [7, 21].

A general construction principle for max-stable processes was provided by [6, 26]: Let $\sum_{i\in\mathbb{N}} \delta_{(U_i,S_i)}$ be a Poisson point process (PPP) on $(0,\infty)\times\mathcal{S}$ with intensity measure $u^{-2}\mathrm{d}u\cdot\nu(\mathrm{d}s)$, where $(\mathcal{S},\mathfrak{S})$ is an arbitrary measurable space and ν a positive measure on $\mathcal{S}$. Further, let $f:\mathcal{S}\times T\to[0,\infty)$ be a non-negative function with $\int_{\mathcal{S}} f(s,t)\nu(\mathrm{d}s) = 1$ for all $t\in T$. Then the process

$$\xi(t) = \max_{i\in\mathbb{N}} U_i f(S_i,t), \quad t\in T, \tag{2}$$

is max-stable and has standard Fréchet margins with distribution function $\exp(-1/x)$ for $x \geq 0$. In this paper, we restrict to two specific choices for f and $(\mathcal{S},\mathfrak{S},\nu)$ and consider processes that admit one of the resulting representations. First, let $\{W(t): t\in T\}$ be a non-negative stochastic process with $\mathrm{E}\,W(t)=1$, $t\in T$, and $W(t_0)=1$ a.s. for some point $t_0\in T$. The latter condition means that $W(t)$ simply describes the multiplicative increment of W w.r.t. the location t_0. For $(\mathcal{S},\mathfrak{S},\nu)$ being the canonical probability space for the sample paths of W and with $f(w,t)=w(t)$, $w\in\mathcal{S}$, $t\in T$, we refer to

$$\xi(t) = \max_{i\in\mathbb{N}} U_i W_i(t), \quad t\in T, \tag{3}$$

as the *incremental representation* of ξ, where $\{W_i\}_{i\in\mathbb{N}}$ are independent copies of W. Since T is an arbitrary index set, the above definition covers multivariate extreme value distributions, i.e. $T=\{t_1,\dots,t_k\}$, as well as max-stable random fields, i.e. $T=\mathbb{R}^d$.
For the second specification, let $\{F(t):\ t\in\mathbb{R}^d\}$ be a stochastic process with sample paths in the space $C(\mathbb{R}^d)$ of non-negative continuous functions, such that

$$\mathrm{E}\int_{\mathbb{R}^d} F(t)\mathrm{d}t = 1. \tag{4}$$

With $S_i=(T_i,F_i)$, $i\in\mathbb{N}$, in $\mathcal{S}=\mathbb{R}^d\times C(\mathbb{R}^d)$, intensity measure $\nu(\mathrm{d}t\times\mathrm{d}g) = \mathrm{d}t\,\mathrm{P}_F(\mathrm{d}g)$ and $f((t,g),s)=g(s-t)$, $(t,g)\in\mathcal{S}$, we obtain the class of *mixed moving maxima (M3) processes*

$$\xi(t) = \max_{i\in\mathbb{N}} U_i F_i(t-T_i), \quad t\in\mathbb{R}^d. \tag{5}$$

These processes are max-stable and stationary on $\mathbb{R}^d$ (see for instance [27]). The function F is called *shape function of* ξ and can also be deterministic (e.g., in case of the Smith process). In Smith's "rainfall-storm" interpretation [26], U_i and T_i are the strength and center point of the ith storm, respectively, and $U_i F_i(t - T_i)$ represents the corresponding amount of rainfall at location t. In this case, $\xi(t)$ is the process of extremal precipitation.

When i.i.d. realizations $\eta_1, \ldots, \eta_n$ of η in the MDA of a max-stable process ξ are observed, a classical approach for parametric inference on ξ is based on generating (approximate) realizations of ξ out of the data $\eta_1, \ldots, \eta_n$ via componentwise block maxima and applying maximum likelihood (ML) estimation afterwards. A clear drawback of this method is that it ignores all information on large values that is contained in the order statistics below the within-block maximum. Further, ML estimation needs to evaluate the multivariate densities while for many max-stable models only the bivariate densities are known in closed form. Thus, composite likelihood approaches have been proposed [20, 5]. In univariate extreme-value theory, the second standard procedure estimates parameters by fitting a certain PPP to the *peaks-over-thresholds* (POT), i.e., to the empirical process of exceedances over a certain critical value [18, 10]. Also in the multivariate framework we can expect to profit from using all extremal data via generalized POT methods instead of aggregated data. In contrast to the ML approach, in this paper, we assume that ξ admits one of the two representations (3) and (5) and we aim at extracting realizations of the processes W and F, respectively, from *single extreme events*. Here, the specification of a single extreme event will depend on the respective representation.
In [12], this concept is applied to derive estimators for the class of Brown-Resnick processes [4, 17], which have the form (3) by construction. With $a(n)$ being a sequence of positive numbers with $\lim_{n\to\infty} a(n) = \infty$, the convergence in distribution

$$\left(\frac{\eta(t_1)}{\eta(t_0)}, \ldots, \frac{\eta(t_k)}{\eta(t_0)} \;\middle|\; \eta(t_0) > a(n) \right) \xrightarrow{\mathrm{D}} \big(W(t_1), \ldots, W(t_k)\big), \tag{6}$$

$t_0, t_1, \ldots, t_k \in T$, $k \in \mathbb{N}$, is established for η being in the MDA of a Brown-Resnick process and with W being the corresponding log-Gaussian random field. A similar approach exists in the theory of homogeneous discrete-time Markov chains. For instance, [25] and [9] investigate the behavior of a Markov chain $\{M(t) : t \in \mathbb{Z}\}$ conditional on the event that $M(0)$ is large. The resulting extremal process is coined the tail chain and turns out to be Markovian again. In this paper, the convergence result (6) is generalized in different aspects. Arbitrary non-negative processes $\{W(t) : t \in T\}$ with $\mathrm{E}\, W(t) = 1$, $t \in T$, are considered, and convergence of the conditional increments of η in the sense of finite-dimensional distributions as well as weak convergence in continuous function spaces is shown (Theorems 2.1 and 2.2). Moreover, in Section 3, similar results are established for M3 processes (5) by considering realizations of η around their (local) maxima. Since one and the same max-stable process ξ might admit both representations (3) and (5) we provide formulae for switching between them in Section 4. Section 5 gives an exemplary outlook on how our results can be applied for statistical inference.

2. Incremental representation

Throughout this section, we suppose that $\{\xi(t):\ t\in T\}$, where T is an arbitrary index set, is normalized to standard Fréchet margins and admits a representation

$$\xi(t)=\max_{i\in\mathbb{N}} U_i V_i(t),\quad t\in T, \tag{7}$$

where $\sum_{i\in\mathbb{N}}\delta_{U_i}$ is a PPP on $(0,\infty)$ with intensity $u^{-2}du$, which we call *Fréchet point process* in the following. The $\{V_i\}_{i\in\mathbb{N}}$ are independent copies of a non-negative stochastic process $\{V(t):\ t\in T\}$ with $\mathrm{E}\,V(t)=1$, $t\in T$. Note that (7) is slightly less restrictive than the representation (3) in that we do not require that $V(t_0)=1$ a.s. for some $t_0\in T$. For any fixed $t_0\in T$, we have

$$\xi(t)\overset{\mathrm{D}}{=}\max_{i\in\mathbb{N}} U_i\left(\mathbf{1}_{P_i=0}V_i^{(1)}(t)+\mathbf{1}_{P_i=1}V_i^{(2)}(t)\right),\quad t\in T, \tag{8}$$

where $\{P_i\}_{i\in\mathbb{N}}$ are i.i.d. Bernoulli variables with parameter $p=\mathrm{P}(V(t_0)=0)$ and the $V_i^{(1)}$ and $V_i^{(2)}$ are independent copies of the process $\{V(t):\ t\in T\}$, conditioned on the events $\{V(t_0)>0\}$ and $\{V(t_0)=0\}$, respectively.

Note that for $k\in\mathbb{N}$, $t_0,\dots,t_k\in T$, the vector $\Xi=(\xi(t_0),\dots,\xi(t_k))$ follows a $(k+1)$-variate extreme-value distribution and its distribution function G can therefore be written as

$$G(\mathbf{x})=\exp(-\mu([\mathbf{0},\mathbf{x}]^C)),\quad \mathbf{x}\in\mathbb{R}^{k+1}, \tag{9}$$

where μ is a measure on $E=[0,\infty)^{k+1}\setminus\{\mathbf{0}\}$, the so-called *exponent measure* of G [21, Prop. 5.8], and $[\mathbf{0},\mathbf{x}]^C=E\setminus[\mathbf{0},\mathbf{x}]$.

The following convergence result provides the theoretical foundation for statistical inference based on the incremental process V.

Theorem 2.1. *Let $\{\eta(t):\ t\in T\}$ be non-negative and in the MDA of some max-stable process ξ that admits a representation (7) and suppose that η is normalized such that (1) holds with $c_n(t)=1/n$ and $b_n(t)=0$ for $n\in\mathbb{N}$ and $t\in T$. Let $a(n)\to\infty$ as $n\to\infty$. For $k\in\mathbb{N}$ and $t_0,\dots,t_k\in T$ we have the convergence in distribution on $\mathbb{R}^{k+1}$*

$$\left(\frac{\eta(t_0)}{a(n)},\frac{\eta(t_1)}{\eta(t_0)},\dots,\frac{\eta(t_k)}{\eta(t_0)}\,\middle|\,\eta(t_0)>a(n)\right)\overset{\mathrm{D}}{\to}\left(Z,\Delta\tilde{\mathbf{V}}^{(1)}\right),\quad n\to\infty,$$

where the distribution of $\Delta\tilde{\mathbf{V}}^{(1)}$ is given by

$$\mathrm{P}(\Delta\tilde{\mathbf{V}}^{(1)}\in d\mathbf{z})=(1-p)\,\mathrm{P}(\Delta\mathbf{V}^{(1)}\in d\mathbf{z})\,\mathrm{E}\big(V^{(1)}(t_0)\big|\Delta\mathbf{V}^{(1)}=\mathbf{z}\big),\quad \mathbf{z}\geq\mathbf{0}. \tag{10}$$

Here, $\Delta \mathbf{V}^{(1)}$ *denotes the vector of increments* $\left(\frac{V^{(1)}(t_1)}{V^{(1)}(t_0)}, \dots, \frac{V^{(1)}(t_k)}{V^{(1)}(t_0)}\right)$ *with respect to* t_0*, and* Z *is an independent Pareto variable.*

Remark 2.1. Note that any process η that satisfies the convergence in (1) for a process ξ with standard Fréchet margins can be normalized such that the norming functions in (1) become $c_n(t) = 1/n$ and $b_n(t) = 0$, $n \in \mathbb{N}$, $t \in T$ [21, Prop. 5.10].

Proof of Theorem 2.1. For $\mathbf{X} = (\eta(t_0), \dots, \eta(t_k))$, which is in the MDA of the random vector $\Xi = (\xi(t_0), \dots, \xi(t_k))$, it follows from [21, Prop. 5.17] that

$$\lim_{m\to\infty} m\,\mathrm{P}(\mathbf{X}/m \in B) = \mu(B), \tag{11}$$

for all elements B of the Borel σ-algebra $\mathcal{B}(E)$ of E bounded away from $\{\mathbf{0}\}$ with $\mu(\partial B) = 0$, where μ is defined by (9). For $s_0 > 0$ and $\mathbf{s} = (s_1, \dots, s_k) \in [0,\infty)^k$, we consider the sets $A_{s_0} = (s_0, \infty) \times [0,\infty)^k$, $A = A_1$ and $B_{\mathbf{s}} = \{\mathbf{x} \in [0,\infty)^{k+1} : (x^{(1)}, \dots, x^{(k)}) \leq x^{(0)}\mathbf{s}\}$ for $\mathbf{s}$ satisfying $\mathrm{P}(\Delta\tilde{\mathbf{V}}^{(1)} \in \partial[\mathbf{0}, \mathbf{s}]) = 0$. Then

$$\big\{\eta(t_0) > s_0 a(n),\ \big(\eta(t_1)/\eta(t_0), \dots, \eta(t_k)/\eta(t_0)\big) \leq \mathbf{s}\big\} = \{\mathbf{X}/a(n) \in B_{\mathbf{s}} \cap A_{s_0}\},$$

since $B_{\mathbf{s}}$ is invariant under multiplication, i.e., $B_{\mathbf{s}} = cB_{\mathbf{s}}$ for any $c > 0$. Thus, we obtain

$$\begin{aligned}
&\mathrm{P}\left(\eta(t_0) > s_0 a(n),\ (\eta(t_1)/\eta(t_0), \dots, \eta(t_k)/\eta(t_0)) \leq \mathbf{s} \,\middle|\, \eta(t_0) > a(n)\right) \\
&\quad = \frac{a(n)\,\mathrm{P}(\mathbf{X}/a(n) \in B_{\mathbf{s}} \cap A \cap A_{s_0})}{a(n)\,\mathrm{P}(\mathbf{X}/a(n) \in A)} \\
&\quad \longrightarrow \frac{\mu(B_{\mathbf{s}} \cap A \cap A_{s_0})}{\mu(A)}, \quad (n \to \infty),
\end{aligned} \tag{12}$$

where the convergence follows from (11), as long as $\mu\{\partial(B_{\mathbf{s}} \cap A \cap A_{s_0})\} = 0$. Let

$$\xi^{(1)}(t) = \max_{i\in\mathbb{N}} U_i^{(1)} V_i^{(1)}(t), \quad t \in T, \tag{13}$$

where $\sum_{i\in\mathbb{N}} \delta_{U_i^{(1)}}$ is a Poisson point process with intensity $(1-p)u^{-2}\mathrm{d}u$ and let $\mu^{(1)}$ be the exponent measure of the associated max-stable random vector $(\xi^{(1)}(t_0), \dots, \xi^{(1)}(t_k))$. Then the choice $A = (1,\infty) \times [0,\infty)^k$ guarantees that $\mu(\cdot \cap A) = \mu^{(1)}(\cdot \cap A)$. Comparing the construction of $\xi^{(1)}$ in (13) with the definition of the exponent measure, we see that $\mu^{(1)}$ is the intensity measure of the Poisson point process $\sum_{i\in\mathbb{N}} \delta_{(U_i^{(1)}V_i^{(1)}(t_0), \dots, U_i^{(1)}V_i^{(1)}(t_k))}$ on E. Hence,

$$\begin{aligned}
\mu(A) &= \int_0^\infty (1-p)u^{-2}\,\mathrm{P}(uV^{(1)}(t_0) > 1)\mathrm{d}u \\
&= (1-p)\int_0^\infty u^{-2}\int_{[u^{-1},\infty)} \mathrm{P}(V^{(1)}(t_0) \in \mathrm{d}y)\mathrm{d}u \\
&= (1-p)\int_0^\infty y\,\mathrm{P}(V^{(1)}(t_0) \in \mathrm{d}y) = (1-p)\,\mathrm{E}\,V^{(1)}(t_0) = 1,
\end{aligned} \tag{14}$$

where the last equality follows from $\mathrm{E}\, V^{(1)}(t_0) = \mathrm{E}\, V(t_0)/(1-p)$. Furthermore, for $s_0 \geq 1$ and $\mathbf{s} \in [0,\infty)^k$ with $\mathrm{P}(\Delta\tilde{\mathbf{V}}^{(1)} \in \partial[\mathbf{0},\mathbf{s}]) = 0$,

$$\begin{aligned}
&\mu(B_{\mathbf{s}} \cap A \cap A_{s_0})/((1-p)\mu(A)) \\
&= \int_0^\infty u^{-2}\, \mathrm{P}\Big(uV^{(1)}(t_0) > s_0,\ \big(uV^{(1)}(t_1),\dots,uV^{(1)}(t_k)\big) \leq \mathbf{s}uV^{(1)}(t_0)\Big)\mathrm{d}u \\
&= \int_0^\infty \int_{[s_0u^{-1},\infty)} u^{-2}\, \mathrm{P}\Big(V^{(1)}(t_0) \in \mathrm{d}y \Big| \Delta\mathbf{V}^{(1)} \leq \mathbf{s}\Big)\, \mathrm{P}(\Delta\mathbf{V}^{(1)} \leq \mathbf{s})\mathrm{d}u \\
&= \int_{[\mathbf{0},\mathbf{s}]} \int_{[0,\infty)} y s_0^{-1} \cdot \mathrm{P}\Big(V^{(1)}(t_0) \in \mathrm{d}y \Big| \Delta\mathbf{V}^{(1)} = \mathbf{z}\Big)\, \mathrm{P}(\Delta\mathbf{V}^{(1)} \in \mathrm{d}\mathbf{z}) \\
&= s_0^{-1} \int_{[\mathbf{0},\mathbf{s}]} \mathrm{E}\Big(V^{(1)}(t_0) \Big| \Delta\mathbf{V}^{(1)} = \mathbf{z}\Big)\, \mathrm{P}(\Delta\mathbf{V}^{(1)} \in \mathrm{d}\mathbf{z}). \qquad (15)
\end{aligned}$$

Equation (15) shows that the convergence in (12) holds for all continuity points $\mathbf{s} \in [0,\infty)^k$ of the distribution function of $\Delta\mathbf{V}^{(1)}$. Since $s_0 \geq 1$ was arbitrary, this concludes the proof.

Remark 2.2.

1. If $V^{(1)}(t_0)$ is stochastically independent of the increments $\Delta\mathbf{V}^{(1)}$, we simply have $\mathrm{P}(\Delta\tilde{\mathbf{V}}^{(1)} \in d\mathbf{z}) = \mathrm{P}(\Delta\mathbf{V}^{(1)} \in d\mathbf{z})$.
2. If $p = \mathrm{P}(V(t_0) = 0) = 0$, the exponent measure μ of any finite-dimensional vector $\Xi = (\xi(t_0),\dots,\xi(t_k))$, $t_0,\dots,t_k \in T$, $k \in \mathbb{N}$, satisfies the condition $\mu\left(\{0\} \times [0,\infty)^k\right) = 0$, and following Proposition 2.1, the incremental representation of Ξ according to (3) is given by $\Xi = \max_{i\in\mathbb{N}} U_i \cdot (1, \Delta\tilde{\mathbf{V}}_i)^\top$, where $\Delta\tilde{\mathbf{V}}_i$, $i \in \mathbb{N}$, are independent copies of $\Delta\tilde{\mathbf{V}} = \Delta\tilde{\mathbf{V}}^{(1)}$.
3. If ξ admits a representation (3), we have $\mathrm{P}(\Delta\tilde{\mathbf{V}}^{(1)} \in d\mathbf{z}) = \mathrm{P}(\Delta\mathbf{V} \in d\mathbf{z})$, which shows that (6) is indeed a special case of Theorem 2.1.

Remark 2.3. In the above theorem, the sequence $a(n)$ of thresholds is only assumed to converge to ∞, as $n \to \infty$, ensuring that $\{\eta(t_0) > a(n)\}$ becomes a rare event. For statistical applications $a(n)$ should also be chosen such that the number of exceedances

$$N(n) = \sum_{i=1}^n \mathbf{1}\{\eta_i(t_0) > a(n)\}$$

converges to ∞ almost surely, where $(\eta_i)_{i\in\mathbb{N}}$ is a sequence of independent copies of η. By the Poisson limit theorem, this is equivalent to the additional assumption that $\lim_{n\to\infty} a(n)/n = 0$, since in that case $n\,\mathrm{P}(\eta(t_0) > a(n)) = n/a(n) \to \infty$, as $n \to \infty$.

Remark 2.4. [12] consider Hüsler-Reiss distributions [15, 16] and obtain their limiting results by conditioning on certain extremal events $A \subset E$. They show that various choices of A are sensible in the Hüsler-Reiss case, leading to different limiting distributions of the increments of η. In case ξ is a Brown-Resnick process and $A = (1,\infty) \times [0,\infty)^k$ the assertions of Theorem 2.1 and [12, Thm. 3.3] coincide.

Example 2.1. (*Extremal Gaussian process [23].*) A commonly used class of stationary yet non-ergodic max-stable processes on $\mathbb{R}^d$ is defined by

$$\xi(t) = \max_{i\in\mathbb{N}} U_i Y_i(t), \quad t \in \mathbb{R}^d, \tag{16}$$

where $\sum_{i\in\mathbb{N}} \delta_{U_i}$ is a Fréchet point process, $Y_i(t) = \max(0, \tilde{Y}_i(t))$, $i \in \mathbb{N}$, and the $\tilde{Y}_i$ are i.i.d. stationary, centered Gaussian processes with $\mathrm{E}(\max(0, \tilde{Y}_i(t))) = 1$ for all $t \in \mathbb{R}^d$ [23, 3]. Note that in general, a $t_0 \in \mathbb{R}^d$ s.t. $Y_i(t_0) = 1$ a.s. does not exist, i.e., the process admits representation (7) but not representation (3). In particular, for the extremal Gaussian process we have $p = \mathrm{P}(V(t_0) = 0) = 1/2$ and the distribution of the increments in (10) becomes

$$\begin{aligned}\mathrm{P}(\Delta\tilde{\mathbf{V}}^{(1)} \in \mathrm{d}\mathbf{z}) = \frac{1}{2}\,\mathrm{E}\Big[Y(t_0)\,\Big|\,(Y(t_1)/Y(t_0),\ldots,Y(t_k)/Y(t_0)) = \mathbf{z},\, Y(t_0) > 0\Big] \\ \cdot\, \mathrm{P}\Big(\big(Y(t_1)/Y(t_0),\ldots,Y(t_k)/Y(t_0)\big) \in \mathrm{d}\mathbf{z}\,\Big|\,Y(t_0) > 0\Big).\end{aligned}$$

While the Hüsler-Reiss distribution is already given by the incremental representation (3), cf. [16], other distributions can be suitably rewritten, provided that the cumulative distribution function and hence the respective exponent measure μ is known.

Proposition 2.1. *Let $\Xi = (\xi(t_0),\ldots,\xi(t_k))$ be a max-stable process on $T = \{t_0,\ldots,t_k\}$ with standard Fréchet margins and suppose that its exponent measure μ is concentrated on $(0,\infty) \times [0,\infty)^k$. Define a random vector $\mathbf{W} = (W^{(1)},\ldots,W^{(k)})$ via its cumulative distribution function*

$$\mathrm{P}(\mathbf{W} \le \mathbf{s}) = \mu(B_{\mathbf{s}} \cap A), \quad \mathbf{s} \in [0,\infty)^k, \tag{17}$$

where $A = (1,\infty)\times[0,\infty)^k$ and $B_{\mathbf{s}} = \{\mathbf{x} \in [0,\infty)^{k+1} : (x^{(1)},\ldots,x^{(k)}) \le x^{(0)}\mathbf{s}\}$. Then, Ξ allows for an incremental representation (3) with $\mathbf{W}_i$, $i \in \mathbb{N}$, being independent copies of $\mathbf{W}$.

Proof. First, we note that (17) indeed defines a valid cumulative distribution function. To this end, consider the measurable transformation

$$T : (0,\infty) \times [0,\infty)^k \to (0,\infty) \times [0,\infty)^k,\ (x_0,\ldots,x_k) \mapsto \left(x_0, \frac{x_1}{x_0},\ldots,\frac{x_k}{x_0}\right).$$

Then, $T(B_{\mathbf{s}} \cap A) = (1,\infty) \times [\mathbf{0},\mathbf{s}]$ and the measure $\mu^T(\cdot) = \mu(T^{-1}((1,\infty) \times \cdot))$ is a probability measure on $[0,\infty)^k$. Since

$$\mu(B_{\mathbf{s}} \cap A) = \mu(T^{-1}((1,\infty) \times [\mathbf{0},\mathbf{s}])) = \mu^T([\mathbf{0},\mathbf{s}]),$$

the random vector $\mathbf{W}$ is well-defined and has law μ^T.

By definition of the exponent measure, we have $\Xi \stackrel{\mathrm{D}}{=} \max_{i\in\mathbb{N}} \mathbf{X}_i$, where $\Pi = \sum_{i\in\mathbb{N}} \delta_{\mathbf{X}_i}$ is a PPP on E with intensity measure μ. Then, the transformed point process $T\Pi = \sum_{i\in\mathbb{N}} \delta_{(X_i^{(0)}, X_i^{(1)}/X_i^{(0)}, \ldots, X_i^{(k)}/X_i^{(0)})}$ has intensity measure

$$\begin{aligned}\tilde{\mu}((c,\infty)\times[\mathbf{0},\mathbf{s}]) &= \mu\left(T^{-1}\left((c,\infty)\times[\mathbf{0},\mathbf{s}]\right)\right)\\ &= \mu(B_{\mathbf{s}}\cap((c,\infty)\times[0,\infty)^k)) = c^{-1}\mu(B_{\mathbf{s}}\cap A)\end{aligned}$$

for any $c > 0$, $\mathbf{s} \in [0,\infty)^k$, where we use the fact that μ, as an exponent measure, has the homogeneity property $c^{-1}\mu(\mathrm{d}\mathbf{x}) = \mu(\mathrm{d}(c\mathbf{x}))$. Thus, $T\Pi$ has the same intensity as $\sum_{i\in\mathbb{N}} \delta_{(U_i,\mathbf{W}_i)}$, where $\sum_{i\in\mathbb{N}} \delta_{U_i}$ is a Fréchet point process and $\mathbf{W}_i$, $i \in \mathbb{N}$, are i.i.d. vectors with law $\mathrm{P}(\mathbf{W} \le \mathbf{s}) = \mu(B_{\mathbf{s}}\cap A)$. Hence, we have

$$\begin{aligned}\Xi &\stackrel{\mathrm{D}}{=} \max_{i\in\mathbb{N}} T^{-1}\left(\left(X_i^{(0)}, X_i^{(1)}/X_i^{(0)}, \ldots, X_i^{(k)}/X_i^{(0)}\right)\right)\\ &\stackrel{\mathrm{D}}{=} \max_{i\in\mathbb{N}} T^{-1}\left((U_i,\mathbf{W}_i)\right) = \max_{i\in\mathbb{N}} U_i\mathbf{W}_i,\end{aligned}$$

which completes the proof.

Example 2.2. (*Symmetric logistic distribution, cf. [14].*) For $T = \{t_0,\ldots,t_k\}$, the symmetric logistic distribution is given by

$$\mathrm{P}(\xi(t_0)\le x_0,\ldots,\xi(t_k)\le x_k) = \exp\left[-\left(x_0^{-q}+\cdots+x_k^{-q}\right)^{1/q}\right], \tag{18}$$

for $x_0,\ldots,x_k > 0$ and $q > 1$. Hence, the density of the exponent measure is

$$\mu(\mathrm{d}x_0,\ldots,\mathrm{d}x_k) = \left(\sum_{i=0}^k x_i^{-q}\right)^{1/q-(k+1)} \left(\prod_{i=1}^k (iq-1)\right)\prod_{i=0}^k x_i^{-q-1}\mathrm{d}x_0\ldots\mathrm{d}x_k.$$

Applying Proposition 2.1, the incremental process W in the representation (3) is given by

$$\mathrm{P}(W(t_1)\le s_1,\ldots W(t_k)\le s_k) = \left(1+\sum_{i=1}^k s_i^{-q}\right)^{1/q-1}.$$

2.1. Continuous sample paths

In this subsection, we provide an analog result to Theorem 2.1, in which convergence in the sense of finite-dimensional distributions is replaced by weak convergence on function spaces. In the following, for a Borel set $U \subset \mathbb{R}^d$, we denote by $C(U)$ and $C^+(U)$ the space of non-negative and strictly positive continuous functions on U, respectively, equipped with the topology of uniform convergence on compact sets.

Theorem 2.2. *Let K be a compact subset of $\mathbb{R}^d$ and $\{\eta(t): \ t \in K\}$ be a process with positive and continuous sample paths in the MDA of a max-stable process $\{\xi(t): \ t \in K\}$ as in (3) in the sense of weak convergence on $C(K)$. In particular, suppose that*

$$\frac{1}{n} \max_{i=1}^{n} \eta_i(\cdot) \xrightarrow{\mathrm{D}} \xi(\cdot), \quad n \to \infty.$$

Let W be the incremental process from (3) and Z a Pareto random variable, independent of W. Then, for any sequence $a(n)$ of real numbers with $a(n) \to \infty$, we have the weak convergence on $(0,\infty) \times C(K)$

$$\left(\frac{\eta(t_0)}{a(n)}, \frac{\eta(\cdot)}{\eta(t_0)} \;\middle|\; \eta(t_0) > a(n) \right) \xrightarrow{\mathrm{D}} (Z, W(\cdot)),$$

as n tends to ∞.

Remark 2.5. Analogously to [28, Thm. 5], weak convergence of a sequence of probability measures P_n, $n \in \mathbb{N}$, to some probability measure P on $C(\mathbb{R}^d)$ is equivalent to weak convergence of $P_n r_j^{-1}$ to $P r_j^{-1}$ on $C([-j,j]^d)$ for all $j \geq 1$, where $r_j : C(\mathbb{R}^d) \to C([-j,j]^d)$ denotes the restriction of a function to the cube $[-j,j]^d$. Hence the assertion of Theorem 2.2 remains valid if the compact set K is replaced by $\mathbb{R}^d$.

Proof of Theorem 2.2. As the process ξ is max-stable and $\eta \in \mathrm{MDA}(\xi)$, similarly to the case of multivariate max-stable distributions (cf. Theorem 2.1), we have that

$$\lim_{u \to \infty} u \,\mathrm{P}(\eta/u \in B) = \mu(B) \tag{19}$$

for any Borel set $B \subset C(K)$ bounded away from 0^K, i.e., $\inf\{\sup_{s \in K} f(s): \ f \in B\} > 0$, and with $\mu(\partial B) = 0$ [7, Cor. 9.3.2], where μ is the *exponent measure* of ξ, defined by

$$\begin{aligned} &\mathrm{P}(\xi(s) \leq x_j, \ s \in K_j, \ j = 1, \ldots, m) \\ &= \exp\left[-\mu\left(\left\{ f \in C(K): \ \sup\nolimits_{s \in K_j} f(s) > x_j \text{ for some } j \in \{1, \ldots, m\} \right\} \right) \right] \end{aligned} \tag{20}$$

for $x_j \geq 0$, $K_j \subset K$ compact. Thus, μ equals the intensity measure of the Poisson point process $\sum_{i \in \mathbb{N}} \delta_{U_i W_i(\cdot)}$. For $z > 0$ and $D \subset C(K)$ Borel, we consider the sets

$$\begin{aligned} A_z &= \{ f \in C(K): \ f(t_0) > z \} \\ B_D &= \{ f \in C(K): f(\cdot)/f(t_0) \in D \} \end{aligned}$$

and $A = A_1$. Note that B_D is invariant w.r.t. multiplication by any positive constant. Then, as $W(t_0) = 1$ a.s., we have $\mu(A_z) = \int_z^\infty u^{-2} \,\mathrm{d}u = z^{-1}$ and for

$s_0 \geq 1$ and any Borel set $D \subset C(K)$ with $\mathrm{P}(W \in \partial D) = 0$, by (19), we get

$$\begin{aligned}
&\mathrm{P}\left\{\eta(t_0)/a(n) > s_0,\ \eta(\cdot)/\eta(t_0) \in D \,\middle|\, \eta(t_0) > a(n)\right\} \\
&= \frac{a(n)\,\mathrm{P}\{\eta(\cdot)/a(n) \in A_{s_0} \cap B_D \cap A\}}{a(n)\,\mathrm{P}\{\eta(\cdot)/a(n) \in A\}} \\
&\overset{n\to\infty}{\longrightarrow} \frac{\mu(B_D \cap A_{s_0})}{\mu(A)} \\
&= \int_{s_0}^{\infty} u^{-2}\,\mathrm{P}\{uW(\cdot) \in B_D\}\,\mathrm{d}u \\
&= s_0^{-1}\,\mathrm{P}\{W(\cdot) \in D\},
\end{aligned}$$

which is the joint distribution of Z and $W(\cdot)$.

Example 2.3. (*Brown-Resnick processes, cf. [4, 17].*) For $T = \mathbb{R}^d$, $d \geq 1$, let $\{Y(t) : t \in T\}$ be a centered Gaussian process with stationary increments, continuous sample paths and $Y(t_0) = 0$ for some $t_0 \in \mathbb{R}^d$. Note that by [1, Thm. 1.4.1] it is sufficient for the continuity of Y that there exist constants $C, \alpha, \delta > 0$, such that

$$\mathrm{E}\,|Y(s) - Y(t)|^2 \leq \frac{C}{|\log \|s-t\||^{1+\alpha}}$$

for all $s, t \in \mathbb{R}^d$ with $\|s-t\| < \delta$. Further let $\gamma(t) = \mathrm{E}(Y(t)-Y(0))^2$ and $\sigma^2(t) = \mathrm{E}(Y(t))^2$, $t \in \mathbb{R}^d$, denote the variogram and the variance of Y, respectively. Then, with a Fréchet point process $\sum_{i\in\mathbb{N}} \delta_{U_i}$ and independent copies Y_i of Y, $i \in \mathbb{N}$, the process

$$\xi(t) = \max_{i\in\mathbb{N}} U_i \exp\left(Y_i(t) - \sigma^2(t)/2\right), \quad t \in \mathbb{R}^d, \tag{21}$$

is stationary and its distribution only depends on the variogram γ. Comparing (21) with the incremental representation (3), the distribution of the increments is given by the log-Gaussian random field $W(t) = \exp\left(Y(t) - \sigma^2(t)/2\right)$, $t \in \mathbb{R}^d$, and Theorem 2.2 applies.

3. Mixed moving maxima representation

A large and commonly used class of max-stable processes is the class of M3 processes (5). Let

$$\Pi_0 = \sum_{i\in\mathbb{N}} \delta_{(U_i.T_i,F_i)} \tag{22}$$

be the corresponding PPP on $(0,\infty)\times\mathbb{R}^d\times C(\mathbb{R}^d)$ with intensity $u^{-2}\mathrm{d}u\,\mathrm{d}t\ \mathrm{P}_F(\mathrm{d}f)$. In the sequel, M3 processes are denoted by

$$M(t) = \max_{i\in\mathbb{N}} U_i F_i(t - T_i), \quad t \in \mathbb{R}^d.$$

The marginal distributions of M are given by

$$\begin{aligned}
&\mathrm{P}(M(t_0) \leq s_0, \ldots, M(t_k) \leq s_k) \\
&= \mathrm{P}\left[\Pi_0\left(\left\{(u,t,f) : \max_{l=0}^{k} u f(t_l - t)/s_l > 1\right\}\right) = 0\right] \\
&= \exp\left(-\int_{C(\mathbb{R}^d)} \int_{\mathbb{R}^d} \max_{l=0}^{k} (f(t_l - t)/s_l)\, \mathrm{d}t \; \mathrm{P}_F(\mathrm{d}f)\right),
\end{aligned} \tag{23}$$

$t_0, \ldots, t_k \in \mathbb{R}^d$, $s_0, \ldots, s_k \geq 0$, $k \in \mathbb{N}$.

In Section 2, we were interested in recovering the incremental process W from processes in the MDA of a max-stable process with incremental representation. In case of M3 processes, the object of interest is clearly the distribution of the shape function F. Thus, in what follows, we provide the corresponding convergence results for processes η in the MDA of an M3 process. We distinguish between processes on $\mathbb{R}^d$ with continuous sample paths and processes on a grid ($\mathbb{Z}^d$). The main idea is to consider η in the neighborhood of its own (local) maximum, conditional on this maximum being large.

3.1. Continuous Case

Let $\{\eta(t) : t \in \mathbb{R}^d\}$ be strictly positive and in the MDA of a mixed moving maxima process M in the sense of weak convergence in $C(\mathbb{R}^d)$. We assume that η is normalized such that the norming functions in (1) are given by $c_n(t) = 1/n$ and $b_n(t) = 0$, for any $n \in \mathbb{N}$ and $t \in \mathbb{R}^d$. Further suppose that the shape function F of M is sample-continuous and satisfies

$$\begin{aligned}
&F(\mathbf{0}) = \lambda \quad a.s., \\
&F(t) \in [0, \lambda) \; \forall t \in \mathbb{R}^d \setminus \{\mathbf{0}\} \quad a.s.
\end{aligned} \tag{24}$$

for some $\lambda > 0$ and

$$\int_{\mathbb{R}^d} \mathrm{E}\left\{\max_{t_0 \in K} F(t_0 - t)\right\} \mathrm{d}t < \infty \tag{25}$$

for any compact set $K \subset \mathbb{R}^d$. Under these assumptions, there is an analog result to Theorem 2.2.

Theorem 3.1. *Let $Q, K \subset \mathbb{R}^d$ be compact such that ∂Q is a Lebesgue null set and let*

$$\tau_Q : \; C(Q) \to \mathbb{R}^d, \; f \mapsto \inf\left(\arg\max_{t \in Q} f(t)\right),$$

where "inf" *is understood in the lexicographic sense. Then, under the above assumptions, for any Borel set $B \subset C(K)$ with $\mathrm{P}(F/\lambda \in \partial B) = 0$, and any* *sequence* $a(n)$ *with* $a(n) \to \infty$ *as* $n \to \infty$, *we have*

$$\lim_{\substack{\{\mathbf{0}\}\in L\nearrow\mathbb{R}^d \\ \text{compact}}} \limsup_{n\to\infty} \mathrm{P}\Big\{\eta\big(\tau_Q(\eta|_Q)+\cdot\big)/\eta(\tau_Q(\eta|_Q)) \in B \;\Big|\; \max_{t\in Q}\eta(t) = \max_{t\in Q\oplus L}\eta(t),\ \max_{t\in Q}\eta(t) \geq a(n)\Big\} = \mathrm{P}\big\{F(\cdot)/\lambda \in B\big\},$$

where $\oplus$ *denotes morphological dilation.*

The same result holds true if we replace $\limsup_{n\to\infty}$ *by* $\liminf_{n\to\infty}$.

Proof. First, we consider a fixed compact set $L \subset \mathbb{R}^d$ large enough such that $K \cup \{\mathbf{0}\} \subset L$ and define

$$A_L = \Big\{f \in C(Q\oplus L):\ \max_{t\in Q} f(t) \geq 1,\ \max_{t\in Q} f(t) = \max_{t\in Q\oplus L} f(t)\Big\}$$

and

$$C_B = \big\{f \in C(Q\oplus L):\ f\big(\tau_Q(f|_Q)+\cdot\big)/f(\tau_Q(f|_Q)) \in B\big\}$$

for any Borel set $B \subset C(K)$. Note that C_B is invariant w.r.t. multiplication by any positive constant. Thus, we get

$$\begin{aligned} &\mathrm{P}\Big\{\eta\big(\tau_Q(\eta|_Q)+\cdot\big)/\eta(\tau_Q(\eta|_Q)) \in B \;\Big|\; \max_{t\in Q}\eta(t) = \max_{t\in Q\oplus L}\eta(t) \geq a(n)\Big\} \\ &= \mathrm{P}\big\{\eta/a(n) \in C_B \,\big|\, \eta/a(n) \in A_L\big\} \\ &= \frac{a(n)\,\mathrm{P}\big\{\eta/a(n) \in C_B,\ \eta/a(n) \in A_L\big\}}{a(n)\,\mathrm{P}\big\{\eta/a(n) \in A_L\big\}}. \end{aligned} \tag{26}$$

By [8, Cor. 9.3.2] and [21, Prop. 3.12] we have

$$\begin{aligned} \limsup_{u\to\infty} u\,\mathrm{P}(\eta/u \in C) &\leq \mu(C), \quad C \subset C(Q\oplus L) \text{ closed}, \\ \liminf_{u\to\infty} u\,\mathrm{P}(\eta/u \in O) &\geq \mu(O), \quad O \subset C(Q\oplus L) \text{ open}, \end{aligned}$$

where C and O are bounded away from 0^K. Here, μ is the intensity measure of the PPP $\sum_{i\in\mathbb{N}} \delta_{U_i F_i(\,\cdot\,-T_i)}$ restricted to $C(Q\oplus L)$. Thus, by adding or removing the boundary, we see that all the limit points of Equation (26) lie in the interval

$$\left[\frac{\mu(C_B\cap A_L) - \mu(\partial(C_B\cap A_L))}{\mu(A_L)+\mu(\partial A_L)}, \frac{\mu(C_B\cap A_L) + \mu(\partial(C_B\cap A_L))}{\mu(A_L)-\mu(\partial A_L)}\right]. \tag{27}$$

We note that A_L is closed and the set

$$A_L^* = \Big\{f \in C(Q\oplus L):\ \tau_Q(f|_Q) \in Q^o,\ \max_{t\in Q} f(t) > \max\big\{1, f(t)\big\}\ \forall t \in Q\oplus L \setminus \{\tau_Q(f|_Q)\}\Big\}$$

is in the interior of A_L (Lemma A.1). Hence, we can assess

$$
\begin{aligned}
\mu(\partial A_L) \leq \quad & \mu(\{f \in C(Q \oplus L):\ \max_{t \in Q} f(t) = 1\}) \\
& + \mu\Bigg(\quad \Bigg(\quad \{f \in C(Q \oplus L):\ \tau_Q(f|_Q) \in \partial Q\} \\
& \qquad \cup \left\{ f \in C(Q \oplus L):\ \operatorname*{arg\,max}_{t \in Q \oplus L} f(t) \text{ is not unique} \right\} \Bigg) \\
& \qquad \cap \left\{ f \in C(Q \oplus L):\ \max_{t \in Q} f(t) = \max_{t \in Q \oplus L} f(t) \geq 1 \right\} \Bigg) \\
\leq\ & 0 + \int_{\partial Q} \int_{\lambda^{-1}}^{\infty} u^{-2}\, \mathrm{d}u\, \mathrm{d}t_0 \\
& + \int_{\mathbb{R}^d \setminus (Q \oplus L)} \int_{\lambda^{-1}}^{\infty} u^{-2}\, \mathrm{P}\left\{ u \max_{t_0 \in Q} F(t_0 - x) \geq 1 \right\} \mathrm{d}u\, \mathrm{d}x. \qquad (28)
\end{aligned}
$$

Here, the equality $\mu(\{f \in C(Q \oplus L) : \max_{t \in Q} f(t) = 1\}) = 0$ holds as $\max_{t \in Q} M(t)$ is Fréchet distributed (cf. [8, Lemma 9.3.4]). Since ∂Q is a Lebesgue null set, the second term on the right-hand side of (28) also vanishes. Thus,

$$
\mu(\partial A_L) \leq \int_{\mathbb{R}^d \setminus (Q \oplus L)} \int_{\lambda^{-1}}^{\infty} u^{-2}\, \mathrm{P}\left\{ u \max_{t_0 \in Q} F(t_0 - x) \geq 1 \right\} \mathrm{d}u\, \mathrm{d}x =: c(L). \qquad (29)
$$

Now, let $B \subset C(K)$ a be Borel set such that $\mathrm{P}(F/\lambda \in \partial B) = 0$. For the set C_B, we obtain that the set

$$
C_B^* = \left\{ f \in C(Q \oplus L):\ \operatorname*{arg\,max}_{f \in Q} f(t) \text{ is unique},\ \frac{f\big(\tau_Q(f|_Q) + \cdot\big)}{f(\tau_Q(f|_Q))} \in B^o \right\}
$$

is in the interior of C_B and that the closure of C_B is a subset of

$$
\begin{aligned}
C_B^* \cup & \left\{ f \in C(Q \oplus L):\ \operatorname*{arg\,max}_{t \in Q} f(t) \text{ is not unique} \right\} \\
& \cup \left\{ f \in C(Q \oplus L):\ f\big(\tau_Q(f|_Q) + \cdot\big) / f(\tau_Q(f|_Q)) \in \partial B \right\}
\end{aligned}
$$

(Lemma A.2 and Lemma A.3). Thus, by (29), we can assess

$$
\begin{aligned}
\mu(\partial(C_B \cap A_L)) & \leq \mu(\partial A_L) + \mu(\partial C_B \cap A_L) \\
& \leq c(L) + \int_{\mathbb{R}^d \setminus (Q \oplus L)} \int_{\lambda^{-1}}^{\infty} u^{-2}\, \mathrm{P}\left\{ u \max_{t_0 \in Q} F(t_0 - x) \geq 1 \right\} \mathrm{d}u\, \mathrm{d}x \\
& \qquad + \int_Q \int_{\lambda^{-1}}^{\infty} u^{-2}\, \mathrm{P}(F/\lambda \in \partial B)\, \mathrm{d}u\, \mathrm{d}t \quad = \quad 2c(L). \qquad (30)
\end{aligned}
$$

Furthermore, we get

$$\begin{aligned}
&\mu(C_B \cap A_L)\\
&= \int_Q \int_{\lambda^{-1}}^\infty u^{-2}\,\mathrm{P}\left\{F(\cdot)/\lambda \in B\right\} \mathrm{d}u\,\mathrm{d}t_0\\
&\quad + \int_{\mathbb{R}^d\setminus(Q\oplus L)} \int_{\lambda^{-1}}^\infty u^{-2}\,\mathrm{P}\Big\{u \max_{t_0\in Q} F(t_0 - x) \geq 1,\\
&\qquad\qquad F\left(\left(\tau_Q(F(\cdot - x)|_Q)\right) + \cdot - x\right) \Big/ \max_{t_0\in Q} F(t_0 - x) \in B,\\
&\qquad\qquad F(t-x)/\max_{t_0\in Q} F(t_0 - x) \leq 1\ \forall t \in Q\oplus L\Big\}\,\mathrm{d}u\,\mathrm{d}x. \qquad (31)
\end{aligned}$$

The second term in (31) is positive and can be bounded from above by $c(L)$. Setting $B = C(K)$, $\mu(A_L)$ can be expressed in an analogous way. Now, we plug in the results of (29), (30) and (31) into (27) to obtain that all the limit points of (26) are in the interval

$$\left[\frac{\lambda\cdot|Q|\cdot\mathrm{P}\left\{F(\cdot)/\lambda\in B\right\} - 2c(L)}{\lambda\cdot|Q| + 2c(L)}, \frac{\lambda\cdot|Q|\cdot\mathrm{P}\left\{F(\cdot)/\lambda\in B\right\} + 3c(L)}{\lambda\cdot|Q| - c(L)}\right].$$

Finally, we note that $c(L)$ can be bounded from above by

$$\int_{\mathbb{R}^d\setminus(Q\oplus L)} \mathrm{E}\left\{\max_{t_0\in Q} F(t_0 - x)\right\}\mathrm{d}x,$$

which vanishes for $L \nearrow \mathbb{R}^d$ because of assumption (25). This yields the assertion of the theorem.

We conclude the treatment of the continuous case with an example of a process η that allows for an application of Theorem 3.1. As η will be composed of a (locally) finite number of shape functions from the M3 construction in (5), η may directly model rainfall data and has therefore the potential for various practical applications.

Example 3.1. Let $\{F(t) :\ t \in \mathbb{R}^d\}$ be a random shape function as defined in (4). For $c, \epsilon > 0$ let $\Pi_{c,\epsilon} = \sum_{i\in\mathbb{N}} \delta_{(U_i.T_i,F_i)}$ be a PPP on $(0,\infty)\times\mathbb{R}^d\times C(\mathbb{R}^d)$ with intensity

$$c\mathbf{1}_{\{u\geq\epsilon\}}u^{-2}\mathrm{d}u\,\mathrm{d}t\ \mathrm{P}_F(\mathrm{d}f).$$

and, for $\kappa > 0$, define a process $\tilde{M} = \tilde{M}_{c,\epsilon,\kappa}$ by

$$\tilde{M}(\cdot) = \kappa \vee \max_{(u,t,f)\in\Pi_{c,\epsilon}} uf(\cdot - t).$$

Then, the following statements hold.

1. $\tilde{M}$ is in the MDA of the M3 process M associated to F in the sense of finite-dimensional distributions.
2. If F satisfies (25), then $\tilde{M}$ is in the MDA of M in the sense of weak convergence on $C(\mathbb{R}^d)$.

For a proof of this example, the reader is referred to Appendix B.

3.2. Discrete Case

Theorem 3.1 allows for estimation of F if the complete sample paths of η are known, at least on a large set $Q \oplus L \subset \mathbb{R}^d$. For many applications, this assumption might be too restrictive. Therefore, we seek after a weaker assumption that only requires to know η on a grid. This needs a modification of the underlying model leading to a discretized mixed moving maxima process.

Let $\{F(t):\ t \in \mathbb{Z}^d\}$ be a measurable stochastic process with values in $[0,\infty)$ and

$$\sum_{t\in\mathbb{Z}^d} \mathrm{E}\, F(t) = 1. \tag{32}$$

Further, let $\Pi_{0,\mathrm{discr}} = \sum_{i\in\mathbb{N}} \delta_{(U_i.T_i,F_i)}$ be a Poisson point process on $(0,\infty) \times \mathbb{Z}^d \times [0,\infty)^{\mathbb{Z}^d}$ with intensity $u^{-2}\mathrm{d}u\, \delta_{\mathbb{Z}^d}(\mathrm{d}t)\ \mathrm{P}_F(\mathrm{d}f)$. Then, the discrete mixed moving maxima process M_{discr} is defined by

$$M_{\mathrm{discr}}(t) = \max_{i\in\mathbb{N}} U_i F_i(t - T_i), \quad t \in \mathbb{Z}^d. \tag{33}$$

The process M_{discr} is max-stable and stationary on $\mathbb{Z}^d$ and has standard Fréchet margins.

Let $\{\eta(t),\ t \in \mathbb{Z}^d\}$ be in the MDA of a discrete mixed moving maxima process M_{discr} in the sense of convergence of finite-dimensional distributions with norming functions $c_n(t) = 1/n$ and $b_n(t) = 0$ in (1), $n \in \mathbb{N}$ and $t \in \mathbb{Z}^d$. Furthermore, we assume that the shape function F satisfies (24) with $\mathbb{R}^d$ being replaced by $\mathbb{Z}^d$. Then, analogously to Theorem 3.1, the following convergence result can be shown.

Theorem 3.2. *Under the above assumptions, for any $k \in \mathbb{N}$, $k+1$ distinct points $t_0,\dots,t_k \in \mathbb{Z}^d$, any Borel sets $B_1,\dots,B_k \subset [0,\infty)$ such that*

$$\mathrm{P}\left\{(F(t_1)/\lambda,\dots,F(t_k)/\lambda) \in \partial(B_1\times\cdots\times B_k)\right\} = 0,$$

and any sequence $a(n)$ with $a(n)\to\infty$ as $n\to\infty$, it holds

$$\begin{aligned}\lim_{\substack{\{\mathbf{0}\}\in L\nearrow\mathbb{Z}^d\\ \text{compact}}} \lim_{n\to\infty} \mathrm{P}\bigl\{\eta(t_0+t_i)/\eta(t_0) \in B_i,\ i=1,\dots,k \ \bigm|\ & \\ \eta(t_0) = \max_{t\in L}\eta(t_0+t),\ \eta(t_0)\geq a(n)\bigr\}& \\ = \mathrm{P}\left\{F(t_i)/\lambda \in B_i,\ i=1,\dots,k\right\}.&\end{aligned}$$

4. Switching between the different representations

In the previous sections we analyzed processes that admit the incremental representations (3) or (7) and, on the other hand, processes of M3 type as in (5). We show that under certain assumptions, we can switch from one representation to the other.

4.1. Incremental representation of mixed moving maxima processes

We distinguish between M3 processes with strictly positive shape functions, for which we can find an incremental representation (3), and general non-negative shape functions, for which only the weaker representation (7) can be obtained.

4.1.1. *Mixed moving maxima processes with positive shape functions*

Theorem 4.1. *Let M be an M3 process on $\mathbb{R}^d$ as in (5) with a shape function F with $F(t) > 0$ for all $t \in \mathbb{R}^d$. Then M admits a representation (3) with $t_0 = 0$ and incremental process W given by*

$$\mathrm{P}(W \in L) = \int_{C^+(\mathbb{R}^d)} \int_{\mathbb{R}^d} \mathbf{1}_{\{f(\cdot - t)/f(-t) \in L\}} f(-t)\,\mathrm{d}t\,\mathrm{P}_F(\mathrm{d}f), \quad L \in \mathcal{B}(C^+(\mathbb{R}^d)). \tag{34}$$

Proof. We consider the two Poisson point processes on $(0,\infty) \times C^+(\mathbb{R}^d)$

$$\Pi_1 = \sum_{i \in \mathbb{N}} \delta_{(U_i F_i(-T_i), F_i(\cdot - T_i)/F_i(-T_i))}, \tag{35}$$

as a transformation of Π_0 in (22), and

$$\Pi_2 = \sum_{i \in \mathbb{N}} \delta_{(U_i', W_i(\cdot))}, \tag{36}$$

with W_i, $i \in \mathbb{N}$, being independent copies of W, and with $\sum_{i\in\mathbb{N}} \delta_{U_i'}$ being a Fréchet point process. Then the intensity measures of Π_1 and Π_2 satisfy

$$\begin{aligned}
&\mathrm{E}\,\Pi_1([z,\infty) \times L) \\
&= \int_{C^+(\mathbb{R}^d)} \int_{\mathbb{R}^d} \int_0^\infty u^{-2} \mathbf{1}_{\{uf(-t) \geq z\}} \mathbf{1}_{\{f(\cdot - t)/f(-t) \in L\}}\,\mathrm{d}u\,\mathrm{d}t\,\mathrm{P}_F(\mathrm{d}f) \\
&= z^{-1} \int_{C^+(\mathbb{R}^d)} \int_{\mathbb{R}^d} \mathbf{1}_{\{f(\cdot - t)/f(-t) \in L\}} f(-t)\,\mathrm{d}t\,\mathrm{P}_F(\mathrm{d}f) \\
&= z^{-1}\,\mathrm{P}(W \in L) \\
&= \mathrm{E}\,\Pi_2([z,\infty) \times L),
\end{aligned}$$

$L \in \mathcal{B}(C^+(\mathbb{R}^d))$, $z > 0$, and hence $\Pi_1 \stackrel{\mathrm{D}}{=} \Pi_2$. The assertion follows from the fact that M is uniquely determined by Π_1 via the relation $M(t) = \max_{(v,g)\in\Pi_1} vg(t)$, $t \in \mathbb{R}^d$.

While the definition of W in (34) is rather implicit, in the following, we provide an explicit construction of the incremental process W, which can also be used for simulation. To this end, let $\sum_{i\in\mathbb{N}} \delta_{U_i''}$ be a Fréchet point process and let the distribution of $(S,G) \in C^+(\mathbb{R}^d) \times \mathbb{R}^d$ be given by

$$\begin{aligned} &\mathrm{P}\big((S,G) \in (B \times L)\big) \\ &= \int_{C^+(\mathbb{R}^d)} \int_{\mathbb{R}^d} \mathbf{1}_{s\in B} \mathbf{1}_{f\in L} \frac{f(-s)}{\int f(r)\mathrm{d}r}\,\mathrm{d}s \left(\int f(r)\mathrm{d}r\right) \mathrm{P}_F(\mathrm{d}f) \\ &= \int_{C^+(\mathbb{R}^d)} \int_{\mathbb{R}^d} \mathbf{1}_{s\in B} \mathbf{1}_{f\in L} f(-s)\,\mathrm{d}s\; \mathrm{P}_F(\mathrm{d}f), \end{aligned} \tag{37}$$

$B \in \mathcal{B}^d$, $L \in \mathcal{B}(C^+(\mathbb{R}^d))$. In other words, $\mathrm{P}_G(\mathrm{d}f) = (\int f(r)\,\mathrm{d}r)\ \mathrm{P}_F(\mathrm{d}f)$ and, conditional on $\{G = f\}$, the density function of the shift S is proportional to $f(-\cdot)$. Putting $W(\cdot) = G(\cdot - S)/G(-S)$, equation (34) is satisfied and with i.i.d. copies W_i, $i \in \mathbb{N}$, of W, we get that $\max_{i\in\mathbb{N}} U_i'' W_i(\cdot)$ is indeed an incremental representation (3) of the mixed moving maxima process M.

Remark 4.1. (*M3 representation of Brown-Resnick processes, cf. [17].*) We consider the following two special cases of mixed moving maxima processes:

1. Let $\Sigma \in \mathbb{R}^{d\times d}$ be a positive definite matrix and let the shape function be given by $F(t) = (2\pi)^{-d/2}|\Sigma|^{-1/2} \exp\left\{-\frac{1}{2} t^\top \Sigma^{-1} t\right\}$, $t \in \mathbb{R}^d$. Then, M becomes the well-known Smith process. At the same time, by (37), $S \sim N(0,\Sigma)$ and $G \equiv F$. Thus

$$\begin{aligned} Y(t) &= \exp\left\{-\tfrac{1}{2}(t-S)^\top \Sigma^{-1}(t-S) + \tfrac{1}{2} S^\top \Sigma^{-1} S\right\} \\ &= \exp\left\{-\tfrac{1}{2} t^\top \Sigma^{-1} t + t^\top \Sigma^{-1} S\right\}. \end{aligned}$$

   Since $\mathrm{E}(t^\top \Sigma^{-1} S)^2 = t^\top \Sigma^{-1} t$, M is equivalent to the Brown-Resnick process in (21) with variogram $\gamma(h) = h^\top \Sigma^{-1} h$.

2. For the one-dimensional Brown-Resnick process ξ in (21) with variogram $\gamma(h) = |h|$, i.e., Y is the exponential of a standard Brownian motion with drift $-|t|/2$, [11] recently showed that the M3 representation is given by $\{F(t) : t \in \mathbb{R}\} = \{Y(t) \mid Y(s) \le 0\ \forall s \in \mathbb{R} : t \in \mathbb{R}\}$, i.e., the shape function is the exponential of a conditionally negative drifted Brownian motion. Having these two representations, it follows that the law of the conditional Brownian motion F, re-weighted by $\int F(t)\mathrm{d}t$ and randomly shifted with density $F(-\cdot)/\int F(t)\mathrm{d}t$, coincides with the law of Y.

4.1.2. *Mixed moving maxima processes with finitely supported shape functions*

Let M be an M3 process on $\mathbb{R}^d$ as in (5). In contrast to Section 4.1.1, where the shape functions are required to take positive values, here, we allow for arbitrary shape functions with values in $[0,\infty)$.

Theorem 4.2. *The M3 process M as in (5) allows for an incremental representation of the form (7), with incremental processes V_i given by*

$$V_i(\cdot) = F_i(\cdot - R_i)/g(R_i).$$

Here R_i, $i \in \mathbb{N}$, are i.i.d. copies of a random vector R with arbitrary density g satisfying $g(t) > 0$ for all $t \in \mathbb{R}^d$, and F_i, $i \in \mathbb{N}$, are i.i.d. copies of the random shape function F.

Proof. With $\sum_{i\in\mathbb{N}} \delta_{U_i}$ being a Fréchet point process, we consider the process

$$\tilde{M}(t) = \max_{i\in\mathbb{N}} U_i F_i(t - R_i)/g(R_i), \qquad t \in \mathbb{R}^d,$$

which clearly is of the form (7). Then,

$$\begin{aligned}
&\mathrm{P}(\tilde{M}(t_0) \le s_0, \ldots, \tilde{M}(t_k) \le s_k) \\
&\quad = \exp\left(- \int_{C(\mathbb{R}^d)} \int_{\mathbb{R}^d} \max_{l=0}^{k} (f(t_l - t)/(g(t)s_l)) g(t) \, \mathrm{d}t \; \mathrm{P}_F(\mathrm{d}f) \right) \\
&\quad = \exp\left(- \int_{C(\mathbb{R}^d)} \int_{\mathbb{R}^d} \max_{l=0}^{k} (f(t_l - t)/s_l)) \, \mathrm{d}t \; \mathrm{P}_F(\mathrm{d}f) \right).
\end{aligned}$$

The right-hand side coincides with the marginal distribution of M, which is given by (23). This concludes the proof.

Decomposing V as in (8) with $t_0 = 0$, we obtain the equality in distribution

$$V^{(1)}(\cdot) \stackrel{\mathrm{D}}{=} \big(F(\cdot - R)/g(R) \;\big|\; -R \in \mathrm{supp}(F)\big).$$

Applying Theorem 2.1 yields

$$\begin{aligned}
&\mathrm{P}\big(\Delta\tilde{\mathbf{V}}^{(1)} \in \mathrm{d}\mathbf{z}\big) \\
&= \mathrm{P}\big(F(-R)/g(R) > 0\big) \cdot \int_0^\infty y \, \mathrm{P}\left(V^{(1)}(0) \in \mathrm{d}y, \; \Delta\mathbf{V}^{(1)} \in \mathrm{d}\mathbf{z}\right) \\
&= \int_{C(\mathbb{R}^d)} \int_{-\mathrm{supp}(f)} g(s) \, \mathrm{d}s \; \mathrm{P}_F(\mathrm{d}f) \\
&\qquad \cdot \int_0^\infty y \int_{C(\mathbb{R}^d)} \int_{-\mathrm{supp}(f)} \mathbf{1}_{f(-t)/g(t) \in \mathrm{d}y} \mathbf{1}_{(f(t_l - t)/f(-t))_{l=1}^k \in \mathrm{d}\mathbf{z}} \\
&\qquad\qquad \cdot g(t) \left(\int_{-\mathrm{supp}(f)} g(s) \mathrm{d}s \right)^{-1} \mathrm{d}t \; \mathrm{P}_F(\mathrm{d}f) \, \mathrm{d}y \\
&= \int_{C(\mathbb{R}^d)} \int_{\mathrm{supp}(f)} g(-s) \, \mathrm{d}s \, \mathrm{P}_F(\mathrm{d}f) \\
&\qquad \cdot \int_{C(\mathbb{R}^d)} \int_{\mathrm{supp}(f)} f(t) \mathbf{1}_{(f(t_l + t)/f(t))_{l=1}^k \in \mathrm{d}\mathbf{z}} \left(\int_{\mathrm{supp}(f)} g(-s) \mathrm{d}s \right)^{-1} \mathrm{d}t \; \mathrm{P}_F(\mathrm{d}f).
\end{aligned} \tag{38}$$

If the shape function F is deterministic, the right-hand side of (38) simplifies to $\int_{\mathrm{supp}(f)} f(t)\mathbf{1}_{(f(t_l+t)/f(t))_{l=1}^k \in \mathrm{d}\mathbf{z}}\,\mathrm{d}t$, i.e., the asymptotic conditional increments of $\eta \in \mathrm{MDA}(M)$ can be seen as a convolution of the shape function's increments with a random shift, whose density is given by the shape function itself. Note in particular, that this distribution is independent of the choice of the density g in Theorem 4.2.

Remark 4.2. Section 4.1.1 considers the subclass of M3 processes with strictly positive shape functions and provides an incremental representation as in (3), which is nicely related to the conditional increments of η due to the property $W(0) = 1$. Section 4.1.2 applies to arbitrary M3 processes but only yields an incremental representation as in (7), for which the incremental process V does not directly represent the conditional increments of η.

4.2. Mixed moving maxima representation of the incremental construction

Theorem 4.3. *Let $\sum_{i\in\mathbb{N}} \delta_{U_i}$ be a Fréchet point process and let W_i, $i \in \mathbb{N}$, be independent copies of a non-negative, sample-continuous process $\{W(t),\ t \in \mathbb{R}^d\}$, satisfying*

$$\lim_{||t||\to\infty} W(t) = 0 \qquad a.s.,$$

$$\mathrm{E}\, W(t) = 1 \qquad \text{for all } t \in \mathbb{R}^d,$$

$$\text{and} \quad \mathrm{E}\left\{\max\nolimits_{t\in K} W(t)\right\} < \infty \qquad \text{for any compact set } K \subset \mathbb{R}^d.$$

Furthermore, let W be Brown-Resnick stationary, i.e., the process ξ, defined by

$$\xi(t) = \max_{i\in\mathbb{N}} U_i W_i(t), \quad t \in \mathbb{R}^d,$$

is stationary with standard Fréchet margins. Then, the following assertions hold:

1. *The random variables*

$$\tau_i = \inf\left\{\arg\sup_{t\in\mathbb{R}^d} W_i(t)\right\} \quad \text{and} \quad \gamma_i = \sup_{t\in\mathbb{R}^d} W_i(t)$$

   are well-defined. Furthermore, $\sum_{i\in\mathbb{N}} \delta_{(U_i\gamma_i, \tau_i, W_i(\cdot+\tau_i)/\gamma_i)}$ is a Poisson point process on $(0,\infty) \times \mathbb{R}^d \times C(\mathbb{R}^d)$ with intensity measure $\Psi(\mathrm{d}u, \mathrm{d}t, \mathrm{d}f) = cu^{-2}\,\mathrm{d}u\ \mathrm{d}t\ \mathrm{P}_{\tilde{F}}(\mathrm{d}f)$ for some $c > 0$ and some probability measure $\mathrm{P}_{\tilde{F}}$.

2. *ξ has an M3 representation with $\mathrm{P}_F(\mathrm{d}f) = \mathrm{P}_{\tilde{F}}(c\,\mathrm{d}f)$ being the probability measure of the shape function F. The constant $c > 0$ is given by*

$$c = \left(\int_{\mathbb{R}^d} \int_{C(\mathbb{R}^d)} f(t)\ \mathrm{P}_{\tilde{F}}(\mathrm{d}f)\,\mathrm{d}t\right)^{-1} \tag{39}$$

and the probability measure $\mathrm{P}_{\tilde F}$ *is defined by*

$$\mathrm{P}_{\tilde F}(A) = \frac{\int_0^\infty y\,\mathrm{P}(W(\cdot+\tau)/y \in A,\ \tau \in K \mid \gamma = y)\ \mathrm{P}_\gamma(\mathrm{d}y)}{\int_0^\infty y\,\mathrm{P}(\tau \in K \mid \gamma = y)\ \mathrm{P}_\gamma(\mathrm{d}y)}$$

for any Borel set $A \subset C(\mathbb{R}^d)$ *and any compact set* $K \subset \mathbb{R}^d$*, where* τ *and* γ *are defined as* τ_i *and* γ_i*, respectively, replacing* W_i *by* W*, and* P_γ *is the probability measure belonging to* γ*.*

Proof. 1. Analogously to the proof of [17, Thm. 14].

2. From the first part it follows that

$$\Phi_0 = \sum_{i\in\mathbb{N}} \delta_{(U_i\gamma_i/c,\,\tau_i,\,c\cdot W_i(\cdot+\tau_i)/\gamma_i)}$$

is a PPP with intensity measure $\Psi_0(\mathrm{d}u, \mathrm{d}t, \mathrm{d}f) = u^{-2}\,\mathrm{d}u \times \mathrm{d}t \times \mathrm{P}_F(\mathrm{d}f)$ where $\mathrm{P}_F(\mathrm{d}f) = \mathrm{P}_{\tilde F}(c\,\mathrm{d}f)$. Hence, Φ_0 is of the same type as Π_0 from the beginning of Section 3 and

$$\xi(t) = \max_{(y,s,f)\in\Phi_0} yf(\cdot - s), \quad t \in \mathbb{R}^d,$$

is a mixed moving maxima representation. The integrability condition (4) follows from the fact that ξ has standard Fréchet marginals. Thus,

$$\int_{\mathbb{R}^d}\int_{C(\mathbb{R}^d)} cf(t)\ \mathrm{P}_{\tilde F}(\mathrm{d}f)\,\mathrm{d}t = 1,$$

which implies (39). In order to calculate $\mathrm{P}_{\tilde F}$, let $A \in \mathcal{B}(C(\mathbb{R}^d))$ and $K \in \mathcal{B}^d$ be compact. The first part of this Theorem implies that

$$\Psi([1,\infty) \times K \times A) = c \cdot |K| \cdot \mathrm{P}_{\tilde F}(A).$$

Therefore,

$$\mathrm{P}_{\tilde F}(A) = \frac{\Psi([1,\infty)\times K \times A)}{\Psi([1,\infty)\times K \times C(\mathbb{R}^d))}, \tag{40}$$

and both the enumerator and the denominator are finite. For the enumerator, we get

$$\begin{aligned}
&\Psi([1,\infty)\times K\times A)\\
&= \int_0^\infty u^{-2}\int_{u^{-1}}^\infty \mathrm{P}(W(\cdot+\tau)/\gamma \in A,\ \tau\in K \mid \gamma = y)\ \mathrm{P}_\gamma(\mathrm{d}y)\,\mathrm{d}u\\
&= \int_0^\infty \int_{y^{-1}}^\infty u^{-2}\,\mathrm{d}u \cdot \mathrm{P}(W(\cdot+\tau)/y \in A,\ \tau\in K \mid \gamma = y)\ \mathrm{P}_\gamma(\mathrm{d}y)\\
&= \int_0^\infty y\,\mathrm{P}(W(\cdot+\tau)/y \in A,\ \tau\in K \mid \gamma = y)\ \mathrm{P}_\gamma(\mathrm{d}y).
\end{aligned}$$

Thus, by (40),

$$\mathrm{P}_{\tilde{F}}(A) = \frac{\int_0^\infty y\,\mathrm{P}(W(\cdot+\tau)/y \in A,\ \tau \in K \mid \gamma = y)\ \mathrm{P}_\gamma(\mathrm{d}y)}{\int_0^\infty y\,\mathrm{P}(\tau \in K \mid \gamma = y)\ \mathrm{P}_\gamma(\mathrm{d}y)},$$

which completes the proof.

5. Outlook: Statistical applications

In univariate extreme value theory, a standard method for estimating the extreme value parameters fits all data exceeding a high threshold to a certain Poisson point process. This peaks-over-threshold approach has been generalized in [22] to the multivariate setting. Therein, generalized multivariate Pareto distributions are obtained as the max-limit of some multivariate random vector in the MDA of an extreme value distribution by conditioning on the event that at least one of the components is large. Conditioning on the same extremal events, the recent contribution [13] analyzes the asymptotic distribution of exceedance counts of stationary sequences.
Here, we have suggested conditioning a stochastic process $\eta(t) : t \in T\}$ in the MDA of a max-stable process $\{\xi(t) : t \in T\}$ such that it converges to the incremental processes W in (3) or the shape functions F in (5). In this section we provide several examples how these theoretical results can be used for statistical inference. The approach is based on a multivariate peaks-over-threshold method for max-stable processes, though the definition of extreme events differs from that in [22, 13].
In the sequel, suppose that $\eta_1, \dots, \eta_n$, $n \in \mathbb{N}$, are independent observations of the random process η, already normalized to standard Pareto margins.

5.1. Incremental representation

For a max-stable process ξ that admits an incremental representation

$$\xi(t) = \max_{i \in \mathbb{N}} U_i W_i(t), \quad t \in T, \tag{41}$$

as in (3), the statistical merit of the convergence results in Theorem 2.1 and Theorem 2.2 is the "deconvolution" of U and W which allows to substitute estimation of ξ by estimation of the process W. As only the single *extreme* events converge to W, we define the index set of extremal observations as

$$I_1(n) = \big\{ i \in \{1, \dots n\} :\ \eta_i(t_0) > a(n) \big\},$$

for some fixed $t_0 \in T$. The set $\{\eta_i(\cdot)/\eta_i(t_0) :\ i \in I_1(n)\}$ then represents a collection of independent random variables that approximately follow the distribution of W. Thus, once the representation in (41) is known, both parametric and non-parametric estimation for the process W is feasible. For statistical inference it is necessary that the number of extremal observations $|I_1(n)|$ converges to ∞, as $n \to \infty$. This is achieved by choosing the sequence of thresholds $a(n)$ according to Remark 2.3.

Example 5.1. (*Symmetric logistic distribution, cf. Example 2.2.*) The dependence parameter $q \geq 1$ of the symmetric logistic distribution (18) can be estimated by perceiving the conditional increments of η in the MDA as realizations of W and maximizing the likelihood

$$\begin{aligned} &\mathrm{P}\left(W(t_1) \in \mathrm{d}s_1, \dots W(t_k) \in \mathrm{d}s_k \,\middle|\, q\right) \\ &= \left(1 + \sum_{i=1}^{k} s_i^{-q}\right)^{1/q-(k+1)} \left(\prod_{i=1}^{k}(iq-1)\right) \prod_{i=0}^{k} s_i^{-q-1}\, \mathrm{d}s_1 \dots \mathrm{d}s_k. \end{aligned}$$

Example 5.2. (*Brown-Resnick processes.*) Recall that the Brown-Resnick processes in Example 2.3 admit a representation (3) with log-Gaussian incremental process $W(t) = \exp\left\{Y(t) - \sigma^2(t)/2\right\}, t \in \mathbb{R}^d$. Hence, standard estimation procedures for Gaussian vectors or processes can be applied for statistical inference. [12] explicitly construct several new estimators of the variogram γ based on the incremental representation, which also covers Hüsler-Reiss distributions, and they provide some basic performance analyses.

5.2. Mixed moving maxima representation

Similarly, in case of the mixed moving maxima representation

$$M(t) = \max_{i \in \mathbb{N}} U_i F_i(t - S_i), \quad t \in \mathbb{R}^d, \tag{42}$$

the convergence results of Theorem 3.1 can be used to estimate F (or $F_1 = F/\lambda$) on some compact domain K instead of estimating M directly. Here, the index set T of the observed processes $\{\eta_i(t) : t \in T\}$, $i = 1, \dots, n$, can be identified with $Q \oplus L$ from Theorem 3.1. The set L should be sufficiently large such that it is reasonable to assume that the components $\{U_i F_i(\cdot - S_i) : S_i \notin Q \oplus L\}$ hardly affect the process M on $Q \oplus K$ (that is, $\mu(C_B \cap A_L)/\mu(A_L) \approx \mathrm{P}(F(\cdot) - \lambda \in B)$ in the proof of Theorem 3.1). At the same time, a large set Q leads to a rich set of usable observations $\widetilde{F}_1^{(i)} = \eta_i(\tau_Q(\eta_i) + \cdot)/\eta_i(\tau_Q(\eta_i)),\ i \in I_2(n)$, where

$$I_2(n) = \left\{ i \in \{1, \dots, n\} : \max_{t \in Q} \eta_i(t) = \max_{t \in Q \oplus L} \eta_i(t) \geq a(n) \right\}.$$

The resulting processes $\widetilde{F}_1^{(i)}$, $i \in I_2(n)$, can be interpreted as independent samples from an approximation to F_1. This approach can be expected to be particularly promising in case of F having a simple distribution or even being deterministic.

Example 5.3. (*M3 processes with deterministic shape functions.*) Some examples of mixed moving maxima processes have already been analyzed for statistical inference by [8] who use normal, exponential and t densities as shape functions. More precisely, they consider M3 models with

$$F_1(t) = \exp\left\{-\frac{\beta^2 t^2}{2}\right\}, \qquad \lambda = \frac{\beta}{\sqrt{2\pi}}, \tag{43}$$

$$F_1(t) = \exp\{-\beta|t|\}, \qquad \lambda = \frac{\beta}{2}, \tag{44}$$

$$\text{and} \quad F_1(t) = \left(1 + \frac{\beta^2 t^2}{\nu}\right)^{-\frac{\nu+1}{2}}, \qquad \lambda = \frac{\beta\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu}\Gamma\left(\frac{\nu}{2}\right)} \quad \nu > 0, \tag{45}$$

all parametrized by $\beta > 0$. [8] introduce consistent and asymptotically normal estimators based on the interpretation of β as a dependence parameter. From the samples $\widetilde{F}_1^{(i)}$, $i \in I_2(n)$, we get a new estimator

$$\widehat{F}_1 = \frac{1}{|I_2(n)|} \sum_{i \in I_2} \widetilde{F}_1^{(i)}$$

for F_1. Applying this estimator, β can be estimated by a least squares fit of (43)–(45) to $\widehat{F}_1$ at some locations $t_1, \ldots t_m \in K$. Note that in case of the normal model (43) and the exponential model (44), the logarithm of the shape function F_1 depends linearly on β^2 and β, respectively, and $\log \widehat{F}_1$ can be fitted by ordinary least squares.

Example 5.4. (*Brown-Resnick processes.*) The mixed moving maxima representation can also be employed for estimation of Brown-Resnick processes although the distribution of F is much more sophisticated than the one of W in the incremental representation (cf. [11, 19]). A relation between the shape function F and the variogram γ of the Brown-Resnick process can be obtained via the *extremal coefficient function* $\theta(\cdot)$. For a stationary, max-stable process ξ with identically distributed marginals, [24] defined the extremal coefficient function θ via the relation

$$\mathrm{P}(\xi(0) \le u, \xi(h) \le u) = \mathrm{P}(\xi(0) \le u)^{\theta(h)}, \quad h \in \mathbb{R}^d.$$

For mixed moving maxima processes, we have

$$\theta(h) = \mathrm{E} \int_{\mathbb{R}^d} \{F(t) \vee F(t+h)\} \,\mathrm{d}t = \frac{\mathrm{E} \int_{\mathbb{R}^d} \{F_1(t) \vee F_1(t+h)\} \,\mathrm{d}t}{\mathrm{E} \int_{\mathbb{R}^d} F_1(t) \,\mathrm{d}t} \tag{46}$$

and, at the same time, for Brown-Resnick processes [17],

$$\theta(h) = 2\Phi\left(\sqrt{\gamma(h)}/2\right), \tag{47}$$

where Φ is the standard Gaussian distribution function. Identifying (46) with (47) and plugging in the samples $\widetilde{F}_1^{(i)}$, $i \in I_2(n)$, we get the variogram estimator

$$\widehat{\gamma}(h) = \left\{ 2\Phi^{-1} \left(\frac{\sum_{i \in I_2(n)} \int_{\widetilde{K}} \widetilde{F}_1^{(i)}(t) \vee \widetilde{F}_1^{(i)}(t+h) \,\mathrm{d}t}{2 \sum_{i \in I_2(n)} \int_{\widetilde{K}} \widetilde{F}_1^{(i)}(t) \,\mathrm{d}t} \right) \right\}^2,$$

where $\widetilde{K}$ is a large set such that $\widetilde{K}, \widetilde{K}+h \subset K$.

Appendix A. Auxiliary Results for the Proof of Theorem 3.1

Lemma A.1. *A_L is closed. The set A_L^* is in the interior of A_L.*

Proof. The first assertion is obvious. For the second one, let $f^* \in A_L^*$. Then, we have $f^*(\tau_Q(f^*|_Q)) =: \alpha > 1$. Furthermore, there is $\delta > 0$ such that $B_\delta(\tau_Q(f^*|_Q)) = \{t \in \mathbb{R}^d : \ ||t - \tau_Q(f|_Q)|| < \delta\} \in Q^o$ and we have

$$\beta := \sup_{t \in Q \oplus L \setminus B_\delta(\tau_Q(f^*|_Q))} f^*(t) - \max_{t \in Q} f^*(t) < 0. \tag{48}$$

Now, we choose $\varepsilon < \min\{\frac{\alpha-1}{2}, \frac{|\beta|}{2}\}$ and show that $B_\varepsilon(f^*) = \{f \in C(Q \oplus L) : ||f - f^*||_\infty < \varepsilon\} \subset A_L$. This holds, as for any $f \in B_\varepsilon(f^*)$, we have

$$f(\tau_Q(f|_Q)) \geq f(\tau_Q(f^*|_Q)) > \alpha - \varepsilon > \frac{1+\alpha}{2} > 1$$

and

$$\max_{t \in Q} f(t) \leq \max_{t \in Q \oplus L} f(t) = \max\left\{ \max_{t \in Q \oplus L \setminus Q^o} f(t), \ \max_{t \in Q} f(t) \right\}$$
$$\leq \max\left\{ \beta + \alpha + \varepsilon, \ \max_{t \in Q} f(t) \right\} \leq \max\left\{ \alpha - \varepsilon, \ \max_{t \in Q} f(t) \right\} = \max_{t \in Q} f(t),$$

which means equality.

Lemma A.2. *The set C_B^* is in the interior of C_B.*

Proof. Let $f^* \in C_B^*$. Then, $t^* = \arg\max_{t \in Q} f^*(t)$ is well-defined and necessarily, as $f \geq 0$,

$$\alpha := f^*(t^*) \in (0, ||f^*||_\infty]. \tag{49}$$

Since $f^*(t^* + \cdot)/f^*(t^*) \in B^o$, there is some $\varepsilon > 0$ such that

$$\left\{ f \in C(K) : \ \left|\left| f^*(t^* + \cdot) \middle/ f^*(t^*) - f \right|\right|_\infty < \varepsilon \right\} \subset B. \tag{50}$$

Furthermore, f^* is uniformly continuous on the compact set $Q \oplus L$, i.e. there exists some $\delta > 0$ such that

$$\sup_{s,t \in Q \oplus L, \ ||s-t|| < \delta} |f^*(s) - f^*(t)| < \frac{\varepsilon}{3}\alpha. \tag{51}$$

Then, as $\arg\max_{t\in Q} f^*(t)$ is unique, we have that

$$\beta := \max_{t\in Q\setminus\{t\in\mathbb{R}^d:\ ||t-t^*||<\delta\}} f^*(t) - f^*(t^*) \in [-\alpha, 0). \tag{52}$$

Choose $\varepsilon^* < \min\left\{\frac{|\beta|}{2\alpha}, \frac{\varepsilon}{6}\frac{\alpha}{||f^*||_\infty}\right\}$. We will show that $B_{\varepsilon^*\alpha}(f^*) = \{f \in C(Q\oplus L) : ||f - f^*||_\infty < \varepsilon^*\alpha\} \subset C_B$. To this end, let $f_0 \in B_{\varepsilon^*\alpha}(f^*)$. Then, because of Equation (52) and $\varepsilon^*\alpha < \frac{|\beta|}{2}$, we have that $||t_0 - t^*|| \leq \delta$ for $t_0 = \tau_Q(f_0|_Q)$. Therefore,

$$\begin{aligned}
&\sup_{t\in K}\left|\frac{f^*(t^*+\cdot)}{f^*(t^*)} - \frac{f_0(t_0+\cdot)}{f_0(t_0)}\right| \\
&\leq \sup_{t\in K}\left|\frac{f^*(t^*+\cdot)}{f^*(t^*)} - \frac{f^*(t_0+\cdot)}{f^*(t^*)}\right| + \sup_{t\in K}\left|\frac{f^*(t_0+\cdot)}{f^*(t^*)} - \frac{f^*(t_0+\cdot)}{f_0(t_0)}\right| \\
&\quad + \sup_{t\in K}\left|\frac{f^*(t_0+\cdot)}{f_0(t_0)} - \frac{f_0(t_0+\cdot)}{f_0(t_0)}\right| \qquad \leq \qquad \frac{\varepsilon}{3} + \frac{\varepsilon^*}{1-\varepsilon^*}\frac{||f^*||_\infty}{\alpha} + \frac{\varepsilon^*}{1-\varepsilon^*},
\end{aligned} \tag{53}$$

where we used Equation (51) and the fact that $f_0 \in B_{\varepsilon^*\alpha}(f^*)$. Equations (49) and (52) and the choice of ε^* yield that $\varepsilon^*/(1-\varepsilon^*) \leq (\varepsilon^*||f^*||_\infty)/((1-\varepsilon^*)\alpha) \leq 2\varepsilon^*||f^*||_\infty/\alpha < \varepsilon/3$, i.e. each summand on the right-hand side of (53) is smaller than $\varepsilon/3$. Thus, $f_0(t_0+\cdot)/f_0(t_0) \in \{f \in C(K) :\ ||f^*(t^*+\cdot)/f^*(t^*) - f||_\infty < \varepsilon\} \subset B$ by Equation (50) and $f_0 \in C_B$.

Lemma A.3. *The closure of C_B is a subset of*

$$\begin{aligned}
B^* \cup &\left\{f \in C(Q\oplus L) :\ \arg\max_{t\in Q} f(t) \text{ is not unique}\right\} \\
&\cup \left\{f \in C(Q\oplus L) :\ f(\tau_Q(f|_Q)+\cdot)\Big/ f(\tau_Q(f|_Q)) \in \partial B\right\}.
\end{aligned}$$

Proof. Let $\{f_n\} \subset C_B$ be a sequence converging uniformly to some $f^* \in C(Q\oplus L)$. We have to verify that $f^*(\tau_Q(f^*|_Q)+\cdot)/f^*(\tau_Q(f^*|_Q)) \in B \cup \partial B$ if $\arg\max_{t\in Q} f^*(t)$ is unique. Analogously to the proof of Lemma A.2 we can show that for any $\varepsilon_2 > 0$ there is some $\varepsilon_1 > 0$ such that

$$\begin{aligned}
&||f - f^*||_{\infty,Q\oplus L} < \varepsilon_1 \\
\Longrightarrow \quad &\left|\left|\frac{f(\tau_Q(f|_Q)+\cdot)}{f(\tau_Q(f|_Q))} - \frac{f^*(\tau_Q(f^*|_Q)+\cdot)}{f^*(\tau_Q(f^*|_Q))}\right|\right|_{\infty,K} < \varepsilon_2.
\end{aligned}$$

Thus, $f_n(\tau_Q(f_n|_Q)+\cdot)/f_n(\tau_Q(f_n|_Q))$ converges to $f^*(\tau_Q(f^*|_Q)+\cdot)/f^*(\tau_Q(f^*|_Q))$ in $C(K)$. Hence, as $B\cup\partial B$ is closed, $f^*(\tau_Q(f^*|_Q)+\cdot)/f^*(\tau_Q(f^*|_Q)) \in B\cup\partial B$.

Appendix B. Proof of Example 3.1

Proof of Example 3.1. Let $\tilde{M}_j$, $j \in \mathbb{N}$, be independent copies of the process $\tilde{M}$ and consider

$$M_n(\cdot) = \frac{1}{cn} \max_{i=1}^{n} \tilde{M}_i(\cdot).$$

Further, suppose that $L \subset \mathbb{R}^d$ is an arbitrary compact set. Note that by Remark 2.5 it suffices to show weak convergence of $M_n \xrightarrow{\mathrm{D}} M$, $n \to \infty$, on $C(L)$.

To prove the first assertion, note that, for $t_0, \ldots, t_k \in \mathbb{R}^d$, $s_0, \ldots, s_k \geq 0$, $k \in \mathbb{N}$, we have

$$\begin{aligned}
&\mathrm{P}(M_n(t_0) \leq s_0, \ldots, M_n(t_k) \leq s_k) \\
&= \left[\mathbf{1}_{\kappa \leq \min_{l=0}^k cns_l} \cdot \mathrm{P}\left\{ \Pi_{c,\epsilon}\left(\left\{ (u,t,f) : \max_{l=0}^{k} uf(t_l - t)/(cns_l) > 1 \right\} \right) = 0 \right\} \right]^n \\
&= \mathbf{1}_{\kappa \leq \min_{l=0}^k cns_l} \cdot \exp\left(-n \int_{C(\mathbb{R}^d)} \int_{\mathbb{R}^d} \min\left\{ \frac{1}{\epsilon}, \max_{l=0}^{k} \frac{f(t_l - t)}{cns_l} \right\} c \,\mathrm{d}t \; \mathrm{P}_F(\mathrm{d}f) \right) \\
&\longrightarrow \exp\left(- \int_{C(\mathbb{R}^d)} \int_{\mathbb{R}^d} \max_{l=0}^{k} (f(t_l - t)/s_l) \,\mathrm{d}t \; \mathrm{P}_F(\mathrm{d}f) \right),
\end{aligned} \tag{54}$$

as $n \to \infty$, where the convergence holds due to monotone convergence. The right-hand side of (54) coincides with the marginal distribution of M (cf. (23)).

For convergence of M_n to M in the sense of weak convergence in $C^+(L)$ endowed with the topology of uniform convergence, it remains to show that the sequence of restricted processes $\{M_n|_L : n \in \mathbb{N}\}$ is tight. To this end, by [2, Thm. 7.3], it suffices to verify that for any $\varepsilon > 0$, $\eta \in (0,1)$, there exist $\delta > 0$, $n_0 \in \mathbb{N}$ such that

$$\mathrm{P}\left\{ \sup_{||s-t||<\delta} |M_n(s) - M_n(t)| \geq \varepsilon \right\} \leq \eta, \qquad n \geq n_0.$$

By Equation (25), we can choose $R > 0$ such that

$$\int_{\mathbb{R}^d \setminus (L \oplus B_R(\mathbf{0}))} \mathrm{E}\left(\sup_{t \in L} F(t-s) \right) \mathrm{d}s < \frac{\varepsilon\eta}{2}, \tag{55}$$

where $B_R(\mathbf{0}) = \{x \in \mathbb{R}^d : \ ||x|| \leq R\}$. Furthermore, (25) implies that $\mathrm{E}\left(\sup_{t\in K} F(s)\right) < \infty$ for any compact set $K \subset \mathbb{R}^d$. Therefore, as each realization of F is uniformly continuous on $B_{R+d(L)}(\mathbf{0})$, where $d(L) = \sup_{s_1,s_2\in L} ||s_1 - s_2||$ denotes the diameter of L, dominated convergence yields

$$\lim_{\delta \searrow 0} \mathrm{E}\left(\sup_{s,t \in B_{R+d(L)}(\mathbf{0}),\ ||s-t||<\delta} |F(s) - F(t)| \right) = 0.$$

In particular, we can choose $\delta > 0$ such that

$$\mathrm{E}\left(\sup_{s_1,s_2 \in B_{R+d(L)}(\mathbf{0}),\ ||s_1-s_2||<\delta} |F(s_1) - F(s_2)| \right) < \frac{\varepsilon\eta}{2|L \oplus B_R(\mathbf{0})|}. \tag{56}$$

Then, we get

$$\begin{aligned}
&\mathrm{P}\left\{ \sup_{||s_1-s_2||<\delta,\ s_1,s_2\in L} |M_n(s_1) - M_n(s_2)| \geq \varepsilon \right\} \\
&\leq n\,\mathrm{P}\left\{ \sup_{||s_1-s_2||<\delta,\ s_1,s_2\in L} |\tilde{M}_n(s_1) - \tilde{M}_n(s_2)| \geq cn\varepsilon \right\} \\
&\leq n\Bigg(\mathrm{P}\Bigg\{ \Pi\Bigg(\Bigg\{ (u,t,f):\ t \in L \oplus B_R(\mathbf{0}), \\
&\qquad\qquad \sup_{\substack{s_1,s_2\in B_{R+d(L)}(\mathbf{0}),\\ ||s_1-s_2||<\delta}} |f(s_1) - f(s_2)| > \frac{cn\varepsilon}{u} \Bigg\} \Bigg) > 0 \Bigg\} \\
&\quad + \mathrm{P}\Bigg\{ \Pi\Bigg(\Bigg\{ (u,t,f):\ t \in \mathbb{R}^d \setminus (L \oplus B_R(\mathbf{0})), \sup_{s\in L} |f(s-t)| > \frac{cn\varepsilon}{u} \Bigg\} \Bigg) > 0 \Bigg\} \Bigg) \\
&\leq n\Bigg(1 - \exp\Bigg\{ - \int_{L\oplus B_R(\mathbf{0})} \int_{\epsilon}^{\infty} u^{-2} \\
&\qquad\qquad \cdot \mathrm{P}\Bigg(\sup_{\substack{s_1,s_2\in B_{R+d(L)}(\mathbf{0}),\\ ||s_1-s_2||<\delta}} |F(s_1) - F(s_2)| > \frac{cn\varepsilon}{u} \Bigg) \,\mathrm{d}u\, c\,\mathrm{d}t \Bigg\} \\
&\quad + 1 - \exp\Bigg\{ - \int_{\mathbb{R}^d\setminus(L\oplus B_R(\mathbf{0}))} \int_{\epsilon}^{\infty} u^{-2}\, \mathrm{P}\Bigg(\sup_{s\in L} |F(s-t)| > \frac{cn\varepsilon}{u} \Bigg) \,\mathrm{d}u\, c\,\mathrm{d}t \Bigg\} \Bigg) \\
&\leq n\Bigg(1 - \exp\Bigg(- \frac{|L \oplus B_R(\mathbf{0})|}{n\varepsilon}\, \mathrm{E}\Bigg\{ \sup_{\substack{s_1,s_2\in B_{R+d(L)}(\mathbf{0}),\\ ||s_1-s_2||<\delta}} |F(s_1) - F(s_2)| \Bigg\} \Bigg) \\
&\quad + 1 - \exp\Bigg(- \frac{1}{n\varepsilon} \int_{\mathbb{R}\setminus(L\oplus B_R(\mathbf{0}))} \mathrm{E}\Bigg\{ \sup_{s\in L} |F(s-t)| \Bigg\} \,\mathrm{d}t \Bigg) \Bigg) \\
&\leq n\left(1 - \exp\left(-\frac{\eta}{2n} \right) + 1 - \exp\left(-\frac{\eta}{2n} \right) \right) \quad \leq \quad \eta,
\end{aligned}$$

where we used Equation (56) and (55). Thus, the sequence of processes $\{M_n|_L : n \in \mathbb{N}\}$ is tight.

Acknowledgements

The authors are grateful to Zakhar Kabluchko for useful suggestions and hints. S. Engelke has been financially supported by Deutsche Telekom Stiftung. A. Malinowski has been financially supported the German Science Foundation (DFG), Research Training Group 1644 'Scaling problems in Statistics'.

M. Oesting and M. Schlather have been financially supported by Volkswagen Stiftung within the project 'WEX-MOP'.

References

[1] Adler, R. J. and Taylor, J. E. (2007). *Random fields and geometry.* Springer Monographs in Mathematics. Springer, New York.

[2] Billingsley, P. (1999). *Convergence of Probability Measures* 2nd ed. John Wiley & Sons Inc., New York.

[3] Blanchet, J. and Davison, A. C. (2011). Spatial modeling of extreme snow depth. *Ann. Appl. Stat.* **5,** 1699–1725.

[4] Brown, B. M. and Resnick, S. I. (1977). Extreme values of independent stochastic processes. *J. Appl. Probab.* **14,** 732–739.

[5] Davison, A. C. and Gholamrezaee, M. M. (2012). Geostatistics of extremes. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **468,** 581–608.

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

[7] de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory.* Springer, New York.

[8] de Haan, L. and Pereira, T. T. (2006). Spatial extremes: models for the stationary case. *Ann. Statist.* **34,** 146–168.

[9] Ehlert, A., Fiebig, U.-R., Janssen, A. and Schlather, M. (2011). Joint Extremal Behavior of Hidden and Observable Time Series with an Application to GARCH Processes. Available from `http://arxiv.org/abs/1107.0493`.

[10] Embrechts, P., Mikosch, T. and Klüppelberg, C. (1997). *Modelling Extremal Events: for Insurance and Finance.* Springer, London.

[11] Engelke, S., Kabluchko, Z. and Schlather, M. (2011). An equivalent representation of the Brown-Resnick process. *Statist. Probab. Lett.* **81,** 1150 – 1154.

[12] Engelke, S., Malinowski, A., Kabluchko, Z. and Schlather, M. (2012). Estimation of Hüsler-Reiss distributions and Brown-Resnick processes. Available from `http://arxiv.org/abs/1207.6886`.

[13] Falk, M. and Tichy, D. (2012). Asymptotic conditional distribution of exceedance counts. *Adv. Appl. Probab.* **44,** 270–291.

[14] Gumbel, E. J. (1960). Distributions des valeurs extrêmes en plusieurs dimensions. *Publ. Inst. Statist. Univ. Paris* **9,** 171–173.

[15] Hüsler, J. and Reiss, R.-D. (1989). Maxima of normal random vectors: between independence and complete dependence. *Statist. Probab. Lett.* **7,** 283–286.

[16] Kabluchko, Z. (2011). Extremes of independent Gaussian processes. *Extremes* **14,** 285–310.

[17] Kabluchko, Z., Schlather, M. and de Haan, L. (2009). Stationary max-stable fields associated to negative definite functions. *Ann. Probab.* **37,** 2042–2065.

[18] Leadbetter, M. (1991). On a basis for 'peaks over threshold' modeling. *Stat. Probab. Lett.* **12,** 357 – 362.

[19] Oesting, M., Kabluchko, Z. and Schlather, M. (2012). Simulation of Brown-Resnick processes. *Extremes* **15,** 89–107.

[20] Padoan, S. A., Ribatet, M. and Sisson, S. A. (2010). Likelihood-based inference for max-stable processes. *J. Amer. Statist. Assoc.* **105,** 263–277.

[21] Resnick, S. I. (2008). *Extreme Values, Regular Variation and Point Processes.* Springer, New York.

[22] Rootzén, H. and Tajvidi, N. (2006). Multivariate generalized Pareto distributions. *Bernoulli* **12,** 917–930.

[23] Schlather, M. (2002). Models for stationary max-stable random fields. *Extremes* **5,** 33–44.

[24] Schlather, M. and Tawn, J. A. (2003). A dependence measure for multivariate and spatial extreme values: Properties and inference. *Biometrika* **90,** 139–156.

[25] Segers, J. (2007). Multivariate regular variation of heavy-tailed markov chains. *Technical report.* Institut de statistique, Université catholique de Louvain. Available from `http://arxiv.org/abs/math/0701411`.

[26] Smith, R. (1990). Max-stable processes and spatial extremes. Unpublished manuscript.

[27] Wang, Y. and Stoev, S. A. (2010). On the structure and representations of max-stable processes. *Adv. Appl. Probab.* **42,** 855–877.

[28] Whitt, W. (1970). Weak Convergence of Probability Measures on the Function Space $C[0,\infty)$. *Ann. Math. Statist.* **41,** 939–944.