Conditioned limit laws for inverted max-stable processes

Ioannis Papastathopoulos [a,b,*], Jonathan A. Tawn [c]

[a] *School of Mathematics and Maxwell Institute, University of Edinburgh, Edinburgh, EH9 3FD, United Kingdom*

[b] *The Alan Turing Institute for Data Science, British Library, London, NW1 2DB, United Kingdom*

[c] *Department of Mathematics and Statistics, Lancaster University, Lancaster, LA1 4YF, United Kingdom*

ABSTRACT

Max-stable processes are widely used to model spatial extremes. These processes exhibit asymptotic dependence meaning that the large values of the process can occur simultaneously over space. Recently, inverted max-stable processes have been proposed as an important new class for spatial extremes which are in the domain of attraction of a spatially independent max-stable process but instead they cover the broad class of asymptotic independence. To study the extreme values of such processes we use the conditioned approach to multivariate extremes that characterises the limiting distribution of appropriately normalised random vectors given that at least one of their components is large. The current statistical methods for the conditioned approach are based on a canonical parametric family of location and scale norming functions. We study broad classes of inverted max-stable processes containing processes linked to the widely studied max-stable models of Brown–Resnick and extremal-t, and identify conditions for the normalisations to either belong to the canonical family or not. Despite such differences at an asymptotic level, we show that at practical levels, the canonical model can approximate well the true conditional distributions.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

Extreme environmental events, such as hurricanes, heatwaves, flooding and droughts, can cause havoc to the people affected and typically result in large financial losses. The impact of this type of event is often exacerbated by the event being severe over a large spatial region. The statistical modelling of spatial extremes is a rapidly evolving area [6] and is crucial to understanding, visualising and predicting the extremes of stochastic processes. The approach that is currently most used for modelling spatial extreme values assumes the environmental process is a max-stable process [8]. The most widely used max-stable processes are the Brown–Resnick [2,18,15], and extremal-t [26,27] processes. These include the Smith process [34] and extremal Gaussian process [33] as special cases.

Max-stable processes are the only non-trivial limits of point-wise normalised maxima of independent and identically distributed realisations of stochastic processes. When max-stable processes are observed at a finite number of locations their joint distribution is a multivariate extreme value distribution, which is underpinned by the assumption of the original variables satisfying the dependence structure conditions of multivariate regular variation [30]. Max-stable processes have marginal generalised extreme value distributions [3] and a complex non-negative dependence structure which has a restricted form. To understand this restriction, let $\{X_M(s), \ s \in \mathbb{R}^2\}$ be a spatial max-stable process with continuous marginal

* Corresponding author at: School of Mathematics and Maxwell Institute, University of Edinburgh, Edinburgh, EH9 3FD, United Kingdom.
  E-mail addresses: i.papastathopoulos@ed.ac.uk (I. Papastathopoulos), j.tawn@lancaster.ac.uk (J.A. Tawn).

distribution functions F_s and corresponding inverse denoted by $F_s^{\leftarrow}$. Then, if for any $s_1, s_2 \in \mathbb{R}^2$ where $X_M(s_1)$ and $X_M(s_2)$ are not independent, it follows that the dependence coefficient

$$\chi = \lim_{p \to 1} \Pr \left\{ X_M(s_2) > F_{s_2}^{\leftarrow}(p) \mid X_M(s_1) > F_{s_1}^{\leftarrow}(p) \right\},$$

is positive. This property is termed asymptotic dependence, and it implies that the largest marginal values at different locations can occur simultaneously. Furthermore, if we denote the joint distribution function of $\{X_M(s_1), X_M(s_2)\}$ by F_{s_1, s_2} then

$$F_{s_1, s_2} \{ F_{s_1}^{\leftarrow}(p), F_{s_2}^{\leftarrow}(p) \} = p^{\theta(s_1, s_2)}$$

for all $0 < p \le 1$ where $1 \le \theta(s_1, s_2) \le 2$, so for max-stable processes this implies that the spatial dependence properties, measured by θ are independent of the severity of the event, measured by p [6].

Wadsworth and Tawn [38] introduced the class of inverted max-stable processes and these are used in Davison et al. [5] and Thibaud et al. [37]. Any inverted max-stable process $\{X(s), s \in \mathbb{R}^2\}$ with unit exponential margins, i.e., for all $s \in \mathbb{R}^2$ and $x > 0$, $\Pr(X(s) < x) = 1 - \exp(-x)$, can be represented by

$$X(s) = 1/X_F(s) \quad s \in \mathbb{R}^2,$$

where $\{X_F(s), s \in \mathbb{R}^2\}$ is a max-stable process with unit Fréchet margins, i.e., for all $s \in \mathbb{R}^2$ and $x > 0$, $\Pr(X_F(s) < x) = \exp(-1/x)$. Thus, for all $s_1, s_2 \in \mathbb{R}^2$, the dependence structure between large $X(s_1)$ and $X(s_2)$ is equivalent to the dependence structure between small $X_F(s_1)$ and $X_F(s_2)$ and hence differs from the max-stable form. As all max-stable processes are non-negatively associated, so are all inverted max-stable processes. Furthermore, for all $s_1, s_2 \in \mathbb{R}^2$, with $s_1 \ne s_2$, for $x > 0$

$$\Pr \{ X(s_1) > x, X(s_2) > x \} = \exp\{ -x/\eta(s_1, s_2) \} = \{ \Pr(X > x) \}^{1/\eta(s_1, s_2)},$$

where $\eta := \eta(s_1, s_2) \in [1/2, 1)$ is the coefficient of tail dependence [23]. It follows that in general margins the inverted max-stable distribution is

$$F_{s_1, s_2} \{ F_{s_1}^{\leftarrow}(p), F_{s_2}^{\leftarrow}(p) \} = 1 - 2(1 - p) + (1 - p)^{1/\eta} \sim p^{2 - (1-p)^{-1+1/\eta}} \quad \text{as } p \uparrow 1,$$

so unlike for the max-stable distribution the spatial dependence, measured by $2 - (1 - p)^{-1+1/\eta}$, is not independent of the severity p. In fact, all non-perfectly dependent inverted max-stable processes are in the domain of attraction of spatially independent max-stable processes, meaning that their point-wise normalised maxima are independent, i.e., for all $s_1, s_2 \in \mathbb{R}^2$, with $s_1 \ne s_2$,

$$\lim_{n \to \infty} \Pr \left(\max_{i=1,\dots,n} X_i(s_1) - \ln n < x_1, \max_{i=1,\dots,n} X_i(s_2) - \ln n < x_2 \right) = \prod_{i=1}^{2} \exp\{ -\exp(-x_i) \}, \tag{1}$$

for any $x_1, x_2 \in \mathbb{R}$, where $\{X_i(s), s \in \mathbb{R}^2\}$, for $i = 1, \dots, n$ denotes a sequence of independent and identically distributed inverted max-stable processes with unit exponential margins.

To reveal the extremal dependence structure for asymptotically independent random variables, alternative asymptotic properties have been studied. Ledford and Tawn [22,23], Resnick [31] and Wadsworth and Tawn [39] explore how the joint survivor decays as both arguments tend to the upper end point. A weakness with this approach is that it fails to describe the behaviour of the $X(s_2)$ values that occur with the largest values of $X(s_1)$. Instead a conditioned approach is required which looks at a more subtle normalisation for $X(s_2)$ that focuses on the region of the joint distribution which is most likely when conditioning on variable $X(s_1)$ being large. This is the approach we take in this paper.

For a bivariate random variable (X, Y) with unit exponential margins and general dependence structure, the conditioned extremes limit theory of Heffernan and Tawn [13] and Heffernan and Resnick [12] is equivalent to the assumption that there exist location and scaling norming functions $a : \mathbb{R}_+ \mapsto \mathbb{R}$ and $b : \mathbb{R}_+ \mapsto \mathbb{R}_+$, such that, for any $x > 0$ and $z \in \mathbb{R}$,

$$\lim_{u \to \infty} \Pr \{ X - u > x, \{ Y - a(X) \}/b(X) < z \mid X > u \} = \exp(-x) G(z), \tag{2}$$

where G is a non-degenerate distribution function. To ensure a, b and G are uniquely defined the condition $\lim_{z \to \infty} G(z) = 1$ is required, so G places no mass at $+\infty$ but some mass is allowed at $-\infty$.

The key development in this paper is deriving the forms of a, b and G in representation (2) for the broad class of inverted max-stable processes. For positively dependent random variables, Heffernan and Tawn [13] found that, for all the standard copula models studied by Joe [17] and Nelsen [25], the norming functions $a(x)$ and $b(x)$, fell into the simple canonical parametric family

$$a(x) = \alpha x \quad \text{and} \quad b(x) = x^{\beta}, \tag{3}$$

where $\alpha \in [0, 1]$ and $\beta \in (-\infty, 1)$ with Keef et al. [19] identifying additional joint constraints. The case $\alpha = 1$ and $\beta = 0$ corresponds to $\chi > 0$, whereas any other combination of α and β gives $\chi = 0$. The statistical use of the canonical family (3)

has been the subject of criticism since the functions a and b seem to be 'proof by example' rather than a general result [35]. Heffernan and Resnick [12] note that under assumption (2) and in unit exponential marginals,

$$\lim_{t\to\infty} b(t+x)/b(t) = \psi_1(x) \quad \text{and} \quad \lim_{t\to\infty} \{a(t+x) - a(t)\}/b(t) = \psi_2(x),$$

for any $x \in \mathbb{R}$, where ψ_1 and ψ_2 are real functions. For the canonical family (3), it readily follows that $(\psi_1, \psi_2) = (1, 0)$ if $\beta > 0$ and $(\psi_1, \psi_2) = (1, \alpha x)$ if $\beta = 0$ and this condition is also satisfied by a range of regularly varying functions. Until our paper no examples have been published that satisfy these broader conditions than the canonical form (3). We will show that the inverted max-stable process, with the Brown–Resnick model requires this more general normalisation but that with extremal-t model does not. This is practically important as these two dependence structures are the most widely used.

For simplicity we focus on bivariate characterisations of the inverted max-stable process, with all joint distributions following inverted bivariate extreme value distributions. For the rest of the paper we denote by (X, Y) a bivariate random variable with inverted bivariate extreme value distribution with unit exponential margins. We find classes of this family where the normalisation required to achieve property (2) either falls in the canonical family (3) or is in the more general form, with the class determined by the behaviour of the spectral measure of the underlying max-stable process near its lower end point.

The statistical conditioned model of Heffernan and Tawn [13] assumes that the limiting relationship (2) holds exactly for all values $X > u$ for a suitably high threshold u with the functions a and b in canonical form (3). This model has been found to fit well in various applications [29,20,14,9,28]. Our identification of the existence of the new classes that do not fall in canonical family (3) questions the validity of the generic use of the canonical family for statistical modelling. Therefore, we compare the new models with the current statistical approach of Heffernan and Tawn [13] and show, through simulation, that at practical levels the canonical family gives good approximations to the conditional distribution of the inverted max-stable model with Brown–Resnick dependence and highlight examples where a good approximation does not hold.

A key question is why have we made the restriction of unit exponential margins on the inverted max-stable process. In the style of copula methods [25] we assume identical margins. In Section 2.2 we discuss the implications of the marginal choice, how our results tie in with others who use different marginals, and in particular show how once the limit relationship has been derived for unit exponential margins alternative limit results follow immediately for other non-identically distributed marginal choices.

The paper is structured as follows. In Section 2 we present the classes of max-stable and inverted max-stable distributions and explore the implications of deriving results on general margins. In Section 3 we present the conditional representation of the class of inverted max-stable distributions with spectral densities of the associated max-stable process being regularly varying and Γ-varying spectral densities at their lower end-point. In Section 4 we discuss the spatial extension of our results. Our derivations and proofs are included in the Appendix.

2. Bivariate inverted max-stable distributions

2.1. Max-stable and inverted max-stable distributions

Max-stable distributions arise naturally as the only non-degenerate limit distributions of appropriately normalised component-wise maxima of random vectors. In unit Fréchet margins, and for $x, y > 0$, a bivariate max-stable distribution function is defined by

$$F(x, y) = \exp\{-V(x, y)\} = \exp\left[-\int_0^1 \max\{w/x, (1-w)/y\}\, dH(w)\right], \tag{4}$$

where V is termed the exponent measure and H is an arbitrary finite measure on $[0, 1]$, known as the spectral measure, with total mass 2, satisfying the marginal moment constraint $\int_0^1 w\, dH(w) = 1$. Coles and Tawn [4] showed that if F has a density, then $H(w)$ has spectral density $h(w)$ on the interior $(0, 1)$ and can have mass $H(\{k\})$, $k = 0, 1$, on each of $\{0\}$ and $\{1\}$, given by

$$h(w) = -\frac{\partial^2 V}{\partial x \partial y}(w, 1-w) \quad 0 < w < 1,$$

$$H(\{0\}) = -y^2 \lim_{x\to 0} \frac{\partial V}{\partial y}(x, y), \quad \text{and} \quad H(\{1\}) = -x^2 \lim_{y\to 0} \frac{\partial V}{\partial x}(x, y).$$

As the class of bivariate max-stable distributions does not admit a finite dimensional parameterisation, a natural method for modelling the spectral measure H of expression (4) relies on constructing parametric sub-classes of models that are flexible enough to approximate any member from the class [4,1]. Two such sub-models are the Brown–Resnick and the extremal-t max-stable distributions which have exponent measures, for $x, y > 0$,

$$V(x, y) = \frac{1}{x}\Phi\left\{\frac{\lambda}{2} + \frac{1}{\lambda}\ln\left(\frac{y}{x}\right)\right\} + \frac{1}{y}\Phi\left\{\frac{\lambda}{2} + \frac{1}{\lambda}\ln\left(\frac{x}{y}\right)\right\}, \tag{5}$$

$$V(x, y) = \frac{1}{x}T_{v+1}\left[\frac{(y/x)^{1/v} - \rho}{\{(1-\rho^2)/(v+1)\}^{1/2}}\right] + \frac{1}{y}T_{v+1}\left[\frac{(x/y)^{1/v} - \rho}{\{(1-\rho^2)/(v+1)\}^{1/2}}\right], \tag{6}$$

respectively, where $\lambda > 0$, $\rho \in (-1, 1)$, $\nu > 0$, Φ is the cumulative distribution function of the standard normal distribution and $T_{\nu+1}$ is the distribution function of the standard-t distribution with $\nu + 1$ degrees of freedom. These are the exponent measures of the pairwise distributions for the Brown–Resnick and extremal-t max-stable models respectively. The parameters λ, ρ and ν control the strength of dependence. In particular, increasing values of ρ and decreasing values of λ and ν imply stronger dependence between X and Y.

Given a max-stable distribution with exponent measure V as in Eq. (4), the bivariate random variable (X, Y) follows the inverted max-stable distribution with unit exponential margins if, for $x, y > 0$, its joint survivor function is,

$$\Pr(X > x, Y > y) = \exp\{-V(1/x, 1/y)\}. \tag{7}$$

As

$$\Pr(X > x, Y > x) = \exp(-x/\eta) = \{\Pr(X > x)\}^{1/\eta}, \quad \eta = 1/V(1, 1) \in [1/2, 1],$$

the inverted max-stable distributions are either perfectly dependent when $V(1, 1) = 1$ or asymptotically independent when $V(1, 1) > 1$. Specifically η is the coefficient of tail dependence [23]. This property explains the independence in limit (1) if $\eta < 1$.

2.2. Conditional representation with different marginal distributions

We study the conditional extremal dependence for the inverted max-stable distribution through the limiting representation (2) for unit exponential margins. Here we compare this approach to that used by others and show how our results can simply be extended to arbitrary non-identically distributed marginals.

Conditioned limit theory (2) has been studied with the margins (X, Y) taken as identically distributed Gumbel variables [13], which is asymptotically equivalent to unit exponential margins but mathematically less clean, whereas Eastoe and Tawn [9] take identically distributed generalised Pareto distributions. In contrast Heffernan and Resnick [12] work with X in the domain of attraction of the generalised extreme value distribution, but do not impose any constraint on Y, and find that it is not always possible to achieve an affine normalisation as in limit (2) after marginal transformation to identical margins. Kulik and Soulier [21] consider (X, Y) with margins that have regularly varying tails, with location function $a(x) = 0$.

Next we show how once the limit relationship has been derived for unit exponential margins alternative limit results follow immediately for other marginal choices, and vice versa. The key difference is that if a non-trivial location and scale affine transformation are required in exponential marginals, then the class of affine transformations is not sufficient for regularly varying marginal distributions. In particular only in the case where $b(\cdot) = 1$ in exponential margins does the affine norming of Kulik and Soulier [21] hold for margins that have regularly varying tails.

Let (X, Y) be a bivariate random variable with common unit exponential margins and assume that limit (2) holds. Here we consider what this representation then implies for the extremal conditioned distribution of $Y_{H_2} \mid X_{H_1}$, where the bivariate random variable (X_{H_1}, Y_{H_2}) has continuous marginal distribution functions H_1 and H_2 respectively and has identical copula to (X, Y). For $i = 1, 2$, let $K_i(y) = -\ln\{1 - H_i(y)\}$ for $y \in \mathbb{R}$ and denote its inverse by $K_i^\leftarrow(y) = H_i^\leftarrow\{1 - \exp(-y)\}$ for $y > 0$. Then $(X_{H_1}, Y_{H_2}) = \{K_1^\leftarrow(X), K_2^\leftarrow(Y)\}$ have the required joint distribution and from limit (2) there exist functions $a : \mathbb{R}_+ \mapsto \mathbb{R}$ and $b : \mathbb{R}_+ \mapsto \mathbb{R}_+$ such that for all $z \in \mathbb{R}$

$$\lim_{u \to \infty} \Pr\left(Y_{H_2} < K_2^\leftarrow\left[a\{K_1(X_{H_1})\} + b\{K_1(X_{H_1})\}z\right] \mid X_{H_1} > u\right) = G(z) \tag{8}$$

for non-degenerate $G(z)$. Therefore, if we can find a location-scale normalisation when working with identical unit exponential margins then limit (8) shows that these results directly provide the appropriate conditioned limit for non-identically distributed marginals. Furthermore, limit (8) shows that in general margins a location-scale normalisation is not always possible even when it can be achieved with unit exponential margins. Of course the converse is true, but as we will see in Section 3 for the class of inverted max-stable distributions limit (2) holds with unit exponential margins, so limit (8) is useful to give the conditioned limits in other marginals.

To help understand the implications of limit (8) it is helpful to focus on specific forms for H_1 and H_2. Kulik and Soulier [21] work with regularly varying tails. Focusing on the specific case of Pareto margins with $H_i(y) = 1 - y^{-\alpha_i}$ for $\alpha_i > 0$, $y > 1$ and $i = 1, 2$, limit (8) then becomes

$$\lim_{u \to \infty} \Pr\left\{Y_{H_2} < \exp\left([a\{\alpha_1 \ln(X_{H_1})\} + b\{\alpha_1 \ln(X_{H_1})\}z]/\alpha_2\right) \,\Big|\, X_{H_1} > u\right\} = G(z)$$

so a location-scale normalisation in these margins can only be achieved with $b(y) \equiv 1$ and then only a scaling is required. Thus studying limit (2) with regularly varying tails only requires a scaling but critically cannot cover any cases where the b scaling function when using unit exponential margins differs from $b(y) \equiv 1$. As we will see in Section 3, for the class of inverted max-stable distributions with unit exponential margins we have many classes where $b(y) \neq 1$. Therefore, to reveal the full structure of the conditioned extremal behaviour of inverted max-stable distributions revealed by limit (2), it is essential to work in marginal variables which are tail equivalent to the unit exponential. Hence for the remainder of the paper we work exclusively with unit exponential margins acknowledging that these results apply directly with different margins using limit (8).

3. Conditional representations

3.1. Conditional distribution preliminaries

A range of results are available to study the conditional limit (2). In particular, Heffernan and Resnick [12], Resnick and Zeber [32] and Wadsworth et al. [40] show that under various conditions, all of the following limits are identical to $G(z)$ in limit (2):

$$\lim_{u \to \infty} \Pr \left\{ \frac{Y - a(u)}{b(u)} < z \Big| X > u \right\} = \lim_{u \to \infty} \Pr \left\{ \frac{Y - a(X)}{b(X)} < z \Big| X > u \right\} = \lim_{u \to \infty} \Pr \left\{ \frac{Y - a(u)}{b(u)} < z \Big| X = u \right\}. \tag{9}$$

These conditions hold for all inverted max-stable distributions if the associated spectral measure of the max-stable process places no point mass on the interval $(0, 1)$, i.e., when (X, Y) have a joint density. So we can use any of these limits to derive the forms of a, b and G. Motivated by statistical considerations, we find that the last expression is most simple to use. However, this is the most restrictive in general as it requires the assumption of a joint density. As all the parametric max-stable models have joint densities so do the associated inverted max-stable distributions. The requirement for a joint density is not restrictive here. For using this third limit form it is helpful to note that the conditional survivor function is

$$\Pr(Y > y \mid X = x) = -V_1(1, x/y) \exp\{x - xV(1, x/y)\} \quad y > 0, \tag{10}$$

where $V_1(x, y) = \partial V(x, y)/\partial x$.

3.2. Known representations

Heffernan and Tawn [13] explored the conditional representation (2) for the class of inverted max-stable distributions subject to the assumption that H places no mass at 0, i.e., $H(\{0\}) = 0$, and that the spectral density satisfies $h(w) \sim sw^t$ as $w \to 0$ with $s > 0$ and $t > -1$. They found that the normalisation (3) required to give a non-degenerate limiting conditional law has $\alpha = 0$, $\beta = (t + 1)/(t + 2)$ and the limit is of Weibull type, i.e.,

$$\lim_{u \to \infty} \Pr\left(Y < X^\beta z \mid X > u\right) = 1 - \exp\left\{-\frac{sz^{t+2}}{(t + 1)(t + 2)}\right\}, \quad \text{for } z > 0. \tag{11}$$

Examples of inverted max-stable models satisfying (11) include those with logistic and Dirichlet dependence structure [4].

3.3. Regular variation at lower tail of spectral measure

We derive the conditional representation (2) for the class of inverted max-stable distributions covering more general spectral measures than those studied in Section 3.2, deriving the normalising functions, $a(x)$ and $b(x)$, and the limit distribution G. Here we assume that the spectral measure H places all its mass in the sub-region, $[w_\ell, w_u]$ say, of $[0, 1]$, can have mass at w_ℓ and has a spectral density that is regularly varying at w_ℓ, i.e.,

$$h(w) \sim \mathcal{L}(w - w_\ell)(w - w_\ell)^t \quad \text{as } w \downarrow w_\ell \in [0, 1/2), \tag{12}$$

$t > -1$ and $\mathcal{L}(w)$ slowly varying at 0, i.e., $\mathcal{L}(wx)/\mathcal{L}(w) \to 1$, as $w \to 0$, for all $x > 0$. This new formulation has three key differences from Section 3.2: restricted domain, mass on the lower boundary, and more general slowly varying functions than simply $\mathcal{L}$ tending to a constant at the lower boundary.

We have allowed for mass on the lower boundary motivated by the properties of the extremal-t distribution (6), for which the spectral measure places mass at $\{0\}$. We explore the assumption of possible mass on the lower end point w_ℓ of H, but no point mass at any other $w \in (w_\ell, w_u)$. The results in the rest of the paper even hold if there is mass at some countable set of points in (w_ℓ, w_u), with modified proofs, but we avoid this unnecessary generalisation.

Having mass at w_ℓ with $w_\ell > 0$, implies that (X, Y) do not have a joint density everywhere, as there is a singular component on the boundary of the sample space of (X, Y). If this boundary is avoided then the derivation of property (10) is still valid, and so calculations are straightforward. In what follows we derive results for the cases where $H(\{w_\ell\}) = 0$ and $H(\{w_\ell\}) > 0$ separately.

Lemma 1. *Let w_ℓ, w_u, be the lower and upper end points, respectively, of the spectral measure H of an inverted max-stable distribution* (7), *i.e.,*

$$w_\ell = \inf\{0 \le w < 1/2 : H(w) \ge 0\}, \qquad w_u = \sup\{1/2 < w \le 1 : H(w) \le 2\},$$

and assume that, apart from the points w_ℓ and w_u, for which $H(w_\ell)$, $H(w_u) \ge 0$, the spectral measure is absolutely continuous with respect to the Lebesgue measure. Then, if $y/(x + y) \downarrow w_\ell$, for $y = y(x)$ a function of x, as $x \to \infty$,

$$V_1(1, x/y) \to w_\ell H(\{w_\ell\}) - 1. \tag{13}$$

In Lemma 1 the case of perfect positive dependence between X and Y, i.e., $w_\ell = w_u = 1/2$, is excluded since there can be no possible normalisation such that G in expression (2) is non-degenerate. Proposition 1 gives the asymptotic form of the log-conditional survivor function of $Y \mid X = x$, for large x.

Proposition 1. *Under the conditions of Lemma 1 and for $h(w)$ as in expression (12), for $y/(x+y) \downarrow w_\ell \in [0, 1/2)$ for $y = y(x)$ a function of x, as $x \to \infty$, we obtain that $\ln \Pr(Y > y \mid X = x)$ is asymptotically equivalent to*

(i) *for $H(\{w_\ell\}) = 0$,*

$$-x\mathcal{L}\left(\frac{y}{x+y} - w_\ell\right)\left(\frac{y}{x+y} - w_\ell\right)^{t+2} \Big/ \{(1 - w_\ell)(t + 1)(t + 2)\}, \tag{14}$$

(ii) *for $H(\{w_\ell\}) > 0$,*

$$\ln\{1 - w_\ell H(\{w_\ell\})\} - (x + y)\left(\frac{y}{x+y} - w_\ell\right)H(\{w_\ell\}). \tag{15}$$

As there is no contribution from the spectral density in expression (15), a general form for the normalisation can be obtained directly. General forms of the normalising functions $a(x)$ and $b(x)$ cannot be obtained from representation (14) without additional assumptions so it is helpful to consider a condition on the slowly varying function $\mathcal{L}$ at 0. Heffernan and Tawn [13] results were based on the case of $\mathcal{L}$ having a finite right limit at 0, Corollary 1 extends this.

Corollary 1. *Under the conditions of Proposition 1, limit (2) holds as follows*

(i) *for $H(\{w_\ell\}) = 0$, then $a(x) = \{w_\ell/(1 - w_\ell)\}x$ and $b(x) = x^{(t+1)/(t+2)}\mathcal{L}\left\{x^{-1/(t+2)}\right\}^{-1/(t+2)}$ for $x > 0$ and assuming that*

$$\lim_{w \to 0^+} \frac{\mathcal{L}\left\{w\mathcal{L}(w)^{-\tau}\right\}}{\mathcal{L}(w)} = 1, \tag{16}$$

for all $\tau \in (0, 1)$, then

$$G(z) = 1 - \exp\left\{-\frac{(1 - w_\ell)^{3+2t}z^{t+2}}{(t + 1)(t + 2)}\right\} \quad \text{for } z > 0.$$

(ii) *For $H(\{w_\ell\}) > 0$, then $a(x) = \{w_\ell/(1 - w_\ell)\}x$ and $b(x) = 1$ for $x > 0$, then*

$$G(z) = 1 - \{1 - w_\ell H(\{w_\ell\})\}\exp\{-(1 - w_\ell)H(\{w_\ell\})z\} \quad \text{for } z > 0,$$

and $G(z) = 0$, for $z < 0$, so $G(\{0\}) = w_\ell H(\{w_\ell\})$.

Condition (16) is satisfied by a range of slowly varying functions, including those studied in Section 3.1 as well as by functions that approach ∞ when the argument tends to zero. Examples for $\mathcal{L}(w)$, with $\lim_{w \to 0^+} \mathcal{L}(w) = \infty$, satisfying condition (16) include $\ln_\kappa(-\ln w)$, $\kappa \in \mathbb{N}_0$, $\exp\{(-\ln w)^\nu\}$, $\nu \in (0, 1/2)$ and $\exp\{-\ln w/\ln(-\ln w)\}$, where $\ln_k$ is the iterated logarithm function defined recursively by $\ln_k x = \ln_{k-1} \ln x$, $\ln_0 x = x$ and $\ln_1 x = \ln x$.

Remark 1. For $\lim_{w \to 0^+} \mathcal{L}(w) = s > 0$, all cases of norming functions (i)–(ii) in Corollary 1 reduce, after absorbing s into the limiting law, to the parametric class of Heffernan–Tawn, i.e., $a(x) + b(x)z = \alpha x + x^\beta z$, where $\alpha = w_\ell/(1 - w_\ell) \in [0, 1)$ for $w_\ell \in [0, 1/2)$ and $\beta = (t + 1)/(t + 2) \in [0, 1)$ for $t \geq -1$. This is the first example to be known with both α and β a function of parameters.

Remark 2. When $\lim_{w \to 0^+} \mathcal{L}(w) = \infty$ and subject to condition (16), the additional factor $\mathcal{L}\{x^{-1/(t+2)}\}^{-1/(t+2)}$ enters in the scaling function and reduces the rate of increase of $b(x)$ to ∞, as $x \to \infty$.

Another interesting case which is satisfied by many max-stable models that appear in the literature, is when $w_\ell = 0$ and $H(\{w_\ell\}) > 0$ in case (ii) of Corollary 1. In this case, $a(x) = 0$ and $b(x) = 1$ for all $x > 0$ and the random variables X and Y, conditionally on $X > u$, are nearly independent in the terminology of Ledford and Tawn [23], with exact independence occurring when the limit distribution is unit exponential, i.e., when $H(\{w_\ell\}) = 1$. Two such max-stable models come from the extremal-t [26] and Gaussian–Gaussian [38] processes, for which the exponent measures of their bivariate distributions are given by expressions (6) and

$$V(x, y) = \frac{1}{2}\left(\frac{1}{x} + \frac{1}{y}\right) + \frac{1}{2}\int_{\mathbb{R}^2}\left\{\frac{\phi_2(u)^2}{x^2} - 2\rho(h)\frac{\phi_2(u)\phi_2(h - u)}{xy} + \frac{\phi_2(h - u)^2}{y^2}\right\}^{1/2} du, \tag{17}$$

respectively, $h \in \mathbb{R}_+^2$, $\rho(h) \in [-1, 1]$ is a valid correlation function and ϕ_2 is the density of the standard bivariate normal distribution with correlation $\rho(h)$. The corresponding mass on the lower end point $w_\ell = 0$ of models (6) and (17) is

$$H(\{0\}) = T_{v+1}\left[-\rho\left(\frac{v + 1}{1 - \rho^2}\right)^{1/2}\right] \quad \text{and} \quad H(\{0\}) = \frac{1 - \rho(h)}{2},$$

Table 1

The mass of the spectral measure on {0} of bivariate exponent measures, from top to bottom, of mixed, asymmetric and asymmetric mixed logistic [36], Schlather [33] and Marshall and Olkin [24] distributions. The final column shows the parameter space, Θ, of the model.

$V(x, y)$	$H(\{0\})$	Θ
$\left(\frac{1}{x} + \frac{1}{y}\right) - \frac{\theta}{x+y}$	$1 - \theta$	$\theta \in (0, 1)$
$\frac{1-\theta}{x} + \frac{1-\phi}{y} + \left\{(\theta/x)^{1/\alpha} + (\phi/y)^{1/\alpha}\right\}^{\alpha}$	$1 - \phi$	$0 \leq \theta, \phi, \alpha \leq 1$
$\frac{1}{x} + \frac{1}{y} - \frac{1}{xy}\left(\frac{1}{x} + \frac{1}{y}\right)^{-2}\left(\frac{\theta+\phi}{x} + \frac{2\phi+\theta}{y}\right)$	$1 - \phi - \theta$	$\theta, \theta + 3\phi > 0$ and $\theta + \phi, \theta + 2\phi \leq 1$
$\frac{1}{2}\left(\frac{1}{x} + \frac{1}{y}\right)\left[1 + \left\{1 - \frac{2(1+\rho)xy}{(x+y)^2}\right\}^{1/2}\right]$	$(1 - \rho)/2$	$\rho \in (-1, 1)$
$\alpha\left(\frac{1}{x} + \frac{1}{y}\right) + (1 - \alpha)\max\{1/x, 1/y\}$	α	$0 \leq \alpha \leq 1$

respectively. Table 1 gives a collection of other max-stable models, including the Schlather distribution (6), placing positive mass on {0}.

3.4. Inverted Brown–Resnick model

In this section we focus on the limiting conditional representation of the inverted Brown Resnick model (5). This model has the same bivariate copula as the Hüsler–Reiss distribution [16] and the Smith model [34]. The spectral measure of the Brown–Resnick max-stable distribution places no mass on any $0 \leq w \leq 1$, and the spectral density satisfies

$$h(w) \sim \frac{\exp(-\lambda^2/8)}{\lambda (2\pi)^{1/2}} w^{-3/2} \exp\left\{-(\ln w)^2/(2\lambda^2)\right\} \quad \text{as } w \to 0. \tag{18}$$

This corresponds to a different form than expression (12) or its more general forms of the slowly varying function $\mathcal{L}$. In particular, the spectral density is Γ-varying (see Section 3.5) at 0 with auxiliary function

$$f(w) = -\lambda^2 w/\ln w.$$

As Proposition 2 shows, this example leads to a different form for the normalising functions $a(x)$ and $b(x)$ than the ones considered by Heffernan and Tawn [13].

Proposition 2. *Assume that* (X, Y) *follows the inverted max-stable distribution* (7) *with exponent measure* (5)*. Limit* (2) *holds with, for* $x > 0$ *and* $z \in \mathbb{R}$,

$$a(x) = x \exp\left\{-\lambda (2\ln x)^{1/2} + \frac{\lambda \ln \ln x}{(2\ln x)^{1/2}} + \frac{\lambda^2}{2}\right\}, \qquad b(x) = a(x)/(\ln x)^{1/2}, \tag{19}$$

and

$$G(z) = 1 - \exp\left\{-\frac{\lambda}{(8\pi)^{1/2}} \exp\left(\sqrt{2}z/\lambda\right)\right\}. \tag{20}$$

Remark 3. Limit distribution (20) is of reverted Gumbel type, which is different from the limits in Corollary 1. The rate of convergence to the limit is order $\ln \ln u/(\ln u)^{1/2}$.

A natural question that arises from this counter-example relates to how well can the conditioned dependence model of Heffernan and Tawn [13] with canonical family (3) approximate the conditional distribution of $Y \mid X > u$, for large u, when the random vector (X, Y) follows the inverted max-stable distribution with Brown–Resnick dependence structure and unit exponential margins. To facilitate comparisons between the two models, Fig. 1 shows the graphs of $a(x)/x$ and $\ln b(x)/\ln x$ where $a(x)$ and $b(x)$ are given by expression (19), for several values of the dependence parameter λ and a range of x values above the 0.95 unit exponential quantile. Both plots show that $a(x)/x$ and $\ln b(x)/\ln x$ are approximately constant for large x so that the canonical class of norming functions is likely to approximate well $a(x)$ and $b(x)$ by αx and x^β, respectively. Subsequently, we simulated data from the inverted max-stable distribution with Brown–Resnick dependence and fitted the conditioned dependence model using: (i) the canonical family (3) and (ii) the model implied by the norming functions (19), treating the functions (19) as a parametric model for the growth of Y given large X. Our comparisons are based on the differences between the conditional quantile estimates of $Y \mid X = x$ from the two models. For both models, similar to Heffernan and Tawn [13], we used for the limiting law G in expression (2) the false working assumption of a normal distribution with mean and variance parameters. We considered two values for the dependence parameter, i.e., $\lambda = 1.3$ (weak dependence) and $\lambda = 0.3$ (strong dependence). For each λ, 10^2 samples of size 10^3 were generated from the inverted max-stable distribution with Brown–Resnick dependence and the 0.025, 0.5 and 0.975 conditional quantile estimates of
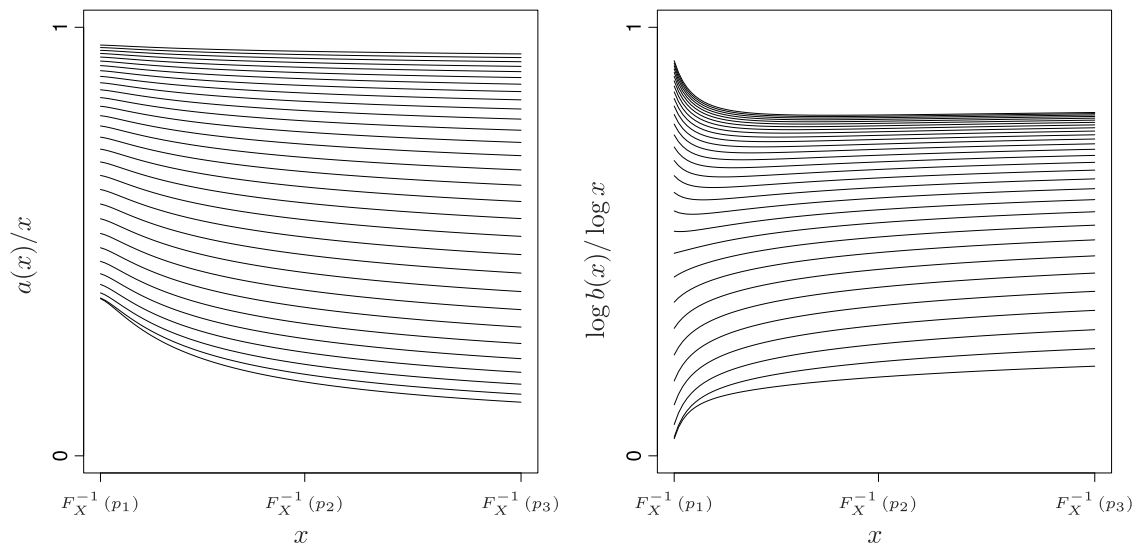
Fig. 1. Plots of $a(x)/x$ and $\ln b(x)/\ln x$, $x > F_X^{-1}(0.95)$, where $a(x)$ and $b(x)$ are given by expression (19), for different values of λ ranging from 0.01 (bottom curves) to 20 (top curves). The inverse of the unit exponential distribution function is $F_X^{\leftarrow}(p) = -\ln(1-p)$, for $p \in (0,1)$. The values of p_1, p_2 and p_3 are 0.95, $1 - 10^{-7}$ and $1 - 10^{-13}$, respectively.
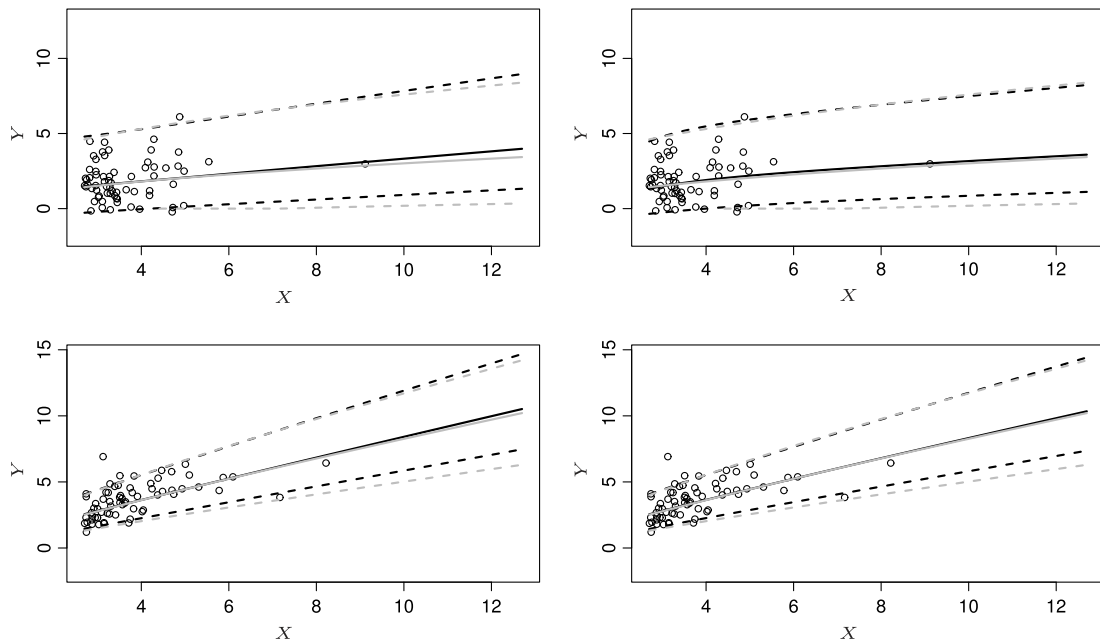

Fig. 2. Conditional exceedances above the 0.935 unit exponential quantile from a simulated sample of size 10^3 from the inverted max-stable distribution with Brown–Resnick dependence and exponential margins for $\lambda = 1.3$ (top) and $\lambda = 0.3$ (bottom). The black lines correspond to averaged estimates from 100 simulations of the 0.025, 0.5 and 0.975 conditional quantiles of $Y \mid X = x$, $x > F_X^{-1}(0.935)$, using the Heffernan–Tawn model (3) (left) and the model constructed by the theoretical functions in expression (19) (right). Grey lines correspond to the theoretical 0.025, 0.5 and 0.975 conditional quantiles.

$Y \mid X = x$, for $x > F_X^{-1}(0.935)$, were computed from the two model fits, i.e., model (3) and the model defined by expression (19). The conditional quantile estimates are of the form $\hat{a}(x) + \hat{b}(x)\hat{z}_p$, where $\hat{a}(x)$, $\hat{b}(x)$ are maximum likelihood estimates and $\hat{z}_p$ is the pth empirical quantile of $\hat{Z} = \{Y - \hat{a}(x)\}/\hat{b}(x)$, for large x. Fig. 2 shows the averaged estimates of the conditional quantiles along with the theoretical conditional quantiles. Both models estimate the true conditional quantiles well and their behaviour is almost indistinguishable. This shows that the canonical model is flexible enough to approximate the conditional distribution of the inverted max-stable distribution with Brown–Resnick dependence.

3.5. Γ-variation at lower tail of spectral density

Having identified a new form for the tail of the spectral density for the Brown–Resnick max-stable model, we consider in this section the log-conditional survivor function of $Y \mid X = x$, under the assumption

$$h(w) \sim g(w - w_\ell), \quad \text{as } w \downarrow w_\ell, \tag{21}$$

where $g(w)$ is a Γ_f-varying function at 0 [7]. A function h is Γ-varying at 0 with auxiliary function f, short-hand $h \in \Gamma_f(0^+)$, if for all $w \in \mathbb{R}$, $\lim_{s \to 0^+} h\{s + wf(s)\}/h(s) = \exp(w)$.

Similar to Section 3.3, we consider the assumption of possible mass at the lower end point w_ℓ. Our findings are based on the assumptions of a differentiable spectral density h and Lemma 2.

Lemma 2. Let $g : \mathbb{R}^+ \mapsto \mathbb{R}^+ \in \Gamma_f(0^+)$ and $U(w) \in R_\nu(0^+)$, $\nu \in \mathbb{R}$. Assume further that there exists an $\epsilon > 0$ such that U and g are $C^\infty(0, \epsilon)$ functions with

$$\lim_{w \to 0^+} \frac{g'(w)}{g(w)^2} \int_0^w g(s) \, ds \tag{22}$$

existing. Then

(i) $U(w)g(w) \in \Gamma_f(0^+)$.
(ii) Define $f(w) = g(w)/g'(w)$, $w > 0$. Then, f is an auxiliary function for g.
(iii) For $f(w)$ as in (ii),

$$\left(\int_0^w U(s)g(s)\mathrm{d}s \right) / \{U(w)f(w)g(w)\} = 1 - \frac{(Uf)'(w)}{U(w)} + \frac{\{(Uf)'f\}'(w)}{U(w)} - \cdots \tag{23}$$

$$= 1 + o(1), \quad \text{as } w \to 0^+. \tag{24}$$

Proposition 3 gives the asymptotic form of the log-conditional survivor function.

Proposition 3. Under the conditions of Lemma 1 and for $h(w)$ as in expression (21), for $y/(x+y) \downarrow w_\ell \in [0, 1/2)$ for $y = y(x)$ function of x, as $x \to \infty$, we obtain that $\ln \Pr(Y > y \mid X = x)$ is asymptotically equivalent to

(i) for $H(\{w_\ell\}) = 0$,

$$-(x+y)f^2\left(\frac{y}{x+y} - w_\ell\right)h\left(\frac{y}{x+y}\right), \tag{25}$$

(ii) for $H(\{w_\ell\}) > 0$,

$$\ln\{1 - w_\ell H(\{w_\ell\})\} - (x+y)\left(\frac{y}{x+y} - w_\ell\right)H(\{w_\ell\}). \tag{26}$$

As an example, we explore a new class of spectral densities that are more flexible than the spectral density (18) of the inverted max-stable distribution with Brown–Resnick dependence. Specifically, consider for $\gamma > 0$, $\delta \in \mathbb{R}$ and $\kappa > 0$, the $\Gamma_f(0^+)$-varying function

$$h(w) \sim w^\delta \exp\left(-\kappa w^{-\gamma}\right) \quad \text{as } w \to 0^+, \tag{27}$$

with auxiliary function

$$f(w) = (\kappa\gamma)^{-1} w^{1+\gamma}. \tag{28}$$

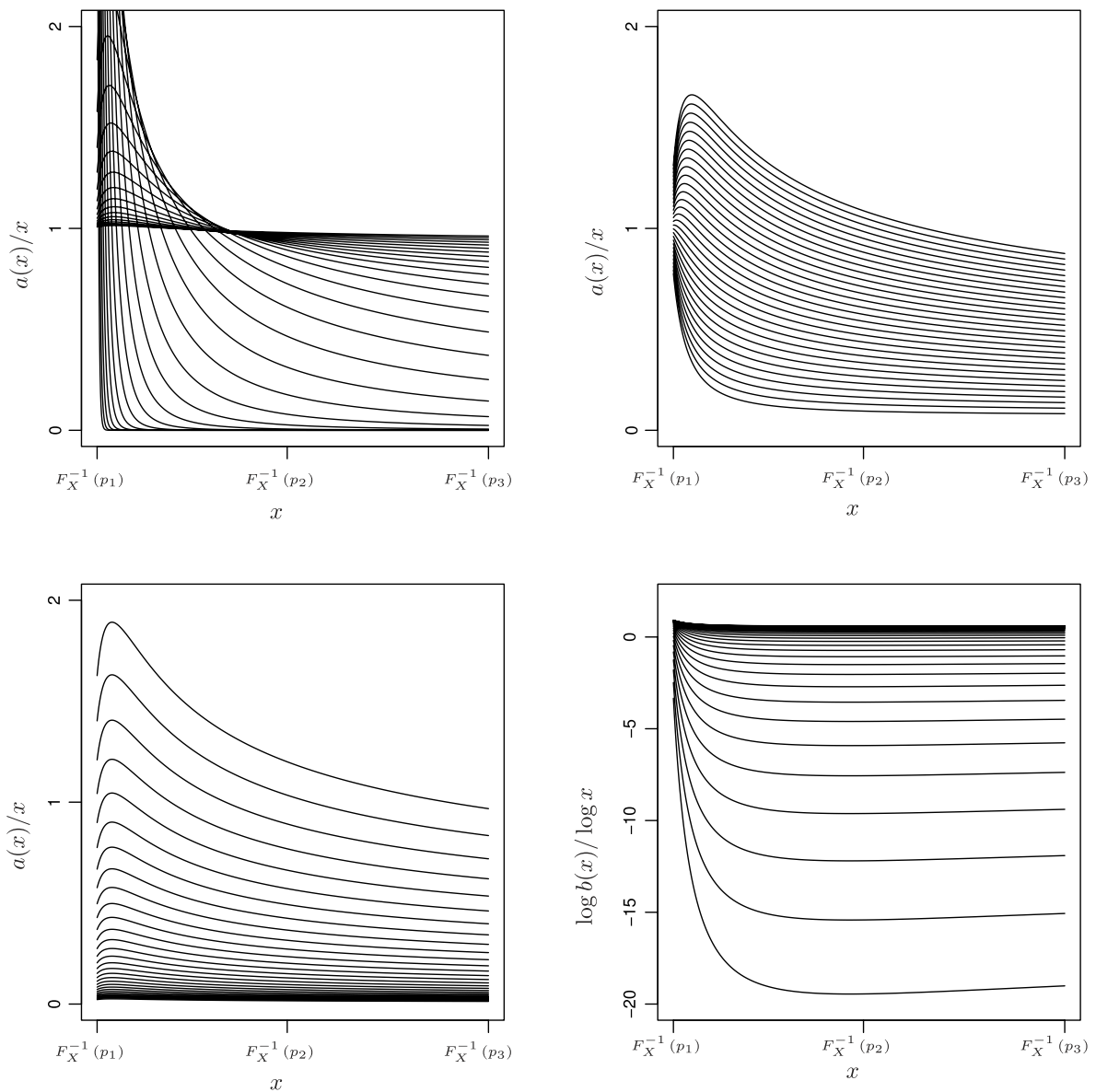Proposition 4 gives the normalising functions and limiting conditional distribution for this example.

Fig. 3. Plots of $a(x)/x$ for different values of the parameters γ (top left), κ (top right) and δ (bottom left) and $\ln b(x)/\ln x$ (bottom right), $x > F_X^{-1}(0.95)$. The inverse of the unit exponential distribution function is $F_X^{\leftarrow}(p) = -\ln(1-p)$, for $p \in (0, 1)$. The values of p_1, p_2 and p_3 are 0.95, $1 - 10^{-7}$ and $1 - 10^{-13}$, respectively.

Proposition 4. *Assume that* (X, Y) *follows the inverted max-stable distribution* (7) *with spectral measure H with* $w_\ell = 0$ *and* $H(\{0\}) = 0$, *and spectral density h that satisfies expression* (27). *Limit* (2) *holds with, for* $x > 0$ *and* $z \in \mathbb{R}$

$$a(x) = x\kappa^{1/\gamma} (\ln x)^{-1/\gamma} \left[1 + \gamma^{-2} \{\delta + 2(1 + \gamma)\} \frac{\ln \ln x}{\ln x} \right], \qquad b(x) = x (\ln x)^{-1-1/\gamma}, \tag{29}$$

and

$$G(z) = 1 - \exp\left[-\left\{ \kappa^{(\delta+2)/\gamma} / \gamma^2 \right\} \exp\left(\gamma\kappa^{-1/\gamma} z \right) \right]. \tag{30}$$

Remark 4. Similarly with the inverted max-stable with Brown–Resnick dependence, the limiting conditional distribution (30) is of reverted Gumbel type and the norming functions (29) do not belong to the Heffernan–Tawn parametric family. The rate of convergence to the limit is order $\ln \ln u / \ln u$.

Fig. 3 shows the graphs of functions $a(x)/x$ and $\ln b(x)/\ln x$, for several values of the parameters γ, δ, κ, and large x. Large values of γ correspond to strong dependence between X and Y so that $a(x)/x$ is nearly constant and equal to 1 for large x. Small values of γ correspond to independence with sharp decrease of $a(x)/x$ to 0 as x increases. Intermediate values of γ correspond to mild-moderate dependence of X and Y with $a(x)/x$ having a turning point and decaying as x increases. Parameters κ and δ, seem to have similar effect with larger values corresponding to increasing dependence. Last, $\ln b(x)/\ln x$ is approximately constant with large x. Comparing with the canonical family (3), the degree of approximation of $a(x)$ and $b(x)$ by αx and x^{β} is not especially good for $a(x)$. To our knowledge, there is no current parametric model for the max-stable process with spectral density satisfying expression (27).

4. Spatial extensions

Engelke et al. [10] have found that considerable insights can be gained by studying max-stable processes conditioned on the process being extreme at a fixed point, $s^* \in \mathbb{R}^2$ say. For $\{X(s); s \in \mathbb{R}^2\}$ an inverted max-stable process with unit exponential margins a similar approach requires the existence of location and scaling norming functions $a_s : \mathbb{R}_+ \mapsto \mathbb{R}$ and $b_s : \mathbb{R}_+ \mapsto \mathbb{R}_+$, for all $s \in \mathbb{R}^2$, such that, for any $x > 0$ and $\{z_s \in \mathbb{R}, s \in \mathbb{R}^2\}$,

$$\lim_{u \to \infty} \Pr\left[X(s^*) - u > x, \frac{X(s) - a_s\{X(s^*)\}}{b_s\{X(s^*)\}} < z_s, s \in \mathbb{R}^2 \setminus \{s^*\} \mid X(s^*) > u \right] = \exp(-x)G(\mathbf{z}), \tag{31}$$

where G is an infinite-dimensional joint distribution function that is non-degenerate in all univariate margins and $\mathbf{z} = \{z_s, s \in \mathbb{R}^2 \setminus \{s^*\}\}$.

From our working in this paper we have derived the forms of (a_s, b_s) and G_s, the marginal of G, as these features are determined uniquely by the bivariate joint distributions of the inverted max-stable process. Thus all that remains to fully characterise limit (31) is to derive the infinite-dimensional dependence structure of G.

Here a complication arises as for most max-stable process models only the bivariate marginal distributions are available in closed form [6]. Maybe future work can explore approaches that Engelke et al. [10] have used to get around this issue for max-stable process. However, for the moment we restrict ourselves to consider the trivariate distributions of two well-known max-stable processes which have closed form margins in all dimensions, the Brown–Resnick process [11,15] and the extremal-t process [26]. In this case we find that the associated bivariate limit $G_{s_1,s_2}(z_1, z_2) = G_{s_1}(z_1)G_{s_2}(z_2)$, thus G factorises; corresponding to asymptotic conditional independence in the terminology of Heffernan and Tawn [13]. It follows that the infinite-dimensional G will also factorise. Thus in this case all the dependence structures of the inverted max-stable processes are absorbed in the location-scale functions. Identifying which classes of max-stable process possess this asymptotic conditional independence for G is an interesting line of future research.

Acknowledgments

Appendix

A.1. Proof of Lemma 1

For $s/(s + t) \in (w_\ell, w_u)$, the partial derivative of the exponent measure is equal to

$$\begin{aligned}
\frac{\partial V(s, t)}{\partial s} &= \frac{\partial}{\partial s}\left[\int_{\frac{s}{s+t}}^{w_u} (w/s)\mathrm{d}H(w) + \int_{w_\ell}^{\frac{s}{s+t}} \{(1-w)/t\}\mathrm{d}H(w) \right] \\
&= \frac{\partial}{\partial s}\left[\int_{\left[\frac{s}{s+t}, w_u\right)} (w/s)h(w)\mathrm{d}w + \int_{\left(w_\ell, \frac{s}{s+t}\right]} \{(1-w)/t\}h(w)\mathrm{d}w \right] - (w_u/s^2)H(\{w_u\}) \\
&= -\int_{\left[\frac{s}{s+t}, w_u\right)} (w/s^2)h(w)\mathrm{d}w - (w_u/s^2)H(\{w_u\}),
\end{aligned}$$

where the last step follows from the Fundamental Theorem of Calculus. Under the assumption of $y/(x+y) \downarrow w_\ell$, as $x \to \infty$, it follows that

$$V_1(1, x/y) \to -\int_{w_\ell}^{w_u} wh(w)\mathrm{d}w - w_u H(\{w_u\}), \tag{32}$$

as $x \to \infty$. Using the moment constraint (4) we have that

$$\int_{w_\ell}^{w_u} wh(w)\mathrm{d}w = \int_{(w_\ell, w_u)} wh(w)\mathrm{d}w = 1 - w_\ell H(\{w_\ell\}) - w_u H(\{w_u\}),$$

which yields Eq. (13), after combining with Eq. (32).

A.2. Proof of Proposition 1

Working similarly to the proof of Lemma 1, we get, after combining Eqs. (10) and (13), that for $c(x, y) = y/\{x + y\}$, $d(x, y) = \{x + y\}/\{(1 - w_\ell)y - w_\ell x\}$, the log-conditional survivor, $\ln\{\Pr(Y > y \mid X = x)\}$, is equal to

(i) for $H(\{w_\ell\}) = 0$,

$$(x + y)\int_{w_\ell}^{c(x,y)} w\mathcal{L}(w - w_\ell)(w - w_\ell)^t \, \mathrm{d}w - y\int_{w_\ell}^{c(x,y)} \mathcal{L}(w - w_\ell)(w - w_\ell)^t \, \mathrm{d}w$$

$$= (x + y)\int_{d(x,y)}^{\infty} \mathcal{L}(1/s)s^{-(t+3)} \, \mathrm{d}s + \{w_\ell x - (1 - w_\ell)y\}\int_{d(x,y)}^{\infty} \mathcal{L}(1/s)s^{-(t+2)} \, \mathrm{d}s.$$

From the assumption $c(x, y) \to w_\ell$ as $x \to \infty$ we conclude that $d(x, y) \to \infty$ for any w_ℓ, as, we have, from Karamata's theorem [30, pg. 17], that for $t > -1$, the last expression is asymptotically equivalent to

$$\frac{(x + y)}{(t + 2)}\{d(x, y)\}^{-(t+2)}\mathcal{L}\{1/d(x, y)\} + \frac{\{w_\ell x - (1 - w_\ell)y\}}{(t + 1)}\{d(x, y)\}^{-(t+1)}\mathcal{L}\{1/d(x, y)\},$$

as $x \to \infty$, which simplifies to expression (14),

(ii) for $H(\{w_\ell\}) > 0$,

$$\ln\{1 - w_\ell H(\{w_\ell\})\} + x - xV(1, x/y)$$

$$= \ln\{1 - w_\ell H(\{w_\ell\})\} + x - x\left[1 - \int_{w_\ell}^{c(x,y)} w \, \mathrm{d}H(w) + (y/x)\int_{w_\ell}^{c(x,y)} (1 - w_\ell) \, \mathrm{d}H(w)\right]$$

$$\sim \ln\{1 - w_\ell H(\{w_\ell\})\} + \{w_\ell x - (1 - w_\ell)y\}H(\{w_\ell\}), \quad \text{as } x \to \infty. \tag{33}$$

A.3. Proof of Proposition 2

Let ϕ be the probability density function of the standard normal distribution. Assuming $y \to \infty$ as $x \to \infty$ with $y/x \to 0$, we have from expression (10), Lemma 1 and Mill's ratio, that for large x, the log-conditional survivor, $\ln \Pr(Y > y \mid X = x)$, is approximately equal to

$$x - x\left[1 - \frac{\phi\{\frac{\lambda}{2} + \frac{1}{\lambda}\ln(x/y)\}}{\frac{\lambda}{2} + \frac{1}{\lambda}\ln(x/y)}\right] + y\frac{\phi\{\frac{\lambda}{2} - \frac{1}{\lambda}\ln(x/y)\}}{\frac{\lambda}{2} - \frac{1}{\lambda}\ln(x/y)}$$

$$= x\frac{\phi\{\frac{\lambda}{2} + \frac{1}{\lambda}\ln(x/y)\}}{\frac{\lambda}{2} + \frac{1}{\lambda}\ln(x/y)}\left[1 + \frac{y}{x}\frac{\phi\{\frac{\lambda}{2} - \frac{1}{\lambda}\ln(x/y)\}}{\phi\{\frac{\lambda}{2} + \frac{1}{\lambda}\ln(x/y)\}}\frac{\{\frac{\lambda}{2} + \frac{1}{\lambda}\ln(x/y)\}}{\{\frac{\lambda}{2} - \frac{1}{\lambda}\ln(x/y)\}}\right]$$

$$= x\frac{\phi\{\frac{\lambda}{2} + \frac{1}{\lambda}\ln(x/y)\}}{\frac{\lambda}{2} + \frac{1}{\lambda}\ln(x/y)}\left[1 + \frac{\{\frac{\lambda}{2} + \frac{1}{\lambda}\ln(x/y)\}}{\{\frac{\lambda}{2} - \frac{1}{\lambda}\ln(x/y)\}}\right]$$

$$\doteq -c\,(xy)^{1/2}\frac{\phi\{\frac{1}{\lambda}\ln(y/x)\}}{\{\frac{1}{\lambda}\ln(y/x)\}^2}\left[1 + O\{(\ln x)^{-1}\}\right], \tag{34}$$

where $c = \lambda \exp\left(-\lambda^2/8\right)$. Now, let $z \in \mathbb{R}$ and $y = a(x) + b(x)z$, where $a(x)$ and $b(x)$ are given by Eqs. (19). We have, as $x \to \infty$,

$$(xy)^{1/2} = x\exp\left\{\lambda^2/4 - \frac{\lambda}{\sqrt{2}}(\ln x)^{1/2}\right\}\left[1 + O\left\{\frac{\ln\ln x}{(\ln x)^{1/2}}\right\}\right], \tag{35}$$

$$\left\{\frac{1}{\lambda}\ln(x/y)\right\}^2 = 2\ln x\left[1 + O\{(\ln x)^{-1/2}\}\right], \tag{36}$$

$$\phi\left\{\frac{1}{\lambda}\ln(x/y)\right\} = (2\pi)^{-1/2}\exp\left[-\ln x - \frac{\lambda^2}{8} + \frac{\lambda}{\sqrt{2}}(\ln x)^{1/2} + \ln\ln x + \frac{\sqrt{2}z}{\lambda}\right]\left[1 + O\left\{\frac{(\ln\ln x)^2}{\ln x}\right\}\right]. \tag{37}$$

Combining Eqs. (34)–(37) we get

$$\Pr\{Y < a(x) + b(x)z \mid X = x\} = 1 - \exp\left[-\frac{\lambda}{(8\pi)^{1/2}}\exp\left\{\sqrt{2}z/\lambda\right\}\right] + O\left\{\frac{\ln\ln x}{(\ln x)^{1/2}}\right\}.$$

Last, direct application of statement (9) yields the result of Proposition 2.

A.4. Proof of Lemma 2

(i) First, for any auxiliary function f, we have that $\lim_{t\to 0^+} f(t)/t = 0$ (see [7] Lemma 1.5.1). Next, for $U(w) = w^\nu \mathcal{L}(w)$, where $\mathcal{L} \in R_0(0^+)$ and $\nu \in \mathbb{R}$,

$$\lim_{t\to 0^+}\frac{U\{t + wf(t)\}g(\{t + wf(t)\})}{U(t)g(t)} = \lim_{t\to 0^+}\frac{\mathcal{L}\left[t\{1 + wf(t)/t\}\right]}{\mathcal{L}(t)}\frac{g\{t + wf(t)\}}{g(t)}\{1 + wf(t)/t\},$$

$$\to \exp(w), \quad w > 0.$$

(ii) Theorem 1.5.2 in de Haan [7] asserts that any function f_0 that satisfies

$$f_0(w) \sim \left(\int_0^w g(s)ds\right)/g(w) \quad \text{as } w \to 0^+, \tag{38}$$

is an auxiliary function for g. Additionally, Theorem 1.5.4 in de Haan [7] states that if $g \in \Gamma_f(0^+)$, then, as $w \to 0^+$

$$\int_0^w \{g(s)\}^2\, ds \sim \frac{1}{2}g(w)\int_0^w g(s)\, ds. \tag{39}$$

Given that limit (22) exists it follows by l'Hôpital's rule that

$$\lim_{w\to 0^+}\frac{1}{2}g(w)\int_0^w g(s)\,ds \Big/ \int_0^w \{g(s)\}^2\,ds = \lim_{w\to 0^+}\left(\frac{1}{2}g(w)^2 + \frac{1}{2}g'(w)\int_0^w g(s)\,ds\right)\Big/g(w)^2, \tag{40}$$

since the right hand side limit in (40) exists due to limit (22) existing. However, by property (39), we have the left hand side limit in (40) is equal to 1 and so it follows from equality (40) that as $w \to 0^+$,

$$\int_0^w g(s)\,ds \sim \{g(w)\}^2/g'(w). \tag{41}$$

Combining expressions (38) and (41) we obtain $f_0(w) \sim f(w)$, as $w \to 0^+$, where $f(w) = g(w)/g'(w)$. Hence, for $w > 0$, the function $f(w)$ is up to asymptotic equivalence equal to f_0.

(iii) Define $f(w) = g(w)/g'(w)$, $w > 0$. We have

$$\int_0^w U(s)g(s)\,ds = \int_0^w U(s)f(s)g'(s)\,ds$$

$$= U(w)f(w)g(w) - \int_0^w (Uf)'(s)f(s)g'(s)\,ds$$

$$= U(w)f(w)g(w) - (Uf)'(w)f(w)g'(w) + \int_0^w \left\{(Uf)'f\right\}'(s)f(s)g'(s)\,ds,$$

which gives expression (23), after continuation of integration by parts and division by $U(w)f(w)g(w)$. Last, expression (24) follows from de Haan's theorem (38) and case (i) of Lemma 2.

A.5. Proof of Proposition 3

We only consider case (i) and note that case (ii) is identical to Proposition 1. Define $c(x, y) = y/(x + y)$ and $l(x, y) = c(x, y) - w_\ell$. For $c(x, y) \to w_\ell$, $l(x, y) = c(x, y) - w_\ell \to 0$, as $x \to \infty$, and with T as in expression (7), we have that $\ln\Pr(Y > y \mid X = x)$ is equal to

$$(x + y)\int_0^{l(x,y)} sg(s)\,ds + \{w_\ell x - (1 - w_\ell)y\}\int_0^{l(x,y)} g(s)\,ds. \tag{42}$$

Using the asymptotic expansion (23), up to first order, we have that the two integrals in expression (42) are, as $x \to \infty$, asymptotically equivalent to

$$\int_0^{l(x,y)} sg(s)\,ds \doteq l(x, y)f\{l(x, y)\}g\{l(x, y)\}\left[1 - \frac{f\{l(x, y)\} - l(x, y)f'\{l(x, y)\}}{l(x, y)}\right], \tag{43}$$

$$\int_0^{l(x,y)} g(s)\,ds \doteq f\{l(x, y)\}g\{l(x, y)\}\left[1 - f'\{l(x, y)\}\right], \tag{44}$$

for $U(s) = s \in R_1(0^+)$ and $U(s) = 1 \in R_0(0^+)$, respectively. Combining expressions (42)–(44) we get that (42) is asymptotically equivalent to

$$
f\{l(x,y)\} g\{l(x,y)\} \left(-\{w_\ell x - (1-w_\ell)y\} \left[f'\{l(x,y)\} - \frac{f\{l(x,y)\} - l(x,y)f'\{l(x,y)\}}{l(x,y)} \right] \right)
$$

$$
= f\{l(x,y)\} g\{l(x,y)\} \left[-\{w_\ell x - (1-w_\ell)y\} \frac{(x+y)f\{l(x,y)\}}{w_\ell x - (1-w_\ell)y} \right]
$$

$$
= -(x+y)f^2 \left(\frac{y}{x+y} - w_\ell \right) h \left(\frac{y}{x+y} \right),
$$

which completes the proof.

A.6. Proof of Proposition 4

Assuming $y \to \infty$ as $x \to \infty$ with $y/x \to 0$, we have from expressions (10), (25), (27), (28) and Lemma 1, that for large x, the log-conditional survivor, $\ln \Pr(Y > y \mid X = x)$, is approximately equal to

$$
-\frac{x}{(\kappa\gamma)^2} (y/x)^{\delta+2(1+\gamma)} \exp\left\{ -\kappa (y/x)^{-\gamma} \right\}. \tag{45}
$$

Now, let $z \in \mathbb{R}$ and $y = a(x) + b(x)z$, where $a(x)$ and $b(x)$ are given by Eqs. (29). We have, as $x \to \infty$,

$$
(y/x)^{\delta+2(1+\gamma)} = \left(\frac{\ln x}{\kappa} \right)^{-\{\delta+2(1+\gamma)\}/\gamma} \left\{ 1 + O\left(\frac{\ln \ln x}{\ln x} \right) \right\}, \tag{46}
$$

$$
\exp\left\{ -\kappa (y/x)^{-\gamma} \right\} = x^{-1} \left(\frac{\ln x}{\kappa} \right)^{\{\delta+2(1+\gamma)\}/\gamma} \exp\left(\gamma\kappa^{-1/\gamma}z \right) \left[1 + O\left\{ \left(\frac{\ln \ln x}{\ln x} \right)^2 \right\} \right]. \tag{47}
$$

Combining Eqs. (45)–(47) we get

$$
\Pr\{Y < a(x) + b(x)z \mid X = x\} = 1 - \exp\left[-\left\{ \kappa^{(\delta+2)/\gamma}/\gamma^2 \right\} \exp\left(\gamma\kappa^{-1/\gamma}z \right) \right] + O\left(\frac{\ln \ln x}{\ln x} \right).
$$

Last, direct application of statement (9) yields the result of Proposition 4.

References

[1] F. Ballani, M. Schlather, A construction principle for multivariate extreme value distributions, Biometrika 98 (2011) 633–645.
[2] B.M. Brown, S.I. Resnick, Extreme values of independent stochastic processes, J. Appl. Probab. 14 (1977) 732–739.
[3] S.G. Coles, An Introduction to Statistical Modeling of Extreme Values, Springer–Verlag, London, 2001.
[4] S.G. Coles, J.A. Tawn, Modelling extreme multivariate events, J. R. Stat. Soc. Ser. B Stat. Methodol. 53 (1991) 377–392.
[5] A.C. Davison, R. Huser, E. Thibaud, Geostatistics of dependent and asymptotically independent extremes, Math. Geosci. 45 (2013) 511–529.
[6] A.C. Davison, S. Padoan, M. Ribatet, Statistical modelling of spatial extremes (with discussion), Statist. Sci. 27 (2012) 161–186.
[7] L. de Haan, On Regular Variation and its Application to the Weak Convergence of Sample Extremes, in: Mathematical Centre Tracts, vol. 32, Mathematisch Centrum, Amsterdam, 1970.
[8] L. de Haan, A spectral representation for max-stable processes, Ann. Probab. 12 (1984) 1194–1204.
[9] E.F. Eastoe, J.A. Tawn, The distribution for the cluster maxima of exceedances of sub-asymptotic thresholds, Biometrika 99 (2012) 43–55.
[10] S. Engelke, A. Malinowski, Z. Kabluchko, M. Schlather, Estimation of Hüsler-Reiss distributions and Brown–Resnick processes, J. R. Stat. Soc. Ser. B Stat. Methodol. 77 (2015) 239–265.
[11] M.G. Genton, Y. Ma, H. Sang, On the likelihood function of Gaussian max-stable processes, Biometrika 98 (2011) 481–488.
[12] J.E. Heffernan, S.I. Resnick, Limit laws for random vectors with an extreme component, Ann. Appl. Probab. 17 (2007) 537–571.
[13] J.E. Heffernan, J.A. Tawn, A conditional approach for multivariate extreme values (with discussion), J. R. Stat. Soc. Ser. B Stat. Methodol. 66 (3) (2004) 497–546.
[14] S. Hilal, S.-H. Poon, J.A. Tawn, Hedging the black swan: conditional heteroskedasticity and tail dependence in S&P500 and VIX, J. Bank. Finance 35 (2011) 2374–2387.
[15] R. Huser, A.C. Davison, Composite likelihood estimation for the Brown-Resnick process, Biometrika 100 (2013) 511–518.
[16] J. Hüsler, R.-D. Reiss, Maxima of normal random vectors: between independence and complete dependence, Statist. Probab. Lett. 7 (1989) 283–286.
[17] H. Joe, Multivariate Models and Dependence Concepts, Chapman & Hall, London, 1997.
[18] Z. Kabluchko, M. Schlather, L. de Haan, Stationary max-stable fields associated to negative definite functions, Ann. Probab. 37 (2009) 2042–2065.
[19] C. Keef, I. Papastathopoulos, J.A. Tawn, Estimation of the conditional distribution of a multivariate variable given that one of its components is large: additional constraints for the Heffernan and Tawn model, J. Multivariate Anal. 115 (2013) 396–404.
[20] C. Keef, C. Svensson, J.A. Tawn, Spatial dependence in extreme river flows and precipitation for Great Britain, J. Hydrol. 378 (2009) 240–252.
[21] R. Kulik, P. Soulier, Heavy tailed time series with extremal independence, Extremes 18 (2015) 1–27.
[22] A.W. Ledford, J.A. Tawn, Statistics for near independence in multivariate extreme values, Biometrika 83 (1996) 169–187.
[23] A.W. Ledford, J.A. Tawn, Modelling dependence within joint tail regions, J. R. Stat. Soc. Ser. B Stat. Methodol. 59 (1997) 475–499.
[24] A.W. Marshall, I. Olkin, A multivariate exponential distribution, J. Amer. Statist. Assoc. 62 (1967) 30–44.
[25] R.B. Nelsen, An Introduction to Copulas, second ed., in: Springer Series in Statistics, Springer, New York, 2006.
[26] A.K. Nikoloulopoulos, H. Joe, H. Li, Extreme value properties of multivariate t copulas, Extremes 12 (2009) 129–148.
[27] T. Opitz, Extremal t processe: elliptical domain of attraction and a spectral representation, J. Multivariate Anal. 122 (2013) 409–413.
[28] I. Papastathopoulos, J.A. Tawn, Stochastic ordering under conditional modelling of extreme values: drug-induced liver injury, J. Roy. Statist. Soc. Ser. C 64 (2015) 299–317.

[29] M. Paulo, H. van der Voet, J. Wood, G. Marion, J. van Klaveren, Analysis of multivariate extreme intakes of food chemicals, Food Chem. Toxicol. 44 (7) (2006) 994–1005.
[30] S.I. Resnick, Extreme Values, Regular Variation and Point Processes, Springer–Verlag, New York, 1987.
[31] S.I. Resnick, Hidden regular variation, second order regular variation and asymptotic independence, Extremes 5 (2002) 303–336.
[32] S.I. Resnick, D. Zeber, Transition kernels and the conditional extreme value model, Extremes 17 (2) (2014) 263–287.
[33] M. Schlather, Models for stationary max-stable random fields, Extremes 5 (2002) 33–44.
[34] R.L. Smith, Max-stable processes and spatial extremes, Technical Report, University of North Carolina, 1990.
[35] R.L. Smith, Discussion of paper by J. Heffernan, and J.A. Tawn, J. R. Stat. Soc. Ser. B Stat. Methodol. 66 (2004) 530–532.
[36] J.A. Tawn, Bivariate extreme value theory: models and estimation, Biometrika 75 (1988) 397–415.
[37] E. Thibaud, R. Mutzner, A.C. Davison, Threshold modeling of extreme spatial rainfall, Water Resour. Res. 49 (2013) 4633–4644.
[38] J.L. Wadsworth, J.A. Tawn, Dependence modelling for spatial extremes, Biometrika 99 (2012) 1–20.
[39] J.L. Wadsworth, J.A. Tawn, A new representation for multivariate tail probabilities, Bernoulli 19 (2013) 2689–2714.
[40] J.L. Wadsworth, J.A. Tawn, A.C. Davison, D.M. Elton, Modelling across extremal dependence classes, J. R. Stat. Soc. Ser. B Stat. Methodol. (2016) http://dx.doi.org/10.1111/rssb.12157.