Bernoulli **17**(1), 2011, 226–252
DOI: 10.3150/10-BEJ271

Conditioning on an extreme component: Model consistency with regular variation on cones

BIKRAMJIT DAS[1] and SIDNEY I. RESNICK[2]

[1]*RiskLab, Department of Mathematics, ETH Zürich, 8092 Zürich, Switzerland.*
E-mail: bikram@math.ethz.ch
[2]*School of Operations Research and Information Engineering, Cornell University, Ithaca, NY 14853, USA.*
E-mail: sir1@cornell.edu

Multivariate extreme value theory assumes a multivariate domain of attraction condition for the distribution of a random vector. This necessitates that each component satisfies a marginal domain of attraction condition. An approximation of the joint distribution of a random vector obtained by conditioning on one of the components being extreme was developed by Heffernan and Tawn [12] and further studied by Heffernan and Resnick [11]. These papers left unresolved the consistency of different models obtained by conditioning on different components being extreme and we here provide clarification of this issue. We also clarify the relationship between these conditional distributions, multivariate extreme value theory and standard regular variation on cones of the form $[0,\infty]\times(0,\infty]$.

Keywords: asymptotic independence; conditional extreme value model; domain of attraction; regular variation

1. Introduction

Classical multivariate extreme value theory (abbreviated as MEVT) captures the extremal dependence structure between components under a robust multivariate domain of attraction condition which requires that each marginal distribution belongs to the (maximum) domain of attraction (hereafter abbreviated as DOA) of some univariate extreme value distribution. Extremal dependence has been well studied, both in the case of asymptotic dependence and asymptotic independence [6,7,15,16,18,24–29]. An innovative approach was provided by Heffernan and Tawn [12], who approximated multivariate distributions by assuming that only one of the components was in an extreme value domain of attraction and that this component was extreme. Their approach allowed for a variety of examples of different types of asymptotic dependence and asymptotic independence. Their statistical ideas were given a more mathematical framework by Heffernan and Resnick [11] after slight changes in the assumptions which make the theory more probabilistically viable.

1350-7265

In [11], a bivariate random vector (X, Y) is considered, where the distribution of Y is in the DOA of an extreme value distribution G_γ, where, for $\gamma \in \mathbb{R}$,

$$G_\gamma(x) = \exp\{-(1+\gamma x)^{-1/\gamma}\}, \qquad 1+\gamma x > 0. \tag{1.1}$$

For $\gamma = 0$, the distribution function is interpreted as $G_0(x) = \exp\{-e^{-x}\}, x \in \mathbb{R}$. Instead of conditioning on Y being large, their theory was developed under the equivalent assumption of the existence of a vague limit for the modified joint distribution of a suitably scaled and centered (X, Y). The precise definition (Definition 1.1) is given in Section 1.2 and defines the *conditional extreme value* (CEV) model. The CEV model differs from classical MEVT and does not assume that the distribution of (X, Y) is in a multivariate DOA. Only one of the marginal distributions is assumed to be in the univariate DOA of an extreme value distribution.

The CEV model is useful in two contexts. In the first, the MEV model holds, but asymptotic independence makes it difficult to compute probabilities of risk regions; in this case, the CEV model, if applicable, provides a supplementary assumption to the MEV model and thus may provide better risk estimates. Therefore, both the MEV and CEV models are assumed to hold. This is the way in which hidden regular variation may be used; see [28] for some background. In the other context, we do not assume that (X, Y) is in a multivariate domain of attraction and the CEV assumptions may still hold; the CEV model is then a standalone model. In a study of Internet traffic data [3,17], one variable was found to be not in any univariate DOA and hence MEVT was not applicable, but the conditional model was still valid.

In Section 3, we complete the study of the relationship between multivariate extreme value theory and conditioned limit theory begun in [11]. The connection is through the theory of regular variation on cones. The defining relation of MEVT can be standardized to produce standard regular variation on the cone $[0,\infty]^2 \setminus \{(0,0)\}$. The limit relation in conditioned limit theory can *sometimes* be standardized to regular variation on the smaller cone $[0,\infty] \times (0,\infty]$. We explain the precise circumstances when the CEV model can be standardized to regular variation.

Section 2 studies a consistency question for conditional models related to one raised in [12] and its discussion following the paper. In practice, for a vector (X, Y), one has a choice of whether to condition on X being large or Y being large and, depending on the choice, different models are potentially possible. We show that if conditional approximations are possible no matter which variable is chosen as the conditioning variable, then, in fact, the joint distribution is in a classical multivariate DOA of an extreme value law. A related issue is when the CEV model can be extended to a classical MEV model; Section 4 provides conditions for this. Section 5 relates *hidden regular variation* [18,26] and the CEV model under the assumption of multivariate extreme value DOA for (X, Y) with asymptotic independence. Finally, Section 6 presents some examples in order to demonstrate features of the conditioned models and the final section supplies some deferred proofs.

1.1. Notation

Below, we list some commonly used notation and provide some references.

$\mathbb{R}_+^d$	$[0,\infty)^d$. Also, similarly denote $\overline{\mathbb{R}}_+^d = [0,\infty]^d$, $\overline{\mathbb{R}}^d = [-\infty,\infty]^d$.
$\mathbb{E}^*$	A nice subset of the compactified finite-dimensional Euclidean space, often denoted $\mathbb{E}$ with different subscripts and superscripts, as required.
$\mathcal{E}^*$	The Borel σ-field of the subspace $\mathbb{E}^*$.
$\mathbb{M}_+(\mathbb{E}^*)$	The class of Radon measures on Borel subsets of $\mathbb{E}^*$.
$f^\leftarrow$	The left-continuous inverse of a monotone function f. For an increasing function, $f^\leftarrow(x) = \inf\{y : f(y) \geq x\}$. For a decreasing function, $f^\leftarrow(x) = \inf\{y : f(y) \leq x\}$.
RV_ρ	The class of regularly varying functions with index ρ; see [1,6,10,29,32].
Π	The class of Π-varying functions; see [1,29].
$\mathbb{E}^{(\gamma)}$	$\{x : 1+\gamma x > 0\}$ for $\gamma \in \mathbb{R}$.
$\overline{\mathbb{E}}^{(\gamma)}$	The closure on the right of the interval $\mathbb{E}^{(\gamma)}$.
$\overline{\overline{\mathbb{E}}}^{(\gamma)}$	The closure on both sides of the interval $\mathbb{E}^{(\gamma)}$.
$\mathbb{E}^{(\lambda,\gamma)}$	$\overline{\overline{\mathbb{E}}}^{(\lambda)} \times \overline{\overline{\mathbb{E}}}^{(\gamma)} \setminus \{(-\frac{1}{\lambda}, -\frac{1}{\gamma})\}$.
$\mathbb{E}$	Usually $[0,\infty]^2 \setminus \{\mathbf{0}\}$.
$\mathbb{E}_0$	Usually $(0,\infty]^2$.
$\mathbb{E}_\sqcap$	$[0,\infty] \times (0,\infty]$. Similarly, $\mathbb{E}_\sqsupset = (0,\infty] \times [0,\infty]$.
$\xrightarrow{v}$	Vague convergence of measures; see [13,22].
G_γ	An extreme value distribution given by (1.1) with parameter $\gamma \in \mathbb{R}$.
$D(G_\gamma)$	The DOA of the extreme value distribution G_γ; in other words, the set of F's satisfying (1.6). For $\gamma > 0$, $F \in D(G_\gamma)$ is equivalent to $1 - F \in RV_{-1/\gamma}$.

1.2. Model setup and basic assumptions

Our model assumptions follow those of [11].

Definition 1.1 (Conditional extreme value model). *Suppose that* $(X, Y) \in \mathbb{R}^2$ *is a random vector and that there exist functions* $\alpha(t) > 0$, $a(t) > 0$, $\beta(t)$, $b(t) \in \mathbb{R}$, *a constant* $\gamma \in \mathbb{R}$ *and a non-null Radon measure* μ *on Borel subsets of* $[-\infty,\infty] \times \overline{\mathbb{E}}^{(\gamma)}$ *such that*

$$\text{(a)}\quad t\mathbf{P}\left(\left(\frac{X-\beta(t)}{\alpha(t)}, \frac{Y-b(t)}{a(t)}\right) \in \cdot\right) \xrightarrow{v} \mu(\cdot) \qquad \text{in } \mathbb{M}_+\big([-\infty,\infty] \times \overline{\mathbb{E}}^{(\gamma)}\big). \tag{1.2}$$

Assume that μ *satisfies the following conditional non-degeneracy conditions: for each* $y \in \mathbb{E}^\gamma$,

$$\text{(b)}\quad \mu\big([-\infty,x] \times (y,\infty]\big) \text{ is not a degenerate distribution in } x, \tag{1.3}$$

$$\text{(c)}\quad \mu\big([-\infty,x] \times (y,\infty]\big) < \infty. \tag{1.4}$$

Additionally, we assume that

$$\text{(d)} \quad H(x) := \mu\big([-\infty, x] \times (0, \infty]\big) \text{ is a probability distribution.} \tag{1.5}$$

We say that (X, Y) *follows a conditional extreme value model* (*abbreviated CEV model*) *if conditions* (1.2)–(1.5) *hold. We write* $(X, Y) \in CEV(\alpha, \beta, a, b, \gamma)$ *and often ignore the parameters for generic usage.*

Since convergence in (1.2) holds in $\mathbb{M}_+([-\infty, \infty] \times \overline{\mathbb{E}}^{(\gamma)})$, it also holds without the X variable and so if the marginal distribution of Y is F, then $F \in D(G_\gamma)$ for $\gamma \in \mathbb{R}$, as defined in (1.1); that is, as $t \to \infty$,

$$t\big(1 - F\big(a(t)y + b(t)\big)\big) = t\mathbf{P}\bigg(\frac{Y - b(t)}{a(t)} > y\bigg) \to (1 + \gamma y)^{-1/\gamma}, \qquad 1 + \gamma y > 0. \tag{1.6}$$

Also, conditions (1.2), (1.3) and (1.4) imply that if $(x, 0)$ is a continuity point of $\mu(\cdot)$, then

$$\mathbf{P}\bigg(\frac{X - \beta(t)}{\alpha(t)} \le x \bigg| Y > b(t)\bigg) \to H(x) = \mu\big([-\infty, x] \times (0, \infty]\big) \qquad \text{as } t \to \infty, \tag{1.7}$$

that is, a conditioned limit holds. This accounts for the name *conditional extreme value model* and we can think of Y, the variable in a univariate DOA, as the *conditioning variable*. Under the above assumptions, a convergence to types argument [11] yields properties of the scaling and centering functions: there exist functions $\psi_1, \psi_2 : \mathbb{R}_+ \mapsto \mathbb{R}$ such that for $c > 0$,

$$\lim_{t\to\infty} \frac{\alpha(tc)}{\alpha(t)} = \psi_1(c), \qquad \lim_{t\to\infty} \frac{\beta(tc) - \beta(t)}{\alpha(t)} = \psi_2(c). \tag{1.8}$$

This implies that $\psi_1(c) = c^\rho$ for some $\rho \in \mathbb{R}$ ([6], Theorem B.1.3), and either $\psi_2 \equiv 0$ or $\psi_2(c) = k(c^\rho - 1)/\rho$ for some $k \neq 0$ ([6], Theorem B.2.1).

1.3. Comparison with the model proposed by Heffernan and Tawn

The model discussed in [11] was motivated by ideas of Heffernan and Tawn [12]. The basic premise is that in classical MEVT, probabilities of extreme sets (values which are very high or very low) are calculated under the existence of a joint extreme value limit. However, in practice, we sometimes observe that only a subset of the components is extreme or, alternatively, we are interested in regions where all extreme values do not occur together. Let us look at a description of Heffernan and Tawn's model with $d = 2$ for simplicity and ease of comparison with the formulation used in this paper.

(1) Assume that $\mathbf{X} = (X_1, X_2)$ is a random vector with joint distribution F_X and marginal distributions F_1 and F_2. Also, assume that we have n i.i.d. copies of $\mathbf{X}$.
(2) Assume that C is an extreme set in the sense that, for any element in C, at least one of its components is extreme. Define

$$C_i = C \cap \{\mathbf{x} \in \mathbb{R}^d : F_i(x_i) > F_j(x_j)\}, \qquad i = 1, 2, \ j \neq i.$$

Also, define $\upsilon_{X_i} = \inf_{\mathbf{x}\in C_i}(x_i), i = 1, 2$, and assume that each $\mathbf{P}(X_i > \upsilon_{X_i})$ is close to 1, making the C_i's extreme and hence making C an extreme set. We then write

$$\mathbf{P}(\mathbf{X} \in C) = \mathbf{P}(\mathbf{X} \in C_1) + \mathbf{P}(\mathbf{X} \in C_2) = \sum_{i=1}^{2} \mathbf{P}(\mathbf{X} \in C_i | X_i > \upsilon_{X_i})\mathbf{P}(X_i > \upsilon_{X_i}).$$

(3) The X_i's are marginally assumed to be extreme-valued. A generalized Pareto distribution is fitted to each of the marginals above a threshold, in the case above, υ_{X_i}; see [23]. Below the threshold, the marginals are approximated by an empirical distribution. Denote the estimate of the marginal distributions by $\hat{F}_i$.

(4) All the marginals are transformed to Gumbel marginals using the transformation

$$Y_i = -\log[-\log\{\hat{F}_i(x_i)\}], \qquad i = 1, 2.$$

(5) In order to estimate $\mathbf{P}(\mathbf{Y} \in C_i | Y_i > \upsilon_{y_i})$ (the transformed case), a conditioned limit is assumed, as follows: there exist normalizing vectors $a_1(y), a_2(y), b_1(y), b_2(y) \in \mathbb{R}$ such that

$$\lim_{y_i\to\infty} \mathbf{P}\big(Y_j \le a_j(y_i)y_j + b_j(y_i) | Y_i = y_i\big) = G_i(y_j), \qquad i = 1, 2, y_j \in \mathbb{R}. \tag{1.9}$$

(6) The parameters $a_i(y)$ and $b_i(y)$ are estimated by assuming a parametric structural form; see [12] for details.

For the model defined in [11] and used in this paper:

(1) We start with the same assumption (1).

(2) We focus on one of the extreme sets C_1 and C_2; without loss of generality, assume this is C_2. We assume only one of the marginals X_2 is extreme-valued. In [12], all of the marginals are extreme-valued.

(3) Instead of fitting an exact GPD over a threshold for the marginal distribution, we assume that $X_2 \in D(G)$, in the sense of (1.6).

(4) Instead of Gumbel marginals, we transform to Pareto marginals for X_2, which facilitates the use of tools from standard regular variation theory. Thus, the transformation here is $X_2^* = 1/(1 - F(X_2))$ and X_1 remains unchanged.

(5) In order to estimate $\mathbf{P}((X_1, X_2^*) \in C_2 | X_2^* > x_2)$, a conditioned limit is assumed: there exist normalizations $\alpha(x_2) > 0, \beta(x_2) \in \mathbb{R}$ such that

$$\lim_{x_2\to\infty} \mathbf{P}\big(X_1 \le \alpha(x_2)x_1 + \beta(x_2) | X_2^* > x_2\big) = G(x_1), \qquad x_1 \in \mathbb{R}. \tag{1.10}$$

This is equivalent to (1.2) when $Y = X_2^*$ has been standardized to a Pareto margin.

A technique for estimating model parameters has been discussed in [9]. Further, if one makes precise what version of the conditional distribution is being used in (1.9), then, as expected, it is shown in [30] that (1.9) implies (1.10), but (1.10) may hold without (1.9) holding.

2. A consistency result for conditional extreme value models

The CEV model defined in Section 1.2 is not symmetric in X and Y. So, given bivariate data, which component should serve as the conditioning variable? A similar issue was raised in [12] and [11]. Heffernan and Tawn [12] considered (X, Y) in a multivariate DOA with asymptotic independence, introduced the supplementary assumption that a conditional model was also valid and raised the question of criteria for deciding which variable to make the conditioning variable. If either variable could be made the conditioning variable, then they considered self-consistency of the two conditional models. Assuming densities, they provided a natural constraint of equality of joint limiting densities under each model for the common region where both models were defined. We consider a related problem without assuming that (X, Y) has a distribution in a multivariate domain.

Definition 1.1 does not assume that the distribution of (X, Y) is in a multivariate DOA. Suppose that $X \sim F_X, Y \sim F_Y$. Assume that $(X, Y) \in CEV(\alpha, \beta, a, b, \gamma)$ with limit measure $\mu_{X,Y>}(\cdot)$ and $F_Y \in D(G_\gamma)$, and also $(Y, X) \in CEV(c, d, \chi, \phi, \lambda)$ with limit measure $\mu_{Y,X>}(\cdot)$ and $F_X \in D(G_\lambda)$. Assuming both conditional models implies that (X, Y) is in the DOA of a bivariate extreme value distribution G. If the limit distribution G is not a product measure, then $\mu_{X,Y>}$ and $\mu_{Y,X>}$ are equal up to linear transformation on subsets that are defined on the intersection of the domains of both measures. Recall that if, marginally, $F_X \in D(G_\lambda)$ and $F_Y \in D(G_\gamma)$, then we do not necessarily have $(X, Y) \in D(G)$ for a bivariate extreme value distribution G; see [31]. The precise consistency statement is next; the proof is deferred to Section 7.

Theorem 2.1. *Suppose we have a bivariate random vector $(X, Y) \in \mathbb{R}^2$, non-negative functions $\alpha(\cdot), a(\cdot)$, $\chi(\cdot), c(\cdot)$ and real-valued functions $\beta(\cdot), b(\cdot), \phi(\cdot), d(\cdot)$ such that $(X, Y) \in CEV(\alpha, \beta, a, b, \gamma)$, that is,*

$$t\mathbf{P}\left[\left(\frac{X-\beta(t)}{\alpha(t)}, \frac{Y-b(t)}{a(t)}\right) \in \cdot\right] \xrightarrow{v} \mu_{X,Y>}(\cdot) \qquad \text{in } \mathbb{M}_+\big([-\infty,\infty] \times \overline{\mathbb{E}}^{(\gamma)}\big) \tag{2.1}$$

and $(Y, X) \in CEV(c, d, \chi, \phi, \lambda)$, that is,

$$t\mathbf{P}\left(\left(\frac{X-\phi(t)}{\chi(t)}, \frac{Y-d(t)}{c(t)}\right) \in \cdot\right) \xrightarrow{v} \mu_{Y,X>}(\cdot) \qquad \text{in } \mathbb{M}_+\big(\overline{\mathbb{E}}^{(\lambda)} \times [-\infty,\infty]\big) \tag{2.2}$$

for $\lambda, \gamma \in \mathbb{R}$, where both $\mu_{X,Y>}$ and $\mu_{Y,X>}$ satisfy conditional non-degeneracy conditions corresponding to (1.3) and (1.4). Then (X, Y) is in the DOA of a multivariate extreme value distribution on $\mathbb{E}^{(\lambda,\gamma)}$, that is,

$$t\mathbf{P}\left(\left(\frac{X-\phi(t)}{\chi(t)}, \frac{Y-b(t)}{a(t)}\right) \in \cdot\right) \xrightarrow{v} \mu_{X,Y}(\cdot) \qquad \text{in } \mathbb{M}_+\big(\mathbb{E}^{(\lambda,\gamma)}\big), \tag{2.3}$$

where $\mu_{X,Y}(\cdot)$ is a non-null Radon measure on $\mathbb{E}^{(\lambda,\gamma)}$.

Remark 2.1. Theorem 2.1 does not impose a restriction on the scaling and centering functions of X and Y, which means that the joint conditional convergences (2.1) and (2.2) impose sufficient regularity so that (X, Y) belongs to a joint DOA. Equation (2.3) says that $\mu_{X,Y}$ is the exponent measure of an extreme value distribution G. The following are further consequences from the proof of Theorem 2.1:

(1) if (X, Y) is not asymptotically independent, then we get $\alpha \sim k_1\chi$ and $c \sim k_2 a$ for some non-zero constants k_1, k_2, hence $\mu_{X,Y>}$ and $\mu_{Y,X>}$ are equal up to linear transformations;
(2) if (X, Y) is asymptotically independent, then $\lim_{t\to\infty} \alpha(t)/\chi(t) = 0$, $\lim_{t\to\infty} c(t)/a(t) = 0$.

Consistency: Standard regularly varying case. We were led to Theorem 2.1 by considering the special case of standard regular variation where (X, Y) satisfies $(X, Y) \in CEV(\alpha(t) = t, \beta(t) = 0, a(t) = t, b(t) = 0, \gamma = 1)$, $(Y, X) \in CEV(\alpha(t) = t, \beta(t) = 0, a(t) = t, b(t) = 0, \gamma = 1)$ and the vague convergence in (1.2) is regular variation on the cone $\mathbb{E}_\sqcap = [0, \infty] \times (0, \infty]$ ([4,28], [27], page 173). We can show [2] that if

$$tP[t^{-1}(X, Y) \in \cdot] \xrightarrow{v} \mu_{X,Y>}(\cdot) \qquad \text{in } \mathbb{M}_+(\mathbb{E}_\sqcap), \tag{2.4}$$

$$tP[t^{-1}(X, Y) \in \cdot] \xrightarrow{v} \mu_{Y,X>}(\cdot) \qquad \text{in } \mathbb{M}_+(\mathbb{E}_\sqsupset), \tag{2.5}$$

where $\mu_{X,Y>}$ and $\mu_{Y,X>}$ satisfy the conditional non-degeneracy conditions (1.3) and (1.4), then (X, Y) is standard regularly varying on $\mathbb{E} := [0, \infty]^2 \setminus \{\mathbf{0}\}$, that is,

$$t\mathbf{P}\left[\left(\frac{X}{t}, \frac{Y}{t}\right) \in \cdot\right] \xrightarrow{v} \mu_{X,Y}(\cdot) \qquad \text{in } \mathbb{M}_+(\mathbb{E}), \tag{2.6}$$

where $\mu_{X,Y}$ is a Radon measure on $\mathbb{E}$ such that

$$\mu_{X,Y}|_{\mathbb{E}_\sqcap}(\cdot) = \mu_{X,Y>}(\cdot) \qquad \text{on } \mathbb{E}_\sqcap \quad \text{and} \quad \mu_{X,Y}|_{\mathbb{E}_\sqsupset}(\cdot) = \mu_{Y,X>}(\cdot) \qquad \text{on } \mathbb{E}_\sqsupset.$$

A proof and discussion of the absolutely continuous case is in [2].

Example 1. Suppose that (X, Y) is a bivariate random variable with joint density

$$f_{X,Y}(x, y) = \frac{4x}{(x^2 + y)^3} + \frac{4y}{(x + y^2)^3}, \qquad x \geq 1, y \geq 1.$$

The following hold as $t \to \infty$:

$$t^2 f_X(tx) \to \frac{2}{x^2}, \qquad t^2 f_Y(ty) \to \frac{2}{y^2}, \qquad x, y > 0,$$

$$t^{5/2} f_{X,Y}(tx, \sqrt{t}y) \to \frac{4y}{(x + y^2)^3} =: g_1(x, y) \in L_1(\mathbb{E}_\sqsupset),$$

$$t^{5/2} f_{X,Y}(\sqrt{t}x, ty) \to \frac{4x}{(x^2 + y)^3} =: g_1(x, y) \in L_1(\mathbb{E}_\sqcap).$$

This means that the conditions of Theorem 2.1 hold. Note that we here have identical Pareto marginals. We are thus led to the analogs of (1.2) on different cones:

$$t\mathbf{P}\left(\frac{X}{\sqrt{t}} \le x, \frac{Y}{t} > y\right) \to \frac{1}{y} - \frac{1}{y+x^2}, \qquad x \ge 0, y > 0,$$

$$t\mathbf{P}\left(\frac{X}{t} > x, \frac{Y}{\sqrt{t}} \le y\right) \to \frac{1}{x} - \frac{1}{x+y^2}, \qquad x > 0, y \ge 0,$$

$$t\mathbf{P}\left(\left(\frac{X}{t}, \frac{Y}{t}\right) \in ([0,x] \times [0,y])^c\right) \to \frac{1}{x} + \frac{1}{y}, \qquad x > 0, y > 0.$$

3. The CEV model and standard regular variation

As remarked after Theorem 2.1, questions about the general conditional model are effectively analyzed by starting with standard regular variation on the cones $\mathbb{E}_{\sqsupset}$ or $\mathbb{E}_{\sqcap}$. It is theoretically useful to know when standardization of the conditional extreme value model is possible. A partial answer appears in [11], Section 2.4, and we consider this issue in more detail, starting with a review and definition of *standardization* [6,7,27,29].

3.1. Standardization

Standardization is the process of marginally transforming a random vector $\mathbf{X}$ into a different vector $\mathbf{Z}^*$, $\mathbf{X} \mapsto \mathbf{Z}^*$, so that the distribution of $\mathbf{Z}^*$ is standard regularly varying on a cone $\mathbb{E}^*$; that is, for some Radon measure $\mu^*(\cdot)$,

$$t\mathbf{P}[t^{-1}\mathbf{Z}^* \in \cdot] \xrightarrow{v} \mu^*(\cdot) \qquad \text{in } \mathbb{M}_+(\mathbb{E}^*). \tag{3.1}$$

Depending on the cone, one or more components of $\mathbf{Z}^*$ are asymptotically Pareto. For classical multivariate extreme value theory, each component is asymptotically Pareto and $\mathbb{E}^* = \mathbb{E} = [\mathbf{0}, \infty] \setminus \{\mathbf{0}\}$. The technique is used in classical multivariate extreme value theory to characterize multivariate domains of attraction and dates back to at least [7]; see also [6,19,20,27] and [29], Chapter 5. Standardization is analogous to the copula transformation, but is better suited to studying limit relations [14].

In Cartesian coordinates, the limit measure in (3.1) has the scaling property

$$\mu^*(c\cdot) = c^{-1}\mu^*(\cdot), \qquad c > 0. \tag{3.2}$$

This scaling in Cartesian coordinates translates to a product limit when expressed in polar coordinates. An angular measure exists, allowing the characterization of limits

$$\mu^*\left\{\mathbf{x} : \|\mathbf{x}\| > r, \frac{\mathbf{x}}{\|\mathbf{x}\|} \in \Lambda\right\} = r^{-1} S(\Lambda)$$

for Borel subsets Λ of the unit sphere in $\mathbb{E}^*$.

In classical multivariate extreme value theory, S is a finite measure which we may take to be a probability measure without loss of generality. However, when $\mathbb{E}^* = \mathbb{E}_\sqcap$, S is *not* necessarily finite because absence of the horizontal axis boundary in $\mathbb{E}_\sqcap$ implies the unit sphere is not compact.

Here is an explicit description of *standardization*. Suppose that $\mathbf{X} = (X_1, X_2, \dots, X_d)$ is a random vector in $\mathbb{R}^d$ which satisfies

$$tP\left[\left(\frac{X_1 - \beta_1(t)}{\alpha_1(t)}, \frac{X_2 - \beta_2(t)}{\alpha_2(t)}, \dots, \frac{X_d - \beta_d(t)}{\alpha_d(t)}\right) \in \cdot\right] \xrightarrow{v} \mu(\cdot) \qquad \text{in } \mathbb{M}_+(\mathfrak{D}) \tag{3.3}$$

for some $\mathfrak{D} \subset \overline{\mathbb{R}}^d$, $\alpha_i(t) > 0$, $\beta_i(t) \in \mathbb{R}$ for $i = 1, \dots, d$. Suppose that we have $\mathbf{f} = (f_1, \dots, f_d)$ such that, for $i = 1, \dots, d$:

(a) f_i : range of $X_i \to (0, \infty)$;
(b) f_i is monotone;
(c) $\nexists K > 0$ such that $|f_i| \le K$.

Then $\mathbf{f}$ *standardizes* $\mathbf{X}$ if $\mathbf{Z}^* = \mathbf{f}(\mathbf{X}) = (f_i(X_i), i = 1, \dots, d)$ satisfies (3.1). We call $\mathbf{f}$ the *standardizing function* and say (3.1) is the *standardization* of (3.3).

For the conditional model defined in Definition 1.1 in Section 1.2, where F, the distribution of Y, satisfies $F \in D(G_\gamma)$, we can always use $b(\cdot) = (1/(1-F))^{\leftarrow}(\cdot)$ to standardize Y and $Y^* = b^{\leftarrow}(Y)$ is the standardization of Y; see [11].

3.2. When can the conditional extreme value model be standardized?

Suppose that (X, Y) satisfies Definition 1.1 and, in particular, (1.2) holds. Standardization in (1.2) is possible *unless* $(\psi_1, \psi_2) = (1, 0)$, which is equivalent to the limit measure being a product measure [11]. The converse is also true. Consequently, when the limit measure is not a product measure, we can reduce to standard regular variation on the cone $\mathbb{E}_\sqcap$ and, conversely, we can think of the general conditional model as a transformation of standard regular variation on $\mathbb{E}_\sqcap$.

We begin by showing that when we have standardized convergence on $\mathbb{E}_\sqcap$, the limit measure cannot be a product measure.

Lemma 3.1. *Suppose that (X, Y) is standard regularly varying on the cone $\mathbb{E}_\sqcap$ such that*

$$tP[t^{-1}(X, Y) \in \cdot] \xrightarrow{v} \mu(\cdot) \qquad \text{in } \mathbb{M}_+(\mathbb{E}_\sqcap) \tag{3.4}$$

for some non-null Radon measure $\mu(\cdot)$ on $\mathbb{E}_\sqcap$ satisfying the conditional non-degeneracy conditions as in (1.3) and (1.4). Then $\mu(\cdot)$ cannot be a product measure.

Proof. If μ is a product measure, then we have

$$\mu\big([0, x] \times (y, \infty]\big) = G(x) y^{-1} \qquad \text{for } x \ge 0,\ y > 0 \tag{3.5}$$

for some finite distribution function G on $[0,\infty)$. Now, (3.4) implies that μ is homogeneous of order -1, that is,

$$\mu(c\Lambda)=c^{-1}\mu(\Lambda) \qquad \forall c>0, \tag{3.6}$$

where Λ is a Borel subset of $\mathbb{E}_{\sqcap}$. Therefore, using (3.5),

$$\mu\bigl(c([0,x]\times(y,\infty])\bigr)=\mu\bigl([0,cx]\times(cy,\infty]\bigr)=G(cx)(cy)^{-1}=c^{-1}G(cx)y^{-1}.$$

Moreover, using (3.5) and (3.6), $\mu(c([0,x]\times(y,\infty]))=c^{-1}G(x)y^{-1}$ and, therefore, $G(cx)=G(x)$ $\forall c>0, x>0$. Hence, for fixed $y\in\mathbb{E}^{(\gamma)}$, $c>0$, $x>0$,

$$\mu\bigl([0,cx]\times(y,\infty]\bigr)=G(cx)y^{-1}=G(x)y^{-1}=\mu\bigl([0,x]\times(y,\infty]\bigr).$$

Thus, μ becomes a degenerate distribution in x, contradicting our conditional non-degeneracy assumptions and, consequently, $\mu(\cdot)$ cannot be a product measure. □

Suppose we have a general CEV model as in Definition 1.1 with product limit measure. We show this CEV model cannot be *standardized* to regular variation on some cone $\mathfrak{C}\subset\mathbb{E}$ ($\mathfrak{C}=\mathbb{E}_{\sqcap}$ in our case). Since Definition 1.1 implies that Y can always be standardized, in the following, we assume that Y^* is the standardized version of Y and we only consider the problem of standardizing X.

Theorem 3.2. *Suppose that $X\in\mathbb{R}$, $Y^*>0$ are random variables such that for functions $\alpha(\cdot)>0$, $\beta(\cdot)\in\mathbb{R}$, we have, as $t\to\infty$,*

$$t\mathbf{P}\biggl[\biggl(\frac{X-\beta(t)}{\alpha(t)},\frac{Y^*}{t}\biggr)\in\cdot\biggr]\xrightarrow{v}G\times\nu_1(\cdot) \qquad \text{in } \mathbb{M}_+\bigl([-\infty,\infty]\times(0,\infty]\bigr), \tag{3.7}$$

where $\nu_1(x,\infty]=x^{-1}$, $x>0$, and G is some finite, non-degenerate distribution on $\mathbb{R}$. Then there does not exist a standardizing function, $f(\cdot)$: range of $X\mapsto(0,\infty)$, in the sense of the discussion after (3.3), such that

$$t\mathbf{P}[t^{-1}(f(X),Y^*)\in\cdot]\xrightarrow{v}\mu(\cdot) \qquad \text{in } \mathbb{M}_+(\mathbb{E}_{\sqcap}), \tag{3.8}$$

where μ satisfies the conditional non-degeneracy conditions.

Proof. Note that Y^* is already standardized here. Suppose that there exists a standardization function $f(\cdot)$ such that (3.8) holds. Without loss of generality, assume $f(\cdot)$ to be non-decreasing. This implies that for μ-continuity points (x,y), we have

$$t\mathbf{P}[f(X)\le tx, Y^*>ty]\to\mu\bigl((-\infty,x]\times(y,\infty]\bigr) \qquad (t\to\infty),$$

which is equivalent to

$$t\mathbf{P}\biggl(\frac{X-\beta(t)}{\alpha(t)}\le\frac{f^{\leftarrow}(xt)-\beta(t)}{\alpha(t)},\frac{Y^*}{t}>y\biggr]\to\mu\bigl((-\infty,x]\times(y,\infty]\bigr) \qquad (t\to\infty). \tag{3.9}$$

Since $\mu((-\infty, x] \times (y, \infty]) < \infty$ and is non-degenerate in x, we have, as $t \to \infty$, that

$$\big(f^{\leftarrow}(xt) - \beta(t)\big)/\alpha(t) \to h(x) \tag{3.10}$$

for some non-decreasing function $h(\cdot)$ which has at least two points of increase. Thus, (3.9) and (3.10) imply that $\mu((-\infty, x] \times (y, \infty]) = G(h(x)) \times y^{-1}$. Hence, $\mu(\cdot)$ is a product measure which, by Lemma 3.1, is not possible. □

Summary. There follows a summary describing when standardization is possible and the relationship of standardization to the limit measure being a product. Part 2 is proved in Section 7. Statistical methods for detecting when a CEV model is appropriate and whether the limit measure is a product are given in [3]:

(1) Suppose that (X, Y) satisfy Definition 1.1 so that the limits in (1.8) hold. If $(\psi_1, \psi_2) \neq (1, 0)$, then there exists a standardizing function $\mathbf{f} = (f_1, f_2)$ such that $(X^*, Y^*) = (f_1(X), f_2(Y))$ is standard regularly varying on $\mathbb{E}_\sqcap$,

$$t\mathbf{P}[t^{-1}(f_1(X), f_2(Y)) \in \cdot] \xrightarrow{v} \mu^{**}(\cdot) \qquad \text{in } \mathbb{M}_+(\mathbb{E}_\sqcap)$$

and μ^{**} is a non-null Radon measure satisfying the conditional non-degeneracy conditions.

(2) Conversely, suppose that we have a bivariate random vector $(X^*, Y^*) \in \mathbb{R}_+^2$ satisfying

$$t\mathbf{P}[t^{-1}(X^*, Y^*) \in \cdot] \xrightarrow{v} \mu^{**}(\cdot) \qquad \text{in } \mathbb{M}_+(\mathbb{E}_\sqcap),$$

where μ^{**} is a non-null Radon measure satisfying the conditional non-degeneracy conditions. Consider functions $\alpha(\cdot) > 0, \beta(\cdot) \in \mathbb{R}$ such that (1.8) holds with $(\psi_1, \psi_2) \neq (1, 0)$. There then exist functions $a(\cdot) > 0, b(\cdot) \in \mathbb{R}$ satisfying (1.6) and $\lambda(\cdot) \in \mathbb{R}$, $\gamma \in \mathbb{R}$ such that

$$t\mathbf{P}\left[\left(\frac{\lambda(X^*) - \beta(t)}{\alpha(t)}, \frac{b(Y^*) - b(t)}{a(t)}\right) \in \cdot\right] \xrightarrow{v} \tilde{\mu}(\cdot) \qquad \text{in } \mathbb{M}_+\big([-\infty, \infty] \times \overline{\mathbb{E}}^{(\gamma)}\big), \tag{3.11}$$

where $\tilde{\mu}$ is a non-null Radon measure in $[-\infty, \infty] \times \overline{\mathbb{E}}^{(\gamma)}$ satisfying the conditional non-degeneracy conditions and $b(Y^*) \in D(G_\gamma)$.

Remark 3.1. The previous summary applies to attempts to produce a standard pair by marginal transformations. If one waives the requirement that only marginal transformations be used, more is possible. Suppose that H is a non-degenerate probability and, in $\mathbb{M}_+([-\infty, \infty] \times (0, \infty])$,

$$t\mathbf{P}\left[\left(\frac{X - \beta(t)}{\alpha(t)}, \frac{Y^*}{t}\right) \in \cdot\right] \xrightarrow{v} H \times \nu_1(\cdot).$$

Define $X^* = ((X - \beta(Y^*))Y^*)/\alpha(Y^*)$. Then [11] in $\mathbb{M}_+([-\infty, \infty] \times (0, \infty])$,

$$t\mathbf{P}\left(\frac{X^*}{t} \leq x, \frac{Y^*}{t} > y\right) \to \int_0^{1/y} H(xv)\,\mathrm{d}v = \frac{1}{x}\int_0^{x/y} H(s)\,\mathrm{d}s \qquad (t \to \infty).$$

The limit measure is homogeneous of order -1 and, thus, a transformation of (X, Y^*) to a standard regularly varying pair exists, even when we have a limit measure which is a product. Note that this transformation is more complex than just a marginal transformation and is not in the sense of the discussion after (3.3).

3.3. A characterization of regular variation on $\mathbb{E}_\sqcap$

The CEV model with limit measure which is not a product can always be standardized to give regular variation on $\mathbb{E}_\sqcap$, so we would like useful characterizations of such regular variation. Standard regular variation on $\mathbb{E}$ was characterized by [5] in terms of one-dimensional regular variation of max-linear combinations and [26] provides a characterization of hidden regular variation in $\mathbb{E}$ and $\mathbb{E}_0$ in terms of max- and min-linear combinations of the random vector. The following are comparable results for $\mathbb{E}_\sqcap$.

Proposition 3.3. *Suppose that $(X, Y) \in \mathbb{R}_+^2$ is a random vector and $X > 0$ almost surely. The following are equivalent*:

(1) *(X, Y) is standard multivariate regularly varying on $\mathbb{E}_\sqcap$ with limit measure satisfying the non-degeneracy conditions* (1.3) *and* (1.4);
(2) *for all $a \in (0, \infty]$, we have*

$$\lim_{t\to\infty} t\mathbf{P}\big(t^{-1}\min(aX, Y) > y\big) = c(a)y^{-1}, \qquad y > 0, \tag{3.12}$$

for some non-constant, non-decreasing function $c : (0, \infty] \to (0, \infty)$.

Proof. (2) $\Rightarrow$ (1): Assume that (3.12) holds for some function $c : (0, \infty] \to (0, \infty)$. Then, for $x \geq 0$, $y > 0$,

$$\begin{aligned} t\mathbf{P}\left(\frac{X}{t} \leq x, \frac{Y}{t} > y\right) &= t\mathbf{P}\left(\frac{Y}{t} > y\right) - t\mathbf{P}\left(\frac{X}{t} > x, \frac{Y}{t} > y\right) \\ &= t\mathbf{P}(X > 0, Y > ty) - t\mathbf{P}\big((y/x)X > ty, Y > ty\big) \\ &= t\mathbf{P}\big(\min(a_1X, Y) > ty\big) - t\mathbf{P}\big(\min\big((y/x)X, Y\big) > ty\big) \qquad (a_1 := \infty) \\ &\to c(\infty)y^{-1} - c(y/x)y^{-1} \qquad (t \to \infty) \\ &=: \mu([0, x] \times (y, \infty]). \end{aligned}$$

Since $c(\cdot)$ is non-decreasing and non-constant, μ is a non-null Radon measure on $\mathbb{E}_\sqcap$ and we have our result. The non-degeneracy of μ follows from the fact that $c(\cdot)$ is a non-constant function.

(1) $\Rightarrow$ (2): Assume now that (X, Y) is standard multivariate regularly varying on $\mathbb{E}_\sqcap$. Hence, there exists a non-degenerate Radon measure μ on $\mathbb{E}_\sqcap$ such that

$$\lim_{t\to\infty} t\mathbf{P}\left(\frac{X}{t} \leq x, \frac{Y}{t} > y\right) = \mu\big([0, x] \times (y, \infty]\big)$$

and, for any $a \in (0, \infty]$,

$$t\mathbf{P}\left(\frac{\min(aX, Y)}{t} > y\right) = t\mathbf{P}\left(\frac{X}{t} > \frac{y}{a}, \frac{Y}{t} > y\right) \to \mu\left(\left(\frac{y}{a}, \infty\right] \times (y, \infty]\right) \qquad (t \to \infty)$$
$$= y^{-1}\mu\left(\left(\frac{1}{a}, \infty\right] \times (1, \infty]\right) =: c(a)y^{-1},$$

by defining $c(a) = \mu((a^{-1}, \infty] \times (1, \infty])$ and using the homogeneity property (3.2). Note that the conditional non-degeneracy of μ implies that c is non-constant and non-decreasing. □

The condition "$X > 0$ almost surely" in Proposition 3.3 can be removed if we assume that $\lim_{t\to\infty} t\mathbf{P}(Y > t) \to 1$.

3.4. Polar coordinates

Section 3.2 shows that when the limit measure is not a product measure, we can transform (X, Y) to (X^*, Y^*) such that

$$\mathbf{P}[t^{-1}(X^*, Y^*) \in \cdot] \xrightarrow{v} \mu^{**}(\cdot) \qquad \text{in } \mathbb{M}_+(\mathbb{E}_\sqcap). \tag{3.13}$$

Hence, μ^{**} satisfies (3.2) and, when written in polar coordinates, has a spectral form [11], Section 3.2. We summarize some useful facts. For convenience, take the norm $\|(x, y)\| = |x| + |y|$, $(x, y) \in \mathbb{R}^2$, although any other norm would suffice. A standard homogeneity argument [29], Chapter 5, yields, for $r > 0$ and Λ a Borel subset of $[0, 1)$,

$$\begin{aligned}\mu^{**}&\left\{(x, y) \in [0, \infty] \times (0, \infty] : x + y > r, \frac{x}{x + y} \in \Lambda\right\} \\ &= r^{-1}\mu^{**}\left\{(x, y) \in [0, \infty] \times (0, \infty] : x + y > 1, \frac{x}{x + y} \in \Lambda\right\} =: r^{-1}S(\Lambda),\end{aligned} \tag{3.14}$$

where S is a Radon measure on $[0, 1)$. For $x > 0$, $y > 0$, we get, from (3.15),

$$\mu^{**}\big([0, x] \times (y, \infty]\big) = y^{-1}\int_0^{x/(x+y)}(1 - w)S(\mathrm{d}w) - x^{-1}\int_0^{x/(x+y)} wS(\mathrm{d}w). \tag{3.15}$$

S need not be a finite measure on $[0, 1)$, but to guarantee that

$$H^{**}(x) := \mu^{**}\big([0, x] \times (1, \infty]\big) \tag{3.16}$$

is a probability measure, we can see by taking $x \to \infty$ in (3.15) that we need

$$\int_0^1 (1 - w)S(\mathrm{d}w) = 1. \tag{3.17}$$

Conclusion: The class of conditional limits $H^{**}(x) = \lim_{t\to\infty} P[X^*/t \le x | Y^* > t]$ or limits μ^{**} in (3.13) is indexed by Radon measures S on $[0, 1)$ satisfying condition (3.17).

Example 2 (Finite angular measure). If S is uniform on $[0,1)$, $S(\mathrm{d}w) = 2\,\mathrm{d}w$, then (3.17) is satisfied and we have

$$\mu^{**}\big([0,x]\times(y,\infty]\big) = \frac{x}{y(x+y)}.$$

Putting $y = 1$, we get the Pareto distribution $H^{**}(x) = 1 - (1+x)^{-1}$ for $x > 0$.

Example 3 (Infinite angular measure). The infinite measure $S(\mathrm{d}w) = (1-w)^{-1}\,\mathrm{d}w$ satisfies equation (3.17) and we have

$$\mu^{**}\big([0,x]\times(y,\infty]\big) = \frac{1}{y} + \frac{1}{x}\log\left(1 - \frac{x}{x+y}\right).$$

Putting $y = 1$ yields $H^{**}(x) = 1 - x^{-1}\log(1+x), x > 0$, and H^{**} is a continuously increasing probability distribution function. One way to get a class of infinite angular measures satisfying (3.17) is to take $S(\mathrm{d}w) = (1-w)^{-1}F(\mathrm{d}w)$ for probability measures $F(\cdot)$ on $[0,1)$.

4. Extending the CEV model to a multivariate extreme value model

The CEV model assumes the existence of a vague limit in a smaller subset of Euclidean space than that required by classical MEVT. Given a CEV model, when can it be extended to a MEVT model? If such an extension of the CEV model is possible, then X will also have a distribution in a DOA, so this will be assumed. The following is a sufficient condition for such an extension.

Proposition 4.1. *Suppose that (X,Y) satisfy Definition 1.1 and, in particular, (1.2)–(1.5). Assume that $X \in D(G_\lambda)$ for some $\lambda \in \mathbb{R}$ so that there exist functions $\chi(t) > 0, \phi(t) \in \mathbb{R}$ such that*

$$t\mathbf{P}\left(\frac{X-\phi(t)}{\chi(t)} > x\right) \to (1+\lambda x)^{-1/\lambda}, \qquad 1+\lambda x > 0.$$

If $\lim_{t\to\infty}\alpha(t)/\chi(t)$ exists and is finite and both $\lim_{t\to\infty}\beta(t), \lim_{t\to\infty}\phi(t)$ exist ($\le\infty$) and are equal, then (X,Y) is in the domain of attraction of a multivariate extreme value distribution on $\mathbb{E}^{(\lambda,\gamma)}$; that is, for a Radon measure $\mu_{X,Y}(\cdot)$ on $\mathbb{E}^{(\lambda,\gamma)}$,

$$t\mathbf{P}\left[\left(\frac{X-\phi(t)}{\chi(t)}, \frac{Y-b(t)}{a(t)}\right) \in \cdot\right] \xrightarrow{v} \mu_{X,Y}(\cdot) \qquad \text{in } \mathbb{M}_+\big(\mathbb{E}^{(\lambda,\gamma)}\big).$$

Proof. For $\lambda > 0$, the proof is a consequence of cases 1 and 2 of Theorem 2.1. The other cases can be proven similarly. □

We now discuss the extension of the CEV model to MEVT after first standardizing (X,Y) to (X^*,Y^*), which is regularly varying on $\mathbb{E}_\sqcap$. We consider extending regular variation on $\mathbb{E}_\sqcap$ to an *asymptotically tail equivalent* regular variation on $\mathbb{E}$, a notion we define next.

Definition 4.1 (Tail equivalence in multivariate regular variation [18]). *If* $\mathbf{X}$ *and* $\mathbf{Y}$ *are* $\mathbb{R}_+^d$*-valued random vectors, then* $\mathbf{X}$ *and* $\mathbf{Y}$ *are tail equivalent on a cone* $\mathfrak{C} \subset \overline{\mathbb{R}}_+^d$ *if there exists a scaling function* $b(t) \uparrow \infty$ *such that*

$$t\mathbf{P}[\mathbf{X}/b(t) \in \cdot] \xrightarrow{v} \nu(\cdot) \quad \textit{and} \quad t\mathbf{P}[\mathbf{Y}/b(t) \in \cdot] \xrightarrow{v} c\nu(\cdot)$$

in $M_+(\mathfrak{C})$ *for some* $c > 0$ *and non-null Radon measure* ν *on* $\mathfrak{C}$*. We write* $X \overset{te(\mathfrak{C})}{\sim} Y$.

Proposition 4.2. *Suppose that* (X^*, Y^*) *is standard regularly varying on* $\mathbb{E}_\sqcap$ *with limit measure* $\nu_\sqcap$ *and angular measure* $S_\sqcap$ *on* $[0, 1)$*. The following are equivalent*:

(1) $S_\sqcap$ *is finite on* $[0, 1)$;
(2) *there exists a random vector* $(X^\#, Y^\#)$ *defined on* $\mathbb{E} = [0, \infty]^2 \setminus \{\mathbf{0}\}$ *such that*

$$(X^\#, Y^\#) \overset{te(\mathbb{E}_\sqcap)}{\sim} (X^*, Y^*)$$

and $(X^\#, Y^\#)$ *is multivariate regularly varying on* $\mathbb{E}$ *with limit measure* ν *such that* $\nu|_{\mathbb{E}_\sqcap} = \nu_\sqcap$.

Proof. (1) $\Rightarrow$ (2): Define the polar coordinate transformation $(R, \Theta) = (X^* + Y^*, \frac{X^*}{X^*+Y^*})$. From Section 3.4 and (3.15), for $r > 0$ and Λ a Borel subset of $[0, 1)$, as $t \to \infty$,

$$t\mathbf{P}\left[\frac{R}{t} > r, \Theta \in \Lambda\right] \to r^{-1} S_\sqcap(\Lambda) = \nu_\sqcap\left\{(x, y) \in \mathbb{E}_\sqcap : x + y > r, \frac{x}{x+y} \in \Lambda\right\}.$$

Since $S_\sqcap$ is finite on $[0, 1)$, the distribution of Θ is finite on $[0, 1)$. Assume that $S_\sqcap[0, 1) = 1$ so that it is a probability measure and extend the measure $S_\sqcap$ to $[0, 1]$ by putting $S_\sqcap(\{1\}) = 0$. Define R_0 and Θ_0 to be independent. Θ_0 has distribution given by the extended $S_\sqcap$ on $[0, 1]$ and R_0 has the standard Pareto distribution. Define $(X^\#, Y^\#) = (R_0\Theta_0, R_0(1 - \Theta_0))$, so $(X^\#, Y^\#)$ is regularly varying on $\mathbb{E}$ with standard scaling and limit measure ν, where $\nu|_{\mathbb{E}_\sqcap} = \nu_\sqcap$.

(2) $\Rightarrow$ (1): Referring to (3.15), note that $S_\sqcap([0, 1)) = \nu_\sqcap\{(x, y) \in \mathbb{E}_\sqcap : x + y > 1\}$. Since $(X^\#, Y^\#)$ is regularly varying on $\mathbb{E}$, we have

$$t\mathbf{P}(X^\# + Y^\# > t) \to \nu\{(x, y) \in \mathbb{E}_\sqcap : x + y > 1\} < \infty.$$

However,

$$\nu\{(x, y) \in \mathbb{E}_\sqcap : x + y > 1\} = \nu_\sqcap\{(x, y) \in \mathbb{E}_\sqcap : x + y > 1\} = S_\sqcap([0, 1)).$$

Hence, $S_\sqcap$ is finite on $[0, 1)$. □

5. The CEV model and hidden regular variation

A vector (X^*, Y^*) whose distribution is standard bivariate regularly varying on $\mathbb{E}$ possesses *hidden regular variation* (HRV) [26] if there exists a Radon measure $\nu_0 \neq 0$ on $\mathbb{E}_0 = (0, \infty] \times$

$(0, \infty]$ and a non-decreasing function $a_0(t) \uparrow \infty$ with $t/a_0(t) \to \infty$, such that, in $\mathbb{M}_+(\mathbb{E}_0)$,

$$t\mathbf{P}\big(a_0^{-1}(t)(X^*, Y^*) \in \cdot\big) \xrightarrow{v} \nu_0(\cdot).$$

If hidden regular variation holds, then X^* and Y^* must be asymptotically independent,

$$\nu^{**}([\mathbf{0}, (x, y)]^c) = x^{-1} + y^{-1}, \qquad x, y > 0.$$

Built into the definition of HRV is regular variation on $\mathbb{E}$; our formulation of the CEV model, when it can be standardized, does not require regular variation on $\mathbb{E}$, but only on $\mathbb{E}_{\sqcap}$. Therefore, comparisons between HRV and the CEV model must be carefully posed.

Suppose that $(X, Y) \in D(G)$ for a bivariate extreme value distribution G. If X and Y are asymptotically dependent, then the CEV model holds with either X or Y as conditioning variable. The centering and scaling functions for CEV can be the same ones as for MEV. We can standardize $(X, Y) \mapsto (X^*, Y^*)$, so (X^*, Y^*) is standard regularly varying on $\mathbb{E}$ with limit measure ν^{**}, but asymptotic dependence implies that hidden regular variation cannot hold.

It is possible for HRV to hold without a CEV model being valid.

Example 4. Suppose that (X^*, Y^*) are random variables such that for $\alpha > 1$ and $x, y \geq 1$,

$$\mathbf{P}\big[(X^*, Y^*) \in ([\mathbf{0}, (x, y)])^c\big] = [x^{-1} + y^{-1} + (x^\alpha \wedge y^\alpha)^{-1}]/3.$$

Asymptotic independence and HRV hold for (X^*, Y^*) with $a_0(t) = t^{1/\alpha}$. The CEV model does not hold, whatever normalization we choose; if a limit holds, it is degenerate.

The following are comments on the relations between MEVT, the CEV model and HRV: MEVT is equivalent via standardization to regular variation on $\mathbb{E}$. The CEV model, if standardization is possible, is equivalent to regular variation on $\mathbb{E}_{\sqcap}$. Hidden regular variation requires standard regular variation on $\mathbb{E}$ and regular variation of lower order on $\mathbb{E}_0$. For a pair (X^*, Y^*) which is standard regularly varying on $\mathbb{E}$:

- asymptotic dependence of (X^*, Y^*) implies that the CEV model holds, but HRV does not; the requirement that $a_0(t)$ be of lower order than t fails;
- the presence of HRV does not imply that the CEV model holds;
- we conjecture that if CEV holds with asymptotic independence, then HRV must hold; this is evident in several examples, but we have no proof to turn this conjecture into fact.

More on HRV and generalizations to higher dimensions can be found in [21].

6. Examples

This section presents examples that illustrate how the CEV model differs from the usual multivariate extreme value model.

Example 5. This example emphasizes that different scaling and centering functions are required for different cones. We will consider a bivariate random vector which is multivariate regularly varying on $\mathbb{E}$ with asymptotic independence. We then show that it possesses hidden regular variation (see Section 5) and also CEV limits under different scalings. Let X, Y be i.i.d. *Pareto*(1) random variables. Let B be a Bernoulli random variable with $\mathbf{P}(B=0)=\mathbf{P}(B=1)=0.5$, U a Uniform(0, 1) random variable and suppose that X, Y, B, U are all independent. Define

$$\mathbf{Z}=(Z_1, Z_2)=B(UX, X^2)+(1-B)(Y^2, UY).$$

As $t\to\infty$, observe that the following hold:

(i) in $\mathbb{M}_+(\mathbb{E})$,

$$t\mathbf{P}\left[\frac{\mathbf{Z}}{t^2}\in([0,x]\times[0,y])^c\right]\to\frac{1}{2}\left[\frac{1}{\sqrt{x}}+\frac{1}{\sqrt{y}}\right],\qquad x\wedge y>0; \tag{6.1}$$

(ii) in $\mathbb{M}_+(\mathbb{E}_0)$,

$$t\mathbf{P}\left[\frac{\mathbf{Z}}{t}\in(x,\infty]\times(y,\infty]\right]\to\frac{1}{2}\left[\frac{1}{x}+\frac{1}{y}\right],\qquad x\wedge y>0; \tag{6.2}$$

(iii) in $\mathbb{M}_+(\mathbb{E}_\sqcap)$, the limit is not a product measure and we have

$$t\mathbf{P}\left[\left(\frac{Z_1}{t},\frac{Z_2}{t^2}\right)\in[0,x]\times(y,\infty]\right]\to\frac{1}{2}\left[\frac{1}{\sqrt{y}}-\frac{1}{2x}\right]_+,\qquad x\wedge y>0; \tag{6.3}$$

(iv) similarly, in $\mathbb{M}_+(\mathbb{E}_\sqsupset)$, the limit is still not a product measure and we have

$$t\mathbf{P}\left[\left(\frac{Z_1}{t^2},\frac{Z_2}{t}\right)\in(x,\infty]\times[0,y]\right]\to\frac{1}{2}\left[\frac{1}{\sqrt{x}}-\frac{1}{2y}\right]_+,\qquad x\wedge y>0. \tag{6.4}$$

This provides an example for the validity of Theorem 2.1. The example holds even if we ignore the random variable U, but then the distribution of $\mathbf{Z}$ concentrates on two parabolic lines restricted to $[0,\infty)^2$. Also, note that the limit measure for (i) concentrates on the lines through $\mathbf{0}$, the limit measure in (ii) concentrates on the lines through $\boldsymbol{\infty}$ and, in (iii) (and, similarly, (iv)), the limit measure does not concentrate on the boundaries. This final feature can also be observed in Example 6.

Example 6 (Example 1 continued). Recall Example 1. We had a bivariate joint density for (X, Y) and in the different cones, we have convergence with different normalizations (as $t\to\infty$):

$$\text{in } \mathbb{M}_+(\mathbb{E}): t\mathbf{P}\left(\left(\frac{X}{t},\frac{Y}{t}\right)\in([0,x]\times[0,y])^c\right)\to\frac{1}{x}+\frac{1}{y},\qquad x>0, y>0,$$

$$\text{in } \mathbb{M}_+(\mathbb{E}_0): t\mathbf{P}\left(\frac{X}{\sqrt{t}}>x,\frac{Y}{\sqrt{t}}>y\right)\to\frac{1}{x^2}+\frac{1}{y^2},\qquad x>0, y>0,$$

$$\text{in } \mathbb{M}_+(\mathbb{E}_\sqcap): t\mathbf{P}\left(\frac{X}{\sqrt{t}} \le x, \frac{Y}{t} > y\right) \to \frac{1}{y} - \frac{1}{y + x^2}, \qquad x \ge 0, y > 0,$$

$$\text{in } \mathbb{M}_+(\mathbb{E}_\sqsupset): t\mathbf{P}\left(\frac{X}{t} > x, \frac{Y}{\sqrt{t}} \le y\right) \to \frac{1}{x} - \frac{1}{x + y^2}, \qquad x > 0, y \ge 0.$$

Example 7. Suppose that (X, Y) has the following distribution generated by an Archimedean copula:

$$F(x, y) := \frac{(1 - 1/x)(1 - 1/y)}{(1 + 1/(xy))}, \qquad x, y \ge 1.$$

Clearly, X and Y are marginally *Pareto*(1) random variables and, for $x, y > 0$,

$$t\mathbf{P}\big[t^{-1}(X, Y) \in [\mathbf{0}, (x, y)]^c\big] = t\big(1 - F(tx, ty)\big) \to x^{-1} + y^{-1} \qquad (t \to \infty).$$

Hence, asymptotic independence holds and, for $x, y > 0$,

$$t\mathbf{P}[X > t^{1/3}x, Y > t^{1/3}y] \to \frac{1}{xy}\left(\frac{1}{x} + \frac{1}{y}\right) \qquad (t \to \infty),$$

which implies hidden regular variation. We also have the CEV model holding with a limit product measure since

$$t\mathbf{P}[X \le x, Y > ty] \to (1 - x^{-1})y^{-1} \qquad (t \to \infty).$$

Example 8. This example gives a class of limit distributions on $\mathbb{E}_\sqcap$ indexed by probability distributions on $[0, \infty]$. Suppose that R is a Pareto random variable on $[1, \infty)$ with parameter 1 and ξ is a random variable with distribution $G(\cdot)$ on $[0, \infty]$. Assume that ξ and R are independent and define $(X, Y) = (R\xi, R)$. Then, for $y > 0, x \ge 0$ and $ty > 1$,

$$\begin{aligned} t\mathbf{P}\left[\frac{X}{t} \le x, \frac{Y}{t} > y\right] &= t\mathbf{P}\left[\frac{R\xi}{t} \le x, \frac{R}{t} > y\right] = t\int_{ty}^{\infty} \mathbf{P}\left[\xi \le \frac{tx}{r}\right] r^{-2}\,\mathrm{d}r \\ &= \int_y^\infty \mathbf{P}\left[\xi \le \frac{x}{s}\right] s^{-2}\,\mathrm{d}s = \int_y^\infty G\left(\frac{x}{s}\right) s^{-2}\,\mathrm{d}s \\ &= \frac{1}{x}\int_0^{x/y} G(s)\,\mathrm{d}s \\ &= \mu\big([0, x] \times (y, \infty]\big). \end{aligned}$$

This can be expressed in polar coordinates. The angular measure $S(\cdot)$ on $\mathbb{E}_\sqcap$ is

$$S([0, \eta]) = \mu\left\{(u, v) : u + v > 1, \frac{u}{u + v} \le \xi\right\}, \qquad 0 \le \eta < 1.$$

Hence, we have

$$
\begin{aligned}
&t\mathbf{P}\left[\frac{X+Y}{t} > 1, \frac{X}{X+Y} \le \eta\right] \\
&\quad = t\mathbf{P}\left[\frac{R\xi + R}{t} > 1, \frac{R\xi}{R\xi + R} \le \eta\right] = t\mathbf{P}\left[\frac{R(1+\xi)}{t} > 1, \xi \le \frac{\eta}{1-\eta}\right] \\
&\quad = t\int_{0\le s\le \eta/(1-\eta)} \mathbf{P}\left[\frac{R}{t}(1+s) > 1\right] G(\mathrm{d}s) = t\int_{0\le s\le \eta/(1-\eta)} \left(\frac{t}{1+s} \vee 1\right)^{-1} G(\mathrm{d}s) \\
&\quad = \int_{0\le s\le \eta/(1-\eta)} (1+s)G(\mathrm{d}s)
\end{aligned}
$$

for $t > 1/(1-\eta)$. However, the left-hand side goes to $\mu\{(u,v): u+v > 1, \frac{y}{u+v} \le \xi\} = S([0,\eta])$ as $t \to \infty$ and, thus,

$$
S([0,\eta]) = \int_{0\le s\le \eta/(1-\eta)} (1+s)G(\mathrm{d}s), \qquad 0 \le \eta < 1.
$$

Hence, S is a finite angular measure if and only if G has first moment.

7. Proofs

In this section, we provide proofs of some of the results given in the previous sections.

7.1. Proof of Theorem 2.1

Assume that $\lambda > 0$, $\gamma > 0$; other cases can be dealt with similarly. From (2.1) and (2.2), respectively, we get

$$
t\mathbf{P}\left(\frac{Y-b(t)}{a(t)} > y\right) \to (1+\gamma y)^{-1/\gamma}, \qquad 1+\gamma y > 0, \tag{7.1}
$$

$$
t\mathbf{P}\left(\frac{X-\phi(t)}{\chi(t)} > x\right) \to (1+\lambda x)^{-1/\lambda}, \qquad 1+\lambda x > 0. \tag{7.2}
$$

Hence, for $(x,y) \in \mathbb{E}^{(\lambda)} \times \mathbb{E}^{(\gamma)}$, which are continuity points of the limit measures $\mu_{X,Y>}$ and $\mu_{Y,X>}$,

$$
\begin{aligned}
Q_t(x,y) &:= t\mathbf{P}\left[\left(\frac{X-\phi(t)}{\chi(t)}, \frac{Y-b(t)}{a(t)}\right) \in ([-\infty,x]\times[-\infty,y])^c\right] \\
&= t\mathbf{P}\left[\frac{X-\phi(t)}{\chi(t)} > x\right] + t\mathbf{P}\left[\frac{Y-b(t)}{a(t)} > y\right]
\end{aligned} \tag{7.3}
$$

$$
- t\mathbf{P}\left[\frac{X-\phi(t)}{\chi(t)} > x, \frac{Y-b(t)}{a(t)} > y\right]
$$
$$
= A_t(x) + B_t(y) + C_t(x, y) \qquad \text{(say)}.
$$

It suffices to show that $Q_t(x, y)$ has a limit and that the limit is non-degenerate in (x, y) (using a generalized version of [27], Lemma 6.1). As $t \to \infty$, we have the limits for $A_t(x)$ and $B_t(y)$ from (7.2) and (7.1). Clearly, $0 \le C_t(x, y) \le \min(A_t(x), B_t(y))$ and these inequalities also hold for any limit of Q_t.

From [11], Proposition 1, there exist functions $\psi_1(\cdot), \psi_2(\cdot), \psi_3(\cdot), \psi_4(\cdot)$ such that

$$
\lim_{t\to\infty} \frac{\alpha(tz)}{\alpha(t)} = \psi_1(z) = z^{\rho_1}, \qquad \lim_{t\to\infty} \frac{\beta(tz)-\beta(t)}{\alpha(t)} = \psi_2(z), \tag{7.4}
$$
$$
\lim_{t\to\infty} \frac{c(tz)}{c(t)} = \psi_3(z) = z^{\rho_2}, \qquad \lim_{t\to\infty} \frac{d(tz)-d(t)}{c(t)} = \psi_4(z) \tag{7.5}
$$

for $z > 0$ and ρ_1, ρ_2 real. Temporarily assume that ρ_1 and ρ_2 are positive. Either $\psi_2(z) = 0$, which implies that $\lim_{t\to\infty} \beta(t)/\alpha(t) = 0$ (from [1], Theorem 3.1.12(a,c)) or $\psi_2(z) = k(z^{\rho_1} - 1)/\rho_1$ for $k \neq 0$, which means that $\lim_{t\to\infty} \beta(t)/\alpha(t) = k/\rho_1$ ([6], Proposition B.2.2). Hence, allowing the constant k to be zero as well, we can write both cases as $\lim_{t\to\infty} \beta(t)/\alpha(t) = k_1/\rho_1$ for some $k_1 \in \mathbb{R}$. Similarly, we have $\lim_{t\to\infty} d(t)/c(t) = k_2/\rho_2$ for some $k_2 \in \mathbb{R}$.

Additionally, marginal DOA conditions for X, Y yield ($z > 0$, $w > 0$)

$$
\lim_{t\to\infty} \frac{b(tz)-b(t)}{a(t)} = \frac{z^\gamma - 1}{\gamma}, \qquad \lim_{t\to\infty} \frac{\phi(tw)-\phi(t)}{\chi(t)} = \frac{w^\lambda - 1}{\lambda}, \tag{7.6}
$$

which imply

$$
\lim_{t\to\infty} \frac{a(tz)}{a(t)} = z^\gamma, \qquad \lim_{t\to\infty} \frac{\chi(tw)}{\chi(t)} = w^\lambda. \tag{7.7}
$$

Observe that

$$
\begin{aligned}
C_t(x, y) &= t\mathbf{P}\left[\frac{X-\phi(t)}{\chi(t)} > x, \frac{Y-b(t)}{a(t)} > y\right] \\
&= t\mathbf{P}\left[\frac{X-\beta(t)}{\alpha(t)} > \left(x + \frac{\phi(t)}{\chi(t)}\right)\frac{\chi(t)}{\alpha(t)} - \frac{\beta(t)}{\alpha(t)}, \frac{Y-b(t)}{a(t)} > y\right]
\end{aligned} \tag{7.8}
$$

and also

$$
C_t(x, y) = t\mathbf{P}\left[\frac{X-\phi(t)}{\chi(t)} > x, \frac{Y-d(t)}{c(t)} > \left(y + \frac{b(t)}{a(t)}\right)\frac{a(t)}{c(t)} - \frac{d(t)}{c(t)}\right]. \tag{7.9}
$$

From [6], Proposition B.2.2, we have that

$$
b(t)/a(t) \to 1/\gamma \quad \text{and} \quad \phi(t)/\chi(t) \to 1/\lambda. \tag{7.10}
$$

We analyze $C_t(x,y)$ for the different cases. First, we will show that at least one of the limits $\lim_{t\to\infty}\frac{\chi(t)}{\alpha(t)}$ and $\lim_{t\to\infty}\frac{a(t)}{c(t)}$ must exist. Suppose both do not exist. We have, for $(x,y)\in\mathbb{E}^{(\lambda)}\times\mathbb{E}^{(\gamma)}$, which are continuity points of the limit measures $\mu_{X,Y>}$ and $\mu_{Y,X>}$,

$$t\mathbf{P}\left[\frac{X-\beta(t)}{\alpha(t)}>x,\frac{Y-b(t)}{a(t)}>y\right]\to\mu_{X,Y>}\big((x,\infty]\times(y,\infty]\big), \tag{7.11}$$

$$t\mathbf{P}\left[\frac{X-\phi(t)}{\chi(t)}>x,\frac{Y-d(t)}{c(t)}>y\right]\to\mu_{Y,X>}\big((x,\infty]\times(y,\infty]\big). \tag{7.12}$$

Now, (7.11) implies that

$$t\mathbf{P}\left[\frac{X-\phi(t)}{\chi(t)}\frac{\chi(t)}{\alpha(t)}+\frac{\phi(t)-\beta(t)}{\alpha(t)}>x,\frac{Y-d(t)}{c(t)}\frac{c(t)}{a(t)}+\frac{d(t)-b(t)}{a(t)}>y\right]$$
$$\to\mu_{X,Y>}\big((x,\infty]\times(y,\infty]\big),$$

which is equivalent to

$$t\mathbf{P}\left[\frac{X-\phi(t)}{\chi(t)}>\frac{\alpha(t)}{\chi(t)}\left(x-\frac{\phi(t)-\beta(t)}{\alpha(t)}\right),\frac{Y-d(t)}{c(t)}>\frac{a(t)}{c(t)}\left(y-\frac{d(t)-b(t)}{a(t)}\right)\right]$$
$$\to\mu_{X,Y>}\big((x,\infty]\times(y,\infty]\big).$$

From (7.12), we also have that the left-hand side of the previous line has a limit

$$t\mathbf{P}\left[\frac{X-\phi(t)}{\chi(t)}>\frac{\alpha(t)}{\chi(t)}\left(x-\frac{\phi(t)-\beta(t)}{\alpha(t)}\right),\frac{Y-d(t)}{c(t)}>\frac{a(t)}{c(t)}\left(y-\frac{d(t)-b(t)}{a(t)}\right)\right]$$
$$\to\mu_{Y,X>}\big((f(x),\infty]\times(g(y),\infty]\big)$$

for some $(f(x),g(y))$, assumed to be a continuity point of the limit $\mu_{Y,X>}$, if and only if, as $t\to\infty$, the following two limits hold:

$$\frac{\alpha(t)}{\chi(t)}\left(x-\frac{\phi(t)-\beta(t)}{\alpha(t)}\right)\to f(x), \tag{7.13}$$

$$\frac{a(t)}{c(t)}\left(y-\frac{d(t)-b(t)}{a(t)}\right)\to g(y). \tag{7.14}$$

For $\mu_{Y,X>}$ to be non-degenerate, f and g should be non-constant and we should also have $\mu_{X,Y>}((x,\infty]\times(y,\infty])=\mu_{Y,X>}((f(x),\infty]\times(g(y),\infty])$. Considering (7.13) and (7.14), we can see that the limit as $t\to\infty$ exists if and only if $\lim_{t\to\infty}a(t)/c(t)$ and $\lim_{t\to\infty}\chi(t)/\alpha(t)$ exists.

We conclude that $\lim_{t\to\infty}\chi(t)/\alpha(t)\in[0,\infty]$ and consider the following cases:

- *Case* 1: $\lim_{t\to\infty}\chi(t)/\alpha(t)=\infty$. Consider (7.8) and note that

$$\left(x+\frac{\phi(t)}{\chi(t)}\right)\frac{\chi(t)}{\alpha(t)}-\frac{\beta(t)}{\alpha(t)}\to\left(x+\frac{1}{\lambda}\right)\times\infty-\frac{k_1}{\rho_1}=\infty,$$

which entails that $\lim_{t\to\infty} C_t(x, y) = \mu_{X,Y>}(\{\infty\} \times (y, \infty]) = 0$. Hence,

$$\lim_{t\to\infty} Q_t(x, y) = (1 + \lambda x)^{-1/\lambda} + (1 + \gamma y)^{-1/\gamma}.$$

- *Case 2*: $\lim_{t\to\infty} \chi(t)/\alpha(t) = M \in (0, \infty)$. From (7.8), we have

$$\left(x + \frac{\phi(t)}{\chi(t)}\right)\frac{\chi(t)}{\alpha(t)} - \frac{\beta(t)}{\alpha(t)} \to \left(x + \frac{1}{\lambda}\right) \times M - \frac{k_1}{\rho_1} = f(x) \qquad \text{(say).}$$

  Therefore,

$$\lim_{t\to\infty} C_t(x, y) = \mu_{X,Y>}\big((f(x), \infty] \times (y, \infty]\big) \le (1 + \lambda y)^{-1/\lambda}$$

  with strict inequality holding for some x because of the non-degeneracy condition (1.3) for $\mu_{X,Y>}$. Hence,

$$\lim_{t\to\infty} Q_t(x, y) = (1 + \lambda x)^{-1/\lambda} + (1 + \gamma y)^{-1/\gamma} - \mu_{X,Y>}\big((f(x), \infty] \times (y, \infty]\big).$$

- *Case* 3: $\lim_{t\to\infty} \chi(t)/\alpha(t) = 0$. In this case, (7.8) leads to a degenerate limit in x for $C_t(x, y)$ and putting $M_1 = k/\rho_1$, we get

$$\lim_{t\to\infty} C_t(x, y) = \mu_{X,Y>}\big((M_1, \infty] \times (y, \infty]\big) =: f_1(y) \le (1 + \gamma y)^{-1/\gamma}.$$

  So, consider (7.9).

  (1) If $\lim_{t\to\infty} a(t)/c(t)$ exists in $(0, \infty]$, then we can use a similar technique as in case 1 or 2 to obtain a non-degenerate limit for $Q_t(x, y)$.

  (2) If $\lim_{t\to\infty} a(t)/c(t) = 0$, then for some $M_2 \in \mathbb{R}$,

$$\lim_{t\to\infty} C_t(x, y) = \mu_{Y,X>}\big((x, \infty] \times (M_2, \infty]\big) =: f_2(x) \le (1 + \lambda x)^{-1/\lambda}.$$

  Therefore, we have, for any $(x, y) \in \mathbb{E}^{(\lambda)} \times \mathbb{E}^{(\gamma)}$ which are continuity points of the limit measures $\mu_{X,Y>}$ and $\mu_{Y,X>}$,

$$f_1(y) = \mu_{X,Y>}\big((M_1, \infty] \times (y, \infty]\big) = \mu_{Y,X>}\big((x, \infty] \times (M_2, \infty]\big) = f_2(x).$$

  It is now easy to check that for any $(x, y) \in \mathbb{E}^{(\lambda)} \times \mathbb{E}^{(\gamma)}$ which are continuity points of the limit measures $\mu_{X,Y>}$ and $\mu_{Y,X>}$, we have $f_1(y) = f_2(x) = 0$. Hence, $C_t(x, y) \to 0$ and thus $Q_t(x, y)$ has a non-degenerate limit.

This proves the result. For general $\rho_1, \rho_2 \in \mathbb{R}$, we can follow the same steps to get to the result by considering cases when ρ_i is greater than, less than or equal to zero, for each $i = 1, 2$.

7.2. Proof of the summary following Theorem 3.2

(1) This part has been dealt with in [11], Section 2.4.

(2) First, simplify the problem. For (x, y), a continuity point of $\mu(\cdot)$,

$$t\mathbf{P}\left[\frac{\lambda(X^*)-\beta(t)}{\alpha(t)} \le x, \frac{b(Y^*)-b(t)}{a(t)} > y\right] \to \tilde{\mu}([-\infty, x] \times (y, \infty]) \qquad (t \to \infty)$$

is equivalent, as $t \to \infty$, to

$$\begin{aligned} t\mathbf{P}\left(\frac{\lambda(X^*)-\beta(t)}{\alpha(t)} \le x, \frac{Y^*}{t} > y\right) &\to \tilde{\mu}\big([-\infty, x] \times (h(y), \infty]\big) \\ &=: \mu^*\big([-\infty, x] \times (y, \infty]\big), \end{aligned} \tag{7.15}$$

where

$$h(y) = \begin{cases} (1+\gamma y)^{1/\gamma}, & \gamma \ne 0, \\ \mathrm{e}^y, & \gamma = 0. \end{cases} \tag{7.16}$$

Hence, (3.11) is equivalent to

$$t\mathbf{P}\left[\left(\frac{\lambda(X^*)-\beta(t)}{\alpha(t)}, \frac{Y^*}{t}\right) \in \cdot\right] \xrightarrow{v} \mu^*(\cdot)$$

and μ^* is a non-null Radon measure on $[-\infty, \infty] \times \overline{\mathbb{E}}^{(\gamma)}$ satisfying the conditional non-degeneracy conditions. Hence, our proof will show the existence of $\lambda(\cdot)$ satisfying (7.15). Now, note that (1.8) implies that $\alpha(\cdot) \in RV_\rho$ for some $\rho \in \mathbb{R}$ and $\psi_1(x) = x^\rho$ ([29], page 14). The function $\psi_2(\cdot)$ may be identically equal to 0 or

$$\psi_2(x) = \begin{cases} k(x^\rho - 1)/\rho & \text{if } \rho \ne 0, x > 0, \\ k \log x & \text{if } \rho = 0, x > 0 \end{cases} \tag{7.17}$$

for $k \ne 0$ ([6], page 373). We have assumed that $(\psi_1, \psi_2) \ne (1, 0)$. We will consider three cases: $\rho > 0$, $\rho = 0$, $\rho < 0$.

Case 1: $\rho > 0$. First, suppose that $\psi_2 \equiv 0$. Since $\alpha(\cdot) \in RV_\rho$, there exists $\tilde{\alpha}(\cdot) \in RV_\rho$ which is ultimately differentiable and strictly increasing and $\alpha \sim \tilde{\alpha}$ ([6], page 366). Thus, $\tilde{\alpha}^{\leftarrow}$ exists. Additionally, from [1], Theorem 3.1.12(a), we have that $\beta(t)/\alpha(t) \to 0$. Hence, for $x > 0$, as $t \to \infty$, we have

$$\frac{\tilde{\alpha}(tx)+\beta(t)}{\alpha(t)} = \frac{\tilde{\alpha}(tx)}{\tilde{\alpha}(t)} \cdot \frac{\tilde{\alpha}(t)}{\alpha(t)} + \frac{\beta(t)}{\alpha(t)} \to x^\rho$$

and inverting, we get, for $z > 0$,

$$\tilde{\alpha}^{\leftarrow}\big(\alpha(t)z + \beta(t)\big)/t \to z^{1/\rho} \qquad (t \to \infty).$$

Thus, we have

$$tP\left[\frac{\tilde{\alpha}(X^*)-\beta(t)}{\alpha(t)}\le x, \frac{Y^*}{t}>y\right] = tP\left[\frac{X^*}{t}\le\frac{\tilde{\alpha}^{\leftarrow}(\alpha(t)x+\beta(t))}{t}, \frac{Y^*}{t}>y\right] \\ \to \mu^{**}\big([0,x^{1/\rho}]\times(y,\infty]\big).$$

Set $\lambda(\cdot)=\tilde{\alpha}(\cdot)$ and this defines $\tilde{\mu}$.

Next, suppose that $\psi_2\neq 0$. Therefore,

$$\psi_2(x)=\lim_{t\to\infty}\big(\beta(tx)-\beta(t)\big)/\alpha(t)=k(x^\rho-1)/\rho,$$

that is, $\beta(\cdot)\in RV_\rho$ and $k>0$. There exists $\tilde{\beta}$ which is ultimately differentiable, strictly increasing and such that $\tilde{\beta}\sim\beta$ ([6], page 366). Thus, $\tilde{\beta}^{\leftarrow}$ exists. We then have, for $x>0$, as $t\to\infty$,

$$\begin{aligned}\frac{\tilde{\beta}(tx)-\beta(t)}{\alpha(t)} &= \frac{\tilde{\beta}(tx)-\beta(tx)}{\alpha(t)}+\frac{\beta(tx)-\beta(t)}{\alpha(t)}\\ &= \frac{\tilde{\beta}(tx)-\beta(tx)}{\beta(tx)}\frac{\beta(tx)}{\alpha(tx)}\frac{\alpha(tx)}{\alpha(t)}+\frac{\beta(tx)-\beta(t)}{\alpha(t)}\\ &\to (1-1)\cdot x^\rho/\rho+k(x^\rho-1)/\rho=k(x^\rho-1)/\rho.\end{aligned}$$

Inverting, we get, as $t\to\infty$,

$$\tilde{\beta}^{\leftarrow}\big(\alpha(t)x+\beta(t)\big)/t\to(1+\rho x/k)^{1/\rho}.$$

Thus, we have

$$tP\left[\frac{\tilde{\beta}(X^*)-\beta(t)}{\alpha(t)}\le x, \frac{Y^*}{t}>y\right] = tP\left[\frac{X^*}{t}\le\frac{\tilde{\beta}^{\leftarrow}(\alpha(t)x+\beta(t))}{t}, \frac{Y^*}{t}>y\right] \\ \to \mu^{**}\left(\left[0,\left(1+\frac{\rho x}{k}\right)^{1/\rho}\right]\times(y,\infty]\right).$$

Here, we can set $\lambda(\cdot)=\tilde{\beta}(\cdot)$ and this defines $\tilde{\mu}$.

Case 2: $\rho=0$. We have $\psi_1(x)=1$, $\psi_2(x)=k\log x$ for $x>0$ and some $k\in\mathbb{R}$. By assumption, $(\psi_1,\psi_2)\neq(1,0)$ and hence $k\neq 0$. First, assume that $k>0$, which means that $\beta\in\Pi_+(\alpha)$. There exists $\tilde{\beta}(\cdot)$ which is continuous, strictly increasing and $\beta-\tilde{\beta}=\mathrm{o}(\alpha)$ ([8], page 1031). If $\beta(\infty)=\tilde{\beta}(\infty)=\infty$, then, for $x>0$,

$$\frac{\tilde{\beta}(tx)-\beta(t)}{\alpha(t)}=\frac{\tilde{\beta}(tx)-\beta(tx)}{\alpha(tx)}\frac{\alpha(tx)}{\alpha(t)}+\frac{\beta(tx)-\beta(t)}{\alpha(t)}\to 0+k\log x$$

and, inverting, we get for $z\in\mathbb{R}$, as $t\to\infty$, $\tilde{\beta}^{\leftarrow}(\alpha(t)z+\beta(t))/t\to\exp\{z/k\}$. Thus, we have

$$tP\left(\frac{\tilde{\beta}(X^*)-\beta(t)}{\alpha(t)}\le x, \frac{Y^*}{t}>y\right) = tP\left(\frac{X^*}{t}\le\frac{\tilde{\beta}^{\leftarrow}(\alpha(t)x+\beta(t))}{t}, \frac{Y^*}{t}>y\right) \\ \to \mu\big([0,\mathrm{e}^{k/x}]\times(y,\infty]\big).$$

If $\beta(\infty) = \tilde{\beta}(\infty) = B < \infty$, define

$$\beta^*(t) = \frac{1}{B - \tilde{\beta}(t)}, \qquad \alpha^*(t) = \frac{\alpha(t)}{(B - \tilde{\beta}(t))^2}$$

and we have that $\beta^* \in \Pi_+(\alpha^*)$, $\beta^*(t) \to \infty$ and $(B - \tilde{\beta}(t))/\alpha(t) \to \infty$ ([10], page 25). Hence, we have reduced the problem to the previous case, which implies that

$$t\mathbf{P}\left(\frac{\beta^*(X^*) - \beta^*(t)}{\alpha^*(t)} \le x, \frac{Y^*}{t} > y\right) \to \mu\big([0, \mathrm{e}^{k/x}] \times (y, \infty]\big)$$

or, equivalently,

$$t\mathbf{P}\left(\frac{\tilde{\beta}(X^*) - \tilde{\beta}(t)}{\alpha(t)} \le \frac{x}{1 + \alpha(t)x/(B - \tilde{\beta}(t))}, \frac{Y^*}{t} > y\right) \to \mu\big([0, \mathrm{e}^{k/x}] \times (y, \infty]\big),$$

and since $B - \tilde{\beta}(t)/\alpha(t) \to \infty$ implies $\alpha(t)/B - \tilde{\beta}(t) \to 0$, we can write

$$t\mathbf{P}\left(\frac{\tilde{\beta}(X^*) - \tilde{\beta}(t)}{\alpha(t)} \le x, \frac{Y^*}{t} > y\right) \to \mu\big([0, \mathrm{e}^{k/x}] \times (y, \infty]\big),$$

which implies, since $\beta - \tilde{\beta} = \mathrm{o}(\alpha)$, that

$$t\mathbf{P}\left(\frac{\tilde{\beta}(X^*) - \beta(t)}{\alpha(t)} \le x, \frac{Y^*}{t} > y\right) \to \mu\big([0, \mathrm{e}^{k/x}] \times (y, \infty]\big).$$

We have thus produced the required transformation $\lambda(\cdot) = \tilde{\beta}(\cdot)$.

The case for which $k < 0$, that is, $\beta \in \Pi_-(\alpha)$, can be proven similarly.

Case 3: $\rho < 0$. This case is similar to the case for $\rho > 0$ and is therefore omitted.

Acknowledgments

This research was partially supported by ARO Contract W911NF-07-1-0078 at Cornell University. We wish to thank the conscientious referees for their helpful suggestions.

References

[1] Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1987). *Regular Variation*. Cambridge: Cambridge University Press. MR0898871

[2] Das, B. (2009). The conditional extreme value model and related topics. Ph.D. thesis, Cornell University, Ithaca, NY.

[3] Das, B. and Resnick, S.I. (2009). Detecting a conditional extreme value model. *Extremes* DOI: 10.1007/s10687-009-0097-3. Available at http://arxiv.org/abs/0902.2996.

[4] Davydov, Y., Molchanov, I. and Zuyev, S. (2007). Stable distributions and harmonic analysis on convex cones. *C. R. Math. Acad. Sci. Paris* **344** 321–326. MR2308120

[5] de Haan, L. (1978). A characterization of multidimensional extreme-value distributions. *Sankhyā Ser. A* **40** 85–88. MR0545467

[6] de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction*. New York: Springer. MR2234156

[7] de Haan, L. and Resnick, S.I. (1977). Limit theory for multivariate sample extremes. *Z. Wahrsch. Verw. Gebiete* **40** 317–337. MR0478290

[8] de Haan, L. and Resnick, S.I. (1979). Conjugate π-variation and process inversion. *Ann. Probab.* **7** 1028–1035. MR0548896

[9] Fougères, A. and Soulier, P. (2009). Estimation of conditional laws given an extreme component. Available at http://arXiv.org/0806.2426v2.

[10] Geluk, J.L. and de Haan, L. (1987). *Regular Variation, Extensions and Tauberian Theorems*. *CWI Tract* **40**. Amsterdam: Stichting Mathematisch Centrum, Centrum voor Wiskunde en Informatica.

[11] Heffernan, J.E. and Resnick, S.I. (2007). Limit laws for random vectors with an extreme component. *Ann. Appl. Probab.* **17** 537–571. MR2308335

[12] Heffernan, J.E. and Tawn, J.A. (2004). A conditional approach for multivariate extreme values (with discussion). *J. R. Stat. Soc. Ser. B Stat. Methodol.* **66** 497–546. MR2088289

[13] Kallenberg, O. (1983). *Random Measures*, 3rd edition. Berlin: Akademie.

[14] Klüppelberg, C. and Resnick, S.I. (2008). The Pareto copula, aggregation of risks and the emperor's socks. *J. Appl. Probab.* **45** 67–84. MR2409311

[15] Ledford, A.W. and Tawn, J.A. (1996). Statistics for near independence in multivariate extreme values. *Biometrika* **83** 169–187. MR1399163

[16] Ledford, A.W. and Tawn, J.A. (1997). Modelling dependence within joint tail regions. *J. Roy. Statist. Soc. Ser. B* **59** 475–499. MR1440592

[17] López-Oliveros, L. and Resnick, S.I. (2009). Extremal dependence analysis of network sessions. *Extremes* DOI: 10.1007/s10687-009-0096-4. Available at http://arxiv.org/pdf/0905.1983v1.

[18] Maulik, K. and Resnick, S.I. (2005). Characterizations and examples of hidden regular variation. *Extremes* **7** 31–67. MR2201191

[19] Mikosch, T. (2005). How to model multivariate extremes if one must? *Statist. Neerlandica* **59** 324–338. MR2189776

[20] Mikosch, T. (2006). Copulas: Tales and facts. *Extremes* **9** 3–20. MR2327880

[21] Mitra, A. and Resnick, S.I. (2010). Hidden regular variation: Detection and estimation. Available at http://people.orie.cornell.edu/~sid.

[22] Neveu, J. (1977). Processus ponctuels. In *École d'Été de Probabilités de Saint-Flour, VI–1976*. *Lecture Notes in Math.* **598** 249–445. Berlin: Springer. MR0474493

[23] Pickands, J. (1975). Statistical inference using extreme order statistics. *Ann. Statist.* **3** 119–131. MR0423667

[24] Pickands, J. (1981). Multivariate extreme value distributions. In *43rd Sess. Int. Statist. Inst. Buenos Aires* 859–878. MR0820979

[25] Ramos, A. and Ledford, A.W. (2009). A new class of models for bivariate joint tails. *J. Roy. Statist. Soc. Ser. B* **71**.

[26] Resnick, S.I. (2002). Hidden regular variation, second order regular variation and asymptotic independence. *Extremes* **5** 303–336. MR2002121

[27] Resnick, S.I. (2007). *Heavy Tail Phenomena: Probabilistic and Statistical Modeling*. New York: Springer. MR2271424

[28] Resnick, S.I. (2008). Multivariate regular variation on cones: Application to extreme values, hidden regular variation and conditioned limit laws. *Stochastics* **80** 269–298. MR2402168

[29] Resnick, S.I. (2008). *Extreme Values, Regular Variation and Point Processes*. New York: Springer. MR0900810

[30] Resnick, S.I. and Zeber, D. (2010). Foundations of conditional extreme value theory. To appear.

[31] Schlather, M. (2001). Examples for the coefficient of tail dependence and the domain of attraction of a bivariate extreme value distribution. *Statist. Probab. Lett.* **53** 325–329. MR1841635

[32] Seneta, E. (1976). *Regularly Varying Functions*. *Lecture Notes in Math.* **508**. New York: Springer. MR0453936

Received May 2008 and revised January 2010