

# LIVING ON THE MULTI-DIMENSIONAL EDGE: SEEKING HIDDEN RISKS USING REGULAR VARIATION

BY BIKRAMJIT DAS\*, ABHIMANYU MITRA<sup>†</sup> AND SIDNEY RESNICK<sup>†</sup>

*ETH Zurich, Cornell University and Cornell University*

Multivariate regular variation plays a role assessing tail risk in diverse applications such as finance, telecommunications, insurance and environmental science. The classical theory, being based on an asymptotic model, sometimes leads to inaccurate and useless estimates of probabilities of joint tail regions. This problem can be partly ameliorated by using *hidden regular variation* [Resnick, 2002, Mitra and Resnick, 2010]. We offer a more flexible definition of hidden regular variation that provides improved risk estimates for a larger class of risk tail regions.

**1. Introduction.** Daily we observe environmental, technological and financial phenomena possessing inherent risks. There are financial risks from large investment losses; environmental risks from health hazards resulting from high concentrations of atmospheric pollutants; hydrological risks from river floods. Risk analysis requires estimation of tail probabilities that provide measures of such risks. The mathematical framework of multivariate regular variation provides tools to compute tail probabilities associated with such risks; see Resnick [2007], Joe and Li [2010], Cai et al. [2011]. These tools have limitations which we begin to address in this paper.

Consider a non-negative random vector  $\mathbf{Z} = (Z^1, Z^2, \dots, Z^d)$  called a *risk vector*. The distribution of  $\mathbf{Z}$  has multivariate regular variation if there exist a function  $b(t) \uparrow \infty$  and a non-negative non-degenerate Radon measure  $\mu(\cdot)$  on  $\mathbb{E} = [0, \infty]^d \setminus \{(0, 0, \dots, 0)\}$  such that as  $t \rightarrow \infty$ ,

$$(1) \quad tP \left[ \frac{\mathbf{Z}}{b(t)} \in \cdot \right] \xrightarrow{v} \mu(\cdot),$$

where  $\xrightarrow{v}$  denotes vague convergence in  $\mathbb{M}_+(\mathbb{E})$ , the set of all Radon measures on  $\mathbb{E}$  [Resnick, 2007, page 172]. Note that (1) effectively assumes tail

---

\*B. Das was partially supported by the IRTG, Zurich-Berlin on *Stochastic Models for Complex Processes*.

<sup>†</sup>S. Resnick and A. Mitra were partially supported by ARO Contract W911NF-10-1-0289 at Cornell University.

*AMS 2000 subject classifications:* Primary 60F99, 62G32; secondary 60G70

*Keywords and phrases:* Regular variation, spectral measure, asymptotic independence, risk sets, vague convergence

equivalence of the marginal components [Resnick, 2007, Section 6.5.6], so while (1) is valuable as a theoretical foundation it must be modified for applications.

The asymptotic relation (1) allows the limit measure  $\mu(\cdot)$  to be used for approximating tail probabilities. For example, approximation of the probability of the event  $\{Z^i > x^i \text{ for some } i\}$  for large thresholds  $x^i, i = 1, 2, \dots, d$  requires the ability to compute  $\mu(\{(z^1, \dots, z^d) \in \mathbb{E} : z^i > w^i \text{ for some } i\})$  for  $w^i > 0, i = 1, 2, \dots, d$ . Such approximations of tail probabilities are sensitive to degeneracies in the limit measure  $\mu(\cdot)$ . For example, when *asymptotic independence* is present as in Gaussian copula models, the limit measure  $\mu(\cdot)$  in (1) concentrates on the coordinate axes  $\mathbb{L}_i := \{\mathbf{x} \in \mathbb{R}^d : x^j = 0 \ \forall j \neq i\}, i = 1, \dots, d$ , and  $\mu(\{(z^1, z^2, \dots, z^d) \in \mathbb{E} : z^i > w^1, z^j > w^2\}) = 0$  for any  $1 \leq i < j \leq d$  and  $w^1, w^2 > 0$ . Consequently, we would approximate the joint tail probability

$$P(Z^i > x^1, Z^j > x^2) \approx 0$$

for large thresholds  $x^1, x^2$  and conclude risk contagion is absent. This conclusion may be naive and hence the concept of *hidden regular variation* (HRV) was introduced [Resnick, 2002] which offered a refinement of this approximation; see Maulik and Resnick [2005], Heffernan and Resnick [2005], Mitra and Resnick [2010] and the seminal concept of *coefficient of tail dependence* in Ledford and Tawn [1996, 1998].

The definition of hidden regular variation offers some strengths but also has weaknesses. The existing definition provides insight only in the presence of a restricted class of degeneracies in the limit measure  $\mu(\cdot)$  in (1); namely when  $\mu(\cdot)$  concentrates either on the coordinate axes, or the coordinate planes or similar coordinate hyperplanes in higher dimensions. However, other degeneracies in  $\mu(\cdot)$  are possible; for example,  $\mu(\cdot)$  may be concentrated on the diagonal  $\{(z^1, z^2, \dots, z^d) \in \mathbb{E} : z^1 = z^2 = \dots = z^d\}$ , a condition called asymptotic full dependence. To deal with such degeneracies and situations where  $\mu$  may place zero mass on large portions of the state space, we define in Section 3 hidden regular variation on cones. For us, a cone  $\mathbb{C}$  in  $\mathbb{R}^d$  is a set  $\mathbb{C} \subset \mathbb{R}^d$  satisfying  $\mathbf{x} \in \mathbb{C}$  implies  $t\mathbf{x} \in \mathbb{C}$  for  $t > 0$ .

In practice, different risk assessment problems require calculating tail probabilities for different kinds of events. Hidden regular variation, as previously defined, may be a natural choice for some calculations but not for others. For example, suppose  $(Z^1, Z^2) \in \mathbb{R}_+^2$  is a risk vector and we must calculate risk probabilities of the form  $P(|Z^1 - Z^2| > w)$  for large thresholds  $w > 0$ . If we use multivariate regular variation when the limit measure  $\mu(\cdot)$  in (1) is concentrated on the diagonal  $\{(z^1, z^2) \in \mathbb{E} : z^1 = z^2\}$ , the tail probability  $P(|Z^1 - Z^2| > w)$  must be approximated as zero for large thresholds

$w > 0$ . The existing notion of hidden regular variation designed to help when  $\mu$  concentrates on the axes cannot offer a refinement in this case. Example 5.2 of Section 5 illustrates how a more general theory overcomes this difficulty. This along with other examples in Section 5 emphasize the need for the theory of hidden regular variation on general cones.

The conditional extreme value (CEV) model [Heffernan and Tawn, 2004, Heffernan and Resnick, 2007, Das and Resnick, 2011] provides one alternative approach to multivariate extreme value modeling. In standard form, the CEV model can also be formulated as regular variation on a particular cone in  $\mathbb{E}$  and this is discussed in Section 4. Also in Section 4, we consider non-standard regular variation from the point of view of regular variation on a sequence of cones. Non-standard regular variation is essential in practice since in applications we cannot assume tail-equivalence of all marginals in a multivariate model as is done in (1).

Section 6 discusses how to fit our model of HRV on cones to data as well as estimation techniques of tail probabilities using our model. We have adapted ideas previously used in multivariate heavy tail analysis and this discussion is not comprehensive, merely a feasibility display. In particular we have not performed data analyses. We close our discussion with concluding remarks in Section 7 and give some deferred results and proofs in Section 8.

Our definition of HRV relies on a notion of convergence of measures called  $\mathbb{M}^*$ -convergence that is similar to  $\mathbb{M}_0$  convergence of Hult and Lindskog [2006]. This  $\mathbb{M}^*$ -convergence is developed in Section 2 where we also discuss reasons to abandon the standard practice of defining regular variation through vague convergence on a compactification of  $\mathbb{R}^d$ .

1.1. *Notation.* We briefly discuss some frequently used notation and concepts.

1.1.1. *Vectors, norms and topology.* Bold letters are used to denote vectors, with capital letters reserved for random vectors and small letters for non-random vectors, e.g.,  $\mathbf{x} = (x^1, x^2, \dots, x^d) \in \mathbb{R}^d$ . We also denote

$$\mathbf{0} = (0, 0, \dots, 0), \quad \mathbf{1} = (1, 1, \dots, 1), \quad \boldsymbol{\infty} = (\infty, \infty, \dots, \infty).$$

Operations on and between vectors are understood componentwise. For example, for vectors  $\mathbf{x}, \mathbf{z}$ ,

$$\mathbf{x} \leq \mathbf{z} \quad \text{means} \quad x^i \leq z^i, \quad i = 1, \dots, d.$$

For a set  $A \subset [0, \infty)^d$  and  $\mathbf{x} \in A$ , use  $[\mathbf{0}, \mathbf{x}]^c$  to mean  $[\mathbf{0}, \mathbf{x}]^c = A \setminus [\mathbf{0}, \mathbf{x}] = \{\mathbf{y} \in A : \bigvee_{i=1}^d y^i / x^i > 1\}$ . When we use the notation  $[\mathbf{0}, \mathbf{x}]^c$ , the set  $A$  should be clear from the context.

For the  $i$ -th largest component of  $\mathbf{x}$ , we write  $x^{(i)}$ , that is,  $x^{(1)} \geq x^{(2)} \geq \dots \geq x^{(d)}$ . Thus a superscripted number  $i$  denotes the  $i$ -th component of a vector, whereas a superscripted  $(i)$  denotes the  $i$ -th largest component in the vector.

Operations with  $\infty$  are understood using the conventions:

$$\begin{aligned} \infty + \infty = \infty, \quad \infty - \infty = 0, & \quad \text{for } x \in \mathbb{R}, \infty + x = \infty - x = \infty, \\ 0 \cdot \infty = 0, & \quad \text{for } x > 0, x \cdot \infty = \infty, \quad \text{for } x < 0, x \cdot \infty = -\infty. \end{aligned}$$

Fix a norm on  $\mathbb{R}^d$  and denote the norm of  $\mathbf{x}$  as  $\|\mathbf{x}\|$ . Let  $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$  be the metric induced by the norm and, as usual, for  $A \subset \mathbb{R}^d$ , set  $d(\mathbf{x}, A) = \inf_{\mathbf{y} \in A} d(\mathbf{x}, \mathbf{y})$ . When attention is focused on the set  $\mathbb{C}$ , and  $A \subset \mathbb{C}$ , the  $\delta$ -dilation or swelling of  $A$  in  $\mathbb{C}$  is  $A^\delta := \{x \in \mathbb{C} : d(x, A) < \delta\}$ . The topology on  $\mathbb{R}^d$  is the usual norm topology referred to as the Euclidean topology and the topology on a subset of  $\mathbb{R}^d$  is the relative topology induced by the Euclidean topology.

Two sets  $A$  and  $B$  in  $\mathbb{R}^d$  are *bounded away* from each other if  $\bar{A} \cap \bar{B} = \emptyset$ , where  $\bar{A}$  and  $\bar{B}$  are the closures of  $A$  and  $B$ .

1.1.2. *Cones.* We denote by  $\mathbb{E} = [0, \infty]^d \setminus \{\mathbf{0}\}$  and  $\mathbb{D} = [0, \infty)^d \setminus \{\mathbf{0}\}$ , the one point puncturing of the compactified and uncompactified versions of  $\mathbb{R}_+^d$ . The symbols  $\mathbb{E}$  and  $\mathbb{D}$  may appear with superscripts denoting subsets of the compactified and uncompactified  $\mathbb{R}_+^d$  respectively. For example,  $\mathbb{E}^{(l)} = \{\mathbf{x} \in [0, \infty]^d : x^{(l)} > 0\}$  and  $\mathbb{D}^{(l)} = \{\mathbf{x} \in [0, \infty)^d : x^{(l)} > 0\}$ , where  $x^{(l)}$  is the  $l$ -th largest component of  $\mathbf{x}$ . Note  $\mathbb{E}^{(1)} = \mathbb{E}$ .

A set  $\mathbb{C} \subset \mathbb{R}^d$  is a cone if  $\mathbf{x} \in \mathbb{C}$  implies  $t\mathbf{x} \in \mathbb{C}$  for all  $t > 0$ . Cones in the Euclidean space are usually denoted by mathematical bold symbols  $\mathbb{C}, \mathbb{D}, \mathbb{E}, \mathbb{F}$ , etc. Since one typically deals with non-negative risk vectors, we focus on the case where  $\mathbb{C} \subset [0, \infty)^d$ . Call a subset  $\mathbb{F} \subset \mathbb{C}$  a closed cone in  $\mathbb{C}$  if  $\mathbb{F}$  is a closed subset of  $\mathbb{C}$  as well as a cone. Example:  $\mathbb{F} = \{\mathbf{0}\}$  or when  $d = 2$ ,  $\mathbb{F} = \{(t, 0) : t \geq 0\}$ . The complement of the closed cone  $\mathbb{F}$  in  $\mathbb{C}$  is an open cone in  $\mathbb{C}$ ; that is, the complement of  $\mathbb{F}$  is an open subset in  $\mathbb{C}$  as well as a cone.

Fix a closed cone  $\mathbb{C} \subset [0, \infty)^d$  containing  $\mathbf{0}$  and suppose  $\mathbb{F} \subset \mathbb{C}$  be a closed cone in  $\mathbb{C}$  containing  $\mathbf{0}$ . Then  $\mathbb{O} := \mathbb{C} \setminus \mathbb{F}$  is an open cone and  $\mathbb{C}$  and  $\mathbb{F}$  are complete separable metric spaces under the metric  $d(\cdot, \cdot)$ . Let  $\mathcal{C}$  denote the Borel  $\sigma$ -algebra of  $\mathbb{C}$ . Clearly  $\mathbb{O}$  is again a separable metric space (not necessarily complete) equipped with the  $\sigma$ -algebra  $\mathcal{O} = \{B \subset \mathbb{O} : B \in \mathcal{C}\}$ .

Examples: (i)  $\mathbb{C} = [0, \infty)^d$ ,  $\mathbb{F} = \{\mathbf{0}\}$  and then  $\mathbb{O} = \mathbb{D} = [0, \infty)^d \setminus \{\mathbf{0}\}$ . (ii)  $d = 2$ ,  $\mathbb{C} = [0, \infty)^2$ ,  $\mathbb{F} = \{(0, x) : x \geq 0\} \cup \{(y, 0) : y \geq 0\}$ ,  $\mathbb{O} = (0, \infty)^2$ . (iii)  $d = 2$ ,  $\mathbb{C} = [0, \infty)^2$ ,  $\mathbb{F} = \{(x, 0) : x \geq 0\}$ , and  $\mathbb{O} = \mathbb{D}_\square := [0, \infty) \times (0, \infty)$ .

1.1.3. *Regularly varying functions.* A function  $U : [0, \infty) \mapsto [0, \infty)$  is regularly varying with index  $\beta \in \mathbb{R}$  if for all  $x > 0$ ,

$$\lim_{t \rightarrow \infty} \frac{U(tx)}{U(t)} = x^\beta.$$

We write  $U \in RV_\beta$ . See Resnick [2008], de Haan and Ferreira [2006], Bingham et al. [1987].

1.1.4. *Vague convergence of measures.* We express vague convergence of Radon measures as  $\xrightarrow{v}$  ([Resnick, 2007, page 173], Kallenberg [1983]) and weak convergence of probability measures as  $\Rightarrow$  [Billingsley, 1999, page 14]. Denote the set of non-negative Radon measures on a space  $\mathbb{S}$  as  $\mathbb{M}_+(\mathbb{S})$  and the set of all non-negative continuous functions with compact support from  $\mathbb{S}$  to  $\mathbb{R}_+$  as  $C_K^+(\mathbb{S})$ .

Vague convergence on  $\mathbb{E}$  has traditionally been used when defining multivariate regular variation. We explain in the next section why continuing with this practice is problematic and what should be done.

## 2. $\mathbb{M}^*$ -convergence of measures and regular variation on cones.

This paper requires a definition of multivariate regular variation on cones of the Euclidean space which differs from the traditional definition through *vague convergence* of measures. Following Hult and Lindskog [2006], we define regular variation based on a notion of convergence of measures we call  $\mathbb{M}^*$ -convergence.

2.1. *Problems with compactification of  $\mathbb{R}^d$ .* Multivariate regular variation on  $[0, \infty)^d$  is usually defined using vague convergence of Radon measures on  $\mathbb{E} = [0, \infty]^d \setminus \{\mathbf{0}\}$  [Resnick, 2007]. The reason for compactifying  $[0, \infty)^d$  and then removing  $\mathbf{0}$  is this makes sets bounded away from  $\{\mathbf{0}\}$  relatively compact (cf. [Resnick, 2007, Section 6.1.3]) and since Radon measures put finite mass on relatively compact sets, this theory is suitable for estimating probabilities of tail regions.

The theory of hidden regular variation may require removal of more than just a point. Furthermore, compactifying from  $[0, \infty)^d$  to  $[0, \infty]^d$  introduces problems. For one thing, it is customary to rely heavily on the polar coordinate transform

$$\mathbf{x} \mapsto \left( \|\mathbf{x}\|, \frac{\mathbf{x}}{\|\mathbf{x}\|} \right)$$

which is only defined on  $[0, \infty)^d \setminus \{\mathbf{0}\}$  and if the state space  $[0, \infty]^d \setminus \{\mathbf{0}\}$  is used an awkward kluge [Resnick, 2007, page 176] is required to show the equivalence of regular variation in polar and Cartesian coordinates. A

workaround is only possible because the limit measure  $\mu$  in (1) puts zero mass on lines through infinity  $\{\mathbf{x} : \bigvee_{i=1}^d x^i = \infty\}$  but absence of mass on lines through  $\infty$  does not necessarily persist for regular variation on other cones [Mitra and Resnick, 2010].

Also, compactification introduces counterintuitive geometric properties. For example, the topology on  $[0, \infty]^d$  can be defined through a homeomorphic map  $[0, \infty]^d \mapsto [0, 1]^d$ , such as

$$\mathbf{z} = (z^1, z^2, \dots, z^d) \mapsto (z^1/(1+z^1), \dots, z^d/(1+z^d)).$$

Restrict attention to  $d = 2$  and consider two parallel lines in  $[0, \infty]^2$  with the same positive and finite slope. These lines both converge to the same point  $(\infty, \infty)$  and therefore in the compactified space, these two parallel lines are not bounded away from each other. Interestingly, this is not the case if the lines are horizontal or vertical.

To see the impact that parallel lines not being bounded away from each other can have recall one of the motivational examples from Section 1 with  $d = 2$ , where the limit measure  $\mu(\cdot)$  in (1) is concentrated on the diagonal  $\mathbb{D}\text{IAG} := \{(z^1, z^2) \in \mathbb{E} : z^1 = z^2\}$  and we need to approximate the tail probability  $P(|Z^1 - Z^2| > w)$  for a large threshold  $w > 0$ . Of course, if we use multivariate regular variation as in (1) to approximate  $P(|Z^1 - Z^2| > w)$ , we approximate  $P(|Z^1 - Z^2| > w)$  as zero. If  $P[Z^1 = Z^2] < 1$ , this approximation is crude. Following the usual definition of HRV, we remove the diagonal  $\mathbb{D}\text{IAG}$  and define regular variation on the sub-cone  $(\mathbb{D}\text{IAG})^c := \{(z^1, z^2) \in \mathbb{E} : z^1 \neq z^2\}$ . Since we seek to approximate  $P(|Z^1 - Z^2| > w)$ , we are interested in the set  $A_{>w} := \{(z^1, z^2) \in \mathbb{E} : |z^1 - z^2| > w\}$ . If we define HRV on the sub-cone  $(\mathbb{D}\text{IAG})^c$  as an asymptotic property using vague convergence, we need the set  $A_{>w}$  to be relatively compact in the sub-cone  $(\mathbb{D}\text{IAG})^c$ . However, if the sub-cone  $(\mathbb{D}\text{IAG})^c$  is endowed with the relative topology from the topology on  $[0, \infty]^2$ ,  $A_{>w}$  is not relatively compact since the boundaries of  $A_{>w}$  are the two parallel lines  $\{(z^1, z^2) \in \mathbb{E} : z^1 - z^2 = w\}$  and  $\{(z^1, z^2) \in \mathbb{E} : z^1 - z^2 = -w\}$ , which are both parallel to the diagonal  $\mathbb{D}\text{IAG}$ . In the topology of  $[0, \infty]^d$ , the boundaries of the set  $A_{>w}$  are not bounded away from the diagonal  $\mathbb{D}\text{IAG}$  and hence by Proposition 6.1 of [Resnick, 2007, page 171], the set  $A_{>w}$  is not relatively compact in  $(\mathbb{D}\text{IAG})^c$ .

As already observed, horizontal or vertical parallel lines are bounded away from each other in  $[0, \infty]^2$ . If the limit measure  $\mu(\cdot)$  in (1) concentrates on the axes, the traditional definition of HRV [Resnick, 2002] removes the axes and defines hidden regular variation on the cone  $(0, \infty]^2$ . However, risk regions of interest of the form  $(z^1, \infty] \times (z^2, \infty]$  are still relatively compact and we do not encounter a problem as above..

Thus, we conclude that a flexible theory of hidden regular variation on general cones of  $[0, \infty)^d$  requires considering the possibility that compactification and vague convergence be abandoned. However, without compactification, how do we guarantee risk sets corresponding to tail events are relatively compact and their probabilities approximable by asymptotic methods? A theory based on  $\mathbb{M}^*$ -convergence of measures sidesteps many difficulties.

**2.2.  $\mathbb{M}^*$ -convergence of measures.** We follow ideas of Hult and Lindskog [2006] who removed only a fixed point from a closed set, whereas we remove a closed cone.

As in Section 1.1.2 we fix a closed cone  $\mathbb{C} \subset [0, \infty)^d$  containing  $\mathbf{0}$  and  $\mathbb{F} \subset \mathbb{C}$  is a closed cone in  $\mathbb{C}$  containing  $\mathbf{0}$ . Set  $\mathbb{O} := \mathbb{C} \setminus \mathbb{F}$ , which is an open cone in  $\mathbb{C}$ . Let  $\mathcal{C}$  be the Borel  $\sigma$ -algebra of  $\mathbb{C}$  and the  $\sigma$ -algebra in  $\mathbb{O}$  is  $\mathcal{O} = \{B \subset \mathbb{O} : B \in \mathcal{C}\}$ . Denote by  $\mathcal{C}_{\mathbb{F}}$  the set of all bounded, continuous real-valued functions  $f$  on  $\mathbb{C}$  such that  $f$  vanishes on  $\mathbb{F}^r := \{\mathbf{x} \in \mathbb{C} : d(\mathbf{x}, \mathbb{F}) < r\}$  for some  $r > 0$ . The class of Borel measures on  $\mathcal{O}$  that assign finite measure to all  $D \in \mathcal{O}$  which are bounded away from  $\mathbb{F}$  is called  $\mathbb{M}^*(\mathbb{C}, \mathbb{O})$ . Equivalently,  $\mu \in \mathbb{M}^*(\mathbb{C}, \mathbb{O})$  if and only if  $\mu$  is finite on  $\mathbb{C} \setminus \mathbb{F}^r$  for all  $r > 0$ .

**DEFINITION 2.1** ( $\mathbb{M}^*(\mathbb{C}, \mathbb{O})$ -convergence). For  $\mu, \mu_n \in \mathbb{M}^*(\mathbb{C}, \mathbb{O})$ ,  $n \geq 1$ ,  $\mu_n$  converges to  $\mu$  in  $\mathbb{M}^*(\mathbb{C}, \mathbb{O})$  if

$$(2) \quad \lim_{n \rightarrow \infty} \mu_n(B) = \mu(B),$$

for all  $B \in \mathcal{O}$  with  $\mu(\partial B) = 0$  and  $B$  bounded away from  $\mathbb{F}$ . We write  $\mu_n \xrightarrow{*} \mu$  in  $\mathbb{M}^*(\mathbb{C}, \mathbb{O})$  as  $n \rightarrow \infty$ .

We can metrize the space  $\mathbb{M}^*(\mathbb{C}, \mathbb{O})$ . One method: For  $\mu, \nu \in \mathbb{M}^*(\mathbb{C}, \mathbb{O})$  define

$$(3) \quad d_{\mathbb{M}^*}(\mu, \nu) := \int_0^\infty e^{-r} \frac{d_P(\mu^{(r)}, \nu^{(r)})}{1 + d_P(\mu^{(r)}, \nu^{(r)})} dr,$$

where  $\mu^{(r)}, \nu^{(r)}$  are the restrictions of  $\mu, \nu$  to  $\mathbb{C} \setminus \mathbb{F}^r$  and  $d_P$  is the Prohorov metric [Prohorov, 1956].

Following Hult and Lindskog [2006],  $(\mathbb{M}^*(\mathbb{C}, \mathbb{O}), d_{\mathbb{M}^*})$  is a complete separable metric space and the expected analogue of the Portmanteau theorem [Billingsley, 1999] holds: For  $\mu_n \in \mathbb{M}^*(\mathbb{C}, \mathbb{O})$ ,  $n \geq 0$ , the following are equivalent:

1.  $\mu_n \xrightarrow{*} \mu_0$  in  $\mathbb{M}^*(\mathbb{C}, \mathbb{O})$  as  $n \rightarrow \infty$ .
2. For each  $f \in \mathcal{C}_{\mathbb{F}}$ ,  $\int f d\mu_n \rightarrow \int f d\mu_0$  as  $n \rightarrow \infty$ .
3.  $\limsup_{n \rightarrow \infty} \mu_n(A) \leq \mu_0(A)$  and  $\liminf_{n \rightarrow \infty} \mu_n(G) \geq \mu_0(G)$  for all closed sets  $A \in \mathcal{O}$  and open sets  $G \in \mathcal{O}$  such that  $\overline{G} \cap \mathbb{F} = \emptyset$ .

**3. Regular and hidden regular variation on cones.** We define regular variation on a nested sequence of cones, where each cone is a subset of the previous one. Each cone in the sequence possesses a different regular variation, which remains hidden while studying regular variation on the bigger cones in the sequence.

3.1. *Regular variation.* We use the concepts of Section 2 to define regular variation.

DEFINITION 3.1. Suppose  $\mathbb{F} \subset \mathbb{C} \subset [0, \infty)^d$  and  $\mathbb{F}$  and  $\mathbb{C}$  are closed cones containing  $\mathbf{0}$ . A random vector  $\mathbf{Z} \in \mathbb{C}$  has a distribution with a regularly varying tail on  $\mathbb{O} = \mathbb{C} \setminus \mathbb{F}$ , if there exist a function  $b(t) \uparrow \infty$  and a non-zero measure  $\nu(\cdot) \in \mathbb{M}^*(\mathbb{C}, \mathbb{O})$  such that as  $t \rightarrow \infty$ ,

$$(4) \quad tP \left[ \frac{\mathbf{Z}}{b(t)} \in \cdot \right] \xrightarrow{*} \nu(\cdot) \quad \text{in } \mathbb{M}^*(\mathbb{C}, \mathbb{O}).$$

When there is no danger of confusion, we sometimes use the notation  $\mathbb{M}^*(\mathbb{O})$  to mean  $\mathbb{M}^*(\mathbb{C}, \mathbb{O})$  and sometimes abuse language and say the distribution is regularly varying on  $\mathbb{O}$ .

Definition 3.1 implies there exists  $\alpha > 0$  such that  $b(\cdot) \in RV_{1/\alpha}$  and that  $\nu$  has the scaling property:

$$(5) \quad \nu(c \cdot) = c^{-\alpha} \nu(\cdot), \quad c > 0.$$

This can be derived as in [Hult and Lindskog, 2006, Theorem 3.1]. We define standard multivariate regular variation, hidden regular variation and the conditional extreme value model in terms of Definition 3.1 and attempt to relate each to the way these ideas were first proposed on modifications of compactified spaces.

Examples:

1. Let  $\mathbb{C} = [0, \infty)^d$  and  $\mathbb{F} = \{\mathbf{0}\}$  and  $\mathbb{D} = \mathbb{O} = [0, \infty)^d \setminus \{\mathbf{0}\}$ . Regular variation on  $\mathbb{D}$  is equivalent to regular variation defined in (1) on  $\mathbb{E}$ . The definition in (1) precludes  $\mu$  having mass on the lines through  $\infty$ . See Appendix 8.1.
2. Let  $d = 2$ ,  $\mathbb{C} = [0, \infty)^2$  and  $\mathbb{F} = \{(x, 0) : x \geq 0\} \cup \{(0, y), y \geq 0\}$  and  $\mathbb{O} = (0, \infty)^2$ , the first quadrant with both the  $x$  and  $y$  axes removed. This is the restriction to  $[0, \infty)^2$  of the cone used in the definition of hidden regular variation in Resnick [2002]. For  $d > 2$ , other examples of  $\mathbb{F}$  are in Mitra and Resnick [2010] and Subsection 8.2 provides a comparison between regular variation defined in (4) on

$\mathbb{D}^{(l)} = [0, \infty)^d \setminus \{\mathbf{x} \in [0, \infty)^d : x^{(l)} > 0\}$  and regular variation defined using (1) on  $\mathbb{E}^{(l)} = [0, \infty)^d \setminus \{\mathbf{x} \in \mathbb{E} : x^{(l)} > 0\}$  where recall  $x^{(l)}$  is the  $l$ th largest component of  $\mathbf{x}$ . The two notions are equivalent provided there is no mass on  $\mathbb{E}^{(l)} \setminus \mathbb{D}^{(l)}$ .

3. Suppose  $d = 2$  and  $\mathbb{C} = [0, \infty)^d$  and  $\mathbb{F} = \{(x, 0) : x \geq 0\}$ . Then  $\mathbb{O} = \{(x, y) \in [0, \infty)^2 : y > 0\} = \mathbb{D}_{\square}$ , the first quadrant with the  $x$ -axis removed. This is the restriction to  $[0, \infty)^2$  of the cone used in the definition of the conditional extreme value model [Heffernan and Tawn, 2004, Heffernan and Resnick, 2007, Das and Resnick, 2011].
4. Suppose  $d = 2$  and  $\mathbb{C} = [0, \infty)^d$  and  $\mathbb{F} = \{(x, x) : x \geq 0\}$ . Then  $\mathbb{O}$  is the first quadrant with the diagonal removed. This example is suitable for discussing asymptotic full dependence (Resnick [2008, page 294], Resnick [2007, page 195]) and is considered in Example 5.2.

### 3.2. Spectral measures, unit spheres and semi-parametric representations.

Regular variation on  $\mathbb{E}$  using the vague convergence definition as in (1) allows a polar coordinate transformation  $\mathbf{x} \mapsto (\|\mathbf{x}\|, \mathbf{x}/\|\mathbf{x}\|)$ . Assuming  $b(\cdot) \in RV_{1/\alpha}$ , the limit measure has the scaling property and when this is expressed in polar coordinates yields the version of (1)

$$tP\left[\frac{\|\mathbf{Z}\|}{b(t)} \in dx, \frac{\mathbf{Z}}{\|\mathbf{Z}\|} \in d\mathbf{a}\right] \xrightarrow{v} \alpha x^{-\alpha-1} dx \times S^*(d\mathbf{a})$$

where  $S^*$  is a finite measure on  $\partial\mathfrak{N} = \{\mathbf{a} \in \mathbb{E} : \|\mathbf{a}\| = 1\}$ , the unit sphere. Fixing  $S^*(\partial\mathfrak{N}) = c$ , we define  $S(\cdot) = S^*(\cdot)/c$  which becomes a probability measure on  $\partial\mathfrak{N}$  called the spectral or angular measure. So in polar coordinates, the limit measure  $\mu$  in (1) has a semi-parametric product structure depending on the parameter  $\alpha$  and the measure  $S$ .

In  $\mathbb{E}$ , the unit sphere  $\partial\mathfrak{N} = \{\mathbf{x} \in \mathbb{E} : \|\mathbf{x}\| = 1\} = \{\mathbf{x} \in \mathbb{E} : d(\mathbf{x}, \mathbf{0}) = 1\}$  is compact. However, this may no longer be true when moving to other sub-cones. For instance in  $(0, \infty]^2$  the usual unit sphere is not relatively compact. While the polar coordinate transformation still allows this semi-parametric representation for other cones, the analogue of  $S$  is no longer necessarily finite and this is a problem for inference. We explain next how to use a change of coordinates different from the polar coordinate transformation which always produces a finite spectral measure analogue. Heffernan and Resnick [2007] and Mitra and Resnick [2010] consider alternatives to the polar coordinate transformation that twist limit measures into a semi-parametric form.

Proceed using the context of Definition 3.1. Assume  $\mathbb{C}, \mathbb{F}$  and  $\mathbb{O}$  are defined as in Definition 3.1 and define  $\mathfrak{N}_{\mathbb{O}} = \{\mathbf{x} \in \mathbb{O} : d(\mathbf{x}, \mathbb{F}) \geq 1\}$  and  $\mathfrak{N}_{\mathbb{O}}$  is a subset

of  $\mathbb{O}$  bounded away from  $\mathbb{F}$  and  $\cup_{\{\theta>0\}}\theta\mathbb{N}_{\mathbb{O}} = \mathbb{O}$ . From Definition 3.1, we have  $0 < \nu(\mathbb{N}_{\mathbb{O}}) < \infty$ . The scaling property (5) implies the scaling function  $b(t)$  in (4) can be chosen so that so that  $\nu(\mathbb{N}_{\mathbb{O}}) = 1$ . Define the related set  $\partial\mathbb{N}_{\mathbb{O}} = \{\mathbf{x} \in \mathbb{O} : d(\mathbf{x}, \mathbb{F}) = 1\}$ .

Examples: (i)  $\mathbb{C} = [0, \infty)^d$  and  $\mathbb{F} = \{\mathbf{0}\}$  and  $d(\mathbf{x}, \mathbb{F}) = \|\mathbf{x}\|$  and  $\partial\mathbb{N}_{\mathbb{O}} = \{\mathbf{x} : \|\mathbf{x}\| = 1\}$ . (ii)  $d = 2$ ,  $\|\mathbf{x}\| = x^1 \vee x^2$ ,  $\mathbb{C} = [0, \infty)^2$  and  $\mathbb{F} = \{(x, 0) : x \geq 0\} \cup \{(0, y), y \geq 0\}$  and  $\mathbb{O} = (0, \infty)^2$ . Then  $\partial\mathbb{N}_{\mathbb{O}} = \{\mathbf{x} : d(\mathbf{x}, \mathbb{F}) = 1\} = \{\mathbf{x} : x^1 \wedge x^2 = 1\}$ . (iii)  $d = 2$ ,  $\|\mathbf{x}\| = x^1 \vee x^2$ , and  $\mathbb{C} = [0, \infty)^d$  and  $\mathbb{F} = \{(x, 0) : x \geq 0\}$ . Then  $\partial\mathbb{N}_{\mathbb{O}} = \{\mathbf{x} : d(\mathbf{x}, \mathbb{F}) = 1\} = \{(x, y) : x \geq 0, y = 1\}$ .

We transform to an appropriate coordinate system in which the limit measure  $\nu$  in (4) is a product of two components: a one-dimensional Pareto measure and a probability measure defined on  $\partial\mathbb{N}_{\mathbb{O}}$  called the hidden spectral or angular measure. To do this note two properties of the distance function  $d(\cdot, \mathbb{F})$ :

- (i) Since  $\mathbb{F}$  is a closed subset of  $\mathbb{C}$ ,  $d(\mathbf{x}, \mathbb{F}) > 0$  for all  $\mathbf{x} \in \mathbb{O}$  (else  $\mathbf{x} \in \mathbb{F}$ ); and
- (ii) Since  $\mathbb{F}$  is a cone,  $\theta \cdot \mathbb{F} = \mathbb{F}$  for  $\theta > 0$ . Hence,  $d(\theta\mathbf{x}, \mathbb{F}) = d(\theta\mathbf{x}, \theta\mathbb{F}) = \theta d(\mathbf{x}, \mathbb{F})$ , that is  $d(\cdot, \mathbb{F})$  is homogeneous of order 1.

A lemma is necessary for the decomposition of limit measure  $\nu(\cdot)$ . For a set  $A \in [0, \infty)^d$  we set  $(A)_1 = \{x^1 : \mathbf{x} \in A\}$ .

LEMMA 3.2. *Suppose  $h : \mathbb{O} \mapsto (0, \infty) \times \partial\mathbb{N}_{\mathbb{O}}$  is a continuous bijection satisfying*

- (i) *For every measurable  $A \subset (0, \infty) \times \partial\mathbb{N}_{\mathbb{O}}$  with  $\overline{(A)_1} \cap \{0\} = \emptyset$ ,  $h^{-1}(A)$  is bounded away from  $\mathbb{F}$ .*
- (ii) *For every measurable  $B \subset \mathbb{O}$  with  $B$  bounded away from  $\mathbb{F}$ ,  $\overline{(h(B))_1} \cap \{0\} = \emptyset$ .*

*Then, the following statements are equivalent:*

- (i) *As  $t \rightarrow \infty$ ,*

$$\mu_t(\cdot) \xrightarrow{*} \mu(\cdot) \quad \text{in } \mathbb{M}^*(\mathbb{C}, \mathbb{O}).$$

- (ii) *For all measurable  $A \subset (0, \infty) \times \partial\mathbb{N}_{\mathbb{O}}$  such that  $\overline{(A)_1} \cap \{0\} = \emptyset$  and  $\mu \circ h^{-1}(\partial A) = 0$ ,*

$$\mu_t \circ h^{-1}(A) \rightarrow \mu \circ h^{-1}(A).$$

PROOF. The proof follows the steps of the proof of Theorem 2.5 of Hult and Lindskog [2006].  $\square$

Now by applying Lemma 3.2 with  $h : \mathbf{x} \mapsto (d(\mathbf{x}, \mathbb{F}), \mathbf{x}/d(\mathbf{x}, \mathbb{F}))$ , we are able to decompose  $\nu$  as follows.

PROPOSITION 3.3. *Regular variation on  $\mathbb{O}$  as given in (4) is equivalent to*

$$(6) \quad tP \left[ \left( \frac{d(\mathbf{Z}, \mathbb{F})}{b(t)}, \frac{\mathbf{Z}}{d(\mathbf{Z}, \mathbb{F})} \right) \in A \right] \rightarrow c\nu_{(\alpha)} \times S_{\mathbb{O}}(A)$$

for all measurable  $A \subset (0, \infty) \times \partial\mathfrak{N}_{\mathbb{O}}$  such that  $\overline{(A)_1} \cap \{0\} = \emptyset$  and  $\nu \circ h^{-1}(\partial A) = 0$  where  $c > 0$ ,  $S_{\mathbb{O}}(\cdot)$  is a probability measure on  $\partial\mathfrak{N}_{\mathbb{O}}$  and  $\nu_{(\alpha)}(\cdot)$  is the Pareto measure given by  $\nu_{(\alpha)}((x, \infty)) = x^{-\alpha}$  for  $x > 0$ . Call  $S_{\mathbb{O}}(\cdot)$  the spectral measure on  $\mathbb{O}$ ; it is related to  $\nu(\cdot)$  by the relation

$$(7) \quad S_{\mathbb{O}}(\Lambda) = \frac{\nu \left( \left\{ \mathbf{x} \in \mathbb{O} : d(\mathbf{x}, \mathbb{F}) \geq 1, \frac{\mathbf{x}}{d(\mathbf{x}, \mathbb{F})} \in \Lambda \right\} \right)}{\nu(\mathbf{x} \in \mathbb{O} : d(\mathbf{x}, \mathbb{F}) \geq 1)}.$$

Since  $(0, \infty) \times \partial\mathfrak{N}_{\mathbb{O}}$  is not a cone, we have not phrased the convergence in (6) as  $\mathbb{M}^*$  convergence as in (4). To do so would require more reworking of the convergence theory in Hult and Lindskog [2006].

COROLLARY 3.4. *The convergence in (4) is equivalent to the following two conditions:*

- (i) *The distribution of  $d(\mathbf{Z}, \mathbb{F})$  is regularly varying on  $(0, \infty)$  following Definition 3.1 with  $\mathbb{C} = [0, \infty)$ .*
- (ii) *The conditional distribution of  $\mathbf{Z}/d(\mathbf{Z}, \mathbb{F})$  given  $d(\mathbf{Z}, \mathbb{F}) > t$ , converges weakly,*

$$P[\mathbf{Z}/d(\mathbf{Z}, \mathbb{F}) \in \cdot \mid d(\mathbf{Z}, \mathbb{F}) > t] \Rightarrow S_{\mathbb{O}}(\cdot) \quad (t \rightarrow \infty).$$

REMARK 3.5. We make a few remarks about Proposition 3.3.

- (i) On the role of the distance function: Proposition 3.3 emphasizes that the spectral probability measure  $S_{\mathbb{O}}(\cdot)$  is dependent on the choice of distance function  $d(\cdot, \cdot)$ . Corollary 3.4 allows us to use the distance function  $d(\cdot, \cdot)$  to detect regular variation on  $\mathbb{O}$ ; see Section 6. However, extending the distance function to a compactified space such as  $[0, \infty]^d$  is difficult and this provides another reason why we deviated from the standard discussion of regular variation using compactified spaces and vague convergence.
- (ii) Connections to prior treatments:
  - (a) Proposition 3.1 of Mitra and Resnick [2010] decomposes the limit measure  $\mu^{(l)}(\cdot)$  on  $\mathbb{E}^{(l)}$  (see (29) below) by applying the transformation  $T : \mathbf{x} \mapsto (x^{(l)}, \mathbf{x}/x^{(l)})$ , where  $x^{(l)}$  is the  $l$ -th largest

component of  $\mathbf{x}$ . If we choose  $\mathbb{C} = [0, \infty)^d$ ,  $\mathbb{F} = \{\mathbf{x} \in \mathbb{C} : x^{(l)} = 0\}$  and  $\mathbb{O} = \{\mathbf{x} \in \mathbb{C} : x^{(l)} > 0\}$ , and choose the  $L_\infty$ -norm when defining  $d(\cdot, \cdot)$ , then  $d(\mathbf{x}, \mathbb{F}) = x^{(l)}$  and our Proposition 3.3 gives a version of Proposition 3.1 in Mitra and Resnick [2010].

- (b) For considering regular variation on the cone  $\mathbb{E}_\cap := [0, \infty] \times (0, \infty]$ , Heffernan and Resnick [2007, Proposition 4] give a decomposition of their limit measure  $\mu_*(\cdot)$  by applying the transformation  $T : (x, y) \mapsto (y, x/y)$ . If we choose  $\mathbb{C} = [0, \infty)^2$ ,  $\mathbb{F} = [0, \infty) \times \{0\}$  and  $\mathbb{O} = [0, \infty) \times (0, \infty) =: \mathbb{D}_\cap$ , and define  $d(\cdot, \cdot)$  using the  $L_\infty$ -norm, then  $d((x, y), \mathbb{F}) = y$  and our Proposition 3.3 connects with Heffernan and Resnick [2007, Proposition 4].
- (c) Proposition 3.3 also relates to the usual polar coordinate characterization of multivariate regular variation on  $\mathbb{E}$  as in Resnick [2007, page 173]. Set  $\mathbb{F} = \{\mathbf{0}\}$  which is a closed cone in  $\mathbb{C} = [0, \infty)^d$  and  $d(\mathbf{x}, \{\mathbf{0}\}) = \|\mathbf{x}\|$ . See also Proposition 8.2.

**3.3. Hidden regular variation.** As in Definition 3.1, consider  $\mathbb{F} \subset \mathbb{C} \subset [0, \infty)^d$  with  $\mathbb{F}$  and  $\mathbb{C}$  closed cones containing  $\{\mathbf{0}\}$ . Suppose  $\mathbb{F}_1$  is another subset of  $\mathbb{C}$  that is a closed cone containing  $\{\mathbf{0}\}$ . Then  $\mathbb{F} \cup \mathbb{F}_1$  is also a closed cone containing  $\{\mathbf{0}\}$ . Set  $\mathbb{O} = \mathbb{C} \setminus \mathbb{F}$  and  $\mathbb{O}_1 = \mathbb{C} \setminus (\mathbb{F} \cup \mathbb{F}_1)$ .

**DEFINITION 3.6.** The distribution of a random vector  $\mathbf{Z} \in \mathbb{C}$  that is regularly varying on  $\mathbb{O}$  with scaling function  $b(t)$  in (4) possesses hidden regular variation (HRV) on  $\mathbb{O}_1$  if

1. The distribution of  $\mathbf{Z}$  is also regularly varying on  $\mathbb{O}_1$  with scaling function  $b_1(t)$  and limit measure  $\nu_1$  and
2.  $b(t)/b_1(t) \rightarrow \infty$  as  $t \rightarrow \infty$ .

Observe that the condition  $b(t)/b_1(t) \rightarrow \infty$  implies  $\nu$  puts zero mass on  $\mathbb{O}_1$ . (See Resnick [2002].) From Definition 3.6 and (5), it follows that there exists  $\alpha_1 \geq \alpha$  such that  $b_1(\cdot) \in RV_{1/\alpha_1}$  and on  $\mathbb{O}_1$ , the limit measure  $\nu_1(\cdot)$  in (4) satisfies the scaling property

$$(8) \quad \nu_1(c \cdot) = c^{-\alpha_1} \nu_1(\cdot), \quad c > 0.$$

Example: Let  $\mathbf{Z} = (Z^1, Z^2)$  be iid unit Pareto random variables. Then the distribution of  $\mathbf{Z}$  is regularly varying on  $\mathbb{D}$  with  $b(t) = t$  and possesses HRV on  $(0, \infty)^2$  with  $b_1(t) = \sqrt{t}$ . Somewhat more generally, if  $\mathbf{Z} = (Z^1, Z^2)$  has regular variation on  $\mathbb{D} = [0, \infty)^2 \setminus \{\mathbf{0}\}$ , then in the presence of asymptotic independence, HRV offers non-zero estimates of joint tail probabilities

$P[Z^1 > x, Z^2 > y]$  for large thresholds  $x, y > 0$ , whereas regular variation on  $\mathbb{D}$  estimates joint tail probabilities as zero. Other examples are considered in Section 5.

REMARK 3.7. We make a few remarks on Definition 3.6.

- (i) There is flexibility in choosing  $\mathbb{O}_1$  and this flexibility is useful for defining HRV on a sequence of cones. Cones can be chosen based on the risk regions whose probabilities are required. For example, if  $d = 2$ , we choose the cones  $\mathbb{D}$  and  $\mathbb{D}^{(2)} := \{\mathbf{z} \in [0, \infty)^d : z^1 \wedge z^2 > 0\}$  if we need the probability that components of the risk vector simultaneously exceed thresholds.
- (ii) *Differences with existing notions of hidden regular variation:* Previous considerations of HRV relied on vague convergence and compactification and were applied to specific choices of cones. Resnick [2002], Heffernan and Resnick [2005], and Maulik and Resnick [2005] consider HRV on  $(0, \infty]^2$  and Mitra and Resnick [2010] consider the cone  $\mathbb{E}^{(l)} = \{\mathbf{x} \in [0, \infty)^d : x^{(l)} > 0\}$ . The choice of these specific cones may not provide sufficient flexibility and generality. For example, to estimate  $P[|X - Y| > x]$  when asymptotic full dependence is present, such cones considered previously are of no help. See Example 5.2.

3.3.1. *Where to seek HRV.* Suppose the distribution of  $\mathbf{Z}$  is regularly varying on  $\mathbb{O}$  and that the limit measure  $\nu$  in (4) gives zero mass to a subset  $R$  of  $\mathbb{O}$ . Using the asymptotic property of regular variation to estimate  $P[\mathbf{Z} \in tR]$  for large  $t$ , means such an estimate is 0. So we seek another regular variation on a subset of  $R$  which is of lower order.

Thus, when seeking HRV our focus is on subsets of  $\mathbb{O}$  where the limit measure  $\nu(\cdot)$  gives zero mass. A systematic way to find HRV is facilitated by the following simple remark.

PROPOSITION 3.8. *In Definition 3.1, the support of the limit measure  $\nu$  is a closed cone  $F_\nu \subset \mathbb{C}$  containing  $\mathbf{0}$ .*

PROOF. Let  $\text{supp}(\nu)$  denote the support of  $\nu$ . By definition,  $\text{supp}(\nu)$  is closed. Let  $\mathbf{x} \in \text{supp}(\nu)$  and we show for  $t > 0$  that  $t\mathbf{x} \in \text{supp}(\nu)$ . For small  $\delta$ , by (5)

$$\begin{aligned} \nu((t\mathbf{x} - \delta\mathbf{1}, t\mathbf{x} + \delta\mathbf{1}) \cap \mathbb{C}) &= \nu\left(t\left(\left(\mathbf{x} - \frac{\delta}{t}\mathbf{1}, \mathbf{x} + \frac{\delta}{t}\mathbf{1}\right) \cap \mathbb{C}\right)\right) \\ &= t^{-\alpha}\nu\left(\left(\mathbf{x} - \frac{\delta}{t}\mathbf{1}, \mathbf{x} + \frac{\delta}{t}\mathbf{1}\right) \cap \mathbb{C}\right) > 0 \end{aligned}$$

since  $\mathbf{x} \in \text{supp}(\nu)$ . □

So  $\mathbb{F}_\nu = \text{supp}(\nu)$  is a union of rays emanating from the origin. It is also true that

$$\text{supp}(\nu) = \{t \cdot \text{supp}(S_\mathbb{O}), t \geq 0\}.$$

When seeking HRV on a cone smaller than  $\mathbb{O}$ , we conclude that

$$\mathbb{O}_\nu := \mathbb{C} \setminus (\mathbb{F} \cup \mathbb{F}_\nu)$$

is the largest possible sub-cone of  $\mathbb{O}$  where we might find a different regular variation. In practice, guided by the type of risk region whose probability we need to estimate, we find a closed cone  $\mathbb{F}_1 \subset \mathbb{C}$  containing  $\mathbf{0}$  such that  $\mathbb{F}_1 \supset \mathbb{F}_\nu$  and set  $\mathbb{O}_1 = \mathbb{C} \setminus (\mathbb{F} \cup \mathbb{F}_1)$  and then seek regular variation on  $\mathbb{O}_1$ . Possibly, but not necessarily  $\mathbb{F}_1 = \text{supp}(\nu)$ . Examples are in Mitra and Resnick [2010] and Section 5.

*3.3.2. Regular variation on a sequence of cones.* Having found regular variation on  $\mathbb{O}$  with HRV on  $\mathbb{O}_1$ , we ask: should we stop here? There might be a subcone  $\mathbb{O}_2$  of  $\mathbb{O}_1$ , where  $\nu_1(\cdot)$  gives zero mass and hence, there might exist a different regular variation on  $\mathbb{O}_2$ .

To proceed further, as before remove the support of  $\nu_1$  from  $\mathbb{O}_1$  and consider the set

$$\mathbb{O}_{\nu_1} := \mathbb{O}_1 \setminus \text{supp}(\nu_1)$$

and  $\mathbb{O}_{\nu_1}$  is the largest possible sub-cone of  $\mathbb{O}_1$  where we might find a different regular variation. So we choose  $\mathbb{F}_2 \supset \text{supp}(\nu_1)$  and set  $\mathbb{O}_2 = \mathbb{C} \setminus (\mathbb{F} \cup \mathbb{F}_1 \cup \mathbb{F}_2)$  and seek regular variation on  $\mathbb{O}_2$  with scaling function  $b_2(t)$  such that  $b_1(t)/b_2(t) \rightarrow \infty$  as  $t \rightarrow \infty$ . This last condition guarantees the regular variation on  $\mathbb{O}_2$  is of lower order than the regular variation on either  $\mathbb{O}$  or  $\mathbb{O}_1$  and hidden from both higher order regular variations. This process of discovery is continued as long as on each new cone regular variation is found. Example 5.3 shows this discovery process may lead to an infinite sequence of cones.

From our definition of HRV, at each stage of the discovery process we have some flexibility in choosing the next cone where we seek HRV. Example 5.3 shows it may be impractical to analyze HRV on every possible cone as the discovery process may lead to an infinite sequence of cones. A more practical approach is to decide on a particular finite sequence of cones based on the risk regions of interest; see Remark 3.7(i). For example, if we are interested in estimating joint tail probabilities, we might consider only the sequence of cones  $\mathbb{D} = [0, \infty)^d \setminus \{\mathbf{0}\} \supset \mathbb{D}^{(2)} = \{\mathbf{x} \in [0, \infty)^d : x^{(2)} > 0\} \supset \dots \supset \mathbb{D}^{(d)} = \{\mathbf{x} \in [0, \infty)^d : x^{(d)} > 0\}$ ; cf. Mitra and Resnick [2010].

#### 4. Remarks on other models of multivariate regular variation.

Despite the fact that most common examples in heavy tail analysis start by analyzing convergence on the cone  $[0, \infty)^d \setminus \{\mathbf{0}\}$ , this need not always be the case. For example, the standard case of the conditional extreme value (CEV) model [Heffernan and Resnick, 2007, Das and Resnick, 2011], is regular variation (4) with  $b(t) = t$  on the cone  $\mathbb{D}_\square := \mathbb{O} = [0, \infty) \times (0, \infty)$  with  $\mathbb{C} = [0, \infty)^2$  and  $\mathbb{F} = [0, \infty) \times \{0\}$ .

4.1. *The CEV model.* The conditional extreme value (CEV) model, suggested in Heffernan and Tawn [2004], is an alternative to classical multivariate extreme value theory (MEVT). In contrast with classical MEVT which implies all marginals are in a maximal domain of attraction, in the CEV model only a particular subset of the random vector is assumed to be in a maximal domain of attraction. For convenience, restrict attention to  $d = 2$ .

The CEV model as formulated in Heffernan and Resnick [2007], Das and Resnick [2011] allows variables to be centered as well as scaled. To make comparison with models of regular variation on the first quadrant easy, we recall the vague convergence definition using only scaling functions.

DEFINITION 4.1. Suppose  $\mathbf{Z} := (\xi, \eta) \in \mathbb{R}_+^2$  is a random vector and there exist functions  $a_1(t), a_2(t) > 0$  and a non-null Radon measure  $\mu$  on Borel subsets of  $\mathbb{E}_\square := [0, \infty) \times (0, \infty]$  such that in  $\mathbb{M}_+(\mathbb{E}_\square)$

$$(9) \quad tP \left[ \left( \frac{\xi}{a_1(t)}, \frac{\eta}{a_2(t)} \right) \in \cdot \right] \xrightarrow{v} \mu(\cdot).$$

Additionally assume that  $\mu$  satisfies the following non-degeneracy conditions:

- (a)  $\mu([0, x] \times (y, \infty))$  is a non-degenerate distribution in  $x$ ,
- (b)  $\mu([0, x] \times (y, \infty)) < \infty$ .

Also assume that

- (c)  $H(x) := \mu([0, x] \times (1, \infty))$  is a probability distribution.

In such a case  $\mathbf{Z}$  satisfies a conditional extreme value model and we write  $\mathbf{Z} \in CEV(a_1, a_2)$ .

The general CEV model, provided the limit measure is not a product, can be standardized to have standard regular variation on the cone  $\mathbb{E}_\square$  [Das and Resnick, 2011, pg. 236]. Following the theme of Examples 1 and 2 at the beginning of Section 3, if no mass exists on the lines through  $\infty$ , and

$a_1 = a_2$ , then the vague convergence definition of the CEV model on  $\mathbb{E}_\square$  is the same as the  $\mathbb{M}^*$  definition on regular variation  $\mathbb{D}_\square$ .

The issue of mass on the lines through  $\infty$  is significant since if mass on these lines is allowed, there is a statistical identifiability problem in the sense that in  $\mathbb{E}_\square$  it is possible to have two different limits in (9) under two different normalizations and under one normalization, there is mass on the lines through  $\infty$  and with the other, such mass is absent. Restricting to  $\mathbb{D}_\square$  resolves the identifiability problem as in this space, limits are unique. See Example 5.4.

4.2. *Non-standard regular variation.* Standard multivariate regular variation on  $\mathbb{E}$  requires the same normalizing function to scale all components and is a convenient starting place for theory but unrealistic for applications as it makes all one dimensional marginal distributions tail equivalent. Non-standard regular variation [Resnick, 2007, Section 6.5.6] allows different normalizing functions for vector components and hence permits each marginal distribution tail to have a different tail index. When the components of the risk vector have different tail indices, non-standard regular variation is sensitive to the different tail strengths. On  $\mathbb{E}$  or  $\mathbb{D}$ , non-standard regular variation takes the form

$$(10) \quad tP[(Z^i/a_i(t), i = 1, \dots, d) \in \cdot] \rightarrow \nu(\cdot),$$

for scaling functions  $a_i(t) \uparrow \infty$ ,  $i = 1, \dots, d$ , where convergence is vague for  $\mathbb{E}$  and in  $\mathbb{M}^*$  for  $\mathbb{D}$ . If convergence is in  $\mathbb{E}$  and there is no mass on the lines through  $\infty$ , the difference between convergence in  $\mathbb{E}$  and  $\mathbb{D}$  evaporates.

In cases where  $a_i(t)/a_{i+1}(t) \rightarrow 0$ , it is sometimes possible to compare the information in non-standard regular variation with what can be obtained from HRV. Sometimes HRV provides more detailed information. Consider the following

EXAMPLE 4.2. Suppose  $X_1, X_2, X_3$  are independent random variables where  $X_1$  is Pareto(1),  $X_2$  is Pareto(3) and  $X_3$  is Pareto(4). Further assume that  $B$  is a Bernoulli(1/2) random variable independent of  $X_1, X_2, X_3$ . Define

$$\mathbf{Z} = (Z^1, Z^2) := B(X_1, X_3) + (1 - B)(X_2, X_2).$$

Non-standard regular variation on  $\mathbb{E}$  or  $\mathbb{D}$  is given by

$$tP \left[ \left( \frac{Z^1}{t}, \frac{Z^2}{t^{1/3}} \right) \in dx dy \right] \rightarrow \frac{1}{2} x^{-2} dx \cdot \epsilon_0(dy) + \frac{1}{2} \epsilon_0(dx) \cdot 3y^{-4} dy,$$

where  $\epsilon_0$  indicates the point measure at 0 and the limit measure concentrates on the two axes. Now HRV can also be sought under such non-standard regular variation and we get

$$tP \left[ \left( \frac{Z^1}{t^{3/7}}, \frac{(Z^2)^3}{t^{3/7}} \right) \in dx dy \right] \rightarrow \frac{1}{2} x^{-2} dx 4y^{-5} dy$$

on the space  $\mathbb{E}^{(2)}$ . Note that, this form of regular variation completely ignores the existence of the completely dependent component of  $\mathbf{Z}$  given by  $(X_2, X_2)$ . Alternatively, if we pursue regular variation and HRV on a sequence of cones as defined in this paper then we observe the following convergences as  $t \rightarrow \infty$ :

1. On  $\mathbb{D}$ , we have

$$tP \left[ \left( \frac{Z^1}{t}, \frac{Z^2}{t} \right) \in dx dy \right] \rightarrow \frac{1}{2} x^{-2} dx \epsilon_0(dy).$$

2. In the next step on  $\mathbb{D} \setminus \{\text{x-axis}\}$ , we have

$$tP \left[ \left( \frac{Z^1}{t^{1/3}}, \frac{Z^2}{t^{1/3}} \right) \in dx dy \right] \rightarrow \frac{1}{2} 3x^{-4} dx \epsilon_x(dy).$$

3. Next, on  $\mathbb{D} \setminus [\{\text{x-axis}\} \cup \{\text{diagonal}\}]$ , we have

$$tP \left[ \left( \frac{Z^1}{t^{1/4}}, \frac{Z^2}{t^{1/4}} \right) \in dx dy \right] \rightarrow \frac{1}{2} \epsilon_0(dx) 4y^{-5} dy.$$

4. Finally, on  $\mathbb{D} \setminus [\{\text{x-axis}\} \cup \{\text{diagonal}\} \cup \{\text{y-axis}\}]$ , we have

$$tP \left[ \left( \frac{Z^1}{t^{1/5}}, \frac{Z^2}{t^{1/5}} \right) \in dx dy \right] \rightarrow \frac{1}{2} x^{-2} dx 4y^{-5} dy.$$

Clearly, this analysis captures the structure of  $\mathbf{Z}$  better than what non-standard regular variation along with HRV in the classical set-up.

Thus, in Example 4.2, HRV provides more information than non-standard regular variation. This is not always the case and sometimes non-standard regular variation is better suited to explaining the structure of a model.

EXAMPLE 4.3. [Resnick, 2007, Example 6.3] Suppose  $X$  is a standard Pareto random variable and  $\mathbf{Z} = (X, X^2)$ . Using the obvious non-standard scaling for the coordinates leads to

$$(11) \quad tP \left[ \left( \frac{X}{t}, \frac{X^2}{t^2} \right) \in \cdot \right] \xrightarrow{v} \nu_{(1)} \circ T^{-1}, \quad (\text{as } t \rightarrow \infty)$$

where  $T : (0, \infty] \rightarrow (0, \infty] \times (0, \infty]$  is defined by  $T(x) = (x, x^2)$  and  $\nu_{(1)}(dx) = x^{-2}dx$ ,  $x > 0$ . The limit measure concentrates on  $\{(x, x^2) : x > 0\}$ . Using the same normalization on both coordinates is not so revealing. With the heavier normalization, we have on  $\mathbb{M}^*([0, \infty)^2, [0, \infty)^2 \setminus \{\mathbf{0}\})$ ,

$$(12) \quad tP[\mathbf{Z}/t^2 \in \cdot] \xrightarrow{*} \epsilon_0 \times \nu_{(1)}, \quad (\text{as } t \rightarrow \infty)$$

and with the lighter normalization and  $x > 0$ ,  $y > 0$ ,

$$P[X/t > x, X^2/t > y] \rightarrow x^{-1}, \quad \forall y > 0,$$

so that

$$(13) \quad tP[\mathbf{Z}/t \in \cdot] \xrightarrow{v} \nu_{(1)} \times \epsilon_\infty, \quad \text{on } \mathbb{E}.$$

Neither (12) nor (13) approach the explanatory power of (11). Moreover, since our modeling excludes points at infinity, the convergence in (13) is not equivalent to any  $\mathbb{M}^*$ -convergence.

**5. Examples.** The definition of HRV given in this paper changes somewhat the definition of convergence but more importantly allows general cones compared with the existing notion of HRV [Resnick, 2002, Mitra and Resnick, 2010]; see Remark 3.7 and equation (8.2) of this paper. Here we provide examples to illustrate use and highlight subtleties. The examples give risk sets, whose probabilities can be approximated by using our general concept of HRV, and not the existing notion. See also [Mitra and Resnick, 2010].

**EXAMPLE 5.1. Diversification of risk:** Suppose, we invest in two financial instruments  $I_1$  and  $I_2$  and for a given time horizon future losses associated with each unit of  $I_1$  and  $I_2$  are  $\xi$  and  $\eta$  respectively. Let  $\mathbf{Z} = (\xi, \eta)$  be regularly varying on  $\mathbb{O} = \mathbb{D} = [0, \infty)^2 \setminus \{\mathbf{0}\}$  with limit measure  $\nu(\cdot)$ . We earn risk premia of  $\$l_1$  and  $\$l_2$  for investing in each unit of  $I_1$  and  $I_2$ .

Suppose we invest in  $a_1$  units of  $I_1$  and  $a_2$  units of  $I_2$ . A possible risk measure for this portfolio is  $P[a_1\xi > x, a_2\eta > y]$  for two large thresholds  $x$  and  $y$  and this risk measure quantifies tail-dependence of  $\xi$  and  $\eta$ . The greater the tail-dependence between  $\xi$  and  $\eta$ , the greater should be our reserve capital requirement so that we guard against the extreme situation that both investments go awry. For this risk measure, the best circumstance is if risk contagion is absent; that is,  $\xi$  and  $\eta$  are asymptotically independent so that the measure  $\nu(\cdot)$  is concentrated on the axes [Resnick, 2007, page 192] since then  $P[a_1\xi > x, a_2\eta > y]$  is estimated to be zero if HRV is absent

and even if HRV exists on the cone  $(0, \infty)^2$  according to Definition 3.6, the estimate of  $P[a_1\xi > x, a_2\eta > y]$  for large thresholds  $x$  and  $y$  should be small compared to the case where  $\xi$  and  $\eta$  are not asymptotically independent.

In place of asymptotic independence, suppose instead that  $\mathbf{Z}$  is regularly varying on  $\mathbb{D}$  as in Definition 3.1 with limit measure  $\nu(\cdot)$  and  $\nu(\cdot)$  has support  $\{(u, v) \in \mathbb{D} : 2u \leq v\} \cup \{(u, v) \in \mathbb{D} : u \geq 2v\}$  so that  $\nu(\cdot)$  puts zero mass on

$$\begin{aligned} \text{CONE} &:= \{(u, v) \in \mathbb{D} : 2u > v, 2v > u\} \\ (14) \quad &= \bigcup_{x>0, y>0} \{(u, v) : 2u - v > x, 2v - u > y\}. \end{aligned}$$

We can still build a portfolio of two financial instruments that have asymptotically independent risks as follows.

Define two new financial instruments  $W_1 = 2I_1 - I_2$  (buy two units of  $I_1$  and sell a unit of  $I_2$ , assuming such transactions are allowed) and  $W_2 = 2I_2 - I_1$ . The loss associated with  $W_1$  is  $L_{W_1} := 2\xi - \eta$  and the loss for  $W_2$  is  $L_{W_2} := 2\eta - \xi$ . We earn the same risk premia  $a_1l_1 + a_2l_2$  in the following two cases:

- (i) invest in  $a_1$  units of  $I_1$  and  $a_2$  units of  $I_2$ , or
- (ii) invest in  $c_1 = (2a_1 + a_2)/3$  units of  $W_1$  and  $c_2 = (a_1 + 2a_2)/3$  units of  $W_2$ .

A measure of risk of the portfolio in (i) is  $P[a_1\xi > x, a_2\eta > y]$  and since as  $t \rightarrow \infty$ ,

$$tP[a_1\xi/b(t) > x, a_2\eta/b(t) > y] \rightarrow \nu(((x, y), \infty)) > 0,$$

since asymptotic independence is absent, we expect the risk probability to be not too small. However, the risk measure for (ii) should be rather small as we now explain. Let  $T(u, v) = (2u - v, 2v - u)$  and the risk measure for (ii) is

$$P[c_1(2\xi - \eta) > x, c_2(2\eta - \xi) > y] = P[T(\xi, \eta) > (x/c_1, y/c_2)].$$

Since as  $t \rightarrow \infty$ ,

$$tP[T(\xi, \eta) > (b(t)x/c_1, b(t)y/c_2)] \rightarrow \nu\left\{T^{-1}\left(\left((x/c_1, y/c_2), \infty\right)\right)\right\}$$

and

$$T^{-1}\left(\left((x/c_1, y/c_2), \infty\right)\right) \subset \text{CONE},$$

the risk measure for (ii) is small and the losses are asymptotically independent. Hence, investment portfolio (ii) above achieves more diversification of risk than portfolio (i), and both earn the same amount of risk premium.

How do we construct a risk vector  $\mathbf{Z}$  whose distribution is regularly varying and whose limit measure  $\nu$  concentrates on  $\mathbb{D} \setminus \text{CONE}$ ? Suppose  $R_1, R_2, U_1, U_2, B$  are independent and  $R_1$  is Pareto(1),  $R_2$  is Pareto(2),  $U_1 \sim U((0, 1/3) \cup (2/3, 1))$ ,  $U_2 \sim U(1/3, 2/3)$  and  $B$  is Bernoulli with values 0, 1 with equal probability. Define

$$\mathbf{Z} = BR_1(U_1, 1 - U_1) + (1 - B)R_2(U_2, 1 - U_2).$$

Because of Proposition 3.3 applied with the  $L_1$  metric,  $\mathbf{Z}$  is regularly varying on  $\mathbb{D}$  with index 1 and angular measure concentrating on  $[0, 1/3] \cup [2/3, 1]$ . Hence the limit measure  $\nu$  concentrates on  $\mathbb{D} \setminus \text{CONE}$ . At scale  $t^{1/2}$ ,  $\mathbf{Z}$  is also regularly varying on  $\text{CONE}$  with uniform angular measure.

**EXAMPLE 5.2. Asymptotic full dependence:** For convenience, restrict this example to  $d = 2$ . HRV [Resnick, 2002] was designed to deal with asymptotic independence [Resnick, 2007, page 322] where  $\nu(\cdot)$  concentrates on the axes. For asymptotic full dependence, the limit measure  $\nu(\cdot)$  concentrates on the diagonal  $\text{DIAG} := \{(z, z) : z \geq 0\}$  [Resnick, 2007, page 195] and previous treatments could not deal with this or related degeneracies where the limit measure  $\nu(\cdot)$  concentrates on a finite number of rays other than the axes.

Consider the following example. Suppose,  $X_1, X_2, X_3$  are iid with common distribution Pareto(2). Let  $B$  be a Bernoulli random variable independent of  $\{X_i : i = 1, 2, 3\}$  and  $P[B = 0] = P[B = 1] = 1/2$ . Construct the random vector  $\mathbf{Z}$  as

$$\mathbf{Z} = (\xi, \eta) = B((X_1)^2, (X_1)^2) + (1 - B)(X_2, X_3)$$

and  $\mathbf{Z}$  is regularly varying on  $\mathbb{O} = \mathbb{D}$  with scaling function  $b(t) = t$  and limit measure  $\nu$  where,

$$\nu([\mathbf{0}, (u, v)]^c) = \frac{1}{2}(u \wedge v)^{-1} \quad (u, v) \in \mathbb{D}.$$

The measure  $\nu(\cdot)$  concentrates on  $\text{DIAG}$  and satisfies  $\nu(\{(u, v) \in \mathbb{D} : |u - v| > x\}) = 0$ , for  $x > 0$  and as a result, we estimate as zero risk probabilities like  $P(|\xi - \eta| > x)$  for large thresholds  $x$ . We gain more precision from HRV.

Define the cone

$$\mathbb{O}_1 = [0, \infty)^2 \setminus (\{\mathbf{0}\} \cup \text{DIAG}) = \{(u, v) \in \mathbb{D} : |u - v| > 0\}.$$

The distribution of  $\mathbf{Z}$  has HRV on  $\mathbb{O}_1$  with scaling function  $b_1(t) = t^{1/2}$  and limit measure

$$\nu_1(du, dv) = u^{-3}\epsilon_{((0, \infty) \times \{0\})}(du, dv) + u^{-3}\epsilon_{(\{0\} \times (0, \infty))}(du, dv), \quad (u, v) \in \mathbb{O}_1,$$

the measure that concentrates mass on the axes and which is restricted to  $\mathbb{O}_1$ . This measure results from the second summand in  $\mathbf{Z}$ , namely  $(1 - B)(X_2, X_3)$ ; the first summand  $B(X_1^2, X_1^2)$  contributes nothing to the limit due to the restriction to  $\mathbb{O}_1$ . So for instance, for  $(u, v) \in \mathbb{O}_1$ ,  $x > 0$ ,

$$\nu_1([\mathbf{0}, (u, v)]^c \cap \{(u, v) \in \mathbb{O}_1 : |u - v| > x\}) = \frac{1}{2} \left( (u \vee x)^{-2} + (v \vee x)^{-2} \right)$$

and for some large  $t > 0$ , letting  $u \downarrow 0$ ,  $v \downarrow 0$ , we see

$$P(|\xi - \eta| > x) \approx \frac{1}{t} \left( \frac{x}{b_1(t)} \right)^{-2}.$$

Statistical estimates of the risk region probability replace  $b_1(t)$  by a statistic as in Section 6.

The fact that  $\nu_1(\cdot)$  concentrates on the axes suggests seeking a further HRV property on a cone smaller than  $\mathbb{O}_1$ . If one needs risk probabilities of the form  $P(\xi - \eta > x, \eta > y)$  for large thresholds  $x$  and  $y$ , we seek HRV property, say,  $\mathbb{O}^2 = \{(u, v) \in \mathbb{O}_1 : u, v > 0\}$  or a sub-cone of  $\mathbb{O}^2$ .

As an example of why risk probabilities like  $P(\xi - \eta > x)$  arise, imagine investing in financial instruments  $I_1$  and  $I_2$  that have risks  $\xi$  and  $\eta$  per unit of investment and suppose these risks have asymptotic full dependence.

For any  $a_1, a_2, c > 0$ , asymptotic full dependence of  $\xi$  and  $\eta$  implies  $P(a_1\xi + a_2\eta > x)$  should be bigger than  $P(c(\xi - \eta) > x)$ , provided  $x$  is large. So, if  $l_1 > l_2$ , it is less risky to invest in the financial instrument  $I_1 - I_2$  rather than investing in both  $I_1$  and  $I_2$ . Obviously, investing in the financial instrument  $I_1 - I_2$  requires us to measure risks associated with this portfolio, which leads to the need to evaluate  $P(\xi - \eta > x)$  for large thresholds  $x$ .

In summary, this example shows how a more flexible definition of HRV possibly allows computation of risk probabilities in the presence of asymptotic full dependence.

**EXAMPLE 5.3. Infinite sequence of cones:** HRV was originally defined for  $d = 2$  and for two cones [Resnick, 2002] and then extended to a finite sequence of cones [Mitra and Resnick, 2010]. In this paper our definition of HRV allows the possibility that we progressively find HRV on an infinite sequence of cones. We exhibit an example for  $d = 2$  where this is indeed the case. An infinite sequence of cones may create problems for risk estimation which we discuss afterwards.

Suppose  $\{X_i, i \geq 1\}$  are iid random variables with common Pareto(1) distribution. Let  $\{Y_1, Y_2\}$  be iid with common Pareto(2) distribution and

suppose  $\{B_i, i \geq 1\}$  is an infinite sequence of random variables with  $P(B_i = 1) = 1 - P(B_i = 0) = 2^{-i}$  and  $\sum_{i=1}^{\infty} B_i = 1$ . (For instance, let  $T$  be the index of the first success in an iid sequence of Bernoulli trials and then set  $B_i = 1_{[T=i]}, i \geq 1$ .) Assume that  $\{X_i, i \geq 1\}$ ,  $\{Y_1, Y_2\}$  and  $\{B_i : i \geq 1\}$  are mutually independent. Define the random vector  $\mathbf{Z}$  as

$$\mathbf{Z} = (\xi, \eta) = B_1(Y_1, Y_2) + \sum_{i=1}^{\infty} B_{i+1} \left( (X_i)^{\frac{1}{2-2^{-(i-1)}}}, 2^{i-1}(X_i)^{\frac{1}{2-2^{-(i-1)}}} \right).$$

So,  $\mathbf{Z}$  has regular variation on the cone  $\mathbb{O} = \mathbb{O}_0 = \mathbb{D} = [0, \infty)^d \setminus \{\mathbf{0}\}$  with index of regular variation  $\alpha = 1$ , scaling function  $b(t) = t$ , limit measure  $\nu(\cdot)$  concentrating on the diagonal  $\text{DIAG} := \{(x, x) : x \in [0, \infty)\}$ . So, we remove the diagonal and find HRV on  $\mathbb{O}_1 = \mathbb{D} \setminus \text{DIAG}$  with  $b_1(t) = t^{2/3}$ ,  $\alpha_1 = 3/2$  and limit measure  $\nu_1$  concentrating on the ray  $\{(x, 2x) : x \geq 0\}$ . Progressively seeking HRV on successive cones, we find at the  $(i+1)$ -th step of our analysis,  $\mathbf{Z}$  has regular variation on the cone

$$(15) \quad \mathbb{O}_i = \mathbb{D} \setminus \left[ \bigcup_{j=1}^i \{(x, 2^{j-1}x) : x \in [0, \infty)\} \right]$$

with the limit measure  $\nu_i(\cdot)$  and the index of regular variation  $\alpha_i = 2 - 2^{-i}$ . The limit measure  $\nu_i(\cdot)$  concentrates on  $\{(x, 2^i x) : x \in [0, \infty)\}$ .

Selection of cones must be guided by the type of risk probability needed. Consider trying to estimate  $P(\xi - \eta > x)$  for large thresholds  $x$  using the cones  $\mathbb{O}_i, i \geq 0$  given in (15). At the  $(i+1)$ st stage, using cone  $\mathbb{O}_i$ , the limit measure  $\nu_i(\cdot)$  puts zero mass on the cone  $\{(u, v) : u > v\}$ . So, even after a million HRV steps, we will estimate  $P(\xi - \eta > x)$  for large thresholds  $x$  as zero, which is clearly wrong due to the definition of  $\mathbf{Z}$  since

$$P[\xi - \eta > x] \geq P[B_1 = 1]P[Y_1 - Y_2 > x] > 0.$$

An alternative procedure seeks regular variation on the cone  $\{(u, v) : u > v\}$  and this leads to somewhat more reasonable estimates of  $P(\xi - \eta > x)$  for large thresholds  $x$  since in this case the regular variation with the Pareto(2) variables is captured.

The moral of the story is that the choice of sequence of cones when defining HRV should be guided by the kind of risk sets considered. For example, if we are only interested in joint tail probabilities, a possible choice of sequence of cones is  $\mathbb{D} = \mathbb{D}^{(1)} \supset \mathbb{D}^{(2)} \supset \dots \supset \mathbb{D}^{(d)}$ , where

$$\mathbb{D}^{(l)} = \{\mathbf{x} \in [0, \infty)^d : x^{(l)} > 0\},$$

and recall  $x^{(l)}$  is the  $l$ -th largest component of  $\mathbf{x}$ . This special case is discussed in Mitra and Resnick [2010].

EXAMPLE 5.4. **The CEV model and mass on the lines through  $\infty$ .** If we consider the CEV model on  $\mathbb{E}_\square$ , there can exist two different limits in (9) under two different normalizations. This problem disappears if we restrict convergence to  $\mathbb{D}_\square$ .

Suppose  $Y$  is Pareto(1) and  $B$  is a Bernoulli random variable with  $P[B = 1] = P[B = 0] = 1/2$ . Define

$$\mathbf{Z} = (\xi, \eta) = B(Y, Y) + (1 - B)(\sqrt{Y}, Y).$$

Then the following convergences both hold in  $\mathbb{E}_\square$ : for  $x \geq 0, y > 0$ ,

$$\begin{aligned} \nu_1([0, x] \times (y, \infty]) &:= \lim_{t \rightarrow \infty} tP \left[ \left( \frac{\xi}{t}, \frac{\eta}{t} \right) \in [0, x] \times (y, \infty] \right] \\ &= \frac{1}{2} \left( \frac{1}{y} - \frac{1}{x} \right)_+ + \frac{1}{2y}. \end{aligned} \tag{16}$$

$$\begin{aligned} \nu_2([0, x] \times (y, \infty]) &:= \lim_{t \rightarrow \infty} tP \left[ \left( \frac{\xi}{\sqrt{t}}, \frac{\eta}{t} \right) \in [0, x] \times (y, \infty] \right] \\ &= \frac{1}{2} \left( \frac{1}{y} - \frac{1}{x^2} \right)_+. \end{aligned} \tag{17}$$

So  $\mathbf{Z}$  follows a CEV model on  $\mathbb{E}_\square$  with two different scalings. Note that  $\nu_1$  does not put any mass on lines through  $\infty$ , but  $\nu_2$  does:  $\nu_2(\{\infty\} \times (y, \infty]) = 1/2y$ . If we restrict convergence to  $\mathbb{D}_\square$ , limits are unique and  $\mathbf{Z}$  is regularly varying on  $\mathbb{D}_\square$  with limit measure  $\nu_1$  given by (16) restricted to  $\mathbb{D}_\square$ .

**6. Estimating the spectral measure and its support.** We have defined regular variation on a big cone  $\mathbb{O}_0 \subset \mathbb{R}^d$  along with hidden regular variation in a nested sequence of subcones  $\mathbb{O}_0 \supset \mathbb{O}_1 \supset \mathbb{O}_2 \supset \dots$ . We now propose strategies for deciding whether HRV is consistent with a given data set and, if so, how to estimate probabilities of sets pertaining to joint occurrence of extreme or high values. We proceed as follows:

1. Specify a fixed finite sequence of cones pertinent to the problem, and seek HRV sequentially on these cones. We discuss this in Section 6.1 which follows ideas proposed in [Mitra and Resnick, 2010, Section 3].
2. If the sequence of cones is not clear, proceed by estimating the support of the limit measure at each step, removing it, and seeking HRV on the complement of the support. Then the hidden limit measure is estimated using semi-parametric techniques similar to 6.1.

6.1. *Specified sequence of cones.* Suppose  $\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_n$  are iid random vectors in  $\mathbb{C} \subset [0, \infty)^d$  whose common distribution has a regularly varying

tail on  $\mathbb{O}$  according to Definition 3.1 with normalizing function  $b(\cdot)$  and limit measure  $\nu(\cdot)$ . Also assume that we have a specified sequence of cones  $\mathbb{O} = \mathbb{O}_0 \supset \mathbb{O}_1 \supset \mathbb{O}_2 \supset \dots$  where we seek regular variation. Such a sequence of cones is known and fixed.

We provide an estimate for the limit measure of regular variation on  $\mathbb{O}$  and the same method can be applied to find limit measures for hidden regular variation on the subcones.

Now, according to Corollary 3.4 regular variation on  $\mathbb{O}$  as above is equivalent to assuming  $P[d(\mathbf{Z}, \mathbb{F}) > x]$  is regularly varying at  $\infty$  with some exponent  $\alpha > 0$  and normalizing function  $b(t)$  which we take as  $b(t) := F_d^{\leftarrow}(1 - 1/t)$  where  $d(\mathbf{Z}, \mathbb{F})$  has distribution function  $F_d$  and

$$(18) \quad P \left[ \frac{\mathbf{Z}}{d(\mathbf{Z}, \mathbb{F})} \in \cdot \mid d(\mathbf{Z}, \mathbb{F}) > t \right] \Rightarrow S_{\mathbb{O}}(\cdot) \quad (t \rightarrow \infty),$$

in  $P(\partial\mathcal{N}_{\mathbb{O}})$ , the class of all probability measures on  $\partial\mathcal{N}_{\mathbb{O}} = \{\mathbf{x} \in \mathbb{O} : d(\mathbf{x}, \mathbb{F}) = 1\}$ . Thus we estimate  $\nu$  by estimating  $\alpha$  and  $S_{\mathbb{O}}$  separately. Considering  $d_1^{\mathbb{F}} := d(\mathbf{Z}_1, \mathbb{F}), \dots, d_n^{\mathbb{F}} = d(\mathbf{Z}_n, \mathbb{F})$  as iid samples from a regularly varying distribution on  $(0, \infty)$ , the exponent  $\alpha$  can be estimated using the Hill, Pickands or QQ estimators [Resnick, 2007].

Here is an outline of how to obtain an empirical estimator of  $S_{\mathbb{O}}$  following [Resnick, 2007].

**PROPOSITION 6.1.** *Assume the common distribution of the iid random vectors  $\mathbf{Z}_1, \dots, \mathbf{Z}_n$  satisfies Definition 3.1 and (4). As  $n \rightarrow \infty$ ,  $k \rightarrow \infty$ ,  $n/k \rightarrow \infty$ , we have in  $P(\partial\mathcal{N}_{\mathbb{O}})$ ,*

$$(19) \quad S_n(\cdot) := \frac{\sum_{i=1}^n \epsilon_{(d_i^{\mathbb{F}}/b(n/k), \mathbf{Z}_i/d_i^{\mathbb{F}})}((1, \infty) \times \cdot)}{\sum_{i=1}^n \epsilon_{d_i^{\mathbb{F}}/b(n/k)}(1, \infty)} \Rightarrow S_{\mathbb{O}}(\cdot).$$

**PROOF.** Since  $\{d_i^{\mathbb{F}}, 1 \leq i \leq n\}$  are iid regularly varying random variables from some distribution  $F_d$  on  $(0, \infty)$  with norming function  $b(t)$  which can be chosen to be  $b(t) = F_d^{\leftarrow}(1 - 1/t)$ , by Theorem 8.1. Thus for  $x > 0$ ,  $\frac{n}{k} P[d_i^{\mathbb{F}}/b(n/k) > x] \rightarrow cx^{-\alpha}$ , and from Resnick [2007, page 139] this is equivalent to  $\frac{1}{k} \sum_{i=1}^n \epsilon_{d_i^{\mathbb{F}}/b(n/k)}(1, \infty) \Rightarrow c$ , and to prove (19), it suffices to show in  $M_+(\partial\mathcal{N}_{\mathbb{O}})$ ,

$$(20) \quad \frac{1}{k} \sum_{i=1}^n \epsilon_{\mathbf{Z}_i/d_i^{\mathbb{F}}}(\cdot) 1_{[d_i^{\mathbb{F}}/b(n/k) > 1]} \Rightarrow S_{\mathbb{O}}(\cdot).$$

The counting function on the left of (20) only counts  $\mathbf{Z}_i/d_i^{\mathbb{F}}$  such that  $d_i^{\mathbb{F}}/b(n/k) > 1$ . The distribution of such random elements is  $P[\mathbf{Z}_i/d_i^{\mathbb{F}} \in \cdot | d_i^{\mathbb{F}}/b(n/k) > 1]$ , [Resnick, 2008, page 212] and (18) holds. Using Resnick [2007, Theorem 5.3ii, page 139] and the style of argument in Resnick [2008, page 213], we get (20).  $\square$

The estimator  $S_n$  of  $S_{\mathbb{O}}$  in Proposition 6.1 relies on  $b(t)$  which is typically unknown but  $b(n/k)$  can be estimated. Order  $d_1^{\mathbb{F}}, \dots, d_n^{\mathbb{F}}$  as  $d_{(1)}^{\mathbb{F}} \geq \dots \geq d_{(n)}^{\mathbb{F}}$  and  $d_{(k+1)}^{\mathbb{F}}/b(n/k) \xrightarrow{P} 1$  so  $d_{(k+1)}^{\mathbb{F}}$  is a consistent estimator of  $b(n/k)$  as  $n \rightarrow \infty, k \rightarrow \infty, n/k \rightarrow \infty$  [Resnick, 2007, page 81]. Hence we replace  $b(n/k)$  by  $d_{(k+1)}^{\mathbb{F}}$  and propose the estimator  $\hat{S}_n$  for  $S_{\mathbb{O}}$  in  $S_n$  as follows:

$$\hat{S}_n(\cdot) := \frac{\sum_{i=1}^n \epsilon_{\{d_i^{\mathbb{F}}/d_{(k+1)}^{\mathbb{F}}, \mathbf{Z}_i/d_i^{\mathbb{F}}\}}((1, \infty) \times \cdot)}{\sum_{i=1}^n \epsilon_{\{d_i^{\mathbb{F}}/d_{(k+1)}^{\mathbb{F}}\}}(1, \infty)} = \frac{1}{k} \sum_{i=1}^n 1_{[d_i^{\mathbb{F}}/d_{(k+1)}^{\mathbb{F}} > 1]} \epsilon_{\mathbf{Z}_i/d_i^{\mathbb{F}}}(\cdot).$$

PROPOSITION 6.2. *As  $n \rightarrow \infty, k \rightarrow \infty, n/k \rightarrow \infty$ ,  $\hat{S}_n \Rightarrow S_{\mathbb{O}}$  in  $P(\partial\mathcal{N}_{\mathbb{O}})$ .*

PROOF. The proof is a consequence of the continuous mapping theorem and Proposition 6.1. For details see, for instance, [de Haan and Resnick, 1993].  $\square$

Thus when  $\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_n$  are iid random vectors in  $\mathbb{C} \subset [0, \infty)^d$  which have a regularly varying distribution on  $\mathbb{O}$ , we can estimate both  $\alpha$  and the spectral measure  $S_{\mathbb{O}}$ . If we have a specified finite sequence of cones  $\mathbb{O} := \mathbb{O}_0 \supset \mathbb{O}_1 \supset \mathbb{O}_2 \supset \dots \supset \mathbb{O}^m$ , then we sequentially estimate the limit measure by separately estimating the spectral measure and the index  $\alpha_i$ .

6.2. *Support estimation.* Suppose  $\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_n$  are iid random vectors in  $\mathbb{C} \subset [0, \infty)^d$  whose common distribution is regularly varying on  $\mathbb{O}$  according to Definition 3.1 with normalizing function  $b(\cdot)$  and limit measure  $\nu(\cdot)$ . Without a sequence of cones where hidden regular variation can be sought, the task of exploring for appropriate cones where HRV may exist is challenging. One clear strategy is to identify the support of  $\nu$ , which we call  $\text{supp}(\nu)$ , and then seek HRV on the complement of the support. Since

$$(21) \quad \text{supp}(\nu) = \{t \cdot \text{supp}(S_{\mathbb{O}}), t \geq 0\}.$$

it suffices to determine the support of  $S_{\mathbb{O}}$ .

We propose estimating the support of the spectral measure  $S_{\mathbb{O}}$  with a point cloud; that is, a discrete random closed set.

PROPOSITION 6.3. *Suppose  $\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_n$  are iid random vectors in  $\mathbb{C} \subset [0, \infty)^d$  whose common distribution is regularly varying on  $\mathbb{O}$  with normalizing function  $b(\cdot)$  and limit measure  $\nu(\cdot)$ . As  $n \rightarrow \infty, k \rightarrow \infty, n/k \rightarrow \infty$ ,*

$$(22) \quad \text{supp}_{k,n} = \left\{ \frac{\mathbf{Z}_i}{d_i^{\mathbb{F}}} : d_i^{\mathbb{F}} > d_{(k+1)}^{\mathbb{F}}, i = 1, \dots, n \right\} \Rightarrow \text{supp}(S_{\mathbb{O}}).$$

Convergence in (22) occurs in the space of closed sets under the Fell topology or the space of compact sets in the Hausdorff topology [Molchanov, 2005].

PROOF. To show (22), it suffices from [Molchanov, 2005, Proposition 6.10, page 87] to show for any  $h \in C_K^+(\partial\mathfrak{N}_{\mathbb{O}})$  that

$$(23) \quad E\left(\sup_i \left\{ h\left(\frac{\mathbf{Z}_i}{d_i^{\mathbb{F}}}\right) : d_i^{\mathbb{F}} > d_{(k+1)}^{\mathbb{F}}, 1 \leq i \leq n \right\}\right) \rightarrow \sup_x \{h(x) : x \in \text{supp}(S_{\mathbb{O}})\}.$$

From Proposition 6.2,

$$\hat{S}_n(\cdot) := \frac{1}{k} \sum_{i=1}^n 1_{[d_i^{\mathbb{F}}/d_{(k+1)}^{\mathbb{F}}]} \epsilon_{\mathbf{Z}_i/d_i^{\mathbb{F}}}(\cdot) \Rightarrow S_{\mathbb{O}},$$

and from the continuous mapping theorem, for any  $h \in C_K^+(\partial\mathfrak{N}_{\mathbb{O}})$ , we get in  $P(\mathbb{R})$ , the class of probability measures on  $\mathbb{R}$ ,

$$(24) \quad \hat{S}_n \circ h^{-1} = \frac{1}{k} \sum_{i=1}^n 1_{[d_i^{\mathbb{F}}/d_{(k+1)}^{\mathbb{F}}]} \epsilon_{h(\mathbf{Z}_i/d_i^{\mathbb{F}})}(\cdot) \Rightarrow S_{\mathbb{O}} \circ h^{-1}.$$

If  $F_n, n \geq 0$  are probability measures on  $\mathbb{R}$  with bounded support and  $F_n \Rightarrow F_0$  then

$$(25) \quad x_{F_n} := \sup_x \{x : F_n(x) < 1\} \rightarrow \sup_x \{x : F(x) < 1\} =: x_F.$$

Applying this remark to (24) and using the continuous mapping theorem yields as  $n \rightarrow \infty, k \rightarrow \infty, n/k \rightarrow \infty$ ,

$$(26) \quad \sup_i \left\{ h\left(\frac{\mathbf{Z}_i}{d_i^{\mathbb{F}}}\right) : d_i^{\mathbb{F}} > d_{(k+1)}^{\mathbb{F}}, 1 \leq i \leq n \right\} \Rightarrow \sup_x \{h(x) : x \in \text{supp}(S_{\mathbb{O}})\}.$$

Since  $h \in C_K^+(\partial\mathfrak{N}_{\mathbb{O}})$  is always bounded above, use dominated convergence applied to convergence in distribution to get the desired (23).  $\square$

Proposition 6.3 provides an estimate of  $\text{supp}(S_{\mathbb{O}})$  and hence  $\text{supp}(\nu)$ . In principle, we can remove the estimated support from  $\mathbb{O}$  and look for hidden regular variation in the complement. How well this works in practice remains to be seen. For one thing, the estimated support set of  $S_{\mathbb{O}}$  is always discrete meaning that the estimated support of  $\nu$  is a finite set of rays. With a large data set, we might be able to get a fair idea about the support of the distribution and where to look for further hidden regular variation. If there were reason to believe or hope that the support of  $S_{\mathbb{O}}$  were convex, our estimation procedure could be modified by taking the convex hull of the points in (22).

**7. Conclusion.** Our treatment of regular variation on cones which is determined by the support of the limit measures unifies under one theoretical umbrella several related concepts: asymptotic independence, asymptotic full dependence, and the conditional extreme value model. Our approach highlights the structural similarities of these concepts while making plain in what ways the cases differ. Furthermore, the notion of  $\mathbb{M}^*$ -convergence introduced in Section 2.2 provides a tool to deal with the generalized notion of regular variation given here. Generalizing this notion of convergence and analyzing its properties admits potential for further research.

It is always an ambitious undertaking to statistically identify lower order behavior and this project has not attempted data analysis or tested the feasibility of the statistical methods discussed in Section 6. It is clear further work is required, particularly for the case where the support of the limit measures must be identified from data. One can imagine that for high dimensional data whose dimension is of the order of hundreds, sophistication is required to pursue successive cones where regular variation exists.

## 8. Appendix.

8.1. *Regular variation on  $\mathbb{E} := [0, \infty]^d \setminus \{\mathbf{0}\}$  vs  $\mathbb{D} = [0, \infty)^d \setminus \{\mathbf{0}\}$ .* We verify that the traditional notion of multivariate regular variation given in (1) on  $\mathbb{E}$  is equivalent to Definition 3.1 if we choose  $\mathbb{C} = [0, \infty)^d$  and  $\mathbb{F} = \{\mathbf{0}\}$ . This yields  $\mathbb{O} = \mathbb{D}$ .

**THEOREM 8.1.** *Regular variation on  $\mathbb{D}$  according to Definition 3.1 is equivalent to the traditional notion of multivariate regular variation given in (1) and the limit measures  $\nu(\cdot)$  of (4) and  $\mu(\cdot)$  of (1) are equal on  $\mathbb{D}$ . Moreover,  $\mu(\cdot)$  puts zero measure on  $\mathbb{E} \setminus \mathbb{D}$ .*

**PROOF.** First we show that the standard notion of multivariate regular variation on  $\mathbb{E}$  given in (1) implies (4) in  $\mathbb{M}^*(\mathbb{D})$ . Let  $\nu(\cdot)$  be a measure

on  $\mathbb{D}$  such that  $\nu(\cdot) = \mu(\cdot)$ . From Resnick [2007, page 176], we get that  $\mu(\mathbb{E} \setminus \mathbb{D}) = 0$ . So, since  $\mu(\cdot) \neq 0$  and non-degenerate,  $\nu(\cdot) \neq 0$  and non-degenerate.

For  $B \subset \mathbb{D}$ , note that  $\partial B = \bar{B} \setminus B^\circ$  is defined with respect to the relative topology on  $\mathbb{D}$  and hence  $\partial B \subset \mathbb{D}$ . Thus,  $\nu(\partial B) = 0$  implies  $\mu(\partial B) = 0$ . Also, since  $[0, \infty]^d$  is a compact space, any set  $B \subset \mathbb{D}$  bounded away from  $\{\mathbf{0}\}$  is a relatively compact set in  $\mathbb{E}$  [Resnick, 2007, page 171, Proposition 6.1]. Therefore, by definition,  $\nu(\cdot) \in \mathbb{M}^*([0, \infty)^d, \mathbb{D})$  and by (1), for any  $B \subset \mathbb{D}$  bounded away from  $\{\mathbf{0}\}$  and  $\nu(\partial B) = 0$ ,

$$tP \left[ \frac{\mathbf{Z}}{b(t)} \in B \right] \rightarrow \mu(B) = \nu(B).$$

So, (4) holds with  $\mathbb{C} = [0, \infty)^d$ ,  $\mathbb{O} = \mathbb{D}$ , and  $b$  is the same as in (1) and  $\nu$  is the restriction of  $\mu$  to  $\mathbb{D}$ .

Conversely, we show that Definition 3.1 and (4) with  $\mathbb{O} = \mathbb{D}$  implies the traditional notion of multivariate regular variation on  $\mathbb{E}$  in (1). Define a measure  $\mu(\cdot)$  on  $\mathbb{E}$  as  $\mu(\cdot) = \nu(\cdot \cap \mathbb{D})$ . A relatively compact set  $B$  of  $\mathbb{E}$  must be bounded away from  $\{\mathbf{0}\}$  [Resnick, 2007, page 171, Proposition 6.1]. So, from definition of  $\mu(\cdot)$ , it is Radon. Note that  $\partial B = \bar{B} \setminus B^\circ$  is defined with respect to the topology on  $\mathbb{E}$ , but  $\partial(B \cap \mathbb{D})$  is defined with respect to the relative topology on  $\mathbb{D}$ . Also from the definition of  $\mu(\cdot)$ ,  $\mu(\partial B) = 0$  implies  $\nu(\partial(B \cap \mathbb{D})) = \nu(\partial B \cap \mathbb{D}) = \mu(\partial B) = 0$ . Therefore, from (4), for any relatively compact set  $B$  of  $\mathbb{E}$  such that  $\mu(\partial B) = 0$ , as  $t \rightarrow \infty$ ,

$$tP \left[ \frac{\mathbf{Z}}{b(t)} \in B \right] = tP \left[ \frac{\mathbf{Z}}{b(t)} \in B \cap \mathbb{D} \right] \rightarrow \nu(B \cap \mathbb{D}) = \mu(B).$$

The first equality above holds since  $\mathbf{Z} \in [0, \infty)^d$ . Hence, vague convergence in (1) holds with the same  $b$  as in (4) and with  $\mu$  as the extension of  $\nu$  from  $\mathbb{D}$  to  $\mathbb{E}$ .  $\square$

Regular variation on  $\mathbb{D}$  can also be expressed in terms of the polar coordinate transformation. As at the beginning of Section 8.1, set  $\mathbb{C} = [0, \infty)^d$ ,  $\mathbb{F} = \{\mathbf{0}\}$  and  $\mathbb{O} = \mathbb{D}$ .

**PROPOSITION 8.2.** *Regular variation on  $\mathbb{O}$  as given in Definition 3.1 is equivalent to the following condition:*

$$(27) \quad tP \left[ \left( \frac{\|\mathbf{Z}\|}{b(t)}, \frac{\mathbf{Z}}{\|\mathbf{Z}\|} \right) \in A \right] \rightarrow \nu_{(\alpha)} \times S_{\mathbb{O}}(A)$$

for all measurable  $A \subset (0, \infty) \times \partial\mathfrak{N}_\mathbb{O}$  such that  $\overline{A^1} \cap \{0\} = \emptyset$  and  $\nu \circ h^{-1}(\partial A) = 0$ , where  $A^1$  is the projection of  $A$  on its first coordinate,  $h(\cdot)$  is a function defined by  $h : \mathbf{x} \mapsto \left( \|\mathbf{x}\|, \frac{\mathbf{x}}{\|\mathbf{x}\|} \right)$ ,  $\partial\mathfrak{N}_\mathbb{O} = \{\mathbf{x} \in \mathbb{O} : \|\mathbf{x}\| = 1\}$ ,  $S_\mathbb{O}(\cdot)$  is a probability measure on  $\partial\mathfrak{N}_\mathbb{O}$  and  $\nu_{(\alpha)}(\cdot)$  is a Pareto measure given by  $\nu_{(\alpha)}((x, \infty)) = x^{-\alpha}$  for  $x > 0$ . The probability measure  $S_\mathbb{O}(\cdot)$  is called the spectral measure and is related to  $\nu(\cdot)$  by the relation

$$(28) \quad S_\mathbb{O}(\Lambda) = \nu \left( \left\{ \mathbf{x} \in \mathbb{O} : \|\mathbf{x}\| \geq 1, \frac{\mathbf{x}}{\|\mathbf{x}\|} \in \Lambda \right\} \right).$$

PROOF. This is a special case of Proposition 3.3.  $\square$

8.2. *Regular variation on  $\mathbb{E}^{(l)} = [0, \infty)^d \setminus \{\mathbf{x} \in \mathbb{E} : x^{(l)} > 0\}$  vs  $\mathbb{D}^{(l)} = [0, \infty)^d \setminus \{\mathbf{x} \in [0, \infty)^d : x^{(l)} > 0\}$ .* Recall  $x^{(l)}$  is the  $l$ -th largest component of  $\mathbf{x}$ ,  $l = 1, 2, \dots, d$ . Hidden regular variation using  $\mathbb{E}^{(l)}$  is considered in Mitra and Resnick [2010]. Unlike the situation in subsection 8.1, here limit measures can put mass on  $\mathbb{E}^{(l)} \setminus \mathbb{D}^{(l)}$  as found in Mitra and Resnick [2010]. We compare regular variation in  $\mathbb{E}^{(l)}$  using the traditional vague convergence definition in which the vague convergence in (1) is assumed to hold in  $\mathbb{M}_+(\mathbb{E}^{(l)})$  with regular variation given in (4) in  $\mathbb{M}^*(\mathbb{C}, \mathbb{O})$  where  $\mathbb{C} = [0, \infty)^d$ ,  $\mathbb{F} = \{\mathbf{x} \in [0, \infty)^d : x^{(l)} = 0\}$  and  $\mathbb{O} = \mathbb{D}^{(l)} = \mathbb{C} \setminus \mathbb{F}$ .

**THEOREM 8.3.** *Regular variation on  $\mathbb{M}^*(\mathbb{D}^{(l)})$  is equivalent to the traditional vague convergence notion of regular variation in  $\mathbb{M}_+(\mathbb{E}^{(l)})$  if the limit measure  $\mu(\cdot)$  given in the  $\mathbb{M}_+(\mathbb{E}^{(l)})$  analogue of (1) does not give any mass to the set  $\mathbb{E}^{(l)} \setminus \mathbb{D}^{(l)}$ . In this case, the limit measures  $\nu(\cdot)$  of (4) and  $\mu(\cdot)$  of (1) are equal on  $\mathbb{D}^{(l)}$ .*

PROOF. Suppose, for a random vector  $\mathbf{Z}$ , there exist a function  $b^{(l)}(t) \uparrow \infty$  and a non-negative non-degenerate Radon measure  $\mu^{(l)}(\cdot) \neq 0$  on  $\mathbb{E}^{(l)}$ , such that in  $\mathbb{M}_+(\mathbb{E}^{(l)})$

$$(29) \quad tP \left[ \frac{\mathbf{Z}}{b^{(l)}(t)} \in \cdot \right] \xrightarrow{v} \mu^{(l)}(\cdot).$$

and the limit measure  $\mu^{(l)}(\cdot)$  does not give any mass to  $\mathbb{E}^{(l)} \setminus \mathbb{D}^{(l)}$ . Define a measure  $\chi(\cdot)$  on  $\mathbb{D}^{(l)}$  as  $\chi(\cdot) = \mu^{(l)}(\cdot)$ . Since  $\mu^{(l)}(\cdot) \neq 0$  is non-negative, non-degenerate and  $\mu^{(l)}(\mathbb{E}^{(l)} \setminus \mathbb{D}^{(l)}) = 0$ , the measure  $\chi(\cdot) \neq 0$  is non-negative and non-degenerate. The subsets of  $\mathbb{D}^{(l)}$  bounded away from  $(\mathbb{D}^{(l)})^c = [0, \infty)^d \setminus \mathbb{D}^{(l)}$  are relatively compact in  $\mathbb{E}^{(l)}$ . Therefore, using the fact that  $\mu^{(l)}(\cdot)$  is Radon and the definition of  $\chi(\cdot)$ , it follows that  $\chi(\cdot)$  gives finite measure to

sets bounded away from  $(\mathbb{D}^{(l)})^c$ . From the definition of  $\mathbb{M}^*$ -convergence, it follows that  $\mathbf{Z}$  satisfies (4) with the scaling function  $b(\cdot) = b^{(l)}(\cdot)$  and the limit measure  $\nu(\cdot) = \chi(\cdot)$ .

Conversely, suppose a random vector  $\mathbf{Z}$  satisfies (4) in  $\mathbb{M}^*(\mathbb{C}, \mathbb{O})$  with  $\mathbb{C} = [0, \infty)^d$  and  $\mathbb{O} = \mathbb{D}^{(l)}$ . Define a measure  $\mu(\cdot)$  on  $\mathbb{E}^{(l)}$  as  $\mu(\cdot) = \nu(\cdot \cap \mathbb{D}^{(l)})$ . Since  $\nu(\cdot) \neq 0$  and is non-negative and non-degenerate, so is  $\mu(\cdot)$ . A subset of  $\mathbb{E}^{(l)}$  is relatively compact in  $\mathbb{E}^{(l)}$  iff it is bounded away from  $\{\mathbf{x} \in [0, \infty]^d : x^{(l)} = 0\}$ . Since  $\nu(\cdot)$  gives finite mass to sets bounded away from  $(\mathbb{D}^{(l)})^c$ , from the definition of  $\mu(\cdot)$ , it follows that  $\mu(\cdot)$  is a Radon measure. From the description of the compact sets in  $\mathbb{E}^{(l)}$ , it follows that  $\mathbf{Z}$  also satisfies (29) with the scaling function  $b^{(l)}(\cdot) = b(\cdot)$  and the limit measure  $\mu^{(l)}(\cdot) = \mu(\cdot)$  [Resnick, 2007, page 52, Theorem 3.2].  $\square$

The set  $\mathbb{E}^{(l)} \setminus \mathbb{D}^{(l)} = \{\mathbf{x} \in \mathbb{E}^{(l)} : \|\mathbf{x}\| = \infty\}$  is the union of the lines through  $\infty$ . We emphasize that there exist examples of random vectors  $\mathbf{Z}$  which satisfy (29) and the limit measure  $\mu^{(l)}(\cdot)$  gives positive measure on the set  $\mathbb{E}^{(l)} \setminus \mathbb{D}^{(l)}$  [Mitra and Resnick, 2010].

8.3. *Regular variation on  $\mathbb{E}_\square = [0, \infty] \times (0, \infty)$  vs  $\mathbb{D}_\square = [0, \infty) \times (0, \infty)$ .*  
Recall CEV model from Section 4.1.

PROPOSITION 8.4. *The following are equivalent:*

(i)  $\mathbf{Z} \in \text{CEV}(b_1, b_2)$  with limit measure  $\mu(\cdot)$  and  $b_1 \sim b_2$  with

$$(30) \quad \mu([0, \infty] \times \{\infty\} \cup \{\infty\} \times (0, \infty]) = 0.$$

(ii)  $\mathbf{Z}$  is regularly varying on  $\mathbb{D}_\square$  according to (4) with normalizing function  $b_1$  and limit measure  $\nu$  which does not concentrate on  $\{0\} \times (0, \infty)$ .

Also, if either of (i) or (ii) holds then  $\mu(\cdot) = \nu(\cdot)$  on  $\mathbb{D}_\square$ .

PROOF. (i) implies (ii): Since  $b_1 \sim b_2$ , (9) implies  $\mathbf{Z} \in \text{CEV}(b_1, b_1)$ . Now, (9) implies that for all relatively compact Borel sets  $B$  in  $\mathbb{D}_\square \subset \mathbb{E}_\square$  with  $\mu(\partial B) = 0$ ,

$$tP \left[ \frac{\mathbf{Z}}{b_1(t)} \in B \right] \rightarrow \mu(B)$$

as  $t \rightarrow \infty$ . Clearly  $B$  is bounded away from  $\mathbb{F}$ . Also  $\mu$  is non-null and satisfies (30). Thus  $\nu(\cdot) = \mu(\cdot)|_{\mathbb{D}_\square}$  is non-negative and non-degenerate on  $\mathbb{D}_\square$ . Hence  $\mathbf{Z}$  is regularly varying on  $\mathbb{D}_\square$  with limit measure  $\nu$ . The non-degeneracy condition (a) for the CEV model implies that  $\mu$  cannot concentrate on  $\{0\} \times (0, \infty)$ . Conversely, if (ii) implies (i) extend  $\nu$  to a measure  $\mu$  on  $\mathbb{E}_\square$  which satisfies (30).  $\square$

REMARK 8.5. We can drop the condition that  $\mu$  does not concentrate on  $\{0\} \times (0, \infty)$  in statement (ii) of Proposition 8.4, if we drop condition (a) from Definition 4.1 of the CEV model.

### References.

- P. Billingsley. *Convergence of Probability Measures*. John Wiley & Sons Inc., New York, second edition, 1999. ISBN 0-471-19745-9. A Wiley-Interscience Publication.
- N.H. Bingham, C.M. Goldie, and J.L. Teugels. *Regular Variation*. Cambridge University Press, 1987.
- J.-J. Cai, J.H.J. Einmahl, and L. de Haan. Estimation of extreme risk regions under multivariate regular variation. *Annals of Statistics*, 39(3):1803–1826, 2011.
- B. Das and S.I. Resnick. Conditioning on an extreme component: Model consistency with regular variation on cones. *Bernoulli*, 17(1):226–252, 2011.
- L. de Haan and A. Ferreira. *Extreme Value Theory: An Introduction*. Springer-Verlag, New York, 2006.
- L. de Haan and S.I. Resnick. Estimating the limit distribution of multivariate extremes. *Stochastic Models*, 9(2):275–309, 1993. ISSN 0882-0287.
- J.E. Heffernan and S.I. Resnick. Hidden regular variation and the rank transform. *Adv. Appl. Prob.*, 37(2):393–414, 2005.
- J.E. Heffernan and S.I. Resnick. Limit laws for random vectors with an extreme component. *Annals of Applied Probability*, 17(2):537–571, 2007.
- J.E. Heffernan and J.A. Tawn. A conditional approach for multivariate extreme values (with discussion). *Journal of the Royal Statistical Society, Series B*, 66(3):497–546, 2004.
- H. Hult and F. Lindskog. Regular variation for measures on metric spaces. *Publ. Inst. Math. (Beograd) (N.S.)*, 80(94):121–140, 2006. ISSN 0350-1302.
- H. Joe and H. Li. Tail risk of multivariate regular variation. *Methodology and Computing in Applied Probability*, page in press, 2010.
- O. Kallenberg. *Random Measures*. Akademie-Verlag, Berlin, third edition, 1983. ISBN 0-12-394960-2.
- A.W. Ledford and J.A. Tawn. Statistics for near independence in multivariate extreme values. *Biometrika*, 83(1):169–187, 1996. ISSN 0006-3444.
- A.W. Ledford and J.A. Tawn. Concomitant tail behaviour for extremes. *Adv. in Appl. Probab.*, 30(1):197–215, 1998. ISSN 0001-8678.
- K. Maulik and S.I. Resnick. Characterizations and examples of hidden regular variation. *Extremes*, 7(1):31–67, 2005.
- A. Mitra and S.I. Resnick. Hidden regular variation: detection and estimation. 2010. URL <http://arxiv.org/abs/1001.5058>.
- I. Molchanov. *Theory of Random Sets*. Probability and its Applications (New York). Springer-Verlag London Ltd., London, 2005.
- Yu.V. Prohorov. Convergence of random processes and limit theorems in probability theory. *Teor. Veroyatnost. i Primenen.*, 1:177–238, 1956. ISSN 0040-361x.
- S.I. Resnick. Hidden regular variation, second order regular variation and asymptotic independence. *Extremes*, 5(4):303–336, 2002.
- S.I. Resnick. *Heavy Tail Phenomena: Probabilistic and Statistical Modeling*. Springer Series in Operations Research and Financial Engineering. Springer-Verlag, New York, 2007.
- S.I. Resnick. *Extreme Values, Regular Variation and Point Processes*. Springer Series in

Operations Research and Financial Engineering. Springer, New York, 2008. Reprint of the 1987 original.

RISKLAB, DEPARTMENT OF MATHEMATICS,  
ETH ZURICH,  
RÄMISTRASSE 101,  
8092 ZURICH, SWITZERLAND  
E-MAIL: bikram@math.ethz.ch

SCHOOL OF OPERATIONS RESEARCH AND  
INFORMATION ENGINEERING,  
CORNELL UNIVERSITY,  
ITHACA, NY 14853, USA  
E-MAIL: am492@cornell.edu; sir1@cornell.edu