Stochastic Systems

MODELS WITH HIDDEN REGULAR VARIATION: GENERATION AND DETECTION

By Bikramjit Das *

SUTD

and

By Sidney I. Resnick†

Cornell University

We review definitions of multivariate regular variation (MRV) and hidden regular variation (HRV) for distributions of random vectors and then summarize methods for generating models exhibiting both properties. We also discuss diagnostic techniques that detect these properties in multivariate data and indicate when models exhibiting both MRV and HRV are plausible fits for the data. We illustrate our techniques on simulated data and also two real Internet data sets.

1. Introduction. Data exhibiting heavy tails appear naturally in many contexts, for example hydrology [1], finance [26], insurance [11], Internet traffic and telecommunication [4] and risk assessment [7, 16]. Often the observed data are multi-dimensional with heavy tailed marginal distributions and come from complex systems and we must study the dependence structure among the components.

The study of multivariate heavy-tailed models is facilitated by the ability to generate such models. Moreover, a generation technique helps in stress-testing worst-case scenarios. In the first part of this paper we consider several generation techniques and discuss their strengths and weaknesses.

A second theme of this paper is the development of diagnostics for detecting and identifying multivariate heavy tailed models prior to estimating model parameters. The second part of this paper deals with this.

1.1. *Outline.* The mathematical framework for the study of multivariate heavy tails is regular variation of measures. We provide a careful review of the definitions of multivariate regular variation (MRV) and hidden regular variation (HRV) in Section 1.2 and list the notations we use in Section 1.3. In Section 2 we discuss methods for generating regularly varying models on $\mathbb{E} = [0,\infty)^2 \setminus \{(0,0)\}$ and $\mathbb{E}_0 = (0,\infty)^2$ when the asymptotic limit measures are specified. The described methods are relatively easy to implement.

In Section 3 we discuss how to create models that exhibit both MRV and HRV. Both MRV and HRV are asymptotic models with curious properties which are often ignored or misinterpreted

*B. Das was supported by SRG-ESD-2012-047.

†S. Resnick was supported by Army MURI grant W911NF-12-1-0385 to Cornell University. Resnick acknowledges hospitality, space and support from SUTD during a visit January 2014.

MSC 2010 subject classifications: 28A33, 60G17, 60G51, 60G70

Keywords and phrases: regular variation, multivariate heavy tails, hidden regular variation, tail estimation, conditional extreme value model

when attempting to generate finite samples exhibiting such properties. We review three model generation methods that yield the asymptotic properties of both MRV on $\mathbb{E}$ and HRV on $\mathbb{E}_0$ and discuss characteristics of each method. These methods are called (i) the mixture method, (ii) the multiplication method and (iii) the additive method. We give particular attention to the recently proposed additive generation method of [27] and show that there are identifiability issues in the sense that asymptotic parameters may not be coming from the anticipated summand of the representation. Accompanying simulation examples illustrate our discussion.

Section 4 gives techniques for detecting when data is consistent with a model exhibiting MRV and HRV. These techniques rely on the fact that under broad conditions, if a vector $\boldsymbol{X}$ has a multivariate regularly varying distribution on a cone $\mathbb{C}$, then under a *generalized polar coordinate transformation* (see (1.4)), the transformed vector satisfies a conditional extreme value (CEV) model for which detection techniques exist from [6]. This methodology goes beyond one dimensional techniques such as checking one dimensional marginal distributions are heavy tailed or checking one dimensional functions of the data vector such as maximum and minimum component are heavy tailed.

In Section 5, we give two examples of our detection and model estimation techniques applied to Internet downloads and HTTP response data.

1.2. *Regularly varying distributions on cones..* We review material from [8, 15, 18] describing the framework for the definition of MRV and HRV and then specialize to two dimensions.

Let $\mathbb{X}$ be a metric space with metric $d(\boldsymbol{x}, \boldsymbol{y})$ satisfying

$$d(c\boldsymbol{x}, c\boldsymbol{y}) = cd(\boldsymbol{x}, \boldsymbol{y}), \qquad c > 0, (\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{X} \times \mathbb{X}. \tag{1.1}$$

If $d(\cdot, \cdot)$ is defined by a norm, (1.1) is satisfied. Hence in finite dimensional Euclidean space, (1.1) can always be satisfied. A flexible framework for discussing regular variation is measure convergence defined by $\mathbb{M}$-convergence [8, 18]) on a closed cone $\mathbb{C} \subset \mathbb{X}$ with a closed cone $\mathbb{C}_0 \subset \mathbb{C}$ deleted. The concept of a cone requires specifying a definition of scalar multiplication $(c, \boldsymbol{x}) \mapsto c\boldsymbol{x}$ from $\mathbb{R}_+ \times \mathbb{X} \mapsto \mathbb{X}$. In this paper, the metric space is Euclidean and scalar multiplication is the usual one. A cone $\mathbb{C}$ is closed under scalar multiplication: If $\boldsymbol{x} \in \mathbb{C}$ then $c\boldsymbol{x} \in \mathbb{C}$ for $c > 0$. A subset $\Lambda \subset \mathbb{C} \setminus \mathbb{C}_0$ is *bounded away from* $\mathbb{C}_0$ if $d(\Lambda, \mathbb{C}_0) > 0$. The two cases of most interest are

1. $\mathbb{C} = \mathbb{R}_+^2$ and $\mathbb{C}_0 = \{\mathbf{0}\}$. Then $\mathbb{E} := \mathbb{C} \setminus \mathbb{C}_0 = \mathbb{R}_+^2 \setminus \{\mathbf{0}\}$ is the space for defining $\mathbb{M}$-convergence appropriate for regular variation of distributions of positive random vectors.
2. $\mathbb{C} = \mathbb{R}_+^2$ and $\mathbb{C}_0 = \{\boldsymbol{x} : \wedge_{i=1}^2 x_i = 0\} :=$ [axes]. Then $\mathbb{E}_0 := \mathbb{C} \setminus \mathbb{C}_0$, the first quadrant without its axes, is the space for defining $\mathbb{M}$-convergence appropriate for HRV.

A random vector $\boldsymbol{Z} \geqslant \mathbf{0}$ is regularly varying on $\mathbb{C} \setminus \mathbb{C}_0$ if there exists a regularly varying function $b(t) \in RV_{1/\alpha}$, $\alpha > 0$ called the *scaling function* and a measure $\nu(\cdot) \in \mathbb{M}(\mathbb{C} \setminus \mathbb{C}_0)$ called the *limit or tail measure* such that as $t \to \infty$,

$$t\mathbf{P}[\boldsymbol{Z}/b(t) \in \cdot\,] \to \nu(\cdot), \tag{1.2}$$

in $\mathbb{M}(\mathbb{C} \setminus \mathbb{C}_0)$, the set of measures on $\mathbb{C} \setminus \mathbb{C}_0$ which are finite on sets bounded away from $\mathbb{C}_0$ [8, 15, 18]. We write $\boldsymbol{Z} \in MRV(\alpha, b(t), \nu, \mathbb{C} \setminus \mathbb{C}_0)$. Since $b(t) \in RV_{1/\alpha}$, $\nu(\cdot)$ has a scaling property

$$\nu(c\cdot) = c^{-\alpha}\nu(\cdot), \qquad c > 0. \tag{1.3}$$

When $\mathbb{C} = \mathbb{R}_+^2$, $\mathbb{C}_0 = \{\mathbf{0}\}$ and ν satisfies $\nu(\boldsymbol{x}, \boldsymbol{\infty}) = 0$ for all $\boldsymbol{x} > \mathbf{0}$ so that ν concentrates on the axes, we say $\boldsymbol{Z}$ possesses *asymptotic independence* [9, 23, 24]. It is convenient to translate (1.2) and (1.3) using generalized polar coordinates [8, 18]. Set $\aleph_{\mathbb{C}_0} = \{\boldsymbol{x} \in \mathbb{C} \setminus \mathbb{C}_0 : d(\boldsymbol{x}, \mathbb{C}_0) = 1\}$, the locus of points at distance 1 from the deleted region $\mathbb{C}_0$. Define $\text{GPOLAR} : \mathbb{C} \setminus \mathbb{C}_0 \mapsto (0, \infty) \times \aleph_{\mathbb{C}_0}$ by

$$\text{GPOLAR}(\boldsymbol{x}) = \Big(d(\boldsymbol{x}, \mathbb{C}_0), \frac{\boldsymbol{x}}{d(\boldsymbol{x}, \mathbb{C}_0)}\Big) \tag{1.4}$$

Then (1.2) and (1.3) are equivalent to

$$t\mathbf{P}[\text{GPOLAR}(\boldsymbol{Z})/b(t) \in \cdot\,] \to \nu_\alpha \times S(\cdot) = \nu \circ \text{GPOLAR}^{-1}, \tag{1.5}$$

in $\mathbb{M}\big((0, \infty) \times \aleph_{\mathbb{C}_0}\big)$ where $\nu_\alpha(x, \infty) = x^{-\alpha}$, $x > 0$, $\alpha > 0$ and $S(\cdot)$ is a probability measure on $\aleph_{\mathbb{C}_0}$; see [8, 18]. One should note that the transformation GPOLAR depends on the cone $\mathbb{C}_0$; this dependence should be understood from the context.

We focus on regular variation for $p = 2$ and the two choices of $\mathbb{C}$ and $\mathbb{C}_0$ which yield the spaces

1. $\mathbb{E} := \mathbb{R}_+^2 \setminus \{\mathbf{0}\}$.
2. $\mathbb{E}_0 = \mathbb{R}_+^2 \setminus \{\boldsymbol{x} : x_1 \wedge x_2 = 0\} =: \mathbb{R}_+^2 \setminus [\text{axes}]$.

Then $\boldsymbol{Z}$ is regularly varying on $\mathbb{E}$ and has *hidden regular variation* (HRV) on $\mathbb{E}_0$ if there exist $0 < \alpha \leqslant \alpha_0$, scaling functions $b(t) \in RV_{1/\alpha}$ and $b_0 \in RV_{1/\alpha_0}$ with $b(t)/b_0(t) \to \infty$ and limit measures ν, ν_0 such that

$$\boldsymbol{Z} \in \text{MRV}(\alpha, b(t), \nu, \mathbb{E}) \cap \text{MRV}(\alpha_0, b_0(t), \nu_0, \mathbb{E}_0)$$

so that unpacking the notation we get,

$$t\mathbf{P}[\boldsymbol{Z}/b(t) \in \cdot\,] \to \nu(\cdot), \quad \text{in } \mathbb{M}(\mathbb{E}) \tag{1.6}$$

and

$$t\mathbf{P}[\boldsymbol{Z}/b_0(t) \in \cdot\,] \to \nu_0(\cdot), \quad \text{in } \mathbb{M}(\mathbb{E}_0). \tag{1.7}$$

On $\mathbb{E}$ we may take $\aleph_{\mathbf{0}} = \{\boldsymbol{x} : \|\boldsymbol{x}\| = 1\}$ for a convenient choice of $d(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{x} - \boldsymbol{y}\|$ and on $\mathbb{E}_0$,

$$\aleph_{[\text{axes}]} := \{\boldsymbol{x} \in \mathbb{E} : x_1 \wedge x_2 = 1\}$$

is the appropriate unit sphere. Then using GPOLAR (1.6) and (1.7) become,

$$t\mathbf{P}\big[\big(\|\boldsymbol{Z}\|/b(t), \boldsymbol{Z}/\|\boldsymbol{Z}\|\big) \in \cdot\,\big] \to \nu_\alpha \times S(\cdot), \quad \text{in } \mathbb{M}((0, \infty) \times \aleph_{\mathbf{0}}) \tag{1.8}$$

and

$$t\mathbf{P}\Big[\Big(\frac{Z_1 \wedge Z_2}{b_0(t)}, \frac{\boldsymbol{Z}}{Z_1 \wedge Z_2}\Big) \in \cdot\Big] \to \nu_{\alpha_0} \times S_0(\cdot) \qquad \text{in } \mathbb{M}\big((0, \infty) \times \aleph_{[\text{axes}]}\big) \tag{1.9}$$

and S and S_0 are probability measures on $\aleph_{\mathbf{0}}$ and $\aleph_{[\text{axes}]}$ respectively. Note

$$\Big(\frac{\boldsymbol{z}}{z_1 \wedge z_2}\Big) = \begin{cases} (1, z_2/z_1), & \text{if } z_1 \leqslant z_2, \\ (z_1/z_2, 1), & \text{if } z_2 < z_1 \end{cases}$$

and

$$\aleph_{[\text{axes}]} = \big([1,\infty) \times \{1\}\big) \cup \big(\{1\} \times [1,\infty)\big).$$

So we may rewrite (1.9) as two statements: For $x \geqslant 1$,

$$tP[\frac{Z_1}{b_0(t)} > r, \frac{Z_2}{Z_1} > x] \to r^{-\alpha_0} S_0\{(1,z) : z > x\} =: r^{-\alpha_0} p\bar{G}_1(x), \tag{1.10}$$

$$tP[\frac{Z_2}{b_0(t)} > r, \frac{Z_1}{Z_2} > x] \to r^{-\alpha_0} S_0\{(z,1) : z > x\} =: r^{-\alpha_0} q\bar{G}_2(x), \tag{1.11}$$

where $p := S_0\{\{1\} \times [1,\infty)\}$, $q := S_0\{[1,\infty) \times \{1\}\} = 1 - p$ and G_1, G_2 are probability distributions on $[1,\infty)$. We also have

$$tP[\frac{Z_1 \wedge Z_2}{b_0(t)} > r, \left(\frac{Z_1}{Z_2} \bigvee \frac{Z_2}{Z_1}\right) > x] \to r^{-\alpha_0}\left(p\bar{G}_1(x) + q\bar{G}_2(x)\right). \tag{1.12}$$

Traditionally [23], regular variation on $\mathbb{E}$ has been studied using the one point uncompactification, vague convergence and the polar coordinate transform $\boldsymbol{x} \mapsto (\|\boldsymbol{x}\|, \boldsymbol{x}/\|\boldsymbol{x}\|)$. On $\mathbb{E}$ this works fine because $\{\boldsymbol{x} \in \mathbb{E} : \|\boldsymbol{x}\| = 1\}$ is compact and lines through $\boldsymbol{\infty}$ cannot carry mass. However, on $\mathbb{E}_0$ the traditional unit sphere $\{\boldsymbol{x} \in \mathbb{E}_0 : \|\boldsymbol{x}\| = 1\}$ is no longer compact. Hence, Radon measures on $\{\boldsymbol{x} \in \mathbb{E}_0 : \|\boldsymbol{x}\| = 1\}$ may not be finite and for estimation problems the approach relying on vague convergence is a dead end if estimation of a possibly infinite measure is required. More details on why an approach without compactification is desirable are in [8, 15, 18]. We emphasize it is difficult to discuss MRV on $\mathbb{E}_0$ with the conventional unit sphere and it is preferable to use $\aleph_{[\text{axes}]}$.

1.3. *Basic notation.* Here is a notation and concept summary.

RV_β	Regularly varying functions with index $\beta > 0$. We can and do assume such functions are continuous and strictly increasing.
$\mathbb{E}$	$\mathbb{R}^2 \setminus \{\mathbf{0}\}$.
[axes]	$\{0\} \times \mathbb{R}_+ \cup \mathbb{R}_+ \times \{0\}$.
$\mathbb{E}_0$	$\mathbb{R}^2 \setminus$ [axes].
$\mathbb{M}(\mathbb{C} \setminus \mathbb{C}_0)$	The set of all non-zero measures on $\mathbb{C} \setminus \mathbb{C}_0$ which are finite on subsets bounded away from $\mathbb{C}_0$.
$\mathcal{C}(\mathbb{C} \setminus \mathbb{C}_0)$	Continuous, bounded, positive functions on $\mathbb{C} \setminus \mathbb{C}_0$ whose supports are bounded away from $\mathbb{C}_0$. Without loss of generality [18], we may assume the functions are uniformly continuous.
$\mu_n \to \mu$	Convergence in $\mathbb{M}(\mathbb{C} \setminus \mathbb{C}_0)$ means $\mu_n(f) \to \mu(f)$ for all $f \in \mathcal{C}(\mathbb{C} \setminus \mathbb{C}_0)$. See [8, 15, 18].
$\aleph_{\mathbb{C}}$	$\{\boldsymbol{x} : d(\boldsymbol{x}, \mathbb{C}) = 1\}$.
$\aleph_{\mathbf{0}}$	$\{\boldsymbol{x} \in \mathbb{E} : d(\boldsymbol{x}, \{\mathbf{0}\}) = 1\}$.
$\aleph_{[\text{axes}]}$	$\{\boldsymbol{x} \in \mathbb{E}_0 : d(\boldsymbol{x}, [\text{axes}]) = 1\} = \{1\} \times [1,\infty) \cup [1,\infty) \times \{1\}$.
MRV	multivariate regular variation; for this paper, it means regular variation on $\mathbb{E}$.
HRV	hidden regular variation; for this paper, it means regular variation on $\mathbb{E}_0$.
GPOLAR	Polar co-ordinate transformation relative to the deleted cone $\mathbb{C}_0$, $\text{GPOLAR}(\boldsymbol{x}) = \big(d(\boldsymbol{x}, \mathbb{C}_0), \boldsymbol{x}/d(\boldsymbol{x}, \mathbb{C}_0)\big)$. See [8, 18].
$\boldsymbol{X} \perp\!\!\!\perp \boldsymbol{Y}$	The random elements $\boldsymbol{X}, \boldsymbol{Y}$ are independent.

2. Generating Regularly Varying Models. We outline schemes for generating regular variation. These schemes generate the full totality of asymptotic limits but not the full totality of pre-asymptotic models; so there can be many other ways to get the same asymptotic models.

2.1. *Generating regular variation on* $\mathbb{E}$.. The easiest way to obtain a regularly varying model on $\mathbb{E}$ with scaling function $b(t)$ and limit measure $\nu(\cdot) = \nu_\alpha \times S \circ \text{GPOLAR}$ is as follows: Suppose R is a random element of $(0,\infty)$ with a regularly varying tail and scaling function $b(t)$:

$$t\mathbf{P}[R/b(t) > x] \to x^{-\alpha}, \quad x > 0,\, \alpha > 0.$$

Let $\boldsymbol{\Theta}$ be a random element of $\aleph_{\mathbf{0}}$ with distribution S

$$\mathbf{P}[\boldsymbol{\Theta} \in \cdot\,] = S(\cdot)$$

and which is independent of R. Then $\boldsymbol{Z} := R\boldsymbol{\Theta} = \text{GPOLAR}^{\leftarrow}(R, \boldsymbol{\Theta})$ is regularly varying on $\mathbb{E}$ with limit measure $\nu = \nu_\alpha \times S \circ \text{GPOLAR}$ on $\mathbb{E}$ because (1.8) and consequently (1.6) hold. Note GPOLAR is defined relative to the deleted cone $\{\mathbf{0}\}$.

2.2. *Generating regular variation on $\mathbb{E}_0$ (and sometimes also on $\mathbb{E}$).*. As suggested in [19], we may follow the same scheme as in Section 2.1. Let R_0 be a random element of $(0,\infty)$ that is regularly varying with index α_0 and scaling function $b_0(t)$. Let $\boldsymbol{\Theta}_0$ be a random element of $\aleph_{[\text{axes}]}$ with distribution S_0 and independent of R_0. Then $\boldsymbol{Z} = R_0\boldsymbol{\Theta}_0 = \text{GPOLAR}^{\leftarrow}(R_0, \boldsymbol{\Theta}_0)$ is regularly varying with scaling function $b_0(t)$ and limit measure $\nu_0 := \nu_{\alpha_0} \times S_0 \circ \text{GPOLAR}^{-1}$ on $\mathbb{E}_0$ because (1.9) and therefore (1.7) hold.

In practice we specify the measure S_0 on $\aleph_{[\text{axes}]}$ as follows: Let G_1, G_2 be two probability measures on $(1,\infty)$ and define

$$\boldsymbol{\Theta}_0 = B(\Theta_1, 1) + (1-B)(1, \Theta_2) \tag{2.1}$$

where B, Θ_1, Θ_2 are independent, B is a Bernoulli switching variable with $P[B=1] = p = 1-P[B=0]$ and Θ_i has distribution G_i, $i = 1,2$. So G_1 smears probability mass on the horizontal line emanating from $(1,1)$ and G_2 does the same thing for the vertical line.

For estimation purposes, note for $s > 1$ that

$$\bar{G}_1(s) = G_1(s,\infty) = \nu_0\{\boldsymbol{x} \in \mathbb{E}_0 : x_1/x_2 > s\}, \tag{2.2}$$

$$\bar{G}_2(s) = G_2(s,\infty) = \nu_0\{\boldsymbol{x} \in \mathbb{E}_0 : x_2/x_1 > s\}. \tag{2.3}$$

Depending on the moments of G_i, $i = 1,2$, it may be possible to extend the regular variation constructed on $\mathbb{E}_0$ to $\mathbb{E}$ so that the marginals Z_1, Z_2 individually have tails which are regularly varying. This means [19]

$$\nu_0\{\boldsymbol{x} \in \mathbb{E}_0 : \|\boldsymbol{x}\| > 1\} < \infty,$$

which occurs when

$$\bigvee_{i=1}^{2} \int_1^\infty s^{\alpha_0 - 1}\bar{G}_i(s)ds < \infty,$$

and is thus a somewhat restricted case. Regular variation on $\mathbb{E}_0$ by itself does not in general imply one dimensional regular variation of the marginals. Also if the tails of G_i are heavier than the tail of R, we can have regular variation on $\mathbb{E}_0$ with index α_0 but the tails of Z_1 and Z_2 may be regularly varying with a smaller index α. Full discussion is in [19].

3. Generating models that have both multivariate regular variation on $\mathbb{E}$ and HRV on $\mathbb{E}_0$.. We summarize several methods for generating models possessing both MRV on $\mathbb{E}$ and HRV on $\mathbb{E}_0$.

3.1. *Mixture method.*. This method [19, 23] expresses the random vector $\boldsymbol{Z}$ as

$$\boldsymbol{Z} = B\boldsymbol{Y} + (1-B)\boldsymbol{V},$$

a mixture where $\boldsymbol{Y}$ gives the regular variation on $\mathbb{E}$ and $\boldsymbol{V}$ gives the regular variation on $\mathbb{E}_0$. Since HRV implies that MRV on $\mathbb{E}$ must include asymptotic independence [22, 23], we need $\boldsymbol{Y}$ to model MRV with index α on $\mathbb{E}$ and have asymptotic independence. So we take $\boldsymbol{Y}$ to concentrate on [axes] and

$$\boldsymbol{Y} = B_1(\xi_1, 0) + (1-B_2)(0, \xi_2) \tag{3.1}$$

where B_1, ξ_1, ξ_2 are independent, B_1 is a Bernoulli switching variable and

$$t\mathbf{P}[\xi_i/b(t) > x] \to x^{-\alpha}, \quad x > 0,\, \alpha > 0,\, t \to \infty. \tag{3.2}$$

Construct $\boldsymbol{V}$ by the scheme of Section 2.2 to be regularly varying on $\mathbb{E}_0$ with limit measure ν_0 and scaling function $b_0(t)$. The resulting $\boldsymbol{Z}$ has both MRV on $\mathbb{E}$ and HRV on $\mathbb{E}_0$:

$$\boldsymbol{Z} \in \mathrm{MRV}(\alpha, b(t), \nu, \mathbb{E}) \cap \mathrm{MRV}(\alpha_0, b_0(t), \nu_0, \mathbb{E}_0).$$

3.2. *Additive method..* Weller and Cooley [27] advocate an additive model of the form

$$\boldsymbol{Z} = \boldsymbol{Y} + \boldsymbol{V},$$

where $\boldsymbol{Y} \in \mathrm{MRV}(\alpha, b(t), \nu, \mathbb{E})$ and $\boldsymbol{V}$ has HRV and $\boldsymbol{V} \in \mathrm{MRV}(\alpha_0, b_0(t), \nu_0, \mathbb{E}_0)$ and $\boldsymbol{Y} \perp\!\!\!\perp \boldsymbol{V}$. They argue there are advantages for estimating the parameters and the additive model overcomes the undesireable and usually unrealistic feature of the mixture method that points are installed directly on the axes. However, as we will see, the additive model does not always successfully separate the HRV piece in a way that is identifiable.

Simple case: $\boldsymbol{Y}$ has no HRV and there is a finite hidden angular measure.. We start with the simplest result.

PROPOSITION 3.1. *Suppose*

1. *$\boldsymbol{Y}$ has the structure given in (3.1) (so that $\boldsymbol{Y}$ has no HRV) and (3.2) holds.*
2. *$\boldsymbol{V}$ has MRV on $\mathbb{E}$ (not $\mathbb{E}_0$) with index $\alpha_0 \geqslant \alpha$, scaling function $b_0(t) = o(b(t))$, limit measure $\nu_0 \in M(\mathbb{E})$ and no asymptotic independence. Regular variation of $\boldsymbol{V}$ on $\mathbb{E}$ has the consequence that for $i = 1, 2$,*

$$t\mathbf{P}[V_i > b_0(t)x] \to c_i x^{-\alpha_0},\, x > 0,\, t \to \infty,\, c_i \geqslant 0,\, c_1 \vee c_2 > 0. \tag{3.3}$$

Then $\boldsymbol{Z} := \boldsymbol{Y} + \boldsymbol{V}$ has

1. *MRV on $\mathbb{E}$: $\boldsymbol{Z} \in MRV(\alpha, b(t), \nu, \mathbb{E})$ and $\boldsymbol{Z}$ has asymptotic independence.*
2. *HRV on $\mathbb{E}_0$: $\boldsymbol{Z} \in MRV(\alpha_0, b_0(t), \nu_{0|_{\mathbb{E}_0}}, \mathbb{E}_0)$. The limit measure $\nu_{0|_{\mathbb{E}_0}}$ is ν_0 restricted to $\mathbb{E}_0$ and*

$$\nu_0\{\boldsymbol{x} \in \mathbb{E}_0 : \|\boldsymbol{x}\| \geqslant 1\} < \infty. \tag{3.4}$$

The last condition means the hidden limit measure ν_0 has finite spectral measure with respect to the conventional unit sphere since $\boldsymbol{V}$ has MRV on $\mathbb{E}$. So the construction in Proposition 3.1 yields only a special case of HRV since there are many cases where (3.4) fails.

PROOF. The statement about MRV on $\mathbb{E}$ can be deduced from known results, eg. Resnick [23, p. 230], Jessen and Mikosch [17], Resnick [21]. (Note, it would not be enough to assume $\boldsymbol{V} \in \mathrm{MRV}(\alpha_0, b_0, \nu_0, \mathbb{E}_0)$.) To prove HRV of $\boldsymbol{Z}$ on $\mathbb{E}_0$, we apply criterion (ii) of the Portmanteau Theorem 2.1 in [18] and let $f \in \mathcal{C}((0,\infty)^2)$ and without loss of generality suppose f is bounded by a constant $\|f\|$, uniformly continuous and

$$f(\boldsymbol{x}) = 0, \quad \text{if } x_1 \wedge x_2 < \eta,$$

for some $\eta > 0$. Uniform continuity of f means that the modulus of continuity

$$\omega_f(\delta) := \sup\{|f(\boldsymbol{x}) - f(\boldsymbol{y})| : \|\boldsymbol{x} - \boldsymbol{y}\| < \delta\} \to 0, \quad (\delta \to 0).$$

Since $\boldsymbol{V}$ has MRV on $\mathbb{E}$ we have

$$t\mathbf{E}f(\boldsymbol{V}/b_0(t)) \to \nu_0(f),$$

and so it suffices to show as $t \to \infty$.

$$t\mathbf{E}f\Big(\frac{\boldsymbol{Y}+\boldsymbol{V}}{b_0(t)}\Big) - t\mathbf{E}f\Big(\frac{\boldsymbol{V}}{b_0(t)}\Big) \to 0. \tag{3.5}$$

Because of the special structure of $\boldsymbol{Y}$, the absolute value of the difference on the previous line is bounded by

$$\begin{aligned}
&\frac{t}{2}\mathbf{E}\Big|f\Big(\frac{\xi_1+V_1}{b_0(t)}, \frac{V_2}{b_0(t)}\Big) - f(\boldsymbol{V}/b_0(t))\Big| + \frac{t}{2}\mathbf{E}\Big|f\Big(\frac{V_1}{b_0(t)}, \frac{\xi_2+V_2}{b_0(t)}\Big) - f(\boldsymbol{V}/b_0)\Big| \\
=&I + II.
\end{aligned}$$

For $\delta < \eta$, write

$$2I = t\mathbf{E}|\cdot|1_{[\xi_1/b_0(t)<\delta]} + t\mathbf{E}|\cdot|1_{[\xi_1/b_0(t)>\delta]} = 2Ia + 2Ib.$$

To keep both terms of the difference from being zero we write

$$\begin{aligned}
2Ia = t\mathbf{E}&\Big|f\Big(\frac{\xi_1+V_1}{b_0(t)}, \frac{V_2}{b_0(t)}\Big) - f(\boldsymbol{V}/b_0)\Big|1_{[\xi_1<b_0(t)\delta, V_1>b_0(t)(\eta-\delta)]} \\
\leqslant&\omega_f(\delta)t\mathbf{P}[V_1 > b_0(t)(\eta-\delta)] \to \omega_f(\delta)c_1(\eta-\delta)^{-\alpha_0} \quad (t \to \infty), \\
\to&0 \quad (\delta \to 0),
\end{aligned}$$

where we used (3.3).

For $2Ib$, in order to keep both terms of the difference from being zero, we write,

$$\begin{aligned}
2Ib =&\frac{t}{2}\mathbf{E}\Big|f\Big(\frac{\xi_1+V_1}{b_0(t)}, \frac{V_2}{b_0(t)}\Big) - f(\boldsymbol{V}/b_0)\Big|1_{[\xi_1>b_0(t)\delta, V_2>b_0(t)\delta]} \\
\leqslant&2\|f\|t\mathbf{P}[V_2 > b_0(t)\delta]\mathbf{P}[\xi_1 > b_0(t)\delta]
\end{aligned}$$

and as $t \to \infty$ this is

$$\sim 2\|f\|c_2\delta^{-\alpha_0}\mathbf{P}[\xi_1 > b_0(t)\delta] \to 0 \quad (t \to \infty).$$

We handle II similarly. □

EXAMPLE 3.1. Suppose $\boldsymbol{Y}$ has the structure given in (3.1) where ξ_1, ξ_2 are iid Pareto distributed with index α. Assume $\boldsymbol{V} = R_0\boldsymbol{\Theta}_0$ where R_0 is Pareto distributed index $\alpha_0 > \alpha$ and $\boldsymbol{\Theta}_0$ has the structure given in (2.1) where $\Theta_i = 1 + \mathbf{E}_i$ and E_1, E_2 are two standard iid exponential random variables. Then $\boldsymbol{V} = R_0\boldsymbol{\Theta}_0 \in \text{MRV}(\alpha_0, b_0(t), \nu_0, \mathbb{E})$ and

$$\nu_0 = \nu_{\alpha_0} \times \mathbf{P}[\boldsymbol{\Theta} \in \cdot\,] \circ \text{GPOLAR}^{-1}.$$

This construction makes the marginals of $\boldsymbol{V} = (V_1, V_2)$ regularly varying with index α_0 which is consistent with $\boldsymbol{V}$ being MRV on $\mathbb{E}$ rather than just $\mathbb{E}_0$:

$$\begin{aligned}\mathbf{P}[V_1 > x] =& p\mathbf{P}[R(1+E_1) > x] + q\mathbf{P}[R > x]\\ \sim& px^{-\alpha_0}\mathbf{E}\big((1+E_1)^{\alpha_0}\big) + qx^{-\alpha_0},\end{aligned}$$

(where the $\sim$ results from an application of Breiman's theorem [2] on products)

$$=(const)x^{-\alpha_0}.$$

Here $p = 1 - q = P(\boldsymbol{\Theta}_0 \in ((1,\infty) \times \{1\}))$.

To check whether we can get identify the distributions of $\boldsymbol{Y}$ and $\boldsymbol{V}$ from a data sample of $\boldsymbol{Z} = \boldsymbol{Y} + \boldsymbol{V}$, we simulate data following this model for three different choices of α while keeping α_0 fixed. We then check whether we can estimate back the values of α and α_0. In all the three cases $\alpha_0 = 2$ with $\Theta_1 \stackrel{d}{=} \Theta_2$ with $(\Theta_1 - 1)$ following an iid standard exponential distribution and $p = 0.5$. In each case we simulate 10000 iid samples from $\boldsymbol{Z}$. Then we create Hill plots for the marginals of Z_1 and Z_2 to identify the value of α. To detect the hidden part we create a Hill plot for $\min(Z_1, Z_2)$ to find the value of α_0. Referencing (1.12), we also make a QQ plot of $\max(Z_1/Z_2, Z_2/Z_1)$ for the 100 highest values of $\min(Z_1, Z_2)$ against the quantiles of standard exponential which is the distribution of Θ_1 and Θ_2. We discuss the cases below.

Case 1: $\alpha = 1$. The top panel of Figure 1 indicates that we can identify the tails of $\boldsymbol{Z}$ to be heavy tailed. The correct index $\alpha = 1$ is slightly overestimated. The Hill plot of $\min(Z_1, Z_2)$ also indicates HRV on $\mathbb{E}_0$ with index close to $\alpha_0 = 2$. The QQ plot of $\max(Z_1/Z_2, Z_2/Z_1)$ thresholded by the 100 largest values of $\min(Z_1, Z_2)$ against standard exponential shows a decent fit.

Case 2: $\alpha = 1.5$. The top panel of Figure 2 again indicates that we can identify the tails of $\boldsymbol{Z}$ to be heavy tailed. The index α is again overestimated, this time more than in the previous case, perhaps because of the closeness of α to α_0. The Hill plot of $\min\{Z_1, Z_2\}$ also indicates HRV on $\mathbb{E}_0$ with index close to $\alpha_0 = 2$. The QQ plot of $\max\{Z_1/Z_2, Z_2/Z_1\}$ thresholded by the 100 largest values of $\min\{Z_1, Z_2\}$ against standard exponential shows a decent fit again.

Case 3: $\alpha = 0.5$. In this case too, the top panel of Figure 3 indicates heavy tailed behavior of $\boldsymbol{Z}$. The Hill plot of $\min(Z_1, Z_2)$ also indicates hidden regular variation. The indices $\alpha = 0.5$ and $\alpha_0 = 2$ are reasonably estimated here, presumably because the original values of α and α_0 are far apart. However, the exponential QQ plot of $\max\{Z_1/Z_2, Z_2/Z_1\}$ for the 100 largest values of $\min\{Z_1, Z_2\}$ struggles to indicate an exponential fit.

□

What happens if $\boldsymbol{Y}$ *has HRV but* $\boldsymbol{V}$ *has no HRV.* In Proposition 3.1, we can remove the restriction that $\boldsymbol{Y} = (Y_1, Y_2)$ concentrates on the axes at the expense of a tail condition on $\boldsymbol{Y}$ that guarantees the tails of $\boldsymbol{V}$ and $\boldsymbol{Y}$ do not interact in such as way as to obscure the fact that the hidden angular measure of $\boldsymbol{Z}$ is that of $\boldsymbol{V}$. Continue to suppose $\boldsymbol{Z} = \boldsymbol{Y} + \boldsymbol{V}$ with $\boldsymbol{Y} \perp\!\!\!\perp \boldsymbol{V}$.

Proposition 3.2. *Suppose*

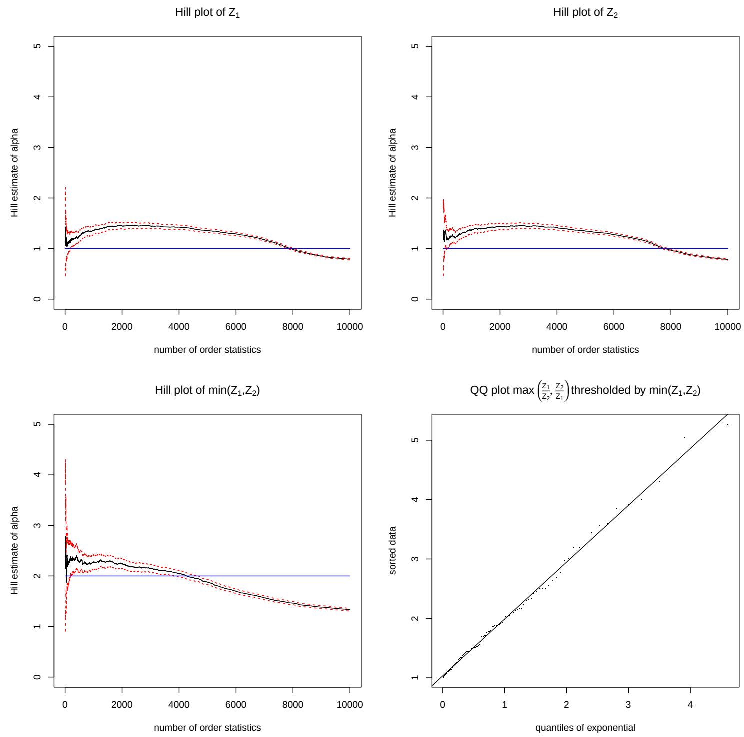

FIG 1. *Exploratory plots for Example 3.1, case 1, with $\alpha = 1, \alpha_0 = 2$. Top panel: Hill plots for the marginals Z_1 and Z_2. Bottom left: Hill plot for $\min\{Z_1, Z_2\}$. Bottom right: exponential QQ plot of $\max\{Z_1/Z_2, Z_2/Z_1\}$ thresholded by the 100 largest values of $\min\{Z_1, Z_2\}$.*

1. *$\boldsymbol{Y} \in MRV(\alpha, b(t), \nu, \mathbb{E})$ and exhibits asymptotic independence.*
2. *$\boldsymbol{V}$ has MRVon $\mathbb{E}$ (not $\mathbb{E}_0$) with index $\alpha_0 \geqslant \alpha$, scaling function $b_0(t) = o(b(t))$, limit measure $\nu_0 \in M(\mathbb{E})$ with no asymptotic independence so that*

$$t\mathbf{P}[\boldsymbol{V}/b_0(t) \in \cdot\,] \to \nu_0 \quad \text{in } \mathbb{M}(\mathbb{E}).$$

3. *The interaction of the tails of $\boldsymbol{Y}$ and $\boldsymbol{V}$ is controlled by the condition*

$$t\mathbf{P}[Y_1 \wedge Y_2 > b_0(t)x] \to 0, \quad t \to \infty,\, x > 0. \tag{3.6}$$

Then $\boldsymbol{Z} = \boldsymbol{Y} + \boldsymbol{V}$ has

1. *$MRV(\alpha, b(t), \nu, \mathbb{E})$ and asymptotic independence.*
2. *HRV on $\mathbb{E}_0$ with index α_0, scaling function $b_0(t)$, limit measure ν_0 restricted to $\mathbb{E}_0$.*

Remarks: For the $\boldsymbol{Y}$ defined in Proposition 3.1, $Y_1 \wedge Y_2 = 0$ so (3.6) is automatic. If Y_1, Y_2 are iid with $\mathbf{P}[Y_i > x] \in RV_{-\alpha}$, $\boldsymbol{Y}$ itself has HRV [22, 23] with index $\alpha_0 = 2\alpha$ and condition (3.6) is needed to guarantee the HRV of $\boldsymbol{Z}$ comes from $\boldsymbol{V}$ and not $\boldsymbol{Y}$. Condition (3.6) is equivalent in this case to

$$\frac{(\mathbf{P}[Y_1 > x])^2}{\mathbf{P}[V_1 \wedge V_2 > x]} \to 0, \quad (x \to \infty). \tag{3.7}$$

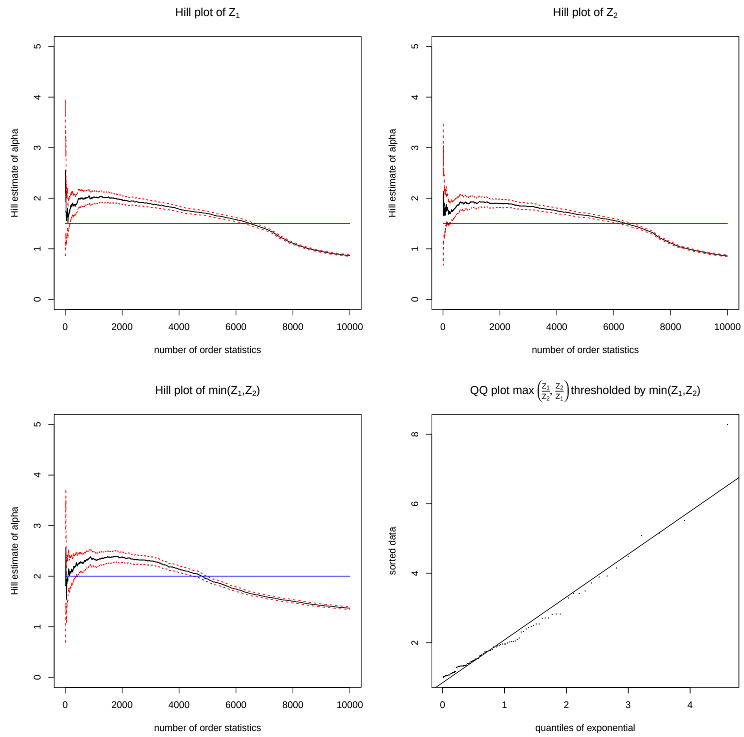

FIG 2. *Exploratory plots for Example 3.1, case 2, with* $\alpha = 1.5, \alpha_0 = 2$. *Top panel: Hill plots for the marginals* Z_1 *and* Z_2. *Bottom left: Hill plot for* $\min\{Z_1, Z_2\}$. *Bottom right: exponential QQ plot of* $\max\{Z_1/Z_2, Z_2/Z_1\}$ *thresholded by the 100 largest values of* $\min\{Z_1, Z_2\}$.

and it is sufficient that

$$\frac{\alpha_0}{2} < \alpha < \alpha_0.$$

This is seen by noting that for Y_1, Y_2 iid index α, (3.6) is

$$\begin{aligned} t\Big(\mathbf{P}[Y_1 > b_0(t)x]\Big)^2 =& t\Big(\mathbf{P}[Y_1 > b\big(b^{\leftarrow}(b_0(t))\big)x]\Big)^2 \\ =& \frac{t}{b^{\leftarrow}(b_0(t))}\Big(b^{\leftarrow}(b_0(t))\mathbf{P}[Y_1 > b\big(b^{\leftarrow}(b_0(t))\big)x]\Big)^2 \end{aligned}$$

and since $b^{\leftarrow}(b_0(t)) \to \infty$ and $b(\cdot)$ is the scaling function of Y_1, this is asymptotic to

$$\sim \frac{t}{b^{\leftarrow}(b_0(t))} x^{-2\alpha}.$$

We need $\lim_{t\to\infty} t/b^{\leftarrow}(b_0(t)) = 0$ and unwinding this condition yields (3.7).

PROOF. As in Proposition 3.1, we focus on the HRV claim. Again assume $f \in \mathcal{C}((0,\infty)^2)$ and f is bounded by $\|f\|$, uniformly continuous with

$$f(\boldsymbol{x}) = 0, \qquad \text{if } x_1 \wedge x_2 < \eta,$$

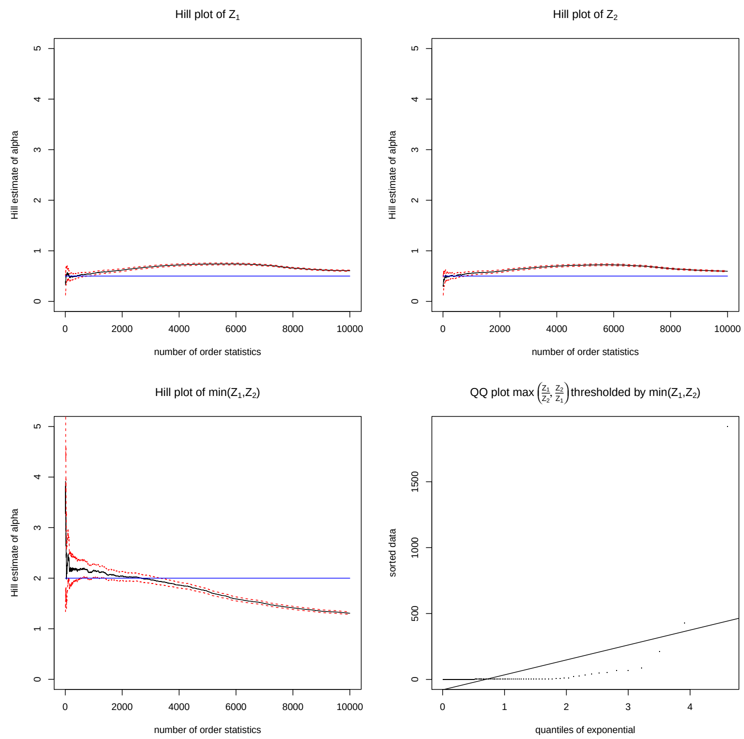

FIG 3. *Exploratory plots for Example 3.1, case 3, with $\alpha = 0.5, \alpha_0 = 2$. Top panel: Hill plots for the marginals Z_1 and Z_2. Bottom left: Hill plot for $\min\{Z_1, Z_2\}$. Bottom right: exponential QQ plot of $\max\{Z_1/Z_2, Z_2/Z_1\}$ thresholded by the 100 maximum values of $\min(Z_1, Z_2)$.*

for some $\eta > 0$. We need to show (3.5). For any small $\delta > 0$ with $\delta < \eta$, the absolute value of the difference in (3.5) is

$$tE|\cdot|1_{[Y_1 \vee Y_2 > b_0(t)\delta]} + tE|\cdot|1_{[Y_1 \vee Y_2 < b_0(t)\delta, V_1 \wedge V_2 > b_0(t)(\eta-\delta)]} = I + II,$$

since for the second term, the only way the difference can be non-zero is if $\boldsymbol{V}$ is sufficiently large. Term II is dominated by

$$\begin{aligned} II \leqslant & \omega_f(\delta) t\mathbf{P}[V_1 \wedge V_2 > b_0(t)(\eta - \delta)] \\ \sim & \omega_f(\delta)(const)(\eta - \delta)^{-\alpha_0}, \quad (t \to \infty) \\ \to & 0, \qquad \text{as } \delta \to 0. \end{aligned}$$

For I we have

$$\begin{aligned} I \leqslant & tE|\cdot|\Big(1_{[Y_1 \wedge Y_2 > b_0(t)\delta]]} + 1_{[Y_1 > b_0(t)\delta, Y_2 \leqslant b_0(t)\delta]]} + 1_{[Y_2 > b_0(t)\delta, Y_1 < b_0(t)\delta]]}\Big) \\ = & Ia + Ib + Ic. \end{aligned}$$

The term Ia can be quickly killed,

$$Ia \leqslant 2\|f\| t\mathbf{P}[Y_1 \wedge Y_2 > b_0(t)\delta] \to 0, \quad (t \to \infty)$$

from (3.6). The term Ib is dominated by

$$\begin{aligned} Ib \leqslant & 2\|f\| t\mathbf{P}[Y_1 > b_0(t)\delta, V_2 > b_0(t)(\eta-\delta)] \\ = & 2\|f\| t\mathbf{P}[V_2 > b_0(t)(\eta-\delta)]\mathbf{P}[Y_1 > b_0(t)\delta] \\ \sim & \|f\|(\eta-\delta)^{-\alpha_0}\mathbf{P}[Y_1 > b_0(t)\delta] \quad (t\to\infty), \\ \to & 0 \qquad (t\to\infty). \end{aligned}$$

Term Ic is handled similarly. □

What happens if $\boldsymbol{Y}$ has no HRV but $\boldsymbol{V}$ has HRV. A problem with the additive model is the tail weights contributing to MRV on $\mathbb{E}$ and HRV on $\mathbb{E}_0$ can be confounded between $\boldsymbol{Y}$ and $\boldsymbol{V}$ and it is possible for $\boldsymbol{V}$ to have MRV on $\mathbb{E}$, HRV on $\mathbb{E}_0$ but the hidden measure of $\boldsymbol{Z} = \boldsymbol{Y} + \boldsymbol{V}$ is not the hidden measure of $\boldsymbol{V}$.

To focus on the influence of $\boldsymbol{V}$, we again assume $\boldsymbol{Y}$ has the structure (3.1) used in Proposition 3.1.

PROPOSITION 3.3. *Suppose*

1. *$\boldsymbol{Y}$ has form (3.1) where ξ_1, ξ_2 are iid, each with distributions having regularly varying tails with index α and scaling function $b(t)$.*
2. *$\boldsymbol{V}$ has both MRV on $\mathbb{E}$ and HRV on $\mathbb{E}_0$:*
   (a) *$\boldsymbol{V} \in MRV(\alpha_*, b_*(t), \nu, \mathbb{E})$ and has asymptotic independence.*
   (b) *$\boldsymbol{V} \in MRV(\alpha_0, b_0(t), \nu_0, \mathbb{E}_0)$.*
3. *The parameters α, α_*, α_0 are related by $\alpha \leqslant \alpha_* \leqslant \alpha_0$ and the scaling functions $b(t), b_*(t), b_0(t)$ satisfy $b_*(t) = o(b(t))$, $b_0(t) = o(b_*(t))$.*
4. *Define a scaling function $h(t)$ through its inverse $h^{\leftarrow}(t)$ by*

$$h^{\leftarrow}(t) =: b^{\leftarrow}(t) b_*^{\leftarrow}(t) \sim (const)\frac{1}{\mathbf{P}[\xi_1 > t]\mathbf{P}[V_1 > t]}. \tag{3.8}$$

Then

1. *If*

$$h(t)/b_0(t) \to \infty, \tag{3.9}$$

   $\boldsymbol{Z} \in MRV(\alpha, b(t), \nu, \mathbb{E})$ with asymptotic independence and has HRV on $\mathbb{E}_0$ with index $\alpha + \alpha_*$ and limit measure (different than the hidden measure of $\boldsymbol{V}$):

$$\nu_{\boldsymbol{Z},hidden} := \frac{1}{2}\big(\nu_\alpha \times \nu_{\alpha_*} + \nu_{\alpha_*} \times \nu_\alpha\big). \tag{3.10}$$

   A sufficient condition for (3.9) is $\alpha_* < \alpha_0 - \alpha$.
2. *If*

$$h(t)/b_0(t) \to 0, \tag{3.11}$$

   then $\boldsymbol{Z} \in MRV(\alpha, b(t), \nu, \mathbb{E}) \cap MRV(\alpha_0, b_0(t), \nu_0, \mathbb{E}_0)$ and $\boldsymbol{Z}$ has asymptotic independence and has HRV and the hidden limit measure ν_0 of $\boldsymbol{Z}$ is the hidden measure of $\boldsymbol{V}$. A sufficient condition for (3.11) is $\alpha_* > \alpha_0 - \alpha$

3. If

$$h(t)/b_0(t) \to c \in (0,\infty), \tag{3.12}$$

then $\boldsymbol{Z} \in MRV(\alpha, b(t), \nu, \mathbb{E})$ with asymptotic independence and $\boldsymbol{Z}$ has HRV with index $\alpha+\alpha_*$ and hidden measure which is a sum of the measure given in (3.10) and ν_0, the hidden measure of $\boldsymbol{V}$,

$$\nu_{\boldsymbol{Z}} = \frac{1}{2}\big(\nu_\alpha \times \nu_{\alpha_*} + \nu_{\alpha_*} \times \nu_\alpha\big) + \nu_0. \tag{3.13}$$

A sufficient condition for (3.11) is $\alpha_* = \alpha_0 - \alpha$.

Proof. Begin with the following observations for all cases: As $t \to \infty$,

$$t\mathbf{P}\Big[\Big(\frac{\xi_1}{h(t)}, \frac{V_2}{h(t)}\Big) \in \cdot\,\Big] \to \nu_\alpha \times \nu_{\alpha_*} \tag{3.14}$$

$$t\mathbf{P}\Big[\Big(\frac{V_1}{h(t)}, \frac{\xi_2}{h(t)}\Big) \in \cdot\,\Big] \to \nu_{\alpha_*} \times \nu_\alpha \tag{3.15}$$

in $\mathbb{M}((0,\infty)^2)$. To see this, write for $x>0$, $y>0$,

$$\begin{aligned} t\mathbf{P}[\xi_1 > h(t)x, V_2 > h(t)y] &= t\mathbf{P}[\xi_1 > b \circ b^{\leftarrow}(h)x]\mathbf{P}[V_2 > b_* \circ b_*^{\leftarrow}(h)y] \\ &= \frac{t}{b^{\leftarrow}(h) b_*^{\leftarrow}(h)} b^{\leftarrow}(h)\mathbf{P}[\xi_1 > b \circ b^{\leftarrow}(h)x] \\ &\qquad b_*^{\leftarrow}(h)\mathbf{P}[V_2 > b_* \circ b_*^{\leftarrow}(h)y] \\ &\sim \frac{t}{b^{\leftarrow}(h) b_*^{\leftarrow}(h)} x^{-\alpha} y^{-\alpha_*} \\ &\sim \nu_\alpha(x,\infty)\nu_{\alpha_*}(y,\infty). \end{aligned}$$

The proof of (3.15) is the same.

Now assume $f \in \mathcal{C}((0,\infty)^2)$ and f is bounded by $\|f\|$, uniformly continuous with

$$f(\boldsymbol{x}) = 0, \qquad \text{if } x_1 \wedge x_2 < \eta,$$

for some $\eta > 0$. Write

$$tEf\Big(\frac{\boldsymbol{Y}+\boldsymbol{V}}{h(t)}\Big) = \frac{t}{2}Ef\Big(\frac{\xi_1+V_1}{h(t)}, \frac{V_2}{h(t)}\Big) + \frac{t}{2}Ef\Big(\frac{V_1}{h(t)}, \frac{\xi_2+V_2}{h(t)}\Big) = A + B \tag{3.16}$$

For case (1) where (3.9) holds, we get a limit for A by writing

$$\begin{aligned} tE\Big|f\Big(\frac{\xi_1+V_1}{h(t)}, \frac{V_2}{h(t)}\Big) - f\Big(\frac{\xi_1}{h(t)}, \frac{V_2}{h(t)}\Big)\Big| &= tE\Big|\cdot\Big|1_{[V_1 < h(t)\delta, \xi_1 > h(t)(\eta-\delta), V_2 > h(t)\eta]} \\ &\quad + tE\Big|\cdot\Big|1_{V_1 > h(t)\delta, V_2 > h(t)\eta]} = I + II. \end{aligned}$$

Now

$$I \leqslant \omega_f(\delta) t\mathbf{P}[\xi_1 > h(t)(\eta-\delta), V_2 > h(t)\eta]$$

$$\to \omega_f(\delta)\nu_\alpha((\eta-\delta),\infty)\nu_{\alpha_*}(\eta,\infty)$$

from (3.14)

$$\to 0 \qquad (\delta \to 0).$$

We can control II by observing

$$\begin{aligned} II \leqslant & 2\|f\| t\mathbf{P}[V_1 > h(t)\delta, V_2 > h(t)\eta] \\ \leqslant & 2\|f\| \frac{t}{b_0^{\leftarrow}(h(t))} b_0^{\leftarrow}(h(t))\mathbf{P}[V_1 \wedge V_2 > b_0 \circ b_0^{\leftarrow}(h(t))\delta \wedge \eta] \\ \to & 0 \qquad (t \to \infty), \end{aligned}$$

from (3.9). The second term of (3.10) comes from B in a similar way to the derivation of A, relying on (3.15). This completes case (1) where (3.9) holds.

For Case (2) when (3.11) holds, replace $h(t)$ with $b_0(t)$ in (3.16) and focus on A. We compare with $f(\boldsymbol{V}/b_0(t))$:

$$\begin{aligned} & t\mathbf{E}\Big| f\Big(\frac{\xi_1+V_1}{b_0(t)}, \frac{V_2}{b_0(t)}\Big) - f\Big(\frac{\boldsymbol{V}}{b_0(t)}\Big)\Big| \\ & = t\mathbf{E}|\cdot|1_{[\xi_1 < b_0(t)\delta, V_1 > b_0(\eta-\delta), V_2 > b_0\eta]} + t\mathbf{E}|\cdot|1_{[\xi_1 > b_0(t)\delta, V_2 > b_0\eta]} = I + II. \end{aligned}$$

Since $t\mathbf{E}f(\boldsymbol{V}/b_0(t)) \to \int f d\nu_0$, we only have to show that both I and II go to zero. For I we have

$$\begin{aligned} I \leqslant & \omega_f(\delta)t\mathbf{P}[V_1 \wedge V_2 > b_o(t)(\eta-\delta)\wedge\eta] \to \omega_f(\delta)((\eta-\delta)\wedge\eta)^{-\alpha_0} \\ \to & 0 \qquad (\delta \to 0). \end{aligned}$$

Also using (3.11),

$$\begin{aligned} II \leqslant 2\|f\| & \frac{t}{b^{\leftarrow}(b_0)b_*^{\leftarrow}(b_0)}\Big(b^{\leftarrow}(b_0)\mathbf{P}[\xi_1 > b \circ b^{\leftarrow}(b_0)\delta] b_*^{\leftarrow}(b_0)\mathbf{P}[V_2 > b_* \circ b_*^{\leftarrow}(b_0)\eta]\Big) \\ & \sim (const)\frac{t}{b^{\leftarrow}(b_0)b_*^{\leftarrow}(b_0)} = (const)\frac{t}{h^{\leftarrow}(b_0)} \to 0. \end{aligned}$$

We can deal with the term B similarly so this completes treatment of Case (2).

Now consider Case (3) where (3.12) holds. Again replace $h(t)$ by $b_0(t)$ in (3.16) and consider A. Write

$$\begin{aligned} 2A = & t\mathbf{E}f\Big(\frac{\xi_1+V_1}{h(t)}, \frac{V_2}{h(t)}\Big)\Big(1_{[\xi_1 \leqslant b_0(t)\delta]} + 1_{[\xi_1 > b_0(t)\delta]}\Big) \\ = & t\mathbf{E}f\Big(\frac{\xi_1+V_1}{h(t)}, \frac{V_2}{h(t)}\Big) - f\Big(\frac{\boldsymbol{V}}{h(t)}\Big)1_{[\xi_1 \leqslant b_0(t)\delta]} \\ & + t\mathbf{E}f\Big(\frac{\xi_1+V_1}{h(t)}, \frac{V_2}{h(t)}\Big) - f\Big(\frac{\xi_1}{h(t)}, \frac{V_2}{h(t)}\Big)1_{[\xi_1 > b_0(t)\delta]} \\ & + tEf(\boldsymbol{V}/b_0(t))1_{[\xi_1 \leqslant b_0(t)\delta]} + t\mathbf{E}f\Big(\frac{\xi_1}{h(t)}, \frac{V_2}{h(t)}\Big)1_{[\xi_1 > b_0(t)\delta]} \\ = & a + b + c + d. \end{aligned}$$

We have $c \to \int f(\boldsymbol{x})\nu_0(d\boldsymbol{x})$ since $\mathbf{P}[\xi_1 \leqslant b_0(t)\delta] \to 1$. For d note

$$d = t\mathbf{E}f\Big(\frac{\xi_1}{h(t)}, \frac{V_2}{h(t)}\Big) \to \int f d\nu_\alpha \times \nu_{\alpha*}$$

using (3.14) and the fact that (3.12) is equivalent to $h^{\leftarrow}(t)/b_0^{\leftarrow}(t) \to c^{-1}$. Take the absolute value of a and add to the indicator the event $[V_1 > b_0(t)(\eta - \delta)]$ (otherwise both terms in the difference are zero) and

$$\begin{aligned}|a| \leqslant & \omega_f(\delta)t\mathbf{P}[V_1 \wedge V_2 > b_0(t)(\eta - \delta)]\\ \to & \omega_f(\delta)(\eta - \delta)^{-\alpha_0} \quad (t \to \infty)\\ \to & 0 \qquad (\delta \to 0).\end{aligned}$$

For b write

$$\begin{aligned}|b| \leqslant & t\mathbf{E}\Big|f\Big(\frac{\xi_1 + V_1}{h(t)}, \frac{V_2}{h(t)}\Big) - f\Big(\frac{\xi_1}{h(t)}, \frac{V_2}{h(t)}\Big)\Big|1_{[\xi_1 > b_0(t)\delta, V_1 \leqslant b_0(t)\delta]}\\ & + t\mathbf{E}|\cdot|1_{[\xi_1 > b_0(t)\delta, V_1 > b_0(t)\delta]} = |b1| + |b2|.\end{aligned}$$

We dominate $|b1|$ by using the modulus of continuity

$$|b1| \leqslant \omega_f(\delta)t\mathbf{P}[\xi_1 > b_0(t)\delta, V_2 > b_0(t)\eta]$$

where we added the condition on V_2 because otherwise, the probability would be zero due to the support of f being bounded away from the axes. Let $t \to \infty$, apply (3.14) and condition (3.12) and then let $\delta \to 0$. Dominate $|b2|$ by

$$\begin{aligned}|b2| \leqslant & 2\|f\|\mathbf{P}[\xi > b_0(t)\delta]t\mathbf{P}[V_1 \wedge V_2 > b_0(t)\delta]\\ \sim & (const)\delta^{-\alpha_0}\mathbf{P}[\xi > b_0(t)\delta] \to 0 \quad (t \to \infty).\end{aligned}$$

The terms involving B are handled similarly. □

EXAMPLE 3.2. We illustrate instances of the three cases given in Proposition 3.3. We simulate data samples from three different regimes as discussed in the Proposition 3.3 and estimate back the parameters of the additive model from which the data was generated.

Case 1: $\alpha_* < \alpha_0 - \alpha$. Let $\alpha = 0.5$, $\alpha_* = 1$, $\alpha_0 = 2$ and then $\alpha_* = 1 < 1.5 = \alpha_0 - \alpha$. Let $\boldsymbol{Y}$ have the form (3.1) where ξ_1, ξ_2 are iid Pareto random variables with parameter $\alpha = 0.5$. For $\boldsymbol{V}$ it is simplest to take $\boldsymbol{V} = (V_1, V_2)$ iid Pareto $\alpha^* = 1$ random variables and hence we do so. Then α_0 is the index of $V_1 \wedge V_2$ and so $\alpha_0 = 2$. It is easy to see that $\boldsymbol{Z} = \boldsymbol{Y} + \boldsymbol{V} \in \text{MRV}(\alpha = 0.5, t^2, \epsilon_{\{0\}} \times \nu_{1/2} + \nu_{1/2} \times \epsilon_{\{0\}}, \mathbb{E})$ with asymptotic independence of the marginals.
To verify that $\boldsymbol{Z} \in \text{MRV}(\alpha + \alpha_*, t^{1/(\alpha+\alpha_*)}, \nu_{\boldsymbol{Z},\text{hidden}}, \mathbb{E}_0) = \text{MRV}(3/2, t^{2/3}, \nu_{\boldsymbol{Z},\text{hidden}}, \mathbb{E}_0)$ ab initio, take $\boldsymbol{z} > \boldsymbol{0}$ and then

$$\begin{aligned}t\mathbf{P}[\boldsymbol{Z} > t^{2/3}\boldsymbol{z}] = & \frac{t}{2}\mathbf{P}[\xi_1 + V_1 > t^{2/3}z_1, V_2 > t^{2/3}z_2]\\ & + \frac{t}{2}\mathbf{P}[V_1 > t^{2/3}z_1, \xi_2 + V_2 > t^{2/3}z_2] = I + II.\end{aligned}$$

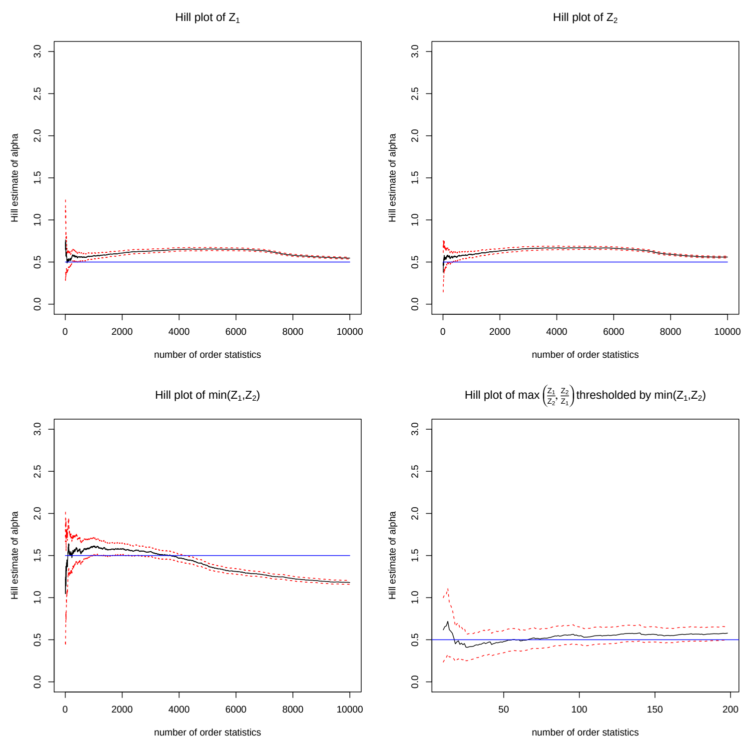

FIG 4. *Exploratory plots for Example 3.2, Case 1, with* $\alpha = 0.5, \alpha^* = 1, \alpha_0 = 2$. *Top panel: Hill plots for the marginals* Z_1 *and* Z_2. *Bottom left: Hill plot for* $\min\{Z_1, Z_2\}$. *Bottom right: Hill plot for* $\max\{Z_1/Z_2, Z_2/Z_1\}$ *thresholded by the 200 largest values of* $\min\{Z_1, Z_2\}$.

Focus on I as treatment of II is almost the same. We have

$$\begin{aligned}2I \sim & t(t^{2/3}z_1)^{-1/2}(t^{2/3}z_2)^{-1} = tt^{-2/3}t^{-1/3}z_1^{-1/2}z_2^{-1} \\ = & z_1^{-1/2}z_2^{-1},\end{aligned}$$

which is the first piece of the limit in (3.10).
Hence we can check that the limit measure $\nu_{\boldsymbol{Z},\text{hidden}}$ in (3.10) has density

$$\frac{1}{4}z_1^{-3/2}z_2^{-2} + \frac{1}{4}z_1^{-2}z_2^{-3/2}, \quad z_1 > 0, z_2 > 0$$

from which one can readily compute G_1 from (1.10) for $s > 1$ as

$$\bar{G}_1(s) = \nu_{\boldsymbol{Z},\text{hidden}}\{\boldsymbol{z} \in \mathbb{E}_0 : z_1/z_2 > s\} = (const)s^{-1/2}.$$

A similar calculation will lead to $G_2(s) = (const)s^{-1/2}, s > 1$ meaning both G_1 and G_2 have regularly varying tail distributions with index $1/2$. In fact they are both Pareto $(1/2)$ distributions. We generate 10000 iid samples following the construction of $\boldsymbol{Z} = \boldsymbol{Y} + \boldsymbol{V}$ described above and check whether we can estimate the regular variation index $\alpha = 0.5$, the hidden regular variation index $\alpha + \alpha^* = 1.5$ and the tail index of G_1 and G_2 from the sample. Figure

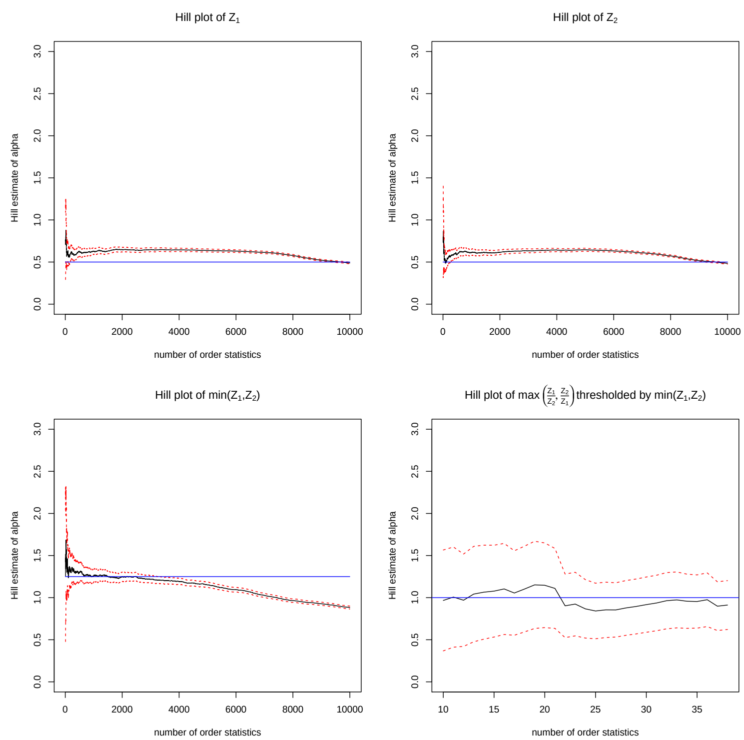

FIG 5. *Exploratory plots for Example 3.2, Case 2, with* $\alpha = 0.5, \alpha^* = 1, \alpha_0 = 1.25$. *Top panel: Hill plots for the marginals* Z_1 *and* Z_2. *Bottom left: Hill plot for* $\min\{Z_1, Z_2\}$. *Bottom right: Hill plot for* $\max\{Z_1/Z_2, Z_2/Z_1\}$ *thresholded by the 200 largest values of* $\min\{Z_1, Z_2\}$.

4 shows Hill plots for Z_1 and Z_2 in the top panel, both of which indicate that the marginals are heavy tailed with parameter $\alpha = 0.5$. The Hill plot of $\min\{Z_1, Z_2\}$ correctly identifies the HRV parameter $\alpha + \alpha^* = 1.5$. The final Hill plot of $\max\{Z_1/Z_2, Z_2/Z_1\}$ for the 200 highest order statistics of $\min\{Z_1, Z_2\}$ clearly indicates a heavy tail with a tail index of $1/2$ for both G_1 and G_2. Note since $G_1 = G_2$, (1.12) allows doing the estimation using the thresholded maxima of the component ratios.

Case 2: $\alpha + \alpha_* > \alpha_0$. Let $\alpha = 0.5$, $\alpha_* = 1$, $\alpha_0 = 1.25$ and then $\alpha_* = 1 > 0.75 = \alpha_0 - \alpha$. We generate $\boldsymbol{Y}$ in exactly the same way as in Case 1. For $\boldsymbol{V}$ we generate R, a Pareto $\alpha_0 = 1.25$ random variable, B a Bernoulli $(1/2)$ random variable and θ a Pareto $\alpha^* = 1$ random variable. Now define:

$$\boldsymbol{V} = BR(\theta, 1) + (1 - B)R(1, \theta).$$

As in Case 1, $\boldsymbol{Z} = \boldsymbol{Y} + \boldsymbol{V} \in \mathrm{MRV}(\alpha = 0.5, t^2, \epsilon_{\{0\}} \times \nu_{1/2} + \nu_{1/2} \times \epsilon_{\{0\}}, \mathbb{E})$ and furthermore $\boldsymbol{Z} = \boldsymbol{Y} + \boldsymbol{V} \in \mathrm{MRV}(\alpha_0, t^{1/\alpha_0}, \mathbb{E}_0) = \mathrm{MRV}(1.25, t^{1/1.25}, \mathbb{E}_0)$. Moreover by construction we have $G_1(s) = G_2(s) = s^{-1}, s > 1$. Of course this is also clear from Proposition 3.3.

We generate 10000 iid samples using the construction of $\boldsymbol{Z} = \boldsymbol{Y} + \boldsymbol{V}$ and from this sample we estimate the regular variation index $\alpha = 0.5$, the hidden regular variation index $\alpha_0 = 1.25$ and the tail index of G_1 and G_2 which is 1. The top panels in Figure 5 display Hill plots for Z_1 and Z_2 that indicate the same tail index of $\alpha = 0.5$. The Hill plot for $\min\{Z_1, Z_2\}$ correctly indicates a tail index of $\alpha_0 = 1.25$. Finally, the Hill plot of $\max\{Z_1/Z_2, Z_2/Z_1\}$ for the 200 highest order statistics of $\min\{Z_1, Z_2\}$ indicates a tail index of $\alpha^* = 1$ for both $G_1 \equiv G_2$.

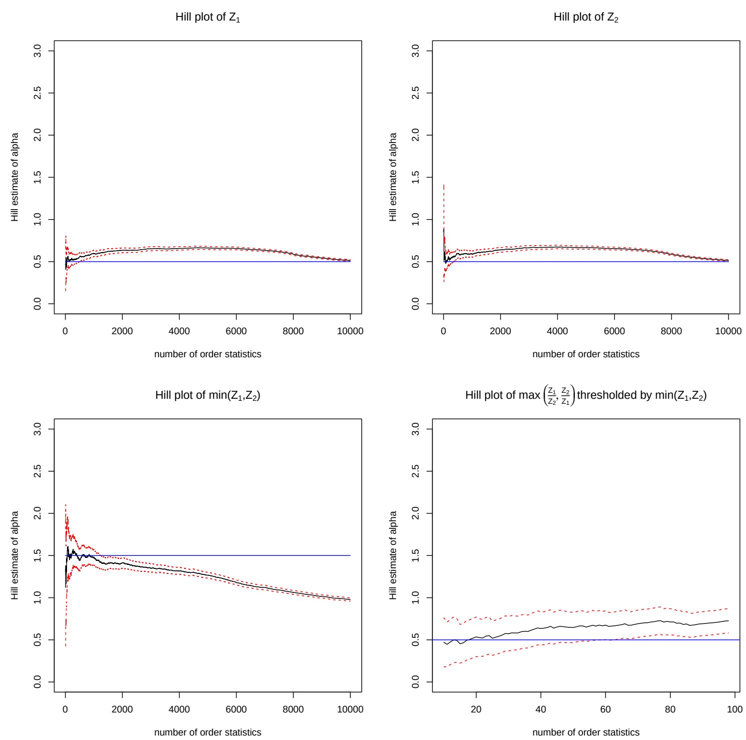

FIG 6. *Exploratory plots for Example 3.2, Case 3, with $\alpha = 0.5, \alpha^* = 1, \alpha_0 = 1.5$. Top panel: Hill plots for the marginals Z_1 and Z_2. Bottom left: Hill plot for $\min\{Z_1, Z_2\}$. Bottom right: Hill plot for $\max\{Z_1/Z_2, Z_2/Z_1\}$ thresholded by the 200 largest values of $\min\{Z_1, Z_2\}$.*

Case 3: $\alpha + \alpha_* = \alpha_0$. Let $\alpha = 0.5$, $\alpha_* = 1$, $\alpha_0 = 1.5$ which satisfies $\alpha + \alpha_* = 1.5 = \alpha_0$. We generate $\boldsymbol{Y}$ as in Case 1 or 2 and generate $\boldsymbol{V}$ using the method of Case 2, except that now R is generated from a Pareto $\alpha_0 = 1.5$ distribution. We verify that $\boldsymbol{Z} = \boldsymbol{Y} + \boldsymbol{V} \in \mathrm{MRV}(\alpha = 0.5, t^2, \epsilon_{\{0\}} \times \nu_{1/2} + \nu_{1/2} \times \epsilon_{\{0\}}, \mathbb{E})$ and $\boldsymbol{Z} = \boldsymbol{Y} + \boldsymbol{V} \in \mathrm{MRV}(1.5, t^{1/1.5}, \nu_{\boldsymbol{Z}}, \mathbb{E}_0)$. Getting the distribution of G_1 and G_2 is more difficult in this case since the hidden limit measure for $\boldsymbol{Z}$ is more complicated as can be seen in (3.13). A careful calculation shows that G_1 and G_2 have regularly varying tails with index 0.5.

We generate 10000 iid samples of $\boldsymbol{Z} = \boldsymbol{Y} + \boldsymbol{V}$ using this model . In Figure 6 the Hill plots for Z_1 and Z_2 are in the neighborhood of $\alpha = 0.5$ and the Hill plot for $\min\{Z_1, Z_2\}$ correctly indicates a tail index of $\alpha_0 = 1.5$ The Hill plot of $\max\{Z_1/Z_2, Z_2/Z_1\}$ for the 200 highest order statistics of $\min\{Z_1, Z_2\}$ indicates a tail index of $\alpha^* = 0.5$ for both $G_1 \equiv G_2$ which was what we were expecting.

□

4. Detection and estimation: regular variation and hidden regular variation. What diagnostic tools exist to help us verify that multivariate data come from a distribution possessing regular variation on some domain? Since regular variation is only an asymptotic tail property, the task of deciding to use a multivariate regularly varying model is challenging.

Suppose we have $\boldsymbol{Z} = (Z_1, Z_2)$ multivariate regularly varying on $\mathbb{E} = [0,\infty)^2 \setminus \{\mathbf{0}\}$. Under the transformation GPOLAR as defined in (1.4), $\|\boldsymbol{Z}\|$ is regularly varying with some tail index α and (1.5) holds. Diagnostics that investigate if $\boldsymbol{Z}$ is regularly varying often reduce the data to one dimension for instance by taking norms or max-linear combinations of $\boldsymbol{Z}$ [23, Chapter 8] and then apply one dimensional heavy tail diagnostics such as Hill or QQ plotting. We propose further diagnostics for the viability of a multivariate regularly varying model using the GPOLAR transformation since GPOLAR converts a regularly varying model to a *conditional extreme value* (CEV) model for which detection techniques exist [6].

4.1. *Detecting multivariate regular variation using the CEV model.* The *conditional extreme value model* [5, 6, 13] requires at least one of the marginals of the distribution be in the domain of attraction of an extreme value distribution. In this section we discuss a modified version of the CEV model for bivariate random vectors in the non-negative orthant where convergences are described according to the notion of $\mathbb{M}$-convergence [8, 18]. Define

$$\mathbb{E}_{\sqsupset} := (0,\infty) \times [0,\infty) = [0,\infty)^2 \setminus ([0,\infty) \times \{0\}) \,.$$

Definition 4.1. Suppose $(\xi,\eta) \in \mathbb{R}_+^2$ is a random vector and there exist functions $a(t) \to \infty$, $b(t) > 0$ for $t > 0$ and a non-null measure $\mu \in \mathbb{M}(\mathbb{E}_{\sqsupset})$ such that in

$$tP\left[\left(\frac{\xi}{a(t)}, \frac{\eta}{b(t)}\right) \in \ \cdot\ \right] \to \mu(\cdot), \qquad \text{in } \mathbb{M}(\mathbb{E}_{\sqsupset}). \tag{4.1}$$

Additionally assume that

(a) $\mu((r,\infty] \times [0,s])$ is a non-degenerate measure in $s \in [0,\infty)$ for any fixed $r > 0$, and,
(b) $H(s) := \mu((1,\infty \times [0,s]))$ is a probability distribution.

Then we say (ξ,η) satisfies a conditional extreme value model and write $(\xi,\eta) \in \mathrm{CEV}(a,b,\mu)$.

Remark 4.1. The definition has some consequences [13, Section 2]:

1. Convergence in (4.1) implies that ξ is regularly varying with some tail index $\alpha > 0$. Consequently $a(t) \in RV_{1/\alpha}$.
2. The limit μ is a product measure of the form
$$\mu((r,\infty) \times [0,s]) = r^{-\alpha} H(s) =: \nu_\alpha(r,\infty) H(s)$$
for all $(r,s) \in \mathbb{E}_{\sqsupset}$ if and only if
$$\lim_{t\to\infty} \frac{b(tc)}{b(t)} = 1.$$
3. If $a(t) = b(t), t > 0$ then (ξ,η) is multivariate regularly varying on $\mathbb{E}_{\sqsupset}$ with limit measure μ. (In such a case μ cannot be a product measure).

REMARK 4.2. Statistical plots that check whether bivariate data can be modelled by a CEV model were derived in [5] and are based on the Hillish, Pickandsish and Kendall's Tau statistics. If data is generated from a CEV model, these statistics tend to a constant as the sample size increases. We concentrate on the Hillish and Pickandsish statistics for this paper. We will further specialize to the case where μ is a product measure $\mu = \nu_\alpha \times H$ for reasons that will be clear in the next subsection.

Suppose $(\xi_i, \eta_i); 1 \leqslant i \leqslant n$ are iid samples in $\mathbb{R}_+^2$ and $(\xi_1, \eta_1) \in \mathrm{CEV}(a, b, \mu)$ for some $a(t) \to \infty, b(t) > 0$ and $\mu \in \mathbb{M}(\mathbb{E}_\sqsupset)$. We use the following notation:

$\xi_{(1)} \geqslant \ldots \geqslant \xi_{(n)}$	The decreasing order statistics of $\xi_1, \ldots, \xi_n$.
$\eta_i^*,\ 1 \leqslant i \leqslant n$	The η-variable corresponding to $\xi_{(i)}$, also called the concomitant of $\xi_{(i)}$.
$N_i^k = \sum_{l=i}^{k} \mathbf{1}_{\{\eta_l^* \leqslant \eta_i^*\}}$	Rank of η_i^* among $\eta_1^*, \ldots, \eta_k^*$. We write $N_i = N_i^k$.
$\eta_{1:k}^* \leqslant \eta_{2:k}^* \leqslant \ldots \leqslant \eta_{k:k}^*$	The increasing order statistics of $\eta_1^*, \ldots, \eta_k^*$.

Hillish statistic. For $1 \leqslant k \leqslant n$, the *Hillish statistic* is

$$\mathrm{Hillish}_{k,n} = \mathrm{Hillish}_{k,n}(\xi, \eta) := \frac{1}{k} \sum_{j=1}^{k} \log \frac{k}{j} \log \frac{k}{N_j^k} \tag{4.2}$$

PROPOSITION 4.1 (Proposition 2.2 and Proposition 2.3 [6]). *Suppose (ξ_i, η_i); $1 \leqslant i \leqslant n$ are iid observations from the CEV(a, b, μ) model as in Definition 4.1 and suppose H is continuous. If $k = k(n) \to \infty$, $n \to \infty$ and $k/n \to 0$, then*

$$\mathit{Hillish}_{k,n} \xrightarrow{P} \int_1^\infty \int_1^\infty \mu((r^{\frac{1}{\alpha}}, \infty) \times [0, H^{\leftarrow}(s^{-1})]) \frac{dr}{r} \frac{ds}{s} =: I_\mu. \tag{4.3}$$

Moreover μ is a product measure if and only if both

$$\mathit{Hillish}_{k,n}(\xi, \eta) \xrightarrow{P} 1 \quad \text{and} \quad \mathit{Hillish}_{k,n}(\xi, -\eta) \xrightarrow{P} 1.$$

Proof. The proof follows from Propositions 2.2 and 2.3 in [6]. The only difference here is the use of measure μ instead of μ^* and the roles of the first and the second components are switched.

Pickandsish statistic. This statistic gives another way to check the suitability of the CEV assumption and to detect a product measure in the limit. The Pickandsish statistic is based on ratios of differences of ordered concomitants and is patterned on the Pickands estimate for the scale parameter of an extreme value distribution (Pickands [20], de Haan and Ferreira [9, page 83], Resnick [23, page 93]). For notational convenience for $s \leqslant t$ write $\eta_{s:t}^* := \eta_{\lceil s \rceil : \lceil t \rceil}^*$. We define the Pickandsish statistic for $0 < q < 1$ as

$$\mathrm{Pickandsish}_{k,n}(q) := \frac{\eta_{qk:k}^* - \eta_{qk/2:k/2}^*}{\eta_{qk:k}^* - \eta_{qk/2:k}^*}. \tag{4.4}$$

Proposition 4.2 (Proposition 2.4 and Corollary 2.5 [6]). *Suppose (ξ_i, η_i); $1 \leqslant i \leqslant n$ are iid observations from the CEV(a, b, μ) model as in Definition 4.1. Assume that $k = k(n) \to \infty$, $n \to \infty$ and $k/n \to 0$. Then*

$$\mathit{Pickandsish}_{k,n}(q) \xrightarrow{P} \frac{H^{\leftarrow}(q)(1-2^{\rho})}{H^{\leftarrow}(q) - H^{\leftarrow}(q/2)}, \tag{4.5}$$

provided $H^{\leftarrow}(q) - H^{\leftarrow}(q/2) \neq 0$. Here $\rho = (\log(c))^{-1} \log\left(\lim_{t\to\infty} \frac{b(tc)}{b(t)}\right)$. Moreover, μ is a product measure if and only if

$$\mathit{Pickandsish}_{k,n}(q) \xrightarrow{P} 0$$

for some $0 < q < 1$ where $H^{\leftarrow}(q) - H^{\leftarrow}(q/2) \neq 0$.

Proof. The proof follows from Proposition 2.4 in [6]. The second part is immediate from (4.5).

4.2. *Relating MRV and CEV.* We have methods to detect a CEV model and indicate when the limit is a product measure. What is the connection with multivariate regular variation? This connection is given in (1.6)–(1.9). Regular variation of a vector $\boldsymbol{Z}$ on $\mathbb{E}$ and $\mathbb{E}_0$ with scaling functions $b(t) \in RV_{1/\alpha}$ and $b_0(t) \in RV_{1/\alpha_0}$ respectively with $0 < \alpha \leqslant \alpha_0$ is equivalent to:

$$tP\big[\big(\|\boldsymbol{Z}\|/b(t), \boldsymbol{Z}/\|\boldsymbol{Z}\|\big) \in \cdot\,\big] \to \nu_\alpha \times S(\cdot), \qquad \text{in } \mathbb{M}((0,\infty)\times\aleph_{\mathbf{0}}) \tag{4.6}$$

and

$$tP\Big[\Big(\frac{Z_1 \wedge Z_2}{b_0(t)}, \frac{\boldsymbol{Z}}{Z_1 \wedge Z_2}\Big) \in \cdot\Big] \to \nu_{\alpha_0} \times S_0(\cdot) \qquad \text{in } \mathbb{M}\big((0,\infty)\times\aleph_{[\text{axes}]}\big). \tag{4.7}$$

If $\aleph_{\mathbf{0}}$ and $\aleph_{[\text{axes}]}$ were subsets of $[0,\infty)$ we could conclude that (4.6) and (4.7) describe CEV models and modest changes, described in the next two results, allow use of the CEV model diagnostics.

Proposition 4.3. *Suppose $\boldsymbol{Z}$ is a random element of $\mathbb{R}_+^2$. Fix a norm for $\boldsymbol{z} \in \mathbb{R}_+^2 : \|(z_1, z_2)\| = z_1 + z_2$. Then $\boldsymbol{Z} \in MRV(\alpha, b(t), \nu, \mathbb{E})$ (which means (1.8) also holds) if and only if $\left(\|Z\|, \frac{Z_1}{\|Z\|}\right) \in CEV(b, 1, \mu)$ with limit measure $\mu = \nu_\alpha \times \bar{S}$ where $\bar{S}(A) = S((x,y) \in \aleph_{\mathbf{0}} : x \in A)$ for any $A \in \mathcal{B}[0,\infty)$.*

Proof. The proof is easily deducible from the relationship between S and $\bar{S}$.

Proposition 4.4. *Suppose $\boldsymbol{Z} \geqslant 0$ is regularly varying on $\mathbb{E}$ with function $b(t) \in RV_{1/\alpha}$. Then $\boldsymbol{Z}$ exhibits HRV on $\mathbb{E}_0$ with scaling function $b_0(t) \in RV_{1/\alpha_0}, \alpha_0 \geqslant \alpha$ if and only if*

$$\left(Z_1 \wedge Z_2, \left(\frac{Z_1}{Z_2} \bigvee \frac{Z_2}{Z_1}\right)\right) \in CEV(b_0, 1, \mu_0)$$

with limit measure given by $\mu_0 = \nu_{\alpha_0} \times (pG_1 + (1-p)G_2)$ where $G_1(s) = S_0([1,s]\times\{1\})$ and $G_2(s) = S_0(\{1\}\times[1,s])$ for $s \geqslant 1$ and $G_1(s) = G_2(s) = 0, s \leqslant 1$.

Proof. The proof follows from the connection between S_0 and G_1, G_2. $\square$

5. Testing for MRV and HRV: data examples. Here we analyze data sets to see whether a multivariate regularly varying model is a valid assumption. We also look for asymptotic independence and if it exists we test for the existence of hidden regular variation.

EXAMPLE 5.1. **Boston University: HTTP downloads.** The first data set is obtained from

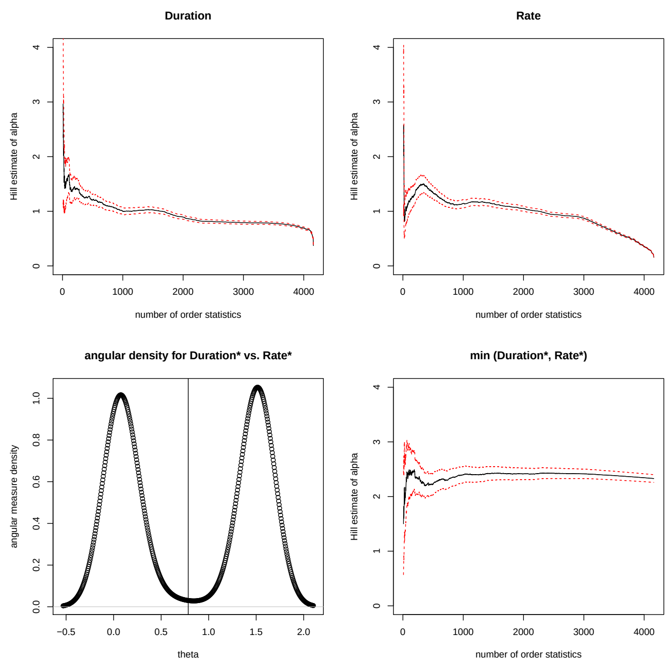

FIG 7. *BU dataset. Top panel: Hill plots of tail parameters for* D *and* R*. Bottom left plot: angular density of* (D^*, R^*)*. Bottom right plot: Hill plot for* $\min(D^*, R^*)$*.*

a now classical Boston University study [3] which suggested self-similarity and heavy tails in web-traffic data. Our dataset was created from HTTP downloads in sessions initiated by logins at a Boston University computer laboratory. It consists of 8 hours 20 minutes worth of downloads in February 1995 after applying an aggregation rule to downloads to associate machine triggered actions with human requests and is discussed in [12, page 176]. The result of the aggregation is 4161 downloads which are characterized by the following variables:

- S = the size of the download in kilobytes,
- D = the duration of the download in seconds,
- R = throughput of the download; that is, $= S/D$.

Previous studies [23, page 299, 316] indicate heavy tailed behavior of all three variables and asymptotic independence between D and R. We concentrate on the variables (D, R) so our data is $\{(D_i, R_i); 1 \leqslant i \leqslant 4161\}$. Moreover the rank-transformed variables are denoted:

$$D_i^* = \sum_{j=1}^{4161} \mathbf{1}_{\{D_i \geqslant D_j\}}, \qquad R_i^* = \sum_{j=1}^{4161} \mathbf{1}_{\{R_i \geqslant R_j\}}.$$

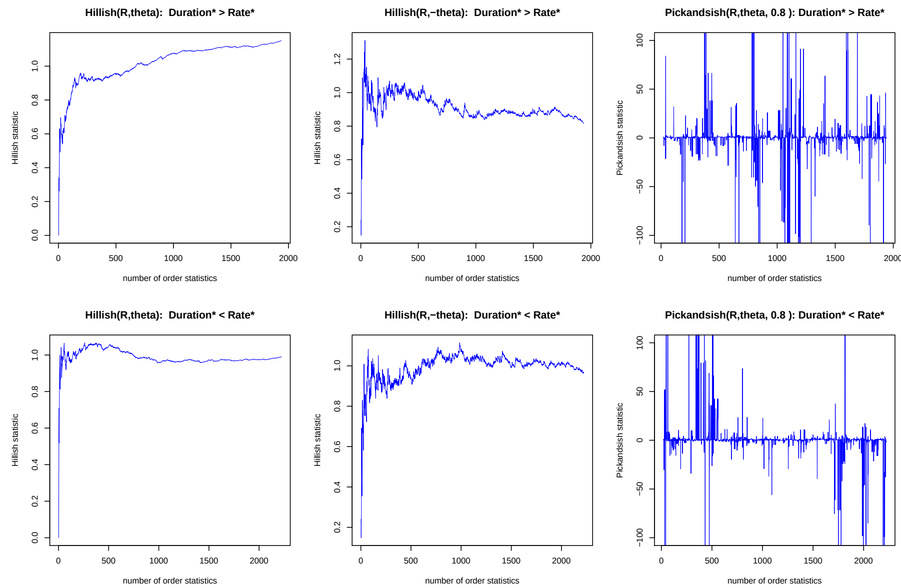

Fig 8. *BU dataset. Top panel ($D^* > R^*$): Hillish plots for (A, θ_1) and $(A, -\theta_1)$ and Pickandsish plot for (A, θ_1) at $q = 0.8$. Bottom panel ($D^* < R^*$): Hillish plots for (A, θ_2) and $(A, -\theta_2)$ and Pickandsish plot for (A, θ_2) at $q = 0.8$.*

for $1 \leqslant i \leqslant 4161$ with the generic rank-transformed variables denoted D^* and R^* respectively.

In Figure 7 we plot Hill estimates of the tail parameters of D and R for increasing number of order statistics of their respective univariate data values. Both plots are consistent with D and R being heavy tailed with tail parameters α_D and α_R greater than 1. (This is confirmed [10, 23, 25] by altHill and QQ plots (*not shown*) showing $\hat{\alpha}_D = 1.4$ and $\hat{\alpha}_R = 1.2$.) The angular density plot of (D^*, R^*) shows data concentration near 0 and $\pi/2$ consistent with asymptotic independence of the quantities. Asymptotic independence does not automatically imply HRV so we check for HRV

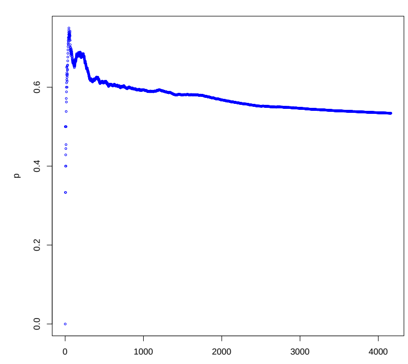

Fig 9. *BU dataset. Proportion of data with $D_i^* > R_i^*$ for order statistics of $A_i = \min\{D_i^*, R_i^*\}$.*

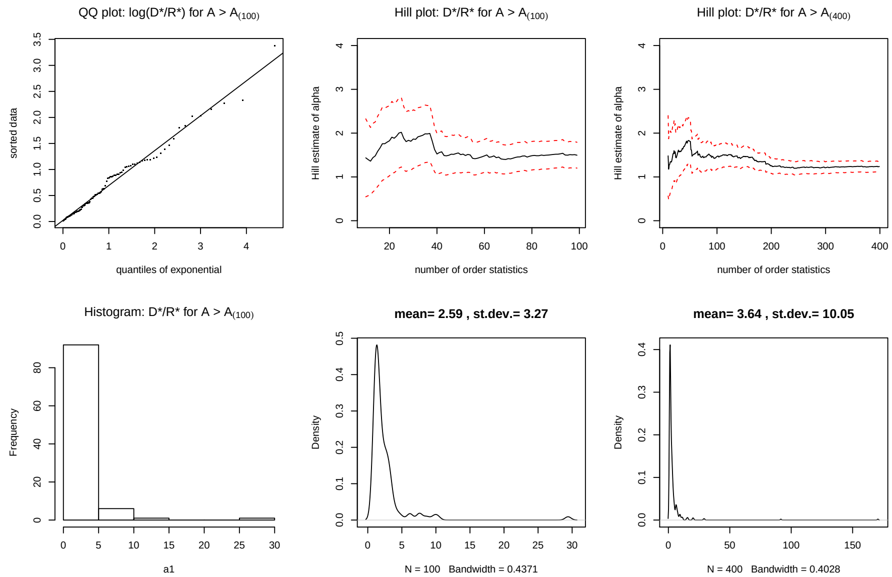

FIG 10. *BU dataset. Top panel: QQ plot of* $\log(D^*/R^*)$ *when* $A_i > A_{(100)}$ *and Hill plots of* D^*/R^* *when* $A_i > A_{(100)}$ *and* $A_i > A_{(400)}$. *Bottom panel: Histogram of* D^*/R^* *when* $A_i > A_{(100)}$ *and kernel density estimates of* D^*/R^* *when* $A_i > A_{(100)}$ *and* $A_i > A_{(400)}$.

on $\mathbb{E}_0$.

1. Is the variable $A = \min\{D^*, R^*\}$ regularly varying with parameter greater than 1? The bottom right plot in Figure 7 plots Hill estimates for increasing number of order statistics of A and stabilizes between 2 and 3 indicating the desired heavy tail behavior.
2. For $D^* > R^*$, we check whether $(A, \theta_1) := (\min\{D^*, R^*\}, \frac{D^*}{R^*})$ follows a CEV model. In the top panel of Figure 8, the Hillish plots of (A, θ_1) and $(A, -\theta_1)$ are close to 1 near the left side of their plots. Moreover we observe that the Pickandsish estimate at $q = 0.8$ also remains near 0. From Propositions 4.1 and 4.2, this is consistent with $(A, \theta_1) \in \mathrm{CEV}(b_0, 1, \mu_0)$ with a limit measure of the CEV being a product measure.
3. For $D^* < R^*$, we similarly check whether $(A, \theta_2) := (\min\{R^*, D^*\}, \frac{R^*}{D^*})$ follows a CEV model. In the bottom panel of Figure 8 we observe that the Hillish plots of (A, θ_2) and $(A, -\theta_2)$ are close to 1 near the left side of their plots. We also observe that the Pickandsish estimate at $q = 0.8$ remains near 0. Hence we again conclude that the evidence is consistent with $(A, \theta_2) \in \mathrm{CEV}(b_0, 1, \mu_0)$ with a limit measure of the CEV being a product measure.

The rank transformation causes (D^*, R^*) to be standard regularly varying with $\alpha = 1$ and Proposition 4.4 implies (D^*, R^*) has hidden regular variation on $\mathbb{E}_0$ if (and only if)

$$(A, \theta) := \left(\min\{D^*, R^*\}, \max\left\{\frac{D^*}{R^*}, \frac{R^*}{D^*}\right\}\right) \in \mathrm{CEV}(b_0, 1, \mu_0)$$

for some function b_0. We proceed by testing the following:

Thus modeling the joint distribution of (D, R) using MRV and HRV is consistent with the data. The next step is to estimate the distributions of $\theta_1 \sim G_1$ and $\theta_2 \sim G_2$ as well as q defined

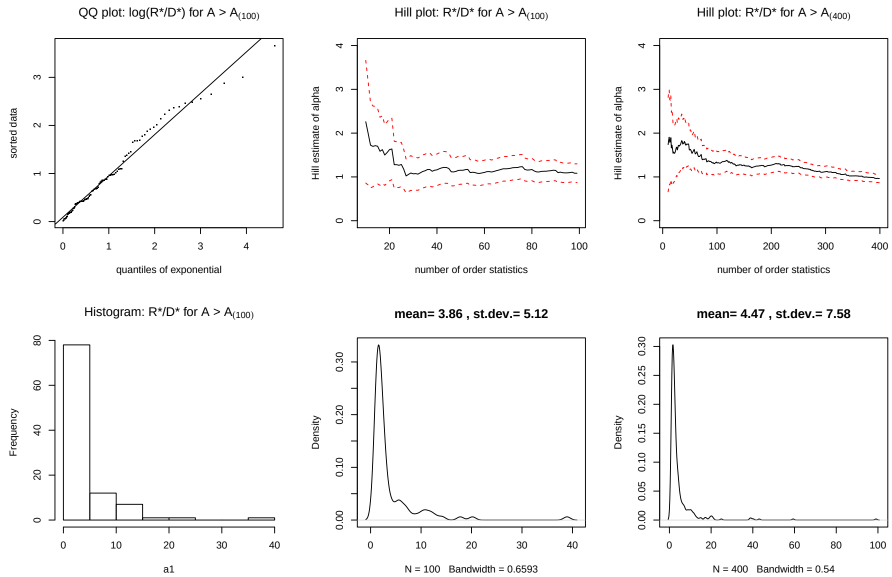

FIG 11. *BU dataset. Top panel: QQ plot of* $\log(R^*/D^*)$ *when* $A_i > A_{(100)}$ *along with Hill plots of* R^*/D^* *when* $A_i > A_{(100)}$ *and* $A_i > A_{(400)}$*. Bottom panel: Histogram of* R^*/D^* *when* $A_i > A_{(100)}$ *and kernel density estimates of* R^*/D^* *when* $A_i > A_{(100)}$ *and* $A_i > A_{(400)}$*.*

in Proposition 4.4. Figure 9 plots $\hat{q}_k = \frac{1}{k}\sum_{i=1}^{4161} \mathbf{1}_{\{D_i^* > R_i^*, A_i > A_{(k)}\}}, \quad k = 2, \ldots, 4161$, where $A_i = \min\{R_i^*, D_i^*\}$ and $A_{(1)} \geqslant A_{(2)} \geqslant \ldots$ form order statistics from $A_i; 1 \leqslant i \leqslant 4161$. Observing Figure 9 for k near 0, an estimate of q is $\hat{q} = 0.6$.

To find the distribution of θ_1 we make a standard exponential QQ plot of $\log(D_i^*/R_i^*)$ where $A_i = \min(D_i^*, R_i^*) > A_{(100)}$, which serves as an exploratory diagnostic for heavy tails. We also create Hill plots for D_i^*/R_i^* where $A_i > A_{(k)}$ for two choices of k. The top panels of Figure 10 give the QQ plot for $k = 100$ (left) and the Hill plots for $k = 100$ and 400 (middle and right). The bottom panels in Figure 10 have a histogram of D_i^*/R_i^* for $A_i > A_{(100)}$ (left) and kernel density plots of D_i^*/R_i^* for $A_i > A_{(100)}$ (middle) and $A_i > A_{(100)}$ (right). The plots indicate G_1 is heavy tailed with an index between 1.5 and 2 and we can provide a density estimate for the distribution of θ_1.

We create the same set of plots for finding G_2 in Figure 11 which also indicates towards a similar conclusion of heavy tailed behavior for G_2 with an index close to but less than 2.

EXAMPLE 5.2. **UNC Chapel Hill HTTP response data.** A *response* is the data transfer resulting from an HTTP request. The data set [14] consists of 21,828 thresholded responses bigger than 100 kilobytes measured between 1:00pm and 5:00pm on 25th April, 2001. We use similar notation as in Example 5.1.

- $S =$ HTTP response size; total size of packets transferred in kilobytes,
- $D =$ the elapsed duration between first and last packets in seconds of the response,
- $R =$ throughput of the response $= S/D$.

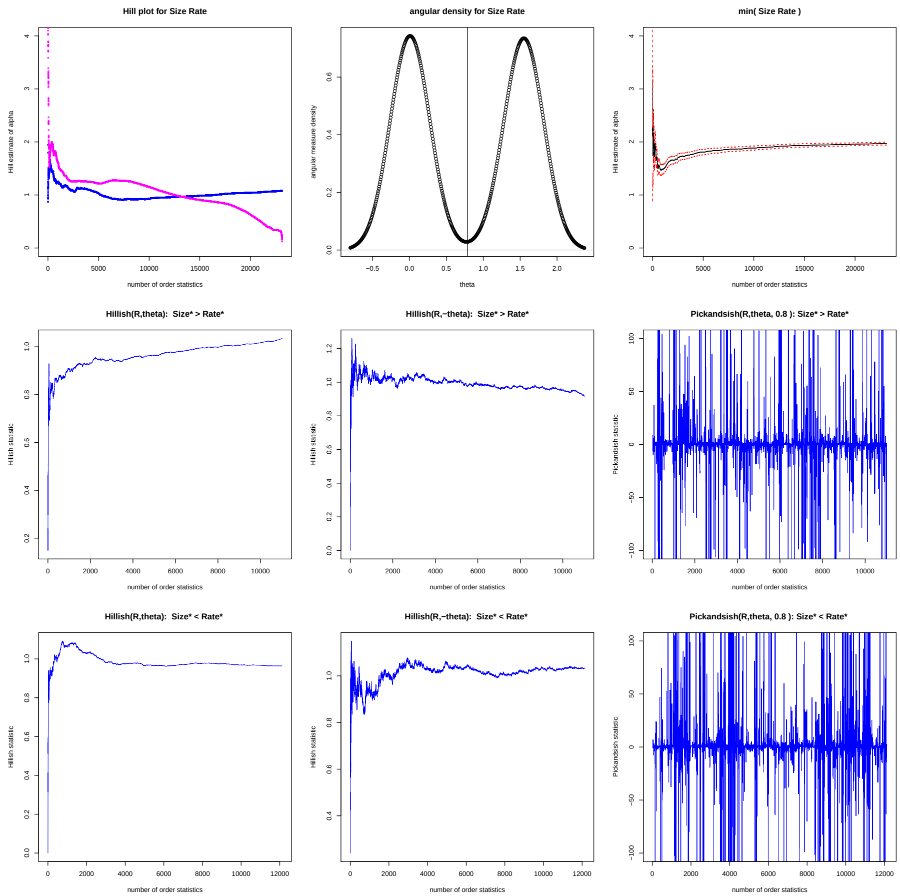

FIG 12. *UNC HTTP responses dataset. Top panel: (Left:) Hill plots of tail parameters for S(blue), R(magenta); (Middle:) angular density of* (S^*, R^*)*; (Right:) Hill plot for* $\min(S^*, R^*)$*). Middle panel (*$S^* > R^*$*): Hillish plots for* (A, θ_1) *and* $(A, -\theta_1)$ *and Pickandsish plot for* (A, θ_1) *at* $q = 0.8$*. Bottom panel (*$S^* < R^*$*): Hillish plots for* (A, θ_2) *and* $(A, -\theta_2)$ *and Pickandsish plot for* (A, θ_2) *at* $q = 0.8$*.*

Thus, the data set consists of $\{(S_i, D_i, R_i); 1 \leqslant i \leqslant 21828\}$. Our interest is in the variables (S, R) which exhibit heavy tails and asymptotic independence [14]. Denote the rank-transformed variables:

$$\Big(S_i^* = \sum_{j=1}^{21828} \mathbf{1}_{\{S_i \geqslant S_j\}}, R_i^* = \sum_{j=1}^{21828} \mathbf{1}_{\{R_i \geqslant R_j\}}\Big), \quad 1 \leqslant i \leqslant 21828,$$

with the generic rank-transformed variables denoted S^* and R^* respectively. The top left plots in Figure 12 give Hill plots of the tail indices of the distributions of S and R and suggest these indices are between 1 and 2. Asymptotic independence of S, R is exhibited in the angular density plot (top middle plot) for (S^*, R^*).

We next inquire if HRV exists on $\mathbb{E}_0$. The Hill plot for $\min(S^*, R^*)$ on the upper right panel of Figure 12 gives a tail estimate $\hat{\alpha}_0$ clearly greater than 1 and is consistent with HRV. We transform

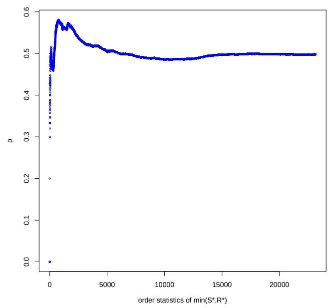

FIG 13. *UNC HTTP responses. Proportion of data with $S_i^* > R_i^*$ for order statistics of $A_i = \min\{S_i^*, R_i^*\}$.*

the data $\{(S^*, R^*); 1 \leqslant i \leqslant 21828\}$ with the transformation GPOLAR) to obtain:

$$(A, \theta) := \text{GPOLAR}(S^*, R^*) = \left(\min\{S^*, R^*\}, \max\left\{\frac{S^*}{R^*}, \frac{R^*}{S^*}\right\}\right).$$

From Proposition 4.4 we know $(A, \theta) \in \text{CEV}(b_0, 1, \mu_0)$ for some function b_0 and measure μ_0 on $\mathbb{E}_0$. For both the cases $S^* > R^*$ (see middle panels in Figure 12) and $S^* < R^*$ (see bottom panels in Figure 12), we employ the Hillish and Pickandsish diagnostics to check consistency of $(A, \theta_1) := (\min\{S^*, R^*\}, S^*/R^*)$ and $(A, \theta_2) := (\min\{S^*, R^*\}, R^*/S^*)$ with the CEV model with product limit measure. The Hillish plots are reasurringly hovering at height 1 and the Pickandsish plots center at 0.

So we have accumulated evidence that the data is consistent with an HRV model on $\mathbb{E}_0$. Now we proceed to provide some estimates on the structure of the hidden angular measure, which boils down to estimating three things

1. The proportion q appearing in μ_0 in Proposition 4.4: this can be estimated by

$$\hat{q}_k = \frac{1}{k}\sum_{i=1}^{21828} \mathbf{1}_{\{S_i^* > R_i^*, A_i > A_{(k)}\}}, \quad k = 2, \ldots, 21,828.$$

   where $A_i = \min\{S_i^*, R_i^*\}$ and $A_{(1)} \geqslant A_{(2)} \geqslant \ldots$ form order statistics from $A_i; 1 \leqslant i \leqslant 21,828$ as in Figure 13. Looking at the plot for k near zero, we can estimate $\hat{p} = 0.55$.
2. The distribution of $\theta_1 \sim G_1$: see Figure 14. First we make a standard exponential QQ plot of $\log(S_i^*/R_i^*)$ when $A_i > A_{(100)}$. This acts as a diagnostic for heavy-tails. This plot clearly indicates against heavy-tails as does a Hill plot of S_i^*/R_i^* when $A_i > A_{(100)}$. A histogram and kernel density estimate plot of (S_i^*/R_i^*) for $A_i > A_{(100)}$ points towards a light-tailed distribution.
3. The distribution of $\theta_2 \sim G_2$: see Figure 15. As before, first we make a standard exponential QQ plot of $\log(R_i^*/S_i^*)$ when $A_i > A_{(100)}$, and the points nicely hug a straight line which indicates presence of heavy-tails. The Hill plots of R_i^*/S_i^* when $A_i > A_{(100)}$ and $A_i > A_{(400)}$ provide an estimate of the tail index to be between 1 and 1.5. The histograms and kernel density estimates seem to support that the distribution of G_2 is heavy-tailed.

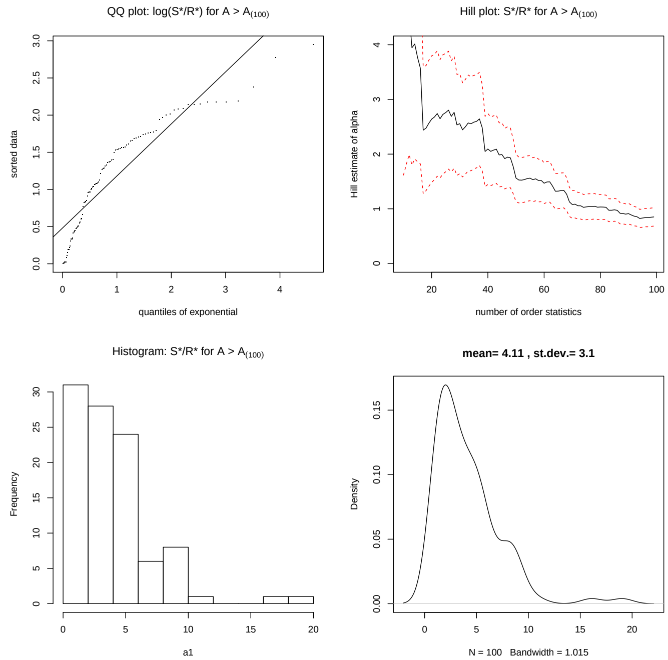

FIG 14. *UNC HTTP responses: Top: QQ plot of* $\log(S^*/R^*)$ *when* $A_i > A_{(100)}$ *along with Hill plots of* S^*/R^* *when* $A_i > A_{(100)}$ *and* $A_i > A_{(400)}$*. Bottom: Histogram and kernel density estimates of* S^*/R^* *when* $A_i > A_{(100)}$

6. Conclusion. In this paper we have discussed different techniques to generate models which exhibit both regular variation and hidden regular variation. We have seen some simulated examples where we can estimate the parameters of both MRV and HRV but there are also examples where it is difficult to correctly estimate parameters. We restricted ourselves to the two dimensional non-negative orthant here, but clearly some of the generation techniques can be extended to higher dimensions. Moreover, the detection techniques for HRV on $\mathbb{E}_0$ using the CEV model can also be extended to detect HRV on other types of cones especially in two dimensions but perhaps even more. Overall this paper serves as a starting point for methods of generating and detecting multivariate heavy tailed models having tail dependence explained by HRV.

REFERENCES

[1] P. L. Anderson and M. M. Meerschaert. Modeling river flows with heavy tails. *Water Resources Research*, 34 (9):2271–2280, 1998. ISSN 1944-7973. . URL http://dx.doi.org/10.1029/98WR01449.

[2] L. Breiman. On some limit theorems similar to the arc-sin law. *Theory Probab. Appl.*, 10:323–331, 1965.

[3] M. Crovella and A. Bestavros. Self-similarity in world wide web traffic: evidence and possible causes. In *Proceedings of the ACM SIGMETRICS '96 International Conference on Measurement and Modeling of Computer Systems*, volume 24, pages 160–169, 1996.

[4] M. Crovella, A. Bestavros, and M.S. Taqqu. Heavy-tailed probability distributions in the world wide web. In M.S. Taqqu R. Adler, R. Feldman, editor, *A Practical Guide to Heavy Tails: Statistical Techniques for Analysing Heavy Tailed Distributions*. Birkhäuser, Boston, 1999.

[5] B. Das and S.I. Resnick. Conditioning on an extreme component: Model consistency with regular variation on cones. *Bernoulli*, 17(1):226–252, 2011. ISSN 1350-7265. .

[6] B. Das and S.I. Resnick. Detecting a conditional extreme value model. *Extremes*, 14(1):29–61, 2011.

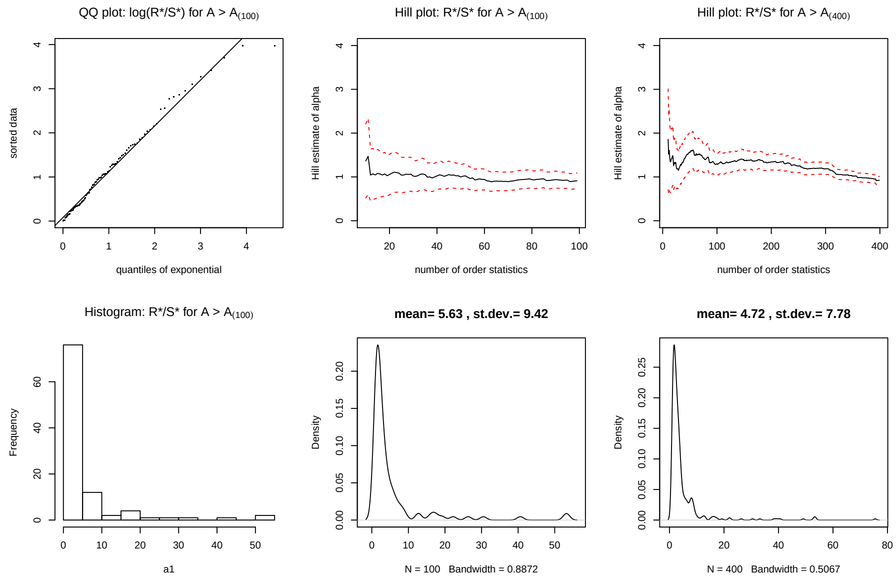

FIG 15. *UNC HTTP responses. Top: QQ plot of* $\log(R^*/S^*)$ *when* $A_i > A_{(100)}$ *and Hill plots of* R^*/S^* *when* $A_i > A_{(100)}$ *and* $A_i > A_{(400)}$*. Bottom: Histogram of* R^*/S^* *when* $A_i > A_{(100)}$ *and kernel density estimates of* R^*/S^* *for* $A_i > A_{(100)}$ *and* $A_i > A_{(400)}$.

[7] B. Das, P. Embrechts, and V. Fasen. Four theorems and a financial crisis. *The International Journal of Approximate Reasoning*, 54(6):701–716, 2013.

[8] B. Das, A. Mitra, and S. Resnick. Living on the multi-dimensional edge: Seeking hidden risks using regular variation. *Advances in Applied Probability*, 45(1):139–163, 2013. ArXiv e-prints 1108.5560.

[9] L. de Haan and A. Ferreira. *Extreme Value Theory: An Introduction*. Springer-Verlag, New York, 2006.

[10] H. Drees, L. de Haan, and S.I. Resnick. How to make a Hill plot. *Ann. Statist.*, 28(1):254–274, 2000.

[11] P. Embrechts, C. Kluppelberg, and T. Mikosch. *Modelling Extreme Events for Insurance and Finance*. Springer-Verlag, Berlin, 1997.

[12] C.A. Guerin, H. Nyberg, O. Perrin, S.I. Resnick, H. Rootzén, and C. Stărică. Empirical testing of the infinite source poisson data traffic model. *Stochastic Models*, 19(2):151–200, 2003.

[13] J.E. Heffernan and S.I. Resnick. Limit laws for random vectors with an extreme component. *Ann. Appl. Probab.*, 17(2):537–571, 2007. ISSN 1050-5164. .

[14] F. Hernandez-Campos, K. Jeffay, C. Park, J. S. Marron, and S.I. Resnick. Extremal dependence: internet traffic applications. *Stoch. Models*, 21(1):1–35, 2005. ISSN 1532-6349.

[15] 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. . URL http://dx.doi.org/10.2298/PIM0694121H.

[16] R. Ibragimov, D.. Jaffee, and J. Walden. Divesification disasters. *Journal of Financial Economics*, 99(2):333–348, 2011. .

[17] A.H. Jessen and T. Mikosch. Regularly varying functions. *Publ. Inst. Math. (Beograd) (N.S.)*, 80(94):171–192, 2006.

[18] F. Lindskog, S.I. Resnick, and J. Roy. Regularly varying measures on metric spaces: Hidden regular variation and hidden jumps. Technical report, School of ORIE, Cornell University, 2013. URL http://arxiv.org/abs/1307.5803.

[19] K. Maulik and S.I. Resnick. Characterizations and examples of hidden regular variation. *Extremes*, 7(1):31–67, 2005.

[20] J. Pickands. Statistical inference using extreme order statistics. *Ann. Statist.*, 3:119–131, 1975.

[21] S.I. Resnick. Point processes, regular variation and weak convergence. *Adv. Applied Probability*, 18:66–138, 1986.

[22] S.I. Resnick. Hidden regular variation, second order regular variation and asymptotic independence. *Extremes*, 5(4):303–336 (2003), 2002. ISSN 1386-1999.

[23] S.I. Resnick. *Heavy Tail Phenomena: Probabilistic and Statistical Modeling.* Springer Series in Operations Research and Financial Engineering. Springer-Verlag, New York, 2007. ISBN: 0-387-24272-4.

[24] S.I. Resnick. *Extreme Values, Regular Variation and Point Processes.* Springer, New York, 2008. ISBN 978-0-387-75952-4. Reprint of the 1987 original.

[25] S.I. Resnick and C. Stărică. Smoothing the Hill estimator. *Adv. Applied Probab.*, 29:271–293, 1997.

[26] R.L. Smith. Statistics of extremes, with applications in environment, insurance and finance. In B. Finkenstadt and H. Rootzén, editors, *SemStat: Seminaire Europeen de Statistique, Exteme Values in Finance, Telecommunications, and the Environment*, pages 1–78. Chapman-Hall, London, 2003.

[27] G.B. Weller and D. Cooley. A sum characterization of hidden regular variation with likelihood inference via expectation-maximization. *Biometrika (to appear)*, 2013.

Bikramjit Das
Engineering Systems and Design,
Singapore University of Technology and Design,
Singapore 138682
E-mail: bikram@sutd.edu.sg

Sidney I. Resnick
School of ORIE,
Cornell University,
Ithaca, NY 14853 USA
E-mail: sir1@cornell.edu