Stochastic Systems
2015, Vol. 5, No. 2, 195–238
DOI: 10.1214/14-SSY141

MODELS WITH HIDDEN REGULAR VARIATION: GENERATION AND DETECTION

By Bikramjit Das[*,‡,¶] and Sidney I. Resnick[†,§,¶]

Singapore University of Technology and Design[*], *Cornell University*[†]

We review the notions of multivariate regular variation (MRV) and hidden regular variation (HRV) for distributions of random vectors and then discuss methods for generating models exhibiting both properties concentrating on the non-negative orthant in dimension two. Furthermore we suggest 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, as well as two real Internet data sets.

1. Introduction. This paper discusses methods for constructing multivariate heavy-tailed models on particular sub-cones of $\mathbb{R}_+^2 = [0,\infty)^2$ by adapting techniques from multivariate regular variation theory. We evaluate two different methods for generating models exhibiting multiple heavy-tailed regimes. The results obtained in the different regimes are governed by the sub-cone that serves as the state-space, the choice of scaling function, and often the interaction between different regimes. We also discuss statistical detection methods which validate that data is consistent with particular multivariate heavy-tailed models; these methods are adapted from those developed for the *conditional extreme value model* (CEV) [8]. We discuss *multivariate regular variation* (MRV) on the cones $\mathbb{R}_+^2 \setminus \{\mathbf{0}\}$ and $(0,\infty)^2$. When regular variation exists on both cones, the regular variation on the smaller cone $(0,\infty)^2$ is called *hidden regular variation* (HRV).

Data that may be modeled by distributions having heavy tails appear in many contexts, for example, hydrology [1], finance [31], insurance [14], Internet traffic [6], social networks and random graphs [3, 13, 28, 30] and risk

Received March 2014.

[‡]B. Das was supported by SRG-ESD-2012-047 and MOE-2013-T2-1-158.

[§]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.

[¶]The authors acknowledge with thanks the referees' detailed comments which have improved the content and presentation of the paper.

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.

management [9, 19]. Often the observed data are multidimensional and generated by complex systems. Empirical evidence often indicates heavy-tailed marginal distributions and the dependence structure between the various components must be discerned.

Analysis of multivariate heavy-tailed models is facilitated by knowledge of model generation methods. Generation methods often lead to efficient simulation algorithms, suitable statistical models, appropriate estimation techniques, and control policies. In a risk-assessment setting, generation techniques help in stress-testing worst-case scenarios.

We pursue two broad themes in this paper: First, we adapt a standard general model generation technique based on the polar coordinate transform [26, p. 198] that produces tractable models in the particular cases of regular variation on $\mathbb{E} := \mathbb{R}_+^2 \setminus \{\mathbf{0}\}$ and $\mathbb{E}_0 := (0,\infty)^2$. We discuss proposals that produce models with both MRV and HRV and outline their strengths and weaknesses. Second, we discuss diagnostics for exploratory detection and identification of multivariate heavy-tailed models prior to estimating model parameters. These ideas can be adapted for dimensions higher than two but we do not pursue this adaptation here.

1.1. *Motivation.* Suppose we have a vector $\boldsymbol{Z} = (Z_1, Z_2)$ giving the risk of two assets and we wish to estimate the probability $\mathbb{P}[\boldsymbol{Z} \in \mathcal{R}]$ of a remote *risk region* beyond the range of observed data. A solution based on the asymptotic assumption of data being heavy tailed is to assume convergence of the measures

$$n\mathbb{P}\left[\frac{\boldsymbol{Z}}{b(n)} \in \cdot\right] \to \nu(\cdot) \tag{1.1}$$

to a limit measure $\nu(\cdot)$ for some limit function $b(n) \to \infty$ as $n \to \infty$. Temporarily ignore technical issues such as the meaning of the arrow "$\to$", how we estimate $\nu(\cdot)$, how to get $b(n)$, and which sets are allowed to be inserted in place of $(\cdot)$ in (1.1). A quasi-solution in this simple view is to estimate with

$$\mathbb{P}[\boldsymbol{Z} \in \mathcal{R}] \approx \frac{1}{n}\hat{\nu}(\mathcal{R}/\hat{b}(n)),$$

where hats indicate quantities needing estimates. We wonder if risk contagion is present and if both components can be simultaneously large; that is, whether $\mathbb{P}[Z_1 > x_1, Z_2 > x_2] > 0$? For models with *asymptotic independence*

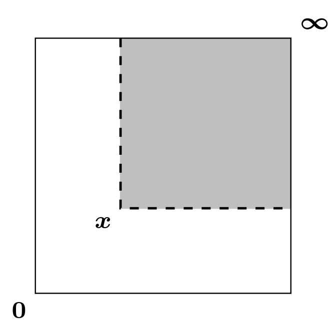

FIG 1. *Remote risk region* $\mathcal{R} = (\boldsymbol{x}, \boldsymbol{\infty})$.

(which means $\nu((0,\infty^2)) = 0$; see Section 1.4), we estimate this probability as 0 because $\nu(\cdot)$ concentrates all its mass on the axes. Does this mean that the risk is actually zero or merely that we used the wrong asymptotic regime for the estimation? Perhaps the scaling function $b(t)$ was too big. Should the state space for the risk estimation problem be $\mathbb{R}_+^2$ or $(0,\infty)^2$?

In general if $\mathcal{R}$ is disjoint from the support of $\nu(\cdot)$, our risk estimate is zero and we wonder if we chose an asymptotic method ill-suited to the purpose. The idea behind hidden regular variation is if the support of $\nu(\cdot)$ is small (eg. the two co-ordinate axes from $\mathbf{0}$) and does not contain $\mathcal{R}$, we further concentrate on the complement of the support (eg. $(0,\infty)^2$) for a second, more appropriate regular variation property where the limit measure has support intersecting $\mathcal{R}$.

1.2. *Outline.* The mathematical framework for the study of multivariate heavy tails is regular variation of measures. The theory is flexible when given for closed subcones of metric spaces [21], but we specialize to subcones of $\mathbb{R}_+^2$ where statistical results are most readily exhibited. Statistical extensions to higher dimensions will be discussed elsewhere. We list needed notation in Section 1.3 for reference. The definitions of multivariate regular variation (MRV) and hidden regular variation (HRV) are reviewed in Section 1.4 where general concepts are adapted for subcones in two dimensions. In Section 2, assuming asymptotic limit measures are specified, we adapt the standard multiplicative method for generating regularly varying models based on the generalized polar coordinate transform [10, 21, 26] to $\mathbb{E} = \mathbb{R}_+^2 \setminus \{\mathbf{0}\}$ and $\mathbb{E}_0 = (0,\infty)^2$ producing relatively tractable models.

In Section 3 we discuss generation of models that exhibit both MRV and HRV. When both MRV and HRV are present, one must be careful to properly take into account their interaction since otherwise estimation procedures will be misinterpreted. We review two model generation methods that yield both MRV on $\mathbb{E}$ and HRV on $\mathbb{E}_0$ and discuss properties of each method. These methods are (i) the mixture method, and (ii) the additive method. We give particular attention to the recently proposed additive generation method (ii) of [32] and study why interaction of MRV and HRV means 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 [8]. Our methodology is more reliable than one dimensional techniques such as checking if one dimensional marginal distributions are heavy tailed or checking whether one dimensional functions of the data vector such as the maximum and the 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. Concluding comments are in Section 6 and Section 7 contains proofs of the propositions in Section 3.

1.3. *Basic notation.* We summarize some notation and concepts here. For this paper, we have dimension $d = 2$ unless otherwise specified. We use bold letters to denote vectors, with capital letters for random vectors and small letters for non-random vectors, e.g., $\boldsymbol{y} = (y_1, y_2) \in \mathbb{R}^2$. We also define $\mathbf{0} = (0,0)$ and $\boldsymbol{\infty} = (\infty, \infty)$. Vector operations are always understood component-wise, e.g., for vectors $\boldsymbol{x}$ and $\boldsymbol{y}$, $\boldsymbol{x} \leqslant \boldsymbol{y}$ means $x_i \leqslant y_i$ for $i = 1, 2$. Some additional notation follows with explanations that are amplified in subsequent sections. Detailed discussions are in the references provided.

RV_β	Regularly varying functions with index $\beta > 0$; that is, functions $f : \mathbb{R}_+ \mapsto \mathbb{R}_+$ satisfying $\lim_{t\to\infty} f(tx)/f(t) = x^\beta$, for $x > 0$. We can and do assume such functions are continuous and strictly increasing. See [2, 11, 27].
$\mathbb{E}$	$\mathbb{R}^2 \setminus \{\mathbf{0}\}$.
[axes]	$(\{0\} \times \mathbb{R}_+) \cup (\mathbb{R}_+ \times \{0\})$.
$\mathbb{E}_0$	$\mathbb{R}_0^2 \setminus$ [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 the *forbidden zone* $\mathbb{C}_0$.
$\mathcal{C}(\mathbb{C} \setminus \mathbb{C}_0)$	Continuous, bounded, positive functions on $\mathbb{C} \setminus \mathbb{C}_0$ whose supports are bounded away from the *forbidden zone* $\mathbb{C}_0$. Without loss of generality [21], 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 [10, 18, 21] and Definition 1.1.
$\boldsymbol{X} \perp\!\!\!\perp \boldsymbol{Y}$	The random elements $\boldsymbol{X}$ and $\boldsymbol{Y}$ are independent.
$d(\boldsymbol{x}, \boldsymbol{y})]$	Euclidean metric in $\mathbb{R}^2$. So $d(\boldsymbol{x}, \boldsymbol{y}) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$.
$d(\boldsymbol{x}, \mathbb{C})$	$\inf_{\boldsymbol{y}\in\mathbb{C}} d(\boldsymbol{x}, \boldsymbol{y})$ for $\boldsymbol{x} \in \mathbb{E}$ and $\mathbb{C} \subset \mathbb{E}$.
$\aleph_{\mathbb{C}}$	$\{\boldsymbol{x} : d(\boldsymbol{x}, \mathbb{C}) = 1\}$.

$\aleph_{\mathbf{0}}$	$\{\boldsymbol{x} \in \mathbb{E} : d(\boldsymbol{x}, \{\mathbf{0}\}) = 1\}$.
$\aleph_{[\text{axes}]}$	$\{\boldsymbol{x} \in \mathbb{E}_0 : d(\boldsymbol{x}, [\text{axes}]) = 1\} = \{1\} \times [1, \infty) \cup [1, \infty) \times \{1\}$.
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 forbidden zone $\mathbb{C}_0$, $\text{GPOLAR}(\boldsymbol{x}) = (d(\boldsymbol{x}, \mathbb{C}_0), \boldsymbol{x}/d(\boldsymbol{x}, \mathbb{C}_0))$. See [10, 21].

1.4. *Regularly varying distributions on cones.* We review material from [10, 18, 21] describing the framework for the definition of MRV and HRV specialized to two dimensions. Note that the convergence concept used for defining regular variation is $\mathbb{M}$-convergence which is slightly different from vague convergence traditionally used in such cases. Reasons for preferring $\mathbb{M}$-convergence are discussed in Remark 1.1 below and in [10, 21].

Consider $\mathbb{R}_+^2$ as a metric space with Euclidean metric $d(\boldsymbol{x}, \boldsymbol{y})$. A subset $\mathbb{C} \subset \mathbb{R}_+^2$ is a *cone* if it is closed under positive scalar multiplication: if $\boldsymbol{x} \in \mathbb{C}$ then $c\,\boldsymbol{x} \in \mathbb{C}$ for $c > 0$. A proper framework for discussing regular variation is measure convergence defined by $\mathbb{M}$-convergence [10, 21] on a closed cone $\mathbb{C} \subset \mathbb{R}_+^2$ with a closed cone $\mathbb{C}_0 \subset \mathbb{C}$ deleted. Call the deleted cone $\mathbb{C}_0$ the *forbidden zone*. The two cases of interest in this paper 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. The forbidden zone is the origin $\{\mathbf{0}\}$.
2. $\mathbb{C} = \mathbb{R}_+^2$ and $\mathbb{C}_0 = \{\boldsymbol{x} : \wedge_{i=1}^2 x_i = 0\} =: [\text{axes}]$. Then $\mathbb{E}_0 := \mathbb{C} \setminus \mathbb{C}_0 = (0, \infty)^2$, the first quadrant without its axes, is the space for defining $\mathbb{M}$-convergence appropriate for HRV. The forbidden zone is the set of axes emanating from the origin in the positive direction.

Let $\mathbb{M}(\mathbb{C} \setminus \mathbb{C}_0)$ be the set of Borel measures on $\mathbb{C} \setminus \mathbb{C}_0$ which are finite on sets bounded away from the forbidden zone $\mathbb{C}_0$ [10, 18, 21]. We think of sets bounded away from the forbidden zone $\mathbb{C}_0$ as *tail regions*. We now formulate $\mathbb{M}$-convergence which becomes the basis for our definition of multivariate regular variation.

Definition 1.1. For $\mu_n, \mu \in \mathbb{M}(\mathbb{C} \setminus \mathbb{C}_0)$ we say $\mu_n \to \mu$ in $\mathbb{M}(\mathbb{C} \setminus \mathbb{C}_0)$ if $\int f \mathrm{d}\mu_n \to \int f \mathrm{d}\mu$ for all bounded, continuous, non-negative f on $\mathbb{C} \setminus \mathbb{C}_0$ whose support is bounded away from $\mathbb{C}_0$.

DEFINITION 1.2. A random vector $\boldsymbol{Z} \geqslant \mathbf{0}$ is regularly varying on $\mathbb{C} \setminus \mathbb{C}_0$ with index $\alpha > 0$ if there exists $b(t) \in RV_{1/\alpha}$, 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 $t \to \infty$,

$$t\mathbb{P}[\boldsymbol{Z}/b(t) \in \cdot\,] \to \nu(\cdot), \quad \text{in } \mathbb{M}(\mathbb{C} \setminus \mathbb{C}_0). \tag{1.2}$$

We write $\boldsymbol{Z} \in MRV(\alpha, b(t), \nu, \mathbb{C} \setminus \mathbb{C}_0)$ to emphasize that regular variation depends on an index α, scaling function $b \in RV_{1/\alpha}$, limit measure ν, and state space $\mathbb{C} \setminus \mathbb{C}_0$. Since $b(t) \in RV_{1/\alpha}$, the limit measure $\nu(\cdot)$ has a scaling property,

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

If $\mathbb{C} = \mathbb{R}_+^2$, $\mathbb{C}_0 = \{\mathbf{0}\}$ and ν satisfies $\nu((0,\infty)^2) = 0$ so that ν concentrates on the axes, then $\boldsymbol{Z}$ possesses *asymptotic independence* [11, 26, 27]. This means the probability of simultaneous occurrence of large values on both co-ordinates is estimated to be zero.

1.4.1. *Regular variation under polar coordinate transformation.* It is convenient to transform (1.2) and (1.3) using generalized polar coordinates [10, 21]. We define this for general cones of the form $\mathbb{C} \setminus \mathbb{C}_0$, but will restrict our attention eventually to $\mathbb{E}$ and $\mathbb{E}_0$.

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 forbidden zone $\mathbb{C}_0$. Define GPOLAR : $\mathbb{C} \setminus \mathbb{C}_0 \mapsto (0,\infty) \times \aleph_{\mathbb{C}_0}$ by

$$\mathrm{GPOLAR}(\boldsymbol{x}) = \left(d(\boldsymbol{x}, \mathbb{C}_0), \frac{\boldsymbol{x}}{d(\boldsymbol{x}, \mathbb{C}_0)} \right). \tag{1.4}$$

Consequently, the inverse $\mathrm{GPOLAR}^{\leftarrow} : (0,\infty) \times \aleph_{\mathbb{C}_0} \mapsto \mathbb{C} \setminus \mathbb{C}_0$ of the GPOLAR function is

$$\mathrm{GPOLAR}^{\leftarrow}(r, \theta) = r\theta. \tag{1.5}$$

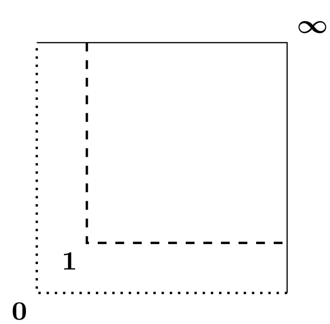

FIG 2. $\mathbb{E}_0$ = *First quadrant minus axes;* $\aleph_{[axes]}$ *is the dark dashed lines.*

Then (1.2) and (1.3) are equivalent to

$$t\mathbb{P}\left[\mathrm{GPOLAR}\left(\frac{\boldsymbol{Z}}{b(t)}\right) \in \cdot\,\right] \to (\nu_\alpha \times S)(\cdot) = (\nu \circ \mathrm{GPOLAR}^{\leftarrow})(\cdot)\,, \tag{1.6}$$

in $\mathbb{M}\left((0,\infty) \times \aleph_{\mathbb{C}_0}\right)$ where $\nu_\alpha(x,\infty) = x^{-\alpha}$, $x > 0$, $\alpha > 0$ and $S(\cdot)$ is a probability measure on $\aleph_{\mathbb{C}_0}$ [10, 21]. The transformation GPOLAR depends on the forbidden zone $\mathbb{C}_0$ and this dependence should be understood from the context.

(i) For $\mathbb{E} = \mathbb{R}_+^2 \setminus \{\mathbf{0}\}$, we have $\aleph_{\mathbf{0}} = \{\boldsymbol{x} \in \mathbb{E} : d(\boldsymbol{x}, \{\mathbf{0}\}) = 1\}$.

(ii) For $\mathbb{E}_0 = \mathbb{R}_+^2 \setminus \{\boldsymbol{x} : x_1 \wedge x_2 = 0\} =: \mathbb{R}_+^2 \setminus [\text{axes}] = (0,\infty)^2$, the appropriate unit sphere is $\aleph_{[\text{axes}]} := \{\boldsymbol{x} \in \mathbb{E} : x_1 \wedge x_2 = 1\}$. See Figure 2.

1.4.2. *MRV and HRV.* Consider simultaneous existence of regular variation on both the big cone $\mathbb{E}$ and the smaller cone $\mathbb{E}_0$. We provide equivalent polar-coordinate conditions for this simultaneous existence. These definitions and conditions help us untangle origins of various limit measures obtained while generating multivariate models in Sections 2 and 3 and aid in formulation of model detection techniques in Section 4.

Definition 1.3. The vector $\boldsymbol{Z}$ is regularly variation on $\mathbb{E}$ and has *hidden regular variation* on $\mathbb{E}_0$ if there exist $0 < \alpha \leqslant \alpha_0$, scaling functions $b(t) \in RV_{1/\alpha}$ and $b_0(t) \in RV_{1/\alpha_0}$ with $b(t)/b_0(t) \to \infty$ and limit measures ν, ν_0 such that

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

Unpacking the notation we obtain the two regular variation limits

$$t\mathbb{P}[\boldsymbol{Z}/b(t) \in \cdot\,] \to \nu(\cdot) \qquad \text{in } \mathbb{M}(\mathbb{E}), \tag{1.7}$$

$$t\mathbb{P}[\boldsymbol{Z}/b_0(t) \in \cdot\,] \to \nu_0(\cdot) \qquad \text{in } \mathbb{M}(\mathbb{E}_0). \tag{1.8}$$

Using GPOLAR, separately for the two cones (1.7) and (1.8) we get for any norm $\|\cdot\|$,

$$t\mathbb{P}\left[(\|\boldsymbol{Z}\|/b(t), \boldsymbol{Z}/\|\boldsymbol{Z}\|) \in \cdot\,\right] \to \nu_\alpha \times S(\cdot) \qquad \text{in } \mathbb{M}((0,\infty) \times \aleph_{\mathbf{0}}), \tag{1.9}$$

$$t\mathbb{P}\left[\left(\frac{Z_1 \wedge Z_2}{b_0(t)}, \frac{\boldsymbol{Z}}{Z_1 \wedge Z_2}\right) \in \cdot\,\right] \to \nu_{\alpha_0} \times S_0(\cdot) \qquad \text{in } \mathbb{M}\left((0,\infty) \times \aleph_{[\text{axes}]}\right) \tag{1.10}$$

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

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

and

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

So we may rewrite (1.10) as two statements: for $x \geqslant 1$,

$$t\mathbb{P}\left[\frac{Z_1}{b_0(t)} > r, \frac{Z_2}{Z_1} > x\right] \to r^{-\alpha_0} S_0\{(1,z) : z > x\} =: r^{-\alpha_0} p\bar{G}_1(x), \tag{1.11}$$

$$(1.12)\quad t\mathbb{P}\left[\frac{Z_2}{b_0(t)} > r, \frac{Z_1}{Z_2} > x\right] \to r^{-\alpha_0} S_0\{(z,1) : z > x\} =: r^{-\alpha_0} q\bar{G}_2(x),$$

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 $[1,\infty)$ and $\bar{G}_i = 1 - G_i$, $i = 1, 2$. In terms of G_1, G_2, p and q, (1.10) is equivalent to

$$(1.13)\quad t\mathbb{P}\left[\frac{Z_1 \wedge Z_2}{b_0(t)} > r, \left(\frac{Z_1}{Z_2} \bigvee \frac{Z_2}{Z_1}\right) > x\right] \to r^{-\alpha_0}\left(p\bar{G}_1(x) + q\bar{G}_2(x)\right).$$

and we will often use (1.13) in place of (1.10).

REMARK 1.1. Traditionally, regular variation on $\mathbb{E}$ relied on vague convergence, the polar coordinate transform $\boldsymbol{x} \mapsto (\|\boldsymbol{x}\|, \boldsymbol{x}/\|\boldsymbol{x}\|)$ and Radon measures being finite on relatively compact sets; see [26]. In order to make the natural tail regions appearing in practice to be relatively compact, the theory required *one point uncompactification* of a compactified version of $\mathbb{E}$; see [26] for further details. On $\mathbb{E} = [0,\infty)^2$ 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 (eg. $\nu_0(\boldsymbol{x}, \boldsymbol{\infty}) = (x_1 x_2)^{-1}$) and for estimation problems the approach relying on vague convergence is a dead end if estimation of a possibly infinite measure is required. Since $\aleph_{[\text{axes}]}$ is bounded away from the forbidden zone, limit measures on this set are finite. By using $\aleph_{[\text{axes}]}$ as the new unit sphere and the equivalent convergence condition (1.13) we are able to estimate the hidden measure (at least a transformed version) by estimating probability measures irrespective of whether the hidden angular measure is finite or not. More details on why $\mathbb{M}$-convergence, an approach without compactification, is desirable are in [10, 18, 21].

2. Generating regularly varying models. In this section we discuss methods for generating regularly varying models on cones in $\mathbb{R}_+^2$. We adapt a standard scheme for generating regularly varying distributions. This approach generates 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. This approach is a *multiplicative method* relying on the polar co-ordinate transformation of the random vectors.

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 $b(t)$:

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

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

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

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 (1.9) and consequently (1.7) hold. Reminder: GPOLAR is defined relative to the deleted forbidden zone $\{\mathbf{0}\}$ in this case.

2.2. *Generating regular variation on $\mathbb{E}_0$ (and sometimes also on $\mathbb{E}$).* As suggested in [22], 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 $\nu_0 := (\nu_{\alpha_0} \times S_0) \circ \text{GPOLAR}$ on $\mathbb{E}_0$ because (1.10) and therefore (1.8) hold. Reminder: GPOLAR in this case is defined relative to the deleted forbidden zone consisting of [axes].

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 $\mathbb{P}[B = 1] = p = 1 - \mathbb{P}[B = 0]$ 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 [22]

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

which occurs when
$$\bigvee_{i=1}^{2} \int_1^\infty s^{\alpha_0 - 1} \bar{G}_i(s) \mathrm{d}s < \infty,$$
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. Moreover, 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 [22].

3. Generating models that have both MRV on $\mathbb{E}$ and HRV on $\mathbb{E}_0$.

Section 2 discussed multiplicative methods to generate regularly varying models separately on $\mathbb{E}$ and $\mathbb{E}_0$. Here we pursue the first theme of the paper, namely methods that generate models with *both* regular variation on $\mathbb{E}$ and hidden regular variation on $\mathbb{E}_0$. We give two methods for generating such models that we call *the mixture method* and *the additive method*. The mixture method is somewhat easier to analyze and the majority of the section is devoted to the nuances of using the additive method under different convergence regimes in $\mathbb{E}$ and $\mathbb{E}_0$. We pay particular attention to interaction between limit measures which are generated by separately generating regularly varying random vectors in $\mathbb{E}$ and $\mathbb{E}_0$.

3.1. *Mixture method.* This method [22, 26] 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 $\mathbb{E}_0$ and B is a Bernoulli mixing variable with $\mathbb{P}[B = 1] = 1 - \mathbb{P}[B = 0]$. Since HRV implies that MRV on $\mathbb{E}$ must induce asymptotic independence [25, 26], 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 the set [axes] and
$$\boldsymbol{Y} = B_1(\xi_1, 0) + (1 - B_1)(0, \xi_2) \tag{3.1}$$
where B_1, ξ_1, ξ_2 are independent, B_1 is a Bernoulli variable with $\mathbb{P}[B_1 = 1] = \mathbb{P}[B_1 = 0] = 1/2$ and
$$t\mathbb{P}[\xi_i/b(t) > x] \to x^{-\alpha}, \quad x > 0,\ \alpha > 0,\ t \to \infty. \tag{3.2}$$
Construct $\boldsymbol{V} = (V_1, V_2)$ by the scheme of Section 2.2 to be regularly varying on $\mathbb{E}_0$ with limit measure ν_0 and scaling function $b_0(t)$. Provided that $\lim_{t\to\infty} t\mathbb{P}[V_i > b(t)] = 0$ or equivalently that $\lim_{t\to\infty} \mathbb{P}[V_i > t]/\mathbb{P}[\xi_i > t] = 0$, $i = 1, 2$, 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.* In their paper, [32] advocate an additive model of the form

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

where $\boldsymbol{Y} \in \mathrm{MRV}(\alpha, b(t), \nu, \mathbb{E})$, $\boldsymbol{V}$ has HRV with $\boldsymbol{V} \in \mathrm{MRV}(\alpha_0, b_0(t), \nu_0, \mathbb{E}_0)$ and $\boldsymbol{Y} \perp\!\!\!\perp \boldsymbol{V}$. The idea is the limit measure of $\boldsymbol{Z}$ on $\mathbb{E}$ should come from $\boldsymbol{Y}$ and the limit measure ν_0 on $\mathbb{E}_0$ should come from $\boldsymbol{V}$. They argue that this representation has advantages for parameter estimation and the additive model overcomes the undesirable and usually unrealistic feature of the mixture method where generated points are installed directly on the axes. However, while this additive model is a nice pre-asymptotic model, it is does not always successfully separate the limit measure on $\mathbb{E}$ and the hidden measure on $\mathbb{E}_0$ in an identifiable manner.

We consider the additive model under three separate scenarios. Proofs for the claims under these scenarios are deferred to Section 7.

(a) $\boldsymbol{Y}$ has form (3.1) and therefore has MRV but not HRV and $\boldsymbol{V}$ has MRV on $\mathbb{E}$ without asymptotic independence and therefore has HRV on $\mathbb{E}_0$. In this case $\boldsymbol{Z}$ is regularly varying on $\mathbb{E}$ with asymptotic independence and has HRV with scaling function, index and limit measure same as that of $\boldsymbol{V}$.
(b) $\boldsymbol{Y}$ is not necessarily of form (3.1) but has MRV while $\boldsymbol{V}$ has MRV on $\mathbb{E}$ (and hence HRV on $\mathbb{E}_0$) but not asymptotic independence. There is a tail condition (3.5) as well. Then $\boldsymbol{Z}$ is regularly varying on $\mathbb{E}$ with asymptotic independence and has HRV on $\mathbb{E}_0$ with scaling function, index and limit measure same as that of $\boldsymbol{V}$.
(c) $\boldsymbol{Y}$ has MRV but not HRV and $\boldsymbol{V}$ has both MRV on $\mathbb{E}$ and HRV on $\mathbb{E}_0$. Identifiability issues are rampant for this case.

Case (a): $\boldsymbol{Y}$ has no HRV and $\boldsymbol{V}$ has MRV on $\mathbb{E}$. We start with this simplest result. The following proposition summarizes the findings.

PROPOSITION 3.1. *Suppose $\boldsymbol{Y}$ and $\boldsymbol{V}$ are non-negative random vectors such that*

(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 $\nu_0 \in \mathbb{M}(\mathbb{E})$ and no asymptotic independence. Regular variation of $\boldsymbol{V}$ on $\mathbb{E}$ has the consequence that for $i = 1, 2$,

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

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*

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

Remark 3.1. Condition (2) in Proposition 3.1 is equivalent to the hidden limit measure ν_0 having 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.

Consider an example to make the ideas clearer.

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 $\Theta_i = 1 + E_i$ and E_1, E_2 are two standard iid exponential random variables. Then $\boldsymbol{V} = R_0\boldsymbol{\Theta}_0 \in \mathrm{MRV}(\alpha_0, b_0(t), \nu_0, \mathbb{E})$ and the limit measure of $\boldsymbol{V}$ is

$$\nu_0 = (\nu_{\alpha_0} \times \mathbb{P}[\boldsymbol{\Theta}_0 \in \cdot\,]) \circ \mathrm{GPOLAR}.$$

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 $\mathbb{E}_0$:

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

where the $\sim$ can be verified either directly or by applying Breiman's theorem [4] on products. Here $p = 1 - q = \mathbb{P}(\boldsymbol{\Theta}_0 \in ((1,\infty) \times \{1\}))$.

To check whether we can 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 three cases $\alpha_0 = 2$ with $\Theta_1 \stackrel{\mathrm{d}}{=} \Theta_2$ and $(\Theta_1 - 1), (\Theta_2 - 1)$ are iid standard exponential distributions, 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

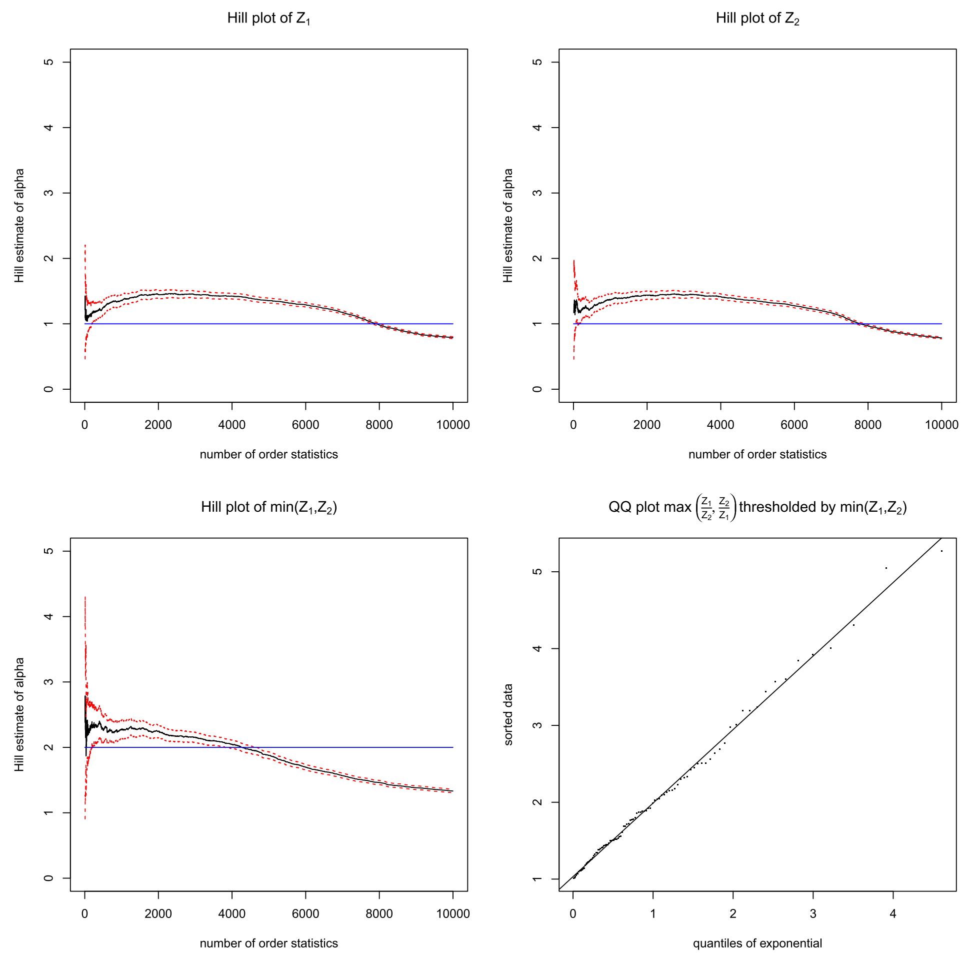

Fig 3. *Exploratory plots for Example 3.1, case 1, with* $\alpha = 1, \alpha_0 = 2$. *Top panel: Hill plots for the marginals* Z_1 *and* Z_2. *Bottom left: Hill plot for* $\min\{Z_1, Z_2\}$. *Bottom right: exponential QQ plot of* $\max\{Z_1/Z_2, Z_2/Z_1\}$ *thresholded by the 100 largest values of* $\min\{Z_1, Z_2\}$.

of α_0. Referencing (1.13), 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 3 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 $\max(Z_1/Z_2, Z_2/Z_1)$

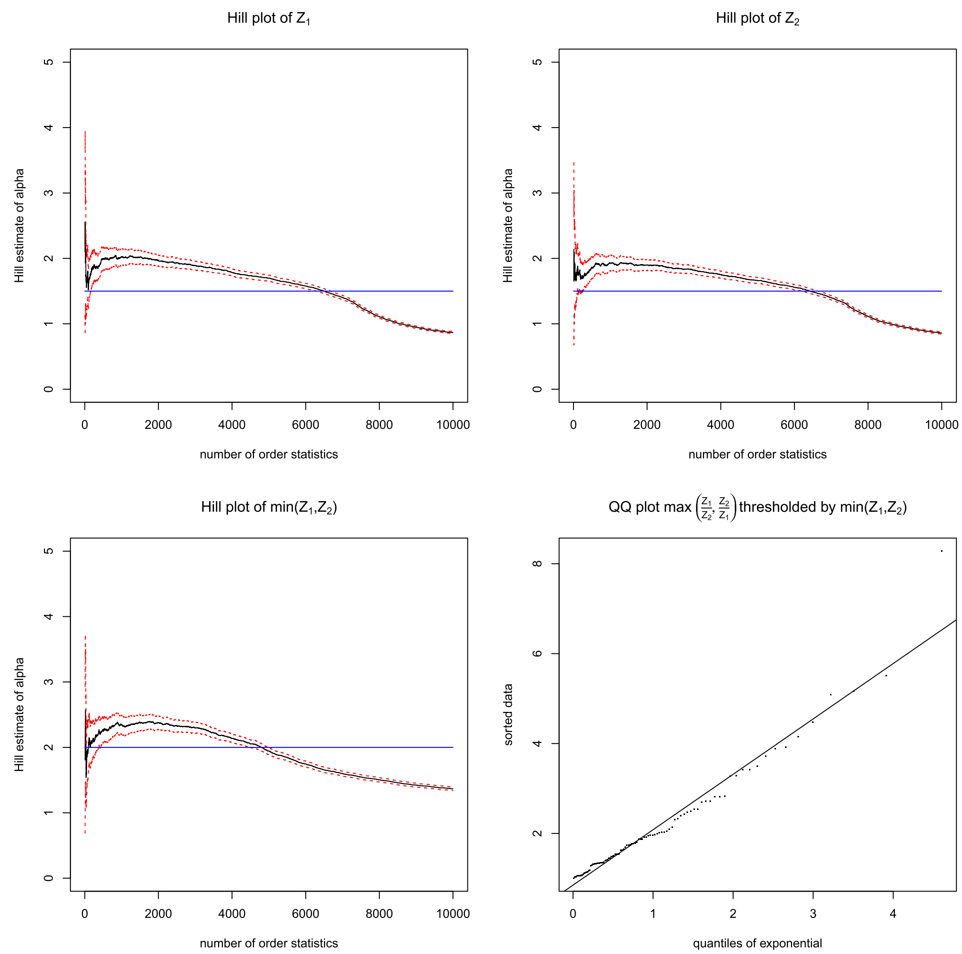

FIG 4. *Exploratory plots for Example 3.1, case 2, with* $\alpha = 1.5, \alpha_0 = 2$. *Top panel: Hill plots for the marginals* Z_1 *and* Z_2. *Bottom left: Hill plot for* $\min\{Z_1, Z_2\}$. *Bottom right: exponential QQ plot of* $\max\{Z_1/Z_2, Z_2/Z_1\}$ *thresholded by the 100 largest values of* $\min\{Z_1, Z_2\}$.

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 4 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.

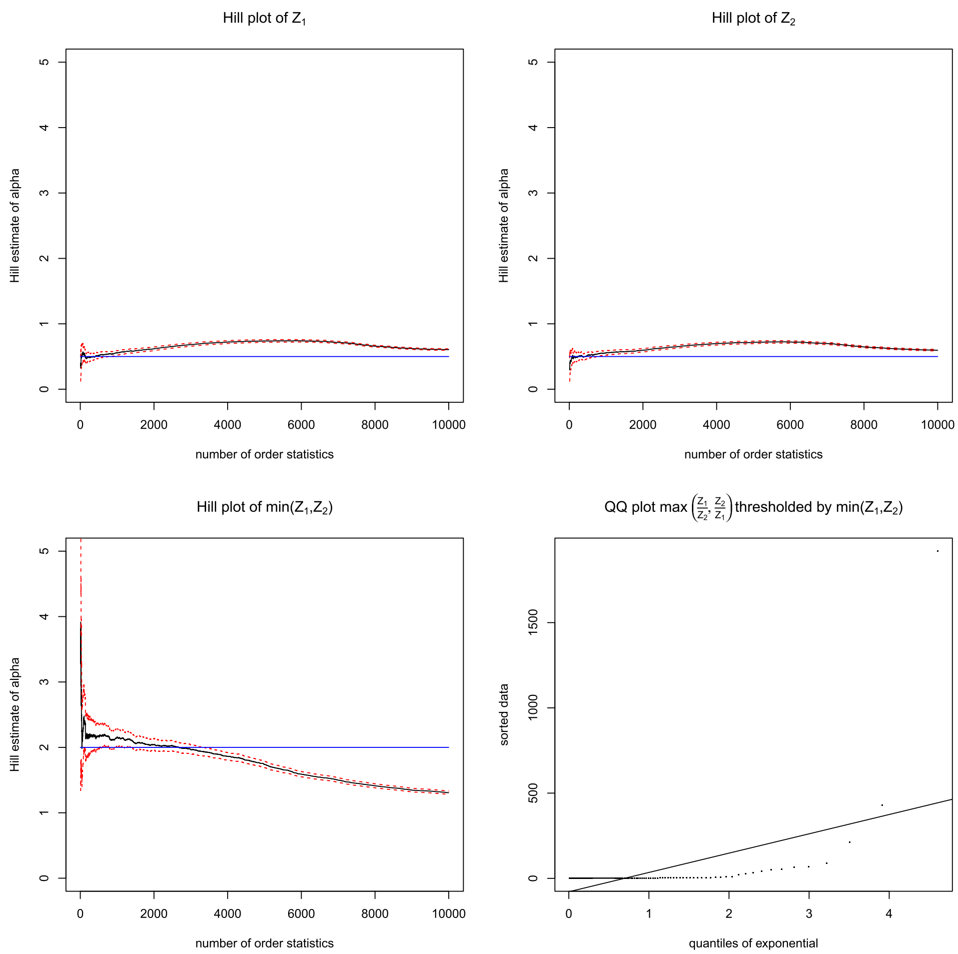

FIG 5. *Exploratory plots for Example 3.1, case 3, with* $\alpha = 0.5, \alpha_0 = 2$. *Top panel: Hill plots for the marginals* Z_1 *and* Z_2. *Bottom left: Hill plot for* $\min\{Z_1, Z_2\}$. *Bottom right: exponential QQ plot of* $\max\{Z_1/Z_2, Z_2/Z_1\}$ *thresholded by the 100 maximum values of* $\min(Z_1, Z_2)$.

- **Case 3:** $\alpha = 0.5$. In this case too, the top panel of Figure 5 indicates heavy-tailed $\boldsymbol{Z}$. 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.

$\square$

Case (b): $\boldsymbol{Y}$ has MRV and $\boldsymbol{V}$ has MRV. For Case (b) we remove the restriction in Proposition 3.1 that $\boldsymbol{Y} = (Y_1, Y_2)$ concentrates on the axes. However, to guarantee that the tails of $\boldsymbol{V}$ and $\boldsymbol{Y}$ do not interact in such a way so as to obscure the fact that the hidden measure of $\boldsymbol{Z}$ is that of $\boldsymbol{V}$ we need a tail condition comparing the tails of $\boldsymbol{Y}$ with $\boldsymbol{V}$. Continue to suppose $\boldsymbol{Y} \perp\!\!\!\perp \boldsymbol{V}$.

PROPOSITION 3.2. *Suppose $\boldsymbol{Y}$ and $\boldsymbol{V}$ are non-negative random vectors such that*

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 $\nu_0 \in \mathbb{M}(\mathbb{E})$ with no asymptotic independence so that*
$$t\mathbb{P}[\boldsymbol{V}/b_0(t) \in \cdot\,] \to \nu_0(\cdot) \quad \text{in } \mathbb{M}(\mathbb{E}).$$
3. *The interaction of the tails of $\boldsymbol{Y}$ and $\boldsymbol{V}$ is controlled by the condition*
$$t\mathbb{P}[Y_1 \wedge Y_2 > b_0(t)x] \to 0, \quad t \to \infty,\, x > 0. \tag{3.5}$$

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 $\mathbb{E}_0$.*

REMARK 3.2. For $\boldsymbol{Y}$ defined in Proposition 3.1, $Y_1 \wedge Y_2 = 0$ so (3.5) is automatic. If Y_1, Y_2 are iid with $\mathbb{P}[Y_i > x] \in RV_{-\alpha}$, $\boldsymbol{Y}$ itself has HRV [25, 26] with index 2α and condition (3.5) is needed to guarantee that the HRV in $\boldsymbol{Z}$ comes from $\boldsymbol{V}$ and not $\boldsymbol{Y}$. Condition (3.5) is equivalent in this case to

$$\frac{(\mathbb{P}[Y_1 > x])^2}{\mathbb{P}[V_1 \wedge V_2 > x]} \to 0, \quad (x \to \infty). \tag{3.6}$$

and it is sufficient that
$$\frac{\alpha_0}{2} < \alpha < \alpha_0.$$

This is seen by noting that for Y_1, Y_2 iid regularly varying with index α, (3.5) is

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

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

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

We need $\lim_{t\to\infty} t/(b^{\leftarrow}(b_0(t)))^2 = 0$ and unwinding this condition yields (3.6).

Case (c): $\boldsymbol{Y}$ does not have HRV and $\boldsymbol{V}$ has both MRV on $\mathbb{E}$ and HRV on $\mathbb{E}_0$. A problem with the additive model is that 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 $\boldsymbol{Y}$ satisfies (3.1) as 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. *For specificity, specify the scaling functions $b(t), b_*(t)$ by*

$$b(t) = \left(\frac{1}{\mathbb{P}[\xi_1 > \cdot]}\right)^{\leftarrow}(t) \quad \text{and} \quad b_*(t) = \left(\frac{1}{\mathbb{P}[V_1 > \cdot]}\right)^{\leftarrow}(t)$$

and define another scaling function $h(t)$ through its inverse $h^{\leftarrow}(t)$ by

$$h^{\leftarrow}(t) =: b^{\leftarrow}(t) b_*^{\leftarrow}(t) \sim \frac{1}{\mathbb{P}[\xi_1 > t]\mathbb{P}[V_1 > t]}. \tag{3.7}$$

Then

1. *If*

$$h(t)/b_0(t) \to \infty, \tag{3.8}$$

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

measure of $\boldsymbol{V}$*):*

$$\nu_{\boldsymbol{Z},hidden} := \frac{1}{2}\left(\nu_\alpha \times \nu_{\alpha_*} + \nu_{\alpha_*} \times \nu_\alpha\right). \tag{3.9}$$

A sufficient condition for (3.8) *is* $\alpha_* < \alpha_0 - \alpha$.

2. *If*

$$h(t)/b_0(t) \to 0, \tag{3.10}$$

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.* $\boldsymbol{Z}$ *has HRV on* $\mathbb{E}_0$ *with hidden limit measure* ν_0 *and scaling function* $b_0(t)$ *equal to those of* $\boldsymbol{V}$*. A sufficient condition for* (3.10) *is* $\alpha_* > \alpha_0 - \alpha$.

3. *If*

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

then $\boldsymbol{Z} \in MRV(\alpha, b(t), \nu, \mathbb{E})$ *with asymptotic independence and* $\boldsymbol{Z}$ *has HRV with index* $\alpha + \alpha_*$*, scaling function* $b_0(t)$ *and hidden measure* $\nu_{\boldsymbol{Z}}$ *which is a linear combination of the measure given in* (3.9) *and* ν_0*, the hidden measure of* $\boldsymbol{V}$*,*

$$\nu_{\boldsymbol{Z}} = \frac{1}{2}C_0\left(\nu_\alpha \times \nu_{\alpha_*} + \nu_{\alpha_*} \times \nu_\alpha\right) + \nu_0, \tag{3.12}$$

where $C_0 = c^{1/\alpha_0}$*. A sufficient condition for* (3.10) *is* $\alpha_* = \alpha_0 - \alpha$.

See Section 7 for the proof. We discuss an example to clarify ideas here.

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 $\boldsymbol{Z} = \boldsymbol{Y} + \boldsymbol{V} \in \mathrm{MRV}(\alpha = 0.5, t^2, \epsilon_{\{0\}} \times \nu_{1/2} + \nu_{1/2} \times \epsilon_{\{0\}}, \mathbb{E})$ with asymptotic independence of the marginals. By Proposition 3.2 we have

$$\boldsymbol{Z} \in \mathrm{MRV}(\alpha + \alpha_*, t^{\frac{1}{(\alpha+\alpha_*)}}, \nu_{\boldsymbol{Z},\mathrm{hidden}}, \mathbb{E}_0) = \mathrm{MRV}(1.5, t^{\frac{2}{3}}, \nu_{\boldsymbol{Z},\mathrm{hidden}}, \mathbb{E}_0).$$

We can check that the limit measure $\nu_{\boldsymbol{Z},\text{hidden}}$ in (3.9) 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 (1.11) 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 6 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 $G_1 = G_2$, (1.13) 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:

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

We have, $\boldsymbol{Z} = \boldsymbol{Y} + \boldsymbol{V} \in \text{MRV}(\alpha = 0.5, t^2, \epsilon_{\{0\}} \times \nu_{1/2} + \nu_{1/2} \times \epsilon_{\{0\}}, \mathbb{E})$ and furthermore $\boldsymbol{Z} = \boldsymbol{Y} + \boldsymbol{V} \in \text{MRV}(\alpha_0, t^{\frac{1}{\alpha_0}}, \mathbb{E}_0) = \text{MRV}(1.25, t^{\frac{1}{1.25}}, \mathbb{E}_0)$. 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 7 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 $G_1 \equiv G_2$.

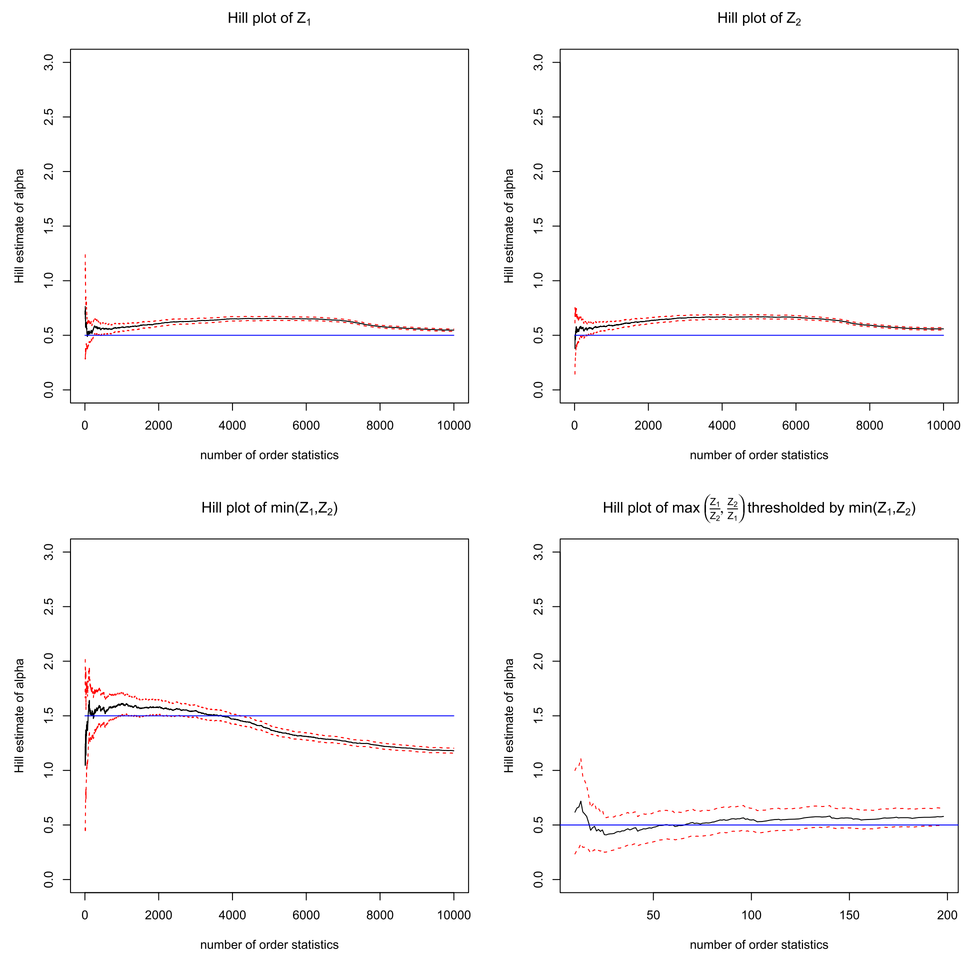

FIG 6. *Exploratory plots for Example 3.2, Case 1, with* $\alpha = 0.5, \alpha^* = 1, \alpha_0 = 2$. *Top panel: Hill plots for the marginals* Z_1 *and* Z_2. *Bottom left: Hill plot for* $\min\{Z_1, Z_2\}$. *Bottom right: Hill plot for* $\max\{Z_1/Z_2, Z_2/Z_1\}$ *thresholded by the 200 largest values of* $\min\{Z_1, Z_2\}$.

- **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 (3.12). A careful calculation shows that G_1 and G_2 have regularly varying tails with index 0.5.

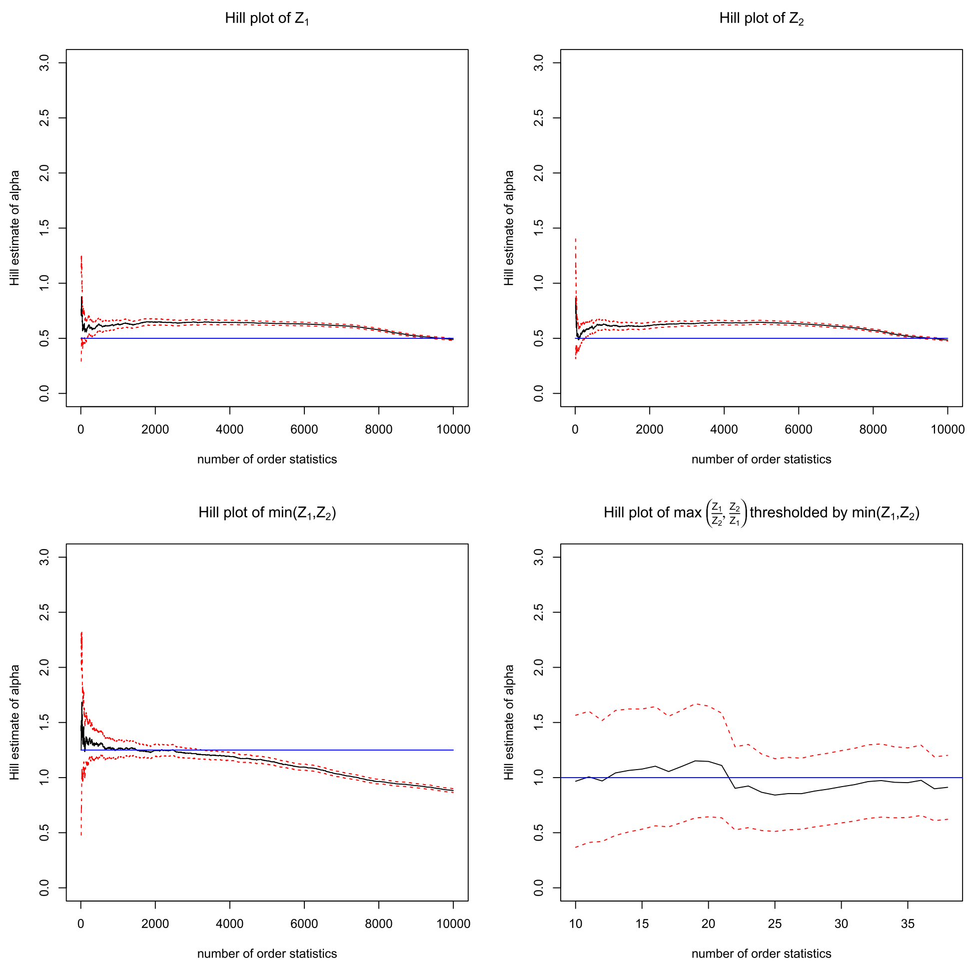

FIG 7. *Exploratory plots for Example 3.2, Case 2, with* $\alpha = 0.5, \alpha^* = 1, \alpha_0 = 1.25$. *Top panel: Hill plots for the marginals* Z_1 *and* Z_2. *Bottom left: Hill plot for* $\min\{Z_1, Z_2\}$. *Bottom right: Hill plot for* $\max\{Z_1/Z_2, Z_2/Z_1\}$ *thresholded by the 200 largest values of* $\min\{Z_1, Z_2\}$.

We generate 10000 iid samples of $\boldsymbol{Z} = \boldsymbol{Y} + \boldsymbol{V}$ using this model. In Figure 8 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.

$\square$

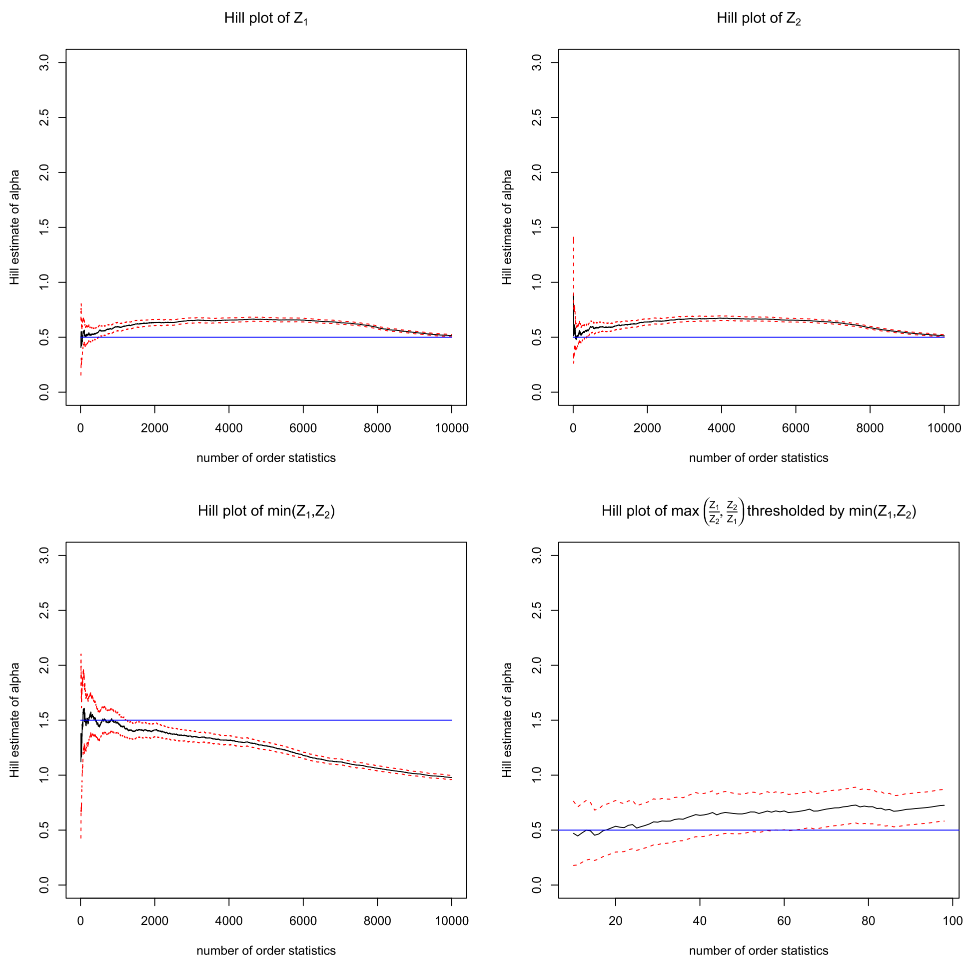

Fig 8. *Exploratory plots for Example 3.2, Case 3, with* $\alpha = 0.5, \alpha^* = 1, \alpha_0 = 1.5$. *Top panel: Hill plots for the marginals* Z_1 *and* Z_2. *Bottom left: Hill plot for* $\min\{Z_1, Z_2\}$. *Bottom right: Hill plot for* $\max\{Z_1/Z_2, Z_2/Z_1\}$ *thresholded by the 200 largest values of* $\min\{Z_1, Z_2\}$.

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 (1.6) 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 $\boldsymbol{Z}$ [26, 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 [8].

4.1. *Detecting multivariate regular variation using the CEV model.* The *conditional extreme value model* [7, 8, 16] requires at least one of the marginal distributions 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 whose first components are non-negative and where convergences are described by $\mathbb{M}$-convergence [10, 21]. Define

$$\mathbb{E}_{\sqsupset} := (0, \infty) \times \mathbb{R}.$$

Definition 4.1. Suppose $(\xi, \eta) \in \mathbb{R}_+ \times \mathbb{R}$ 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) $\mu((r, \infty] \times \cdot\,)$ is a non-degenerate measure for any fixed $r > 0$, and,
(b) $H(\cdot) := \mu((1, \infty) \times \cdot\,)$ is a probability distribution.

Then we say (ξ, η) satisfies a conditional extreme value model and write $(\xi, \eta) \in \mathrm{CEV}(a, b, \mu)$. Note that assumption (b) in the definition is a normalization, and the theory would work assuming H to be a finite measure.

Remark 4.1. The definition has some consequences [16, 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

$$\mu((r, \infty) \times (-\infty, s]) = r^{-\alpha} H(s) =: \nu_\alpha(r, \infty) H(s)$$

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 $\mathbb{E}_{\sqcap}$ 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 [7] 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.

Let $(\xi_i, \eta_i); 1 \leqslant i \leqslant n$ be 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 $\mu \in \mathbb{M}(\mathbb{E}_{\sqcap})$. 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 [8]). *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*

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

Moreover μ is a product measure if and only if both

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

The proof follows from Propositions 2.2 and 2.3 in [8]. 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 [23], de Haan and Ferreira [11, page 83], Resnick [26, 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

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

PROPOSITION 4.2 (Proposition 2.4 and Corollary 2.5 [8]). *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*

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

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*

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

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

The proof follows from Proposition 2.4 in [8]. 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 (1.7)–(1.10). 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:

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

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

If $\aleph_{\mathbf{0}}$ and $\aleph_{[\text{axes}]}$ were subsets of $\mathbb{R}$ 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* (1.9) *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* $A \in \mathcal{B}[0, \infty)$.

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* $G_1(s) = G_2(s) = 0, s \leqslant 1$.

Proposition 4.3 is easily deducible from the relationship between S and $\bar{S}$ and Proposition 4.4 follows from the connection between S_0 and G_1, G_2.

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. The relevant plots for this section appear at the end of the paper.

EXAMPLE 5.1 (Boston University: HTTP downloads.). The first data set is obtained from a now classical Boston University study [5] 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 [15, 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 [26, 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 $\{(D_i, R_i); 1 \leqslant i \leqslant 4161\}$.

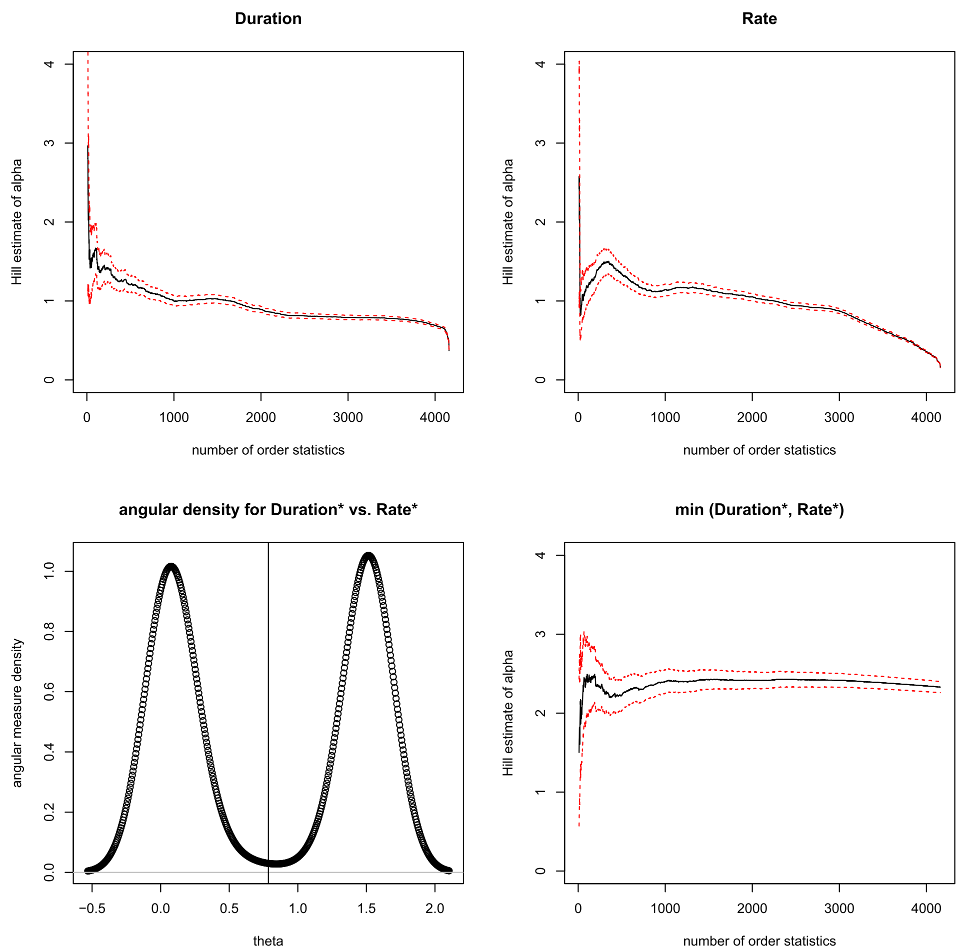

FIG 9. *BU dataset. Top panel: Hill plots of tail parameters for D and R. Bottom left plot: angular density of* (D^*, R^*). *Bottom right plot: Hill plot for* $\min(D^*, R^*)$.

Moreover the rank-transformed variables are denoted:

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

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

In Figure 9 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 [12, 26, 29]

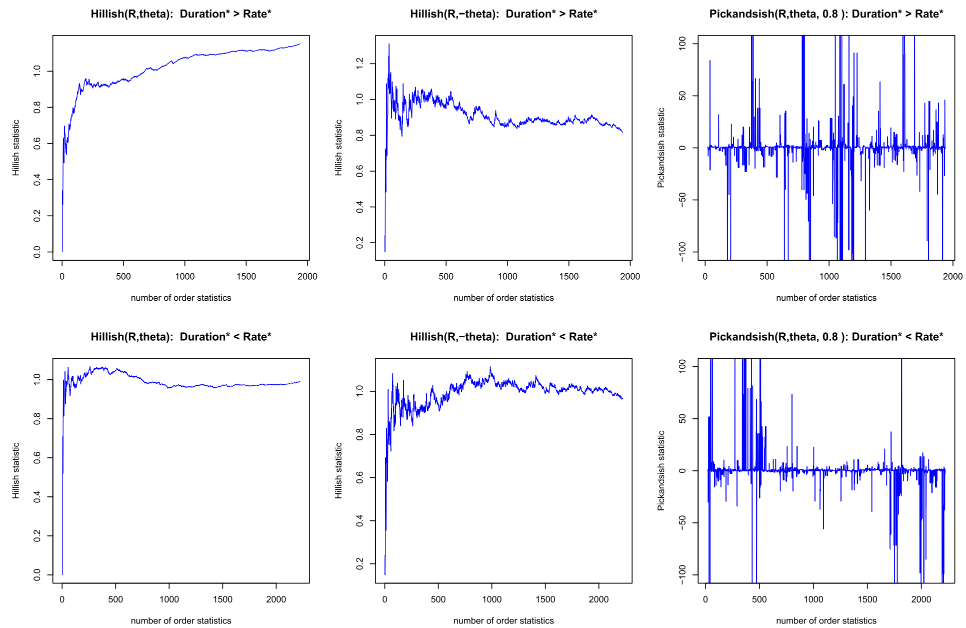

FIG 10. *BU dataset. Top panel ($D^* > R^*$): Hillish plots for (A, θ_1) and $(A, -\theta_1)$ and Pickandsish plot for (A, θ_1) at $q = 0.8$. Bottom panel ($D^* < R^*$): Hillish plots for (A, θ_2) and $(A, -\theta_2)$ and Pickandsish plot for (A, θ_2) at $q = 0.8$.*

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 on $\mathbb{E}_0$. We proceed by testing the following:

1. Is the variable $A = \min\{D^*, R^*\}$ regularly varying with parameter greater than 1? The bottom right plot in Figure 9 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 10, 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 10 we observe

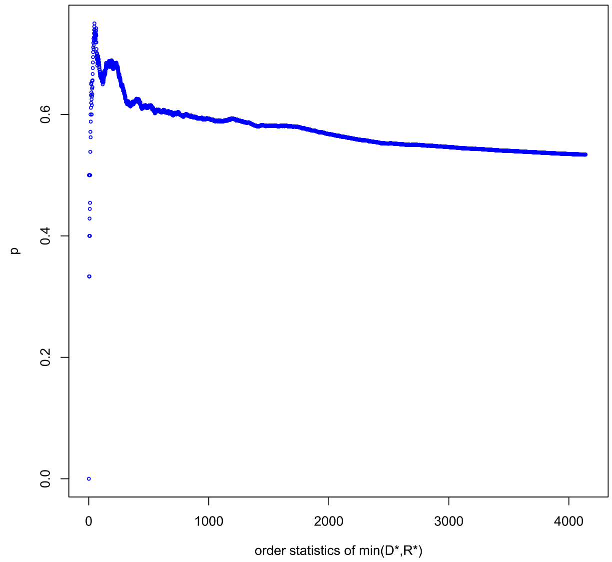

FIG 11. *BU dataset. Proportion of data with* $D_i^* > R_i^*$ *for order statistics of* $A_i = \min\{D_i^*, R_i^*\}$.

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 $(A, \theta_2) \in \text{CEV}(b_0, 1, \mu_0)$ 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)

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

for some function b_0.

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 in Proposition 4.4. Figure 11 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 11 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^*

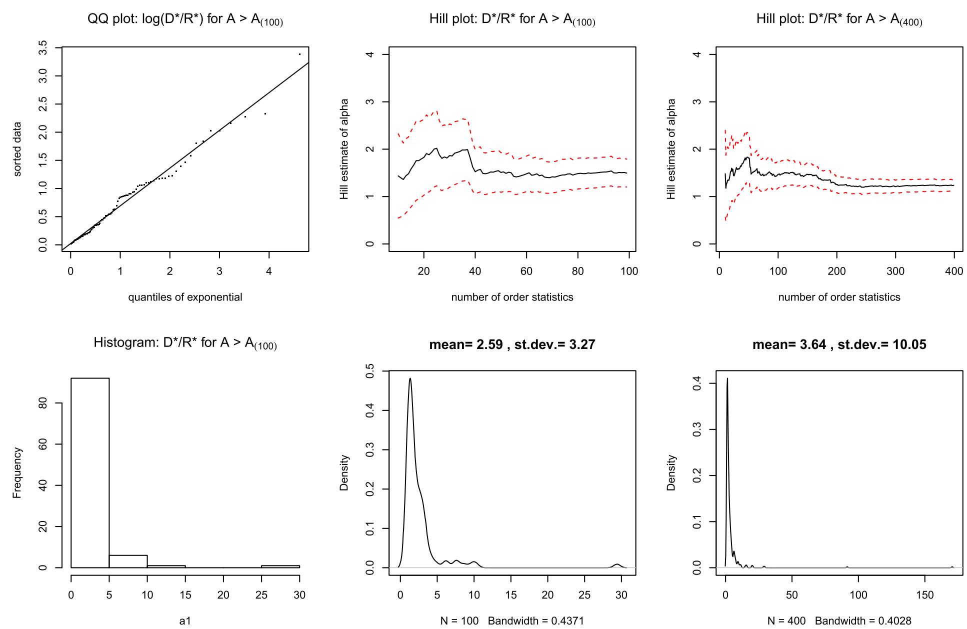

FIG 12. *BU dataset. Top panel: QQ plot of* $\log(D^*/R^*)$ *when* $A_i > A_{(100)}$ *and Hill plots of* D^*/R^* *when* $A_i > A_{(100)}$ *and* $A_i > A_{(400)}$. *Bottom panel: Histogram of* D^*/R^* *when* $A_i > A_{(100)}$ *and kernel density estimates of* D^*/R^* *when* $A_i > A_{(100)}$ *and* $A_i > A_{(400)}$.

where $A_i > A_{(k)}$ for two choices of k. The top panels of Figure 12 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 12 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 13 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 [17] 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,

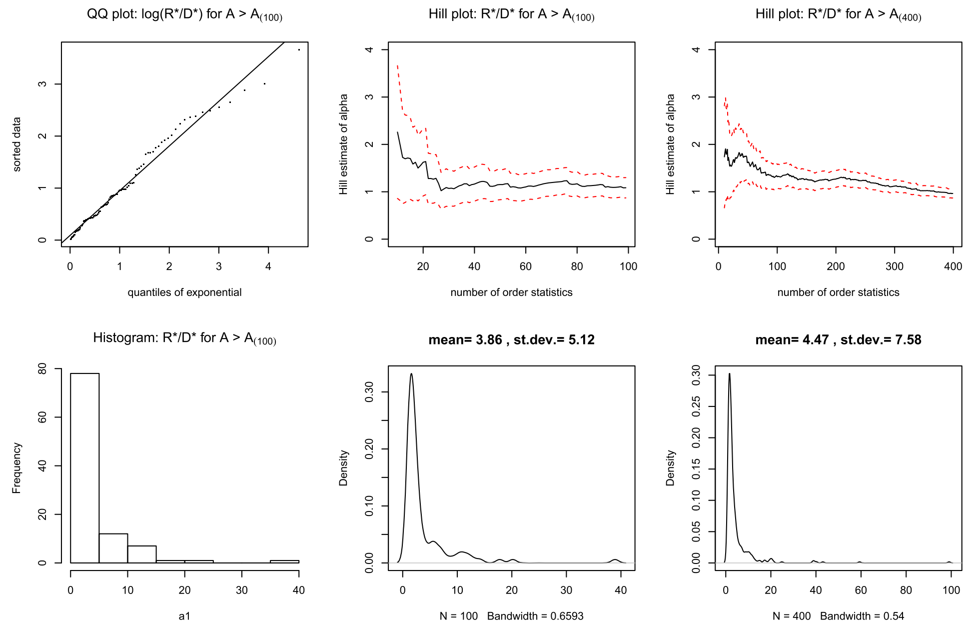

FIG 13. *BU dataset. Top panel: QQ plot of* $\log(R^*/D^*)$ *when* $A_i > A_{(100)}$ *along with Hill plots of* R^*/D^* *when* $A_i > A_{(100)}$ *and* $A_i > A_{(400)}$*. Bottom panel: Histogram of* R^*/D^* *when* $A_i > A_{(100)}$ *and kernel density estimates of* R^*/D^* *when* $A_i > A_{(100)}$ *and* $A_i > A_{(400)}$*.*

- R = throughput of the response = S/D.

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 [17]. Denote the rank-transformed variables:

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

with the generic rank-transformed variables denoted S^* and R^* respectively. The top left plots in Figure 14 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 14 gives a tail estimate $\hat{\alpha}_0$ clearly greater than 1 and is consistent with HRV. We transform the data $\{(S^*, R^*); 1 \leqslant i \leqslant 21828\}$ with the transformation GPOLAR) to obtain:

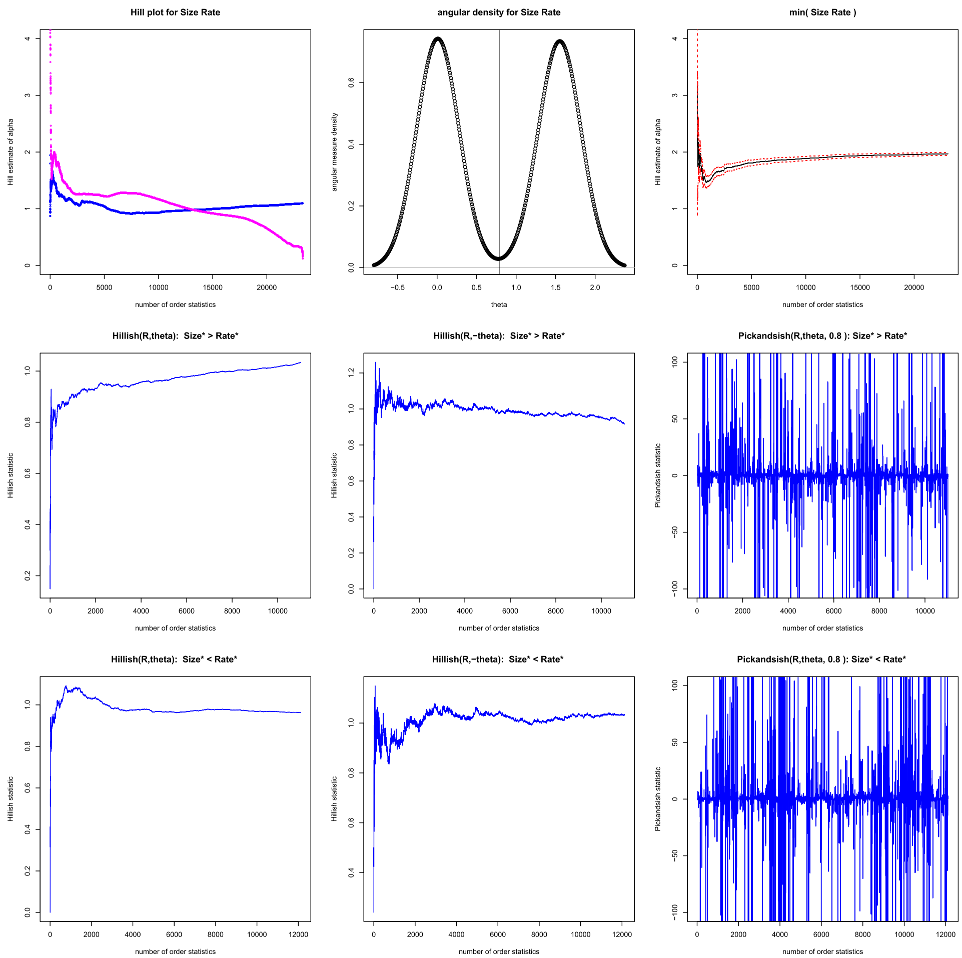

FIG 14. *UNC HTTP responses dataset. Top panel: (Left:) Hill plots of tail parameters for* S*(blue),* R*(magenta); (Middle:) angular density of* (S^*, R^*)*; (Right:) Hill plot for* $\min(S^*, R^*)$*). Middle panel (*$S^* > R^*$*): Hillish plots for* (A, θ_1) *and* $(A, -\theta_1)$ *and Pickandsish plot 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* $q = 0.8$*.*

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

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 14) and $S^* < R^*$ (see bottom panels in Figure 14), 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

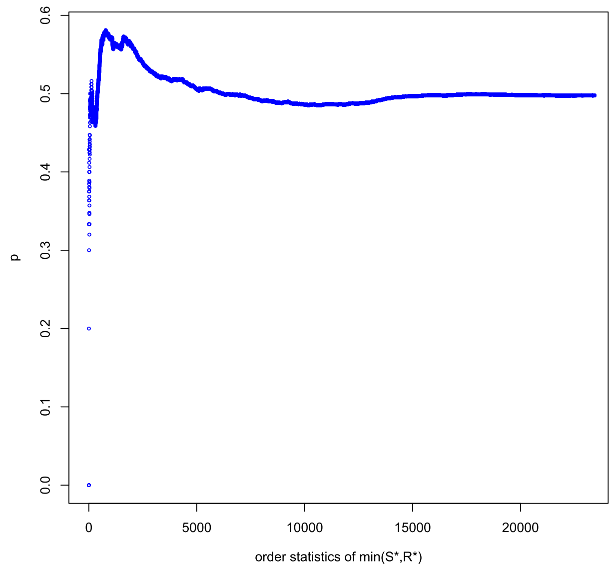

FIG 15. *UNC HTTP responses. Proportion of data with* $S_i^* > R_i^*$ *for order statistics of* $A_i = \min\{S_i^*, R_i^*\}$.

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

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

   where $A_i = \min\{S_i^*, R_i^*\}$ and $A_{(1)} \geqslant A_{(2)} \geqslant \dots$ form order statistics from $A_i; 1 \leqslant i \leqslant 21,828$ as in Figure 15. 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 16. 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.

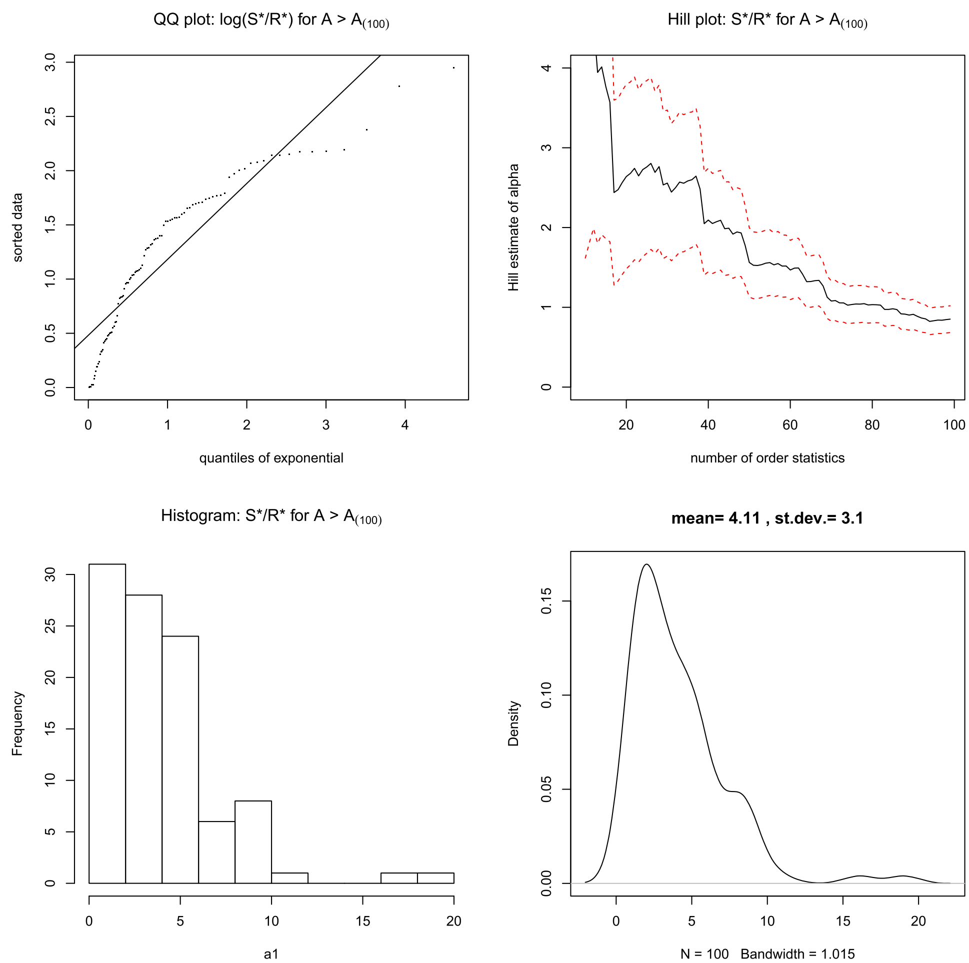

FIG 16. *UNC HTTP responses: Top: QQ plot of* $\log(S^*/R^*)$ *when* $A_i > A_{(100)}$ *along with Hill plots of* S^*/R^* *when* $A_i > A_{(100)}$ *and* $A_i > A_{(400)}$. *Bottom: Histogram and kernel density estimates of* S^*/R^* *when* $A_i > A_{(100)}$

3. The distribution of $\theta_2 \sim G_2$: see Figure 17. 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.

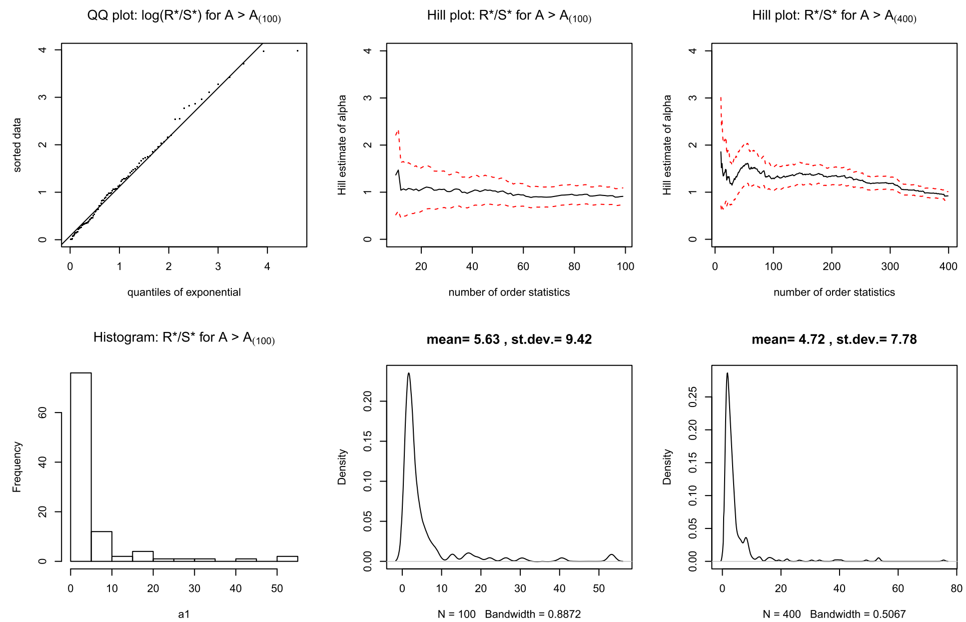

FIG 17. *UNC HTTP responses. Top: QQ plot of* $\log(R^*/S^*)$ *when* $A_i > A_{(100)}$ *and Hill plots of* R^*/S^* *when* $A_i > A_{(100)}$ *and* $A_i > A_{(400)}$*. Bottom: Histogram of* R^*/S^* *when* $A_i > A_{(100)}$ *and kernel density estimates of* R^*/S^* *for* $A_i > A_{(100)}$ *and* $A_i > A_{(400)}$*.*

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.

7. Proofs. This section contains proofs for Propositions 3.1 and 3.2.

PROOF OF PROPOSITION 3.1. The statement about MRV on $\mathbb{E}$ can be deduced from known results, eg. Resnick [26, p. 230], Jessen and Mikosch [20], Resnick [24]. (Note, assuming $\boldsymbol{V} \in \mathrm{MRV}(\alpha_0, b_0, \nu_0, \mathbb{E}_0)$ would not be enough here.) To prove HRV of $\boldsymbol{Z}$ on $\mathbb{E}_0$, we apply criterion (ii) of the

Portmanteau Theorem 2.1 in [21] and let $f \in \mathcal{C}((0,\infty)^2)$ and without loss of generality assume that f is uniformly continuous, bounded by a constant $\|f\|$ and

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

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

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

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

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

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

$$t\mathbf{E}f\left(\frac{\boldsymbol{Y}+\boldsymbol{V}}{b_0(t)}\right) - t\mathbf{E}f\left(\frac{\boldsymbol{V}}{b_0(t)}\right) \to 0. \tag{7.1}$$

Because of the special structure of $\boldsymbol{Y}$, we have a bound for the absolute value of the difference in the previous line as

$$\begin{aligned}\left|t\mathbf{E}f\left(\frac{\boldsymbol{Y}+\boldsymbol{V}}{b_0(t)}\right) - t\mathbf{E}f\left(\frac{\boldsymbol{V}}{b_0(t)}\right)\right| \leqslant & \frac{t}{2}\mathbf{E}\left|f\left(\frac{\xi_1+V_1}{b_0(t)}, \frac{V_2}{b_0(t)}\right) - f\left(\frac{\boldsymbol{V}}{b_0(t)}\right)\right| \\ & + \frac{t}{2}\mathbf{E}\left|f\left(\frac{V_1}{b_0(t)}, \frac{\xi_2+V_2}{b_0(t)}\right) - f\left(\frac{\boldsymbol{V}}{b_0(t)}\right)\right| \\ = & \frac{1}{2}\, I + \frac{1}{2}\, II.\end{aligned}$$

For any small $\delta > 0$ with $\delta < \eta$, write

$$\begin{aligned} I = & t\mathbf{E}\left|f\left(\frac{\xi_1+V_1}{b_0(t)}, \frac{V_2}{b_0(t)}\right) - f\left(\frac{\boldsymbol{V}}{b_0(t)}\right)\right| 1_{\left[\frac{\xi_1}{b_0(t)} \leqslant \delta\right]} \\ & + t\mathbf{E}\left|f\left(\frac{\xi_1+V_1}{b_0(t)}, \frac{V_2}{b_0(t)}\right) - f\left(\frac{\boldsymbol{V}}{b_0(t)}\right)\right| 1_{\left[\frac{\xi_1}{b_0(t)} > \delta\right]} \\ = & Ia + Ib.\end{aligned}$$

Since $f(x_1, x_2) = 0$ if $x_1 \wedge x_2 < \eta$, if if $V_1/b_0(t) < \eta - \delta$, then both quantities in Ia are zero. So we can write

$$Ia = t\mathbf{E}\left|f\left(\frac{\xi_1+V_1}{b_0(t)}, \frac{V_2}{b_0(t)}\right) - f\left(\frac{\boldsymbol{V}}{b_0(t)}\right)\right| 1_{\left[\frac{\xi_1}{b_0(t)} \leqslant \delta, \frac{V_1}{b_0(t)} > (\eta-\delta)\right]}$$

$$
\begin{aligned}
&\leqslant \omega_f(\delta) t\mathbb{P}[V_1 > b_0(t)(\eta-\delta)]\\
&\to c_1\,\omega_f(\delta)(\eta-\delta)^{-\alpha_0} \quad (t\to\infty),\\
&\to 0 \quad (\delta \to 0),
\end{aligned}
$$

where we used (3.3). Following a similar argument as in Ia, if $V_2/b_0(t) < \delta$, then both quantities in $2Ib$ are zero. Hence we write,

$$
\begin{aligned}
Ib &= t\mathbf{E}\left| f\left(\frac{\xi_1+V_1}{b_0(t)}, \frac{V_2}{b_0(t)}\right) - f\left(\frac{\boldsymbol{V}}{b_0(t)}\right)\right| 1_{\left[\frac{\xi_1}{b_0(t)}>\delta, \frac{V_2}{b_0(t)}>\delta\right]}\\
&\leqslant 2\|f\| t\mathbb{P}[V_2 > b_0(t)\delta]\mathbb{P}[\xi_1 > b_0(t)\delta]\\
&\to 0 \quad (t\to\infty),
\end{aligned}
$$

since as $t\to\infty$ we have $t\mathbb{P}[V_2 > b_0(t)\delta] \to c_2\delta^{-\alpha_0}$ and $\mathbb{P}[\xi > b_0(t)\delta] \to 0$. We handle II similarly to show that $II \to 0$ completing the proof. □

Proof of Proposition 3.2. As in Proposition 3.1, we focus on the HRV claim. Again assume that $f \in \mathcal{C}((0,\infty)^2)$ where f is uniformly continuous, bounded by a constant $\|f\|$ and

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

for some $\eta > 0$. We need to show (7.1). For any small $\delta > 0$ with $\delta < \eta$, the absolute value of the difference in (7.1) is bounded by

$$
\begin{aligned}
t\mathbf{E}&\left| f\left(\frac{\boldsymbol{Y}+\boldsymbol{V}}{b_0(t)}\right) - f\left(\frac{\boldsymbol{V}}{b_0(t)}\right)\right| 1_{\left[\frac{Y_1\vee Y_2}{b_0(t)}>\delta\right]}\\
&\qquad + t\mathbf{E}\left| f\left(\frac{\boldsymbol{Y}+\boldsymbol{V}}{b_0(t)}\right) - f\left(\frac{\boldsymbol{V}}{b_0(t)}\right)\right| 1_{\left[\frac{Y_1\vee Y_2}{b_0(t)}<\delta, \frac{V_1\wedge V_2}{b_0(t)}>(\eta-\delta)\right]}\\
&\qquad = I + II,
\end{aligned}
$$

since for the second term, the only way the difference can be non-zero is if both V_1 and V_2 are sufficiently large. Observe that

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

For I we have

$$
I \leqslant Ia + Ib + Ic.
$$

where

$$Ia = t\mathbf{E}\left| f\left(\frac{\boldsymbol{Y}+\boldsymbol{V}}{b_0(t)}\right) - f\left(\frac{\boldsymbol{V}}{b_0(t)}\right)\right| 1_{\left[\frac{Y_1 \wedge Y_2}{b_0(t)} > \delta\right]}$$

$$Ib = t\mathbf{E}\left| f\left(\frac{\boldsymbol{Y}+\boldsymbol{V}}{b_0(t)}\right) - f\left(\frac{\boldsymbol{V}}{b_0(t)}\right)\right| 1_{\left[\frac{Y_1}{b_0(t)} > \delta, \frac{Y_2}{b_0(t)} \leqslant \delta\right]}$$

$$Ic = t\mathbf{E}\left| f\left(\frac{\boldsymbol{Y}+\boldsymbol{V}}{b_0(t)}\right) - f\left(\frac{\boldsymbol{V}}{b_0(t)}\right)\right| 1_{\left[\frac{Y_1}{b_0(t)} \leqslant \delta, \frac{Y_2}{b_0(t)} > \delta\right]}$$

The term Ia can be quickly killed,

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

from (3.5). The term Ib also tends to zero since

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

The term Ic is handled similarly. □

Proof of Proposition 3.3. Begin with the following observations for all cases: As $t \to \infty$,

$$t\mathbb{P}\left[\left(\frac{\xi_1}{h(t)}, \frac{V_2}{h(t)}\right) \in \cdot\ \right] \to (\nu_\alpha \times \nu_{\alpha_*})(\cdot), \tag{7.2}$$

$$t\mathbb{P}\left[\left(\frac{V_1}{h(t)}, \frac{\xi_2}{h(t)}\right) \in \cdot\ \right] \to (\nu_{\alpha_*} \times \nu_\alpha)(\cdot) \tag{7.3}$$

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

$$\begin{aligned} \lim_{t\to\infty} & t\mathbb{P}[\xi_1 > h(t)x, V_2 > h(t)y] \\ &= \lim_{s\to\infty} h^{\leftarrow}(s)\mathbb{P}[\xi_1 > sx]\mathbb{P}[V_2 > sy] \\ &= \lim_{u\to\infty} b^{\leftarrow}(u)\mathbb{P}[\xi_1 > ux] \lim_{v\to\infty} b_*^{\leftarrow}(v)\mathbb{P}[V_2 > vx] \\ &= \lim_{u\to\infty} u\mathbb{P}[\xi_1 > b(u)x] \lim_{v\to\infty} v\mathbb{P}[V_2 > b_*(v)x] \\ &= \nu_\alpha(x,\infty)\nu_{\alpha_*}(y,\infty). \end{aligned}$$

The convergence in (7.3) can be similarly proved.

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

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

for some $\eta > 0$. Write

$$\text{(7.4)} \quad t\mathbf{E}f\left(\frac{\boldsymbol{Y}+\boldsymbol{V}}{h(t)}\right) = \frac{1}{2}t\,\mathbf{E}f\left(\frac{\xi_1+V_1}{h(t)}, \frac{V_2}{h(t)}\right) + \frac{1}{2}t\,\mathbf{E}f\left(\frac{V_1}{h(t)}, \frac{\xi_2+V_2}{h(t)}\right)$$

$$\text{(7.5)} \quad = \frac{1}{2}A + \frac{1}{2}B.$$

Case 1: In this case (3.8) holds. For any small $\delta > 0$ with $\delta < \eta$ we get a limit for A by writing

$$\begin{aligned}
&\left| t\mathbf{E}f\left(\frac{\xi_1+V_1}{h(t)}, \frac{V_2}{h(t)}\right) - t\mathbf{E}f\left(\frac{\xi_1}{h(t)}, \frac{V_2}{h(t)}\right)\right| \\
&\leqslant t\mathbf{E}\left| f\left(\frac{\xi_1+V_1}{h(t)}, \frac{V_2}{h(t)}\right) - f\left(\frac{\xi_1}{h(t)}, \frac{V_2}{h(t)}\right)\right| \\
&= t\mathbf{E}\left| f\left(\frac{\xi_1+V_1}{h(t)}, \frac{V_2}{h(t)}\right) - f\left(\frac{\xi_1}{h(t)}, \frac{V_2}{h(t)}\right)\right| 1_{\left[\frac{V_1}{h(t)}\leqslant\delta, \frac{\xi_1}{h(t)}>(\eta-\delta), \frac{V_2}{h(t)}>\eta\right]} \\
&\qquad + t\mathbf{E}\left| f\left(\frac{\xi_1+V_1}{h(t)}, \frac{V_2}{h(t)}\right) - f\left(\frac{\xi_1}{h(t)}, \frac{V_2}{h(t)}\right)\right| 1_{\left[\frac{V_1}{h(t)}>\delta, \frac{V_2}{h(t)}>\eta\right]} \\
&= I + II.
\end{aligned}$$

Now

$$\begin{aligned}
I &\leqslant \omega_f(\delta)t\mathbb{P}[\xi_1 > h(t)(\eta-\delta), V_2 > h(t)\eta] \\
&\to \omega_f(\delta)\nu_\alpha((\eta-\delta),\infty)\nu_{\alpha_*}(\eta,\infty) \quad (\text{using } (7.2)) \\
&\to 0 \qquad (\delta \to 0).
\end{aligned}$$

We can control II by observing

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

from (3.8). The second term in (3.9) can be seen as a limit of B in a similar fashion as in the derivation of A, relying on (7.3). This completes case (1) where (3.8) holds.

Case 2: This is the case when (3.10) holds. Replace $h(t)$ with $b_0(t)$ in (7.4) and focus on the term A. We compare it with $t\mathbf{E}f(\boldsymbol{V}/b_0(t))$:

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

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

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

Also using (3.10),

$$\begin{aligned}
II &\leqslant 2\|f\| \frac{t\, b^{\leftarrow}(b_0(t)\mathbb{P}[\xi_1 > b\circ b^{\leftarrow}(b_0(t))\delta]\, b_*^{\leftarrow}(b_0(t))\,\mathbb{P}[V_2 > b_*\circ b_*^{\leftarrow}(b_0(t))\eta]}{b^{\leftarrow}(b_0(t))b_*^{\leftarrow}(b_0(t))} \\
&\sim (constant)\frac{t}{b^{\leftarrow}(b_0(t))b_*^{\leftarrow}(b_0(t))} \qquad \text{(for large } t) \\
&= (constant)\frac{t}{h^{\leftarrow}(b_0(t))} \to 0 \qquad (t\to\infty).
\end{aligned}$$

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

Case 3: Finally, assume (3.11). Again, replace $h(t)$ by $b_0(t)$ in (7.4) and consider the term A. Write for small $\delta > 0$ with $\delta < \eta$

$$\begin{aligned}
A &= t\mathbf{E}f\left(\frac{\xi_1+V_1}{b_0(t)}, \frac{V_2}{b_0(t)}\right)\left(1_{\left[\frac{\xi_1}{b_0(t)}\leqslant\delta\right]} + 1_{\left[\frac{\xi_1}{b_0(t)}>\delta\right]}\right) \\
&= t\mathbf{E}\left[f\left(\frac{\xi_1+V_1}{b_0(t)}, \frac{V_2}{b_0(t)}\right) - f\left(\frac{\boldsymbol{V}}{b_0(t)}\right)\right] 1_{\left[\frac{\xi_1}{b_0(t)}\leqslant\delta\right]}
\end{aligned}$$

$$
\begin{aligned}
&+ t\mathbf{E}\left[f\left(\frac{\xi_1+V_1}{b_0(t)},\frac{V_2}{b_0(t)}\right)-f\left(\frac{\xi_1}{b_0(t)},\frac{V_2}{b_0(t)}\right)\right]1_{\left[\frac{\xi_1}{b_0(t)}>\delta\right]}\\
&+ t\mathbf{E}f\left(\frac{\boldsymbol{V}}{b_0(t)}\right)1_{\left[\frac{\xi_1}{b_0(t)}\leqslant\delta\right]}+t\mathbf{E}f\left(\frac{\xi_1}{b_0(t)},\frac{V_2}{b_0(t)}\right)1_{\left[\frac{\xi_1}{b_0(t)}>\delta\right]}\\
&= I+II+III+IV.
\end{aligned}
$$

We have $III \to \int f(\boldsymbol{x})\nu_0(\mathrm{d}\boldsymbol{x})$ since $\mathbb{P}[\xi_1 \leqslant b_0(t)\delta] \to 1$ as $t\to\infty$. For IV note that since $f(\boldsymbol{x})=0$ if $x_1\wedge x_2<\eta$, with $\delta<\eta$,

$$
\begin{aligned}
IV &= t\mathbf{E}f\left(\frac{\xi_1}{b_0(t)},\frac{V_2}{b_0(t)}\right)1_{\left[\frac{\xi_1}{b_0(t)}>\delta\right]}\\
&= t\mathbf{E}f\left(\frac{\xi_1}{b_0(t)},\frac{V_2}{b_0(t)}\right)\\
&= \frac{t}{h^{\leftarrow}(b_0(t))}h^{\leftarrow}(b_0(t))\mathbf{E}f\left(\frac{\xi_1}{h(h^{\leftarrow}(b_0(t)))},\frac{V_2}{h(h^{\leftarrow}(b_0(t)))}\right)\\
&\to C_0\int f\,\mathrm{d}(\nu_\alpha\times\nu_{\alpha*})
\end{aligned}
$$

using (7.2) and the fact that (3.11) is equivalent to $t/h^{\leftarrow}(b_0(t))\to c^{1/\alpha_0}=:$ C_0 as $t\to\infty$. Take the absolute value of I and add to it the indicator of the events $[V_1>b_0(t)(\eta-\delta)]$ and $[V_2>b_0(t)\eta]$ (since otherwise both terms in the difference are zero) and

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

For II write for some small $\epsilon>0$,

$$
\begin{aligned}
|II| \leqslant{}& t\mathbf{E}\left|f\left(\frac{\xi_1+V_1}{b_0(t)},\frac{V_2}{b_0(t)}\right)-f\left(\frac{\xi_1}{b_0(t)},\frac{V_2}{b_0(t)}\right)\right|1_{\left[\frac{\xi_1}{b_0(t)}>\delta,\frac{V_1}{b_0(t)}\leqslant\epsilon\right]}\\
&+t\mathbf{E}\left|f\left(\frac{\xi_1+V_1}{b_0(t)},\frac{V_2}{b_0(t)}\right)-f\left(\frac{\xi_1}{b_0(t)},\frac{V_2}{b_0(t)}\right)\right|1_{\left[\frac{\xi_1}{b_0(t)}>\delta,\frac{V_1}{b_0(t)}>\epsilon\right]}\\
={}& |IIa|+|IIb|.
\end{aligned}
$$

For $|IIa|$ we add the condition $[V_2>b_0(t)\eta]$ to avoid both terms in the difference being zero and get

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

$$
\begin{aligned}
&= \omega_f(\epsilon)\frac{t}{h^{\leftarrow}(b_0(t))}h^{\leftarrow}(b_0(t))\mathbb{P}[\xi_1 > h(h^{\leftarrow}(b_0(t)))\delta, V_2 > h(h^{\leftarrow}(b_0(t)))\eta] \\
&\to \omega_f(\epsilon)\cdot C_0 \cdot ((\nu_\alpha \times \nu_{\alpha^*})((\delta,\infty)\times(\eta,\infty))) \qquad (t\to\infty) \\
&\to 0 \qquad (\epsilon \to 0)
\end{aligned}
$$

applying (7.2) and using condition (3.11) which is equivalent to $\frac{t}{h^{\leftarrow}(b_0(t))} \to c^{1/\alpha_0} = C_0$ as $t \to \infty$. On the other hand we can dominate $|IIb|$ after adding the condition $[V_2 > b_0(t)\delta]$ to avoid both terms in the difference from being zero, to get

$$
\begin{aligned}
|IIb| &\leqslant 2\|f\|\mathbb{P}[\xi_1 > b_0(t)\delta]\, t\,\mathbb{P}[V_1 \wedge V_2 > b_0(t)(\delta\wedge\epsilon)] \\
&\sim (constant)(\delta\wedge\epsilon)^{-\alpha_0}\mathbb{P}[\xi_1 > b_0(t)\delta] \to 0 \quad (t\to\infty).
\end{aligned}
$$

The terms involving B are handled similarly. □

REFERENCES

[1] Anderson, P. L. and Meerschaert, M. M. 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] Bingham, N. H., Goldie, C. M., and Teugels, J. L. *Regular Variation*. Cambridge University Press, 1987. MR0898871

[3] Bollobás, B., Borgs, C., Chayes, J., and Riordan, O. Directed scale-free graphs. In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms (Baltimore, 2003)*, pages 132–139, New York, 2003. ACM. MR1974912

[4] Breiman, L. On some limit theorems similar to the arc-sin law. *Theory Probab. Appl.*, 10: 323–331, 1965. MR0184274

[5] Crovella, M. and Bestavros, A. 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.

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

[7] Das B. and Resnick, S. I. Conditioning on an extreme component: Model consistency with regular variation on cones. *Bernoulli*, 17 (1): 226–252, 2011. ISSN 1350-7265. MR2797990

[8] Das, B. and Resnick, S. I. Detecting a conditional extreme value model. *Extremes*, 14 (1): 29–61, 2011. MR2775870

[9] Das, B., Embrechts, P., and Fasen, V. Four theorems and a financial crisis. *The International Journal of Approximate Reasoning*, 54 (6): 701–716, 2013. MR3068113

[10] Das, B., Mitra, A., and Resnick, S. Living on the multi-dimensional edge: Seeking hidden risks using regular variation. *Advances in Applied Probability*, 45 (1): 139–163, 2013. http://arxiv.org/abs/1108.5560. MR3077544

[11] de Haan, L. and Ferreira, A. *Extreme Value Theory: An Introduction*. Springer-Verlag, New York, 2006. MR2234156

[12] Drees, H., de Haan, L., and Resnick, S. I. How to make a Hill plot. *Ann. Statist.*, 28 (1): 254–274, 2000. MR1762911

[13] Durrett, R. T. *Random Graph Dynamics*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 2010. ISBN 978-0-521-15016-3. MR2656427

[14] Embrechts, P., Kluppelberg, C., and Mikosch, T. *Modelling Extreme Events for Insurance and Finance*. Springer-Verlag, Berlin, 1997. MR1458613

[15] Guerin, C. A., Nyberg, H., Perrin, O., Resnick, S. I., Rootzén, H., and Stărică, C. Empirical testing of the infinite source poisson data traffic model. *Stochastic Models*, 19 (2): 151–200, 2003. MR1979141

[16] Heffernan, J. E. and Resnick, S. I. Limit laws for random vectors with an extreme component. *Ann. Appl. Probab.*, 17 (2): 537–571, 2007. ISSN 1050-5164. MR2308335

[17] Hernandez-Campos, F., Jeffay, K., Park, C., Marron, J. S., and Resnick, S. I. Extremal dependence: Internet traffic applications. *Stoch. Models*, 21 (1): 1–35, 2005. ISSN 1532-6349. MR2124357

[18] Hult, H. and Lindskog, F. Regular variation for measures on metric spaces. *Publ. Inst. Math. (Beograd) (N.S.)*, 80(94): 121–140, 2006. ISSN 0350-1302. http://dx.doi.org/10.2298/PIM0694121H. MR2281910

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

[20] Jessen, A. H. and Mikosch, T. Regularly varying functions. *Publ. Inst. Math. (Beograd) (N.S.)*, 80 (94): 171–192, 2006. MR2281913

[21] Lindskog, F., Resnick, S. I., and Roy, J. Regularly varying measures on metric spaces: Hidden regular variation and hidden jumps. *Probability Surveys*, 11: 270–314, 2014. http://arxiv.org/abs/1307.5803. MR3271332

[22] Maulik, K. and Resnick, S. I. Characterizations and examples of hidden regular variation. *Extremes*, 7 (1): 31–67, 2005. MR2201191

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

[24] Resnick, S. I. Point processes, regular variation and weak convergence. *Adv. Applied Probability*, 18: 66–138, 1986. MR0827332

[25] Resnick, S. I. Hidden regular variation, second order regular variation and asymptotic independence. *Extremes*, 5 (4): 303–336 (2003), 2002. ISSN 1386-1999. MR2002121

[26] Resnick, S. I. *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. MR2271424

[27] Resnick, S. I. *Extreme Values, Regular Variation and Point Processes*. Springer, New York, 2008. ISBN 978-0-387-75952-4. Reprint of the 1987 original. MR2364939

[28] Resnick, S. I. and Samorodnitsky, G. Tauberian theory for multivariate regularly varying distributions with application to preferential attachment networks. *Extremes*, 2015. http://arxiv.org/pdf/1406.6395. http://dx.doi.org/10.007/s10687-015-0216-2. MR3377813

[29] Resnick, S. I. and Stărică, C. Smoothing the Hill estimator. *Adv. Applied Probab.*, 29: 271–293, 1997. MR1432940

[30] Samorodnitsky, G., Resnick, S., Towsley, D., Davis, R., Willis, A., and Wan, P. Nonstandard regular variation of in-degree and out-degree in the preferential attachment model. *Journal of Applied Probability*, 53 (1), 2016. http://arxiv.org/pdf/1405.4882.pdf.

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

[32] Weller, G. B. and Cooley, D. A sum characterization of hidden regular variation with likelihood inference via expectation-maximization. *Biometrika*, 101 (1): 17–36, 2014. ISSN 0006-3444. http://dx.doi.org/10.1093/biomet/ast046. MR3180655

Bikramjit Das
Engineering Systems and Design
Singapore University of Technology and Design
Singapore 487372
E-mail: bikram@sutd.edu.sg

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