HIDDEN REGULAR VARIATION AND THE RANK TRANSFORM

JANET HEFFERNAN AND SIDNEY RESNICK

Abstract. Random vectors on the positive orthant whose distributions possess *hidden regular variation* are a subclass of those whose distributions are multivariate regularly varying with asymptotic independence. The concept is an elaboration of the *coefficient of tail dependence* of Ledford and Tawn (1996, 1997). We show that the rank transform which brings unequal marginals to the standard case also preserves the hidden regular variation. We discuss applications of the results to two examples involving flood risk and Internet data.

1. Introduction

A refinement of the class of multivariate regularly varying distributions, which we call *hidden regular variation*, is a semi-parametric subfamily of the full family of distributions possessing multivariate regular variation and asymptotic independence. Various cases of *hidden regular variation* have received considerable attention recently as part of the program to distinguish statistically asymptotic independence from dependence. See Campos et al. (2004), Coles et al. (1999), de Haan and de Ronde (1998), Draisma et al. (2001), Heffernan (2000), Ledford and Tawn (1996, 1997), Peng (1999), Poon et al. (2003), Resnick (2004), Stărică (1999, 2000). In particular, *hidden regular variation* is based on the analysis of the coefficient of tail dependence of Ledford and Tawn (1996, 1997).

Treatments of hidden regular variation, placing it in relation to the concepts of asymptotic independence and second order regular variation and giving characterizations and examples were given in Resnick (2002) and Maulik and Resnick (2003). Here we discuss how the rank transformation yielding standard form regular variation preserves the hidden regular variation.

1.1. **Outline.** The rest of this section reviews notation (Subsection 1.2) and the polar coordinate transformation (Subsection 1.3). Section 2 defines multivariate regular variation and hidden regular variation for heavy tailed vectors without distributionally equal components. We review the rank method for estimating the spectral measure in Section 3 and show that the rank transform can also be used to estimate the hidden spectral measure. Section 4 gives some remarks on estimation of the hidden spectral measure since the limit result of Section 3 involves a function which, typically, would not be known. Section 5 gives a high level view of several problems where existence of hidden regular variation allows for improved accuracy when estimating probabilities of extreme events. Two examples are given in Section 6, one requiring estimation of the risk of flooding and a second where the problem is to ascertain whether response size and transfer rate in Internet traffic exhibit asymptotic independence.

1.2. **Notation.** For simplicity we will generally assume that random vectors have non-negative components. Set

$$\mathbb{E} := [0,\infty]^d \setminus \{\mathbf{0}\}$$

so that the origin is excluded from $\mathbb{E}$. Compact subsets of $\mathbb{E}$ are compact sets of $[0,\infty]^d$ which do not intersect the origin; see the discussion in Resnick (2002). In some applications, for instance in finance, it is natural

2000 *Mathematics Subject Classification.* Primary 60G70; secondary 62H05, 62H20 , 60E05.

Key words and phrases. heavy tails, regular variation, Pareto tails, coefficient of tail dependence, hidden regular variation, rank transform, asymptotic independence, Internet traffic, flood risk.

Sidney Resnick's research was partially supported by NSF grant DMS-0303493 and Grant MSPF-02G-183 from the Mathematical Sciences Program of NSA. Grateful acknowledgement is also made for support for a visit to Lancaster University November 17–22, 2003, funded by the Lancaster University Small Grant Scheme.

to consider the cone $[-\infty,\infty]^d \setminus \{\mathbf{0}\}$. We leave it to the reader to make the modest changes necessary to generalize to this case, by considering the orthants individually.

Vectors are denoted by bold letters, capitals for random vectors and lower case for non-random vectors. For example: $\boldsymbol{x} = (x^{(1)}, \ldots, x^{(d)}) \in \mathbb{R}^d$. Operations between vectors should be interpreted componentwise so that for two vectors $\boldsymbol{x}$ and $\boldsymbol{z}$

$$\begin{aligned}
&\boldsymbol{x} < \boldsymbol{z} \text{ means } x^{(i)} < z^{(i)},\ i = 1, \ldots, d, && \boldsymbol{x} \le \boldsymbol{z} \text{ means } x^{(i)} \le z^{(i)},\ i = 1, \ldots, d,\\
&\boldsymbol{x} = \boldsymbol{z} \text{ means } x^{(i)} = z^{(i)},\ i = 1, \ldots, d, && \boldsymbol{z}\boldsymbol{x} = (z^{(1)}x^{(1)}, \ldots, z^{(d)}x^{(d)}),\\
&\boldsymbol{x} \bigvee \boldsymbol{z} = (x^{(1)} \vee z^{(1)}, \ldots, x^{(d)} \vee z^{(d)}), && \frac{\boldsymbol{x}}{\boldsymbol{z}} = \Big(\frac{x^{(1)}}{z^{(1)}}, \ldots, \frac{x^{(d)}}{z^{(d)}}\Big),
\end{aligned}$$

and so on. Also define $\mathbf{0} = (0, \ldots, 0)$. For a real number c, denote as usual $c\boldsymbol{x} = (cx^{(1)}, \ldots, cx^{(d)})$. We denote the rectangles (or the higher dimensional intervals) by

$$[\boldsymbol{a}, \boldsymbol{b}] = \{\boldsymbol{x} \in \mathbb{E} : \boldsymbol{a} \le \boldsymbol{x} \le \boldsymbol{b}\}.$$

Higher dimensional rectangles with one or both endpoints open are defined analogously, for example,

$$(\boldsymbol{a}, \boldsymbol{b}] = \{\boldsymbol{x} \in \mathbb{E} : \boldsymbol{a} < \boldsymbol{x} \le \boldsymbol{b}\}.$$

Complements are taken with respect to $\mathbb{E}$, so that for $\boldsymbol{x} > \mathbf{0}$,

$$[\mathbf{0}, \boldsymbol{x}]^c = \mathbb{E} \setminus [\mathbf{0}, \boldsymbol{x}] = \{\boldsymbol{y} \in \mathbb{E} : \bigvee_{i=1}^{d} \frac{y^{(i)}}{x^{(i)}} > 1\}.$$

For $i = 1, \ldots, d$, define the basis vectors $\boldsymbol{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0)$ so that the axes originating at $\mathbf{0}$ are $\mathbb{L}_i := \{t\boldsymbol{e}_i, t > 0\}$, $i = 1, \ldots, d$. Then define the cone

$$\mathbb{E}_0 = \mathbb{E} \setminus \bigcup_{i=1}^{d} \mathbb{L}_i = \{\boldsymbol{s} \in \mathbb{E} :\ \text{For some } 1 \le i < j \le d, s^{(i)} \wedge s^{(j)} > 0\}.$$

If $d = 2$, we have $\mathbb{E}_0 = (0, \infty]^2$. The cone $\mathbb{E}_0$ consists of points of $\mathbb{E}$ such that at most $d - 2$ coordinates are 0.

1.3. **The polar coordinate transformation.** It is sometimes illuminating to consider multivariate regular variation for the distribution of a random vector after a polar coordinate transformation. Suppose $\|\cdot\| : \mathbb{R}^d \mapsto [0, \infty)$ is a *norm* on $\mathbb{R}^d$. The most useful norms for us are the usual Euclidean L_2 norm, the L_p norm for $p > 0$ and the L_∞ norm: $\|\boldsymbol{x}\| = \bigvee_{i=1}^{n} |x^{(i)}|$. Assume the norm has been scaled so that $\|\boldsymbol{e}_i\| = 1$ for $i = 1, \ldots, d$. Given a chosen norm $\|\cdot\|$, the points at unit distance from the origin $\mathbf{0}$ are

$$\aleph := \{\boldsymbol{x} \in \mathbb{E} : \|\boldsymbol{x}\| = 1\}.$$

For the purpose of hidden regular variation, we need to look at smaller subcone $\mathbb{E}_0$ of $\mathbb{E}$ and restriction of $\aleph$ to $\mathbb{E}_0$ is denoted by $\aleph_0 = \aleph \cap \mathbb{E}_0$. Recall norms on $\mathbb{R}^d$ are all topologically equivalent in that convergence in one norm implies convergence in another.

For a fixed norm, define the polar coordinate transformation $T : [0, \infty)^d \setminus \{\mathbf{0}\} \mapsto (0, \infty) \times \aleph$ by

$$T(\boldsymbol{x}) = \left(\|\boldsymbol{x}\|, \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}\right) =: (r, \boldsymbol{a}),$$

and the inverse transformation $T^{\leftarrow} : (0, \infty) \times \aleph \mapsto [0, \infty]^d_+ \setminus \{\mathbf{0}\}$ by

$$T^{\leftarrow}(r, \boldsymbol{a}) = r\boldsymbol{a}.$$

Think of $\boldsymbol{a} \in \aleph$ as defining a direction and r telling how far in direction $\boldsymbol{a}$ to proceed. Since we excluded $\mathbf{0}$ from the domain of T, both T and $T^{\leftarrow}$ are continuous bijections.

When $d = 2$, it is customary, but not obligatory, to write $T(\boldsymbol{x}) = (r, \theta)$, where $\boldsymbol{x} = (r\cos\theta, r\sin\theta)$, with $0 \le \theta \le \pi/2$, rather than the more consistent notation $T(\boldsymbol{x}) = \big(r, (\cos\theta, \sin\theta)\big)$. For a random vector $\boldsymbol{X}$ in $\mathbb{R}^d$ we sometimes write $T(\boldsymbol{X}) = (R_X, \boldsymbol{\Theta}_X)$. When there is no risk of confusion, we suppress the subscript.

2. Multivariate regular variation and hidden regular variation.

We start with the definition of regular variation with unequal components.

2.1. **Regular variation on $\mathbb{E}$.** Suppose $\boldsymbol{Z}$ is a d-dimensional random vector on $[0,\infty)^d$. The distribution of $\boldsymbol{Z}$ is regularly varying (with unequal components) if there exist functions $b^{(j)}(t)\uparrow\infty$, as $t\to\infty$ such that for a Radon measure ν on $\mathbb{E}$ we have

$$tP[\Big(\frac{Z^{(j)}}{b^{(j)}(t)}, j=1,\ldots,d\Big)\in\cdot] = tP[\frac{\boldsymbol{Z}}{\boldsymbol{b}(t)}\in\cdot]\xrightarrow{v}\nu \tag{2.1}$$

on $\mathbb{E}$ (cf. de Haan and Omey (1984), Greenwood and Resnick (1979)). We assume the marginal convergences satisfy

$$tP[\frac{Z^{(j)}}{b^{(j)}(t)}>x]\to\nu_{\alpha^{(j)}}(x,\infty]=x^{-\alpha^{(j)}}, \tag{2.2}$$

where $\alpha^{(j)}>0$, $j=1,\ldots,d$. Then $b^{(j)}(t)\in RV_{1/\alpha^{(j)}}$ and we can and do assume each $b^{(j)}(t)$ is continuous and strictly increasing.

2.2. **Standard form of regular variation.** By a change of variables, regular variation with unequal components can be converted into regular variation where the marginal distributions of the each $Z^{(j)}$ are tail equivalent. The relation (2.1) is equivalent to (cf. for example, Resnick (1987, page 277))

$$tP[\Big(\frac{{b^{(j)}}^{\leftarrow}(Z^{(j)})}{t}, j=1,\ldots,d\Big)\in\cdot]\xrightarrow{v}\nu_{\text{standard}}(\cdot) \tag{2.3}$$

on $\mathbb{E}$ where ν_{standard} satisfies the homogeneity condition

$$\nu_{\text{standard}}(t\cdot)=t^{-1}\nu_{\text{standard}}(\cdot) \tag{2.4}$$

on $\mathbb{E}$. Relation (2.3) is the *standard* form of regular variation where each component of the random vector is normalized by the same function. The measures ν_{standard} and ν are related by the relation

$$\nu([\mathbf{0},\boldsymbol{x}]^c)=\nu_{\text{standard}}([0,\boldsymbol{x}^{\boldsymbol{\alpha}}]^c),\quad \boldsymbol{x}\in\mathbb{E} \tag{2.5}$$

where we continue to use our vector conventions so that

$$\boldsymbol{x}^{\boldsymbol{\alpha}}=\Big((x^{(1)})^{\alpha^{(1)}},\ldots,(x^{(d)})^{\alpha^{(d)}}\Big).$$

After transforming to polar coordinates, the homogeneity property in (2.4) becomes

$$\nu_{\text{standard}}\{\boldsymbol{x}\in\mathbb{E}:\|\boldsymbol{x}\|>r,\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}\in\cdot\}=cr^{-1}S(\cdot), \tag{2.6}$$

where $c>0$ and S is a probability measure on Borel subsets of $\aleph$.

2.3. **Hidden regular variation.** We say that the distribution of $\boldsymbol{Z}$ has *hidden regular variation* if in addition to (2.1) or (2.3) we have the following property after transforming to the standard case: There must exist a function $b_0(t)\in RV_{1/\alpha_0}$ with $b_0(t)\uparrow\infty$, $\alpha_0\geq 1$ and

$$\lim_{t\to\infty}\frac{t}{b_0(t)}=\infty, \tag{2.7}$$

such that on $\mathbb{E}_0$

$$tP[\Big(\frac{{b^{(j)}}^{\leftarrow}(Z^{(j)})}{b_0(t)}, j=1,\ldots,d\Big)\in\cdot]\xrightarrow{v}\nu_0, \tag{2.8}$$

for some Radon measure ν_0 on $\mathbb{E}_0$. Note that (2.8) is equivalent to

$$tP[\Big(\frac{Z^{(j)}}{b^{(j)}(b_0(t))}, j=1,\ldots,d\Big)\in\cdot]\xrightarrow{v}\tilde{\nu}_0, \tag{2.9}$$

on $\mathbb{E}_0$ where ν_0 and $\tilde{\nu}_0$ are related by

$$\tilde{\nu}_0(\boldsymbol{x},\boldsymbol{\infty}]=\nu_0(\boldsymbol{x}^{\boldsymbol{\alpha}},\boldsymbol{\infty}],\quad \boldsymbol{x}\in\mathbb{E}_0. \tag{2.10}$$

The measure ν_0 is also homogeneous on $\mathbb{E}_0$

$$\nu_0(t\cdot) = t^{-\alpha_0}\nu_0(\cdot)$$

but ν_0 can be either finite or infinite on E_0 (Maulik and Resnick (2003), Resnick (2002)) and when we transform ν_0 to polar coordinates, we get

$$\nu_0\{\boldsymbol{x} \in \mathbb{E} : \|\boldsymbol{x}\| > r, \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|} \in \cdot\} = r^{-\alpha_0} S_0(\cdot), \tag{2.11}$$

where S_0 is a Radon measure on Borel subsets of

$$\aleph_0 = \aleph \bigcap \mathbb{E}_0.$$

Since the region

$$\aleph_{\text{INV}} := \{\boldsymbol{x} \in \mathbb{E}_0 : \bigwedge_{j=1}^{d} x^{(j)} \geq 1\}$$

is a compact subset of $\mathbb{E}_0$ and hence will always have finite hidden measure, we can (and do) always choose $b_0(t)$ so that,

$$\nu_0(\aleph_{\text{INV}}) = 1. \tag{2.12}$$

Recall that the presence of hidden regular variation implies that the vector $\boldsymbol{Z}$ possesses *asymptotic independence* (Resnick (2002)), which in our case is equivalent to

$$\nu_{\text{standard}}(\mathbb{E}_0) = 0.$$

This means the probability of two components of the vector being simultaneously large is negligible compared to the probability of one component being large. The motivation (Ledford and Tawn (1996, 1997)) behind the concept of hidden regular variation is to create a tractable subclass of the distributions possessing asymptotic independence which would allow statistical analysis.

3. Rank methods for estimating the spectral and hidden spectral measures.

We now review and extend a method based on ranks for estimating the spectral measure S which is discussed by Huang (1992). This rank method overcomes discomfort inherent in multivariate methods which first require one to estimate tail indices $\alpha^{(j)}, j = 1, \ldots, d$. Assume we wish to estimate the spectral measure S and the hidden spectral measure S_0 on the basis of a sample $\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_n$ of size n.

3.1. Review of diagnostics based on ranks for estimating the spectral measure S.

Continue to suppose multivariate regular variation (without tail equivalent marginal distributions) given by (2.1). It follows from Resnick (1986) (see also Resnick (1987, Exercise 3.5.7, page 161)) that the empirical measures converge

$$\nu_n(\cdot) := \frac{1}{k}\sum_{i=1}^{n} \epsilon_{\left(\frac{Z_i^{(1)}}{b^{(1)}(n/k)}, \ldots, \frac{Z_i^{(d)}}{b^{(d)}(n/k)}\right)}(\cdot) \xrightarrow{P} \nu(\cdot) \tag{3.1}$$

in $M_+(\mathbb{E})$, the space of Radon measures on $\mathbb{E}$. Here $k = k(n)$ is a function of n satisfying $k(n) \to \infty$ but $k/n \to 0$ so that k/n is a vanishing proportion of the sample size. For the sample $(Z_1^{(j)}, \ldots, Z_n^{(j)})$ of jth components, let

$$Z_{(1)}^{(j)} \geq \cdots \geq Z_{(n)}^{(j)}$$

be the order statistics in decreasing order starting with the biggest. Taking the marginal convergences in (3.1) and inverting yields for each $j = 1, \ldots, d$

$$\frac{Z^{(j)}_{(\lceil kt^{(j)} \rceil)}}{b^{(j)}(n/k)} \xrightarrow{P} (t^{(j)})^{-1/\alpha_j}, \tag{3.2}$$

in $D(0,\infty]$. Because convergence is to a constant limit, we may append this to (3.1) to get

$$\Big(\nu_n(\cdot), (\frac{Z^{(j)}_{(\lceil kt^{(j)}\rceil)}}{b^{(j)}(n/k)}; j=1,\dots,d)\Big) \Rightarrow \Big(\nu, \big((t^{(j)})^{-1/\alpha_j}; j=1,\dots,d\big)\Big) \tag{3.3}$$

in $M_+(\mathbb{E}) \times D(0,\infty] \times \cdots \times D(0,\infty]$.

Recall (2.5) and convert (3.3) into

$$\Big(\nu_n([\mathbf{0},\boldsymbol{x}]^c), (\frac{Z^{(j)}_{(\lceil kt^{(j)}\rceil)}}{b^{(j)}(n/k)}; j=1,\dots,d)\Big) \Rightarrow \Big(\nu([\mathbf{0},\boldsymbol{x}]^c), \big((t^{(j)})^{-1/\alpha_j}; j=1,\dots,d\big)\Big) \tag{3.4}$$

and then apply the almost surely continuous map

$$(\nu([\mathbf{0},\boldsymbol{x}]^c), \boldsymbol{t}) \mapsto \nu([\mathbf{0},\boldsymbol{t}]^c)$$

to (3.4) to get

$$\nu_n\left(\big[\mathbf{0}, (\frac{Z^{(j)}_{(\lceil kt^{(j)}\rceil)}}{b^{(j)}(n/k)}; j=1,\dots,d)\big)\big]^c\right) \Rightarrow \nu([\mathbf{0}, \boldsymbol{t}^{-1/\boldsymbol{\alpha}}]^c). \tag{3.5}$$

Unpack the left side of (3.5). We have

$$\begin{aligned}\nu_n\left(\big[\mathbf{0}, (\frac{Z^{(j)}_{(\lceil kt^{(j)}\rceil)}}{b^{(j)}(n/k)}; j=1,\dots,d)\big]^c\right) &= \frac{1}{k}\sum_{i=1}^n 1_{[\frac{Z^{(j)}_i}{b^{(j)}(n/k)} \le \frac{Z^{(j)}_{(\lceil kt^{(j)}\rceil)}}{b^{(j)}(n/k)}; j=1,\dots,d]^c} \\ &= \frac{1}{k}\sum_{i=1}^n 1_{[Z^{(j)}_i \le Z^{(j)}_{(\lceil kt^{(j)}\rceil)}; j=1,\dots,d]^c}. \end{aligned} \tag{3.6}$$

Define the anti-ranks for $j=1,\dots,d$

$$r^{(j)}_i = \sum_{l=1}^n 1_{[Z^{(j)}_l \ge Z^{(j)}_i]}$$

to be the number of j-th components bigger than or equal to $Z^{(j)}_i$. Rephrase (3.6) as

$$\frac{1}{k}\sum_{i=1}^n 1_{[r^{(j)}_i \ge kt^{(j)};\ j=1,\dots,d]^c}.$$

Change variables $\boldsymbol{s} \mapsto \boldsymbol{t}^{-1}$ to get

$$\frac{1}{k}\sum_{i=1}^n 1_{[r^{(j)}_i \ge k(s^{(j)})^{-1};\ j=1,\dots,d]^c} \Rightarrow \nu\Big([\mathbf{0}, \boldsymbol{s}^{1/\boldsymbol{\alpha}}]^c\Big)$$

or

$$\frac{1}{k}\sum_{i=1}^n 1_{[\frac{k}{r^{(j)}_i} \le (s^{(j)});\ j=1,\dots,d]^c} \Rightarrow \nu\Big([\mathbf{0}, \boldsymbol{s}^{1/\boldsymbol{\alpha}}]^c\Big)$$

or

$$\hat{\nu}_{\text{standard},n} =: \frac{1}{k}\sum_{i=1}^n \epsilon_{\Big(\frac{k}{r^{(j)}_i};\ j=1,\dots,d\Big)} \Rightarrow \nu_{\text{standard}} \tag{3.7}$$

in $M_+(\mathbb{E})$ where we used (2.10). A polar coordinate transformation of the points

$$\{\Big(\frac{k}{r^{(j)}_i};\ j=1,\dots,d\Big);\ i=1,\dots,n\}$$

allows one to estimate the spectral measure S. Suppose the polar coordinates of $(1/r_i^{(j)}, j = 1, \ldots, d)$ are $(R_i, \boldsymbol{\Theta}_i)$. Then the empirical measure

$$\frac{\sum_{i=1}^n 1_{[kR_i \geq 1]} \epsilon_{\boldsymbol{\Theta}_i}}{\sum_{i=1}^n 1_{[kR_i \geq 1]}} = \frac{\sum_{i=1}^n 1_{[R_i \geq k^{-1}]} \epsilon_{\boldsymbol{\Theta}_i}}{\sum_{i=1}^n 1_{[R_i \geq k^{-1}]}} \Rightarrow S \tag{3.8}$$

is a consistent estimator of S.

3.2. **Finding the hidden spectral measure with the rank transform.** Now suppose both regular variation represented by (2.1) and hidden regular variation represented by (2.8) hold. How can we use the rank transform to estimate the hidden measure S_0?

To find the hidden angular measure, we expect we have to use points

$$\{\Big(\frac{k}{r_i^{(j)}};\ j = 1, \ldots, d\Big);\ i = 1, \ldots, n\}$$

thresholded at a lower level than in (3.8). Since $b_0(t)/t \to 0$ (see (2.7)), it seems plausible to use the points

$$\Big\{\Big(\frac{\frac{k}{r_i^{(j)}}}{b_0(n/k)/(n/k)};\ j = 1, \ldots, d\Big);\ i = 1, \ldots, n\Big\} = \Big\{\Big(\frac{n/r_i^{(j)}}{b_0(n/k)};\ j = 1, \ldots, d\Big);\ i = 1, \ldots, n\Big\}.$$

This scheme yields the hidden measure.

Proposition 1. *Assume $\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_n$ is an iid sample from a distribution on $[0, \infty)^d$ which possesses both regular and hidden regular variation so that* (2.1) *and* (2.8) *hold. Then we have*

$$\frac{1}{k} \sum_{i=1}^n \epsilon_{\left(\frac{n/r_i^{(j)}}{b_0(n/k)};\ j=1,\ldots,d\right)} \Rightarrow \nu_0, \tag{3.9}$$

in $M_+(\mathbb{E}_0)$ where recall ν_0 is given in (2.8).

Proof. The proof mimics the scheme followed for using ranks to estimate ν or S. Observe for $\boldsymbol{x} \in \mathbb{E}_0$ that

$$\begin{aligned}\frac{1}{k} \sum_{i=1}^n \epsilon_{\left(\frac{n/r_i^{(j)}}{b_0(n/k)};\ j=1,\ldots,d\right)}[\boldsymbol{x}, \boldsymbol{\infty}] =& \frac{1}{k} \sum_{i=1}^n 1_{[n(x^{(j)})^{-1} b_0^{-1}(n/k) \geq r_i^{(j)};\ j=1,\ldots,d]} \\ =& \frac{1}{k} \sum_{i=1}^n 1_{\left[\frac{Z_i^{(j)}}{b^{(j)}(b_0(n/k))} \geq \frac{Z^{(j)}_{(\lceil n(x^{(j)} b_0(n/k))^{-1} \rceil)}}{b^{(j)}(b_0(n/k))};\ j=1,\ldots,d\right]}.\end{aligned}$$

We claim (see below for the proof) that for each $j = 1, \ldots, d$,

$$\frac{Z^{(j)}_{(\lceil n(x^{(j)} b_0(n/k))^{-1} \rceil)}}{b^{(j)}(b_0(n/k))} \xrightarrow{P} (x^{(j)})^{1/\alpha^{(j)}}. \tag{3.10}$$

Using this to scale the convergence in (2.9) we get

$$\frac{1}{k} \sum_{i=1}^n \epsilon_{\left(\frac{n/r_i^{(j)}}{b_0(n/k)};\ j=1,\ldots,d\right)}[\boldsymbol{x}, \boldsymbol{\infty}] \Rightarrow \tilde{\nu}_0[\boldsymbol{x}^{\mathbf{1}/\boldsymbol{\alpha}}, \boldsymbol{\infty}] = \nu_0[\boldsymbol{x}, \boldsymbol{\infty}],$$

where we used (2.10) for the last equality. This suffices to prove the result modulo the claim. □

Lemma 1. *Assume $\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_n$ is an iid sample from a distribution on $[0, \infty)^d$ which possesses both regular and hidden regular variation so that* (2.1) *and* (2.8) *hold. Then* (3.10) *holds in $D[0, \infty)$ for each $j = 1, \ldots, d$.*

Proof. We have for each $j = 1, \ldots, d$,

$$\frac{b_0(n/k)}{n} \sum_{i=1}^n \epsilon_{Z_i^{(j)}/b^{(j)}(b_0(n/k))} \Rightarrow \nu_{\alpha^{(j)}}$$

in $M_+(0,\infty]$ using Resnick (1986) or Resnick (1987, Exercise 3.5.7, page 161). In particular

$$\frac{b_0(n/k)}{n}\sum_{i=1}^n \epsilon_{Z_i^{(j)}/b^{(j)}(b_0(n/k))}(x^{-1},\infty] \Rightarrow x^{\alpha^{(j)}}$$

in $D[0,\infty)$. This is a sequence of non-decreasing functions converging to a continuous limit and so the inverse functions converge as well. This yields the statement of the Lemma. □

If one converts (3.9) to polar coordinates in order to estimate S_0 one gets the analogue of (3.8). As in (3.8), set $\mathbf{\Theta}_i$ for the polar angular coordinate of $(1/r_i^{(j)}, 1\le j\le d)$ and let R_i be the norm of $(1/r_i^{(j)}, 1\le j\le d)$. Then assuming S_0 is finite (otherwise one has to restrict $\mathbf{\Theta}_i$ to a compact subset of $\aleph_0$)

$$\frac{\sum_{i=1}^n 1_{[R_i\ge n^{-1}b_0(n/k)]}\epsilon_{\mathbf{\Theta}_i}}{\sum_{i=1}^n 1_{[R_i\ge n^{-1}b_0(n/k)]}} \Rightarrow S_0, \tag{3.11}$$

in $M_+(\mathbb{E}_0)$. Since $b_0(n/k)$ is unknown for statistical purposes, it must be estimated before we can regard (3.11) as a suitable estimate of S_0.

4. Estimation of the hidden measure.

For this section, recall

$$\aleph_{\text{INV}} := \{\boldsymbol{x}\in\mathbb{E}_0 : \bigwedge_{j=1}^d x^{(j)} \ge 1\}$$

for the set of vectors all of whose components are at least 1. Define

$$m_i = \bigwedge_{j=1}^d \frac{1}{r_i^{(j)}}, \quad i=1,\dots,n$$

and further suppose

$$m_{(1)} \ge m_{(2)} \ge \cdots \ge m_{(n)}$$

is the ordering of $m_1,\dots,m_n$ with the biggest first. We have the following result.

Proposition 2. *Assume $\boldsymbol{Z}_1,\dots,\boldsymbol{Z}_n$ is an iid sample from a distribution on $[0,\infty)^d$ which possesses both regular and hidden regular variation so that* (2.1) *and* (2.8) *hold and continue to assume that $\nu_0(\aleph_{\text{INV}})=1$. Then we have in $M_+(\mathbb{E}_0)$*

$$\widehat{\nu_0} := \frac{1}{k}\sum_{i=1}^n \epsilon_{(\frac{1/r_i^{(j)}}{m_{(k)}}, 1\le j\le d)} \Rightarrow \nu_0. \tag{4.1}$$

Thus we have removed the unknown $b_0(n/k)$ and replaced it by a random variable.

Proof. On $D[0,\infty)$ we have from Proposition 1 and continuous mapping that

$$\eta_n(t) := \frac{1}{k}\sum_{i=1}^k \epsilon_{\frac{n}{b_0(n/k)}\wedge_{j=1}^d \frac{1}{r_i^{(j)}}}(t^{-1},\infty] \Rightarrow \nu_0\{\boldsymbol{x} : \wedge_{j=1}^d x^{(j)} \ge t^{-1}\} = t^{\alpha_0}\nu_0(\aleph_{\text{INV}}) = t^{\alpha_0} =: \eta_\infty(t).$$

Therefore we also have in $D[0,\infty)$ that the inverse processes converge:

$$\eta_n^{\leftarrow}(s) \Rightarrow \eta_\infty^{\leftarrow}(s) = s^{1/\alpha_0}.$$

Unpack the left hand side. We have

$$\begin{aligned}\eta_n^{\leftarrow}(s) &= \inf\{u : \eta_n(u)\ge s\}\\ &= \inf\{u : \sum_{i=1}^n \epsilon_{\frac{n}{b_0(n/k)}m_i}(u^{-1},\infty] \ge ks\}\end{aligned}$$

$$
\begin{aligned}
&= \Big(\sup\{v : \sum_{i=1}^n \epsilon_{\frac{n}{b_0(n/k)} m_i}(v,\infty] \geq ks\} \Big)^{-1} \\
&= \frac{b_0(n/k)}{n} \Big(\sup\{w : \sum_{i=1}^n \epsilon_{m_i}(w,\infty] \geq ks\} \Big)^{-1} \\
&= \frac{b_0(n/k)}{n} m^{-1}_{([ks])}.
\end{aligned}
$$

Therefore we see that

$$
\frac{n}{b_0(n/k)} m_{([ks])} \Rightarrow s^{-1/\alpha_0} \tag{4.2}
$$

in $D(0,\infty]$.

The rest is a scaling argument. Couple (4.2) with (3.9) and compose to get in $D(\mathbb{E}_0)$

$$
\frac{1}{k}\sum_{i=1}^n \epsilon_{\left(\frac{n/r_i^{(j)}}{b_0(n/k)},\, j=1,\ldots,d\right)} [\frac{n}{b_0(n/k)} m_{(k)} \boldsymbol{x}, \boldsymbol{\infty}] = \frac{1}{k}\sum_{i=1}^n \epsilon_{\left(\frac{1/r_i^{(j)}}{m_{(k)}},\, j=1,\ldots,d\right)} [\boldsymbol{x}, \boldsymbol{\infty}] \Rightarrow \nu_0(\boldsymbol{x}, \boldsymbol{\infty}]
$$

as required. □

This suggests a way forward around the problem of the unknown function $b_0(n/k)$ in (3.11): We replace $n^{-1}b_0(n/k)$ by $m_{(k)}$. We can then write the analogue of (3.11). If ν_0 is infinite, let $\aleph_0(K)$ be a convenient compact subset of $\aleph_0$. For $d = 2$ where $\aleph$ can be parameterized as $\aleph = [0, \pi/2]$ and $\aleph_0 = (0, \pi/2)$, we can set $\aleph_0(K) = [\delta, \pi/2 - \delta]$ for some small $\delta > 0$. Then we have from Proposition 2,

$$
\frac{\sum_{i=1}^n 1_{[R_i \geq m_{(k)}, \boldsymbol{\Theta}_i \in \aleph_0(K)]} \epsilon_{\boldsymbol{\Theta}_i}}{\sum_{i=1}^n 1_{[R_i \geq m_{(k)}, \boldsymbol{\Theta}_i \in \aleph_0(K)]}} \Rightarrow S_0\big(\cdot \bigcap \aleph_0(K)\big). \tag{4.3}
$$

If ν_0 is finite, we can replace $\aleph_0(K)$ with $\aleph_0$ as was done in (3.11).

Thus, to summarize, we proceed as follows when estimating S_0:

(1) Replace the heavy tailed multivariate sample $\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_n$ by the n vectors of anti-ranks $\mathbf{r}_1, \ldots, \mathbf{r}_n$, where
$$
r_i^{(j)} = \sum_{l=1}^n 1_{[Z_l^{(j)} \geq Z_i^{(j)}]}; \quad j = 1, \ldots, d;\ i = 1, \ldots, n.
$$
(2) Compute the normalizing factors
$$
m_i = \bigwedge_{j=1}^d \frac{1}{r_i^{(j)}}; \quad i = 1, \ldots, n,
$$
and their order statistics
$$
m_{(1)} \geq \cdots \geq m_{(n)}.
$$
(3) Compute the polar coordinates $\{(R_i, \boldsymbol{\Theta}_i); i = 1, \ldots, n\}$ of $\{(1/r_i^{(j)}; j = 1, \ldots, d);\ i = 1, \ldots, n\}$.
(4) Estimate S_0 using the $\boldsymbol{\Theta}_i$ corresponding to $R_i \geq m_{(k)}$.

For estimating S_0 these results offer some advice on how to pick thresholds to estimate both S and S_0 but of course one must still choose k. The Stărică scaling device (Resnick (2003), Stărică (1999, 2000)) seems to offer some guidance, though nothing, presently, is known about its theoretical properties.

5. Why estimation of the hidden measure matters.

Hidden regular variation is designed to produce a model subclass of the multivariate regularly varying distributions possessing asymptotic independence. This model subclass is better suited to estimating very small probabilities that the random vector falls in jointly remote regions well beyond the range of the observed data. (See de Haan and de Ronde (1998), Ledford and Tawn (1996, 1997).) In this section we review how hidden regular variation can help in the estimation of small probabilities.

For example, consider the following problems.

5.1. **Estimate the probability of non-compliance.** Supppose the vector $\boldsymbol{Z} = (Z^{(1)}, \ldots, Z^{(d)})$ represents concentrations of a specific pollutant at d locations. (Alternatively, $\boldsymbol{Z}$ could represent concentations of different pollutants at a single site.) Environmental agencies set standards by insisting that critical levels $\boldsymbol{t}_0 = (t_0^{(1)}, \ldots, t_0^{(d)})$ not be exceeded at each of the d sites; that is, that $\boldsymbol{Z} \le \boldsymbol{t}_0$. Non-compliance is represented by the event

$$[\text{non-compliance}] = [\boldsymbol{Z} \le \boldsymbol{t}_0]^c = \bigcup_{j=1}^{d} [Z^{(j)} > t_0^{(j)}].$$

Non-compliance results in a fine or withdrawal of government support; it has various economic and political implications, none of which is desirable. How do we estimate the probability of non-compliance?

We assume only that (2.1) and (2.8) hold and that $d = 2$ for simplicity. We observe that the probability of non-compliance is

$$P\{\bigcup_{j=1}^{d} [Z^{(j)} > t_0^{(j)}]\} = \sum_{j=1}^{2} P[Z^{(j)} > t_0^{(j)}] - P[Z^{(1)} > t_0^{(1)}, Z^{(2)} > t_0^{(2)}]. \tag{5.1}$$

Assuming only asymptotic independence is present, one would be inclined to neglect the joint probability on the right since it is negligible compared with the univariate probabilities. However, as the following outline shows, the assumption of hidden regular variation allows for reasonable estimation of the joint probability.

Assume an iid random sample $\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_n$. For the univariate probabilities we have

$$\begin{aligned}
\sum_{j=1}^{2} P[Z^{(j)} > t_0^{(j)}] =& \sum_{j=1}^{2} P[\frac{Z^{(j)}}{b^{(j)}(\frac{n}{k})} > \frac{t_0^{(j)}}{b^{(j)}(\frac{n}{k})}] \\
\approx& \frac{k}{n} \nu \left(\left[\mathbf{0}, \left(\frac{t_0^{(1)}}{b^{(1)}(n/k)}, \frac{t_0^{(2)}}{b^{(2)}(n/k)} \right) \right]^c \right) \\
=& \frac{k}{n} \left(\left(\frac{t_0^{(1)}}{b^{(1)}(n/k)} \right)^{-\alpha_1} + \left(\frac{t_0^{(2)}}{b^{(2)}(n/k)} \right)^{-\alpha_2} \right) \\
\approx& \frac{k}{n} \left(\left(\frac{t_0^{(1)}}{\widehat{b^{(1)}}(n/k)} \right)^{-\widehat{\alpha_1}} + \left(\frac{t_0^{(2)}}{\widehat{b^{(2)}}(n/k)} \right)^{-\widehat{\alpha_2}} \right),
\end{aligned}$$

where

$$\widehat{b^{(j)}}(n/k) = Z_{(k)}^{(j)},$$

the k-th largest of the j-th components of the data and $\widehat{\alpha_j}$ is an estimate of α_j obtained from the one-dimensional sample of j-th components. For example, $\widehat{\alpha_j}$ could be the Hill estimator or the MLE.

For the multivariate tail probability on the right side of (5.1), we estimate

$$\begin{aligned}
P[Z^{(1)} > t_0^{(1)}, Z^{(2)} > t_0^{(2)}] =& P[\frac{Z^{(j)}}{b^{(j)}(b_0(\frac{n}{k}))} > \frac{t_0^{(j)}}{b^{(j)}(b_0(\frac{n}{k}))}; j = 1, 2] \\
\approx& \frac{k}{n} \tilde{\nu}_0 \left(\left(\left(\frac{t_0^{(1)}}{b^{(1)}(b_0(\frac{n}{k}))}, \frac{t_0^{(2)}}{b^{(2)}(b_0(\frac{n}{k}))} \right), \boldsymbol{\infty} \right] \right)
\end{aligned}$$

from (2.9) and from (2.10) this is

$$\begin{aligned}
=& \frac{k}{n} \nu_0 \left(\left(\left(\left(\frac{t_0^{(1)}}{b^{(1)}(b_0(\frac{n}{k}))} \right)^{\alpha_1}, \left(\frac{t_0^{(2)}}{b^{(2)}(b_0(\frac{n}{k}))} \right)^{\alpha_2} \right), \boldsymbol{\infty} \right] \right) \\
\approx& \frac{k}{n} \widehat{\nu}_0 \left(\left(\left(\left(\frac{t_0^{(1)}}{\widehat{b^{(1)} \circ b_0}(\frac{n}{k})} \right)^{\hat{\alpha}_1}, \left(\frac{t_0^{(2)}}{\widehat{b^{(2)} \circ b_0}(\frac{n}{k})} \right)^{\hat{\alpha}_2} \right), \boldsymbol{\infty} \right] \right).
\end{aligned} \tag{5.2}$$

For our estimate of $\widehat{b^{(j)} \circ b_0}(\frac{n}{k})$ we use $(j = 1, 2)$

$$\widehat{b^{(j)} \circ b_0}(\frac{n}{k}) = Z^{(j)}_{([1/m_{(k)}])}. \tag{5.3}$$

To see why this is the appropriate estimator of the composition, note that from (4.2) (with $s = 1$) and (3.10) we have jointly

$$\Big(\frac{Z^{(j)}_{(\lceil n(xb_0(n/k))^{-1}\rceil)}}{b^{(j)}(b_0(n/k))}, \frac{n}{b_0(n/k)} m_{(k)}\Big) \xrightarrow{P} \big(x^{1/\alpha_j}, 1\big) \tag{5.4}$$

in $D(0,\infty) \times [0,\infty)$. Therefore by scaling, using the map $(x(\cdot), t) \mapsto x(t\cdot)$, we get from (5.4) that

$$\frac{Z^{(j)}_{(\lceil n((nm_{(k)}/b_0(n/k))b_0(n/k))^{-1}\rceil)}}{b^{(j)}(b_0(n/k))} = \frac{Z^{(j)}_{(\lceil 1/m_{(k)}\rceil)}}{b^{(j)}(b_0(n/k))} \xrightarrow{P} 1.$$

Hence the choice given in (5.3)

Although the rank method of estimating ν_0 obviates the need to compute the α's, the α-estimates are needed for estimating probabilities of remote events.

The calculation leading to (5.2) will be used in Subsection 6.2 to assess the risk of flooding. The data analyzed there are precipitation in successive hours. For these bivariate data, simultaneous exceedance of a threshold represents the the greatest risk of flooding due to rainfall.

5.2. **Estimate the probability of a failure region.** In certain water resource problems a failure region is of the form

$$A := \{\boldsymbol{z} \in \mathbb{E} : f(\boldsymbol{z}, t_0) > 0\}$$

and it is required to estimate

$$P[\boldsymbol{Z} \in A]$$

(Bruun and Tawn (1998), de Haan and de Ronde (1998)). For example, the failure probability could be of the form

$$P[a^{(1)}Z^{(1)} + a^{(2)}Z^{(2)} > t_0],$$

where $Z^{(1)}$ represents still water level and $Z^{(2)}$ represents wave height and $a^{(j)} > 0$, $j = 1, 2$. If one only assumes asymptotic independence, then one is tempted to approximate the failure probability with $\sum_{j=1}^{2} P[a^{(j)}Z^{(j)} > t_0]$.

As before, assume estimation of the failure probability is based on an iid sample $\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_n$ and that hidden regular variation is present. We may write $\boldsymbol{a} \cdot \boldsymbol{Z} = a^{(1)}Z^{(1)} + a^{(2)}Z^{(2)}$ and then

$$P[\boldsymbol{a} \cdot \boldsymbol{Z} > t_0] = P[\boldsymbol{a} \cdot \boldsymbol{Z} > t_0, \boldsymbol{Z} \in \mathbb{E}_0] + P[\boldsymbol{a} \cdot \boldsymbol{Z} > t_0, \boldsymbol{Z} \in \mathbb{E} \setminus \mathbb{E}_0].$$

The second probability is approximately

$$\frac{k}{n} \sum_{j=1}^{2} P[a^{(j)} \frac{Z^{(j)}}{b^{(j)}(n/k)} > \frac{t_0}{b^{(j)}(n/k)}] \approx \frac{k}{n} \sum_{j=1}^{2} \Big(\frac{t_0}{a^{(j)} b^{(j)}(n/k)}\Big)^{-\alpha_j} \approx \frac{k}{n} \sum_{j=1}^{2} \Big(\frac{t_0}{a^{(j)} Z^{(j)}_{(k)}}\Big)^{-\widehat{\alpha_j}}.$$

The probability on $\mathbb{E}_0$ is approximately

$$\frac{k}{n} \tilde{\nu}_0\{\boldsymbol{z} \in \mathbb{E}_0 : \sum_{j=1}^{2} a^{(j)} b^{(j)}(b_0(n/k)) z^{(j)} > t_0\} \approx \frac{k}{n} \widehat{\tilde{\nu}}_0\{\boldsymbol{z} \in \mathbb{E}_0 : \sum_{j=1}^{2} a^{(j)} Z^{(j)}_{([1/m_{(k)}])} z^{(j)} > t_0\}$$

and

$$\widehat{\tilde{\nu}}_0(\boldsymbol{x}, \boldsymbol{\infty}] = \widehat{\nu}_0(\boldsymbol{x}^{\hat{\boldsymbol{\alpha}}}, \boldsymbol{\infty}].$$

Recall that $\widehat{\nu_0}$ is given in (4.1) and $\hat{\boldsymbol{\alpha}} = (\hat{\alpha}_1, \ldots, \hat{\alpha}_2)$ can be estimated using the Hill or MLE estimator applied to the marginal one-dimensional samples. Finally we give a reminder that

$$\boldsymbol{x}^{\boldsymbol{\alpha}} = (x_1^{\alpha_1}, x_2^{\alpha_2}).$$

6. Two examples.

This section illustrates the usefulness of the our theoretical results for the two contexts of Internet traffic studies and flood risk analysis.

6.1. **Internet HTTP response data.** Our first example analyzes HTTP response data describing Internet transmissions observed during a four hour period from 1:00pm-5:00pm on 26th April 2001 at the University of North Carolina. The data sets were obtained from the University of North Caroline Computer Science Distributed and Real-Time Systems Group under the direction of Don Smith and Kevin Jeffay. Interest in this subject was stimulated by Steve Marron in his Mary Upson lectures at Cornell University in fall 2001.

Internet file transfers are subject to delays and although one expects larger file transfers to encounter more delays, this is overly simplistic (Campos et al. (2004)). Large file transfers, while comparatively rare, comprise a significant fraction of all the bytes transferred on the Internet and hence are important for understanding the impact of diverse networking technologies such as routing, congestion control, and server design on end-user performance measures. For HTTP (web browsing) responses, the joint behavior of large values of three variables – size of response, time duration of response, and throughput (syonym: rate = size / time) – can be considered. All three quantities are typically heavy tailed but size and rate tend to be asymptotically independent. See also Maulik et al. (2002), Resnick (2003). We consider here the existence of hidden regular variation for the variables (size, throughput).

The dataset consists of responses (bytes) in the stated time period whose size is in excess of 100 000 bytes coupled with the time required for transmission (seconds). There were 21829 such transmissions. As opposed to the next example on flood risk, interest here is not specifically in estimating probabilities of rare events such as joint threshold exceedance, but rather in understanding the underlying structure of the transmission process. This is intended to aid network engineers in their development of realistic models of Internet traffic processes, with which they simulate network behaviour. To this end, we focus on the establishment of asymptotic independence and hidden regular variation between transmission size and throughput, and then on obtaining an estimate of the hidden spectral measure characterizing the joint tail of the distribution of these variables. This would suggest that a suitable model for (size, throughput) could be obtained from a mixture model as in Maulik and Resnick (2003).

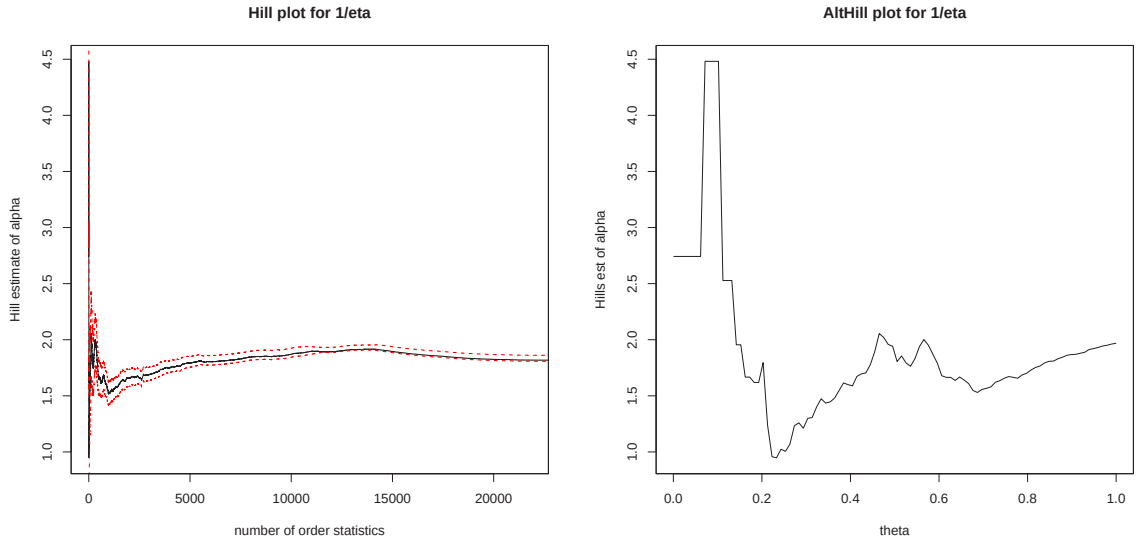

Figure 1. Hill plots on two scales for estimating the coefficient of tail dependence η.

Although estimation of the marginal distributions is not required for the examination of the dependence structure, we first estimate the marginal tail parameters to establish the heavy tailed behaviour of the transmission rate and size variables. We used Hill plots and AltHill plots (not shown) to choose values of $k = 150, 250$ for the size and rate variables respectively. Estimates of α were relatively stable around the values of $\alpha_1 = 1.8$ and $\alpha_2 = 2.1$ in the ranges $[50, 3000]$ and $[50, 400]$ for these variables respectively.

The next stage is to establish whether these Internet data exhibit asymptotic independence. We calculate the coefficient of tail dependence η of Ledford and Tawn (1996, 1997) using the Hill estimator for the shape parameter of the distribution of componentwise minima taken after rank transformation to standard Pareto margins. The Hill and AltHill plots used to inform our choice of k for this estimation are shown in Figure 1. These plots show estimates of $\alpha = 1/\eta$ constructed using different numbers of order statistics. The plots show the estimated value of η to be stable at around 0.6 for k in the range $[50, 400]$. This value is consistent with asymptotic independence and weak dependence between transmission size and rate at finite levels.

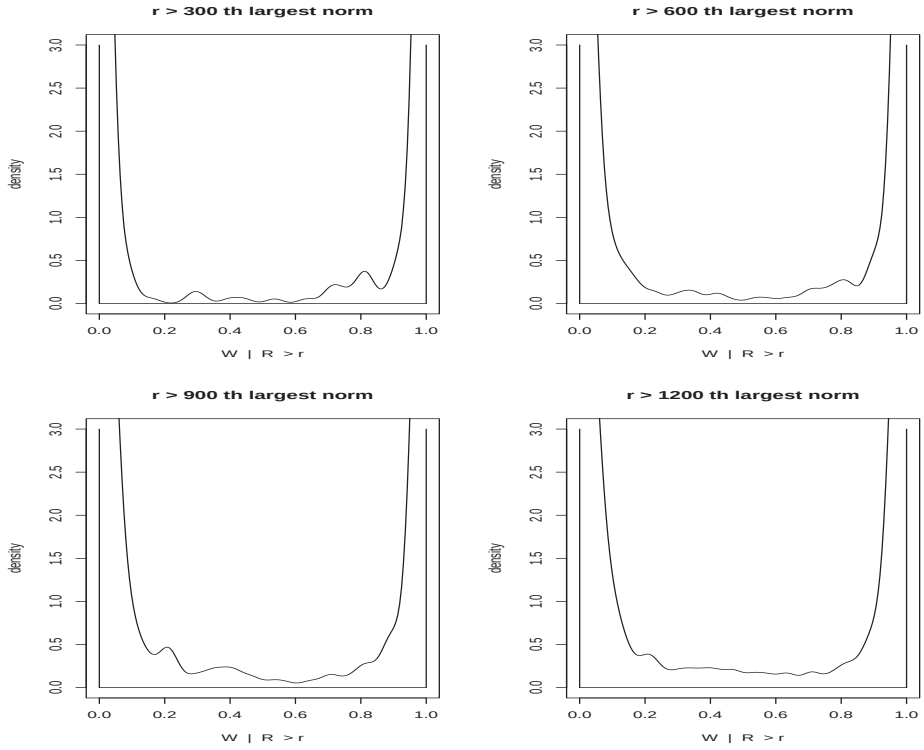

FIGURE 2. Density estimates of the spectral measure S for the Internet transmission size and rate data.

Another visual confirmation of asymptotic independence is to calculate the empirical measure (3.8) for different values of k. Figure 2 shows kernel density estimates corresponding to these estimates for $k = 300, 600, 900$ and 1200. These plots clearly show the mass concentrating towards the endpoints of the interval [0,1] as we use fewer and fewer order statistics. This gives further compelling evidence of asymptotic independence.

Having established that our variables are both heavy tailed and asymptotically independent variables, we can now employ steps 1)-4) given at the end of Section 4 to estimate the hidden measure. Implementation of these steps involves choice of k, and we use the Stărică scaling device (Resnick (2003), Stărică (1999, 2000)) to aid this choice. This tool is motivated by the homogeneity property (2.4) of the spectral measure ν_{standard}. Recall from (3.7), that $\hat{\nu}_{\text{standard},n}$ is the empirical estimate of ν_{standard} and contains the parameter

k which must be selected. Since ν_{standard} is homogeneous, Stărică argues that a good choice of k is that for which the following approximation is true:

$$\hat{\nu}_{\text{standard},n}(uA) \approx u^{-1}\hat{\nu}_{\text{standard},n}(A)$$

for u in a neighborhood of 1 and the set $A = \{\boldsymbol{x} \in \mathbb{E} : \|\boldsymbol{x}\| > 1\}$. For the Internet data, the value of k for which the scaling ratio

$$\hat{\nu}_{\text{standard},n}(uA)/\{u^{-1}\hat{\nu}_{\text{standard},n}(A)\} \tag{6.1}$$

is most stable around the value 1 for u in the interval [0.5,1.5] is $k = 1112$. The resulting a value of $m_{[1112]}$ corresponds to the 0.45 quantile of the radial components. Figure 3 shows this scaling ratio calculated using $k = 1112$ for u between 0 and 10.

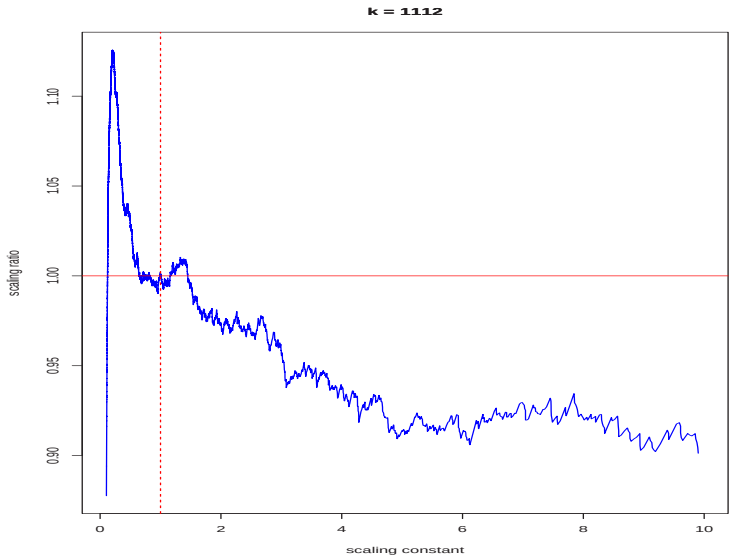

FIGURE 3. Stărică scaling plot for the Internet transmission size and rate data. Horizontal axis shows scaling constant u and vertical axis shows scaling ratio (6.1) for $k = 1112$.

Following this suggestion for the choice of k, we plot the estimated hidden measure for $k = 1000, 1150, 1200$ and 1250 in Figure 4. These plots show stability of the estimated measure for these values of k. Again, since the measure may be infinite, we have bounded the interval on which we estimate the measure away from 0 and 1 and show the kernel density estimate on the interval [0.1,0.9]. An edge correction has been applied so that the density integrates to 1 on this interval. All plots show the hidden measure to be bimodal with peaks around 0.2 and 0.85.

6.2. **Risk from flooding in Eskdalemuir.** For our next example, we illustrate the estimation of risk from flooding using hourly rainfall measurements (units: 0.1 mm) from Eskdalemuir in the south west of Scotland, during the years 1970–1986 inclusive. The dataset has also been considered in Nadarajah et al. (1998). The rainfall process exhibits both serial dependence and diurnal cycles. To avoid complications arising from these features we focus on consecutive pairs of hourly observations taken from 1100 to 1200 and from 1200 to 1300 daily. The choice of this particular pair of consecutive hours is arbitrary but does not materially affect the analysis which we now describe. This gives us a series of 6209 observations of the random pair $\boldsymbol{Z} = (Z^{(1)}, Z^{(2)})$, which we assume i.i.d. Around 80% of the observations are zero, corresponding to dry hours. The data are plotted in Figure 5.

Interest in extreme rainfall generally results from an intention to understand and protect against rainfall which can result in flooding. Four factors contribute to the likelihood of flooding following a rain event:

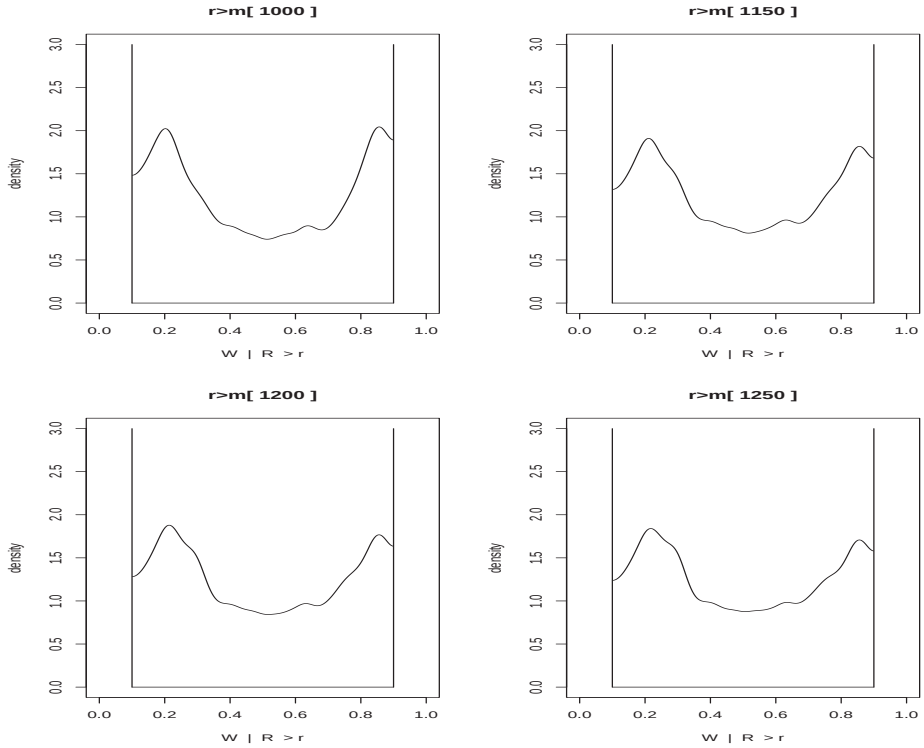

FIGURE 4. Density estimates of the hidden spectral measure S_0 for the Internet transmission size and rate data.

rainfall intensity and duration (meteorological), ground saturation and rainfall catchment response (hydrological). We focus on the meteorological factors. A historical study of extreme rainfall events in the UK (Hand (2002)) shows clearly that the rainfall events with the highest hourly precipitation rates in the UK are the shortest events, typically lasting less than one hour and exclusively resulting from "convective" rain. This type of precipitation is typically intense, of short duration, and is often accompanied by thunder and lighting. This finding suggests that rainfall amounts may be asymptotically independent from one hour to the next, as the intensity of such precipitation events is rarely sustained for long enough to make the rainfall counts extreme in both hours. Our aim in this example is to estimate the probability of flooding, ie rainfall in excess of threshold t_0 in consecutive hours for large values of t_0.

Diagnostic plots (not shown) such as the Hill, Alt Hill and QQ plots (Beirlant et al. (1996), Drees et al. (2000), Kratz and Resnick (1996), Resnick (2003), Resnick and Stărică (1997)) suggest that $k = 300$ is appropriate for estimation of the marginal parameters α_1 and α_2. These were estimated using the Hill estimator as 1.66 and 1.57 for 1100-1200 and 1200-1300 respectively. We checked for asymptotic independence by estimating S using the estimator in (3.8) for a variety of values of k. Plots of this empirical measure (not shown) show the movement of the mass to the endpoints of its support as k decreases. We confirmed this finding by estimating the coefficient of tail dependence of Ledford and Tawn (1996, 1997) using the Hill estimator for the shape parameter of the distribution of componentwise minima taken after rank transformation to standard Fréchet margins. The estimated value of this parameter was stable at around 0.83 for k in the range [10,500]. This value is consistent with asymptotic independence but with reasonably strong dependence at finite levels.

We chose the value of k to use for the dependence estimation (4.1), using the Stărică scaling plot (Resnick (2003), Stărică (1999, 2000)). This device suggests a value of $k = 210$ for this pair of variables. We use the above values of $\hat{\alpha}_1$, $\hat{\alpha}_2$ and $k = 210$ to obtain estimates of the probability of joint exceedance of $t_0 = 4, 5, 6$

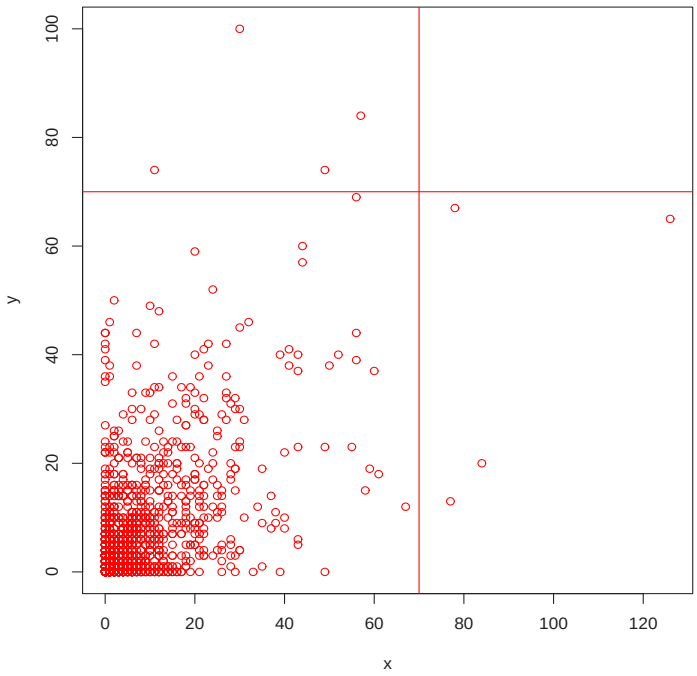

FIGURE 5. Scatter plot of 1 hour rainfall measurements $\boldsymbol{Z} = (Z^{(1)}, Z^{(2)})$ (0.1 mm) from Eskdalemuir, 1970-1986: the lines show the exceedance thresholds $t_0 = 70$, which are not jointly exceeded by any pair in the dataset.

and $7mm$ using our method as 6.0e-5, 3.2e-5, 2.2e-5 and 1.6e-5 respectively. For such values of t_0, empirical exceedance rates using the observed counts of joint exceedances are based on too few data to be reliable. In particular, for $t_0 = 70$ (indicated by lines on plot in Figure 5) there are no observed joint exceedances of this threshold. Our estimate of $\hat{\nu}_0$ uses a greater proportion of the data, then scales by k/n, according to expression (5.2).

An additional benefit of our proposed estimation scheme is the ability to visualise the hidden spectral measure S_0. We follow steps 1)-4) given at the end of Section 4 to construct points which we treat as a sample from the hidden angular distribution. Figure 6 shows the kernel density estimate of the angular components constructed using the L1 norm, and radial components above $m_{[k]}$ where $k = 210$. Since the hidden spectral measure may be infinite, we have bounded the support of our estimate away from the ends of the interval [0,1], and show the density estimate on the interval [0.1,0.9]. An edge correction has been applied to ensure that the density integrates to one on this interval. This density estimate is relatively stable for k in the range [100,300]. Note that $m_{[210]}$ corresponds to the 0.84 quantile of the distribution of radial components.

REFERENCES

J. Beirlant, P. Vynckier, and J. Teugels. Tail index estimation, Pareto quantile plots, and regression diagnostics. *J. Amer. Statist. Assoc.*, 91(436):1659–1667, 1996. ISSN 0162-1459.

J.T. Bruun and J.A. Tawn. Comparison of approaches for estimating the probability of coastal flooding. *J. R. Stat. Soc., Ser. C, Appl. Stat.*, 47(3):405–423, 1998.

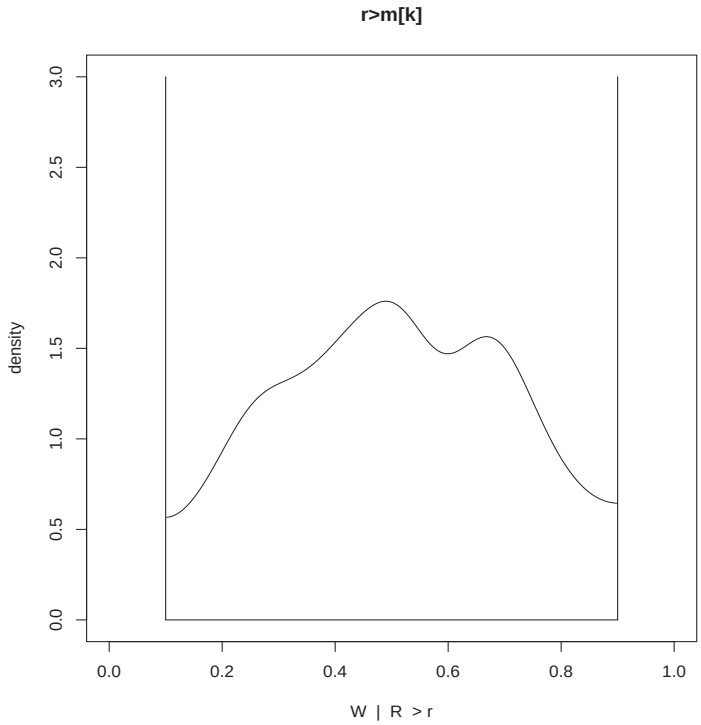

FIGURE 6. Density estimate of the hidden spectral measure S_0.

F. H. Campos, J.S. Marron, S.I. Resnick, and K. Jaffay. Extremal dependence: Internet traffic applications. Web available at http://www.orie.cornell.edu/∼sid. To appear: Stochastic Models., 2004.

S. Coles, J. Heffernan, and J. Tawn. Dependence measures for extreme value analyses. *Extremes*, 2(4):339–365, 1999.

L. de Haan and J. de Ronde. Sea and wind: multivariate extremes at work. *Extremes*, 1(1):7–46, 1998.

L. de Haan and E. Omey. Integrals and derivatives of regularly varying functions in $\mathbb{R}^d$ and domains of attraction of stable distributions II. *Stochastic Process. Appl.*, 16(2):157–170, 1984. ISSN 0304-4149.

G. Draisma, H. Drees, A. Ferreira, and L. De Haan. Tail dependence in independence. Technical report, Erasmus University Rotterdam, Econometric Institute, EUR, PO Box 1736, 3000DR Rotterdam, 2001.

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

P. Greenwood and S. Resnick. A bivariate stable characterization and domains of attraction. *Journal of Multivariate Analysis*, 9:206–221, 1979.

W. Hand. Numerical weather prediction: a historical study of rainfall events in the 20th century. Forecasting Research Technical Report No. 384, UK Met Office, 2002.

J.E. Heffernan. A directory of coefficients of tail dependence. *Extremes*, 3(3):279–290, 2000. ISSN 1386-1999.

Xin Huang. *Statistics of Bivariate Extreme values*. Ph.D. thesis, Tinbergen Institute Research Series 22, Erasmus University Rotterdam, Postbus 1735, 3000DR, Rotterdam, The Netherlands, 1992.

M. Kratz and S.I. Resnick. The qq–estimator and heavy tails. *Stochastic Models*, 12:699–724, 1996.

A.W. Ledford and J.A. Tawn. Statistics for near independence in multivariate extreme values. *Biometrika*, 83(1): 169–187, 1996. ISSN 0006-3444.

A.W. Ledford and J.A. Tawn. Modelling dependence within joint tail regions. *J. Roy. Statist. Soc. Ser. B*, 59(2): 475–499, 1997. ISSN 0035-9246.

K. Maulik and S.I. Resnick. Characterizations and examples of hidden regular variation. Technical report, School of ORIE, Cornell University, 2003. Web available at www.orie.cornell.edu/∼sid; to appear: *Extremes*.

K. Maulik, S.I. Resnick, and H. Rootzén. Asymptotic independence and a network traffic model. *J. Appl. Probab.*, 39(4):671–699, 2002. ISSN 0021-9002.

S. Nadarajah, C. W. Anderson, and J. A. Tawn. Ordered multivariate extremes. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 60(2):473–496, 1998. ISSN 1369-7412.

L. Peng. Estimation of the coefficient of tail dependence in bivariate extremes. *Statist. Probab. Lett.*, 43(4):399–409, 1999. ISSN 0167-7152.

S.-H Poon, M. Rockinger, and J. Tawn. Modelling extreme-value dependence in international stock markets. *Statist. Sinica*, 13(4):929–953, 2003. ISSN 1017-0405. Statistical applications in financial econometrics.

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

S.I. Resnick. *Extreme Values, Regular Variation and Point Processes*. Springer-Verlag, New York, 1987.

S.I. Resnick. Hidden regular variation, second order regular variation and asymptotic independence. *Extremes*, 5(4): 303–336, 2002.

S.I. Resnick. Modeling data networks. In B. Finkenstadt and H. Rootzen, editors, *SemStat: Seminaire Europeen de Statistique, Exteme Values in Finance, Telecommunications, and the Environment*, pages 287–372. Chapman-Hall, London, 2003.

S.I. Resnick. The extremal dependence measure and asymptotic independence. *Stochastic Models*, 20(2):205–227, 2004.

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

C. Stărică. Multivariate extremes for models with constant conditional correlations. *J. Empirical Finance*, 6:515–553, 1999.

C. Stărică. Multivariate extremes for models with constant conditional correlations. In P. Embrechts, editor, *Extremes and Integrated Risk Management*, pages 515–553. Risk Books, London, 2000.

Janet Heffernan, Department of Mathematics and Statistics, Lancaster University, Lancaster, LA1 4YF, UK
E-mail address: j.heffernan@lancaster.ac.uk

Sidney Resnick, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY 14853
E-mail address: sid@orie.cornell.edu