Extremes (2013) 16:407–428
DOI 10.1007/s10687-012-0166-x

Extreme value theory with operator norming

Mark M. Meerschaert · Hans-Peter Scheffler · Stilian A. Stoev

Received: 7 July 2012 / Revised: 3 December 2012 /
Accepted: 14 December 2012 / Published online: 27 January 2013

Abstract A new approach to extreme value theory is presented for vector data with heavy tails. The tail index is allowed to vary with direction, where the directions are not necessarily along the coordinate axes. Basic asymptotic theory is developed, using operator regular variation and extremal integrals. A test is proposed to judge whether the tail index varies with direction in any given data set.

Keywords Operator regular variation · Heavy tails · Directional extremes · Spectral representation · Parametric bootstrap · Hetero-ouracity

AMS 2000 Subject Classifications Primary—60G70 · 60G52 · 60F17 · Secondary—62G32 · 62G09 · 62G10

Mark M. Meerschaert was partially supported by NSF grant DMS-1025486 and NIH grant R01-EB012079.

Stilian A. Stoev was partially supported by the NSF grant DMS–1106695 at the University of Michigan.

M. M. Meerschaert
Department of Statistics & Probability, Michigan State University, East Lansing, MI 48824, USA
e-mail: mcubed@stt.msu.edu
URL: http://www.stt.msu.edu/users/mcubed/

H.-P. Scheffler
Fachbereich Mathematik, Universität Siegen, 57068 Siegen, Germany
e-mail: scheffler@mathematik.uni-siegen.de
URL: http://www.uni-siegen.de/fb6/src/scheffler/

S. Stoev (✉)
Department of Statistics, University of Michigan, Ann Arbor, MI 48109, USA
e-mail: sstoev@umich.edu
URL: http://www.stat.lsa.umich.edu/~sstoev

1 Introduction

This paper develops a new approach to extreme value theory for vector data with heavy tails. The primary goal is to allow the tail index to vary with direction, in a coordinate-free setting. The main technical tools are vector regular variation with norming by linear operators (Balkema 1973; Meerschaert and Scheffler 2001; Meerschaert 1988; Resnick 1987), and extremal integrals (de Haan 1984; Stoev and Taqqu 2005; Kabluchko 2009). Taking the point of view of directional extremes, data are projected onto each radial direction, and the maximum in each direction is considered. This leads to an extremal limit process indexed by the direction. The extremal limit theory employs operator norming, allowing the tail index to vary with the direction. Continuous mapping arguments yield a useful comparison of extreme behavior in different directions, leading to a useful test for variations in the tail index i.e., *hetero-ouracity*. Distributions with different tail exponents in different directions will be said to have hetero-ouracity (from the greek word for tail – $o\upsilon\rho\acute{\alpha}$). This paper was motivated by the observation that vector data with heavy tails need not have the same tail index in every direction (Meerschaert and Scalas 2006; Mittnik and Rachev 1999; Reeves et al. 2008), and that it can be necessary to consider rotated coordinate systems (which need not be orthogonal) to detect variations in tail behavior (Meerschaert and Scheffler 2003; Painter et al. 2002).

A simple example illustrates our general approach. Suppose that $X, X_1, X_2, X_3, \ldots$ are independent and identically distributed (iid) random variables whose tail distribution $\bar{F}(x) = \mathbb{P}(X > x)$ varies regularly at infinity with index $-\alpha$ for some $\alpha > 0$. Then $n\bar{F}(c_n^{-1}x) \to Cx^{-\alpha}$, as $x \to \infty$, for some $C > 0$ and some regularly varying sequence (c_n) with index $-1/\alpha$. Let $M_n = \max\{X_1, \ldots, X_n\}$ and note that for $x > 0$

$$\mathbb{P}(c_n M_n \leq x) = \left(1 - \frac{n\bar{F}(c_n^{-1}x)}{n}\right)^n \to \exp(-Cx^{-\alpha})$$

as $n \to \infty$, so that the normalized maximum $c_n M_n$ converges in law to a Fréchet random variable Y with $\mathbb{P}(Y \leq x) = \exp(-Cx^{-\alpha}),\ x > 0$. For vector data, the same argument shows that the random variables $\langle X_i, \theta\rangle$ for any direction vector $\theta \neq 0$ are attracted to an α-Fréchet limit if the tail function $\bar{F}_\theta(x) = \mathbb{P}(\langle X_i, \theta\rangle > x)$ varies regularly with index $-\alpha$. Of course it is possible that the tail index $\alpha = \alpha(\theta)$ varies with the direction θ. This paper develops a general theory for such directional extremes, where the tail index is allowed to vary with the direction. By considering the joint convergence over all directions θ, a functional limit theorem is established, using norming by linear operators to retain the full tail information. Then continuous mapping arguments can be used to compare extremes in different directions.

A functional limit theorem for directional extremes is proven in Section 2. Section 3 develops operator max-stability properties of the limit process, and simulation methods based on a Poisson point process representation. Section 4 considers the special, scalar norming case where the tail index is the same in every direction, and gives a representation of the limit process in terms of extremal integrals. A useful test for hetero-ouracity is developed in Section 5, to determine whether the data can be treated with scalar norming, with the same tail index in every radial direction. Section 6 reports the results of a small simulation study to validate the practical utility of this test, and Section 7 applies the test to two different data sets, one from finance and another from hydrology, to see whether the test can detect a difference in the tail index.

2 Limit theory

Our theory of directional extremes is based on the notion of operator regular variation, which was developed and used extensively for the study of sums of independent random vectors (Meerschaert and Scheffler 2001) whose tail index can vary with direction. A probability distribution μ on $\mathbb{R}^d$ is said to be *operator regularly varying* at infinity if

$$n\mu(A_n^{-1}\cdot) \xrightarrow{v} \phi(\cdot), \quad \text{as } n \to \infty, \tag{1}$$

where the linear operators $A_n \to 0$ in norm, and the limit ϕ is a *full* (not supported on any lower dimensional subspace) Borel measure that assigns finite mass to *tail sets* (sets bounded away from the origin). The vague convergence $\xrightarrow{v}$ means that $n\mu(A_n^{-1}B) \longrightarrow \phi(B)$ for any Borel tail set $B \subset \mathbb{R}^d$, with $\phi(\partial B) = 0$. It follows from Eq. 1 that

$$t\phi(B) = \phi(t^{-E}B), \quad \text{for all } t > 0 \text{ and } B \in \mathcal{B}(\mathbb{R}^d \setminus \{\mathbf{0}\}), \tag{2}$$

for some *exponent* matrix E, where $t^{-E} = \exp(-E\log(t))$, and $\exp(A) = I + A + A^2/2! + \cdots$ is the usual matrix exponential (Meerschaert and Scheffler 2001, Proposition 6.1.2). The normalizing sequence of operators A_n may be chosen so that

$$A_{[tn]}A_n^{-1} \to t^{-E} \quad \text{as } n \to \infty, \quad \forall t > 0, \tag{3}$$

where $[\cdot]$ denotes the greatest integer function (Meerschaert and Scheffler 2001, Theorem 6.1.24). This extends scalar-normed multivariate regular variation (Resnick 2007) in which $A_n = c_n I$, a scalar multiple of the identity matrix. If $A_n = c_n I$, then $E = (1/\alpha)I$ for some $\alpha > 0$, and Eq. 3 shows that the sequence c_n varies regularly with index $-1/\alpha$. Then $\mathbb{P}(\|X\| > x)$ varies regularly with index $-\alpha$. In fact, operator regular variation allows one to treat in a unified way distributions with different tail exponents along different coordinate axes that need not be the original axes nor orthogonal. In the general case, Theorem 6.4.15 in Meerschaert and Scheffler (2001) yields a tail index function $\alpha(\theta)$ such that, for any direction vector $\theta \neq 0$, for any $\varepsilon > 0$, there exists an x_0 such that

$$x^{-\alpha(\theta)-\varepsilon} < \mathbb{P}(|\langle X, \theta\rangle| > x) < x^{-\alpha(\theta)+\varepsilon} \tag{4}$$

for all $x \geq x_0$. For example, if A_n is a diagonal matrix, then E is also diagonal with positive entries $1/\alpha_1 \leq \cdots \leq 1/\alpha_d$, and $\alpha(\theta) = \min\{\alpha_i : \theta_i \neq 0\}$, where θ_i is the ith component of the direction vector θ, so that the heaviest tail dominates.

The limit measure ϕ in Eq. 2 has a convenient spectral representation (Meerschaert and Scheffler 2001, Theorem 6.1.7): Let $\|\cdot\|_E$ be a norm on $\mathbb{R}^d$ depending on E

in such a way that: (i) for all $x \neq \mathbf{0}$, $t \mapsto \|t^E x\|$ is strictly increasing in $t > 0$; and (ii) $(t, x) \mapsto t^E x$ is a homeomorphism from $(0, \infty) \times \mathbb{S}_E$ onto $\mathbb{R}^d \setminus \{\mathbf{0}\}$, where $\mathbb{S}_E := \{x \in \mathbb{R}^d : \|x\|_E = 1\}$. The existence of such a norm is guaranteed by Jurek and Mason (1993, Proposition 3.4.3), see also Meerschaert and Scheffler (2001, Lemma 6.1.5). Then the operator scaling property (2) allows us to write

$$\phi(A) = \int_{\mathbb{S}_E} \int_0^\infty \mathbf{1}_A(t^E\theta) \frac{\mathrm{d}t}{t^2} \lambda(\mathrm{d}\theta)\,, \tag{5}$$

where the *spectral measure* λ is defined as follows

$$\lambda(C) := \phi\{r^E\theta : r > 1, \theta \in C\}. \tag{6}$$

Let $X, X_1, \ldots, X_n$ be iid random vectors in $\mathbb{R}^d$ whose distribution μ is operator regularly varying, so that Eq. 1 holds. Consider an arbitrary direction $\theta \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ and define the directional maximum

$$M_n(\theta) := \max_{i=1,\ldots,n} \langle X_i, \theta \rangle. \tag{7}$$

We view $M_n(\theta)$ as a stochastic process indexed by θ, and we establish a limit theorem for $\{M_n(A_n^*\theta) : \theta \neq \mathbf{0}\}$, where A_n is from Eq. 1, and A_n^* denotes its transpose. If $A_n = c_n I$ is scalar, then $M_n(A_n^*\theta) = c_n M_n(\theta)$, and our results reduce to the normalized maxima of the iid random variables $\langle X_i, \theta \rangle$. The process convergence will allow us to compare directional extremes using continuous mapping arguments.

Introduce the half-spaces

$$B(r, \theta) = \{x \in \mathbb{R}^d : \langle x, \theta \rangle > r\}\,, \quad r \in \mathbb{R},\ \theta \in \mathbb{R}^d \setminus \{\mathbf{0}\}, \tag{8}$$

and notice that $B(r, \theta)$ is a *tail set* if $r > 0$, i.e., bounded away from the origin. Let $C(\mathbb{R}^d \setminus \{\mathbf{0}\}) \equiv C(\mathbb{R}^d \setminus \{\mathbf{0}\}; \mathbb{R})$ denote the set of continuous functions on $\mathbb{R}^d \setminus \{\mathbf{0}\}$, equipped with the topology induced by the uniform convergence of functions on all compact sets. This topology is generated, for example, by the metric

$$\rho(f, g) := \sum_{n=1}^{\infty} \max_{\|\theta\| \in [1/n, n]} |f(\theta) - g(\theta)| \wedge 2^{-n},$$

which turns $C(\mathbb{R}^d \setminus \{\mathbf{0}\})$ into a complete separable metric space.

The next theorem is the main result of this section. It shows that the rescaled directional extremes process $\{M_n(A_n^*\theta)\}$ converges weakly in $C(\mathbb{R}^d \setminus \{\mathbf{0}\})$ to a process $\{Y(\theta)\}$. The finite-dimensional distributions of the limit process will be given by

$$\begin{aligned} \mathbb{P}\{Y(\theta_1) \le r_1, \ldots, Y(\theta_m) \le r_m\} &= F_{\theta_1,\ldots,\theta_m}(r_1, \ldots, r_m) \\ &:= \exp\left\{-\phi\left(\bigcup_{j=1}^m B(r_j, \theta_j)\right)\right\}, \end{aligned} \tag{9}$$

when all $r_j > 0$, and $F_{\theta_1,\ldots,\theta_m}(r_1, \ldots, r_m) := 0$ when any $r_j < 0$. This defines the joint distribution function on a dense set D, and we extend to all $(r_1, \ldots, r_m)$ by taking the right-continuous limit.

Theorem 2.1 *Let M_n be as in Eq. 7 where X_i are iid with μ and Eq. 1 holds. Then*

$$\{M_n(A_n^*\theta)\}_{\theta\in\mathbb{R}^d\setminus\{\mathbf{0}\}} \xrightarrow{d} \{Y(\theta)\}_{\theta\in\mathbb{R}^d\setminus\{\mathbf{0}\}} \text{ as } n\to\infty \text{ in } C(\mathbb{R}^d\setminus\{\mathbf{0}\}) \tag{10}$$

where $\mathcal{Y}=\{Y(\theta)\}_{\theta\in\mathbb{R}^d\setminus\{\mathbf{0}\}}$ is defined by Eq. 9.

The proof of Theorem 2.1 requires a few simple lemmas. For all $\theta_j\in\mathbb{R}^d\setminus\{\mathbf{0}\}$, $j=1,\dots,m$, consider the distribution functions

$$F_{n,\theta_1,\dots,\theta_m}(r_1,\dots,r_m) := \mathbb{P}\{M_n(A_n^*\theta_j)\le r_j, \forall j=1,\dots,m\}. \tag{11}$$

Lemma 2.2 *Under the assumptions of Theorem 2.1, for all $\theta_j\in\mathbb{R}^d\setminus\{\mathbf{0}\}$, $j=1,\dots,m$, we have that $F_{\theta_1,\dots,\theta_m}$ in Eq. 9 is a valid distribution function and*

$$F_{n,\theta_1,\dots,\theta_m} \xrightarrow{w} F_{\theta_1,\dots,\theta_m}, \quad \text{as } n\to\infty. \tag{12}$$

Proof First suppose that $r_j>0$ for all $j=1,\dots,m$. Since $\langle A_nX_i,\theta\rangle=\langle X_i,A_n^*\theta\rangle$, in view of Eqs. 7 and 11, we obtain

$$\begin{aligned} F_{n,\theta_1,\dots,\theta_m}(r_1,\dots,r_m) &= \mathbb{P}\{\langle A_nX_i,\theta_j\rangle\le r_j,\ \forall j=1,\dots,m,\ i=1,\dots,n\} \\ &= \mathbb{P}\left\{A_nX_i\in\bigcap_{j=1}^m B(r_j,\theta_j)^c,\ \forall i=1,\dots,n\right\} \\ &= \mathbb{P}\{A_nX\in B^c\}^n = \left(1-\frac{n\mathbb{P}\{A_nX\in B\}}{n}\right)^n, \end{aligned} \tag{13}$$

where $B^c:=\bigcap_{j=1}^m B(r_j,\theta_j)^c=(\bigcup_{j=1}^m B(r_j,\theta_j))^c$. Note that $B=\bigcup_{j=1}^m B(r_j,\theta_j)$ is a tail set. Lemma 6.1.27 in Meerschaert and Scheffler (2001) shows that $B(r,\theta)$ is a continuity set for any $r>0$ and $\theta\ne\mathbf{0}$. Thus, in view of Eqs. 1 and 13, we obtain

$$F_{n,\theta_1,\dots,\theta_m}(r_1,\dots,r_m) \longrightarrow \exp(-\phi(B)) \equiv F_{\theta_1,\dots,\theta_m}(r_1,\dots,r_m),$$

as $n\to\infty$.

Since ϕ is full, it follows from Eq. 2 that $\mathcal{R}e(\lambda)>0$ for all eigenvalues λ of E, so that $n^{-E}\to 0$ as $n\to\infty$ in operator norm (Meerschaert and Scheffler 2001, Lemma 6.1.4). This, together with the scaling relation (2), implies that the measure ϕ assigns infinite mass to any neighborhood of the origin. Extend ϕ to a σ-finite Borel measure on $\mathbb{R}^d$ by setting $\phi\{0\}=0$. Then the regular variation condition (1) implies that $n\mu(A_n^{-1}B)\to\infty$ for any Borel set B that contains an open neighborhood of the origin. If $r_j<0$ for some $j=1,\dots,m$, then B contains a neighborhood of the origin, so $n\mathbb{P}\{A_nX\in B\}\to\infty$, and hence

$$F_{n,\theta_1,\dots,\theta_m}(r_1,\dots,r_m) = \left(1-\frac{n\mathbb{P}\{A_nX\in B\}}{n}\right)^n \to 0 = F_{\theta_1,\dots,\theta_m}(r_1,\dots,r_m). \tag{14}$$

Since we have established convergence for all $(r_1,\dots,r_m)$ in a dense subset of $\mathbb{R}^m$, and since $F_{\theta_1,\dots,\theta_m}(\infty,\dots,\infty)=1$, it follows that Eq. 12 holds. In particular, $F_{\theta_1,\dots,\theta_m}$ is a valid distribution function. □

Lemma 2.3 *There exists a stochastic process* $\mathcal{Y} = \{Y(\theta)\}_{\theta \in \mathbb{R}^d \setminus \{\mathbf{0}\}}$ *with finite-dimensional distributions as in Eq. 9.*

Proof By Eq. 9, we have that $F_{\theta_1,\ldots,\theta_{m-1},\theta_m}(r_1,\ldots,r_{m-1},\infty) = F_{\theta_1,\ldots,\theta_{m-1}}(r_1,\ldots, r_{m-1})$, for all $r_j > 0$ and $\theta_j \in \mathbb{R}^d \setminus \{\mathbf{0}\}$, $j = 1,\ldots,m-1$ and the latter relation is invariant to permutations. Therefore, $\{F_{\theta_1,\ldots,\theta_m} : \theta_1,\ldots,\theta_m \in \mathbb{R}^d \setminus \{\mathbf{0}\}, m \in \mathbb{N}\}$ is a projective system of distributions, and the statement follows from the Kolmogorov consistency theorem. □

Lemma 2.4 *The distributions of the processes* $\{M_n\}_{n\in\mathbb{N}}$ *are tight in* $C(\mathbb{R}^d \setminus \{\mathbf{0}\})$.

Proof Consider the modulus of continuity

$$\omega_n(\delta) := \sup_{\theta,\theta' \in \mathbb{R}^d \setminus \{\mathbf{0}\} : \|\theta-\theta'\|<\delta} |M_n(A_n^*\theta) - M_n(A_n^*\theta')|, \tag{15}$$

where $\|\cdot\|$ stands for the Euclidean norm in $\mathbb{R}^d$. By Theorem 3.1.1 in Khoshnevisan (2002), it is enough to show that

$$\lim_{\delta\downarrow 0} \limsup_{n\to\infty} \mathbb{P}\{\omega_n(\delta) > \epsilon\} = 0\,, \text{ for all } \epsilon > 0\,. \tag{16}$$

Observe that for all $a_i, b_i \in \mathbb{R}$, $i = 1,\ldots,n$, $n \in \mathbb{N}$, we have

$$\left|\bigvee_{i=1}^n a_i - \bigvee_{i=1}^n b_i\right| \le \bigvee_{i=1}^n |a_i - b_i|\,. \tag{17}$$

Now, by Eqs. 15 and 17,

$$\begin{aligned}\omega_n(\delta) &\le \sup_{\|\theta-\theta'\|<\delta} \bigvee_{i=1}^n |\langle A_n X_i, \theta\rangle - \langle A_n X_i, \theta'\rangle| = \sup_{\|\theta-\theta'\|<\delta} \bigvee_{i=1}^n |\langle A_n X_i, (\theta-\theta')\rangle| \\ &\le \delta \bigvee_{i=1}^n \|A_n X_i\| =: \delta\, Z_n\,, \end{aligned} \tag{18}$$

almost surely. Introduce the tail sets $B(r) = \{x \in \mathbb{R}^d : \|x\| > r\}$ and observe that by Eq. 1,

$$\begin{aligned}\mathbb{P}\{Z_n \le r\} &= \mathbb{P}\{A_n X_i \in B(r)^c, \forall i = 1,\ldots,n\} \\ &= \left(1 - \frac{n\mu(A_n^{-1}B(r))}{n}\right)^n \to e^{-\phi(B(r))}, \quad \text{as } n\to\infty,\end{aligned} \tag{19}$$

for all but countably many r's. In view of Eq. 18, we have

$$\limsup_{n\to\infty} \mathbb{P}\{\omega_n(\delta) > \epsilon\} \le \limsup_{n\to\infty} \mathbb{P}\{Z_n > \epsilon/\delta\} =: g(\delta).$$

Notice that the function $g(\delta)$ is bounded and non-decreasing and by Eq. 19, we have $g(\delta) = 1 - e^{-\phi(B(\epsilon/\delta))}$, for all but countably many $\delta > 0$. This, since $\phi(B(\epsilon/\delta)) \downarrow 0$, as $\delta \downarrow 0$, implies $g(\delta) \downarrow 0$, and hence Eq. 16. □

Proof of Theorem 2.1 Lemmas 2.2 and 2.3 show that Eq. 10 holds in the sense of convergence of the finite–dimensional distributions, where $\mathcal{Y}$ is a *bona fide* stochastic process. Lemma 2.4 and Prokhorov's theorem imply that the convergence in Eq. 10 holds in $(C(\mathbb{R}^d \setminus \{\mathbf{0}\}), \rho)$. □

The next result gives a Poisson representation of the extremal limit process, akin to the de Haan spectral representation of a max-stable process (de Haan 1984; de Haan and Ferreira 2006; Kabluchko 2009; Stoev and Taqqu 2005). This construction uses the spectral decomposition (5) and (6).

Proposition 2.5 *Let* $0 < \Gamma_1 < \Gamma_2 < \cdots$ *be the arrival times of a Poisson process with constant rate* $\lambda(\mathbb{S}_E) > 0$*, and take* $\Lambda_i, i \in \mathbb{N}$ *iid random vectors on* $\mathbb{S}_E$ *with distribution* $\lambda(\cdot)/\lambda(\mathbb{S}_E)$*, independent of* (Γ_i)*. Then*

$$\{Y(\theta)\}_{\theta\in\mathbb{R}^d\setminus\{\mathbf{0}\}} \stackrel{d}{=} \left\{\bigvee_{i\in\mathbb{N}} \langle \Gamma_i^{-E}\Lambda_i, \theta\rangle\right\}_{\theta\in\mathbb{R}^d\setminus\{\mathbf{0}\}} \tag{20}$$

where E *is the exponent from Eq. 2, and* λ *is the spectral measure from Eq. 5.*

Proof Write $\epsilon_i = \Gamma_i^{-E}\Lambda_i$, and use the disintegration formula (5) to see that $\mathcal{N} = \{\epsilon_i\}_{i\in\mathbb{N}}$ is Poisson point process on $\mathbb{R}^d \setminus \{\mathbf{0}\}$ with intensity measure ϕ. Indeed, this follows, from the fact that $\mathcal{N}$ is a measurable transformation of the Poisson point process $\{(\Gamma_i, \Lambda_i),\ i \in \mathbb{N}\}$ on $(0, \infty) \times \mathbb{S}_E$ (see e.g., Propositions 3.7 & 3.8 in Resnick 1987). Let $\widetilde{Y}(\theta) := \sup_{i\in\mathbb{N}}\langle\epsilon_i, \theta\rangle$ for any $\theta \in \mathbb{R}^d \setminus \{\mathbf{0}\}$. For arbitrary $\theta_j \in \mathbb{R}^d \setminus \{\mathbf{0}\}$, $r_j > 0,\ j = 1, \ldots, m$, we have that

$$\begin{aligned}
\mathbb{P}\{\widetilde{Y}(\theta_j) \le r_j,\ \forall j = 1, \ldots, m\} &= \mathbb{P}\{\langle\epsilon_i, \theta_j\rangle \le r_j,\ \forall i \in \mathbb{N},\ j = 1, \ldots, m\} \\
&= \mathbb{P}\left\{\mathcal{N} \subset \bigcap_{j=1}^m B(\theta_j, r_j)^c\right\} \\
&= \mathbb{P}\left\{\mathcal{N} \cap \left(\bigcup_{j=1}^m B(r_j, \theta_j)\right) = \emptyset\right\} \\
&= \exp\left\{-\phi\left(\bigcup_{j=1}^m B(r_j, \theta_j)\right)\right\}
\end{aligned}$$

and apply Eq. 9 to finish the proof. □

Theorem 2.1 can be employed, along with continuous mapping arguments, to compare extremes in different directions. Define

$$V_n(\theta) := \max_{1\le i\le n}\{|\langle X_i, \theta\rangle|\} \equiv \max_{1\le i\le n}\{\langle X_i, \theta\rangle \vee \langle X_i, -\theta\rangle\}$$

and let $Y^{|\cdot|}(\theta) := Y(\theta) \vee Y(-\theta)$.

Proposition 2.6 *Under the assumptions of Theorem* 2.1, *we have, as $n \to \infty$,*

$$\{V_n(A_n^*\theta)\}_{\theta\in\mathbb{R}^d\setminus\{\mathbf{0}\}} \xrightarrow{d} \{Y^{|\cdot|}(\theta)\}_{\theta\in\mathbb{R}^d\setminus\{\mathbf{0}\}}, \quad in\ C(\mathbb{R}^d\setminus\{\mathbf{0}\}), \tag{21}$$

as well as

$$\Big(\max_{\|\theta\|=1} V_n(A_n^*\theta),\ \min_{\|\theta\|=1} V_n(A_n^*\theta)\Big) \xrightarrow{d} \big(Y^{(\max)}, Y^{(\min)}\big) \tag{22}$$

where $Y^{(\max)} := \max_{\|\theta\|=1} Y^{|\cdot|}(\theta)$, $Y^{(\min)} := \min_{\|\theta\|=1} Y^{|\cdot|}(\theta)$. *Moreover,* $Y^{(\min)} > 0$ *almost surely.*

Proof Relations (21) and (22) follow from Theorem 2.1 by a simple continuous mapping argument. Use Proposition 2.5 to see that

$$\mathbb{P}\{Y^{(\min)} = 0\} = \mathbb{P}\left\{\min_{\|\theta\|=1} \bigvee_{i\in\mathbb{N}} |\langle\Gamma_i^{-E}\Lambda_i, \theta\rangle| = 0\right\}$$

and note that the event on the right-hand side is equivalent to the event that every sample point of the Poisson point process $\mathcal{N} = \{\epsilon_i\}_{i\in\mathbb{N}}$ with $\epsilon_i = \Gamma_i^{-E}\Lambda_i$ and full control measure ϕ lies on some lower dimensional subspace $\{x \in \mathbb{R}^d : |\langle x, \theta\rangle| = 0\}$ for some $\|\theta\| = 1$. Since ϕ assigns zero measure to any lower dimensional subspace, an easy conditioning argument based on Slyvniak's formula (see e.g. Proposition 2.3 in Garcia and Kurtz 2008) shows that the sample points of this Poisson point process are a.s. not all contained in the same lower dimensional subspace. Then it follows that $Y^{(\min)} > 0$ almost surely. □

It follows from Eq. 2 with ϕ full dimensional that every eigenvalue of E has positive real part. Write these in strictly increasing order $0 < a_1 = 1/\alpha_1 < a_2 = 1/\alpha_2 < \cdots < a_p = 1/\alpha_p$, so that $\alpha_1 > \cdots > \alpha_p$ are the tail indices of different directional extremes in Eq. 4 (not counting multiplicities) where $1 \le p \le d$.

Define the sample counterparts to $Y^{(\max)}$ and $Y^{(\min)}$, that is, let

$$V_n^{(\max)} := \max_{\|\theta\|=1} V_n(\theta) \equiv \max_{1\le i\le n} \|X_i\| \quad \text{and} \quad V_n^{(\min)} := \min_{\|\theta\|=1} V_n(\theta) \tag{23}$$

The next result gives consistent estimators of the largest and smallest tail index.

Proposition 2.7 *Under the above assumptions, we have that, as $n \to \infty$,*

$$\frac{\log V_n^{(\max)}}{\log n} \xrightarrow{\mathbb{P}} 1/\alpha_p \tag{24}$$

as well as

$$\frac{\log V_n^{(\min)}}{\log n} \xrightarrow{\mathbb{P}} 1/\alpha_1. \tag{25}$$

Proof For the proof we need a technical tool called the spectral decomposition of $\mathbb{R}^d$ with respect to the exponent E. See Meerschaert and Scheffler (2001), Section 4.3 and page 406 for details. Let $W_1, \ldots, W_p$ denote the spectral decomposition of $\mathbb{R}^d$ with respect to E and assume without loss of generality that (A_n) is spectrally compatible with $-E$. For $i = 1, \dots, p$ let $\bar{L}_i = W_1 \oplus \cdots \oplus W_i$. Since $(A_n^{-1})^*$ is regularly varying with exponent E^* and spectrally compatible with E^*, Proposition 4.3.14 in Meerschaert and Scheffler (2001) shows that the conclusions of Theorem 4.3.1 in Meerschaert and Scheffler (2001) hold with $\bar{L}_i$.

We first prove Eq. 24. Fix any $\delta > 0$ and note that

$$\mathbb{P}\left\{\left|\frac{\log V_n^{(\max)}}{\log n} - a_p\right| > \delta\right\} \le \mathbb{P}\left\{V_n^{(\max)} > n^{a_p+\delta}\right\} + \mathbb{P}\left\{V_n^{(\max)} < n^{a_p-\delta}\right\}.$$

Choose $\theta_0 \in \bar{L}_p \setminus \bar{L}_{p-1}$, $\|\theta_0\| = 1$ and note that, in view of Theorem 4.3.1 in Meerschaert and Scheffler (2001), for any $0 < \varepsilon < \delta$, there exists a constant $C > 0$ such that, if we write $(A_n^{-1})^*\theta_0 = r_n\theta_n$, for some $r_n > 0$ and $\|\theta_n\| = 1$, we have $C^{-1}n^{a_p-\varepsilon} \le r_n$ for all $n \ge 1$.

Then we get

$$\begin{aligned}\mathbb{P}\left\{V_n^{(\max)} < n^{a_p-\delta}\right\} &\le \mathbb{P}\left\{V_n(\theta_0) < n^{a_p-\delta}\right\}\\ &= \mathbb{P}\left\{V_n(A_n^*\theta_n) < r_n^{-1}n^{a_p-\delta}\right\}\\ &\le \mathbb{P}\left\{V_n(A_n^*\theta_n) < Cn^{\varepsilon-\delta}\right\}.\end{aligned}$$

Since $\|\theta_n\| = 1$, every subsequence (n') contains a further subsequence $(n'') \subset (n')$ such that $\theta_n \to \theta$ along (n''). Then, by Proposition 2.6 we have $V_n(A_n^*\theta_n) \xrightarrow{d} Y^{|\cdot|}(\theta)$ along (n''). Since $\mathbb{P}\{Y^{|\cdot|}(\theta) < r\} \to 0$ as $r \to 0$, given $\varepsilon_1 > 0$ there exists $\rho > 0$ such that $\mathbb{P}\{Y^{|\cdot|}(\theta) < \rho\} < \varepsilon_1/2$. Now choose $n_0 \ge 1$ such that $Cn^{\varepsilon-\delta} < \rho$ and

$$\left|\mathbb{P}\left\{V_n(A_n^*\theta_n) < \rho\right\} - P\left\{Y^{|\cdot|}(\theta) < \rho\right\}\right| < \varepsilon_1/2$$

for all $n'' \ge n_0$. Then we get $\mathbb{P}\{V_n(A_n^*\theta_n) < Cn^{\varepsilon-\delta}\} < \varepsilon_1$ for all $n'' \ge n_0$. Hence

$$P\left\{V_n(A_n^*\theta_n) < Cn^{\varepsilon-\delta}\right\} \to 0 \quad \text{along } (n'')$$

and therefore $\mathbb{P}\{V_n^{(\max)} < n^{a_p-\delta}\} \to 0$ as $n \to \infty$.

Using Theorem 4.3.1 in Meerschaert and Scheffler (2001) again, for any $0 < \varepsilon < \delta$ there exists a $C > 0$ such that $\|(A_n^{-1})^*\theta\| \le C^{-1}n^{a_p+\varepsilon}$ for all $n \ge 1$ and all $\|\theta\| = 1$. Write $(A_n^{-1})^*\theta = r_n\theta_n$ as before. Then we get

$$\begin{aligned}\mathbb{P}\left\{V_n^{(\max)} > n^{a_p+\varepsilon}\right\} &= \mathbb{P}\left\{\max_{\|\theta\|=1} V_n(A_n^*\theta_n) > r_n^{-1}n^{a_p+\delta}\right\}\\ &\le \mathbb{P}\left\{\max_{\|\theta\|=1} V_n(A_n^*\theta) > Cn^{\delta-\varepsilon}\right\}.\end{aligned}$$

Since by Proposition 2.6 the sequence $(\max_{\|\theta\|=1} V_n(A_n^*\theta))_n$ is tight, Eq. 24 follows easily.

Similarly, for the proof of Eq. 25, note that for any $\delta > 0$ we have

$$\mathbb{P}\left\{\left|\frac{\log V_n^{(\min)}}{\log n} - a_1\right| > \delta\right\} \le \mathbb{P}\left\{V_n^{(\min)} > n^{a_1+\delta}\right\} + \mathbb{P}\left\{V_n^{(\min)} < n^{a_1-\delta}\right\}.$$

As before, choose $\theta_0 \in V_1$, $\|\theta_0\| = 1$. Then, if we write $(A_n^{-1})^*\theta_0 = r_n\theta_n$, for any $0 < \varepsilon < \delta$ there exists $C > 0$ such that $r_n \le C^{-1}n^{a_1+\varepsilon}$ for all $n \ge 1$. Now

$$\begin{aligned}\mathbb{P}\big\{V_n^{(\min)} > n^{a_1+\delta}\big\} &\le \mathbb{P}\big\{V_n(\theta_0) > n^{a_1+\delta}\big\}\\ &= \mathbb{P}\big\{V_n(A_n^*\theta_n) > r_n^{-1}n^{a_1+\delta}\big\}\\ &\le \mathbb{P}\Big\{\max_{\|\theta\|=1} V_n(A_n^*\theta) > Cn^{\delta-\varepsilon}\Big\}\end{aligned}$$

and use tightness again to see that $\mathbb{P}\big\{V_n^{(\min)} > n^{a_1+\delta}\big\} \to 0$ as $n \to \infty$.

Finally, if we write $(A_n^{-1})^*\theta = r_n\theta_n$ again, given $0 < \varepsilon < \delta$, there exists a $C > 0$ such that $r_n = \|(A_n^{-1})^*\theta\| \ge C^{-1}n^{a_1-\varepsilon}$ for all $\|\theta\| = 1$ and all $n \ge 1$. Then we have

$$\begin{aligned}\mathbb{P}\big\{V_n^{(\min)} < n^{a_1-\delta}\big\} &= P\Big\{\min_{\|\theta\|=1} V_n(A_n^*\theta_n) < r_n^{-1}n^{a_1-\delta}\Big\}\\ &\le \mathbb{P}\Big\{\min_{\|\theta\|=1} V_n(A_n^*\theta) < Cn^{\varepsilon-\delta}\Big\}.\end{aligned}$$

Now, in view of Proposition 2.6 we have $\min_{\|\theta\|=1} V_n(A_n^*\theta) \xrightarrow{d} Y^{(\min)}$ where $Y^{(\min)} > 0$ almost surely. This easily implies $\mathbb{P}\big\{V_n^{(\min)} < n^{a_1-\delta}\big\} \to 0$ and the proof is complete. □

3 The extremal limit process

In this section, we investigate properties of the extremal limit process $\mathcal{Y} := \{Y(\theta)\}$ in Theorem 2.1. The first result shows that the random field $\mathcal{Y}$ is operator max-scaling.

Proposition 3.1 *Let $\mathcal{Y}_k = \{Y_k(\theta)\}_{\theta\in\mathbb{R}^d\setminus\{\mathbf{0}\}}$ for $k = 1, \ldots, n$ be independent copies of the process $\mathcal{Y}$ in Eq. 9. Then*

$$\{Y_1(\theta) \vee \cdots \vee Y_n(\theta)\}_{\theta\in\mathbb{R}^d\setminus\{\mathbf{0}\}} \stackrel{d}{=} \{Y(n^{E^*}\theta)\}_{\theta\in\mathbb{R}^d\setminus\{\mathbf{0}\}} \tag{26}$$

for all $n \in \mathbb{N}$.

Proof Since $n^{-E}B(r,\theta) = \{n^{-E}x : \langle x,\theta\rangle > r\} = \{y : \langle y, n^{E*}\theta\rangle > r\} = B(r, n^{E^*}\theta)$, we may write

$$\begin{aligned}\mathbb{P}\left\{\bigvee_{k=1}^{n} Y(\theta_j) \le r_j \ \forall\, 1 \le j \le m\right\} &= \exp\left\{-n\phi\left(\bigcup_{j=1}^{m} B(r_j,\theta_j)\right)\right\}\\ &= \exp\left\{-\phi\left(n^{-E}\bigcup_{j=1}^{m} B(r_j,\theta_j)\right)\right\}\\ &= \exp\left\{-\phi\left(\bigcup_{j=1}^{m} B(r_j, n^{E^*}\theta_j)\right)\right\}\\ &= \mathbb{P}\big\{Y\big(n^{E^*}\theta_j\big) \le r_j \ \forall\, 1 \le j \le m\big\}\end{aligned}$$

using Eq. 9, the operator scaling property (2) of ϕ, and the independence of the $\mathcal{Y}_k$'s. □

Remark 3.2 The operator scaling Eq. 26 implies that $Y(\theta_i)$ is Fréchet max-stable with index $\alpha_i = 1/a_i$ for any eigenvector $E\theta_i = a_i\theta_i$. To see this, first note that $r^{E^*}\theta_i = r^{a_i}\theta_i$ for any $r > 0$. Then observe that $B(r, c\theta) = B(r/c, \theta)$ in Eq. 8, and so it follows from Eq. 9 that $Y(c\theta) = cY(\theta)$ almost surely for all $c > 0$ and all $\theta \in \mathbb{R}^d \setminus \{\mathbf{0}\}$. Hence Eq. 26 yields

$$Y_1(\theta_i) \vee \cdots \vee Y_n(\theta_i) \stackrel{d}{=} n^{1/\alpha_i} Y(\theta_i) \quad \text{for all } n = 1, 2, 3, \ldots$$

so that $Y(\theta_i)$ is max-stable. In fact, Eqs. 8 and 9 imply that $P(Y(\theta_i) \leq r) = \exp(-Cr^{-\alpha_i})$ for any $r > 0$, where $C = \phi(B(1, \theta_i))$. We could formulate our extreme value theory for θ on the unit sphere, but our approach reveals the operator scaling property in Proposition 3.1.

As in the case of scalar norming, the Poisson representation (20) is particularly useful for computer simulations. To that end, consider the truncated maximum

$$Y^{(n)}(\theta) := \bigvee_{i=1}^{n} \langle \Gamma_i^{-E}\Lambda_i, \theta\rangle \quad \text{and let also} \quad Y^{(\infty)}(\theta)$$
$$:= \lim_{n\to\infty} Y^{(n)}(\theta) \equiv \bigvee_{i=1}^{\infty} \langle \Gamma_i^{-E}\Lambda_i, \theta\rangle.$$

Proposition 3.3 *For all $K \subset S^{d-1} := \{\theta \in \mathbb{R}^d : \|\theta\| = 1\}$, $\epsilon > 0$ and $\delta \in (0, 1)$, there exists $n_{\epsilon,\delta}$ such that for all $n \geq n_{\epsilon,\delta}$*

$$\mathbb{P}\{Y^{(\infty)}(\theta) \neq Y^{(n)}(\theta),\ \textit{for some } \theta \in K\} \leq \delta^n + \mathbb{P}\Big\{\inf_{\theta\in K} Y^{(\infty)}(\theta) \leq \epsilon\Big\}. \qquad (27)$$

Proof Let $\epsilon > 0$ and observe that if $Y^{(\infty)}(\theta) > \epsilon$ and $\langle \Gamma_i^{-E}\Lambda_i, \theta\rangle \leq \epsilon$, $\forall\, i \geq n+1$, then $Y^{(\infty)}(\theta) = Y^{(n)}(\theta) > \epsilon$. Therefore,

$$\mathbb{P}\{Y^{(\infty)}(\theta) = Y^{(n)}(\theta),\ \forall\theta \in K\} \geq \mathbb{P}\Big\{\sup_{i\geq n+1} \|\Gamma_i^{-E}\Lambda_i\| \leq \epsilon,\ \inf_{\theta\in K} Y^{(\infty)}(\theta) > \epsilon\Big\},$$

and hence with $c_E = (\sup_{\theta\in\mathbb{S}_E}\|\theta\|)^{-1}$, we have

$$\mathbb{P}\{Y^{(\infty)}(\theta) \neq Y^{(n)}(\theta),\ \exists\,\theta \in K\} \leq \mathbb{P}\Big\{\sup_{i\geq n+1} \|\Gamma_i^{-E}\| > c_E\epsilon\Big\}$$
$$+\,\mathbb{P}\Big\{\inf_{\theta\in K} Y^{(\infty)}(\theta) \leq \epsilon\Big\}.$$

By Theorem 2.2.4 in Meerschaert and Scheffler (2001), we have that for all $t \geq 1$, $\|t^{-E}\| \leq C_a t^{-a}$, where $\|\cdot\|$ stands for the operator norm induced by the Euclidean norm in $\mathbb{R}^d$ and where $0 < a < \min_{\lambda\in\mathrm{spec}(E)} \mathcal{R}e(\lambda)$ is strictly less than the smallest real part of the eigenvalues of E. Since $a > 0$ and $\Gamma_{n+1} \le \Gamma_i$, for all $i \ge n+1$, we obtain

$$\mathbb{P}\left\{\bigvee_{i=n+1}^{\infty} \|\Gamma_i^{-E}\Lambda_i\| > \epsilon\right\} \le \mathbb{P}\{\Gamma_{n+1}^{-a} > c_{E,a}\epsilon\}$$
$$= \mathbb{P}\{\Gamma_{n+1}/(n+1) < (c_{E,a}\epsilon)^{-1/a}/(n+1)\},$$

where $c_{E,a} := (C_a \sup_{\theta\in\mathbb{S}_E}\|\theta\|)^{-1}$. Recall, however, that $\Gamma_{n+1} = E_1 + \cdots + E_{n+1}$, where the E_i's are iid Exponential with mean $1/\lambda(\mathbb{S}_E)$. Thus, by the Cramér's large deviations theorem (see e.g. Theorem 2.2.3, Remark (c), p. 27, and Exercise 2.2.23 in Dembo and Zeitouni 1998), we obtain

$$\mathbb{P}\left\{\frac{\Gamma_{n+1}}{(n+1)} < \frac{(c_{E,a}\epsilon)^{-1/a}}{(n+1)}\right\}$$
$$\le 2\exp\left\{\frac{\lambda(\mathbb{S}_E)}{(c_{E,a}\epsilon)^{1/a}} + (n+1)(1-\log n + \log(\lambda(\mathbb{S}_E)(c_{E,a}\epsilon)^{-1/a}))\right\}.$$

Since $\log(1/n) \to -\infty$, as $n \to \infty$, the last inequality implies Eq. 27. □

Remark 3.4 If supp(ϕ) contains a neighborhood of the origin, take $K = S^{d-1}$ in Eq. 27. Since ϕ is full and assigns finite mass to sets not containing a neighborhood of the origin, the convex hull of the point process $\mathcal{N}$ is almost surely a polyhedron with finite number of extremal points. The process $Y^{(\infty)}(\cdot)$ depends only on this convex hull, hence $Y^{(n)}(\cdot)$ equals $Y^{(\infty)}(\cdot)$ for n sufficiently large. Then the finite maximum $Y^{(n)}(\cdot)$ can be used to simulate the exact values of arbitrary statistics of $Y(\cdot)$ with high probability. For vector data where every component is positive, it is natural for ϕ to be concentrated on $[0,\infty)^d$, and then one should choose $K = [0,\infty)^d \cap S^{d-1}$.

4 Scalar norming

If $A_n = c_n I$, a scalar multiple of the identity, then $M_n(A_n^*\theta) = c_n M_n(\theta)$, $E = (1/\alpha)I$ for some $\alpha > 0$, condition (1) is scalar-normed multivariate regular variation, and $Y(\theta)$ has an α-Fréchet distribution. In this case, we can use the Euclidean norm $\|\cdot\|$ and the corresponding sphere S^{d-1} in the disintegration formula (5). Then we can develop a representation theorem for the limit process $\{Y(\theta)\}_{\theta\in\mathbb{R}^d\setminus\{\mathbf{0}\}}$ in Theorem 2.1 using extremal integrals.

Following (de Haan 1984; Kabluchko 2009; Stoev and Taqqu 2005), any separable-in-probability α-Fréchet max-stable process $\xi = \{\xi(t)\}_{t\in T}$ can be represented as

$$\{\xi(t)\}_{t\in T} \stackrel{d}{=} \left\{{}^{e}\!\!\int_D f_t(u) M_\alpha(\mathrm{d}u)\right\}_{t\in T} \tag{28}$$

where $f_t \in L_+^\alpha(D,\lambda)$ are non-negative deterministic functions defined on a Borel measure space $(D, \mathcal{B}(D), \lambda)$ with $\int_D f_t^\alpha \mathrm{d}\lambda < \infty$, and ${}^{e}\!\!\int_D f \mathrm{dM}_\alpha$ is an *extremal* stochastic integral with respect to the random α-Fréchet sup-measure M_α on $(D, \mathcal{B}(D))$ with control measure λ, so that M_α is σ-sup-additive rather than additive, and it assigns independent α-Fréchet variables to disjoint measurable sets, with scale coefficients controlled by the deterministic measure λ:

$$\mathbb{P}\{M_\alpha(B) \le x\} = \exp\{-\lambda(B)x^{-\alpha}\}, \quad x > 0.$$

The *spectral functions* f_t in Eq. 28 yield the finite-dimensional distributions of the process:

$$\mathbb{P}\{\xi(t_j) \le x_j,\ \forall j = 1, \ldots, m\} = \exp\left\{-\int_D \left(\max_{1\le j\le m} \frac{f_{t_j}^\alpha(u)}{x_j^\alpha}\right)\lambda(\mathrm{d}u)\right\},\ (x_j > 0) \tag{29}$$

(see also Proposition 5.11 in Resnick 1987). The extremal integral representation (28) is called the *spectral representation* of ξ. The next result provides the spectral representation of the directional process $\mathcal{Y}$.

Proposition 4.1 *Suppose that Eq. 1 holds with scalar $A_n = c_n I$. Then c_n is regularly varying with exponent $-1/\alpha$ for some $\alpha > 0$, and the directional process in Eq. 10 is α-Fréchet with the following extremal integral representation:*

$$\{Y(\theta)\}_{\theta\in\mathbb{R}^d\setminus\{\mathbf{0}\}} \stackrel{d}{=} \left\{ {}^{e}\!\!\int_{S^{d-1}} \langle u, \theta\rangle_+ M_\alpha(\mathrm{d}u)\right\}_{\theta\in\mathbb{R}^d\setminus\{\mathbf{0}\}} \tag{30}$$

where M_α is an α-Fréchet sup-measure on $(S^{d-1}, \mathcal{B}(S^{d-1}))$ with control measure λ, the spectral measure from Eq. 6.

Proof If Eq. 1 holds with scalar $A_n = c_n I$, then c_n is regularly varying with exponent $-1/\alpha$ for some $\alpha > 0$ by Meerschaert and Scheffler (2001, Proposition 6.1.37). Let $\theta_j \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ and $r_j > 0,\ j = 1, \ldots, m$ be arbitrary. In view of Eqs. 9 and 5, we have

$$\begin{aligned}
&\mathbb{P}\big\{Y(\theta_j) \le r_j,\ \forall j = 1, \ldots, m\big\}\\
&\quad= \exp\left\{-\int_{S^{d-1}}\int_0^\infty \mathbf{1}_{\cup_{j=1}^m B(r_j,\theta_j)}(t^{1/\alpha}\theta)\frac{\mathrm{d}t}{t^2}\lambda(\mathrm{d}\theta)\right\}\\
&\quad= \exp\left\{-\int_{S^{d-1}}\int_0^\infty \Big(\max_{1\le j\le m}\mathbf{1}_{B(r_j,\theta_j)}(t^{1/\alpha}\theta)\Big)\frac{\mathrm{d}t}{t^2}\lambda(\mathrm{d}\theta)\right\},
\end{aligned} \tag{31}$$

where we used the facts that $t^E = t^{1/\alpha} I$ and $\mathbf{1}_{\cup_{j=1}^m B(r_j,\theta_j)}(\cdot) = \max_{1\le j\le m}\mathbf{1}_{B(r_j,\theta_j)}(\cdot)$. By focusing on the inner integral in the right–hand side of Eq. 31 and making the change of variables $\tau := t^{-1}$, we obtain

$$\int_0^\infty \Big(\max_{1\le j\le m}\mathbf{1}_{B(r_j,\theta_j)}(t^{1/\alpha}\theta)\Big)\frac{\mathrm{d}t}{t^2} = \int_0^\infty \Big(\max_{1\le j\le m}\mathbf{1}_{B(r_j,\theta_j)}(\tau^{-1/\alpha}\theta)\Big)\mathrm{d}\tau.$$

Observe that the integrand equals 0 or 1, and it is equal to 1 if and only if $\tau^{-1/\alpha}\langle\theta, \theta_j\rangle > r_j$, for some $j = 1, \ldots, m$, or equivalently, if $\tau < \max_{1\le j\le m}\langle\theta, \theta_j\rangle_+^\alpha / r_j^\alpha$ (with $\max\emptyset = 0$, by convention.) Thus,

$\int_0^\infty (\max_{1\le j\le m} \mathbf{1}_{B(r_j,\theta_j)}(\tau^{-1}\theta))\mathrm{d}\tau = \max_{1\le j\le m} \langle\theta,\theta_j\rangle_+^\alpha / r_j^\alpha$, which implies that the right–hand side of Eq. 31 equals

$$\exp\left\{-\int_{S^{d-1}}\left(\max_{1\le j\le m}\frac{\langle\theta,\theta_j\rangle_+^\alpha}{r_j^\alpha}\right)\lambda(\mathrm{d}\theta)\right\}.$$

In view of Eq. 29 the last expression is precisely that of the finite–dimensional distributions of a max–stable process with spectral representation as in Eq. 30. □

5 Testing for hetero-ouracity

A natural and important question is whether the tail index of a given data set varies with direction, so that operator normalization should be used. We say that such distributions possess *hetero-ouracity* (from the greek words for different tails). We address this question here with a formal hypothesis test. As in Proposition 2.7, we consider the distinct real parts $0 < 1/\alpha_1 \le \cdots \le 1/\alpha_p$ of the eigenvalues of the scaling matrix E from Eq. 2, so that $\alpha_1 > \cdots > \alpha_p > 0$ are the distinct tail indices Eq. 4 of the data. We assume iid data $X_1, X_2, \ldots$ whose underlying distribution μ is operator regularly varying with index E, so that Eq. 1 holds. Our goal is to test whether the data has different tail indices in different directions. Under the null hypothesis $\mathcal{H}_0$: *scalar norming*, the norming operators in Eq. 1 are of the form $A_n = c_n I$, where c_n is a regularly varying sequence with index $-1/\alpha$, $\alpha = \alpha_1 = \alpha_p$ (same tail index in every direction). By Proposition 2.6, we then have

$$\left\{c_n V_n(\theta)\right\}_{\theta\in\mathbb{R}^d\setminus\{\mathbf{0}\}} \stackrel{d}{\longrightarrow} \{Y^{|\cdot|}(\theta)\}_{\theta\in\mathbb{R}^d\setminus\{\mathbf{0}\}}, \quad \text{in } C(\mathbb{R}^d\setminus\{\mathbf{0}\}), \quad \text{as } n\to\infty. \tag{32}$$

Moreover, Eq. 22 in this particular case reads

$$c_n(V_n^{(\max)}, V_n^{(\min)}) \stackrel{d}{\longrightarrow} (Y^{(\max)}, Y^{(\min)}), \tag{33}$$

where $V_n^{(\max)}$ and $V_n^{(\min)}$ are as in Eq. 23. Since $Y^{(\min)} > 0$ almost surely by Proposition 2.6, applying continuous mapping again to Eq. 33 yields

$$\frac{V_n^{(\max)}}{V_n^{(\min)}} \stackrel{d}{\longrightarrow} \frac{Y^{(\max)}}{Y^{(\min)}}, \quad \text{as } n\to\infty. \tag{34}$$

For this convergence to hold, it is essential that $\alpha_1 = \alpha_p$, so that $V_n^{(\max)}$ and $V_n^{(\min)}$ are of the same order. In fact, it follows immediately from Proposition 2.7 that

$$V_n^{(\max)}/V_n^{(\min)} \stackrel{\mathbb{P}}{\to} \infty \quad \text{as } n\to\infty$$

in the case where $\alpha_p < \alpha_1$.

Relation (34) (valid under the null hypothesis) and Proposition 2.7 will help us design a test for scalar norming, with asymptotic power equal to 1 when the tail index varies. To obtain asymptotically accurate rejection regions, however, we need to first construct consistent estimates for the quantiles of $Y^{(\max)}/Y^{(\min)}$ under the null hypothesis. The following two results are required for this purpose.

Proposition 5.1 *Let λ_n and λ be finite measures on S^{d-1} such that $\lambda_n \stackrel{w}{\to} \lambda$, and that $\alpha_n \to \alpha > 0$ as $n \to \infty$. Then*

$$\{Y_n(\theta)\}_{\theta \in S^{d-1}} \stackrel{d}{\longrightarrow} \{Y(\theta)\}_{\theta \in S^{d-1}}, \quad \text{as } n \to \infty \text{ in } (C(S^{d-1}, \mathbb{R}), \|\cdot\|) \tag{35}$$

where $Y_n(\theta) := {}^e\!\!\int_{S^{d-1}} \langle u, \theta\rangle_+ M_{\alpha_n}^{(n)}(\mathrm{du})$ is a continuous version of the directional process whose α_n-Fréchet random sup–measure $M_{\alpha_n}^{(n)}$ has control measures λ_n, and similarly $Y(\theta) := {}^e\!\!\int_{S^{d-1}} \langle u, \theta\rangle_+ M_\alpha(\mathrm{du})$ has an α-Fréchet random sup–measure M_α with control measure λ.

Proof For all $r_j > 0$ and $\theta_j \in S^{d-1}$, $j = 1, \ldots, m$ by Eq. 29, we have

$$\mathbb{P}\{Y_n(\theta_j) \le r_j,\ j = 1, \ldots, m\} = \exp\left\{-\int_{S^{d-1}} g_{\theta,r}^{\alpha_n}(u)\lambda_n(\mathrm{du})\right\} \tag{36}$$

and similarly

$$\mathbb{P}\{Y(\theta_j) \le r_j,\ j = 1, \ldots, m\} = \exp\left\{-\int_{S^{d-1}} g_{\theta,r}^{\alpha}(u)\lambda(\mathrm{du})\right\}, \tag{37}$$

where $g_{\theta,r}(u) = \max_{1 \le j \le n} \langle u, \theta_j\rangle_+ / r_j$. Observe that $u \mapsto g_{\theta,r}^{\alpha_n}(u)$ are bounded and continuous functions, such that $\sup_{\|u\|=1} |g_{\theta,r}^{\alpha_n}(u) - g_{\theta,r}^{\alpha}(u)| \to 0$, as $n \to \infty$. Therefore, $\lambda_n \stackrel{w}{\to} \lambda$ implies that the probabilities in Eq. 36 converge to those in Eq. 37, as $n \to \infty$. This shows that Eq. 35 is valid in the sense of convergence of the finite-dimensional distributions.

The tightness of the laws of $\mathcal{Y}_n = \{Y_n(\theta)\}_{\theta \in S^{d-1}}$, $n \in \mathbb{N}$ follows as in the proof of Lemma 2.4. Indeed, by applying Eq. 17 to relation (20) above, we obtain

$$\begin{aligned} |Y_n(\theta) - Y_n(\theta')| &\le \bigvee_{i=1}^{\infty} |\langle \theta - \theta', \Gamma_i^{-1/\alpha_n} \Lambda_i^{(n)}\rangle| \\ &\le \|\theta - \theta'\| \bigvee_{i=1}^{\infty} \Gamma_i^{-1/\alpha_n} = \|\theta - \theta'\| \Gamma_1^{-1/\alpha_n}, \end{aligned} \tag{38}$$

where $0 < \Gamma_1 < \Gamma_2 < \cdots$ is a Poisson point process with constant intensity $\lambda_n(S^{d-1})$ on $(0, \infty)$ and the $\Lambda_i^{(n)}$'s are iid random variables on S^{d-1} with distribution $\lambda_n(\cdot)/\lambda_n(S^{d-1})$. Since $\alpha_n \to \alpha > 0$, as $n \to \infty$, relation (38) readily implies Eq. 16 for the modulus of continuity of $\mathcal{Y}_n$ and thus the desired tightness. □

The next result involves the notion of weak convergence of random measures, which we briefly recall. Following Section 3.3.5 in Resnick (2007), let $M_+(S^{d-1})$ be the set of finite Borel measures on S^{d-1}. $M_+(S^{d-1})$ can be equipped with a metric ρ_{M_+} which metrizes the vague (equivalently, weak) convergence of measures. In fact, ρ_{M_+} can be chosen in such a way that $(M_+(S^{d-1}), \rho_{M_+})$ becomes a complete separable metric space. By a random measure on S^{d-1} we will understand a Borel measurable mapping $\xi : \Omega \to M_+(S^{d-1})$, i.e. a random element taking values in the metric space $M_+(S^{d-1})$. For a sequence of random measures ξ_n, we write $\xi_n \Rightarrow \xi$ if the law of ξ_n converges to that of ξ in $M_+(S^{d-1})$ as $n \to \infty$.

Proposition 5.2 *Let $X_i,\ i=1,\dots,n$ be independent random vectors with distribution μ, that satisfies Eq. 1 with $A_n = c_n I$ and spectral measure λ as in Eq. 6. For all $u>0$, consider the random measure*

$$\widehat{\lambda}_u(A) := \frac{\sum_{i=1}^n \mathbf{1}_A(X_i/\|X_i\|)\mathbf{1}_{(u,\infty)}(\|X_i\|)}{\sum_{i=1}^n \mathbf{1}_{(u,\infty)}(\|X_i\|)} \tag{39}$$

using the convention $\widehat{\lambda}_u \equiv 0$ if $\|X_i\| \le u$ for all $i=1,\dots,n$. If $u_n \to \infty$ and $c_n u_n \to 0$, then

$$\widehat{\lambda}_{u_n} \Longrightarrow \lambda(\cdot)/\lambda(S^{d-1}), \qquad \text{as } n\to\infty \text{ in } M_+(S^{d-1}). \tag{40}$$

Proof The result readily follows from a slight modification of Theorem 6.2, part (9) in Resnick (2007). Since we are assuming scalar-normed multivariable regular variation, with spectral measure λ as in Eq. 6, the convergence (6.18) on p. 180 of Resnick (2007) holds in $M_+((0,\infty]\times S^{d-1})$. Namely, with $R_i := \|X_i\|$ and $\Theta_i := X_i/\|X_i\|$, we have:

$$\xi_n := \frac{1}{k}\sum_{i=1}^n \epsilon_{(R_i/b(n/k),\Theta_i)} \Longrightarrow \xi := c\nu_\alpha \times \lambda, \tag{41}$$

where $\nu_\alpha([x,\infty)) = x^{-\alpha},\ x>0$, λ is the spectral measure of μ on S^{d-1}, and $k = k(n) \to \infty$ in such a way that $n/k(n)\to\infty$. Write $b(n) := c_n^{-1} = \ell(n)n^{1/\alpha}$ where $\ell(n)$ is slowly varying. Since $c_n u_n \to 0$ and $u_n\to\infty$, one can choose $k=k(n)$ such that $u_n \sim b(n/k) = \ell(n/k)(n/k)^{1/\alpha}$ as $n\to\infty$.

Then, for all measurable sets $A\subset S^{d-1}$,

$$\widehat{\lambda}_{u_n}(A) = \frac{\xi_n((1,\infty)\times A)}{\xi_n((1,\infty)\times S^{d-1})}. \tag{42}$$

Observe that since ν_α in Eq. 41 has no atoms, for all $A\subset S^{d-1}$ that are continuity sets of λ, the set $(1,\infty)\times A$ is also a continuity set of the limit measure in Eq. 41. Therefore, for all measurable continuity sets A of λ, relation (41) implies that, as $n\to\infty$,

$$(\xi_n((1,\infty)\times A), \xi_n((1,\infty)\times S^{d-1})) \xrightarrow{\mathbb{P}} c(\lambda(A),\lambda(S^{d-1})), \tag{43}$$

where the last convergence is in probability since the limit in Eq. 41 is deterministic. In view of Eq. 42, Eq. 43 implies that $\widehat{\lambda}_{u_n}(A) \overset{\mathbb{P}}{\to} \lambda(A)/\lambda(S^{d-1})$ as $n\to\infty$, for all $\lambda-$continuity sets A. This yields the desired weak convergence (40). □

We are now ready to propose a test for hetero-ouracity. Let $\mathcal{H}_0$: *scalar norming* (same tail index in all directions). Under $\mathcal{H}_0$, Eq. 34 holds. On the other hand, when the tail indices vary, Proposition 2.7 implies that the ratio in Eq. 34 tends to infinity. We therefore reject $\mathcal{H}_0$ at a level $\beta\in(0,1)$ if

$$\frac{V_n^{(\max)}}{V_n^{(\min)}} > c_{1-\beta}(\lambda,\alpha), \tag{44}$$

where $c_{1-\beta}(\lambda, \alpha)$ is the $(1-\beta)$ quantile of the distribution of $Y^{(\max)}/Y^{(\min)}$, that is,

$$c_{1-\beta}(\lambda, \alpha) = F^{\leftarrow}_{\frac{Y^{(\max)}}{Y^{(\min)}}}(1-\beta) := \inf\{c > 0 \, : \, \mathbb{P}\{Y^{(\max)}/Y^{(\min)} \le c\} \ge 1-\beta\}.$$

Given a real data set, one must estimate the tail exponent α and the spectral measure λ. Proposition 5.2 provides a consistent estimate of the spectral measure λ in Eq. 6 under the null hypothesis $\mathcal{H}_0$. For the tail index, one can apply Proposition 2.7 to get a consistent estimate of the tail index α, on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$. If desired, any other consistent tail estimator can be applied (e.g., the simple Hill estimator). To justify the use of these estimators in the proposed test, we develop the following parametric bootstrap procedure. Consider the parametric bootstrap version of the directional process $\mathcal{Y}_n^* = \{Y_n^*(\theta)\}_{\theta \in S^{d-1}}$ defined by

$$Y_n^*(\theta) := {}^{e}\!\!\int_{S^{d-1}} \langle u, \theta \rangle_+ M^*_{\widehat{\alpha}_n}(\mathrm{d}u), \tag{45}$$

where $M^*_{\widehat{\alpha}_n}$ is an $\widehat{\alpha}_n$-Fréchet random sup-measure with control measure $\widehat{\lambda}_n$. The sup-measure $M^*_{\widehat{\alpha}_n}$ and the process $\mathcal{Y}_n^*$ are defined on a *different* probability space $(\Omega^*, \mathcal{F}^*, \mathbb{P}^*)$. Without loss of generality, assume that $\mathcal{Y}_n^*$ has continuous paths, and introduce the quantities

$$Y_n^{*(\max)} := \sup_{\|\theta\|=1} Y_n^*(\theta) \quad \text{and} \quad Y_n^{*(\min)} := \inf_{\|\theta\|=1} (Y_n^*(\theta) \vee Y_n^*(-\theta)). \tag{46}$$

The next result justifies our bootstrap procedure.

Proposition 5.3 *Under the above assumptions,*

$$\mathcal{L}_{\mathbb{P}^*}\left(\frac{Y_n^{*(\max)}}{Y_n^{*(\min)}}\right) \overset{\mathbb{P}}{\Longrightarrow} \mathcal{L}_{\mathbb{P}^*}\left(\frac{Y^{*(\max)}}{Y^{*(\min)}}\right) \tag{47}$$

as $n \to \infty$, where $Y^{*(\max)}$ and $Y^{*(\min)}$ are as in Eq. 46 but with $M^*_{\widehat{\alpha}_n}$ in Eq. 45 replaced by an α-Fréchet sup-measure M^*_α with control measure λ. The convergence in probability in Eq. 47 is viewed in the space $(\Omega, \mathcal{F}, \mathbb{P})$, while '$\Rightarrow$' therein denotes weak convergence of probability distributions in the space $(\Omega^*, \mathcal{F}^*, \mathbb{P}^*)$.

Proof To establish convergence in probability (47), it suffices to show that for any integer sequence $n_k \to \infty$, there exists a further sub-sequence $n' \to \infty$ such that

$$\frac{Y_{n'}^{*(\max)}(\cdot, \omega)}{Y_{n'}^{*(\min)}(\cdot, \omega)} \Longrightarrow \frac{Y^{*(\max)}(\cdot)}{Y^{*(\min)}(\cdot)}, \quad \text{for } \mathbb{P}\text{–almost all } \omega \in \Omega. \tag{48}$$

Note that $Y^{*(\max)}/Y^{*(\min)}$ is supported on the probability space $(\Omega^*, \mathcal{F}^*, \mathbb{P}^*)$ and it does not depend on ω. Now since

$$\widehat{\alpha}_n \overset{\mathbb{P}}{\to} \alpha \quad \text{and} \quad \widehat{\lambda}_n \overset{\mathbb{P}}{\Rightarrow} \lambda(\cdot)/\lambda(S^{d-1})$$

as $n \to \infty$, for every $n_k \to \infty$, there exists a further sub-sequence $n' \to \infty$, such that

$$\widehat{\alpha}_{n'}(\omega) \longrightarrow \alpha \quad \text{and} \quad \widehat{\lambda}_{n'}(\omega) \longrightarrow \lambda(\cdot)/\lambda(S^{d-1}), \quad \text{for } \mathbb{P}\text{–almost all } \omega \in \Omega. \tag{49}$$

Fix such an $\omega \in \Omega$, and view $Y_{n'}^{(\max)}(\cdot, \omega)/Y_{n'}^{(\min)}(\cdot, \omega)$ as random variables on the probability space $(\Omega^*, \mathcal{F}^*, \mathbb{P}^*)$. Observe that

$$\frac{Y_{n'}^{*(\max)}(\cdot, \omega)}{Y_{n'}^{*(\min)}(\cdot, \omega)} = \Psi(Y_n^*(\cdot, \omega)),$$

where $\Psi(f) := (\sup_{\|\theta\|=1} f(\theta))/(\inf_{\|\theta\|=1} f(\theta) \vee f(-\theta))$. The functional $\Psi : C(S^{d-1}) \to \mathbb{R}$ is continuous on the support of the law of $\mathcal{Y}^* = \{Y^*(\theta)\}_{\theta \in S^{d-1}}$ since $Y^{*(\min)} > 0$. Therefore, Proposition 5.1 and the continuous mapping theorem applied to the functional Ψ show that Eq. 48 holds. This is true for $\mathbb{P}$-almost all $\omega \in \Omega$. Therefore, since the sequence $n_k \to \infty$ was arbitrary, the convergence (48) implies Eq. 47. □

Corollary 5.4 *Suppose that $c_{1-\beta}(\alpha, \lambda)$ is a continuity point of the quantile function $F^{\leftarrow}_{Y^{(\max)}/Y^{(\min)}}$ of $Y^{(\max)}/Y^{(\min)}$. Then, $c_{1-\beta}(\widehat{\alpha}_n, \widehat{\lambda}_n) \xrightarrow{\mathbb{P}} c_{1-\beta}(\alpha, \lambda)$, as $n \to \infty$.*

Proof As in the proof of Proposition 5.3, we apply the method of selecting a $\mathbb{P}$-almost surely converging sub-sequence. The result then follows from the continuous mapping theorem applied to Eq. 47 with the help of Proposition 0.1 on page 5 in Resnick (1987) (see also Proposition 2.2, page 20 in Resnick 2007). □

To conclude this section, we summarize the steps of our test:

(1) Compute consistent statistics $\widehat{\alpha}_n$ and $\widehat{\lambda}_n$ from the sample $X_i,\ i = 1, \ldots, n$.
(2) Apply the simulation methods from Section 3 to obtain N independent bootstrap samples from $Y_n^{*(\max)}/Y_n^{*(\min)}$ using (1).
(3) Approximate $c_{1-\beta}(\widehat{\alpha}_n, \widehat{\lambda}_n)$ with its empirical quantile $\widehat{c}_{1-\beta}(\widehat{\alpha}_n, \widehat{\lambda}_n)$ based on (2).
(4) Compute the statistic $V_n^{(\max)}/V_n^{(\min)}$ from the sample $X_i,\ i = 1, \ldots, n$ and test the hypothesis $\mathcal{H}_0$ according to the rule Eq. 44 with $c_{1-\beta}(\alpha, \lambda)$ replaced by $\widehat{c}_{1-\beta}(\widehat{\alpha}_n, \widehat{\lambda}_n)$.

Corollary 5.4 ensures that this procedure yields asymptotically accurate estimates of $c_{1-\beta}(\alpha, \lambda)$. Equation 34 shows that under the null hypothesis $\mathcal{H}_0$: *scalar norming*, the test is asymptotically consistent, provided $c_{1-\beta}(\alpha, \lambda)$ is a continuity point of the quantile function of the test statistic. Furthermore, Proposition 2.7 implies that under the alternative of hetero-ouracity (variation in tail index), the test has an asymptotic power of 1.

Remark 5.5 Scalar norming implies but is not equivalent to having the same tail exponent in all directions. If the matrix exponent E has non-trivial nilpotent part, or multiple eigenvalues with equal real parts but different imaginary parts, then the same tail index can pertain in every direction, but with operator norming, the tail behavior can vary between coordinates. Such subtle scenarios are hard to detect in practice since one tail can fall off like $x^{-\alpha}$ and the other can fall off like $x^{-\alpha} \log x$.

6 Simulation

This section reports the results of a small simulation study to validate the testing procedure outlined in Section 5. We simulate data from an operator Pareto distribution in $\mathbb{R}^3$ with uniform spectral density on the Euclidean unit sphere. Namely, the distribution μ is the law of the random vector

$$X = U^{-E}\Theta,$$

where $U \sim \text{Uniform}(0, 1)$ and $\Theta \sim \text{Uniform}(S^{d-1})$ are independent. The exponent matrix E is diagonal: $E := \text{diag}(1/\alpha_1, 1/\alpha_2, 1/\alpha_3)$ with $0 < \alpha_3 \le \alpha_2 \le \alpha_1$. We simulate independent samples from this distribution and test for variation in the tail index as follows. The corresponding MATLAB code can be downloaded from Stoev et al. (2012).

(i) Under the null hypothesis (scalar norming), the spectral measure estimate $\widehat{\lambda}_{u_n}$ from Proposition 5.2 depends on the number of sample points with $\|X_i\| > u_n$. For simplicity, we choose $u_n \equiv 1$, so that all the simulated data is used to estimate the spectral measure.
(ii) The tail estimate $\widehat{\alpha}_n$ is computed using the plain Hill estimator for the sample $\|X_i\|,\ i = 1, \ldots, n$, which under the null is heavy tailed with exponent α. The threshold for the Hill estimator is $k := [n/2]$, where n is the sample size, so that the upper half of the data is used to estimate the tail index.
(iii) The spectral measure λ is estimated using Eq. 39, with $u_n = 1$.

Remark 6.1 In step (i) we have chosen to use all the simulated data. This is justifiable because of the operator Pareto model, which assumes a power law distribution for the entire data set. For real data analysis, it is often advisable to consider the largest order statistics of the data, where the power law behavior should be evident (assuming that

Table 1 Empirical rejection probabilities for our test (at level $\beta = 0.1$) under the null and eight alternatives

n	α_*								
	0.1	0.3	0.5	0.9	1.0	1.1	1.5	2.0	3.0
100	1.000	0.844	0.512	0.116	0.095	0.090	0.232	0.399	0.633
1000	1.000	1.000	0.924	0.125	0.096	0.119	0.431	0.816	0.952
10000	1.000	1.000	0.999	0.162	0.107	0.150	0.754	0.953	0.962

Observe that in the case $\alpha_* = 1$ (null hypothesis), these probabilities approximate the Type I error, while in the rest of the cases they approximate the power of the test under various alternatives

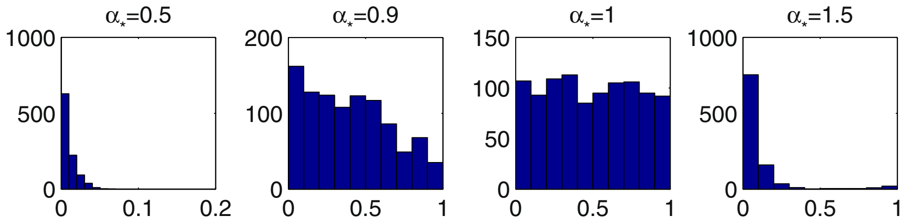

Fig. 1 Histograms of the p-values based on 1000 replicates of the test. The first two panels (*left* to *right*) correspond to the alternatives $\alpha_1 = \alpha_*, \alpha_2 = \alpha_3 = 1$, with $\alpha_* = 0.5$ and 0.9, respectively. The third panel corresponds to the null and the right-most panel to the alternative $\alpha_1 = \alpha_2 = 1$, and $\alpha_3 = \alpha_* = 1.5$

a power law tail model is appropriate). Under the null hypothesis, the data can be ordered in terms of the vector norm. Then the behavior of the test for sample sizes $n = 100$, 1000, and 10 000 in our simulation study can serve as a proxy to the *large sample* behavior of the test for heavy tailed data, where the threshold is chosen to grow with the sample size, and the number of data points with $\|X_i\| \geq u_n$ equals 100, 1000, or 10 000.

We tested nine scenarios, where two of the tail exponents α_i's were set equal to 1, and the third tail exponent α_* varied over the set $\{0.1, 0.3, 0.5, 0.9, 1, 1.1, 1.5, 2, 3\}$. Table 1 shows empirical rejection probabilities for 1000 independent replications of the test. Observe that under the null hypothesis ($\alpha_* = 1$), the test is essentially exact, with Type I error equal to the nominal level $\beta = 0.1$. The power of the test increases as α_* departs from 1. Rejection probabilities are highest for the alternative $\alpha_* = 0.1$, the case in which the ratio between tail indices is largest. Figure 1 illustrates the range of p-values in several cases. Under the null hypothesis, the distribution of the p-values is nearly uniform, which confirms the accuracy of the parametric bootstrap.

7 Applications

We applied the test for variations in tail index from Section 5 to two bivariate data sets studied in the literature. The first one consists of $n = 2853$ consecutive daily log returns for exchange rates of the Deutschmark versus the US dollar, and Japanese Yen versus the US dollar (Meerschaert and Scheffler 2003; Nolan et al. 2001). A scatter-plot of the data set is shown in Fig. 2 (top left). A multivariate stable model with the same tail index $\alpha = 1.65$ was fitted to this data in Nolan et al. (2001). An alternative model was proposed in Meerschaert and Scheffler (2003), with a heavier tail with index $\alpha_2 = 1.65$ along the 45° axis, and a lighter tail with $\alpha_1 \approx 2.0$ along the −45° axis. The second data set describes flow through a simulated fracture network, obtained from a large simulation study for site characterization of a proposed nuclear waste repository in Sweden (Meerschaert and Scheffler 2003; Painter et al. 2002). The two variables are travel time τ in years, and inverse velocity β in years/meter. Here the power law behavior is quite clear, and exhibits as a long straight

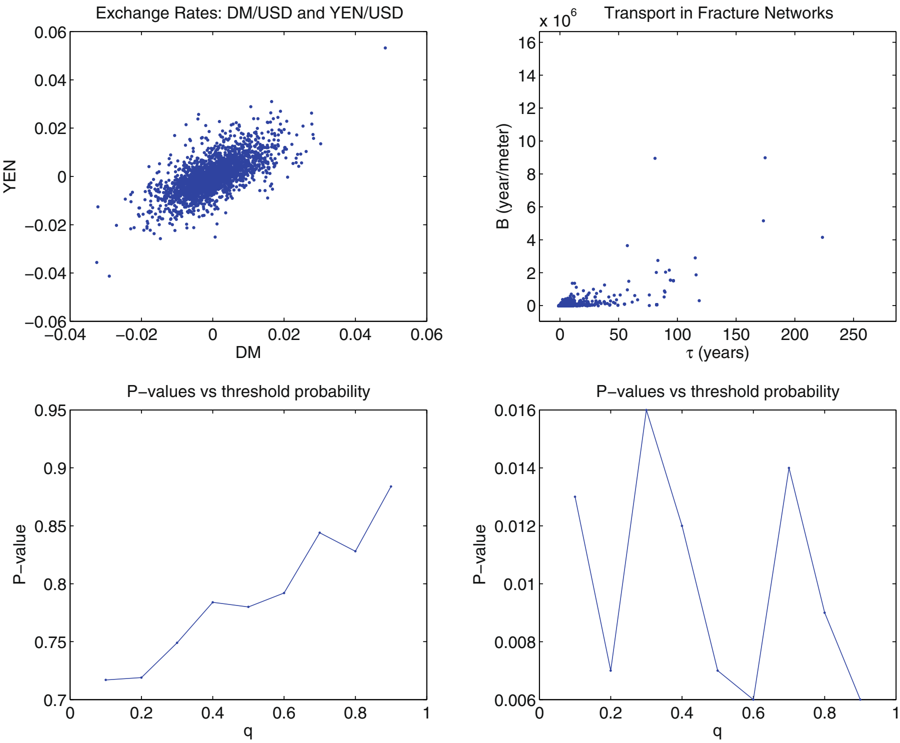

Fig. 2 *Top left*: Log returns of daily exchange rates. *Top right*: Flow data from a fracture network. *Bottom left* and *right*: The p-values of the proposed test for variations in the tail index, applied to the corresponding data sets above, versus the fraction q of largest (in norm) data points used to estimate the spectral measure. The null hypothesis (same tail index in every direction) is conclusively rejected for the fracture data, but not for the exchange rate data

line on a log-log plot of the sorted data for each variable, allowing an easy estimate of $\alpha_2 = 1.05$ for β (vertical axis) and $\alpha_1 = 1.4$ for τ (horizontal axis).

The estimation of the tail exponent under the null hypothesis assumption is one of the most delicate problems in practice. For these two data sets, we used the previously obtained estimates of $\alpha = 1.65$ for the exchange rate data set and $\alpha = 1.05$ for the fracture transport data. Since the heaviest tail dominates, the lower index is appropriate, and this is also consistent with the simulation study in Section 6. The test uses Proposition 5.2 to estimate the spectral measure, and because the test may be sensitive to the number of upper order statistics used, we consider a range of thresholds $u = u(q)$, corresponding largest $100q$ % of the observations, ordered in terms of the vector norm. The resulting p-values are shown in Fig. 2 (bottom), obtained using the parametric bootstrap with 1000 independent replications. Observe that for the exchange rate data set, the p-values are high, and our test fails to reject the null hypothesis of scalar norming. On the other hand, in the case of fracture transport, there is strong evidence in favor of operator norming.

References

Balkema, A.: Monotone transformations and limit laws. Mathematical Centre Tracts, vol. 45. Mathematisch Centrum, Amsterdam (1973)
de Haan, L.: A spectral representation for max–stable processes. Ann. Probab. **12**(4), 1194–1204 (1984)
de Haan, L., Ferreira, A.: Extreme value theory. Springer Series in Operations Research and Financial Engineering. An introduction. Springer, New York (2006)
Dembo, A., Zeitouni, O.: Large deviations techniques and applications. Applications of Mathematics (New York), vol. 38. 2nd edn. Springer, New York (1998)
Garcia, N.L., Kurtz, T.G.: Spatial point processes and the projection method. In: In and Out of Equilibrium 2, Progress in Probability, vol. 60, pp. 271–298. Springer, New York (2008)
Jurek, Z.J., Mason, J.D.: Operator-limit distributions in probability theory. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. Wiley, New York (1993)
Kabluchko, Z.: Spectral representations of sum- and max-stable processes. Extremes **12**(4), 401–424 (2009)
Khoshnevisan, D.: Multiparameter processes: an introduction to random fields. Springer Monographs in Mathematics. Springer, New York (2002)
Meerschaert, M.M.: Regular variation in $\mathbb{R}^k$. Proc. Am. Math. Soc. **102**, 341–348 (1988)
Meerschaert, M.M., Scheffler, H.-P.: Limit Distributions for Sums of Independent Random Vectors: Heavy Tails in Theory and Practice. Wiley, New York (2001)
Meerschaert, M.M., Scheffler, H.-P.: Nonparametric methods for heavy tailed vector data: a survey with applications from finance and hydrology. In: Recent Advances and Trends in Nonparametric Statistics, pp. 265–279. Elsevier, Amsterdam (2003)
Meerschaert, M.M., Scalas, E.: Coupled continuous time random walks in finance. Physica A **370**, 114–118 (2006)
Mittnik, S., Rachev, S.: Asset and Option Pricing with Alternative Stable Models. Wiley, New York (1999)
Nolan, J.P., Panorska, A.K., McCulloch, J.H.: Estimation of stable spectral measures. Math. Comput. Model. **34**, 1113–1122 (2001)
Painter, S., Cvetkovic, V., Selroos, J.: Power-law velocity distributions in fracture networks: numerical evidence and implications for tracer transport. Geophys. Res. Lett. **29**(14), 20-1–20-4 (2002)
Reeves, D.M., Benson, D.A., Meerschaert, M.M., Scheffler, H.-P.: Transport of conservative solutes in simulated fracture networks 2. Ensemble solute transport and the correspondence to operator-stable limit distributions. Water Resour. Res. **44**, W05410 (2008)
Resnick, S.I.: Extreme Values, Regular Variation and Point Processes. Springer, New York (1987)
Resnick, S.I.: Heavy-tail phenomena. Springer Series in Operations Research and Financial Engineering. Springer, New York (2007)
Stoev, S., Taqqu, M.S.: Extremal stochastic integrals: a parallel between max–stable processes and $\alpha-$stable processes. Extremes **8**, 237–266 (2005)
Stoev, S., Meerschaert, M.M., Scheffler, H.-P.: MATLAB Code for Testing in Operator Normalized Extremes (2012). http://www.stat.lsa.umich.edu/~sstoev/prm/oEVT-code.zip