Extremes of scale mixtures of multivariate time series

Helena Ferreira Department of Mathematics, University of Beira Interior, Covilhã, Portugal (helena.ferreira@ubi.pt)

Marta Ferreira Center of Mathematics of Minho University, Braga, Portugal (msferreira@math.uminho.pt)

Abstract: Factor models have large potencial in the modeling of several natural and human phenomena. In this paper we consider a multivariate time series $\mathbf{Y}_n$, $n \geq 1$, rescaled through random factors $\mathbf{T}_n$, $n \geq 1$, extending some scale mixture models in the literature. We analyze its extremal behavior by deriving the maximum domain of attraction and the multivariate extremal index, which leads to new ways to construct multivariate extreme value distributions. The computation of the multivariate extremal index and the characterization of the tail dependence show an interesting property of these models. More precisely, however much it is the dependence within and between factors $\mathbf{T}_n$, $n \geq 1$, the extremal index of the model is unit whenever $\mathbf{Y}_n$, $n \geq 1$, presents cross-sectional and sequential tail independence. We illustrate with examples of thinned multivariate time series and multivariate autoregressive processes with random coefficients. An application of these latter to financial data is presented at the end.

Keywords: multivariate extreme value theory, factor models, tail dependence

2000 Mathematics Subject Classification: 60G70

1 Introduction

Factor models have been used in the modeling of data within hydrology (Nadarajah [26, 27] 2006/2009, Nadarajah and Masoom [28] 2008), storm insurance (Lescourret and Robert, [22] 2006), soil erosion in crops (Todorovic and Gani [36] 1987, Alpuim and Athayde [2] 1990), reliability (Alpuim and Athayde [2] 1990, Kotz *et al.* [19] 2000), economy (Arnold, [3] 1983) and finance (Ferreira and Canto e Castro, [13] 2010).

Let $\mathbf{X}_n = (X_{n1}, \ldots, X_{nd})$, $n \geq 1$, be a d-variate sequence, such that $X_{nj} = Y_{nj}T_{nj}$, $j = 1, \ldots, d$, where

(a) $\mathbf{Y} = \{(Y_{n1}, \ldots, Y_{nd})\}_{n \geq 1}$ is a stationary sequence such that, Y_{nj} has a Pareto-type distribution F_{Y_j}, $j = 1, \ldots, d$, i.e., for each $j = 1, \ldots, d$, there exists a positive constant β_j for which

$$F_{Y_j}(x) = 1 - x^{-\beta_j} l_{Y_j}(x), \tag{1}$$

with l_{Y_j} a slowly varying function, i.e., $l_{Y_j}(ax)/l_{Y_j}(x) \to 1$, as $x \to \infty$, for all $a > 0$,

(b) $\mathbf{T} = \{(T_{n1}, \ldots, T_{nd})\}_{n \geq 1}$ is a stationary sequence, independent of $\mathbf{Y}$, with support $\mathbb{R}_+^d$ and such that $E(T_{nj}^{\epsilon_j}) < \infty$, for some $\epsilon_j > \beta_j$, $j = 1, \ldots, d$.

This work is concerned with the extremal behavior of the multivariate time series $\mathbf{X}_n$, extending most of the factor models mentioned above. More precisely, we derive the max-domain of attraction (Section 2), calculate the multivariate extremal index (Section 3) and characterize the tail dependence (Section 4).

The product $Y_{nj}T_{nj}$ can be seen as a random normalization of Y_{nj} by T_{nj}, which is often required when modeling extremal behavior. For instance, if Y_{nj} is the rate of an extreme event and T_{nj} its average cost, then $Y_{nj}T_{nj}$ can be interpreted as the total cost of the extreme event. Products of two independent

random variables where one of them is regularly varying have been addressed from both theoretical and applied points of view (Maulik *et al.* [24] 2002, Lescourret and Robert [22] 2006, Nadarajah [26] 2006 and references therein).

Our motivation to the probabilistic study of extremes of multivariate sequences of products was originated from some particular models. Consider, for instance that T_{nj} are Bernoulli distributed. Then $\mathbf{X}_n$ provides a model for multivariate data subjected to missing values. Extremes of univariate sequences with random missing values have been considered in Weissman and Cohen ([37], 1995) as a particular case of some mixture models. Additional results on extremes of incomplete samples can be found in Mladenovic and Piterbarg ([25], 2006) and Tan and Wang ([34], 2012).

Li ([23], 2009) analyzed the tail dependence of the scale mixture $\mathbf{X}_n$ when $\mathbf{Y}_n = (Y_{n1}, \ldots, Y_{nd})$ has multivariate extreme value distribution with standard Fréchet margins and $T_{nj} = T_n$, $j = 1, \ldots, d$. Here we consider scale mixtures of multivariate sequences which are very flexible models for data exhibiting tail dependence and asymptotic tail independence such as, respectively, ARMAX and pARMAX sequences (Ferreira and Ferreira [15]). We give particular emphasis to a model in which $\beta_j = \alpha/\gamma_j$, $\alpha, \gamma_j > 0$, $j = 1, \ldots, d$, generalizing the results of Lescourret and Robert ([22], 2006) (Section 5). An application to financial data will be provided at the end (Section 6).

2 Preliminary results and max-domain of attraction

We start with some properties of $\{\mathbf{X}_n\}_n \geq 1$, that will be used along the paper. We use notation $r_j = E(T_{nj}^{\beta_j})$ along the paper.

Proposition 2.1. *For each $j = 1, \ldots, d$, $\{X_{nj}\}_{n \geq 1}$ is a stationary sequence having Pareto-type distribution.*

Proof First, observe that

$$\lim_{x \to \infty} \frac{P(Y_{nj}T_{nj} > x)}{P(Y_{nj} > x)} = \lim_{x \to \infty} \frac{\int x^{-\beta_j} z^{\beta_j} l_{Y_j}(x/z) dP_{T_{nj}}(z)}{x^{-\beta_j} l_{Y_j}(x)} = r_j,$$

where the last step is due to the dominated convergence theorem and by using the Potter bounds of regularly varying functions (Bingham *et al.*, [5] 1987; Theorem 1.5.6.). Therefore, for large x,

$$1 - F_{X_j}(x) = P(X_{nj} > x) = x^{-\beta_j} l_{Y_j}(x) r_j (1 + o(1)) := x^{-\beta_j} l_{X_j}(x), \tag{2}$$

where it is immediately seen that l_{X_j} is a slowly varying function. $\qquad \square$

In the sequel we denote $U_{X_j}(x)$ and $U_{Y_j}(x)$ the quantile functions, $F_{X_j}^{-1}(1 - 1/x)$ and $F_{Y_j}^{-1}(1 - 1/x)$, respectively.

Given (1) and applying the the Bruyn conjugate concept (Beirlant *et al.* [4] 2004, Proposition 2.5), we have that, for large x,

$$U_{Y_j}(x) = x^{1/\beta_j} l_{U_{Y_j}}(x) = x^{1/\beta_j} l_{Y_j}^{1/\beta_j}(x^{1/\beta_j})(1 + o(1)), \tag{3}$$

where $l_{U_{Y_j}}$ is a slowly varying function. Using again the the Bruyn conjugate concept and by (2), we can state, for large x,

$$U_{X_j}(x) = x^{1/\beta_j} l_{U_{X_j}}(x) = x^{1/\beta_j} l_{X_j}^{1/\beta_j}(x^{1/\beta_j})(1 + o(1)) = x^{1/\beta_j} \left(l_{Y_j}(x^{1/\beta_j}) r_j \right)^{1/\beta_j} (1 + o(1)),$$

where $l_{U_{X_j}}$ is a slowly varying function. Therefore, and considering (3), we have for large x,

$$U_{X_j}(x) = x^{1/\beta_j} l_{U_{Y_j}}(x) r_j^{1/\beta_j}(1 + o(1)) = U_{Y_j}(r_j x)(1 + o(1)). \tag{4}$$

Proposition 2.2. *The upper tail copula function of* **X** *is given by*

$$\Lambda_{\mathbf{X}}(x_1,...,x_d) = E\left(\Lambda_{\mathbf{Y}}\left(\frac{T_1^{\beta_1}x_1}{r_1},\ldots,\frac{T_d^{\beta_d}x_d}{r_d}\right)\right),$$

with $(x_1,...,x_d) \in \overline{\mathbb{R}}_+^d = [0,\infty]^d \backslash \{(\infty,\ldots,\infty)\}$, *where* $\mathbf{T} = (T_1,\ldots,T_d)$ *is a random vector distributed as* $\mathbf{T}_n = (T_{n1},\ldots,T_{nd})$ *and provided that the upper tail copula function of* $\mathbf{Y}_n$ *exists, i.e., the limit*

$$\Lambda_{\mathbf{Y}}(x_1,...,x_d) = \lim_{t\to\infty} tP\left(\bigcap_{j=1}^{d}\{Y_{1j} > U_{Y_j}(t/x_j)\}\right) \tag{5}$$

is finite.

Proof The upper tail copula function of **X** is defined by

$$\Lambda_{\mathbf{X}}(x_1,...,x_d) = \lim_{t\to\infty} tP\left(\bigcap_{j=1}^{d}\{X_{1j} > U_{X_j}(t/x_j)\}\right)$$

The result follows immediately by applying (4) and the dominated convergence theorem, since

$$
\begin{aligned}
&\lim_{t\to\infty} tP\left(\bigcap_{j=1}^{d}\{X_{1j} > U_{X_j}(t/x_j)\}\right)\\
=\ & \lim_{t\to\infty} t\int P\left(\bigcap_{j=1}^{d}\left\{Y_{1j} > U_{Y_j}\left(\frac{r_j\,t}{z_j^{\beta_j}\,x_j}\right)\right\}\right) dP_{(T_{11},\ldots,T_{1d})}(z_1,\ldots,z_d)\\
=\ & \int \Lambda_{\mathbf{Y}}\left(\frac{z_1^{\beta_1}x_1}{r_1},\ldots,\frac{z_d^{\beta_d}x_d}{r_d}\right) dP_{(T_{11},\ldots,T_{1d})}(z_1,\ldots,z_d)\\
=\ & E\left(\Lambda_{\mathbf{Y}}\left(\frac{T_1^{\beta_1}x_1}{r_1},\ldots,\frac{T_d^{\beta_d}x_d}{r_d}\right)\right). \qquad \square
\end{aligned} \tag{6}
$$

The results of Proposition 2.1 and 2.2 are related to the Breiman's theorem (see, e.g., Proposition 7.5 in Resnick 2007).

Remark 2.3. *If* $\beta_j = \beta$ *and* $T_{nj} = T$, $j = 1,\ldots,d$, *then* $\Lambda_{\mathbf{X}} = \Lambda_{\mathbf{Y}}$, *corresponding to the case considered in Li ([23], 2009). If* $\Lambda_{\mathbf{Y}}(x_1,...,x_d) = \sum_{i=1}^{d} x_j$, *we also obtain* $\Lambda_{\mathbf{X}} = \Lambda_{\mathbf{Y}}$, *for any choice of* $\mathbf{T}$. *In the last section we give particular attention to the case* $\Lambda_{\mathbf{Y}}(x_1,...,x_d) = \bigwedge_{i=1}^{d} x_j$.

Normalized levels $u_{nj}^{(\tau_j)}$ of X_j are such that $n(1 - F_{X_j}(u_{nj}^{(\tau_j)})) \to \tau_j > 0$, as $n \to \infty$, i.e., $u_{nj}^{(\tau_j)} = U_{X_j}(n/\tau_j)(1 + o(1))$, as $n \to \infty$. Thus, by (4) and for large n, X_j has normalized levels

$$u_{nj}^{(\tau_j)} = U_{Y_j}(nr_j/\tau_j)(1 + o(1)).$$

Consider $\mathbf{u}_n(\boldsymbol{x}) = (u_{n1}^{(\tau_1(x_1))},\ldots,u_{nd}^{(\tau_d(x_d))}) = (nx_1,\ldots,nx_d)$ a vector of normalized levels of **X** and, for each $j = 1,\ldots,d$, let $\{\widehat{X}_{nj}\}_{n\geq1}$ be an i.i.d. sequence with the same marginal distribution as $\{X_{nj}\}_{n\geq1}$. Denote $\widehat{\mathbf{M}}_n = (\widehat{M}_{n1},\ldots,\widehat{M}_{nd})$ the vector of the componentwise maxima $\widehat{M}_{nj} = \bigvee_{i=1}^{n} \widehat{X}_{ij}$, $j = 1,\ldots,d$.

Proposition 2.4. *We have* $F_{\mathbf{X}_1}$ *in the domain of attraction of* $G_{\mathbf{X}}$, *that is,*

$$\lim_{n\to\infty} P\left(\widehat{\mathbf{M}}_n \leq \mathbf{u}_n(\boldsymbol{x})\right) = G_{\mathbf{X}}(x_1,...,x_d),$$

4

with

$$G_{\mathbf{X}}(x_1, ..., x_d) = \exp\left\{ -E\left(-\log G_{\mathbf{Y}}\left(\frac{x_1 r_1}{T_1^{\beta_1}}, \ldots, \frac{x_d r_d}{T_d^{\beta_d}} \right) \right) \right\},$$

provided that $F_{\mathbf{Y}_1}$ is in the domain of attraction of $G_{\mathbf{Y}}(x_1, ..., x_d)$, where both $G_{\mathbf{X}}$ and $G_{\mathbf{Y}}$ have unit Fréchet marginals.

Proof Just observe that

$$\lim_{n \to \infty} P\left(\bigcap_{j=1}^{d} \left\{ \widehat{M}_{nj} \le U_{X_j}(nx_j) \right\} \right) = \lim_{n \to \infty} \exp\left\{ -nP\left(\bigcup_{j=1}^{d} \left\{ X_{1j} > U_{X_j}(nx_j) \right\} \right) \right\}.$$

Now the proof runs along the same lines as in (6). $\qquad\qquad\square$

The results of Proposition 2.1 and 2.4 are related to the Breiman's theorem (see, e.g., Proposition 7.5 in Resnick [31] 2007).

In the following, for any vector $\mathbf{z}$ and $A \subset \{1, \ldots, d\}$, $\mathbf{z}_A$ denotes the sub-vector of $\mathbf{z}$ with indices in A.

Example 1. If $(T_1, \ldots, T_d)$ has support in $\{0, 1\}^d$ and $P\left(\bigcap_{i \in J} T_i = 1, \bigcap_{i \in D \setminus J} T_i = 0 \right) = p(J), \emptyset \ne J \subset D = \{1, \ldots, d\}$, then $G_{\mathbf{X}}$ corresponds to a geometric mean of the marginal distributions of $G_{\mathbf{Y}}$. In this case, we have

$$G_{\mathbf{X}}(x_1, \ldots, x_d) = \prod_{\emptyset \ne J \subset D} G_{\mathbf{Y}_J}^{p(J)}(r_1 x_1, \ldots, r_d x_d)_J = \prod_{\emptyset \ne J \subset D} G_{\mathbf{Y}_J}\left(\frac{r_1 x_1}{p(J)}, \ldots, \frac{r_d x_d}{p(J)} \right)_J$$

$$= \prod_{\emptyset \ne J \subset D} G_{\mathbf{Y}_J}\left(\frac{p(\{1\}) x_1}{p(J)}, \ldots, \frac{p(\{d\}) x_d}{p(J)} \right)_J.$$

We illustrate the result with some choices for $G_{\mathbf{Y}}$.

If the stationary sequence $\mathbf{Y}$ has common copula logistic, i.e.,

$$C_{\mathbf{Y}_n}(u_1, \ldots, u_d) = \exp\left\{ -\left(\sum_{j=1}^{d} (-\log u_j)^{-1/\alpha} \right)^{\alpha} \right\}$$

then we obtain for $G_{\mathbf{Y}}$ the logistic distribution and

$$G_{\mathbf{X}}(x_1, \ldots, x_d) = \exp\left\{ -\sum_{\emptyset \ne J \subset D} \left(\sum_{i \in J} (\beta_{Ji} x_i)^{-1/\alpha} \right)^{\alpha} \right\},$$

with $\beta_{Ji} = p(\{i\})/p(J)$, which is an asymmetric logistic distribution already found in Tawn ([35], 1990), by following a different probabilistic approach. More precisely, this parametric multivariate extreme value distribution was presented as the joint survivor function a particular random vector with unit exponential marginals. The parameters β_{Ji} increase the variability within the tail dependence coefficients regarding the departure distribution $G_{\mathbf{Y}}$. For $\mathbf{X} = (X_1, \ldots, X_d)$ with distribution $G_{\mathbf{X}}$, we have, for instance,

$$\Lambda_{(X_i, X_j)}(1, 1) = \lim_{t \to \infty} P(F_{X_i}(X_i) > 1 - 1/t | F_{X_j}(X_j) > 1 - 1/t) = 2 - \sum_{J \in \mathcal{F}_{\{i,j\}}} \left(\beta_{Ji}^{-1/\alpha} + \beta_{Jj}^{-1/\alpha} \right)^{\alpha},$$

where $\mathcal{F}_{\{i,j\}}$ is the family of subsets of D containing $\{i,j\}$, expression that presents a larger number of possibilities than the value $\Lambda_{(Y_i,Y_j)}(1,1) = 2 - 2^\alpha$ of the symmetric logistic distribution. Note that $\sum_{J\in\mathcal{F}_{\{i\}}} \beta_{Ji} = 1$.

Now, suppose that $G_{\mathbf{Y}}(x_1,\ldots,x_d) = \prod_{l=1}^\infty \prod_{k=-\infty}^\infty \bigwedge_{j=1}^d \exp\left(-a_{lkj}x_j^{-1}\right)$, where $\{a_{lkj}, l \geq 1, -\infty < k < \infty, 1 \leq j \leq d\}$ are real non negative constants satisfying $\sum_{l=1}^\infty \sum_{k=-\infty}^\infty a_{lkj} = 1$, $j = 1,\ldots,d$. This is the attractor MEV when $\mathbf{Y}$ is an M4 process (Smith and Weissman [33], 1996) and we have

$$G_{\mathbf{X}}(x_1,\ldots,x_d) = \prod_{l=1}^\infty \prod_{k=-\infty}^\infty \prod_{\emptyset \neq J \subset D} \bigwedge_{j\in J} \exp\left(-\frac{a_{lkj}}{\beta_{Jj}}x_j^{-1}\right)$$

and the bivariate tail dependence summarized by

$$\Lambda_{(X_i,X_j)}(1,1) = 2 - \sum_{l=1}^\infty \sum_{k=-\infty}^\infty \sum_{J\in\mathcal{F}_{\{i,j\}}} \left(\frac{a_{lki}}{\beta_{Ji}} \vee \frac{a_{lkj}}{\beta_{Jj}}\right) \leq 2 - \sum_{l=1}^\infty \sum_{k=-\infty}^\infty (a_{lki} \vee a_{lkj}) = \Lambda_{(Y_i,Y_j)}(1,1).\blacktriangledown$$

3   The multivariate extremal index

The extremal index measures the tendency of clusters occurrence, a phenomena commonly observed in real data. In this section we will compute the multivariate extremal index of $\mathbf{X}_n$ (Nandagopalan [29], 1990). We start by analyzing some long range and local dependence conditions that will make easier its calculation.

Proposition 3.1. *If $\mathbf{Y}$ and $\mathbf{T}$ are strong-mixing, then $\mathbf{X}$ is strong-mixing.*

Proof Consider the events A and B, respectively, in σ-algebras generated by $\{\mathbf{X}_1,\ldots,\mathbf{X}_p\}$ and $\{\mathbf{X}_{p+s},\mathbf{X}_{p+s+1},\ldots\}$, i.e., $A \in \sigma(\mathbf{X}_1,\ldots,\mathbf{X}_p)$ and $B \in \sigma(\mathbf{X}_{p+s},\ldots)$. Given the independence between $\mathbf{Y}$ and $\mathbf{T}$, we can state

$$P(A \cap B) = E(P(A' \cap B'|\mathbf{Y})) = E\left(P(A'|\mathbf{Y}_1,\ldots,\mathbf{Y}_p)P(B'|\mathbf{Y}_{p+s},\ldots)\right) + O(\alpha_{\mathbf{T}}(s)),$$

where $A' \in \sigma(\mathbf{T}_1,\ldots,\mathbf{T}_p)$, $B' \in \sigma(\mathbf{T}_{p+s},\ldots)$ and $\alpha_{\mathbf{T}}(s)$ denotes the mixture coefficients of the sequence $\mathbf{T}$. The result follows by Doukhan [10] (1995, Theorem 3 in Section 1.2.2), since

$$E\left(P(A'|\mathbf{Y}_1,\ldots,\mathbf{Y}_p)P(B'|\mathbf{Y}_{p+s},\ldots)\right) \leq E\left(P(A'|\mathbf{Y}_1,\ldots,\mathbf{Y}_p)\right) E\left(P(B'|\mathbf{Y}_{p+s},\ldots)\right) + 4\alpha_{\mathbf{Y}}(s),$$

where $\alpha_{\mathbf{Y}}(s)$ denotes the mixture coefficients of the sequence $\mathbf{Y}$. $\qquad\square$

Consider $\mathbf{u}_n(\boldsymbol{\tau}) = (u_{n1}^{(\tau_1)},\ldots,u_{nd}^{(\tau_d)})$ a vector of normalized levels of $\mathbf{X}$ and $\mathbf{M}_{1,n} \equiv \mathbf{M}_n = (M_{n1},\ldots,M_{nd})$ the vector of the componentwise maxima $M_{nj} = \bigvee_{i=1}^n X_{ij}$, $j = 1,\ldots,d$. As a consequence of the strong-mixing condition, we have

$$\lim_{n\to\infty} P\left(\mathbf{M}_n \leq \mathbf{u}_n^{(\boldsymbol{\tau})}\right) = \lim_{n\to\infty} P\left(M_{n1} \leq u_{n1}^{(\tau_1)},\ldots,M_{nd} \leq u_{nd}^{(\tau_d)}\right)$$

$$= \exp\left\{-\lim_{n\to\infty} nP\left(\mathbf{X}_1 \not\leq \mathbf{u}_n^{(\boldsymbol{\tau})}, \bigcap_{i=2}^{[n/k_n]} \mathbf{X}_i \leq \mathbf{u}_n^{(\boldsymbol{\tau})}\right)\right\}$$

where $k_n \to \infty$ and $n/k_n \to \infty$, as $n \to \infty$. Therefore, $\mathbf{X}$ has multivariate extremal index

$$\theta_{\mathbf{X}}(\tau_1,\ldots,\tau_d) = \frac{-\log\lim_{n\to\infty} P\left(\mathbf{M}_n \leq \mathbf{u}_n^{(\boldsymbol{\tau})}\right)}{-\log\lim_{n\to\infty} P\left(\widehat{\mathbf{M}}_n \leq \mathbf{u}_n^{(\boldsymbol{\tau})}\right)},$$

if and only if $\lim_{n\to\infty} nP\left(\mathbf{X}_1 \not\leq \mathbf{u}_n^{(\boldsymbol{\tau})}, \mathbf{M}_{2,[n/k_n]} \leq \mathbf{u}_n^{(\boldsymbol{\tau})}\right)$ exists and, if so, we have

$$\theta_{\mathbf{X}}(\tau_1,\ldots,\tau_d) = \lim_{n\to\infty} \frac{P\left(\mathbf{X}_1 \not\leq \mathbf{u}_n^{(\boldsymbol{\tau})}, \mathbf{M}_{2,[n/k_n]} \leq \mathbf{u}_n^{(\boldsymbol{\tau})}\right)}{P\left(\mathbf{X}_1 \not\leq \mathbf{u}_n^{(\boldsymbol{\tau})}\right)}. \tag{7}$$

This function relates the two multivariate extreme value distributions arising from the maxima in $\{\mathbf{X}_n\}_{n\geq 1}$ and $\{\widehat{\mathbf{X}}_n\}_{n\geq 1}$. We analyze its values under additional assumptions on the scale sequence $\mathbf{Y}$.

Let

$$\frac{\mathbf{r}\,n}{\boldsymbol{\tau}\,\mathbf{T}_i^{\boldsymbol{\beta}}} = \left(\frac{r_1\,n}{\tau_1\,T_{i1}^{\beta_1}},\ldots,\frac{r_d\,n}{\tau_d\,T_{id}^{\beta_d}}\right)$$

and $\mathbf{U}_{\mathbf{Y}_i}(\mathbf{x}) = (U_{Y_{i1}}(x_1),\ldots,U_{Y_{id}}(x_d))$. About the numerator in (7) we can write

$$nP\left(\mathbf{X}_1 \not\leq \mathbf{u}_n^{(\boldsymbol{\tau})}, \mathbf{M}_{2,[n/k_n]} \leq \mathbf{u}_n^{(\boldsymbol{\tau})}\right)$$

$$= nP\left(\mathbf{Y}_1 \not\leq \mathbf{U}_{\mathbf{Y}_1}\left(\frac{\mathbf{r}\,n}{\boldsymbol{\tau}\,\mathbf{T}_1^{\boldsymbol{\beta}}}\right), \bigcap_{i=2}^{[n/k_n]} \mathbf{Y}_i \leq \mathbf{U}_{\mathbf{Y}_i}\left(\frac{\mathbf{r}\,n}{\boldsymbol{\tau}\,\mathbf{T}_i^{\boldsymbol{\beta}}}\right)\right). \tag{8}$$

In some cases, the sequence $\mathbf{Y}$ is such that

$$\lim_{n\to\infty} nP\left(\mathbf{Y}_1 \not\leq \mathbf{U}_{\mathbf{Y}_i}\left(\frac{n}{\boldsymbol{\tau}}\right), \bigcap_{i=2}^{[n/k_n]} \mathbf{Y}_i \leq \mathbf{U}_{\mathbf{Y}_i}\left(\frac{n}{\boldsymbol{\tau}}\right)\right)$$

$$= \lim_{n\to\infty} nP\left(\mathbf{Y}_1 \not\leq \mathbf{U}_{\mathbf{Y}_i}\left(\frac{n}{\boldsymbol{\tau}}\right), \bigcap_{i=2}^{k} \mathbf{Y}_i \leq \mathbf{U}_{\mathbf{Y}_i}\left(\frac{n}{\boldsymbol{\tau}}\right)\right),$$

for some levels $\mathbf{u}_n^{(\boldsymbol{\tau})} = \mathbf{U}_{\mathbf{Y}}\left(\frac{n}{\boldsymbol{\tau}}\right)$ and for some finite k, which makes easier the calculation of the multivariate extremal index. Situations leading to this easier procedure are highlighted in the next results.

Proposition 3.2. *Consider* $\mathbf{a} = (a_1,\ldots,a_d)$, $\mathbf{b} = (b_1,\ldots,b_d)$, T_{nj} *with support in* $[a_j,b_j]$, $j = 1,\ldots,d$, *and* $\mathbf{v}_n^{(\boldsymbol{\tau}^*(\mathbf{z}))} = \left(U_{Y_1}\left(\frac{r_1 n}{z_1^{\beta_1}\tau_1}\right),\ldots,U_{Y_d}\left(\frac{r_d n}{z_d^{\beta_d}\tau_d}\right)\right)$, *for all* $z_j \in [a_j,b_j]$, $j = 1,\ldots,d$. *If* $\mathbf{Y}$ *satisfies condition* $\mathtt{D}^{(k)}\left(\mathbf{v}_n^{(\boldsymbol{\tau}^*(\mathbf{a}))}, \mathbf{v}_n^{(\boldsymbol{\tau}^*(\mathbf{b}))}\right)$, *defined by*

$$\lim_{n\to\infty} n \sum_{i=k}^{[n/k_n]} P\left(\mathbf{Y}_1 \not\leq \mathbf{v}_n^{(\boldsymbol{\tau}^*(\mathbf{b}))}, \mathbf{Y}_i \leq \mathbf{v}_n^{(\boldsymbol{\tau}^*(\mathbf{a}))}, \mathbf{Y}_{i+1} \not\leq \mathbf{v}_n^{(\boldsymbol{\tau}^*(\mathbf{b}))}\right) = 0, \tag{9}$$

then $\mathtt{D}^{(k)}\left(\mathbf{u}_n^{(\boldsymbol{\tau})}\right)$, *defined by*

$$\lim_{n\to\infty} n \sum_{i=k}^{[n/k_n]} P\left(\mathbf{X}_1 \not\leq \mathbf{u}_n^{(\boldsymbol{\tau})}, \mathbf{X}_i \leq \mathbf{u}_n^{(\boldsymbol{\tau})}, \mathbf{X}_{i+1} \not\leq \mathbf{u}_n^{(\boldsymbol{\tau})}\right) = 0,$$

holds for $\mathbf{X}$, $\boldsymbol{\tau} = (\tau_1,\ldots,\tau_d)$.

Proof Observe that, if we take $\mathbf{u}_n^{(\boldsymbol{\tau})} = (U_{X_1}(n/\tau_1),\ldots,U_{X_d}(n/\tau_d))$, we have

$$\lim_{n\to\infty} n \sum_{i=k}^{[n/k_n]} P\left(\mathbf{X}_1 \not\leq \mathbf{u}_n^{(\boldsymbol{\tau})}, \mathbf{X}_i \leq \mathbf{u}_n^{(\boldsymbol{\tau})}, \mathbf{X}_{i+1} \not\leq \mathbf{u}_n^{(\boldsymbol{\tau})}\right)$$

$$= \lim_{n\to\infty} \int n \sum_{i=k}^{[n/k_n]} P\left(\mathbf{Y}_1 \not\leq \mathbf{v}_n^{(\boldsymbol{\tau}^*(\mathbf{z}_1))}, \mathbf{Y}_i \leq \mathbf{v}_n^{(\boldsymbol{\tau}^*(\mathbf{z}_i))}, \mathbf{Y}_{i+1} \not\leq \mathbf{v}_n^{(\boldsymbol{\tau}^*(\mathbf{z}_{i+1}))}\right) dP_{(\mathbf{T}_1,\mathbf{T}_i,\mathbf{T}_{i+1})}(\mathbf{z}_1,\mathbf{z}_i,\mathbf{z}_{i+1})$$

Now, just observe that

$$P\left(\mathbf{Y}_1 \not\leq \mathbf{v}_n^{(\boldsymbol{\tau}^*(\mathbf{z}_1))}, \mathbf{Y}_i \leq \mathbf{v}_n^{(\boldsymbol{\tau}^*(\mathbf{z}_i))}, \mathbf{Y}_{i+1} \not\leq \mathbf{v}_n^{(\boldsymbol{\tau}^*(\mathbf{z}_{i+1}))}\right)$$

$$\leq \quad P\left(\mathbf{Y}_1 \not\leq \mathbf{v}_n^{(\boldsymbol{\tau}^*(\mathbf{b}))}, \mathbf{Y}_i \leq \mathbf{v}_n^{(\boldsymbol{\tau}^*(\mathbf{a}))}, \mathbf{Y}_{i+1} \not\leq \mathbf{v}_n^{(\boldsymbol{\tau}^*(\mathbf{b}))}\right). \quad \square$$

Under the condition (9), smooth oscillations around $\mathbf{v}_n^{(\boldsymbol{\tau}^*(\mathbf{b}))}$ by $\mathbf{Y}$ (in the sense that its values which are at least lag-k apart no longer exceed $\mathbf{v}_n^{(\boldsymbol{\tau}^*(\mathbf{a}))} > \mathbf{v}_n^{(\boldsymbol{\tau}^*(\mathbf{b}))}$) are not followed by subsequent exceedances of $\mathbf{v}_n^{(\boldsymbol{\tau}^*(\mathbf{b}))}$. In the Example 4, we provide an illustration of such local behavior which leads to the condition $\mathrm{D}^{(k)}\left(\mathbf{u}_n^{(\boldsymbol{\tau})}\right)$ for the scale model $\mathbf{X}$.

In order to avoid restrictions on the support of T_{nj}, $j = 1, \ldots, d$, we propose now a greater restriction to the oscillations of $\mathbf{Y}$ around normalized levels. First observe that

$$P\left(\mathbf{Y}_1 \not\leq \mathbf{v}_n^{(\boldsymbol{\tau}^*(\mathbf{z}_1))}, \mathbf{Y}_i \leq \mathbf{v}_n^{(\boldsymbol{\tau}^*(\mathbf{z}_i))}, \mathbf{Y}_{i+1} \not\leq \mathbf{v}_n^{(\boldsymbol{\tau}^*(\mathbf{z}_{i+1}))}\right) \leq P\left(\mathbf{Y}_1 \not\leq \mathbf{v}_n^{(\boldsymbol{\tau}^*(\mathbf{z}))}, \mathbf{Y}_{i+1} \not\leq \mathbf{v}_n^{(\boldsymbol{\tau}^*(\mathbf{z}))}\right)$$

with

$$\mathbf{v}_n^{(\boldsymbol{\tau}^*(\mathbf{z}))} = \mathbf{U}_{\mathbf{Y}_1}\left(\frac{r_1 n}{(z_{11} \vee z_{i+1,1})^{\beta_1}\tau_1}, \ldots, \frac{r_d n}{(z_{1d} \vee z_{i+1,d})^{\beta_d}\tau_d}\right).$$

Therefore, we can take for $\mathbf{Y}$ a sequence satisfying the following restriction on the local occurrence of exceedances of $\mathbf{v}_n^{(\boldsymbol{\tau})}$ which are lag-k apart:

$$\lim_{n\to\infty} n \sum_{i=k}^{[n/k]} P\left(\mathbf{Y}_1 \not\leq \mathbf{v}_n^{(\boldsymbol{\tau})}, \mathbf{Y}_{i+1} \not\leq \mathbf{v}_n^{(\boldsymbol{\tau})}\right) = 0. \tag{10}$$

This condition limits the size of the clusters of high levels and is satisfied by k-dependent sequences $\mathbf{Y}$, i.e., $\mathbf{Y}_n$ and $\mathbf{Y}_m$ are independent whenever $|n - m| > k$.

Proposition 3.3. *If $\mathbf{Y}$ satisfies (10) for all $\mathbf{v}_n^{(\boldsymbol{\tau})} = \mathbf{U}_{\mathbf{Y}_1}\left(\frac{n}{\boldsymbol{\tau}}\right)$, then $\mathbf{X}$ satisfies $\mathrm{D}^{(k)}\left(\mathbf{u}_n^{(\boldsymbol{\tau})}\right)$, for any sequence of normalized levels $\mathbf{u}_n^{(\boldsymbol{\tau})} = \mathbf{U}_{\mathbf{X}_1}\left(\frac{n}{\boldsymbol{\tau}}\right)$.*

Corollary 3.4. *If $\mathbf{Y}$ is k-dependent, then $\mathbf{X}$ satisfies $\mathrm{D}^{(k)}\left(\mathbf{u}_n^{(\boldsymbol{\tau})}\right)$, for any sequence of normalized levels $\mathbf{u}_n^{(\boldsymbol{\tau})}$.*

If $\mathbf{X}$ is strong-mixing and satisfies $\mathrm{D}^{(k)}\left(\mathbf{u}_n^{(\boldsymbol{\tau})}\right)$ condition, for all $\boldsymbol{\tau} \in \mathbb{R}_+^d$, then it has multivariate extremal index given by

$$\theta_{\mathbf{X}}(\tau_1, \ldots, \tau_d) = \lim_{n\to\infty} \frac{P\left(\mathbf{X}_1 \leq \mathbf{u}_n^{(\boldsymbol{\tau})}, \ldots, \mathbf{X}_{k-1} \leq \mathbf{u}_n^{(\boldsymbol{\tau})}, \mathbf{X}_k \not\leq \mathbf{u}_n^{(\boldsymbol{\tau})}\right)}{P\left(\mathbf{X}_1 \not\leq \mathbf{u}_n^{(\boldsymbol{\tau})}\right)}, \tag{11}$$

whenever this limit exists (Chernick *et al.* [7] 1991, Ferreira [11] 1994) . In this case,

$$\lim_{n\to\infty} P\left(\mathbf{M}_n \leq \mathbf{u}_n^{(\boldsymbol{\tau})}\right) = \exp\left\{-\lim_{n\to\infty} nP\left(\mathbf{X}_1 \leq \mathbf{u}_n^{(\tau)}, \ldots, \mathbf{X}_{k-1} \leq \mathbf{u}_n^{(\tau)}, \mathbf{X}_k \not\leq \mathbf{u}_n^{(\tau)}\right)\right\}. \tag{12}$$

The probability in the second term can be computed throughout the tail behavior of $(\mathbf{Y}_1, \ldots, \mathbf{Y}_k)$. In order to improve such approach, define

$$\Lambda_{(\mathbf{Y}_1, \ldots, \mathbf{Y}_k)}(\mathbf{x}_1, \ldots, \mathbf{x}_k) = \lim_{t\to\infty} t P\left(\mathbf{Y}_1 \not\leq \mathbf{U}_{\mathbf{Y}_1}\left(\frac{t}{\mathbf{x}_1}\right), \ldots, \mathbf{Y}_k \not\leq \mathbf{U}_{\mathbf{Y}_k}\left(\frac{t}{\mathbf{x}_k}\right)\right), \tag{13}$$

with $\mathbf{x}_j \in \overline{\mathbb{R}}_+^d$, $j = 1, \ldots, k$. This definition extends the upper tail dependence concept of Definition 2.3 in Ferreira and Ferreira ([14], 2012).

Note that, for $\mathbf{x}_i$ with null components x_{ij} with $j \in \{1,\ldots,d\}\backslash J_i$, $J_i \neq \emptyset$, $i = 1,\ldots,k$, we have

$$\Lambda_{(\mathbf{Y}_1,\ldots,\mathbf{Y}_k)}(\mathbf{x}_1,\ldots,\mathbf{x}_k) = \Lambda_{((\mathbf{Y}_1)_{J_1},\ldots,(\mathbf{Y}_k)_{J_k})}((\mathbf{x}_1)_{J_1},\ldots,(\mathbf{x}_k)_{J_k}),$$

where $(\mathbf{x}_i)_{J_i}$ denotes the sub-vector of $\mathbf{x}_i$ having components indexed in J_i. On the other hand, if $\mathbf{x}$ has some null component, we have $\Lambda_{\mathbf{Y}_1}(\mathbf{x}) = 0$.

We are going to apply the result in (12) to our model and derive an expression for the multivariate extremal index, which will depend on the upper tail dependence function of $(\mathbf{Y}_1,\ldots,\mathbf{Y}_k)$ given in (13) and also on the dependence between random vectors $\mathbf{T}_1,\ldots,\mathbf{T}_k$.

Proposition 3.5. *If (13) holds for all $\mathbf{x}_1,\ldots,\mathbf{x}_k \in \overline{\mathbb{R}}_+^d$ then*

$$\lim_{n\to\infty} n\, P\left(\mathbf{X}_1 \leq \mathbf{u}_n^{(\boldsymbol{\tau})},\ldots,\mathbf{X}_{k-1} \leq \mathbf{u}_n^{(\boldsymbol{\tau})}, \mathbf{X}_k \not\leq \mathbf{u}_n^{(\boldsymbol{\tau})}\right)$$

$$= E\left(\sum_{\emptyset \subseteq I \subset \{1,\ldots,k-1\}} (-1)^{|I|} \Lambda_{(\mathbf{Y}_1,\ldots,\mathbf{Y}_k)_{I\cup\{k\}}}\left(\frac{\boldsymbol{\tau}\,\mathbf{T}_1^{\boldsymbol{\beta}}}{\mathbf{r}},\ldots,\frac{\boldsymbol{\tau}\,\mathbf{T}_k^{\boldsymbol{\beta}}}{\mathbf{r}}\right)_{I\cup\{k\}}\right).$$

Proof Observe that

$$n\, P\left(\mathbf{X}_1 \leq \mathbf{u}_n^{(\boldsymbol{\tau})},\ldots,\mathbf{X}_{k-1} \leq \mathbf{u}_n^{(\boldsymbol{\tau})}, \mathbf{X}_k \not\leq \mathbf{u}_n^{(\boldsymbol{\tau})}\right)$$

$$= n\left\{P\left(\mathbf{X}_k \not\leq \mathbf{u}_n^{(\boldsymbol{\tau})}\right) - \sum_{\emptyset \neq I \subset \{1,\ldots,k-1\}} (-1)^{|I|+1} P\left(\bigcap_{i\in I}\mathbf{X}_i \not\leq \mathbf{u}_n^{(\boldsymbol{\tau})}, \mathbf{X}_k \not\leq \mathbf{u}_n^{(\boldsymbol{\tau})}\right)\right\}$$

$$= \sum_{\emptyset \subseteq I \subset \{1,\ldots,k-1\}} (-1)^{|I|} n P\left(\bigcap_{i\in I\cup\{k\}} \mathbf{Y}_i \not\leq \mathbf{U}_{\mathbf{Y}_i}\left(\frac{\mathbf{r}\,n}{\boldsymbol{\tau}\,\mathbf{T}_i^{\boldsymbol{\beta}}}\right)\right),$$

which leads to the result. $\qquad\square$

Corollary 3.6. *If $\mathbf{X}$ satisfies strong-mixing and $\mathrm{D}^{(k)}\left(\mathbf{u}_n^{(\boldsymbol{\tau})}\right)$ conditions, for all $\boldsymbol{\tau} \in \mathbb{R}_+^d$, and if (13) holds for all $\mathbf{x}_1,\ldots,\mathbf{x}_k \in \overline{\mathbb{R}}_+^d$ then*

$$\theta_{\mathbf{X}}(\tau_1,\ldots,\tau_d) = 1 - \frac{E\left(\displaystyle\sum_{\emptyset \neq I \subset \{1,\ldots,k-1\}} (-1)^{|I|+1} \Lambda_{(\mathbf{Y}_1,\ldots,\mathbf{Y}_k)_{I\cup\{k\}}}\left(\dfrac{\boldsymbol{\tau}\,\mathbf{T}_1^{\boldsymbol{\beta}}}{\mathbf{r}},\ldots,\dfrac{\boldsymbol{\tau}\,\mathbf{T}_k^{\boldsymbol{\beta}}}{\mathbf{r}}\right)_{I\cup\{k\}}\right)}{E\left(\displaystyle\sum_{\emptyset \neq I \subset \{1,\ldots,d\}} (-1)^{|I|-1} \Lambda_{(\mathbf{Y}_1)_I}\left(\dfrac{\boldsymbol{\tau}\,\mathbf{T}_1^{\boldsymbol{\beta}}}{\mathbf{r}}\right)_I\right)}, \qquad (14)$$

$$\theta_{X_j} = 1 - \frac{1}{r_j} E\left(\sum_{\emptyset \neq I \subset \{1,\ldots,k-1\}} (-1)^{|I|+1} \Lambda_{(Y_{1j},\ldots,Y_{kj})_{I\cup\{k\}}}\left(T_{1j}^{\beta_j},\ldots,T_{kj}^{\beta_j}\right)_{I\cup\{k\}}\right), \qquad (15)$$

Moreover, for all $\mathbf{x} \in \mathbb{R}_+^d$, we have

$$\lim_{n\to\infty} P\left(\mathbf{M}_n \leq \mathbf{U}_{\mathbf{X}}(n\mathbf{x})\right) = \exp\left\{-E\left(\sum_{\emptyset \subseteq I \subset \{1,\ldots,k-1\}} (-1)^{|I|} \Lambda_{(\mathbf{Y}_1,\ldots,\mathbf{Y}_k)_{I\cup\{k\}}}\left(\frac{\mathbf{T}_1^{\boldsymbol{\beta}}}{\mathbf{r}\mathbf{x}},\ldots,\frac{\mathbf{T}_k^{\boldsymbol{\beta}}}{\mathbf{r}\mathbf{x}}\right)_{I\cup\{k\}}\right)\right\}.$$

The result points out that, however much is the sequential dependence in the stationary sequence $\mathbf{T}$, the extremal index will be unit if $\mathbf{Y}_1, \ldots, \mathbf{Y}_k$ are tail independent. This characteristic is illustrated in Example 4 with multivariate pRARMAX models, which are adjusted to a bivariate financial series in the last section.

Example 2. Let $\{\mathbf{W}_n = (W_{n1}, \ldots, W_{nd})\}_{n \geq 1}$ be an i.i.d. sequence of random vectors with independent marginals W_{nj}, $j = 1 \ldots, d$, having distribution function (d.f.) $F_{W_{nj}}(x) = (1 - x^{-\beta_j})^{1/2}$, and $Y_{nj} = W_{n+1,j} \vee W_{nj}$, $j = 1, \ldots, d$. $\mathbf{Y}$ is 2-dependent and thus we can apply the results on the calculation of the multivariate extremal index under the condition $\mathtt{D}^{(2)}\left(\mathbf{u}_n^{(\boldsymbol{\tau})}\right)$. We have $\Lambda_{(\mathbf{Y}_1)_I}(\mathbf{x}_I) = 0$ if $|I| > 1$ and

$$\Lambda_{(\mathbf{Y}_1, \mathbf{Y}_2)}(\mathbf{x}_1, \mathbf{x}_2) = \lim_{t \to \infty} tP\left(\mathbf{Y}_1 \not\leq \mathbf{U}_{\mathbf{Y}_1}\left(\frac{t}{\mathbf{x}_1}\right), \mathbf{Y}_2 \not\leq \mathbf{U}_{\mathbf{Y}_2}\left(\frac{t}{\mathbf{x}_2}\right)\right) = \sum_{j=1}^{d} \frac{1}{2}(x_{1j} \wedge x_{2j}). \qquad (16)$$

Then, by (14),

$$\theta_{\mathbf{X}}(\tau_1, \ldots, \tau_d) = 1 - \frac{E\left(\Lambda_{(\mathbf{Y}_1, \mathbf{Y}_2)}\left(\dfrac{\boldsymbol{\tau}\,\mathbf{T}_1^{\boldsymbol{\beta}}}{\mathbf{r}}, \dfrac{\boldsymbol{\tau}\,\mathbf{T}_2^{\boldsymbol{\beta}}}{\mathbf{r}}\right)\right)}{\displaystyle\sum_{j=1}^{d} E\left(\dfrac{\tau_j\,T_{1j}^{\beta_j}}{r_j}\right)} = 1 - \frac{\dfrac{1}{2}\displaystyle\sum_{j=1}^{d} E\left(\dfrac{\tau_j\,T_{1j}^{\beta_j}}{r_j} \wedge \dfrac{\tau_j\,T_{2j}^{\beta_j}}{r_j}\right)}{\displaystyle\sum_{j=1}^{d} \tau_j}$$

and

$$\theta_j = 1 - \frac{1}{2r_j} E\left(T_{1j}^{\beta_j} \wedge T_{2j}^{\beta_j}\right).$$

Observe that the extremal indexes of the sequences $\{Y_{nj}\}_{n \geq 1}$, $j = 1, \ldots, d$, are all equal to $1/2$ while the presence of the random factors may increase this value for different θ_j, $j = 1, \ldots, d$. Therefore, each marginal sequence may have a different tendency for clustering of high values. We also find that, since the marginals of $\mathbf{Y}_n$ are independent, the dependence structure of $\mathbf{T}_n$ does not affect the clustering of high values of $\mathbf{X}_n$. As expected from (14) and illustrated in the examples of Section 5, only some tail dependence of $\mathbf{Y}_n$ allows to account the dependence of T_{nj}, $j = 1, \ldots, d$, on the value of $\theta_{\mathbf{X}}(\tau_1, \ldots, \tau_d)$. ▼

Example 3. Suppose that $\mathbf{Y}$ is 2-dependent and $\{\mathbf{T}_n\}_{n \geq 1}$ is a sequence of independent vectors with independent marginals having Bernoulli distribution with mean p. By (14), we have

$$\theta_{\mathbf{X}}(\tau_1, \ldots, \tau_d) = 1 - \frac{\displaystyle\sum_{\emptyset \neq I, J \subset \{1, \ldots, d\}} p^{|I|+|J|}(1-p)^{2d-|I|-|J|}\Lambda_{((\mathbf{Y}_1)_I, (\mathbf{Y}_2)_J)}\left(\left(\frac{\boldsymbol{\tau}}{\mathbf{r}}\right)_I, \left(\frac{\boldsymbol{\tau}}{\mathbf{r}}\right)_J\right)}{\displaystyle\sum_{\emptyset \neq J \subset \{1, \ldots, d\}} p^{|J|}(1-p)^{d-|J|}\left(-\log G_{\mathbf{Y}_J}\left(\left(\frac{\boldsymbol{\tau}}{\mathbf{r}}\right)_J\right)\right)}.$$

In the particular case of the previous example in (16), the function above becomes

$$\theta_{\mathbf{X}}(\tau_1, \ldots, \tau_d) = 1 - \frac{\dfrac{p^2}{2}\displaystyle\sum_{j=1}^{d}\frac{\tau_j}{r_j}}{\displaystyle\sum_{j=1}^{d}\tau_j}$$

and $\theta_j = 1 - \frac{p^2}{2r_j}$, $j = 1, \ldots, d$.

In the Example 5 ahead, we illustrate this choice of $\mathbf{T}_n$ to model lost values of $\mathbf{Y}_n$, considering that the marginals Y_{nj}, $j = 1, \ldots, d$, have total dependence. ▼

4 Tail dependence

The bivariate upper tail dependence of two random variables X_i and X_j can be measured through the tail dependence coefficient $\Lambda_{(X_i, X_j)}(1, 1)$ (Sibuya [32] 1960, Joe [18] 1997). Extending this concept to cross-sectional lag-m, s upper tail dependence, $m \geq 1$ and $0 \leq s < d$, for a sequence $\mathbf{X} = \{\mathbf{X}_n\}_{n \geq 1}$, we have

$$\lambda_X^{(m,s)} = \lim_{t \to \infty} tP\left(X_{1,1} > U_{X_1}(t), X_{1+m,1+s} > U_{X_{1+s}}(t)\right). \tag{17}$$

If $\lambda_X^{(m,s)} = 0$ we say that $X_{1,1}$ and $X_{1+m,1+s}$ are upper tail independent. In this case, it is possible that a residual tail dependence captured at penultimate high levels may occur. This is measured through the asymptotic tail independent coefficient $\eta \in (0, 1]$ (Ledford and Tawn [20, 21], 1996/97). Analogously, we can extend this concept to cross-sectional lag-m, s:

$$\lim_{t \to \infty} \frac{P\left(X_{1,1} > U_{X_1}(t/x), X_{1+m,1+s} > U_{X_{1+s}}(t/y)\right)}{P\left(X_{1,1} > U_{X_1}(t), X_{1+m,1+s} > U_{X_{1+s}}(t)\right)} = h_X^{(m,s)}(x, y), \tag{18}$$

for all $x, y \geq 0$, where $h_X^{(m,s)}$ is some non-degenerate function, homogeneous of order $-1/\eta_X^{(m,s)}$, $\eta_X^{(m,s)} \in (0, 1]$ and such that $h_X^{(m,s)}(1, 1) = 1$. In particular, we have

$$P\left(X_{1,1} > U_{X_1}(t), X_{1+m,1+s} > U_{X_{1+s}}(t)\right) = t^{-1/\eta_X^{(m,s)}} l_X^{(m,s)}(t), \tag{19}$$

where $l_X^{(m,s)}$ is a slowly varying function, standing for the relative strength of dependence given a particular value of $\eta_X^{(m,s)}$. Whenever $\eta_X^{(m,s)} = 1$ and $l_X^{(m,s)}$ converge to some constant $0 < a \leq 1$ then $X_{1,1}$ and $X_{1+m,1+s}$ are tail dependent ($\lambda_X^{(m,s)} = a$), otherwise we have asymptotic tail independence with positive association if $\eta_X^{(m,s)} > 1/2$, negative association if $\eta_X^{(m,s)} < 1/2$ and (almost) independence if $\eta_X^{(m,s)} = 1/2$ (perfect if $l_X^{(m,s)} = 1$).

By using the same arguments as in (6), we obtain the following relation for $\lambda_X^{(m,s)}$.

Proposition 4.1. *The lag-m, s upper tail dependence coefficient of $\mathbf{X}$ is given by*

$$\lambda_X^{(m,s)} = E\left(\Lambda_{(Y_{1,1}, Y_{1+m,1+s})}\left(\frac{T_{1,1}^{\beta_1}}{r_1}, \frac{T_{1+m,1+s}^{\beta_{1+s}}}{r_{1+s}}\right)\right), \tag{20}$$

where $\Lambda_{(Y_{1,1}, Y_{1+m,1+s})}(x, y) = \lim_{t \to \infty} tP\left(Y_{1,1} > U_{Y_1}(t/x), Y_{1+m,1+s} > U_{Y_{1+s}}(t/y)\right)$ is the lag-m, s upper tail copula of $\mathbf{Y}$.

Proposition 4.2. *If $(Y_{1,1}, Y_{1+m,1+s})$ satisfies (18) for some function $h_Y^{(m,s)}$ and coefficient $\eta_Y^{(m,s)}$, then*

$$P\left(X_{1,1} > U_{X_1}(t), X_{1+m,1+s} > U_{X_{1+s}}(t)\right) = E\left(h_Y^{(m,s)}\left(\frac{T_{1,1}^{\beta_1}}{r_1}, \frac{T_{1+m,1+s}^{\beta_{1+s}}}{r_{1+s}}\right)\right) t^{-1/\eta_Y^{(m,s)}} l_Y^{(m,s)}(t)(1 + o(1)),$$

for large t, provided that

$$E\left(\left(\frac{r_1}{T_1^{\beta_1}} \vee \frac{r_{1+s}}{T_{1+s}^{\beta_{1+s}}}\right)^{-1/\eta_Y^{(m,s)} + \delta} \vee \left(\frac{r_1}{T_1^{\beta_1}} \vee \frac{r_{1+s}}{T_{1+s}^{\beta_{1+s}}}\right)^{-1/\eta_Y^{(m,s)} - \delta}\right) < \infty. \tag{21}$$

Proof Just observe that, using (18), we have for large t

$$P\left(X_{1,1} > U_{X_1}(t), X_{1+m,1+s} > U_{X_{1+s}}(t)\right)$$

$$= \int P\left(Y_{1,1} > U_{Y_1}\left(\frac{r_1 t}{z_1^{\beta_1}}\right), Y_{1+m,1+s} > U_{Y_{1+s}}\left(\frac{r_{1+s} t}{z_2^{\beta_{1+s}}}\right)\right) dP_{(T_{1,1}, T_{1+m,1+s})}(z_1, z_2)$$

$$= P\left(Y_{1,1} > U_{Y_1}(t), Y_{1+m,1+s} > U_{Y_{1+s}}(t)\right) \int h_Y^{(m,s)}\left(\frac{z_1^{\beta_1}}{r_1}, \frac{z_2^{\beta_{1+s}}}{r_{1+s}}\right) dP_{(T_{1,1}, T_{1+m,1+s})}(z_1, z_2)(1 + o(1)),$$

where the last step is due to (21) and by applying the Potter bounds (Bingham *et al.*, [5] 1987; Theorem 1.5.6.) and the dominated convergence theorem. Now the result result is straightforward from (19). $\quad\square$

Observe that the lag-m, s upper tail dependence or independence of $\mathbf{X}$ is ruled by $\mathbf{Y}$. This is also easily seen by taking logarithm transformation of $\mathbf{X}$ when $\mathbf{T} > \mathbf{0}$. In the asymptotic tail independence case, this is even more evident since, by Proposition 4.2, they have the same lag-m, s asymptotic tail independent coefficient, i.e., $\eta_X^{(m,s)} = \eta_Y^{(m,s)}$, only differing in the relative strength of dependence coming from the slowly varying functions.

If we consider the models in Examples 2 and 3, we derive $\lambda_X^{(m,s)} = 0$, for all $m \geq 1$ and $1 \leq s < d$. Observe that this corresponds to perfect independence of $\mathbf{Y}$ and thus it satisfies (18) and (19), with $\eta_Y^{(m,s)} = 1/2$, $h_Y^{(m,s)}(x,y) = xy$ and $l_Y^{(m,s)}(t) = 1$. Therefore,

$$P\left(X_{1,1} > U_{X_1}(t), X_{1+m,1+s} > U_{X_{1+s}}(t)\right) = t^{-2},\tag{22}$$

i.e., perfect independence of $\mathbf{X}$ too.

5 Some particular cases that extend existing models

In this section we consider our model restricted to a common rescaled factor Y, i.e., by taking $Y_{nj} = Y_n^{\gamma_j}$, $\gamma_j > 0$, $j = 1, \ldots, d$, with $\mathbf{Y}^{(1)} = \{Y_n\}_{n \geq 1}$ a stationary sequence having common Pareto-type d.f. with shape parameter $\alpha > 0$, and $T_{nj} = Z_{nj}^{\gamma_j}$, $j = 1, \ldots, d$, where $\mathbf{Z} = \{(Z_{n1}, \ldots, Z_{nd})\}_{n \geq 1}$ is a stationary sequence, independent of $\mathbf{Y}^{(1)}$, with support $\mathbb{R}_+^d$ and such that, $r_j = E(Z_{nj}^\alpha) < \infty$, $j = 1, \ldots, d$. We will therefore apply the previous results with $\beta_j = \alpha/\gamma_j$, $j = 1, \ldots, d$ and $\Lambda_\mathbf{Y}(x_1, \ldots, x_d) = \bigwedge_{j=1}^d x_j$.

This case extends the heavy-tailed factor model in Lescourret and Robert ([22], 2006), which was used to measure the extreme dependence in loss severities of storm insurance data and where Y was considered a common latent factor corresponding to the intensity of the natural disaster. In addition, if we consider $\gamma = \gamma_j = 1$, $j = 1, \ldots, d$, we obtain an extended version of the scale mixture model of Li ([23], 2009) with several applications to real data (see, e.g., Arnold [3] 1983, Kotz *et al.* [19] 2000 and references therein).

As a consequence of Proposition 2.2, we obtain the following result for the upper tail copula of $\mathbf{X}$:

$$\Lambda_\mathbf{X}(x_1, \ldots, x_d) = E\left(\bigwedge_{j=1}^d \frac{x_j T_j^{\beta_j}}{r_j}\right) = E\left(\bigwedge_{j=1}^d \frac{x_j Z_j^\alpha}{r_j}\right),$$

for all $(x_1, ..., x_d) \in \overline{\mathbb{R}}_+^d$, where $(Z_1, \ldots, Z_d) \stackrel{d}{=} (Z_{n1}, \ldots, Z_{nd})$. From Proposition 2.4, we conclude that $F_{\mathbf{X}_1}$ is in the domain of attraction of

$$G(x_1, \ldots, x_d) = \exp\left\{-E\left(\bigvee_{j=1}^d \frac{x_j^{-1} Z_{nj}^\alpha}{r_j}\right)\right\}.\tag{23}$$

Remark 5.1. *If we assume that $\mathbf{Z}$ is an i.i.d. sequence of random vectors with i.i.d. marginals Z_{nj}, $j = 1 \ldots, d$, having common d.f. H and $r_j = r$, $j = 1, \ldots, d$, we easily derive*

$$E\left(\bigwedge_{j=1}^d \frac{x_j Z_{1j}^\alpha}{r_j}\right) = \frac{1}{r} \sum_{j=1}^d x_j \, E\left(\frac{Z^\alpha}{1-H(Z)} \prod_{i=1}^d \left[1 - H\left((x_j/x_i)^{1/\alpha} Z\right)\right]\right)$$

and

$$E\left(\bigvee_{j=1}^d \frac{x_j^{-1} Z_{1j}^\alpha}{r_j}\right) = \frac{1}{r} \sum_{j=1}^d x_j^{-1} \, E\left(\frac{Z^\alpha}{H(Z)} \prod_{i=1}^d H\left((x_i/x_j)^{1/\alpha} Z\right)\right),$$

extending the bivariate results in Lescourret and Robert ([22], 2006).

In the sequel, we rewrite Proposition 3.2, compute the extremal index and find the attractor MEV within this particular model.

In what concerns the local dependence condition of Proposition 3.2, observe that

$$\left\{ \mathbf{Y}_i^{(1)} \not\leq \mathbf{v}_n^{(\boldsymbol{\tau}^*(\mathbf{z}))} \right\} = \left\{ Y_i > \bigwedge_{j=1}^d \left(\frac{r_j n}{z_{ij}^\alpha \tau_j} \right)^{1/\alpha} \right\},$$

$$\left\{ \mathbf{Y}_i^{(1)} \leq \mathbf{v}_n^{(\boldsymbol{\tau}^*(\mathbf{z}))} \right\} = \left\{ Y_i \leq \bigwedge_{j=1}^d \left(\frac{r_j n}{z_{ij}^\alpha \tau_j} \right)^{1/\alpha} \right\}$$

and

$$a^\alpha \bigvee_{j=1}^d \frac{\tau_j}{r_j} < \bigvee_{j=1}^d \frac{z_{ij}^\alpha \tau_j}{r_j} < b^\alpha \bigvee_{j=1}^d \frac{\tau_j}{r_j}.$$

Therefore, we obtain for this model the following particular result.

Proposition 5.2. *If Z_{nj} has support in $[a, b]$, $j = 1, \ldots, d$, and $\mathbf{Y}^{(1)}$ satisfies condition $\mathrm{D}^{(k)}\left(v_n^{(\tau(a))}, v_n^{(\tau(b))} \right)$ defined by*

$$\lim_{n \to \infty} n \sum_{i=k}^{[n/k_n]} P\left(Y_1 > v_n^{(\tau(b))}, Y_i \leq v_n^{(\tau(a))}, Y_{i+1} > v_n^{(\tau(b))} \right) = 0,$$

with $\tau(a) = \bigvee_{j=1}^d \frac{a^\alpha \tau_j}{r_j} > \tau(b) = \bigvee_{j=1}^d \frac{b^\alpha \tau_j}{r_j}$, $\tau_j > 0$, $j = 1, \ldots, d$, then $\mathrm{D}^{(k)}\left(\mathbf{u}_n^{(\boldsymbol{\tau})} \right)$ holds for $\mathbf{X}$, for each $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_d)$.

Assuming that

$$\Lambda_{(Y_1, \ldots, Y_k)}(x_1, \ldots, x_k) = \lim_{t \to \infty} t P\left(Y_i > \bigcap_{i=1}^k \{ U_{Y_i} (t/x_i) \} \right), \tag{24}$$

exists with $(x_1, \ldots, x_k) \in \overline{\mathbb{R}}_+^k$, then

$$\lim_{n \to \infty} n P\left(\mathbf{X}_1 \leq \mathbf{u}_n^{(\boldsymbol{\tau})}, \ldots, \mathbf{X}_{k-1} \leq \mathbf{u}_n^{(\boldsymbol{\tau})}, \mathbf{X}_k \not\leq \mathbf{u}_n^{(\boldsymbol{\tau})} \right)$$

$$= E\left(\sum_{\emptyset \subseteq I \subset \{1, \ldots, k-1\}} (-1)^{|I|} \Lambda_{(Y_1, \ldots, Y_k)_{I \cup \{k\}}} \left(\bigvee_{j=1}^d \frac{\tau_j Z_{1j}^\alpha}{r_j}, \ldots, \bigvee_{j=1}^d \frac{\tau_j Z_{kj}^\alpha}{r_j} \right)_{I \cup \{k\}} \right),$$

with the convention that $\Lambda_{Y_k}\left(\bigvee_{j=1}^d \frac{\tau_j Z_{kj}^\alpha}{r_j} \right) = \bigvee_{j=1}^d \frac{\tau_j Z_{kj}^\alpha}{r_j}$. Therefore, if $\mathbf{X}$ satisfies strong-mixing and $\mathrm{D}^{(k)}\left(\mathbf{u}_n^{(\boldsymbol{\tau})} \right)$ conditions, for all $\boldsymbol{\tau} \in \mathbb{R}_+^d$, and if (24) holds for all $(x_1, \ldots, x_k) \in \overline{\mathbb{R}}_+^k$, then

$$\theta_{\mathbf{X}}(\tau_1, \ldots, \tau_d) = 1 - \frac{E\left(\sum_{\emptyset \neq I \subset \{1, \ldots, k-1\}} (-1)^{|I|+1} \Lambda_{(Y_1, \ldots, Y_k)_{I \cup \{k\}}} \left(\bigvee_{j=1}^d \frac{\tau_j Z_{1j}^\alpha}{r_j}, \ldots, \bigvee_{j=1}^d \frac{\tau_j Z_{kj}^\alpha}{r_j} \right)_{I \cup \{k\}} \right)}{E\left(\bigvee_{j=1}^d \frac{\tau_j Z_{1j}^\alpha}{r_j} \right)}, \tag{25}$$

the marginal extremal index is, for $j = 1, \ldots, d$,

$$\theta_{X_j} = 1 - \frac{1}{r_j} E \left(\sum_{\emptyset \neq I \subset \{1,\ldots,k-1\}} (-1)^{|I|+1} \Lambda_{(Y_1,\ldots,Y_k)_{I \cup \{k\}}} \left(Z_{1j}^\alpha, \ldots, Z_{kj}^\alpha \right)_{I \cup \{k\}} \right) \tag{26}$$

and, as $n \to \infty$, $P(\mathbf{M}_n \leq \mathbf{U_X}(n\mathbf{x}))$ converges to the limiting MEV

$$\exp \left\{ -E \left(\sum_{\emptyset \subseteq I \subset \{1,\ldots,k-1\}} (-1)^{|I|} \Lambda_{(Y_1,\ldots,Y_k)_{I \cup \{k\}}} \left(\bigvee_{j=1}^d \frac{Z_{1j}^\alpha}{x_j\, r_j}, \ldots, \bigvee_{j=1}^d \frac{Z_{kj}^\alpha}{x_j\, r_j} \right)_{I \cup \{k\}} \right) \right\},$$

for all $(x_1, \ldots, x_d) \in \mathbb{R}_+^d$.

The cross-sectional lag-m, s upper tail dependent coefficient, $m \geq 1$ and $0 \leq s < d$, is given by

$$\lambda_X^{(m,s)} = E \left(\Lambda_{(Y_1, Y_{1+m})} \left(\frac{Z_{1,1}^\alpha}{r_1}, \frac{Z_{1+m,1+s}^\alpha}{r_{1+s}} \right) \right), \tag{27}$$

In case $\mathbf{X}$ is cross-sectional lag-m, s upper tail independent (i.e., $\lambda_X^{(m,s)} = 0$) and if (Y_1, Y_{1+m}) satisfies (18) for some function $h_Y^{(m)}$ and asymptotic tail independent coefficient $\eta_Y^{(m)}$, for large t, then

$$P \left(X_{1,1} > U_{X_1}(t), X_{1+m,1+s} > U_{X_{1+s}}(t) \right) = E \left(h_Y^{(m)} \left(\frac{Z_{1,1}^\alpha}{r_1}, \frac{Z_{1+m,1+s}^\alpha}{r_{1+s}} \right) \right) t^{-1/\eta_Y^{(m)}} l_Y^{(m)}(t)(1 + o(1)). \tag{28}$$

Next, we consider some particular examples.

Example 4. *(Multivariate autoregressive processes with random coefficients)*
Suppose that $Y_n = c(Y_{n-1} \vee W_n)$, $n \geq 1$, where $0 < c < 1$ is a constant and $\{W_n\}_{n \geq 1}$ is an i.i.d. sequence, independent of Y_0, with common d.f. F_W. The max-autoregressive sequence $\mathbf{Y}^{(1)} = \{Y_n\}_{n \geq 1}$, usually denoted ARMAX, has stationary distribution F_Y if and only if $0 < \sum_{j=1}^\infty (1 - F_W(x/c^j)) < \infty$, for some $x > 0$ and, in this case, $F_Y(x) = F_Y(x/c)F_W(x/c)$ (Alpuim, [1] 1989). The Pareto distribution with parameter $\alpha > 0$ is a stationary d.f. of $\mathbf{Y}$ with $F_W(x) = (1 - (cx)^{-\alpha})/(1 - x^{-\alpha})$.

If Z_{nj} has support in $[a, b]$, $j = 1, \ldots, d$, such that $c < a/b$, then $\mathbf{Y}^{(1)}$ satisfies condition $\mathsf{D}^{(2)} \left(v_n^{(\tau(a))}, v_n^{(\tau(b))} \right)$, with $\tau(a)$ and $\tau(b)$ as in Proposition 5.2. Indeed, we have

$$n \sum_{i=2}^{[n/k_n]} P(Y_1 > v_n^{(\tau(b))}, Y_i \leq v_n^{(\tau(a))}, Y_{i+1} > v_n^{(\tau(b))})$$

$$\leq n \sum_{i=2}^{[n/k_n]} P(Y_1 > v_n^{(\tau(b))}, Y_i \leq v_n^{(\tau(a))}, Y_i > v_n^{(\tau(b))}/c)$$

$$+ n \sum_{i=2}^{[n/k_n]} P(Y_1 > v_n^{(\tau(b))}, Y_i \leq v_n^{(\tau(a))}, W_{i+1} > v_n^{(\tau(b))}/c).$$

The first term is null since $c < a/b \Leftrightarrow v_n^{(\tau(a))} < u_n^{(\tau(b))}/c$ and the second term is upper-bounded, successively, by

$$n \frac{n}{k_n} P(Y_1 > v_n^{(\tau(b))}) P(W_{i+1} > v_n^{(\tau(b))}) = n \frac{n}{k_n} (1 - F_Y(v_n^{(\tau(b))})) \left(1 - \frac{F_{Y_1}(v_n^{(\tau(b))})}{F_{Y_1}(v_n^{(\tau(b))}/c)} \right)$$

$$\leq \frac{n^2}{k_n} (1 - F_Y(v_n^{(\tau(b))}))^2 \to 0$$

for any $k_n \to \infty$, as $n \to \infty$. Therefore, if the ARMAX sequence $\mathbf{Y}$ is such that $c < a/b$ and $\mathbf{Z}$ has support in $[a,b]^d$, then $\mathbf{X}$ satisfies $\mathsf{D}^{(2)}\left(\mathbf{u}_n^{(\boldsymbol{\tau})}\right)$, for all $\boldsymbol{\tau} \in \mathbb{R}_+^d$ (Proposition 5.2).

From Ferreira and Canto e Castro ([12], 2008; Proposition 3.6), we obtain the following bivariate upper tail copula function, for $(x,y) \in \overline{\mathbb{R}}_+^d$:

$$\Lambda_{(Y_1, Y_{1+m})}(x,y) = x \wedge y c^{m\alpha}$$

Now we assume that $\mathbf{Z}$ is an i.i.d. sequence of random vectors with i.i.d. marginals Z_{nj}, $j = 1\ldots,d$, having common d.f. H and $r_j = r$, $j = 1,\ldots,d$. After some simple calculations, we derive

$$E\left(\bigvee_{j=1}^d \frac{x_j Z_{1j}^\alpha}{r_j} \wedge c^\alpha \bigvee_{j=1}^d \frac{x_j Z_{2j}^\alpha}{r_j}\right)$$

$$= \quad \frac{1}{r}\sum_{j=1}^d x_j \, E\left(\frac{Z^\alpha}{H(Z)}\prod_{i=1}^d H\left((x_j/x_i)^{1/\alpha}\,Z\right)\left[1 - \prod_{i=1}^d H\left((x_j/x_i)^{1/\alpha}\,Z/c\right)\right]\right)$$

$$+\frac{1}{r}\sum_{j=1}^d x_j \, E\left(\frac{(cZ)^\alpha}{H(Z)}\prod_{i=1}^d H\left((x_j/x_i)^{1/\alpha}\,Z\right)\left[1 - \prod_{i=1}^d H\left((x_j/x_i)^{1/\alpha}\,Zc\right)\right]\right),$$

where $r = r_j$, $j = 1,\ldots,d$.

Therefore, $\mathbf{X}$ has multivariate extremal index given by

$$\theta_{\mathbf{X}}(\tau_1,\ldots,\tau_d) = 1 - \frac{\sum_{j=1}^d \tau_j \left\{ E\left(\frac{Z^\alpha}{H(Z)}\pi(Z)\left[1 - \pi(Z/c)\right]\right) + E\left(\frac{(cZ)^\alpha}{H(Z)}\pi(Z)\left[1 - \pi(Zc)\right]\right)\right\}}{\sum_{j=1}^d \tau_j \, E\left(\frac{Z^\alpha}{H(Z)}\pi(Z)\right)}$$

where $\pi(y) = \prod_{i=1}^d H\left((\tau_j/\tau_i)^{1/\alpha}\,y\right)$, and for each $j = 1,\ldots,d$,

$$\theta_{X_j} = 1 - \frac{E\left(Z^\alpha\left(1 - H(Z/c)\right) + (cZ)^\alpha\left(1 - H(Zc)\right)\right)}{r}.$$

In what concerns the limiting MEV of $\mathbf{X}$, for all $(x_1,\ldots,x_d) \in \mathbb{R}_+^d$, we have

$$\lim_{n\to\infty} P\left(\mathbf{M}_n \le \mathbf{U}_{\mathbf{X}}(n\mathbf{x})\right) = \exp\left\{-\frac{1}{r}\sum_{j=1}^d x_j^{-1}\left[E\left(\frac{Z^\alpha}{H(Z)}\pi^*(Z)\pi^*(Z/c)\right) - E\left(\frac{(cZ)^\alpha}{H(Z)}\pi^*(Z)\left[1 - \pi^*(Zc)\right]\right)\right]\right\}.$$

where $\pi^*(y) = \prod_{i=1}^d H\left((x_i/x_j)^{1/\alpha}\,y\right)$.

Analogously, we obtain the cross-sectional lag-m,s upper tail dependent coefficient given by

$$\lambda_X^{(m,s)} = \frac{1}{r}\,E\left(Z^\alpha H(c^m Z) + (c^m Z)^\alpha\left(1 - H(c^m Z)\right)\right).$$

Particular cases can be obtained by replacing H by any d.f. with support in $[a,b]$, such that $c < a/b$.

Now if we consider, $Y_n = Y_{n-1}^c \vee W_n$, $n \ge 1$, with $0 < c < 1$ and $\{W_n\}_{n\ge 1}$ an i.i.d. sequence with common d.f. F_W, independent of Y_0, we obtain a max-autoregressive sequence with an exponent transformation, denoted pARMAX (Ferreira and Canto e Castro, [12], 2008). A stationary distribution F_Y exists if and only if $0 < \sum_{j=1}^\infty (1 - F_W(x^{c^j})) < \infty$, for some $x > 0$, and is such that $F_Y(x) = F_Y(x^{1/c})F_W(x)$. This relation is satisfied by a Pareto(α) distribution with $F_W(x) = (1 - x^{-\alpha})/(1 - x^{-\alpha/c})$ (see Ferreira and Canto e Castro [13] 2010 and references therein). The pARMAX sequence $\mathbf{Y}^{(1)}$ has lag-m upper tail dependence function $\Lambda_{(Y_1, Y_{1+m})}(x,y) = 0$, for all $m \ge 1$, and thus it is an upper tail independent process. Therefore, in this case, the lag-m,s upper tail dependence function $\Lambda_{(X_{1,1}, X_{1+m,1+s})}(x,y)$ is also

null, for all $m \geq 1$ and $s = 1, \ldots, d$, the extremal index is unit and the limiting MEV is as in the i.i.d. case, i.e., it is given by (23). Moreover, since $\mathbf{Y}^{(1)}$ satisfies (18) with asymptotic tail independent coefficient $\eta_Y^{(m)} = 1/2 \vee c^m$ and $h_Y^{(m)}(x, y) = xy\mathbb{1}_{\{c^m \leq 1/2\}} + y^{1/c^m}\mathbb{1}_{\{c^m > 1/2\}}$, $x, y \geq 0$ (Ferreira and Canto e Castro, [12], 2008), by applying (28), we obtain

$$P\left(X_{1,1} > U_{X_1}(t), X_{1+m,1+s} > U_{X_{1+s}}(t)\right)$$

$$= \left(\mathbb{1}_{\{c^m \leq 1/2\}} + \frac{E\left(Z_{1+m,1+s}^{\alpha/c^m}\right)}{r_{1+s}^{1/c^m}}\mathbb{1}_{\{c^m > 1/2\}}\right) t^{-1/(1/2 \vee c^m)} l_Y^{(m)}(t)(1 + o(1)).$$

Observe that sequence $\mathbf{X}$ generated from an ARMAX recursion $\mathbf{Y}^{(1)}$ corresponds to a multivariate formulation of the RARMAX process introduced in Alpuim and Athayde ([2], 1990), with applications within reliability and various natural phenomena. If sequence $\mathbf{X}$ is generated from a pARMAX recursion $\mathbf{Y}$, we have a multivariate formulation of the pRARMAX process introduced in Ferreira and Canto e Castro ([13], 2010), used in the modeling of financial series. ▼

Example 5. *(Multivariate moving maxima processes with random coefficients)*
Consider $Y_n = W_{n+1} \vee W_n$, $n \geq 1$, where $W = \{W_n\}_{n \geq 1}$ is an i.i.d. sequence with common d.f. $F_W(x) = (1 - x^{-\alpha})^{1/2}$, $x \geq 1$. The sequence $\mathbf{Y}^{(1)}$ is 2-dependent and thus $\mathbf{X}$ satisfies condition $\mathsf{D}^{(2)}(\mathbf{u}_n^\tau)$, for all $\boldsymbol{\tau} \in \mathbb{R}_+^d$, as well as, $\Lambda_{(Y_1, Y_{1+m})}(x, y) = 0$, for $m \geq 2$. It is easily seen that

$$\Lambda_{(Y_1, Y_2)}(x, y) = \frac{1}{2}(x + y - x \vee y) = \frac{1}{2}(x \wedge y)$$

Analogously to the ARMAX example above, we assume that $\mathbf{Z}$ is an i.i.d. sequence of random vectors with i.i.d. marginals Z_{nj}, $j = 1 \ldots, d$, having common d.f. H and $r_j = r$, $j = 1, \ldots, d$, and a similar procedure lead us to

$$E\left(\frac{1}{2}\left(\bigvee_{j=1}^d \frac{x_j Z_{1j}^\alpha}{r_j} \wedge \bigvee_{j=1}^d \frac{x_j Z_{2j}^\alpha}{r_j}\right)\right)$$

$$= \frac{1}{r}\sum_{j=1}^d x_j E\left(\frac{Z^\alpha}{H(Z)}\prod_{i=1}^d H\left((x_j/x_i)^{1/\alpha} Z\right)\left[1 - \prod_{i=1}^d H\left((x_j/x_i)^{1/\alpha} Z\right)\right]\right),$$

where $r = r_j$, $j = 1, \ldots, d$. Hence, we derive the multivariate extremal index

$$\theta_{\mathbf{X}}(\tau_1, \ldots, \tau_d) = 1 - \frac{\sum_{j=1}^d \tau_j E\left(\frac{Z^\alpha}{H(Z)}\prod_{i=1}^d H\left((\tau_j/\tau_i)^{1/\alpha} Z\right)\left[1 - \prod_{i=1}^d H\left((\tau_j/\tau_i)^{1/\alpha} Z\right)\right]\right)}{\sum_{j=1}^d \tau_j E\left(\frac{Z^\alpha}{H(Z)}\prod_{i=1}^d H\left((\tau_j/\tau_i)^{1/\alpha} Z\right)\right)},$$

the univariate extremal index, for $j = 1, \ldots, d$,

$$\theta_{X_j} = 1 - \frac{E\left(Z^\alpha[1 - H(Z)]\right)}{r},$$

and, for the limiting MEV, with $(x_1, \ldots, x_d) \in \mathbb{R}_+^d$,

$$\lim_{n \to \infty} P\left(\mathbf{M}_n \leq \mathbf{U}_{\mathbf{X}}(n\mathbf{x})\right) = \exp\left\{-\frac{1}{r}\sum_{j=1}^d x_j^{-1} E\left(\frac{Z^\alpha}{H(Z)}\left[\prod_{i=1}^d H\left((x_i/x_j)^{1/\alpha} Z\right)\right]^2\right)\right\}.$$

The cross-sectional lag-$1, s$ upper tail dependent coefficient is

$$\lambda_X^{(1,s)} = \frac{1}{2} = \lambda_Y^{(1)}.$$

The case $m \geq 2$ corresponding to perfect independence of $\mathbf{Y}^{(1)}$ has been analyzed in (22).

Observe that each marginal $X_{n,j}$ is a moving maxima process with a random coefficient $Z_{n,j}$, $j = 1, \ldots, d$. Therefore $\mathbf{X}$ corresponds to a multivariate formulation of this latter. For applications of multivariate moving maxima processes, see Zhang ([38], 2009) and references therein.

We are going to consider some particular cases of H.

- (Multivariate Gumbel case) Suppose that H is a Fréchet(δ, ξ) distribution, $\delta, \xi > 0$, i.e., $H(x) = \exp(-\delta\, x^{-\xi})$, $x > 0$, with $\xi > \alpha$. Therefore, $P(Z^\alpha \leq x) = \exp(-\delta\, x^{-\xi/\alpha})$ and $E(Z^\alpha) = e^{\alpha/\xi}\Gamma(1 - \alpha/\xi) < \infty$, since $\xi/\alpha > 1$. We have

$$
E\left(\frac{Z^\alpha}{H(Z)} \prod_{i=1}^{d} H\left((x_i/x_j)^{1/\alpha}\, Z \right) \right) = \frac{1}{\displaystyle\sum_{i=1}^{d}(x_j/x_i)^{\xi/\alpha}} \left[\delta \sum_{i=1}^{d}(x_j/x_i)^{\xi/\alpha} \right]^{\alpha/\xi} \Gamma(1 - \alpha/\xi),
$$

as well as

$$
E\left(\frac{Z^\alpha}{H(Z)} \left[\prod_{i=1}^{d} H\left((x_i/x_j)^{1/\alpha}\, Z \right) \right]^2 \right) = \frac{1}{2\displaystyle\sum_{i=1}^{d}(x_j/x_i)^{\xi/\alpha}} \left[2\delta \sum_{i=1}^{d}(x_j/x_i)^{\xi/\alpha} \right]^{\alpha/\xi} \Gamma(1 - \alpha/\xi).
$$

Therefore,

$$
G(x_1, \ldots, x_d) = \exp\left\{ -\left(\sum_{i=1}^{d} x_i^{-\xi/\alpha} \right)^{\alpha/\xi} \right\},
$$

which is a multivariate Gumbel distribution, and

$$
\lim_{n\to\infty} P\left(\mathbf{M}_n \leq \mathbf{U}_\mathbf{X}(n\mathbf{x}) \right) = \exp\left\{ -2^{\alpha/\xi - 1}\left(\sum_{i=1}^{d} x_i^{-\xi/\alpha} \right)^{\alpha/\xi} \right\}.
$$

Thus the extremal index is constant and is given by

$$
\theta_\mathbf{X}(\tau_1, \ldots, \tau_d) = 2^{\alpha/\xi - 1} = \theta_j,\ j = 1\ldots, d.
$$

- (Thinning model) Suppose that H is a Bernoulli(p) distribution, $0 < p < 1$. Since

$$
E\left(\frac{Z^\alpha}{H(Z)} \prod_{i=1}^{d} H\left((x_i/x_j)^{1/\alpha}\, Z \right) \right) = p(1-p)^{\sum_{i=1}^{d} \mathbb{1}_{\{x_i < x_j\}}}
$$

and

$$
E\left(\frac{Z^\alpha}{H(Z)} \left[\prod_{i=1}^{d} H\left((x_i/x_j)^{1/\alpha}\, Z \right) \right]^2 \right) = p(1-p)^{2\sum_{i=1}^{d} \mathbb{1}_{\{x_i < x_j\}}},
$$

we have

$$
G(x_1, \ldots, x_d) = \exp\left\{ -\sum_{j=1}^{d} x_j^{-1}(1-p)^{\sum_{i=1}^{d} \mathbb{1}_{\{x_i < x_j\}}} \right\},
$$

as well as

$$\lim_{n \to \infty} P\left(\mathbf{M}_n \leq \mathbf{U_X}(n\mathbf{x})\right) = \exp\left\{-\sum_{j=1}^{d} x_j^{-1}(1-p)^{2\sum_{i=1}^{d} \mathbb{1}_{\{x_i < x_j\}}}\right\}.$$

The extremal index is given by the function

$$\theta_{\mathbf{X}}(\tau_1, \ldots, \tau_d) = 1 - \frac{\displaystyle\sum_{j=1}^{d} \tau_j(1-p)^{2\sum_{i=1}^{d} \mathbb{1}_{\{\tau_i > \tau_j\}}}}{\displaystyle\sum_{j=1}^{d} \tau_j(1-p)^{\sum_{i=1}^{d} \mathbb{1}_{\{\tau_i > \tau_j\}}}}$$

and $\theta_j = 1$, $j = 1, \ldots, d$. ▼

6   An application to financial data

The presented results besides give us an insight of the effect of the random factors on the extremal behavior of the model, suggest estimation methods whenever data for $\mathbf{Y}$ and $\mathbf{T}$ are available, through the estimation of the dependence in $\mathbf{Y}$. Do not having real data sets in these conditions, we have chosen to illustrate in this section the application of the multivariate pRARMAX model considered in Example 4 to a bivariate stock market index data. The univariate pRARMAX was introduced in Ferreira and Canto e Castro ([13], 2010) and an adjustment algorithm to real data was implemented and applied to the S&P500 index, for the period April of 1957 to December 1987. Here we also consider the Dow Jones index for the same time period and our analysis is based on the volatility of the stock indexes measured through the square of the log-returns defined by $R_i = \log(P_{i+1})/\log(P_i)$, $i = 1, \ldots, 7731$ ($n = 7731$), where P_i is the closing price of the index at the ith day. The values of R_i^{SP} and $\left(R_i^{\mathsf{SP}}\right)^2$, as well as, R_i^{DJ} and $\left(R_i^{\mathsf{DJ}}\right)^2$, corresponding, respectively, to the log-returns and volatility of the S&P500 index and of the Dow Jones index, are plotted in Figure 1. Observe that they have similar plots. The large peak corresponds to "Black Monday" stock market crash on the 19th October 1987. In Ferreira and Canto e Castro ([13], 2010) was considered a robust regression so that the marginals can be modeled through a Pareto distribution, and thus conducted the analysis on the transformed data $X^{SP} = a\left(R^{SP}\right)^2 + b$, with $a = 13618.3$ and $b = 1.1$. Following the same procedure, we consider $X^{DJ} = a^*\left(R^{DJ}\right)^2 + b^*$, with a^* and b^* estimated in 15154.2 and 1.1, respectively.

An adjustment algorithm for a pRARMAX process, $X_n = Y_n Z_n$, $n \geq 1$, obtained by considering $Y_n = Y_{n-1}^c \vee W_n$ with $0 < c < 1$ and $F_Y(x) = 1 - x^{-\beta}$, where $\{W_n\}_{n \geq 1}$ is an i.i.d. sequence independent of Y_0 and having distribution $F_W(x) = (1 - x^{-\beta})/(1 - x^{-\beta/c})$, and $\{Z_n\}_{n \geq 1}$ is a sequence of standard uniform random variables, was proposed in Ferreira and Canto e Castro ([13] 2010, Section 3). This was used in this latter reference to model the X^{SP} data producing the estimates $\widehat{c} = 0.8$ (although the values $\widehat{c} = 0.85$, 0.75 were also considered) and $\widehat{\beta} = 2$. Now we are going to apply this adjustment algorithm to the X^{DJ} data. In the following we summarize the running steps:

1. Test if X^{DJ} is in the Fréchet domain of attraction (see, e.g., Dietrich *et al.* [9] 2002) and estimate the tail index corresponding to β^{-1} using, e.g., the Hill estimator (Hill, [17] 1975) and moments estimator (Dekkers *et al.*, [8] 1989).

2. Estimate the parameter c, through the estimation of the tail index of $S = \{S_i^{(n)}\}$, $i = 1, \ldots, n-1$, where

$$S_i^{(n)} = \left(\frac{n+1}{n+1-K_n(X_i^{DJ})}\right) \wedge \left(\frac{n+1}{n+1-K_n(X_{i+1}^{DJ})}\right),$$

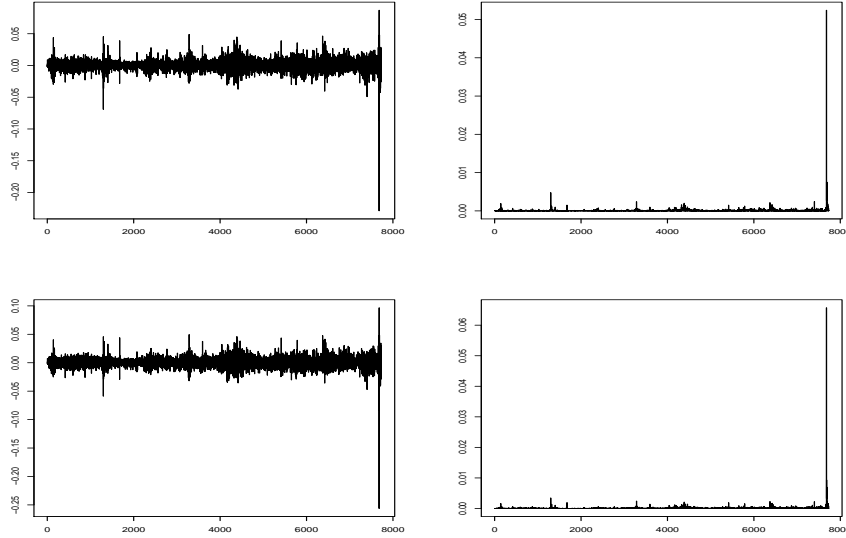with $K_n(x) = n^{-1}\sum_{i=1}^{n} \mathbb{1}_{\{X_i \leq x\}}$.

Figure 1: Left: log-returns (R_i) of the S&P500 (top) and Dow Jones indexes (bottom); Right: volatility (R_i^2) of the S&P500 (top) and Dow Jones indexes (bottom).

3. Apply the criterion: "if $X_i^{DJ} > \left(X_{i-1}^{DJ}\right)^{\widehat{c}}$, with the estimate $\widehat{c}$ obtained in step 2, then X_i^{DJ} comes from the innovation component W_i", to separate the innovations W and test if this sample is also in the Fréchet domain of attraction.

4. Capture the observations corresponding to Z, through the criterion: "if $X_i^{DJ} < \left(X_{i-1}^{DJ}\right)^{\widehat{c}}$ and $X_i^{DJ} \in \mathcal{B}_v = \left\{ t : \frac{\pi_0 t^{-\widehat{\beta}/\widehat{c}} f_W(t)}{\pi_0 t^{-\widehat{\beta}/\widehat{c}} f_W(t) + (1-\pi_0) t^{-\widehat{\beta}/\widehat{c})-1} F_W(t)} \leq v \right\}$, with f_W the density function of W, $\pi_1 = P(Z_i \left(X_{i-1}^{DJ}\right)^{\widehat{c}} > W_i)/P(\left(X_{i-1}^{DJ}\right)^{\widehat{c}} > W_i)$, $\pi_0 = 1 - \pi_1$, and where $\widehat{\beta}$ and $\widehat{c}$ are the estimates obtained in steps 1. and 2., respectively, then, $Z_i = X_i^{DJ}/\left(X_{i-1}^{DJ}\right)^{\widehat{c}}$"; v cannot be too large (larger errors) nor too small (not enough observations to carry out the test of the next step).

5. Test whether the sample of random variables Z captured in the step 4 has distribution $Beta(\widehat{\beta}/\widehat{c} + 1, 1)$ (e.g., Kolmogorov-Smirnov test).

For details, see Ferreira and Canto e Castro ([13] 2010) and references therein.

In Figure 2 it is plotted the sample paths (against the k upper order statistics) of the heavy tailed test statistic (left), the Hill estimator (center) and the moments estimator (right), referred in step 1. The heavy tailed domain of attraction is not rejected for $50 \lesssim k \lesssim 120$ and $200 \lesssim k \lesssim 1180$, which are plausible values within this text (see again Dietrich *et al.* [9] 2002). In what concerns the sample paths of Hill and moments, the value where the paths yield approximately a flat line is at about 0.57 and thus we consider $\widehat{\beta} = 1/0.55 \approx 1.82$.

The Exponential and Pareto qq-plots and the empirical mean excess function plots (Figure 3) indicate a Pareto-type model (see, e.g., Beirlant *et al.* [4] 2004).

In Figure 4 we have the sample paths of the Hill (left) and the moments (right) estimators concerning the sample S (step 2). The plot is somewhat "stable" at about 0.8 and thus we take $\widehat{c} = 0.8$.

Applying the criterium of step 3, we separate the innovations component W. The sample path of the heavy tailed test statistic (Figure 5, left) do not reject this assumption for $k \lesssim 780$, which is again a plausible result to conclude a Fréchet domain of attraction. From the Hill and moments plots (Figure 5, center and right, respectively) we do not discard the same estimate of $\widehat{\beta} \approx 1.82$.
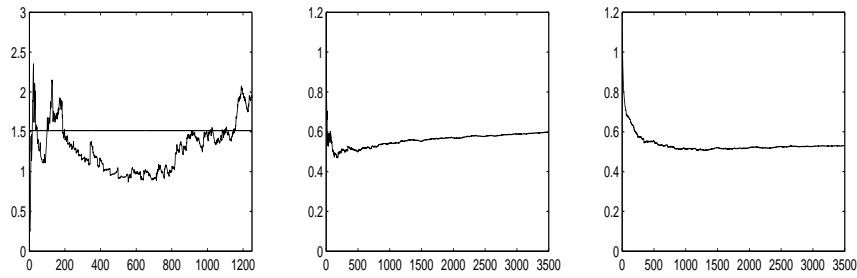
Figure 2: Left: sample path of the heavy tailed test statistic, against k upper order statistics, applied to the data X^{DJ} (horizontal line: critical value above which reject the null hypothesis of a heavy tailed domain of attraction); Center and right, respectively: sample paths of Hill and moments estimators, against k upper order statistics, of the data X^{DJ} concerning the estimation of the tail index β^{-1}.
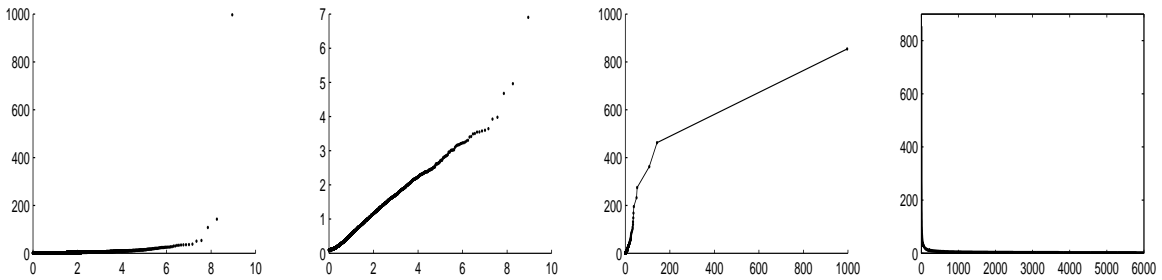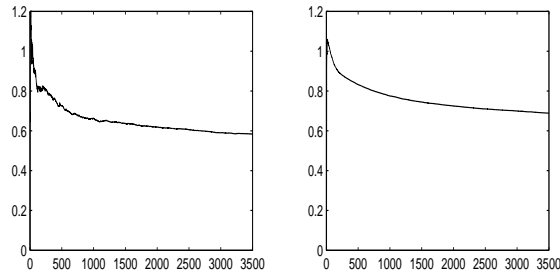

Figure 3: Left to right, respectively: exponential and Pareto qq-plots of X^{DJ}, empirical mean excess function against X^{DJ} and against k upper order statistics.

Figure 4: Left and right, respectively: sample paths of Hill and moments estimators, against k upper order statistics, of S concerning the estimation of c.

According to step 4, we capture the observations corresponding to random coefficients, Z, considering the significance regions $\mathcal{B}_v$ for $v = 0.05, 0.1, \ldots, 0.45$ and $\widehat{c} = 0.8$.

Implementing step 5, we applied the Kolmogorov-Smirnov test for the distribution $\texttt{Beta}(1.82/0.8+1, 1)$ to the sample of random variables Z captured in the step 4. The rejection is obtained for $v \geq 0.3$. By Figure 6, the value $v = 0.15$ seems to be an appropriate choice (with 21 observations captured) and is in agreement with the simulation study conducted in Ferreira and Canto e Castro ([13] 2010, Section 4.1). Therefore, we consider that the model pRARMAX with $c = 0.8$ and $\beta = 1.82$, can be used for the modeling of the transformed data X^{DJ}.

Now we are going to estimate the dependence of $(X_i^{\text{SP}}, X_i^{\text{DJ}})$, $i = 1, \ldots, 7731$, based on the dependence
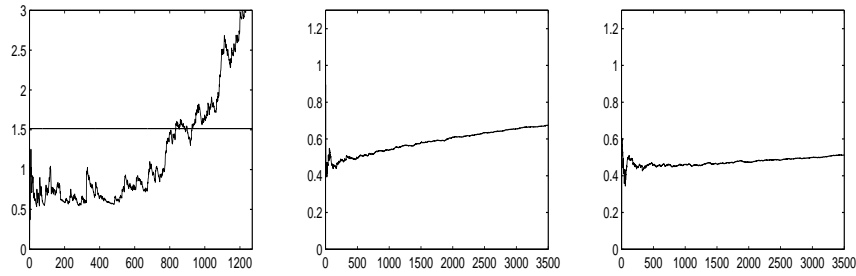
Figure 5: Left: sample path of the heavy tailed test statistic, against k upper order statistics, applied to the separated innovations W, according to step 3 (horizontal line: critical value above which reject the null hypothesis of a heavy tailed domain of attraction); Center and right, respectively: sample paths of Hill and moments estimators, against k upper order statistics, of W concerning the estimation of the tail index β^{-1}.
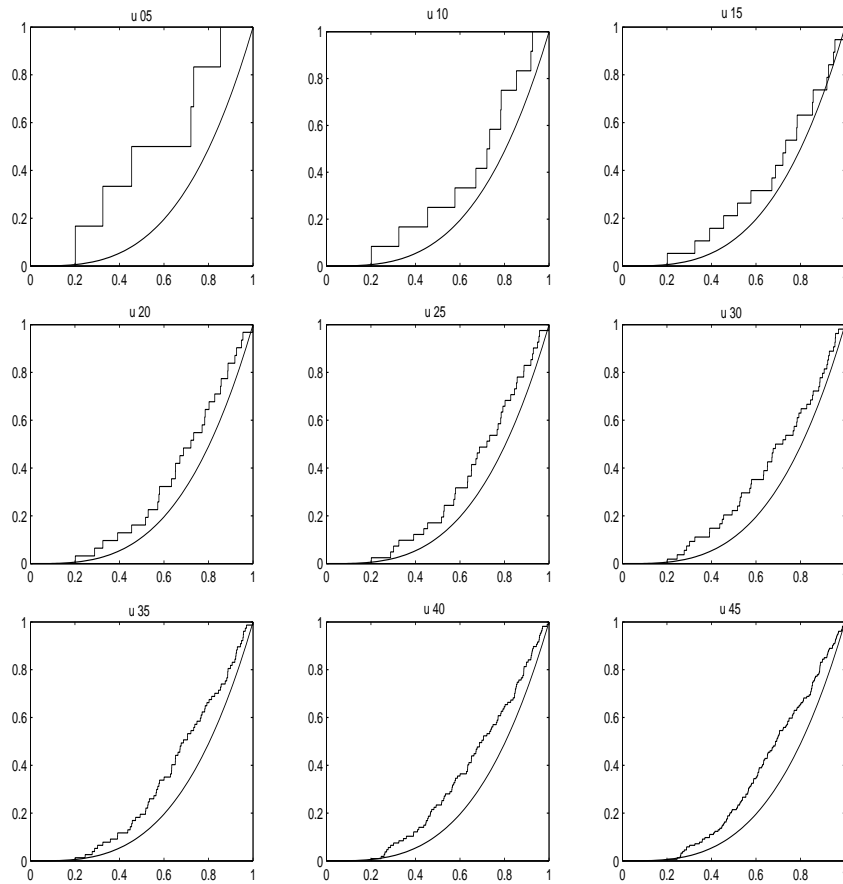

Figure 6: Empirical and theoretical distribution functions of the random coefficients Z captured on step 4, for significance regions with, $\upsilon = 0.05$ ("u 05"),..., $\upsilon = 0.45$ ("u 45").

function A (Pickands [30] 1981) of a GEV distribution G, defined by

$$G(x_1, x_2) = \exp\left\{-(x_1^{-1} + x_2^{-1})A(x_1(x_1 + x_2)^{-1})\right\}.$$

In order to have a dataset more coherent with an MEV model, we take block maxima within each

marginal. We use the Pickands ([30] 1981), the Capérà *et al.* ([6] 1997) and the Hall and Tajvidi ([16] 2004) nonparametric estimators for the function A, and thus free of model assumptions underlying the data. We also consider the Lescourret and Robert estimator ([22], 2006), given by

$$\widehat{A}(w) = \min\left(1, \max\left(\frac{\sum_{i=1}^{n} \max\left((1-w)\widetilde{X_i}^{\mathtt{SP}}, w\widetilde{X_i}^{\mathtt{DJ}}\right)}{\sum_{i=1}^{n}\left((1-w)\widetilde{X_i}^{\mathtt{SP}} + w\widetilde{X_i}^{\mathtt{DJ}}\right)}, w, 1-w\right)\right), \tag{29}$$

with

$$\left(\widetilde{X_i}^{\mathtt{SP}}, \widetilde{X_i}^{\mathtt{DJ}}\right) = \left(\left(\frac{X_i^{\mathtt{SP}}}{X_{n-k:n}^{\mathtt{SP}}}\right)^{\widehat{\beta}_{SP}}, \left(\frac{X_i^{\mathtt{DJ}}}{X_{n-k:n}^{\mathtt{DJ}}}\right)^{\widehat{\beta}_{DJ}}\right) \mathbb{1}_{\{X_i^{\mathtt{SP}} > X_{n-k:n}^{\mathtt{SP}}, X_i^{\mathtt{DJ}} > X_{n-k:n}^{\mathtt{DJ}}\}},$$

where $X_{1:n}^{\mathtt{j}} \leq \ldots \leq X_{n:n}^{\mathtt{j}}$, $j \in \{\mathtt{SP}, \mathtt{DJ}\}$, are the respective order statistics. We have considered $\widehat{A}(w) = (\widehat{A}(w) + \widehat{A}(1-w))/2$ in order to constrain the estimators to be symmetric. This estimator was developed for factor models with a common factor Y, as described in Section 5. The multivariate pRARMAX model (Example 4) was derived within this context. Observe in Figure 7 that the estimators are quite close, specially if we consider the bimonthly maxima, an indication that the multivariate pRARMAX may be a plausible model.
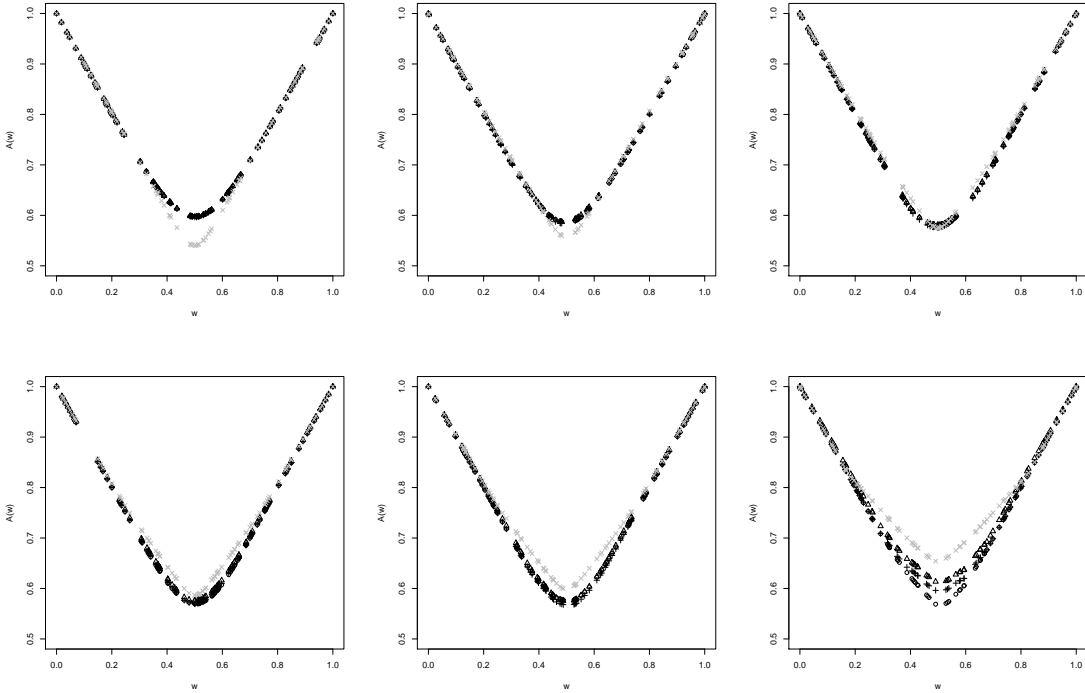

Figure 7: Estimation of the dependence function A in (29) for the bivariate data $\left(X_i^{\mathtt{SP}}, X_i^{\mathtt{DJ}}\right)$, $i = 1, \ldots, 7731$, using the non parametric estimators of Capéraà *et al.* [6] ($\circ$), Pickands [30] ($+$), Hall and Tajvidi [16] ($\triangle$) and using the estimator of Lescourret and Robert [22] ($\times$; grey), considering (top-to-bottom and left-to-right, respectively) the weekly, monthly, bimonthly, three monthly, quarterly and anual maxima.

7   Acknowledgements

References

[1] Alpuim, M.T. (1989). An extremal markovian sequence. Journal of Applied Probability 26 , 219–232.

[2] Alpuim, M. T, Athayde, E. (1990). On the stationary distribution of some extremal Markovian sequences, Journal of Applied Probability 27 (2), 291–302.

[3] Arnold, B.C. (1983). Pareto Distributions, International Cooperative Publishing House, Silver Spring, MD.

[4] Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J. (2004). Statistics of Extremes: Theory and Application. John Wiley, Chichester, UK.

[5] Bingham, N.H., Goldie, C.M., Teugels , J.L. (1987). Regular Variation, Cambridge Univ. Press.

[6] Capéraà , P., Fougères, A.L., Genest, C. (1997). A non-parametric estimation procedure for bivariate extreme value copulas. Biometrika 84, 567–577.

[7] Chernick, M.R., Hsing, T., McCormick W.P. (1991). Calculating the extremal index for a class of stationary sequences. Adv Appl Probab 23, 835–850.

[8] Dekkers, A.L.M., Einmahl, J.H.J., de Haan, L. (1989). A moment estimator for the index of an extreme value distribution. Ann. Statist. 17, 1833–1855.

[9] Dietrich, D., de Haan, L., Hüsler, J. (2002). Testing extreme value conditions. Extremes 5, 71–85.

[10] Doukhan, P. (1995). Mixing: Properties and Examples. Springer-Verlag, New York, USA.

[11] Ferreira, H. (1994). Multivariate extreme values in T-periodic random sequences under mild oscillation restrictions. Stochastic Proc. Appl. 49, 111–125.

[12] Ferreira, M., Canto e Castro, L. (2008). Tail and dependence behaviour of levels that persist for a fixed period of time, Extremes, 11(2), 113–133.

[13] Ferreira, M., Canto e Castro, L. (2010). Modeling rare events through a pRARMAX process. J. Statist. Plann. Inference, 140 11, 3552–3566.

[14] Ferreira, H., Ferreira, M. (2012). On extremal dependence of block vectors. Kybernetika 48(5), 988–1006.

[15] Ferreira, H., Ferreira, M. Extremes of multivariate ARMAX processes. (Submitted).

[16] Hall, P., Tajvidi, N. (2004). Prediction regions for bivariate extreme events. Daley Festschrift issue of the Australian New Zealand Journal of Statistics 46, 99–112.

[17] Hill, B.M. (1975). A simple general approach to inference about the tail of a distribution. Ann. Statist. 3, 1163–1174.

[18] Joe, H. (1997). Multivariate Models and Dependence Concepts. Chapman & Hall, London, UK.

[19] Kotz, S., Balakrishnan, N., Johnson, N.L. (2000). Continuous Multivariate Distributions, Wiley & Sons, New York, USA.

[20] Ledford, A., Tawn, J. (1996). Statistics for near independence in multivariate extreme values. Biometrika 83 , 169–187.

[21] Ledford, A., Tawn, J. (1997). Modelling Dependence within joint tail regions. J. R. Stat. Soc. Ser. B Stat. Methodol. 59, 475–499.

[22] Lescourret, L., Robert, C. (2006). Extreme dependence of multivariate catastrophic losses, Scandinavian Actuarial Journal, 2006(4), 203–225.

[23] Li, H. (2009). Orthant tail dependence of multivariate extreme value distributions, J. Multivariate Anal., 100(1), 243–256.

[24] Maulik, K., Resnick, S., Rootzén, H. (2002). Asymptotic independence and a network traffic model. Journal of Applied Probability 39(4), 671–699.

[25] Mladenovic, P., Piterbarg, V. (2006). On asymptotic distribution of maxima of complete and incomplete samples from stationary sequences. Stochastic Processes and Their Applications 116, 1977–1991.

[26] Nadarajah, S. (2006). Extreme value models with application to drought data. Allgemeines Statistisches Archiv 90(3), 403–418.

[27] Nadarajah, S. (2009). A bivariate pareto model for drought. Stochastic Environmental Research and Risk Assessment 23(6), 811–822.

[28] Nadarajah, S., Masoom, A.M. (2008). Pareto Random Variables for Hydrological Modeling. Water Resources Management 22(10), 1381–1393.

[29] Nandagopalan, S. (1990). Multivariate extremes and estimation of the extremal index, Ph.D. Dissertation, Department of Statistics, University of North Carolina, North Carolina, USA.

[30] Pickands, J. (1981). Multivariate extreme value distributions. Bulletin of the International Statistical Institute: Proceedings of the 43rd Session. Inst., 859–878.

[31] Resnick, S.I. (2007). Heavy Tail Phenomena: Probabilistic and Statistical Modeling. Springer, New York.

[32] Sibuya, M. (1960). Bivariate extreme statistics. Ann. Inst. Statist. Math. 11, 195–210.

[33] Smith, R.L., Weissman, I. (1996). Characterization and estimation of the multivariate extremal index. Manuscript, University of North Carolina, Chapel Hill, North Carolina, USA.

[34] Tan, Z., Wang, Y. (2012). Some asymptotic results on extremes of incomplete samples. Extremes 15, 319–332.

[35] Tawn, J. (1990). Modelling Multivariate Extreme Value Distributions. Biometrika 77(2), 245–253.

[36] Todorovic, P. and Gani, J. (1987). Modeling of the effect of erosion on crop production. Journal of Applied Probability, 24, 787–797.

[37] Weissman, I., Cohen, U. (1995). The extremal index and clustering of high values for derived stationary sequences, Journal of Applied Probability 32, 972–981.

[38] Zhang, Z. (2009). On approximating max-stable processes and constructing extremal copula functions. Stat Infer Stoch Process 12, 89–114.