Statistics and Its Interface Volume 8 (2015) 93–107

Copula structure analysis based on extreme dependence

Claudia Klüppelberg[*,†], Stephan Haug, and Gabriel Kuhn

We introduce a technique to analyse the dependence structure of an elliptical copula with focus on extreme observations. The classical assumption of a linear model for the distribution of a random vector is replaced by the weaker assumption of an elliptical copula in the high risk observations. More precisely, we describe the extreme dependence structure by an elliptical copula, which preserves a 'correlation-like' structure in the extremes. Based on the tail dependence function we estimate the extreme copula correlation matrix, which is then analysed through classical covariance structure analysis techniques. After introducing the new concepts we derive some theoretical results. A simulation study shows that the estimator performs very well even under the complexity of the extreme value problem. Finally, we use our method on real financial data assessing extreme risk dependence.

AMS 2000 subject classifications: Primary 62G32, 62H20, 62H25; secondary 60G70, 62H12.
Keywords and phrases: Extreme dependence, Extreme value statistics, Elliptical copula, Factor analysis, Kendall's tau, Multivariate statistics, Risk analysis, Structure analysis, Tail dependence function.

1. INTRODUCTION

Covariance or correlation structure analysis is a popular method in multivariate statistics to analyse the dependence in the data assuming a latent structure. Classical structure analysis is based on the assumption of normally distributed data, see e.g. [20] or the review paper [2]. Likelihood ratio tests were developed to distinguish between different model hypotheses. But the asymptotic χ^2-distribution is only valid for normally distributed data. However, many data sets exhibit properties contradicting the assumption of normality, see e.g. [5] for a study of financial data. Therefore, a number of extensions have been developed to deal with those kinds of features. On the one hand, scaled normal theory test statistics and estimators were introduced, e.g. in [22], [35] or [40], on the other hand so-called asymptotically distribution-free test statistics were developed in [3]. Concerning the first approach we also like to mention [37], who extended the normal theory methods for structure analysis to the class of elliptical distributions, by suitably scaling the test statistics with an estimator of the kurtosis parameter of the elliptical distribution.

All of the above approaches use in some way the covariance matrix, but for multivariate data it may happen that some margins are well modelled as being normal and some are much more heavy-tailed such that the existence of second moments is not guaranteed.

Motivated by such problems [24] introduced a correlation structure analysis, which does not assume the existence of any moments of the data. Their approach is based on analysing the 'correlation matrix' of an elliptical copula model describing the dependence in the data by Kendall's tau. Recall that this dependence measure is based on the ranks of the data, consequently, it disregards the absolute size of the data. The result is a *robust copula structure analysis*.

In many applications, however, dependence in extremes is a much more important issue than dependence in the mean of the data or its ranks as it is assessed by the classical correlation or by Kendall's tau, respectively. For example, financial risk management is confronted with problems concerning joint extreme losses, and one of its prominent questions is how to measure or understand dependence in the extremes; see e.g. [31]. This focus on dependence in extremes requires a different approach than in [24] and will be developed in this paper. We assess extreme dependence by the well-known concept of a tail dependence function. For such elliptical copulas, which can model extreme dependence, we present how to estimate a copula correlation matrix based on the tail dependence function. Given this estimate we illustrate, how to analyse the structure of the estimated correlation matrix. We will call this new method *extreme copula structure analysis*.

Our paper is organised as follows. We start with a short review of classical factor analysis and afterwards give some definitions and preliminary results on elliptical distributions and elliptical copulas in Section 2. Section 3 introduces the tail dependence function as a copula dependence concept and estimators are developed, which can be used for an extreme copula structure analysis. We also derive asymptotic results like asymptotic normality of our estimators. In Section 4 a simulation study shows that the derived asymptotic

*Corresponding author.
†We thank Jeffrey Dißmann for providing the data, which we analysed in Section 5.

results hold already for a moderate sample size. Finally, we perform a factor analysis based on the copula correlation matrix estimate of a real life data set and give an interpretation of the results. The longer proof of our main result is postponed to Section 6.

Throughout this paper we shall use the following notation. We denote the set of real $d \times m$ matrices by $M_{d,m}(\mathbb{R})$. If $d = m$ we simply write $M_d(\mathbb{R})$. The space of symmetric matrices is denoted by $\mathbb{S}_d$, the positive semi-definite cone by $\mathbb{S}_d^+$ and the positive definite cone by $\mathbb{S}_d^{++}$. The symbol $\mathcal{I}_d$ stands for the $d \times d$ identity matrix and $\det(\mathcal{A})$ for the determinant of a matrix $\mathcal{A} \in M_d(\mathbb{R})$. The transposed of a matrix $\mathcal{A} \in M_d(\mathbb{R})$ will be denoted by $\mathcal{A}^\top$. Moreover, $\operatorname{vecp} : M_{k,l}(\mathbb{R}) \to \mathbb{R}^u$ stands for the operator that stacks the u non-duplicated and non-fixed elements of a patterned matrix below another. For example, in case of a correlation matrix $\mathcal{R}$ we get

$$\text{(1)} \qquad \mathbf{r} \;:=\; \operatorname{vecp}(\mathcal{R}) \in \mathbb{R}^{d(d-1)/2}.$$

Finally, we abbreviate $\overline{\mathbb{R}}_+^d := [0,\infty]^d \backslash \{(\infty,\dots,\infty)\}$, and for $\mathbf{x}, \mathbf{y} \in \overline{\mathbb{R}}_+^d$ we denote by $\mathbf{x} \vee \mathbf{y}$ the componentwise maximum and by $\mathbf{x} \wedge \mathbf{y}$ the componentwise minimum.

2. PRELIMINARIES

2.1 Structure analysis based on the correlation matrix

Classical structure analysis techniques such as factor analysis can be based on the covariance $\mathbb{C}\mathrm{ov}(\mathbf{X}) = \mathbf{\Sigma}$ or the correlation matrix $\mathbb{C}\mathrm{orr}(\mathbf{X}) = \mathcal{R}$ of a random vector $\mathbf{X} \in \mathbb{R}^d$. In the first case the results depend on the scale of $\mathbf{X}$ and, thus, often the correlation matrix is used. We will later on also work with a correlation like dependence measure, the 'copula correlation matrix'.

In classical factor analysis the data $\mathbf{X}$ is assumed to satisfy a linear model $\mathbf{X} \stackrel{d}{=} \boldsymbol{\mu} + \widetilde{\mathcal{L}}\mathbf{f} + \widetilde{\mathcal{V}}\mathbf{e}$, where $\boldsymbol{\mu} = (\mu_1,\dots,\mu_d)^\top$, $\mathbf{f} = (f_1,\dots,f_m)^\top$ $(m < d)$ are non-observable and (usually) uncorrelated factors, $\mathbf{e} = (e_1,\dots,e_d)^\top$ is a noise vector, and $\stackrel{d}{=}$ means equality in distribution. Further, $\widetilde{\mathcal{L}} \in M_{d,m}(\mathbb{R})$ is called loading matrix and $\widetilde{\mathcal{V}}$ is a diagonal matrix with non-negative entries, the specific factor loadings. An often used additional assumption is that $(\mathbf{f}^\top, \mathbf{e}^\top)$ has mean zero and covariance matrix $\mathcal{I}_{m+d}$. Describing the dependence structure of $\mathbf{X}$ through its covariance matrix yields $\mathbf{\Sigma} = \widetilde{\mathcal{L}}\widetilde{\mathcal{L}}^\top + \widetilde{\mathcal{V}}^2$; i.e., the dependence of $\mathbf{X}$ is described through the entries of $\widetilde{\mathcal{L}}$. In terms of the correlation matrix $\mathcal{R}$ we get the following decomposition $\mathcal{R} = \mathcal{L}\mathcal{L}^\top + \mathcal{V}^2$, where $\mathcal{L} = \operatorname{diag}(\mathbf{\Sigma})^{-1/2}\widetilde{\mathcal{L}}$ and $\mathcal{V}^2 = \operatorname{diag}(\mathbf{\Sigma})^{-1/2}\widetilde{\mathcal{V}}^2\operatorname{diag}(\mathbf{\Sigma})^{-1/2}$.

2.2 Elliptical copulas

Elliptical copulas describe the dependence structure in elliptical distributions as well as in their extensions, the meta-elliptical distributions, which were originally introduced in [13]. We start by recalling the definition of an elliptical distribution and refer also to [12] for a comprehensive overview.

Definition 2.1. *A d-dimensional random vector $\mathbf{Z}$ is said to have an* elliptical distribution *with parameters $\boldsymbol{\mu} \in \mathbb{R}^d$ and $\mathbf{\Sigma} = (\sigma_{ij})_{1\le i,j\le d} \in \mathbb{S}_d^+(\mathbb{R})$, if it has the stochastic representation $\mathbf{Z} \stackrel{d}{=} \boldsymbol{\mu} + G\mathcal{A}\mathbf{U}^{(m)}$, where G is a positive random variable, $\mathbf{U}^{(m)} \sim unif(\mathbf{s} \in \mathbb{R}^m : \mathbf{s}^\top\mathbf{s} = 1)$ is independent of G, and $\mathcal{A} \in M_{d,m}(\mathbb{R})$ is a matrix such that $\mathcal{A}\mathcal{A}^\top = \mathbf{\Sigma}$ for some $m \in \mathbb{N}$. In particular, if G has a density, then the density of $\mathbf{Z}$ is of the form*

$$\det(\mathbf{\Sigma})^{-1/2} g\left((\mathbf{z}-\boldsymbol{\mu})^\top \mathbf{\Sigma}^{-1}(\mathbf{z}-\boldsymbol{\mu})\right),$$

where $g(\cdot)$ is a function uniquely determined by the distribution of the generating variable G. We shall use the notation $\mathbf{Z} \sim \mathcal{E}_d(\boldsymbol{\mu}, \mathbf{\Sigma}, G)$. Further, if the first moment exists, then $\mathbb{E}(\mathbf{Z}) = \boldsymbol{\mu}$ and, if the second moment exists, then G can be chosen such that $\mathbb{C}\mathrm{ov}(\mathbf{Z}) = \mathbf{\Sigma}$.

Up to a scaling factor all marginal distributions of a d-variate elliptical distribution are identical. We regain the flexibility of modelling the margins separately, while keeping the dependence structure of an elliptical distribution, by considering meta-elliptical distributions. The dependence structure in a meta-elliptical distribution is described by the corresponding elliptical copula, where a copula $C : [0,1]^d \to [0,1]$ is a d-dimensional distribution function with standard uniform margins, i.e. $C(1,\dots,1,u_j,1,\dots,1) = u_j$ for $j \in \{1,\dots,d\}$. For more technical background information on the copula concept we refer to [32].

Definition 2.2. *Let $\mathbf{Z} \sim \mathcal{E}_d(\boldsymbol{\mu}, \mathbf{\Sigma}, G)$ and define $\mathbf{Z}^* := \operatorname{diag}(\sigma_{11},\dots,\sigma_{dd})^{-1/2}(\mathbf{Z} - \boldsymbol{\mu}) \sim \mathcal{E}_d(\mathcal{R}, G)$ with $\mathcal{R} := (\sigma_{ij}/\sqrt{\sigma_{ii}\sigma_{jj}})_{1\le i,j\le d}$. Then we define the* elliptical copula *$\mathcal{EC}_d(\mathcal{R}, G)$ as the copula of $\mathbf{Z}^* \sim \mathcal{E}_d(,\mathcal{R}, G)$. We shall call $\mathcal{R}$ the* copula correlation matrix.

As a simple consequence of the definition and the fact that copulas are invariant with respect to $\boldsymbol{\mu}$ under strictly increasing transformations $\mathbf{Z}$ and $\mathbf{Z}^*$ have the same copula; see [11, Theorem 2.6]. Hence an elliptical copula is characterised by the generating variable G and the copula correlation matrix $\mathcal{R} =: (\rho_{ij})_{1\le i,j\le d}$.

Our assumptions are based on the following:

(A1) $\mathbf{X}_1,\dots,\mathbf{X}_n$ are i.i.d. with elliptical copula $\mathcal{EC}_d(\mathcal{R}, G)$.
(A2) $\rho_{ii} > 0$, for $i = 1,\dots,d$, and $|\rho_{ij}| < 1$, for $i \neq j$.
(A3) $\lim_{x\to\infty} \mathbb{P}(G > tx)/\mathbb{P}(G > x) = t^{-\nu}$ for all $t > 0$ and some $\nu > 0$. This means that G is regularly varying with index ν, denoted by $G \in RV_{-\nu}$.

3. METHODOLOGY

In this section we will introduce a copula-based dependence concept and the corresponding copula correlation matrix estimate. This estimate will describe the dependence structure in the extremes.

3.1 Dependence concepts

Measuring dependence by correlation or covariance is limited by the fact that they measure only linear dependence. Further, since copulas are invariant under strictly increasing transformations, correlation is not a copula parameter, but depends on the full distribution; see e.g. [11, Example 3.1]. On the other hand, for the proposed structure analysis method, we need a dependence concept, which can at least be linked to correlation.

Although the focus in this paper is on extreme dependence, we shall also need *Kendall's tau*, since we shall use it as a preliminary dependence estimate; for more details see [23].

Definition 3.1. *Kendall's tau τ_{ij} between two different components (X_i, X_j), $i \neq j$, of a random vector $\mathbf{X}$ is defined as*

$$\begin{aligned}\tau_{ij} \quad &:= \quad \mathbb{P}\left((X_i - \tilde{X}_i)(X_j - \tilde{X}_j) > 0\right) \\ &\quad -\mathbb{P}\left((X_i - \tilde{X}_i)(X_j - \tilde{X}_j) < 0\right),\end{aligned}$$

where $(\tilde{X}_i, \tilde{X}_j)$ is an independent copy of (X_i, X_j).

Concerning elliptical copulas the following result will be used, which is given in [13, Theorem 3.1].

Proposition 3.2. *Let $\mathbf{X}$ be a random vector with elliptical copula $\mathcal{EC}_d(\mathcal{R}, G)$ and generating variable $G > 0$. If* $\text{rank}(\mathcal{R}) = 1$ *and G is continuous or, if* $\text{rank}(\mathcal{R}) \geq 2$ *and* $\mathbb{P}(G = 0) = 0$, *then $\tau_{ij} = 2\arcsin(\rho_{ij})/\pi$.*

By Sklar's theorem, the copula C describes the dependence structure in a multivariate distribution model on all levels of the data. It also describes dependence in extremes. As C is a distribution on $[0, 1]^d$ with uniform marginals, extreme values happen near all boundaries and joint extreme dependence between all components happens around the points $(0, \ldots, 0)$ and $(1, \ldots, 1)$. As we will be interested in dependence of large risks, we concentrate on the dependence around the point $(1, \ldots, 1)$.

This can be captured by the following concept; see e.g. equation (1) in [26] or equation (2.3) in [19].

Definition 3.3. *Let $\mathbf{X}$ be a random vector with values in $\overline{\mathbb{R}}_+^d$ and marginal distribution functions F_j for $j = 1, \ldots, d$. We define the* (upper) tail dependence function *of $\mathbf{X}$ as*

(2)
$$\begin{aligned}T(\mathbf{x}) &:= \lim_{t\to 0} t^{-1}\mathbb{P}(1 - F_1(X_1) \leq tx_1, \ldots,\ 1-F_d(X_d) \leq tx_d) \\ &= \lim_{t\to 0} t^{-1}\overline{C}(1 - tx_1, \ldots, 1 - tx_d),\end{aligned}$$

for $\mathbf{x} = (x_1, \ldots, x_d) \in \overline{\mathbb{R}}_+^d$ if the limit exists, where $\overline{C}$ is the survival copula of C.

Remark 3.4. (i) If $T(\mathbf{x}) > 0$ for some $\mathbf{x} > \mathbf{0}$, $\mathbf{X}$ is called asymptotically dependent and asymptotically independent, otherwise. In [18, Theorem 4.3] it is shown that a random vector $\mathbf{X}$ with elliptical copula is asymptotically dependent, if and only if the corresponding generating variable $G \in RV_{-\nu}$ for some $\nu > 0$. By definition, $T(\mathbf{x}) = 0$, if $T^{(X_i,X_j)}(x_i, x_j) = 0$ for some i, j, i.e. $\mathbf{X}$ is asymptotically independent, if some bivariate margins (X_i, X_j) of $\mathbf{X}$ are asymptotically independent. Concerning asymptotic independence we refer to [28], and for a conditional modelling and estimation approach allowing for asymptotic independence in some components and asymptotic dependence in others; see [16]. We will use the assumption of asymptotic dependence for modelling and estimation and, therefore, we omit further discussions about asymptotic independence.

(ii) The bivariate marginal tail dependence function measures the amount of dependence in the upper right corner of the first quadrant of $\mathbb{R}^2$. Thus only positive dependence of (X_i, X_j) will be considered. As a consequence, if the estimated T is close to 0, the data may still be dependent or, for instance, negatively dependent. For an account of negative dependence in the extremes, one can move due to the symmetry of the elliptical copula from the pair (X_i, X_j) to $(X_i, -X_j)$ and interpret the findings appropriately.

From [26, Theorem 4] we know the tail dependence function $T(\mathbf{x})$ corresponding to an elliptical copula $\mathcal{EC}_d(\mathcal{R}, G)$. It is given by

$$T(\mathbf{x}) = \Big(\int\limits_{\mathbf{u}\in S_{d-1}, \mathcal{A}_{1\cdot}\mathbf{u}>0} (\mathcal{A}_{1\cdot}\mathbf{u})^\nu dF_U(\mathbf{u})\Big)^{-1} \int\limits_{\mathbf{u}\in S_{d-1}, \mathcal{A}_{1\cdot}\mathbf{u}>0,\ldots,\mathcal{A}_{d\cdot}\mathbf{u}>0} \bigwedge_{i=1}^{d} x_i(\mathcal{A}_{i\cdot}\mathbf{u})^\nu dF_U(\mathbf{u})\,,$$

where $\mathcal{A}_{i\cdot}$ is the i-th row of $\mathcal{A}$ from Definition 2.1 and F_U is the uniform distribution on the unit sphere $S_{d-1} = \{\mathbf{s} \in \mathbb{R}^d : \mathbf{s}^\top\mathbf{s} = 1\}$ in $\mathbb{R}^d$. For the estimation of $\mathcal{R}$ and ν with focus on extreme dependence we will need a one-to-one relation between the tail dependence function and $\mathcal{R}$ respectively ν. As shown in [26, Remark 5.2] the bivariate version of $T(\mathbf{x})$ reduces to a nice functional form, and depends only on ρ_{ij} and ν (cf. [25, Theorem 1]). We recall this representation in Proposition 3.5. The second equality of this Proposition is due to [8], who derived an expression for the Pickands dependence function $A(x) := 1 - T(x, 1 - x, \nu, \rho)$ of the bivariate t-distribution, and which was shown in [1] to be the same for all elliptical distributions with $G \in RV_{-\nu}$ for some $\nu > 0$.

Proposition 3.5. *Suppose $\mathbf{X}$ has elliptical copula $\mathcal{EC}_d(\mathcal{R}, G)$ and (A2)–(A3) hold. Then the bivariate marginal tail dependence function of $\mathbf{X}$ is given by*

(3)

$$\begin{aligned}
&T_{ij}(x,y)\\
&=\left(x\int_{g_{ij}\left((x/y)^{1/\nu}\right)}^{\pi/2}(\cos\phi)^{\nu}\,\mathrm{d}\phi+y\int_{g_{ij}\left((x/y)^{-1/\nu}\right)}^{\pi/2}(\cos\phi)^{\nu}\,\mathrm{d}\phi\right)\\
&\qquad\left(\int_{-\pi/2}^{\pi/2}(\cos\phi)^{\nu}\,d\phi\right)^{-1}\\
&=x\Big(1-t_{\nu+1}(a(x,y))\Big)+y\Big(1-t_{\nu+1}(a(y,x))\Big)\\
&=:T(x,y,\nu,\rho_{ij}),
\end{aligned}$$

where

$$a(x,y)=\left(\left(\frac{x}{y}\right)^{\frac{1}{\nu}}-\rho_{ij}\right)\sqrt{\frac{\nu+1}{1-\rho_{ij}^2}}$$

and x is the i-th and y the j-th component of $\mathbf{x}\in\mathbb{R}^d$.

Moreover, $g_{ij}(t):=\arctan((t-\rho_{ij})/\sqrt{1-\rho_{ij}^2})$ and $t_{\nu+1}$ denotes the t-distribution with $\nu+1$ degrees of freedom.

Remark 3.6. (i) The case of $\rho_{ij}=1$ can be interpreted as a limit, i.e.

$$T(x,y,\nu,1):=\lim_{\rho_{ij}\to 1}T(x,y,\nu,\rho_{ij}).$$

Then

$$g_{ij}(t)=\lim_{\rho_{ij}\to 1}\arctan\left(\frac{t-\rho_{ij}}{\sqrt{1-\rho_{ij}^2}}\right)=\begin{cases}+\pi/2, & t>1,\\ 0, & t=1,\\ -\pi/2, & t<1,\end{cases}$$

and we obtain $T(x,y,\nu,1)=x\wedge y$ for all $i\neq j\in\{1,\dots,d\}$. Similarly, $T(x,y,\nu,-1)=0$.

(ii) We want to recall that by [26, Theorem 5.1], T is also for arbitrary dimension d completely characterised by the copula correlation matrix $\mathcal{R}$ and the index ν of regular variation of G.

3.2 Extreme copula correlation estimator

From Proposition 3.5 we observe that, for an elliptical copula, T can be expressed as a function of $\mathcal{R}$ and ν. Vice versa, the correlation matrix $\mathcal{R}$ is a function of the tail dependence function and the index ν of regular variation of G. We will exploit this functional relationship for the estimation of $\mathcal{R}$. Using the tail dependence function for estimation of $\mathcal{R}$ focuses on the dependence structure in the upper extremes and does not necessarily model the dependence of the data in other regions in a realistic way. This is in contrast to the classical approach (based on the empirical correlation matrix). However, it allows us to assess the dependence in the extreme risks appropriately.

Given an estimator of T, we can estimate $\mathcal{R}$ and ν; i.e., we estimate the elliptical structure, which is likely to generate the observed extreme dependence. By Proposition 3.5, given an estimator of ν and of all bivariate marginal tail dependence functions, we can estimate the bivariate correlations, i.e. the correlation matrix $\mathcal{R}$. We start with an estimator of the tail dependence function.

Definition 3.7. *Given an i.i.d. sample $\mathbf{X}_1,\dots,\mathbf{X}_n$ with $\mathbf{X}_l=(X_{l,1},\dots,X_{l,d})^\top$ for $l=1,\dots,n$, we define the empirical tail dependence function for $\mathbf{x}=(x_1,\dots,x_d)>\mathbf{0}$ as*

(4) $$\mathbb{T}(\mathbf{x};k)=\frac{1}{k}\sum_{l=1}^{n}\mathbb{1}(1-\mathbb{F}_j(X_{l,j})\le\frac{k}{n}x_j,\ j=1,\dots,d),$$

where $1\le k\le n$, and $\mathbb{F}_j$ denotes the empirical distribution function of $\{X_{l,j}\}_{l=1}^n$ for $1\le j\le d$. Further, we define the empirical bivariate marginal tail dependence function as

(5) $$\begin{aligned}&\mathbb{T}_{ij}(x,y;k)\\&:=\frac{1}{k}\sum_{l=1}^{n}\mathbb{1}(1-\mathbb{F}_i(X_{l,i})\le\frac{k}{n}x\,,1-\mathbb{F}_j(X_{l,j})\le\frac{k}{n}y\,),\end{aligned}$$

where x is at the i-th and y at the j-th component of $\mathbf{x}$.

For details on empirical tail dependence functions see [9], [25], [36], and further references therein. Since $\mathbb{T}$ estimates a tail dependence function, the number k should be small compared to n. Setting $x_j=1$ for $1\le j\le d$ in (4), only the k largest observations of $X_{l,j}$ satisfy $1-\mathbb{F}_j(X_{l,j})\le k/n$, therefore, k can be interpreted as the number of the largest order statistics, which are used for the estimation as is typical in extreme value statistics.

Immediately from representation (3), it follows that $T(a\mathbf{x})=aT(\mathbf{x})$ for every $a>0$, i.e. T is homogeneous of order 1. Hence, for the estimation we follow the convention only to consider points $(x(\theta),y(\theta)):=(\sqrt{2}\cos(\theta),\sqrt{2}\sin(\theta))$ for $\theta\in(0,\pi/2)$, which includes the point $(1,1)$, but also points off the diagonal. By this procedure we obtain more information about the tail dependence of the data than by just considering the point $(1,1)$ on the diagonal.

Recall that our task is now to estimate ν and ρ from $\mathbb{T}$ for each pair of marginals $(X_{l,i},X_{l,j})$ for $1\le i,j\le d$. Obviously, it is not straightforward to estimate two parameters from one curve. We proceed as follows. For estimation of ν we use the approach of [26], which is based on inversion of the tail dependence function with respect to ν. To derive this estimator we need to replace the unknown ρ_{ij} in (3) by an appropriate initial estimator $\widehat{\rho}$. We will choose the estimator (cf. Proposition 3.2)

(6) $$\widehat{\rho}_{ij}^{\tau}:=\sin\left(\frac{\pi}{2}\widehat{\tau}_{ij}\right),$$

where $\widehat{\tau}_{ij}=\binom{n}{2}^{-1}\sum_{1\le l<k\le n}\operatorname{sgn}((X_{k,i}-X_{l,i})(X_{k,j}-X_{l,j}))$ is the empirical version of Kendall's tau τ_{ij}, cf. Definition 3.1. The convergence rate of $\widehat{\tau}_{ij}$, i.e. of $\widehat{\rho}_{ij}^{\tau}$, is $n^{-1/2}$, which is much faster than the convergence rate for any tail dependence function estimator, even when based on all n data points; see e.g. Theorem 5 in [36]. Thus the asymptotic behaviour of the tail index estimator $\widehat{\nu}$ is not changed,

if ρ_{ij} is replaced by $\widehat{\rho}^{\tau}_{ij}$ in the tail dependence function. We also want to recall that this estimator works regardless of the marginal models, which can be heavy- or light-tailed, and different in different components.

The following estimate has been suggested and its properties discussed in some detail in [26]. In contrast to estimators focusing only on the point $(1,1)$ it uses more data, giving high weights to realisations near the diagonal and lower weights away from the diagonal.

Definition 3.8. *Define $T^{\leftarrow\nu}(\cdot|\,x,y,\rho)$ as the inverse of $T(x,y,\nu,\rho)$ (given in (3)) with respect to ν. Using $\widehat{\rho}^{\tau}_{ij}$ estimated as in (6) and $\mathbb{T}_{ij}$ estimated as in (5), define for $i\neq j$*

$$\widehat{Q}_{ij} := \Big\{\theta\in\Big(0,\frac{\pi}{2}\Big) : \mathbb{T}_{ij}(x(\theta),y(\theta);k) < T\left(x(\theta),y(\theta),\left|\frac{\ln(\tan(\theta))}{\ln(\widehat{\rho}^{\tau}_{ij}\vee 0)}\right|,\widehat{\rho}^{\tau}_{ij}\right)\Big\},$$

$$\widehat{Q}^{*}_{ij} := \Big\{\theta\in\Big(0,\frac{\pi}{2}\Big) : |\ln(\tan(\theta))| < \Big(1-k^{-1/4}\Big)\,\tilde{\nu}_{ij}(1,1;k)\,\big|\ln(\widehat{\rho}^{\tau}_{ij}\vee 0)\big|\Big\}$$

$$Q^{*}_{ij} := \Big\{\theta\in\Big(0,\frac{\pi}{2}\Big) : |\ln(\tan(\theta))| < \nu\,|\ln(\rho_{ij}\vee 0)|\Big\},$$

where for $\theta\in\widehat{Q}_{ij}$ we define $\tilde{\nu}_{ij}$ as the estimator of ν based on the empirical bivariate tail dependence function (5)

$$\tilde{\nu}_{ij}(x(\theta),y(\theta);k) := T^{\leftarrow\nu}\left(\mathbb{T}_{ij}(x(\theta),y(\theta);k)|\,x(\theta),y(\theta),\widehat{\rho}^{\tau}_{ij}\right).$$

Further, let w be a non-negative weight function. Then we define the smoothed estimator $\widehat{\nu}$ of ν as

$$(7)\quad \widehat{\nu}(k,w) := \frac{1}{d(d-1)}\sum_{i\neq j}\frac{1}{W\big(\widehat{Q}_{ij}\cap\widehat{Q}^{*}_{ij}\big)}\int_{\theta\in\widehat{Q}_{ij}\cap\widehat{Q}^{*}_{ij}}\tilde{\nu}_{ij}(x(\theta),y(\theta);k)\,W(d\theta),$$

where W is the measure induced by w.

The asymptotic mean-squared error of $\widehat{\nu}(k,w)$ is given in [26, Corollary 1]. From Theorem 1 in [26] we know that for every fixed $x,y>0$ the tail dependence function $T(x,y,\nu,\rho_{ij})$ is strictly decreasing with respect to ν for all $\nu>|\ln(x/y)/\ln(\rho_{ij}\vee 0)|$. Thus the estimator $\tilde{\nu}_{ij}$ is well-defined.

We use now the estimate $\widehat{\nu}$ to define an estimator of the correlation matrix $\mathcal{R}$ via extreme observations. To this end we invert the bivariate tail dependence function with respect to ρ after having plugged in $\widehat{\nu}$. Using (3) it is straightforward to show the following.

Lemma 3.9. *For fixed $x,\,y,\,\nu>0$ and all $\rho\in[-1,1]$, the tail dependence function $T(x,y,\nu,\rho)$ is strictly increasing in ρ and the inverse $T^{\leftarrow\rho}(\cdot|\,x,y,\nu)$ of T with respect to ρ exists.*

By Remark 3.6 (i), for $\nu>0$ we have $T(1,1,\nu,1)=1$ and $T(1,1,\nu,-1)=0$. Hence, we can define

$$(8)\quad \tilde{\rho}_{ij}(1,1;\,k) := T^{\leftarrow\rho}\left(\mathbb{T}_{ij}(1,1;\,k)|\,1,1,\widehat{\nu}(k,w)\right).$$

Since this estimator only employs information at $(x,y)=(1,1)$, it may not be very efficient. Therefore, we define an estimator based on $\mathbb{T}_{ij}(x,y;\,k)$ for other values $(x(\theta),y(\theta))$ for $\theta\in(0,\frac{\pi}{2})$.

The following definition is an analogue of Definition 3.8. To ensure existence and consistency of the estimator, we define the appropriate sets.

Definition 3.10. *Define $T^{\leftarrow\rho}(\cdot|\,x,y,\nu)$ as the inverse of $T(x,y,\nu,\rho)$ (as given in (3)) with respect to ρ. Using $\widehat{\nu}$ estimated as in (7) and $\mathbb{T}_{ij}$ estimated as in (5), define for $i\neq j$*

$$\widehat{U}_{ij} := \Big\{\theta\in\Big(0,\frac{\pi}{2}\Big) : \mathbb{T}_{ij}\left(x(\theta),y(\theta);\,k\right) < T\left(x(\theta),y(\theta),\widehat{\nu}(k,w),e^{-|\ln(\tan\theta)|/\widehat{\nu}(k,w)}\right)\Big\},$$

$$\widehat{U}^{*}_{ij} := \Big\{\theta\in\Big(0,\frac{\pi}{2}\Big) : |\ln(\tan\theta)| < (1-k^{-1/4})\widehat{\nu}(k,w)\left|\ln\left(\tilde{\rho}_{ij}(1,1;\,k)\vee 0\right)\right|\Big\}$$

$$U^{*}_{ij} := Q^{*}_{ij}\,,$$

where for $\theta\in\widehat{U}_{ij}$ we define $\tilde{\rho}_{ij}$ as the estimator of ρ_{ij} based on the empirical bivariate tail dependence function (5)

$$(9)\quad \tilde{\rho}_{ij}(x(\theta),y(\theta);\,k) := T^{\leftarrow\rho}\left(\mathbb{T}_{ij}(x(\theta),y(\theta);\,k)|\,x(\theta),y(\theta),\widehat{\nu}(k,w)\right).$$

Observe that the set U^{*}_{ij} defines for given θ and ν the constraint

$$\rho_{ij}<\rho^{*} := \left((x_i(\theta)\wedge x_j(\theta))/(x_i(\theta)\vee x_j(\theta))\right)^{1/\nu}.$$

By Lemma 3.9 there exists a unique ρ such that

$$T\left(x(\theta),y(\theta),\widehat{\nu}(k,w),\rho\right)=\mathbb{T}_{ij}\left(x(\theta),y(\theta);\,k\right),\quad \theta\in\widehat{U}_{ij}.$$

This implies that the definition in (9) makes sense.

Note further that, by the definition of $\tilde{\rho}_{ij}(1,1;\,k)$ in (8), it always holds that $\pi/4\in\widehat{U}_{ij}$ provided that $\mathbb{T}_{ij}(1,1;\,k)<1$, and we also have $\pi/4\in\widehat{U}^{*}_{ij}$, since

$$(1-k^{-1/4})\widehat{\nu}(k,w)\left|\ln\left(\tilde{\rho}_{ij}(1,1;\,k)\vee 0\right)\right|>0.$$

To ensure consistency we further require $\theta\in\widehat{U}_{ij}$. This implies that the true ρ_{ij} is smaller than $e^{-|\ln(\tan(\theta))|/\widehat{\nu}(k,w)}$ with probability tending to one. The set U^{*}_{ij} is then the true subset of $(0,\pi/2)$, where Lemma 3.9 applies.

Now we can define an estimator for ρ_{ij} as a smooth version of $\tilde{\rho}_{ij}$:

Definition 3.11. *Let w^* be a non-negative weight function and W^* be the measure induced by w^*. Then we define for $i \neq j$ and with $\tilde{\rho}_{ij}$ as in (9)*

$$(10)\quad \widehat{\rho}_{ij}^{T}(k,w^*) := \frac{1}{W^*\big(\widehat{U}_{ij}\cap \widehat{U}_{ij}^*\big)} \int_{\theta\in \widehat{U}_{ij}\cap\widehat{U}_{ij}^*} \tilde{\rho}_{ij}(x(\theta),y(\theta);k)\,W^*(d\theta).$$

Further, we define $\widehat{\rho}_{ii}^{T}(k,w^*) := 1$ for $1 \leq i \leq d$, and $\widehat{\mathcal{R}}_T(k,w^*) := (\widehat{\rho}_{ij}^{T}(k,w^*))_{1\leq i,j\leq d}$. We call $\widehat{\mathcal{R}}_T$ the extreme copula correlation estimator.

The next theorem presents the asymptotic properties of $\widehat{\mathcal{R}}_T(k,w^*)$. To derive these properties we will use the theory developed in [36] about the limit behaviour of $\mathbb{T}_{ij}$ and give a formal proof in Section 6. In order to derive the asymptotic properties we need the following second order condition.

There exists $A(t)\to 0$ as $t\to 0$ such that

$$(11)\quad \lim_{t\to 0}\frac{t^{-1}\mathbb{P}(1-F_1(X_{1,1})\leq tx_1,\ldots,1-F_d(X_{1,d})\leq tx_d)-T(\mathbf{x})}{A(t)} = b(\mathbf{x})$$

holds locally uniformly for all $\mathbf{x}=(x_1,\ldots,x_d)$ in $\overline{\mathbb{R}}_+^d$, and b is some non-constant function.

Remark 3.12. The second order condition (11) holds provided the regularly varying distribution function of the generating random variable G satisfies such a second order condition. More precisely, it is required that there exists some function $\tilde{A}(t)\to 0$ such that for all $x>0$ and some $\beta\leq 0$

$$\lim_{t\to\infty}\frac{P(G>tx)/P(G>t)-x^{-\nu}}{\tilde{A}(t)} = x^{-\nu}\frac{x^{\beta}-1}{\beta}.$$

This entails the second order condition on the tail dependence function (11); cf. [25, Theorem 3].

Theorem 3.13. *Suppose (A1)–(A3) and (11) hold. Further assume that $k=k(n)\to\infty$, $k/n\to 0$ and $\sqrt{k}A(k/n)\to 0$ as $n\to\infty$. Let w^* be a non negative weight function satisfying $\sup_{\theta\in U_{ij}^*} w^*(\theta)<\infty$ for all $i\neq j$ and $\theta\in(0,\pi/2)$, and W^* is the measure induced by w^*. Define*

$$(12)\quad \begin{aligned}\tilde{B}_{ij}(x,y) &:= B_{ij}(x,y) + \\ &- B_{ij}(x,\infty)\frac{\partial}{\partial x}T(x,y,\nu,\rho_{ij}) - B_{ij}(\infty,y)\frac{\partial}{\partial y}T(x,y,\nu,\rho_{ij}),\\ B_{ij}(x,y) &:= B(\infty,\ldots,\infty,x,\infty,\ldots,\infty,y,\infty,\ldots,\infty),\end{aligned}$$

where x is the i-th and y the j-th component and B is a zero mean Wiener process on $\overline{\mathbb{R}}_+^d$ with covariance structure

$$(13)\qquad \mathbb{E}(B(\mathbf{x})B(\mathbf{y})) = T(\mathbf{x}\wedge\mathbf{y}),\quad \mathbf{x},\mathbf{y}\in\overline{\mathbb{R}}_+^d.$$

Set as in (1)

$$\mathbf{r} := \operatorname{vecp}(\mathcal{R}) \quad \text{and} \quad \widehat{\mathbf{r}}_T(k,w^*) := \operatorname{vecp}\left(\widehat{\mathcal{R}}_T(k,w^*)\right),$$

then

$$\sqrt{k}\left(\widehat{\mathbf{r}}_T(k,w^*)-\mathbf{r}\right)\xrightarrow{d}\mathcal{N}_{d(d-1)/2}(\mathbf{0},\mathbf{\Gamma}_T),\quad n\to\infty,$$

where $\mathbf{\Gamma}_T=(\gamma_{ij,kl}^T)_{1\leq i\neq j,k\neq l\leq d}$ with

$$(14)\qquad \gamma_{ij,kl}^T = \sigma_{1;ij,kl}+\sigma_{2;ij,kl}+\sigma_{3;ij,kl}+\sigma_{4;ij,kl}.$$

Setting $h(\theta,\nu,\rho):=T(x(\theta),y(\theta),\nu,\rho)$ we have

$$(15)\quad \begin{aligned}\sigma_{1;ij,kl} = {} & \frac{2}{d^2(d-1)^2W^*(U_{ij}^*)W^*(U_{kl}^*)}\times\prod_{J\in\{ij,kl\}} \\ & \int_{\theta\in U_J^*}\frac{\partial}{\partial\nu}T^{\leftarrow\rho}\left(T_J(x(\theta),y(\theta))\,|\,x(\theta),y(\theta),\nu\right)W^*(d\theta)\\ & \times\Bigg(\sum_{1\leq p<q,r<s\leq d}\frac{1}{W^*(Q_{pq}^*)W^*(Q_{rs}^*)}\int_{\theta_1\in Q_{pq}^*}\int_{\theta_2\in Q_{rs}^*}\\ & \times\frac{\mathbb{E}\left(\tilde{B}_{pq}(x(\theta_1),y(\theta_1))\tilde{B}_{rs}(x(\theta_2),y(\theta_2))\right)}{\frac{\partial}{\partial\nu}h(\theta_1,\nu,\rho_{pq})\frac{\partial}{\partial\nu}h(\theta_2,\nu,\rho_{rs})}\\ & W^*(d\theta_2)W^*(d\theta_1)\Bigg)\end{aligned}$$

and

$$(16)\quad \begin{aligned}\sigma_{2;ij,kl} = {} & \frac{1}{d(d-1)W^*(U_{ij}^*)W^*(U_{kl}^*)} \\ & \sum_{1\leq p<q\leq d}\frac{1}{W^*(Q_{pq}^*)}\Bigg(\int_{\theta_1\in U_{ij}^*}\int_{\theta_2\in U_{kl}^*}\int_{\theta_3\in Q_{pq}^*}\\ & \frac{\partial}{\partial\nu}T^{\leftarrow\rho}\left(T_{ij}(x(\theta_1),y(\theta_1))|x(\theta_1),y(\theta_1),\nu\right)\\ & \times\frac{\mathbb{E}\left(\tilde{B}_{pq}(x(\theta_3),y(\theta_3))\tilde{B}_{kl}(x(\theta_2),y(\theta_2))\right)}{\frac{\partial}{\partial\nu}h(\theta_3,\nu,\rho_{pq})\frac{\partial}{\partial\rho}h(\theta_2,\nu,\rho_{kl})}\\ & W^*(d\theta_3)W^*(d\theta_2)W^*(d\theta_1)\Bigg),\end{aligned}$$

similarly $\sigma_{3;ij,kl}$ (by interchanging the indices 'ij' and 'kl'), and

$$(17)\quad \begin{aligned}\sigma_{4;ij,kl} = {} & \frac{1}{2W^*(U_{ij}^*)W^*(U_{kl}^*)}\int_{\theta_1\in U_{ij}^*}\int_{\theta_2\in U_{kl}^*}\\ & \frac{\mathbb{E}\left(\tilde{B}_{ij}(x(\theta_1),y(\theta_1))\tilde{B}_{kl}(x(\theta_2),y(\theta_2))\right)}{\frac{\partial}{\partial\nu}h(\theta_1,\nu,\rho_{ij})\frac{\partial}{\partial\rho}h(\theta_2,\nu,\rho_{kl})}\\ & W^*(d\theta_2)W^*(d\theta_1).\end{aligned}$$

Using (14), we can define an estimator of $\mathbf{\Gamma}_T$.

Definition 3.14. *We define the estimator of*

$$\mathbf{\Gamma}_T=(\gamma_{ij,kl}^T)_{1\leq i\neq j,k\neq l\leq d} \quad \text{by} \quad \widehat{\mathbf{\Gamma}}_T=(\widehat{\gamma}_{ij,kl}^T)_{1\leq i\neq j,k\neq l\leq d}$$

with

$$(18)\qquad \widehat{\gamma}^T_{ij,kl} := \widehat{\sigma}_{1;ij,kl} + \widehat{\sigma}_{2;ij,kl} + \widehat{\sigma}_{3;ij,kl} + \widehat{\sigma}_{4;ij,kl}.$$

The $\widehat{\sigma}$ are defined in (15)–(17), where ν, ρ_{ij} and ρ_{kl} are replaced by their estimators $\widehat{\nu}(k,w)$, $\widehat{\rho}^T_{ij}(k,w^*)$ and $\widehat{\rho}^T_{kl}(k,w^*)$, respectively, the sets U^* and Q^* are replaced by their estimators $\widehat{U} \cap \widehat{U}^*$ and $\widehat{Q} \cap \widehat{Q}^*$, respectively, and the covariances $\mathbb{E}(\tilde{B}_{ij}(\cdot)\tilde{B}_{kl}(\cdot))$ are replaced by their estimators $\widehat{\mathbb{E}}(\tilde{B}_{ij}(\cdot)\tilde{B}_{kl}(\cdot))$ using (12) and (13) and estimating T by $\mathbb{T}$.

The asymptotic properties of $\mathbb{T}$, $\widehat{\nu}$, $\widehat{\rho}^T_{ij}$ in combination with the delta method yield immediately the following result.

Theorem 3.15. *Under the conditions of Theorem 3.13, the estimator* $\mathrm{vecp}(\widehat{\boldsymbol{\Gamma}}_T)$ *is consistent and asymptotically normal.*

Remark 3.16. It may happen that the correlation matrix estimators (6) or (10) are not positive semi-definite. In this case, we apply some of the methods described in [17] or [34] to project the indefinite correlation matrix to the set of positive semi-definite correlation matrices; see also [24] for details. Considering covariance matrix estimators, which are not positive semi-definite, we project them on $\mathbb{S}^+$ by replacing the negative eigenvalues of the covariance matrix estimator by their absolute values.

Estimation of dependence in extremes is always difficult. The problem of estimating tail dependence lies in its definition as a limit; see (2). For some methods and pitfalls of estimating the tail dependence function $T_{ij}(1,1)$ we refer to [14]. Estimators of the tail dependence are based on a sub-sample using the largest (or smallest) observations. Concerning the optimal choice of the threshold (equivalently the number k of upper order statistics used in the estimation), we refer to [6], [10], [25], [26] and [33]. In our applications we used a heuristical approach to select k, which will be explained in the next section.

4. COPULA STRUCTURE ANALYSIS: A FACTOR ANALYSIS EXAMPLE

In the previous section we have presented an extreme copula correlation estimate $\widehat{\mathcal{R}}_T$ for the copula correlation $\mathcal{R}$ of an elliptical copula, which is based on extreme dependence. Now we want to explain how the structure of this matrix can be analysed. Therefore, we will assume a model for the structure of $\mathcal{R}$ in the extremes. Throughout the rest of the paper we will assume a factor model; i.e., we assume the structure $\mathcal{R}(\boldsymbol{\vartheta}) = \mathcal{L}\mathcal{L}^\top + \mathcal{V}^2$, where $\mathcal{L} \in M_{d,m}(\mathbb{R})$, $\mathcal{V} \in \mathbb{S}^+_d(\mathbb{R})$ is a diagonal matrix, and $\boldsymbol{\vartheta} \in \Theta \subset \mathbb{R}^p$ are the free parameters in $\mathcal{L}$ and $\mathcal{V}$. But this choice is arbitrary. One can choose any parametric model which defines a correlation matrix. We will now use the asymptotic properties of the extreme copula correlation estimator $\widehat{\mathcal{R}}_T$ defined in (10) to derive testing procedures.

In case of the classical covariance structure analysis, there is a vast amount of literature on how to define a suitable test to decide if the true covariance is a member of the assumed model class; see e.g. [2], [3], [35] and [37]. We adapted the main ideas to our setting, that is to measure the discrepancy between the model $\mathcal{R}(\boldsymbol{\vartheta})$ and an estimate $\widehat{\mathcal{R}}$ of the copula correlation matrix. Therefore, we minimise a discrepancy function $F(\cdot,\cdot)$ with respect to $\boldsymbol{\vartheta} \in \Theta$:

$$\min_{\boldsymbol{\vartheta}\in\Theta} F(\widehat{\mathcal{R}}, \mathcal{R}(\boldsymbol{\vartheta}))\,.$$

A suitably scaled version of the discrepancy function should then be asymptotically pivotal; i.e., its asymptotic distribution should be independent of the unknown parameters in the model. There are different choices of discrepancy functions, but for most of them their corresponding asymptotic distribution depends on distributional assumptions about $\mathbf{X}$, like e.g. the normal maximum likelihood discrepancy function

$$\begin{aligned} F_{ML}(\widehat{\mathcal{R}}, \mathcal{R}(\boldsymbol{\vartheta})) \quad &= \quad \log\Big(\det(\mathcal{R}(\boldsymbol{\vartheta}))\Big) - \log\Big(\det(\widehat{\mathcal{R}})\Big) \\ &\quad + \mathrm{tr}\Big(\widehat{\mathcal{R}}\mathcal{R}(\boldsymbol{\vartheta})^{-1}\Big) - d\,. \end{aligned}$$

Since we do not make any assumptions about the marginal distributions of $\mathbf{X}$ we can not work with such a discrepancy function. But we can apply the asymptotic distribution-free method developed in [3]. There a quadratic discrepancy function

$$F(\widehat{\mathcal{R}}, \mathcal{R}(\boldsymbol{\vartheta})|\mathcal{U}) = (\widehat{\mathbf{r}} - \mathbf{r}(\boldsymbol{\vartheta}))^\top \mathcal{U}^{-1} (\widehat{\mathbf{r}} - \mathbf{r}(\boldsymbol{\vartheta}))\,,$$

where $\mathcal{U}$ is a suitably chosen weight matrix, is used to estimate the parameter $\boldsymbol{\vartheta}$. The estimator

$$\widehat{\boldsymbol{\vartheta}} := \mathrm{argmin}\, F(\widehat{\mathcal{R}}, \mathcal{R}(\boldsymbol{\vartheta})|\mathcal{U})$$

is asymptotically normal with mean $\boldsymbol{\vartheta}_0$ and covariance matrix (2.12a) in [3] as long as $\mathrm{vecp}(\widehat{\mathcal{R}})$ is asymptotically normal, which is the case for the extreme copula correlation estimator $\widehat{\mathcal{R}}_T(k,w^*)$ defined in (10).

Now let $\mathbf{r}_0 = \mathrm{vecp}(\mathcal{R}_0)$ be the vectorised correlation matrix and assume that $\mathbf{X}$ has the elliptical copula $\mathcal{EC}_d(\mathcal{R}_0, G)$ at least in the extremes. Then we estimate the copula correlation by $\widehat{\mathcal{R}}_T(k,w^*)$ as shown in Section 3.2. The parameter $\boldsymbol{\vartheta}$ of the structure model is then estimated by minimising the quadratic discrepancy function with weight matrix

$$\widehat{\mathcal{U}}_T := \widehat{\Gamma}_T^{-1} - \widehat{\Gamma}_T^{-1}\widehat{\Delta}(\widehat{\Delta}\widehat{\Gamma}_T^{-1}\widehat{\Delta})^{-1}\widehat{\Delta}^\top\widehat{\Gamma}_T^{-1}\,,$$

where $\widehat{\Delta}$ is an estimator of $\Delta := \frac{\partial \mathbf{r}(\boldsymbol{\vartheta})}{\partial \boldsymbol{\vartheta}}|_{\boldsymbol{\vartheta}=\boldsymbol{\vartheta}_0}$. Using the estimator

$$\widehat{\boldsymbol{\vartheta}}_T := \mathrm{argmin}\, F(\widehat{\mathcal{R}}_T(k,w^*), \mathcal{R}(\boldsymbol{\vartheta})|\widehat{\mathcal{U}}_T)\,,$$

we define the asymptotic distribution-free test statistic

$$\begin{aligned} &\mathrm{ADF}(\widehat{\boldsymbol{\vartheta}}_T, \widehat{\mathcal{R}}_T(k,w^*), \widehat{\Gamma}_T) \\ &\quad = \quad k(\widehat{\mathbf{r}}_T(k,w^*) - \mathbf{r}(\widehat{\boldsymbol{\vartheta}}_T))^\top \widehat{\mathcal{U}}_T (\widehat{\mathbf{r}}_T(k,w^*) - \mathbf{r}(\widehat{\boldsymbol{\vartheta}}_T)) \end{aligned}$$

which is, due to [3, Proposition 4], non-centrally χ^2-distributed with $\mathrm{df} = d(d-1)/2 - p$ degrees of freedom and non-centrality parameter

$$\eta = k(\mathbf{r}_0 - \mathbf{r}(\boldsymbol{\vartheta}_0))^\top \mathcal{U} (\mathbf{r}_0 - \mathbf{r}(\boldsymbol{\vartheta}_0)),$$

for $\mathcal{U} = \Gamma_T^{-1} - \Gamma_T^{-1}\Delta(\Delta\Gamma_T^{-1}\Delta)^{-1}\Delta^\top\Gamma_T^{-1}$. In case $\mathcal{R}_0 = \mathcal{R}(\boldsymbol{\vartheta}_0)$ the non-centrality parameter η is zero and ADF has a central χ^2-distribution. This fact will now be used for testing

$$H_0: \ \mathbf{r}_0 = \mathbf{r}(\boldsymbol{\vartheta}_0) \qquad \text{for some } \boldsymbol{\vartheta}_0 \in \Theta,$$

by assuming $s \geq 1$ nested models

$$\mathbf{r}^{(i)} : \Theta^{(i)} \to \mathbb{R}^{d(d-1)/2},\ \boldsymbol{\vartheta}^{(i)} \mapsto \mathbf{r}^{(i)}(\boldsymbol{\vartheta}^{(i)}), \text{ and } \Theta^{(i)} \subset \mathbb{R}^{p^{(i)}},$$

for $1 \leq i \leq s$, which all have to satisfy the conditions in [3, Proposition 4]. The s models will be nested, if for every $1 \leq i \leq s-1$ and $\boldsymbol{\vartheta}^{(i)} \in \Theta^{(i)}$ there exists some $\boldsymbol{\vartheta}^{(i+1)} \in \Theta^{(i+1)}$ such that $\mathbf{r}^{(i+1)}(\boldsymbol{\vartheta}^{(i+1)}) = \mathbf{r}^{(i)}(\boldsymbol{\vartheta}^{(i)})$. In our factor analysis example these nested models correspond to models with an increasing number m of common factors.

Next consider the null hypotheses

$$H_0^{(i)}: \quad \mathbf{r}_0 = \mathbf{r}^{(i)}(\boldsymbol{\vartheta}_0^{(i)}) \quad \text{for some} \quad \boldsymbol{\vartheta}_0^{(i)} \in \Theta^{(i)}, \quad 1 \leq i \leq s,$$

and assume that some of these null hypotheses are true. Then there exists some $j \in \{1,\ldots,s\}$ such that $H_0^{(i)}$ does not hold for $1 \leq i < j$ and does hold for $j \leq i \leq s$. As we are interested in a structure model, which is likely to explain the observed extreme dependence structure, and is as simple as possible, we have to estimate j, the smallest index, where the null hypothesis holds.

By [3, Proposition 4] the corresponding test statistics

$$\begin{aligned} &\mathrm{ADF}^{(i)}(\widehat{\boldsymbol{\vartheta}}_T, \widehat{\mathcal{R}}_T(k,w^*), \widehat{\Gamma}_T) \\ &:= k \min_{\boldsymbol{\vartheta}\in\Theta^{(i)}} (\widehat{\mathbf{r}}_T(k,w^*) - \mathbf{r}^{(i)}(\widehat{\boldsymbol{\vartheta}}_T))^\top \widehat{\mathcal{U}}_T (\widehat{\mathbf{r}}_T(k,w^*) - \mathbf{r}^{(i)}(\widehat{\boldsymbol{\vartheta}}_T)) \end{aligned}$$

are non-centrally χ^2_{df}-distributed for $1 \leq i < j$ and are χ^2_{df}-distributed for $j \leq i \leq s$. Consequently, we reject a null hypothesis $H_0^{(i)}$, if the corresponding test statistic $\mathrm{ADF}^{(i)}$ is larger than some χ^2_{df}-quantile. Hence, j is the smallest number, where $H_0^{(j)}$ cannot be rejected.

Example 4.1. To analyse the small sample behaviour of the test statistic ADF, we perform a simulation study. We choose a $d = 10$ dimensional setting with $m = 2$ factors, loading matrix

$$\mathcal{L}^\top = 0.9 \cdot \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

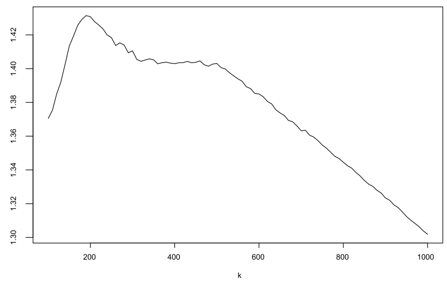

Figure 1. Estimates $\widehat{\nu}(k,w)$ for different values of k for sample size $n = 5\,000$.

and specific factors

$$\mathrm{diag}(\mathcal{V}^2) = 0.19 \cdot \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

Then $\mathcal{L}\mathcal{L}^\top + \mathcal{V}^2 = \mathcal{R}$ is a correlation matrix. The dependence structure is described by the t_ν-copula $\mathcal{EC}_d(\mathcal{R}, G)$, where $G \sim \sqrt{\nu/\chi^2_\nu}$ for $\nu > 0$. The tail parameter was chosen to be $\nu = 1.5$. All marginal distributions were chosen to be standard exponential.

We simulated 500 i.i.d. samples of sizes $n = 1\,000, 5\,000$ and $10\,000$, respectively, of the $t_{1.5}$-copula. For each sample we calculated for the one and two factor models $\mathbf{r}^{(1)}$ respectively $\mathbf{r}^{(2)}$ the corresponding test statistic $\mathrm{ADF}^{(i)}$, $i = 1, 2$, based on the extreme copula correlation estimators (10) and their estimated covariance matrices (18) with weight function taken as a discrete version of

$$w^*(\theta) = 1 - \left(\frac{\theta}{\pi/4} - 1\right)^2, \quad 0 \leq \theta \leq \frac{\pi}{2},$$

which gives maximal weight to data near the diagonal, and falls off to 0 in 0 and $\pi/2$. This assigns reasonable weights, since $w^*(0) = w^*(\pi/2) = 0$ and maximal weight $w^*(\pi/4) = 1$ on the diagonal. The mean-squared error of the estimator $\widehat{\nu}(k, w^*)$ was analysed in [26] through a simulation study (see Figure 3). It was shown that it is advantageous to use $\widehat{\nu}(k, w^*)$ compared to just using the simple estimator $\tilde{\nu}(1, 1, k)$. The number of upper order statistics used for the estimation of the tail dependence function have been $k = 80$, 300 and 500, respectively. These were chosen by a graphical approach shown in Figure 1. The figure shows estimates of ν for different values of k. The idea is that estimates $\widehat{\nu}(k, w)$ should be stable for suitable values of k. Thus we identify the first stable region of the estimates of ν and pick one of the corresponding k. Analysing such figures showed that the above values of k are suitable choices for our simulation study.

To ensure uniqueness of the loadings, we restrict $\mathcal{L}^\top\mathcal{V}^{-2}\mathcal{L}$ to be diagonal, hence we have $m(m-1)/2 = 1$ additional constraints; see [27, Section 2.3]. Using this restriction and the 2-factor setting, $\mathrm{ADF}^{(2)}$ should be asymptotically χ^2_{df}-distributed with $\mathrm{df} = d(d-1)/2 - dm + m(m-1)/2 = 26$

Table 1. The empirical level of the ADF *test for the two factor model for different significance levels*

	$k=80$	$k=300$	$k=500$
$\alpha=0.20$	0.168	0.256	0.279
$\alpha=0.10$	0.116	0.165	0.186
$\alpha=0.05$	0.084	0.113	0.127
$\alpha=0.01$	0.046	0.044	0.050

Table 2. The empirical level of the ADF *test for the three factor model for different significance levels*

	$k=80$	$k=300$	$k=500$
$\alpha=0.20$	0.158	0.127	0.183
$\alpha=0.10$	0.108	0.054	0.080
$\alpha=0.05$	0.067	0.028	0.045
$\alpha=0.01$	0.036	0.005	0.008

degrees of freedom. Therefore, we compare in each case the 500 estimates of $\text{ADF}^{(i)}$, $i=1,2$, with the χ^2_{26}-distribution by a QQ-plot; see Figure 2.

From the second row of Figure 2 we see that the distribution of $\text{ADF}^{(2)}$ fits the χ^2_{26}-distribution rather well even for $n=1\,000$ observations corresponding to $k=80$ upper order statistics. In the one factor cases, depicted in the first row, one clearly recognises the non-centrality parameter in the huge values of the vertical axes, which leads to a rejection of the one-factor model in almost all cases.

For the two factor model ($i=2$) we present the empirical levels of the ADF test at level $\alpha\in\{0.01,0.05,0.1,0.2\}$: in Table 1.

We observe that in almost all cases the empirical level is above the expected level of the test. This indicates that, although we get the correct shape of the distribution, we still have a bias in the estimate of the non-centrality parameter, which leads to an increased rejection rate.

Example 4.2. In the second example we consider a $d=15$ dimensional setting with $m=3$ factors. In this case the loading matrix is equal to

$$\mathcal{L}^\top=\begin{pmatrix}\mathbf{L}^\top & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{L}^\top & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{L}^\top\end{pmatrix}$$

and the specific factor is

$$\operatorname{diag}(\mathcal{V}^2)=\begin{pmatrix}\mathbf{v}^\top & \mathbf{v}^\top & \mathbf{v}^\top\end{pmatrix}$$

with $\mathbf{L}^\top=(0.70\quad 0.70\quad 0.75\quad 0.80\quad 0.80)$ and $\mathbf{v}^\top=(0.51\quad 0.51\quad 0.4375\quad 0.36\quad 0.36)$.

The generating variable G is the same as in the previous example, i.e. we have again the copula of a t_ν-distribution with parameter $\nu=1.5$. We simulated again 500 i.i.d. samples of sizes $n=1\,000$, $5\,000$ and $10\,000$, respectively. For each sample we calculated for the one, two and three factor model $\mathbf{r}^{(1)}$, $\mathbf{r}^{(2)}$ respectively $\mathbf{r}^{(3)}$ the corresponding test statistic $\text{ADF}^{(i)}$, $i=1,2,3$, based on the estimators (10) and (18). The number k of upper order statistics was selected as in Example 4.1. Under the same uniqueness restrictions on the loadings as in Example 4.1 $\text{ADF}^{(3)}$ should be asymptotically χ^2_{df}-distributed with $\text{df}=d(d-1)/2-dm+m(m-1)/2=63$ degrees of freedom. The test statistics $\text{ADF}^{(i)}$ for $i=1,2$ should be asymptotically non-centrally χ^2_{df}-distributed with $\text{df}=90$ and $\text{df}=76$ degrees of freedom, respectively. Figure 3 shows the corresponding QQ plots.

We clearly observe again the high non-centrality parameter in the first two cases on the horizontal axes, leading to a rejection of the one and two factor model in almost all cases.

For the three factor model ($i=3$) we present the empirical levels of the ADF test at level $\alpha\in\{0.01,0.05,0.1,0.2\}$ in Table 2

The empirical levels are in almost all cases lower than the nominal level of the test and especially for $k=500$ quite close to the level of the test.

5. DATA EXAMPLE

Finally we want to apply our method to a financial data set. We consider the same data as in [7]. It contains daily observations of 21 financial indices over a period from January 2001 to December 2009: five equity indices (DAX, STOXX50, S&P500, MSCI-World and MSCI-EE), 14 fixed income indices (iBoxx indices) and two commodity indices (Commodities and Gold). Ten of fourteen iBoxx indices are German government bonds and bonds of Euro nations with different times to maturity. The remaining four bonds are corporate bonds of Euro nations with different ratings. One can observe a very strong correlation between the German government bonds and the iBoxx indices consisting of Euro nation bonds, which have the same time to maturity. This is due to the fact that the Euro indices consisted to a large part of German bonds. High correlations are also observed between indices of the same type but consecutive time to maturities. Therefore we only considered the fixed income indices iBoxx-G-3-5, iBoxx-G-7-10, iBoxx-E-1-3, iBoxx-E-5-7, iBoxx-E-10+, iBoxx-E-AAA, iBoxx-E-AA, iBoxx-E-A and iBoxx-E-BBB.

In a first step we computed log-returns and fitted univariate ARMA(1,1)-GARCH(1,1) models to these return series analogously to [7]. The analysis of the dependence structure in the extremes of the fitted residuals is our goal. Before doing this we recall that we assume an elliptical dependence structure in the extremes, and we should put this to the test. A formal statistical test can be based on the difference between the elliptical tail dependence function and its empirical counterpart, which converges with rate k to some functional of a Gaussian process (cf. [29, Theorem 1]). As the limit process is inaccessible, the high quantile needed to formulate a goodness-of-fit test can in principle be obtained by some bootstrap mechanism. It is, however, well-known

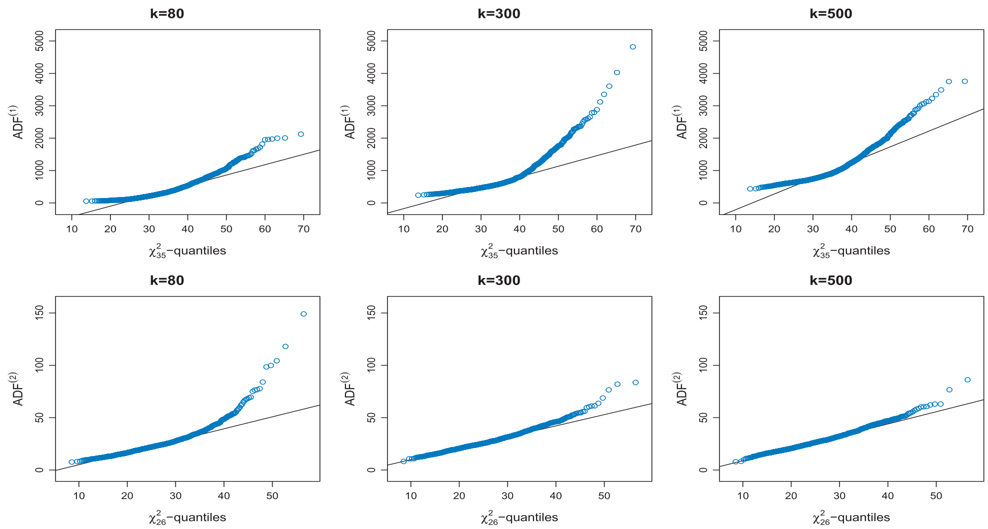

Figure 2. *QQ-plot of ordered estimates* $\mathrm{ADF}^{(i)}$ *for* $n = 1\,000$ *and* $k = 80$ (left), $n = 5\,000$ *and* $k = 300$ (middle)*and* $n = 10\,000$ *and* $k = 500$ (right)*observations. The first row shows the results for the one factor model and the second row for the two factor model.*

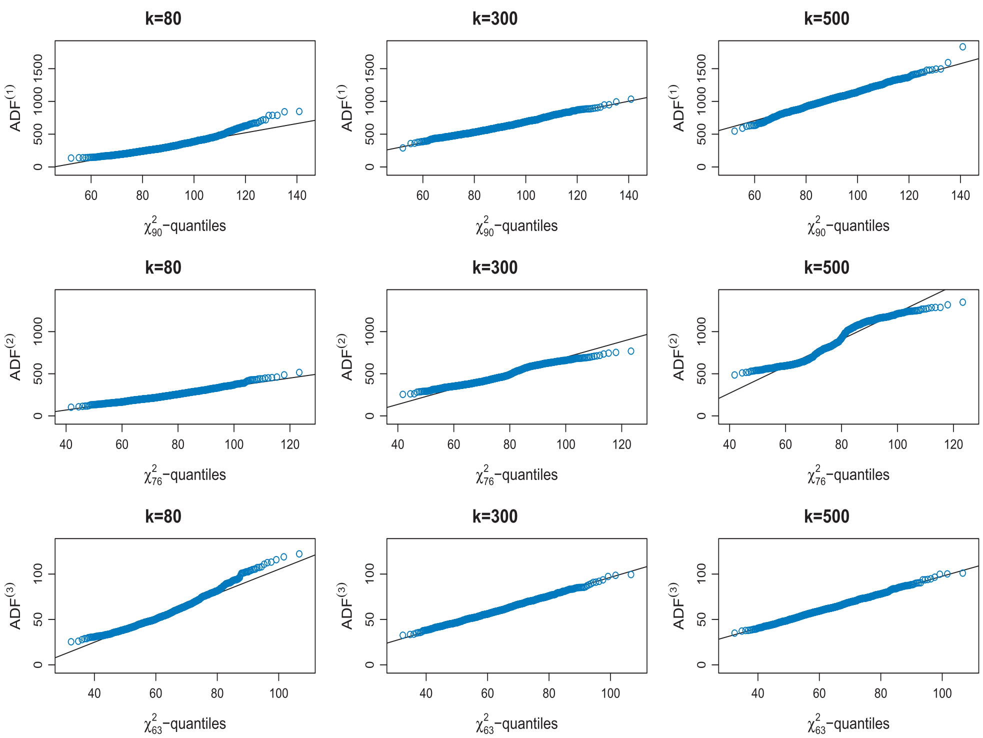

Figure 3. *QQ-plot of ordered estimates* $\mathrm{ADF}^{(i)}$ *for* $n = 1\,000$ *and* $k = 80$ (left), $n = 5\,000$ *and* $k = 300$ (middle)*and* $n = 10\,000$ *and* $k = 500$ (right)*observations. The first row shows the results for one factor model, the second row for the two factor model and the last row for three factor model.*

that the naive bootstrap yields unsatisfactory results for the estimation of high quantiles and more sophisticated bootstrap methods are required (e.g. the wild bootstrap, cf. [39] or [30]). One version of the wild bootstrap, called multiplier bootstrap has been suggested in [4] for the goodness-of-fit of one-parametric tail copulas based on some specific estimation procedure for this one parameter. It should be possible to extend this theory also to develop goodness-of-fit tests for elliptical tail dependence. This is, however, not straightforward and hence a topic for future work. For the data set

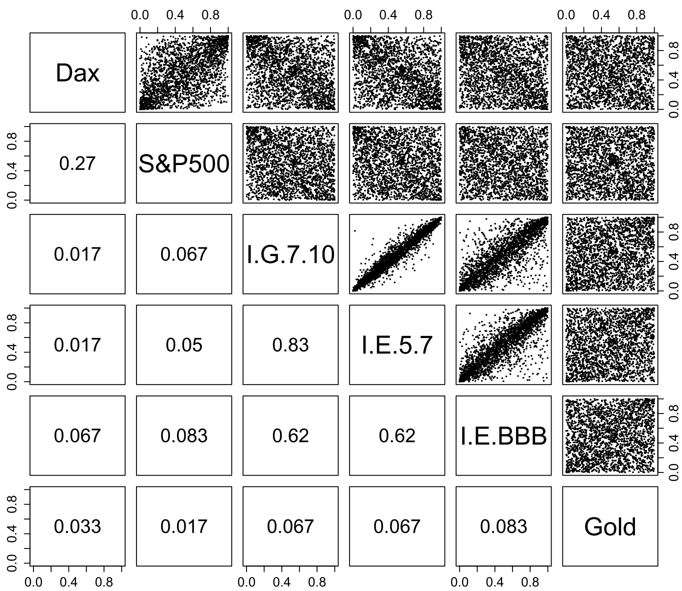

Figure 4. Pairwise scatter plots of pseudo-observations computed from the negative residuals of DAX, S&P500, iBoxx-G-3-5, iBoxx-E-5-7, iBoxx-E-BBB and the Gold index.

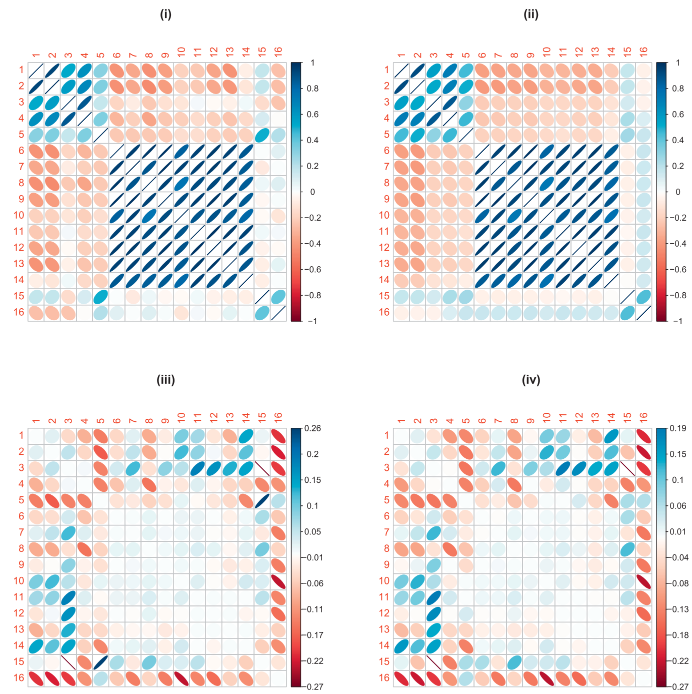

Figure 5. Dependence explained by the optimal copula factor models: (i) extreme dependence $\widehat{\mathcal{R}}_T^{(7)}$, (ii) robust dependence $\widehat{\mathcal{R}}_\tau^{(9)}$; (iii) difference $\widehat{\mathcal{R}}_T^{(7)} - \widehat{\mathcal{R}}_\tau^{(9)}$; (iv) difference between extreme and robust copula correlation $\widehat{\mathcal{R}}_T - \widehat{\mathcal{R}}_\tau$.

of the present paper, before setting to work on the structure analysis, we performed some exploratory tests like checking the shape of the pseudo-observations (see e.g. Section 2 in [15]) computed from the fitted residuals (cf. Figure 4) and the estimation of ν on different marginals of the data set (cf. Figures 5 and 6 in [26] for an assessment of the variability in the estimation of ν). We did not find any indication which contradicted the assumption.

Since we are interested in the dependence structure of large losses, we will perform an extreme copula factor analysis of the negative residuals, so that extreme dependence corresponds to (upper) tail dependence. A pairs plot of pseudo-observations, computed from the negative residuals of some of the indices, is shown in Figure 4. The empirical tail dependence coefficients $\mathbb{T}_{ij}(1,1;60)$ are given in the lower half of the pairs plot.

The copula factor analysis is done by using the function `cop.struc()` from the R package Cop.Struc, which is available on the first author's web site. For the sixteen-dimensional data set, the extreme copula factor model with six factors can still be rejected at a chosen significance level of 0.05, but the model with seven factors cannot be rejected. We denote by $\widehat{\mathcal{R}}_T^{(7)}$ the estimated extreme copula correlation matrix corresponding to the optimal 7-factor model. For comparison, we also carried out a robust copula factor analysis, where we estimate $\mathcal{R}$ by Kendall's tau via (6). In that case we obtained a model with nine factors. Analogously, we denote by $\widehat{\mathcal{R}}_\tau^{(9)}$ the estimated robust copula correlation matrix corresponding to the optimal 9-factor model. Note that our data set combines three different types of indices: equity, fixed income and commodity indices. The fixed income indices can further be divided into German government and Euro nation bonds with different maturities and corporate bonds of Euro nations with different ratings. Therefore, we would have guessed that we need at least six factors to suitably represent the copula correlation matrix.

Figure 5 shows estimates of the extreme copula correlation matrix $\widehat{\mathcal{R}}_T^{(7)}$ and of the robust copula correlation matrix $\widehat{\mathcal{R}}_\tau^{(9)}$ in the top row. The estimator $\widehat{\mathcal{R}}_T^{(7)}$ (which focuses after all on extreme positive dependence) assigns slightly less weight on the negative dependence between equity and fixed income indices compared to the estimator $\widehat{\mathcal{R}}_\tau^{(9)}$ based on Kendall's tau. Further we see a slightly lower correlation between fixed income and commodity indices in $\widehat{\mathcal{R}}_T^{(7)}$ than in $\widehat{\mathcal{R}}_\tau^{(9)}$. In the second row of Figure 5 the estimated differences $\widehat{\mathcal{R}}_T^{(7)} - \widehat{\mathcal{R}}_\tau^{(9)}$ and $\widehat{\mathcal{R}}_T - \widehat{\mathcal{R}}_\tau$ are depicted, where the latter one is the difference between (10) and (6). The main difference between the figures (iii) and (iv) is that the extreme copula factor model assigns much more weight to the dependence between the MSCI East Europe index and the commodity index than the robust copula factor model. This effect cannot be seen for the raw estimates $\widehat{\mathcal{R}}_T$ and $\widehat{\mathcal{R}}_\tau$. Overall, the differences between the copula correlations vary between -0.27 and 0.26 for the factor models, and between -0.27 and 0.19 for the raw estimates.

Finally we compare the estimated loadings for different factors summarised in $\widehat{\mathcal{L}}_T^{(7)}$ and $\widehat{\mathcal{V}}_T^{(7)}$, which define the estimate of the extreme copula correlation

$$\widehat{\mathcal{R}}_T^{(7)} = \widehat{\mathcal{L}}_T^{(7)} \widehat{\mathcal{L}}_T^{(7)^\top} + \widehat{\mathcal{V}}_T^{(7)} .$$

The loadings, which are shown in Figure 6, are obtained by applying the varimax method to the original loadings $\widehat{\vartheta}_T$. Recall that the set of loadings is not unique with respect to orthogonal transformations. The aim of the varimax method, originally introduced in [21] (see also Chapter 6.3 in [27]), is such that the transformed loadings are either rather large or rather small in magnitude. For the investigator such a set of loadings is then easier to interpret. Since the set of possible orthogonal transformations is not limited to the varimax method, different methods may lead to slightly different interpretations. But this is a general drawback of every factor analysis and will not be discussed here any further.

For comparison, Figure 6 also depicts the factor loadings summarised in $\widehat{\mathcal{L}}_\tau^{(9)}$ and $\widehat{\mathcal{V}}_\tau^{(9)}$ based on the robust copula correlation estimator. The first three factors can be identified as the fixed income, equity and commodity factors. For the loadings of factor 4 we observe the first differences comparing the loading of the extreme copula factor $\widehat{\mathcal{L}}_T^{(7)}$ and the robust copula factor $\widehat{\mathcal{L}}_\tau^{(9)}$. Factor 4 puts weight on the equity and commodity indices for $\widehat{\mathcal{L}}_T^{(7)}$. Concerning the equity indices there is special emphasis on the MSCI East Europe index. The loadings for the commodity indices are much larger for the extreme copula factor model than for the robust copula factor model. Therefore, for the given sample the dependence between the MSCI East Europe index and the commodity indices is larger, when focusing on common large losses compared to considering all rank observations. Factor 5 distinguishes between different times to maturity. The short term bonds iBoxx-E-1-3 and iBoxx-E-3-5 get negative loadings and the mid term bond iBoxx-E-5-7 has a loading of roughly zero, while the long term bonds iBoxx-G-7-10 and iBoxx-E-10+ have positive loadings. The sixth factor is again an equity factor. This time higher loadings are given to DAX and STOXX50. We also see some relation to the iBoxx-BBB index, which is not the case for factor 2. The last two factors are hard to interpret. The loadings for factor 7 are all very small. Therefore, one may argue that a model with six factor may be sufficient when taking the finite sample properties of our test statistic into account as demonstrated in Figure 2.

6. PROOF OF THEOREM 3.13

In the bivariate case the following weak convergence result was shown in [36, Theorem 5] under the assumption that the tail dependence function possesses continuous partial derivatives:

(19)

$$\begin{aligned}&\sqrt{k}\,(\mathbb{T}(x_1,x_2;\,k) - T(x_1,x_2))\\ &\stackrel{w}{\to} B(\mathbf{x}) - \frac{\partial}{\partial x_1}T(x_1,x_2)B(x_1,\infty) - \frac{\partial}{\partial x_2}T(x_1,x_2)B(\infty,x_2)\end{aligned}$$

in $\mathcal{B}(\overline{\mathbb{R}}_+^2)$ (see [36] for details on convergence in this space), where B is a zero mean Wiener process with covariance structure $\mathbb{E}(B(x_1,x_2)B(y_1y_2)) = T(x_1\wedge x_2, y_1\wedge y_2)$ for $\mathbf{x},\mathbf{y}\in\overline{\mathbb{R}}_+^2$. We need an extension of this result from the bivariate to a d-dimensional setting. Using the arguments in the proof of [36, Theorem 5] we get

$$\sqrt{k}\,(\mathbb{T}(\mathbf{x};\,k) - T(\mathbf{x})) \stackrel{w}{\to} B(\mathbf{x}) - \sum_{i=1}^d \frac{\partial}{\partial x_i}T(\mathbf{x})B_i(x_i)$$

in $\mathcal{B}(\overline{\mathbb{R}}_+^d)$, where $B_i(x_i) = B(\mathbf{x})$ with $x_l=\infty$ for $l\neq i$ and B is a zero mean Wiener process with covariance structure $\mathbb{E}(B(\mathbf{x})B(\mathbf{y})) = T(\mathbf{x}\wedge\mathbf{y})$ for $\mathbf{x},\mathbf{y}\in\overline{\mathbb{R}}_+^d$. The same result is also given in the proof of Theorem 1 in [29]. In particular, this implies consistency of all bivariate empirical tail dependence functions $\mathbb{T}_{ij}$ and their distributional convergence (centred and scaled) to a Gaussian limit. Recently Bücher and Dette [4] have shown that the assumption of continuous partial derivatives is unnecessarily restrictive and established the same result as in [36, Theorem 5] under weaker smoothness assumptions; see [4, Theorem 2]. The arguments in their proof can also be extended to the d-dimensional setting.

Since $\widehat{\mathbf{r}}_T$ is the image of $\mathbb{T}$ under a certain map ϕ we can again use an extended version of the classical delta-method (see [38, p. 374] for details) to show the asymptotic normality of $\widehat{\mathbf{r}}_T$. First, note that for all $i\neq j$ and for T defined in (3)

$$\inf_{\theta\in Q_{ij}^*}\left|\frac{\partial}{\partial\nu}T(x(\theta),y(\theta),\,\nu,\rho_{ij})\right| > 0,$$
$$\inf_{\theta\in U_{ij}^*}\left|\frac{\partial}{\partial\rho}T(x(\theta),y(\theta),\,\nu,\rho_{ij})\right| > 0$$

and

$$\sup_{\theta\in U_{ij}^*}\left|\frac{\partial}{\partial\nu}T^{\leftarrow\rho}(\cdot\,|\,x(\theta),y(\theta),\,\nu,\rho_{ij})\right| < \infty.$$

Next, define $\mathbb{D}$ as the set of all d-dimensional tail dependence functions, which is a subset of the complete metric space $\mathcal{B}_\infty(\overline{\mathbb{R}}_+^d)$; see [36, Definition 4]. Abbreviate for $\mu\in\mathbb{D}$, with μ_{ij} being the ij-th marginal of μ,

$$\tilde{\nu}_{ij}(\theta,\mu,\rho) := T^{\leftarrow\nu}\left(\mu_{ij}(x(\theta),y(\theta))\;|\;x(\theta),y(\theta),\rho\right)$$

and

$$\tilde{\rho}_{ij}(\theta,\mu,\nu) := T^{\leftarrow\rho}\left(\mu_{ij}(x(\theta),y(\theta))\;|\;x(\theta),y(\theta),\nu\right).$$

Next, define for some correlation matrix $\mathcal{R} = (\rho_{ij})_{1\le i,j\le d}$

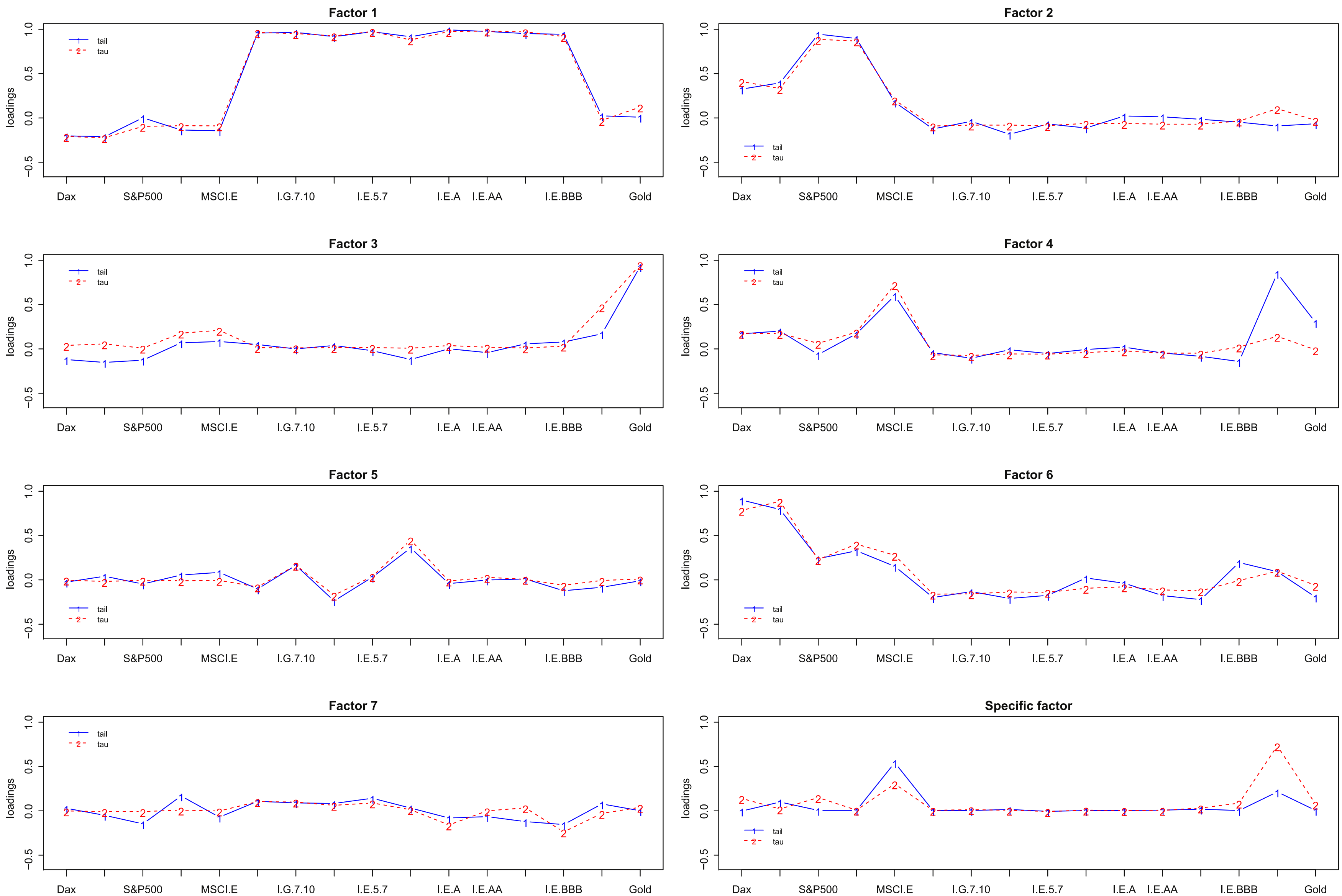

Figure 6. Estimated factors ($\widehat{\mathcal{L}}_T$ and $\widehat{\mathcal{L}}_\tau$) and specific ($\widehat{\mathcal{V}}_T$ and $\widehat{\mathcal{V}}_\tau$) loadings.

$$\nu(\mu,\mathcal{R}) := \frac{1}{d(d-1)}\sum_{i\neq j}\frac{1}{W^*(Q^*_{ij})}\int_{\theta\in Q^*_{ij}}\tilde\nu_{ij}(\theta,\mu,\rho_{ij})\,W^*(d\theta),$$

$$\rho_{ij}(\mu,\mathcal{R}) := \frac{1}{W^*(U^*_{ij})}\int_{\theta\in U^*_{ij}}\tilde\rho_{ij}(\theta,\mu,\nu(\mu,\mathcal{R}))\,W^*(d\theta)\,.$$

Using this notation we get that $\mathbf{r}=\phi(T,\mathcal{R})$ with $\phi: T\to \mathrm{vecp}((\rho_{ij}(T,\mathcal{R}))_{1\le i,j\le d})$. Due to the chain rule ([38, Lemma 3.9.3]) ϕ is Hadamard-differentiable if $\nu(\cdot,\mathcal{R})$ and $\rho_{ij}(\cdot,\mathcal{R})$ are Hadamard-differentiable. But $\nu(\cdot,\mathcal{R})$ is Hadamard-differentiable, since for $t_m\to\infty$ and $h_m\to h\in\mathbb{D}$ for $m\to\infty$, such that $\mu+h_m/t_m\in\mathbb{D}$ for all m, we obtain by Taylor expansion,

$$\begin{aligned}\lim_{m\to\infty} & t_m\left(\nu(\mu+h_m/t_m,\mathcal{R})-\nu(\mu,\mathcal{R})\right)\\ &= \frac{1}{d(d-1)}\sum_{i\neq j}\frac{1}{W^*(Q^*_{ij})}\\ &\quad \int_{\theta\in Q^*_{ij}}\frac{h_{ij}(x(\theta),y(\theta))}{\frac{\partial}{\partial\nu}T\left(x(\theta),y(\theta),\nu(\mu,\mathcal{R}),\rho_{ij}\right)}W^*(d\theta)\\ &=: \nu'_\mu(h),\end{aligned}$$

which obviously is a linear map. Analogously, $\rho_{ij}(\cdot,\mathcal{R})$ is Hadamard differentiable; i.e.,

$$\begin{aligned}&\lim_{m\to\infty} t_m\left(\rho_{ij}(\mu+h_m/t_m,\mathcal{R})-\rho_{ij}(\mu,\mathcal{R})\right)\\ &= \frac{1}{W^*(U^*_{ij})}\int_{\theta\in U^*_{ij}}\left(\frac{h_{ij}(x(\theta),y(\theta))}{\frac{\partial}{\partial\rho}T\left(x(\theta),y(\theta),\nu(\mu),\rho_{ij}\right)}\right.\\ &\left.+\nu'_\mu(h)\frac{\partial}{\partial\nu}T^{\leftarrow\rho}\left(\mu_{ij}(x(\theta),y(\theta))\mid x(\theta),y(\theta),\nu(\mu)\right)\right)W^*(d\theta).\end{aligned}$$

Define

$$\widehat{\mathbf{r}}_T(k,w^*)=\phi\left(\mathbb{T}(\cdot;k),\widehat{\mathcal{R}}_\tau\right)\,.$$

Since $\widehat{\mathcal{R}}_\tau-\mathcal{R}=o_p(1/\sqrt{k})$, the delta method yields

$$\sqrt{k}\left(\widehat{\mathbf{r}}_T(k,w^*)-\mathbf{r}\right)\xrightarrow{d}\phi'(\tilde B,\mathcal{R})\,,$$

where $\tilde B(\mathbf{x}):=B(\mathbf{x})-\sum_{i=1}^d\frac{\partial}{\partial x_i}T(\mathbf{x})B_i(x_i)$, $\mathbf{x}\in\overline{\mathbb{R}}^d_+$. The result then follows using

$$\begin{aligned}&\mathbb{E}\left(\left(\phi'(\tilde B,\mathcal{R})\right)_{ij}\left(\phi'(\tilde B,\mathcal{R})\right)_{kl}\right)\\ &\quad=\sigma_{1;ij,kl}+\sigma_{2;ij,kl}+\sigma_{3;ij,kl}+\sigma_{4;ij,kl},\end{aligned}$$

with $\sigma_{1;ij,kl},\sigma_{2;ij,kl},\sigma_{3;ij,kl},\sigma_{4;ij,kl}$ defined through (15)–(17).

Received 6 November 2013

REFERENCES

[1] Asimit, A. V. and Jones, B. L. Extreme behavior of bivariate elliptical distributions. *Insurance: Mathematics & Economics*, 41:53–61, 2007. MR2324565

[2] Bentler, P. M. and Dudgeon, P. Covariance structure analysis: Statistical practice, theory and directions. *Annual Reviews of Psychology*, 47:563–592, 1996.

[3] Browne, M. Asymptotically distribution-free methods for the analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology*, 37:62–83, 1984. MR0783499

[4] Bücher, A. and Dette, H. Multiplier bootstrap of tail copulas – with applications. *Bernoulli*, 19:1655–1687, 2013. MR3129029

[5] Cont, R. Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance*, 1:223–236, 2001.

[6] Danielsson, J., de Haan, L., Peng, L., and de Vries, C. Using a bootstrap method to choose the sample fraction in tail index estimation. *Journal of Multivariate Analysis*, 76:226–248, 2001. MR1821820

[7] Dißmann, J., Brechmann, E. C., Czado, C., and Kurowicka, D. Selecting and estimating regular vine copulae and application to financial returns. *Computational Statistics and Data Analysis*, 59: 52–69, 2013 MR3000041

[8] Demarta, S. and McNeil, A. J. The *t* copula and related copulas. *International Statistical Review*, 73:111–129, 2005.

[9] Drees, H. and Huang, X. Best attainable rates of convergence for estimators of the stable tail dependence function. *Journal of Multivariate Analysis*, 64:25–47, 1998. MR1619974

[10] Drees, H. and Kaufmann, E. Selecting the optimal sample fraction in univariate extreme value estimation. *Stochastic Processes and Their Applications*, 75:149–172, 1998. MR1632189

[11] Embrechts, P., Lindskog, F., and McNeil, A. J. Modelling dependence with copulas and applications to risk management. In *Handbook of Heavy Tailed Distributions in Finance* (ed. S. Rachev), Chapter 8, 329–384. Amsterdam: Elsevier, 2003.

[12] Fang, K.-T., Kotz, S., and Ng, K. *Symmetric Multivariate and Related Distributions*. London: Chapman and Hall, 1990. MR1071174

[13] Fang, H.-B., Fang, K.-T., and Kotz, S. The meta-elliptical distributions with given marginals. *Journal of Multivariate Analysis*, 82:1–16, 2002. MR1918612

[14] Frahm, G., Junker, M., and Schmidt, R. Estimating the tail-dependence coefficient: Properties and pitfalls. *Insurance: Mathematic & Economics*, 37:80–100, 2005. MR2156598

[15] Haug, S., Klüppelberg, C., and Peng, L. Statistical models and methods for dependence in insurance data. *Journal of the Korean Statistical Society*, 40:125–139, 2011. MR2830506

[16] Heffernan, J. E. and Tawn, J. A. A conditional approach for multivariate extreme values. *Journal of the Royal Statistical Society, Series B*, 66:497–529, 2004. MR2088289

[17] Higham, N. Computing the nearest correlation matrix – A problem from finance. *IMA Journal of Numerical Analysis*, 22:329–343, 2002. MR1918653

[18] Hult, H. and Lindskog, F. Multivariate extremes, aggregation and dependence in elliptical distributions. *Advances in Applied Probability*, 34:587–608, 2002. MR1929599

[19] Joe, H., Li, H., and Nikoloulopoulos, A. K. Tail dependence functions and vine copulas. *Journal of Multivariate Analysis*, 101: 252–270, 2010. MR2557632

[20] Jöreskog, K. G. A general method for analysis of covariance structures. *Biometrika*, 57:239–251, 1970. MR0269024

[21] Kaiser, H. F. The varimax criterion for analytic factor analysis. *Psychometrika*, 23:187–200, 1958.

[22] Kano, Y., Berkane, M., and Bentler, P. M. Covariance structure analysis with heterogeneous kurtosis parameters. *Biometrika*, 77:575–585, 1990. MR1087848

[23] Kendall, M. and Gibbons, J. *Rank Correlation Methods*. London: Edward Arnold, 1990. MR1079065

[24] Klüppelberg, C. and Kuhn, G. Copula structure analysis. *Journal of the Royal Statistical Society, Series B*, 71:737–753, 2009. MR2749917

[25] Klüppelberg, C., Kuhn, G., and Peng, L. Estimating the tail dependence of an elliptical distribution. *Bernoulli*, 13:229–251, 2007. MR2307405

[26] Klüppelberg, C., Kuhn, G., and Peng, L. Semi-parametric models for the multivariate tail dependence function – The asymptotically dependent case. *Scandinavian Journal of Statistics*, 35:701–718, 2008. MR2468871

[27] Lawley, D. and Maxwell, A. *Factor Analysis as a Statistical Method*. New York: Elsevier, 1971. MR0343471

[28] Ledford, A. and Tawn, J. Statistics for near independence in multivariate extreme values. *Biometrika*, 83:169–187, 1996. MR1399163

[29] Li, D. and Peng, L. Goodness-of-fit test for tail dependence functions modeled by elliptical copulas. *Statistics & Probability Letters*, 79:1097–1104, 2009. MR2510774

[30] Mammen, E. Bootstrap, wild bootstrap, and asymptotic normality. *Probability Theory and Related Fields*, 93:439–455, 1992. MR1183886

[31] McNeil, A. J., Frey, R., and Embrechts, P. *Quantitative Risk Management*. Princeton: Princeton University Press, 2005. MR2175089

[32] Nelsen, R. *An Introduction to Copulas*, Volume 139 of Lecture Notes in Statistics. New York: Springer, 1999. MR1653203

[33] Peng, L. A practical way for estimating tail dependence functions. *Statistica Sinica*, 20:365–378, 2010. MR2640699

[34] Rousseeuw, P. J. and Molenberghs, G. Transformation of non positive semidefinite correlation matrices. *Communications in Statistics – Theory and Methods*, 22:965–984, 1993.

[35] Satorra, A. and Bentler, P. A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, 66:507–514, 2001. MR1961910

[36] Schmidt, R. and Stadtmüller, U. Non-parametric estimation of tail dependence. *Scandinavian Journal of Statistics*, 33:307–335, 2006. MR2279645

[37] Shapiro, A. and Browne, M. W. Analysis of covariance structure under elliptical distributions. *Journal of the American Statistical Association*, 82:1092–1097, 1987. MR0922173

[38] van der Vaart, A. and Wellner, J. *Weak Convergence and Empirical Processes*. New York: Springer, 1996. MR1385671

[39] Wu, C. F. J. Jackknife, bootstrap and other resampling methods in regression analysis. *Annals of Statistics*, 14:1261–1295, 1986. MR0868303

[40] Yuan, K.-H., Bentler, P. M., and Chan, W. Structural equation modelling with heavy tailed distributions. *Psychometrika*, 69:421–436, 2004. MR2272458

Claudia Klüppelberg
Zentrum Mathematik
Technische Universität München
Boltzmannstraße 3
D-85748 Garching
Germany
E-mail address: cklu@tum.de
url: www.statistics.ma.tum.de

Stephan Haug
Zentrum Mathematik
Technische Universität München
Boltzmannstraße 3
D-85748 Garching
Germany
E-mail address: haug@tum.de
url: www.statistics.ma.tum.de

Gabriel Kuhn
Zentrum Mathematik
Technische Universität München
Boltzmannstraße 3
D-85748 Garching
Germany