Klüppelberg, Kuhn:

Copula Structure Analysis Based on Robust and Extreme Dependence Measures

Sonderforschungsbereich 386, Paper 507 (2006)

Copula Structure Analysis Based on Robust and Extreme Dependence Measures

CLAUDIA KLÜPPELBERG AND GABRIEL KUHN

Center for Mathematical Sciences, Munich University of Technology, D-85747 Garching, Germany.

gabriel@ma.tum.de

Summary

In this paper we extend the standard approach of correlation structure analysis in order to reduce the dimension of highdimensional statistical data. The classical assumption of a linear model for the distribution of a random vector is replaced by the weaker assumption of a model for the copula. For elliptical copulae a 'correlation-like' structure remains but different margins and non-existence of moments are possible. Moreover, elliptical copulae allow also for a 'copula structure analysis' of dependence in extremes. After introducing the new concepts and deriving some theoretical results we observe in a simulation study the performance of the estimators: the theoretical asymptotic behavior of the statistics can be observed even for a sample of only 100 observations. Finally, we test our method on real financial data and explain differences between our copula based approach and the classical approach. Our new method yields a considerable dimension reduction also in non-linear models.

Keywords: copula structure analysis, correlation structure analysis, covariance structure analysis, dimension reduction, elliptical copula, factor analysis, Kendall's tau, tail copula, tail dependence.

1 Introduction

When analyzing high-dimensional data one is often interested in understanding the dependence structure aiming at a dimension reduction. In the framework of correlation representing linear dependence, *correlation structure analysis* is a classical tool; see Steiger (1994) or Bentler and Dudgeon (1996). Correlation structure analysis is based on the assumption that the correlation matrix of the data satisfies the equation $\mathbf{R} = \mathbf{R}(\boldsymbol{\vartheta})$ for some function $\mathbf{R}(\boldsymbol{\vartheta})$ and a parameter vector $\boldsymbol{\vartheta}$. Typically, a *general linear structure model* is then considered for a d-dimensional random vector $\boldsymbol{X}$, i.e. $\boldsymbol{X} \stackrel{\mathrm{d}}{=} \mathbf{A}\boldsymbol{\xi}$, where $\mathbf{A} = \mathbf{A}(\boldsymbol{\vartheta})$ is a function of a parameter vector $\boldsymbol{\vartheta}$, and $\boldsymbol{\xi}$ represents some (latent) random vector.

The typical goal of correlation structure analysis is to reduce dimension, i.e. to explain the whole dependence structure through lower dimensional parameters summarized in $\boldsymbol{\vartheta}$. One particularly popular method is *factor analysis*, where the data $\boldsymbol{X}$ are assumed to satisfy the linear model $\boldsymbol{X} \stackrel{d}{=} \boldsymbol{\mu} + \mathbf{L}\boldsymbol{f} + \mathbf{V}\boldsymbol{e}$, where $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_d)^T$, $\boldsymbol{f} = (f_1, \ldots, f_m)^T$ $(m < d)$ are non-observable and (usually) uncorrelated *factors* and $\boldsymbol{e} = (e_1, \ldots, e_d)^T$ is some *noise variables*. Further, $\mathbf{L} \in \mathbb{R}^{d \times m}$ is called *loading matrix* and $\mathbf{V}$ is a diagonal matrix with nonnegative entries. An often used additional assumption is that $(\boldsymbol{f}^T, \boldsymbol{e}^T)$ has mean zero and covariance matrix $\mathbf{I}$, the identity matrix. Then, describing the dependence structure of $\boldsymbol{X}$ through its covariance matrix yields $\operatorname{Cov}(\boldsymbol{X}) = \boldsymbol{\Sigma} = \mathbf{L}\mathbf{L}^T + \mathbf{V}^2$, i.e., the dependence of $\boldsymbol{X}$ is described through the entries of $\mathbf{L}$.

Provided that the data are normally distributed this approach of decomposing the correlation structure is justified, since dependence in normal data is uniquely determined by correlation. However, many data sets exhibit properties contradicting the assumption of normality, see e.g. Cont (2001) for a study of financial data. Further, several covariance structure studies based on the normal assumption exhibit problems for nonnormal data, see e.g. Browne (1982, 1984). A modified approach is to assume an elliptical model, and the corresponding methods can be found for instance in Muirhead and Waternaux (1980) and Browne and Shapiro (1987). Browne (1982, 1984) also developed a method being asymptotically free of any distributional assumption, but it was found that acceptable performance of this procedure requires very large sample sizes; see Hu, Bentler, and Kano (1992).

Relaxing more and more the assumptions of classical correlation structure analysis as indicated above, one assumption still remains, namely that $\boldsymbol{X} \stackrel{d}{=} \mathbf{A}(\boldsymbol{\vartheta})\boldsymbol{\xi}$, i.e. $\boldsymbol{X}$ can be described as a linear combination of some (latent) random variables $\boldsymbol{\xi}$ with existing second moments (and existing fourth moments to ensure asymptotic distributional limits of sample covariance estimators). For real multivariate data it may happen that some margins are well modeled as being normal and some are more heavy-tailed (or leptokurtic). Moreover, nonlinear dependence can occur, e.g. in financial portfolios of assets and derivatives. If this happens, it is hard to believe that some linear model is appropriate. Since the primary aim of correlation or covariance structure analysis is to decompose and describe dependence we present a simple method to avoid problems of non-existing moments or different marginal distributions by using *copulae*. A copula is a d-dimensional distribution function with $\operatorname{unif}(0,1)$ margins and, by Sklar's theorem, each distribution function can be described through its margins and its copula separately. We will focus on *elliptical copulae* being the copulae of elliptical distributions, which are very flexible and easy to handle also in high dimensions. As a correlation matrix is a parameter of an elliptical copula, correlation structure analysis can be easily extended to such copulae and we will call this method *copula structure analysis*.

In many applications dependence in extremes is an important issue. 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. McNeil, Frey, and Embrechts (2005). This requires a different approach and is one of the major issue of this paper. We assess extreme dependence by a concept called *tail copula*. For such elliptical copulae, which model extreme dependence, we present a new structure analysis based on the tail copula. This focusses on dependence structure in the extremes.

Our paper is organized as follows. We start with definitions and preliminary results on copulae and elliptical distributions in Section 2. In Section 3 we introduce the new copula structure model and show which (classical) methods can be used for a structure analysis and model selection. In Section 4 we show two copula dependence concepts, one based on Kendall's tau, one on the tail copula, and develop estimators, which can then be used for the copula structure analysis. These concepts lead to different estimates of the copula structure model, and we derive asymptotic results for our estimates.

In Section 5 a simulation study shows that the derived asymptotic results hold already for a rather small simulated sample. Finally, we fit a copula factor model to real data based on both our dependence concepts and give an interpretation of the results. Proofs are summarized in Section 6.

2 Preliminaries

First, we introduce the copula concept. For more technical background information we refer to Nelsen (1999).

Definition 2.1. *A* copula $C : [0,1]^d \to [0,1]$ *is a d-dimensional distribution function with standard uniform margins, i.e.* $C(1,\ldots,1,u_j,1,\ldots,1) = u_j$, $1 \le j \le d$.

The following theorem shows that each multivariate distribution function can be separated in its dependence structure, i.e. the copula, and its margins. This important result is used in essentially all applications of copulae. We shall need the notion of a generalized inverse function. For a distribution function F the *generalized inverse* is defined as

$$F^{\leftarrow}(y) = \inf\{x \in \mathbb{R} \mid F(x) \ge y\}, \quad y \in (0,1).$$

Theorem 2.2 (Sklar's Theorem (1996)). *Let F be a d-dimensional distribution function with margins $F_1,\ldots,F_d$. Then there exists a copula C such that for all $\boldsymbol{x} \in \mathbb{R}^d$*

$$F(x_1,\ldots,x_d) = C\left(F_1(x_1),\ldots,F_d(x_d)\right).$$

The copula C is unique on $\mathrm{Ran}F_1 \times \cdots \times \mathrm{Ran}F_d$.
If F is a continuous d-dimensional distribution function with margins $F_1,\ldots,F_d$, and generalized inverse functions $F_1^{\leftarrow},\ldots,F_d^{\leftarrow}$, then the copula C of F is $C(u_1,\ldots,u_d) = F\left(F_1^{\leftarrow}(u_1),\ldots,F_d^{\leftarrow}(u_d)\right)$.

We will focus on copulae of elliptical distributions, and we first give some definitions and state some properties. For a general treatment of elliptical distributions we refer to Fang, Kotz, and Ng (1990) and to Cambanis, Huang, and Simons (1981). Elliptical copulae and their properties have also been investigated with respect to financial application by Embrechts, Lindskog, and McNeil (2003) or Frahm, Junker, and Szimayer (2003).

Definition 2.3. *A d-dimensional random vector $\boldsymbol{X}$ has an elliptical distribution, if, for some $\boldsymbol{\mu} \in \mathbb{R}^d$, some positive (semi-)definite matrix $\boldsymbol{\Sigma} = (\sigma_{ij})_{1 \le i,j \le d} \in \mathbb{R}^{d \times d}$, a positive random variable G and a random vector $\boldsymbol{U}^{(m)} \sim \mathrm{unif}\{\boldsymbol{s} \in \mathbb{R}^m : \boldsymbol{s}^T\boldsymbol{s} = 1\}$ independent of G it holds that $\boldsymbol{X} \overset{\mathrm{d}}{=} \boldsymbol{\mu} + G\mathbf{A}\boldsymbol{U}^{(m)}$, $\mathbf{A} \in \mathbb{R}^{d \times m}$, $\mathbf{A}\mathbf{A}^T = \boldsymbol{\Sigma}$ and some $m \in \mathbb{N}$. We write $\boldsymbol{X} \sim \mathcal{E}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}, G)$. The random variable G is called* generating variable. *Further, if the first moment exists, then $E\boldsymbol{X} = \boldsymbol{\mu}$, and if the second moment exists, then G can be chosen such that $\mathrm{Cov}\boldsymbol{X} = \boldsymbol{\Sigma}$.*

Definition 2.4. *Let $\boldsymbol{X} \sim \mathcal{E}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \Phi)$ with $\boldsymbol{\Sigma} = (\sigma_{ij})_{1 \le i,j \le d}$. We define the* correlation matrix $\mathbf{R}$ *by $\mathbf{R} := \left(\sigma_{ij}/\sqrt{\sigma_{ii}\sigma_{jj}}\right)_{1 \le i,j \le d}$. If $\boldsymbol{X}$ has finite second moment, then $\mathrm{Corr}\boldsymbol{X} = \mathbf{R}$.*

Definition 2.5. *We define an* elliptical copula *as the copula of an elliptical random vector. Let $\mathbf{R}$ be the correlation matrix corresponding to $\boldsymbol{\Sigma}$. We denote the copula of $\mathcal{E}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}, G)$ by $\mathcal{EC}_d(\mathbf{R}, G)$ and call $\mathbf{R}$ the* copula correlation matrix.

The following corollary shows that the notation $\mathcal{EC}_d(\mathbf{R}, G)$ of elliptical copulae is reasonable. It is a simple consequence of the definition and the fact that copulae are invariant under strictly increasing transformations; see Embrechts et al. (2003, Theorem 2.6).

Corollary 2.6. *An elliptical copula is characterized by the generating variable G and the copula correlation matrix $\mathbf{R}$. The generating variable G is uniquely determined up to some positive constant.*

Based on elliptical copulae, we can now formulate the copula structure model.

3 Copula structure models

First, we give some notations: let $\boldsymbol{\vartheta} \in \Theta \subset \mathbb{R}^p$ be a p-dimensional parameter vector in some parameter space Θ with $\dim(\Theta) \le p$. A *correlation structure model* is then a function

$$\mathbf{R} : \Theta \mapsto \mathbb{R}^{d \times d}, \quad \boldsymbol{\vartheta} \to \mathbf{R}(\boldsymbol{\vartheta}), \tag{3.1}$$

such that $\mathbf{R}(\boldsymbol{\vartheta})$ is a correlation matrix, i.e. $\mathbf{R}(\boldsymbol{\vartheta})$ is positive definite with diagonal $\mathbf{1}$. As we will later also use vector notation, we denote by $\mathrm{vec}[\cdot]$ the column vector formed from the non-duplicated and non-fixed elements of a symmetric matrix. If a matrix $\mathbf{A}$ is not symmetric, then vec[$\mathbf{A}$] denotes the column vector formed from all non-fixed elements of the columns of $\mathbf{A}$. In case of a correlation matrix

$$\boldsymbol{r} \;:=\; \text{vec}[\mathbf{R}] \in \mathbb{R}^{d(d-1)/2}. \tag{3.2}$$

For a general linear correlation structure model, (3.1) corresponds to the following situation: let $\boldsymbol{\xi} \in \mathcal{E}_q(\mathbf{0}, \mathbf{I}, G)$ and let $\mathbf{A} : \Theta \mapsto \mathbb{R}^{d\times q}$, $\boldsymbol{\vartheta} \to \mathbf{A}(\boldsymbol{\vartheta})$, be some matrix valued function and define

$$\boldsymbol{\Sigma} : \Theta \mapsto \mathbb{R}^{d\times d}, \quad \boldsymbol{\vartheta} \to \boldsymbol{\Sigma}(\boldsymbol{\vartheta}) := \mathbf{A}(\boldsymbol{\vartheta})\mathbf{A}(\boldsymbol{\vartheta})^T.$$

Then, (3.1) can be written as $\mathbf{R}(\boldsymbol{\vartheta}) = \text{diag}[\boldsymbol{\Sigma}(\boldsymbol{\vartheta})]^{-1/2}\boldsymbol{\Sigma}(\boldsymbol{\vartheta})\text{diag}[\boldsymbol{\Sigma}(\boldsymbol{\vartheta})]^{-1/2}$.

3.1 The model

As by Definition 2.5 a correlation matrix is a parameter of an elliptical copula, we can extend the usual correlation structure model to elliptical copulae.

Definition 3.1. *Let $\boldsymbol{\vartheta} \in \Theta \subset \mathbb{R}^p$ be a p-dimensional parameter vector, $\mathbf{A} : \Theta \mapsto \mathbb{R}^{d\times q}$ a matrix valued function and $\boldsymbol{\xi} \in \mathcal{E}_q(\mathbf{0}, \mathbf{I}, G)$ a q-dimensional elliptical random vector with continuous generating variable $G > 0$. Further, denote by $C_{\mathbf{A}(\vartheta)\boldsymbol{\xi}}$ the copula of $\mathbf{A}(\boldsymbol{\vartheta})\boldsymbol{\xi} \in \mathbb{R}^d$. We say that the random vector $\boldsymbol{X} \in \mathbb{R}^d$ with copula $C_{\boldsymbol{X}}$ satisfies a* copula structure model*, if*

$$C_{\boldsymbol{X}} = C_{\mathbf{A}(\boldsymbol{\vartheta})\boldsymbol{\xi}} \in \mathcal{EC}_d(\mathbf{R}(\boldsymbol{\vartheta}), G), \tag{3.3}$$

where $\mathbf{R}(\boldsymbol{\vartheta}) := \text{diag}[\boldsymbol{\Sigma}(\boldsymbol{\vartheta})]^{-1/2}\boldsymbol{\Sigma}(\boldsymbol{\vartheta})\text{diag}[\boldsymbol{\Sigma}(\boldsymbol{\vartheta})]^{-1/2}$ *and* $\boldsymbol{\Sigma}(\boldsymbol{\vartheta}) := \mathbf{A}(\boldsymbol{\vartheta})\mathbf{A}(\boldsymbol{\vartheta})^T$.

Remark 3.2. (i) Define by $\boldsymbol{F}^{\leftarrow}(\boldsymbol{u}) := (F_1^{\leftarrow}(u_1), \ldots, F_d^{\leftarrow}(u_d))$ the vector of the inverses of the marginal distribution functions of $\boldsymbol{X}$ and by $\boldsymbol{H}(\boldsymbol{x}) := (H_1(x_1), \ldots, H_d(x_d))$ the vector of the marginal distribution functions of $\mathbf{A}(\boldsymbol{\vartheta})\boldsymbol{\xi}$. Then, (3.3) is equivalent to $\boldsymbol{X} \stackrel{\text{d}}{=} \boldsymbol{F}^{\leftarrow}(\boldsymbol{H}(\mathbf{A}(\boldsymbol{\vartheta})\boldsymbol{\xi}))$, where all operations are componentwise. Hence, the copula model can also be seen as an extension of a correlation structure model for elliptical data: if not only $C_{\boldsymbol{X}} = C_{\mathbf{A}(\boldsymbol{\vartheta})\boldsymbol{\xi}}$ holds but also $\boldsymbol{H} = \boldsymbol{F}$ with existing second moment, then this would be a classical correlation or covariance structure model. For normal $\boldsymbol{\xi}$ it gives back the standard normal model and for elliptical $\boldsymbol{\xi}$ the elliptical model of Browne (1984).

(ii) The classical correlation structure model assumes some (functional) structure for the correlation matrix of the observed data. In the copula structure model this functional structure prevails. The only difference lies in the interpretation of the 'correlation' matrix. In the classical model it represents the linear correlation between the data, in the copula model it represents a dependence parameter which can be interpreted as a 'correlation-like' measure; see Lemma 2.6.

Example 3.3. For classical factor analysis, (3.3) translates to $\boldsymbol{\vartheta} = \text{vec}[\mathbf{L}, \mathbf{V}]$, $\mathbf{R}(\boldsymbol{\vartheta}) = \mathbf{L}\mathbf{L}^T + \mathbf{V}^2$ for some $m < d$, $\mathbf{L} \in \mathbb{R}^{d \times m}$ and a diagonal matrix (with nonnegative entries) $\mathbf{V} \in \mathbb{R}^{d \times d}$. The corresponding copula structure model assumes that there exists $\boldsymbol{\xi} \in \mathcal{E}_{m+d}(\mathbf{0}, \mathbf{I}, G)$ such that

$$C_{\boldsymbol{X}} = C_{(\mathbf{L},\mathbf{V})\boldsymbol{\xi}}. \tag{3.4}$$

We call this identity a *copula factor model*. An example of this copula factor model is the *Credit Metrics* model in the framework of credit risk, see e.g. Bluhm, Overbeck, and Wagner (2003, Section 2.4). There, a factor model $\boldsymbol{X} = (X_1, \ldots, X_d)^T = \mathbf{L}\boldsymbol{f} + \mathbf{V}\boldsymbol{e}$ is assumed for the underlying (latent) variables of a set of credit default indicators $\left(I_{\{X_i < s_i\}}\right)_{1 \le i \le d}$ and $\boldsymbol{X}$ is assumed to be normal. By Frey, McNeil, and Nyfeler (2001, Proposition 2), the distribution of $\left(I_{\{X_i < s_i\}}\right)_{1 \le i \le d}$ is uniquely determined by the single default probabilities $P(I_{\{X_i < s_i\}} = 1)$ and the copula of $\boldsymbol{X}$. Therefore, in this case the assumption of $\boldsymbol{X} = \mathbf{L}\boldsymbol{f} + \mathbf{V}\boldsymbol{e}$ is equivalent to $C_{\boldsymbol{X}} = C_{(\mathbf{L},\mathbf{V})\boldsymbol{\xi}}$ with $\boldsymbol{\xi} \sim \mathcal{N}_{m+d}(\mathbf{0}, \mathbf{I})$. The model extends easily to non-normal $\boldsymbol{X}$.

3.2 Estimation of $\boldsymbol{\vartheta}$

The next step is to estimate a structure model. Let $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ be an iid sequence of d-dimensional random vectors and denote by $\widehat{\mathbf{R}} := \widehat{\mathbf{R}}(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n)$ an estimator of $\mathbf{R}$, a correlation matrix. This estimator can be the empirical correlation or a copula correlation estimator or some other correlation estimator. We review some results from the literature, which we will need for the estimation of the copula structure model later.

Given this estimator $\widehat{\mathbf{R}}$ we want to find some parameter vector $\boldsymbol{\vartheta}$ which fits the assumed structure $\mathbf{R}(\boldsymbol{\vartheta})$ to $\widehat{\mathbf{R}}$ as good as possible. Similarly to (3.2), we define $\widehat{\boldsymbol{r}} := \text{vec}[\widehat{\mathbf{R}}]$ and $\boldsymbol{r}(\boldsymbol{\vartheta}) := \text{vec}[\mathbf{R}(\boldsymbol{\vartheta})]$.

Estimation of $\boldsymbol{\vartheta}$ is based on the minimization of a *discrepancy function* $D = D(\widehat{\boldsymbol{r}}, \boldsymbol{r}(\boldsymbol{\vartheta}))$ which measures the discrepancy between the estimated correlation matrix represented by $\widehat{\boldsymbol{r}}$ and $\boldsymbol{r}(\boldsymbol{\vartheta})$. A discrepancy function D has to satisfy

(i) $D \ge 0$,

(ii) $D(\widehat{\boldsymbol{r}}, \boldsymbol{r}) = 0$ if and only if $\widehat{\boldsymbol{r}} = \boldsymbol{r}$ and

(iii) D is twice differentiable with respect to both $\widehat{\boldsymbol{r}}$ and $\boldsymbol{r}$.

Note that the concept of a discrepancy function (without condition (iii)) is weaker than the concept of a metric, as a discrepancy function D does not have to be symmetric or translation invariant in its arguments, nor does it have to satisfy the triangular inequality.

In the following example we present two classical discrepancy functions, for more details about discrepancy functions, their properties, advantages and drawbacks, we refer to Bentler and Dudgeon (1996) and Steiger (1994). For more details about the quadratic form discrepancy function below see Steiger, Shapiro, and Browne (1985).

Example 3.4. (i) The *normal theory maximum likelihood discrepancy function* is

$$D_{\mathrm{ML}}(\widehat{\boldsymbol{r}}, \boldsymbol{r}(\boldsymbol{\vartheta})) = \ln|\mathbf{R}(\boldsymbol{\vartheta})| + \operatorname{tr}\left(\widehat{\mathbf{R}}\,(\mathbf{R}(\boldsymbol{\vartheta}))^{-1}\right) - \ln|\widehat{\mathbf{R}}| - d. \tag{3.5}$$

This function is the log-likelihood term of $\mathbf{R}(\boldsymbol{\vartheta})$ in case of normal data.

(ii) The *quadratic form* (or *weighted least squares*) *discrepancy function* is

$$D_{\mathrm{QD}}(\widehat{\boldsymbol{r}}, \boldsymbol{r}(\boldsymbol{\vartheta})|\,\boldsymbol{\Upsilon}) = (\widehat{\boldsymbol{r}} - \boldsymbol{r}(\boldsymbol{\vartheta}))^T\,\boldsymbol{\Upsilon}^{-1}\,(\widehat{\boldsymbol{r}} - \boldsymbol{r}(\boldsymbol{\vartheta}))\,, \tag{3.6}$$

where $\boldsymbol{\Upsilon}$ is a positive definite matrix or a consistent estimator of some positive definite matrix $\boldsymbol{\Upsilon}^*$. Note that $D_{\mathrm{QD}}(\cdot,\cdot|\,\boldsymbol{\Upsilon})$ is a metric.

Given some discrepancy function D and some estimator $\widehat{\mathbf{R}}$ of the correlation matrix $\mathbf{R}$, we can define a consistent estimator of $\boldsymbol{\vartheta}$.

Proposition 3.5 (Browne (1984), Proposition 1). *Let $\mathbf{R}_0$ be some correlation matrix, $\boldsymbol{r}_0 := \mathrm{vec}[\mathbf{R}_0] \in \mathbb{R}^{d(d-1)/2}$ and $\Theta \subset \mathbb{R}^p$ a closed and bounded parameter space. Further assume that $\widehat{\boldsymbol{r}}$ is an estimator based on an iid sample $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ of d-dimensional random vectors and let D be a discrepancy function. Assume that $\widehat{\boldsymbol{r}} \xrightarrow{P} \boldsymbol{r}_0$ as $n \to \infty$ and that $\boldsymbol{\vartheta}_0 \in \Theta$ is the unique minimizer of $D(\boldsymbol{r}_0, \boldsymbol{r}(\boldsymbol{\vartheta}))$ in Θ. Assume also that the Jacobian matrix $\partial\boldsymbol{r}(\boldsymbol{\vartheta})/\partial\boldsymbol{\vartheta}^T$ is continuous in $\boldsymbol{\vartheta}$. Define the estimator*

$$\widehat{\boldsymbol{\vartheta}} := \arg\min_{\boldsymbol{\vartheta}\in\Theta} D(\widehat{\boldsymbol{r}}, \boldsymbol{r}(\boldsymbol{\vartheta})). \tag{3.7}$$

Then $\widehat{\boldsymbol{\vartheta}} \xrightarrow{P} \boldsymbol{\vartheta}_0$ as $n \to \infty$.

Of course, if we know the true correlation vector $\boldsymbol{r}_0$ satisfying the structure model $\boldsymbol{r}_0 = \boldsymbol{r}(\boldsymbol{\vartheta}_0)$ for some parameter vector $\boldsymbol{\vartheta}_0$, then $\widehat{\boldsymbol{\vartheta}}$ will always be such that $\boldsymbol{r}_0 = \boldsymbol{r}(\boldsymbol{\vartheta}_0) = \boldsymbol{r}(\widehat{\boldsymbol{\vartheta}})$, independent of the choice of the discrepancy function. We also have $D(\boldsymbol{r}_0, \boldsymbol{r}(\widehat{\boldsymbol{\vartheta}})) = 0$ in this case. Since in practice we neither know the true $\boldsymbol{r}_0$ nor the true structure model, we need a method to find an appropriate model.

3.3 Model selection

First, we show the asymptotic distribution of a certain test statistic, which will later be used for model selection.

Definition 3.6. *Under the settings of Proposition **??**, we define the test statistic*

$$T := n\widehat{D} = nD(\widehat{\boldsymbol{r}}, \boldsymbol{r}(\widehat{\boldsymbol{\vartheta}})) = n\min_{\boldsymbol{\vartheta}\in\Theta} D(\widehat{\boldsymbol{r}}, \boldsymbol{r}(\vartheta)). \tag{3.8}$$

The null hypothesis is that the true correlation vector $\boldsymbol{r}_0$ satisfies a structure model, i.e.

$$H_0: \quad \boldsymbol{r}_0 = \boldsymbol{r}(\boldsymbol{\vartheta}_0) \quad \textit{for some } \boldsymbol{\vartheta}_0 \in \Theta. \tag{3.9}$$

To obtain the limit distribution of T we use a version of Steiger et al. (1985, Theorem 1), adapted to our situation. We replace the regularity condition (R7) of that article by the stronger assumption that the null hypothesis (3.9) holds. The equivalent statement in case of the quadratic form discrepancy function $D_{\mathrm{QD}}(\cdot,\cdot|\boldsymbol{\Upsilon})$ is given in Browne (1984, Corollary 4.1), where it is additionally required that $\boldsymbol{\Upsilon}$ is a consistent estimator of $\boldsymbol{\Gamma}$, the asymptotic covariance matrix of $\widehat{\boldsymbol{r}}$.

Theorem 3.7. *Assume that the conditions of Proposition* **??** *hold and* $\boldsymbol{\vartheta}_0$ *is an interior point of* Θ. *Further assume that* $\sqrt{n}(\widehat{\boldsymbol{r}}-\boldsymbol{r}_0)\xrightarrow{d}\mathcal{N}(\mathbf{0},\boldsymbol{\Gamma})$ *as* $n\to\infty$ *and that the Hessian matrix*

$$2\boldsymbol{\Psi}_0 \;=\; \left.\frac{\partial^2 D(\boldsymbol{r},\boldsymbol{\xi})}{\partial\boldsymbol{\xi}\partial\boldsymbol{\xi}^T}\right|_{\boldsymbol{r}=\boldsymbol{\xi}=\boldsymbol{r}_0} \tag{3.10}$$

is nonsingular and satisfies $\boldsymbol{\Psi}_0=\boldsymbol{\Gamma}^{-1}$. *In case of the quadratic form discrepancy function* $D_{\mathrm{QD}}(\cdot,\cdot|\,\boldsymbol{\Upsilon})$ *defined in* (3.6), *the assumption* (3.10) *is replaced by assuming that* $\boldsymbol{\Upsilon}$ *is a consistent estimator of* $\boldsymbol{\Gamma}$. *Also assume that the* $p\times d$ *Jacobian matrix*

$$\triangle \;=\; \left.\frac{\partial\boldsymbol{r}(\boldsymbol{\vartheta})}{\partial\boldsymbol{\vartheta}}\right|_{\boldsymbol{\vartheta}=\boldsymbol{\vartheta}_0} \tag{3.11}$$

is of full column rank p. *Then, under the null hypothesis* (3.9),

$$T \;=\; n\widehat{D} \;\xrightarrow{d}\; \chi^2_{df}\,,\quad n\to\infty\,, \tag{3.12}$$

where $df=d(d-1)/2-p^*$ *with* $p^*\le p$ *is the number of free parameters of* $\boldsymbol{\vartheta}\in\Theta\subset\mathbb{R}^p$.

Remark 3.8. Under the conditions of Proposition **??**, if $\boldsymbol{\Psi}_0=\boldsymbol{\Gamma}^{-1}$ does not hold, the limiting distribution of T in (3.8) under the null hypothesis (3.9) will not be χ^2_{df}, see Satorra and Bentler (2001) or van Praag, Dijkstra, and van Velzen (1985). In this case,

$$T \;\xrightarrow{d}\; \sum_{j=1}^{df}\kappa_j\zeta_j\,,\quad n\to\infty\,,$$

where the ζ_j are iid χ^2_1 distributed and κ_j are the non-null eigenvalues of the matrix $\mathbf{U}\boldsymbol{\Gamma}$ with

$$\mathbf{U} \;=\; \boldsymbol{\Psi}_0-\boldsymbol{\Psi}_0\triangle(\triangle^T\boldsymbol{\Psi}_0\triangle)^{-1}\triangle^T\boldsymbol{\Psi}_0,$$

where $\triangle$ is given in (3.11). An example for this situation is $D_{\mathrm{ML}}(\widehat{\boldsymbol{\sigma}},\boldsymbol{\sigma}(\boldsymbol{\vartheta}))$ given in (3.5), where $\widehat{\boldsymbol{\sigma}}$ is the vector of a covariance matrix estimator, $\boldsymbol{\sigma}(\boldsymbol{\vartheta})$ is the vector of a covariance structure model and $\widehat{\boldsymbol{\sigma}}$ has an asymptotic covariance matrix different from the asymptotic covariance matrix of the empirical covariance estimator under a normal population.

From now on we will use the quadratic form discrepancy function $D := D_{\mathrm{QD}}$ from Example 3.4(ii), where $\boldsymbol{\Upsilon} = \widehat{\boldsymbol{\Gamma}}$ is an estimator of $\boldsymbol{\Gamma}$. If $\widehat{\boldsymbol{\Gamma}}$ is consistent, Theorem 3.7 applies and by Browne (1984, Corollary 2.1), $\widehat{\boldsymbol{\vartheta}}$ is asymptotically normal with covariance matrix $\left(\triangle^T \boldsymbol{\Gamma}^{-1} \triangle\right)^{-1}$, where $\triangle$ is given in (3.11). Note that, if $\widehat{\boldsymbol{\Gamma}}$ is only consistent and does not have a finite second moment, large sample sizes may be necessary to observe the limiting χ^2-distribution of the test statistic T or the asymptotic normality of $\widehat{\boldsymbol{\vartheta}}$.

To select an appropriate structural model, we consider a set of g models (which all have to satisfy the assumptions of Theorem 3.7)

$$\boldsymbol{r}^{(i)} : \Theta^{(i)} \to \mathbb{R}^{d(d-1)/2}, \quad \boldsymbol{\vartheta}^{(i)} \mapsto \boldsymbol{r}^{(i)}(\boldsymbol{\vartheta}^{(i)}), \quad \text{and } \Theta^{(i)} \subset \mathbb{R}^{p^{(i)}}, \quad 1 \leq i \leq g. \tag{3.13}$$

Further, we require that the g models are *nested*, i.e. for every $1 \leq i \leq g-1$ and $\boldsymbol{\vartheta}^{(i)} \in \Theta^{(i)}$ there exists some $\boldsymbol{\vartheta}^{(i+1)} \in \Theta^{(i+1)}$ such that $\boldsymbol{r}^{(i+1)}(\boldsymbol{\vartheta}^{(i+1)}) = \boldsymbol{r}^{(i)}(\boldsymbol{\vartheta}^{(i)})$. Next, define the null hypotheses

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

and assume that some of these null hypotheses are true. Then there exists some j such that $H_0^{(i)}$ does not hold for $1 \leq i < j$ and does hold for $j \leq i \leq g$. As we are interested in a structure model, which is likely to explain the observed dependence structure, and is as simple as possible and, since the models are nested, we have to estimate j, the smallest index where the null hypothesis holds. By Theorem 3.7, the corresponding test statistics $T^{(i)} := nD(\widehat{\boldsymbol{r}}, \boldsymbol{r}^{(i)}(\vartheta_0^{(i)})) := n \min_{\boldsymbol{\vartheta} \in \Theta^{(i)}} D(\widehat{\boldsymbol{r}}, \boldsymbol{r}^{(i)}(\vartheta))$ are not χ^2 distributed for $1 \leq i < j$ and are χ^2_{df}-distributed for $j \leq i \leq g$ with *df* given in Theorem 3.7. Consequently, we reject a null hypothesis $H_0^{(i)}$, if the corresponding test statistic $T^{(i)}$ is larger than some χ^2_{df}-quantile. Hence, j is the smallest number, where $H_0^{(j)}$ cannot be rejected.

Remark 3.9. (i) Note that classical estimates of $\boldsymbol{\Gamma}$ rely on the estimation of second and fourth moments of $\boldsymbol{X}$. For non-normal or, especially, for heavy-tailed data these estimates often have large sampling variability and in simulation studies it turned out that large samples are necessary for acceptable performance of the test statistics, see e.g. Hu, Bentler, and Kano (1992).

(ii) In general, a unique *true* parameter $\boldsymbol{\vartheta}_0$ need not exist: in the classical factor model (see Example 3.3, where $\mathbf{R} = \mathbf{L}\mathbf{L}^T + \mathbf{V}^2$), $\mathbf{L}$ can always be replaced by $\mathbf{L}^* = \mathbf{L}\mathbf{P}$ for any orthogonal matrix $\mathbf{P}$. By a minor adaption of the parameter space Θ (i.e. $\mathbf{L}^T\mathbf{V}^{-2}\mathbf{L}$ has to be diagonal), $\widehat{\boldsymbol{\vartheta}}$ can be forced to be unique and Proposition **??** applies, see Lawley and Maxwell (1971, Section 2.3). By Lee and Bentler (1980) the degrees of freedom in (3.12) are then increased by the number of additional constraints. For better interpretation, the factors can be rotated after estimation, e.g. with the *varimax* method, for details see Anderson (2003, chapter 14).

(iii) With the correction for uniqueness in (ii) above, the factor model of Example 3.3 satisfies the regularity conditions of Proposition **??** and Theorem 3.7, see Steiger et al. (1985, Section 4) and Browne (1984, Section 5).

(iv) In case of the copula factor model (see Remark 3.2(iii)) we only need to estimate the loading matrix $\mathbf{L} \in \mathbb{R}^{d\times m}$, since $\text{diag}(\mathbf{V}^2) = \mathbf{1} - \text{diag}(\mathbf{L}\mathbf{L}^T)$. Therefore the number of free parameters are dm minus the number of the additional constraints to ensure that $\mathbf{L}^T\mathbf{V}^{-2}\mathbf{L}$ is diagonal, i.e. the degrees of freedom of the limiting χ^2 distribution are $df = d(d-1)/2 - dm + m(m-1)/2$.

(v) For the quadratic form discrepancy function $D(\cdot,\cdot|\widehat{\boldsymbol{\Gamma}})$, where $\widehat{\boldsymbol{\Gamma}}$ is a consistent estimator of $\boldsymbol{\Gamma}$, it can be shown that $T^{(i)}$, $1 \le i < j$, has an approximate noncentral χ^2_{df}-distribution with non-centrality parameter $nD(\boldsymbol{r}_0, \boldsymbol{r}^{(i)}(\boldsymbol{\vartheta}_0^{(i)})|\boldsymbol{\Gamma})$, see Browne (1984, Corollary 4.1).

4 Methodology

As we consider a copula structure model, we need an estimator $\widehat{\mathbf{R}}$ of the copula correlation matrix $\mathbf{R}$, whose limit distribution is $\mathcal{N}(\mathbf{0},\boldsymbol{\Gamma})$ for some non-degenerate covariance matrix $\boldsymbol{\Gamma}$ and a consistent estimator of $\boldsymbol{\Gamma}$. In the following we will introduce two copula based dependence concepts and their corresponding correlation and asymptotic covariance estimators (which are also consistent and asymptotically normal).

4.1 Dependence Concepts

A well known dependence concept is (linear) correlation or covariance, which is limited by the fact that it measures only linear dependence. Further, since correlation is not invariant under non-linear (strictly increasing) transformations, it is not a copula property. As we want for our copula structure analysis a dependence concept which is at least related to correlation we use the following one known as *Kendall's tau*.

This copula-based dependence concept provides a good alternative to the linear correlation as a measure also for non-elliptical distributions, for which linear correlation is an inappropriate measure of dependence and often misleading. Originally, it has been suggested as a robust dependence measure, which makes it also appropriate for heavy-tailed data; for more details see Kendall and Gibbons (1990).

Definition 4.1. *Kendall's tau τ_{ij} between two components (X_i, X_j) of a random vector $\boldsymbol{X}$ is defined as*

$$\tau_{ij} := P\left((X_i - \widetilde{X}_i)(X_j - \widetilde{X}_j) > 0\right) - P\left((X_i - \widetilde{X}_i)(X_j - \widetilde{X}_j) < 0\right),$$

where $(\widetilde{X}_i, \widetilde{X}_j)$ *is an independent copy of* (X_i, X_j)*. Moreover, we call* $\mathbf{T} := (\tau_{ij})_{1 \le i,j \le d}$ *the* Kendall's tau matrix.

Concerning elliptical copulae the following result is given in Lindskog, McNeil, and Schmock (2003, Theorem 2).

Theorem 4.2. *Let* $\boldsymbol{X}$ *be a vector of random variables with elliptical copula* $C \sim \mathcal{EC}_d(\mathbf{R}, G)$ *and continuous generating variable* $G > 0$*, then* $\tau_{ij} = 2 \arcsin(\rho_{ij})/\pi$*.*

Considering extreme observations, we need the concept of regular variation. A textbook treatment of this topic is to be found in Bingham, Goldie, and Teugels (1989), for a multivariate extension we refer to Resnick (1987, 2004) or Basrak, Davis, and Mikosch (2002).

Definition 4.3. *A random variable* G *is called* regularly varying at infinity with index $-\alpha$*,* $0 < \alpha < \infty$*, if* $\lim_{x \to \infty} P(G > tx)/P(G > x) = t^{-\alpha}$*, for all* $t > 0$*. We write* $G \in RV_{-\alpha}$*.*

In financial risk management, one is often interested only in the dependence of extreme observations. By Sklar's theorem, the copula is sufficient to describe dependence in extremes. As C is a uniform distribution on $[0,1]^d$, extreme values happen near the boundaries and extreme dependence happens around the points $(0, \dots, 0)$ and $(1, \dots, 1)$. This can be captured by the following concept.

Definition 4.4. *(i) We define the* upper tail copula *of* $\boldsymbol{X}$ *as*

$$\begin{aligned} \lambda^{\boldsymbol{X}}_{\text{upper}}(\boldsymbol{x}) &= \lambda^{\boldsymbol{X}}_{\text{upper}}(x_1, \dots, x_d) \\ &= \lim_{t \to 0} t^{-1} P\left(1 - F_1(X_1) \le tx_1, \dots,\ 1 - F_d(X_d) \le tx_d\right), \end{aligned} \tag{4.1}$$

for $x_1, \dots, x_d \ge 0$ *if the limit exists.*

(ii) We define the lower tail copula *of* $\boldsymbol{X}$ *as*

$$\lambda^{\boldsymbol{X}}_{\text{lower}}(\boldsymbol{x}) := \lim_{t \to 0} t^{-1} P\left(F_1(X_1) \le tx_1, \dots, F_d(X_d) \le tx_d\right). \tag{4.2}$$

for $x_1, \dots, x_d \ge 0$ *if the limit exists.*

Remark 4.5. Since by symmetry $\lambda^{\boldsymbol{X}}_{\text{lower}}(\boldsymbol{x}) = \lambda^{\boldsymbol{X}}_{\text{upper}}(\boldsymbol{x}) =: \lambda^{\boldsymbol{X}}(\boldsymbol{x})$ holds for elliptical copulae (see Definitions 2.3 and 2.5), we concentrate only on the upper tail copula and call it *tail copula*. Of course, by definition, the tail copula is a copula property. For more details about the tail copula, see Schmidt and Stadtmüller (2005).

Notions like tail copula or tail dependence function go back to Gumbel (1960), Pickands (1981) and Galambos (1987), and they represent the full dependence structure of the model in the extremes. If $\lambda^{\boldsymbol{X}}(\boldsymbol{x}) > 0$ for some $\boldsymbol{x} > \boldsymbol{0}$, $\boldsymbol{X}$ is called *asymptotically dependent* and *asymptotically independent*, otherwise. Assuming elliptical copulae, Hult and Lindskog (2002, Theorem 4.3) show that $\boldsymbol{X}$ is asymptotically dependent if $\boldsymbol{X}$ has an elliptical copula with regularly varying generating variable $G \in RV_{-\alpha}$, $\alpha > 0$. For a textbook treatment of multivariate extremes, see Resnick (1987).

By definition, $\lambda^{\boldsymbol{X}}(\boldsymbol{x}) = 0$ if $\lambda^{(X_i,X_j)}(x_i, x_j) = 0$ for some i, j and $\boldsymbol{x} > \mathbf{0}$, i.e. $\boldsymbol{X}$ is asymptotically independent if some bivariate margin (X_i, X_j) of $\boldsymbol{X}$ is asymptotically independent. Concerning asymptotic independence we refer to Ledford and Tawn (1996, 1997), and for a conditional modeling and estimation approach allowing for asymptotic independence in some components and asymptotic dependence in others, see Heffernan and Tawn (2004). We will use the assumption of asymptotic dependence for modeling and estimation and therefore we omit further discussions about asymptotic independence.

For estimation of $\mathbf{R}$ we only need a representation of the bivariate marginal tail copula (4.1) for elliptical copulae. It follows from Hult and Lindskog (2002, Corollary 3.1), Klüppelberg, Kuhn, and Peng (2005a, Theorem 2.1) and transformation of variable. A representation of the full multivariate version is given in Klüppelberg, Kuhn, and Peng (2005b, Theorem 5.1).

Theorem 4.6. *Suppose $\boldsymbol{X}$ has copula $C_{\boldsymbol{X}} \in \mathcal{EC}_d(\mathbf{R}, G)$ with generating variable $G \in RV_{-\alpha}$, $\alpha > 0$, and copula correlation matrix $\mathbf{R} = (\rho_{ij})_{1\le i,j\le d}$ with $\max|\rho_{ij}| < 1$. Then the* bivariate marginal tail copula *of $\boldsymbol{X}$ is given by*

$$\begin{aligned}\lambda^{\boldsymbol{X}}_{ij}(x, y) &:= \lambda^{\boldsymbol{X}}(\infty, \dots, \infty, x, \infty, \dots, \infty, y, \infty, \dots, \infty) \\ &= \left(x \int_{g_{ij}\left((x/y)^{1/\alpha}\right)}^{\pi/2} (\cos\phi)^{\alpha}\, \mathrm{d}\phi + y \int_{g_{ij}\left((x/y)^{-1/\alpha}\right)}^{\pi/2} (\cos\phi)^{\alpha}\, \mathrm{d}\phi \right) \left(\int_{-\pi/2}^{\pi/2} (\cos\phi)^{\alpha}\, d\phi \right)^{-1} \\ &=: \lambda(x, y, \alpha, \rho_{ij}), \end{aligned} \tag{4.3}$$

where x is the i-th, y the j-th component and $g_{ij}(t) := \arctan\left((t - \rho_{ij})/\sqrt{1-\rho_{ij}^2}\right)$.

Remark 4.7. The case of $\rho := \rho_{ij} = 1$ can be interpreted as a limit, i.e.

$$\lambda(x, y, \alpha, 1) := \lim_{\rho\to 1} \lambda(x, y, \alpha, \rho).$$

Then

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

and we obtain $\lambda(x, y, \alpha, 1) = x \wedge y$. Similarly, $\lambda(x, y, \alpha, -1) = 0$.

This bivariate marginal tail copula $\lambda^{\boldsymbol{X}}_{ij}$ given in (4.3) measures the amount of dependence in the upper right quadrant of (X_i, X_j). Note that by Klüppelberg et al. (2005b, Theorem 5.1), $\lambda^{\boldsymbol{X}}$ is completely characterized by the copula correlation matrix $\mathbf{R}$ and the index α of regular variation of G.

By Theorems 4.2 and 4.6 we see that for an elliptical copula the correlation matrix $\mathbf{R}$ is a function of Kendall's tau or of the tail copula with the index α of regular variation of G. In Sections 4.2 and 4.3 we will invoke this functional relationship for the estimation of $\mathbf{R}$. The two approaches differ in their interpretation: estimating $\mathbf{R}$ via Kendall's tau fits a robust dependence structure of the data to an elliptical copula. Using the tail copula for estimation of $\mathbf{R}$ fits only the dependence structure in the upper extremes to an elliptical copula and does not necessarily fit the dependence of the data in other regions. Of course, copula structure analysis can be applied to any copula correlation estimator with a certain limiting behavior as given by Theorem 3.7. Using Kendall's tau for estimation can then be seen as a robust extension of the usual correlation structure analysis, whereas using the tail copula provides a structure analysis of dependence in the extremes. The next two sections explain the estimation procedures and give asymptotic results.

4.2 Copula correlation estimator based on Kendall's tau

The first method is based on Kendall's tau, which can be used for estimating the correlation matrix $\mathbf{R}$ by Theorem 4.2. For a general treatment of U-statistics see Lee (1990); the results we use go back to Hoeffding (1948).

Definition 4.8. *Given an iid sample $\boldsymbol{X}_1, \dots, \boldsymbol{X}_n$, $\boldsymbol{X}_i = (X_{i,1}, \dots, X_{i,d})^T$, we define the estimator $\widehat{\mathbf{T}} = (\widehat{\tau}_{ij})_{1 \le i,j \le d}$ of Kendall's tau matrix $\mathbf{T}$ by $\widehat{\tau}_{ii} = 1$ for $i = 1, \dots, d$ and*

$$\widehat{\tau}_{ij} \;=\; \binom{n}{2}^{-1} \sum_{1 \le l < k \le n} \operatorname{sign}\left((X_{k,i} - X_{l,i})\,(X_{k,j} - X_{l,j}) \right), \quad 1 \le i \ne j \le d.$$

Estimating the copula correlation matrix via Kendall's tau yields the following result. Its proof can be found in Section 6.

Theorem 4.9. *Let $\boldsymbol{X}_1, \boldsymbol{X}_2, \dots$ be an iid sequence of d-dimensional random vectors with elliptical copula $C_{\boldsymbol{X}} \in \mathcal{EC}_d(\mathbf{R}, G)$ with continuous G. Further, define*

$$\widehat{\mathbf{R}}_\tau = (\widehat{\rho}^{\,\tau}_{ij})_{1 \le i,j \le d} \;:=\; \sin\left(\frac{\pi}{2} \widehat{\mathbf{T}} \right), \tag{4.4}$$

where the 'sin' *is used componentwise and define $\widehat{\boldsymbol{r}}_\tau := \operatorname{vec}[\widehat{\mathbf{R}}_\tau]$ and $\boldsymbol{r} := \operatorname{vec}[\mathbf{R}]$. Then,*

$$\sqrt{n}\,(\widehat{\boldsymbol{r}}_\tau - \boldsymbol{r}) \;\xrightarrow{d}\; \mathcal{N}_{d(d-1)/2}(\mathbf{0}, \boldsymbol{\Gamma}_\tau), \quad n \to \infty,$$

holds, where $\boldsymbol{\Gamma}_\tau = (\gamma^\tau_{ij,kl})_{1 \le i \ne j, k \ne l \le d}$ with

$$\gamma^\tau_{ij,kl} \;=\; \pi^2 \cos\left(\frac{\pi}{2} \tau_{ij} \right) \cos\left(\frac{\pi}{2} \tau_{kl} \right) (\tau_{ij,kl} - \tau_{ij}\tau_{kl}) \quad \textit{and} \tag{4.5}$$

$$\tau_{ij,kl} \;=\; E\Big(E\left(\operatorname{sign}\left[(X_{1,i} - X_{2,i})(X_{1,j} - X_{2,j}) \right] \,\big|\, \boldsymbol{X}_1 \right) E\left(\operatorname{sign}\left[(X_{1,k} - X_{2,k})(X_{1,l} - X_{2,l}) \right] \,\big|\, \boldsymbol{X}_1 \right) \Big).$$

By (4.5), an estimator of $\mathbf{\Gamma}_\tau = (\gamma^\tau_{ij,kl})_{1\le i\ne j, k\ne l\le d}$ can be defined by its empirical version.

Definition 4.10. *Given an iid sample $\boldsymbol{X}_1,\dots,\boldsymbol{X}_n$, $\boldsymbol{X}_i = (X_{i,1},\dots,X_{i,d})^T$, we define the estimator $\widehat{\mathbf{\Gamma}}_\tau = (\widehat{\gamma}^\tau_{ij,kl})_{1\le i\ne j,k\ne l\le d}$, where*

$$\widehat{\gamma}^\tau_{ij,kl} \;:=\; \pi^2\cos\left(\frac{\pi}{2}\widehat{\tau}_{ij}\right)\cos\left(\frac{\pi}{2}\widehat{\tau}_{kl}\right)\left(\widehat{\tau}_{ij,kl}-\widehat{\tau}_{ij}\widehat{\tau}_{kl}\right) \quad \textit{and} \tag{4.6}$$

$$\widehat{\tau}_{ij,kl} \;:=\; \frac{1}{n(n-1)^2}\sum_{p=1}^{n}\left[\left(\sum_{q=1,\,q\ne p}^{n}\operatorname{sign}\left((X_{p,i}-X_{q,i})(X_{p,j}-X_{q,j})\right)\right)\times \right.$$
$$\left.\times\left(\sum_{q=1,\,q\ne p}^{n}\operatorname{sign}\left((X_{p,k}-X_{q,k})(X_{p,l}-X_{q,l})\right)\right)\right]. \tag{4.7}$$

The following result is also proved in Section 6.

Theorem 4.11. *The estimator* $\mathrm{vec}[\widehat{\mathbf{\Gamma}}_\tau]$ *is consistent and asymptotically normal.*

4.3 Copula correlation estimator based on the tail copula

The second estimation method is based on the tail copula. If someone is interested in the dependence structure of the extreme data and assumes an elliptical copula, (4.1) shows how $\lambda^{\boldsymbol{X}}$ can be expressed as a function of $\mathbf{R}$ and α. By estimation of $\lambda^{\boldsymbol{X}}$ one can estimate $\mathbf{R}$ and α (i.e. the elliptical structure), which is likely to generate the observed extreme dependence.

We use an approach closely related to Klüppelberg et al. (2005b); i.e. we use the tail copula for the estimation of $\mathbf{R}$ and α. By Theorem 4.6 we need an estimator of α and of all bivariate marginal tail copulae. We start with an empirical tail copula estimator, for details see Klüppelberg et al. (2005a, 2005b) (and references therein) and estimate $\mathbf{R}$ and α from this.

Definition 4.12. *Given an iid sample $\boldsymbol{X}_1,\dots,\boldsymbol{X}_n$, $\boldsymbol{X}_l = (X_{l,1},\dots,X_{l,d})^T$, we define the empirical tail copula estimator for $\boldsymbol{x}=(x_1,\dots,x_d) > \mathbf{0}$ as*

$$\widehat{\lambda}^{\mathrm{emp}}(\boldsymbol{x};k) \;=\; \frac{1}{k}\sum_{l=1}^{n} I\left(1-\widehat{F}_j(X_{lj}) \le \frac{k}{n}x_j,\; j=1,\dots,d\right), \tag{4.8}$$

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

$$\widehat{\lambda}^{\mathrm{emp}}_{ij}(x,y;k) \;:=\; \widehat{\lambda}^{\mathrm{emp}}(\infty,\dots,\infty,x,\infty,\dots,\infty,y,\infty,\dots,\infty)$$
$$:=\; \frac{1}{k}\sum_{l=1}^{n} I\left(1-\widehat{F}_i(X_{li}) \le \frac{k}{n}x\,, 1-\widehat{F}_j(X_{lj}) \le \frac{k}{n}y\right), \tag{4.9}$$

where x is at the i-th and y at the j-th component, respectively.

Since $\widehat{\lambda}^{\text{emp}}$ estimates the tail copula, the number k should be small in comparison to n. Setting $x_j = 1$, $1 \le j \le d$, only the k largest observations of $X_{l,j}$ satisfy $1 - \widehat{F}_j(X_{lj}) \le k/n$, 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 theory.

Immediately by definition (4.1), $\lambda^{\boldsymbol{X}}$ is homogenous of order 1, and, for large k and n, also $\widehat{\lambda}_{ij}^{\text{emp}}$ is (see (4.8)). Consequently, setting $\theta = \arctan(y/x)$, i.e. $(x, y) = (c\cos\theta, c\sin\theta)$ for some constant $c > 0$, we have $\widehat{\lambda}_{ij}^{\text{emp}}(x, y; k) = \widehat{\lambda}_{ij}^{\text{emp}}(\sqrt{2}\cos\theta, \sqrt{2}\sin\theta; k^*)$ for some appropriate k^*. Hence, for the estimation, we follow the convention and only consider points $(x, y) = (\sqrt{2}\cos\theta, \sqrt{2}\sin\theta)$, $\theta \in (0, \pi/2)$.

For estimation of α we use the approach of Klüppelberg et al. (2005b), which is based on inversion of the tail copula with respect to α.

Definition 4.13. *Define $\lambda^{\leftarrow\alpha}(\cdot; x, y, \rho)$ as the inverse of $\lambda(x, y, \alpha, \rho)$ (given in (4.1)) with respect to α and, using $\widehat{\rho}_{ij}^{\tau}$ given in (4.4) and $\widehat{\lambda}^{\text{emp}}$ given in (4.9), define for $i \neq j$*

$$\begin{aligned}
\widehat{Q}_{ij} &:= \left\{\theta \in \left(0, \frac{\pi}{2}\right) : \widehat{\lambda}_{ij}^{\text{emp}}(\sqrt{2}\cos\theta, \sqrt{2}\sin\theta; k) < \right. \\
&\qquad\qquad \left. < \lambda\left(\sqrt{2}\cos\theta, \sqrt{2}\sin\theta, \left|\frac{\ln(\tan\theta)}{\ln(\widehat{\rho}_{ij}^{\tau} \vee 0)}\right|, \widehat{\rho}_{ij}^{\tau}\right)\right\}, \\
\widehat{Q}_{ij}^* &:= \left\{\theta \in \left(0, \frac{\pi}{2}\right) : |\ln(\tan\theta)| < \left(1 - k^{-1/4}\right)\widetilde{\alpha}_{ij}(1, 1; k)\left|\ln(\widehat{\rho}_{ij}^{\tau} \vee 0)\right|\right\} \quad \textit{and} \\
Q_{ij}^* &:= \left\{\theta \in \left(0, \frac{\pi}{2}\right) : |\ln(\tan\theta)| < \alpha\,|\ln(\rho_{ij} \vee 0)|\right\},
\end{aligned}$$

where for $\theta \in \widehat{Q}_{ij}$ we define $\widetilde{\alpha}_{ij}$ as the estimator of α based on the empirical bivariate tail copula (4.9)

$$\begin{aligned}
&\widetilde{\alpha}_{ij}(\sqrt{2}\cos\theta, \sqrt{2}\sin\theta; k) \\
&\quad := \lambda^{\leftarrow\alpha}\left(\widehat{\lambda}_{ij}^{\text{emp}}(\sqrt{2}\cos\theta, \sqrt{2}\sin\theta; k); \sqrt{2}\cos\theta, \sqrt{2}\sin\theta, \widehat{\rho}_{ij}^{\tau}\right).
\end{aligned} \tag{4.10}$$

Further, let w be a nonnegative weight function. Then we define the smoothed estimator $\widehat{\alpha}$ of α as

$$\widehat{\alpha}(k, w) := \frac{2}{d(d-1)} \sum_{1 \le i < j \le d} \frac{1}{W\left(\widehat{Q}_{ij} \cap \widehat{Q}_{ij}^*\right)} \int_{\theta \in \widehat{Q}_{ij} \cap \widehat{Q}_{ij}^*} \widetilde{\alpha}_{ij}(\sqrt{2}\cos\theta, \sqrt{2}\sin\theta; k)\, W(d\theta), \tag{4.11}$$

where W is the measure induced by w.

To define an estimator of $\mathbf{R}$ via extreme observations, we invert the bivariate tail copula with respect to ρ. Using (4.3) it is straightforward to show the following.

Lemma 4.14. *For fixed x, y, $\alpha > 0$ define $\rho^* := \left((x \wedge y)/(x \vee y)\right)^{1/\alpha}$. Then, for all $\rho < \rho^*$, $\frac{\partial}{\partial\rho}\lambda(x, y, \alpha, \rho) > 0$ holds and the inverse $\lambda^{\leftarrow\rho}(\cdot; x, y, \alpha)$ of λ with respect to ρ exists.*

By Remark 4.7, $\lambda(1,1,\alpha,1)=1$ and $\lambda(1,1,\alpha,-1)=0$ for $\alpha>0$. Hence, we can define

$$\widetilde{\rho}_{ij}(1,1;\,k) \quad := \quad \lambda^{\leftarrow\rho}\left(\widehat{\lambda}^{\mathrm{emp}}_{ij}(1,1;\,k);\,1,1,\widehat{\alpha}(k,w)\right). \tag{4.12}$$

Since this estimator only employs information at $(x,y)=(1,1)$, it may not be very efficient. Therefore, we define an estimator based on $\widehat{\lambda}^{\mathrm{emp}}_{ij}(x,y;\,k)$ for other values $(x,y)=(\sqrt{2}\cos\theta\,,\,\sqrt{2}\sin\theta)\in\mathbb{R}^2_+$.

To ensure existence and consistency of the estimator, we define the following sets and give some explanations below:

$$\begin{aligned}
\widehat{U}_{ij} \quad &:= \quad \Big\{\theta\in\Big(0,\frac{\pi}{2}\Big):\ \widehat{\lambda}^{\mathrm{emp}}_{ij}\left(\sqrt{2}\cos\theta,\sqrt{2}\sin\theta;\,k\right)< \\
&\qquad\qquad <\ \lambda\left(\sqrt{2}\cos\theta,\sqrt{2}\sin\theta,\widehat{\alpha}(k,w),e^{-|\ln(\tan\theta)|/\widehat{\alpha}(k,w)}\right)\Big\}, \\
\widehat{U}^*_{ij} \quad &:= \quad \Big\{\theta\in\Big(0,\frac{\pi}{2}\Big):\ |\ln(\tan\theta)|<(1-k^{-1/4})\widehat{\alpha}(k,w)\left|\ln\left(\widetilde{\rho}_{ij}(1,1;\,k)\vee 0\right)\right|\Big\} \text{ and} \\
U^*_{ij} \quad &:= \quad \Big\{\theta\in\Big(0,\frac{\pi}{2}\Big):\ |\ln(\tan\theta)|<\alpha\,|\ln(\rho_{ij}\vee 0)|\Big\}.
\end{aligned}$$

By Lemma 4.14 there exists a unique ρ such that

$$\lambda\left(\sqrt{2}\cos\theta,\sqrt{2}\sin\theta,\widehat{\alpha}(k,w),\rho\right) \quad = \quad \widehat{\lambda}^{\mathrm{emp}}_{ij}\left(\sqrt{2}\cos\theta,\sqrt{2}\sin\theta;\,k\right),\qquad \theta\in\widehat{U}_{ij}.$$

Hence, we can define

$$\begin{aligned}
&\widetilde{\rho}_{ij}(\sqrt{2}\cos\theta,\sqrt{2}\sin\theta;\,k) \\
&\quad :=\ \lambda^{\leftarrow\rho}_{ij}\left(\widehat{\lambda}^{\mathrm{emp}}_{ij}(\sqrt{2}\cos\theta,\sqrt{2}\sin\theta;\,k);\,\sqrt{2}\cos\theta,\sqrt{2}\sin\theta,\widehat{\alpha}(k,w)\right),\qquad \theta\in\widehat{U}_{ij}.
\end{aligned} \tag{4.13}$$

Note that, by the definition of $\widetilde{\rho}_{ij}(1,1;\,k)$ in (4.12), it always holds that $\pi/4\in\widehat{U}_{ij}$ provided that $\widehat{\lambda}^{\mathrm{emp}}_{ij}(1,1;\,k)<1$. Hence, if $\widehat{U}_{ij}=\emptyset$, we can replace it by $\widehat{U}_{ij}:=\{\pi/4\}$ and also replace $\widehat{U}^*_{ij}:=\{\pi/4\}$. To ensure consistency we further require $\theta\in\widehat{U}^*_{ij}$. This implies that the true ρ_{ij} is smaller than $e^{-|\ln(\tan\theta)|/\widehat{\alpha}(k,w)}$ with probability tending to one. The set U^*_{ij} is then the true set of $\theta\in(0,\pi/2)$, where Lemma 4.14 applies.

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

Definition 4.15. *Let w^* be a nonnegative weight function and W^* be the measure induced by w^*. Then we define for $i\neq j$ and with* (4.13)

$$\widehat{\rho}^{\lambda}_{ij}(k,w^*) \quad := \quad \frac{1}{W^*\big(\widehat{U}_{ij}\cap\widehat{U}^*_{ij}\big)}\int_{\theta\in\widehat{U}_{ij}\cap\widehat{U}^*_{ij}}\widetilde{\rho}_{ij}(\sqrt{2}\cos\theta,\sqrt{2}\sin\theta;k)\,W^*(d\theta). \tag{4.14}$$

Further, define $\widehat{\rho}^{\lambda}_{ii}(k,w^*):=1$, $1\leq i\leq d$, and $\widehat{\mathbf{R}}_{\lambda}(k,w^*):=\big(\widehat{\rho}^{\lambda}_{ij}(k,w^*)\big)_{1\leq i,j\leq d}$.

The next theorem shows the asymptotic properties of $\widehat{\mathbf{R}}_{\lambda}(k,w^*)$. We use the theory developed in Schmidt and Stadtmüller (2005) and give a formal proof in Section 6.

Theorem 4.16. *Suppose the following regularity conditions hold:*

(C1) $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ *are iid with copula* $C_{\boldsymbol{X}} \in \mathcal{EC}_d(\mathbf{R}, G)$, $G \in RV_{-\alpha}$ *for* $\alpha > 0$ *and* $\max_{i \neq j} |\rho_{ij}| < 1$.

(C2) *There exists* $A(t) \to 0$ *such that for* $i \neq j$

$$\lim_{t \to 0} \frac{t^{-1} P(1 - F_i(X_i) \leq tx, 1 - F_j(X_j) \leq ty) - \lambda(x, y, \alpha, \rho_{ij})}{A(t)} \quad = \quad b_{ij}^{(C2)}(x, y)$$

uniformly on $\mathcal{S}_2 := \{\boldsymbol{s} \in \mathbb{R}^2 : \boldsymbol{s}^T \boldsymbol{s} = 1\}$, *where* $b_{ij}^{(C2)}(x, y)$ *is some non-constant function.*

(C3) $k = k(n) \to \infty$, $k/n \to 0$ *and* $\sqrt{k} A(k/n) \to 0$ *as* $n \to \infty$.

Let w^* *be a nonnegative weight function with* $\sup_{\theta \in U_{ij}^*} w^*(\theta) < \infty$ *for all* $i \neq j$, λ'^{ρ} *and* λ'^{α} *denotes the derivative of* λ *with respect to* ρ *and* α, *respectively, and* $(\lambda^{\leftarrow \rho})'^{\alpha}$ *denotes the derivative of* $\lambda^{\leftarrow \rho}$ *with respect to* α. *Define*

$$\begin{aligned} \widetilde{B}_{ij}(x, y) &:= B_{ij}(x, y) - B_{ij}(x, \infty) \frac{\partial}{\partial x} \lambda_{ij}(x, y) - B_{ij}(\infty, y) \frac{\partial}{\partial y} \lambda_{ij}(x, y), \qquad (4.15) \\ B_{ij}(x, y) &:= B(\infty, \ldots, \infty, x, \infty, \ldots, \infty, y, \infty, \ldots, \infty), \end{aligned}$$

where x *is the* i*-th,* y *the* j*-th component and* B *is a centered tight continuous Gaussian random field on* $\overline{\mathbb{R}}^d$ *with covariance structure*

$$E(B(\boldsymbol{x}) B(\boldsymbol{y})) \quad = \quad \lambda^{\boldsymbol{X}}(\boldsymbol{x} \wedge \boldsymbol{y}), \quad \boldsymbol{x}, \boldsymbol{y} \in [0, \infty]^d, \qquad (4.16)$$

where $\boldsymbol{x} \wedge \boldsymbol{y}$ *is taken componentwise. Set as before* $\boldsymbol{r} := \mathrm{vec}[\mathbf{R}]$ *and* $\widehat{\boldsymbol{r}}_\lambda(k, w^*) := \mathrm{vec}[\widehat{\mathbf{R}}_\lambda(k, w^*)]$, *then*

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

where $\Gamma_\lambda = (\gamma_{ij,kl}^\lambda)_{1 \leq i \neq j, k \neq l \leq d}$ *with*

$$\gamma_{ij,kl}^\lambda \quad = \quad \sigma_\alpha + \sigma_{ij,\alpha} + \sigma_{kl,\alpha} + \sigma_{ij,kl}, \qquad (4.17)$$

and

$$\begin{aligned} \sigma_\alpha \;=\; & \frac{2}{d^2 (d-1)^2 W^*(U_{ij}^*) W^*(U_{kl}^*)} \qquad (4.18) \\ & \times \prod_{J \in \{ij, kl\}} \int_{\theta \in U_J^*} (\lambda^{\leftarrow \rho})'^{\alpha} \left(\lambda_J^{\boldsymbol{X}}(\sqrt{2} \cos\theta, \sqrt{2} \sin\theta), \sqrt{2} \cos\theta, \sqrt{2} \sin\theta, \alpha \right) W^*(d\theta) \\ & \times \left(\sum_{1 \leq p < q, r < s \leq d} \frac{1}{W(Q_{pq}^*) W(Q_{rs}^*)} \right. \\ & \times \left. \int_{\theta_1 \in Q_{pq}^*} \int_{\theta_2 \in Q_{rs}^*} \frac{E\left(\widetilde{B}_{pq}(\sqrt{2} \cos\theta_1, \sqrt{2} \sin\theta_1) \widetilde{B}_{rs}(\sqrt{2} \cos\theta_2, \sqrt{2} \sin\theta_2) \right)}{\lambda'^{\alpha}(\cos\theta_1, \sin\theta_1, \alpha, \rho_{pq}) \lambda'^{\alpha}(\cos\theta_2, \sin\theta_2, \alpha, \rho_{rs})} W(d\theta_2) W(d\theta_1) \right), \end{aligned}$$

$$\sigma_{ij,\alpha} = \frac{1}{d(d-1)W^*(U^*_{ij})W^*(U^*_{kl})} \sum_{1\le p<q\le d} \frac{1}{W(Q^*_{pq})} \times \tag{4.19}$$
$$\times \Bigg(\int_{\theta_1\in U^*_{ij}} \int_{\theta_1\in U^*_{kl}} \int_{\theta_3\in Q^*_{pq}} (\lambda^{\leftarrow\rho})'^{\alpha} \left(\lambda^{\boldsymbol{X}}_{ij}(\sqrt{2}\cos\theta_1, \sqrt{2}\sin\theta_1), \sqrt{2}\cos\theta_1, \sqrt{2}\sin\theta_1, \alpha \right) \times$$
$$\times \frac{E\left(\widetilde{B}_{pq}(\sqrt{2}\cos\theta_3, \sqrt{2}\sin\theta_3)\widetilde{B}_{kl}(\sqrt{2}\cos\theta_2, \sqrt{2}\sin\theta_2)\right)}{\lambda'^{\alpha}(\cos\theta_3, \sin\theta_3, \alpha, \rho_{pq})\lambda'^{\rho}(\cos\theta_2, \sin\theta_2, \alpha, \rho_{kl})} W^*(d\theta_3)W^*(d\theta_2)W(d\theta_1) \Bigg),$$

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

$$\sigma_{ij,kl} = \frac{1}{2W^*(U^*_{ij})W^*(U^*_{kl})} \tag{4.20}$$
$$\times \int_{\theta_1\in U^*_{ij}} \int_{\theta_2\in U^*_{kl}} \frac{E\left(\widetilde{B}_{ij}(\sqrt{2}\cos\theta_1, \sqrt{2}\sin\theta_1)\widetilde{B}_{kl}(\sqrt{2}\cos\theta_2, \sqrt{2}\sin\theta_2)\right)}{\lambda'^{\alpha}(\cos\theta_1, \sin\theta_1, \alpha, \rho_{ij})\lambda'^{\rho}(\cos\theta_2, \sin\theta_2, \alpha, \rho_{kl})} W^*(d\theta_2)W^*(d\theta_1).$$

Remark 4.17. If condition (C3) in Theorem 4.16 is replaced by

(C3') $k = k(n) \to \infty$, $k/n \to 0$ and $\sqrt{k}A(k/n) \to b^{(C3)} \in (-\infty, \infty)$ as $n \to \infty$,

an asymptotic bias occurs in vec$[\widehat{\mathbf{R}}_\lambda(k, w^*)]$. Using the delta method it immediately follows that

$$\sqrt{k}\left(\widehat{\boldsymbol{r}}_\lambda(k, w^*) - \boldsymbol{r}\right) \xrightarrow{d} \mathcal{N}_{d(d-1)/2}(\boldsymbol{b}_\rho + \boldsymbol{b}_\alpha, \boldsymbol{\Gamma}_\lambda),$$

where $\boldsymbol{\Gamma}_\lambda$ is given in (4.17), $\boldsymbol{b}_\rho = \mathrm{vec}[(b_{ij,\rho})_{1\le i,j\le d}]$, $\boldsymbol{b}_\alpha = \mathrm{vec}[(b_{ij,\alpha})_{1\le i,j\le d}]$,

$$b_{ij,\rho} = \frac{1}{W(U^*_{ij})} \int_{\theta\in U^*_{ij}} \frac{b^{(C3)} b^{(C2)}_{ij}(\sqrt{2}\cos\theta, \sqrt{2}\sin\theta)}{\lambda'^{\rho}(\sqrt{2}\cos\theta, \sqrt{2}\sin\theta, \alpha, \rho_{ij})} W^*(d\theta), \ i \ne j, \quad \text{and}$$

$$b_{ij,\alpha} = \frac{1}{W(U^*_{ij})} \int_{\theta\in U^*_{ij}} (\lambda^{\leftarrow\rho})'^{\alpha} \left(\lambda^{\boldsymbol{X}}_{ij}\left(\sqrt{2}\cos\theta, \sqrt{2}\sin\theta\right), \sqrt{2}\cos\theta, \sqrt{2}\sin\theta, \alpha \right) W^*(d\theta) \times$$
$$\times \frac{2}{d(d-1)} \sum_{1\le p<q\le d} \frac{1}{W(Q^*_{pq})} \int_{\theta\in Q^*_{pq}} \frac{b^{(C3)} b^{(C2)}_{pq}(\sqrt{2}\cos\theta, \sqrt{2}\sin\theta)}{\lambda'^{\alpha}(\sqrt{2}\cos\theta, \sqrt{2}\sin\theta, \alpha, \rho_{pq})} W(d\theta).$$

Using (4.17), we can define an estimator of $\boldsymbol{\Gamma}_\lambda$.

Definition 4.18. *We define the estimator of* $\boldsymbol{\Gamma}_\lambda = (\gamma^\lambda_{ij,kl})_{1\le i\ne j, k\ne j}$ *by* $\widehat{\boldsymbol{\Gamma}}_\lambda = (\widehat{\gamma}^\lambda_{ij,kl})_{1\le i\ne j, k\ne j}$ *with*

$$\widehat{\gamma}^\lambda_{ij,kl} = \widehat{\sigma}_\alpha + \widehat{\sigma}_{ij,\alpha} + \widehat{\sigma}_{kl,\alpha} + \widehat{\sigma}_{ij,kl}, \tag{4.21}$$

the $\widehat{\sigma}$ are defined in (4.18)–(4.20), where α, ρ_{ij} and ρ_{kl} are replaced by their estimators $\widehat{\alpha}(k, w)$, $\widehat{\rho}^\lambda_{ij}(k, w^*)$ and $\widehat{\rho}^\lambda_{kl}(k, w^*)$, 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 $E\left(\widetilde{B}_{ij}(\cdot)\widetilde{B}_{kl}(\cdot)\right)$ are replaced by their estimators $\widehat{E}\left(\widetilde{B}_{ij}(\cdot)\widetilde{B}_{kl}(\cdot)\right)$ using (4.15) and (4.16) and estimating $\lambda^{\boldsymbol{X}}$ by $\widehat{\lambda}^{\mathrm{emp}}$.

The asymptotic properties of $\widehat{\lambda}^{\text{emp}}$, $\widehat{\alpha}$, $\widehat{\rho}_{ij}^{\lambda}$ in combination with the delta method yield immediately the following result.

Theorem 4.19. *Under the regularity conditions (C1)–(C3), the estimator* $\text{vec}[\widehat{\mathbf{\Gamma}}_\lambda]$ *is consistent and asymptotically normal.*

Estimation of dependence in extremes is always a difficult topic, for some methods of estimation of $\lambda_{ij}^{\boldsymbol{X}}(1,1)$ and pitfalls we refer to Frahm, Junker, and Schmidt (2005). The problem of estimating tail dependence lies in its definition as a limit, see (4.1). Estimators of the tail dependence are based on a sub-sample using the largest (or smallest) observations. Concerning the optimal choice of the threshold, we refer to Danielsson, de Haan, Peng, and de Vries (2001), Drees and Kaufmann (1998) and to Klüppelberg et al. (2005a, 2005b).

Remark 4.20. It may happen that the correlation matrix estimators (4.4) or (4.14) are not positive definite. In this case we use the approach of Higham (2002), i.e. we replace $\widehat{\mathbf{R}}$ by the (positive definite) correlation matrix $\mathbf{R}^*$ solving

$$\left\|\widehat{\mathbf{R}} - \mathbf{R}^*\right\|_2 \quad = \quad \min\left\{\left\|\widehat{\mathbf{R}} - \mathbf{R}\right\|_2 : \mathbf{R} \text{ is a correlation matrix}\right\},$$

where $\|\mathbf{R}\|_2 = \sum_{i,j} \rho_{ij}^2$ is the Euclidean or *Frobenius* norm of a matrix $\mathbf{R} = (\rho_{ij})_{1\le i,j\le d}$. Let $\mathbf{R}$ have spectral decomposition $\mathbf{R} = \mathbf{Q}\mathbf{D}\mathbf{Q}^T$ with $\mathbf{Q}$ orthogonal and $\mathbf{D} = \text{diag}(\kappa_1, \ldots, \kappa_d)$. By Higham (2002, Theorem 3.1 and 3.2), $P_U(\mathbf{R}) := \mathbf{R} - \text{diag}(\mathbf{R} - \mathbf{I})$ is the projection of $\mathbf{R}$ to the set of symmetric matrices with diagonal $\mathbf{1}$ and $P_S(\mathbf{R}) := \mathbf{Q}\,\text{diag}(\max(\kappa_i, 0))\,\mathbf{Q}^T$ is the projection of $\mathbf{R}$ to the set of positive definite matrices, respectively. Then, Higham (2002, Algorithm 3.3) calculates $\mathbf{Y}_i$ converging to $\mathbf{R}^*$ with respect to the Frobenius norm as $i \to \infty$:

$$\begin{aligned}
&\triangle\mathbf{S}_0 - \mathbf{0},\ \mathbf{Y}_0 = \widehat{\mathbf{R}} \\
&\text{for } i = 1, 2, \ldots \\
&\qquad \mathbf{Z}_i = \mathbf{Y}_{i-1} - \triangle\mathbf{S}_{i-1} \\
&\qquad \mathbf{X}_i = P_S(\mathbf{Z}_i) \\
&\qquad \triangle\mathbf{S}_i = \mathbf{X}_i - \mathbf{Z}_i \\
&\qquad \mathbf{Y}_i = P_U(\mathbf{X}_i) \\
&\text{end.}
\end{aligned}$$

Considering covariance matrices, we do not need the projection P_U. Hence, if we observe not positive definite covariance estimators (4.6) or (4.21), we project them to the set of positive definite matrices by $P_S(\widehat{\mathbf{\Gamma}})$.

5 The new methods at work

Using the estimators (4.4) and (4.6) or (4.14) and (4.21) together with the quadratic form discrepancy function (3.6), we can now apply copula structure analysis. In the following, we consider the copula factor model, i.e. we choose the setting $C_{\boldsymbol{X}} = C_{(\mathbf{L},\mathbf{V})\boldsymbol{\xi}}$, where $\mathbf{L} \in \mathbb{R}^{d\times m}$, $\mathbf{V} \in \mathbb{R}^{d\times d}$ is a diagonal matrix with nonnegative entries and $\boldsymbol{\xi} \in \mathcal{E}_{m+d}(\mathbf{0}, \mathbf{I}, G)$; also see Remark 3.2(iii).

As for the test statistic T based on the quadratic form discrepancy function (3.6) we first compare in a simulation study T to its limiting χ^2-distribution. Therefore, we define by

$$T^{\tau}_{\mathrm{QD}} \;:=\; n \min_{\boldsymbol{\vartheta}\in\Theta} D_{\mathrm{QD}}\left(\widehat{\boldsymbol{r}}_{\tau}, \boldsymbol{r}(\boldsymbol{\vartheta}) | \widehat{\boldsymbol{\Gamma}}_{\tau}\right)$$

the quadratic form test statistic obtained from the Kendall's tau based estimators $\widehat{\boldsymbol{r}}_{\tau} = \mathrm{vec}[\widehat{\mathbf{R}}_{\tau}]$ and $\widehat{\boldsymbol{\Gamma}}_{\tau}$ given in (4.4) and (4.6), respectively.

Similarly,

$$T^{\lambda}_{\mathrm{QD}} \;:=\; k \min_{\boldsymbol{\vartheta}\in\Theta} D_{\mathrm{QD}}\left(\widehat{\boldsymbol{r}}_{\lambda}, \boldsymbol{r}(\boldsymbol{\vartheta}) | \widehat{\boldsymbol{\Gamma}}_{\lambda}\right),$$

where k is the number of the largest order statistics used for estimation, $\widehat{\boldsymbol{r}}_{\lambda} = \mathrm{vec}[\widehat{\mathbf{R}}_{\lambda}]$ and $\widehat{\boldsymbol{\Gamma}}_{\lambda}$ given in (4.14) and (4.21), respectively. As a weight function we choose a discrete version of

$$w(\theta) \;=\; 1 - \left(\frac{\theta}{\pi/4} - 1\right)^2, \quad 0 \le \theta \le \frac{\pi}{2}, \tag{5.1}$$

both for the estimation of α and $\mathbf{R}$ given a copula $C \in \mathcal{EC}(\mathbf{R}, G)$, $G \in RV_{-\alpha}$ and $\alpha > 0$.

We also compare the copula factor model to the classical factor model $\boldsymbol{X} = (\mathbf{L}, \mathbf{V})\boldsymbol{\xi}$, $\boldsymbol{\xi} \in \mathcal{E}_{m+d}(\mathbf{0}, \mathbf{I}, G)$. To this end we define

$$T^{\mathrm{emp}}_{\mathrm{QD}} \;:=\; n \min_{\boldsymbol{\vartheta}\in\Theta} D_{\mathrm{QD}}\left(\widehat{\boldsymbol{r}}_{\mathrm{emp}}, \boldsymbol{r}(\boldsymbol{\vartheta}) | \widehat{\boldsymbol{\Gamma}}_{\mathrm{emp}}\right),$$

where $\widehat{\boldsymbol{r}}_{\mathrm{emp}} = \mathrm{vec}[\widehat{\mathbf{R}}_{\mathrm{emp}}]$ is the vector of the standard empirical correlation estimator with its asymptotic covariance matrix estimator $\widehat{\boldsymbol{\Gamma}}_{\mathrm{emp}}$ under normal assumptions, for details see Browne and Shapiro (1986).

The parameter $\boldsymbol{\vartheta}$ is then estimated also in three different ways, denoted by $\widehat{\boldsymbol{\vartheta}}_{\tau}$, $\widehat{\boldsymbol{\vartheta}}_{\lambda}$ and $\widehat{\boldsymbol{\vartheta}}_{\mathrm{emp}}$, by minimizing T^{τ}_{QD}, $T^{\lambda}_{\mathrm{QD}}$ and $T^{\mathrm{emp}}_{\mathrm{QD}}$, respectively.

Example 5.1. [Model selection by χ^2-tests]
To see the performance of the quadratic form test statistics T^{τ}_{QD} and $T^{\lambda}_{\mathrm{QD}}$, we perform a simulation study. We choose a $d = 10$ dimensional setting with $m = 2$ factors and loadings as given in Table 1. Then $\mathbf{L}\mathbf{L}^T + \mathbf{V}^2 = \mathbf{R}$ is a correlation matrix.

Define a multivariate t_α-copula as the copula of the random vector $G\boldsymbol{N}$, where $G \sim \sqrt{\alpha/\chi^2_\alpha}$, $\alpha > 0$, is independent of $\boldsymbol{N} \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$. Note that the t_α-copula is elliptical and, since $G \in RV_{-\alpha}$, its tail copula satisfies (4.3). Choose $\alpha = 3$, then $G\boldsymbol{N}$ has finite second moment, but its fourth moment does not exist. Hence, classical factor analysis cannot be applied to $G\boldsymbol{N}$, see Proposition **??** and Theorem 3.7. Also, if the model with

component	1	2	3	4	5	6	7	8	9	10
$\mathbf{L}_{\cdot,1}$	.9	.9	.9	.9	.9	0	0	0	0	0
$\mathbf{L}_{\cdot,2}$	0	0	0	0	0	.9	.9	.9	.9	.9
$\mathrm{diag}(\mathbf{V}^2)$	.19	.19	.19	.19	.19	.19	.19	.19	.19	.19

Table 1: Factor loadings of Example 5.1

$\alpha < 8$ is considered, which has finite fourth moment but non-existing eight moment, the estimator of $\mathbf{\Gamma}$ will only be consistent and large sample sizes may be necessary to observe the limiting χ^2 distribution of the test statistic T. As the test statistics T^{τ}_{QD} and T^{λ}_{QD} are based on the copula of the sample, they are not affected by the existence or non-existence of moments.

We simulate 500 iid samples of length $n = 1\,000$ of the t_3-copula, calculate the Kendall's tau based estimators (4.4) and (4.6) and estimate T^{τ}_{QD} from these. To ensure uniqueness of the loadings, we use the restriction that $\mathbf{L}^T\mathbf{V}^{-2}\mathbf{L}$ is diagonal, hence we have $m(m-1)/2 = 1$ additional constraints, see Lawley and Maxwell (1971, Section 2.3). Using this restriction and the 2-factor setting, T^{τ}_{QD} should be (for a large sample) χ^2_{df} distributed with $df = d(d-1)/2 - dm + m(m-1)/2 = 26$ degrees of freedom; see Theorem 3.7. Therefore, we compare the 500 estimates of T^{τ}_{QD} with the χ^2_{26}-distribution by a *QQ*-plot, see Figure 1, left plot. From this plot we see that the distribution of T^{τ}_{QD} fits the χ^2_{26}-distribution quite well. Similarly, we estimate T^{λ}_{QD} based on the tail copula estimators (4.14) and (4.17) with weight function (5.1) using the same samples as for T^{τ}_{QD} and based on the $k = 100$ largest observations; see Figure 1, right plot. Also here we observe a reasonable fit to the χ^2_{26}-distribution – not as good as before since the estimators are calculated from a smaller (sub)sample. Note that under the assumption of $m = 1$ factor the corresponding T^{τ}_{QD}'s and T^{λ}_{QD}'s were always larger than 600, which clearly rejects the 1-factor hypothesis.

Example 5.2. [Oil-currency data]
In this example we consider an 8-dimensional set of data, *(oil, s&p500, gbp, usd, chf, jpy, dkk, sek)*, i.e. we are interested in the dependence structure between the oil-price, the S&P500 index and some currency exchange rates with respect to euro. Each time series consists of 4 904 daily logreturns from May, 1985 to June, 2004. To this data set we fit a copula factor model using the $T^{\text{emp}}_{\text{QD}}$, T^{τ}_{QD} and T^{λ}_{QD} statistics for estimation and model selection. Estimation of T^{λ}_{QD} is based on the $k = 300$ largest observations. The values of these test statistics, based on different numbers of factors are given in Table 2. To estimate the number of factors, we use a 95% confidence test, i.e. we reject the null hypothesis of having a m-factor model if the test statistic T is larger than the 95%-quantile of the χ^2_{df}-distribution. This yields 4 factors under the empirical, Kendall's tau based and tail copula based test statistics.

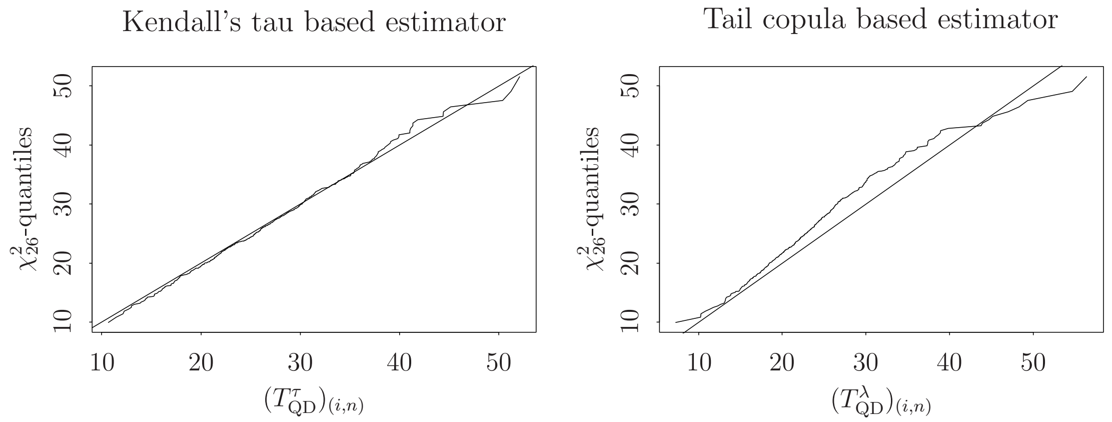

Figure 1: *QQ*-plot of ordered estimates $\widehat{T}$ against the χ^2_{26}-quantiles.
Left plot: T^{τ}_{QD} obtained from Kendall's tau based estimators (4.4) and (4.6).
Right plot: $T^{\lambda}_{\mathrm{QD}}$ obtained from tail copula based estimators (4.14) and (4.17).

number of factors	*df*	$T^{\mathrm{emp}}_{\mathrm{QD}}$	T^{τ}_{QD}	$T^{\lambda}_{\mathrm{QD}}$	$\chi^2_{df,0.95}$
2	13	298.5	252.7	52.7	22.36
3	7	33.7	17.4	24.0	14.07
4	2	2.3	3.3	0.9	5.99

Table 2: Test statistics $T^{\mathrm{emp}}_{\mathrm{QD}}$, T^{τ}_{QD} and $T^{\lambda}_{\mathrm{QD}}$ of oil-currency data under different number of factors.

Applying factor analysis based on the different correlation estimates (and their asymptotic covariance estimates) yield different results; see Figure 2. The first four plots show the loadings of the four factors, obtained from the empirical correlation estimator, Kendall's tau based and tail copula based estimator. The last plot shows the loadings of the specific factors for all three correlation estimators.

We want to emphasize that, although we have plotted the factors in the same figures, the factors obtained by the three different estimation methods are not known and may have different interpretations. We call them *empirical factors*, *Kendall's tau factors* and *tail copula factors*.

For the first factor all loadings of the different correlation estimators behave very similar with respect to factor 1, which has a weight close to one for usd. Hence, factor one can be interpreted as the *usd-factor*. It also can be seen that this factor has a positive weight for all currencies, but not for the oil-price and s&p500 (almost 0 or very small negative), and the largest dependence is observed for gbp, and jpy.

For factor 2 we observe for all correlation estimators a large weight on Swiss Francs chf, so we call it *chf-factor*. We observe that the empirical and Kendall's tau factor has almost no (or only little) correlation with oil, s&p500, gbp, usd and jpy. The weights on dkk and sek are larger and also moderate for gpd for the tail copula factor indicating that extreme dependence between all European currencies is present.

Considering factor 3, we see for the empirical and Kendall's tau factor a large loading for sek and dkk with only little impact on the other components. If scandinavian currencies were merged, then only a specific factor would remain. The tail copula factor indicates moderate dependence between oil and gbp.

From factor 4 we observe for the empirical factor a loading close to one for the oil-price and loadings close to 0 for the rest of the factors. This indicates that a 3-factor model is sufficient in this case. In combination with the model selection procedure as seen in Table 2 this indicates that the distribution of $T_{\mathrm{QD}}^{\mathrm{emp}}$ is far away from a χ^2 distribution. For the Kendall's tau factor there is some dependence between the European currencies and the usd. The tail copula factor behaves different: there is dependence observed between large positive jumps of s&p500 and large negative jumps of the oil price which would not be detected when only considering the other correlation estimators.

Finally, we give an interpretation of the specific factors, where we find the correlation which is not explained through the common factors. For the empirical factor oil is completely explained by factor 4, which is the specific factor for oil, and s&p500 has a loading close to one, showing there is (almost) no correlation to oil and the other currencies. For the Kendall's tau factor, oil and s&p500 are uncorrelated and uncorrelated from the rest. Contrary, for the tail copula factor, oil and s&p500 are not uncorrelated from the common factors. Oil has a rather large specific loading factor, but s&p500 is explained to a large extend by factor 4.

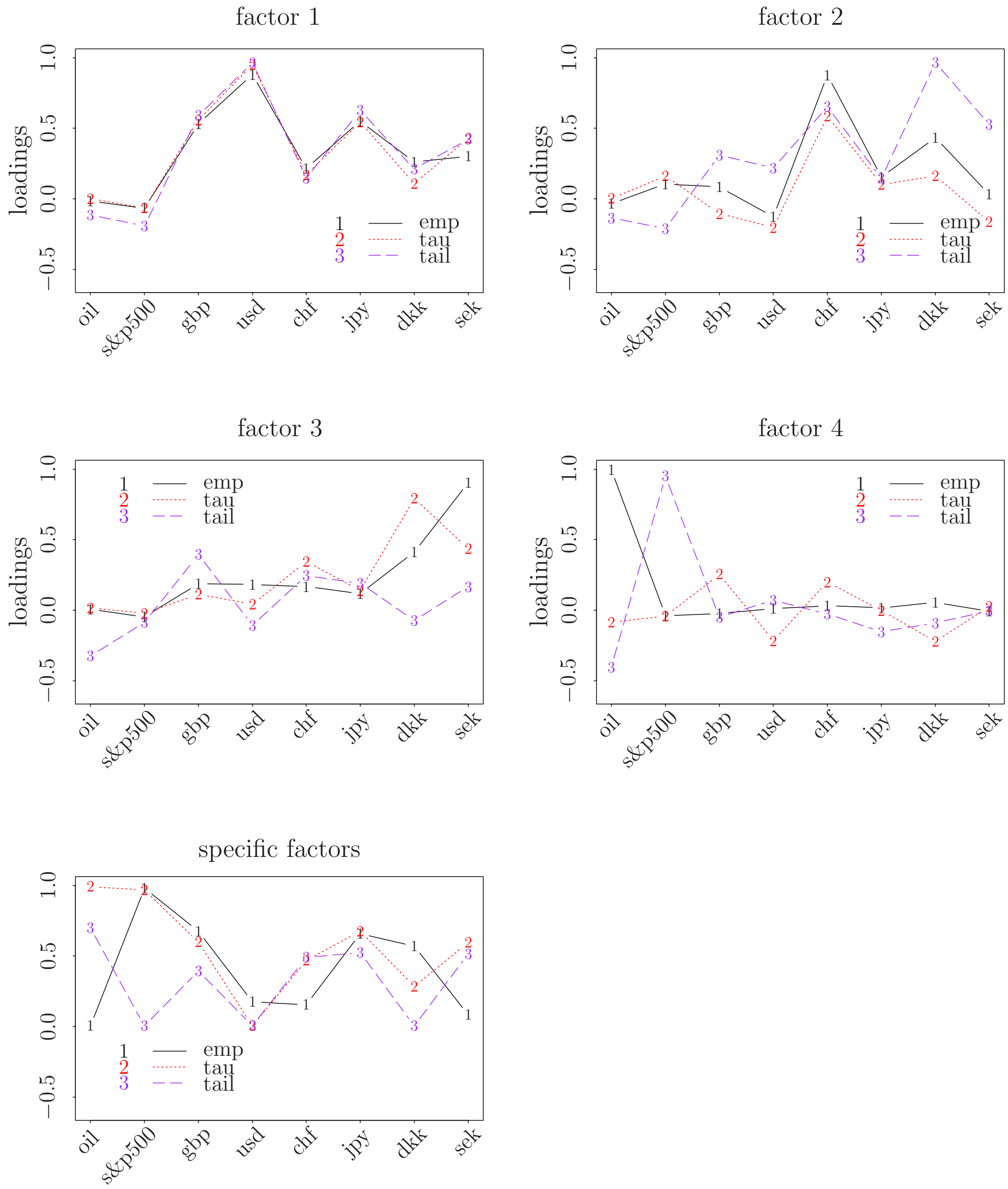

Figure 2: Oil-currency data: factor analysis based on 4 factors and different statistics, "emp" for the loadings $\widehat{\boldsymbol{\vartheta}}_{\text{emp}}$, "tau" for $\widehat{\boldsymbol{\vartheta}}_{\tau}$ and "tail" for $\widehat{\boldsymbol{\vartheta}}_{\lambda}$.

Upper row: loadings of factor 1 (left) and 2 (right).

Middle row: loadings of factor 3 (left) and 4 (right).

Lower row: specific factors $\text{diag}(\mathbf{V}^2)$.

6 Proofs

Proof of Theorem 4.9: Define $\widehat{\boldsymbol{t}} := \text{vec}[\widehat{\mathbf{T}}]$ and $\boldsymbol{t} := \text{vec}[\mathbf{T}]$. Since $\widehat{\boldsymbol{t}}$ is a vector of U-statistics, and, obviously,

$$E\left(\text{sign}\left((X_{1,i} - X_{2,i})\,(X_{1,j} - X_{2,j})\right)^2\right) \;<\; \infty, \quad i \neq j,$$

Lee (1990, Chapter 3, Theorem 2) applies (together with the remark at the end of p.7 therein that all results also hold for random vectors). The covariance structure is stated in Lee (1990, Section 1.4, Theorem 1), hence

$$\sqrt{n}(\widehat{\boldsymbol{t}} - \boldsymbol{t}) \;\stackrel{d}{\longrightarrow}\; \mathcal{N}_{d(d-1)/2}\left(\mathbf{0}, 4\boldsymbol{\Upsilon}\right), \quad n \to \infty,$$

where $\boldsymbol{\Upsilon} = (\tau_{ij,kl} - \tau_{ij}\tau_{kl})_{1\le i\neq j,k\neq l\le d}$ and $\tau_{ij,kl}$ is given in (4.5). Note that the Jacobian matrix $\mathbf{D} := \partial\left(\sin\left(\boldsymbol{t}\pi/2\right)\right)/\partial\boldsymbol{t}$ is a diagonal matrix with

$$\text{diag}(\mathbf{D}) \;=\; \frac{\pi}{2}\cos\left(\frac{\pi}{2}\boldsymbol{t}\right).$$

Hence, by the delta method (see Casella and Berger (2001, Section 5.5.4)),

$$\sqrt{n}\left(\widehat{\boldsymbol{r}} - \boldsymbol{r}\right) \;\stackrel{d}{\longrightarrow}\; \mathcal{N}_{d(d-1)/2}\left(\mathbf{0}, 4\mathbf{D}^T\boldsymbol{\Upsilon}\mathbf{D}\right), \quad n \to \infty,$$

and the proof is complete.

Proof of Theorem 4.11: We first consider $\widehat{\tau}_{ij,kl}$ and rewrite it as a linear combination of some U-statistics. Define for $1 \le a < b < c \le n$

$$\begin{aligned}
\Phi_2^{ij,kl}(\boldsymbol{x}_a, \boldsymbol{x}_b) &:= \text{sign}\left[(x_{a,i} - x_{b,i})(x_{a,j} - x_{b,j})\right]\text{sign}\left[(x_{a,k} - x_{b,k})(x_{a,l} - x_{b,l})\right] \\
\Phi_{abc}^{ij,kl} &:= \Phi_{abc}^{ij,kl}(\boldsymbol{x}_a, \boldsymbol{x}_b, \boldsymbol{x}_c) \\
&:= \text{sign}\left[(x_{a,i} - x_{b,i})(x_{a,j} - x_{b,j})\right]\text{sign}\left[(x_{a,k} - x_{c,k})(x_{a,l} - x_{c,l})\right] \quad \text{and} \\
\Phi_3^{ij,kl}(\boldsymbol{x}_a, \boldsymbol{x}_b, \boldsymbol{x}_c) &:= \frac{1}{6}\left(\Phi_{abc}^{ij,kl} + \Phi_{acb}^{ij,kl} + \Phi_{bac}^{ij,kl} + \Phi_{bca}^{ij,kl} + \Phi_{cab}^{ij,kl} + \Phi_{cba}^{ij,kl}\right).
\end{aligned}$$

Hence, $\Phi_2^{ij,kl}$ and $\Phi_3^{ij,kl}$ are symmetric in their arguments. Next, define

$$\begin{aligned}
\widehat{u}_2^{ij,kl} &:= \frac{2}{n(n-1)}\sum_{1\le a<b\le n}\Phi_2^{ij,kl}(\boldsymbol{X}_a, \boldsymbol{X}_b) \quad \text{and} \\
\widehat{u}_3^{ij,kl} &:= \frac{6}{n(n-1)(n-2)}\sum_{1\le a<b<c\le n}\Phi_3^{ij,kl}(\boldsymbol{X}_a, \boldsymbol{X}_b, \boldsymbol{X}_c),
\end{aligned}$$

and note that both are U-statistics. Obviously,

$$E\left(\left(\Phi_2^{ij,kl}(\boldsymbol{X}_1, \boldsymbol{X}_2)\right)^2\right) < \infty \quad \text{and} \quad E\left(\left(\Phi_3^{ij,kl}(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{X}_3)\right)^2\right) < \infty,$$

therefore, by Lee (1990, Chapter 3, Theorem 2), the vector of all $\widehat{u}_2^{ij,kl}$ and $\widehat{u}_2^{ij,kl}$ is consistent and asymptotically normal. Since

$$\widehat{\tau}_{ij,kl} \quad = \quad \frac{1}{n(n-1)^2}\left(\frac{n(n-1)}{2}\widehat{u}_2^{ij,kl} + \frac{n(n-1)(n-2)}{6}\widehat{u}_3^{ij,kl}\right),$$

$\widehat{\tau}_{ij,kl}$ is a linear combination of U-statistics and is therefore also consistent and asymptotically normal. The result then follows using the delta method.

Proof of Theorem 4.16: First, by homogeneity,

$$\lambda(\sqrt{2}\cos\theta_2, \sqrt{2}\sin\theta_2, \alpha, \rho) = \sqrt{2}\lambda(\cos\theta_2, \sin\theta_2, \alpha, \rho)$$

holds. Let '$\xrightarrow{w}$' denote weak convergence in the space of all functions $f : \overline{\mathbb{R}}_+^n \to \mathbb{R}$ which are locally uniformly-bounded on every compact subset of $\overline{\mathbb{R}}_+^n$. Next, extending Schmidt and Stadtmüller (2005, Theorem 6) from the bivariate to the d-dimensional setting, we have

$$\sqrt{k}\left(\widehat{\lambda}^{\text{emp}}(\boldsymbol{x};\, k) - \lambda^{\boldsymbol{X}}(\boldsymbol{x})\right) \xrightarrow{w} B(\boldsymbol{x}) - \sum_{i=1}^{d} \frac{\partial}{\partial x_i}\lambda^{\boldsymbol{X}}(\boldsymbol{x}) B_i(x_i),$$

where $B_i(x) = B(\infty, \ldots, \infty, x, \infty, \ldots, \infty)$, x is the i-th component and B is a zero mean Wiener process with covariance structure $E(B(\boldsymbol{x})B(\boldsymbol{y})) = \lambda^{\boldsymbol{X}}(\boldsymbol{x} \wedge \boldsymbol{y})$.

To show asymptotic normality we use an extended version of the classical delta-method, for details see van der Vaart and Wellner (1996, p.374). First, note that for all $i \neq j$ and for λ defined in (4.3)

$$\begin{aligned}
&\inf_{\theta \in Q_{ij}^*} |\lambda'^{\alpha}(\sqrt{2}\cos\theta, \sqrt{2}\sin\theta,\ \alpha, \rho_{ij})| > 0,\\
&\inf_{\theta \in U_{ij}^*} |\lambda'^{\rho}(\sqrt{2}\cos\theta, \sqrt{2}\sin\theta,\ \alpha, \rho_{ij})| > 0 \quad \text{and}\\
&\sup_{\theta \in U_{ij}^*} \left|(\lambda^{\leftarrow\rho})'^{\alpha}(\sqrt{2}\cos\theta, \sqrt{2}\sin\theta,\ \alpha, \rho_{ij})\right| < \infty.
\end{aligned}$$

Next, define $\mathcal{TC}$ as the set of all d-dimensional tail copulae. By Schmidt and Stadtmüller (2005, Theorem 1(iii)) a tail copula is Lipschitz-continuous, hence $\mathcal{TC}$ is a subset of a topological vector space. Abbreviate for λ defined in (4.3) and $\mu \in \mathcal{TC}$ with μ_{ij} being the ij-th marginal of μ

$$\begin{aligned}
\widetilde{\alpha}_{ij}(\theta, \mu, \rho) &:= \lambda^{\leftarrow\alpha}\left(\mu_{ij}(\sqrt{2}\cos\theta, \sqrt{2}\sin\theta); \sqrt{2}\cos\theta, \sqrt{2}\sin\theta, \rho\right), \quad \text{and}\\
\widetilde{\rho}_{ij}(\theta, \mu, \alpha) &:= \lambda^{\leftarrow\rho}\left(\mu_{ij}(\sqrt{2}\cos\theta, \sqrt{2}\sin\theta);\ \sqrt{2}\cos\theta, \sqrt{2}\sin\theta, \alpha\right).
\end{aligned}$$

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

$$\begin{aligned}
\alpha(\mu, \mathbf{R}) &:= \frac{1}{d(d-1)} \sum_{i\neq j} \frac{1}{W(Q^*_{ij})} \int_{\theta\in Q^*_{ij}} \widetilde{\alpha}_{ij}(\theta, \mu, \rho_{ij})\, W(d\theta), \\
\rho_{ij}(\mu, \mathbf{R}) &:= \frac{1}{W^*(U^*_{ij})} \int_{\theta\in U^*_{ij}} \widetilde{\rho}_{ij}(\theta, \mu, \alpha(\mu, \mathbf{R}))\, W^*(d\theta), \quad \text{and} \\
\boldsymbol{r}(\mu, \mathbf{R}) &:= \operatorname{vec}\left[(\rho_{ij}(\mu, \mathbf{R}))_{1\le i,j\le d}\right].
\end{aligned}$$

Write $\alpha(\mu) := \alpha(\mu, \mathbf{R})$ and note that $\alpha(\mu)$ is Hadamard-differentiable, i.e. let $t_m \stackrel{m\to\infty}{\longrightarrow} \infty$ and $h_m \stackrel{m\to\infty}{\longrightarrow} h \in \mathcal{TC}$ such that $\mu + h_m/t_m \in \mathcal{TC}$ for all m. Then, using Taylor expansion,

$$\begin{aligned}
&\lim_{m\to\infty} t_m\left(\alpha(\mu + h_m/t_m) - \alpha(\mu)\right) \\
&= \frac{1}{d(d-1)} \sum_{i\neq j} \frac{1}{W(Q^*_{ij})} \int_{\theta\in Q^*_{ij}} \frac{h_{ij}(\sqrt{2}\cos\theta, \sqrt{2}\sin\theta)}{\lambda'^{\alpha}\left(\sqrt{2}\cos\theta, \sqrt{2}\sin\theta, \alpha(\mu), \rho_{ij}\right)}\, W(d\theta) \\
&=: \alpha'_\mu(h),
\end{aligned}$$

which obviously is a linear map. Analogously, $\rho_{ij}(\mu) := \rho_{ij}(\mu, \mathbf{R})$ is Hadamard differentiable, i.e.

$$\begin{aligned}
&\lim_{m\to\infty} t_m\left(\rho_{ij}(\mu + h_m/t_m) - \rho_{ij}(\mu)\right) \\
&= \frac{1}{W^*(U^*_{ij})} \int_{\theta\in U^*_{ij}} \frac{h_{ij}(\sqrt{2}\cos\theta, \sqrt{2}\sin\theta)}{\lambda'^{\rho}\left(\sqrt{2}\cos\theta, \sqrt{2}\sin\theta, \alpha(\mu), \rho_{ij}\right)} + \\
&\qquad + \alpha'_\mu(h)(\lambda^{\leftarrow\rho})'^{\alpha}\left(\mu_{ij}(\sqrt{2}\cos\theta, \sqrt{2}\sin\theta);\, \sqrt{2}\cos\theta, \sqrt{2}\sin\theta, \alpha(\mu)\right) W^*(d\theta) \\
&=: \rho'_{ij;\mu}(h),
\end{aligned}$$

and similarly for $\boldsymbol{r}(\mu, \mathbf{R})$. Since $\widehat{\mathbf{R}}_\tau - \mathbf{R} = o_p(1/\sqrt{k})$,

$$\widehat{\rho}_{ij}(k, w^*) = \rho_{ij}\left(\widehat{\lambda}^{\text{emp}}(\cdot; k), \widehat{\mathbf{R}}_\tau\right),$$

and similarly for $\widehat{\boldsymbol{r}}(k, w^*)$, the delta method yields

$$\sqrt{k}\left(\widehat{\boldsymbol{r}}(k, w^*) - \boldsymbol{r}\right) \stackrel{w}{\longrightarrow} \boldsymbol{r}'_{\lambda\boldsymbol{x}}(\widetilde{B}).$$

The result then follows using

$$E\left(\left(\boldsymbol{r}'_{\lambda\boldsymbol{x}}(\widetilde{B})\right)_{ij}\left(\boldsymbol{r}'_{\lambda\boldsymbol{x}}(\widetilde{B})\right)_{kl}\right) = \sigma_\alpha + \sigma_{ij,\alpha} + \sigma_{kl,\alpha} + \sigma_{ij,kl},$$

with σ_α, $\sigma_{ij,\alpha}$, $\sigma_{kl,\alpha}$, $\sigma_{ij,kl}$ defined through (4.18)–(4.20).

References

Anderson, T. (2003). *An introduction to multivariate statistical analysis* (Third ed.). Wiley Series in Probability and Statistics. Hoboken, NJ: Wiley.

Basrak, B., R. Davis, and T. Mikosch (2002). Regular variation of GARCH processes. *Stochastic Process. Appl. 99*(1), 95–115.

Bentler, P. and P. Dudgeon (1996). Covariance structure analysis: Statistical practice, theory and directions. *Annu. Rev. Psychol. 47*, 563–593.

Bingham, N., C. Goldie, and J. Teugels (1989). *Regular variation*, Volume 27 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press.

Bluhm, C., L. Overbeck, and C. Wagner (2003). *An introduction to credit risk modeling*. Chapman&Hall, London.

Browne, M. (1982). Covariance structures. In D. Hawkins (Ed.), *Topics in applied multivariate analysis*. Cambridge University Press.

Browne, M. (1984). Asymptotically distribution-free methods for the analysis of covariance structures. *British J. Math. Statist. Psych. 37*(1), 62–83.

Browne, M. and A. Shapiro (1986). The asymptotic covariance matrix of sample correlation coefficients under general conditions. *Linear Algebra Appl. 82*, 169–176.

Browne, M. and A. Shapiro (1987). Adjustments for kurtosis in factor analysis with elliptically distributed errors. *J. Roy. Statist. Soc. Ser. B 49*(3), 346–352.

Cambanis, S., S. Huang, and G. Simons (1981). On the theory of elliptically contoured distributions. *J. Multivariate Anal. 11*(3), 368–385.

Casella, G. and R. Berger (2001). *Statistical inference* (Second ed.). Belmont, CA: Duxbury Press.

Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance 1*(2), 223–236.

Danielsson, J., L. de Haan, L. Peng, and C. de Vries (2001). Using a bootstrap method to choose the sample fraction in tail index estimation. *J. Multivariate Anal. 76*(2), 226–248.

Drees, H. and E. Kaufmann (1998). Selecting the optimal sample fraction in univariate extreme value estimation. *Stochastic Process. Appl. 75*(2), 149–172.

Embrechts, P., F. Lindskog, and A. McNeil (2003). Modelling dependence with copulas and applications to risk management. In S. Rachev (Ed.), *Handbook of heavy tailed distributions in finance*, Chapter 8, pp. 329–384. Amsterdam: Elsevier.

Fang, K., S. Kotz, and K. Ng (1990). *Symmetric multivariate and related distributions*, Volume 36 of *Monographs on Statistics and Applied Probability*. London: Chapman and Hall.

Frahm, G., M. Junker, and R. Schmidt (2005). Estimating the tail-dependence coefficient: properties and pitfalls. *Insurance Math. Econom. 37*(1), 80–100.

Frahm, G., M. Junker, and A. Szimayer (2003). Elliptical copulas: applicability and limitations. *Statist. Probab. Lett. 63*(3), 275–286.

Frey, R., A. McNeil, and M. Nyfeler (2001, October). Copulas and credit models. *RISK*, 111–114.

Galambos, J. (1987). *The asymptotic theory of extreme order statistics* (Second ed.). Melbourne: Krieger Publishing.

Gumbel, E. (1960). Bivariate exponential distributions. *J. Amer. Statist. Assoc. 55*, 698–707.

Heffernan, J. and J. Tawn (2004). A conditional approach for multivariate extreme values. *J. R. Stat. Soc. Ser. B 66*(3), 497–546.

Higham, N. (2002). Computing the nearest correlation matrix—a problem from finance. *IMA J. Numer. Anal. 22*(3), 329–343.

Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statistics 19*, 293–325.

Hu, L., P. Bentler, and Y. Kano (1992). Can test statistics in covariance structure analysis be trusted? *Psychological Bulletin 112*, 351–362.

Hult, H. and F. Lindskog (2002). Multivariate extremes, aggregation and dependence in elliptical distributions. *Adv. in Appl. Probab. 34*(3), 587–608.

Kendall, M. and J. Gibbons (1990). *Rank correlation methods* (Fifth ed.). London: Edward Arnold.

Klüppelberg, C., G. Kuhn, and L. Peng (2005a). Estimating the tail dependence of an elliptical distribution. Submitted.

Klüppelberg, C., G. Kuhn, and L. Peng (2005b). Multivariate tail copula: modeling and estimation. Submitted.

Lawley, D. and A. Maxwell (1971). *Factor analysis as a statistical method* (Second ed.). New York: Elsevier Publishing.

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

Lee, A. (1990). *U-statistics*, Volume 110 of *Statistics: Textbooks and Monographs*. New York: Marcel Dekker Inc.

Lee, S. and P. Bentler (1980). Some asymptotic properties of constrained generalized least squares estimation in covariance structure modules. *South African Statist. J. 14*(2), 121–136.

Lindskog, F., A. McNeil, and U. Schmock (2003). Kendall's tau for elliptical distributions. In G. Bol, G. Nakhaeizadeh, S. Rachev, T. Ridder, and K. Vollmer (Eds.), *Credit risk – measurement, evaluation and management*. Physica-Verlag Heidelberg.

McNeil, A., R. Frey, and P. Embrechts (2005). *Quantitative risk management*. Princeton Series in Finance. Princeton, NJ: Princeton University Press.

Muirhead, R. and C. Waternaux (1980). Asymptotic distributions in canonical correlation analysis and other multivariate procedures for nonnormal populations. *Biometrika 67*(1), 31–43.

Nelsen, R. (1999). *An introduction to copulas*, Volume 139 of *Lecture Notes in Statistics*. New York: Springer-Verlag.

Pickands, J. (1981). Multivariate extreme value distributions. In *Proceedings of the 43rd session of the International Statistical Institute, Vol. 2 (Buenos Aires, 1981)*, Volume 49, pp. 859–878 and 894–902.

Resnick, S. (1987). *Extreme values, regular variation, and point processes*, Volume 4 of *Applied Probability. A Series of the Applied Probability Trust*. New York: Springer-Verlag.

Satorra, A. and P. Bentler (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika 66*(4), 507–514.

Schmidt, R. and Stadtmüller (2005). Non-parametric estimation of tail dependence. Scandinavian Journal of Statistics. To appear.

Sklar, A. (1996). Random variables, distribution functions, and copulas—a personal look backward and forward. In *Distributions with fixed marginals and related topics (Seattle, WA, 1993)*, Volume 28 of *IMS Lecture Notes Monogr. Ser.*, pp. 1–14. Hayward, CA: Inst. Math. Statist.

Steiger, J. (1994). Factor analysis in the 1980's and the 1990's: some old debates and some new developments. In I. Borg and P. Mohler (Eds.), *Trends and perspectives in empirical social research*, pp. 201–224. Walter de Gruyter, Berlin.

Steiger, J., A. Shapiro, and M. Browne (1985). On the multivariate asymptotic distribution of sequential chi-square statistics. *Psychometrika 50*(3), 253–264.

van der Vaart, A. and J. Wellner (1996). *Weak convergence and empirical processes*. Springer Series in Statistics. New York: Springer-Verlag.

van Praag, B., T. Dijkstra, and J. van Velzen (1985). Least-squares theory based on general distributional assumptions with an application to the incomplete observations problem. *Psychometrika 50*(1), 25–36.