Copula Structure Analysis

Claudia Klüppelberg * Gabriel Kuhn*

January 19, 2009

Abstract

In this paper we extend the standard approach of correlation structure analysis for dimension reduction 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. After introducing the new concept 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 small sample sizes. Finally, we show our method at work for a financial data set 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.

AMS 2000 Subject Classifications: primary: 62H05, 62H12, 62H25, 62P05, 62P15, 62P25;
secondary: 91B28, 91B70, 91B82

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

*Center for Mathematical Sciences, Munich University of Technology, D-85747 Garching, Germany, email: {cklu , gabriel}@ma.tum.de, www.ma.tum.de/stat/

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 random vector $\boldsymbol{X} \in \mathbb{R}^d$, 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 of a multivariate data set 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 $\boldsymbol{X} \stackrel{\mathrm{d}}{=} \boldsymbol{\mu} + \mathbf{L}\boldsymbol{f} + \mathbf{V}\boldsymbol{e}$, where $\boldsymbol{\mu} = (\mu_1, \dots, \mu_d)^\top$ is a location parameter, $\boldsymbol{f} = (f_1, \dots, f_m)^\top$ for $m < d$ is a vector of non-observable and (usually) uncorrelated *factors* and $\boldsymbol{e} = (e_1, \dots, e_d)^\top$ is a vector of *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 $(\boldsymbol{f}^\top, \boldsymbol{e}^\top)$ has mean zero and covariance matrix $\mathbf{I}$, the identity matrix. Then, describing the dependence structure of $\boldsymbol{X}$ through its covariance matrix yields $\mathrm{Cov}\boldsymbol{X} = \boldsymbol{\Sigma} = \mathbf{L}\mathbf{L}^\top + \mathbf{V}^2$, 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 normality assumption, 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 $\boldsymbol{X} \stackrel{\mathrm{d}}{=} \mathbf{A}(\boldsymbol{\vartheta})\boldsymbol{\xi}$, i.e. $\boldsymbol{X} \in \mathbb{R}^d$ can be described as a linear combination of some (latent) random variables $\boldsymbol{\xi} \in \mathbb{R}^q$ for $q \geq d$ 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, as is common 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 $\text{unif}(0, 1)$ margins and, by Sklar's theorem, each distribution function can be described through its margins and its copula separately. We focus on *elliptical copulae* being the copulae of elliptical distributions, which are very flexible and easy to handle also in high dimensions. As the correlation matrix is a parameter of an elliptical copula, correlation structure analysis can be extended to such copulae and we will call this method *copula structure analysis*.

The main advantage of our method is that we only need iid data to ensure consistency and asymptotic normality of the estimated factor loadings as well as the asymptotic χ^2-distribution of the test statistic for model selection (i.e. for the estimation of the number of latent factors). We require an elliptical copula only to allow for a meaningful interpretation of the analysis, the asymptotics of the estimators are not affected by the specific copula of the data's distribution. Furthermore, a simulation study shows that our approach works well for reasonably large sample sizes in the sense that the distribution of the test statistic is close to its asymptotic $\chi 2$-distribution. This is in contrast to other methods of correlation structure analysis or generalized structural equation models, which either need large sample sizes or distributional assumptions. Moreover, they show quite unstable behavior, if these assumptions are not met.

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 the copula dependence concept based on Kendall's tau and develop estimators, which can then be used for the copula structure analysis. We also derive asymptotic results for our estimates.

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

2 Elliptical copulae versus elliptical distributions

First, we give a short summary of the copula concept. For more technical background information we refer to Nelsen (1999).

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

Its importance is based on Sklar's Theorem, which ensures that every d-dimensional distribution function F with margins $F_1,\dots,F_d$ has representation

$$F(x_1,\dots,x_d) = C\left(F_1(x_1),\dots,F_d(x_d)\right), \quad \boldsymbol{x} = (x_1,\dots,x_d)^\top \in \mathbb{R}^d,$$

for some copula C. The copula C is unique on $\mathrm{Ran}F_1 \times \cdots \times \mathrm{Ran}F_d$. If F has continuous margins $F_1,\dots,F_d$, then the copula C of F is for $(u_1,\dots,u_d)^\top \in [0,1]^d$

$$C(u_1,\dots,u_d) = F\left(F_1^{\leftarrow}(u_1),\dots,F_d^{\leftarrow}(u_d)\right).$$

Recall that for a distribution function F on $\mathbb{R}$ the *generalized inverse* is defined as

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

We will focus on copulae of elliptical distributions, and we first give definitions and state some properties. For a general treatment of elliptical distributions we refer to Fang, Kotz, and Ng (1990), Fang, Fang, and Kotz (2002) and to Cambanis, Huang, and Simons (1981).

Definition 2.1. *(a) A random vector $\boldsymbol{X} \in \mathbb{R}^d$ has an elliptical distribution, if for $\boldsymbol{\mu} \in \mathbb{R}^d$, a 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 $\boldsymbol{U}^{(q)} \sim \mathrm{unif}\{\boldsymbol{s} \in \mathbb{R}^q : \boldsymbol{s}^\top\boldsymbol{s} = 1\}$ independent of G the distribution of $\boldsymbol{X}$ satisfies*

$$\boldsymbol{X} \stackrel{\mathrm{d}}{=} \boldsymbol{\mu} + G\mathbf{A}\boldsymbol{U}^{(q)} \quad \textit{with} \quad \mathbf{A} \in \mathbb{R}^{d\times q} \textit{ and } \mathbf{A}\mathbf{A}^\top = \boldsymbol{\Sigma}. \tag{2.1}$$

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 $\mathrm{Cov}\boldsymbol{X} = \boldsymbol{\Sigma}$.*
(b) We define the correlation matrix $\mathbf{R}$ *of $\boldsymbol{X}$ 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 $\mathrm{Corr}\boldsymbol{X} = \mathbf{R}$.*
(c) We define an elliptical copula *as the copula of $\boldsymbol{X} \sim \mathcal{E}_d(\boldsymbol{\mu},\boldsymbol{\Sigma},G)$, denoted by $\mathcal{EC}_d(\mathbf{R},G)$. We call $\mathbf{R}$ the* copula correlation matrix *and G the* generating variable.

Note that the notion $\mathcal{EC}_d(\mathbf{R},G)$ for an elliptical copula makes sense, since it is characterized by the generating variable G (unique up to a multiplicative constant) and the copula correlation matrix $\mathbf{R}$. This follows as a simple consequence of the definition and the fact that copulae are invariant under strictly increasing transformations.

In general, the dimension q of $\boldsymbol{U}^{(q)}$ can be arbitrary compared to d, both for the distribution and, by Sklar's theorem, for its copula. However, for $q < d$ the random vector

$\boldsymbol{X}$ lies in a q-dimensional subspace of $\mathbb{R}^d$, which transfers to the copula (cf. the uniqueness part of Sklar's theorem). We avoid such sophistications and restrict ourselves throughout to $q \geq d$.

Remark 2.2. (a) With $\boldsymbol{\mu} = \mathbf{0}$ and $\mathbf{A} = \mathbf{I}_d$, the random vector $\boldsymbol{X} \stackrel{\mathrm{d}}{=} G\boldsymbol{U}^{(d)}$ is *spherical.* Then $\boldsymbol{X}$ has characteristic function

$$E\big(e^{i\boldsymbol{t}^\top \boldsymbol{X}}\big) = E\big(e^{i\boldsymbol{t}^\top G\boldsymbol{U}^{(q)}}\big) = \psi(\boldsymbol{t}^\top \boldsymbol{t})\,, \quad \boldsymbol{t} \in \mathbb{R}^d\,.$$

Denoting $\|\boldsymbol{t}\| = \sqrt{\boldsymbol{t}^\top \boldsymbol{t}}$, we also have

$$\big(G, \boldsymbol{U}^{(d)}\big) = \left(\|G\boldsymbol{U}^{(d)}\|, \frac{G\boldsymbol{U}^{(d)}}{\|G\boldsymbol{U}^{(d)}\|}\right)\,.$$

(b) If $\boldsymbol{X}$ is elliptic as defined in (2.1) with $\boldsymbol{\mu} = \mathbf{0}$, then it has characteristic function

$$E\big(e^{i\boldsymbol{t}^\top \boldsymbol{X}}\big) = e^{i\boldsymbol{t}^\top \boldsymbol{\mu}} E\big(e^{-i(\mathbf{A}^\top \boldsymbol{t})^\top G\boldsymbol{U}^{(q)}}\big) = e^{i\boldsymbol{t}^\top \boldsymbol{\mu}} \psi\left(\boldsymbol{t}^\top \boldsymbol{\Sigma} \boldsymbol{t}\right)\,, \quad \boldsymbol{t} \in \mathbb{R}^d\,.$$

Moreover, this implies

$$\big(\Lambda, \Xi\big) := \left(\sqrt{(\boldsymbol{X}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{X}-\boldsymbol{\mu})}, \frac{\Sigma^{-1/2}(\boldsymbol{X}-\boldsymbol{\mu})}{\sqrt{(\boldsymbol{X}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{X}-\boldsymbol{\mu})}}\right) \stackrel{\mathrm{d}}{=} \big(G, \boldsymbol{U}^{(q)}\big)$$

and Λ^2 and Ξ are independent. □

In the next example we present some classical elliptical distributions and their copulae.

Example 2.3. [Normal variance mixture model and its copula]
(a) Let $\boldsymbol{X} \stackrel{\mathrm{d}}{=} \boldsymbol{\mu} + \sqrt{W}\mathbf{A}\boldsymbol{Z}$ with $\boldsymbol{\mu} \in \mathbb{R}^d$, $\mathbf{A} \in \mathbb{R}^{d\times q}$ a matrix of rank $d \leq q$, $\mathbf{Z} \in \mathbb{R}^q$ a standard normal vector and $W > 0$ a random variable, independent of $\boldsymbol{Z}$. Then $\boldsymbol{X}$ is a *normal variance mixture model* with characteristic function

$$E\big(e^{i\boldsymbol{t}^\top \mathbf{X}}\big) = e^{i\boldsymbol{t}^\top \boldsymbol{\mu}} E\big(e^{-(W/2)\boldsymbol{t}^\top \boldsymbol{\Sigma} \boldsymbol{t}}\big) = e^{i\boldsymbol{t}^\top \boldsymbol{\mu}} \Psi(\boldsymbol{t}^\top \boldsymbol{\Sigma}\, \boldsymbol{t})\,, \quad \boldsymbol{t} \in \mathbb{R}^d\,,$$

i.e. $\boldsymbol{X}$ is elliptic. We calculate $\Lambda^2 = (\sqrt{W}\mathbf{A}\boldsymbol{Z})^\top \boldsymbol{\Sigma}^{-1}(\sqrt{W}\mathbf{A}\boldsymbol{Z}) = W\boldsymbol{Z}^\top \boldsymbol{Z}$, where $\boldsymbol{Z}^\top \boldsymbol{Z}$ is χ^2_d-distributed and independent of W.

(b) In the situation of part (a), if W has an inverse gamma distribution with parameters $(\frac{\nu}{2}, \frac{\nu}{2})$, then $\frac{\nu}{W} \sim \chi^2_\nu$, which implies that $\frac{\Lambda^2}{d} \sim \frac{\nu \chi^2_d}{d\chi^2_\nu}$, which is $F(d,\nu)$-distributed. Moreover, we have $\mathbf{X} - \boldsymbol{\mu} \stackrel{\mathrm{d}}{=} \dfrac{\mathbf{A}\boldsymbol{Z}}{\sqrt{\chi^2_\nu/\nu}} \sim \boldsymbol{t}_\nu(\mathbf{0}, \boldsymbol{\Sigma})$; i.e. $\mathbf{X} - \boldsymbol{\mu}$ is a $\boldsymbol{t}$-distributed vector with ν degrees of freedom. Further, if $\nu > 2$, then $\mathbf{X} - \boldsymbol{\mu}$ has covariance matrix $\dfrac{\nu}{\nu-2}\boldsymbol{\Sigma}$.

(c) Let $\boldsymbol{Z} = (Z_1, \dots, Z_d)$ be standard normal and denote by Φ the standardnormal distribution function in $\mathbb{R}$. Then $(\Phi(Z_1), \dots, \Phi(Z_d))$ is a Gaussian copula. Let $\boldsymbol{X} = (X_1, \dots, X_d)$ be $\boldsymbol{t}_\nu(\mathbf{0}, \boldsymbol{\Sigma})$ and denote by t_ν the t-distribution function in $\mathbb{R}$ with ν degrees of freedom, then $(t_\nu(X_1), \dots, t_\nu(X_d))$ is a $\boldsymbol{t}_\nu$ copula. □

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. A *correlation structure model* is then a function

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

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 vec$[\cdot]$ the column vector formed from the non-duplicated and non-fixed elements of a symmetric matrix. Fixed elements are known, whereas non-fixed elements have to be estimated from the data. 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 with all elements non-fixed

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

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

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

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

3.1 The model

As by Definition 2.1 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, $\mathbf{A} : \Theta \to \mathbb{R}^{d \times q}$ a matrix valued function and $\boldsymbol{\xi} \sim \mathcal{E}_q(\mathbf{0}, \mathbf{I}, G)$ a q-dimensional elliptical random vector with continuous generating variable $G > 0$ and $q \geq d$. 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 $\mathbf{R}(\boldsymbol{\vartheta}) := \mathrm{diag}[\boldsymbol{\Sigma}(\boldsymbol{\vartheta})]^{-1/2}\boldsymbol{\Sigma}(\boldsymbol{\vartheta})\mathrm{diag}[\boldsymbol{\Sigma}(\boldsymbol{\vartheta})]^{-1/2}$ *and* $\boldsymbol{\Sigma}(\boldsymbol{\vartheta}) := \mathbf{A}(\boldsymbol{\vartheta})\mathbf{A}(\boldsymbol{\vartheta})^\top$.

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 $\boldsymbol{X} \stackrel{\mathrm{d}}{=} \boldsymbol{F}^{\leftarrow}(\boldsymbol{H}(\mathbf{A}(\boldsymbol{\vartheta})\boldsymbol{\xi}))$, 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 is a classical correlation or covariance structure model. For normal $\boldsymbol{\xi}$ it gives back the classical normal model and for elliptical $\boldsymbol{\xi}$ the elliptical model of Browne (1984).

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 Definition 2.1(c).

Example 3.2. (a) For classical factor analysis, (3.3) translates to $\boldsymbol{\vartheta} = \mathrm{vec}[\mathbf{L}, \mathbf{V}]$, $\mathbf{R}(\boldsymbol{\vartheta}) = \mathbf{L}\mathbf{L}^\top + \mathbf{V}^2$ 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 $\boldsymbol{\xi} \sim \mathcal{E}_q(\mathbf{0}, \mathbf{I}, G)$ with $q = m + d$ such that

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

We call this identity a *copula factor model.*

(b) Generalized covariance structure analysis aims at a model

$$\boldsymbol{\Sigma}(\boldsymbol{\vartheta}) \quad = \quad \mathbf{F}_1\mathbf{P}_1\mathbf{F}_1^\top + \cdots + \mathbf{F}_m\mathbf{P}_m\mathbf{F}_m^\top,$$

where $\mathbf{F}_k = F_{k,1}\cdots F_{k,n(k)}$ for arbitrary matrices $F_{k,i}$ and the matrices $\mathbf{P}_k$ are symmetric, $\boldsymbol{\vartheta} = \mathrm{vec}[F_{1,1}, \ldots, F_{1,n(1)}, F_{2,1}, \ldots, F_{m,n(m)}, \mathbf{P}_1, \ldots, \mathbf{P}_m]$, and some values in all matrices can be fixed. Note that the above copula factor model can be seen as a special case of a corresponding generalized copula covariance model. For more details about latent structure models like *explanatory factor analysis*, *confirmatory factor analysis* and *general structural equation models* see e.g. Loehlin (2001), Bollen (1989), Everitt (1984), or Fuller (1987) for structure models in the context of observable variables. Further note that such particular generalized covariance structure models are provided by statistical software packages like SAS and LISREL. They also allow, in particular for ordinal data, to estimate the Kendall's tau matrix.

(c) Another well know correlation structure model is the *latent variable model.* Assume a factor model $\boldsymbol{X} = \mathbf{L}\boldsymbol{f} + \mathbf{V}\boldsymbol{e}$, where $\boldsymbol{X} \in \mathbb{R}^d$ is assumed to have mean $\mathbf{0}$ and correlation matrix $\mathbf{R}$. For $1 \le i \le d$ let $\boldsymbol{D}^i := (D^i_{-1}, D^i_0, \ldots, D^i_n)$ with $-\infty = D^i_{-1} < D^i_0 < \cdots < D^i_n = \infty$ be deterministic cut-off levels and define

$$Y_i = j \quad \Longleftrightarrow \quad D^i_{j-1} < X_i \le D^i_j \quad \text{for } j = \{1, \ldots, n\}.$$

Then $(X_i, \boldsymbol{D}^i)_{i=1,\ldots,d}$ is called the *latent variable model* for the state vector $(Y_1, \ldots, Y_d)$. It is straightforward to show that the distribution of $(Y_1, \ldots, Y_d)$ is uniquely determined by the copula C and the marginal distributions $p_{ij} := P(X_i \le D^i_j)$, cf. Frey and McNeil (2003, Proposition 3.2). Therefore, concerning the distribution of $\boldsymbol{Y}$, the assumption of $\boldsymbol{X} \stackrel{\mathrm{d}}{=} \mathbf{L}\boldsymbol{f} + \mathbf{V}\boldsymbol{e}$ is equivalent to $C_{\boldsymbol{X}} = C_{(\mathbf{L},\mathbf{V})\boldsymbol{\xi}}$ for $\boldsymbol{\xi} \in \mathbb{R}^{m+d}$. □

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 random vectors in $\mathbb{R}^d$ and denote by $\widehat{\mathbf{R}} := \widehat{\mathbf{R}}(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n)$ an arbitrary estimator of the correlation matrix $\mathbf{R}$ of $\boldsymbol{X}$ as for instance the empirical correlation or a copula 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 well as possible. Similarly to (3.2), we define $\widehat{\boldsymbol{r}} := \mathrm{vec}[\widehat{\mathbf{R}}]$ and $\boldsymbol{r}(\boldsymbol{\vartheta}) := \mathrm{vec}[\mathbf{R}(\boldsymbol{\vartheta})]$.

We estimate $\boldsymbol{\vartheta}$ by minimizing the *quadratic (or weighted least squares) discrepancy function* defined as

$$D(\widehat{\boldsymbol{r}}, \boldsymbol{r}(\boldsymbol{\vartheta}) | \boldsymbol{\Upsilon}) \quad = \quad (\widehat{\boldsymbol{r}} - \boldsymbol{r}(\boldsymbol{\vartheta}))^\top \boldsymbol{\Upsilon}^{-1} (\widehat{\boldsymbol{r}} - \boldsymbol{r}(\boldsymbol{\vartheta})), \tag{3.5}$$

where $\boldsymbol{\Upsilon}$ is a positive definite matrix or a consistent estimator of some positive definite matrix. For more details, see Steiger, Shapiro, and Browne (1985).

Alternatively, an often used discrepancy function is the one derived from the normal maximum likelihood function, see e.g. Lawley and Maxwell (1971). In our set-up we take the normal maximum likelihood estimator as initial value for the numerical optimization procedure to follow. For more details about discrepancy functions, their properties, advantages and drawbacks, we refer to Bentler and Dudgeon (1996) and Steiger (1994).

Given a 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 3.3 (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 bounded and closed parameter space. Further assume that $\widehat{\boldsymbol{r}}$ is an estimator based on an iid sample $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ in $\mathbb{R}^d$ 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* $D(\boldsymbol{r}_0, \boldsymbol{r}(\boldsymbol{\vartheta})|\,\boldsymbol{\Upsilon})$ *in* Θ*. Assume also that the Jacobian matrix* $\partial \boldsymbol{r}(\boldsymbol{\vartheta})/\partial \boldsymbol{\vartheta}$ *is continuous in* $\boldsymbol{\vartheta}$*. Define the estimator*

$$\widehat{\boldsymbol{\vartheta}} \quad := \quad \arg\min_{\boldsymbol{\vartheta}\in\Theta} D(\widehat{\boldsymbol{r}}, \boldsymbol{r}(\boldsymbol{\vartheta})|\,\boldsymbol{\Upsilon}). \tag{3.6}$$

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

Under the assumptions of Proposition 3.3, we define the test statistic

$$T_{\boldsymbol{\Upsilon}} \quad := \quad n\widehat{D}_{\boldsymbol{\Upsilon}} \;=\; nD(\widehat{\boldsymbol{r}}, \boldsymbol{r}(\widehat{\boldsymbol{\vartheta}})|\,\boldsymbol{\Upsilon}) \;=\; n \min_{\boldsymbol{\vartheta}\in\Theta} D(\widehat{\boldsymbol{r}}, \boldsymbol{r}(\boldsymbol{\vartheta})|\,\boldsymbol{\Upsilon}), \tag{3.7}$$

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

$$H_0: \quad \boldsymbol{r}_0 \;=\; \boldsymbol{r}(\boldsymbol{\vartheta}_0) \quad \text{for some } \boldsymbol{\vartheta}_0 \in \Theta. \tag{3.8}$$

To obtain the limit distribution of $T_{\boldsymbol{\Upsilon}}$ for the quadratic discrepancy function (3.5), we apply the following result.

Theorem 3.4. *(Browne (1984, Corollary 4.1)) Assume that the conditions of Proposition 3.3 hold and that* $\boldsymbol{\vartheta}_0$ *is an interior point of* Θ*. Furthermore, assume that* $\sqrt{n}(\widehat{\boldsymbol{r}} - \boldsymbol{r}_0) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \boldsymbol{\Gamma})$ *and that* $\widehat{\boldsymbol{\Gamma}}$ *is a consistent estimator of* $\boldsymbol{\Gamma}$*, the asymptotic covariance matrix of* $\widehat{\boldsymbol{r}}$*. Finally, assume that the* $p \times d$ *Jacobian matrix*

$$\triangle \;=\; \left.\frac{\partial \boldsymbol{r}(\boldsymbol{\vartheta})}{\partial \boldsymbol{\vartheta}}\right|_{\boldsymbol{\vartheta}=\boldsymbol{\vartheta}_0} \tag{3.9}$$

has full column rank p*. Then, under the null hypothesis* (3.8)*,*

$$T_{\widehat{\boldsymbol{\Gamma}}} \;=\; n\widehat{D}_{\widehat{\boldsymbol{\Gamma}}} \;\xrightarrow{d}\; \chi^2_{df}\,, \quad n \to \infty\,, \tag{3.10}$$

where $df = d(d-1)/2 - p$ *and* p *is the dimension of* $\boldsymbol{\vartheta}$*.*

3.3 Model selection

To select an appropriate structural model, we consider a set of g models (which all have to satisfy the assumptions of Theorem 3.4)

$$\boldsymbol{r}^{(i)} : \Theta^{(i)} \to \mathbb{R}^{d(d-1)/2}, \quad \boldsymbol{\vartheta}^{(i)} \mapsto \boldsymbol{r}^{(i)}(\boldsymbol{\vartheta}^{(i)}), \quad \text{and } \Theta^{(i)} \subset \mathbb{R}^{p^{(i)}}, \quad 1 \le i \le g. \tag{3.11}$$

Further, we require that the g models are *nested*, i.e. for every $1 \le i \le g-1$ 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

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

and assume that at least one of these null hypotheses holds true; i.e. 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, we have to estimate j, the smallest index where the null hypothesis is not rejected. By Theorem 3.4, the corresponding test statistics

$$T^{(i)} := nD(\widehat{\boldsymbol{r}}, \boldsymbol{r}^{(i)}(\boldsymbol{\vartheta}_0^{(i)}) \mid \widehat{\boldsymbol{\Gamma}}) := n \min_{\boldsymbol{\vartheta} \in \Theta^{(i)}} D(\widehat{\boldsymbol{r}}, \boldsymbol{r}^{(i)}(\boldsymbol{\vartheta}) \mid \widehat{\boldsymbol{\Gamma}})$$

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 3.4. 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.

For a proof of the asymptotic normality of $\widehat{\boldsymbol{\vartheta}}$ with covariance matrix $(\Delta^\top \boldsymbol{\Gamma}^{-1} \Delta)^{-1}$ with Δ as in (3.9) we refer to Browne (1984, Corollary 2.1). If $\widehat{\boldsymbol{\Gamma}}$ is consistent, then $T^{(i)}$, $1 \leq i < j$, 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). Regarding the limiting distribution of $T^{(j)}$, if $\widehat{\boldsymbol{\Gamma}}$ is not consistent, see Satorra and Bentler (2001) or van Praag, Dijkstra, and van Velzen (1985).

In general, a unique *true* parameter $\boldsymbol{\vartheta}_0$ need not exist: in the classical factor model (see Example 3.2, where $\mathbf{R} = \mathbf{L}\mathbf{L}^\top + \mathbf{V}^2$), $\mathbf{L}$ 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. $\mathbf{L}^\top \mathbf{V}^{-2} \mathbf{L}$ has to be diagonal), $\widehat{\boldsymbol{\vartheta}}$ can be forced to be unique and Proposition 3.3 applies; see Lawley and Maxwell (1971, Section 2.3). By Lee and Bentler (1980) the degrees of freedom in (3.10) 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). With this correction for uniqueness, the factor model of Example 3.2(a) satisfies the regularity conditions of Proposition 3.3 and Theorem 3.4, see Steiger et al. (1985, Section 4) and Browne (1984, Section 5).

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

For the computation of $\widehat{\boldsymbol{\vartheta}}$ as the minimizer of the quadratic discrepancy function as in (3.6) and the test statistic $T_{\boldsymbol{\Upsilon}}$ from (3.7) we used the statistical software package `R` and the optimization routine `optim` with the `Nelder-Mead` method therein. By adding appropriate penalty terms to the discrepancy functions, we took both side conditions into account, i.e. that $\mathbf{L}^\top \mathbf{V}^{-2} \mathbf{L}$ is diagonal and $\operatorname{diag}(\widehat{\mathbf{L}}\widehat{\mathbf{L}}^\top + \widehat{\mathbf{V}}^2) = \mathbf{1}$. As starting values for the optimization algorithm, we took the loadings derived from the standard factor analysis routine `factanal` (which uses the normal maximum likelihood discrepancy function).

4 Methodology

As we consider a copula structure model, according to Theorem 3.4 we need an estimator $\widehat{\mathbf{R}}$ of the copula correlation matrix $\mathbf{R}$, whose limit distribution is $\mathcal{N}(\mathbf{0}, \mathbf{\Gamma})$ for some non-degenerate covariance matrix $\mathbf{\Gamma}$, and a consistent estimator of $\mathbf{\Gamma}$. In the following we will introduce a copula based dependence concept and its 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 a dependence concept which is related to correlation and a copula property we use *Kendall's tau*.

Kendall's tau τ_{ij} between two components (X_i, X_j) of a random vector $\boldsymbol{X} \in \mathbb{R}^d$ 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 $(\widetilde{X}_i, \widetilde{X}_j)$ are the corresponding components of an independent copy $\widetilde{\boldsymbol{X}}$ of $\boldsymbol{X}$. Moreover, we call $\mathbf{T} := (\tau_{ij})_{1 \leq i,j \leq d}$ the *Kendall's tau matrix*.

This dependence concept provides a valid 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; see Embrechts, McNeil, and Straumann (2002) for a readable discussion. Originally, Kendall's tau 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).

Concerning elliptical copulae the following result is given in Fang, Fang, and Kotz (2002), see also Hult and Lindskog (2002).

Theorem 4.1. *(Fang, Fang, and Kotz (2002, Theorem 3.1)) Let $\boldsymbol{X}$ be a vector of random variables with elliptical copula $C \in \mathcal{EC}_d(\mathbf{R}, G)$ and absolutely continuous generating variable $G > 0$, then $\tau_{ij} = 2\arcsin(\rho_{ij})/\pi$.*

Theorem 4.1 implies for an elliptical copula with continuous G that the correlation matrix $\mathbf{R}$ is a function of Kendall's tau. In Section 4.2 we will invoke this functional relationship for the estimation of $\mathbf{R}$. Of course, copula structure analysis can be applied to any copula correlation estimator with a certain limiting behavior as required in Theorem 3.4. Using Kendall's tau for estimation can then be seen as a robust extension of the usual correlation structure analysis.

Another alternative candidate for the linear correlation could be Spearman's ρ^S. For a bivariate normal vector (X_1, X_2) with correlation ρ it is well-known that

$$\rho^S(X_1, X_2) = \frac{6}{\pi} \arcsin \frac{\rho}{2} .$$

Unfortunately, simple formulas for Spearman's ρ^S for other than the Gaussian distribution are to our knowledge not available. Moreover, it has been shown in Lindskog et al. (2003) that ρ^S in the case of an elliptical copula depends on the distribution of the generating variate G. Consequently, there can be no one-to-one relationship between Spearman's ρ^S and ρ.

The next section explains the estimation procedures and presents asymptotic results.

4.2 Estimating the copula correlation matrix

Our method is based on Kendall's tau, which can be used by Theorem 4.1 for estimating the correlation matrix $\mathbf{R}$. The properties of its empirical version can be derived from general results on U-statistics; see Lee (1990). The following results go back to Hoeffding (1948).

Given an iid sample $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ with $\boldsymbol{X}_i = (X_{i,1}, \ldots, X_{i,d})^\top$, 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 $i = 1, \ldots, d$ and

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

Note that a naive implementation of this estimator is numerically slow, since it requires a computation time of order $\mathcal{O}(d^2 n^2)$ for sample size n. For faster algorithms with computation time $\mathcal{O}(d^2 n \ln n)$ and additional adjustments for duplicate entries see e.g. Christensen (2005).

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

Theorem 4.2. *Let $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots$ be an iid sequence in $\mathbb{R}^d$ satisfying the conditions of Theorem 4.1. Further, define*

$$\widehat{\mathbf{R}}_\tau = (\widehat{\rho}^{\tau}_{ij})_{1 \le i,j \le d} \;:=\; \sin\left(\frac{\pi}{2} \widehat{\mathbf{T}} \right), \tag{4.1}$$

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

$$\sqrt{n}\,(\widehat{\boldsymbol{r}}_\tau - \boldsymbol{r}) \quad \stackrel{d}{\longrightarrow} \quad \mathcal{N}_{d(d-1)/2}(\mathbf{0}, \boldsymbol{\Gamma}_\tau)\,, \qquad n \to \infty\,, \tag{4.2}$$

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

$$\begin{aligned}\gamma^\tau_{ij,kl} &= \pi^2 \cos\left(\frac{\pi}{2}\tau_{ij}\right)\cos\left(\frac{\pi}{2}\tau_{kl}\right)(\tau_{ij,kl} - \tau_{ij}\tau_{kl}) \quad \textit{and} \\ \tau_{ij,kl} &= E\Big(E\left(\text{sign}\,[(X_{1,i}{-}X_{2,i})(X_{1,j}{-}X_{2,j})]\ \big|\boldsymbol{X}_1\right)E\left(\text{sign}\,[(X_{1,k}{-}X_{2,k})(X_{1,l}{-}X_{2,l})]\ \big|\boldsymbol{X}_1\right)\Big).\end{aligned} \tag{4.3}$$

It is worth mentioning that also for iid $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots$, which do not have an elliptical copula, the limit relation (4.2) still holds, provided $\boldsymbol{r}$ is replaced by vec$[\mathbf{T}\pi/2]$. This means that the asymptotic properties of the Kendall's tau based correlation estimator do not depend on the copula class of the $\boldsymbol{X}_i$'s.

An estimator of $\boldsymbol{\Gamma}_\tau = (\gamma^\tau_{ij,kl})_{1\le i\ne j,k\ne l\le d}$ can be defined by the empirical version of (4.3).

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

$$\widehat{\gamma}^\tau_{ij,kl} := \pi^2 \cos\left(\frac{\pi}{2}\widehat{\tau}_{ij}\right)\cos\left(\frac{\pi}{2}\widehat{\tau}_{kl}\right)(\widehat{\tau}_{ij,kl} - \widehat{\tau}_{ij}\widehat{\tau}_{kl}) \quad \text{and} \tag{4.4}$$

$$\begin{aligned}\widehat{\tau}_{ij,kl} := \frac{1}{n(n-1)^2}\sum_{p=1}^n \Bigg[&\left(\sum_{q=1,\,q\ne p}^n \text{sign}\,((X_{p,i}{-}X_{q,i})(X_{p,j}{-}X_{q,j}))\right) \\ &\times\left(\sum_{q=1,\,q\ne p}^n \text{sign}\,((X_{p,k}{-}X_{q,k})(X_{p,l}{-}X_{q,l}))\right)\Bigg].\end{aligned} \tag{4.5}$$

Similarly to the remark after the Definition of Kendall's tau, a naive implementation of this estimator without numerical improvements requires a computation time of the order $\mathcal{O}(d^4n^3)$; we also used this simple one for the simulation and data example in section 5.

The following result is proved in Section 6.

Theorem 4.3. *Under the assumptions of Theorem 4.2* vec$[\widehat{\boldsymbol{\Gamma}}_\tau]$ *is consistent and asymptotically normal.*

Unfortunately, both the Kendall's tau based estimated correlation matrix (4.1) as well as its estimated asymptotic covariance matrix may sometimes not be positive definite. In such a case, the estimator can be replaced by its projection into the class of correlation or covariance matrices, respectively. However, this may lead only to a positive semi-definite matrix.

Using, for instance, the Euclidean norm $\|\mathbf{\Gamma}\|^2 = \sum_{i,j} \gamma_{i,j}^2$, the projection $\mathbf{\Gamma}^*$ of $\widehat{\mathbf{\Gamma}}$ to the class of covariance matrices is

$$\mathbf{\Gamma}^* \;=\; \arg\inf \left\{ \|\widehat{\mathbf{\Gamma}} - \mathbf{\Gamma}\| \,:\, \mathbf{\Gamma} \text{ is symmetric and positive definite} \right\}.$$

It can be shown that $\mathbf{\Gamma}^*$ is obtained by replacing the negative eigenvalues of $\widehat{\mathbf{\Gamma}}$ by 0, see Higham (2002).

In case of the correlation estimator, the projection $\mathbf{R}^*$ of $\widehat{\mathbf{R}}$ to the class of correlation matrices is

$$\mathbf{R}^* \;=\; \arg\inf \left\{ \|\widehat{\mathbf{R}} - \mathbf{R}\| \,:\, \mathbf{R} \text{ is symmetric and positive definite with diagonal } 1 \right\}.$$

An algorithm for computation of $\mathbf{R}^*$ iteratively replaces negative eigenvalues by 0 and then replaces the diagonal of the resulting matrix by 1, also see Higham (2002).

Note that the discrepancy function D requires a strictly positive definite covariance estimator $\widehat{\mathbf{\Gamma}}$. Further, the standard factor analysis routine (using the normal maximum likelihood discrepancy function) also requires strictly positive definite correlation matrices. Therefore, since we use the standard routine for computation of the starting values, we also need strictly positive definite correlation matrices. A simple pragmatic approach in this case would be to replace the negative eigenvalues not by 0 but by some small $\varepsilon > 0$.

5 The new method at work

Using the estimators (4.1) and (4.4) together with the quadratic discrepancy function (3.5), 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 $\boldsymbol{\xi} \sim \mathcal{E}_{m+d}(\mathbf{0}, \mathbf{I}, G)$; see also Example 3.2(a).

For the test statistic T, as defined in (3.7), based on the quadratic discrepancy function (3.5) we first compare in a simulation study the empirical distribution of T to its limiting χ^2-distribution as formulated in Theorem 3.4. Therefore, we define

$$T_\tau \;:=\; n \min_{\boldsymbol{\vartheta}\in\Theta} D\left(\widehat{\boldsymbol{r}}_\tau, \boldsymbol{r}(\boldsymbol{\vartheta}) \,|\, \widehat{\mathbf{\Gamma}}_\tau\right) \tag{5.1}$$

with the Kendall's tau based correlation estimators $\widehat{\boldsymbol{r}}_\tau = \mathrm{vec}[\widehat{\mathbf{R}}_\tau]$ as given in (4.1) and its estimated covariance matrix $\widehat{\mathbf{\Gamma}}_\tau$ as defined in (4.4), respectively.

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

$$T_\rho \;:=\; n \min_{\boldsymbol{\vartheta}\in\Theta} D\left(\widehat{\boldsymbol{r}}_{\mathrm{emp}}, \boldsymbol{r}(\boldsymbol{\vartheta}) \,|\, \widehat{\mathbf{\Gamma}}_{\mathrm{emp}}\right),$$

where $\widehat{\boldsymbol{r}}_{\text{emp}} = \text{vec}[\widehat{\mathbf{R}}_{\text{emp}}]$ is the vector of the empirical correlations. Furthermore, if we denote by $\boldsymbol{\Gamma}_{\text{emp}}$ the asymptotic covariance matrix of the normal limit in the central limit theorem analogous to (4.2) for $\widehat{\boldsymbol{r}}_{\text{emp}}$ provided $\boldsymbol{X}$ is normal, then $\widehat{\boldsymbol{\Gamma}}_{\text{emp}}$ denotes the maximum likelikhood estimator of $\boldsymbol{\Gamma}_{\text{emp}}$. It is a disadvantage of the classical correlation estimator $\widehat{\boldsymbol{r}}_{\text{emp}}$ that the covariance matrix $\boldsymbol{\Gamma}$ depends on the distribution of $\boldsymbol{X}$; for details see Browne and Shapiro (1986).

The parameter $\boldsymbol{\vartheta}$ is then estimated also in two different ways, denoted by $\widehat{\boldsymbol{\vartheta}}_\tau$ and $\widehat{\boldsymbol{\vartheta}}_\rho$, by minimizing T_τ and T_ρ, respectively.

5.1 Model selection by χ^2-tests

To see the performance of the quadratic test statistics T_τ from (5.1), we perform a simulation study. We choose a $d = 10$ dimensional setting with $m = 2$ factors and loadings as given in Table 1. Then $\mathbf{L}\mathbf{L}^\top + \mathbf{V}^2 = \mathbf{R}$ is a correlation matrix.

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
diag($\mathbf{V}^2$)	.19	.19	.19	.19	.19	.19	.19	.19	.19	.19

Table 1: Factor loadings of the simulation example

Recall that a multivariate $\boldsymbol{t}_\nu$-copula is the copula of the random vector $\sqrt{W}\boldsymbol{Z}$, where $\nu/W \sim \chi^2_\nu$ ($\nu > 0$) is independent of $\boldsymbol{Z} \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$; cf. Examples 2.3(b,c).

For $2 < \nu \leq 4$ the random vector $\sqrt{W}\boldsymbol{Z}$ has finite second moment, but its fourth moment does not exist. Hence, classical factor analysis applied on $\sqrt{W}\boldsymbol{Z}$ may only lead to a consistent estimate of $\boldsymbol{\vartheta}$, but model selection is not justified by Proposition 3.3 and Theorem 3.4.

Also note that for $4 < \nu < 8$ any estimator of the linear correlation's asymptotic covariance matrix $\boldsymbol{\Gamma}$ may only be consistent, but not asymptotically normal, and large sample sizes may be necessary to observe the limiting χ^2 distribution of the test statistic $T_{\widehat{\boldsymbol{\Gamma}}}$. Moreover, $\boldsymbol{\Gamma}_{\text{emp}}$ is the asymptotic covariance matrix of $\widehat{\boldsymbol{r}}_{\text{emp}}$ only, if $\sqrt{W}\boldsymbol{Z}$ is normal. On the other hand, the test statistic T_τ is not affected by the existence or non-existence of moments; cf. Theorems 4.2 and 4.3.

We simulate 500 iid samples of length $n = 100$ of the $\boldsymbol{t}_3$-copula, calculate the Kendall's tau based estimators (4.1) and (4.4) and calculate the test statistic T_τ from these.

In case of the 1-factor with $m = 1$, the empirical distribution of the test statistic T_τ was far off the expected χ^2_{df}-distribution (with $df = 35$ under the null hypothesis) rejecting obviously the 1-factor model on almost any confidence level.

In case of a 2-factor setting, to ensure uniqueness of the loadings, we use the restriction that $\mathbf{L}^\top\mathbf{V}^{-2}\mathbf{L}$ 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 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 3.4. Therefore, we compare the 500 estimates of T_τ with the χ^2_{26}-distribution by a QQ-plot, see Figure 1, left plot. From this plot we see that the empirical distribution of T_τ fits the χ^2_{26}-distribution quite well. The right plot of Figure 1 shows the same situation with sample size $n = 1\,000$ showing an almost perfect fit to the χ^2_{26}-distribution. We further compare some theoretical asymptotic acceptance rates α with the rates observed from the 500 samples, see Table 2. There, it can be seen that the observed acceptance rates are very close to the asymptotically expected rates. In case of sample size 100, the empirical rates are below the theoretical as also the QQ-plot is below the diagonal, and, in case of the sample size 1 000, the empirical rates are slightly above the theoretical as also the QQ-plot is in this region above the diagonal.

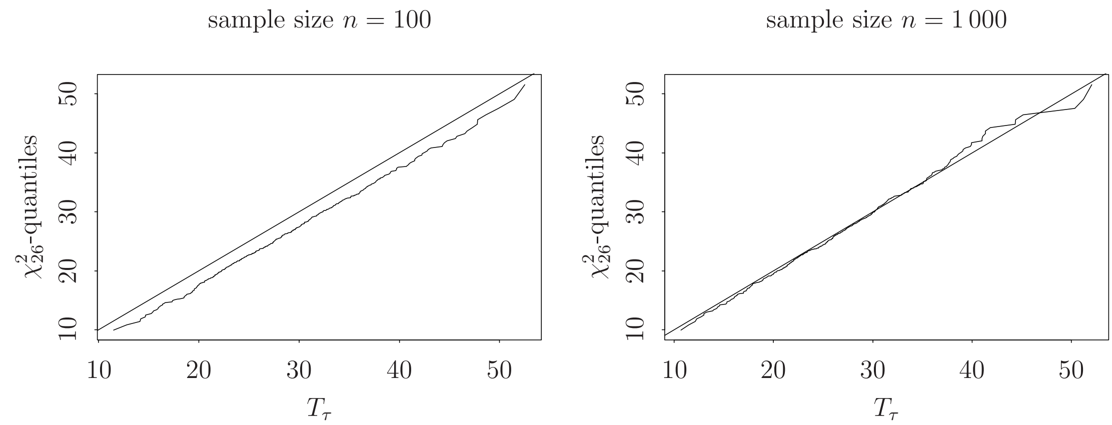

Figure 1: QQ-plots of ordered estimates T_τ against the χ^2_{26}-quantiles.

α	$\chi^2_{26,\alpha}$	$\widehat{F}_{100}(\chi^2_{26,\alpha})$	$\widehat{F}_{1\,000}(\chi^2_{26,\alpha})$
0.80	31.79	0.710	0.818
0.85	33.43	0.772	0.850
0.90	35.56	0.834	0.906
0.95	38.89	0.916	0.962
0.99	45.64	0.980	0.994

Table 2: Empirical acceptance rates based on the empirical distribution $\widehat{F}_n$ ($n = 100$ and $n = 1\,000$) estimated from the 500 simulated T_τ-statistics from figure 1. We denote by $\chi^2_{26,\alpha}$ the α-quantile of the χ^2-distribution with 26 degrees of freedom.

5.2 The oil-index-currency data

Before we apply our method to real life data, we have to discuss possibilities for verifying or falsifying the assumption of an elliptical copula as this is our basic requirement on the data. Unfortunately, there is no straightforward answer to this problem. Existing methods either test if data are elliptically (or spherically) distributed (cf. Manzotti, Pérez, and Quiroz (2002) and references therein) or test for certain parametric copula families like the normal, the $\boldsymbol{t}$ or the Archimedian copula families. For the suggested parametric tests it is, however, not clear, whether the dependence structure of a data set is really appropriately modeled by the chosen copula family. Information criterions, such as Akaike's information criterion (AIC) are not able to provide any understanding about the power of the decision rule employed. Alternatively, goodness-of-fit (GOF) tests have been suggested by various researchers and we refer to Panchenko (2005) and Berg (2007) for discussions of existing methods. GOF tests have the advantage that they are able to reject or fail to reject a parametric copula model.

Unfortunately, this does not provide a solution to our problem as an elliptical copula is not parameterized by a few parameters, but by the generating random variable G and the correlation matrix $\mathbf{R}$. At the moment we have to leave the important problem of formally testing for an elliptical copula for future research, and provide below only some reasoning, why the assumption of an elliptical copula may be an acceptable working hypothesis for our data.

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 log-returns from May 1985 to June 2004.

We start with an application of both Panchenko's and Berg and Bakken's GOF test to the oil-index-currency data based on the null hypotheses of a normal copula and of a $\boldsymbol{t}$-copula as the most prominent examples in the literature. For the whole dataset of 4 904 data, both, the normal copula and the $\boldsymbol{t}$-copula are rejected; the normal copula, however, at a much higher significance level than the $\boldsymbol{t}$-copula. Now, using a subsample of the last 904 observations the normal copula is still rejected, whereas the $\boldsymbol{t}$-copula is not rejected. Of course, the rejection of the normal copula also implies a rejection of the normal distribution indicating that usage of the standard estimation procedure could indeed lead to the selection of a wrong model. What remains is the $\boldsymbol{t}$-copula, which has completely different features than the normal copula as the generating random variable G is heavy-tailed, and can model dependence in large observations.

Certainly an important problem in our data is the fact that they cover almost 30 years of development. As market parameters have undoubtedly changed during this time it is not to be expected that the very same distributional model may hold for our data over the whole time interval. One simple indicator is, for instance, the currency change to the euro in 2002. For our data, before the change to euro we took the deutsch mark (dem) as reference currency and transfered it to euro by taking the fixed official exchange rate 1.95583 dem/eur. Although there is no guarantee, we believe that the rather weak assumption of an elliptical copula without having to specify a particular distribution provides a reasonable and robust working hypothesis for our data.

To our data set we fit a copula factor model using the T_ρ and T_τ statistics for estimation and model selection. The values of these test statistics, based on different numbers of factors are given in Table 3.

number of factors	df	T_ρ	T_τ	$\chi^2_{df,0.95}$
2	13	298.5	252.7	22.36
3	7	33.7	17.4	14.07
4	2	2.3	3.3	5.99

Table 3: Test statistics T_ρ and T_τ of oil-index-currency data under different number of factors.

To estimate the number of factors, we use a 95% confidence test, i.e. we reject the null hypothesis of having an 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 correlation and the Kendall's tau based test statistics.

Applying a 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 and the Kendall's tau based estimator. The last plot shows the loadings of the specific factors for both correlation estimators.

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

For the first factor the 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 both correlation estimators a large weight on Swiss Francs chf, so we call it *chf-factor*. We observe that the empirical factor and the Kendall's tau factor has almost no (or very little) correlation with oil, s&p500, gbp, usd and jpy. The weights on dkk and sek are larger.

Considering factor 3, we see for both factors 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.

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 3 it indicates that the distribution of T_ρ is far away from a χ^2 distribution. For the Kendall's tau factor there is some (weak) dependence indicated between the European currencies and the usd. Using this method, the oil-price factor is correctly categorized as specific factor.

Finally, we give an interpretation of the specific factors, where we find the correlation which is not explained through common factors. For the empirical factor, the oil-price 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 the oil-price and the other currencies. For the Kendall's tau factor, oil and s&p500 are uncorrelated and uncorrelated from the rest.

As pointed out by a referee, the χ^2 approximation of Theorem 3.4 can be improved by various methods. One was suggested in Browne and Shapiro (1987) correcting for kurtosis in the data, another possibility is the approximation of the χ_2 distribution by a normal distribution using the Bartlett correction term. Since we have found reasonable interpretations of the factors, and the emphasis of this paper is on the presented robust method, we have refrained from further sophistications for model selection.

6 Proofs

Proof of Theorem 4.2. Define $\widehat{\boldsymbol{t}} := \mathrm{vec}[\widehat{\mathbf{T}}]$ and $\boldsymbol{t} := \mathrm{vec}[\mathbf{T}]$. Since $\widehat{\boldsymbol{t}}$ is a vector of U-statistics and, obviously, for $i \neq j$

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

Lee (1990, Section 1.3, 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

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

where $\boldsymbol{\Upsilon} = \left(\tau_{ij,kl} - \tau_{ij}\tau_{kl}\right)_{1 \le i \neq j, k \neq l \le d}$ and $\tau_{ij,kl}$ is given in (4.3). Note that the Jacobian matrix $\partial\left(\sin\left(\boldsymbol{t}\pi/2\right)\right)/\partial \boldsymbol{t}$ is diagonal with entries $\frac{\pi}{2}\cos(\boldsymbol{t}\pi/2)$. Hence, by the delta method (see Casella and Berger (2001, Section 5.5.4)),

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

and the proof is complete. □

Proof of Theorem 4.3. We first consider $\widehat{\tau}_{ij,kl}$ and rewrite it as a linear combination of U-statistics. Define for $1 \le a < b < c \le n$

$$\begin{aligned}
\mathbb{U}_2^{ij,kl}(\boldsymbol{x}_a, \boldsymbol{x}_b) &:= \operatorname{sign}\left[(x_{a,i} - x_{b,i})(x_{a,j} - x_{b,j})\right] \operatorname{sign}\left[(x_{a,k} - x_{b,k})(x_{a,l} - x_{b,l})\right] , \\
\mathbb{U}_{abc}^{ij,kl} &:= \mathbb{U}_{abc}^{ij,kl}(\boldsymbol{x}_a, \boldsymbol{x}_b, \boldsymbol{x}_c) \\
&:= \operatorname{sign}\left[(x_{a,i} - x_{b,i})(x_{a,j} - x_{b,j})\right] \operatorname{sign}\left[(x_{a,k} - x_{c,k})(x_{a,l} - x_{c,l})\right] , \\
\mathbb{U}_3^{ij,kl}(\boldsymbol{x}_a, \boldsymbol{x}_b, \boldsymbol{x}_c) &:= \frac{1}{6}\left(\mathbb{U}_{abc}^{ij,kl} + \mathbb{U}_{acb}^{ij,kl} + \mathbb{U}_{bac}^{ij,kl} + \mathbb{U}_{bca}^{ij,kl} + \mathbb{U}_{cab}^{ij,kl} + \mathbb{U}_{cba}^{ij,kl}\right) .
\end{aligned}$$

Hence, $\mathbb{U}_2^{ij,kl}$ and $\mathbb{U}_3^{ij,kl}$ are symmetric in their arguments. Next, define

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

and note that both are U-statistics. Obviously,

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

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. □

Acknowledgement

We would like to thank Daniel Berg for allowing us to use his programs for goodness-of-fit tests for copulae.

References

Anderson, T. (2003). *An Introduction to Multivariate Statistical Analysis.* (Third ed.). Hoboken, NJ: Wiley.

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

Berg, D. (2007). Copula goodness-of-fit testing: An overview and power comparison. Preprint. Norwegian Computing Centre. Available at *http://www.danielberg.no/dunder/research.php*.

Bollen, K. A. (1989). *Structural Equations With Latent Variables.* New York: Wiley.

Browne, M. W. (1982). Covariance structures. In D. M. Hawkins (Ed.), *Topics in Applied Multivariate Analysis*, pp. 72–141. Cambridge: Cambridge University Press.

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

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

Browne, M. W. 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. Multiv., Anal. 11*(3), 368–385.

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

Christensen, D. (2005). Fast algorithms for the calculation of Kendall's τ. *Computational Statistics 20*, 51–62.

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

Embrechts, P., A. McNeil, and D. Straumann (2002). Correlation and dependence in risk management: properties and pitfalls. In M. Dempster (Ed.), *Risk Management: Value at Risk and Beyond*, pp. 176–223. Cambridge: Cambridge University Press.

Everitt, B. S. (1984). *An Introduction to Latent Variable Models.* London: Chapman & Hall.

Fang, H. B., K. T. Fang, and S. Kotz (2002). The meta-elliptical distributions with given marginals. *J. Multiv, Anal. 82*, 1–16.

Fang, K., S. Kotz, and K. Ng (1990). *Symmetric Multivariate and Related Distributions.* London: Chapman & Hall.

Frey, R. and A. McNeil (2003). Dependence modelling, model risk and model calibration in models of portfolio credit risk. *J. of Risk 6*(1), 59–92.

Fuller, W. A. (1987). *Measurement Error Models*. New York: Wiley.

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. Statist. 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: Arnold.

Lawley, D. and A. Maxwell (1971). *Factor Analysis as a Statistical Method* (Second ed.). New York: Elsevier.

Lee, A. (1990). *U-Statistics*. New York: Marcel Dekker.

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*. Heidelberg: Physica.

Loehlin, J. C. (2001). *Latent Variable Models: An Introduction to Factor, Path, and Structural Analysis* (Third ed.). Mahwah, NJ: Lawrence Erlbaum.

Manzotti, F., F. Pérez, and A. Quiroz (2002). A statistic for testing the null hypothesis of elliptical symmetry. *J. Multiv. Anal. 81*, 274–285.

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*. Lecture Notes in Statistics No 139. New York: Springer.

Panchenko, V. (2005). Goodness-of-fit test for copulas. *Physica A 355*(1), 176–182.

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

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

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.

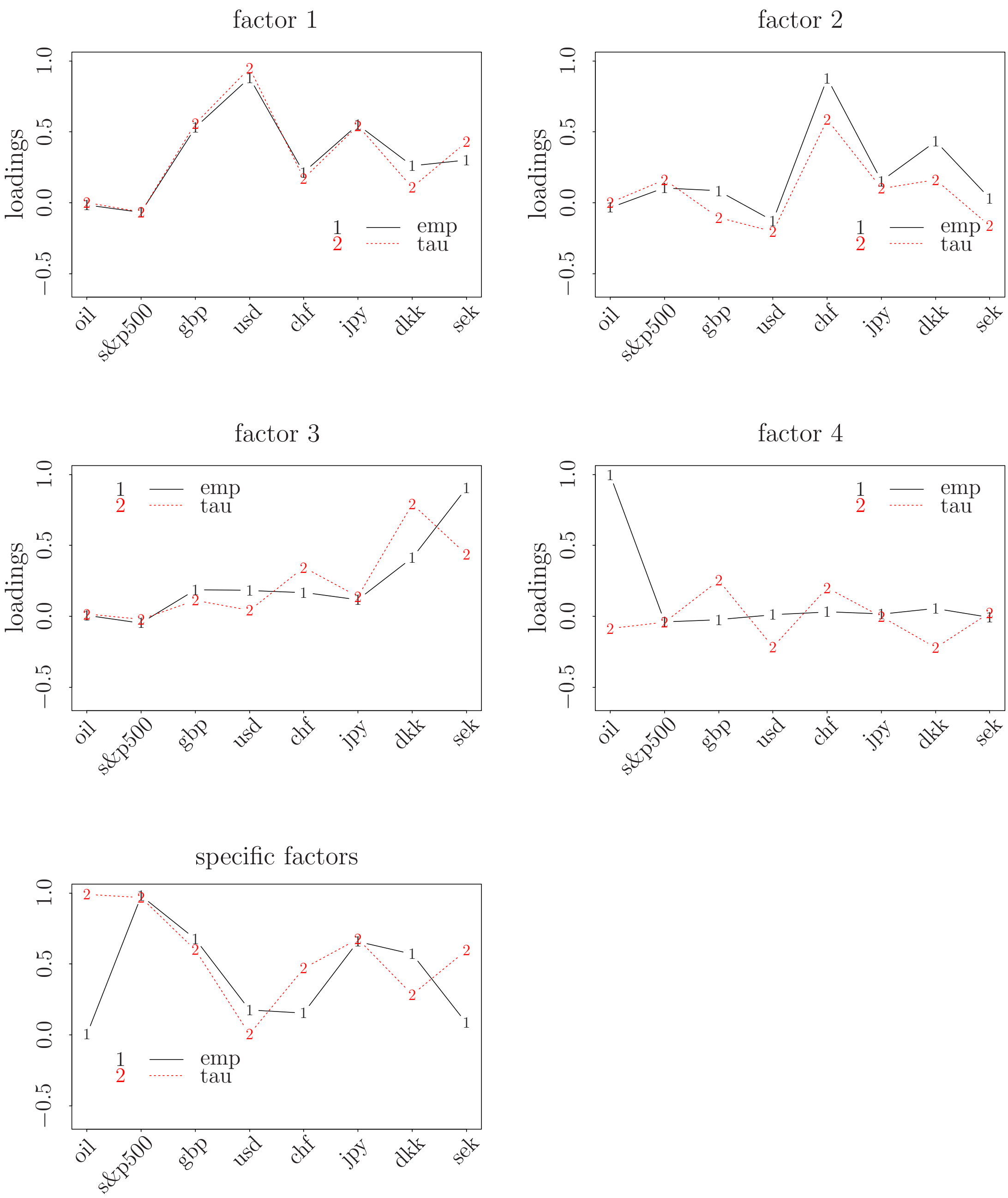

Figure 2: Oil-currency data: factor analysis based on 4 factors and different statistics, "emp" for the loadings $\widehat{\boldsymbol{\vartheta}}_\rho$, and "tau" for $\widehat{\boldsymbol{\vartheta}}_\tau$.
Upper row: loadings of factor 1 (left) and 2 (right).
Middle row: loadings of factor 3 (left) and 4 (right).
Lower row: specific factors diag($\mathbf{V}^2$).