Journal of Multivariate Analysis 138 (2015) 53–73

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Structured factor copula models: Theory, inference and computation

Pavel Krupskii *, Harry Joe

Department of Statistics, University of British Columbia, Vancouver BC, Canada V6T 1Z4

ARTICLE INFO

Article history:
Received 15 March 2014
Available online 13 November 2014

AMS 2000 subject classifications:
62H25
60H99

Keywords:
Bi-factor model
Conditional independence
Dependence measure
Factor analysis
Tail asymmetry
Tail dependence
Truncated vine

ABSTRACT

In factor copula models for multivariate data, dependence is explained via one or several common factors. These models are flexible in handling tail dependence and asymmetry with parsimonious dependence structures. We propose two structured factor copula models for the case where variables can be split into non-overlapping groups such that there is homogeneous dependence within each group. A typical example of such variables occurs for stock returns from different sectors. The structured models inherit most of dependence properties derived for common factor copula models. With appropriate numerical methods, efficient estimation of dependence parameters is possible for data sets with over 100 variables. We apply the structured factor copula models to analyze a financial data set, and compare with other copula models for tail inference. Using model-based interval estimates, we find that some commonly used risk measures may not be well discriminated by copula models, but tail-weighted dependence measures can discriminate copula models with different dependence and tail properties.

1. Introduction

Modeling high-dimensional data is a challenging task requiring flexible and tractable models. Models based on multivariate normality or Gaussianity are widely used in different applications due to their simplicity and tractability. In these models, special correlation structures are used to reduce the number of dependence parameters to a linear function of the dimension. A typical example is a Gaussian factor model where one or several common factors define the dependence structure for all of the variables. Factor copula models proposed in Krupskii and Joe [19] are extensions of the Gaussian factor model allowing greater flexibility when modeling non-Gaussian dependence. In particular, strong tail dependence and tail asymmetry can be accommodated. In data sets with a large number of variables, data can come from different sources or be clustered in different groups, for example, stock returns from different sectors or grouped item response data in psychometrics; thus dependence within each group and among different groups can be qualitatively different, and structured factor models can make use of the group information.

In psychometrics, sometimes a bi-factor correlation structure is used when variables or items can be split into non-overlapping groups; see for example Gibbons and Hedeker [10] and Holzinger and Swineford [12]. In a Gaussian bi-factor model, there is one common Gaussian factor which defines dependence between different groups, and one or several independent group-specific Gaussian factors which define dependence within each group. An alternative way to model dependence for grouped data is a nested model where the dependence in groups is modeled via dependent group-specific factors and the observed variables are assumed to be conditionally independent given these group-specific factors. The

* Corresponding author.
E-mail address: KrupskiiPV@yandex.ru (P. Krupskii).

http://dx.doi.org/10.1016/j.jmva.2014.11.002
0047-259X/

nested model is similar to Gaussian models with multilevel covariance structure; see Muthen [25]. Despite the simplicity, these two models have the same drawbacks as a common Gaussian factor model—they do not account for tail asymmetry and tail dependence.

In this paper, we propose copula extensions for bi-factor and nested Gaussian models. The extensions are called *structured factor copula models*. The proposed models contain 1- and 2-factor copula models introduced in Krupskii and Joe [19] as special cases, while allowing flexible dependence structure both for within group and between group dependence. As a result, the models can be suitable for modeling high-dimensional data sets consisting of several groups of variables with homogeneous dependence in each group.

The proposed multivariate copula models are built from a sequence of bivariate copulas in a similar way to vine copulas. Let $F_{\mathbf{X}}$ be the multivariate cumulative distribution function (cdf) of a random d-dimensional vector $\mathbf{X} = (X_1, \ldots, X_d)$, and let F_{X_j} be the cdf of X_j for $j = 1, \ldots, d$. The copula $C_{\mathbf{X}}$, corresponding to $F_{\mathbf{X}}$, is a multivariate uniform cdf such that $F_{\mathbf{X}}(x_1, \ldots, x_d) = C_{\mathbf{X}}(F_{X_1}(x_1), \ldots, F_{X_d}(x_d))$. By Sklar [30], $C_{\mathbf{X}}$ is unique if $F_{\mathbf{X}}$ is continuous. Copula functions allow for different types of dependence structure and are popular for modeling non-Gaussian dependence, including stock returns, insurance and hydrology data; see for example see Patton [29], McNeil et al. [23], Salvadori et al. [24] and others.

The proposed structured copula models are special cases of truncated-vine copula models with latent variables. In a vine model, bivariate linking copulas are applied to conditional cdfs to sequentially construct a multivariate distribution. The resulting vine model or pair-copula construction allows great flexibility in modeling different types of dependence structure by choosing appropriate linking copulas; see Kurowicka and Joe [20] and Brechmann and Czado [5] for more details. We show that depending on the choice of bivariate copulas in the structured copula models, different types of strength of dependence in the tails can be accommodated, similar to common factor copula models.

For a large number of variables which divide naturally into non-overlapping groups, it could be convenient to first separately model each group of variables followed by a method to combine the smaller models into a bigger model. Our structured copula models are one way to do this. The grouped t-copula of Demarta and McNeil [7] can handle groups but can only accommodate reflection symmetry. Another approach is hierarchical Kendall copulas in Brechmann [4]; it makes use of conditional independence given some group aggregation variables. Kendall functions only have simple form for exchangeable Archimedean copulas, so that hierarchical Kendall copulas are only convenient for exchangeable dependence within groups. Also nested Archimedean copulas (Section 4.1 of Joe [15]) are too parsimonious and have the property of exchangeable dependence within groups.

The details in this paper are given for continuous response variables, but the structured copula models can also be developed for discrete ordinal variables or mixed discrete/continuous variables. Factor copula models for item response are studied in Nikoloulopoulos and Joe [27], and if the items can be classified into non-overlapping groups, then the bi-factor or nested factor copula models are candidates when there is tail asymmetry or tail dependence.

The rest of the paper is organized as follows. In Section 2 we define bi-factor and nested copula models including a special case of Gaussian copulas, and compare the properties of these models with those of 1- and 2-factor copula models in Section 3. Section 4 has details on numerical maximum likelihood with a modified Newton–Raphson algorithm. Section 5 has a resampling method to obtain model-based interval estimates of the portfolio risk measures of Value-at-Risk and conditional tail expectation. In Section 6, we apply different copula–GARCH models to a financial data set and compare estimates of the Value-at-Risk, conditional tail expectations as well as some other tail-based quantities. The results show that structured factor copula models can parsimoniously estimate the dependence structure of the data. Value-at-Risk and other risk measures, which are widely used in financial applications, cannot efficiently differentiate models with different tail properties, and tail-weighted dependence measures are a better match to the fit of copula models based on the Akaike information criterion. Section 7 concludes with a discussion of future research.

2. Structured factor copula models

Common factor models assume that d observed variables are conditionally independent given $1 \leq p \ll d$ latent variables that affect each observed variables; for identifiability, the latent variables are assumed to be independent. Structured factor models assume that there is structure to the observed variables and each latent variable is linked to a subset of the observed variables. For Gaussian structured factor models, this corresponds to many structured zeros in the matrix of loadings; in this case, with fewer parameters in the loading matrix compared with the common factor model, and the p latent variables could be dependent, as in the oblique factor model of Harris and Kaiser [11] and McDonald [22]. With a large d, structured Gaussian factor models are also parsimonious models to parameterize the correlation matrix in $O(d)$ parameters (instead of $d(d-1)/2$ parameters). The main goal of this section is to present the copula version of two Gaussian structured factor models; for the extension, the parameters of the Gaussian structured factor models are converted to a set of correlations and partial correlations that are algebraically independent and that have a truncated vine structure, and then the correlations and partial correlations are replaced by bivariate copulas. Similar copula extensions exist for other structured factor models.

A specific case of structured factor models occurs when variables can be divided into non-overlapping groups. Assume that we have G groups of variables and there are d_g variables in the gth group, $g = 1, \ldots, G$. Let $U_{ij} \sim U(0, 1), i = 1, \ldots, d_g$, and suppose variables $U_{1g}, \ldots, U_{d_g g}$ belong to the gth group. Denote the joint cdf of $\mathbf{U} = (U_{11}, \ldots, U_{d_1 1}, \ldots, U_{1G}, \ldots, U_{d_G G})$ by $C_{\boldsymbol{U}}$. Let $d = \sum_{g=1}^{G} d_g$ be the total number of variables.

Table 1
Some copula notation; assuming $U, V, V_0, V_1 \sim U(0,1)$, V_0, V_1 are independent.

Notation	Definition
$C_{U,V}(u,v)$	copula cdf for (U,V)
$C_{U\|V}(u\|v) := \partial C_{U,V}(u,v)/\partial v$	conditional cdf of U given V
$c_{U,V}(u,v) := \partial^2 C_{U,V}(u,v)/\partial u \partial v$	copula pdf for (U,V)
$C_{U,V_1;V_0}(C_{U\|V_0}(u\|v_0), v_1)$	copula linking the conditional cdfs $C_{U\|V_0}(u\|v_0)$ and $C_{V_1\|V_0}(v_1\|v_0) = v_1$
$C_{U\|V_1;V_0}(C_{U\|V_0}(u\|v_0)\|v_1) := \partial C_{U,V_1;V_0}(C_{U\|V_0}(u\|v_0), v_1)/\partial v_1$	conditional copula cdf of $U\|V_0$ given V_1
$c_{U,V_1;V_0}(C_{U\|V_0}(u\|v_0), v_1) := \partial^2 C_{U,V_1;V_0}(C_{U\|V_0}(u\|v_0), v_1)/\partial v_1 \partial v_2$	copula pdf for $(U\|V_0, V_1)$

We consider two classes of structured factor copula models. The first model is an extension of the bi-factor model and we call it the bi-factor copula model. The second model is an extension of the oblique factor model with a blocked loading matrix where the dependent latent variables satisfy a 1-factor structure.

Factor copulas have appeared in a number of contexts with finance applications. There are factor copula models in McNeil et al. [23] and Hull and White [14] but these are not as general as those in Krupskii and Joe [19]. The bi-factor copula model can be considered as a special case of a p-factor copula in Krupskii and Joe [19] where appropriate linking copulas to latent variables are set to conditional independence based on the group structure. We briefly introduce one and two-factor copula models and some copula notation in Section 2.1 and then define bi-factor and nested copula models in Sections 2.2 and 2.3.

2.1. One and two-factor copula models

Assume we have d variables $U_1, \ldots, U_d \sim U(0,1)$ and let $C_{\mathbf{U}}$ be the joint cdf of the vector $\mathbf{U} = (U_1, \ldots, U_d)$. In the one-factor copula model the variables $U_1, \ldots, U_d$ are assumed to be conditionally independent given a latent factor $V_1 \sim U(0,1)$ and

$$C_{\mathbf{U}}(u_1, \ldots, u_d) = \int_0^1 \prod_{j=1}^d F_{U_j|V_1}(u_j|v_1)dv_1 = \int_0^1 \prod_{j=1}^d C_{U_j|V_1}(u_j|v_1)dv_1, \tag{1}$$

where $C_{U_j,V_1}(u_j, v_1)$ is the copula cdf linking U_j and V_1, and $C_{U_j|V_1}(u_j|v_1) := \partial C_{U_j,V_1}(u_j, v_1)/\partial v_1$ is a conditional cdf of C_{U_j,V_1}.

In the two-factor copula model the variables $U_1, \ldots, U_d$ are assumed to be conditionally independent given two independent latent factors $V_1, V_2 \sim U(0,1)$ and

$$C_{\mathbf{U}}(u_1, \ldots, u_d) = \int_0^1 \int_0^1 \prod_{j=1}^d C_{U_j|V_2;V_1}(C_{U_j|V_1}(u_j|v_1)|v_2)dv_1dv_2, \tag{2}$$

where $C_{U_j,V_2;V_1}(C_{U_j|V_1}(u_j|v_1), v_2)$ is the copula cdf linking the conditional distributions $C_{U_j|V_1}(\cdot|v_1)$ and $C_{V_2|V_1}(\cdot|v_1)$, and $C_{U_j|V_2;V_1}(C_{U_j|V_1}(u_j|v_1)|v_2) := \partial C_{U_j,V_2;V_1}(C_{U_j|V_1}(u_j|v_1), v_2)/\partial v_2$. In the next sections we will be using similar copula notation for bivariate copula cdfs, conditional cdfs and pdfs; these notation is summarized in Table 1.

It is seen from (1) and (2), that the joint cdf in the models can be expressed in terms of bivariate linking copula cdf and conditional cdfs of copulas. Tail properties and dependence properties of these linking copulas reflect the properties of the joint distribution $C_{\mathbf{U}}$; see Krupskii and Joe [19] for details.

2.2. Bi-factor copula model

Consider a model with one common (global) factor and G group-specific factors, such that the $G+1$ factors or latent variables are mutually independent. Assume that within the gth group, $U(0,1)$ distributed random variables $U_{1g}, \ldots, U_{d_g g}$ are conditionally independent given V_0 and V_g, where $V_0, V_1, \ldots, V_G$ are independent and identically distributed (i.i.d.) $U(0,1)$ random variables. We also assume that U_{ig} in group g does not depend on $V_{g'}$ for $g' \neq g$. The bi-factor copula model is therefore an extension of the two-factor copula model where $G = 1$.

Let C_{U_{ig},V_0} be the copula cdf of (U_{ig}, V_0) and let $C_{U_{ig}|V_0}$ the corresponding conditional distribution. Let $C_{U_{ig},V_g;V_0}$ be the copula for the conditional univariate distributions $C_{U_{ig}|V_0}(\cdot|v_0)$ and $C_{V_g|V_0}(\cdot|v_0)$, with $C_{V_g|V_0}$ being the $U(0,1)$ cdf from the independence of $V_0, V_1, \ldots, V_g$. By Sklar's theorem, in general, the copula $C_{U_{ig},V_g;V_0}$ depends on v_0. However, similar to Krupskii and Joe [19] and the vine copula literature [20], we make the simplifying (modeling) assumption that $C_{U_{ig},V_g;V_0}$ does not depend on v_0. This is not a strong assumption as we are dealing with latent variables, and in a special case of the Gaussian model, copulas for conditional distributions do not depend on the values of the conditioning variables.

The bivariate copulas C_{U_{ig},V_0} (for common factor) and $C_{U_{ig},V_g;V_0}$ (for group-specific factor) are assumed to be absolutely continuous with respective copula densities c_{U_{ig},V_0} and $c_{U_{ig},V_g;V_0}$.

With a vector $\mathbf{u} = (u_{11}, \ldots, u_{d_1 1}, \ldots, u_{1G}, \ldots, u_{d_G G})$, then using the above conditional independence assumptions, we get:

$$\Pr(U_{ig} \le u_{ig}, i = 1, \ldots, d_g, g = 1, \ldots, G) = \int_{[0,1]^{G+1}} \prod_{g=1}^{G} \prod_{i=1}^{d_g} \Pr(U_{ig} \le u_{ig} | V_0 = v_0, V_g = v_g) dv_1 \cdots dv_G dv_0$$

and

$$\begin{aligned} \Pr(U_{ig} \le u_{ig} | V_0 = v_0, V_g = v_g) &= \frac{\partial}{\partial v_g} \Pr(U_{ig} \le u_{ig}, V_g \le v_g | V_0 = v_0) \\ &= \frac{\partial}{\partial v_g} C_{U_{ig}, V_g; V_0}(C_{U_{ig}|V_0}(u_{ig}|v_0), v_g) =: C_{U_{ig}|V_g; V_0}(C_{U_{ig}|V_0}(u_{ig}|v_0)|v_g). \end{aligned}$$

Hence

$$\begin{aligned} C_{\boldsymbol{U}}(\mathbf{u}) &= \int_0^1 \int_{[0,1]^G} \prod_{g=1}^{G} \prod_{i=1}^{d_g} C_{U_{ig}|V_g; V_0}(C_{U_{ig}|V_0}(u_{ig}|v_0)|v_g) dv_1 \cdots dv_G\, dv_0 \\ &= \int_0^1 \prod_{g=1}^{G} \left\{ \int_0^1 \left[\prod_{i=1}^{d_g} C_{U_{ig}|V_g; V_0}(C_{U_{ig}|V_0}(u_{ig}|v_0)|v_g) \right] dv_g \right\} dv_0; \end{aligned}$$

$$c_{\boldsymbol{U}}(\mathbf{u}) = \int_0^1 \prod_{g=1}^{G} \left\{ \left[\prod_{i=1}^{d_g} c_{U_{ig}, V_0}(u_{ig}, v_0) \right] \int_0^1 \left[\prod_{i=1}^{d_g} c_{U_{ig}, V_g; V_0}(C_{U_{ig}|V_0}(u_{ig}|v_0), v_g) \right] dv_g \right\} dv_0. \quad (3)$$

It is seen that the joint density is represented as a one-dimensional integral of a function which in turn is a product of G one-dimensional integrals. As a result, $(G+1)$-dimensional numerical integration can be avoided. The model has $d = d_1 + \cdots + d_G$ bivariate linking copulas both for the first and second factors, or $2d$ linking copulas in total. The marginal distribution of (3) for a single group g of variables is a 2-factor copula model.

For the parametric version of this model, there is a parameter $\theta_{i,g}$ for C_{U_{ig}, V_0} and a parameter $\gamma_{i,g}$ for $C_{U_{ig}|V_g; V_0}$; $\theta_{i,g}$ and $\gamma_{i,g}$ could be vectors. The parameter vector for the density in (3) is $\boldsymbol{\theta} = (\theta_{i,g}, \gamma_{i,g} : i = 1, \ldots, d_g, g = 1, \ldots, G)$.

2.3. Nested copula model

Consider the case of G dependent factors without a common factor. Assume that for a fixed $g = 1, \ldots, G$, $U(0, 1)$ distributed random variables $U_{1g}, \ldots, U_{d_g g}$ are conditionally independent given $V_g \sim U(0, 1)$, and the joint cdf of $\boldsymbol{V} = (V_1, \ldots, V_G)$ is given by the copula $C_{\boldsymbol{V}}$. We also assume that U_{ig} in group g does not depend on $V_{g'}$ for $g' \neq g$. That is, we have G groups of variables and G latent factors where the gth latent factor defines dependence structure in the gth group. The nested copula model is therefore an extension of the one-factor copula model where $G = 1$.

Let C_{U_{ig}, V_g} be the copula cdf of (U_{ig}, V_g) and $C_{U_{ig}|V_g}$ be the corresponding conditional distribution. With a vector $\mathbf{u} = (u_{11}, \ldots, u_{d_1 1}, \ldots, u_{1G}, \ldots, u_{d_G G})$ we get:

$$C_{\boldsymbol{U}}(\mathbf{u}) = \int_{[0,1]^G} \left\{ \prod_{g=1}^{G} \prod_{i=1}^{d_g} C_{U_{ig}|V_g}(u_{ig}|v_g) \right\} c_{\boldsymbol{V}}(v_1, \ldots, v_G) dv_1 \cdots dv_G, \quad (4)$$

where C_{U_{ig}, V_g} is the absolutely continuous copula linking U_{ig} and V_g. This is a copula version of the oblique Gaussian factor model where each observed variable loads on exactly one latent variable.

We additionally assume that $V_1, \ldots, V_G$ are conditionally independent given another latent variable V_0, that is the joint distribution of $\boldsymbol{V}$ has one-factor copula structure. Then we get:

$$c_{\boldsymbol{V}}(v_1, \ldots, v_G) = \int_0^1 \left\{ \prod_{g=1}^{G} c_{V_g, V_0}(v_g, v_0) \right\} dv_0$$

where C_{V_g, V_0} is the absolutely continuous copula linking V_g and V_0. It implies that

$$\begin{aligned} C_{\boldsymbol{U}}(\mathbf{u}) &= \int_0^1 \int_{[0,1]^G} \left\{ \prod_{g=1}^{G} \prod_{i=1}^{d_g} C_{U_{ig}|V_g}(u_{ig}|v_g) \right\} \left\{ \prod_{g=1}^{G} c_{V_g, V_0}(v_g, v_0) \right\} dv_1 \cdots dv_G\, dv_0 \\ &= \int_0^1 \left\{ \prod_{g=1}^{G} \int_0^1 \left[c_{V_g, V_0}(v_g, v_0) \prod_{i=1}^{d_g} C_{U_{ig}|V_g}(u_{ig}|v_g) \right] dv_g \right\} dv_0; \end{aligned}$$

$$c_{\boldsymbol{U}}(\mathbf{u}) = \int_0^1 \left\{ \prod_{g=1}^{G} \int_0^1 \left[c_{V_g, V_0}(v_g, v_0) \prod_{i=1}^{d_g} c_{U_{ig}, V_g}(u_{ig}, v_g) \right] dv_g \right\} dv_0. \quad (5)$$

The total number of bivariate linking copulas in the model (5) is $d_1 + \cdots + d_G + G = d + G$ (d copulas C_{U_{ig},V_g} and G copulas C_{V_g,V_0}). The marginal distribution of (5) for a single group g of variables is a 1-factor copula model.

In this setting, there exist a common factor V_0, say the current state of economy, which drives some other factors $V_1, \ldots, V_G$, say some unobservable parameters reflecting the situation in different stock sectors. Each factor, in turn, defines a dependence structure of a group of variables (such as stocks in a common sector).

For the parametric version of this model, there is a parameter θ_g for C_{V_g,V_0} and a parameter $\gamma_{i,g}$ for C_{U_{ig},V_g}. The parameter vector for (5) is $\boldsymbol{\theta} = (\theta_g, \eta_{i,g} : i = 1, \ldots, d_g, g = 1, \ldots, G)$.

2.4. Special case of Gaussian copulas

In this subsection, we consider the Gaussian bi-factor model and show that if all the bivariate linking copulas are Gaussian, then the nested factor model is a special case of the bi-factor model.

Let Φ, ϕ be the standard normal cdf and density respectively, and let Φ_2 be the bivariate normal cdf. If (Z_1, Z_2) is bivariate normal with zero means, unit variances and correlation ρ, then $[Z_2|Z_1 = z_1] \sim N(\rho z_1, 1 - \rho^2)$ so that from the bivariate Gaussian copula $\Phi_2(\Phi^{-1}(u), \Phi^{-1}(v); \rho)$, the conditional cdf in Table 1 is $\Phi([\Phi^{-1}(u) - \rho\Phi^{-1}(v)]/(1 - \rho^2)^{1/2})$. Suppose C_{U_{ig},V_0} and $C_{U_{ig},V_g;V_0}$ are bivariate Gaussian copulas with parameters φ_{ig} and $\gamma_{ig} = \eta_{ig}/(1-\varphi_{ig}^2)^{1/2}$ respectively, $g = 1, \ldots, G$. Here η_{ig} is a correlation of $Z_{ig} = \Phi^{-1}(U_{ig})$ and $W_g = \Phi^{-1}(V_g)$ so that the independence of V_0, V_g implies that γ_{ig} is the partial correlation of Z_{ig} and W_g given $W_0 = \Phi^{-1}(V_0)$ (in general $\rho_{ZW_2;W_1} = [\rho_{ZW_2} - \rho_{ZW_1}\rho_{W_2W_1}]/[(1 - \rho_{ZW_1}^2)(1 - \rho_{W_2W_1}^2)]^{1/2}$). Hence,

$$\begin{aligned} C_{U_{ig}|V_g;V_0}(C_{U_{ig}|V_0}(u|v_0)|v_g) &= \Phi\left(\left[\frac{\Phi^{-1}(u) - \varphi_{ig}\Phi^{-1}(v_0)}{(1-\varphi_{ig}^2)^{1/2}} - \gamma_{ig}\Phi^{-1}(v_g)\right] \Big/ \sqrt{1-\gamma_{ig}^2}\right) \\ &= \Phi\left(\frac{\Phi^{-1}(u) - \varphi_{ig}\Phi^{-1}(v_0) - \gamma_{ig}(1-\varphi_{ij}^2)^{1/2}\Phi^{-1}(v_g)}{\sqrt{(1-\varphi_{ig}^2)(1-\gamma_{ig}^2)}}\right). \end{aligned}$$

With $z_{ig} = \Phi^{-1}(u_{ig})$, $i = 1, \ldots, d_g$, $g = 1, \ldots, G$, the cdf for the bi-factor model becomes

$$\begin{aligned} &F(z_{11}, \ldots, z_{d_1 1}, \ldots, z_{1G}, \ldots, z_{d_G G}) \\ &\quad := C(\Phi(z_{11}), \ldots, \Phi(z_{d_1 1}), \ldots, \Phi(z_{1G}), \ldots, \Phi(z_{d_G G})) \\ &\quad = \int_{-\infty}^{\infty} \prod_{g=1}^{G} \left\{ \int_{-\infty}^{\infty} \prod_{i=1}^{d_g} \Phi\left(\frac{z_{ig} - \varphi_{ig}w_0 - \gamma_{ig}(1-\varphi_{ig}^2)^{1/2}w_g}{\sqrt{(1-\varphi_{ig}^2)(1-\gamma_{ig}^2)}}\right) \cdot \phi(w_g)dw_g \right\} \cdot \phi(w_0)dw_0. \end{aligned}$$

Hence this model is the same as a multivariate Gaussian model with a bi-factor correlation structure because this multivariate cdf comes from the representation:

$$Z_{ig} = \varphi_{ig}W_0 + \eta_{ig}W_g + \sqrt{1 - \varphi_{ig}^2 - \eta_{ig}^2}\,\epsilon_{ig},$$

where $W_0, W_g, \epsilon_{ig}, g = 1, \ldots, G, i = 1, \ldots, d_g$ are i.i.d. N(0, 1) random variables. From the linear representation, one can write the joint cdf $F(\mathbf{z}) = F(z_{11}, \ldots, z_{d_1 1}, \ldots, z_{1G}, \ldots, z_{d_G G})$ as follows:

$$F(\mathbf{z}) = \int_{-\infty}^{\infty} F_{\mathbf{Z}|W_0}(z_{11}, \ldots, z_{d_1 1}, \ldots, z_{1G}, \ldots, z_{d_G G}|w_0)\phi(w_0)dw_0 = \int_{-\infty}^{\infty} \prod_{g=1}^{G} \Phi_{d_g|W_0}(z_{1g}, \ldots, z_{d_g g}|w_0)\phi(w_0)dw_0,$$

where $\Phi_{d_g|W_0}$ is the conditional cdf of $(Z_{1g}, \ldots, Z_{d_g g})$ given W_0, $g = 1, \ldots, G$. Here we use the fact the variables from different groups are conditionally independent given W_0. In the gth group variables are conditionally independent given W_0 and W_g. With $\eta_{ig} = \gamma_{ig}(1 - \varphi_{ig}^2)^{1/2}$ and $1 - \varphi_{ig}^2 - \eta_{ig}^2 = (1 - \varphi_{ig}^2)(1 - \gamma_{ig}^2)$, it implies

$$\begin{aligned} \Phi_{d_g|W_0}(z_{1g}, \ldots, z_{d_g g}|w_0) &= \int_{-\infty}^{\infty} \prod_{i=1}^{d_g} \Pr\{Z_{ig} \le z_{ig}|w_0, w_g\}\phi(w_g)dw_g \\ &= \int_0^1 \prod_{i=1}^{d_g} \Phi\left(\frac{z_{ig} - \varphi_{ig}w_0 - \gamma_{ig}(1-\varphi_{ig}^2)^{1/2}w_g}{\sqrt{(1-\varphi_{ig}^2)(1-\gamma_{ig}^2)}}\right)\phi(v_g)dw_g. \end{aligned}$$

It implies that $\mathbf{Z}$ has a multivariate Gaussian distribution and

$$\begin{aligned} &\mathrm{Cor}(Z_{i_1 g}, Z_{i_2 g}) = \varphi_{i_1 g}\,\varphi_{i_2 g} + \eta_{i_1 g}\,\eta_{i_2 g}, \quad i_1 \ne i_2, \\ &\mathrm{Cor}(Z_{i_1 g_1}, Z_{i_2 g_2}) = \varphi_{i_1 g_1}\,\varphi_{i_2 g_2}, \quad g_1 \ne g_2. \end{aligned}$$

The number of parameters in the Gaussian bi-factor structure is $2d - N_1 - N_2$, where N_1 is the number of groups of size 1 and N_2 is the number of groups of size 2. For a group g of size 1 with variable j, W_g is absorbed with ϵ_{ig} because η_{ig} would not be identifiable. For a group g of size 2 with variable indices i_1, i_2, the parameters η_{i_1g} and η_{i_2g} appear only in the correlation for variables i_1, i_2 and this correlation is $\varphi_{i_1g}\varphi_{i_2g} + \eta_{i_1g}\eta_{i_2g}$. Since only the product $\eta_{i_1g}\eta_{i_2g}$ appears, one of η_{i_1g}, η_{i_2g} can be taken as 1 without loss of generality. For the bi-factor copula with non-Gaussian linking copulas, near non-identifiability can occur when there are groups of size 2; in this case, one of the linking copulas to the group latent variable can be fixed (say at comonotonicity) for a group of size 2.

A special case of the bi-factor copula model with Gaussian copulas can be defined as follows. Assume that

$$\xi_g = \varphi_g W_0 + \sqrt{1-\varphi_g^2}\, W_g, \qquad Z_{ig} = \varphi_{ig}^*\xi_g + \sqrt{1-(\varphi_{ig}^*)^2}\,\epsilon_{ig},$$

where $W_0, W_g, \epsilon_{ig}, g = 1, \ldots, G, i = 1, \ldots, d_g$, are i.i.d. N(0, 1) random variables. The $d \times (1 + G)$ loading matrix of $\mathbf{Z} = (Z_{11}, \ldots, Z_{d_11}, \ldots, Z_{1G}, \ldots, Z_{d_GG})^T$ on $W_0, W_1, \ldots, W_G$:

$$\boldsymbol{A} = \begin{pmatrix} \varphi_{11}^*\varphi_1 & \varphi_{11}^*\sqrt{1-\varphi_1^2} & \mathbf{0} & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \varphi_{d_11}^*\varphi_1 & \varphi_{d_11}^*\sqrt{1-\varphi_1^2} & \mathbf{0} & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_{1G}^*\varphi_G & 0 & \mathbf{0} & \varphi_{1G}^*\sqrt{1-\varphi_G^2} \\ \vdots & \vdots & \vdots & \vdots \\ \varphi_{d_GG}^*\varphi_G & 0 & \mathbf{0} & \varphi_{d_GG}^*\sqrt{1-\varphi_G^2} \end{pmatrix}.$$

Then $\mathbf{Z}$ has a multivariate Gaussian distribution and

$$\mathrm{Cor}(Z_{i_1g}, Z_{i_2g}) = \varphi_{i_1g}^*\,\varphi_{i_2g}^* = \varphi_{i_1g}^*\,\varphi_{i_2g}^*\,\varphi_g^2 + \varphi_{i_1g}^*\,\varphi_{i_2g}^*(1-\varphi_g^2), \quad i_1 \neq i_2,$$
$$\mathrm{Cor}(Z_{i_1g_1}, Z_{i_2g_2}) = \varphi_{i_1g_1}^*\,\varphi_{i_2g_2}^*\,\varphi_{g_1}\,\varphi_{g_2}, \quad g_1 \neq g_2.$$

Note that the nested Gaussian model is a special case of a bi-factor model with one common factor for all groups. It is seen, that if $\varphi_{ig} = \varphi_{ig}^*\varphi_g$ and $\eta_{ig} = \varphi_{ig}^*\sqrt{1-\varphi_g^2}$, then we get the bi-factor model. Nevertheless, in general, the nested copula model is not a special case of a bi-factor copula model.

The number of parameters in the Gaussian nested-factor structure is $d + G$ for $G \geq 3$ groups and $d + 1$ for two groups, because in the case of two groups, $\varphi_1\varphi_2$ occur only as a product in the correlation of variables in different groups.

3. Tail and dependence properties of the structured factor copula model

In this section, we summarize some results on positive dependence and tail order of bivariate margins of the structured factor copulas in Section 2. The lower and upper tail order from Hua and Joe [13] can be used to summarize the strength of dependence in the joint lower and upper tail respectively, and the difference of the two tail orders can indicate the direction of tail asymmetry. The lower tail order of a m-variate copula $C_{1:m}$ is κ_L if $C_{1:m}(u\mathbf{1}_m) \sim \ell_L(u)u^{\kappa_L}$ as $u \to 0$ where $\ell_L(u)$ is a slowly varying function (such as a constant or a power of $-\log u$). Similarly the upper tail order κ_U is such that $\widehat{C}_{1:m}(u\mathbf{1}_m) \sim \ell_U(u)u^{\kappa_U}$ as $u \to 0$, where $\widehat{C}_{1:m}$ is the copula of $(1-U_1, \ldots, 1-U_m)$ when $(U_1, \ldots, U_m) \sim C_{1:m}$. The tail orders have the property that $\kappa_L \geq 1$ and $\kappa_U \geq 1$, with a smaller value corresponding to more dependence in the tail. The strongest tail dependence corresponds to $\kappa_L = 1$ or $\kappa_U = 1$ (the usual tail dependence). Tail orthant independence corresponds to $\kappa_L = m$ or $\kappa_U = m$ with the slowly varying function being a constant, and this is called tail quadrant independence for $m = 2$. Intermediate tail dependence corresponds to $1 < \kappa_L < m$ or $1 < \kappa_U < m$.

From now on we will assume that all bivariate linking copulas are twice continuously differentiable functions on $(0, 1)^2$. Note that, for the model (3), within each group, variables are independent given the group-specific factor and the common factor V_0. In other words, the dependence structure is a two-factor copula model. Two variables from different groups are independent given the common factor V_0 and so the dependence is the same as a one-factor copula model. For the model (5), properties derived for bivariate margins depend on the choice of copula $C_{\boldsymbol{V}}$. If we choose two variables from the same group, we get the same marginal distribution as in 1-factor copula model. However, the case when the variables are selected from different groups requires special attention. Without loss of generality, consider the pair (U_{11}, U_{12}) with U_{11} from group 1 and U_{12} from group 2. Let C_{1V_1} and C_{2V_2} be shorthand for $C_{U_{11}V_1}$ and $C_{U_{12}V_2}$ respectively. Denote the cdf of (U_{11}, U_{12}) by C_{12} and the cdf of (V_1, V_2) by C_{V_1,V_2}. Let κ_L be the lower tail order of C_{12}. It follows from (4), with two groups of size 1, that

$$C_{12}(u_1, u_2) = \int_0^1\int_0^1 C_{1|V_1}(u_1|v_1)C_{2|V_2}(u_2|v_2)c_{V_1,V_2}(v_1, v_2)dv_1dv_2. \tag{6}$$

The conclusions about the tail order depend on some positive dependence conditions for pairs (V_1, V_2), (U_{11}, V_1), (U_{12}, V_2). A bivariate distribution F_{12} with univariate margins F_1, F_2 is *positive quadrant dependent (PQD)* if $F_{12} \geq F_1F_2$ pointwise. A bivariate distribution F_{12} with conditional distribution $F_{1|2}(\cdot|x_2)$ has the first variable stochastically increasing (decreasing) in the second variable if $1 - F_{1|2}(x_1|x_2)$ is increasing (decreasing, respectively) in x_2 for all x_1.

The first result is useful to show positive dependence for the nested copula model.

Proposition 1. *Let both $C_{1|V_1}$ and $C_{2|V_2}$ be stochastically increasing or stochastically decreasing conditional cdfs, and C_{V_1,V_2} is a copula with positive quadrant dependence. Then C_{12} in (6) is a PQD copula.*

Proof. Using the integration by parts formula (three times), we get:

$$
\begin{aligned}
C_{12}(u_1, u_2) &= \int_0^1 C_{1|V_1}(u_1|v_1)\left\{C_{2|V_2}(u_2|1) - \int_0^1 \frac{\partial C_{2|V_2}(u_2|v_2)}{\partial v_2} C_{V_2|V_1}(v_2|v_1)dv_2\right\} dv_1 \\
&= u_1C_{2|V_2}(u_2|1) - \int_0^1 \frac{\partial C_{2|V_2}(u_2|v_2)}{\partial v_2}\left\{v_2C_{1|V_1}(u_1|1) - \int_0^1 \frac{\partial C_{1|V_1}(u_1|v_1)}{\partial v_1} C_{V_1,V_2}(v_1, v_2)dv_1\right\} dv_2 \\
&= u_1C_{2|V_2}(u_2|1) + u_2C_{1|V_1}(u_1|1) - C_{1|V_1}(u_1|1)C_{2|V_2}(u_2|1) + I_{12}
\end{aligned}
$$

where

$$
I_{12} := \int_0^1\int_0^1 \frac{\partial C_{1|V_1}(u_1|v_1)}{\partial v_1} \cdot \frac{\partial C_{2|V_2}(u_2|v_2)}{\partial v_2} \cdot C_{V_1,V_2}(v_1, v_2)dv_1dv_2.
$$

Using the PQD assumption for $C_{V_1V_2}$, and the stochastic monotonicity assumption for $\frac{\partial C_{1|V_1}(u_1|v_1)}{\partial v_1}$ and $\frac{\partial C_{2|V_2}(u_2|v_2)}{\partial v_2}$,

$$
I_{12} \geq \int_0^1 v_1\frac{\partial C_{1|V_1}(u_1|v_1)}{\partial v_1}dv_1 \int_0^1 v_2\frac{\partial C_{2|V_2}(u_2|v_2)}{\partial v_2}dv_2 = [C_{1|V_1}(u_1|1) - u_1][C_{2|V_2}(u_2|1) - u_2].
$$

Therefore,

$$
\begin{aligned}
C_{12}(u_1, u_2) \geq\ & u_1C_{2|V_2}(u_2|1) + u_2C_{1|V_1}(u_1|1) - C_{1|V_1}(u_1|1)C_{2|V_2}(u_2|1) \\
&+ [C_{1|V_1}(u_1|1) - u_1][C_{2|V_2}(u_2|1) - u_2] = u_1u_2. \quad \square
\end{aligned}
$$

The stochastic monotonicity is not a very restrictive assumption as many parametric bivariate copula families used in applications have stochastically increasing or decreasing conditional cdfs. Typical examples of copulas with stochastically increasing conditional cdfs include the normal copula with a positive correlation parameter $\rho > 0$ and Frank copula with a positive dependence parameter $\theta > 0$, the Gumbel and BB1 copulas; see Joe [15] for more details on these copulas. The Gaussian copula with $\rho < 0$ and Frank copula with $\theta < 0$ have stochastically decreasing conditional cdfs.

We next indicate how the above result is used for the nested copula model. Suppose in (5) that all of the bivariate linking copulas $C_{V_g|V_0}$ and $C_{U_{ig}|V_g}$ satisfy the positive dependence condition of stochastically increasing. By Proposition 1 of Krupskii and Joe [19], $C_{V_{g_1},V_{g_2}}$ is PQD for any $g_1 \neq g_2$ because of the 1-factor copula structure for $V_0, V_1, \ldots, V_G$. From the above proposition of $U_{ig_1}, U_{i'g_2}$ are random variables in two different groups, then they are PQD.

If C_{12} is PQD and $c_{V_1,V_2}(v_1, v_2) \leq K$ for $v_1, v_2 \in [0, 1]$ for $K > 0$, then $C_{12}(u_1, u_2) \leq Ku_1u_2$ and $C_{12}(u, u)/u^2 \leq K$ which implies $\kappa_L = 2$; similarly, the upper tail order equals two in this case. Hence, positive dependence of U_{11}, U_{12}, V_1, V_2 and a bounded density for c_{V_1,V_2} means that (U_{11}, U_{12}) has tail quadrant independence. The next results apply with stronger dependence in the tails.

Proposition 2. *Let $\lim_{u\to 0} C_{j|V_j}(u|hu) = t_j(h)$ and assume that $C_{j|V_j}(u|v)$ is a continuous function of u and v on $(0, 1)^2, j = 1, 2$. Assume $\lim_{h\to 0} t_j(h) = t_{j0} > 0, j = 1, 2$. In addition, assume that the density $c_{V_1,V_2}(v_1, v_2)$ is a continuous function of v_1 and v_2, and that $c_{V_1,V_2}(w_1u, w_2u) \geq k(w_1, w_2)/u^\alpha$ for small enough $u > 0$, where $\alpha \in [0, 1]$ and k is a positive continuous function of w_1, w_2. Then the tail order κ_L of C_{12} in (6) is at most $2 - \alpha$. A similar result holds for the upper tail order.*

Proof. It follows from (6) that:

$$
C_{12}(u, u) = u^2\int_0^{1/u}\int_0^{1/u} C_{1|V_1}(u|h_1u)C_{2|V_2}(u|h_2u)c_{V_1,V_2}(h_1u, h_2u)dh_1dh_2.
$$

For any $\epsilon > 0$ we can find $h_j(\epsilon) > 0$ such that $|t_j(h_j) - t_{j0}| < \epsilon$ for $h_j \leq h_j(\epsilon), j = 1, 2$. Denote $h^*(\epsilon) = \min\{h_1(\epsilon), h_2(\epsilon)\}$. By the assumption, there exists $u(\epsilon) > 0$ such that $|C_{j|V_j}(u|h^*(\epsilon)u) - t_j(h^*(\epsilon))| < \epsilon$ for $0 < u \leq u(\epsilon)$. It implies that $|C_{j|V_j}(u|h^*(\epsilon)u) - t_{j0}| < 2\epsilon$ for $u \leq u(\epsilon)$. Due to uniform continuity of $C_{j|V_j}$ on $[0, 1] \times [0, 1]$ we can find $u_{0j} > 0, h_j^-, h_j^+$ such that $h_j^- < h^*(\epsilon) < h_j^+$ and $|C_{j|V_j}(u|h^*(\epsilon)u) - C_{j|V_j}(u|h_ju)| < \epsilon$ for $0 < u \leq u_{0j}$ and $h_j^- \leq h_j \leq h_j^+$. Therefore $|C_{j|V_j}(u|h_ju) - t_{j0}| < 3\epsilon$ for $u < \tilde{u} = \min\{u_{01}, u_{02}, u^*(\epsilon)\}$ and $\tilde{h} = \max\{h_1^-, h_2^-\} < h_j < h^*(\epsilon)$. Let $\epsilon = \min\{t_{10}, t_{20}\}/6$.

Due to the continuity of $k(w_1, w_2)$, there are constants $K_V > 0$ and $\tilde{h}^* > \tilde{h}$ such that $c_{V_1,V_2}(h_1u, h_2u) \geq K_V/u^\alpha$ for $\tilde{h} < h_j < \tilde{h}^*, j = 1, 2$. Then

$$\begin{aligned} C_{12}(u, u) &\geq u^2 \int_{\tilde{h}}^{\tilde{h}^*} \int_{\tilde{h}}^{\tilde{h}^*} C_{1|V_1}(u|h_1u)C_{2|V_2}(u|h_2u)c_{V_1,V_2}(h_1u, h_2u)dh_1dh_2 \\ &\geq u^2 \int_{\tilde{h}}^{\tilde{h}^*} \int_{\tilde{h}}^{\tilde{h}^*} \frac{t_{10}}{6}\frac{t_{02}}{6}\frac{K_V}{u^\alpha}\, dh_1dh_2 \geq u^{2-\alpha} \cdot \frac{K_V(\tilde{h}^* - \tilde{h})^2 t_{10}t_{20}}{36} \end{aligned}$$

and hence the lower tail order of C_{12} is less or equal than $2 - \alpha$. □

Remark 1. The condition on the limit $\lim_{u\to 0} C_{j|V_j}(u|hu)$ implies that C_{j,V_j} is a copula with the lower tail dependence, such as the Student, reflected Gumbel or BB1 copula.

Remark 2. Suppose the lower tail order of C_{V_1,V_2} is κ with a slowly varying function $\ell(u)$, and there is a tail order function $b_\kappa(w_1, w_2)$ such that $C_{V_1,V_2}(w_1u, w_2u) \sim u^\kappa \ell(u)b_\kappa(w_1, w_2)$ as $u \to 0$. Hua and Joe [13] showed, under the condition of continuity and ultimate monotonicity in the lower tail, that this implies $c_{V_1,V_2}(w_1u, w_2u) \sim u^{\kappa-2}\ell(u) \cdot \partial^2 b_\kappa(w_1, w_2)/\partial w_1\partial w_2$ as $u \to 0$. Hence the assumption on c_{V_1,V_2} in the above proposition is essentially that the tail order of C_{V_1,V_2} is at most $2-\alpha$. In other words, C_{V_1,V_2} is a copula with intermediate tail dependence if $0 < \alpha < 1$.

The condition on c_{V_1,V_2} with $\alpha = 1$ implies C_{V_1,V_2} is a lower tail dependent copula. It follows from the proposition, with tail dependent copulas $C_{V_1,V_2}, C_{1,V_1}, C_{2,V_2}$ we get tail dependence for C_{12}; this result also follows from a main theorem in Joe et al. [16] because the pairs (U_{11}, V_1), (V_1, V_2), (V_2, U_{22}) are the edges of the first tree of a vine (the vine representations of the bi-factor and nested factor copulas are given in the Appendix).

Also, if $C_{j|V_j}(u_j|v_j) \leq u_jv_{0j}$ for some $v_{0j} > 0$ if u_j is small enough (that is, the tail order of $C_{j|V_j}$ equals two), then we get

$$C_{12}(u, u) \leq u^2 v_{01}v_{02} \int_0^1 \int_0^1 c_{V_1,V_2}(v_1, v_2)dv_1dv_2 = u^2 v_{01}v_{02}.$$

If in addition C_{12} is PQD (conditions of Proposition 1 are satisfied), then C_{12} is a copula with tail quadrant independence. Hence, tail quadrant independence can be obtained by choosing linking copulas with tail order equal to two.

Proposition 3. *Assume that C_{j,V_j} is such that $C_{j|V_j}$ is stochastically increasing for $j = 1, 2$. Then the tail order κ_L of C_{12} in (6) is not less than the lower tail order of $C^*_{12}(u_1, u_2) = \int_0^1 C_{1|V_1}(u_1|v)C_{2|V_2}(u_2|v)dv$. Denote the lower tail order of the latter copula by κ^*_L. In addition, if for small enough $v > 0$ and some $m \geq 0, K_c > 0$ the inequality $vc_{V_1,V_2}(v, vq) \geq K_c\, q^m$ holds for any $q \in (0, 1)$, then $\kappa_L = \kappa^*_L$.*

Proof. Write $C_{12}(u, u) = C^-_{12}(u, u) + C^+_{12}(u, u)$, where the double integral over $[0, 1]^2$ for C_{12} is split into an integral over $v_1 \leq v_2$ and $v_1 > v_2$ respectively for C^-_{12} and C^+_{12}. Then, with $C_{2|V_2}$ stochastically increasing,

$$\begin{aligned} C^-_{12}(u, u) &= \int\int_{v_1\leq v_2} C_{1|V_1}(u|v_1)C_{2|V_2}(u|v_2)c_{V_1,V_2}(v_1, v_2)dv_1dv_2 \\ &\leq \int_0^1 \int_0^{v_2} C_{1|V_1}(u|v_1)C_{2|V_2}(u|v_1)c_{V_1,V_2}(v_1, v_2)dv_1dv_2 \\ &\leq \int_0^1 \int_0^1 C_{1|V_1}(u|v_1)C_{2|V_2}(u|v_1)c_{V_1,V_2}(v_1, v_2)dv_1dv_2 \\ &= \int_0^1 C_{1|V_1}(u|v_1)C_{2|V_2}(u|v_1) \left\{\int_0^1 c_{V_1,V_2}(v_1, v_2)dv_2\right\} dv_1 \\ &= \int_0^1 C_{1|V_1}(u|v_1)C_{2|V_2}(u|v_1)dv_1 = C^*_{12}(u, u). \end{aligned}$$

Similarly, with $C_{1|V_1}$ stochastically increasing,

$$C^+_{12}(u, u) = \int\int_{v_1>v_2} C_{1|V_1}(u|v_1)C_{2|V_2}(u|v_2)c_{V_1,V_2}(v_1, v_2)dv_1dv_2 \leq C^*_{12}(u, u)$$

and therefore $C_{12}(u, u) = C^-_{12}(u, u) + C^+_{12}(u, u) \leq 2C^*_{12}(u, u)$. It implies that $\kappa_L \geq \kappa^*_L$.

Now we prove the opposite inequality $\kappa_L \leq \kappa^*_L$ using the second assumption. Denote $\tilde{u} = \ln u$. For any $\epsilon > 0$ we have:

$$C_{12}(u, u) = \tilde{u}^2 \int_0^\infty \int_0^\infty C_{1|V_1}(u|u^{s_1})C_{2|V_2}(u|u^{s_2})c_{V_1,V_2}(u^{s_1}, u^{s_2})u^{s_1+s_2}ds_1ds_2$$

$$
\begin{aligned}
&\geq \tilde{u}^2 \int_0^\infty \int_{s_1}^{s_1+\epsilon} C_{1|V_1}(u|u^{s_1}) C_{2|V_2}(u|u^{s_2}) c_{V_1,V_2}(u^{s_1}, u^{s_2}) u^{s_1+s_2} ds_1 ds_2 \\
&\geq K_c \tilde{u}^2 \int_0^\infty \int_{s_1}^{s_1+\epsilon} C_{1|V_1}(u|u^{s_1}) C_{2|V_2}(u|u^{s_2}) u^{s_2} u^{m(s_2-s_1)} ds_1 ds_2 \\
&\geq K_c \tilde{u}^2 \int_0^\infty \int_{s_1}^{s_1+\epsilon} C_{1|V_1}(u|u^{s_1}) C_{2|V_2}(u|u^{s_2}) u^{s_1+\epsilon} u^{m\epsilon} ds_1 ds_2 \\
&\geq K_c\, \epsilon \tilde{u}^2 u^{\epsilon(m+1)} \int_0^\infty C_{1|V_1}(u|u^{s_1}) C_{2|V_2}(u|u^{s_1}) u^{s_1} ds_1 = K_c\, \epsilon \tilde{u} u^{\epsilon(m+1)} C^*_{12}(u, u).
\end{aligned} \tag{7}
$$

It implies that $\kappa_L \leq \kappa_L^* + (m+1)\epsilon$ for any $\epsilon > 0$ and hence $\kappa_L \leq \kappa_L^*$. As a result, $\kappa_L = \kappa_L^*$. $\square$

The condition on the density in Proposition 3 implies C_{V_1,V_2} is a lower tail dependent copula. In the Appendix we show, that the reflected Gumbel, BB1 and Student t_ν copula satisfy this condition; that is, it is something that can be readily checked and is not the most general sufficient condition. Under this condition, the nested copula model with a tail dependent copula C_{V_1,V_2} has the same tail order as the corresponding copula $C^*_{1,2}$ in a 1-factor copula model if C_{1,V_1} and C_{2,V_2} are stochastically increasing copulas. In particular, if C_{j,V_j} is a Gumbel copula for $j = 1, 2$, then we get intermediate lower tail dependence in the model.

In the next proposition we show that the increasing in concordance and stochastically increasing property can be obtained in a nested copula model under assumptions similar to factor copula models. If C_1, C_2 are two bivariate copulas, then C_2 is larger than C_1 in the concordance ordering if $C_2 \geq C_1$ pointwise.

Proposition 4. *Consider C_{12} in (6). Assume that C_{2,V_2} is fixed, $C_{V_2|V_1}$ is stochastically increasing and that $C_{2|V_2}$ is stochastically increasing (respectively decreasing). (a) As C_{1,V_1} increases in the concordance ordering, then C_{12} is increasing (respectively decreasing) in concordance. (b) If $C_{V_1|1}$ is stochastically increasing then $C_{2|1}$ is stochastically increasing (respectively decreasing).*

Proof. Suppose C_{1,V_1} is parameterized by a parameter θ and C_{2,V_1} is fixed. The increasing in concordance assumption implies that $C_{1,V_1}(\cdot;\theta_2) - C_{1,V_1}(\cdot;\theta_1) \geq 0$ for $\theta_1 < \theta_2$. Using the integration by parts formula we get:

$$
\begin{aligned}
C_{12}(u_1, u_2; \theta) &= u_1 \int_0^1 C_{2|V_2}(u_2|v_2) c_{V_1,V_2}(1, v_2) dv_2 - \int_0^1 \int_0^1 C_{1,V_1}(u_1, v_1; \theta)\, C_{2|V_2}(u_2|v_2) \frac{\partial c_{V_1,V_2}(v_1, v_2)}{\partial v_1} dv_1 dv_2 \\
&= u_1 \int_0^1 C_{2|V_2}(u_2|v_2) c_{V_1,V_2}(1, v_2) dv_2 \\
&\quad + \int_0^1 \int_0^1 C_{1,V_1}(u_1, v_1; \theta)\, \frac{\partial C_{2|V_2}(u_2|v_2)}{\partial v_2} \frac{\partial C_{V_2|V_1}(v_2|v_1)}{\partial v_1} dv_1 dv_2.
\end{aligned} \tag{8}
$$

With the assumption of twice continuous differentiability, $\partial C_{2|V_2}(u_2|v_2)/\partial v_2$ and $\partial C_{V_2|V_1}(v_2|v_1)/\partial v_1$ are continuous functions of v_1 and v_2 for $v_1, v_2 \in (0, 1)$ but can be unbounded at 0 or 1. Nevertheless, the integrand is an integrable function since

$$
\begin{aligned}
&\int_0^1 \int_0^1 \left| C_{1,V_1}(u_1, v_1; \theta) \cdot \frac{\partial C_{2|V_2}(u_2|v_2)}{\partial v_2} \frac{\partial C_{V_2|V_1}(v_2|v_1)}{\partial v_1} \right| dv_1 dv_2 \\
&\leq \left| \int_0^1 \int_0^1 \frac{\partial C_{2|V_2}(u_2|v_2)}{\partial v_2} \frac{\partial C_{V_2|V_1}(v_2|v_1)}{\partial v_1} dv_1 dv_2 \right| \\
&= \left| \int_0^1 [C_{V_2|V_1}(v_2|1) - C_{V_2|V_1}(v_2|0)] \cdot \frac{\partial C_{2|V_2}(u_2|v_2)}{\partial v_2} dv_2 \right| \\
&\leq \left| \int_0^1 \frac{\partial C_{2|V_2}(u_2|v_2)}{\partial v_2} dv_2 \right| = \left| C_{2|V_1}(u_2|0) - C_{2|V_1}(u_2|1) \right|.
\end{aligned}
$$

Therefore the formula (8) is valid.

To complete the proof of (a), for $\theta_2 > \theta_1$ we have:

$$
C_{1,2}(u_1, u_2; \theta_2) - C_{1,2}(u_1, u_2; \theta_1) = \int_0^1 \int_0^1 [C_{1,V_1}(u_1, v; \theta_2) - C_{1,V_1}(u_1, v; \theta_1)] \cdot \frac{\partial C_{2|V_2}(u_2|v_2)}{\partial v_2} \frac{\partial C_{V_2|V_1}(v_2|v_1)}{\partial v_1} dv_1 dv_2.
$$

Since $C_{1,V_1}(u_1, v; \theta_2) \geq C_{1,V_1}(u_1, v; \theta_1)$, $\partial C_{V_2|V_1}(v_2|v_1)/\partial v_1 \leq 0$ and $\partial C_{2|V_1}(u_2|v)/\partial v \leq (\geq)\ 0$ by the assumption of stochastically increasing (decreasing), we get $C_{1,2}(u_1, u_2; \theta_2) \geq (\leq)\ C_{1,2}(u_1, u_2; \theta_1)$ respectively, that is $C_{1,2}$ is increasing (decreasing) in concordance.

Table 2
Lower tail and dependence properties for $C_{1,2}$ depending on the choice of linking copulas in nested and bi-factor copula models.

Nested copula model, variables from the same group (V is the group-specific factor) Bi-factor copula model, variables from different groups (V is a common factor) U_1, U_2 are conditionally independent given V	
$C_{U_1\|V}, C_{U_2\|V}$ are SI	C_{U_1,U_2} is PQD
$C_{U_1,V}$ increases in concordance, $C_{U_2\|V}$ is SI	C_{U_1,U_2} increases in concordance
$C_{V\|U_1}, C_{U_2\|V}$ are SI	$C_{U_2\|U_1}$ is SI
$C_{U_1,V}, C_{U_2,V}$ are tail dependent	C_{U_1,U_2} is tail dependent
$C_{U_1,V}, C_{U_2,V}$ are both Gumbel or Gaussian[a] copulas	C_{U_1,U_2} has intermediate tail dependence
$C_{U_1,V}$ or $C_{U_2,V}$ is tail quadrant independent	C_{U_1,U_2} is tail quadrant independent
Nested copula model, variables from different groups (V_1, V_2 are the group-specific factors) U_1, U_2 are conditionally independent given V_1, V_2; V_1, V_2 are conditionally independent given V_0	
$C_{U_1\|V_1}, C_{U_2\|V_2}$ are SI, C_{V_1,V_2} is PQD	C_{U_1,U_2} is PQD
C_{U_1,V_1} increases in concordance, $C_{U_2\|V_2}, C_{V_2\|V_1}$ are SI	C_{U_1,U_2} increases in concordance
$C_{V_1\|U_1}, C_{U_2\|V_2}, C_{V_2\|V_1}$ are SI	$C_{U_2\|U_1}$ is SI
$C_{U_1,V_1}, C_{U_2,V_2}, C_{V_1,V_2}$ are tail dependent	C_{U_1,U_2} is tail dependent
C_{U_1,V_1}, C_{U_2,V_2} are tail dependent, and C_{V_1,V_2} has intermediate tail dependence	C_{U_1,U_2} has intermediate tail dependence
C_{U_1,V_1}, C_{U_2,V_2} are both Gumbel or Gaussian[a] copulas and C_{V_1,V_2} is tail dependent	C_{U_1,U_2} has intermediate tail dependence
C_{U_1,V_1} or C_{U_2,V_1} or C_{V_1,V_2} is tail quadrant independent	C_{U_1,U_2} is tail quadrant independent
Bi-factor copula model, variables from the same group U_1, U_2 are conditionally independent given V_0 and V_1 (V_0 is a common factor, V_1 is the-group specific factor)	
$C_{U_1\|V_0}, C_{U_2\|V_0}, C_{U_1\|V_1;V_0}, C_{U_2\|V_1;V_0}$ are SI	C_{U_1,U_2} is PQD
$C_{U_1,V_1;V_0}$ increases in concordance, $C_{U_2\|V_1;V_0}$ is SI	C_{U_1,U_2} increases in concordance
C_{U_1,V_0}, C_{U_2,V_0} or $C_{U_1,V_1;V_0}, C_{U_2,V_1;V_0}$ are tail dependent	C_{U_1,U_2} is tail dependent
$C_{U_1,V_0}, C_{U_2,V_0}, C_{U_1,V_1;V_0}, C_{U_2,V_1;V_0}$ are all Gumbel or Gaussian[a] copulas	C_{U_1,U_2} has intermediate tail dependence
$C_{U_1,V_0}, C_{U_1,V_1;V_0}$ or $C_{U_2,V_0}, C_{U_2,V_1;V_0}$ are tail quadrant independent	C_{U_1,U_2} is tail quadrant independent

[a] with positive correlation parameter.

Similarly, for (b), with $u_1 \in (0, 1)$, both parts of (8) can be differentiated with respect to u_1 twice to get

$$\frac{\partial^2 C_{1,2}(u_1, u_2; \theta)}{\partial u_1^2} = \frac{\partial C_{2|1}(u_2|u_1; \theta)}{\partial u_1} = \int_0^1 \int_0^1 \frac{\partial C_{V_1|1}(v|u_1; \theta)}{\partial u_1} \cdot \frac{\partial C_{2|V_1}(u_2|v_2)}{\partial v_2} \cdot \frac{\partial C_{V_2|V_1}(v_2|v_1)}{\partial v_1} dv_1 dv_2.$$

Assuming $C_{V_1|1}$ and $C_{V_2|V_1}$ are stochastically increasing we get $\partial C_{V_1|1}(v_1|u_1; \theta)/\partial u_1 \le 0$ and $\partial C_{V_2|V_1}(v_2|v_1; \theta)/\partial u_1 \le 0$. In addition, $\partial C_{2|V_1}(u_2|v_2; \theta)/\partial v \le (\ge)\ 0$ by the assumption of stochastically increasing (decreasing), then $\partial C_{2|1}(u_2|u_1; \theta)/\partial u_1 \le (\ge$, respectively) 0, that is, $C_{2|1}$ is stochastically increasing (decreasing). □

Dependence and tail properties of the nested and bi-factor copula models are summarized in Table 2. One can see that with a proper choice of bivariate linking copulas, different types of dependence and tail structures can be obtained. This is important as some preliminary analysis can be done before fitting the model to data in order to summarize dependence properties of the data set. The linking copulas in the model can then be selected to get a similar dependence structure; more details are given in Section 6.

4. Computational details for factor copula models

In this section we provide more details on the log-likelihood and maximum likelihood estimation of parameters in different factor copula models, including structured factor copula models.

4.1. Log-likelihood maximization in factor copula models

Suppose each bivariate linking copula in (3) or (5) has a parameter and $\boldsymbol{\theta}$ is the vector of all dependence parameters in the $2d$ or $d+G$ bivariate linking copulas. For multivariate data $(u_{i1}, \ldots, u_{id})$, $i = 1, \ldots, n$, that have been converted to have $U(0, 1)$ margins, the log-likelihood is:

$$\ell_n = \sum_{i=1}^{n} \log c_{\mathbf{U}}(u_{i1}, \ldots, u_{id}; \boldsymbol{\theta}). \tag{9}$$

When $\boldsymbol{\theta}$ is fixed, each term of the form (3) or (5) in the log-likelihood can be evaluated via Gauss–Legendre quadrature. With a relabeled vector of data $\mathbf{u} = (u_{11}, \ldots, u_{d_1 1}, \ldots, u_{1G}, \ldots, u_{d_G G})$, the copula density for the bi-factor copula model is evaluated as:

$$c_{\mathbf{U}}(\mathbf{u}; \boldsymbol{\theta}) \approx \sum_{i_1=1}^{n_q} w_{i_1} \prod_{g=1}^{G} \left\{ \left[\prod_{j=1}^{d_g} c_{U_{jg},V_0}(u_{jg}, x_{i_1}) \right] I_g(\mathbf{u}, x_{i_1}) \right\},$$

where

$$I_g(\mathbf{u}, x_{i_1}) \approx \sum_{i_2=1}^{n_q} w_{i_2} \left[\prod_{j=1}^{d_g} c_{U_{jg},V_g;V_0}\big(C_{U_{jg}|V_0}(u_{jg}|x_{i_1}), x_{i_2}\big) \right]$$

and $\{x_k\}$ are the quadrature nodes, $\{w_k\}$ are the quadrature weights, and n_q is the number of quadrature points. Similarly, the copula density for the nested copula model is evaluated as:

$$c_{\mathbf{U}}(\mathbf{u}; \boldsymbol{\theta}) \approx \sum_{i_1=1}^{n_q} w_{i_1} \prod_{g=1}^{G} I_g^*(\mathbf{u}, x_{i_1}),$$

where

$$I_g^*(\mathbf{u}, x_{i_1}) \approx \sum_{i_2=1}^{n_q} w_{i_2} c_{V_g,V_0}(x_{i_2}, x_{i_1}) \prod_{j=1}^{d_g} c_{U_{jg},V_g}(u_{jg}, x_{i_2}).$$

It is seen, that multidimensional summation is not required for the approximation, so that computational complexity is reduced. The number of quadrature points n_q between 25 and 30 tends to give a good approximation of these integrals and the resulting maximum likelihood estimates.

Maximizing the log-likelihood is the same as minimizing the negative log-likelihood and the latter is typically the numerical approach so that the Hessian of the negative log-likelihood at the global minimum is the inverse of the observed Fisher information matrix. For numerical minimization, quasi-Newton or modified Newton–Raphson algorithms can be used. For this purpose, one requires the first and second order partial derivatives of the density $c_U(\mathbf{u}; \boldsymbol{\theta})$ with respect to the dependence parameter vector $\boldsymbol{\theta}$.

The partial derivatives can be evaluated numerically by computing difference quotients of the log-likelihood function. The Hessian is obtained numerically through an updating method, such as the Broyden, Fletcher, Goldfarb and Shanno (BFGS) method. The algorithm with numerical derivatives is usually referred as a quasi-Newton method [26]. However, when d and the dimension of $\boldsymbol{\theta}$ become larger, multiple computations of the log-likelihood are needed and the algorithm becomes very slow to converge because of the steps needed for evaluating a numerical Hessian.

To overcome this difficulty for arbitrarily large d, we obtain analytical expressions for the gradient and Hessian. Then the Newton–Raphson method can be used, and the numerical minimization of the log-likelihood can work for large d and large dimension of $\boldsymbol{\theta}$, with a quadratic rate of convergence after iterations get close to a local or global minimum. Using the differentiation under the integral sign, one can see that the first and second order derivatives of the bivariate linking copulas with respect to their dependence parameters and arguments are required to find the gradient and Hessian of the log-likelihood. See the Appendix for the required (analytical) partial derivatives of the density and conditional distribution of the bivariate linking copulas. The partial derivatives of (3) or (5) with respect to the parameters are evaluated at the same time with Gauss–Legendre quadrature.

It is important to make sure that the value of the likelihood increases at each iteration. This is one of the modification steps for the modified Newton–Raphson method. However, when minimizing the nonlinear negative log-likelihood function of many parameters, the function value can increase and the algorithm can fail to converge especially if starting points are not close to the global minimum point. This happens if the Hessian is not a positive definite matrix so that there are some negative eigenvalues. To modify the algorithm, an eigenvalue decomposition of the Hessian matrix can be obtained and negative eigenvalues in the decomposition can be replaced by small positive numbers. With the adjusted positive definite Hessian matrix, the iterations will move to a local minimum of the negative log-likelihood and not a local maximum. The step size of the modified algorithm should be controlled so that parameters do not exceed lower and/or upper boundaries and it is not too large in any iteration.

Note that for each group of size 1 in a bi-factor model there is no group latent variable. Assume the gth group consists of a single variable U_{1g} only. To avoid overparameterization, the dependence parameter for the copula $C_{U_{1g},V_g;V_0}$ can be set to independence. In addition, if the gth group has two variables U_{1g}, U_{2g} in a bi-factor Gaussian model, the correlation parameter of the copula $C_{U_{1g},V_g;V_0}$ or $C_{U_{2g},V_g;V_0}$ can be set to 1 as there is redundance in this case as well. In a bi-factor model with other copulas the likelihood can be flat if there are some groups of size 2 so that dependence parameter for the copula $C_{U_{1g},V_g;V_0}$ can be set to comonotonic dependence similar to a Gaussian model.

For the algorithm, some good starting points may be required to obtain the global minimum of the negative log-likelihood. These starting points can be obtained from a stepwise optimization when dependence parameters are estimated in steps. For a nested copula model, parameters for the copulas C_{i,V_g}, $i = 1, \ldots, d_j$, can be estimated using data from the gth group. Within the group, data are modeled using a 1-factor copula model so that estimation is fast and stable. The parameters of the copulas C_{V_g,V_0} are estimated at the second step with the other parameters set equal to their estimates obtained at the first step. For a bi-factor copula model, the estimation can be done for each group separately. For each group we have a 2-factor copula model but with a smaller number of dependence parameters so that the estimation is much faster.

Alternatively, good starting points can be obtained from nested or bi-factor Gaussian model estimates, after conversion to parameter values to match Spearman's rho or Kendall's tau. Then for both models, stepwise estimates can be used as starting points when all dependence parameters are estimated simultaneously with the modified Newton–Raphson algorithm. The convergence of the algorithm is fast when good starting points are used and the sample size is large enough so that the log-likelihood is not as flat with many local maxima/minima.

In the bi-factor copula model, if bivariate Student t_ν copulas are used to link to the common latent variable, then the conditional distributions are univariate $t_{\nu+1}$ and this is needed in the copula $C_{jV_g;V_0}$ for linking variable j in group g given the common latent variable V_0. The speed of the log-likelihood evaluation is much faster when monotone interpolation is used for the univariate $t_{\nu+1}$ cdf based on its values on a fixed grid, say, at the quantiles in the set

$$\{0.0001, 0.0002, 0.0005, 0.001, 0.002, 0.005, 0.01(0.01)0.99, 0.995, 0.998, 0.999, 0.9995, 0.9998, 0.9999\};$$

references for monotone interpolation are Fritsch and Carlson [9] and Kahaner et al. [18].

4.2. Asymptotic covariance matrix of 2-stage copula–GARCH parameter estimates

For copula models with financial asset returns, it is common to use the copula–GARCH model (see, for example Jondeau and Rockinger [17], Aas et al. [2], Lee and Long [21] and others). For univariate marginals, the AR(1)–GARCH(1,1) model with symmetric Student t innovations is quite general for individual log-returns. At any time t, the jth (for $j = 1, \ldots, d$) GARCH innovations are assumed to be standardized t_{ν_j} with mean 0 and variance 1 with $\nu_j > 2$, and the vector of d innovations has a joint distribution based on the parametric copula family $C(\cdot; \boldsymbol{\theta})$. We assume parameters of the d univariate GARCH models are such that the time series are stationary.

In this section, we outline a resampling method to get the asymptotic covariance matrix of the parameters of the copula–GARCH model based on two-stage parameter estimation. The procedure can apply to any parametric copula model on the GARCH innovations of d dependent financial time series. Parameter estimates in copula models (including the factor copula models) are computed in two stages so that standard errors obtained from maximizing the copula likelihood (9) do not reflect the variability of GARCH parameter estimates. The simplest way to get standard errors for the two-stage estimation procedure is to use appropriate bootstrap methods. Let n be the original sample size and d is a number of log-returns. We use the following steps to get a bootstrap distribution for the maximum likelihood estimates.

1. Compute GARCH parameter estimates $\hat{\boldsymbol{\eta}}_1, \ldots, \hat{\boldsymbol{\eta}}_d$ using the original data, separately for the d returns.
2. For the jth return, convert GARCH-filtered residuals $\mathbf{R}_j = (R_{j1}, \ldots, R_{jn})^T$ to uniform data using the probability integral transform: $\mathbf{U}_j = (U_{j1}, \ldots, U_{jn})^T$, where $U_{ji} = T_{\nu_j}(R_{ji})$ and T_{ν_j} is the cdf of the Student distribution with ν_j degrees of freedom, $i = 1, \ldots, n$. Alternatively, the vector $\mathbf{U}_j$ can be obtained using uniform scores (as in Aas et al. [2]). The latter approach can provide less sensitivity to the assumption of innovations having a Student t distribution. Our experience is that the two methods give very similar results.
3. Compute copula parameter estimates $\hat{\boldsymbol{\theta}}$ from the d-dimensional data set $\mathbf{U}_{n\times d} = (\mathbf{U}_1, \ldots, \mathbf{U}_d)$, using the procedure in Section 4.1.
4. For the bth bootstrap sample, resample the filtered residuals as d-vectors at different time points (see Pascual et al. [28] for more details on bootstrap for GARCH parameter estimates).
5. Use the resampled filtered data and estimated GARCH parameters $\hat{\boldsymbol{\eta}}_1, \ldots, \hat{\boldsymbol{\eta}}_d$, to get a bootstrap sample of log-returns $\mathbf{r}^{(b)} = (\mathbf{r}_1^{(b)}, \ldots, \mathbf{r}_d^{(b)})$, where $\mathbf{r}_j^{(b)} = (r_{j1}^{(b)}, \ldots, r_{jn}^{(b)})$.
6. From a bootstrap sample $\mathbf{r}^{(b)}$, compute GARCH parameter estimates $\hat{\boldsymbol{\eta}}_1^{(b)}, \ldots, \hat{\boldsymbol{\eta}}_d^{(b)}$ and copula parameters $\hat{\boldsymbol{\theta}}^{(b)}$.
7. Repeat steps 4 to 6 for $b = 1, \ldots, B$, where B is a number of bootstrap samples. For example, B can be chosen to be between 1000 and 5000. Then one has a $B \times n_p$ matrix where n_p is the total number of parameters in the vectors $\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_d, \boldsymbol{\theta}$.

From a bootstrap distribution of the two-stage likelihood estimates one can compute standard errors and confidence intervals for $\hat{\boldsymbol{\eta}}_1, \ldots, \hat{\boldsymbol{\eta}}_d$, and $\hat{\boldsymbol{\theta}}$ as well as for the model-based estimates of different quantities which are functions of these parameter vectors. For example, to compute a confidence interval for the model-based Value-at-Risk estimate, for each bootstrap estimate $\hat{\boldsymbol{\eta}}_1^{(b)}, \ldots, \hat{\boldsymbol{\eta}}_d^{(b)}, \hat{\boldsymbol{\theta}}^{(b)}$, one can simulate a large data set of log-returns to compute portfolio VaR and hence get a bootstrap distribution; see the next section.

5. Interval estimation of VaR and CTE for copula–GARCH

In order to assess the comparison of different parametric copula models and how well they perform for tail-based inference, we propose model-based interval estimates of two risk-measures that are popular among financial analysts. Our approach with these risk measures is different from Aas and Berg [1] and others.

The first measure, the Value-at-Risk (VaR) is defined as a quantile of the distribution of a portfolio return. To explain the ideas, we assume an equally weighted portfolio of the d assets. Let $(\bar{r}_1, \ldots, \bar{r}_n)$ be the portfolio returns for n consecutive time units (such as trading days) and let $\hat{F}_{\bar{R}}$ be the corresponding empirical cdf. The $100\alpha\%$ VaR of the portfolio can be estimated as follows:

$$\widehat{\mathrm{VaR}}_\alpha = \{\inf \bar{r} : \hat{F}_{\bar{R}} \ge \alpha\}.$$

Common α values are 0.01, 0.05, 0.95 and 0.99. With a small α, VaR_α represents the maximal possible loss for investors, who buy the portfolio, that can occur with the probability not less than $100\alpha\%$. Similarly, with a large α, VaR_α represents maximal possible loss for investors, who short sell the portfolio, that can occur with the probability not less than $100(1-\alpha)\%$. Thus, both lower and upper quantiles are important for assessing risks related to the portfolio.

The second measure is called the conditional tail expectation (CTE) and it is defined as a conditional mean of a portfolio return given that the return falls below or exceeds some threshold. The lower (upper) CTE at level r^* can be estimated as follows:

$$\widehat{\mathrm{CTE}}^-(r^*) = \frac{\sum_{i^*:\,\bar R_{i^*}\le r^*} \bar R_{i^*}}{\sum_i I\{\bar R_i \le r^*\}}, \qquad \widehat{\mathrm{CTE}}^+(r^*) = \frac{\sum_{i^*:\,\bar R_{i^*}\ge r^*} \bar R_{i^*}}{\sum_i I\{\bar R_i \ge r^*\}}.$$

The lower CTE^- is used for small quantiles near 0.01, and the upper CTE^+ is used for large quantiles near 0.99. Unlike VaR, this risk measure estimates the expected loss of a portfolio if this loss occurs. Usually, the threshold r^* is set equal to the Value-at-Risk at a certain level α.

For each value of α, we compute one value of VaR_α for the whole data set using the stationarity of the log-returns. With a model $C(\cdot;\boldsymbol\theta)$ for copula–GARCH, the steps to obtain 95% model-based confidence intervals for VaR and CTE are as follows. Using the procedure in Section 4.2, one has

$$\hat{\boldsymbol\eta}_1^{(b)},\dots,\hat{\boldsymbol\eta}_d^{(b)},\hat{\boldsymbol\theta}^{(b)}, \quad b=1,\dots,B. \tag{10}$$

For each b, simulate a d-dimensional copula–GARCH time series of length N (N can be bigger than the original sample size n) with parameter vector in (10) and obtain the (equally-weighted) portfolio $(\bar r_1^{(b)},\dots,\bar r_N^{(b)})$ and make this into the bth row of a $B\times N$ matrix.

1. Each row b of the matrix can be considered as a realization of a stationary time series, so that quantiles (that is, $\mathrm{VaR}_\alpha^{(b)}$ for several α) can be computed for the series as well as some values of $\mathrm{CTE}^{(b)}(r^*)$.
2. With B series, we have B different realizations of $\mathrm{VaR}_\alpha^{(b)}$ and $\mathrm{CTE}^{(b)}(r^*)$ for a fixed α.
3. The middle interval containing 95% of the $\mathrm{VaR}_\alpha^{(b)}$ (respectively, $\mathrm{CTE}^{(b)}(r^*)$) values can be considered as a 95% confidence interval for VaR_α (respectively, $\mathrm{CTE}(r^*)$) that is model-based.

Note that these interval estimates account for the uncertainty in the parameter estimates in using the parametric model. The model-based estimates of VaR and CTE are functions of $\boldsymbol\eta_1,\dots,\boldsymbol\eta_d,\boldsymbol\theta$ that involve high-dimensional integrals, and hence they are estimated via Monte Carlo simulation (see the Appendix for simulation from nested factor and bi-factor copula models). The reason for obtaining model-based estimates of the portfolio VaR and CTE is so that in an empirical application, we can compare the effects of different models $C(\cdot;\boldsymbol\theta)$ that are structured copula models with bivariate linking copula families that have quite different tail characteristics.

6. Empirical study

In this section, we use some copula–GARCH models to analyze a financial data set. We consider S&P 500 stock returns from Health Care sector, 51 stocks in total, time period consists of the years 2010 and 2011. The sample size $n = 503$ days. The returns in this sector can be subdivided into 5 groups: health care distributors and services (27 stocks), health care equipment and services (6 stocks), biotechnology (6 stocks), managed health care (8 stocks) and pharmaceuticals (4 stocks).

For the copula–GARCH model, we apply AR(1)–GARCH(1, 1) model with symmetric Student t innovations to fit univariate marginals for log-returns. GARCH-filtered data are then transformed to uniform scores and different copula models are applied to model the joint dependence. Parameters in the model are estimated in two steps as given in Section 4. Copula choice is very important in the model and therefore we do a preliminary analysis to get summary of dependence structure of the data. More details on selecting appropriate linking copulas for the data set are provided in the next section.

6.1. Assessing strength of dependence in the tails

To choose appropriate copulas in the model, we employ some measures of dependence to assess strength of dependence in the lower and upper tails for each pair of returns from our data set and to summarize the dependence structure of the data. We use results from Section 3 to find linking copulas with tail properties that are in good agreement with the estimated strength of dependence in the tails.

To estimate dependence in the tails, one can use tail dependence coefficients but these quantities are defined as limits and cannot be estimated well unless the sample size is very large. Instead, we use tail-weighted measures of dependence, as described in Krupskii and Joe [19]. These measures are defined as correlations of transformed data where more weight is put in the joint tail. With variables transformed to U_1, U_2 that are $U(0,1)$ random variables, good choices of lower and upper tail-weighted dependence measures are:

$$\varrho_L = \mathrm{Cor}((1-2U_1)^6, (1-2U_2)^6 | U_1 < 0.5, U_2 < 0.5),$$
$$\varrho_U = \mathrm{Cor}((2U_1-1)^6, (2U_2-1)^6 | U_1 > 0.5, U_2 > 0.5).$$

Table 3
Overall and group averages of $\hat{\varrho}_L,\ \hat{\varrho}_U,\ \hat{\rho}_S$ and $\hat{\delta}_L,\ \hat{\delta}_U,\ \hat{\delta}_\rho$ (for the bi-factor Gaussian model); GARCH-filtered log-returns from S&P500 index, health care sector, years 2010–2011.

	All	Group 1	Group 2	Group 3	Group 4	Group 5
$\hat{\varrho}_L$	0.43	0.46	0.41	0.60	0.47	0.36
$\hat{\varrho}_U$	0.25	0.29	0.20	0.47	0.25	0.27
$\hat{\rho}_S$	0.50	0.53	0.50	0.72	0.52	0.53
$\hat{\delta}_L$	−0.17	−0.17	−0.16	−0.11	−0.19	−0.06
$\hat{\delta}_U$	0.00	−0.01	0.05	0.02	0.03	0.02
$\hat{\delta}_\rho$	−0.02	−0.01	−0.02	−0.01	0.00	−0.01

The weighting function $a(u) = (1-2u)^k$ downweights data points that are far from the joint tail and is more sensitive to different types of tail behavior if k is large. However, with a larger k, variability of the empirical estimates increases as well, and the choice $k = 6$ keeps balance between quite high sensitivity and relatively low variability.

The empirical versions involve sample correlations after transformation and truncation, and the model-based versions involves numerical integration. To compare the accuracy of different models in terms of assessing dependence in the tails, we compute empirical estimates $\hat{\varrho}_L,\ \hat{\varrho}_U$ for each pair of GARCH-filtered log-returns converted to uniform scores. In addition, as a measure of overall monotone dependence, we compute empirical estimates of the Spearman rank correlation coefficient $\hat{\rho}_S$.

For a given copula model, the model-based estimates $\varrho_L^m,\ \varrho_U^m,\ \rho_S^m$ are computed as a function of the MLE. We also compute the differences between the model-based and empirical estimates: $\hat{\delta}_L = \varrho_L^m - \hat{\varrho}_L$, $\hat{\delta}_U = \varrho_U^m - \hat{\varrho}_U$, $\hat{\delta}_\rho = \rho_S^m - \hat{\rho}_S$ for each pair of bivariate marginal distributions.

For $d = 51$, the number of pairs is $51 \times 25 = 1275$. Therefore to summarize dependence structure of the data set, we compute the average of $\hat{\varrho}_L, \hat{\varrho}_U, \hat{\rho}_S$ for all pairs of uniform scores, as well as for all pairs within each of the 5 groups. We denote the overall averages by $\hat{\varrho}_L(\text{all}),\ \hat{\varrho}_U(\text{all})$, and group averages by $\hat{\varrho}_L(g),\ \hat{\varrho}_U(g)$ for the gth group respectively. To evaluate the accuracy of assessing dependence in the tails by a given model, we compute the following quantities:

$$
\begin{array}{lll}
\hat{\delta}_L(\text{all}), \hat{\delta}_U(\text{all}) \quad (\hat{\delta}_L(g), \hat{\delta}_U(g)) & : & \text{the overall (within the } g\text{th group) averages of } \hat{\delta}_L, \hat{\delta}_U \text{ respectively;} \\
\hat{\delta}_\rho(\text{all}) \quad \hat{\delta}_\rho(g) & : & \text{the overall (within the } g\text{th group) average of } \hat{\delta}_\rho;
\end{array}
$$

The averaged differences allow the summarization of information in a few numbers and reduced variability when constructing confidence intervals. We compute $\hat{\varrho}_L(\text{all}), \hat{\varrho}_U(\text{all}), \hat{\varrho}_L(g), \hat{\varrho}_U(g), g = 1, \ldots, 5$, for our data and use the bi-factor Gaussian model as a benchmark to compute $\hat{\delta}_L(\text{all}), \hat{\delta}_U(\text{all}), \hat{\delta}_L(g), \hat{\delta}_U(g), g = 1, \ldots, 5$, for this model. The results are presented in Table 3.

It is seen dependence in the lower tail is stronger for all groups of log-returns which means models with reflection symmetric dependence structure, including the Gaussian and Student t models, may not be suitable for modeling these data. Large negative values $\hat{\delta}_L(\text{all}), \hat{\delta}_L(g), g = 1, \ldots, 5$, indicate that for all 5 groups dependence in the lower tail is much stronger compared to that of the model with Gaussian linking copulas. It implies that copulas with the lower tail dependence can be more suitable for the data. According to the results of Section 3 (see Table 2), one can use bivariate tail asymmetric linking copulas (copulas linking a latent variable and a group variable) with the lower tail dependence to get bivariate marginals with lower tail dependence for each of the 5 groups.

At the same time, the values $\hat{\delta}_U(\text{all}), \hat{\delta}_U(g), g = 1, \ldots, 5$, are quite close to zero and therefore dependence the upper tail is comparable to that of the Gaussian copula. It implies that one can use linking copulas with at least intermediate upper tail dependence to get bivariate marginals with at least intermediate upper tail dependence for each of the 5 groups (see Table 2). To summarize, one can see that for bivariate copulas linking a latent variable and group variables one can select an asymmetric copula with the lower tail dependence and at least intermediate upper tail dependence for all of the 5 groups in the data set.

6.2. Assessing adequacy of different models

For the dependence of the d innovations, we fit different nested and bi-factor models. For comparisons and sensitivity analyses, we include models that may not be appropriate based on the preliminary analyses in the preceding section. For each of these models we choose the same linking copula family to model dependence within each of the five groups. This is done for illustration purposes to show how the model choice affects the estimates of VaR, CTE and some dependence measures. Notice that according to the results from the previous section, dependence structure is quite similar in all 5 groups (lower tail dependence and intermediate upper tail dependence). In general, when the dependence structure and tail behavior is quite different in different groups, one can use different copula families in these groups.

For comparisons, we use bi-factor and nested Gaussian models as well as the multivariate Student t distribution with nested and bi-factor correlation structure (the latter behave similar to (5) and (3) with bivariate Student t copulas and are computationally faster for likelihood calculations). More specifically, we fit the following models.

Table 4
Overall and group averages of $\hat{\delta}_L$, $\hat{\delta}_U$, $\hat{\delta}_\rho$, and the scaled AIC value for different models; GARCH-filtered log-returns from S&P500 index, health care sector, years 2010–2011, $n = 503$.

	Overall	Group 1	Group 2	Group 3	Group 4	Group 5	AIC/n
	bi-factor Gaussian model						
$\hat{\delta}_L$	−0.17	−0.17	−0.16	−0.11	−0.19	−0.06	
$\hat{\delta}_U$	0.00	−0.01	0.05	0.02	0.03	0.02	−35.2
$\hat{\delta}_\rho$	−0.02	−0.01	−0.02	−0.01	0.00	−0.01	
	bi-factor Student t model						
$\hat{\delta}_L$	−0.11	−0.11	−0.10	−0.06	−0.15	−0.01	
$\hat{\delta}_U$	0.06	0.05	0.11	0.07	0.07	0.07	−37.5
$\hat{\delta}_\rho$	0.01	0.01	0.00	0.01	0.00	0.01	
	bi-factor Frank copula model						
$\hat{\delta}_L$	−0.30	−0.27	−0.25	−0.23	−0.30	−0.16	
$\hat{\delta}_U$	−0.13	−0.10	−0.04	−0.10	−0.09	−0.08	−36.7
$\hat{\delta}_\rho$	0.00	0.04	0.03	0.01	0.03	0.02	
	1-factor reflected BB1 copula model						
$\hat{\delta}_L$	−0.01	−0.02	−0.01	−0.24	−0.03	−0.07	
$\hat{\delta}_U$	0.05	0.04	0.00	−0.10	−0.01	0.08	−34.0
$\hat{\delta}_\rho$	−0.01	−0.01	−0.08	−0.22	−0.05	0.09	
	nested reflected Gumbel/reflected BB1 copula model						
$\hat{\delta}_L$	0.05	0.05	0.05	0.05	0.04	0.13	
$\hat{\delta}_U$	−0.04	0.00	0.03	0.03	−0.01	0.02	−36.7
$\hat{\delta}_\rho$	−0.02	0.00	−0.02	−0.01	0.00	−0.01	
	2-factor BB1/Frank copula model						
$\hat{\delta}_L$	−0.04	−0.04	−0.06	−0.23	−0.06	−0.03	
$\hat{\delta}_U$	0.00	0.02	−0.02	−0.20	−0.05	−0.03	−37.0
$\hat{\delta}_\rho$	0.04	0.04	−0.04	0.02	0.00	−0.05	
	bi-factor BB1/Frank copula model						
$\hat{\delta}_L$	0.01	0.02	0.02	−0.01	0.00	0.07	
$\hat{\delta}_U$	0.03	0.03	0.09	−0.03	0.08	0.07	−37.4
$\hat{\delta}_\rho$	−0.04	0.00	−0.01	0.00	−0.01	−0.02	

1. Nested and bi-factor models with Frank copulas at both levels. This is tail quadrant independent copula and so it is unsuitable for modeling data with tail dependence. We use this for comparison purpose to show that effect on tail inference with tail quadrant independent versus tail dependent bivariate linking copulas.
2. 1-factor model with reflected BB1 copulas with asymmetric tail dependence. This model can be used to model asymmetric dependence. We also fitted with BB1 copulas, but for this data set, reflected BB1 provided a better fit.
3. 2-factor model with BB1 copulas linking to the first factor and Frank copula linking to the second factor. This is more flexible model than the BB1 1-factor model because the Frank copula allows for negative conditional dependence.
4. Nested model with reflected Gumbel copulas to model dependence between groups and reflected BB1 copulas to model dependence within groups. This is an extension of the 1-factor BB1 model that accounts for group structure.
5. Bi-factor model with BB1 copulas for linking the common factor and Frank copulas for the second group-specific factor. This is an extension of the 2-factor BB1/Frank model that accounts for group structure.

For the algebraic forms of the Frank, BB1 and Gumbel copulas, see Joe [15]. The reflected form of a bivariate copula family $C(u, v; \theta)$ is $\widehat{C}(u, v; \theta) = u + v - 1 + C(1 - u, 1 - v; \theta)$; that is, if $(U, V) \sim C$, then $(1 - U, 1 - V) \sim \widehat{C}$.

We compute $\hat{\delta}_L(\text{all})$, $\hat{\delta}_U(\text{all})$, $\hat{\delta}_L(g)$, $\hat{\delta}_U(g)$, $g = 1, \ldots, 5$, for the above models; the results are presented in Table 4. We do not include nested Gaussian, Student t and Frank models as the results for these models and for the corresponding bi-factor models are quite close. The averaged values of $\hat{\delta}_L$ are negative for all groups for bi-factor Gaussian and Student t models. It implies dependence in the lower tail is underestimated by these models. The Frank copula model is even worse, as it heavily underestimates dependence in both tails. At the same time, models with tail asymmetric dependence structure perform better. Nevertheless, both the 1-factor reflected BB1 copula model and the 2-factor BB1/Frank copula model underestimate dependence in both tails in the third group. In addition, Spearman's rho is also underestimated by the 1-factor model (so that this dependence structure is too parsimonious), unlike other models with a group structure that give quite accurate estimates of the Spearman's rho in all groups. The reason is that dependence in the third group is significantly stronger than in the other groups and factor copula models assume homogeneous dependence across all groups. As a result, the estimated strength of dependence in the tails as well as overall dependence is mostly defined by the first very large group. With only 5 additional parameters, the nested reflected Gumbel/reflected BB1 copula model does better than 1-factor reflected BB1 copula model, with slightly overestimated dependence in the lower tail in the last group. The bi-factor BB1/Frank copula model is the best one as it assesses the strength of dependence in both tails reasonably well in all groups.

Now we compare the proposed models with vine copulas in terms of AIC using the algorithm of Dißmann et al. [8] and in the VineCopula R package. The regular vine model allows great flexibility to approximate the joint dependence of a multivariate data set by selecting bivariate linking copulas similar to the structured copula models. The following values of AIC/n with $n = 503$ were obtained for the regular vine model truncated after the first, second and third levels: -28.8, -33.9, -36.0. Models 4 and 5 yields $\mathrm{AIC}/n = -36.7$ and $\mathrm{AIC}/n = -37.4$ respectively. It is seen, that the proposed structured copula models do better in terms of AIC comparing to truncated regular vine models and a higher level of truncation is required for the regular vine to get AIC which is comparable to that of the nested and bi-factor copula models. The linking copulas and dependence structure in vine models are sequentially selected to maximize components of the likelihood. The vine models are less interpretable and do not use the information of the sectors; for our data set, stocks in the same sector are not always neighbors in the first tree of the vine. Also the choice of linking copulas in the structured copula models is based on the assessed strength of dependence in the tail.

The improved fits from structured copula models comes with additional computational time. With a personal computer with an Intel Core i5-2410M CPU at 2.3 GHz, some timings are: Multivariate Student t or Gaussian with bi-factor or nested factor structure: less than one minute; reflected Gumbel / reflected BB1 nested copula model: 29 min (12 iterations, 107 parameters); BB1/Frank bi-factor model: 48 min (23 iterations, 153 parameters); other nested factor and bi-factor models converge faster; 1-truncated, 2-truncated and 3-truncated regular vines: 1.5, 2.0 and 2.3 min respectively.

In the next subsection we do a more detailed analysis of the financial data set. In particular, we compute Value-at-Risk (VaR) and conditional tail expectations (CTE) for different models and compare the model-based estimates with the corresponding empirical estimates of these risk measures. Because the VaR/CTE are numerically more intensive with bootstrapping and Monte Carlo simulations, for further comparisons, we exclude the first and the last group, leaving 3 groups and 20 stocks.

6.3. VaR and CTE for different models

We consider an equally weighted portfolio of 20 stocks from the second, third and fourth groups of the data set considered in the previous section: health care equipment and services (6 stocks), biotechnology (6 stocks) and pharmaceuticals (8 stocks).

We use bi-factor Gaussian, Student t and Frank copula models (the corresponding nested models have very similar performance so that they are not included). In addition, we use the 1-factor reflected BB1 copula model, the nested model with reflected Gumbel/reflected BB1 and the bi-factor model with BB1/Frank.

For each model, using the procedure outlined in Section 5, we compute the model-based 95% confidence intervals for overall and group averages for ϱ_L, ϱ_U, ρ_S. Again, the averaged values allow reduced variability of the estimates and narrower confidence intervals despite the sample size not being very large. With smaller confidence intervals, models with different types of dependence structure can be differentiated more efficiently. We use $B = 2000$ bootstrap samples to obtain the intervals; in fact, the results stabilize when $B \geq 1000$.

GARCH-filtered log-returns are used to compute the corresponding empirical values. The results are presented in Table 5, and these should be mainly considered as a diagnostic assessment of fits of different copula models. It is seen, that all tail symmetric models underestimate dependence in the lower tail. In addition, the Student t model overestimates dependence in the upper tail and the bi-factor Frank copula model underestimate dependence in both tails. At the same time, the 1-factor model with a reflected BB1 copula better estimates dependence in the tails and the two other tail asymmetric models are more conservative as they slightly overestimate dependence in the tails. Nevertheless, Spearman's rho is significantly underestimated in all groups by the 1-factor model unlike other models with a group structure.

Next, we compute empirical estimates as well as the 95% confidence intervals for VaR_α with $\alpha = 0.01, 0.05, 0.95, 0.99$ and for $\mathrm{CTE}^-(-0.03)$, $\mathrm{CTE}^-(-0.02)$, $\mathrm{CTE}^+(0.02)$, $\mathrm{CTE}^+(0.03)$. We use $B = 2000$ bootstrap samples and simulate copula–GARCH time series of length $N = 20\,000$ to compute VaR and CTE for each sample and therefore to obtain the confidence intervals. The values ± 0.02, ± 0.03 for CTE approximately correspond to the lower and upper 5% and 1% empirical quantiles of the portfolio returns. The results are presented in Tables 6 and 7, as another diagnostic assessment of fits of different copula models. It is seen that VaR is not very sensitive to the model choice. Only $\mathrm{VaR}_{0.01}$ can detect weaker dependence for Frank copula as expressed in larger lower tail quantiles. Nevertheless, the reflection symmetric Student t and Gaussian models do reasonably well in terms of VaR risk measure despite the fact the dependence structure is misspecified by these models. A possible reason is that smaller quantiles are required to detect weaker dependence of these models in the lower tail. Otherwise the inferences on VaR and CTE are mainly dominated by the fit in the middle. However, VaR_α cannot be empirically estimated well with very small or large values of α unless the sample size is very large.

Conditional tail expectation is underestimated in the lower tail by the reflection symmetric copula models, but all models do reasonably well in the upper tail. However, the Student t model significantly overestimates the strength of dependence in the upper tail according to the tail-weighted dependence measures, whereas Frank copula model underestimates dependence in the upper tail. CTE can thus still not be very sensitive to the model misspecification.

Table 5
Overall and group estimated averages of ϱ_L, ϱ_U, ρ_S and the model-based 95% confidence intervals (**intervals that do not contain the empirical value are shown in bold font**); GARCH-filtered log-returns of stocks in the health care sector of the S&P500 index, years 2010–2011.

Model	ϱ_L(all)	ϱ_L (group 2)	ϱ_L(group 3)	ϱ_L(group 4)
Empirical estimate	0.42	0.41	0.60	0.47
bi-factor Gaussian	**(0.19, 0.29)**	**(0.21, 0.28)**	**(0.45, 0.53)**	**(0.23, 0.30)**
bi-factor Student *t*	**(0.30, 0.35)**	**(0.28, 0.35)**	**(0.50, 0.58)**	**(0.30, 0.37)**
bi-factor Frank	**(0.12, 0.15)**	**(0.14, 0.18)**	**(0.33, 0.40)**	**(0.14, 0.19)**
1-factor reflected BB1	(0.41, 0.50)	(0.34, 0.46)	(0.48, 0.63)	(0.38, 0.49)
Nested rGumbel/rBB1	**(0.44, 0.51)**	(0.41, 0.51)	(0.59, 0.69)	(0.44, 0.55)
bi-factor BB1/Frank	(0.37, 0.49)	(0.35, 0.49)	(0.51, 0.65)	(0.40, 0.54)
Model	ϱ_U (all)	ϱ_U (group 2)	ϱ_U(group 3)	ϱ_U(group 4)
Empirical estimate	0.24	0.20	0.47	0.25
bi-factor Gaussian	(0.19, 0.29)	**(0.21, 0.28)**	(0.45, 0.53)	(0.23, 0.30)
bi-factor Student *t*	**(0.30, 0.35)**	**(0.28, 0.35)**	**(0.50, 0.58)**	**(0.30, 0.37)**
bi-factor Frank	**(0.12, 0.15)**	**(0.14, 0.18)**	**(0.33, 0.40)**	**(0.14, 0.19)**
1-factor reflected BB1	(0.20, 0.30)	(0.11, 0.23)	(0.30, 0.53)	(0.13, 0.28)
Nested rGumbel/rBB1	(0.18, 0.24)	(0.18, 0.29)	(0.43, 0.57)	(0.20, 0.31)
bi-factor BB1/Frank	(0.22, 0.34)	(0.20, 0.34)	(0.38, 0.54)	(0.24, 0.37)
Model	ρ_S (all)	ρ_S (group 2)	ρ_S(group 3)	ρ_S(group 4)
Empirical estimate	0.50	0.50	0.72	0.52
bi-factor Gaussian	(0.38, 0.52)	(0.43, 0.52)	(0.68, 0.74)	(0.45, 0.54)
bi-factor Student *t*	(0.47, 0.54)	(0.45, 0.53)	(0.69, 0.76)	(0.48, 0.56)
bi-factor Frank	(0.46, 0.54)	(0.48, 0.56)	(0.69, 0.75)	(0.51, 0.59)
1-factor reflected BB1	(0.44, 0.51)	**(0.34, 0.44)**	**(0.57, 0.67)**	**(0.39, 0.49)**
Nested rGumbel/rBB1	(0.44, 0.52)	(0.43, 0.52)	(0.67, 0.73)	(0.47, 0.55)
bi-factor BB1/Frank	(0.43, 0.50)	(0.44, 0.53)	(0.68, 0.74)	(0.47, 0.55)

Table 6
Empirical estimates of VaR_α for $\alpha = 0.01, 0.05, 0.95, 0.99$ and the model-based 95% confidence intervals; GARCH-filtered log-returns for stocks from health care sector of the S&P500 index, years 2010–2011.

Model	$\text{VaR}_{0.01}$	$\text{VaR}_{0.05}$	$\text{VaR}_{0.95}$	$\text{VaR}_{0.99}$
Empirical estimate	−0.035	−0.020	0.019	0.029
bi-factor Gaussian	(−0.039, −0.024)	(−0.024, −0.016)	(0.017, 0.025)	(0.026, 0.039)
bi-factor Student *t*	(−0.040, −0.026)	(−0.024, −0.016)	(0.018, 0.025)	(0.028, 0.041)
bi-factor Frank	**(−0.030, −0.020)**	(−0.022, −0.015)	(0.017, 0.023)	(0.022, 0.032)
1-factor reflected BB1	(−0.045, −0.030)	(−0.025, −0.017)	(0.017, 0.024)	(0.026, 0.038)
Nested rGumbel/rBB1	(−0.046, −0.031)	(−0.025, −0.017)	(0.017, 0.023)	(0.025, 0.037)
bi-factor BB1/Frank	(−0.045, −0.029)	(−0.025, −0.017)	(0.017, 0.023)	(0.027, 0.039)

Table 7
Empirical estimates of $\text{CTE}^-(r^*)$, for $r^* = -0.03, -0.02$, and $\text{CTE}^+(r^*)$, for $r^* = 0.02, 0.03$, and the model-based 95% confidence intervals; GARCH-filtered log-returns for stocks from the health care sector in the S&P500 index, years 2010–2011.

Model	$\text{CTE}^-(-0.03)$	$\text{CTE}^-(-0.02)$	$\text{CTE}^+(0.02)$	$\text{CTE}^+(0.03)$
empirical estimate	−0.043	−0.030	0.028	0.039
bi-factor Gaussian	**(−0.041, −0.036)**	**(−0.029, −0.026)**	(0.026, 0.029)	(0.036, 0.041)
bi-factor Student *t*	**(−0.041, −0.037)**	(−0.030, −0.026)	(0.026, 0.030)	(0.037, 0.041)
bi-factor Frank	**(−0.041, −0.032)**	**(−0.026, −0.022)**	**(0.023, 0.026)**	(0.032, 0.039)
1-factor reflected BB1	(−0.045, −0.040)	(−0.033, −0.029)	(0.026, 0.030)	(0.036, 0.042)
nested rGumbel/rBB1	(−0.046, −0.040)	(−0.033, −0.029)	(0.025, 0.029)	(0.035, 0.041)
bi-factor BB1/Frank	(−0.045, −0.039)	(−0.033, −0.029)	(0.027, 0.031)	(0.038, 0.044)

Note that the 1-factor copula with reflected BB1 provides good estimates for VaR and CTE as the strength of dependence in the tails is estimated quite well by the model. After removing the first large group of returns, the model improves its performance, however the overall dependence is still underestimated as indicated by poor estimates of Spearman's rho. In conclusion, one can see that structured copula models specify both overall dependence and dependence in the tails quite well and provide good estimates for the considered risk measures. The improved fit of the bi-factor copulas as seen in the AIC values matches improved model-based estimates of tail-weighted dependence measures; that is, tail-weighted dependence measures are more sensitive than portfolio VaR/CTE in differentiating models that fit less well based on AIC. That is, some models with incorrectly specified dependence structure (in the middle or tails) still do reasonably well in terms of VaR and CTE. This indicates that very small (or large) values of α for VaR and, respectively, thresholds for CTE may be required to efficiently discriminate models with different tail properties and thus one needs a very large sample size to get a good estimates for VaR and CTE for a model validation. On the other hand, tail-weighted dependent measures can discriminate the copula models without such a large sample size.

7. Concluding remarks

The structured factor copula models allow the modeling of dependence for multivariate data sets when there are several non-overlapping groups of variables with homogeneous dependence in each group. These models contain bi-factor and nested Gaussian models as special cases. Tail dependence and tail asymmetry can be accommodated by choosing appropriate bivariate linking copulas. The number of dependence parameters in the models is a linear function of dimension d and, with appropriate numerical methods, two-stage maximum likelihood estimation is efficient for d up to over 100. Resampling to get model-based confidence intervals is however much more numerically intensive.

For simplicity of presentation in Section 6, we used a common parametric bivariate linking copula family for the group latent variables. To increase the log-likelihood or decrease AIC, more choices could be considered for each edge of the truncated vine associated with the structured copula model. In practice, one has to decide whether or not to continue trying to find parametric models that can yield smaller AIC values. Diagnostic procedures based on Spearman's rho and tail-weighted dependence measures can be used to assess adequacy of fit; bad fits can suggest dependence structures that are less parsimonious or choices of linking copulas to better match tail behavior seen in the multivariate data. The example in Section 6 and other similar examples show that the best model(s) based on simple assessments of adequacy of fit does (do) not always match with the best model(s) based on AIC.

The models can be further extended to a dynamic setting when dependence structure of the multivariate data set can change in time. To account for dynamic dependence, copulas with dependence parameters evolving in time can be used as in Patton [29] or alternatively regime switching copula can be applied; see, for example, Chollette et al. [6]. As a possible application, the dynamic structured copula model would allow to forecast financial portfolio returns and quantify possible risks of the portfolio, while accounting for (i) strong dependence in the tails and (ii) possible tail asymmetry of a multivariate distribution of the portfolio components. Future research includes studying the dependence properties of such an extended dynamic structured copula model. Based on the model prediction, possible strategies for investors could be investigated in order to minimize the risks of a portfolio.

Acknowledgments

This research has been supported by an NSERC Discovery Grant 8698. We are grateful to the referees for detailed comments leading to an improved presentation.

Appendix

A.1. Bivariate copulas satisfying assumptions of Proposition 3

We check that the condition $vc(v, qv) \geq K_c q^m$, $q \in (0, 1)$, $K_c > 0$, $m \geq 0$, for small enough v from Proposition 3 is satisfied for many bivariate copulas that are used in applications.

1. Reflected Gumbel copula with a dependence parameter $\theta > 1$. Define $t_1 = -\ln(1-u_1)$, $t_2 = -\ln(1-u_2)$, $s = t_1^\theta + t_2^\theta$, $r = s^{1/\theta}$. It can be shown the density of the reflected Gumbel copula is as follows:

$$c(u_1, u_2) = (r + \theta - 1)r^{1-2\theta} \exp\{-r\}(t_1 t_2)^{\theta-1}/[(1-u_1)(1-u_2)].$$

Let $u_1 = v$ and $u_2 = qv$. It follows that $t_1 \geq v$ and $t_2 \geq qv$ and as $v \to 0$ $t_1 \leq 2v$, $t_2 \leq 2vq$. It implies $r \leq 2v(1+q^\theta)^{1/\theta} \leq v2^{1/\theta+1}$. For the reflected Gumbel copula we have:

$$vc(v, vq) \geq v(\theta-1)[v2^{1/\theta+1}]^{1-2\theta} \exp\{-v2^{1/\theta+1}\}(t_1 t_2)^{\theta-1} \geq K_c v v^{1-2\theta}(v^2 q)^{\theta-1} = K_c q^{\theta-1}$$

where $K_c = (\theta-1)2^{(1-2\theta)(1/\theta+1)} \exp\{-2^{1/\theta+1}\}$. Hence the assumption of Proposition 3 is satisfied with $m = \theta - 1$.

2. BB1 copula with dependence parameters $\theta > 0$ and $\delta > 1$. Define $t_1 = u_1^{-\theta}$, $t_2 = u_2^{-\theta}$, $s = (t_1-1)^\delta + (t_2-1)^\delta$ and $r = s^{1/\delta}$. The density of the BB1 copula is given by the formula:

$$c(u_1, u_2) = (1+r)^{-(1/\theta+2)} r^{1-2\delta}[\theta(\delta-1) + (\theta\delta+1)r][(t_1-1)(t_2-1)]^{\delta-1}/(u_1 u_2)^{\theta+1}.$$

Let $u_1 = v$ and $u_2 = vq$. As $v \to 0$ we have $t_j \to \infty$, $j = 1, 2$, $r \to \infty$ and therefore for small enough v we have $r + 1 < 2r$ and $t_j - 1 > t_j/2$. It follows that $s \leq t_1^\delta + t_2^\delta = (qv)^{-\delta\theta}(1+q^{\theta\delta}) \leq 2(qv)^{-\delta\theta}$ and $r \leq 2^{1/\delta}(qv)^{-\theta}$. As a result, for the BB1 copula we have:

$$\begin{aligned} vc(v, vq) &\geq v(2r)^{-(1/\theta+2)} r^{1-2\delta}[(\theta\delta+1)r](t_1 t_2/4)^{\delta-1}(u_1 u_2)^{-\theta-1} = K_c^* v r^{-1/\theta-2\delta}(v^2 q)^{-\theta(\delta-1)-\theta-1} \\ &\geq K_c^* 2^{-(1/(\delta\theta)+2)} v(qv)^{1+2\theta\delta}(v^2 q)^{-\theta\delta-1} = K_c q^{\theta\delta} \end{aligned}$$

where $K_c^* = (\theta\delta + 1)2^{-1/\theta-2\delta}$ and $K_c = K_c^* 2^{-(1/(\theta\delta)+2)} = (\theta\delta + 1)2^{-(2+1/(\theta\delta))(\delta+1)}$. Hence the assumption of Proposition 3 is satisfied with $m = \theta\delta$.

3. Student t_ν copula with dependence parameters $\rho \in (-1; 1)$ and shape parameter $\nu > 0$. Define $t_1 = T_\nu^{-1}(u_1)$, $t_2 = T_\nu^{-1}(u_2)$ and $d^* = 1 + (t_1^2 - 2\rho t_1 t_2 + t_2^2)/[\nu(1-\rho^2)]$. The density of the Student copula is given by the formula:

$$c(u_1, u_2) = \left[\left(1 + \frac{t_1^2}{\nu}\right)\left(1 + \frac{t_2^2}{\nu}\right)\right]^{0.5\nu+0.5} \left(\frac{\sqrt{\nu}\Gamma(0.5\nu)}{\Gamma(0.5\nu+0.5)}\right)^2 \frac{[d^*]^{-0.5\nu-1}}{2\sqrt{1-\rho^2}}.$$

Let $u_1 = v, u_2 = vq$. Using the asymptotic behavior of Student quantiles, as $v \to 0$, we get $-k_1^- v^{-1/\nu} \le t_1 \le -k_1^+ v^{-1/\nu}$, $-k_2^-(vq)^{-1/\nu} \le t_2 \le -k_2^+(vq)^{-1/\nu}$ for some positive constants $k_1^-, k_1^+, k_2^-, k_2^+$. It implies $d^* \le 1 + (t_1 + t_2)^2/[\nu(1-\rho^2)] \le 1 + (2t_2)^2/[\nu(1-\rho^2)] \le 5t_2^2/[\nu(1-\rho^2)] \le k_3(vq)^{-2/\nu}$, where $k_3 = 5(k_2^-)^2/[\nu(1-\rho^2)]$. Denote $k_4 = \left(\frac{\sqrt{\nu}\Gamma(0.5\nu)}{\Gamma(0.5\nu+0.5)}\right)^2 \frac{1}{2\sqrt{1-\rho^2}}$.
For the Student copula we have:

$$\begin{aligned} vc(v, vq) &\ge k_4 v[(t_1t_2)^2/\nu^2]^{0.5\nu+0.5}[d^*]^{-0.5\nu-1} \ge k_4 v[(k_1^+k_2^+)^2(v^2q)^{-2/\nu}/\nu^2]^{0.5\nu+0.5}[k_3(vq)^{-2/\nu}]^{-0.5\nu-1} \\ &= K_c v v^{-2-2/\nu} q^{-1-1/\nu}(vq)^{1+2/\nu} = K_c q^{1/\nu} \end{aligned}$$

where $K_c = k_4[k_1^+k_2^+/\nu]^{\nu+1}k_3^{-0.5\nu-1}$. Hence the assumption of Proposition 3 is satisfied with $m = 1/\nu$.

A.2. Derivatives of bivariate linking copulas

In order to get the gradient and Hessian of the negative log-likelihood for numerical minimization via a modified Newton–Raphson method, the following derivatives are analytically needed. These derivatives are required when differentiating under the integral sign using the chain rule formula.

In a nested copula model (5) the following derivatives are required:

$$\frac{\partial c_{V_g,V_0}(v_g, v_0; \theta_g)}{\partial \theta_g}, \quad \frac{\partial c_{U_{ig},V_g}(u_{ig}, v_g; \eta_{ig})}{\partial \eta_{ig}}, \quad \frac{\partial^2 c_{V_g,V_0}(v_g, v_0; \theta_g)}{\partial \theta_g \partial \theta_g^T}, \quad \frac{\partial^2 c_{U_{ig},V_g}(u_{ig}, v_g; \eta_{ig})}{\partial \eta_{ig}\partial \eta_{ig}^T},$$

where θ_g, η_{ig} are dependence parameters of the copulas C_{V_g,V_0} and C_{U_{ig},V_g} respectively. In a bi-factor model (3) the required derivatives are as follows:

$$\frac{\partial c_{U_{ig},V_0}(u_{ij}, v_0; \theta_{ig})}{\partial \theta_{ig}}, \quad \frac{\partial c_{U_{ig},V_g;V_0}(w_{ig}, v_g; \gamma_{ig})}{\partial \gamma_{ig}}, \quad \frac{\partial c_{U_{ig},V_g;V_0}(w_{ig}, v_g; \gamma_{ig})}{\partial w_{ig}}, \quad \frac{\partial C_{U_{ig}|V_0}(u_{ig}|v_0; \theta_{ig})}{\partial \theta_{ig}},$$
$$\frac{\partial^2 c_{U_{ig},V_0}(u_{ig}, v_0; \theta_{ig})}{\partial \theta_{ig}\partial \theta_{ig}^T}, \quad \frac{\partial^2 c_{U_{ig},V_g;V_0}(w_{ig}, v_g; \gamma_{ig})}{\partial \gamma_{ig}\partial \gamma_{ig}^T}, \quad \frac{\partial^2 c_{U_{ig},V_g;V_0}(w_{ig}, v_g; \gamma_{ig})}{\partial w_{ig}^2},$$
$$\frac{\partial^2 C_{U_{ig}|V_0}(u_{ig}|v_0; \theta_{ig})}{\partial \theta_{ig}\partial \theta_{ig}^T}, \quad \frac{\partial^2 c_{U_{ig},V_g;V_0}(w_{ig}, v_g; \gamma_{ig})}{\partial w_{ig}\,\partial \gamma_{ig}},$$

where θ_{ig}, γ_{ig} are dependence parameters of the copula C_{U_{ig},V_0} and $C_{U_{ig},V_g;V_0}$ respectively.

For the common bivariate copula families, the density along with the partial derivatives are available in a closed form. As such, the required derivatives of the $c_{\boldsymbol{U}}(\mathbf{u})$ can be represented as integrals of some explicit functions and can be computed using numerical integration.

A.3. Structured factor models as truncated vines

Simulation from the structured factor copula models is straightforward after recognizing that they are special cases of truncated vines rooted at latent variables.

For the bi-factor copula model with G groups of variables, let $V_0, V_1, \ldots, V_G$ be the independent latent $U(0, 1)$ variables and let U_{ig} be the observed variables for $i = 1, \ldots, d_g, g = 1, \ldots, G$. This can be represented as a 2-truncated regular-vine. The edges of tree 1 are $[V_0, V_1], \ldots, [V_0, V_G]$ and $[V_0, U_{ig}]$ for $i = 1, \ldots, d_g, g = 1, \ldots, G$; there is a total of $G + d$ edges in this tree. For tree 2, the edges are $[V_1, V_g|V_0]$ for $g = 2, \ldots, G$ and $[V_g, U_{ig}|V_0]$ for $i = 1, \ldots, d_g, g = 1, \ldots, G$; there is a total of $G - 1 + d$ edges in this tree; see Fig. 1.

For the nested factor copula model with G groups of variables, let $V_0, V_1, \ldots, V_G$ be the dependent latent $U(0, 1)$ variables with a 1-factor structure, and let U_{ig} be the observed variables. This can be represented as a 1-truncated regular vine with the edges of tree denoted as $[V_0, V_1], \ldots, [V_0, V_G]$ and $[V_g, U_{ig}]$ for $i = 1, \ldots, d_g, g = 1, \ldots, G$; there is a total of $G + d$ edges; see Fig. 2.

From these truncated vines, the joint density of the observed and latent variable can be obtained from the vine density result in Bedford and Cooke [3], and then integration over the latent variables leads to (3) and (5).

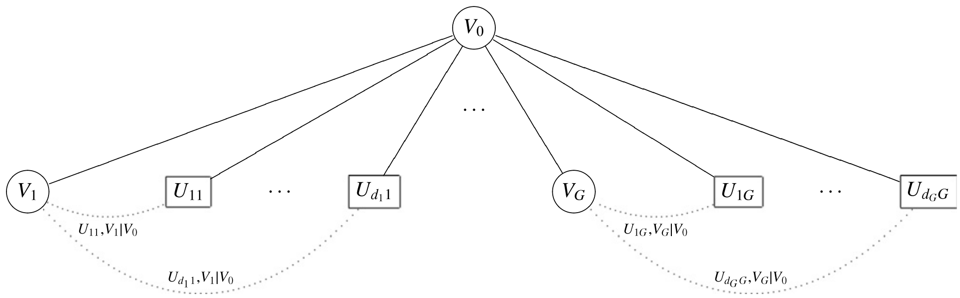

Fig. 1. Bi-factor model with G groups, d_j variables in the jth group.

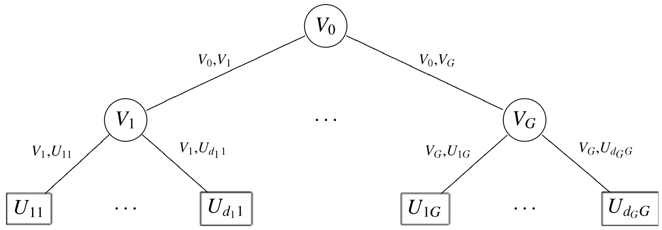

Fig. 2. Nested model with G groups, d_j variables in the jth group.

A.4. Simulating data from a nested copula model

Assume we need to simulate a random vector $\mathbf{U} = (U_{11}, \ldots, U_{d_1 1}, \ldots, U_{1G}, \ldots, U_{d_G G})$ from the model (5) with G groups of size $d_g, g = 1, \ldots, G$. Below is a simple algorithm for simulating data from the model:

1. Simulate $G + 1$ independent random variables $V_0, W_1, \ldots, W_G \sim U(0, 1)$;
2. Use the inverse conditional cdf $C^{-1}_{V_g|V_0}$ to simulate group latent variables: $V_g = C^{-1}_{V_g|V_0}(W_g|V_0), g = 1, \ldots, G$;
3. Simulate $d_1 + \cdots + d_g$ independent random variables $W_{11}, \ldots, W_{d_1 1}, \ldots, W_{1G}, \ldots, W_{d_G G} \sim U(0, 1)$;
4. Use the inverse conditional cdf $C^{-1}_{U_{ig}|V_g}$ to simulate variables within the gth group: $U_{ig} = C^{-1}_{U_{ig}|V_g}(W_{ig}|V_g), i = 1, \ldots, d_g$.

For some bivariate copulas the inverse conditional cdfs are available in a closed form. For other copulas, such as the Gumbel or BB1 copula, the inverse conditional cdfs can be computed quickly using numerical methods.

A.5. Simulating data from a bi-factor copula model

Assume we need to simulate a random vector $\mathbf{U} = (U_{11}, \ldots, U_{d_1 1}, \ldots, U_{1G}, \ldots, U_{d_G G})$ from the model (3) with G groups of size $d_g, g = 1, \ldots, G$. We use the following algorithm:

1. Simulate $1 + G + d_1 + \cdots + d_g$ independent random variables $V_0, V_1, \ldots, V_g, W_{11}, \ldots, W_{d_1 1}, \ldots, W_{1G}, \ldots, W_{d_G G} \sim U(0, 1)$;
2. Use the inverse conditional cdfs $C^{-1}_{U_{ig}|V_g;V_0}$ and $C^{-1}_{U_{ig}|V_0}$ to simulate $V_{ig} = C^{-1}_{U_{ig}|V_g;V_0}(W_{ig}|V_g)$, $U_{ig} = C^{-1}_{U_{ig}|V_0}(V_{ig}|V_0)$ for $i = 1, \ldots, d_g, g = 1, \ldots, G$.

References

[1] K. Aas, D. Berg, Models for construction of multivariate dependence - a comparison study, Eur. J. Finance 15 (2009) 639–659.
[2] K. Aas, C. Czado, A. Frigessi, H. Bakken, Pair-copula constructions of multiple dependence, Insurance Math. Econom. 44 (2009) 182–198.
[3] T. Bedford, R.M. Cooke, Probability density decomposition for conditionally dependent random variables modeled by vines, Ann. Math. Artif. Intell. 32 (2001) 245–268.
[4] E.C. Brechmann, Hierarchical Kendall copulas: properties and inference, Canad. J. Statist. 42 (2014) 78–108.
[5] E.C. Brechmann, C. Czado, Risk management with high-dimensional vine copulas: an analysis of the Euro Stoxx 50, Stat. Risk Model. 30 (2013) 307–342.
[6] L. Chollette, A. Heinen, A. Valdesogo, Modeling international financial returns with a multivariate regime-switching copula, J. Financ. Econom. 7 (2009) 437–480.
[7] S. Demarta, A.J. McNeil, The t copula and related copulas, Int. Stat. Rev. 73 (2005) 111–129.
[8] J. Dißmann, E.C. Brechmann, C. Czado, D. Kurowicka, Selecting and estimating regular vine copulae and application to financial returns, Comput. Statist. Data Anal. 59 (2013) 52–69.
[9] F.N. Fritsch, R.E. Carlson, Monotone piecewise cubic interpolation, SIAM J. Numer. Anal. 17 (1980) 238–246.
[10] R. Gibbons, D. Hedeker, Full-information item bi-factor analysis, Psychometrika 57 (1992) 423–436.
[11] C.W. Harris, H.F. Kaiser, Oblique factor analytic solutions by orthogonal transformations, Psychometrika 29 (1964) 347–362.
[12] K.J. Holzinger, F. Swineford, The bi-factor method, Psychometrika 2 (1937) 41–54.
[13] L. Hua, H. Joe, Tail order and intermediate tail dependence of multivariate copulas, J. Multivariate Analysis 102 (2011) 1454–1471.

[14] J. Hull, A. White, Valuation of a CDO and an nth to default CDS without Monte Carlo simulation, J. Derivatives 12 (2004) 8–23.
[15] H. Joe, Multivariate Models and Dependence Concepts, Chapman & Hall, London, 1997.
[16] H. Joe, H. Li, A. Nikoloulopoulos, Tail dependence functions and vine copulas, J. Multivariate Analysis 101 (2010) 252–270.
[17] E. Jondeau, M. Rockinger, The copula-GARCH model of conditional dependencies: an international stock market application, J. Int. Money and Finance 25 (2006) 827–853.
[18] D. Kahaner, C.B. Moler, S. Nash, Numerical Methods and Software, Prentice Hall, Englewood Cliffs, NJ, 1989.
[19] P. Krupskii, H. Joe, Factor copula models for multivariate data, J. Multivariate Anal. 120 (2013) 85–101.
[20] D. Kurowicka, H. Joe, Dependence Modeling: Vine Copula Handbook, World Scientific, Singapore, 2011.
[21] T.H. Lee, X. Long, Copula-based multivariate GARCH model with uncorrelated dependent errors, J. Econometrics 150 (2009) 207–218.
[22] R.P. McDonald, Factor Analysis and Related Methods, Lawrence Erlbaum Associates, Hillsdale, NJ, 1985.
[23] A.J. McNeil, R. Frey, P. Embrechts, Quantitative Risk Management, Princeton University Press, Princeton, NJ, 2005.
[24] G. Salvadori, C. De Michele, N.T. Kottegoda, R. Rosso, Extremes in Nature. An Approach using Copulas, Springer, Dordrecht, 2007.
[25] B.O. Muthen, Multilevel covariance structure analysis, Sociol. Methods Res. 22 (1994) 376–398.
[26] J.C. Nash, Compact Numerical Methods for Computers: Linear Algebra and Function Minimisation, second ed., Hilger, New York, 1990.
[27] A.K. Nikoloulopoulos, H. Joe, Factor copula models for item response data, Psychometrika (2015).
[28] L. Pascual, J. Romo, E. Ruiz, Bootstrap prediction for returns and volatilities in GARCH models, Comput. Statist. Data Anal. 50 (2006) 2293–2312.
[29] A. Patton, Modeling asymmetric exchange rate dependence, Internat. Econom. Rev. 47 (2006) 527–556.
[30] A. Sklar, Fonctions de répartition à n dimensions et leurs marges, Publ. Inst. Statist. Univ. Paris 8 (1959) 229–231.