

Hierarchical Kendall copulas: Properties and inference

Eike Christian Brechmann*

September 20, 2013

Abstract

While there is substantial need for dependence models in higher dimensions, most existing models quickly become rather restrictive and barely balance parsimony and flexibility. Hierarchical constructions may improve on that by grouping variables in different levels. In this paper, the new class of hierarchical Kendall copulas is proposed and discussed. Hierarchical Kendall copulas are built up by flexible copulas specified for groups of variables, where aggregation is facilitated by the Kendall distribution function, the multivariate analog to the probability integral transform for univariate random variables. After deriving properties of the general model formulation, particular focus is given to inference techniques of hierarchical Kendall copulas with Archimedean components, for which closed-form analytical expressions can be derived. A substantive application to German stock returns finally shows that hierarchical Kendall copulas perform very well for real data, out-of- as well as in-sample.

Keywords: multivariate copula, hierarchical copula, Kendall distribution function

1 Introduction

Dependence modeling using copulas has made significant progress in the last years. A d -dimensional copula is a multivariate distribution function on $[0, 1]^d$ with uniform margins. Their central role in dependence modeling is due to the famous theorem by Sklar (1959), which states that every multivariate distribution function can be expressed in terms of a copula and the univariate marginal distribution functions. More precisely, let $F_{\mathbf{X}}$ be the d -dimensional distribution function of the random vector $\mathbf{X} = (X_1, \dots, X_d)'$ with univariate marginal distribution functions F_{X_1}, \dots, F_{X_d} . Then there exists a copula C such that for all $\mathbf{x} = (x_1, \dots, x_d)' \in [-\infty, \infty]^d$,

$$F_{\mathbf{X}}(\mathbf{x}) = C(F_{X_1}(x_1), \dots, F_{X_d}(x_d)). \quad (1.1)$$

The copula C is unique if F_{X_1}, \dots, F_{X_d} are continuous.

Many of the standard, and also of the newly proposed, copula models however turn out to be rather restrictive in higher dimensions, which makes it virtually impossible to use them for very large data sets as required, e.g., in financial or spatial applications. While standard multivariate elliptical copulas such as the Gaussian and the Student's t require the specification of the full correlation matrix and can only account for symmetric dependence, multivariate Archimedean copulas are even more restrictive by assuming exchangeability and imposing that

*Center for Mathematical Sciences, Technische Universität München, Boltzmannstr. 3, D-85747 Garching, Germany. E-mail: eike.brechmann@tum.de. Phone: +49 89 289-17425. Fax: +49 89 289-17435.

all multivariate margins are the same. One common procedure to approach such problems therefore is grouping data, e.g., by industry sectors or nationality. Such copula models include the grouped Student’s t copula by Daul et al. (2003), elliptical copulas with clustered correlation matrix (see, e.g., Gregory and Laurent (2004)) and hierarchical Archimedean copulas, which were initially proposed by Joe (1997). In particular, hierarchical structures such as the latter two are very appealing and received considerable attention lately (see, e.g., Hofert (2010)).

A major issue of any copula model is to find a good balance between parsimony and flexibility. While elliptical copulas require an enormous number of parameters for specifying the correlation matrix (the number of parameters grows quadratically with the dimension), Archimedean and also hierarchical Archimedean copulas are much more parsimonious, since the number of parameters is at most linear in the dimension. However, such restrictions may be severe, since hierarchical Archimedean copulas are at the same time limited to the class of Archimedean copulas as building blocks. Similarly, an elliptical copula with clustered correlation matrix has to satisfy positive definiteness constraints and is limited to an elliptical dependence structure, which in particular implies tail symmetry.

Another class of copulas that recently attracted increasing attention are pair copula constructions as proposed by Aas et al. (2009). Such pair copula constructions use bivariate copulas of arbitrary types as building blocks and are also available for applications in larger dimensions. They however are non-hierarchical models and therefore may also severely suffer from extreme numbers of parameters. While means to counteract these problems are now investigated (Brechmann et al. 2012), we focus here on hierarchical constructions, which are inherently more parsimonious and easier to interpret in terms of within- and between-group dependence.

The purpose of this paper is to introduce the new class of hierarchical Kendall copulas as a flexible, but yet parsimonious dependence model. It is built up by copulas for groups (clusters) of variables in different hierarchical levels. In particular—and in contrast to hierarchical Archimedean copulas—the choice of copulas and their parameters is not restricted. With pair copula constructions the model shares the property that building blocks can be copulas of arbitrary types. Hierarchical Kendall copulas therefore provide a new and attractive option to model dependence patterns between large numbers of variables.

The name “hierarchical Kendall copula” is chosen to stress the central role of the Kendall distribution function in the model formulation. The Kendall distribution function is the multivariate analog to the probability integral transform for univariate random variables. It is used to aggregate the (dependence) information of a group of variables. It was first studied by Genest and Rivest (1993) in the bivariate case and in more detail by Barbe et al. (1996). Other accounts on it can be found, amongst others, in Imlahi et al. (1999), Genest and Rivest (2001) and Nelsen et al. (2003) as well as in the copula goodness-of-fit literature (see, e.g., Wang and Wells (2000) and Genest et al. (2006)).

It has been shown by Genest et al. (1995) that the only copula that gives a valid multivariate distribution for non-overlapping multivariate marginals is the independence copula. That is, if in Equation (1.1) non-overlapping multivariate distribution functions instead of univariate ones are plugged into the copula C , this copula can only be the independence copula. Marco and Ruiz-Rivas (1992) state conditions how a distribution function with specified multivariate marginals can be constructed; the easiest case being that margins are max-infinitely divisible, which includes distributions based on Archimedean copulas. Hierarchical Kendall copulas circumvent such issues through aggregation facilitated by the Kendall distribution function.

The model, which we call hierarchical Kendall copula, has previously been mentioned by Anjos and Kolev (2005), who however do not further develop the model in terms of statistical properties and inference. The work presented here is completely independent of theirs and develops properties and inference techniques including sampling algorithms for hierarchical Kendall copulas, with particular focus on Archimedean cluster components.

The features of hierarchical dependence models and of hierarchical Kendall copulas in particular are attractive to different areas of applications. In finance and insurance, risk capital needs to be aggregated over different levels of business lines and operating entities, which introduces a natural hierarchy with different dependencies across levels. Also in other financial areas, there is a need for such models. For instance, hierarchical Archimedean copulas have previously been used by Hofert and Scherer (2011) for the pricing of CDOs. For the purpose of market risk portfolio management, a substantial 30-dimensional application to German stock returns is presented in this paper, showing the need for careful dependence modeling and a good out-of- as well as in-sample performance of hierarchical Kendall copulas.

The model is however not limited to applications in finance and insurance, but may be used in any area that deals with some kind of clustered data such as geographic or temporal clusters. For instance, in hydrology Kendall distribution functions are used to characterize multivariate return periods (see Salvadori et al. (2011)) and hierarchical Kendall copulas may be used to relate different return periods to each other.

The remainder of the paper is organized as follows. The new model is introduced and discussed in Section 2. Section 3 then treats statistical inference techniques for hierarchical Kendall copulas, while the financial application is presented in Section 4. Section 5 concludes.

2 Hierarchical Kendall copulas

A central part of the definition of hierarchical Kendall copulas, which will be given below, is the notion of the Kendall distribution function, which is therefore treated first. Subsequently, hierarchical Kendall copulas are defined and their properties are discussed, in particular in comparison to hierarchical Archimedean copulas.

2.1 Kendall distribution functions

Kendall distribution functions were first studied in two dimensions by Genest and Rivest (1993) and studied in more generality by Barbe et al. (1996). For $\mathbf{U} := (U_1, \dots, U_d)' \sim C$, where C is a d -dimensional copula, the Kendall distribution function $K^{(d)}$ is defined as

$$K^{(d)}(t) := P(C(\mathbf{U}) \leq t), \quad t \in [0, 1]. \quad (2.1)$$

It holds that $t \leq K^{(d)}(t) \leq 1$ for $t \in [0, 1]$ as well as $K^{(d)}(0-) = 0$.

In this paper, it is assumed that copulas are absolutely continuous with continuous Kendall distribution functions. Because the Kendall distribution function is the univariate distribution function of the random variable $Z := C(\mathbf{U})$, we then have $K^{(d)}(Z) \sim U(0, 1)$, which is the multivariate probability integral transform of \mathbf{U} . Another interpretation is that $K^{(d)}$ describes the distribution of the level sets of a copula

$$L_C(z) = \{\mathbf{u} \in [0, 1]^d : C(\mathbf{u}) = z\}, \quad z \in (0, 1). \quad (2.2)$$

It has been shown by Genest and Rivest (1993) that bivariate Archimedean copulas are uniquely characterized by their Kendall distribution functions. Genest et al. (2011) recently extended this result to the trivariate case and strongly conjecture that this holds in general.

The computation of the Kendall distribution function for a given copula is however complicated in general. Imlahi et al. (1999) provide the recursive formula

$$K^{(d)}(t) = K^{(d-1)}(t) + \int_t^1 \int_{C_{u_1}^{-1}(t)}^1 \dots \int_{C_{u_1, \dots, u_{d-2}}^{-1}(t)}^1 \int_0^{C_{u_1, \dots, u_{d-1}}^{-1}(t)} c(u_1, \dots, u_d) du_d \dots du_1, \quad (2.3)$$

where $K^{(d)}$ denotes the Kendall distribution function of the d -dimensional copula C with density c and $K^{(d-1)}$ that of the $(d-1)$ -dimensional margin of the first $d-1$ variables. Furthermore, the formula uses the notion of the copula quantile function, as studied by Imlahi et al. (1999) and Chakak and Ezzerg (2000). Define $C_{u_1, \dots, u_{d-1}}(\cdot) := C(u_1, \dots, u_{d-1}, \cdot)$, then the copula quantile function is the inverse $C_{u_1, \dots, u_{d-1}}^{-1}$. It holds that

$$C(u_1, \dots, u_{d-1}, C_{u_1, \dots, u_{d-1}}^{-1}(z)) = z$$

for $z \in (0, 1)$. For ease of notation, we define $C_{u_1, \dots, u_r}(\cdot) := C(u_1, \dots, u_r, \cdot, 1, \dots, 1)$ for $r = 1, \dots, d-2$, and $C_{\emptyset}^{-1}(z) := z$ for $z \in (0, 1)$.

Equation (2.3) requires high-dimensional integration and availability of the copula quantile function in closed form. For general copulas, it is therefore not possible to easily determine the Kendall distribution function in closed form. A convenient exception are however Archimedean copulas. For a d -dimensional Archimedean copula (see McNeil and Nešlehová (2009)) with generator¹ φ ,

$$C(u_1, \dots, u_d) = \varphi^{-1}(\varphi(u_1) + \dots + \varphi(u_d)), \quad (2.4)$$

Barbe et al. (1996) showed that the Kendall distribution function is given in terms of the generator φ and higher order derivatives of its inverse φ^{-1} as

$$K^{(d)}(t) = t + \sum_{i=1}^{d-1} \frac{(-1)^i}{i!} \varphi(t)^i (\varphi^{-1})^{(i)}(\varphi(t)), \quad (2.5)$$

where $t \in (0, 1]$.

Example 2.1 (Kendall distribution function). *Figure 1 shows Kendall distribution functions $K^{(d)}$ (2.5) of the independence copula ($\varphi(t) = -\log(t)$) and the Gumbel copula ($\varphi(t) = (-\log(t))^\theta$, $\theta \geq 1$) with parameter $\theta = 2$ (corresponding to a Kendall's τ of 0.5) for different dimensions d . Both Kendall distribution functions are, for fixed t , increasing with the dimension d . The rate of increase in the case of the Gumbel copula with medium positive dependence is however much weaker. The practical implications of this property on our model will be discussed in Section 3.3.*

In the following, the superscript of the Kendall distribution function, which indicates the dimension of the associated random vector, will be omitted for reasons of readability.

¹The continuous and strictly decreasing function $\varphi : [0, 1] \rightarrow [0, \infty)$ with $\varphi(1) = 0$ generates a d -dimensional Archimedean copula if and only if its inverse φ^{-1} is d -monotone on $[0, \infty)$, that is, φ^{-1} is differentiable up to the order $d-2$ on $[0, \infty)$, it holds that $(-1)^k (\varphi^{-1})^{(k)}(x) \geq 0$ for $k = 0, 1, \dots, d-2$ and for any $x \in [0, \infty)$, and $(-1)^{d-2} (\varphi^{-1})^{(d-2)}$ is non-increasing and convex on $[0, \infty)$ (see McNeil and Nešlehová (2009)).

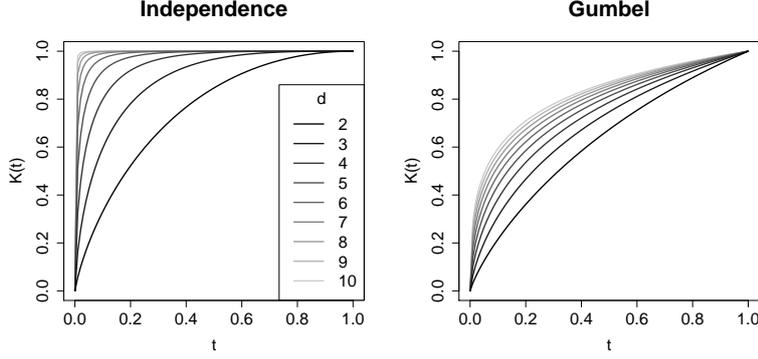


Figure 1: Kendall distribution functions (2.5) of the independence copula (left panel) and the Gumbel copula with parameter $\theta = 2$ (right panel) for different dimensions d .

2.2 Model formulation and properties

The constructive definition of the new dependence model class of hierarchical Kendall copulas is now stated. Although the model has previously been formulated by Anjos and Kolev (2005), it has—to the best of our knowledge—not yet been treated in detail or used for statistical inference. We choose the name “hierarchical Kendall copula” to stress the central role the Kendall distribution function plays in the model formulation.

Definition 2.2 (Hierarchical Kendall copula). *Let $U_1, \dots, U_n \sim U(0, 1)$ and let C_0, C_1, \dots, C_d be copulas of dimensions d, n_1, \dots, n_d , respectively, where $n_i \geq 1$, $i = 1, \dots, d$, and $n = \sum_{i=1}^d n_i$. Further, let K_1, \dots, K_d denote the Kendall distribution functions corresponding to C_1, \dots, C_d , respectively. We define $m_i = \sum_{j=1}^i n_j$ for $i = 1, \dots, d$, and $m_0 = 0$ as well as $\mathbf{U}_i := (U_{m_{i-1}+1}, \dots, U_{m_i})'$ and $V_i := K_i(C_i(\mathbf{U}_i))$ for $i = 1, \dots, d$. Under the assumptions that*

\mathcal{A}_1 : $\mathbf{U}_1, \dots, \mathbf{U}_d$ are mutually independent conditionally on $(V_1, \dots, V_d)'$, and

\mathcal{A}_2 : the conditional distribution of $\mathbf{U}_i | (V_1, \dots, V_d)'$ is the same as the conditional distribution of $\mathbf{U}_i | V_i$ for all $i = 1, \dots, d$, that is, $F_{\mathbf{U}_i | V_1, \dots, V_d} = F_{\mathbf{U}_i | V_i} \forall i \in \{1, \dots, d\}$,

the random vector $(U_1, \dots, U_n)'$ is said to be distributed according to the hierarchical Kendall copula $C_{\mathcal{K}}$ with nesting copula C_0 and cluster copulas C_1, \dots, C_d if

(i) $\mathbf{U}_i \sim C_i \forall i \in \{1, \dots, d\}$,

(ii) $(V_1, \dots, V_d)' \sim C_0$.

The distribution function $C_{\mathcal{K}}$ of $(U_1, \dots, U_n)'$ will be characterized in terms of its density below. First, we discuss the construction, which is illustrated in Figure 2, in more detail and provide examples.

The intuition behind the two assumptions \mathcal{A}_1 and \mathcal{A}_2 is that, given the information of the nesting variables V_1, \dots, V_d , the clusters $\mathbf{U}_1, \dots, \mathbf{U}_d$ are independent of each other and also of other nesting variables, since the dependence among the clusters is explained through the “representatives” V_1, \dots, V_d . In other words, V_1, \dots, V_d can be interpreted as unobserved factors, whose joint behavior determines the dependence of the different clusters. In finance, such factors may be, e.g., industry sectors.

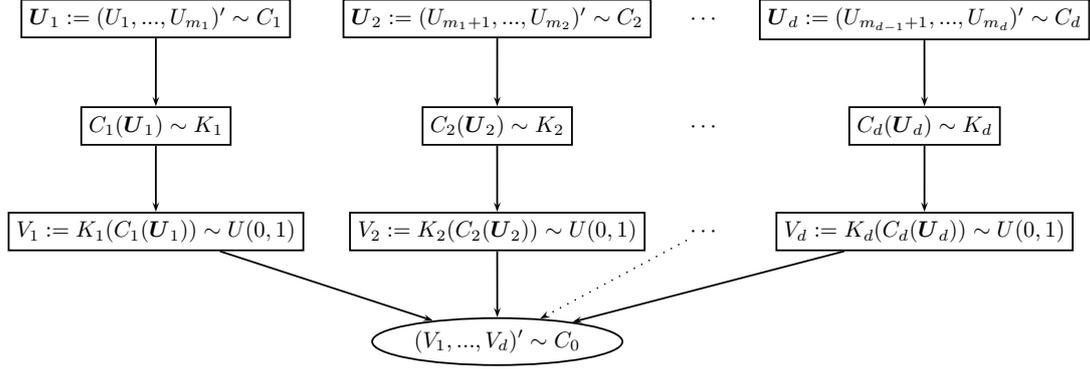


Figure 2: Illustration of a hierarchical Kendall copula (see Definition 2.2).

Note that C_0 is in general not the copula of $(U_1, \dots, U_n)'$ but of $(V_1, \dots, V_d)'$, which are uniform random variables due to $C_i(\mathbf{U}_i) \sim K_i$ for all $i = 1, \dots, d$. The nesting copula C_0 and the cluster copulas C_1, \dots, C_d can be chosen independently. They can be arbitrary copulas such as common Archimedean or elliptical copulas or from any other class of copulas. A special case of hierarchical Kendall copulas is the upper Fréchet-Hoeffding bound.

Example 2.3 (Upper Fréchet-Hoeffding bound). *Let $C_{\mathcal{K}}$ be a hierarchical Kendall copula, where the clusters are perfectly positively dependent and the nesting copula also is the upper Fréchet-Hoeffding bound $C_0(v_1, \dots, v_d) = M^{(d)}(v_1, \dots, v_d) := \min\{v_1, \dots, v_d\}$. Since $C_i = M^{(n_i)}$ and $K_i(t) = t$ for all $i = 1, \dots, d$ and $t \in [0, 1]$, it holds that $V_i = \min\{U_{m_{i-1}+1}, \dots, U_{m_i}\}$ and therefore $C_{\mathcal{K}} = M^{(n)}$. In other words, the upper Fréchet-Hoeffding bound belongs to the class of hierarchical Kendall copulas.*

The Kendall distribution functions K_1, \dots, K_d are used to summarize the information contained in the clusters (transformation of n_i -variate to univariate random vector). While one may also think of other transformations, we believe that Kendall distribution functions are particularly useful for this purpose, since they have a range of properties that are reasonable to demand of an aggregating function. The properties mainly follow from the definitions of a copula and the Kendall distribution function (2.1).

Proposition 2.4 (Properties of aggregation with Kendall distribution functions). *Let K be the Kendall distribution function of a p -dimensional copula C . Then the aggregation function $K \circ C : [0, 1]^p \rightarrow [0, 1]$, $\mathbf{u} \mapsto K(C(\mathbf{u}))$ has the following properties:*

(i) *The function $K \circ C$ is monotone increasing, that is,*

$$K \circ C(u_1, \dots, u_p) \leq K \circ C(v_1, \dots, v_p)$$

for every $u_1, \dots, u_p, v_1, \dots, v_p \in [0, 1]$ such that $u_i \leq v_i$, $i = 1, \dots, p$.

(ii) *If $K(0) = 0$, then $K \circ C$ is grounded, that is, $K \circ C(u_1, \dots, u_p) = 0$ if any $u_i = 0$, $i \in \{1, \dots, p\}$, in particular $K \circ C(0, \dots, 0) = 0$.*

(iii) *It holds that $K \circ C(1, \dots, 1) = 1$.*

(iv) *In the degenerate case that $p = 1$, the function $K \circ C$ reduces to the identity function, that is, $K \circ C(u_1) = u_1$ for $u_1 \in [0, 1]$.*

(v) If C is exchangeable, then $K \circ C$ is commutative, that is,

$$K \circ C(u_1, \dots, u_p) = K \circ C(u_{\pi(1)}, \dots, u_{\pi(p)})$$

for every $u_1, \dots, u_p \in [0, 1]$ and any permutation $\pi : \{1, \dots, p\} \rightarrow \{1, \dots, p\}$.

The condition $K(0) = 0$ in the second property is fulfilled in particular by all absolutely continuous copulas. Examples of exchangeable copulas are Archimedean (2.4) and elliptical copulas with equi-correlation matrix.

The nesting copula C_0 then essentially models the comovement of the cluster copula level sets (2.2). This can be seen as a proxy for the strength of dependence in the clusters, since the dimensionality of the single clusters is “normalized” through the Kendall distribution functions. Other variables such as components of an elliptical distribution do not contain such specific information that summarizes the information in a single variable. Thus, the transformation using the Kendall distribution function reasonably summarizes the (dependence) information of a multivariate random vector in the spirit of the univariate probability integral transform and with the appealing characteristics stated in Proposition 2.4. For the particular purpose of risk aggregation, an alternative hierarchical dependence model for sums of random variables was recently proposed by Arbenz et al. (2012).

In the case of Archimedean clusters, we can give a particularly convenient representation of the multivariate distribution of U_1, \dots, U_n .

Remark 2.5 (Hierarchical Kendall copula with Archimedean clusters). *Let $\mathbf{U} := (U_1, \dots, U_n)$ be distributed according to a hierarchical Kendall copula $C_{\mathcal{K}}$, where C_1, \dots, C_d are Archimedean with generators $\varphi_1, \dots, \varphi_d$, respectively. According to McNeil and Nešlehová (2009), it holds for all $i = 1, \dots, d$,*

$$(\varphi_i(U_{m_{i-1}+1}), \dots, \varphi_i(U_{m_i}))' \stackrel{d}{=} R_i \mathbf{S}^{(i)}, \quad (2.6)$$

where $\mathbf{S}^{(i)} = (S_1^{(i)}, \dots, S_{n_i}^{(i)})'$ is uniformly distributed on the unit simplex $\{\mathbf{x} \geq 0 : \sum_{j=1}^{n_i} x_j = 1\} \subset [0, 1]^{n_i}$, and the radial part $R_i = \sum_{j=1}^{n_i} \varphi_i(U_{m_{i-1}+j})$ is independent of $\mathbf{S}^{(i)}$ and has distribution F_{R_i} , which can be determined through the inverse Williamson transform of φ_i^{-1} .

As a result we can represent the random vector \mathbf{U} as

$$(U_1, \dots, U_n)' \stackrel{d}{=} (\varphi_1^{-1}(R_1 S_1^{(1)}), \dots, \varphi_1^{-1}(R_1 S_{n_1}^{(1)}), \varphi_2^{-1}(R_2 S_1^{(2)}), \dots, \varphi_2^{-1}(R_2 S_{n_2}^{(2)}), \dots, \varphi_d^{-1}(R_d S_{n_d}^{(d)}))', \quad (2.7)$$

where $R_i = \varphi_i(K_i^{-1}(V_i))$ for $i = 1, \dots, d$, since by the definition of C_i , V_i and R_i ,

$$V_i = K_i(C_i(\mathbf{U}_i)) = K_i\left(\varphi_i^{-1}\left(\sum_{j=1}^{n_i} \varphi_i(U_{m_{i-1}+j})\right)\right) = K_i(\varphi_i^{-1}(R_i)).$$

In other words, if all clusters are Archimedean, dependence among clusters is introduced solely through the dependence between the radial variables of the different clusters. In particular, if the nesting copula C_0 is also Archimedean with generator φ_0 and corresponding radial variable R_0 , we have for $i = 1, \dots, d$ that $V_i = \varphi_0^{-1}(R_0 S_i^{(0)})$, where $S_1^{(0)}, \dots, S_d^{(0)}$ are uniformly distributed on the d -dimensional unit simplex. That is, the radial variables of the clusters, R_i , can be expressed through R_0 and uniform random variables on the simplex.

Equation (2.7) also motivates to speak of a “grouped Archimedean copula” similar to the grouped Student’s t copula by Daul et al. (2003).

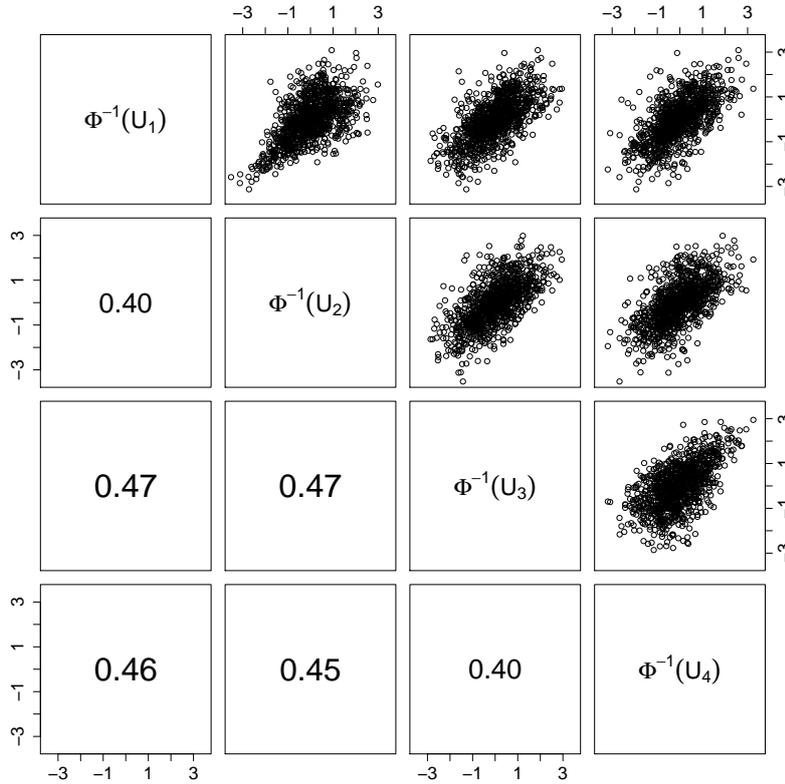


Figure 3: A sample of size 1000 from a four-dimensional hierarchical Kendall copula with Clayton and Gumbel clusters and Frank nesting copula. Margins are transformed to standard normal. The lower triangle shows pairwise empirical Kendall's τ values.

We now provide an illustrative example of a hierarchical Kendall copula with Archimedean clusters.

Example 2.6 (Hierarchical Kendall copula with Archimedean clusters). *Let $C_{\mathcal{K}}$ be a four-dimensional hierarchical Kendall copula with $n_1 = n_2 = 2$. The bivariate cluster copulas are chosen as Clayton (Archimedean with $\varphi(t) = t^{-\theta} - 1$, $\theta > 0$) with parameter $\theta = 1.33$ and Gumbel with parameter $\theta = 1.67$ (both corresponding to a Kendall's τ of 0.4). The nesting copula is set as a Frank (also Archimedean with $\varphi(t) = -\log((1 - e^{-\theta t})/(1 - e^{-\theta}))$, $\theta \in \mathbb{R} \setminus \{0\}$) with parameter $\theta = 11.41$ (Kendall's τ of 0.7). A sample of size 1000 from this hierarchical Kendall copula is shown in Figure 3. It shows the typical features of lower tail dependence for the pair $(U_1, U_2)'$ (Clayton copula) and of upper tail dependence for the pair $(U_3, U_4)'$ (Gumbel copula). The between-cluster dependence looks rather tail-symmetric as implied by the Frank copula.*

The illustration in Example 2.6 also provides an example where the between-cluster dependence is stronger than the within-cluster dependence. This case cannot be modeled using hierarchical Archimedean copulas as will be discussed in Section 2.3.

The two-level construction given in Definition 2.2 can also be extended to an arbitrary number of levels.

Remark 2.7 (Hierarchical Kendall copula with k levels). Let $U_1, \dots, U_n \sim U(0, 1)$ and let d_j , $j = 1, \dots, k-1$, denote the number of clusters per level j , such that $d_1 \geq d_2 \geq \dots \geq d_{k-1}$. Further, let the nesting copula C_0 be d_{k-1} -dimensional and let the nested cluster copulas $C_i^{(j)}$, $j = 1, \dots, k-1$, $i = 1, \dots, d_j$, be of dimension $n_i^{(j)} \geq 1$, where $n = \sum_{i=1}^{d_1} n_i^{(1)}$ and $d_{j-1} = \sum_{i=1}^{d_j} n_i^{(j)}$ for $j = 2, \dots, k-1$. The index i runs over the d_j clusters of each level j . The Kendall distribution function corresponding to $C_i^{(j)}$, $j = 1, \dots, k-1$, $i = 1, \dots, d_j$, is denoted by $K_i^{(j)}$ and we define $m_i^{(j)} = \sum_{\ell=1}^i n_\ell^{(j)}$ for $i = 1, \dots, d_j$, and $m_0^{(j)} = 0$. Under independence assumptions as in Definition 2.2, we then say that the random vector $(U_1, \dots, U_n)'$ is distributed according to the k -level hierarchical Kendall copula C_K with nesting copula C_0 and cluster copulas $C_i^{(j)}$, $j = 1, \dots, k-1$, $i = 1, \dots, d_j$, if

$$(i) \mathbf{U}_i := (U_{m_{i-1}^{(1)}+1}, \dots, U_{m_i^{(1)}})' \sim C_i^{(1)} \quad \forall i \in \{1, \dots, d_1\},$$

$$(ii) V_i^{(1)} := K_i^{(1)}(C_i^{(1)}(\mathbf{U}_i)) \quad \forall i \in \{1, \dots, d_1\},$$

(iii) for $j = 2, \dots, k-1$:

$$(a) \mathbf{V}_i^{(j-1)} := (V_{m_{i-1}^{(j)}+1}^{(j-1)}, \dots, V_{m_i^{(j)}}^{(j-1)})' \sim C_i^{(j)} \quad \forall i \in \{1, \dots, d_j\},$$

$$(b) V_i^{(j)} := K_i^{(j)}(C_i^{(j)}(\mathbf{V}_i^{(j-1)})) \quad \forall i \in \{1, \dots, d_j\},$$

$$(iv) (V_1^{(k-1)}, \dots, V_{d_{k-1}}^{(k-1)})' \sim C_0.$$

In particular, the clusters $\mathbf{U}_1, \dots, \mathbf{U}_{d_1}$ at the lowest level ($j = 1$) are assumed to be mutually independent given the “representatives” $V_i^{(j)}$, $j = 1, \dots, k-1$, $i = 1, \dots, d_j$. In the case of $k = 3$, we define

$$\mathbf{U}_i^{(1)} := (\mathbf{U}'_{m_{i-1}^{(2)}+1}, \dots, \mathbf{U}'_{m_i^{(2)}})', \quad i = 1, \dots, d_2.$$

Then the assumptions are

$\mathcal{A}_1^{(1)}$: $\mathbf{U}_1^{(1)}, \dots, \mathbf{U}_{d_2}^{(1)}$ are mutually independent conditionally on $(V_1^{(2)}, \dots, V_{d_2}^{(2)})'$;

$\mathcal{A}_2^{(1)}$: the conditional distribution of $\mathbf{U}_i^{(1)} | (V_1^{(2)}, \dots, V_{d_2}^{(2)})'$ is the same as the conditional distribution of $\mathbf{U}_i^{(1)} | V_i^{(2)}$ for all $i = 1, \dots, d_2$;

$\mathcal{A}_1^{(2)}$: $\mathbf{U}_{m_{i-1}^{(2)}+1}^{(2)}, \dots, \mathbf{U}_{m_i^{(2)}}^{(2)}$ are mutually independent conditionally on $(V_i^{(2)}, \mathbf{V}_i^{(1)})'$ for all $i = 1, \dots, d_2$;

$\mathcal{A}_2^{(2)}$: the conditional distribution of $\mathbf{U}_{m_{i-1}^{(2)}+j}^{(2)} | (V_i^{(2)}, \mathbf{V}_i^{(1)})'$ is the same as the conditional distribution of $\mathbf{U}_{m_{i-1}^{(2)}+j}^{(2)} | (V_i^{(2)}, V_{m_{i-1}^{(2)}+j}^{(1)})'$ for all $i = 1, \dots, d_2$, $j = 1, \dots, n_i^{(2)}$.

Their interpretation is essentially the same as for the assumptions \mathcal{A}_1 and \mathcal{A}_2 stated in Definition 2.2.

An example of a three-level hierarchical Kendall copula is shown in Figure 4. For simplicity and illustrative reasons, we restrict our exposition here to the case of $k = 2$ hierarchical levels. It will be sketched how to generalize all derivations and methods, which are described in the following, to the general k level case.

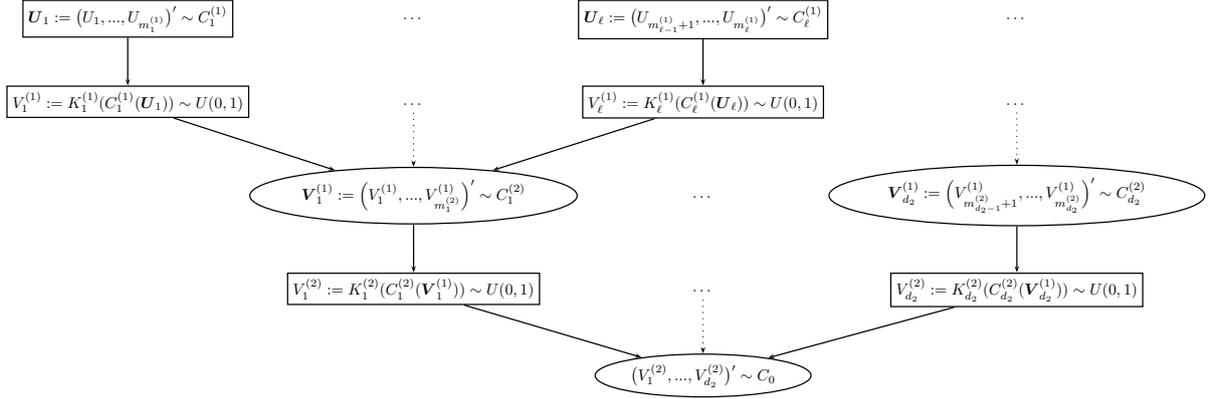


Figure 4: Illustration of a three-level hierarchical Kendall copula (see Remark 2.7). Here $m_1^{(2)} = \ell$.

The two independence assumptions \mathcal{A}_1 and \mathcal{A}_2 of Definition 2.2 provide a natural structure for hierarchical dependence models and allow to derive the joint density function of a hierarchical Kendall copula as stated in the following theorem. Densities of the copulas C_0, \dots, C_d are denoted by c_0, \dots, c_d , respectively.

Theorem 2.8 (Joint density of a hierarchical Kendall copula). *Let $\mathbf{U} = (U_1, \dots, U_n)'$ be distributed according to a hierarchical Kendall copula $C_{\mathcal{K}}$ with nesting copula C_0 and cluster copulas C_1, \dots, C_d . The joint density function $c_{\mathcal{K}}$ of $C_{\mathcal{K}}$ is then given as follows:*

$$c_{\mathcal{K}}(\mathbf{u}) = c_0(K_1(C_1(\mathbf{u}_1)), \dots, K_d(C_d(\mathbf{u}_d))) \prod_{i=1}^d c_i(\mathbf{u}_i), \quad (2.8)$$

where $\mathbf{u} = (u_1, \dots, u_n)'$ and $\mathbf{u}_i = (u_{m_{i-1}+1}, \dots, u_{m_i})'$, $i = 1, \dots, d$.

Proof: It is

$$C_{\mathcal{K}}(\mathbf{u}) = P(\mathbf{U} \leq \mathbf{u}) = \int_{[0,1]^d} P(\mathbf{U} \leq \mathbf{u} | V_1 = v_1, \dots, V_d = v_d) c_0(v_1, \dots, v_d) dv_1 \dots dv_d. \quad (2.9)$$

By applying assumption \mathcal{A}_1 first and then assumption \mathcal{A}_2 , we obtain

$$\begin{aligned} P(\mathbf{U} \leq \mathbf{u} | V_1 = v_1, \dots, V_d = v_d) &= \prod_{i=1}^d P(\mathbf{U}_i \leq \mathbf{u}_i | V_1 = v_1, \dots, V_d = v_d) \\ &= \prod_{i=1}^d P(\mathbf{U}_i \leq \mathbf{u}_i | V_i = v_i). \end{aligned}$$

Using this result, Equation (2.9) simplifies to

$$C_{\mathcal{K}}(\mathbf{u}) = \int_{[0,1]^d} \left[\prod_{i=1}^d F_{\mathbf{U}_i | V_i}(\mathbf{u}_i | v_i) \right] c_0(v_1, \dots, v_d) dv_1 \dots dv_d. \quad (2.10)$$

Further, we denote by $\mathbf{U}_{i,-m_i}$, $i \in \{1, \dots, d\}$, the $(n_i - 1)$ -dimensional sub-vector of $\mathbf{U}_i = (U_{m_{i-1}+1}, \dots, U_{m_i})'$ with element U_{m_i} removed. Since $V_i = K_i(C_i(\mathbf{U}_i)) \sim U(0,1)$ and according to a change of variables, it then holds that

$$\begin{aligned} f_{\mathbf{U}_{i,-m_i} | V_i}(\mathbf{u}_{i,-m_i} | v_i) &= f_{\mathbf{U}_{i,-m_i}, V_i}(\mathbf{u}_{i,-m_i}, v_i) \\ &= c_i(\mathbf{u}_{i,-m_i}, C_{i; \mathbf{u}_{i,-m_i}}^{-1}(K_i^{-1}(v_i))) \frac{\partial}{\partial v_i} C_{i; \mathbf{u}_{i,-m_i}}^{-1}(K_i^{-1}(v_i)), \end{aligned} \quad (2.11)$$

if $v_i \leq K_i(C_i(\mathbf{u}_{i,-m_i}, 1))$. This yields

$$F_{U_i|V_i}(\mathbf{u}_i|v_i) = \int_0^{u_{m_i-1}} \dots \int_0^{u_{m_i-1}+1} f_{U_{i,-m_i}|V_i}(\mathbf{w}_{i,-m_i}|v_i) 1_{\{C_{i;\mathbf{w}_{i,-m_i}}^{-1}(K_i^{-1}(v_i)) \leq u_{m_i}\}} d\mathbf{w}_{i,-m_i}.$$

Plugging this expression into Equation (2.10) and substituting v_i by $w_{m_i} = C_{i;\mathbf{w}_{i,-m_i}}^{-1}(K_i^{-1}(v_i))$ for $i = 1, \dots, d$ then leads to

$$C_{\mathcal{K}}(\mathbf{u}) = \int_0^{u_n} \dots \int_0^{u_1} \left[\prod_{i=1}^d c_i(\mathbf{w}_i) \right] c_0(K_1(C_1(\mathbf{w}_1)), \dots, K_d(C_d(\mathbf{w}_d))) d\mathbf{w},$$

where we used that $v_i = K_i(C_i(\mathbf{w}_i))$, $i = 1, \dots, d$. Taking derivatives with respect to u_1, \dots, u_n therefore gives the desired result. \square

Remark 2.9 (Joint density of a k -level hierarchical Kendall copula). *The arguments of Theorem 2.8 can be iterated to derive the joint density of a k -level hierarchical Kendall copula. By first conditioning on the aggregated variables of level $k-1$, $V_1^{(k-1)}, \dots, V_{d_{k-1}}^{(k-1)}$, then on those of level $k-2$ up to level 1, an expression similar to Equation (2.8) is obtained.*

For instance, the density of the three-level hierarchical Kendall copula can be derived along the lines of the proof of Theorem 2.8 as

$$\begin{aligned} c_{\mathcal{K}}(\mathbf{u}) &= c_0(v_1^{(2)}, \dots, v_{n_2}^{(2)}) \prod_{i=1}^{d_2} \left(c_i^{(2)}(\mathbf{v}_i^{(1)}) \prod_{j=1}^{n_i^{(2)}} c_{m_{i-1}^{(2)}+j}^{(1)}(\mathbf{u}_{m_{i-1}^{(2)}+j}) \right) \\ &= c_0(v_1^{(2)}, \dots, v_{n_2}^{(2)}) \prod_{i=1}^{d_2} c_i^{(2)}(\mathbf{v}_i^{(1)}) \prod_{i=1}^{d_1} c_i^{(1)}(\mathbf{u}_i), \end{aligned} \quad (2.12)$$

where $\mathbf{v}_i^{(1)} = (v_{m_{i-1}^{(2)}+1}^{(1)}, \dots, v_{m_i^{(2)}}^{(1)})'$, $i = 1, \dots, d_2$, with components $v_i^{(1)} = K_i^{(1)}(C_i^{(1)}(\mathbf{u}_i))$, $i = 1, \dots, d_1$. Further, $v_i^{(2)} = K_i^{(2)}(C_i^{(2)}(\mathbf{v}_i^{(1)}))$, $i = 1, \dots, d_2$.

This means that the density of a three-level hierarchical Kendall copula also conveniently decomposes into the product of copula densities, where the arguments are obtained through the repeated application of Kendall distribution functions.

Another important special case of hierarchical Kendall copulas can easily be stated using Theorem 2.8.

Example 2.10 (Independence copula). *Let $C_{\mathcal{K}}$ be a hierarchical Kendall copula, where both cluster and nesting copulas are independence copulas. Since the independence copula has density equal to 1, it follows that $c_{\mathcal{K}}(\mathbf{u}) = 1$. This means that the independence copula also belongs to the class of hierarchical Kendall copulas.*

Theorem 2.8 also allows to formulate the following corollary which summarizes the marginal properties of hierarchical Kendall copulas.

Corollary 2.11 (Margins of a hierarchical Kendall copula). *The same notation as in Theorem 2.8 is used.*

- (i) *Bivariate margins: Let $k, \ell \in \{1, \dots, n\}$, $k \neq \ell$. W.l.o.g. $k < \ell$.*

(a) If U_k and U_ℓ are in the same cluster i , their marginal distribution function $C_{\mathcal{K},k\ell}$ is

$$C_{\mathcal{K},k\ell}(u_k, u_\ell) := C_i(1, \dots, 1, u_k, 1, \dots, 1, u_\ell, 1, \dots, 1).$$

(b) If U_k and U_ℓ are in different clusters i and j , respectively, their marginal distribution function $C_{\mathcal{K},k\ell}$ is

$$C_{\mathcal{K},k\ell}(u_k, u_\ell) := \int_0^{u_k} \int_0^{u_\ell} \int_{[0,1]^{n_i+n_j-2}} c_{0,ij}(K_i(C_i(\mathbf{w}_i)), K_j(C_j(\mathbf{w}_j))) \times c_i(\mathbf{w}_i) c_j(\mathbf{w}_j) d\mathbf{w}_{i,-k} d\mathbf{w}_{j,-\ell} dw_\ell dw_k, \quad (2.13)$$

where $c_{0,ij}$ is the density of the bivariate (i, j) -margin of C_0 .

(ii) *Multivariate margins: The marginal distribution of the cluster U_i is C_i .*

More general multivariate margins involving variables from different clusters can be derived as in Equation (2.13).

Remark 2.12 (Mixture representation). *As a consequence of Corollary 2.11 (i)(b), bivariate marginal distributions where the variables are in different clusters can be regarded as a kind of continuous mixture of the nesting copula C_0 . The density of $C_{\mathcal{K},k\ell}$ as defined above is given by*

$$c_{\mathcal{K},k\ell}(u_k, u_\ell) = \int_{[0,1]^{n_i+n_j-2}} c_{0,ij}(K_i(C_i(\mathbf{u}_i)), K_j(C_j(\mathbf{u}_j))) c_i(\mathbf{u}_i) c_j(\mathbf{u}_j) d\mathbf{u}_{i,-k} d\mathbf{u}_{j,-\ell},$$

where the mixing density weights are given by the product $c_i(\mathbf{u}_i)c_j(\mathbf{u}_j)$.

This representation complements the results of Remark 2.5. Equation (2.7) shows that hierarchical Kendall copulas with Archimedean cluster copulas can be represented as transformed mixtures of uniform distributions on unit simplices.

As noted above, the multivariate distribution of U_1, \dots, U_n defined through a hierarchical Kendall copula is in general not the copula C_0 but given through its density (see Equation (2.9)). We showed that the important special cases of independence as well as of comonotonicity are hierarchical Kendall copulas (see Examples 2.3 and 2.10), while in general dependence between clusters ranges between these cases and can also be negative. It is yet an open question which other common multivariate distributions can be represented as hierarchical Kendall copulas with non-trivial cluster sizes (at least one $n_i > 1$, that is, $d < n$). In particular, hierarchical Archimedean copulas are different from hierarchical Kendall copulas—with positive and negative implications as discussed next.

2.3 Comparison with hierarchical Archimedean copulas

The popular class of hierarchical Archimedean copulas also allows for a nested modeling of clusters of variables (see, e.g., Joe (1997), Hofert (2010), Savu and Tiede (2010) and Okhrin et al. (2013)). In contrast to hierarchical Kendall copulas, hierarchical Archimedean copulas are however limited to Archimedean copulas as building blocks, while hierarchical Kendall copulas can be built up by any possible copula. Furthermore, hierarchical Archimedean copulas require stronger within-cluster dependence, which results in parameter restrictions if generators are the same. This is not the case for hierarchical Kendall copulas (see Example 2.6).

Archimedean copulas have many useful properties. For instance, their relationship to Laplace transforms can be used to conveniently sample from hierarchical Archimedean copulas as described, e.g., in Hofert (2010, 2011). Hierarchical Kendall copulas also benefit from many of these properties. It will be shown in Section 3.1 that closed-form sampling of hierarchical Kendall copulas is feasible, when cluster copulas are Archimedean. Hierarchical Kendall copulas with Archimedean cluster copulas are further particularly easy to estimate, since Kendall distribution functions are known in closed form for Archimedean copulas. For that reason they also provide a closed-form joint density function, which is numerically tractable even in higher dimensions (see Equation (2.12) for the case of three hierarchical levels). Although the general density expression of hierarchical Archimedean copulas is in general hardly accessible (see Savu and Trede (2010)), Hofert and Pham (2013) recently derived a tractable formula for the case of a moderate number of nesting levels.

In contrast to hierarchical Archimedean copulas, multivariate margins of hierarchical Kendall copulas are however not directly available (see Corollary 2.11). For instance, if $(U_1, U_2)' \sim C_1$, $(U_3, U_4)' \sim C_2$ and the nesting copula in both hierarchical copula models is C_0 , then in the hierarchical Archimedean copula the margin $(U_1, U_3)'$ is distributed according to C_0 , while in the case of the hierarchical Kendall copula this marginal distribution has to be obtained using integration as in Equation (2.13).

Example 2.13 (Bivariate margin of hierarchical Archimedean and Kendall copulas). *Let U_1, \dots, U_4 be distributed according to a hierarchical Kendall copula or hierarchical Archimedean copula with Gumbel cluster and nesting copulas: as above, $(U_1, U_2)' \sim C_1$, $(U_3, U_4)' \sim C_2$ and C_0 denotes the nesting copula. Figure 5 illustrates the marginal density of $(U_1, U_3)'$ in both cases. It shows contour plots of the marginal density of $(\Phi^{-1}(U_1), \Phi^{-1}(U_3))'$, where Φ^{-1} denotes the inverse standard normal distribution function. Parameters are chosen as $\theta_1 = 3$ (Kendall's τ of 0.67) and $\theta_2 = 4$ (Kendall's τ of 0.75) for the bivariate cluster copulas C_1 and C_2 , respectively, and $\theta_0 = 2$ (Kendall's τ of 0.5) for the nesting copula C_0 . The difference between the distributions is minor. However, the contour plot corresponding to the hierarchical Archimedean copula is slightly sharper in the upper right corner, implying a stronger joint tail behavior.*

Note that in the example the parameters of the cluster copulas are larger than that of the nesting copula, as required for the hierarchical Archimedean copula to yield a valid multivariate distribution ($\theta_0 \leq \min\{\theta_1, \theta_2\}$). This is however not needed for the hierarchical Kendall copula.

Finally, the nesting copula C_0 of a hierarchical Kendall copula is also not closed under addition and removal of cluster components U_ℓ , which is contrary to hierarchical Archimedean copulas. This is because the Kendall distribution function is not independent with respect to the dimension (see Expression 2.3). That is, if a random variable U_{n+1} is added to cluster $i \in \{1, \dots, d\}$, the transformation K_i and thus V_i change, even if C_i is Archimedean; similarly if a random variable is removed from a cluster.

3 Inference for hierarchical Kendall copulas

We now discuss appropriate statistical inference techniques for hierarchical Kendall copulas. While hierarchical Kendall copulas could in principle be built up by any possible copula and the developed methodology holds in general, we focus in the following on the case of Archimedean clusters (see Remark 2.5 and Example 2.6). In this case, analytical derivations are possible due

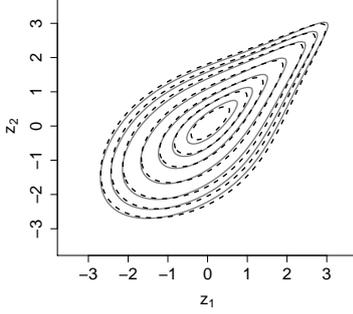


Figure 5: Contour plots of the marginal density of $(\Phi^{-1}(U_1), \Phi^{-1}(U_3))'$, where U_1, \dots, U_4 are distributed according to a hierarchical Kendall copula (black dashed line) or hierarchical Archimedean copula (gray line) with Gumbel cluster and nesting copulas.

to the tractability of the Kendall distribution function (2.5) and the copula itself (2.4). We first treat simulation, then estimation and model selection.

3.1 Simulation

The following general simulation procedure describes how to sample from a given hierarchical Kendall copula with clusters that are not necessarily Archimedean.

Algorithm 3.1 (Simulation of hierarchical Kendall copulas). *Let $C_{\mathcal{K}}$ be a hierarchical Kendall copula with nesting copula C_0 and cluster copulas C_1, \dots, C_d . Further, let K_i^{-1} denotes the inverse of the Kendall distribution function K_i for $i = 1, \dots, d$.*

- (i) *Sample $(v_1, \dots, v_d)'$ from C_0 .*
- (ii) *Set $z_i := K_i^{-1}(v_i) \forall i \in \{1, \dots, d\}$.*
- (iii) *Sample $(u_{m_{i-1}+1}, \dots, u_{m_i})'$ from $(U_{m_{i-1}+1}, \dots, U_{m_i})' | C_i(U_{m_{i-1}+1}, \dots, U_{m_i}) = z_i$ for $i = 1, \dots, d$.*
- (iv) *Return $\mathbf{u} := (u_1, \dots, u_n)'$.*

Remark 3.2 (Simulation of k -level hierarchical Kendall copulas). *The above procedure for two levels can be iterated to simulate from a k -level hierarchical Kendall copula (see Remark 2.7). In the following, samples of random variables will be denoted in the corresponding lower case letters. Let $(v_1^{(k-1)}, \dots, v_{d_{k-1}}^{(k-1)})'$ be sampled from C_0 . Then,*

- (i) *sample from $(V_{m_{i-1}+1}^{(j-1)}, \dots, V_{m_i}^{(j-1)})' | C_i^{(j)}(V_{m_{i-1}+1}^{(j-1)}, \dots, V_{m_i}^{(j-1)}) = (K_i^{(j)})^{-1}(v_i^{(j)})$ for $j = k - 1, \dots, 2$ and $i = 1, \dots, d_j$.*
- (ii) *and finally from $(U_{m_{i-1}+1}^{(1)}, \dots, U_{m_i}^{(1)})' | C_i^{(1)}(U_{m_{i-1}+1}^{(1)}, \dots, U_{m_i}^{(1)}) = (K_i^{(1)})^{-1}(v_i^{(1)})$ for $i = 1, \dots, d_1$.*

Given that we can simulate from the copula C_0 , sampling from hierarchical Kendall copulas thus amounts to the more general question of sampling from a distribution $\mathbf{U} | C(\mathbf{U}) = z$, where C is the copula of a marginally uniform random vector $\mathbf{U} := (U_1, \dots, U_d)'$ and $z \in (0, 1)$. In other

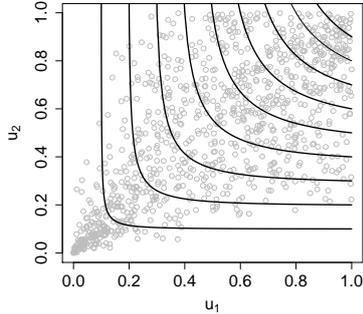


Figure 6: A sample from a bivariate Clayton copula with parameter $\theta = 2$ (Kendall's τ of 0.5) and level sets $L_C(z)$ at levels $z = 0.1, \dots, 0.9$ (solid lines).

words, we want to sample from a multivariate distribution given a specific level set (2.2) at level z as illustrated in Figure 6. This problem is discussed in the following two sections. First, we solve the issue using the conditional inverse method, for which expressions of conditional distribution functions are provided, which are shown to be available in closed form for Archimedean copulas. An alternative solution based on the representation given in Equation (2.6) is provided thereafter. Closed-form sampling procedures for Plackett and Archimax copulas, which include the large class of extreme value copulas, can be found in Brechmann (2013). There, different approximate methods such as rejection sampling for the case that the cluster copulas do not admit a closed-form approach are also discussed and evaluated.

Note that inversion of the Kendall distribution function, as required in Step (ii) of Algorithm 3.1, is numerically feasible for an Archimedean copula, since its Kendall distribution function is known in closed form and can be computed efficiently (see Hofert et al. (2012) for explicit functional expressions of $(\varphi^{-1})^{(i)}$ for common Archimedean generators).

3.1.1 Conditional inverse method

A common method to generate samples from a multivariate distribution is the conditional inverse method (see Devroye (1986)). For this, we need to determine the iterative conditional distribution functions of $\mathbf{U}|C(\mathbf{U}) = z$, that is, of $U_1|C(\mathbf{U}) = z$, $U_2|(U_1 = u_1, C(\mathbf{U}) = z)$, ..., $U_{d-1}|(U_1 = u_1, \dots, U_{d-2} = u_{d-2}, C(\mathbf{U}) = z)$. The corresponding conditional distribution functions are denoted by $F_{j|1, \dots, j-1}(\cdot|u_1, \dots, u_{j-1}, z)$ and densities by $f_{j|1, \dots, j-1}(\cdot|u_1, \dots, u_{j-1}, z)$ for $j = 1, \dots, d-1$, respectively. Then we obtain observations $(u_1, \dots, u_d)'$ from $\mathbf{U}|C(\mathbf{U}) = z$ by generating w_1, \dots, w_{d-1} independently from the uniform distribution and setting $u_j := F_{j|1, \dots, j-1}^{-1}(w_j|u_1, \dots, u_{j-1}, z)$ for $j = 1, \dots, d-1$. For $j = d$ it is $u_d := C_{u_1, \dots, u_{d-1}}^{-1}(z)$. The problem therefore is to determine the conditional distribution functions, which are generally not given in closed form.

Theorem 3.3 (Conditional distributions). *Let $\mathbf{U} \sim C$, then for all $j = 1, \dots, d-1$,*

$$F_{j|1, \dots, j-1}(u|u_1, \dots, u_{j-1}, z) = \frac{\int_{C_{u_1, \dots, u_{j-1}}^{-1}(z)}^u g_j(u_j) du_j}{\int_{C_{u_1, \dots, u_{j-1}}^{-1}(z)}^1 g_j(u_j) du_j}, \quad u \in (C_{u_1, \dots, u_{j-1}}^{-1}(z), 1), \quad (3.1)$$

where

$$g_j(u_j) = \int_{C_{u_1, \dots, u_j}^{-1}(z)}^1 \cdots \int_{C_{u_1, \dots, u_{d-2}}^{-1}(z)}^1 c(u_1, \dots, u_{d-1}, C_{u_1, \dots, u_{d-1}}^{-1}(z)) \frac{\partial}{\partial z} C_{u_1, \dots, u_{d-1}}^{-1}(z) du_{d-1} \dots du_{j+1}. \quad (3.2)$$

Proof: The idea is to derive the conditional density $f_{j|1, \dots, j-1}(\cdot | u_1, \dots, u_{j-1}, z)$ and then integrate to obtain the distribution function. We begin by observing that

$$f_{j|1, \dots, j-1}(u_j | u_1, \dots, u_{j-1}, z) = \frac{f_{U_1, \dots, U_j, C(\mathbf{U})}(u_1, \dots, u_{j-1}, u_j, z)}{f_{U_1, \dots, U_{j-1}, C(\mathbf{U})}(u_1, \dots, u_{j-1}, z)}. \quad (3.3)$$

According to a change of variables as in Equation (2.11), the numerator can then be rewritten as

$$\begin{aligned} & f_{U_1, \dots, U_j, C(\mathbf{U})}(u_1, \dots, u_j, z) \\ &= \int_{C_{u_1, \dots, u_j}^{-1}(z)}^1 \cdots \int_{C_{u_1, \dots, u_{d-2}}^{-1}(z)}^1 f_{U_1, \dots, U_{d-1}, C(\mathbf{U})}(u_1, \dots, u_{d-1}, z) du_{d-1} \dots du_{j+1} \\ &= \int_{C_{u_1, \dots, u_j}^{-1}(z)}^1 \cdots \int_{C_{u_1, \dots, u_{d-2}}^{-1}(z)}^1 c(u_1, \dots, u_{d-1}, C_{u_1, \dots, u_{d-1}}^{-1}(z)) \frac{\partial}{\partial z} C_{u_1, \dots, u_{d-1}}^{-1}(z) du_{d-1} \dots du_{j+1} \\ &= g_j(u_j), \end{aligned}$$

where g_j is defined in Equation (3.2) and dependence on u_1, \dots, u_{j-1}, z is suppressed for ease of notation. Further, the denominator of Equation (3.3) then reads as

$$f_{U_1, \dots, U_{j-1}, C(\mathbf{U})}(u_1, \dots, u_{j-1}, z) = \int_{C_{u_1, \dots, u_{j-1}}^{-1}(z)}^1 g_j(u_j) du_j.$$

By integration, we obtain the expression for the conditional distribution function (3.1). \square

Evidently, the conditional distribution functions given in Equation (3.1) in general do not allow for explicit expressions. In particular, if the copula quantile function C^{-1} is not available in closed form such as for the Gaussian copula, the expression in Equation (3.1) hardly simplifies. In the case of Archimedean copulas, the conditional distribution functions can be obtained as particularly convenient expressions.

Lemma 3.4 (Conditional distributions of Archimedean copulas). *Let $\mathbf{U} \sim C$, where C is a d -dimensional Archimedean copula with generator φ , then for all $j = 1, \dots, d-1$,*

$$F_{j|1, \dots, j-1}(u | u_1, \dots, u_{j-1}, z) = \left(1 - \frac{\varphi(u)}{\varphi(z) - \sum_{1 \leq i < j} \varphi(u_i)} \right)^{d-j}, \quad u \in (C_{u_1, \dots, u_{j-1}}^{-1}(z), 1). \quad (3.4)$$

This result can be derived using the general formula provided in Theorem 3.3. As noted by an anonymous referee, a considerably simpler proof can be formulated by exploiting the representation (2.6) of Archimedean copulas and properties of the Dirichlet distribution. The proof is given in the next section in Remark 3.8 using a characterization result of Proposition 3.6 below.

Lemma 3.4 then allows to use the conditional inverse method for Archimedean copulas, for which the conditional distribution functions can easily be inverted in closed form.

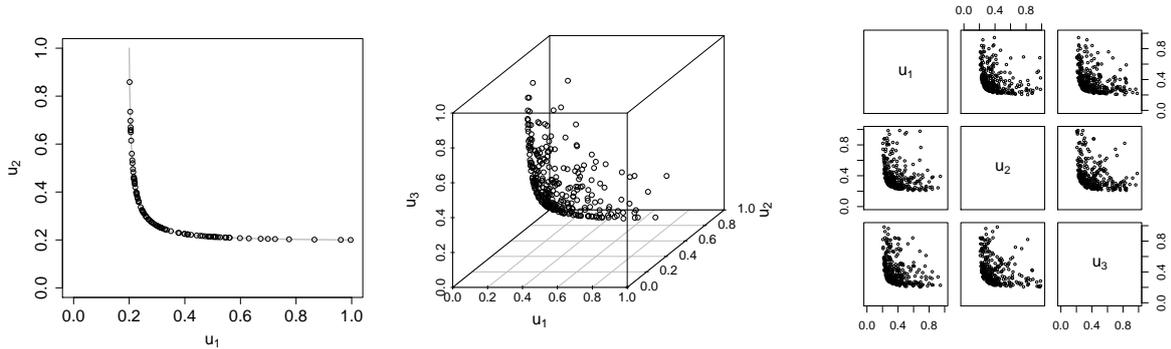


Figure 7: Left panel: scatter plot of a sample of a bivariate Clayton copula with parameter $\theta = 2$ at $z = 0.2$. Middle and right panel: 3D scatter plot and pairwise scatter plots of a sample of a trivariate Clayton copula with parameter $\theta = 2$ at $z = 0.2$.

Algorithm 3.5 (Conditional inverse method for Archimedean copulas). *Let C be an Archimedean copula with generator φ and $z \in (0, 1)$.*

- (i) *Sample w_1, \dots, w_{d-1} independently from the uniform distribution.*
- (ii) *For $j = 1, \dots, d - 1$: $u_j := \varphi^{-1}((1 - w_j^{1/(d-j)})(\varphi(z) - \sum_{1 \leq i < j} \varphi(u_i)))$.*
- (iii) *Set $u_d := \varphi^{-1}(\varphi(z) - \sum_{1 \leq i < d} \varphi(u_i))$.*
- (iv) *Return $(u_1, \dots, u_d)'$.*

For illustration Figure 7 shows scatter plots of samples from bivariate and trivariate Clayton copulas with parameter $\theta = 2$ (Kendall's τ of 0.5).

As a side note, we observe that Algorithm 3.5 can in particular be used to sample from a given Archimedean copula. For this, draw an additional uniform observation w_d and set $z := K^{-1}(w_d)$ prior to performing Steps (ii) and (iii). An equivalent version of this result has previously been stated in Wu et al. (2007).

3.1.2 Projected distribution

Clearly the condition $C(\mathbf{U}) = z$ on the distribution of \mathbf{U} means that we are in fact investigating a $(d-1)$ -dimensional distribution, namely the distribution of \mathbf{U} projected to the $(d-1)$ -dimensional level set $L_C(z) \subset [0, 1]^d$. This distribution is however not easily tractable in general.

Again in the case of Archimedean copulas some convenient results are obtainable.

Proposition 3.6 (Projected distribution of Archimedean copulas). *Let $\mathbf{U} \sim C$, where C is a d -dimensional Archimedean copula with generator φ , then it holds for $z \in (0, 1)$ that*

$$[\mathbf{U} | C(\mathbf{U}) = z] \stackrel{d}{=} (\varphi^{-1}(S_1 \varphi(z)), \dots, \varphi^{-1}(S_d \varphi(z)))', \quad (3.5)$$

where $\mathbf{S} = (S_1, \dots, S_d)'$ is uniformly distributed on the unit simplex.

Proof: According to Equation (2.6), we have the representation $(\varphi(U_1), \dots, \varphi(U_d))' \stackrel{d}{=} R\mathbf{S}$, where $R = \sum_{j=1}^d \varphi(U_j) = \varphi(C(\mathbf{U}))$ is the radial part, which is independent of \mathbf{S} . Fixing the level set

$L_C(z)$ is therefore equivalent to setting $R = \varphi(z)$, so that we obtain Equation (3.5). \square

Equation (3.5) for the projected distribution of an Archimedean copula is particularly appealing, since it does not depend on the radial variable and its distribution, which may not be available in closed form. This result can then be used to provide another sampling algorithm for $\mathbf{U}|C(\mathbf{U}) = z$, which can be shown to be equivalent to Algorithm 3.5, when using explicit expressions for the observations $(s_1, \dots, s_d)'$ from \mathbf{S} in terms of uniform random variables (see Hering (2011, Lemma 3.1.8)).

Algorithm 3.7 (Projected distribution sampling for Archimedean copulas). *Let C be an Archimedean copula with generator φ and $z \in (0, 1)$.*

- (i) *Sample $(s_1, \dots, s_d)'$ from \mathbf{S} .*
- (ii) *For $j = 1, \dots, d$: $u_j := \varphi^{-1}(s_j \varphi(z))$.*
- (iii) *Return $(u_1, \dots, u_d)'$.*

As stated above, it is an open problem to determine this projected distribution for general copulas. Analogous sampling methods to Algorithm 3.7 could be used then.

Finally, Proposition 3.6 also enables us to formulate a simple proof of Lemma 3.4.

Remark 3.8 (Proof of Lemma 3.4). *We show that it holds for an Archimedean copula C and all $j = 1, \dots, d - 1$ that*

$$F_{j|1, \dots, j-1}(u|u_1, \dots, u_{j-1}, z) = \left(1 - \frac{\varphi(u)}{\varphi(z) - \sum_{1 \leq i < j} \varphi(u_i)}\right)^{d-j}, \quad u \in (C_{u_1, \dots, u_{j-1}}^{-1}(z), 1).$$

According to Proposition 3.6, it holds that

$$\begin{aligned} F_{j|1, \dots, j-1}(u|u_1, \dots, u_{j-1}, z) &= P(U_j \leq u | U_1 = u_1, \dots, U_{j-1} = u_{j-1}, C(\mathbf{U}) = z) \\ &= P\left(S_j \geq \frac{\varphi(u)}{\varphi(z)} \mid S_1 = \frac{\varphi(u_1)}{\varphi(z)}, \dots, S_{j-1} = \frac{\varphi(u_{j-1})}{\varphi(z)}\right), \end{aligned}$$

where $\mathbf{S} = (S_1, \dots, S_d)'$ follows a uniform distribution on the unit simplex, which can be represented as a Dirichlet distribution with all parameters equal to 1. It then holds that (see Fang et al. (1990, Theorem 1.6))

$$\frac{S_j}{1 - s_1 - \dots - s_{j-1}} \Big| (S_1 = s_1, \dots, S_{j-1} = s_{j-1}) \sim \text{Beta}(1, d - j), \quad j = 1, \dots, d - 1.$$

Further, the distribution function of the $\text{Beta}(1, d - j)$ distribution is $F_{\text{Beta}}(s; 1, d - j) = 1 - (1 - s)^{d-j}$. Therefore, we obtain

$$\begin{aligned} &P\left(S_j \geq \frac{\varphi(u)}{\varphi(z)} \mid S_1 = \frac{\varphi(u_1)}{\varphi(z)}, \dots, S_{j-1} = \frac{\varphi(u_{j-1})}{\varphi(z)}\right) \\ &= 1 - F_{\text{Beta}}\left(\frac{\frac{\varphi(u)}{\varphi(z)}}{1 - \frac{\varphi(u_1)}{\varphi(z)} - \dots - \frac{\varphi(u_{j-1})}{\varphi(z)}}; 1, d - j\right) \\ &= \left(1 - \frac{\varphi(u)}{\varphi(z) - \sum_{1 \leq i < j} \varphi(u_i)}\right)^{d-j}, \end{aligned}$$

as claimed in Lemma 3.4.

3.2 Estimation

In light of Sklar's theorem (1.1), it is common in dependence modeling to transform data $(x_{j1}, \dots, x_{jn})'$, $j = 1, \dots, N$, to $[0, 1]^n$ using the marginal distribution functions F_{X_i} , $i = 1, \dots, n$, that is, we compute $u_{ji} = F_{X_i}(x_{ji})$. In most cases, F_{X_i} will however be unknown so that this transformation needs to be based on a parametric or non-parametric estimate, which introduces uncertainty into the modeling.

When a parametric modeling of the margins is chosen, parameters of margins and dependence model can either be estimated jointly or, when this is not feasible, sequentially using the estimation method of inference functions for margins (IFM) by McLeish and Small (1988) and Joe and Xu (1996). In the IFM method, first marginal parameters are estimated and then dependence parameters given the estimated margins \hat{F}_{X_i} , $i = 1, \dots, n$. That is, parameters of a hierarchical Kendall copula are estimated based on $\hat{u}_{ji} = \hat{F}_{X_i}(x_{ji})$, $j = 1, \dots, N$, $i = 1, \dots, n$, by maximizing the log likelihood, which conveniently decomposes into separate sums (see Theorem 2.8) and which is straightforward to evaluate for Archimedean cluster copulas with Kendall distribution function given in Equation (2.5).

The asymptotic covariance matrix is given by the inverse Godambe information matrix, which unfortunately is typically very cumbersome to compute. To see this, note that the joint density (2.8) depends on the parameters of a cluster copula both through the density of the cluster copula as well as through the arguments of the nesting copula. For such situations, Joe and Xu (1996) propose a jackknife estimate of the asymptotic covariance. In our application in Section 4, the margins will be time-dependent. In this case, a stationary block bootstrap can be used to calculate approximate standard errors (see Politis and Romano (1994) and Gonçalves and White (2004)).

As an approximation to maximum likelihood estimates of the dependence parameters we additionally propose a sequential approach and evaluate and compare the finite sample behavior of the estimators in an extensive simulation study.

3.2.1 Sequential estimation

The hierarchical construction given in Definition 2.2 directly leads to a sequential estimation procedure of hierarchical Kendall copulas, which avoids higher dimensional maximum likelihood estimation.

Algorithm 3.9 (Sequential estimation of hierarchical Kendall copulas). *Let $(u_{j1}, \dots, u_{jn})'_{j=1, \dots, N}$ be a sample of a hierarchical Kendall copula $C_{\mathcal{K}}$ as defined in Definition 2.2 (possibly after appropriate transformation of the margins). Further, let $\theta_0, \theta_1, \dots, \theta_d$ be the parameters of the copulas C_0, C_1, \dots, C_d , respectively. We then obtain corresponding sequential estimates $\hat{\theta}_i$, $i = 0, \dots, d$, as follows.*

(i) *For each $i \in \{1, \dots, d\}$ estimate θ_i based on $(u_{j, m_{i-1}+1}, \dots, u_{j, m_i})'_{j=1, \dots, N}$ by maximum likelihood.*

(ii) *Estimate θ_0 based on the pseudo observations*

$$\hat{v}_{ji} := K_i(C_i(u_{j, m_{i-1}+1}, \dots, u_{j, m_i}; \hat{\theta}_i); \hat{\theta}_i), \quad i = 1, \dots, d, \quad j = 1, \dots, N, \quad (3.6)$$

by maximum likelihood.

Clearly, this two-step estimation procedure directly generalizes to a k -step estimation approach for k -level hierarchical Kendall copulas (see Remark 2.7). Resulting estimates may be used as starting values for a joint maximum likelihood estimation of the dependence parameters.

3.2.2 Simulation study

In order to examine the finite sample behavior of the estimation procedures discussed above, we perform a large scale Monte Carlo study. For this, we simulate from a four-dimensional hierarchical Kendall copula (two bivariate clusters; margins are assumed to be known) and then estimate the parameters according to the following methods:

- Sequential estimation;
- MLE with known starting values (true parameters);
- MLE with sequentially estimated starting values.

The cluster copulas C_1 and C_2 are chosen as Clayton, Gumbel or Frank; the nesting copula C_0 as Gaussian, Student's t (ten degrees of freedom), Clayton, Gumbel or Frank. Parameters are determined according to Kendall's τ values of 0.4 and 0.7. Sample sizes are 250, 500 and 1000 and the number of repetitions is 100. Estimation accuracy is compared based on the mean squared error of the estimated nesting copula parameter θ_0 (transformed to Kendall's τ values) as shown in Figure 8 for the case of Clayton and Gumbel cluster copulas. Results of the other cases are available from the author upon request. An illustrative sample of size 1000 for the case of Clayton and Gumbel cluster copulas (Kendall's τ of 0.4) and Frank nesting copula (Kendall's τ of 0.7) is shown in Figure 3.

It turns out that results are essentially independent of the chosen cluster copulas (not shown here). Similarly, the choice of parameters of the cluster copulas does not have a great influence on the results, while larger nesting parameters mean more accurate results. Overall, there is hardly any difference between the three estimation procedures. In particular, this means that sequential estimation provides good starting values for joint estimation of the dependence parameters.

3.3 Model selection

In practical applications, the clusters $(U_{m_{i-1}+1}, \dots, U_{m_i})'$, $i \in \{1, \dots, d\}$, have to be identified. In cases where they are not given from the data (e.g., industry sectors in financial data; see Section 4), common clustering techniques (see, e.g., Hastie et al. (2009)) can be used. If a multi-level hierarchical Kendall copula is considered, hierarchical clustering methods may be particularly helpful. In hierarchical clustering, the use of an appropriate metric to measure the closeness between (groups of) variables is essential. Inspired by Mantegna (1999), who uses the linear correlation coefficient, we propose to use the following metric between variables j and k ,

$$d(j, k) = \sqrt{1 - \hat{\rho}_{jk}^S}, \quad (3.7)$$

where $\hat{\rho}_{jk}^S$ is the empirical Spearman's ρ between the variables. Thus, the stronger the dependence between variables j and k , the smaller is $d(j, k)$. Obviously, $d(j, k) = 0$ if variables j and k are comonotonic, that is, if the observations are the same. Further, it holds that $d(j, k) = d(k, j)$ (symmetry) and $d(j, k) \leq d(j, \ell) + d(\ell, k)$ for another variable ℓ .

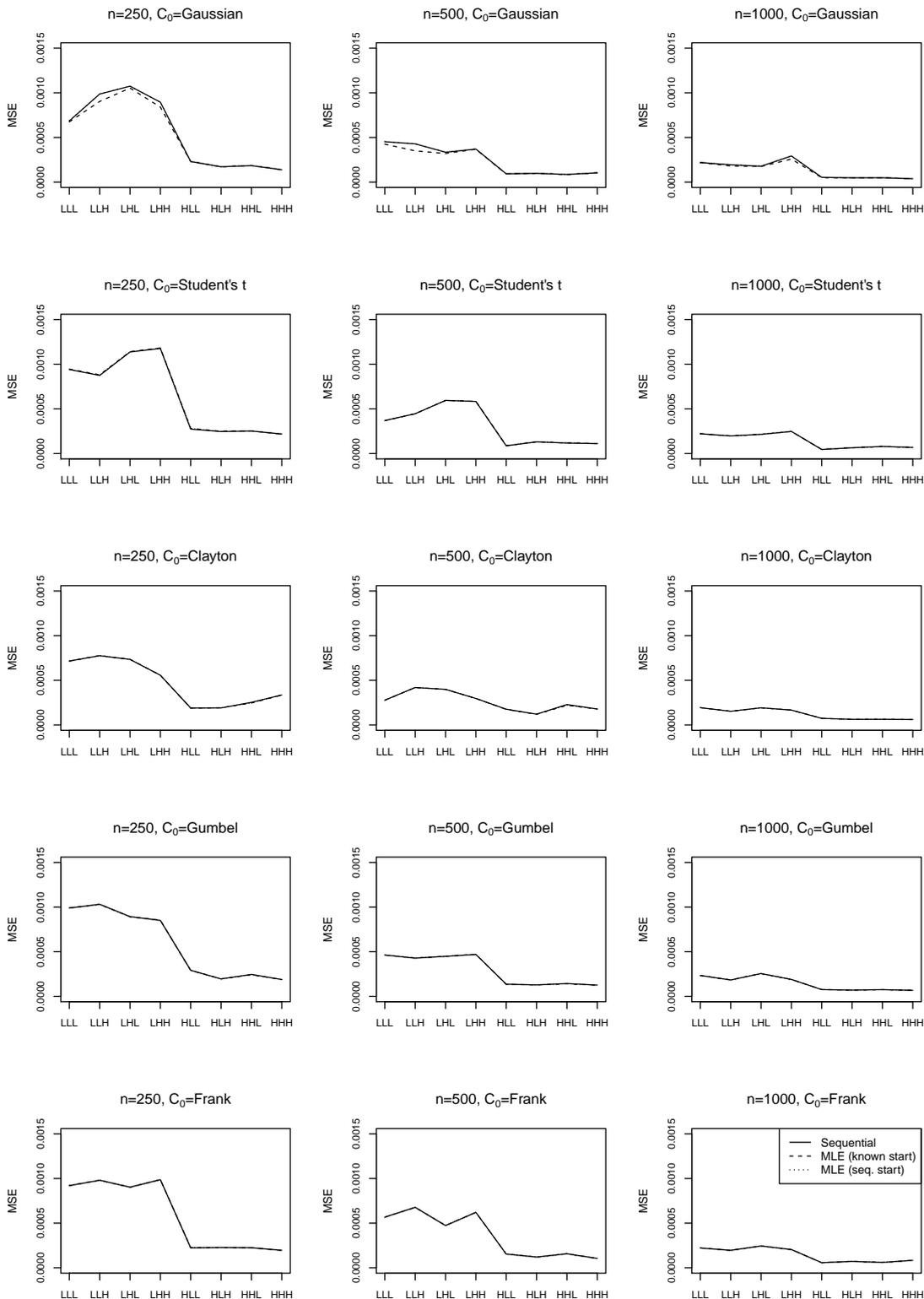


Figure 8: Mean squared errors (MSEs) of estimated parameters in terms of Kendall's τ values of the three estimation procedures for different sample sizes n and nesting copulas C_0 . Cluster copula 1: Clayton. Cluster copula 2: Gumbel. Notation for the x -axis: (τ_0, τ_1, τ_2) , where $L := 0.4$ and $H := 0.7$. The range of the y -axes is chosen such that MSEs in each column are comparable.

Hierarchical clustering further requires the choice of a linkage criterion to determine the distance between groups of variables. Classical average linkage clustering simply uses the mean distance between the elements of the groups. In the setting of hierarchical Kendall copulas, it is however more natural to form the pseudo observations \hat{v}_{ji} (3.6) based on purely empirical versions of the copula C_i and the Kendall distribution function K_i (see, e.g., Barbe et al. (1996)) and then compute the distance (3.7) between them, since this respects the model definition. This measure of association between multivariate random vectors, along with others, is also discussed by Grothe et al. (2011). It however does not ensure that the closeness between grouped variables is monotone decreasing with increasing level of the merger. The process of the merger is typically illustrated in a binary tree, which is called dendrogram and represents the closeness between cluster members (see Figure 11 below).

Recalling the discussion in Example 2.1, the size of clusters has to be carefully chosen because Kendall distributions may become almost degenerate at 0 for very large clusters. In most practical situations, this is however not an issue, since already under medium positive dependence the convergence to the constant function at 1 is very slow (see the right panel of Figure 1).

In the next step, copulas have to be selected for the clusters. Due to the hierarchical nature of the model, higher order levels depend on copulas in lower levels, so that a careful selection of cluster copulas is necessary. A possible approach is a stepwise selection similar to the sequential estimation procedure outlined in Algorithm 3.9, that is, the nesting copula C_0 is selected based on the pseudo observations (3.6). This is similar to selection approaches of hierarchical Archimedean copulas (Okhrin et al. 2013) and of pair copula constructions (Dißmann et al. 2013). Common criteria for the selection of copulas are information criteria such as the AIC. Since the selection based on pseudo observations however induces uncertainty in the selection of the nesting copula, we perform a misspecification study.

3.3.1 Copula misspecification

To analyze the effect of misspecification of the cluster and nesting copulas, we resume the setting of Section 3.2.2 and simulate samples of size 1000 from four-dimensional hierarchical Kendall copulas with cluster copulas C_1 and C_2 chosen as Clayton, Gumbel or Frank, and the nesting copula C_0 as Gaussian, Student's t (ten degrees of freedom), Clayton, Gumbel or Frank. Parameters are again determined according to Kendall's τ values of 0.4 and 0.7 and the number of repetitions is 100.

In addition, we simulate from a range of alternative multivariate copulas to investigate how well these copulas can be approximated by a hierarchical Kendall copula. We consider four-dimensional regular vine pair copula constructions (see Dißmann et al. (2013) and the R package `VineCopula` by Schepsmeier et al. (2013)) with unconditional pair copulas chosen as $C_{1,2} = C_1$, $C_{2,3} = C_0$ and $C_{3,4} = C_2$ and conditional pair copulas chosen as $C_{1,3|2} = C_{2,4|3} = C_{1,4|2,3} = C_0$, whose parameters are determined according to a decreasing value of Kendall's τ compared to the unconditional copula $C_{2,3}$ ². Such a pair copula construction mimics, to some extent, a hierarchical dependence model, but without having explicit between-cluster dependence. Furthermore, we simulate from hierarchical Archimedean copulas with Clayton, Gumbel and Frank cluster and nesting copulas. This is however only possible if $\tau_0 \leq \min\{\tau_1, \tau_2\}$, since the between-cluster dependence cannot be stronger than the within-cluster dependence. Finally, we also consider

²If $\tau_{j,k|D}$ denotes the Kendall's τ corresponding to the pair copula $C_{j,k|D}$ and τ_0 is the Kendall's τ of $C_{2,3} = C_0$, then we choose $\tau_{1,3|2} = \tau_{2,4|3} = 2\tau_0/3$ and $\tau_{1,4|2,3} = \tau_0/3$

four-dimensional Gaussian and Student's t copulas (ten degrees of freedom) with correlation matrices clustered according to the respective Kendall's τ values for within- and between-cluster dependence³.

The effect of misspecification is then examined in terms of the Kullback-Leibler divergence between the true (simulated) model and the alternative models fitted by maximum likelihood. The results are illustrated in Figures 9 and 10. First of all, they show that the hierarchical Kendall copulas provide a very good fit if the cluster copulas are identified correctly. This means that the effect of the uncertainty with respect to the selection of the nesting copula based on pseudo observations is not severe. These results also hold when the true model is a hierarchical Archimedean copula, which apparently can be well approximated by a hierarchical Kendall copula (see Figure 5). Clearly, differences between the models become more distinct with increasing dependence in terms of Kendall's τ . In particular, the Clayton copula is rather different from the other copulas (see the shape of the scatter plots in Figure 3) and therefore harder to approximate by a misspecified model. As a result, the elliptical copulas are also best approximated by hierarchical Kendall copulas with Gumbel and Frank components. In particular, hierarchical Kendall copulas with elliptical nesting copula and Gumbel or Frank cluster copulas are very close in terms of the Kullback-Leibler divergence even if clusters are heterogeneous. The non-hierarchical pair copula constructions are naturally more difficult to approximate by hierarchical Kendall copulas. Especially when the cluster copulas C_1 and C_2 are selected in the same way as the pair copulas $C_{1,2}$ and $C_{3,4}$, respectively, hierarchical Kendall copulas may however be quite close in terms of the Kullback-Leibler divergence. Conversely, a pair copula construction may also quite well approximate a hierarchical Kendall copula especially when the dependence is weak and if the vine copula is well chosen (note that only the results of the best fitting among the 15 considered vine copulas are shown in Figures 9 and 10). The selection of vine copulas in higher dimensions is however a major problem (see Czado et al. (2013)) and the models typically do not stay as parsimonious and as easily interpretable as hierarchical Kendall copulas. Finally, the results show that also elliptical and hierarchical Archimedean copulas can sometimes provide good approximations to hierarchical Kendall copulas in this four-dimensional setting. It is, however, to be expected that the differences among the models become larger in higher dimensions.

4 Application

Finance is a major field, where copulas are used for dependence modeling (see, e.g., Cherubini et al. (2004)). Often financial data exhibits some kind of clustering structure such as industry sectors and national stock markets. For such data, hierarchical Kendall copulas are very suitable. To investigate the usefulness of this newly proposed class of dependence models and to illustrate the presented inference techniques, the most important German stock market index DAX is analyzed.

The DAX is composed of 30 major German stocks. For these we identified ten industry sectors: financials (Allianz, Commerzbank, Deutsche Bank, Deutsche Börse, Munich Re), chemicals (BASF, Bayer, K+S, Linde), healthcare (Fresenius, Fresenius Medical Care, Merck),

³In order to ensure positive definiteness, Kendall's τ values have to be adapted when between-cluster dependence is 0.7. Then, within-cluster dependence is set to either 0.6 for both clusters or to 0.7 and 0.5 for the different clusters.

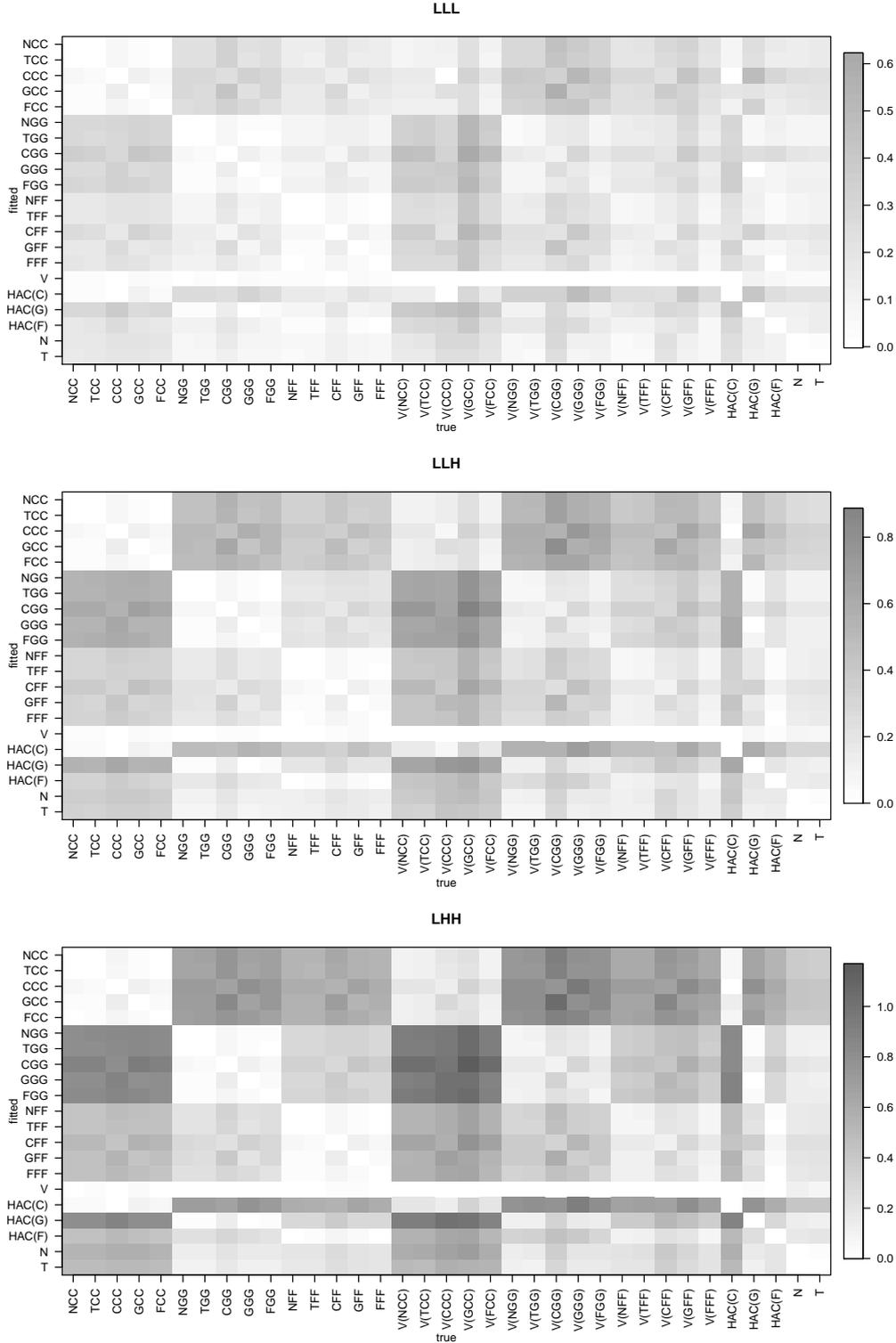


Figure 9: Illustration of mean Kullback-Leibler divergences (part 1): light colors indicate a small divergence, dark colors a large divergence. Notation for x - and y -axes: (C_0, C_1, C_2) with Gaussian (N), Student's t (T), Clayton (C), Gumbel (G), and Frank (F) copulas; vine copulas are indicated by 'V', hierarchical Archimedean copulas by 'HAC'. Columns correspond to the true (simulated) models, rows to the fitted models (only the best fitting among the 15 considered vine copulas is shown). Notation for panel titles: (τ_0, τ_1, τ_2) , where $L := 0.4$ and $H := 0.7$.

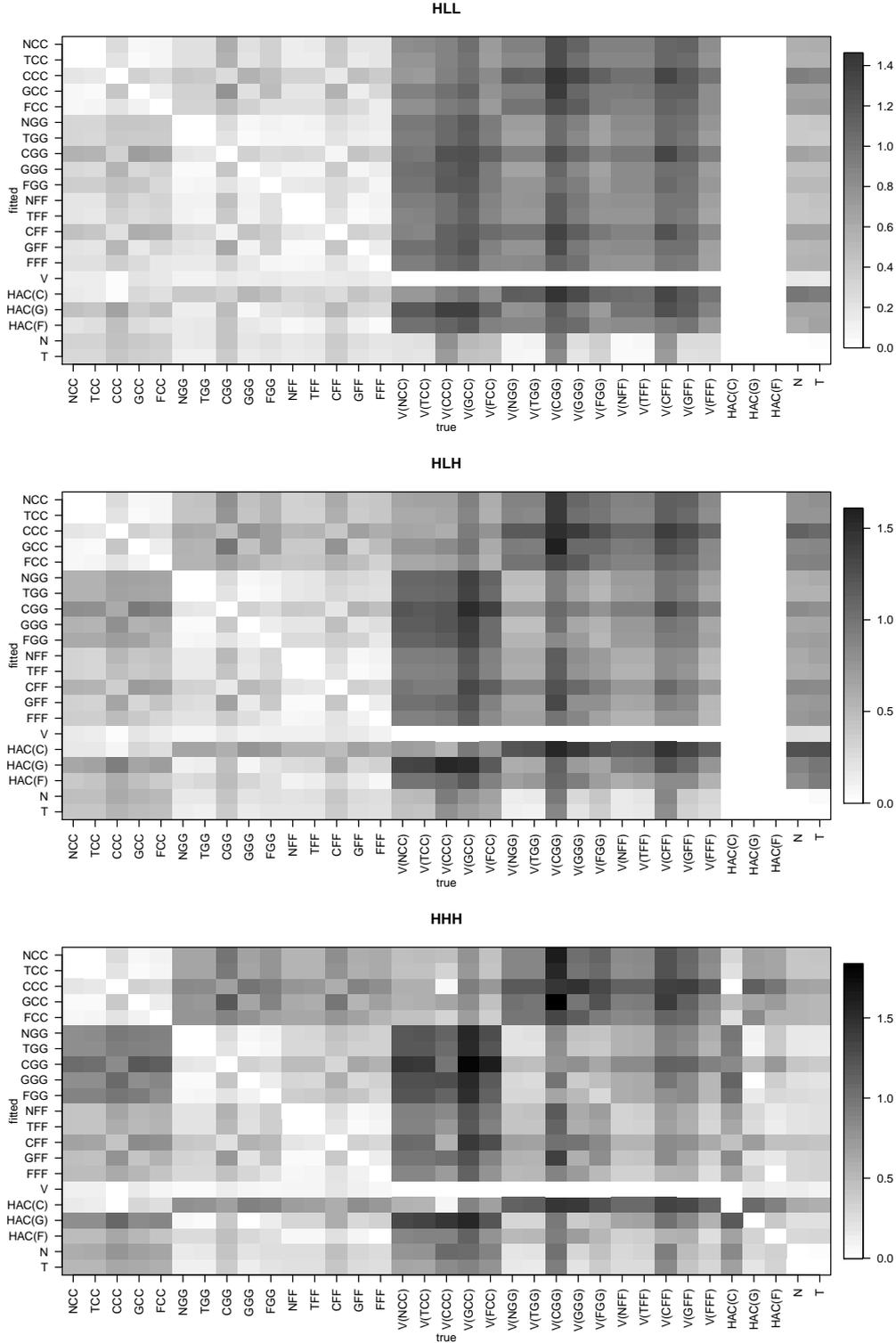


Figure 10: Illustration of mean Kullback-Leibler divergences (part 2): light colors indicate a small divergence, dark colors a large divergence. Notation for x - and y -axes: (C_0, C_1, C_2) with Gaussian (N), Student's t (T), Clayton (C), Gumbel (G), and Frank (F) copulas; vine copulas are indicated by 'V', hierarchical Archimedean copulas by 'HAC'. Columns correspond to the true (simulated) models, rows to the fitted models (only the best fitting among the 15 considered vine copulas is shown). Notation for panel titles: (τ_0, τ_1, τ_2) , where $L := 0.4$ and $H := 0.7$.

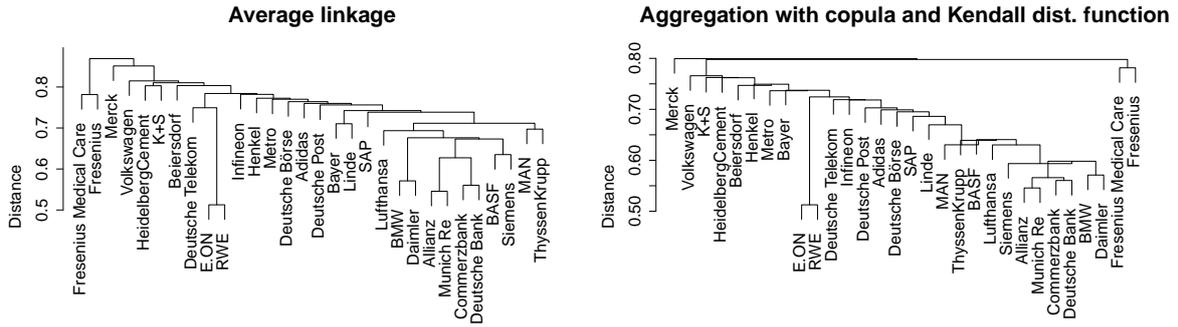


Figure 11: Dendrograms of the DAX constituents according to average linkage (left panel) and aggregation using the empirical Kendall distribution function of clusters (right panel).

Sectors	Fin.	Chem.	Healthc.	Auto.	Ind.	Retail	IT	Util.	Transp.
Size	5	4	3	3	3	4	3	2	2
Mean pairw. Kendall's τ	0.41	0.33	0.21	0.39	0.38	0.26	0.28	0.56	0.29
Estimated deg. of freedom	8.80	10.70	22.96	12.65	8.13	10.07	8.74	4.63	7.22

Table 1: Sector size, mean pairwise empirical Kendall's τ and estimated degrees of freedom of a Student's t copula for each cluster.

automobile (BMW, Daimler, Volkswagen), industrials (MAN, Siemens, ThyssenKrupp), retail and consumer goods (Adidas, Beiersdorf, Henkel, Metro), IT and communications (Deutsche Telekom, Infineon, SAP), utilities (E.ON, RWE), transportation and logistics (Deutsche Post, Lufthansa), and building materials (HeidelbergCement).

For all 30 stocks, more than six years of log returns (January 2005 to July 2011) are considered, where the time series are split into a training set of 1158 observations and a testing set of 500 observations for out-of-sample validation of our models. As it is common for copula modeling in finance, we preliminarily fit time series models to the marginal time series and then work with standardized residuals which are transformed to marginally uniform data by the probability integral transform (inference functions for margins method; see Section 3.2). In particular, marginal GARCH(1,1)-models with Student's t innovations are chosen, which have been validated with appropriate tests. Using this data, we illustrate the cluster selection procedures described in Section 3.3 and perform hierarchical clustering with the metric (3.7) and average linkage as well as aggregation using the empirical Kendall distribution function of clusters. The resulting dendrograms are shown in Figure 11. For instance, the utility and the healthcare sector can easily be identified. This is not the case for the chemical and the IT companies.

Table 1 shows the mean pairwise empirical Kendall's τ and the estimated degrees of freedom of a multivariate Student's t copula for each cluster. Evidently, within-sector dependence is variable, since some clusters are more homogeneous than others. Also strong tail dependence, as indicated by small degrees of freedom, cannot be found in all clusters.

We then fitted different hierarchical Kendall copulas to the training data set. Results (log likelihood, AIC, BIC) are reported in Table 2. As cluster copulas, we considered three different Archimedean copulas to account for different dependence structures as typically observed in financial data: Clayton with lower tail dependence, Gumbel with upper tail dependence and

Model	# Par.	Seq. Est. Log Lik.	Joint MLE Log Lik.	AIC	BIC
(Clayton, Student's t)	55	6656.50	6677.73	-13245.46	-12967.47
(Gumbel, Student's t)	55	6989.32	6992.29	-13874.57	-13596.58
(Frank, Student's t)	55	7185.70	7190.29	-14270.58	-13992.58
(Clayton, Clayton)	10	5452.34	5471.09	-10922.17	-10871.63
(Gumbel, Gumbel)	10	5860.74	5862.93	-11705.85	-11655.31
(Frank, Frank)	10	6003.60	6005.56	-11991.12	-11940.58
Multivar. Gaussian	435	-	8487.71	-16105.41	-13906.73
Multivar. Student's t	436	-	8906.14	-16940.28	-14736.54
Regular vine	509	-	9512.29	-18006.58	-15433.87

Table 2: Estimation results based on the training set. Notation of models: (cluster copulas, nesting copula). AIC and BIC values are based on the joint MLE.

Frank with no tail dependence. For the nesting copulas, we also investigated Gaussian and Student's t copulas, where the fits of the Gaussian copula to the aggregated pseudo observations of the different sectors were always inferior to that of the Student's t copula and are therefore not displayed here. The different specified models are shown in the first column of Table 2. The finite sample behavior of the estimates is investigated in Section 3.2.2.

For comparison, we also fitted classical multivariate Gaussian and Student's t copulas as well as a regular vine pair copula construction using the selection and estimation algorithm by Dißmann et al. (2013), which is implemented in the R package `VineCopula` by Schepsmeier et al. (2013). Pair copulas are selected from the following list: Gaussian, Student's t, Clayton, Gumbel and Frank as well as rotations by 90, 180 and 270 degrees of the tail-asymmetric Clayton and Gumbel copulas. A 30-dimensional hierarchical Archimedean copula could however not be fitted due to the dependence restrictions of hierarchical Archimedean copulas (see Section 2.3): while there is moderate dependence within some clusters (see Table 1), there still is considerable and heterogeneous dependence among clusters (pairwise Kendall's τ values of aggregated data are ranging between 0.15 and 0.51), which cannot be modeled using a hierarchical Archimedean copula.

The hierarchical Kendall copulas therefore benefit from not having this restriction on within- and between-cluster dependence. Moreover, ten-dimensional Student's t nesting copulas appear more reasonable than Archimedean nesting copulas (with only one parameter) due to the varying pairwise between-cluster dependence. With respect to cluster copulas, the tail-symmetric Frank copula is, according to the AIC, superior to tail-asymmetric ones (Clayton, Gumbel). For the two hierarchical Kendall copulas with Frank cluster copulas, parameter estimates and their standard errors according to the stationary bootstrap by Politis and Romano (1994) with an average block length of 20 observations can be found in Table 3, showing that there is significant within- and between-sector dependence. In particular, the within-sector dependence is quite variable and therefore can also not be appropriately fitted by an exchangeable 30-dimensional Archimedean copula. For instance, the maximum likelihood parameter estimate of a 30-dimensional Frank copula is 2.096, of which most cluster copula parameters are significantly different.

In comparison to standard multivariate Gaussian and Student's t copulas, hierarchical Kendall copulas perform quite well, in particular when taking into account the enormous number of parameters of these models. The number of parameters of elliptical copulas could be reduced significantly by using clustered correlation matrices. Nevertheless, these have to be fitted carefully

	(Frank, Student's t)		(Frank, Frank)	
	Estimate (Kendall's τ)	Std. Error	Estimate (Kendall's τ)	Std. Error
Financials	3.816 (0.374)	0.210	3.946 (0.384)	0.210
Chemicals	2.915 (0.300)	0.197	3.047 (0.311)	0.199
Healthcare	1.810 (0.195)	0.191	2.003 (0.214)	0.182
Automobile	3.775 (0.371)	0.237	3.897 (0.380)	0.238
Industrials	3.634 (0.360)	0.303	3.880 (0.379)	0.304
Retail	2.189 (0.232)	0.168	2.394 (0.252)	0.167
IT and comm.	2.401 (0.253)	0.218	2.664 (0.277)	0.215
Utilities	6.829 (0.555)	0.544	6.912 (0.558)	0.543
Transportation	2.669 (0.278)	0.325	2.981 (0.306)	0.329
Between-sector	0.248–0.704 (0.160–0.497)	0.020–0.043	2.979 (0.305)	0.165
Degrees of freedom	19.597	1.559	-	-

Table 3: Parameter estimates and their estimated standard errors of the hierarchical Kendall copulas with Frank cluster copulas (based on the training set). Kendall's τ values corresponding to the parameter estimates are given in brackets. For the entries of the Student's t correlation matrix ranges are reported.

in order to satisfy positive definiteness constraints. Overall the regular vine copula provides the best fit, since it constitutes the most flexible model. It is however even less parsimonious than the elliptical copulas and is not straightforward to interpret, especially not in terms of sectoral dependence. Given that the highly parameterized multivariate elliptical copulas and regular vine pair copula constructions can be regarded as the current state-of-the-art models for financial return data, we focus on these models and investigate if the more parsimonious hierarchical Kendall copulas are competitive with them. The good in-sample results obtained so far are in line with the misspecification study in Section 3.3.1, where hierarchical Kendall copulas with elliptical nesting and Frank or Gumbel cluster copulas are shown to be reasonably close to multivariate elliptical models and, to some extent, to pair copula constructions.

4.1 Value-at-Risk forecasting

In finance, interest is however not so much in a good in-sample fit but rather in out-of-sample validation. A typical exercise for this is Value-at-Risk (VaR) forecasting. If the distribution of returns is continuous, the $(1-\alpha)$ -VaR is the α -quantile of the distribution. For risk management, this value needs to be predicted on a daily basis, which we do for the testing set of 500 days using moving windows of length 1158.

Forecasts are typically evaluated in terms of exceedances, that is, the event that the predicted VaR is exceeded by the observed return. For 500 forecasts, on average $500 \times \alpha$ exceedances are expected. Whether the number of exceedances (“unconditional coverage”) and their occurrences (should be independent; both properties: “conditional coverage”) are appropriate can be evaluated using a range of tests, so-called backtests, that have been proposed in the literature: the proportion of failures test of unconditional coverage by Kupiec (1995) (UC), the Markov test of independence by Christoffersen (1998) (IND1), the joint test of conditional coverage by Christoffersen (1998) (CC1), the mixed Kupiec test of conditional coverage by Haas (2001) (CC2), the Weibull test of independence by Christoffersen and Pelletier (2004) (IND2), and the duration-based GMM test of conditional coverage by Candelon et al. (2011) (CC3 and CC4

Model	Level	# Exceed.	UC	IND1	IND2	CC1	CC2	CC3	CC4
Independence copula	99%	103	<i>0.00</i>	<i>0.04</i>	<i>0.00</i>	<i>0.00</i>	<i>0.00</i>	<i>0.00</i>	<i>0.00</i>
	95%	135	<i>0.00</i>	0.15	<i>0.03</i>	<i>0.00</i>	<i>0.00</i>	<i>0.00</i>	<i>0.00</i>
	90%	157	<i>0.00</i>	0.17	<i>0.03</i>	<i>0.00</i>	<i>0.00</i>	<i>0.00</i>	<i>0.00</i>
Gaussian copula	99%	5	1.00	0.75	0.79	0.95	0.35	0.93	1.00
	95%	26	0.84	0.09	0.74	0.23	0.35	0.98	1.00
	90%	52	0.77	0.46	0.87	0.73	<i>0.03</i>	0.29	0.73
Student's t copula	99%	4	0.64	0.80	0.21	0.87	0.93	0.75	0.95
	95%	24	0.84	0.12	0.73	0.29	0.39	0.89	1.00
	90%	53	0.66	0.53	0.85	0.74	0.05	0.31	0.73
Regular vine pair copula construction	99%	5	1.00	0.75	0.81	0.95	0.35	0.93	1.00
	95%	27	0.69	0.08	0.70	0.20	0.35	0.95	1.00
	90%	53	0.66	0.53	0.87	0.74	<i>0.04</i>	0.24	0.66
Hierarchical	99%	2	0.13	0.90	0.46	0.31	0.35	0.43	0.82
Kendall copula	95%	31	0.23	<i>0.04</i>	0.76	0.06	0.21	0.55	0.92
(Frank, Student's t)	90%	57	0.31	0.52	0.81	0.48	<i>0.03</i>	0.22	0.64

Table 4: P -values of VaR backtests for hypotheses of independence and (un)conditional coverage. We expect 5/25/50 exceedances at the 99%/95%/90%-level, respectively.

with orders 2 and 5, respectively).

Here, the Value-at-Risk of an equally weighted portfolio of the 30 DAX stocks is forecasted. Backtesting results of the following five different models can be found in Table 4:

- Independence copula (for comparison);
- Gaussian and Student's t copulas;
- Regular vine pair copula construction;
- Hierarchical Kendall copula with Frank cluster copulas and with Student's t nesting copula (best fit among the considered hierarchical Kendall copulas).

In summary, none of the hypotheses of independence and (un)conditional coverage can consistently be rejected for any of the VaR levels and for any of the models—except for the multivariate independence copula, as expected. The weak lack of conditional coverage at the 90% level, as detected by the mixed Kupiec test of Haas (2001), is not supported by the other tests.

This shows that hierarchical Kendall copulas are as good as the common Gaussian and Student's t copulas and also as the more flexible pair copula constructions when it comes to out-of-sample validation. In particular, the hierarchical Kendall copula with Frank cluster copulas, which we consider here, is very parsimonious and allows for closed-form calculations and very efficient simulations due to its Archimedean clusters. In contrast to elliptical and vine copulas, it is also directly interpretable in terms of within- and between-sector dependence.

5 Discussion

In this paper we introduce and discuss the new class of hierarchical Kendall copulas. By grouping variables at different hierarchical levels, it provides an appealing construction principle for high-dimensional dependence models. It is shown that the important special cases of independence

as well as of comonotonicity belong to this model class. For Archimedean cluster copulas, a stochastic representation is given and differences to hierarchical Archimedean copulas are investigated. Most importantly, the density of hierarchical Kendall copulas is derived.

Thereafter, statistical inference techniques for hierarchical Kendall copulas are developed. In particular, simulation algorithms are provided, with focus on Archimedean clusters, for which convenient closed-form expressions are derived. The availability of the density of hierarchical Kendall copulas renders feasible estimation using maximum likelihood techniques. Finally, we show that a model with Archimedean cluster copulas and Student's t nesting copula performs very well in a substantial financial application.

Nevertheless, it is an open research question to derive sampling and estimation methods for the case that copulas and Kendall distribution functions are not available in closed form. In particular, the popular class of elliptical copulas is of interest here.

Acknowledgement

The author gratefully acknowledges the helpful comments of the referees, which further improved the manuscript. He also likes to thank his supervisor Claudia Czado as well as Harry Joe, Johanna Nešlehová, Matthias Scherer and Michael Smith for instructive remarks and discussions. Furthermore, he acknowledges the support of the TUM Graduate School's International School of Applied Mathematics as well as of Allianz Deutschland AG. The numerical computations were performed on a Linux cluster supported by DFG grant INST 95/919-1 FUGG.

References

- Aas, K., C. Czado, A. Frigessi, and H. Bakken (2009). Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics* 44(2), 182–198.
- Anjos, U. d. and N. Kolev (2005). An application of Kendall distributions. *Review of Business and Economics* L(1), 95–102.
- Arbenz, P., C. Hummel, and G. Mainik (2012). Copula based hierarchical risk aggregation through sample reordering. *Insurance: Mathematics and Economics* 51(1), 122–133.
- Barbe, P., C. Genest, K. Ghoudi, and B. Rémillard (1996). On Kendall's process. *Journal of Multivariate Analysis* 58(2), 197–229.
- Brechmann, E. C. (2013). Sampling from hierarchical Kendall copulas. *Journal de la Société Française de Statistique* 154(1), 192–209.
- Brechmann, E. C., C. Czado, and K. Aas (2012). Truncated regular vines in high dimensions with applications to financial data. *Canadian Journal of Statistics* 40(1), 68–85.
- Candelon, B., G. Colletaz, C. Hurlin, and S. Tokpavi (2011). Backtesting Value-at-Risk: A GMM duration-based test. *Journal of Financial Econometrics* 9(2), 314–343.
- Chakak, A. and M. Ezzerg (2000). Bivariate contours of copula. *Communications in Statistics - Simulation and Computation* 29(1), 175–185.
- Cherubini, U., E. Luciano, and W. Vecchiato (2004). *Copula Methods in Finance*. Chichester: John Wiley & Sons.

- Christoffersen, P. F. (1998). Evaluating interval forecasts. *International Economic Review* 39, 841–862.
- Christoffersen, P. F. and D. Pelletier (2004). Backtesting Value-at-Risk: A duration-based approach. *Journal of Financial Econometrics* 2(1), 84–108.
- Czado, C., E. C. Brechmann, and L. Gruber (2013). Selection of vine copulas. In P. Jaworski, F. Durante, and W. Härdle (Eds.), *Copulae in Mathematical and Quantitative Finance*. Berlin: Springer.
- Daul, S., E. De Giorgi, F. Lindskog, and A. McNeil (2003). The grouped t-copula with an application to credit risk. *Risk* 16(11), 73–76.
- Devroye, L. (1986). *Non-Uniform Random Variate Generation*. New York: Springer.
- Dißmann, J., E. C. Brechmann, C. Czado, and D. Kurowicka (2013). Selecting and estimating regular vine copulae and application to financial returns. *Computational Statistics & Data Analysis* 59(1), 52–69.
- Fang, K.-T., S. Kotz, and K. Ng (1990). *Symmetric Multivariate and Related Distributions*. London: Chapman & Hall.
- Genest, C., J. Nešlehová, and J. Ziegel (2011). Inference in multivariate Archimedean copulas. *Test* 20, 223–292.
- Genest, C., J. J. Quesada Molina, and J. A. Rodríguez Lallena (1995). De l'impossibilité de construire des lois à marges multidimensionnelles données à partir de copules. *Comptes Rendus de l'Académie des Sciences, Series I, Mathematics* 320, 723–726.
- Genest, C., J.-F. Quessy, and B. Rémillard (2006). Goodness-of-fit procedures for copula models based on the probability integral transform. *Scandinavian Journal of Statistics* 33, 337–366.
- Genest, C. and L.-P. Rivest (1993). Statistical inference procedures for bivariate Archimedean copulas. *Journal of the American Statistical Association* 88(423), 1034–1043.
- Genest, C. and L.-P. Rivest (2001). On the multivariate probability integral transformation. *Statistics & Probability Letters* 53(4), 391–399.
- Gonçalves, S. and H. White (2004). Maximum likelihood and the bootstrap for nonlinear dynamic models. *Journal of Econometrics* 119, 199–220.
- Gregory, J. and J.-P. Laurent (2004). In the core of correlation. *Risk* 17(10), 87–91.
- Grothe, O., F. Schmid, J. Schnieders, and J. Segers (2011). Measuring association between random vectors. Working paper, University of Cologne.
- Haas, M. (2001). New methods in backtesting. Technical report, Research Center Caesar.
- Hastie, T., R. Tibshirani, and J. H. Friedman (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2nd ed.). New York: Springer.
- Hering, C. (2011). *Estimation techniques and goodness-of-fit tests for certain copula classes in large dimensions*. Ph. D. thesis, Universität Ulm.
- Hofert, M. (2010). Construction and sampling of nested Archimedean copulas. In P. Jaworski, F. Durante, W. Härdle, and T. Rychlik (Eds.), *Copula Theory and Its Applications*. Berlin: Springer.

- Hofert, M. (2011). Efficiently sampling nested Archimedean copulas. *Computational Statistics & Data Analysis* 55(1), 57–70.
- Hofert, M., M. Mächler, and A. J. McNeil (2012). Likelihood inference for Archimedean copulas in high dimensions under known margins. *Journal of Multivariate Analysis* 110, 133–150.
- Hofert, M. and D. Pham (2013). Densities of nested Archimedean copulas. *Journal of Multivariate Analysis* 118, 37–52.
- Hofert, M. and M. Scherer (2011). CDO pricing with nested Archimedean copulas. *Quantitative Finance* 11(5), 775–787.
- Imlahi, L., M. Ezzerg, and A. Chakak (1999). Estimación de la curva mediana de una cópula $C(x_1, \dots, x_m)$. *Revista de la Real Academia de Ciencias Exactas, Físicas y Naturales. Serie A: Matemáticas* 93(2), 241–250.
- Joe, H. (1997). *Multivariate Models and Dependence Concepts*. London: Chapman & Hall.
- Joe, H. and J. Xu (1996). The estimation method of inference functions for margins for multivariate models. Technical Report 166, Department of Statistics, University of British Columbia.
- Kupiec, P. H. (1995). Techniques for verifying the accuracy of risk measurement models. *The Journal of Derivatives* 3(2), 73–84.
- Mantegna, R. N. (1999). Hierarchical structure in financial markets. *European Physical Journal B* 11(1), 193–197.
- Marco, J. M. and C. Ruiz-Rivas (1992). On the construction of multivariate distributions with given nonoverlapping multivariate marginals. *Statistics & Probability Letters* 15, 259–265.
- McLeish, D. L. and C. G. Small (1988). *The Theory and Applications of Statistical Inference Functions*. New York: Springer.
- McNeil, A. J. and J. Nešlehová (2009). Multivariate Archimedean copulas, d -monotone functions and ℓ_1 -norm symmetric distributions. *Annals of Statistics* 37(5B), 3059–3097.
- Nelsen, R. B., J. J. Quesada Molina, J. A. Rodríguez Lallena, and M. Úbeda Flores (2003). Kendall distribution functions. *Statistics & Probability Letters* 65(3), 263–268.
- Okhrin, O., Y. Okhrin, and W. Schmid (2013). On the structure and estimation of hierarchical archimedean copulas. *Journal of Econometrics* 173(2), 189–204.
- Politis, D. N. and J. P. Romano (1994). The stationary bootstrap. *Journal of the American Statistical Association* 89, 1303–1313.
- Salvadori, G., C. De Michele, and F. Durante (2011). On the return period and design in a multivariate framework. *Hydrology and Earth System Sciences* 15(11), 3293–3305.
- Savu, C. and M. Trede (2010). Hierarchical Archimedean copulas. *Quantitative Finance* 10, 295–304.
- Schepsmeier, U., J. Stöber, E. C. Brechmann, and B. Gräler (2013). *VineCopula: Statistical inference of vine copulas*. R package.
- Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de L'Université de Paris* 8, 229–231.

- Wang, W. and M. T. Wells (2000). Model selection and semiparametric inference for bivariate failure-time data. *Journal of the American Statistical Association* 95, 62–72.
- Wu, F., E. Valdez, and M. Sherris (2007). Simulating from exchangeable Archimedean copulas. *Communications in Statistics - Simulation and Computation* 36(5), 1019–1034.