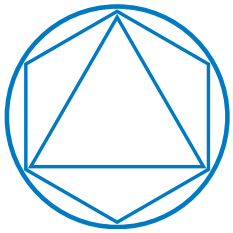
Hierarchical Kendall Copulas and the Modeling of Systemic and Operational Risk

Eike Christian Brechmann

2013



Fakultät für Mathematik
Technische Universität München
85748 Garching bei München

Technische Universität München
Fakultät für Mathematik

Hierarchical Kendall Copulas and the Modeling of Systemic and Operational Risk

Eike Christian Brechmann

Vollständiger Abdruck der von der Fakultät für Mathematik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:		Univ.-Prof. Dr. Rudi Zagst
Prüfer der Dissertation:	1.	Univ.-Prof. Claudia Czado, Ph.D.
	2.	Prof. Dorota Kurowicka, Ph.D. Technische Universiteit Delft, Niederlande
	3.	Prof. Harry Joe, Ph.D. University of British Columbia, Kanada (nur schriftliche Beurteilung)

Die Dissertation wurde am 27.06.2013 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik am 14.10.2013 angenommen.

Zusammenfassung

In der vorliegenden Arbeit untersuchen wir statistische Abhängigkeitsmodellierung mittels eines hierarchischen Ansatzes: Um Modellflexibilität und -sparsamkeit auszugleichen, wird Abhängigkeit in Form von Copulas für Gruppen von Variables in verschiedenen hierarchischen Ebenen spezifiziert, und werden Informationen über Ebenen hinweg durch die Kendall-Verteilungsfunktionen der Copulas aggregiert. Da Kendall-Verteilungsfunktionen die multivariaten Gegenstücke zur univariaten Wahrscheinlichkeitsintegral-Transformation sind, ahmt unser Ansatz klassische Copula-Modellierung mit univariaten Rändern nach. Das sich ergebende Abhängigkeitsmodell nennen wir „hierarchische Kendall-Copula“, untersuchen seine Eigenschaften und vergleichen es mit alternativen Modellen. Für die statistische Inferenz entwickeln wir geeignete Instrumente und Techniken. Während Likelihood-basierte Methoden aufgrund eines expliziten Ausdrucks der Dichte praktikabel sind, ist das Simulieren von hierarchischen Kendall-Copulas besonders anspruchsvoll. Wir untersuchen das Simulationsproblem detailliert und leiten geschlossene Lösungen für bestimmte Copula-Klassen her. Für den allgemeinen Fall werden approximative Methoden eingeführt und sorgfältig evaluiert.

Zwei wichtige Arten finanziellen Risikos werden in dieser Arbeit betrachtet: systemisches und operationelles Risiko. Zur Einschätzung der systemische Relevanz von Finanzinstituten schlagen wir vor, die Vernetzung der Institute im Market mittels multivariater Copulas zu analysieren. Für diese leiten wir neue bedingte Simulationsverfahren her, die wir nutzen, um einen Stress-Test des Marktes für Credit Default Swaps durchzuführen.

Schließlich entwickeln wir ein flexibles Abhängigkeitsmodell für quantitatives operationelles Risikomanagement. Die Modellbestandteile werden bezüglich relevanter Eigenschaften untersucht und geeignete Empfehlungen werden abgegeben. Anhand von Daten über operationelle Schäden von italienischen Banken sind wir dann in der Lage, die Auswirkungen der Modellierungsentscheidungen auf das operationelle Risikokapital abzuschätzen.

Abstract

In this thesis, we study statistical dependence modeling using a hierarchical approach: To balance model flexibility and parsimony, dependence is specified in terms of copulas for groups of variables in different hierarchical levels, and information across levels is aggregated by the Kendall distribution functions of the copulas. As Kendall distribution functions are the multivariate analogs of the univariate probability integral transform, our approach mimics classical copula modeling with univariate margins. We call the resulting dependence model "hierarchical Kendall copula", investigate its properties and compare it with alternative models. For the statistical inference, we develop appropriate tools and techniques. While likelihood-based methods are feasible due to an explicit expression of the density, sampling from hierarchical Kendall copulas is particularly challenging. We explore the sampling problem in detail and derive closed-form solutions for certain classes of copulas. For the general case, approximate methods are introduced and carefully evaluated.

Two important types of financial risk are considered in this thesis: systemic and operational risk. For the assessment of the systemic relevance of financial institutions, we propose to analyze the interconnectedness of the institutions in the market using multivariate copulas. For these, we derive new conditional sampling procedures, which we exploit to conduct a stress test of the market for credit default swaps.

Finally, we develop a flexible dependence model for quantitative operational risk management. The model components are investigated with regard to a range of relevant properties and appropriate recommendations are given. Based on operational loss data from Italian banks, we are then able to assess the effect of the modeling decisions on the operational risk capital.

Acknowledgments

First of all, I would like to thank Prof. Claudia Czado for the excellent supervision over the last three years. My work greatly benefited from our fruitful discussions, her constructive advice and her continuous encouragement. I very much appreciate that she provided me the freedom to develop my own ideas. Furthermore, I am particularly grateful to her for giving me the opportunity to participate in many scientific conferences and workshops and to present my work there. This allowed me to exchange with many great researchers, who I like to thank for helpful discussions and feedback.

It is also a particular pleasure for me to thank Prof. Harry Joe for the kind invitation to Vancouver. Working with him was very inspiring and fruitful. I also would like to thank him and Prof. Dorota Kurowicka for acting as referees of this thesis.

My special thanks go to Prof. Sandra Paterlini for the invitation to Modena and for the pleasant collaboration on developing the operational risk model. Likewise, I would like to thank Prof. Carole Bernard for inviting me to Waterloo and introducing me to the topic of systemic risk assessment in the first place.

Moreover, I am grateful to my colleagues at the Chair of Mathematical Statistics for the valuable exchange of ideas and for the enjoyable time we spent together.

Financial support through a scholarship from Allianz Deutschland AG and through the TUM Graduate School's International School of Applied Mathematics is gratefully acknowledged. Many of the numerical computations for this thesis were performed on a Linux cluster supported by DFG grant INST 95/919-1 FUGG.

Last but not least, I would like to sincerely thank my family and, most importantly, Susanne for her loving support.

Contents

1 Introduction

The modeling of dependencies among quantities of interest is an important topic in many areas such as finance and actuarial science but also in the natural and social sciences. Only the accurate measurement of joint probabilities allows for a diligent assessment and management of critical events. Especially joint tail probabilities, which characterize the joint behavior of variables in extreme situations, need to be thoroughly evaluated, as they importantly influence decision making.

Classically, the multivariate normal distribution has been central to statistical dependence modeling. Dependencies are then specified in terms of correlation coefficients, which however only measure the linear dependence of variables. Moreover, dependencies in the tails are not appropriately accounted for (see McNeil et al. (2005)). Today, it is therefore common to use copulas for dependence modeling. According to the famous theorem of Sklar (1959), any multivariate distribution function can be expressed in terms of its marginal distribution functions and a copula, which is a multivariate distribution function on the unit hypercube with uniformly distributed margins, and which contains all information on the dependence structure. From a modeling perspective, this hence allows to construct flexible multivariate distributions by individually combining different margins and a suitably chosen copula, which specifies the between-variable dependence structure.

While many different copulas with appealing properties are available and well-investigated in the bivariate case (see Joe (1997) and Nelsen (2006)), standard multivariate copulas are often rather restrictive or have an excessive number of parameters. The Gaussian copula, which is derived from the multivariate normal distribution, therefore still is a popular choice in higher-dimensional applications—but is also frequently criticized for its limitations (see Salmon (2009) for a critical discussion about the role of the Gaussian copula in the financial crisis of 2007–2009). While the Student's t copula, which is similarly derived from the multivariate Student's t distribution, may add some flexibility, in particular with respect to the handling of the tails, the need for more flexible dependence models is strong. However, not only flexibility but also parsimony is of particular importance here, since it ensures that a model stays interpretable and computationally tractable also in higher dimensions.

One of the most promising approaches to flexible multivariate dependence modeling is the concept of pair copula constructions, as originally proposed by Joe (1996, 1997). A vine copula, which is a graph theoretical model to define such a pair copula construction (see Bedford and Cooke (2001, 2002)), is built up by a quadratic number of bivariate copulas as building blocks. Since the bivariate copulas can be of arbitrary types, highly flexible multivariate copulas can be constructed. Vine copulas are however generally non-parsimonious, so that model selection tools such as truncation are needed to reduce the model complexity (see Brechmann et al. (2012)). Moreover, the interpretation of the intertwined model components may be difficult especially in higher dimensions.

In this thesis, we explore an alternative approach: In order to balance model flexibility and parsimony, we propose a hierarchical construction, which yields an inherently more parsimonious model. Variables are grouped in different hierarchical levels and the distribution of the groups (or clusters) of variables is specified in terms of lower-dimensional copulas. This within-group information is aggregated into univariate quantities, in terms of which between-group dependence is then quantified at the next level. In particular, we propose to use the Kendall distribution function to aggregate the groups, since it is the multivariate analog to the probability integral transform for univariate random variables and therefore naturally mimics the theorem of Sklar (1959) for multivariate margins. For this reason, we refer to the model as "hierarchical Kendall copula".

Obviously, such hierarchical Kendall copulas are straightforward to interpret in terms of within- and between-group dependence. Furthermore, the lower-dimensional copulas as building blocks can be copulas of arbitrary types as in a vine copula, so that within- and between-group dependence can be specified quite flexibly. Of course, such an approach is particularly appealing, when variables exhibit a natural hierarchical structure. Nevertheless, even if this is not the case, hierarchical Kendall copulas (with appropriately selected groups) may be used as a parsimonious and potentially flexible multivariate dependence model.

As a newly proposed statistical model, the properties and special cases of hierarchical Kendall copulas are investigated and illustrations are given. Most importantly, we derive the density of a hierarchical Kendall copula, which is of convenient form and therefore facilitates likelihood-based inference. Although Archimedean copulas themselves are typically inappropriate for higher-dimensional dependence modeling, they are attractive choices for the copulas of groups of variables in a hierarchical Kendall copula, since calculations then turn out to be rather straightforward.

Sampling is however challenging even in the case of Archimedean building blocks. The problem of sampling from hierarchical Kendall copulas essentially boils down to sampling from a random vector given that it lies in a particular level set of its copula, which is generally a difficult problem. We derive closed-form solutions for Archimedean copulas, for the copula by Plackett (1965) as well as for Archimax copulas (see Capéraà et al. (2000)), of which the popular extreme value copulas are a special case. For other copulas, we propose three approximate approaches, which are compared in a simulation study. Furthermore, we develop tools for model selection of hierarchical Kendall copulas and analyze the effect of copula misspecification.

As noted above, an accurate assessment of dependencies is very important in many areas. In light of the financial crisis of 2007–2009 and the Western sovereign debt crisis, this especially applies to the banking and insurance sector, which is heavily reliant on a prudent risk management. In this thesis, we analyze both classical and more recently proposed multivariate copulas for the modeling of two types of financial risk that have attracted considerable attention recently: systemic and operational risk. In addition, we also look at the market risk of stock portfolios.

The risk of a loss due to changes in the market price of a portfolio of assets is referred to as market risk. Clearly, the market risk is strongly influenced by the interdependencies of the assets: the more diversified the portfolio, the smaller the market risk. To measure these interdependencies, a statistical model needs to be set up. This model is then used

to determine the required market risk capital to be held in order to withstand extreme losses. We consider an equity portfolio of the 30 constituents of the most important German stock market index DAX and analyze it using hierarchical and non-hierarchical copulas. Especially hierarchical models can conveniently exploit the grouping of stocks according to industry sectors. In a forecasting study, we show how to forecast the one day ahead portfolio Value-at-Risk on a daily basis using the respective underlying multivariate copula. This allows for an assessment whether the chosen dependence model produces adequate risk capital figures. It is shown that this is, in fact, the case for appropriately selected hierarchical Kendall copulas.

The notion of systemic relevance of financial institutions is central in the discussion about lessons learned from the financial crisis (see Financial Stability Board et al. (2009)). Both the banking as well as the insurance industry are dealing with this issue, in an effort to identify systemically important institutions and reduce the systemic risk in the international financial market. As systemic risk is closely related to the interconnectedness of institutions, we develop copula-based methods for stress testing in order to analyze contagion effects among financial institutions. For this purpose, we derive new conditional simulation algorithms for the individual Student's t copula by Luo and Shevchenko (2010) and for Archimedean and vine copulas, which then also facilitate conditional sampling from a hierarchical Kendall copula. Such a hierarchical copula arises as a natural dependence model in our case study of credit default swap spreads of 38 important financial institutions from all over the world. Using different multivariate copulas (also non-hierarchical ones), we then carry out a systemic risk stress testing exercise and gain new insights into the systemic relevance of the institutions.

Finally, operational risk covers a diverse range of risks, which are mainly due to failed or inappropriate internal processes (see Basel Committee on Banking Supervision (2006)), and which are typically classified with respect to a range of different event types. Examples are losses incurred through fraud or system failures. Similar to market or credit risk, financial institutions are required to set aside capital to cover such losses. Because of data heterogeneity and scarcity, the measurement of operational risk is however difficult. In particular, it is still not fully analyzed how dependencies among different business lines and event types are characterized and how they influence risk capital figures. Therefore, we develop a model for quantitative operational risk management, which explicitly takes into account data scarcity and allows to flexibly model heterogeneous pairwise dependence (in the tails) of the losses. We carefully discuss the modeling challenges and identify reasonable choices for the model components. Especially the individual Student's t copula turns out to constitute an appealing model in the proposed framework. Using real-world data from Italian banks, we are then able to determine the impact of explicit dependence modeling among operational loss categories on risk capital figures.

Outline of the thesis

In **Chapter 2** we provide the necessary background for the rest of the thesis. We state the definition of a copula and of related quantities, which allow to characterize the dependence among random variables. In the following, we discuss relevant classes of copulas and state

their properties: elliptical copulas, the individual Student's t copula, Archimedean copulas, extreme value and Archimax copulas, the Plackett copula, and finally vine copulas, which are treated in detail.

Hierarchical Kendall copulas are introduced in **Chapter 3**, which is mainly based on Brechmann (2013a). We first discuss different choices of aggregation functions and argue why we believe that the Kendall distribution function is a reasonable choice for the purpose of aggregating groups of random variables. After stating the model definition, we investigate properties and special cases of hierarchical Kendall copulas and derive an explicit expression for the density. Furthermore, illustrative examples are provided and the model is compared with hierarchical Archimedean copulas, which are constructed in a similar way.

In the next step, statistical inference techniques for hierarchical Kendall copulas are developed. A general sampling algorithm is stated and appropriate estimation methods are discussed. In particular, we propose a sequential estimation algorithm, which is evaluated in a simulation study. To stabilize numerical calculations, we suggest to use a simple transformation in the aggregation step. Subsequently, we treat tools for the selection of appropriate groups using hierarchical clustering, for which we propose a suitable metric, and for the sequential selection of copulas. The risk of copula misspecification is investigated in a simulation study. Finally, the results of the in-sample market risk analysis of the stock market index DAX are presented.

Chapter 4 is devoted entirely to the problem of sampling from hierarchical Kendall copulas. It is mainly based on Brechmann (2013b) with some material taken from Brechmann (2013a). The problem can be solved using top-down and bottom-up approaches. We first discuss top-down sampling from a general perspective and derive closed-form solutions for Archimedean and Archimax copulas as well as for the Plackett copula. Alternatively, a procedure for rejection-like sampling is proposed. In the following, we introduce and discuss two methods for bottom-up sampling: sample reordering and density resampling, which can also be used for hierarchical Kendall copulas with arbitrary building blocks.

Since the two bottom-up sampling methods and top-down rejection-like sampling are approximate approaches, we assess them in a simulation study. Based on a range of evaluation criteria, the methods and different choices of control parameters are compared and recommendations for the practical use of them are derived. With the sampling procedures at hand, the market risk study of the previous chapter is then continued. It is shown how to forecast the one day ahead portfolio Value-at-Risk, in terms of which different dependence models are compared according to appropriate tests.

In **Chapter 5**, which is mainly based on Brechmann, Hendrich, and Czado (2013), we develop methods for stress testing financial institutions to assess their systemic relevance. As the proposed methodology requires the conditional simulation from copulas, we discuss appropriate approaches for the copulas considered in this thesis. While the cases of the Gaussian and of the Student's t copula are well-known, we derive new methods for the individual Student's t copula and for Archimedean and vine copulas as well as for hierarchical Kendall copulas. In the case study, the methodology is then used to conduct a systemic risk stress test of 38 important financial institutions.

Chapter 6 is based Brechmann, Czado, and Paterlini (2013) and presents our new model for quantitative operational risk management. We develop a zero-inflated depen-

dence model and carefully discuss the different model components. In particular, the copula choice is examined in terms of four relevant properties that a flexible model for operational losses should exhibit. Further, the zero-inflation components of the model have to be modeled by a multivariate binary distribution, for which we also propose a copula approach. The impact of the modeling decisions is then investigated in terms of risk capital figures for operational losses of Italian banks.

Finally, **Chapter 7** provides a brief conclusion and mentions two specific directions of future research: the relationship of hierarchical Kendall copulas to multivariate return periods and the hierarchical dependence modeling using a factor approach.

2 Preliminaries

In this chapter, we present a range of concepts that are used throughout the thesis. Most importantly, we define copulas and describe important properties. Thereby, we mainly follow the reference books by Joe (1997) and by Nelsen (2006). We then discuss popular classes of copulas and finally introduce vine copulas as a mean to construct flexible higher-dimensional copulas.

2.1 Copulas and dependence measures

A d-dimensional copula is a multivariate distribution function on the unit hypercube, $[0,1]^d$, with uniformly distributed margins. Copulas arise as the natural tool for statistical dependence modeling through Sklar's Theorem.

Theorem 2.1 (Sklar, 1959)**.** Let $\boldsymbol{X} = (X_1, ..., X_d)' \sim F$, where $X_j \sim F_j,\ j = 1, ..., d$. Then there exists a d-dimensional copula C such that

$$F(\boldsymbol{x}) = C(F_1(x_1), ..., F_d(x_d)), \quad \boldsymbol{x} := (x_1, ..., x_d)' \in (\mathbb{R} \cup \{-\infty, \infty\})^d. \tag{2.1}$$

If $F_1, ..., F_d$ are continuous, then C is unique. Conversely, if C is a d-dimensional copula and $F_1, ..., F_d$ are distribution functions, then F defined by (2.1) is a d-dimensional distribution function with marginal distribution functions $F_1, ..., F_d$.

Proof: See Nelsen (2006, Theorem 2.10.9). □

Sklar's Theorem hence establishes the link between multivariate distribution functions and their univariate margins. All information about the dependence among the variables is captured by the copula. In this thesis, we assume that F is absolutely continuous and $F_1, ..., F_d$ are strictly increasing. Then it holds for the d-dimensional density f of F and the univariate densities f_j of $F_j,\ j = 1, ..., d$, that

$$f(\boldsymbol{x}) = c(F_1(x_1), ..., F_d(x_d)) \prod_{j=1}^{d} f_j(x_j), \quad \boldsymbol{x} \in (\mathbb{R} \cup \{-\infty, \infty\})^d, \tag{2.2}$$

where c is the density of the copula C.

If the components of $\boldsymbol{X}$ are independent, the corresponding copula is called the *independence copula*.

Example 2.2 (Independence copula)**.** The independence copula is defined as

$$\Pi(\boldsymbol{u}) = \prod_{j=1}^{d} u_j, \quad \boldsymbol{u} := (u_1, ..., u_d)' \in [0,1]^d.$$

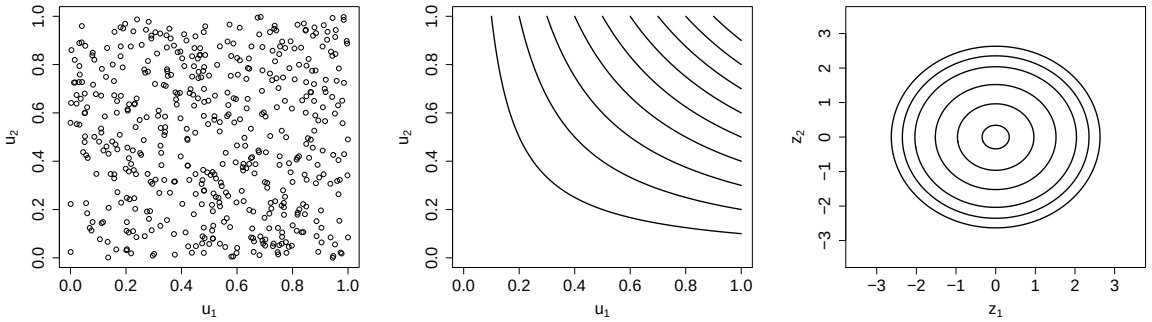


Figure 2.1: A scatter plot of a sample from the bivariate independence copula (left panel), contour lines of the bivariate independence copula (middle panel), and contour lines of its density combined with standard normal margins (right panel). At the points on a contour line, the copula and its density with standard normal margins are constant, respectively (see Appendix A for more details).

It obviously has uniform margins and density π given by

$$\pi(\boldsymbol{u}) = 1, \quad \boldsymbol{u} \in [0,1]^d.$$

The independence copula and its density are illustrated in Figure 2.1.

It holds that, if $\boldsymbol{X}$ has the copula C, then $X_1, ..., X_d$ are independent if and only if $C = \Pi$. □

The two important boundary cases of counter- and comonotonicity (perfect negative and positive dependence, respectively) are given through the Fréchet-Hoeffding bounds.

Theorem 2.3 (Fréchet-Hoeffding bounds.)**.** Let C be a d-dimensional copula. Then it holds

$$W(\boldsymbol{u}) \leq C(\boldsymbol{u}) \leq M(\boldsymbol{u}) \quad \forall \boldsymbol{u} \in [0,1]^d,$$

where

$$W(\boldsymbol{u}) = \max\{u_1 + ... + u_d - d + 1, 0\},$$

and

$$M(\boldsymbol{u}) = \min\{u_1, ..., u_d\}. \tag{2.3}$$

Proof: See Nelsen (2006, Theorem 2.10.12) □

The upper Fréchet-Hoeffding bound M is a copula, the *comonotonicity copula*. If $\boldsymbol{U} = (U_1, ..., U_d)' \sim M$, then it holds that $P(U_1 = U_2 = ... = U_d) = 1$. The lower Fréchet-Hoeffding bound W is however only a copula if $d = 2$. In this case, it is called the *countermonotonicity copula* and it holds that $P(U_1 = -U_2) = 1$ if $(U_1, U_2)' \sim W$. Nevertheless, W is the best possible lower bound in any dimension (see Nelsen (2006, Theorem 2.10.13)).

Copulas can also be characterized with respect to different notions of symmetry. Here, we consider two such notions: reflection symmetry and exchangeability.

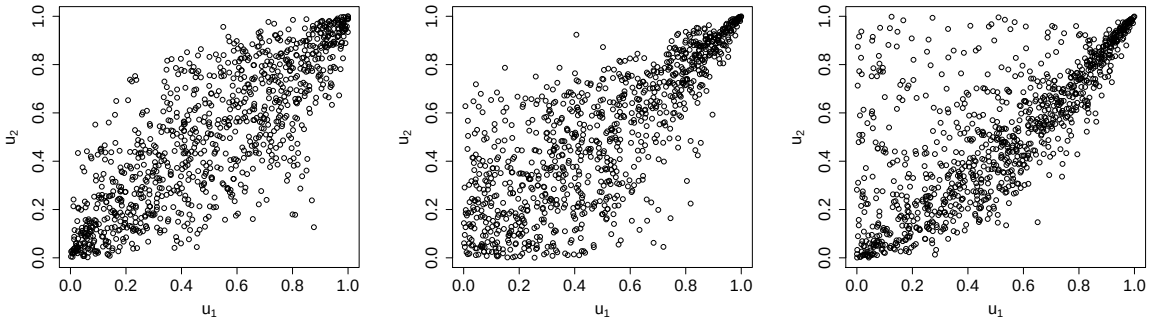


Figure 2.2: Scatter plots of data simulated from a reflection symmetric and exchangeable copula (left panel), from an exchangeable copula, which is not reflection symmetric (middle panel), and from a copula, which is neither reflection symmetric nor exchangeable (right panel).

Definition 2.4 (Reflection symmetry)**.** A copula C is called *reflection symmetric* (or *radially symmetric*) if it follows from $\boldsymbol{U} \sim C$ that also $\mathbf{1} - \boldsymbol{U} \sim C$, where $\mathbf{1} := (1, ..., 1)'$.

Definition 2.5 (Exchangeability)**.** A copula C is called *exchangeable* (or *permutation symmetric*) if $\boldsymbol{U} \sim C$ implies that also $(U_{\sigma(1)}, ..., U_{\sigma(d)})' \sim C$ for any permutation $\sigma : \{1, ..., d\} \to \{1, ..., d\}$.

In the bivariate case, we will refer to an exchangeable copula also as *symmetric* copula, because it holds for an exchangeable bivariate copula that

$$C(u_1, u_2) = C(u_2, u_1) \quad \forall (u_1, u_2)' \in [0, 1]^2.$$

The two notions of symmetry are illustrated in Figure 2.2.

To summarize the dependence information among two variables $(X_1, X_2)' \sim F$ in single numbers, it is common to use association measures. The most common ones are *Kendall's* τ (Kendall, 1938) and *Spearman's* ρ_S (Spearman, 1904). Both depend only on the copula C of $(X_1, X_2)'$ (see Nelsen (2006, Theorems 5.1.3 and 5.1.6)).

Remark 2.6 (Kendall's τ and Spearman's ρ_S)**.** Kendall's τ is given by

$$\tau(C) = 4 \int_{[0,1]^2} C(u_1, u_2)\, dC(u_1, u_2) - 1, \tag{2.4}$$

and Spearman's ρ_S by

$$\rho_S(C) = 12 \int_{[0,1]^2} C(u_1, u_2)\, du_1\, du_2 - 3 = 12 \int_{[0,1]^2} u_1 u_2\, dC(u_1, u_2) - 3. \tag{2.5}$$

It holds that $\tau(C) \in [-1, 1]$ and $\rho_S(C) \in [-1, 1]$. □

Due to these relationships, parameters of copulas are often calibrated according to a specific value of Kendall's τ or Spearman's ρ_S.

The dependence in the tails of the joint distribution can be characterized using the *lower* and *upper tail dependence coefficients*, which are also purely copula-based measures (see Joe (1993) and Nelsen (2006, Theorem 5.4.2)). They measure the strength of dependence in the lower-left and upper-right quadrant of $[0,1]^2$, respectively.

Remark 2.7 (Tail dependence coefficients)**.** The lower tail dependence coefficient is given as

$$\lambda_L(C) = \lim_{t \downarrow 0} P(X_2 \leq F_2^{-1}(t) | X_1 \leq F_1^{-1}(t)) = \lim_{t \downarrow 0} P(U_2 \leq t | U_1 \leq t) = \lim_{t \downarrow 0} \frac{C(t,t)}{t}, \quad (2.6)$$

where we used that, according to the probability integral transform, $U_j = F(X_j) \sim U(0,1)$, $j = 1, 2$, and $(U_1, U_2)' \sim C$. Similarly, the upper tail dependence coefficient is given as

$$\begin{aligned} \lambda_U(C) &= \lim_{t \uparrow 1} P(X_2 > F_2^{-1}(t) | X_1 > F_1^{-1}(t)) = \lim_{t \uparrow 1} P(U_2 > t | U_1 > t) \\ &= \lim_{t \uparrow 1} \frac{1 - 2t + C(t,t)}{1-t} = 2 - \lim_{t \uparrow 1} \frac{1 - C(t,t)}{1-t}. \end{aligned} \quad (2.7)$$

Since $\lambda_L(C)$ and $\lambda_U(C)$ are both probabilities, it holds that $\lambda_L(C), \lambda_U(C) \in [0,1]$. □

A copula C is said to be lower (upper) tail dependent if $\lambda_L(C) > 0$ ($\lambda_U(C) > 0$). Otherwise, the copula is lower (upper) tail independent. If the copula is reflection symmetric, the two tail dependence coefficients coincide. An additional characterization of the tail behavior of copulas using the notion of the tail order is given by Hua and Joe (2011).

Alternative concepts of dependence can be found in Joe (1997, Chapter 2). One such notion is TP_2 dependence, which provides a quite specific characterization of positive dependence.

Definition 2.8 (TP_2 dependence)**.** A bivariate copula density c is called *totally positive of order 2* (TP_2) if

$$c(u_1, u_2)\, c(w_1, w_2) \geq c(u_1, w_2)\, c(w_1, u_2),$$

for all $(u_1, u_2)', (w_1, w_2)' \in [0,1]^2$ with $w_1 > u_1$ and $w_2 > u_2$.

If a copula has a TP_2 density, this means that there is a higher likelihood of observing a pair with low values and one with high values than two pairs with low and high values. TP_2 dependence is a rather strong notion of positive dependence, since it implies a range of other concepts (see Joe (1997, Theorem 2.3)).

If a specific bivariate copula can only model positive dependence, or if it is lower tail dependent but upper tail independent, it can be rotated to obtain a new copula with negative dependence or with upper tail dependence, respectively. There are three possible rotations, which we define counterclockwisely.

- *Rotation by* 90 *degrees:* Let $(U_1, U_2)' \sim C_{90}$. If $(1 - U_1, U_2)' \sim C$, then C_{90} is the copula rotated by 90 degrees. It is given by

$$C_{90}(u_1, u_2) = u_2 - C(1 - u_1, u_2), \quad (u_1, u_2)' \in [0,1]^2.$$

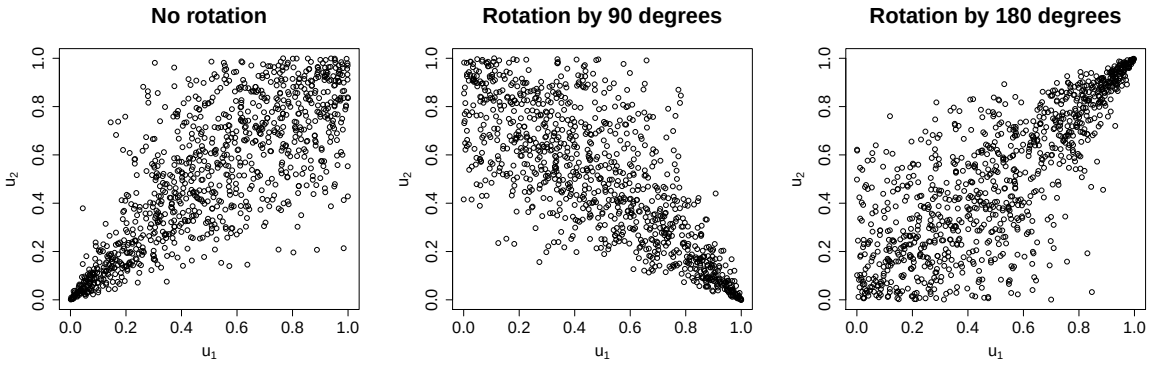


Figure 2.3: Scatter plots of data simulated from a copula (left panel) and its rotations by 90 degrees (middle panel) and 180 degrees (right panel).

- *Rotation by* 180 *degrees:* Let $(U_1, U_2)' \sim C_{180}$. If $(1-U_1, 1-U_2)' \sim C$, then C_{180} is the copula rotated by 180 degrees or the *survival copula*. It is given by

$$C_{180}(u_1, u_2) = u_1 + u_2 - 1 + C(1-u_1, 1-u_2), \quad (u_1, u_2)' \in [0,1]^2.$$

- *Rotation by* 270 *degrees:* Let $(U_1, U_2)' \sim C_{270}$. If $(U_1, 1-U_2)' \sim C$, then C_{270} is the copula rotated by 270 degrees. It is given by

$$C_{270}(u_1, u_2) = u_1 - C(u_1, 1-u_2), \quad (u_1, u_2)' \in [0,1]^2.$$

Clearly, if a copula is reflection symmetric, then it coincides with its survival version. The rotation of copulas is illustrated in Figure 2.3

Finally, we consider one more quantity that is closely related to each copula. In the univariate case it is known that $U_1 = F_1(X_1) \sim U(0,1)$, as already used above. Multivariate distribution functions also provide a mapping to $[0,1]$, but $F(\boldsymbol{X})$ is not uniform in general. Its distribution can be characterized using the notion of the *Kendall distribution function*.

Definition 2.9 (Kendall distribution function)**.** The Kendall distribution function of a copula C is defined as

$$K(z; C) = P(C(\boldsymbol{U}) \leq z), \quad z \in [0,1], \tag{2.8}$$

where $\boldsymbol{U} \sim C$.

It follows that the distribution of $F(\boldsymbol{X})$ is the Kendall distribution function of the corresponding copula C,

$$\begin{aligned} P(F(\boldsymbol{X}) \leq z) &= P(C(F_1(X_1), ..., F_d(X_d)) \leq z) \\ &= P(C(\boldsymbol{U}) \leq z) = K(z; C), \quad z \in [0,1], \end{aligned}$$

where $U_j = F_j(X_j),\ j = 1, ..., d$.

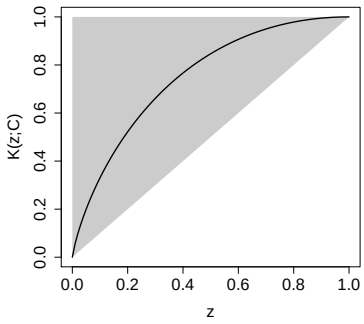


Figure 2.4: Example of a Kendall distribution function. The gray area illustrates the lower and upper bounds of the Kendall distribution function.

Kendall distribution functions were first studied in two dimensions by Genest and Rivest (1993) and in more generality by Barbe et al. (1996). It holds that $\lim_{z\uparrow 0} K(z;C) = 0$ as well as

$$z \leq K(z;C) \leq 1, \quad z \in [0,1]. \tag{2.9}$$

While the second inequality holds simply by the definition of a distribution function, the first is a consequence of Theorem 2.10 and Corollary 2.11 stated below. The lower bound is, in fact, the Kendall distribution function of the comonotonicity copula: If $\boldsymbol{U} \sim M$, then

$$K(z;M) = P(M(\boldsymbol{U}) \leq z) = P(U_1 \leq z) = z, \quad z \in [0,1].$$

On the other hand, the upper bound corresponds to the extreme case of perfect negative dependence, where the Kendall distribution function is constant at 1, that is, $K(z;W) = 1$ for all $z \in [0,1]$. This is illustrated in Figure 2.4, which shows an example of a Kendall distribution function.

It immediately follows from Equation (2.8) that the Kendall distribution function describes the distribution of the *level sets* of a copula,

$$L(z;C) = \{\boldsymbol{u} \in [0,1]^d : C(\boldsymbol{u}) = z\}, \quad z \in (0,1). \tag{2.10}$$

They can be used to illustrate copulas (see Figure 2.1 and Appendix A). In the bivariate case, level sets are also called *contour lines*.

The computation of the Kendall distribution function for a given copula is however complicated in general. Imlahi et al. (1999) provide a recursive formula, for which we need to introduce the notion of the copula quantile function, as studied also in Chakak and Ezzerg (2000). Define $C(\cdot|u_1,...,u_{d-1}) := C(u_1,...,u_{d-1},\cdot)$, then the *copula quantile function* is the inverse $C^{-1}(\cdot|u_1,...,u_{d-1})$. It holds that

$$C(u_1,...,u_{d-1},C^{-1}(z|u_1,...,u_{d-1})) = z$$

for $z \in (0,1)$. For ease of notation, we define $C(\cdot|u_1,...,u_r) := C(u_1,...,u_r,\cdot,1,...,1)$ for $r = 1,...,d-2$, and $C^{-1}(z|\emptyset) := z$ for $z \in (0,1)$. Note that this notation is different from that in Imlahi et al. (1999), where C^{-1} is used to denote the copula level set $L(\cdot;C)$ and the copula quantile function is denoted by ψ.

Theorem 2.10 (Recursive formula of the Kendall distribution function)**.** For the Kendall distribution function of the d-dimensional copula C, it holds for $z \in [0,1]$ that

$$K(z;C) = K(z;C_{1,\dots,d-1}) + \int_z^1 \int_{C^{-1}(z|u_1)}^1 \dots \int_{C^{-1}(z|u_1,\dots,u_{d-2})}^1 \int_0^{C^{-1}(z|u_1,\dots,u_{d-1})} c(u_1,\dots,u_d)\, du_d \dots du_1, \tag{2.11}$$

where c is the copula density and $C_{1,\dots,d-1}$ is the copula of the first $d-1$ variables. If $d=2$, then $C_{1,\dots,d-1} = C_1$ is the distribution function of U_1, which is uniform. Therefore, $K(z;C_1) = z$ for all $z \in [0,1]$.

Proof: See Imlahi et al. (1999, Proposition 1). As an illustration, we show here the case $d=2$. It holds that

$$K(z;C) = P(C(U_1,U_2) \le z) = P(C(U_1,U_2) \le z, U_1 \le z) + P(C(U_1,U_2) \le z, U_1 > z).$$

Since $U_1 \le z$ implies $C(U_1,U_2) \le z$, we get for the first term that

$$P(C(U_1,U_2) \le z, U_1 \le z) = P(U_1 \le z) = z = K(z;C_1).$$

For the second term we calculate

$$\begin{aligned} P(C(U_1,U_2) \le z, U_1 > z) &= \int_z^1 P(C(U_1,U_2) \le z | U_1 = u_1)\, du_1 \\ &= \int_z^1 P(U_2 \le C^{-1}(z|u_1) | U_1 = u_1)\, du_1 \\ &= \int_z^1 \int_0^{C^{-1}(z|u_1)} c(u_1,u_2)\, du_2\, du_1, \end{aligned} \tag{2.12}$$

which proves Equation (2.11) in the bivariate case. □

An immediate corollary of Theorem 2.10 is that the Kendall distribution function of a d-dimensional copula is, for fixed $z \in [0,1]$, increasing in the dimension d.

Corollary 2.11 (Monotonicity of the Kendall distribution function)**.** For the Kendall distribution function of the d-dimensional copula C, it holds that

$$K(z;C) \ge K(z;C_{1,\dots,d-1}) \quad \forall z \in [0,1],$$

where $C_{1,\dots,d-1}$ is the copula of the first $d-1$ variables.

Proof: The integrand in Equation (2.11) is the copula density. Since densities are positive, the integral is positive and it holds that $K(z;C) - K(z;C_{1,\dots,d-1}) \ge 0$. □

It directly follows from this result that $K(z;C) \ge K(z;C_1) = z$, which proves the lower bound in Equation (2.9).

Equation (2.11) 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. A convenient exception are Archimedean copulas, as will be discussed in Section 2.4. There, we will also derive the Kendall distribution function of the multivariate independence copula.

Even in the bivariate case, the calculation of the Kendall distribution function can be challenging if the copula quantile function is not known. At least, Equation (2.11) can be simplified using the following notation for the derivative of a bivariate copula with respect to one of its arguments:

$$C_{1|2}(u_1|u_2) := \frac{\partial C(u_1, u_2)}{\partial u_2}, \quad \text{and} \quad C_{2|1}(u_2|u_1) := \frac{\partial C(u_1, u_2)}{\partial u_1}, \quad (u_1, u_2)' \in [0,1]^2. \tag{2.13}$$

Obviously, it holds that

$$C_{1|2}(u_1|u_2) = P(U_1 \leq u_1 | U_2 = u_2), \quad \text{and} \quad C_{2|1}(u_2|u_1) = P(U_2 \leq u_2 | U_1 = u_1). \tag{2.14}$$

These conditional distribution functions play a major role in the construction of vine copulas (see Section 2.7).

Using this notation and Equation (2.12), we obtain for the Kendall distribution function of a bivariate copula C and $z \in [0,1]$ that

$$K(z; C) = z + \int_z^1 C_{2|1}(C^{-1}(z|u_1)|u_1)\, du_1. \tag{2.15}$$

This is equivalent to an expression provided by Genest and Rivest (2001).

There is also a notable connection between Kendall's τ and the Kendall distribution function.

Remark 2.12 (Kendall's τ and the Kendall distribution function)**.** It holds that

$$\tau(C) \overset{(2.4)}{=} 4E(C(U_1, U_2)) - 1 = 4 \int_0^1 z\, dK(z; C) - 1 = 3 - 4 \int_0^1 K(z; C)\, dz, \tag{2.16}$$

where the last equality is obtained through integration by parts. □

In the following, we now present and discuss the three most popular classes of copulas: elliptical, Archimedean and extreme value copulas. In addition, the individual Student's t copula, which extends the popular standard Student's t copula, and the Plackett copula, which does not belong to either one of these three classes, are introduced. If known, expressions for the Kendall's distribution function, Kendall's τ, Spearman's ρ_S and the tail dependence coefficients are provided. We concentrate on copulas that can model the full range of positive (and negative) dependence, that is, copulas with $\tau(C), \rho_S(C) \in (0,1)$ or even $\tau(C), \rho_S(C) \in (-1,1)$. Table 2.1 at the end of this chapter summarizes the most important properties of the copulas and Appendix A provides graphical illustrations similar to Figure 2.1.

2.2 Elliptical copulas

Elliptical copulas are very popular and often used, since they can be used also in higher dimensions and are straightforward to interpret. They arise by inversion of Sklar's Theorem (2.1), which shows that copulas can be constructed for arbitrary multivariate distribution functions F and marginal distribution functions $F_1, ..., F_d$ as

$$C(\boldsymbol{u}) = F(F_1^{-1}(u_1), ..., F_d^{-1}(u_d)), \quad \boldsymbol{u} \in [0,1]^d. \tag{2.17}$$

Elliptical copulas are obtained by letting F be an elliptical distribution function and $F_1, ..., F_d$ the corresponding margins (see Fang et al. (1990), Frahm et al. (2003) and McNeil et al. (2005)). The density of an elliptical copula is

$$c(\boldsymbol{u}) = \frac{f(F_1^{-1}(u_1), ..., F_d^{-1}(u_d))}{\prod_{j=1}^d f_j(F_j^{-1}(u_j))}, \quad \boldsymbol{u} \in [0,1]^d, \tag{2.18}$$

which holds for any copula constructed according to Equation (2.17). For the Kendall distribution function of elliptical copulas, however no closed-form expression is available, which is mainly due to the fact that an elliptical distribution function typically involves higher-dimensional integration.

It holds that all elliptical distributions, and hence also the derived copulas, are reflection symmetric. Therefore, lower and upper tail dependence coefficients of elliptical copulas coincide. The most popular examples of elliptical copulas are the tail independent *Gaussian* and the tail dependent *Student's t copula*.

Example 2.13 (Gaussian copula)**.** Let $\Phi_{\boldsymbol{\mu},\Sigma}$ denote the distribution function of the multivariate normal distribution $\mathcal{N}_d(\boldsymbol{\mu}, \Sigma)$ with mean $\boldsymbol{\mu} \in \mathbb{R}^d$ and positive definite covariance matrix $\Sigma \in \mathbb{R}^{d\times d}$. Its density is

$$\phi_{\boldsymbol{\mu},\Sigma}(\boldsymbol{x}) = (2\pi)^{-d/2}|\Sigma|^{-1/2} \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right), \quad \boldsymbol{x} \in \mathbb{R}^d. \tag{2.19}$$

Further, let Φ denote the distribution function of the standard normal distribution $\mathcal{N}_1(0,1)$ and write $\Phi_\Sigma := \Phi_{\mathbf{0},\Sigma}$ to shorten notation. Then, the Gaussian copula is defined as

$$C(\boldsymbol{u}; R) = \Phi_R\left(\Phi^{-1}(u_1), ..., \Phi^{-1}(u_d)\right), \quad \boldsymbol{u} \in [0,1]^d,$$

where $R = (\rho_{jk})_{j,k=1,...,d} \in [-1,1]^{d\times d}$ is a correlation matrix. The Gaussian copula therefore has $d(d-1)/2$ parameters, unless a specific structure of the correlation matrix is assumed. In the case of an exchangeable correlation matrix with $\rho_{jk} = \rho \in (-1/(d-1), 1)$ for all $j, k = 1, ..., d,\ j \neq k$, the Gaussian copula itself is exchangeable and has only one parameter. Other parameterizations such as an autoregressive structure are also feasible. If $\rho_{jk} = 1$ ($\rho_{jk} = 0$) for all $j, k = 1, ..., d,\ j \neq k$, then the Gaussian copula corresponds to the comonotonicity (independence) copula. The bivariate countermonotonicity copula is obtained for $\rho_{12} = -1$ if $d = 2$.

Using Equations (2.18) and (2.19), the density of the Gaussian copula is easily derived as

$$\begin{aligned} c(\boldsymbol{u}; R) &= \phi_R\left(\Phi^{-1}(u_1), ..., \Phi^{-1}(u_d)\right) \prod_{j=1}^{d} \phi(\Phi^{-1}(u_j))^{-1} \\ &= |R|^{-1/2} \exp\left(-\frac{1}{2}\boldsymbol{x}'(R^{-1} - I_d)\boldsymbol{x}\right), \quad \boldsymbol{u} \in [0,1]^d, \end{aligned} \tag{2.20}$$

where $x_j = \Phi^{-1}(u_j),\ j = 1, ..., d$. Further, ϕ is the density of the standard normal distribution $\mathcal{N}_1(0,1)$ and ϕ_R that of the $\mathcal{N}_d(\mathbf{0}, R)$ distribution, where $\mathbf{0} := (0, ..., 0)'$.

In the bivariate case, the Gaussian copula has only one parameter: the off-diagonal parameter of the correlation matrix R. It is typically denoted by $\rho = \rho_{12}$. Kendall's τ and Spearman's ρ_S can be expressed in terms of this parameter as

$$\tau(\rho) = \frac{2}{\pi} \arcsin(\rho), \quad \text{and} \tag{2.21}$$

$$\rho_S(\rho) = \frac{6}{\pi} \arcsin\left(\frac{\rho}{2}\right). \tag{2.22}$$

Due to the reflection symmetry of the Gaussian copula, lower and upper tail dependence coefficients are the same. They are both zero,

$$\lambda_L(\rho) = \lambda_U(\rho) = 0.$$

In other words, the Gaussian copula is tail independent. For $\rho \geq 0$, it however has a TP_2 density, which can be easily verified by plugging the density (2.20) for $d = 2$ into the definition of TP_2 dependence (see Definition 2.8).

In addition to reflection symmetry, the Gaussian copula is also symmetric in the bivariate case. Extensions to skew-elliptical distributions are not considered here (see, e.g., Genton (2004)). □

Example 2.14 (Student's t copula)**.** The Student's t copula is an example of a tail dependent elliptical copula—a property that is not surprising given the well-known property that the univariate Student's t distribution has heavier tails than the normal. The distribution function of the multivariate Student's t distribution $\mathcal{T}_d(\boldsymbol{\mu}, \Sigma, \nu)$ with mean $\boldsymbol{\mu} \in \mathbb{R}^d$, positive definite scale matrix $\Sigma \in \mathbb{R}^{d \times d}$ and ν degrees of freedom is denoted by $T_{\boldsymbol{\mu},\Sigma,\nu}$ and its density by $t_{\boldsymbol{\mu},\Sigma,\nu}$. The latter is given as

$$t_{\boldsymbol{\mu},\Sigma,\nu}(\boldsymbol{x}) = (\nu\pi)^{-d/2} |\Sigma|^{-1/2} \frac{\Gamma\left(\frac{\nu+d}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{1}{\nu}(\boldsymbol{x} - \boldsymbol{\mu})'\Sigma^{-1}(\boldsymbol{x} - \boldsymbol{\mu})\right)^{-(\nu+d)/2}, \quad \boldsymbol{x} \in \mathbb{R}^d,$$

where Γ is the gamma function (see Kotz and Nadarajah (2004) and Demarta and McNeil (2005)). We assume that $\nu > 2$, which ensures the existence of first and second moments. As before, we define $T_{\Sigma,\nu} := T_{\mathbf{0},\Sigma,\nu}$ and denote the distribution function of $\mathcal{T}_1(0, 1, \nu)$ by T_ν. Using this notation, the Student's t copula is given as

$$C(\boldsymbol{u}; R, \nu) = T_{R,\nu}\left(T_\nu^{-1}(u_1), ..., T_\nu^{-1}(u_d)\right), \quad \boldsymbol{u} \in [0,1]^d, \tag{2.23}$$

with correlation matrix $R \in [-1,1]^{d\times d}$ and density

$$\begin{aligned} c(\boldsymbol{u}; R, \nu) &= t_{R,\nu}\left(T_\nu^{-1}(u_1), ..., T_\nu^{-1}(u_d)\right) \prod_{j=1}^{d} t_\nu(T_\nu^{-1}(u_j))^{-1} \\ &= |R|^{-1/2} \frac{\Gamma\left(\frac{\nu+d}{2}\right)\Gamma\left(\frac{\nu}{2}\right)^{d-1}}{\Gamma\left(\frac{\nu+1}{2}\right)^{d}} \frac{\prod_{j=1}^{d}\left(1+\frac{x_j^2}{\nu}\right)^{(\nu+1)/2}}{\left(1+\frac{1}{\nu}\boldsymbol{x}'R^{-1}\boldsymbol{x}\right)^{(\nu+d)/2}}, \quad \boldsymbol{u} \in [0,1]^d, \end{aligned}$$

where $x_j = T_\nu^{-1}(u_j),\ j = 1, ..., d$, and t_ν is the density of the univariate $\mathcal{T}_1(0, 1, \nu)$ distribution and $t_{R,\nu}$ that of the $\mathcal{T}_d(\mathbf{0}, R, \nu)$ distribution. The copula has $d(d-1)/2$ correlation parameters and additionally the degrees of freedom parameter ν, that is, $d(d-1)/2+1$ parameters in total. Like the Gaussian copula, the Student's t copula is exchangeable if the correlation matrix is exchangeable. Comonotonicity is obtained if $\rho_{jk} = 1$ for all $j, k = 1, ..., d,\ j \neq k$, and countermonotonicity if $\rho_{12} = -1$ and $d = 2$. The independence copula is not a special case of the Student's t copula with $\nu < \infty$.

In the bivariate case, the Student's t copula is therefore characterized by two parameters: the correlation parameter, which is denoted by $\rho = \rho_{12}$, and ν. While there is no closed-form expression of Spearman's ρ_S in terms of the parameters, Kendall's τ is given by the same formula as for the Gaussian copula (see Equation (2.21)),

$$\tau(\rho, \nu) = \frac{2}{\pi}\arcsin(\rho), \tag{2.24}$$

which does not depend on the degrees of freedom ν. Similar to the Gaussian copula, the bivariate Student's t copula is also symmetric, but, as mentioned above, it has tail dependence. In particular, the lower and upper tail dependence coefficients are the same due to the reflection symmetry and given by

$$\lambda_L(\rho, \nu) = \lambda_U(\rho, \nu) = 2\, T_{\nu+1}\left(-\sqrt{\nu+1}\sqrt{\frac{1-\rho}{1+\rho}}\right). \tag{2.25}$$

It holds for fixed ρ that the larger the degrees of freedom ν are, the weaker the tail dependence is (see also the right panel of Figure 2.5 below). This is due to the fact that the Student's t copula converges to the Gaussian copula if $\nu \to \infty$. Compared to the Gaussian copula (see Figure A.3), the non-zero tail dependence coefficients leads to a sharper shape of the density contour lines of the Student's t copula in Figure A.4. □

Although the Student's t copula is tail dependent, a major disadvantage is that the tail dependence among pairs of variables is symmetric in both tails and governed by only one parameter, which limits the flexibility of the tail behavior. The first issue is a general disadvantage of elliptical copulas due to the reflection symmetry, but, to overcome the second issue, Luo and Shevchenko (2010) recently proposed an extension of the standard Student's t copula, which is discussed in the following section.

2.3 Individual Student's t copula

The *individual Student's t copula* by Luo and Shevchenko (2010) extends the Student's t copula by allowing for multiple degrees of freedom parameters, so that more flexibility

in modeling tail dependencies is achieved. A special case of this individual Student's t copula is the *grouped Student's t copula*, which was previously proposed by Daul et al. (2003).

Before we introduce both copulas, we note that the multivariate Student's t distribution is a variance mixture of normals. If $\boldsymbol{X} \sim \mathcal{T}_d(\boldsymbol{\mu}, \Sigma, \nu)$, it can be represented as

$$\boldsymbol{X} \overset{d}{=} \boldsymbol{\mu} + W\boldsymbol{Z} = \boldsymbol{\mu} + (WZ_1, ..., WZ_d)', \tag{2.26}$$

where $\boldsymbol{Z} := (Z_1, ..., Z_d)' \sim \mathcal{N}_d(\boldsymbol{0}, \Sigma)$ and the mixing variable W is independent of $\boldsymbol{Z}$ and satisfies $\nu/W^2 \sim \chi^2_\nu$. Luo and Shevchenko (2010) generalize this construction and define the individual Student's t distribution and copula.

As before, let $\boldsymbol{Z} \sim \mathcal{N}_d(\boldsymbol{0}, \Sigma)$. Further, let Q be uniformly distributed on $[0, 1]$ and independent of $\boldsymbol{Z}$. For constants $\nu_j > 2,\ j = 1, ..., d$, we then define

$$W_j = \sqrt{\nu_j / F^{-1}_{\chi^2}(Q; \nu_j)}, \quad j = 1, ..., d,$$

where $F^{-1}_{\chi^2}(\cdot; \nu)$ denotes the inverse χ^2 distribution function with ν degrees of freedom. This means that $W_1, ..., W_d$ are perfectly positively dependent. In addition, it holds that $F^{-1}_{\chi^2}(Q; \nu_j) \sim \chi^2_{\nu_j}$, so that we have $\nu_j/W_j^2 \sim \chi^2_{\nu_j}$ for all $j = 1, ..., d$.

The individual Student's t distribution $\mathcal{IT}_d(\boldsymbol{\mu}, \Sigma, \boldsymbol{\nu})$ with mean $\boldsymbol{\mu} \in \mathbb{R}^d$, positive definite scale matrix $\Sigma \in \mathbb{R}^{d\times d}$ and multiple degrees of freedom $\boldsymbol{\nu} = (\nu_1, ..., \nu_d)'$ is then defined as the distribution of the random vector $\boldsymbol{X}$ given by

$$\boldsymbol{X} := \boldsymbol{\mu} + (W_1 Z_1, ..., W_d Z_d)', \tag{2.27}$$

which generalizes Equation (2.26). The univariate margins of $\boldsymbol{X}$ follow univariate Student's t distributions with ν_j degrees of freedom, $j \in \{1, ..., d\}$.

The individual Student's t copula is then obtained by inverting Sklar's Theorem (see Equation (2.17)) for the $\mathcal{IT}_d(\boldsymbol{0}, R, \boldsymbol{\nu})$ distribution with correlation matrix $R \in [-1, 1]^{d\times d}$ and distribution function

$$\begin{aligned}
F_{\boldsymbol{X}}(\boldsymbol{x}; R, \boldsymbol{\nu}) &= \int_0^1 P(X_1 \le x_1, ..., X_d \le x_d | Q = q)\, dq \\
&= \int_0^1 P(W_1 Z_1 \le x_1, ..., W_d Z_d \le x_d | Q = q)\, dq \\
&= \int_0^1 P\left(Z_1 \le \frac{x_1}{w_1(q)}, ..., Z_d \le \frac{x_d}{w_d(q)}\right) dq \\
&= \int_0^1 \Phi_R\left(\frac{x_1}{w_1(q)}, ..., \frac{x_d}{w_d(q)}\right) dq,
\end{aligned}$$

where $w_j(q) = \sqrt{\nu_j / F^{-1}_{\chi^2}(q; \nu_j)},\ j = 1, ..., d$. Hence, the corresponding copula has the form

$$C(\boldsymbol{u}; R, \boldsymbol{\nu}) = \int_0^1 \Phi_R\left(\frac{x_1}{w_1(q)}, ..., \frac{x_d}{w_d(q)}\right) dq, \quad \boldsymbol{u} \in [0, 1]^d,$$

where $x_j = T_{\nu_j}^{-1}(u_j)$. Its density is given for $\boldsymbol{u} \in [0,1]^d$ by

$$\begin{aligned} c(\boldsymbol{u}; R, \boldsymbol{\nu}) = & \int_0^1 \phi_R \left(\frac{x_1}{w_1(q)}, ..., \frac{x_d}{w_d(q)} \right) \left(\prod_{j=1}^d w_j(q) \right)^{-1} dq \\ & \times (\nu\pi)^{d/2} \frac{\Gamma\left(\frac{\nu}{2}\right)^d}{\Gamma\left(\frac{\nu+1}{2}\right)^d} \prod_{j=1}^d \left(1 + \frac{x_j^2}{\nu_j}\right)^{(\nu_j+1)/2}. \end{aligned} \tag{2.28}$$

The limiting cases (co-/countermonotonicity) are the same as for the standard Student's t copula.

Each component of an individual Student's t copula hence has an individual degrees of freedom parameter, so that the individual Student's t copula has a total of $d(d-1)/2+d$ parameters. The standard Student's t copula is obtained when $\nu_1 = \nu_2 = ... = \nu_d$. A special case is also the grouped Student's t copula with fixed degrees of freedom for groups of variables (see Daul et al. (2003)): For example, in the case of two groups of size d_1 and $d_2 = d - d_1$, respectively, it holds that $\nu_1 = ... = \nu_{d_1}$ and $\nu_{d_1+1} = ... = \nu_d$.

The bivariate individual Student's t copula is characterized by the correlation parameter $\rho = \rho_{12}$ and the two degrees of freedom parameters ν_1 and ν_2. Kendall's τ is approximately given by

$$\tau(\rho, \nu_1, \nu_2) \approx \frac{2}{\pi} \arcsin(\rho), \tag{2.29}$$

as for the Gaussian and the standard Student's t copula. According to Daul et al. (2003) and Luo and Shevchenko (2010), the approximation error is typically very small. For Spearman's ρ_S no such approximate expression is known.

The individual Student's t copula is also reflection symmetric. The tail dependence coefficients are given by

$$\lambda_L(\rho, \nu_1, \nu_2) = \lambda_U(\rho, \nu_1, \nu_2) = \Omega(\rho, \nu_1, \nu_2) + \Omega(\rho, \nu_2, \nu_1), \tag{2.30}$$

with

$$\begin{aligned} \Omega(\rho, \nu_1, \nu_2) &= \int_0^\infty f_{\chi^2}(t; \nu_1 + 1)\, \Phi\left(-\frac{B(\nu_1, \nu_2) t^{\nu_1/(2\nu_2)} - \rho t^{1/2}}{\sqrt{1-\rho^2}} \right) dt, \\ B(\nu_1, \nu_2) &= \left(\frac{2^{\nu_2/2}\Gamma((1+\nu_2)/2)}{2^{\nu_1/2}\Gamma((1+\nu_1)/2)} \right)^{1/\nu_2}, \end{aligned}$$

where $f_{\chi^2}(\cdot; \nu)$ denotes the χ^2 density function with ν degrees of freedom. If $\nu = \nu_1 = \nu_2$, this reduces to Equation (2.25) for the standard Student's t copula with ν degrees of freedom. As before, small degrees of freedom indicate stronger tail dependence. It holds that $\lambda_L(\rho, \nu_j^{(1)}, \nu_k) > \lambda_L(\rho, \nu_j^{(2)}, \nu_k)$ if $\nu_j^{(2)} > \nu_j^{(1)} > \nu_k,\ j, k \in \{1, 2\},\ j \neq k$, and similarly for the upper tail dependence coefficient (see Figure 2.5).

In contrast to the Student's t copula, the bivariate individual Student's t copula is not symmetric if $\nu_1 \neq \nu_2$. This is intuitively clear, since ν_1 and ν_2 are parameters attached to the different variables. According to Luo and Shevchenko (2010), this asymmetry is most pronounced in the tails of the copula and less so in the main body of the copula, other

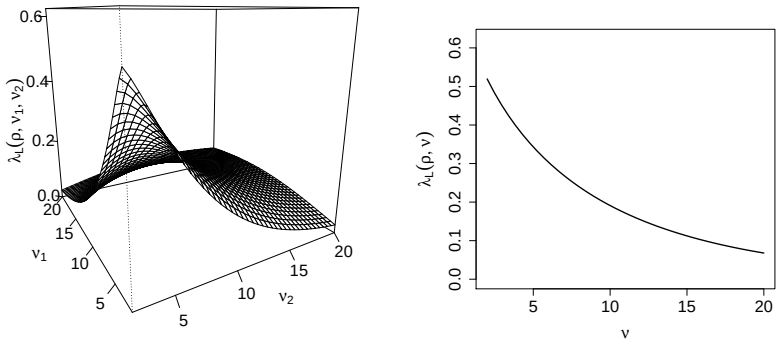


Figure 2.5: Lower tail dependence coefficient of the individual Student's t copula with $\rho = 0.7$ and different choices of degrees of freedom ν_1 and ν_2. The right panel shows the case, where $\nu = \nu_1 = \nu_2$, which corresponds to the standard Student's t copula.

than for example in the right panel of Figure 2.2. The asymmetry is therefore not well visible from the plots in Figure A.5.

To summarize, the individual Student's t copula extends the standard Student's t copula in order to obtain additional flexibility in the tails of a multivariate random vector. It is however also reflection symmetric, so that lower and upper tail dependence coefficients coincide. To overcome this limitation, we consider alternative classes of copulas in the following.

2.4 Archimedean copulas

An *Archimedean copula* is characterized by a generator φ and given by

$$C(\boldsymbol{u}; \varphi) = \varphi^{-1}\left(\varphi(u_1) + ... + \varphi(u_d)\right), \quad \boldsymbol{u} \in [0,1]^d, \tag{2.31}$$

where the generator $\varphi : [0,1] \to [0,\infty)$ is a continuous and strictly decreasing function, which satisfies $\varphi(1) = 0$. According to McNeil and Nešlehová (2009), φ generates a d-dimensional Archimedean copula if and only if its inverse φ^{-1} is *d-monotone* on $[0,\infty)$. This means that

(i) φ^{-1} is differentiable on $[0,\infty)$ up to the order $d-2$,

(ii) $(-1)^k(\varphi^{-1})^{(k)}(t) \geq 0$ for $k = 0, 1, ..., d-2$ and for any $t \in [0,\infty)$, and

(iii) $(-1)^{d-2}(\varphi^{-1})^{(d-2)}$ is non-increasing and convex on $[0,\infty)$.

A generator φ is called *completely monotone* if φ^{-1} has derivatives of all orders and satisfies $(-1)^k(\varphi^{-1})^{(k)}(x) \geq 0$ for $k \geq 0$ and any $x \in [0,\infty)$. Completely monotone generators can generate Archimedean copulas in any dimension (see Kimberling (1974), Joe (1997) and Nelsen (2006)). Due to the central role of the inverse generator φ^{-1}, Archimedean

copulas are also often defined in terms of $\psi := \varphi^{-1}$ in the literature. Here, we will use the parameterization in terms of φ.

An alternative characterization result of Archimedean copulas is provided by McNeil and Nešlehová (2009). They show that

$$(\varphi(U_1), ..., \varphi(U_d))' \stackrel{d}{=} R\boldsymbol{S}, \tag{2.32}$$

where $\boldsymbol{S} = (S_1, ..., S_d)'$ is uniformly distributed on the d-dimensional unit simplex,

$$\mathcal{S}_{d-1} = \{\boldsymbol{x} \in \mathbb{R}^d_{\geq 0} : \textstyle\sum_{j=1}^d x_j = 1\} \subset [0,1]^d. \tag{2.33}$$

Further, the radial part $R = \sum_{j=1}^d \varphi(U_j) \geq 0$ is independent of $\boldsymbol{S}$ and has distribution F_R, which can be determined through the inverse Williamson transform of φ^{-1} (see McNeil and Nešlehová (2009) for more details).

A simple example of an Archimedean copula is the independence copula (see Example 2.2). It can be represented through the generator $\varphi(t) = -\log t$, which is obviously completely monotone.

A d-dimensional Archimedean copula is absolutely continuous if $(\varphi^{-1})^{(d-1)}$ exists and is absolutely continuous on $(0, \infty)$ (see McNeil and Nešlehová (2009, Proposition 4.2)). Its density is given by

$$c(\boldsymbol{u}; \varphi) = (\varphi^{-1})^{(d)} \left(\varphi(u_1) + ... + \varphi(u_d)\right) \prod_{j=1}^d \varphi'(u_j), \quad \boldsymbol{u} \in [0,1]^d. \tag{2.34}$$

This expression requires the calculation of $(\varphi^{-1})^{(d)}$, which is typically very complex. Hofert et al. (2012) provide explicit functional expressions for common Archimedean generators, such as those that will be discussed below.

It immediately follows from Equation (2.31) that the copula quantile function of an Archimedean copula is conveniently given by

$$C^{-1}(z|u_1, ..., u_{d-1}; \varphi) = \varphi^{-1}\left(\varphi(z) - \sum_{j=1}^{d-1} \varphi(u_j)\right), \quad z \in (0,1). \tag{2.35}$$

The Kendall distribution function of a d-dimensional Archimedean copula is of more complicated form. Barbe et al. (1996) and McNeil and Nešlehová (2009) show that it is given in terms of the generator φ and higher order derivatives of its inverse φ^{-1} as

$$K(z; \varphi) = \begin{cases} \frac{(-\varphi(0))^{d-1}}{(d-1)!} (\varphi^{-1})_-^{(d-1)}(\varphi(0)) & \text{if } z = 0, \\ z + \sum_{k=1}^{d-2} \frac{(-\varphi(z))^k}{k!} (\varphi^{-1})^{(k)}(\varphi(z)) + \frac{(-\varphi(z))^{d-1}}{(d-1)!} (\varphi^{-1})_-^{(d-1)}(\varphi(z)) & \text{if } z \in (0,1], \end{cases} \tag{2.36}$$

where $(\varphi^{-1})_-^{(d-1)}$ denotes the left-hand derivative of φ^{-1} of order $d-1$. Again, availability of explicit functional expressions of $(\varphi^{-1})^{(k)}$ for common Archimedean generators (see Hofert et al. (2012)) renders feasible a computationally efficient computation of the Kendall distribution function.

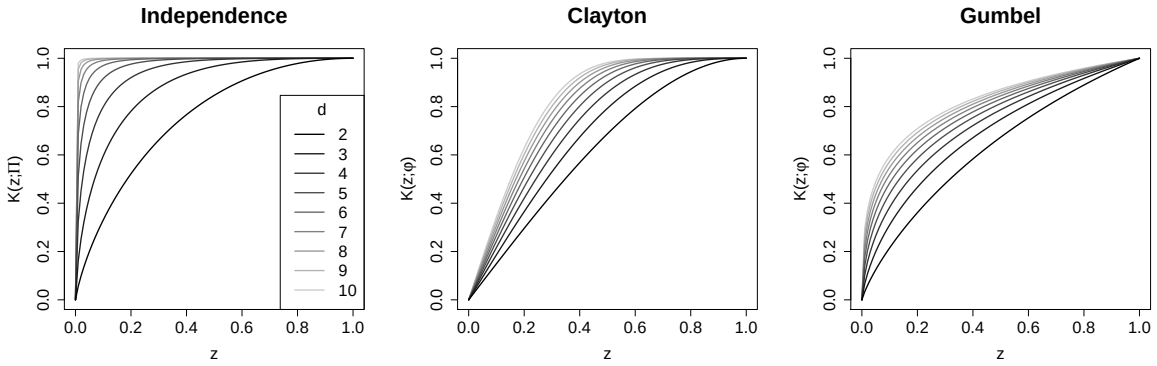


Figure 2.6: Kendall distribution functions of the independence (left panel), the Clayton (middle panel) and the Gumbel copula (right panel) for $d \in \{2, ..., 10\}$. The parameters of the Clayton and the Gumbel copula are chosen according to a Kendall's τ of 0.5.

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

Equation (2.36) now also allows to derive the Kendall distribution function of the independence copula with generator $\varphi(t) = -\log t$ as

$$K(z; \Pi) = z + z \sum_{k=1}^{d-1} \frac{(-1)^k}{k!} (\log z)^k, \quad z \in [0, 1],$$

since $\varphi^{-1}(t) = e^{-t}$ and hence $(\varphi^{-1})^{(k)}(t) = (-1)^k e^{-t}$, so that $(\varphi^{-1})^{(k)}(\varphi(z)) = (-1)^k z$. It depends on the dimension of the copula as illustrated in the left panel of Figure 2.6, which shows the Kendall distribution function of the d-dimensional independence copula for different choices of d.

In the bivariate case, it holds for Archimedean copulas that

$$K(z; \varphi) = z - \frac{\varphi(z)}{\varphi'(z)}, \quad z \in [0, 1], \tag{2.37}$$

which can also directly be derived using Equation (2.15) and noting that $C_{2|1}(u_2|u_1; \varphi) = \varphi'(u_1)/\varphi'(C(u_1, u_2; \varphi))$. Then,

$$K(z; \varphi) = z + \int_z^1 C_{2|1}(C^{-1}(z|u_1; \varphi)|u_1; \varphi)\, du_1 = z + \int_z^1 \frac{\varphi'(u_1)}{\varphi'(z)}\, du_1 = z - \frac{\varphi(z)}{\varphi'(z)},$$

since $\varphi(1) = 0$. Using Equation (2.16), Kendall's τ can therefore be expressed as

$$\tau(\varphi) = 1 + 4 \int_0^1 \frac{\varphi(z)}{\varphi'(z)}\, dz. \tag{2.38}$$

A similar expression for Spearman's ρ_S is not known. The tail dependence coefficients can be written as

$$\lambda_L(\varphi) = \lim_{s\to\infty} \frac{\varphi^{-1}(2s)}{\varphi^{-1}(s)}, \quad \text{and} \quad \lambda_U(\varphi) = 2 - \lim_{s\downarrow 0} \frac{1-\varphi^{-1}(2s)}{1-\varphi^{-1}(s)},$$

which follows by a simple reparameterization of Equations (2.6) and (2.7) with $t = \varphi^{-1}(s)$.

It also follows from Equation (2.31) that Archimedean copulas are exchangeable. This means that all lower dimensional margins of an Archimedean copula have the same distribution. In particular, all pairs of variables are identically distributed. As this is a quite strict assumption, Archimedean copulas are mostly used in the bivariate case—or as bivariate building blocks of vine copulas (see Section 2.7). A non-exchangeable extension are *hierarchical Archimedean copulas*, which are also called *nested Archimedean copulas* (see Section 3.3.1).

In the following, we present four popular Archimedean copulas, which exhibit different properties, especially with respect to their tail behavior: the *Clayton*, the *Gumbel*, the *Frank* and the *Joe copula*. Explicit density expressions are provided for the important bivariate case. In addition, all four copulas have TP_2 densities (if $\theta > 0$ in the case of the Frank copula).

Example 2.15 (Clayton copula)**.** The generator of the Clayton copula (see Clayton (1978) and also Kimeldorf and Sampson (1975) and Cook and Johnson (1981)) is $\varphi(t;\theta) = \theta^{-1}(t^{-\theta} - 1)$. If $\theta > 0$, the copula is completely monotone and given by

$$C(\boldsymbol{u};\theta) = \left(u_1^{-\theta} + ... + u_d^{-\theta} - d + 1\right)^{-1/\theta}, \quad \boldsymbol{u} \in [0,1]^d.$$

The extension to negative parameters is not considered here (φ is d-monotone if $\theta \geq -1/(d-1)$). The limiting cases of the Clayton copula are independence if $\theta \to 0$ and comonotonicity if $\theta \to \infty$.

The Kendall distribution function of the multivariate Clayton copula is illustrated in the middle panel of Figure 2.6. In the bivariate case, the density of the Clayton copula can be obtained as

$$c(u_1, u_2;\theta) = (1+\theta)(u_1 u_2)^{-1-\theta}\left(u_1^{-\theta} + u_2^{-\theta} - 1\right)^{-1/\theta-2}, \quad (u_1, u_2)' \in [0,1]^2.$$

The corresponding Kendall's τ is given by

$$\tau(\theta) = \frac{\theta}{\theta+2}.$$

In terms of tail dependence, it turns out that the Clayton copula is lower tail dependent but upper tail independent. The tail dependence coefficients are

$$\lambda_L(\theta) = 2^{-1/\theta}, \quad \lambda_U(\theta) = 0.$$

The Clayton copula is hence the first example of a reflection asymmetric, and thus tail asymmetric, copula. This is reflected in the shape of the scatter plots and the contour lines in Figure A.6. □

Example 2.16 (Gumbel copula)**.** Unlike the Clayton copula, the Gumbel copula (see Gumbel (1960)) is upper tail dependent. Its generator is defined as $\varphi(t;\theta) = (-\log t)^\theta$, which is completely monotone for $\theta \geq 1$. The d-dimensional Gumbel copula is then given by

$$C(\boldsymbol{u};\theta) = \exp\left(-\left((-\log u_1)^\theta + ... + (-\log u_d)^\theta\right)^{1/\theta}\right), \quad \boldsymbol{u} \in [0,1]^d,$$

and, similar to the Clayton copula, the limiting cases of the Gumbel copula are independence if $\theta = 1$ and comonotonicity if $\theta \to \infty$.

The Kendall distribution function of the d-dimensional Gumbel copula for different choices of d is shown in the right panel of Figure 2.6. The density of the Gumbel copula in the bivariate case can be derived as

$$\begin{aligned} c(u_1,u_2;\theta) = \frac{C(u_1,u_2;\theta)}{u_1u_2} &\frac{((\log u_1)(\log u_2))^{\theta-1}}{((-\log u_1)^\theta + (-\log u_2)^\theta)^{2-1/\theta}} \\ &\times \left(\left((-\log u_1)^\theta + (-\log u_2)^\theta\right)^{1/\theta} + \theta - 1\right), \quad (u_1,u_2)' \in [0,1]^2. \end{aligned}$$

While there is again no known closed-form expression of Spearman's ρ_S, Kendall's τ is given by

$$\tau(\theta) = 1 - \frac{1}{\theta}.$$

Finally, the Gumbel copula is also reflection asymmetric and, as noted above, exhibits upper tail dependence but no lower tail dependence:

$$\lambda_L(\theta) = 0, \quad \lambda_U(\theta) = 2 - 2^{1/\theta},$$

which also translates to the shape of the contour lines in Figure A.7. □

Example 2.17 (Frank copula)**.** The generator of the Frank copula (see Frank (1979)), $\varphi(t;\theta) = -\log((e^{-\theta t}-1)/(e^{-\theta}-1))$, is completely monotone for $\theta > 0$. It defines the copula

$$C(\boldsymbol{u};\theta) = -\frac{1}{\theta}\log\left(1 + \frac{\prod_{j=1}^d (e^{-\theta u_j}-1)}{(e^{-\theta}-1)^{d-1}}\right), \quad \boldsymbol{u} \in [0,1]^d. \tag{2.39}$$

The Frank copula converges to independence and to comonotonicity if $\theta \to 0$ and $\theta \to \infty$, respectively. In the bivariate case, Equation (2.39) also yields a valid copula for $\theta < 0$, so that also negative dependence can be covered. Then, the countermonotonicity copula is the limiting case if $\theta \to -\infty$.

The density of its bivariate version is

$$c(u_1,u_2;\theta) = \theta(e^{-\theta}-1)\frac{e^{-\theta(u_1+u_2)}}{(e^{-\theta}-1+(e^{-\theta u_1}-1)(e^{-\theta u_2}-1))^2}, \quad (u_1,u_2)' \in [0,1]^2.$$

The corresponding Kendall's τ can be derived in terms of the so-called Debye function, which is defined by

$$D_k(x) = \frac{k}{x^k}\int_0^x \frac{t^k}{e^t-1}\,dt, \quad x \in \mathbb{R}\setminus\{0\},\ k \in \mathbb{N}.$$

Then,

$$\tau(\theta) = 1 + \frac{4}{\theta}(D_1(\theta) - 1).$$

The Frank copula is also an example of a copula with a simplified expression for Spearman's ρ_S,

$$\rho_S(\theta) = 1 - \frac{12}{\theta}(D_1(\theta) - D_2(\theta)).$$

Similar to the Gaussian copula, the Frank copula is reflection symmetric in the bivariate case (but not for $d \geq 3$; see Joe (1997, Section 7.1.7)) and does not exhibit any tail dependence,

$$\lambda_L(\theta) = \lambda_U(\theta) = 0.$$

Nevertheless, the shape of the contour lines of the copula is rather non-elliptical in contrast to the Gaussian copula (see Figure A.8). □

Example 2.18 (Joe copula)**.** The Joe copula (see Joe (1993)) is yet another example of an upper tail dependent Archimedean copula. Its generator is $\varphi(t;\theta) = -\log\left(1-(1-t)^\theta\right)$, which implies the copula

$$C(\boldsymbol{u};\theta) = 1 - \left(1 - \prod_{j=1}^{d}\left(1-(1-u_j)^\theta\right)\right)^{1/\theta}, \quad \boldsymbol{u} \in [0,1]^d,$$

for $\theta > 1$. As for the Gumbel copula, the limiting case of the Joe copula for $\theta \to 1$ is the independence copula, while comonotonicity is obtained for $\theta \to \infty$.

For $(u_1, u_2)' \in [0,1]^2$, the density of the bivariate Joe copula is

$$\begin{aligned} c(u_1,u_2;\theta) = &\left((1-u_1)^\theta + (1-u_2)^\theta - (1-u_1)^\theta(1-u_2)^\theta\right)^{1/\theta-2} \\ &\times (1-u_1)^{\theta-1}(1-u_2)^{\theta-1} \\ &\times \left(\theta - 1 + (1-u_1)^\theta + (1-u_2)^\theta - (1-u_1)^\theta(1-u_2)^\theta\right). \end{aligned} \tag{2.40}$$

Using a result by Schepsmeier (2010, Section 2.3.2), the corresponding Kendall's τ can be obtained as

$$\tau(\theta) = \begin{cases} 1 + \frac{2}{2-\theta}\left(\Psi(2) - \Psi\left(\frac{2}{\theta}+1\right)\right) & \text{if } \theta \neq 2, \\ 1 - \Psi'(2) & \text{if } \theta = 2, \end{cases} \tag{2.41}$$

where Ψ is the digamma function, which is defined as the logarithmic derivative of the gamma function. The tail dependence coefficients are the same as for the Gumbel copula and given by

$$\lambda_L(\theta) = 0, \quad \lambda_U(\theta) = 2 - 2^{1/\theta}.$$

Although both the Gumbel and the Joe copula are reflection asymmetric and upper tail dependent Archimedean copulas, the shape of their contour lines is quite different (see Figure A.9), so that it is actually sensible to consider both copulas. □

There are also other popular Archimedean copulas such as the *Ali-Mikhail-Haq copula* (see Ali et al. (1978)), but here we concentrate on the four presented ones, because we believe that they reasonably capture common dependence patterns. A worthwhile extension,

which is not treated here in detail, are the two parameter *BB copulas* by Joe (1997, Section 5.2). They include extensions of the presented copulas, such as the *Clayton-Gumbel (BB1)* or the *Joe-Clayton (BB7) copula*, which are also Archimedean and exhibit different non-zero lower and upper tail dependence coefficients.

2.5 Extreme value and Archimax copulas

Extreme value copulas are the asymptotic limits of component-wise maxima (see Pickands (1981) and the overview by Gudendorf and Segers (2010)). Let $\boldsymbol{X}_i = (X_{i1}, ..., X_{id})', \ i = 1, ..., n$, be n independent copies of a d-dimensional random vector $\boldsymbol{X}$ with copula C_0 and define $M_{n,j} := \max\{X_{1j}, ..., X_{nj}\}$. Then, the copula of $M_n = (M_{n,1}, ..., M_{n,d})'$ is

$$C_{0(n)}(\boldsymbol{u}) := C_0\left(u_1^{1/n}, ..., u_d^{1/n}\right)^n, \quad \boldsymbol{u} \in [0,1]^d.$$

A copula C is called an *extreme value copula* if there exists a copula C_0 such that

$$C_{0(n)}(\boldsymbol{u}) \xrightarrow{n\to\infty} C(\boldsymbol{u}) \quad \forall \boldsymbol{u} \in [0,1]^d. \tag{2.42}$$

The copula C_0 is then said to lie in the domain of attraction of C.

In the bivariate case, an extreme value copula can be uniquely identified by a univariate function, the dependence function by Pickands (1981). Let $A : [0,1] \to [0.5, 1]$ be convex and satisfy $\max\{t, 1-t\} \leq A(t) \leq 1$ for all $t \in [0,1]$. Then,

$$C(u_1, u_2; A) = \exp\left(\log(u_1 u_2)\, A\left(\frac{\log u_2}{\log(u_1 u_2)}\right)\right), \quad (u_1, u_2)' \in [0,1]^2, \tag{2.43}$$

is an extreme value copula. The converse statement is also true: If C is an extreme value copula, then there exists a function A with the above stated properties, such that C can be written as in (2.43).

An extreme value copula C is symmetric if and only if the Pickands dependence function A is symmetric about 0.5, since

$$\begin{aligned} A(t) = A(1-t)\ \forall t \in [0,1] \quad &\Leftrightarrow \quad A\left(\frac{\log u_2}{\log(u_1 u_2)}\right) = A\left(\frac{\log u_1}{\log(u_1 u_2)}\right)\ \forall (u_1, u_2)' \in [0,1]^2 \\ &\Leftrightarrow \quad C(u_1, u_2; A) = C(u_2, u_1; A)\ \forall (u_1, u_2)' \in [0,1]^2. \end{aligned}$$

The bounds on the Pickands dependence function A correspond to the cases of independence and comonotonicity. The independence copula can be represented as an extreme value copula with $A(t) = 1$ for all $t \in [0,1]$. Similarly, $A(t) = \max\{t, 1-t\}$ is the Pickands dependence function of the comonotonicity copula. This is illustrated in Figure 2.7, which also shows an example of an asymmetric Pickands dependence function.

The unique characterization (2.43) of extreme value copulas allows to conveniently derive quantities like the density and the Kendall distribution function in terms of the Pickands dependence function A. Assuming that A is twice differentiable, the density of an extreme value copula is given for $(u_1, u_2)' \in [0,1]^2$ by

$$c(u_1, u_2; A) = \frac{C(u_1, u_2; A)}{u_1 u_2}\left(A(t)^2 + (1-2t)A'(t)A(t) - (1-t)t\left(A'(t)^2 - \frac{A''(t)}{\log(u_1 u_2)}\right)\right),$$

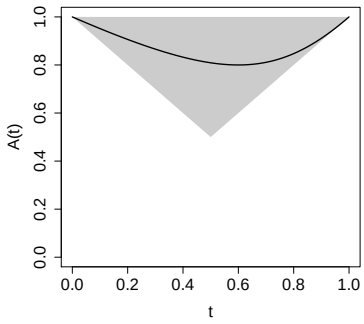


Figure 2.7: Example of an asymmetric Pickands dependence function. The gray area illustrates the lower and upper bounds of the Pickands dependence function.

where $t = t(u_1, u_2) := \log u_2 / \log(u_1 u_2)$. Further, according to Ghoudi et al. (1998), the Kendall distribution function is

$$K(z; A) = z(1 + (\tau(A) - 1) \log z), \quad z \in [0, 1], \tag{2.44}$$

where $\tau(A)$ is the Kendall's τ of an extreme value copula,

$$\tau(A) = \int_0^1 \frac{t(1-t)}{A(t)} \, dA'(t), \tag{2.45}$$

assuming that A' exists. This means that the Kendall distribution function is the same for all extreme value copulas with the same Kendall's τ.

According to Hürlimann (2003), Kendall's τ can alternatively be determined as

$$\tau(A) = \int_0^1 \frac{(2t-1)A(t)A'(t) + t(1-t)A'(t)^2}{A(t)^2} \, dt, \tag{2.46}$$

which does not involve the second derivative of A. Furthermore, Hürlimann (2003) also provides a simplified expression of Spearman's ρ_S in terms of the Pickands dependence function (see also Capéraà et al. (1997)):

$$\rho_S(A) = 12 \int_0^1 \frac{1}{(1 + A(t))^2} \, dt - 3.$$

These expressions for Kendall's τ and Spearman's ρ_S however only seldom lead to closed-form expressions in terms of copula parameters.

Finally, the lower and upper tail dependence coefficients of extreme value copulas take on particularly convenient forms, namely

$$\lambda_L(A) = \begin{cases} 1 & \text{if } A(0.5) = 0.5, \\ 0 & \text{otherwise,} \end{cases} \qquad \lambda_U(A) = 2(1 - A(0.5)),$$

which follows from a straightforward calculation according to Remark 2.7. This means that extreme value copulas are lower tail independent expect for the case of comonotonicity, where $A(0.5) = 0.5$. The strength of the upper tail dependence is determined by the Pickands dependence function A evaluated at 0.5. Except for the boundary cases of independence and comonotonicity, extreme value copulas are therefore not reflection symmetric.

In the literature, a wide range of different extreme value copulas has been proposed either as the asymptotic limit of a common copula according to Equation (2.42) or directly in terms of a Pickands dependence function A, exploiting Equation (2.43). For instance, the Student's t copula (see Example 2.14) lies in the domain of attraction of the so-called *t-EV copula* (see Demarta and McNeil (2005) and Nikoloulopoulos et al. (2009)). Archimedean copulas lie in the domain of attraction of the Gumbel copula (see Example 2.16) if $-\lim_{s\downarrow 0} \frac{s\varphi'(1-s)}{\varphi(1-s)} \in [1, \infty]$ exists (see Genest and Rivest (1989)). The Gumbel copula is actually the only copula that is both an Archimedean and an extreme value copula. Its Pickands dependence function is

$$A(t;\theta) = \left(t^\theta + (1-t)^\theta\right)^{1/\theta}, \quad t \in [0,1], \tag{2.47}$$

which is symmetric about 0.5.

An overview of other extreme value copulas can be found, e.g., in Eschenburg (2013), where it is also shown that many common extreme value copulas all model very similar dependence patterns (see also Genest, Kojadinovic, Nešlehová, and Yan (2011)). Most flexibility is added by allowing for asymmetry. A famous example of such a copula is the *Tawn copula*, which is an extension of the Gumbel copula.

Example 2.19 (Tawn copula)**.** The Pickands dependence function of the copula by Tawn (1988) is

$$A(t;\theta,\psi_1,\psi_2) = (1-\psi_2)(1-t) + (1-\psi_1)t + \left((\psi_1(1-t))^\theta + (\psi_2 t)^\theta\right)^{1/\theta}, \quad t \in [0,1],$$

where $\theta \geq 1$ and $\psi_1, \psi_2 \in [0,1]$. If $\psi_1 = \psi_2 = 1$, the Tawn copula coincides with the Gumbel copula (see Equation (2.47)). The independence copula is a special case for $\theta = 1$ or $\psi_1 = 0$ or $\psi_2 = 0$. Furthermore, if $\theta \to \infty$, the Tawn copula converges to the *Marshall-Olkin copula* (see Marshall and Olkin (1967)), which can be defined in terms of the following Pickands dependence function:

$$A(t;\psi_1,\psi_2) = \max\{1-\psi_1(1-t), 1-\psi_2 t\}, \quad t \in [0,1],$$

where $\psi_1, \psi_2 \in [0,1]$.

The Tawn copula is asymmetric if $\psi_1 \neq \psi_2$. This is illustrated in Figure 2.7, which, in fact, shows the Pickands dependence function of the Tawn copula for $\theta = 2$, $\psi_1 = 1$ and $\psi_2 = 0.5$. In contrast to the asymmetry of the individual Student's t copula, which is strongest in the tails, the asymmetry of the Tawn copula is more distinct and very well visible in Figure A.10. According to Khoudraji (1995), the Tawn copula can also be written as

$$C(u_1,u_2;\theta,\psi_1,\psi_2) = u_1^{1-\psi_1} u_2^{1-\psi_2} C(u_1^{\psi_1}, u_2^{\psi_2};\theta),$$

where $C(\cdot,\cdot;\theta)$ is a Gumbel copula with parameter θ. As this construction yields a valid copula for any choice of $C(\cdot,\cdot;\theta)$, it can also be seen as a general device to define asymmetric copulas.

The choice of the asymmetry parameters ψ_1 and ψ_2 also influences the possible range of Kendall's τ. Since the Kendall's τ of the Tawn copula is increasing in θ and the Marshall-Olkin copula is the boundary case if $\theta \to \infty$, it holds that

$$\tau(\theta,\psi_1,\psi_2) \leq \tau(\psi_1,\psi_2) = \frac{\psi_1\psi_2}{\psi_1+\psi_2-\psi_1\psi_2},$$

where $\tau(\psi_1,\psi_2)$ denotes the Kendall's τ of the Marshall-Olkin copula (see Embrechts et al. (2003)). Similarly, the Spearman's ρ_S of the Tawn copula is bounded above by the Spearman's ρ_S of the Marshall-Olkin copula, which is denoted by $\rho_S(\psi_1,\psi_2)$:

$$\rho_S(\theta,\psi_1,\psi_2) \leq \rho_S(\psi_1,\psi_2) = \frac{3\psi_1\psi_2}{2\psi_1+2\psi_2-\psi_1\psi_2}.$$

While the lower tail dependence coefficient is zero, the upper tail dependence coefficient is also strongly dependent on ψ_1 and ψ_2:

$$\lambda_L(\theta,\psi_1,\psi_2) = 0, \quad \lambda_U(\theta,\psi_1,\psi_2) = \psi_1+\psi_2-(\psi_1^\theta+\psi_2^\theta)^{1/\theta}.$$

For fixed θ, the upper tail dependence coefficient is maximal if $\psi_1 = \psi_2 = 1$, which corresponds the boundary case of the Gumbel copula. □

An extension of both the classes of extreme value and of bivariate Archimedean copulas was proposed by Capéraà et al. (2000). They define an *Archimax copula* for $(u_1,u_2)' \in [0,1]^2$ as

$$C(u_1,u_2;\varphi,A) = \varphi^{-1}\left((\varphi(u_1)+\varphi(u_2))\,A\left(\frac{\varphi(u_2)}{\varphi(u_1)+\varphi(u_2)}\right)\right), \tag{2.48}$$

where φ is a 2-monotone Archimedean generator (see Section 2.4) and A is a Pickands dependence function. If $A(t) = 1$ for all $t \in [0,1]$ (Pickands dependence function of the independence copula), then the Archimax copula (2.48) becomes the bivariate Archimedean copula with generator φ (see Equation (2.31)). Conversely, if $\varphi(t) = -\log t$ (generator of the independence copula), it corresponds to the extreme value copula with Pickands dependence function A (see Equation (2.43)).

Capéraà et al. (2000) show that the Kendall distribution function of an Archimax copula is given by

$$K(z;\varphi,A) = K(z;\varphi) + (z - K(z;\varphi))\,\tau(A), \quad z \in [0,1], \tag{2.49}$$

where $K(\cdot;\varphi)$ is the Kendall distribution function of the corresponding Archimedean copula (see Equation (2.37)) and $\tau(A)$ is the Kendall's τ of the corresponding extreme value copula (see Equation (2.45) or (2.46)). Similarly, the Kendall's τ of an Archimax copula can be derived as

$$\tau(\varphi,A) = \tau(\varphi) + (1-\tau(\varphi))\,\tau(A),$$

where $\tau(\varphi)$ is given in Equation (2.38).

2.6 Plackett copula

The *Plackett copula* (see Plackett (1965) and Mardia (1970)) is an example of a copula that does not belong to either one of the popular classes of elliptical, Archimedean and extreme value copulas, which were discussed before. It is derived through the constant cross-product ratio

$$\frac{C(u_1,u_2;\alpha)(1-u_1-u_2+C(u_1,u_2;\alpha))}{(u_1-C(u_1,u_2;\alpha))(u_2-C(u_1,u_2;\alpha))}=\alpha+1,\quad (u_1,u_2)'\in[0,1]^2,$$

where $\alpha\in(-1,\infty)\setminus\{0\}$. This results in

$$C(u_1,u_2;\alpha)=\frac{1}{2\alpha}\left(1+\alpha(u_1+u_2)-\sqrt{(1+\alpha(u_1+u_2))^2-4\alpha(\alpha+1)u_1u_2}\right). \tag{2.50}$$

If $\alpha\to-1$, the Plackett copula converges to the countermonotonicity copula, while comonotonicity is obtained for $\alpha\to\infty$. The independence copula is the limit as $\alpha\to0$.

The density of the Plackett copula is given by

$$c(u_1,u_2;\alpha)=\frac{(\alpha+1)(1+\alpha(u_1+u_2-2u_1u_2))}{((1+\alpha(u_1+u_2))^2-4\alpha(\alpha+1)u_1u_2)^{3/2}},\quad (u_1,u_2)'\in[0,1]^2, \tag{2.51}$$

and also the copula quantile function is available in closed form as

$$C^{-1}(z|u;\alpha)=z\,\frac{1+\alpha(u-z)}{z+(\alpha+1)(u-z)},\quad z\in(0,1). \tag{2.52}$$

Using this expression, an explicit but cumbersome expression of the Kendall distribution function can be derived (see Appendix B.3).

In contrast to most other copulas, the Plackett copulas does not possess an explicit formula for Kendall's τ but for Spearman's ρ_S (see Nelsen (2006, Exercise 5.8)):

$$\rho_S(\alpha)=\frac{\alpha+2}{\alpha}-\frac{2(\alpha+1)}{\alpha^2}\log(\alpha+1).$$

Similar to the Gaussian and the bivariate Frank copula, the Plackett copula is also reflection symmetric and tail independent, which is visible in Figure A.11. It holds that

$$\lambda_L(\alpha)=\lambda_U(\alpha)=0.$$

It is also a symmetric copula, which directly follows from Equation (2.50).

Another copula that does not belong to either one of the most popular classes of copulas is the *Farlie-Gumbel-Morgenstern (FGM) copula*,

$$C(u_1,u_2;\alpha)=u_1u_2\left(1+\alpha(1-u_1)(1-u_2)\right),\quad (u_1,u_2)'\in[0,1]^2, \tag{2.53}$$

where $\alpha\in[-1,1]$. It can be interpreted as first-order approximation to the Plackett copula (2.50) (see Nelsen (2006, Exercise 3.39)). Its range of Kendall's τ is however rather limited, since

$$\tau(\alpha)=\frac{2\alpha}{9}\in[-2/9,2/9], \tag{2.54}$$

as shown in Nelsen (2006, Example 5.2). It therefore has to be used carefully.

2.7 Vine copulas

Vine copulas are multivariate copulas that are defined through a cascade of bivariate copulas, so-called *pair copulas*. Recursive conditioning ensures that such a *pair copula construction (PCC)* yields a valid multivariate distribution. We first discuss this general idea of constructing multivariate copulas and have a closer look at a critical assumption that is often made. Vines then provide a convenient graphical model to organize PCCs and we briefly treat statistical inference techniques for this model class.

2.7.1 Pair copula constructions

PCCs were originally proposed by Joe (1996, 1997) as so-called mixtures of conditional distributions. Bedford and Cooke (2001, 2002) and Aas et al. (2009) picked up the idea to construct flexible multivariate copulas. We begin with a trivariate example.

Let $\boldsymbol{X} = (X_1, X_2, X_3)' \sim F$. The multivariate density f of $\boldsymbol{X}$ can then be decomposed as

$$f(x_1, x_2, x_3) = f_1(x_1)\, f_{2|1}(x_2|x_1)\, f_{3|1,2}(x_3|x_1, x_2), \quad (x_1, x_2, x_3)' \in \mathbb{R}^3. \tag{2.55}$$

With Sklar's Theorem (2.2) it follows that

$$\begin{aligned} f_{2|1}(x_2|x_1) &= \frac{f_{1,2}(x_1, x_2)}{f_1(x_1)} = \frac{c_{1,2}(F_1(x_1), F_2(x_2))\, f_1(x_1)\, f_2(x_2)}{f_1(x_1)} \\ &= c_{1,2}(F_1(x_1), F_2(x_2))\, f_2(x_2), \end{aligned} \tag{2.56}$$

where $C_{1,2}$ is the bivariate copula of the pair $(1, 2)$. In the same way, it holds that

$$\begin{aligned} f_{3|1,2}(x_3|x_1, x_2) &= \frac{f_{2,3|1}(x_2, x_3|x_1)}{f_{2|1}(x_2|x_1)} \\ &= \frac{c_{2,3;1}(F_{2|1}(x_2|x_1), F_{3|1}(x_3|x_1)|x_1)\, f_{2|1}(x_2|x_1)\, f_{3|1}(x_3|x_1)}{f_{2|1}(x_2|x_1)} \\ &= c_{2,3;1}(F_{2|1}(x_2|x_1), F_{3|1}(x_3|x_1)|x_1)\, f_{3|1}(x_3|x_1) \\ &\overset{(2.56)}{=} c_{2,3;1}(F_{2|1}(x_2|x_1), F_{3|1}(x_3|x_1)|x_1)\, c_{1,3}(F_1(x_1), F_3(x_3))\, f_3(x_3), \end{aligned} \tag{2.57}$$

where $F_{j|1},\ j = 2, 3$, can be expressed in terms of $C_{j|1}$ as defined in Equation (2.13),

$$\begin{aligned} F_{j|1}(x_j|x_1) &= \frac{\partial F_{1,j}(x_1, x_j)}{\partial x_1} \frac{1}{f_1(x_1)} = \frac{\partial C_{1,j}(F_1(x_1), F_j(x_j))}{\partial x_1} \left(\frac{\partial F_1(x_1)}{\partial x_1}\right)^{-1} \\ &= \frac{\partial C_{1,j}(F_1(x_1), F_j(x_j))}{\partial F_1(x_1)} = C_{j|1}(F_j(x_j)|F_1(x_1)). \end{aligned} \tag{2.58}$$

Thus, we have decomposed the density f of $\boldsymbol{X}$ in terms of the marginal densities and the three bivariate copulas $C_{1,2}$, $C_{1,3}$ and $C_{2,3;1}$ with densities $c_{1,2}$, $c_{1,3}$ and $c_{2,3;1}$, respectively. The copula $C_{2,3;1}$ should not be confused with $C_{2,3|1}$, which we use to denote the conditional distribution function of $(U_2, U_3)'|U_1$ (see Examples 2.20 and 2.21 below).

Note that the copula $C_{2,3;1}$ depends on variable 1 not only through its arguments $C_{j|1}$, $j = 1, 2$, but also directly. To facilitate the statistical inference of PCCs, it is typically assumed that

$$C_{2,3;1}(\cdot, \cdot|x_1) = C_{2,3;1}(\cdot, \cdot). \tag{2.59}$$

We treat this so-called *simplifying assumption* in more detail in Section 2.7.2.

Under the simplifying assumption, we can construct a flexible multivariate copula by choosing the pair copulas $C_{1,2}$, $C_{1,3}$ and $C_{2,3;1}$ independently of each other and from different copula classes (see Sections 2.2–2.6). The density c of the corresponding three-dimensional copula C of $\boldsymbol{X}$ is then given for $(u_1, u_2, u_3)' \in [0, 1]^3$ by

$$c(u_1, u_2, u_3) = c_{1,2}(u_1, u_2)\, c_{1,3}(u_1, u_3)\, c_{2,3;1}(C_{2|1}(u_2|u_1), C_{3|1}(u_3|u_1)). \tag{2.60}$$

Clearly, this construction can also be done in higher dimensions. We extend the above example and let $d = 4$. Then we can decompose the density f of $\boldsymbol{X} = (X_1, ..., X_4)'$ as in Equation (2.55),

$$f(x_1, x_2, x_3, x_4) = f_1(x_1)\, f_{2|1}(x_2|x_1)\, f_{3|1,2}(x_3|x_1, x_2)\, f_{4|1,2,3}(x_4|x_1, x_2, x_3), \tag{2.61}$$

where $(x_1, x_2, x_3, x_4)' \in \mathbb{R}^4$. In other words, we add X_4 to the decomposition of the trivariate density (2.55). The first three terms are treated exactly as above, and we similarly proceed with $f_{4|1,2,3}$, where we now assume the simplifying assumption (2.59) to hold. We have that

$$\begin{aligned} f_{4|1,2,3}(x_4|x_1, x_2, x_3) &= \frac{f_{3,4|1,2}(x_3, x_4|x_1, x_2)}{f_{3|1,2}(x_3|x_1, x_2)} \\ &= c_{3,4;1,2}(F_{3|1,2}(x_3|x_1, x_2), F_{4|1,2}(x_4|x_1, x_2))\, f_{4|1,2}(x_4|x_1, x_2), \end{aligned} \tag{2.62}$$

where $f_{4|1,2}$ is decomposed as $f_{3|1,2}$ in (2.57) in terms of copulas $C_{1,4}$ and $C_{2,4;1}$ and the marginal density f_4. Hence, we obtain a PCC in terms of the four marginal densities $f_1, ..., f_4$, three copulas $C_{1,2}$, $C_{1,3}$ and $C_{1,4}$ of unconditional pairs of variables as well as three copulas $C_{2,3;1}$, $C_{2,4;1}$ and $C_{3,4;1,2}$ of conditional pairs. The arguments of the copula $C_{3,4;1,2}$, $F_{j|1,2}$ for $j = 3, 4$, are given as follows:

$$F_{j|1,2}(x_j|x_1, x_2) = C_{j|1,2}(F_j(x_j)|F_1(x_1), F_2(x_2)),$$

where $C_{j|1,2}$ is the distribution function of U_j given U_1 and U_2. As shown, e.g., in Czado (2010), $C_{j|1,2}$ can be determined iteratively in terms of the pair copulas of the decomposition,

$$C_{j|1,2}(u_j|u_1, u_2) = C_{j|2;1}(C_{j|1}(u_j|u_1)|C_{2|1}(u_2|u_1)), \quad (u_1, u_2, u_j)' \in [0, 1]^3, \tag{2.63}$$

where

$$C_{j|2;1} = \frac{\partial C_{2,j;1}(u_2, u_j)}{\partial u_2}, \tag{2.64}$$

as in Equation (2.13). This means that only the pair copulas $C_{1,2}$, $C_{1,j}$ and $C_{2,j;1}$ are required to compute $C_{j|1,2}$ and $F_{j|1,2}$ for $j = 3, 4$.

Using the recursive density decomposition (2.55), we can hence derive d-dimensional PCCs for any $d \geq 3$. This however requires some choices with respect to the order of the variables in the PCC. In fact, we already did a few such choices above: It starts with the decomposition (2.55), where an order of the variables has to be chosen. Then, in (2.57) we could also condition on variable 2 and similarly in (2.62), where even more choices are possible. In order to organize possible decompositions, we therefore discuss vines, which link PCCs to a graph theoretical model.

2.7.2 Simplifying assumption

Although the simplifying assumption (2.59) of PCCs is commonly made to facilitate the statistical inference, a diligent statistical analysis of the assumption was missing until recently. A first discussion on the simplifying assumption of PCCs is due to Hobæk Haff et al. (2010), who provide first examples and results, which were recently refined by Stöber et al. (2013). The Clayton copula (see Example 2.15) is, in fact, an example of a copula, which can be represented as a simplified PCC. We illustrate this result here in the trivariate case.

Example 2.20 (Trivariate Clayton copula)**.** We consider the trivariate Clayton copula $C(\cdot, \cdot, \cdot; \theta)$ with $\theta > 0$ and derive its PCC to see if it is of simplified type. Let $\boldsymbol{U} \sim C(\cdot, \cdot, \cdot; \theta)$, then the distribution of $(U_2, U_3)'$ conditioned on $U_1 = u_1$ is given for $(u_1, u_2, u_3)' \in [0,1]^3$ by

$$C_{2,3|1}(u_2, u_3|u_1; \theta) = \frac{\partial}{\partial u_1} C(u_1, u_2, u_3; \theta) = \left(u_1^{-\theta} + u_2^{-\theta} + u_3^{-\theta} - 2\right)^{-1/\theta} u_1^{-\theta-1},$$

and similarly $C_{j|1}(u_j|u_1; \theta) = (u_1^{-\theta} + u_j^{-\theta} - 1)^{-1/\theta} u_1^{-\theta-1}$ for $j = 2, 3$. Therefore, we obtain

$$C_{j|1}^{-1}(u_j|u_1; \theta) = \left((u_j^{-\theta/(\theta+1)} - 1)u_1^{-\theta} + 1\right)^{-1/\theta}, \quad j = 2, 3.$$

This yields the following expression for the copula of $(U_2, U_3)'|U_1 = u_1$:

$$\begin{aligned} C_{2,3;1}(u_2, u_3|u_1; \theta) &= C_{2,3|1}\left(C_{2|1}^{-1}(u_2|u_1; \theta), C_{3|1}^{-1}(u_3|u_1; \theta)|u_1; \theta\right) \\ &= \left(u_1^{-\theta/(\theta+1)} + u_2^{-\theta/(\theta+1)} - 1\right)^{-(\theta+1)/\theta} \\ &= C\left(u_2, u_3; \tfrac{\theta}{\theta+1}\right), \end{aligned}$$

which is a bivariate Clayton copula with parameter $\theta/(\theta+1)$. In particular, the copula $C_{2,3;1}$ is independent of the value u_1, so that the trivariate Clayton copula is a PCC of simplified type.

According to Stöber et al. (2013), the d-dimensional Clayton copula can, in fact, be represented as a simplified PCC for any $d \geq 3$. It is even shown to be the only Archimedean copula of simplified type, when the generator is twice continuously differentiable on the set where it is positive. This is especially true if $d \geq 4$ (see Section 2.4). □

This however means that all other Archimedean copulas cannot be represented as a simplified PCC. An example involving the trivariate Frank copula is provided in Stöber et al. (2013). An example of a non-Archimedean copula, which is also not of simplified type is given in the following.

Example 2.21 (Trivariate FGM copula)**.** A simple trivariate extension of the bivariate FGM copula (2.53) can be defined as

$$C(u_1, u_2, u_3; \alpha) = u_1 u_2 u_3 (1 + \alpha(1 - u_1)(1 - u_2)(1 - u_3)), \quad (u_1, u_2, u_3)' \in [0, 1]^3, \quad (2.65)$$

where $\alpha \in [-1, 1]$. If $\boldsymbol{U} \sim C(\cdot, \cdot, \cdot; \alpha)$, then U_1, U_2 and U_3 are pairwise independent but not mutually (if $\alpha \neq 0$). The conditional distribution of $(U_2, U_3)'$ given $U_1 = u_1$ is given by

$$C_{2,3|1}(u_2, u_3|u_1; \alpha) = \frac{\partial}{\partial u_1} C(u_1, u_2, u_3; \alpha) = u_2 u_3 (1 + \alpha(1 - 2u_1)(1 - u_2)(1 - u_3)).$$

Since $C_{j|1}(u_j|u_1; \alpha) = u_j$ for $j = 2, 3$ due to the pairwise independence, we obtain for the corresponding copula that

$$\begin{aligned} C_{2,3;1}(u_2, u_3|u_1; \alpha) &= C_{2,3|1}(u_2, u_3|u_1; \alpha) \\ &= u_2 u_3 (1 + \alpha(1 - 2u_1)(1 - u_2)(1 - u_3)) \\ &= C(u_2, u_3; \alpha(u_1)), \end{aligned}$$

which can be recognized as a bivariate FGM copula with parameter $\alpha(u_1) = \alpha(1 - 2u_1)$. In other words, the trivariate FGM copula (2.65) cannot be represented as a simplified PCC. Moreover, we can quantify the extent to which the FGM copula is of non-simplified type. According to Equation (2.54) for the bivariate FGM copula, the Kendall's τ of the conditioned copula $C_{2,3;1}$ can be determined as

$$\tau(\alpha(u_1)) = \frac{2\alpha(1 - 2u_1)}{9},$$

which ranges linearly between $2\alpha/9$ ($u_1 = 0$) and $-2\alpha/9$ ($u_1 = 1$). If $|\alpha| = 1$, then the range is maximal and given by $4/9 \approx 0.44$. □

Acar et al. (2012) also investigate non-simplified PCCs and take a first step in building PCCs of non-simplified structure in three dimensions. Their non-parametric approach is however not straightforward to generalize to d-dimensional PCCs, so that it is yet an open question how to deal with non-simplified PCCs in arbitrary dimensions.

Apart from the Clayton copula, it is also known that the multivariate Gaussian and the Student's t copula are of simplified type (see Stöber et al. (2013)). In particular, the parameters of the copulas $C_{j,k;\ell_1,\ldots,\ell_m}$ of conditional pairs of variables are the corresponding partial correlations $\rho_{j,k;\ell_1,\ldots,\ell_m}$, which can be computed recursively (see Kurowicka and Cooke (2003) and also Lewandowski et al. (2009)), and, in the case of the Student's t copula with ν degrees of freedom, the degrees of freedom of the pair copulas are $\nu + m$.

2.7.3 Vines

Bedford and Cooke (2001, 2002) introduced vines as a graph theoretical model to organize different PCCs. A graph is defined in terms of a set of nodes N and a set of edges E connecting these nodes, that is, $E \subset N \times N$. Vines are based on *trees*, which are particular graphs with a unique sequence of edges between each two nodes (also known as connected acyclic graphs). A *regular vine* is then defined as follows.

Definition 2.22 (Regular vine)**.** A set of linked trees $\mathcal{V} = (T_1, T_2, ..., T_{d-1})$ is called a regular vine (R-vine) on d elements if the following three conditions are satisfied.

(i) T_1 is a tree with nodes $N_1 = \{1, ..., d\}$ and a set of $d-1$ edges denoted by E_1.

(ii) For $i = 2, ..., d-1$, T_i is a tree with nodes $N_i = E_{i-1}$ and edge set E_i.

(iii) For $i = 2, ..., d-1$, if $a = \{a_1, a_2\}$ and $b = \{b_1, b_2\}$ are two nodes in N_i, which are connected by an edge, then exactly one of the a_is equals one of the b_is (proximity condition).

In other words, the proximity condition requires that the edges corresponding to two connected nodes in tree T_i share a common node in tree T_{i-1}. This ensures that the vine decomposition into bivariate copulas, which is given below, is well-defined, that is, that the corresponding PCC is, in fact, a valid decomposition.

Two sub-classes of R-vines have been studied extensively in the literature: *canonical vines (C-vines)* and *drawable vines (D-vines)* (see Kurowicka and Cooke (2006) and Aas et al. (2009)). C-vines are characterized by a root node in each tree T_i, $i \in \{1, ..., d-1\}$, which has degree $d-i$. In other words, the root node is connected to all other nodes of the tree. D-vines, on the other hand, are uniquely characterized through their first tree, which is, in graph theoretical terms, a path. This means that each node is connected to at most two other nodes. Therefore the order of the elements in the first tree defines the complete D-vine tree sequence, while C-vines are defined through the order of the root nodes.

Some more definitions are needed to introduce R-vine copulas: The complete union A_e of an edge $e = \{a, b\} \in E_i$ in tree T_i of an R-vine $\mathcal{V}$ is defined by

$$A_e = \{v \in N_1 : \ \exists e_m \in E_m, \ m = 1, ..., i-1, \ \text{such that } v \in e_1 \in \cdots \in e_{i-1} \in e\}.$$

The *conditioning set* associated with $e = \{a, b\}$ is defined as $D_e := A_a \cap A_b$ and the *conditioned sets* associated with $e = \{a, b\}$ are defined as $\mathcal{C}_{e,a} := A_a \backslash D_e$ and $\mathcal{C}_{e,b} := A_b \backslash D_e$. Bedford and Cooke (2001) showed that the conditioned sets are singletons, and we will therefore refer to edges by their labels $\{j(e), k(e)|D(e)\} := \{\mathcal{C}_{e,a}, \mathcal{C}_{e,b}|D_e\}$. An exemplary R-vine on five elements with edge labels is shown in Figure 2.8.

Given these sets, we can specify an R-vine copula by associating a (conditional) pair copula to each edge of the R-vine.

Definition 2.23 (Regular vine copula)**.** The random vector $\boldsymbol{U} = (U_1, ..., U_d)'$ with uniform margins is said to be distributed according to the d-dimensional R-vine copula $C(\cdot; \mathcal{V}, \mathcal{B}, \boldsymbol{\theta})$ if

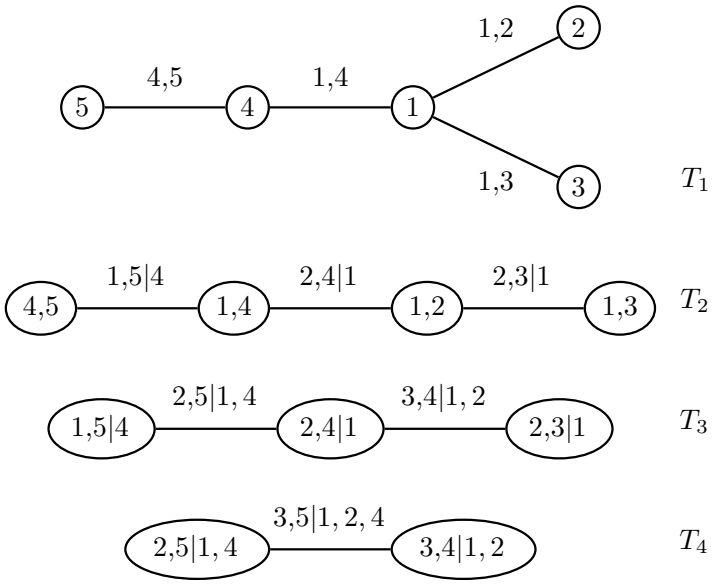


Figure 2.8: An R-vine on five elements with edge labels.

(i) $\mathcal{V}$ is an R-vine on d elements (see Definition 2.22),

(ii) $\mathcal{B} = \mathcal{B}(\mathcal{V}) = \left\{C_{j(e),k(e);D(e)} : e \in E_i,\ i = 1, ..., d-1\right\}$ is a set of $\binom{d}{2} = d(d-1)/2$ copula families identifying the conditional distributions of $(U_{j(e)}, U_{k(e)})'|\boldsymbol{U}_{D(e)}$, where $\boldsymbol{U}_{D(e)} = \{U_\ell : \ell \in D(e)\}$ is the sub-vector of $\boldsymbol{U}$ with indices in $D(e)$, and

(iii) $\boldsymbol{\theta} = \boldsymbol{\theta}(\mathcal{B}(\mathcal{V})) = \left\{\boldsymbol{\theta}_{j(e),k(e);D(e)} : e \in E_i,\ i = 1, ..., d-1\right\}$ is the set of parameters, corresponding to the copulas in $\mathcal{B}(\mathcal{V})$.

Therefore the full specification of an R-vine copula consists of three layers: the regular vine tree structure $\mathcal{V}$, the pair copula families $\mathcal{B} = \mathcal{B}(\mathcal{V})$ and the pair copula parameters $\boldsymbol{\theta} = \boldsymbol{\theta}(\mathcal{B}(\mathcal{V}))$. R-vine copulas that differ in the tree structure or in at least one pair copula family represent in general different statistical models. The density of an R-vine copula can be calculated as described in the following theorem.

Theorem 2.24 (Regular vine copula density)**.** Let $C(\cdot; \mathcal{V}, \mathcal{B}, \boldsymbol{\theta})$ be a d-dimensional R-vine copula. Then its density is given for $\boldsymbol{u} \in [0,1]^d$ by

$$c(\boldsymbol{u}; \mathcal{V}, \mathcal{B}, \boldsymbol{\theta}) = \prod_{i=1}^{d-1} \prod_{e \in E_i} c_{j(e),k(e);D(e)} \left(C_{j(e)|D(e)}(u_{j(e)}|\boldsymbol{u}_{D(e)}), C_{k(e)|D(e)}(u_{j(e)}|\boldsymbol{u}_{D(e)})\right), \quad (2.66)$$

where the copula $C_{j(e),k(e);D(e)}$ has parameter(s) $\boldsymbol{\theta}_{j(e),k(e);D(e)}$ and $C_{\ell|D(e)}$ is the conditional distribution function of $U_\ell|\boldsymbol{U}_{D(e)},\ \ell \in \{j(e), k(e)\}$.

Proof: See Bedford and Cooke (2001, Theorem 3). □

The conditional distribution functions as arguments of the pair copulas can be determined recursively tree-by-tree as in Equation (2.63) using the relationship

$$\begin{aligned} C_{j(e)|D(e)}\left(u_{j(e)}|\boldsymbol{u}_{D(e)}\right) &= C_{\mathcal{C}_{e,a}|D_e}\left(u_{\mathcal{C}_{e,a}}|\boldsymbol{u}_{D_e}\right) \\ &= C_{\mathcal{C}_{a,a_1}|\mathcal{C}_{a,a_2};D_a}\left(C_{\mathcal{C}_{a,a_1}|D_a}\left(u_{\mathcal{C}_{a,a_1}}|\boldsymbol{u}_{D_a}\right) \mid C_{\mathcal{C}_{a,a_2}|D_a}\left(u_{\mathcal{C}_{a,a_2}}|\boldsymbol{u}_{D_a}\right)\right), \end{aligned} \tag{2.67}$$

where $e = \{a, b\}$ with $a = \{a_1, a_2\}$ as before and

$$C_{\mathcal{C}_{a,a_1}|\mathcal{C}_{a,a_2};D_a}\left(u_{\mathcal{C}_{a,a_1}}|u_{\mathcal{C}_{a,a_2}}\right) = \frac{\partial C_{\mathcal{C}_{a,a_1},\mathcal{C}_{a,a_2};D_a}\left(u_{\mathcal{C}_{a,a_1}}, u_{\mathcal{C}_{a,a_2}}\right)}{\partial u_{\mathcal{C}_{a,a_2}}}, \tag{2.68}$$

as in Equation (2.64). Similarly $C_{k(e)|D(e)}$ is obtained. Further, if $i = 1$ (first tree), we have that

$$C_{j(e)|D(e)}(u_{j(e)}|\boldsymbol{u}_{D(e)}) = C_{j(e)|\emptyset}(u_{j(e)}|\boldsymbol{u}_{\emptyset}) = u_{j(e)},$$

and similarly for $C_{k(e)|D(e)}$, so that the pair copulas corresponding to the first R-vine tree are, in fact, unconditional ones. Note that only conditional distributions $C_{j|k;D}$ and $C_{k|j;D}$ of pair copulas $C_{j,k;D}$ in lower order trees are required for the recursive calculation. In the context of vine copulas, these conditional distributions are also often called *h-functions*, of which explicit expressions for many common bivariate copulas are given, e.g., in Aas et al. (2009). More details on the recursion (2.67) can be found in Dißmann et al. (2013).

Theorem 2.24 relates vine copulas to PCCs as introduced in Section 2.7.1. In fact, the four-dimensional PCC derived there corresponds to the vine copula defined through the four-dimensional sub-vine on the elements 1, 2, 3 and 4 of the R-vine in Figure 2.8. This sub-vine corresponds to a C-vine with root node order $1, 2, 3, 4$ as shown in Figure 2.9, which better illustrates the role of the root nodes in each tree.

If $e \in E_i$ in a C-vine with root node order $1, ..., d$, then it holds that $j(e) = i$, $k(e) = i + k$, $k = 1, ..., d - i$, and $D(e) = \{1, ..., i - 1\}$ (see also Czado (2010)). Therefore, we can rewrite the density of an R-vine copula (2.66) for a C-vine copula $C(\cdot; \mathcal{V}, \mathcal{B}, \boldsymbol{\theta})$ and $\boldsymbol{u} \in [0, 1]^d$ as

$$\begin{aligned} &c(\boldsymbol{u}; \mathcal{V}, \mathcal{B}, \boldsymbol{\theta}) \\ &= \prod_{1=1}^{d-1} \prod_{k=1}^{d-i} c_{i,i+k;1,...,i-1}(C_{i|1,...,i-1}(u_i|u_1, ..., u_{i-1}), C_{i+k|1,...,i-1}(u_{i+k}|u_1, ..., u_{i-1})), \end{aligned} \tag{2.69}$$

where the arguments $C_{i+k|1,...,i-1}$, $k = 0, ..., d - i$, are calculated as in Equation (2.67). More precisely, we obtain

$$\begin{aligned} &C_{i+k|1,...,i-1}(u_{i+k}|u_1, ..., u_{i-1}) \\ &\quad = C_{i+k|i-1;1,...,i-2}\left(C_{i+k|1,...,i-2}(u_{i+k}|u_1, ..., u_{i-2})|C_{i-1|1,...,i-2}(u_{i-1}|u_1, ..., u_{i-2})\right), \end{aligned} \tag{2.70}$$

where we have according to Equation (2.68) that

$$C_{i+k|i-1;1,...,i-2}(u_{i+k}|u_{i-1}) = \frac{\partial C_{i-1,i+k;1,...,i-2}(u_{i-1}, u_{i+k})}{\partial u_{i-1}} \tag{2.71}$$

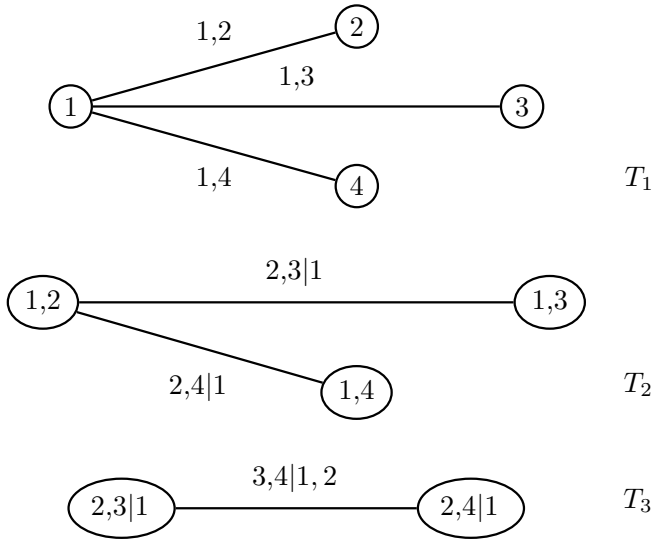


Figure 2.9: A C-vine on four elements with edge labels corresponding to the PCC of Section 2.7.1.

for $i = 2, ..., d-1$ and $k = 0, ..., d-i$.

An alternative principle to define PCCs is using directed acyclic graphs (DAGs) as investigated by Hanea et al. (2006) and Bauer et al. (2012). We do not follow this approach here. More details on vine copulas can also be found in Kurowicka and Cooke (2006), Czado (2010) and Kurowicka and Joe (2011).

2.7.4 Statistical inference

An algorithm for sampling from a given R-vine copula is provided in Dißmann et al. (2013). For the special cases of C- and D-vines, simpler algorithms can be found in Aas et al. (2009). Estimation of vine copula parameters $\boldsymbol{\theta} = \boldsymbol{\theta}(\mathcal{B}(\mathcal{V}))$ can be based on maximum likelihood techniques, since the log likelihood expression is computationally well tractable. Appropriate algorithms for R-vine copulas and C- and D-vine copulas are also given in Dißmann et al. (2013) and Aas et al. (2009), respectively (see also Hobæk Haff (2012, 2013) and Stöber and Schepsmeier (2013)).

A common approach to quickly obtain parameter estimates is a sequential method, which only requires bivariate maximum likelihood estimation. According to this method, first the parameters of the pair copulas in the first tree are estimated. Given the estimated parameters, those of the pair copulas in the second tree are estimated, and so on. Such parameter estimates often provide good starting values for a joint estimation of all copula parameters.

The selection of vine copulas is very demanding: Both an appropriate tree structure $\mathcal{V}$ as well as adequate pair copulas $\mathcal{B} = \mathcal{B}(\mathcal{V})$ need to be chosen. Since the number of different d-dimensional regular vines is excessive ($d!/2 \times 2^{\binom{d-2}{2}}$ as shown by Morales-Nápoles (2011)) and the pair copulas can be chosen from arbitrary classes, the class of vine copulas is so large that an exhaustive examination of all possible models is virtually

impossible, unless d is very small. Recently, Dißmann et al. (2013) proposed a sequential top-down selection procedure, which selects vine trees as well as pair copulas such that the strongest dependencies among the variables in terms of Kendall's τ are captured in the first trees.

Algorithm 2.25 (Sequential R-vine copula selection)**.** Let a sample $(u_{mj})_{m=1,...,N,\, j=1,...,d}$ of size N be given. The margins are assumed to be uniformly distributed.

(i) Calculate the empirical Kendall's τ value denoted by $\widehat{\tau}_{j,k}$ based on $(u_{mj}, u_{mk})_{m=1,...,N}$ for all variable pairs $\{j,k\},\ 1 \leq j < k \leq d$.

(ii) Select the maximum spanning tree in terms of the absolute empirical Kendall's τ values, that is,

$$T_1 = \underset{T=(N,E) \text{ spanning tree}}{\operatorname{argmax}} \quad \sum_{e \in E} |\widehat{\tau}_{j(e),k(e)}|,$$

where a spanning tree is a tree on all nodes $1, ..., d$.

(iii) For each edge $e \in E_1$:

a) Select a copula $C_{j(e),k(e)}$.

b) Estimate the corresponding parameter(s) $\boldsymbol{\theta}_{j(e),k(e)}$.

c) Calculate $C_{j(e)|k(e)}(u_{m,j(e)}|u_{m,k(e)})$ and $C_{k(e)|j(e)}(u_{m,k(e)}|u_{m,j(e)})$ for $m = 1, ..., N$.

(iv) For $i = 2, ..., d-1$:

a) Calculate the empirical Kendall's τ value denoted by $\widehat{\tau}_{j(e),k(e)|D(e)}$ based on $(C_{j(e)|D(e)}(u_{m,j(e)}|\boldsymbol{u}_{m,D(e)}), C_{k(e)|D(e)}(u_{m,k(e)}|\boldsymbol{u}_{m,D(e)}))_{m=1,...,N}$ for all conditional variable pairs $\{i(e), j(e)|D(e)\}$ that can be part of tree T_i, that is, all edges E_P fulfilling the proximity condition.

b) Among the edges E_P, select the maximum spanning tree in terms of the absolute empirical Kendall's τ values, that is,

$$T_i = \underset{T=(N,E) \text{ spanning tree with } E \subset E_P}{\operatorname{argmax}} \quad \sum_{e \in E} |\widehat{\tau}_{i(e),j(e)|D(e)}|.$$

c) For each edge $e \in E_i$:

i. Select a copula $C_{j(e),k(e);D(e)}$.

ii. Estimate the corresponding parameter(s) $\boldsymbol{\theta}_{j(e),k(e);D(e)}$.

iii. According to Equation (2.67), calculate $C_{j(e)|k(e)\cup D(e)}(u_{m,j(e)}|\boldsymbol{u}_{m,k(e)\cup D(e)})$ and $C_{k(e)|j(e)\cup D(e)}(u_{m,k(e)}|\boldsymbol{u}_{m,j(e)\cup D(e)})$ for $m = 1, ..., N$ using $C_{j(e)|k(e);D(e)}$ and $C_{k(e)|j(e);D(e)}$, respectively.

For the selection of pair copulas, often the Akaike information criterion (AIC) by Akaike (1973) or the Bayesian information criterion (BIC) by Schwarz (1978) are proposed (see Brechmann (2010, Section 5.4.4) for a comparison study of different selection criteria). Furthermore, this algorithm can be modified to select a C-vine copula rather than a

general R-vine copula. For this, select each root node such that it maximizes the sum of absolute empirical Kendall's τ values with respect to the other variables (see Czado et al. (2012)).

Sequential selection in this way also leads to the idea of R-vine copula truncation, where pair copulas in higher order trees are set to independence copulas to reduce the model complexity (see Brechmann et al. (2012)). An overview of R-vine selection methods, including a bottom-up vine tree selection algorithm based on partial correlations by Kurowicka (2011) and Bayesian approaches, can be found in Czado et al. (2013). Implementations of a wide range of inference techniques for C- and D-vine copulas are available in the R-package `CDVine` (Schepsmeier and Brechmann, 2011; Brechmann and Schepsmeier, 2013) and for R-vine copulas in the R-package `VineCopula` (Schepsmeier et al., 2012).

Copula	Class	Parameters	Kendall's τ	Spearman's ρ_S	Lower tail dep.	Upper tail dep.
Gaussian	elliptical	$\rho \in (-1, 1)$	$\frac{2}{\pi} \arcsin(\rho)$	$\frac{6}{\pi} \arcsin\left(\frac{\rho}{2}\right)$	0	0
Student's t	elliptical	$\rho \in (-1, 1),\ \nu > 2$	$\frac{2}{\pi} \arcsin(\rho)$	$*$	$2\,T_{\nu+1}\left(-\sqrt{\nu+1}\sqrt{\frac{1-\rho}{1+\rho}}\right)$	
Indiv. Student's t	extended ellip.	$\rho \in (-1, 1),\ \nu_1, \nu_2 > 2$	$\approx \frac{2}{\pi} \arcsin(\rho)$	$*$	see (2.30)	
Clayton	Archimedean	$\theta > 0$	$\frac{\theta}{\theta+2}$	$*$	$2^{-1/\theta}$	0
Gumbel	Arch./extr. value	$\theta \geq 1$	$1 - \frac{1}{\theta}$	$*$	0	$2 - 2^{1/\theta}$
Frank	Archimedean	$\theta \in \mathbb{R} \setminus \{0\}$	$1 + \frac{4}{\theta}(D_1(\theta) - 1)$	$1 - \frac{12}{\theta}(D_1(\theta) - D_2(\theta))$	0	0
Joe	Archimedean	$\theta > 1$	see (2.41)	$*$	0	$2 - 2^{1/\theta}$
Tawn	extreme value	$\theta \geq 1,\ \psi_1, \psi_2 \in [0, 1]$	$*$	$*$	0	$\psi_1 + \psi_2 - (\psi_1^\theta + \psi_2^\theta)^{1/\theta}$
Plackett	–	$\alpha \in (-1, \infty) \setminus \{0\}$	$*$	$\frac{\alpha+2}{\alpha} - \frac{2(\alpha+1)}{\alpha^2} \log(\alpha + 1)$	0	0

Table 2.1: Properties of common bivariate copulas. An asterisk indicates that there is no simple expression known.

3 Hierarchical Kendall copulas

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 chapter, the new class of hierarchical Kendall copulas is proposed and discussed. After deriving properties of the general model formulation, statistical inferences techniques for hierarchical Kendall copulas are developed and evaluated. This chapter is mainly based on Brechmann (2013a).

3.1 Introduction

Although dependence modeling using copulas has made significant progress in the last years, 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, for example, in financial or spatial applications. While standard multivariate elliptical copulas such as the Gaussian and the Student's t (see Section 2.2) require the specification of the full correlation matrix and can only account for symmetric dependence, multivariate Archimedean copulas (see Section 2.4) are even more restrictive by assuming exchangeability and imposing that all multivariate margins are the same. One common procedure to approach such problems therefore is grouping data, for example, by industry sectors or nationality. Such copula models include the grouped Student's t copula by Daul et al. (2003) (see Section 2.3), elliptical copulas with two-factor correlation structure (one group-specific and one overall factor; see, e.g., Gregory and Laurent (2004)) and hierarchical Archimedean copulas, which were initially proposed by Joe (1993, 1997). Especially such hierarchical structures 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. Nevertheless, 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 structured correlation matrix has to satisfy positive definiteness constraints and is limited to an elliptical dependence structure, which in particular implies reflection symmetry.

Vine copulas, which constitute another class of non-hierarchical multivariate copulas (see Section 2.7), gain their flexibility from using a quadratic number of bivariate copulas

of arbitrary types as building blocks. Therefore, they may also severely suffer from extreme numbers of parameters in higher dimensions. While the selection of independence copulas as building blocks or even truncation methods (see Brechmann et al. (2012)) may be used to counteract such problems, we focus here on hierarchical constructions, which are inherently more parsimonious.

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 vine copulas 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 (see Definition 2.9) in the model formulation. The Kendall distribution function is the multivariate analog to the probability integral transform for univariate random variables. In combination with the copula, it is used to aggregate the (dependence) information of a group of variables. As noted in Section 2.1, 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 (each variable belongs to exactly one group) is the independence copula. That is, if in Sklar's Theorem (2.1) 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 copula and 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 for hierarchical Kendall copulas. Sampling from hierarchical Kendall copulas is treated in detail in Chapter 4.

The features of hierarchical dependence models in general 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 (see the discussion in Section 6.4.5). 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 collateralized debt obligations. For the purpose of market risk portfolio management, a substantial 30-dimensional application to German stock returns is presented at the end of this chapter, showing the need for

careful dependence modeling and a good in-sample fit of hierarchical Kendall copulas. The out-of-sample performance is considered in Section 4.5.

Nonetheless, the model is 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 (see Chapter 7 for more details).

The remainder of the chapter is organized as follows. In Section 3.2 we first discuss the general idea of aggregating information of groups of variables to construct a hierarchical dependence model and examine different aggregation functions. The new model based on aggregation through the copula and the Kendall distribution function is then introduced and discussed in Section 3.3. A comparison to hierarchical Archimedean copulas is provided in Section 3.3.1, and statistical inference techniques for hierarchical Kendall copulas are presented in the following. A general sampling algorithm is given in Section 3.4 (with a detailed discussion in Chapter 4), and Section 3.5 treats the estimation of the parameters of a hierarchical Kendall copula. The appropriate selection of the components of a hierarchical Kendall copula is discussed in Section 3.6. The financial application is finally presented in Section 3.7, before Section 3.8 concludes.

3.2 Hierarchical copulas and aggregation functions

Our approach to hierarchical dependence modeling that we will pursue here is the following: In order to balance flexibility and parsimony and also to allow for interpretability of a multivariate dependence model, groups of variables are aggregated into univariate quantities, in terms of which between-group dependence is quantified. Therefore, we investigate how to appropriately summarize the information of a random vector. Since all continuous random variables can be transformed to be uniformly distributed using the probability integral transform, we directly consider a continuous random vector $\boldsymbol{U} := (U_1, ..., U_d)' \in [0,1]^d$, where $U_j \sim U(0,1),\ j = 1, ..., d$, and assume that it is absolutely continuous.

Summarizing multivariate information in a single quantity is a common problem, for example, in the social sciences, engineering and economics, and it exists a diverse literature on the subject (see, amongst others, Calvo et al. (2002), Beliakov et al. (2008) and Grabisch et al. (2009)). We call H an *aggregation function* if it maps a multivariate value $\boldsymbol{u} \in [0,1]^d$ to a univariate one, $H(\boldsymbol{u}) \in \mathbb{R}$. Here, we concentrate on aggregation functions with $H(\boldsymbol{u}) \in [0,1]$. Clearly, if the values of an aggregation function are bounded, it can be standardized to take on values in $[0,1]$. Moreover, we focus on aggregation functions H such that $H(\boldsymbol{U})$ has a continuous distribution function $F_{H(\boldsymbol{U})}$. This guarantees that we can again apply the probability integral transform to $H(\boldsymbol{U})$ and use a continuous copula to specify the between-group dependence.

Making use of such an aggregation function, we can constructively define the corresponding two-level hierarchical copula as follows.

Definition 3.1 (Hierarchical copula)**.** Let $U_1, ..., U_n \sim U(0,1)$ and let $C_0, C_1, ..., C_{n_1}$ be copulas of dimension $n_1, d_1, ..., d_{n_1}$, respectively, where $d_i \geq 1,\ i = 1, ..., n_1$, and $n = \sum_{i=1}^{n_1} d_i$. We define the cumulative sum $m_i = \sum_{j=1}^{i} d_j,\ i = 1, ..., n_1$, and $m_0 = 0$ as well as $\boldsymbol{U}_i := (U_{m_{i-1}+1}, ..., U_{m_i})'$ and $V_i := F_{H_i(\boldsymbol{U}_i)}(H_i(\boldsymbol{U}_i))$ for $i = 1, ..., n_1$. Under the assumptions that

$\mathcal{A}_1$: $\boldsymbol{U}_1, ..., \boldsymbol{U}_{n_1}$ are mutually independent conditionally on $(V_1, ..., V_{n_1})'$, and

$\mathcal{A}_2$: the conditional distribution of $\boldsymbol{U}_i|(V_1, ..., V_{n_1})'$ is the same as the conditional distribution of $\boldsymbol{U}_i|V_i$ for all $i = 1, ..., n_1$,

the random vector $(U_1, ..., U_n)'$ is said to be distributed according to the hierarchical copula $C_{\mathcal{H}}$ with aggregation functions $H_1, ..., H_{n_1}$, cluster copulas $C_1, ..., C_{n_1}$ and nesting copula C_0 if

(i) $\boldsymbol{U}_i \sim C_i\ \forall i \in \{1, ..., n_1\}$,

(ii) $(V_1, ..., V_{n_1})' \sim C_0$.

This definition is straightforward to extend to the case of k hierarchical levels. We consider this in more detail in the next section. For the discussion of aggregation functions, it is sufficient to work with this two-level definition of a hierarchical copula.

The intuition behind the two assumptions $\mathcal{A}_1$ and $\mathcal{A}_2$ is that, given the information of the nesting variables $V_1, ..., V_{n_1}$, the clusters $\boldsymbol{U}_1, ..., \boldsymbol{U}_{n_1}$ are independent of each other and also of other nesting variables, since the dependence among the clusters is explained through the "representatives" $V_1, ..., V_{n_1}$. In other words, $V_1, ..., V_{n_1}$ can be interpreted as unobserved factors, whose joint behavior determines the dependence of the different clusters. In finance, such factors may be, for instance, industry sectors.

Now, the question is what an adequate aggregation function H to summarize the information of a d-dimensional random vector $\boldsymbol{U}$ should look like. We consider the following examples:

(i) The *minimum* $H_{\min}$,

$$H_{\min}(\boldsymbol{u}) = \min\{u_1, ..., u_d\}, \quad \boldsymbol{u} \in [0,1]^d.$$

(ii) The *maximum* $H_{\max}$,

$$H_{\max}(\boldsymbol{u}) = \max\{u_1, ..., u_d\}, \quad \boldsymbol{u} \in [0,1]^d.$$

(iii) The *arithmetic mean* H_{a},

$$H_{\mathrm{a}}(\boldsymbol{u}) = \frac{1}{d}\sum_{j=1}^{d} u_j, \quad \boldsymbol{u} \in [0,1]^d.$$

(iv) The *weighted arithmetic mean* H_{wa} with weights $\omega_j > 0,\ j = 1, ..., d$,

$$H_{\mathrm{wa}}(\boldsymbol{u}) = \frac{1}{\sum_{j=1}^{d} \omega_j}\sum_{j=1}^{d} \omega_j u_j, \quad \boldsymbol{u} \in [0,1]^d.$$

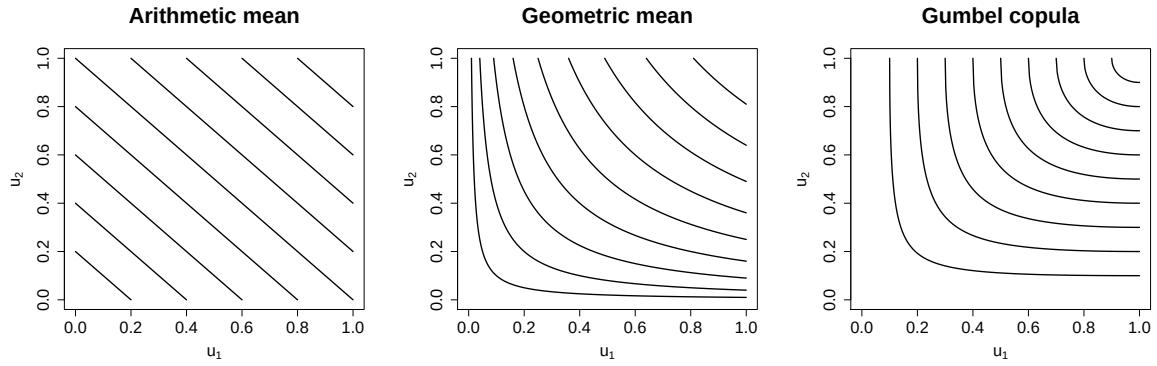


Figure 3.1: Level sets of the arithmetic mean, of the geometric mean and of the Gumbel copula with parameter chosen according to a Kendall's τ of 0.5.

(v) The *geometric mean* H_{g},

$$H_{\mathrm{g}}(\boldsymbol{u}) = \left(\prod_{j=1}^{d} u_j\right)^{1/d}, \quad \boldsymbol{u} \in [0,1]^d.$$

(vi) The *weighted geometric mean* H_{wg} with weights $\omega_j > 0,\ j = 1, ..., d$,

$$H_{\mathrm{wg}}(\boldsymbol{u}) = \left(\prod_{j=1}^{d} u_j^{\omega_j}\right)^{1/\sum_{j=1}^{d} w_j}, \quad \boldsymbol{u} \in [0,1]^d.$$

In contrast to the sum, the product is also a possible aggregation function mapping to $[0,1]$. This also holds for any copula C. Especially, let C be the copula of $\boldsymbol{U}$. Then, we also consider a seventh aggregation function.

(vii) The *copula* H_C,

$$H_C(\boldsymbol{u}) = C(\boldsymbol{u}), \quad \boldsymbol{u} \in [0,1]^d.$$

The level sets of the arithmetic mean, $L(z, H_{\mathrm{a}})$, of the geometric mean, $L(z, H_{\mathrm{g}})$, and of the Gumbel copula, $L(z, H_C) = L(z, C)$, are illustrated in Figure 3.1 (see Equation (2.10)).

In order to examine the usefulness of different aggregation functions, we consider some basic properties.

Definition 3.2 (Properties of aggregation functions)**.** Let $H : [0,1]^d \to [0,1]$ be an aggregation function. Then H satisfies the *boundary conditions* if

(A1) $H(\mathbf{0}) = 0$,

(A2) $H(\mathbf{1}) = 1$, and

(A3) it reduces to the identity function $H(u) = u,\ u \in [0,1]$, in the degenerate case that $d = 1$.

Further, H is called

(A4) *monotone* if $H(\boldsymbol{u}) \leq H(\boldsymbol{v})$ for every $\boldsymbol{u}, \boldsymbol{v} \in [0,1]^d$ such that $u_j \leq v_j$ for all $j = 1, ..., d$.

(A5) *strictly monotone* if $H(\boldsymbol{u}) < H(\boldsymbol{v})$ for every $\boldsymbol{u}, \boldsymbol{v} \in [0,1]^d$ such that $u_j \leq v_j$ for all $j = 1, ..., d$, and at least one $u_j < v_j,\ j \in \{1, ..., d\}$.

(A6) *permutation symmetric* if $H(\boldsymbol{u}) = H(u_{\sigma(1)}, ..., u_{\sigma(p)})$ for every $\boldsymbol{u} \in [0,1]$ and any permutation $\sigma : \{1, ..., d\} \to \{1, ..., d\}$.

All seven examples of aggregation functions that we consider here satisfy the boundary conditions (A1)–(A3). Especially boundary condition (A3) is important, since a degenerate group of size 1, that is, a single random variable, should not be altered by the aggregation function. The considered aggregation functions are also all monotone, which is desirable, since larger input values should result in larger aggregated values. However, $H_{\min}$ and $H_{\max}$ are not strictly monotone. Further, the functions H_{wa}, H_{wg} and H_C are the only aggregation functions, which are not necessarily permutation symmetric. The function H_C is permutation symmetric if and only if the copula C is exchangeable, while H_{wa} and H_{wg} are permutation symmetric if and only if $\omega_1 = \omega_2 = ... = \omega_d$.

For the purpose of aggregating the information of random vectors, it is reasonable to demand strict monotonicity (A5), since an increase in at least one component of the input values should increase the aggregated value and not leave it unchanged. This rules out $H_{\min}$ and $H_{\max}$. Further, if the random vector $\boldsymbol{U}$ is non-exchangeable, the aggregation function H should not be permutation symmetric (A6) in order to reflect this non-exchangeability. This also rules out H_{a} and H_{g}, so that only H_{wa}, H_{wg} and H_C appear to be sensible choices. In contrast to H_C, it is however unclear how to choose the weights of H_{wa} and H_{wg}, such that the non-exchangeability of $\boldsymbol{U}$ is reflected appropriately. In addition, the distribution of the univariate random variables $H_{\text{wa}}(\boldsymbol{U})$ and $H_{\text{wg}}(\boldsymbol{U})$ is difficult to determine even in the bivariate case. For $(U_1, U_2)' \sim C$, it holds that

$$\begin{aligned} P\left(H_{\text{wa}}(U_1, U_2) \leq z\right) &= P\left(\omega_1 U_1 + \omega_2 U_2 \leq z(\omega_1 + \omega_2)\right) \\ &= \int_0^1 P\left(\omega_1 U_1 + \omega_2 U_2 \leq z(\omega_1 + \omega_2) | U_1 = u_1\right) du_1 \\ &= \int_0^1 C_{2|1}\left(z + \frac{\omega_1}{\omega_2}(z - u_1) \mid u_1\right) du_1, \quad z \in [0,1], \end{aligned}$$

where $C_{2|1}$ is defined in Equation (2.13). In the same way, it holds that

$$P\left(H_{\text{wg}}(U_1, U_2) \leq z\right) = \int_0^1 C_{2|1}\left(z \left(\frac{z}{u_1}\right)^{\omega_1/\omega_2} \mid u_1\right) du_1, \quad z \in [0,1].$$

These expressions are similar to the derivation of the Kendall distribution function (see Definition 2.9), which is the distribution function of $H_C(\boldsymbol{U}) = C(\boldsymbol{U})$ (see Equation (2.15)). Although its derivation is generally also rather complicated, there are closed-form expressions for common classes of copulas such as Archimedean copulas (see Equation (2.36)),

extreme value and Archimax copulas (see Equations (2.44) and (2.49)) and the Plackett copula (see Appendix B.3).

Furthermore, using the copula as aggregation function has the attractive interpretation that the definition of the hierarchical copula (see Definition 3.1) mimics Sklar's Theorem (2.1) for multivariate non-overlapping margins, since the Kendall distribution function constitutes the multivariate analog of the univariate probability transform in the sense that $C(\boldsymbol{U}) \sim K(\cdot; C)$ and $K(C(\boldsymbol{U}); C) \sim U(0,1)$. We therefore concentrate on the copula and its Kendall distribution function for the construction of a flexible hierarchical dependence model.

3.3 Model formulation and properties

We now restate Definition 3.1 using copulas and their Kendall distribution functions as aggregation functions. Although this 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 constructive model formulation.

Definition 3.3 (Hierarchical Kendall copula)**.** Let $U_1, ..., U_n \sim U(0,1)$ and let C_0, C_1, ..., C_{n_1} be copulas of dimensions $n_1, d_1, ..., d_{n_1}$, respectively, where $d_i \geq 1,\ i = 1, ..., n_1$, and $n = \sum_{i=1}^{n_1} d_i$. Further, let $K_1, ..., K_{n_1}$ denote the Kendall distribution functions corresponding to $C_1, ..., C_{n_1}$, that is, $K_i(\cdot) := K(\cdot; C_i),\ i = 1, ..., n_1$. We define the cumulative sum $m_i = \sum_{j=1}^{i} d_j,\ i = 1, ..., n_1$, and $m_0 = 0$ as well as $\boldsymbol{U}_i := (U_{m_{i-1}+1}, ..., U_{m_i})'$ and $V_i := K_i(C_i(\boldsymbol{U}_i))$ for $i = 1, ..., n_1$. Under the assumptions that

$\mathcal{A}_1$: $\boldsymbol{U}_1, ..., \boldsymbol{U}_{n_1}$ are mutually independent conditionally on $(V_1, ..., V_{n_1})'$, and

$\mathcal{A}_2$: the conditional distribution of $\boldsymbol{U}_i | (V_1, ..., V_{n_1})'$ is the same as the conditional distribution of $\boldsymbol{U}_i | V_i$ for all $i = 1, ..., n_1$,

the random vector $\boldsymbol{U} := (U_1, ..., U_n)'$ is said to be distributed according to the *hierarchical Kendall copula* $C_{\mathcal{K}}$ with cluster copulas $C_1, ..., C_{n_1}$ and nesting copula C_0 if

(i) $\boldsymbol{U}_i \sim C_i\ \forall i \in \{1, ..., n_1\}$,

(ii) $(V_1, ..., V_{n_1})' \sim C_0$.

The distribution function $C_{\mathcal{K}}$ of $\boldsymbol{U}$ will be characterized in terms of its density below. First, we discuss the construction, which is illustrated in Figure 3.2, in more detail and provide examples. Note that C_0 is in general not the copula of $\boldsymbol{U} = (U_1, ..., U_n)'$ but of $(V_1, ..., V_{n_1})'$, which are uniform random variables due to $C_i(\boldsymbol{U}_i) \sim K_i$ for all $i = 1, ..., n_1$. The nesting copula C_0 then essentially models the co-movement of the copula level sets (2.10) of the different clusters. The dimensionality of the clusters, which can be of different size, is "normalized" through the Kendall distribution functions. Further, the nesting copula C_0 and the cluster copulas $C_1, ..., C_{n_1}$ can be chosen independently. They can be arbitrary copulas, for example, from the classes discussed in Chapter 2. A special case of hierarchical Kendall copulas is the upper Fréchet-Hoeffding bound.

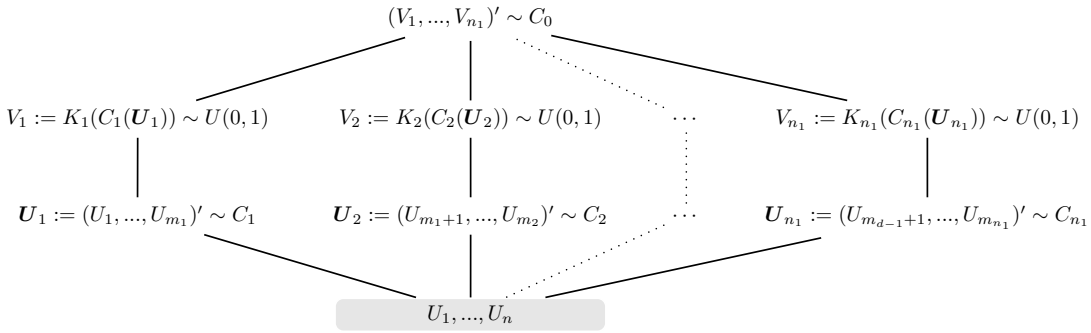


Figure 3.2: Illustration of a two-level hierarchical Kendall copula (see Definition 3.3).

Example 3.4 (Upper Fréchet-Hoeffding bound)**.** Let $C_{\mathcal{K}}$ be a hierarchical Kendall copula, where the clusters are comonotonic and the nesting copula is also the upper Fréchet-Hoeffding bound $C_0(v_1, ..., v_{n_1}) = M(v_1, ..., v_{n_1}) = \min\{v_1, ..., v_{n_1}\}$ (see Theorem 2.3). Since the Kendall distribution function of the comonotonicity copula is the identity function, it holds that $V_i = \min\{U_{m_{i-1}+1}, ..., U_{m_i}\}$ and therefore $C_{\mathcal{K}} = M$. In other words, the upper Fréchet-Hoeffding bound belongs to the class of hierarchical Kendall copulas. □

In the case of Archimedean clusters, we can give a particularly convenient representation of the multivariate distribution of the random vector $\boldsymbol{U}$.

Remark 3.5 (Hierarchical Kendall copula with Archimedean clusters)**.** Let $\boldsymbol{U}$ be distributed according to a hierarchical Kendall copula $C_{\mathcal{K}}$, where the cluster copulas C_1, ..., C_{n_1} are Archimedean with generators $\varphi_1, ..., \varphi_{n_1}$, respectively. According to the characterization by McNeil and Nešlehová (2009) (see Equation (2.32)), it holds for all $i = 1, ..., n_1$, that

$$(\varphi_i(U_{m_{i-1}+1}), ..., \varphi_i(U_{m_i}))' \stackrel{d}{=} R_i \boldsymbol{S}^{(i)},$$

where $\boldsymbol{S}^{(i)} = (S_1^{(i)}, ..., S_{d_i}^{(i)})'$ is uniformly distributed on the d_i-dimensional unit simplex $\mathcal{S}_{d_i-1}$ (see Equation (2.33)) and the radial part $R_i = \sum_{j=1}^{d_i} \varphi_i(U_{m_{i-1}+j})$ is independent of $\boldsymbol{S}^{(i)}$. As a result we can represent the random vector $\boldsymbol{U} = (U_1, ..., U_n)'$ as

$$\boldsymbol{U} \stackrel{d}{=} (\varphi_1^{-1}(R_1 S_1^{(1)}), ..., \varphi_1^{-1}(R_1 S_{d_1}^{(1)}), \varphi_2^{-1}(R_2 S_1^{(2)}), ..., \varphi_2^{-1}(R_2 S_{d_2}^{(2)}), ..., \varphi_{n_1}^{-1}(R_d S_{d_{n_1}}^{(n_1)}))', \tag{3.1}$$

where $R_i = \varphi_i(K_i^{-1}(V_i))$ for $i = 1, ..., n_1$, since by the definitions of C_i, V_i and R_i,

$$\begin{aligned} V_i &= K_i(C_i(U_{m_{i-1}+1}, ..., U_{m_i})) \\ &= K_i\left(\varphi_i^{-1}\left(\textstyle\sum_{j=1}^{d_i} \varphi_i(U_{m_{i-1}+j})\right)\right) = K_i(\varphi_i^{-1}(R_i)). \end{aligned}$$

In other words, if all clusters are Archimedean, dependence among the 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 $i = 1, ..., n_1$ that $V_i = \varphi_0^{-1}(R_0 S_i^{(0)})$, where $S_1^{(0)}, ..., S_{n_1}^{(0)}$ are uniformly distributed on the n_1-dimensional unit simplex $\mathcal{S}_{n_1-1}$. Hence,

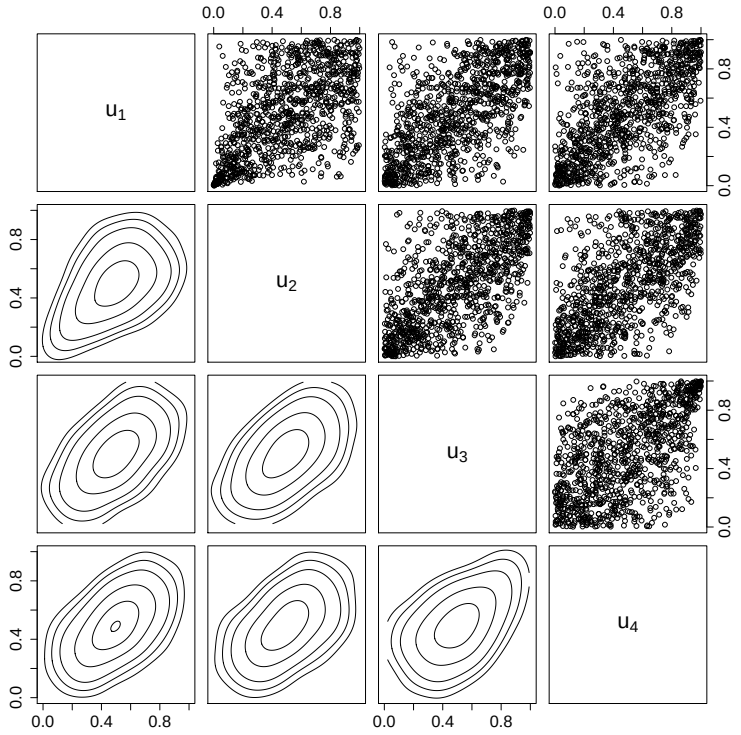


Figure 3.3: A sample of size 1000 from a four-dimensional hierarchical Kendall copula with bivariate Clayton and Gumbel clusters and Frank nesting copula. The lower triangle shows contour lines of the pairwise empirical densities with standard normal margins (the corresponding axes range from -3 to 3).

the radial variables of the clusters, R_i, can be expressed through R_0 and uniform random variables on the simplex.

Equation (3.1) also motivates to speak of a "grouped Archimedean copula" similar to the grouped Student's t copula by Daul et al. (2003) (see Section 2.3). In contrast to the grouped Student's t copula, where between-group dependence depends on the correlation matrix and on the degrees of freedom parameters of the different groups, a hierarchical Kendall copula explicitly controls this between-group dependence through the nesting copula. This is also the case for the hierarchical Archimedean copula, which will be discussed in Section 3.3.1. □

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

Example 3.6 (Hierarchical Kendall copula with Archimedean clusters)**.** Let $C_{\mathcal{K}}$ be a four-dimensional hierarchical Kendall copula with $d_1 = d_2 = 2$. The bivariate cluster copulas are chosen as Clayton (see Example 2.15) with parameter $\theta_1 = 1.33$ and Gumbel (see Example 2.16) with parameter $\theta_2 = 1.67$ (both parameters correspond to a Kendall's τ of 0.4). The nesting copula is set as a Frank (see Example 2.17) with parameter $\theta_0 = 11.41$ (Kendall's τ of 0.7). Figure 3.3 shows a sample of size 1000 from this hierarchical Kendall

copula. It shows the typical features of lower tail dependence for the pair $\boldsymbol{U}_1 = (U_1, U_2)'$ (Clayton copula) and of upper tail dependence for the pair $\boldsymbol{U}_2 = (U_3, U_4)'$ (Gumbel copula). The between-cluster dependence looks reflection symmetric as implied by the Frank copula. Corresponding empirical between-cluster Kendall's τ values lie between 0.45 and 0.47. □

The illustration in Example 3.6 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 3.3.1.

The two-level construction given in Definition 3.3 can also be extended to an arbitrary number of levels.

Remark 3.7 (Hierarchical Kendall copula with k levels)**.** Let $U_1, ..., U_n \sim U(0,1)$ and let $n_j,\ j = 1, ..., k-1$, denote the number of clusters per level j, such that $n_1 \geq n_2 \geq ... \geq n_{k-1}$. Further, let the nesting copula C_0 be n_{k-1}-dimensional and let the nested cluster copulas $C_i^{(j)},\ j = 1, ..., k-1,\ i = 1, ..., n_j$, be of dimension $d_i^{(j)} \geq 1$, where $n = \sum_{i=1}^{n_1} d_i^{(1)}$ and $n_{j-1} = \sum_{i=1}^{n_j} d_i^{(j)}$ for $j = 2, ..., k-1$. The index i runs over the n_j clusters of each level j. The Kendall distribution function corresponding to $C_i^{(j)},\ j = 1, ..., k-1,\ i = 1, ..., n_j$, is denoted by $K_i^{(j)}$, that is, $K_i^{(j)}(\cdot) := K(\cdot; C_i^{(j)})$. We define the cumulative sum $m_i^{(j)} = \sum_{\ell=1}^{i} d_\ell^{(j)}$ for $i = 1, ..., n_j$, and $m_0^{(j)} = 0$. Under independence assumptions as in Definition 3.3, the random vector $\boldsymbol{U} = (U_1, ..., U_n)'$ is said to be distributed according to the *k-level hierarchical Kendall copula* $C_{\mathcal{K}}$ with nested cluster copulas $C_i^{(j)},\ j = 1, ..., k-1,\ i = 1, ..., n_j$, and nesting copula C_0 if

(i) $\boldsymbol{U}_i := (U_{m_{i-1}^{(1)}+1}, ..., U_{m_i^{(1)}})' \sim C_i^{(1)}\ \forall i \in \{1, ..., n_1\}$,

(ii) $V_i^{(1)} := K_i^{(1)}(C_i^{(1)}(\boldsymbol{U}_i))\ \forall i \in \{1, ..., n_1\}$,

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

  a) $\boldsymbol{V}_i^{(j-1)} := (V_{m_{i-1}^{(j)}+1}^{(j-1)}, ..., V_{m_i^{(j)}}^{(j-1)})' \sim C_i^{(j)}\ \forall i \in \{1, ..., n_j\}$,

  b) $V_i^{(j)} := K_i^{(j)}(C_i^{(j)}(\boldsymbol{V}_i^{(j-1)}))\ \forall i \in \{1, ..., n_j\}$,

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

In particular, the clusters $\boldsymbol{U}_1, ..., \boldsymbol{U}_{n_1}$ at the lowest level ($j = 1$) are assumed to be mutually independent given the "representatives" $V_i^{(j)},\ j = 1, ..., k-1,\ i = 1, ..., n_j$.

An example of a three-level hierarchical Kendall copula is shown in Figure 3.4. In this case, we define

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

Then the assumptions are

$\mathcal{A}_1^{(1)}$: $\boldsymbol{U}_1^{(1)}, .., \boldsymbol{U}_{n_2}^{(1)}$ are mutually independent conditionally on $(V_1^{(2)}, ..., V_{n_2}^{(2)})'$;

$\mathcal{A}_2^{(1)}$: the conditional distribution of $\boldsymbol{U}_i^{(1)} | (V_1^{(2)}, ..., V_{n_2}^{(2)})'$ is the same as the conditional distribution of $\boldsymbol{U}_i^{(1)} | V_i^{(2)}$ for all $i = 1, ..., n_2$;

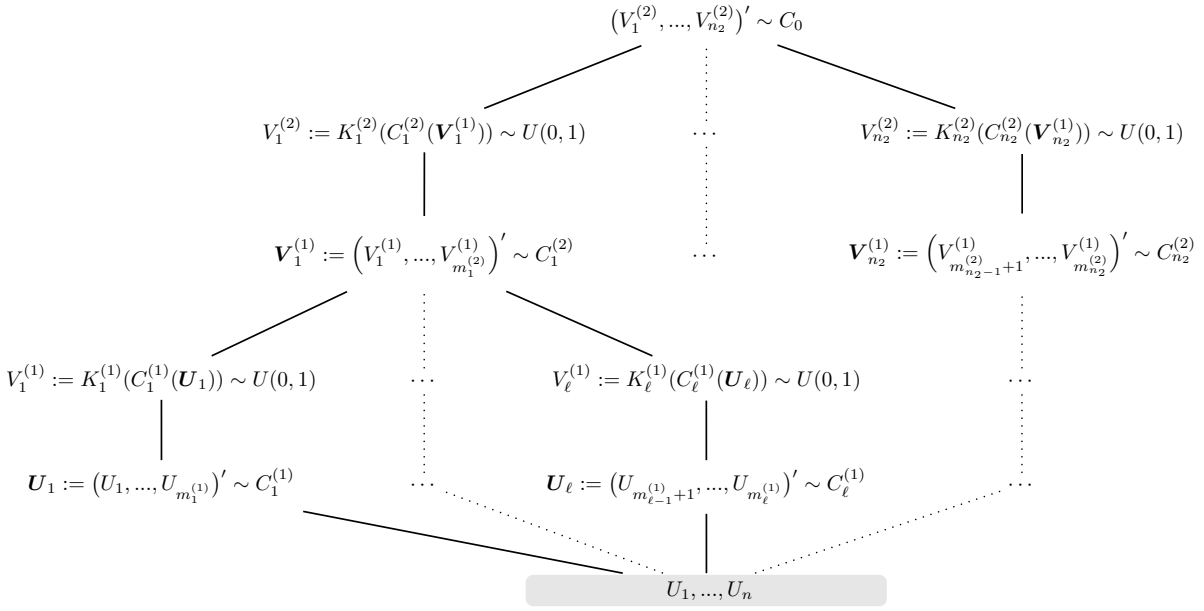


Figure 3.4: Illustration of a three-level hierarchical Kendall copula (see Remark 3.7) with $m_1^{(2)} = \ell$.

$\mathcal{A}_1^{(2)}$: $\boldsymbol{U}_{m_{i-1}^{(2)}+1},...,\boldsymbol{U}_{m_1^{(2)}}$ are mutually independent conditionally on $(V_i^{(2)}, \boldsymbol{V}_i^{(1)\prime})'$ for all $i = 1,...,n_2$;

$\mathcal{A}_2^{(2)}$: the conditional distribution of $\boldsymbol{U}_{m_{i-1}^{(2)}+j}|(V_i^{(2)}, \boldsymbol{V}_i^{(1)\prime})'$ is the same as the conditional distribution of $\boldsymbol{U}_{m_{i-1}^{(2)}+j}|(V_i^{(2)}, V_{m_{i-1}^{(2)}+j}^{(1)})'$ for all $i = 1,...,n_2,\ j = 1,...,d_i^{(2)}$.

Their interpretation is essentially the same as for the assumptions $\mathcal{A}_1$ and $\mathcal{A}_2$ stated in Definition 3.3. Such a three-level model will be considered in Chapter 5.

Also note that a k-level hierarchical Kendall copula could be used to construct higher-dimensional dependence models solely in terms of bivariate copulas. □

For simplicity and illustrative reasons, we mainly 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 developed in the following, to the general k-level case.

The two independence assumptions $\mathcal{A}_1$ and $\mathcal{A}_2$ of Definition 3.3 provide a natural structure for hierarchical dependence models and allow to derive the density function of a hierarchical Kendall copula as stated in the following theorem. Densities of the copulas $C_0,...,C_{n_1}$ are denoted by $c_0,...,c_{n_1}$, respectively.

Theorem 3.8 (Density of a hierarchical Kendall copula)**.** Let $\boldsymbol{U}$ be distributed according to a hierarchical Kendall copula $C_\mathcal{K}$ with cluster copulas $C_1,...,C_{n_1}$ and nesting copula C_0. The density function $c_\mathcal{K}$ of $C_\mathcal{K}$ is then given as follows:

$$c_\mathcal{K}(\boldsymbol{u}) = c_0(K_1(C_1(\boldsymbol{u}_1)),...,K_{n_1}(C_{n_1}(\boldsymbol{u}_{n_1}))) \prod_{i=1}^{n_1} c_i(\boldsymbol{u}_i), \quad \boldsymbol{u} \in [0,1]^n, \tag{3.2}$$

where $\boldsymbol{u}_i := (u_{m_{i-1}+1},...,u_{m_i})',\ i = 1,...,n_1$.

Proof: It holds that

$$\begin{aligned} C_{\mathcal{K}}(\boldsymbol{u}) &= P(\boldsymbol{U} \leq \boldsymbol{u}) \\ &= \int_{[0,1]^{n_1}} P(\boldsymbol{U} \leq \boldsymbol{u} | V_1 = v_1, ..., V_{n_1} = v_{n_1})\, c_0(v_1, ..., v_{n_1})\, dv_1 ... dv_{n_1}. \end{aligned} \tag{3.3}$$

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

$$\begin{aligned} P(\boldsymbol{U} \leq \boldsymbol{u} | V_1 = v_1, ..., V_{n_1} = v_{n_1}) &= \prod_{i=1}^{n_1} P(\boldsymbol{U}_i \leq \boldsymbol{u}_i | V_1 = v_1, ..., V_{n_1} = v_{n_1}) \\ &= \prod_{i=1}^{n_1} P(\boldsymbol{U}_i \leq \boldsymbol{u}_i | V_i = v_i) \\ &= \prod_{i=1}^{n_1} F_{\boldsymbol{U}_i | V_i}(\boldsymbol{u}_i | v_i). \end{aligned}$$

Using this result, Equation (3.3) simplifies to

$$C_{\mathcal{K}}(\boldsymbol{u}) = \int_{[0,1]^{n_1}} \left(\prod_{i=1}^{n_1} F_{\boldsymbol{U}_i | V_i}(\boldsymbol{u}_i | v_i) \right) c_0(v_1, ..., v_{n_1})\, dv_1 ... dv_{n_1}. \tag{3.4}$$

Further, we denote by $\boldsymbol{U}_{i,-m_i}$, $i \in \{1, ..., n_1\}$, the $(d_i - 1)$-dimensional sub-vector of $\boldsymbol{U}_i = (U_{m_{i-1}+1}, ..., U_{m_i})'$ with element U_{m_i} removed, that is, $\boldsymbol{U}_{i,-m_i} = (U_{m_{i-1}+1}, ..., U_{m_i-1})'$. Since $V_i = K_i(C_i(\boldsymbol{U}_i)) \sim U(0,1)$ by definition and according to the change of variables $\boldsymbol{U}_i \mapsto (\boldsymbol{U}'_{i,-m_i}, K_i(C_i(\boldsymbol{U}_i)))'$, it then holds that

$$\begin{aligned} f_{\boldsymbol{U}_{i,-m_i} | V_i}(\boldsymbol{u}_{i,-m_i} | v_i) &= f_{\boldsymbol{U}_{i,-m_i}, V_i}(\boldsymbol{u}_{i,-m_i}, v_i) \\ &= c_i(u_{m_{i-1}+1}, ..., u_{m_i-1}, C_i^{-1}(K_i^{-1}(v_i) | \boldsymbol{u}_{i,-m_i})) \\ &\quad \times \frac{\partial}{\partial v_i} C_i^{-1}(K_i^{-1}(v_i) | \boldsymbol{u}_{i,-m_i}), \end{aligned} \tag{3.5}$$

if $v_i \leq K_i(C_i(u_{m_{i-1}+1}, ..., u_{m_i-1}, 1))$. This yields

$$F_{\boldsymbol{U}_i | V_i}(\boldsymbol{u}_i | v_i) = \int_0^{u_{m_i-1}} ... \int_0^{u_{m_{i-1}+1}} f_{\boldsymbol{U}_{i,-m_i} | V_i}(\boldsymbol{w}_{i,-m_i} | v_i)\, 1_{\{C_i^{-1}(K_i^{-1}(v_i) | \boldsymbol{w}_{i,-m_i}) \leq u_{m_i}\}}\, d\boldsymbol{w}_{i,-m_i}.$$

Plugging this expression for the conditional distribution function of $\boldsymbol{U}_i$ given $V_i = v_i$ into Equation (3.4) and substituting v_i by $w_{m_i} = C_i^{-1}(K_i^{-1}(v_i) | \boldsymbol{w}_{i,-m_i})$ for $i = 1, ..., n_1$ then leads to

$$C_{\mathcal{K}}(\boldsymbol{u}) = \int_0^{u_n} ... \int_0^{u_1} \left(\prod_{i=1}^{n_1} c_i(\boldsymbol{w}_i) \right) c_0(K_1(C_1(\boldsymbol{w}_1)), ..., K_d(C_d(\boldsymbol{w}_d)))\, d\boldsymbol{w},$$

where we used that $v_i = K_i(C_i(\boldsymbol{w}_i))$, $i = 1, ..., n_1$. Taking derivatives with respect to $u_1, ..., u_n$ therefore gives the desired result. □

As the following remark shows, the proof of Theorem 3.8 can also be stated directly in terms of densities.

Remark 3.9 (Alternative proof of Theorem 3.8)**.** Using the notation from above, it holds according to assumptions $\mathcal{A}_1$ and $\mathcal{A}_2$ that

$$\begin{aligned} &f_{\boldsymbol{U}_{1,-m_1},...,\boldsymbol{U}_{n_1,-m_{n_1}}|V_1,...,V_{n_1}}(\boldsymbol{u}_{1,-m_1},...,\boldsymbol{u}_{n_1,-m_{n_1}}|v_1,...,v_{n_1}) \\ &\quad= \prod_{i=1}^{n_1} f_{\boldsymbol{U}_{i,-m_i}|V_1,...,V_{n_1}}(\boldsymbol{u}_{i,-m_i}|v_1,...,v_{n_1}) \\ &\quad= \prod_{i=1}^{n_1} f_{\boldsymbol{U}_{i,-m_i}|V_i}(\boldsymbol{u}_{i,-m_i}|v_i), \end{aligned} \tag{3.6}$$

where the expression for $f_{\boldsymbol{U}_{i,-m_i}|V_i},\ i = 1,...,n_1$, is derived in Equation (3.5) using the change of variables $\boldsymbol{U}_i \mapsto (\boldsymbol{U}'_{i,-m_i}, K_i(C_i(\boldsymbol{U}_i)))'$. Using the inverse of this change of variables, that is, $(\boldsymbol{U}'_{i,-m_i}, V_i)' \mapsto (\boldsymbol{U}'_{i,-m_i}, C_i^{-1}(K_i^{-1}(V_i)|\boldsymbol{U}_{i,-m_i}))'$, then yields the following expression of the density $c_{\mathcal{K}}$ of $\boldsymbol{U}$:

$$\begin{aligned} c_{\mathcal{K}}(\boldsymbol{u}) &= f_{\boldsymbol{U}_1,...,\boldsymbol{U}_{n_1}}(\boldsymbol{u}_1,...,\boldsymbol{u}_{n_1}) \\ &= f_{\boldsymbol{U}_{1,-m_1},V_1,...,\boldsymbol{U}_{n_1,-m_{n_1}},V_{n_1}}(\boldsymbol{u}_{1,-m_1}, K_1(C_1(\boldsymbol{u}_1)),...,\boldsymbol{u}_{n_1,-m_{n_1}}, K_{n_1}(C_{n_1}(\boldsymbol{u}_{n_1}))) \\ &\quad\times \prod_{i=1}^{n_1} \frac{\partial}{\partial u_{m_i}} K_i(C_i(\boldsymbol{u}_i)) \\ &= c_0(K_1(C_1(\boldsymbol{u}_1)),...,K_{n_1}(C_{n_1}(\boldsymbol{u}_{n_1}))) \\ &\quad\times f_{\boldsymbol{U}_{1,-m_1},...,\boldsymbol{U}_{n_1,-m_{n_1}}|V_1,...,V_{n_1}}(\boldsymbol{u}_{1,-m_1},...,\boldsymbol{u}_{n_1,-m_{n_1}}|K_1(C_1(\boldsymbol{u}_1)),...,K_{n_1}(C_{n_1}(\boldsymbol{u}_{n_1}))) \\ &\quad\times \prod_{i=1}^{n_1} \frac{\partial}{\partial u_{m_i}} K_i(C_i(\boldsymbol{u}_i)) \\ &\overset{(3.6)}{=} c_0(K_1(C_1(\boldsymbol{u}_1)),...,K_{n_1}(C_{n_1}(\boldsymbol{u}_{n_1}))) \\ &\quad\times \prod_{i=1}^{n_1} f_{\boldsymbol{U}_{i,-m_i}|V_i}(\boldsymbol{u}_{i,-m_i}|K_i(C_i(\boldsymbol{u}_i))) \frac{\partial}{\partial u_{m_i}} K_i(C_i(\boldsymbol{u}_i)) \\ &\overset{(3.5)}{=} c_0(K_1(C_1(\boldsymbol{u}_1)),...,K_{n_1}(C_{n_1}(\boldsymbol{u}_{n_1}))) \\ &\quad\times \prod_{i=1}^{n_1} c_i(\boldsymbol{u}_i) \frac{\partial}{\partial v_i} C_i^{-1}(K_i^{-1}(v_i)|\boldsymbol{u}_{i,-m_i})\Big|_{v_i=K_i(C_i(\boldsymbol{u}_i))} \frac{\partial}{\partial u_{m_i}} K_i(C_i(\boldsymbol{u}_i)). \end{aligned} \tag{3.7}$$

In the next step, we explicitly calculate the two last terms. For the very last term, we immediately get

$$\frac{\partial}{\partial u_{m_i}} K_i(C_i(\boldsymbol{u}_i)) = F_{\boldsymbol{U}_{i,-m_i}|U_{m_i}}(\boldsymbol{u}_{i,-m_i}|u_{m_i})\, K_i'(C_i(\boldsymbol{u}_i))$$

Further, we have

$$\frac{\partial}{\partial v_i} C_i^{-1}(K_i^{-1}(v_i)|\boldsymbol{u}_{i,-m_i}) = \left(F_{\boldsymbol{U}_{i,-m_i}|U_{m_i}}(\boldsymbol{u}_{i,-m_i}|C_i^{-1}(K_i^{-1}(v_i)|\boldsymbol{u}_{i,-m_i}))\, K_i'(K_i^{-1}(v_i))\right)^{-1},$$

so that

$$\frac{\partial}{\partial v_i} C_i^{-1}(K_i^{-1}(v_i)|\boldsymbol{u}_{i,-m_i})\Big|_{v_i=K_i(C_i(\boldsymbol{u}_i))} = \left(F_{\boldsymbol{U}_{i,-m_i}|U_{m_i}}(\boldsymbol{u}_{i,-m_i}|u_{m_i})\, K_i'(C(\boldsymbol{u}_i))\right)^{-1}.$$

As a result, the two derivatives in Equation (3.7) cancel each other out and we obtain that

$$c_{\mathcal{K}}(\boldsymbol{u}) = c_0(K_1(C_1(\boldsymbol{u}_1)), ..., K_{n_1}(C_{n_1}(\boldsymbol{u}_{n_1}))) \prod_{i=1}^{n_1} c_i(\boldsymbol{u}_i),$$

which proves Theorem 3.8. □

The density of a k-level hierarchical Kendall copula (see Remark 3.7) can be derived similarly.

Remark 3.10 (Density of a k-level hierarchical Kendall copula)**.** The arguments of Theorem 3.8 can be iterated to derive the density of a k-level hierarchical Kendall copula. By first conditioning on the aggregated variables of level $k-1$, $V_1^{(k-1)}, ..., V_{n_{k-1}}^{(k-1)}$, then on those of level $k-2$ and so on up to level 1, an expression similar to Equation (3.2) is obtained.

For instance, the density of the three-level hierarchical Kendall copula can be derived along the lines of the proof of Theorem 3.8 as

$$\begin{aligned} c_{\mathcal{K}}(\boldsymbol{u}) &= c_0(v_1^{(2)}, ..., v_{n_2}^{(2)}) \prod_{i=1}^{n_2} \left(c_i^{(2)}\left(\boldsymbol{v}_i^{(1)}\right) \prod_{j=1}^{d_i^{(2)}} c^{(1)}_{m_{i-1}^{(2)}+j}\left(\boldsymbol{u}_{m_{i-1}^{(2)}+j}\right)\right) \\ &= c_0(v_1^{(2)}, ..., v_{n_2}^{(2)}) \prod_{i=1}^{n_2} c_i^{(2)}\left(\boldsymbol{v}_i^{(1)}\right) \prod_{j=1}^{n_1} c_j^{(1)}\left(\boldsymbol{u}_j\right), \end{aligned} \tag{3.8}$$

where $\boldsymbol{v}_i^{(1)} = (v^{(1)}_{m_{i-1}^{(2)}+1}, ..., v^{(1)}_{m_i^{(2)}})'$, $i = 1, ..., n_2$, with components $v_i^{(1)} = K_i^{(1)}(C_i^{(1)}(\boldsymbol{u}_i))$, $i = 1, ..., n_1$. Further, $v_i^{(2)} = K_i^{(2)}(C_i^{(2)}(\boldsymbol{v}_i^{(1)}))$, $i = 1, ..., n_2$.

This means that the density of a three-level hierarchical Kendall copula also conveniently decomposes into the product of the densities of the (nested) cluster copulas and of the nesting copula, 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 3.8.

Example 3.11 (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 (see Example 2.2), it follows that $c_{\mathcal{K}}(\boldsymbol{u}) = 1$. This means that the independence copula also belongs to the class of hierarchical Kendall copulas. □

Theorem 3.8 also allows to formulate the following corollary, which summarizes the marginal properties of hierarchical Kendall copulas.

Corollary 3.12 (Margins of a hierarchical Kendall copula)**.** The same notation as in Theorem 3.8 is used.

(i) Bivariate margins: Let $k, \ell \in \{1, ..., n\},\ k \neq \ell$. Without loss of generality let $k < \ell$.

a) If U_k and U_ℓ are in the same cluster i, their marginal distribution function $C_{\mathcal{K},k\ell}$ is the bivariate (k,ℓ)-margin of C_i denoted by $C_{i,k\ell}$, that is,

$$C_{\mathcal{K},k\ell}(u_k, u_\ell) = C_{i,k\ell}(u_k, u_\ell) := C_i(1, ..., 1, u_k, 1, ..., 1, u_\ell, 1, ..., 1), \tag{3.9}$$

where $(u_k, u_\ell)' \in [0,1]^2$.

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

$$C_{\mathcal{K},k\ell}(u_k, u_\ell) = \int_0^{u_k}\int_0^{u_\ell}\int_{[0,1]^{d_i+d_j-2}} c_{0,ij}(K_i(C_i(\boldsymbol{w}_i)), K_j(C_j(\boldsymbol{w}_j))) \times c_i(\boldsymbol{w}_i)\, c_j(\boldsymbol{w}_j)\, d\boldsymbol{w}_{i,-k}\, d\boldsymbol{w}_{j,-\ell}\, dw_\ell\, dw_k, \tag{3.10}$$

where $(u_k, u_\ell)' \in [0,1]^2$ and $c_{0,ij}$ is the density of the (i,j)-margin of C_0.

(ii) Multivariate margins: The marginal distribution function of the cluster $\boldsymbol{U}_i$ is C_i.

Proof: Statements (i) a) and (ii) directly follow from Definition 3.3, statement (i) b) from Theorem 3.8. □

More general multivariate margins involving variables from different clusters can be derived as in Equation (3.10).

Remark 3.13 (Mixture representation)**.** As a consequence of Corollary 3.12 (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 for $(u_k, u_\ell)' \in [0,1]^2$ by

$$c_{\mathcal{K},k\ell}(u_k, u_\ell) = \int_{[0,1]^{d_i+d_j-2}} c_{0,ij}(K_i(C_i(\boldsymbol{u}_i)), K_j(C_j(\boldsymbol{u}_j)))\, c_i(\boldsymbol{u}_i)\, c_j(\boldsymbol{u}_j)\, d\boldsymbol{u}_{i,-k}\, d\boldsymbol{u}_{j,-\ell}, \tag{3.11}$$

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

This representation complements the results of Remark 3.5. Equation (3.1) shows that hierarchical Kendall copulas with Archimedean cluster copulas can be represented as transformed mixtures of uniform distributions on unit simplices. □

Clearly, the copula $C_{\mathcal{K},k\ell}$ of two variables in different clusters is not available explicitly (see Equations (3.10) and (3.11)). To better understand the between-cluster dependence, the following theorem provides a characterization result in terms of TP_2 dependence (see Definition 2.8).

Theorem 3.14 (TP_2 dependence of hierarchical Kendall copulas)**.** The copula $C_{\mathcal{K},k\ell}$ of $(U_k, U_\ell)'$, where U_k and U_ℓ are in different clusters i and j, respectively, has a TP_2 density if the bivariate (i,j)-margin of C_0 denoted by $C_{0,ij}$ has a TP_2 density.

Proof: For all $(u_k, u_\ell)', (w_k, w_\ell)' \in [0,1]^2$ with $w_k > u_k$ and $w_\ell > u_\ell$, we have to show that

$$c_{\mathcal{K},k\ell}(u_k, u_\ell)\, c_{\mathcal{K},k\ell}(w_k, w_\ell) \geq c_{\mathcal{K},k\ell}(u_k, w_\ell)\, c_{\mathcal{K},k\ell}(w_k, u_\ell). \tag{3.12}$$

Since C_i and C_j are strictly monotone, it holds that $K_i(C_i(\boldsymbol{w}_i)) > K_i(C_i(\boldsymbol{u}_i))$ and $K_j(C_j(\boldsymbol{w}_j)) > K_j(C_j(\boldsymbol{u}_j))$. Therefore, the TP_2 property of $c_{0,ij}$ yields

$$\begin{aligned} &c_{0,ij}(K_i(C_i(\boldsymbol{u}_i)), K_j(C_j(\boldsymbol{u}_j)))\, c_{0,ij}(K_i(C_i(\boldsymbol{w}_i)), K_j(C_j(\boldsymbol{w}_j))) \\ &\quad \geq c_{0,ij}(K_i(C_i(\boldsymbol{u}_i)), K_j(C_j(\boldsymbol{w}_j)))\, c_{0,ij}(K_i(C_i(\boldsymbol{w}_i)), K_j(C_j(\boldsymbol{u}_j))). \end{aligned} \tag{3.13}$$

According to Equations (3.11) and (3.13), we then have

$$\begin{aligned} &c_{\mathcal{K},k\ell}(u_k, u_\ell)\, c_{\mathcal{K},k\ell}(w_k, w_\ell) \\ &= \int_{[0,1]^{d_i+d_j-2}} \int_{[0,1]^{d_i+d_j-2}} c_{0,ij}(K_i(C_i(\boldsymbol{u}_i)), K_j(C_j(\boldsymbol{u}_j)))\, c_{0,ij}(K_i(C_i(\boldsymbol{w}_i)), K_j(C_j(\boldsymbol{w}_j))) \\ &\qquad \times c_i(\boldsymbol{u}_i)\, c_j(\boldsymbol{u}_j)\, c_i(\boldsymbol{w}_i)\, c_j(\boldsymbol{w}_j)\, d\boldsymbol{u}_{i,-k}\, d\boldsymbol{u}_{j,-\ell}\, d\boldsymbol{w}_{i,-k}\, d\boldsymbol{w}_{j,-\ell} \\ &\geq \int_{[0,1]^{d_i+d_j-2}} \int_{[0,1]^{d_i+d_j-2}} c_{0,ij}(K_i(C_i(\boldsymbol{u}_i)), K_j(C_j(\boldsymbol{w}_j)))\, c_{0,ij}(K_i(C_i(\boldsymbol{w}_i)), K_j(C_j(\boldsymbol{u}_j))) \\ &\qquad \times c_i(\boldsymbol{u}_i)\, c_j(\boldsymbol{w}_j)\, c_i(\boldsymbol{w}_i)\, c_j(\boldsymbol{u}_j)\, d\boldsymbol{u}_{i,-k}\, d\boldsymbol{w}_{j,-\ell}\, d\boldsymbol{w}_{i,-k}\, d\boldsymbol{u}_{j,-\ell} \\ &= c_{\mathcal{K},k\ell}(u_k, w_\ell)\, c_{\mathcal{K},k\ell}(w_k, u_\ell), \end{aligned}$$

which proves the statement (3.12). □

Since the Frank copula with positive parameter has a TP_2 density, this means that $c_{\mathcal{K},13}$, $c_{\mathcal{K},14}$, $c_{\mathcal{K},23}$ and $c_{\mathcal{K},24}$ in Example 3.6 are TP_2.

As noted above, the multivariate distribution of $\boldsymbol{U} = (U_1, ..., U_n)'$ defined through a hierarchical Kendall copula is in general not the copula C_0 but given through Equation (3.3). We showed that the important special cases of independence as well as of comonotonicity are hierarchical Kendall copulas (see Examples 3.4 and 3.11), 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 (that is, there is at least one cluster $i \in \{1, ..., n_1\}$ with more than one element ($d_i > 1$) or, in others words, the number of clusters n_1 is smaller than the number of variables n). For example, when choosing cluster and nesting copulas as Gaussian, then it is clear from the density expression (3.2) that the resulting hierarchical Kendall copula will not be multivariate Gaussian (unless the nesting copula is the independence copula). Similarly, hierarchical Archimedean copulas are different from hierarchical Kendall copulas—with positive and negative implications as discussed next.

3.3.1 Comparison with hierarchical Archimedean copulas

The class of hierarchical (or nested) Archimedean copulas, which extends standard Archimedean copulas (2.31) to non-exchangeability, also allows for a nested modeling of clusters of variables. Borrowing the notation of Definition 3.3, let C_0 be an n_1-dimensional Archimedean copula with generator φ_0 and $C_1, ..., C_{n_1}$ Archimedean copulas of dimension $d_i \geq 1$ and with generator φ_i for $i = 1, ..., n_1$. Then the two-level *partially nested Archimedean copula* is defined as

$$C(\boldsymbol{u}; \varphi_0, \varphi_1, ..., \varphi_{n_1}) = C_0(C_1(\boldsymbol{u_1}), ..., C_{n_1}(\boldsymbol{u}_{n_1})), \quad \boldsymbol{u} \in [0,1]^n. \tag{3.14}$$

Together with the so-called fully nested Archimedean copulas, which are not considered here, this copula forms the class of hierarchical Archimedean copulas (see, e.g., Joe (1993, 1997), McNeil (2008), Hofert (2010), Savu and Trede (2010) and Okhrin et al. (2013)). The extension to k hierarchical levels is straightforward (see Remark 3.7).

The hierarchical Archimedean copula defined in Equation (3.14) can be written in terms of the generators $\varphi_0, ..., \varphi_{n_1}$ as

$$\begin{aligned} &C(\boldsymbol{u}; \varphi_0, \varphi_1, ..., \varphi_{n_1}) \\ &\quad = \varphi_0^{-1}\left(\varphi_0\left(\varphi_1^{-1}\left(\textstyle\sum_{j=1}^{d_1} \varphi_1(u_j)\right)\right) + ... + \varphi_0\left(\varphi_{n_1}^{-1}\left(\textstyle\sum_{j=1}^{d_{n_1}} \varphi_{n_1}(u_{m_{n_1-1}+j})\right)\right)\right). \end{aligned}$$

According to Joe (1993, 1997) and McNeil (2008), a sufficient condition for $C(\cdot; \varphi_0, ..., \varphi_{n_1})$ to be a copula is that the derivative of $\varphi_0 \circ \varphi_i^{-1}$ is completely monotone for all $i = 1, ..., n_1$. If all generators are of the same type, this typically translates to parameter restrictions. For the four copulas presented in Section 2.4 (Clayton, Gumbel, Frank with positive parameter, and Joe), a sufficient condition is that $\theta_0 \leq \min\{\theta_1, ..., \theta_{n_1}\}$, where θ_i is the parameter of the generator φ_i, $i = 0, ..., n_1$. This means that hierarchical Archimedean copulas require stronger within-cluster than between-cluster dependence (see also Joe (1997, Corollary 4.2)). This is not the case for hierarchical Kendall copulas (see Example 3.6). Furthermore, hierarchical Archimedean copulas are, of course, limited to Archimedean copulas as building blocks, while hierarchical Kendall copulas can be built up by any possible copula.

Nevertheless, 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 in McNeil (2008) and Hofert (2010, 2011). Hierarchical Kendall copulas also benefit from many of these properties. It will be shown in Chapter 4 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 density function, which is numerically tractable even in higher dimensions (see Equation (3.8) for the case of three hierarchical levels) and also for other building blocks than Archimedean copulas. The density expression of hierarchical Archimedean copulas is however hardly accessible in general (see Savu and Trede (2010)). This is illustrated in the following example.

Example 3.15 (Density of a hierarchical Archimedean copula)**.** Let $C(\cdot;\varphi_0,\varphi_1,\varphi_2)$ be a four-dimensional hierarchical Archimedean copula with $d_1 = d_2 = 2$. As in Savu and Trede (2010), we derive the corresponding density using the chain rule as

$$
\begin{aligned}
&c(\boldsymbol{u};\varphi_0,\varphi_1,\varphi_2)\\
&\quad= \left.\frac{\partial^2 c_0(v_1,v_2)}{\partial v_1 \partial v_2}\right|_{v_1=C_1(u_1,u_2),\, v_2=C_2(u_3,u_4)} C_{1,2|1}(u_2|u_1)\, C_{1,1|2}(u_1|u_2)\, C_{2,4|3}(u_4|u_3)\, C_{2,3|4}(u_3|u_4)\\
&\qquad+ \left.\frac{\partial c_0(v_1,v_2)}{\partial v_1}\right|_{v_1=C_1(u_1,u_2),\, v_2=C_2(u_3,u_4)} C_{1,2|1}(u_2|u_1)\, C_{1,1|2}(u_1|u_2)\, c_2(u_3,u_4)\\
&\qquad+ \left.\frac{\partial c_0(v_1,v_2)}{\partial v_2}\right|_{v_1=C_1(u_1,u_2),\, v_2=C_2(u_3,u_4)} c_1(u_1,u_2)\, C_{2,4|3}(u_4|u_3)\, C_{2,3|4}(u_3|u_4)\\
&\qquad+ c_0(C_1(u_1,u_2),C_2(u_3,u_4))\, c_1(u_1,u_2)\, c_2(u_3,u_4),
\end{aligned}
$$

where $\boldsymbol{u} \in [0,1]^4$ and $C_{i,k|\ell},\ i \in \{1,2\},\ k,\ell \in \{1,...,4\}$, is the conditional distribution function of $U_k|U_\ell$ (see Equations (2.13) and (2.14)). While the last term resembles the density of a corresponding hierarchical Kendall copula, the other three terms complicate density evaluations. Clearly, additional terms occur in higher dimensions. For this reason, Hofert and Pham (2013) recently proposed an alternative approach using Laplace transforms, which may yield a more tractable density expression at least for the case of a moderate number of nesting levels. □

Multivariate margins of hierarchical Kendall copulas are not directly available (see Corollary 3.12). This is different for hierarchical Archimedean copulas, as it can be directly inferred from Equation (3.14). If variables U_k and U_ℓ are in the same cluster i, then $(U_k,U_\ell)' \sim C_{i,k\ell}$, which is similar to the case of hierarchical Kendall copulas (see Equation (3.9)). On the other hand, if U_k and U_ℓ are in different clusters i and j, respectively, then $(U_k,U_\ell)' \sim C_{0,ij}$, while in the case of a hierarchical Kendall copula this marginal distribution has to be obtained using integration as in Equation (3.10).

Example 3.16 (Bivariate margin of hierarchical Archimedean and Kendall copulas)**.** Let $U_1,...,U_4$ be distributed according to a hierarchical Kendall copula or hierarchical Archimedean copula with bivariate Gumbel cluster and nesting copulas ($d_1 = d_2 = 2$): $(U_1,U_2)' \sim C_1$, $(U_3,U_4)' \sim C_2$ and C_0 denotes the nesting copula. Figure 3.5 illustrates the marginal density of $(U_1,U_3)'$ in both cases. It shows contour lines of the marginal density of the pair $(\Phi^{-1}(U_1),\Phi^{-1}(U_3))'$, which has standard normal margins according to the inverse probability integral transform. 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. Apparently, the difference between the distributions is minor. The contour lines corresponding to the hierarchical Archimedean copula are 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 not needed for the hierarchical Kendall copula.

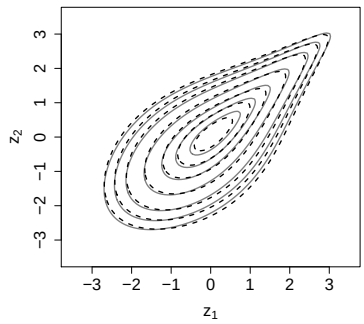


Figure 3.5: Contour lines of the marginal density of $(\Phi^{-1}(U_1), \Phi^{-1}(U_3))'$, where $U_1, ..., U_4$ are distributed according to a hierarchical Kendall copula (black dashed line) or hierarchical Archimedean copula (gray line) with Gumbel cluster and nesting copulas.

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 Theorem 2.10 and Corollary 2.11). That is, if a random variable U_{n+1} is added to cluster $i \in \{1, ..., d\}$, the transformation K_i changes and thus V_i also does so, even if C_i is Archimedean; similarly if a random variable is removed from a cluster.

3.4 Sampling

We now develop inference techniques for hierarchical Kendall copulas. First, simulation is treated, then estimation and model selection. The following general simulation procedure describes how to sample from a given hierarchical Kendall copula.

Algorithm 3.17 (Simulation of hierarchical Kendall copulas)**.** Let $C_{\mathcal{K}}$ be a hierarchical Kendall copula with cluster copulas $C_1, ..., C_{n_1}$ and nesting copula C_0.

(i) Sample $(v_1, ..., v_{n_1})'$ from C_0.

(ii) Set $z_i = K_i^{-1}(v_i)$ for $i = 1, ..., n_1$.

(iii) Sample $\boldsymbol{u}_i$ from $\boldsymbol{U}_i | C_i(\boldsymbol{U}_i) = z_i$ for $i = 1, ..., n_1$.

(iv) Return $\boldsymbol{u} = (\boldsymbol{u}_1', ..., \boldsymbol{u}_{n_1}')'$.

The algorithm proceeds by first sampling the level z_i of the level set of each cluster (top of Figure 3.2) and then sampling from each cluster given this level set $L(z_i; C_i)$ (bottom of Figure 3.2). The procedure is therefore also referred to as top-down sampling. It can easily be generalized to the case of k hierarchical levels.

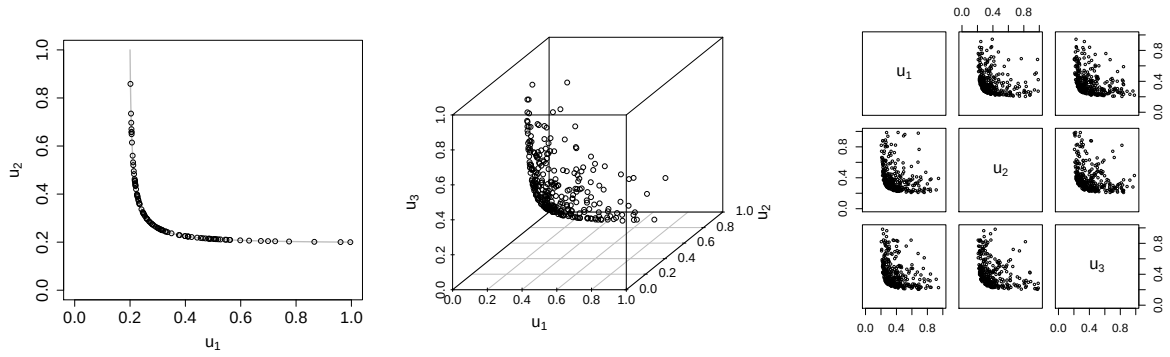


Figure 3.6: Left panel: scatter plot of a sample from a bivariate Clayton copula with parameter $\theta = 2$ (Kendall's τ of 0.5) on the level set at $z = 0.2$. The corresponding contour line is shown in gray. Middle and right panel: three-dimensional scatter plot and pairwise scatter plots of a sample from a trivariate Clayton copula with parameter $\theta = 2$ on the level set at $z = 0.2$.

Remark 3.18 (Simulation of k-level hierarchical Kendall copulas)**.** Algorithm 3.17 for two levels can be iterated to simulate from a k-level hierarchical Kendall copula (see Remark 3.7).

(i) Sample $(v_1^{(k-1)}, ..., v_{n_{k-1}}^{(k-1)})'$ from C_0.

(ii) For $j = k-1, ..., 2$:

  a) Set $z_i^{(j)} = (K_i^{(j)})^{-1}(v_i^{(j)})$ for $i = 1, ..., n_j$.

  b) Sample $\boldsymbol{v}_i^{(j-1)}$ from $\boldsymbol{V}_i^{(j-1)} | C_i^{(j)}(\boldsymbol{V}_i^{(j-1)}) = z_i^{(j)}$ for $i = 1, ..., n_j$.

(iii) Set $z_i^{(1)} = (K_i^{(1)})^{-1}(v_i^{(1)})$ for $i = 1, ..., n_1$.

(iv) Sample $\boldsymbol{u}_i$ from $\boldsymbol{U_i} | C_i^{(1)}(\boldsymbol{U}_i) = z_i^{(1)}$ for $i = 1, ..., n_1$.

(v) Return $\boldsymbol{u} = (\boldsymbol{u}_1', ..., \boldsymbol{u}_{n_1}')'$.

Hence, the algorithm essentially proceeds from the top to the bottom of Figure 3.4. □

Given that simulation from the copula C_0 is feasible, sampling from hierarchical Kendall copulas thus amounts to the more general question of sampling from a distribution $\boldsymbol{U} | C(\boldsymbol{U}) = z$, where C is the copula of a marginally uniform random vector $\boldsymbol{U} := (U_1, ..., U_d)'$ and $z \in (0, 1)$. In other words, we want to sample from a multivariate distribution given a specific level set $L(z; C)$ at level z as illustrated in Figure 3.6. This problem is discussed in detail in Chapter 4, where different methods are discussed and closed-form solutions for Archimedean, Archimax and Plackett copulas are derived. Alternatively, different approximate methods are proposed, where also a bottom-up approach is followed in contrast to the top-down algorithm, which is presented here.

3.5 Estimation

In light of Sklar's Theorem (2.1), it is common in dependence modeling to transform data $(x_{k1}, ..., x_{kn})'$, $k = 1, ..., N$, to $[0,1]^n$ using the marginal distribution functions F_j, $j = 1, ..., n$, that is, we compute $u_{kj} = F_j(x_{kj})$. In most cases, F_j is unknown, so that this transformation needs to be based on a parametric or a non-parametric estimate, which introduces uncertainty into the modeling.

Here, we concentrate on the parametric modeling of the margins. In this case, the parameters of the margins and the 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) (see also Joe (2005)). In the IFM method, first the marginal parameters are estimated and then the dependence parameters given the estimated margins $\widehat{F}_j$, $j = 1, ..., n$.

More precisely, let $C_{\mathcal{K}}$ be a hierarchical Kendall copula with cluster and nesting copulas $C_0, C_1, ..., C_{n_1}$ and density function $c_{\mathcal{K}}$ (3.2). Further, we denote the parameter(s) of copula C_i by $\boldsymbol{\theta}_i$ for $i = 0, ..., n_1$. According to the IFM method, we then set $\widehat{u}_{kj} = \widehat{F}_j(x_{kj})$, $k = 1, ..., N$, $j = 1, ..., n$, and estimate the parameters $\boldsymbol{\theta}_0, \boldsymbol{\theta}_1, ..., \boldsymbol{\theta}_{n_1}$ by maximizing the log likelihood expression $\ell_{\mathcal{K}}$, which conveniently decomposes into separate sums (see Theorem 3.8):

$$
\begin{aligned}
&\ell_{\mathcal{K}}(\boldsymbol{\theta}_0, \boldsymbol{\theta}_1, ..., \boldsymbol{\theta}_{n_1}; (\widehat{u}_{k1}, ..., \widehat{u}_{kn})'_{k=1,...,N}) \\
&\quad = \sum_{k=1}^{N} \log c_{\mathcal{K}}(\widehat{u}_{k1}, ..., \widehat{u}_{kn}; \boldsymbol{\theta}_0, \boldsymbol{\theta}_1, ..., \boldsymbol{\theta}_{n_1}) \\
&\quad = \sum_{k=1}^{N} \log c_0(K_1(C_1(\widehat{\boldsymbol{u}}_{k1}; \boldsymbol{\theta}_1); \boldsymbol{\theta}_1), ..., K_{n_1}(C_{n_1}(\widehat{\boldsymbol{u}}_{kn_1}; \boldsymbol{\theta}_{n_1}); \boldsymbol{\theta}_{n_1}); \boldsymbol{\theta}_0) \\
&\qquad + \sum_{i=1}^{n_1} \sum_{k=1}^{N} \log c_i(\widehat{\boldsymbol{u}}_{ki}; \boldsymbol{\theta}_i) \\
&\quad =: \ell_0(\boldsymbol{\theta}_0, \boldsymbol{\theta}_1, ..., \boldsymbol{\theta}_{n_1}; (\widehat{u}_{k1}, ..., \widehat{u}_{kn})'_{k=1,...,N}) + \sum_{i=1}^{n_1} \ell_i(\boldsymbol{\theta}_i; (\widehat{\boldsymbol{u}}_{ki})_{k=1,...,N}),
\end{aligned}
\tag{3.15}
$$

where $\widehat{\boldsymbol{u}}_{ki} := (\widehat{u}_{km_{i-1}+1}, ..., \widehat{u}_{km_i})'$, $k = 1, ..., N$, $i = 1, ..., n_1$.

The asymptotic covariance matrix of resulting parameter estimates is given by the inverse Godambe information matrix, which is unfortunately typically very cumbersome to compute. To see this, observe that the log likelihood $\ell_{\mathcal{K}}$ (3.15) depends on the parameters of the 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 financial applications, as in Section 3.7, the margins are however often 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)).

The hierarchical construction given in Definition 3.3 and the log likelihood expression (3.15) also directly lead to a sequential estimation procedure of hierarchical Kendall copulas, which avoids higher-dimensional maximum likelihood estimation. This is similar to the

sequential estimation of vine copulas, which only requires bivariate maximum likelihood estimation (see Section 2.7.4).

Algorithm 3.19 (Sequential estimation of hierarchical Kendall copulas)**.** Let $C_{\mathcal{K}}$ be a hierarchical Kendall copula with cluster and nesting copulas $C_0, C_1, ..., C_{n_1}$ and denote the parameter(s) of copula C_i by $\boldsymbol{\theta}_i$ for $i = 0, ..., n_1$. Corresponding sequential estimates $\widehat{\boldsymbol{\theta}}_i,\ i = 0, ..., n_1$, are then obtained as follows.

(i) For each cluster $i \in \{1, ..., n_1\}$ estimate $\boldsymbol{\theta}_i$ based on $(\widehat{\boldsymbol{u}}_{ki})_{k=1,...,N}$ by maximum likelihood, that is, maximize $\ell_i(\boldsymbol{\theta}_i; (\widehat{\boldsymbol{u}}_{ki})_{k=1,...,N})$ as defined in Equation (3.15) with respect to $\boldsymbol{\theta}_i$.

(ii) Estimate $\boldsymbol{\theta}_0$ based on the pseudo observations

$$\widehat{v}_{ki} := K_i(C_i(\widehat{\boldsymbol{u}}_{ki}; \widehat{\boldsymbol{\theta}}_i); \widehat{\boldsymbol{\theta}}_i),\ k = 1, ..., N,\ i = 1, ..., n_1, \tag{3.16}$$

by maximum likelihood, that is, maximize $\ell_0(\boldsymbol{\theta}_0, \widehat{\boldsymbol{\theta}}_1, ..., \widehat{\boldsymbol{\theta}}_{n_1}; (\widehat{u}_{k1}, ..., \widehat{u}_{kn})'_{k=1,...,N})$ as defined in Equation (3.15) with respect to $\boldsymbol{\theta}_0$.

This two-step estimation procedure immediately generalizes to a k-step estimation approach for k-level hierarchical Kendall copulas as defined in Remark 3.7. Resulting estimates may be used as starting values for a joint maximum likelihood estimation of the dependence parameters. In order to evaluate and compare their finite sample behavior in comparison with maximum likelihood estimates, 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;
- Maximum likelihood 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, which is transformed to Kendall's τ, so that values are on a comparable scale. The results are shown in Figure 3.7 for the case of Clayton and Gumbel cluster copulas and in Appendix C for the other five possible combinations of cluster copulas, which yield very similar results. 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 3.3 (see Example 3.6).

It turns out that the results are essentially independent of the chosen cluster copula family. Similarly, the choice of the parameters of the cluster copulas does not have a great

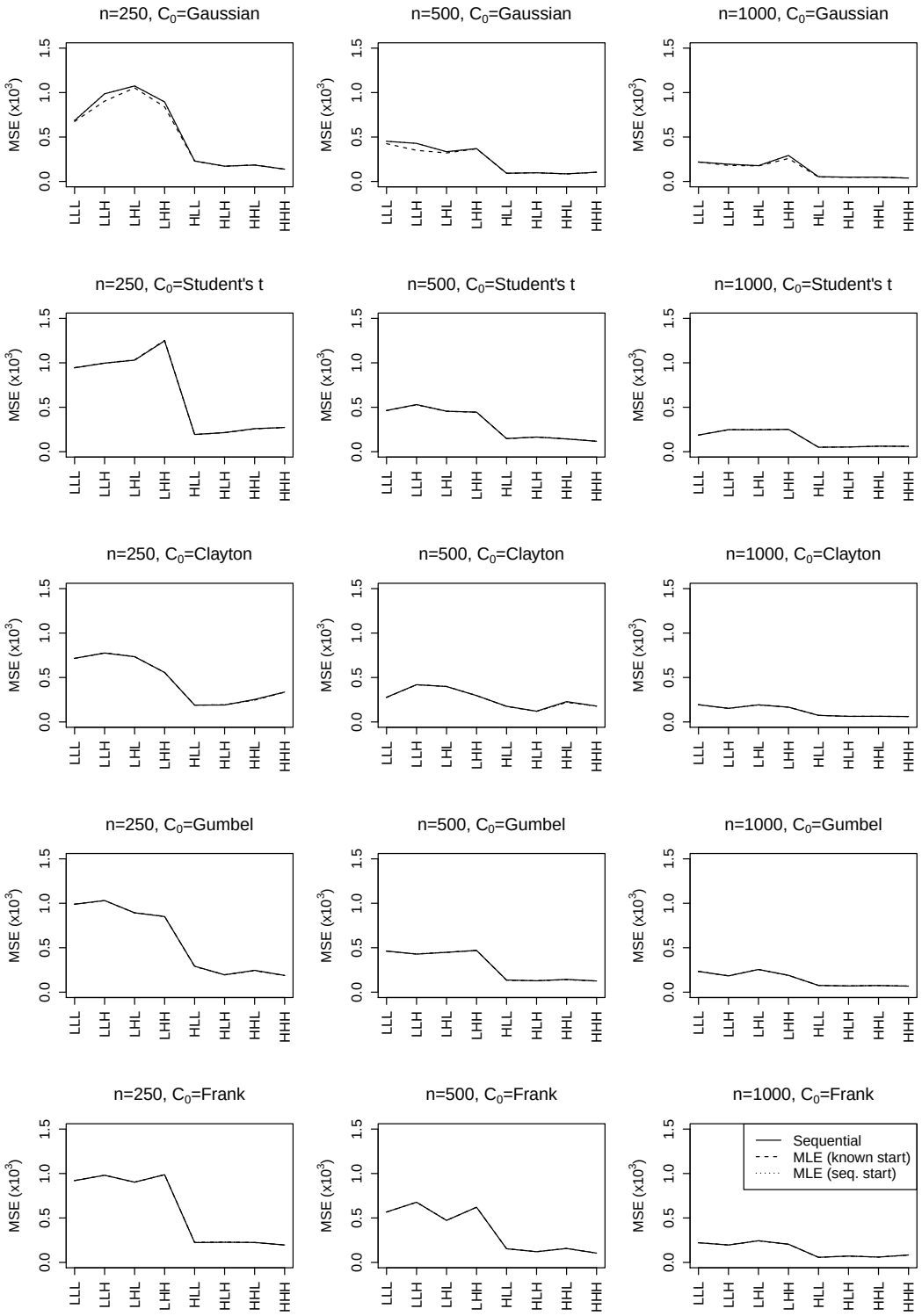


Figure 3.7: Mean squared errors (MSEs) of $\widehat{\theta}_0$ (in terms of Kendall's τ) for the three estimation procedures. Cluster copula 1: Clayton. Cluster copula 2: Gumbel. Notation for the x-axes: (τ_0,τ_1,τ_2), where $L := 0.4$ and $H := 0.7$. The range of the y-axes is chosen such that the MSEs are comparable.

influence on the results, while larger parameters of the nesting copula (stronger dependence) 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 the joint estimation of the dependence parameters.

3.5.1 Stabilizing transformation

A numerical issue in the calculation of the log likelihood $\ell_{\mathcal{K}}$ (3.15) that may occur for larger cluster sizes d_i is due to the shape of the Kendall distribution function. According to Corollary 2.11, the Kendall distribution function of a d-dimensional copula increases for fixed $z \in [0,1]$ in the dimension d. This can lead to boundary issues for values close to zero (see, in particular, the shape of the Kendall distribution functions of the independence and of the Gumbel copula, which are shown in Figure 2.6). As a remedy we propose to apply a stabilizing transformation in the calculation of $v := K(C(\boldsymbol{u}); C)$ for some data $\boldsymbol{u} \in [0,1]^d$.

For this, let $b : [0,1] \to \mathbb{R}$ be a strictly monotone function. This particularly implies that b is invertible with inverse $b^{-1} : \mathbb{R} \to [0,1]$. If we choose b such that $K \circ b^{-1} : \mathbb{R} \to [0,1]$ and $b \circ C : [0,1]^d \to \mathbb{R}$ are numerically stable, then we propose to calculate v by $K(b^{-1}(w); C)$, where $w := b(C(\boldsymbol{u}))$. Obviously, it holds that

$$K(b^{-1}(w); C) = K(b^{-1}(b(C(\boldsymbol{u}))); C) = K(C(\boldsymbol{u}); C) = v.$$

Thus, the essential question is how to find an appropriate stabilizing transformation for a given copula. We first illustrate this in a small example and then propose a general solution.

Example 3.20 (Stabilizing transformation for the independence copula)**.** The value of the independence copula along the diagonal $\boldsymbol{u} = u\mathbf{1},\ u \in [0,1]$, is $\Pi(u, ..., u) = u^d$. If we choose $b(z) = z^{1/d},\ z \in [0,1]$, the transformation normalizes the dimension along this diagonal, since $(b \circ \Pi)(u, ..., u) = b(\Pi(u, ..., u)) = u$. In fact, this also stabilizes the Kendall distribution function in the sense that the calculation of $K \circ b^{-1}$ is more stable (see the left panel of Figure 3.8 in comparison to the left panel of Figure 2.6). □

Motivated by this example, we define the diagonal of a copula C as

$$D(u; C) = C(u\mathbf{1}), \quad u \in [0,1],$$

and propose to use the stabilizing transformation

$$b(z) = D^{-1}(z; C), \quad z \in [0,1],$$

in order to normalize the dimension in the calculation of $K(C(\boldsymbol{u}); C)$ for $\boldsymbol{u} \in [0,1]^d$. The middle and the right panel of Figure 3.8 show the stabilized Kendall distribution functions, $K \circ D$, of the Clayton and the Gumbel copula, respectively. The stabilizing effect is evident in comparison to Figure 2.6.

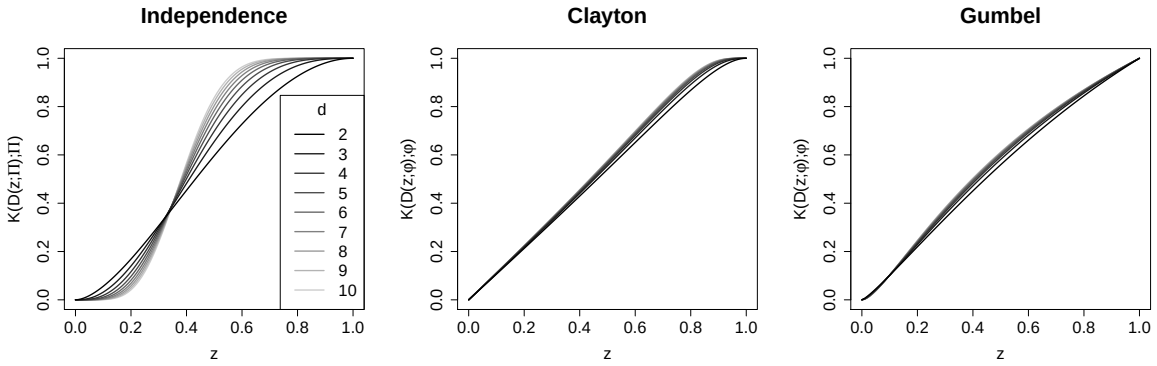


Figure 3.8: Numerically stabilized Kendall distribution functions of the independence (left panel), the Clayton (middle panel) and the Gumbel copula (right panel) for $d \in \{2, ..., 10\}$. The parameters of the Clayton and the Gumbel copula are chosen according to a Kendall's τ of 0.5.

3.6 Model selection

In practical applications, the clusters $\boldsymbol{U}_i = (U_{m_{i-1}+1}, ..., U_{m_i})', \ i = 1, ..., n_1$, have to be identified. In cases where they are not given from the data, such as industry sectors in financial data (see Section 3.7), common clustering techniques can be used (see, e.g., Hastie et al. (2009)). 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. Although it is not necessarily required by a hierarchical Kendall copula that within-cluster dependence is stronger than between-cluster dependence (see Example 3.6), clusters are typically identified as groups of variables that are strongly dependent. We therefore propose to use the following metric between the variables j_1 and j_2, which is inspired by Gower (1966) and Mantegna (1999), who however use the linear correlation coefficient,

$$d(j_1, j_2) = \sqrt{1 - \widehat{\rho}_{S,j_1 j_2}}, \tag{3.17}$$

where $\widehat{\rho}_{S,j_1 j_2}$ is the empirical Spearman's ρ_S (2.5) based on the observations $(x_{kj_1}, x_{kj_2})'$, $k = 1, ..., N$. Thus, the stronger the dependence between the variables j_1 and j_2, the smaller is $d(j_1, j_2)$. Obviously, $d(j_1, j_2) = 0$ if the variables j_1 and j_2 are comonotonic, that is, if $x_{kj_1} = x_{kj_2}$ for all $k = 1, ..., N$. Further, it holds that $d(j_1, j_2) = d(j_2, j_1)$ (symmetry) and $d(j_1, j_2) \leq d(j_1, j_3) + d(j_3, j_2)$ for another variable j_3. The validity of the triangle inequality follows from the fact that $\widehat{\rho}_{S,j_1 j_2}$ is the correlation coefficient of the ranks $r_j(k)$ of the observations $x_{kj}, \ k = 1, ..., N, \ j \in \{j_1, j_2\}$, and it holds that

$$\widehat{\rho}_{S,j_1 j_2} = 1 - \frac{6}{N(N^2 - 1)} \sum_{k=1}^{N} \left(r_{j_1}(k) - r_{j_2}(k) \right)^2 .$$

Therefore, we have that

$$\sqrt{1 - \widehat{\rho}_{S,j_1 j_2}} = \left(\frac{6}{N(N^2 - 1)} \right)^{1/2} \left(\sum_{k=1}^{N} \left(r_{j_1}(k) - r_{j_2}(k) \right)^2 \right)^{1/2},$$

which is, up to a multiplicative factor, the Euclidean distance between the ranks, so that the triangle inequality is trivially satisfied.

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 (3.16) based on purely empirical versions of the copula and the Kendall distribution function and then compute the distance (3.17) between them. This means that we form

$$\widehat{v}_{ki} := \widehat{K}_i(\widehat{C}_i(\widehat{\boldsymbol{u}}_{ki})),\ k = 1, ..., N,\ i = 1, ..., n_1, \tag{3.18}$$

where $\widehat{C}_i$ is the empirical copula of the observations $\widehat{\boldsymbol{u}}_{ki},\ k = 1, ..., N$, which is given by

$$\widehat{C}_i(\boldsymbol{u}_i) = \frac{1}{N}\sum_{k=1}^{N} 1_{\{\widehat{u}_{km_{i-1}+1} \le u_{m_{i-1}+1}, ..., \widehat{u}_{km_i} \le u_{m_i}\}}, \quad \boldsymbol{u}_i \in [0,1]^{d_i}.$$

Further, $\widehat{K}_i$ is the empirical Kendall distribution function based on $\widehat{z}_{ki} = \widehat{C}_i(\widehat{\boldsymbol{u}}_{ki}),\ k = 1, ..., N$, and hence defined as

$$\widehat{K}_i(z) = \frac{1}{N}\sum_{k=1}^{N} 1_{\{\widehat{z}_{ki} \le z\}}, \quad z \in [0,1].$$

Then the distance (3.17) between clusters i_1 and i_2 can be computed based on the empirical Spearman's ρ_S of $\widehat{v}_{ki_1}$ and $\widehat{v}_{ki_2},\ k = 1, ..., N$.

Example 3.21 (Hierarchical clustering)**.** Let $\widehat{u}_{kj},\ k = 1, ..., N, j = 1, ..., 4$, be given observations of a four-dimensional random vector. In a first step, we calculate the distance $d(j_1, j_2)$ given in Equation (3.17) for all $j_1, j_2 = 1, ..., 4,\ j_1 \neq j_2$, based on the observations. Assume that $d(1,2) < \min\{d(1,3), d(1,4), d(2,3), d(2,4), d(3,4)\}$. Therefore, we first group the variables 1 and 2.

To update the distances after the grouping, we calculate pseudo observations $\widehat{v}_{k1},\ k = 1, ..., N$, as in Equation (3.18), where the size of the first cluster is $d_1 = 2$. Then we calculate the empirical Spearman's ρ_S of $\widehat{v}_{k1}$ and $\widehat{u}_{kj},\ k = 1, ..., N$, for $j = 3, 4$ and denote the resulting values by $\widehat{\rho}_{S,\{12\}3}$ and $\widehat{\rho}_{S,\{12\}4}$, respectively. Based on these empirical values, we also get the distances $d(\{12\}, 3)$ and $d(\{12\}, 4)$, respectively.

If now $d(3,4) < \min\{d(\{12\}, 3), d(\{12\}, 4)\}$, then we also group variables 3 and 4. Distances could again be updated by calculating the pseudo observations $\widehat{v}_{k2},\ k = 1, ..., N$, as in Equation (3.18) with cluster size $d_2 = 2$. Calculating the empirical Spearman's ρ_S of $\widehat{v}_{k1}$ and $\widehat{v}_{k2},\ k = 1, ..., N$, gives the between-cluster distance $d(\{12\}, \{34\})$.

If however $d(\{12\}, 3) < \min\{d(3,4), d(\{12\}, 4)\}$, then variable 3 joins the first cluster and the distance between the cluster $\{1, 2, 3\}$ and the variable 4 could be calculated based on the empirical Spearman's ρ_S of $\widehat{v}_{k1}$ and $\widehat{u}_{k4},\ k = 1, ..., N$, where $\widehat{v}_{k1}$ has to be recalculated using Equation (3.18), since now $d_1 = 3$. □

Measuring association between multivariate random vectors through aggregation via the copula and the Kendall distribution function is, along with other association measures,

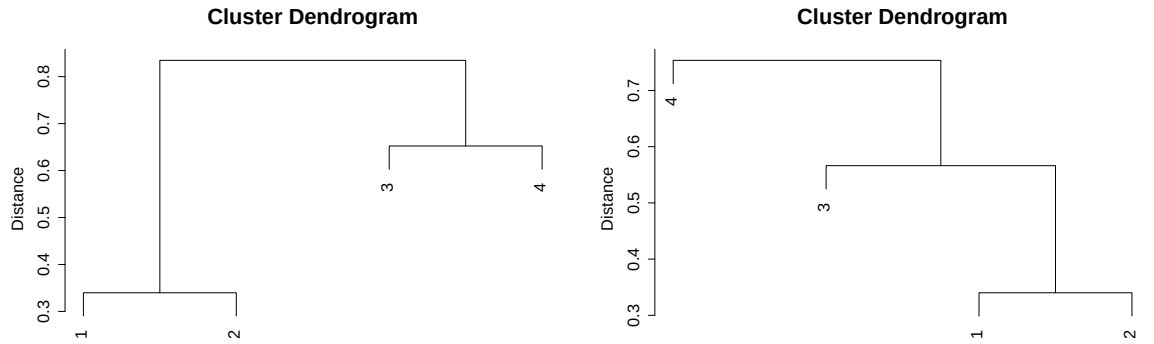


Figure 3.9: Cluster dendograms corresponding to the two cases discussed in Example 3.21.

also discussed by Grothe et al. (2011). In contrast to average linkage, 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 a *dendogram* and represents the closeness between cluster members. Exemplary dendograms illustrating Example 3.21 are shown in Figure 3.9 (see also Sections 3.7 and, especially, 5.3).

Since Kendall distributions may become almost degenerate at 0 for very large clusters (see Corollary 2.11), the size of the clusters should be chosen carefully. This issue is mitigated by the stabilizing transformation proposed in Section 3.5.1. In addition, already under medium positive dependence the convergence to the constant function at 1 is very slow (see Figure 2.6), so that the numerical issues are minor.

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 the cluster copulas is necessary. A possible approach is a stepwise selection similar to the sequential estimation procedure outlined in Algorithm 3.19, that is, the nesting copula C_0 is selected based on pseudo observations.

Algorithm 3.22 (Sequential selection of hierarchical Kendall copulas)**.** Let n_1 clusters of size $d_i \geq 1,\ i = 1, ..., n_1$, be given. Cluster and nesting copulas of a hierarchical Kendall copula are then sequentially selected as follows.

(i) For each cluster $i \in \{1, ..., n_1\}$ select C_i and estimate its parameter(s) $\boldsymbol{\theta}_i$ based on $(\widehat{\boldsymbol{u}}_{ki})_{k=1,...,N}$.

(ii) Select C_0 and estimate its parameter(s) $\boldsymbol{\theta}_0$ based on the pseudo observations $\widehat{v}_{ki},\ k = 1, ..., N,\ i = 1, ..., n_1$ (see Equation (3.16)).

Typical copula selection criteria are the AIC and the BIC, which are both likelihood-based and penalize the log likelihood for the number of parameters. Goodness-of-fit tests can also be used (see Genest et al. (2009) and Berg (2009)), in particular to verify the fit. Further, Algorithm 3.22 can also be easily generalized to select cluster and nesting copulas of multi-level hierarchical Kendall copulas (see Remark 3.7).

The proposed procedure is similar to selection approaches of hierarchical Archimedean copulas (3.14) (see Okhrin et al. (2013)) and of vine copulas (see Section 2.7.4 and Dißmann et al. (2013)). Since the selection based on pseudo observations however induces uncertainty in the selection of the nesting copula, we perform a misspecification study.

3.6.1 Copula misspecification

To analyze the effect of misspecification of the cluster and nesting copulas, we resume the setting of the simulation study in Section 3.5 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 copulas (see Section 2.7) with first tree pair copulas chosen as $C_{1,2} = C_1$, $C_{2,3} = C_0$ and $C_{3,4} = C_2$ and second and third tree pair copulas chosen as $C_{1,3|2} = C_{2,4|3} = C_{1,4|2,3} = C_0$. The parameters of the first three copulas are determined according to a Kendall's τ of τ_1, τ_0 and τ_2, respectively; the parameters of the latter three are chosen according to a decreasing value of Kendall's τ compared to the copula $C_{2,3}$: If $\tau_{j,k|D}$ denotes the Kendall's τ corresponding to the pair copula $C_{j,k|D}$, then we choose $\tau_{1,3|2} = \tau_{2,4|3} = 2\tau_0/3$ and $\tau_{1,4|2,3} = \tau_0/3$. 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 (3.14) 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 four-dimensional Gaussian and Student's t copulas (ten degrees of freedom) with correlation matrices structured according to the respective Kendall's τ values for within- and between-cluster dependence. In order to ensure positive definiteness of the correlation matrices, Kendall's τ values however 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.

The effect of misspecification is examined in terms of the Kullback-Leibler divergence (see Kullback and Leibler (1951)) between the true (simulated) model and the alternative models fitted by maximum likelihood estimation. The results are illustrated in Figures 3.10 and 3.11. 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 also Figure 3.5).

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, e.g., the shape of the scatter plots in Figure 3.3) and therefore harder to approximate by a misspecified model. As a result, the elliptical copulas are also best

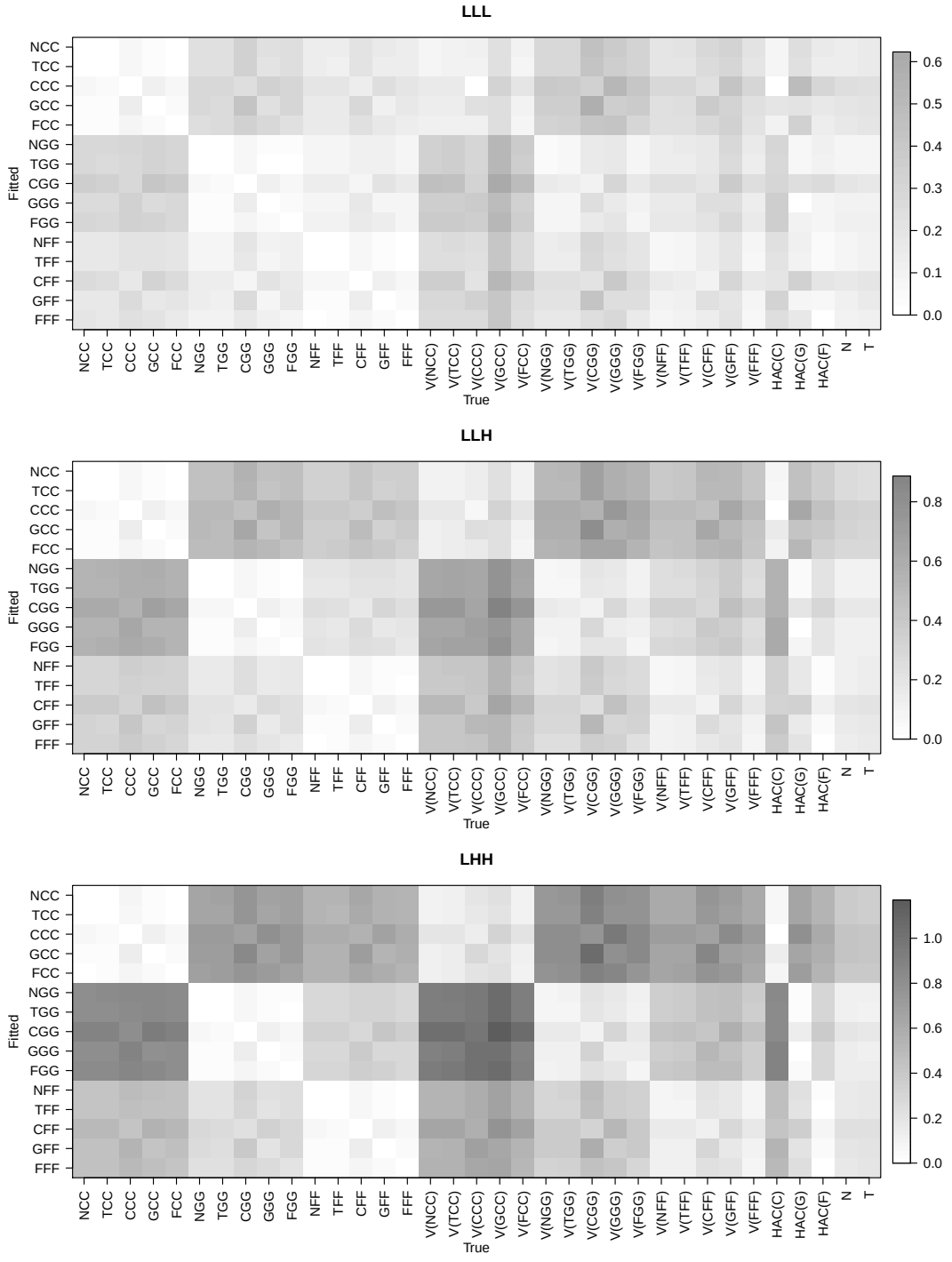


Figure 3.10: Illustration of mean Kullback-Leibler divergences (part 1): light colors indicate a small divergence, dark colors a large divergence (black corresponds to the maximum divergence of 1.839). Notation for the x- and y-axes: (C_0,C_1,C_2) with Gaussian (N), Student's t (T), Clayton (C), Gumbel (G), and Frank (F) copulas; vine copulas are indicated by 'V', hierarchical Archimedean copulas by 'HAC'. Columns correspond to the true models, rows to the fitted models. Notation for the panel titles: (τ_0,τ_1,τ_2), where $L := 0.4$ and $H := 0.7$.

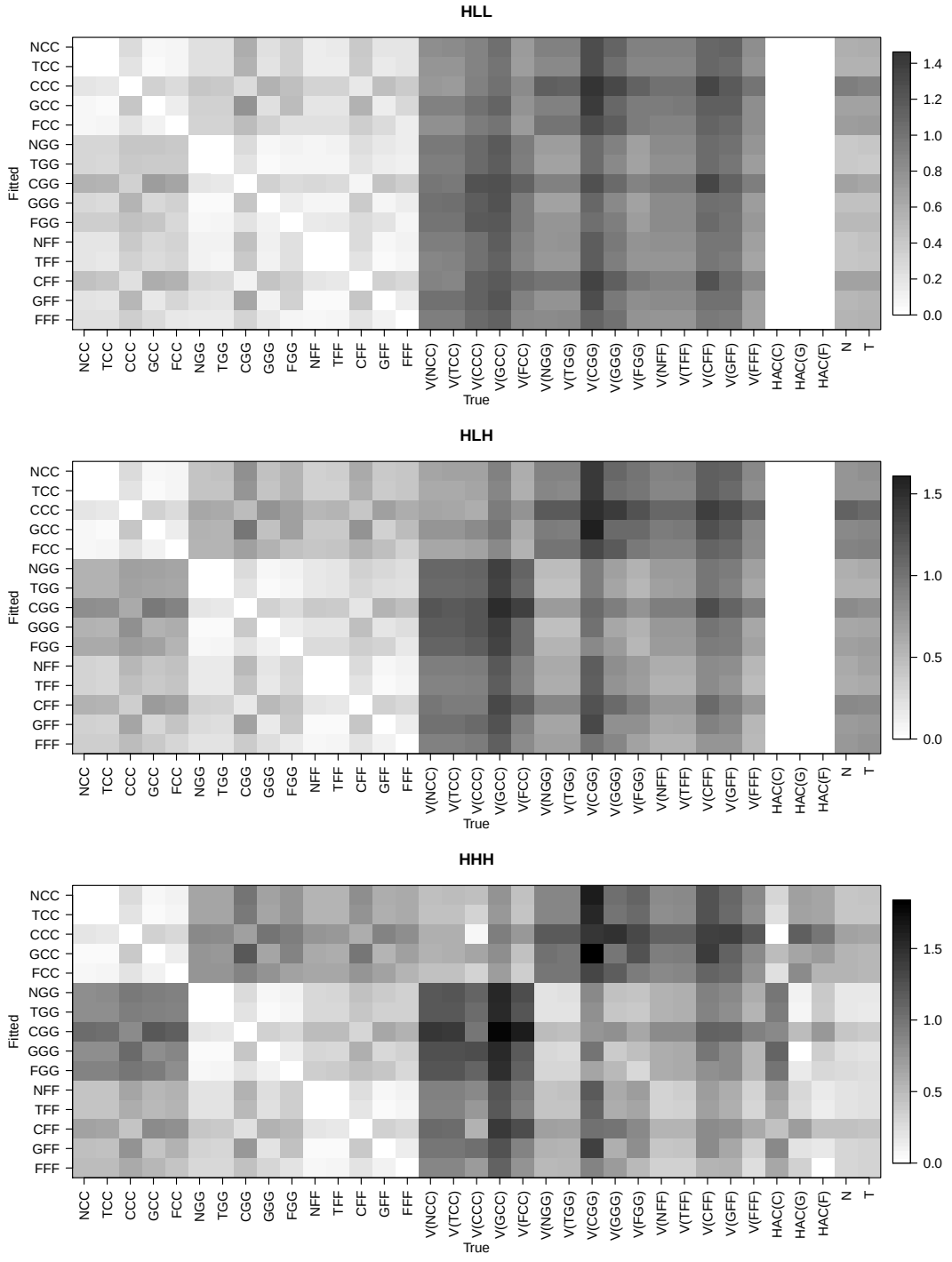


Figure 3.11: Illustration of mean Kullback-Leibler divergences (part 2): light colors indicate a small divergence, dark colors a large divergence (black corresponds to the maximum divergence of 1.839). Notation for the x- and y-axes: (C_0,C_1,C_2) with Gaussian (N), Student's t (T), Clayton (C), Gumbel (G), and Frank (F) copulas; vine copulas are indicated by 'V', hierarchical Archimedean copulas by 'HAC'. Columns correspond to the true models, rows to the fitted models. Notation for the panel titles: (τ_0,τ_1,τ_2), where $L := 0.4$ and $H := 0.7$.

approximated by hierarchical Kendall copulas with Gumbel and Frank components. 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 vine copulas 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.

3.7 Application: Returns of major German stocks

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.

3.7.1 Data

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), 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 4, 2005 to July 22, 2011) are considered, where the time series are split into a training set of $N = 1158$ observations (before August 7, 2009) and a testing set of 500 observations, which will be used for out-of-sample validation of our models in Section 4.5. Three time series of stock prices and corresponding log returns are shown in Figure 3.12.

3.7.2 Marginal modeling

As it is common in finance, we analyze the log returns using copula-GARCH models (see, e.g., Jondeau and Rockinger (2006), Patton (2006) and Liu and Luger (2009)). Following the IFM method (see Section 3.5), we preliminarily fit time series models to the marginal time series $r_{tj},\ t = 1, ..., N,\ j = 1, ..., 30$, and then work with the standardized residuals, which are transformed to marginally uniform data by the probability integral transform. In particular, we choose marginal GARCH(1,1) models with Student's t innovations to capture volatility clustering as observed in Figure 3.12:

$$\begin{aligned} r_{tj} &= \mu_j + \varepsilon_{tj}, \\ \varepsilon_{tj} &= \sigma_{tj} Z_{tj}, \\ \sigma_{tj}^2 &= \omega_j + \alpha_j \varepsilon_{t-1,j}^2 + \beta_j \sigma_{t-1,j}^2, \end{aligned} \tag{3.19}$$

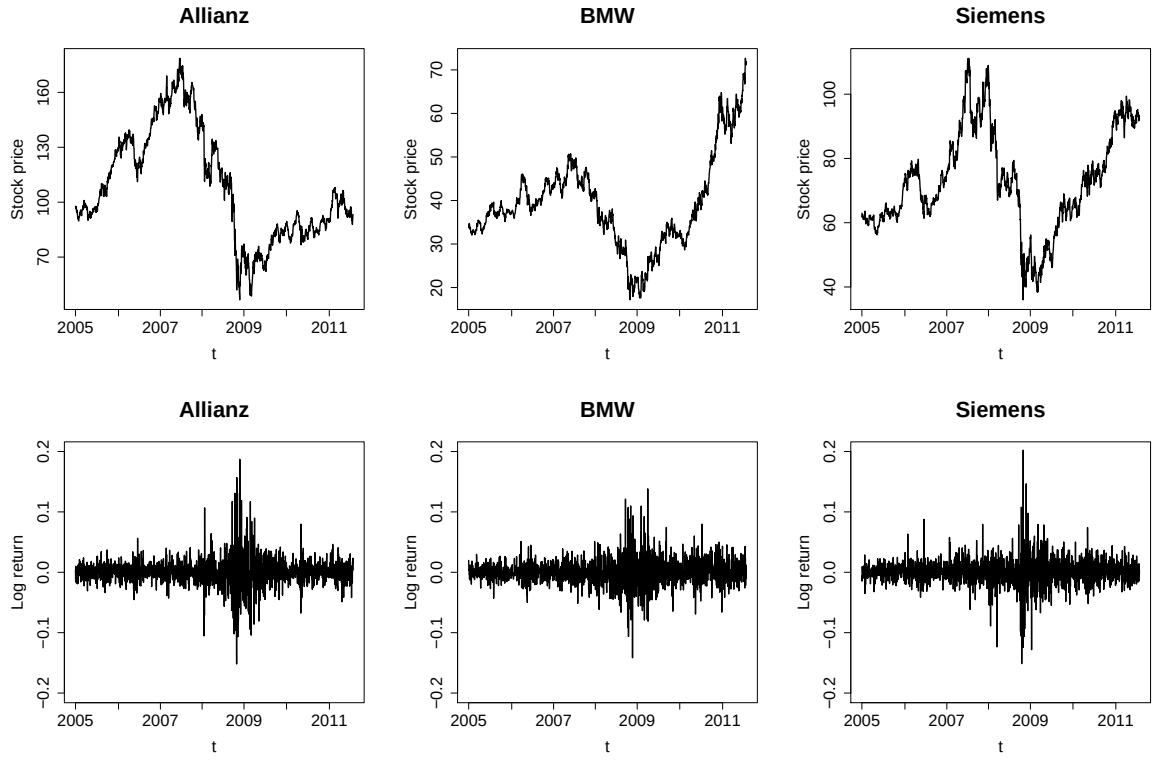


Figure 3.12: Stock prices (top) and corresponding log returns (bottom) of Allianz, BMW and Siemens from January 2005 to July 2011.

where $Z_{tj} \sim \widetilde{\mathcal{T}}_1(\nu_j)$ for $t = 1, ..., N$ and $j = 1, ..., 30$. Here, $\widetilde{\mathcal{T}}_1(\nu)$ denotes the univariate standardized Student's t distribution with ν degrees of freedom and distribution function $\widetilde{T}_\nu$ (in contrast to the Student's t distribution $\mathcal{T}_1(0, 1, \nu)$, which has variance $\nu/(\nu - 2)$).

After estimation of the model parameters μ_j, ω_j, α_j, β_j and ν_j for $j = 1, ..., 30$, the standardized residuals are given by

$$\widehat{z}_{tj} = \frac{r_{tj} - \widehat{\mu}_j}{\widehat{\sigma}_{tj}},$$

and we set $\widehat{u}_{tj} := \widetilde{T}_{\widehat{\nu}_j}(\widehat{z}_{tj})$, $t = 1, ..., N$, $j = 1, ..., 30$. The model fits have been validated with appropriate tests such as the Ljung-Box and the Kolmogorov-Smirnov test.

3.7.3 Dependence modeling

Although cluster selection is not needed here, we now illustrate the selection procedures developed in Section 3.6. Using the transformed standardized residuals $\widehat{u}_{tj}$, $t = 1, ..., N$, $j = 1, ..., 30$, we perform hierarchical clustering with the metric (3.17) and average linkage as well as aggregation using the empirical copulas and Kendall distribution functions of the clusters. The resulting dendograms are shown in Figure 3.13. For instance, the utility and the healthcare sector can easily be identified. This is not the case for the chemical and the IT companies.

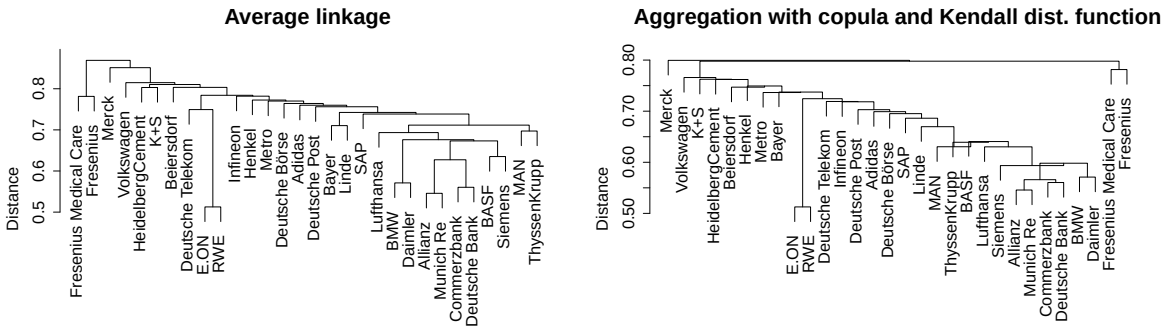


Figure 3.13: Dendograms of the DAX constituents according to average linkage (left panel) and aggregation using the empirical copulas and Kendall distribution functions of the clusters (right panel).

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

Table 3.1: Sector size, mean pairwise empirical Kendall's τ values and estimated degrees of freedom of a Student's t copula for each cluster.

Although both dendograms are quite similar, this is also an example, where hierarchical clustering using the aggregation with the empirical copulas and Kendall distribution functions does not yield clusters that have a monotone decreasing closeness with increasing level of the merger (see the right panel of Figure 3.13).

Table 3.1 shows the mean pairwise empirical Kendall's τ and the estimated degrees of freedom of a multivariate Student's t copula (see Example 2.14) 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 fit different hierarchical Kendall copulas to the training data set. Results (log likelihood, AIC, BIC) are reported in Table 3.2, both for sequential and for joint parameter estimation, of which the finite sample behavior is investigated in the simulation study in Section 3.5. As cluster copulas, we consider 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 Frank without tail dependence (see Section 2.4). For the nesting copulas, we also investigate Gaussian and Student's t copulas, where the fits of the Gaussian copula to the aggregated pseudo observations of the different sectors turn out to be always inferior to those of the Student's t copula and are therefore not displayed here. The different specified models are shown in the first column of Table 3.2.

For comparison, we also fit classical multivariate Gaussian and Student's t copulas with

Copula (cluster, nesting)	# Par.	Seq. est. log lik.	Joint est. log lik.	AIC	BIC
Hier. Kendall (Clayton, Student's t)	55	6656.50	6677.73	-13245.46	-12967.47
Hier. Kendall (Gumbel, Student's t)	55	6989.32	6992.29	-13874.57	-13596.58
Hier. Kendall (Frank, Student's t)	55	7185.70	7190.29	-14270.58	-13992.58
Hier. Kendall (Clayton, Clayton)	10	5452.34	5471.09	-10922.17	-10871.63
Hier. Kendall (Gumbel, Gumbel)	10	5860.74	5862.93	-11705.85	-11655.31
Hier. Kendall (Frank, Frank)	10	6003.60	6005.56	-11991.12	-11940.58
Gaussian	435	-	8487.71	-16105.41	-13906.73
Student's t	436	-	8906.14	-16940.28	-14736.54
Grouped Student's t	445	8934.42	-	-16978.85	-14729.62
R-vine (with pairw. indep. tests)	287	-	9204.77	-17835.54	-16384.91
R-vine	509	-	9512.29	-18006.58	-15433.87

Table 3.2: Log likelihoods according to sequential and joint estimation, numbers of parameters, AIC and BIC values of the copulas. AIC and BIC values are based on the joint estimation (if available).

unstructured correlation matrix as well as a grouped Student's t copula with ten different parameters for the degrees of freedom of the sectors (see Section 2.3). In addition, an R-vine copula is selected using the algorithm by Dißmann et al. (2013) (see Algorithm 2.25), where adequate pair copulas are chosen from the following list: Gaussian, Student's t, Clayton, Gumbel and Frank as well as rotations by 90, 180 and 270 degrees of the reflection asymmetric copulas (see Table 2.1 and Figure 2.3). We also consider a more parsimonious model specification, where the bivariate independence copula is potentially selected according to an independence test of each pair. R-vine copula parameters are estimated jointly by maximum likelihood. This is however not the case for the parameters of the grouped Student's t copula because of the numerical complexity of the density expression (2.28). Daul et al. (2003) therefore propose to independently estimate the degrees of freedom for the different sectors (see Table 3.1). But this approach does not allow to estimate the degrees of freedom parameter of the building materials sector, which has only one member. Moreover, we detect a significant improvement in the fit, when the degrees of freedom parameters are estimated jointly given the estimated correlation matrix of the standard Student's t copula. Interestingly, we find that the jointly estimated degrees of freedom are at least 50% higher than the independently estimated ones reported in Table 3.1. This implies a weaker joint tail behavior. Similarly, the estimated degrees of freedom of the standard Student's t copula are rather high: $\widehat{\nu} = 21.41$. The approach by Daul et al. (2003) therefore seems to underestimate the degrees of freedom of the sectors, since it ignores between-sector (tail) dependence, which is typically weaker than within-sector (tail) dependence.

In contrast, a 30-dimensional hierarchical Archimedean copula cannot be fitted due to the dependence restrictions of hierarchical Archimedean copulas (see Section 3.3.1): While there is moderate dependence within some clusters (see Table 3.1), there is still considerable and, especially, heterogeneous dependence among clusters (the pairwise Kendall's τ values of the aggregated pseudo observations range between 0.15 and 0.51), which cannot

	(Clayton, Student's t)		(Gumbel, Student's t)		(Frank, Student's t)	
	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error
Financials	0.757	0.050	1.531	0.036	3.816	0.210
Chemicals	0.568	0.048	1.396	0.033	2.915	0.197
Healthcare	0.323	0.043	1.215	0.028	1.810	0.191
Automobile	0.780	0.055	1.489	0.040	3.775	0.237
Industrials	0.748	0.075	1.505	0.056	3.634	0.303
Retail	0.417	0.037	1.276	0.028	2.189	0.168
IT and comm.	0.396	0.053	1.304	0.033	2.401	0.218
Utilities	1.486	0.146	2.057	0.111	6.829	0.544
Transportation	0.461	0.085	1.325	0.054	2.669	0.325
Between-sector	0.197–0.626	0.023–0.044	0.257–0.733	0.019–0.043	0.248–0.704	0.020–0.043
Deg. of freedom	15.702	1.214	12.790	0.971	19.597	1.559

	(Clayton, Clayton)		(Gumbel, Gumbel)		(Frank, Frank)	
	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error
Financials	0.771	0.051	1.543	0.037	3.946	0.210
Chemicals	0.585	0.049	1.397	0.032	3.047	0.199
Healthcare	0.295	0.042	1.242	0.026	2.003	0.182
Automobile	0.785	0.056	1.501	0.040	3.897	0.238
Industrials	0.799	0.078	1.531	0.058	3.880	0.304
Retail	0.426	0.039	1.293	0.027	2.394	0.167
IT and comm.	0.391	0.058	1.330	0.032	2.664	0.215
Utilities	1.468	0.153	2.075	0.110	6.912	0.543
Transportation	0.458	0.103	1.365	0.054	2.981	0.329
Between-sector	0.464	0.034	1.409	0.028	2.979	0.165

Table 3.3: Parameter estimates and their estimated standard errors for the hierarchical Kendall copulas (based on the training set). For the entries of the correlation matrix of the Student's t copulas, ranges are reported.

be modeled using a hierarchical Archimedean copula.

The hierarchical Kendall copulas therefore benefit from not having this restriction on the 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 the cluster copulas, the Frank copula is, according to the AIC, superior to the copulas with asymmetric tail dependence (Clayton, Gumbel). Hence, the hierarchical Kendall copula with Frank cluster copulas and Student's t nesting copula is the best model selected according to the sequential procedure outlined in Algorithm 3.22. For the six considered hierarchical Kendall 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 and 2000 samples can be found in Tables 3.3 and 3.4, showing that there is significant within- and between-sector dependence.

In comparison to 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 additional flexibility of the grouped Student's t copula leads only to a weak improvement in the model fit and, according to the BIC,

	(Clayton, Student's t)		(Gumbel, Student's t)		(Frank, Student's t)	
	Kendall's τ	Std. Error	Kendall's τ	Std. Error	Kendall's τ	Std. Error
Financials	0.275	0.013	0.347	0.015	0.374	0.016
Chemicals	0.221	0.014	0.284	0.017	0.300	0.017
Healthcare	0.139	0.016	0.177	0.019	0.195	0.019
Automobile	0.281	0.014	0.328	0.018	0.371	0.018
Industrials	0.272	0.019	0.335	0.024	0.360	0.024
Retail	0.173	0.013	0.217	0.017	0.232	0.016
IT and comm.	0.165	0.018	0.233	0.020	0.253	0.021
Utilities	0.426	0.023	0.514	0.026	0.555	0.025
Transportation	0.187	0.027	0.245	0.029	0.278	0.029
Between-sector	0.126–0.431	0.017–0.030	0.166–0.524	0.018–0.031	0.160–0.497	0.018–0.031
Deg. of freedom	15.702	1.214	12.790	0.971	19.597	1.559

	(Clayton, Clayton)		(Gumbel, Gumbel)		(Frank, Frank)	
	Kendall's τ	Std. Error	Kendall's τ	Std. Error	Kendall's τ	Std. Error
Financials	0.278	0.013	0.352	0.015	0.384	0.016
Chemicals	0.226	0.014	0.284	0.016	0.311	0.017
Healthcare	0.128	0.016	0.195	0.017	0.214	0.018
Automobile	0.282	0.014	0.334	0.018	0.380	0.018
Industrials	0.285	0.020	0.347	0.024	0.379	0.023
Retail	0.176	0.013	0.226	0.016	0.252	0.016
IT and comm.	0.164	0.020	0.248	0.018	0.277	0.020
Utilities	0.423	0.025	0.518	0.025	0.558	0.024
Transportation	0.186	0.033	0.267	0.027	0.306	0.028
Between-sector	0.188	0.011	0.290	0.014	0.305	0.014

Table 3.4: Kendall's τ values of the parameter estimates (see Table 2.1) and their estimated standard errors for the hierarchical Kendall copulas (based on the training set). For the entries of the correlation matrix of the Student's t copulas, ranges are reported.

the standard Student's t copula is even superior. The number of parameters of elliptical copulas could be reduced significantly by using correlation matrices structured according to an appropriate factor model. These however have to be fitted carefully in order to satisfy positive definiteness constraints. Overall the R-vine copulas provide the best fit, since they constitute the most flexible models. But if no independence copulas are used as building blocks, the resulting R-vine copula is even less parsimonious than the elliptical copulas. Moreover, the vine copulas are not straightforward to interpret, especially not in terms of sectoral dependence.

Given that the highly parameterized multivariate elliptical and vine copulas 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 in more detail in a Value-at-Risk forecasting study in Section 4.5. The good in-sample results obtained here are in line with the misspecification study in Section 3.6.1, 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 vine copulas.

3.8 Conclusion

In this chapter, we introduce and discuss the new class of hierarchical Kendall copulas. By grouping variables in 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, a general simulation algorithm is provided and parameter estimation methods are discussed. The availability of the density of hierarchical Kendall copulas renders feasible estimation using maximum likelihood techniques. Methods for the selection of clusters and of appropriate cluster and nesting copulas are also proposed. Finally, we show that a model with Archimedean cluster copulas and Student's t nesting copula performs very well in a substantial financial application.

In the next chapter, we develop different approaches to sampling from hierarchical Kendall copulas. Both closed-form procedures as well as approximate methods are derived and compared. At the end of the chapter, the analysis of dependencies among returns of major German stocks is continued.

4 Sampling from hierarchical Kendall copulas

Sampling from a hierarchical Kendall copula involves simulation of a random vector given that it lies in a particular level set of its copula, which is generally a non-trivial problem. We derive closed-form approaches for Archimedean, Archimax and Plackett copulas and propose approximate sampling procedures for cases where no such closed-form approach is available. This chapter is mainly based on Brechmann (2013b). The material on Archimedean copulas is taken from Brechmann (2013a).

4.1 Introduction

Sampling from different classes of copulas has been a major subject of the copula literature of the last years. For instance, Archimedean and nested Archimedean copulas have been treated in Whelan (2004), McNeil (2008) and Hofert (2008, 2011), while solutions for vine copulas have been discussed in Kurowicka and Cooke (2006), Aas et al. (2009) and Dißmann et al. (2013). For an overview, we refer to Mai and Scherer (2012).

According to Algorithm 3.17 given in Section 3.4, sampling from a hierarchical Kendall copula involves the distribution of a random vector given that it lies in a particular level set. Explicit solutions for this distribution are generally hard to find. Nevertheless, we provide general guidelines for this issue and develop closed-form sampling procedures for Archimedean, Archimax and Plackett copulas. In addition, we propose three approximate sampling methods, which are not restricted to any particular copula class: rejection-like sampling, sample reordering, and density resampling, where the latter two will be referred to as bottom-up approaches—in contrast to the top-down procedure of Algorithm 3.17. An overview of the different approaches that are developed is provided in Figure 4.1. The sampling accuracy of the approximate approaches is evaluated in a simulation study and we show how hierarchical Kendall copulas can be used to forecast the Value-at-Risk of a stock portfolio. This continues the investigation of Section 3.7.

The problem of simulating from a random vector given that it lies in a particular level set is not exclusively a problem in sampling from hierarchical Kendall copulas. Recently, Salvadori et al. (2011) proposed a notion of a multivariate return period that identifies copula level sets as the critical sets of extreme events, e.g., in hydrology (see also Gräler et al. (2013) and the discussion in Chapter 7). To simulate events that lie in such a critical set, the top-down sampling approach discussed in this chapter can be used. Especially extreme value copulas such as the asymmetric Tawn copula (see Example 2.19) are of interest in this context. As they are a special case of Archimax copulas, a closed-form sampling approach for extreme value copulas is presented in this chapter.

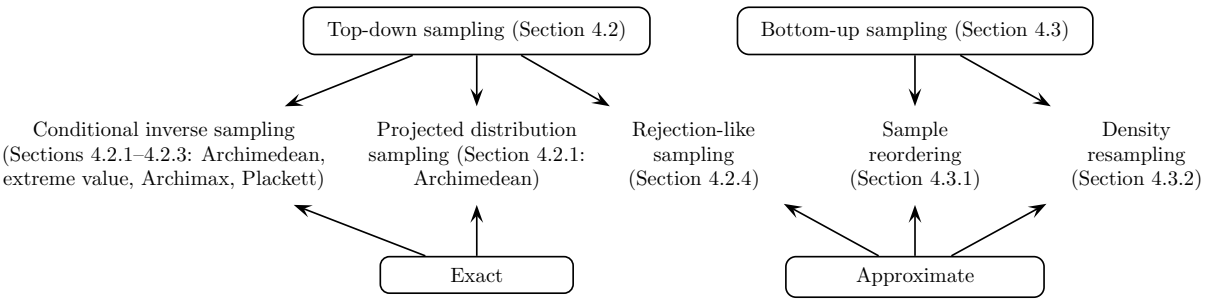


Figure 4.1: Overview of the available sampling methods. Copulas that allow for closed-form approaches are indicated in brackets.

The remainder of the chapter is organized as follows. Section 4.2 presents general guidelines for top-down sampling and develops top-down sampling procedures of Archimedean, Archimax and Plackett copulas. Rejection-like sampling, as an approximate top-down approach, is proposed in Section 4.2.4. Section 4.3 then discusses the two approximate bottom-up simulation algorithms. The three approximate methods are then evaluated and compared in Section 4.4, and the Value-at-Risk forecasting study is presented in Section 4.5. Section 4.6 concludes.

4.2 Top-down sampling

The general top-down sampling procedure is described in Algorithm 3.17. Since it essentially inverts the model formulation in Definition 3.3, we consider it as the canonical approach for sampling from hierarchical Kendall copulas. It amounts to sampling from a random vector given a particular level set (2.10) of its copula. Hence, we need to derive the distribution of $\boldsymbol{U}|C(\boldsymbol{U}) = z$, where $\boldsymbol{U} := (U_1, ..., U_d)' \sim C$ and C is a d-dimensional copula. We present two approaches to this problem: the conditional inverse method and the explicit characterization of the distribution projected to the level set $L(z; C) = \{\boldsymbol{u} \in [0,1]^d : C(\boldsymbol{u}) = z\}$. Solutions for certain classes of copulas are then discussed in the following. Alternatively, approximate rejection-like sampling can be used.

A common method to sample from a multivariate distribution is the conditional inverse method (see, e.g., Devroye (1986)). For this, we need to determine the iterative conditional distribution functions of $\boldsymbol{U}|C(\boldsymbol{U}) = z$, that is, the distribution functions of

$$\begin{aligned} &U_1|C(\boldsymbol{U}) = z, \\ &U_2|(U_1 = u_1, C(\boldsymbol{U}) = z), \\ &\quad\vdots \\ &U_{d-1}|(U_1 = u_1, ..., U_{d-2} = u_{d-2}, C(\boldsymbol{U}) = z). \end{aligned}$$

The distribution function of $U_d|(U_1 = u_1, ..., U_{d-1} = u_{d-1}, C(\boldsymbol{U}) = z)$ does not need to be determined because the value of U_d is uniquely given through the conditioning variables by $u_d = C^{-1}(z|u_1, ..., u_{d-1})$. The corresponding conditional distribution functions are denoted by $F_{U_j|U_1,...,U_{j-1},C(\boldsymbol{U})}(\cdot|u_1, ..., u_{j-1}, z)$ and densities by $f_{U_j|U_1,...,U_{j-1},C(\boldsymbol{U})}(\cdot|u_1, ..., u_{j-1}, z)$ for $j = 1, ..., d-1$, respectively. This yields the following algorithm.

Algorithm 4.1 (Conditional inverse method)**.** Let C be a d-dimensional copula and $z \in (0,1)$.

(i) Sample $w_1, ..., w_{d-1}$ independently from the uniform distribution.

(ii) For $j = 1, ..., d-1$: $u_j = F^{-1}_{U_j|U_1,...,U_{j-1},C(\boldsymbol{U})}(w_j|u_1, ..., u_{j-1}, z)$.

(iii) Set $u_d = C^{-1}(z|u_1, ..., u_{d-1})$.

(iv) Return $(u_1, ..., u_d)'$.

Hence, the problem is to determine the conditional distribution functions, which are generally not given in closed form.

Theorem 4.2 (Conditional distributions)**.** Let $\boldsymbol{U} \sim C$, where C is a d-dimensional copula. Then it holds for all $j = 1, ..., d-1$, $(u_1, ..., u_{j-1})' \in [0,1]^{j-1}$, $z \in (0,1)$ and $u \in (C^{-1}(z|u_1, ..., u_{j-1}), 1)$ that

$$F_{U_j|U_1,...,U_{j-1},C(\boldsymbol{U})}(u|u_1, ..., u_{j-1}, z) = \frac{\int_{C^{-1}(z|u_1,...,u_{j-1})}^{u} g_j(u_1, ..., u_j, z)\, du_j}{\int_{C^{-1}(z|u_1,...,u_{j-1})}^{1} g_j(u_1, ..., u_j, z)\, du_j}, \tag{4.1}$$

where

$$g_j(u_1, ..., u_j, z) = \int_{C^{-1}(z|u_1,...,u_j)}^{1} ... \int_{C^{-1}(z|u_1,...,u_{d-2})}^{1} \widetilde{g}(u_1, ..., u_{d-1}, z)\, du_{d-1}...du_{j+1}, \tag{4.2}$$

and $\widetilde{g}(u_1, ..., u_{d-1}, z) = c(u_1, ..., u_{d-1}, C^{-1}(z|u_1, ..., u_{d-1}))\, \frac{\partial}{\partial z} C^{-1}(z|u_1, ..., u_{d-1})$.

Proof: The idea is to derive the conditional density $f_{U_j|U_1,...,U_{j-1},C(\boldsymbol{U})}(\cdot|u_1, ..., u_{j-1}, z)$ and then integrate to obtain the distribution function. We observe that

$$f_{U_j|U_1,...,U_{j-1},C(\boldsymbol{U})}(u_j|u_1, ..., u_{j-1}, z) = \frac{f_{U_1,...,U_j,C(\boldsymbol{U})}(u_1, ..., u_{j-1}, u_j, z)}{f_{U_1,...,U_{j-1},C(\boldsymbol{U})}(u_1, ..., u_{j-1}, z)}. \tag{4.3}$$

According to the change of variables $\boldsymbol{U} \mapsto (U_1, ..., U_{d-1}, C(\boldsymbol{U}))'$ similar as in Equation (3.5), we have that

$$\begin{aligned} f_{U_1,...,U_{d-1},C(\boldsymbol{U})}(u_1, ..., u_{d-1}, z) &= c(u_1, ..., u_{d-1}, C^{-1}(z|u_1, ..., u_{d-1}))\, \frac{\partial}{\partial z} C^{-1}(z|u_1, ..., u_{d-1}) \\ &= \widetilde{g}(u_1, ..., u_{d-1}, z). \end{aligned}$$

Therefore, the numerator can be rewritten as

$$\begin{aligned} &f_{U_1,...,U_j,C(\boldsymbol{U})}(u_1, ..., u_j, z) \\ &= \int_{C^{-1}(z|u_1,...,u_j)}^{1} f_{U_1,...,U_{j+1},C(\boldsymbol{U})}(u_1, ..., u_{j+1}, z)\, du_{j+1} \\ &= ... = \int_{C^{-1}(z|u_1,...,u_j)}^{1} ... \int_{C^{-1}(z|u_1,...,u_{d-2})}^{1} f_{U_1,...,U_{d-1},C(\boldsymbol{U})}(u_1, ..., u_{d-1}, z)\, du_{d-1}...du_{j+1} \\ &= \int_{C^{-1}(z|u_1,...,u_j)}^{1} ... \int_{C^{-1}(z|u_1,...,u_{d-2})}^{1} \widetilde{g}(u_1, ..., u_{d-1}, z)\, du_{d-1}...du_{j+1} \\ &= g_j(u_1, ..., u_j, z). \end{aligned}$$

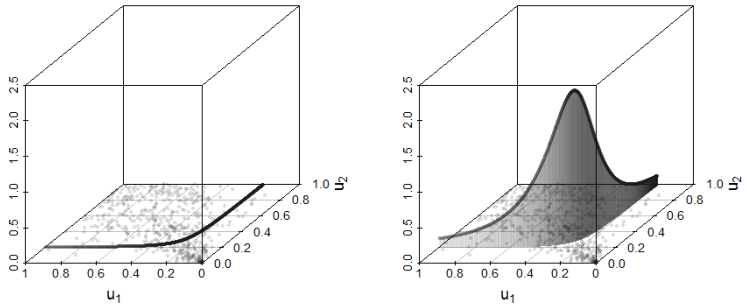


Figure 4.2: Distribution of $\boldsymbol{U}|C(\boldsymbol{U}) = z$ for the Clayton copula. The left panel shows the level set $L(z; C)$ for $z = 0.2$ and the right panel illustrates the univariate probability density on $L(z; C)$. The parameter of the Clayton copula is chosen according to a Kendall's τ of 0.5 ($\theta = 2$).

Further, the denominator of (4.3) then reads as

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

In the case $j = 1$, this is simply the density expression of the Kendall distribution function (see Definition 2.9).

Finally, we obtain the expression for the conditional distribution function (4.1) by integration over (4.3). □

In general, the conditional distribution functions given in Equation (4.1) do not allow for explicit expressions. Especially if the copula quantile function $C^{-1}(\cdot|u_1, \dots, u_j)$ is not available in closed form, such as for common elliptical copulas, Equation (4.1) hardly simplifies.

In such cases, a potential alternative may be to directly consider the distribution of $\boldsymbol{U} \sim C$ projected to the $(d-1)$-dimensional level set $L(z; C) \subset [0, 1]^d$. This $(d-1)$-dimensional distribution of $\boldsymbol{U}|C(\boldsymbol{U}) = z$ is illustrated for the Clayton copula (see Example 2.15) in Figure 4.2.

To derive an explicit characterization of this distribution, it is promising if a stochastic representation of the copula is available (see, e.g., Equation (2.26) for the Student's t copula and Equation (2.32) for Archimedean copulas). In fact, both elliptical and Archimedean copulas can be traced back to the same underlying distribution: the Dirichlet distribution.

Remark 4.3 (Dirichlet distribution)**.** The d-dimensional Dirichlet distribution $D(\boldsymbol{\alpha})$ with parameter vector $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_d)'$, $\alpha_j > 0$, $j = 1, \dots, d$, is a distribution on the d-dimensional unit simplex $\mathcal{S}_{d-1}$ (see Equation (2.33)). We write $\boldsymbol{S} := (S_1, \dots, S_d)' \sim D(\boldsymbol{\alpha})$

if $\boldsymbol{S}$ has density

$$f_D(\boldsymbol{s};\boldsymbol{\alpha}) = \frac{\Gamma\left(\sum_{j=1}^d \alpha_j\right)}{\prod_{j=1}^d \Gamma(\alpha_j)} \prod_{j=1}^d s_j^{\alpha_j - 1}, \quad \boldsymbol{s} := (s_1, ..., s_d)' \in \mathcal{S}_{d-1}.$$

If $d = 2$, the Dirichlet distribution reduces to the beta distribution, of which it is the multivariate generalization (see, e.g., Fang et al. (1990) and Kotz et al. (2000)). □

If $\alpha_1 = ... = \alpha_d = 1$, then the Dirichlet distribution is the uniform distribution on $\mathcal{S}_{d-1}$. McNeil and Nešlehová (2009) show that Archimedean copulas are the survival copulas of random vectors $\boldsymbol{X}$ with so-called ℓ_1-norm symmetric distribution, which can be represented as

$$\boldsymbol{X} \stackrel{d}{=} R\boldsymbol{S}, \tag{4.4}$$

where $\boldsymbol{S} \sim D(1, ..., 1)$ and the radial variable R is a non-negative random variable independent of $\boldsymbol{S}$ (see Fang et al. (1990)). This yields the representation (2.32).

Conversely, an elliptically distributed random vector $\boldsymbol{X}$ with mean $\boldsymbol{\mu} \in \mathbb{R}^d$ and positive definite scale matrix $\Sigma \in \mathbb{R}^{d\times d}$ can be represented as (see Fang and Fang (1988) and Hashorva et al. (2007))

$$\boldsymbol{X} \stackrel{d}{=} \boldsymbol{\mu} + RA'\left(\mathcal{I}_1 S_1^{1/2}, ..., \mathcal{I}_d S_d^{1/2}\right)', \tag{4.5}$$

where $\boldsymbol{S} \sim D(\frac{1}{2}, ..., \frac{1}{2})$ and $A \in \mathbb{R}^{d\times d}$ such that $A'A = \Sigma$. Further, R is a non-negative random variable independent of $\boldsymbol{S}$, and $\mathcal{I}_1, ..., \mathcal{I}_d$ are independent and identically distributed with $P(\mathcal{I}_1 = -1) = P(\mathcal{I}_1 = 1) = 0.5$. They are also independent of $\boldsymbol{S}$ and of R. The random vector $(\mathcal{I}_1 S_1^{1/2}, ..., \mathcal{I}_d S_d^{1/2})'$ is therefore uniformly distributed on the d-dimensional unit sphere $\{\boldsymbol{x} \in \mathbb{R}^d : \sum_{j=1}^d x_j^2 = 1\}$. As introduced in Section 2.2, elliptical copulas are the copulas of elliptically distributed random vectors.

The major difference in the stochastic representations of Archimedean and elliptical copulas in terms of the Dirichlet distribution lies in the random signs $\mathcal{I}_j,\ j = 1, ..., d$. While the radial variable of ℓ_1-norm symmetric random vectors (4.4) uniquely characterizes level sets of the distribution, this is not the case for elliptical distributions (4.5). As a result, we will be able to explicitly derive the distribution of $\boldsymbol{U}|C(\boldsymbol{U}) = z$, when C is an Archimedean copula, but not when it is elliptical.

4.2.1 Archimedean copulas

Closed-form top-down sampling procedures for an Archimedean copula $C(\cdot;\varphi)$ with generator φ can be derived both using the conditional inverse method and by explicitly characterizing the distribution of $\boldsymbol{U}|C(\boldsymbol{U};\varphi) = z$, as noted above. We begin with the conditional inverse method, where particularly convenient expressions can be obtained for the conditional distribution functions $F_{U_j|U_1,...,U_{j-1},C(\boldsymbol{U})},\ j = 1, ..., d-1$ (see Theorem 4.2).

Lemma 4.4 (Conditional distributions of Archimedean copulas)**.** Let $\boldsymbol{U} \sim C(\cdot;\varphi)$, where $C(\cdot;\varphi)$ is a d-dimensional Archimedean copula with generator φ. Then it holds for all $j = 1, ..., d-1$, $(u_1, ..., u_{j-1})' \in [0,1]^{j-1}$, $z \in (0,1)$ and $u \in (C^{-1}(z|u_1, ..., u_{j-1};\varphi), 1)$ that

$$F_{U_j|U_1,...,U_{j-1},C(\boldsymbol{U};\varphi)}(u|u_1, ..., u_{j-1}, z;\varphi) = \left(1 - \frac{\varphi(u)}{\varphi(z) - \sum_{1\leq i<j}\varphi(u_i)}\right)^{d-j}. \tag{4.6}$$

Proof: See Appendix B.1 for two proofs. One involves Theorem 4.2 and the other exploits properties of the Dirichlet distribution (see Remark 4.3). □

Lemma 4.4 allows to use the conditional inverse method to sample from Archimedean copulas, for which the conditional distribution functions can easily be inverted in closed form.

Algorithm 4.5 (Conditional inverse method for Archimedean copulas)**.** Let $C(\cdot;\varphi)$ be a d-dimensional Archimedean copula with generator φ and $z \in (0,1)$.

(i) Sample $w_1, ..., w_{d-1}$ independently from the uniform distribution.

(ii) For $j = 1, ..., d-1$: $u_j = \varphi^{-1}((1 - w_j^{1/(d-j)})(\varphi(z) - \sum_{1\leq i<j}\varphi(u_i)))$.

(iii) Set $u_d = \varphi^{-1}(\varphi(z) - \sum_{1\leq i<d}\varphi(u_i))$.

(iv) Return $(u_1, ..., u_d)'$.

In Step (iii) the copula quantile function as defined in Equation (2.35) is used. Illustrations of samples from bivariate and trivariate Clayton copulas generated according to this algorithm can be found in Figure 3.6.

As noted above, we can also explicitly consider the distribution of $\boldsymbol{U}$ on a level set. The following result is an immediate consequence of the stochastic representation of Archimedean copulas in terms of ℓ_1-norm symmetric distributions (see Equations (2.32) and (4.4)).

Proposition 4.6 (Projected distribution of Archimedean copulas)**.** Let $\boldsymbol{U} \sim C(\cdot;\varphi)$, where $C(\cdot;\varphi)$ is a d-dimensional Archimedean copula with generator φ. Then it holds for $z \in (0,1)$ that

$$(\boldsymbol{U}|C(\boldsymbol{U};\varphi) = z) \overset{d}{=} \left(\varphi^{-1}(S_1\varphi(z)), ..., \varphi^{-1}(S_d\varphi(z))\right)', \tag{4.7}$$

where $\boldsymbol{S} = (S_1, ..., S_d)'$ is uniformly distributed on the unit simplex $\mathcal{S}_{d-1}$, that is, $\boldsymbol{S} \sim D(1, ..., 1)$ (see Remark 4.3).

Proof: According to Equation (2.32), we have the representation $(\varphi(U_1), ..., \varphi(U_d))' \overset{d}{=} R\boldsymbol{S}$, where $R = \sum_{j=1}^{d}\varphi(U_j) = \varphi(C(\boldsymbol{U};\varphi))$ is the radial part, which is independent of $\boldsymbol{S}$. Fixing the level set $L(z;\varphi)$ is therefore equivalent to setting $R = \varphi(z)$, so that we obtain Equation (4.7). □

Equation (4.7) for the distribution of an Archimedean copula projected to a level set 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 be used to provide an alternative sampling algorithm for $\boldsymbol{U}|C(\boldsymbol{U};\varphi) = z$, which can be shown to be equivalent to Algorithm 4.5, when using explicit expressions for the observations $(s_1, ..., s_d)'$ from $\boldsymbol{S}$ in terms of uniform random variables (see Hering (2011, Lemma 3.1.8)).

Algorithm 4.7 (Projected distribution sampling for Archimedean copulas)**.** Let $C(\cdot;\varphi)$ be an Archimedean copula with generator φ and $z \in (0,1)$.

(i) Sample $(s_1, ..., s_d)'$ from $\boldsymbol{S}$.

(ii) For $j = 1, ..., d$: $u_j = \varphi^{-1}(s_j\varphi(z))$.

(iii) Return $(u_1, ..., u_d)'$.

Algorithms 4.5 and 4.7 can both also be used to sample from an Archimedean copula $C(\cdot;\varphi)$. For this, we need to sample the level of the level set from the corresponding Kendall distribution function $K(\cdot;\varphi)$ (see Equation (2.36)). This can be done, for example, by independently drawing an additional uniform observation w_d and setting $z = K^{-1}(w_d;\varphi)$ prior to the other steps. An equivalent version of this result has previously been stated in Wu et al. (2007), which has been restated by Hering (2011) in the setting of the work by McNeil and Nešlehová (2009), as we use it in Algorithm 4.7.

4.2.2 Extreme value and Archimax copulas

For bivariate extreme value and Archimax copulas as introduced in Section 2.5, a variant of the conditional inverse method (see Algorithm 4.1) can be used in order to circumvent that there is no known closed-form copula quantile function in general. We first state it for the more general class of Archimax copulas (2.48) and then explicitly formulate it for the special case of extreme value copulas (2.43).

Let $(U_1, U_2)' \sim C(\cdot,\cdot;\varphi, A)$, where $C(\cdot,\cdot;\varphi, A)$ is φ is an Archimedean generator and A is a Pickands dependence function. Furthermore, define

$$T := \varphi(U_2)/(\varphi(U_1) + \varphi(U_2)) \quad \text{and} \quad Z := C(U_1, U_2;\varphi, A).$$

Instead of directly sampling from $U_1|Z = z$, the idea is to first sample t from $T|Z = z$ and then solve the equations $\varphi(u_2)/(\varphi(u_1) + \varphi(u_2)) = t$ and $C(u_1, u_2;\varphi, A) = z$ for u_1 and u_2, which hence constitute observations from $(U_1, U_2)'|Z = z$.

Assuming that all required derivatives exist, it holds according to Capéraà et al. (2000, Proposition 5.1) that

$$P(T \le t, Z \le z) = K(z;\varphi)\left(t + t(1-t)\frac{A'(t)}{A(t)}\right) + (z - K(z;\varphi))\,\tau_A(t), \quad t, z \in [0,1],$$

where $K(z;\varphi) = z - \varphi(z)/\varphi'(z)$ is the Kendall distribution function of the corresponding Archimedean copula with generator φ (see Equation (2.37)) and

$$\tau_A(t) = \int_0^t \frac{s(1-s)}{A(s)}\, dA'(s).$$

In the case $t = 1$, this is the Kendall's τ of the extreme-value copula with Pickands dependence function A, that is, $\tau(A) = \tau_A(1)$ (see Equation (2.45) or (2.46)). It follows that the distribution function $F_{T|Z}$ of $T|Z = z$ is given by

$$F_{T|Z}(t|z;\varphi,A) := P(T \le t|Z = z) = \frac{k(z;\varphi)\left(t + t(1-t)\frac{A'(t)}{A(t)}\right) + (1 - k(z;\varphi))\,\tau_A(t)}{k(z;\varphi) + (1 - k(z;\varphi))\,\tau(A)},$$

where $k(z;\varphi) = K'(z;\varphi) = \varphi(z)\varphi''(z)/(\varphi'(z))^2$ is the density of the Kendall distribution function of the bivariate Archimedean copula with generator φ.

Following the idea outlined above, this gives the following conditional simulation algorithm for Archimax copulas.

Algorithm 4.8 (Conditional inverse method for Archimax copulas)**.** Let $C(\cdot,\cdot;\varphi,A)$ be an Archimax copula with generator φ and Pickands dependence function A and let $z \in (0,1)$.

(i) Sample w from the uniform distribution.

(ii) Set $t = F^{-1}_{T|Z}(w|z;\varphi,A)$.

(iii) Set $u_1 = \varphi^{-1}\left((1-t)\frac{\varphi(z)}{A(t)}\right)$ and $u_2 = \varphi^{-1}\left(t\frac{\varphi(z)}{A(t)}\right)$.

(iv) Return $(u_1,u_2)'$.

As can be easily verified, the values u_1 and u_2 in Step (iii), in fact, solve the equations $\varphi(u_2)/(\varphi(u_1)+\varphi(u_2)) = t$ and $C(u_1,u_2;\varphi,A) = z$.

In the case of extreme value copulas ($\varphi(t) = -\log t$), the conditional distribution function $F_{T|Z}$ simplifies to

$$F_{T|Z}(t|z;A) = \frac{\log z\left(t + t(1-t)\frac{A'(t)}{A(t)}\right) - (1+\log z)\,\tau_A(t)}{\log z - (1+\log z)\,\tau(A)}.$$

Algorithm 4.8 can then be restated for extreme value copulas as follows.

Algorithm 4.9 (Conditional inverse method for extreme value copulas)**.** Let $C(\cdot,\cdot;A)$ be an extreme value copula with Pickands dependence function A and let $z \in (0,1)$.

(i) Sample w from the uniform distribution.

(ii) Set $t = F^{-1}_{T|Z}(w|z;A)$.

(iii) Set $u_1 = z^{(1-t)/A(t)}$ and $u_2 = z^{t/A(t)}$.

(iv) Return $(u_1,u_2)'$.

Figure 4.3 shows samples from different Tawn copulas (see Example 2.19) generated according to this algorithm. It illustrates the effect of asymmetry on the contour line and on the samples on the contour line.

In both algorithms, inverses of $F_{T|Z}$ have to be obtained numerically, since a closed-form inversion is not feasible in general. Moreover, Algorithm 4.8 can, of course, also be used for sampling from Archimedean copulas if $A(t) = 1$ for all $t \in [0,1]$ (see Section 4.2.1).

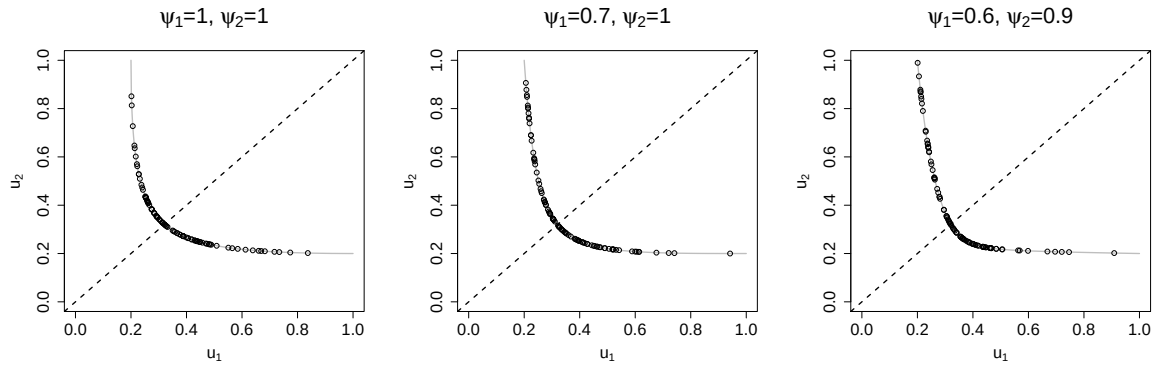


Figure 4.3: Scatter plots of samples from different Tawn copulas with a Kendall's τ of 0.5 on the level set at $z = 0.2$. The corresponding contour lines are shown in gray. The dashed line is the angle bisector.

4.2.3 Plackett copula

The conditional inverse method for the Plackett copula $C(\cdot, \cdot; \alpha)$ with parameter $\alpha \in (-1, \infty) \setminus \{0\}$ can be derived through a direct application of Theorem 4.2. The calculation of the conditional distribution function $F_{U_1|C(U_1,U_2;\alpha)}(\cdot|z;\alpha)$ for level $z \in (0,1)$ is rather cumbersome but provides a closed-form expression. Its derivation can be found in Appendix B.4. Algorithm 4.1 is then straightforward to apply. The copula quantile function, as needed in Step (iii) of the conditional inverse algorithm, is given in Equation (2.52).

For the FGM copula (see Equation (2.53)), a similar derivation as for the Plackett copula is possible, but it does not lead to such a convenient closed-form expression of the conditional distribution function.

While for Gaussian copulas no closed-form sampling procedure is known, the Plackett copula may be used as an approximate substitute in the bivariate case, since the properties of the Gaussian and the Plackett copula are quite similar (see Example 2.13, Section 2.6, and, especially, Figures A.3 and A.11).

4.2.4 Rejection-like sampling

For classes of copulas, for which no closed-form solutions are available, such as the elliptical copulas, approximate rejection-like sampling may be used to generate approximate observations from $\boldsymbol{U}|C(\boldsymbol{U}) = z$: Instead of sampling from $\boldsymbol{U}|C(\boldsymbol{U}) = z$, we choose a small number $\varepsilon > 0$ and sample from $\boldsymbol{U}|(z - \varepsilon < C(\boldsymbol{U}) < z + \varepsilon)$. This is illustrated in the left panel of Figure 4.4.

Algorithm 4.10 (Rejection-like sampling)**.** Let C be a d-dimensional copula, $z \in (0,1)$ and $\varepsilon > 0$.

(i) Sample $\boldsymbol{u}$ from C.

(ii) If $|C(\boldsymbol{u}) - z| < \varepsilon$, return $\boldsymbol{u}$. Otherwise go back to Step (i).

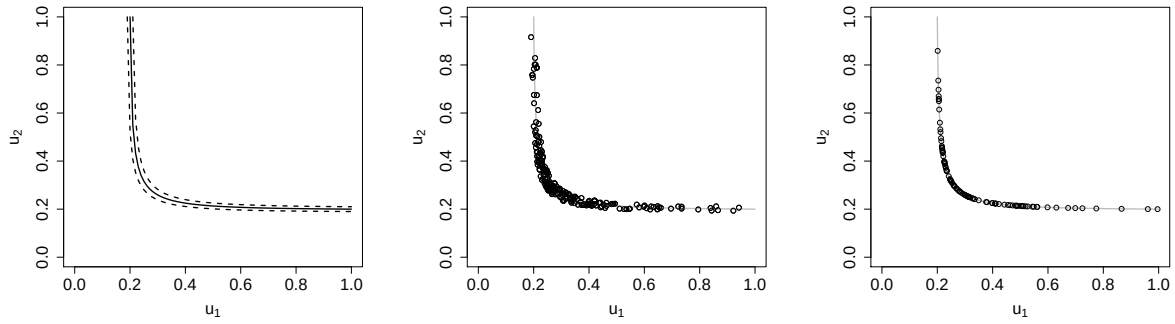


Figure 4.4: Left panel: level set $L(z; C)$ at $z = 0.2$ of the Clayton copula with parameter $\theta = 2$ (Kendall's τ of 0.5). The dashed lines illustrate the non-rejection area $(z-\varepsilon,\ z+\varepsilon)$ for $\varepsilon = 0.01$. Middle and right panel: scatter plots of corresponding samples generated through rejection-like sampling and closed-form conditional inverse sampling (see Algorithm 4.5).

The middle panel of Figure 4.4 shows an exemplary sample from a Clayton copula for $z = 0.2$. The error ε is chosen as 0.01. Notice in particular the difference to the closed-form solution shown in the right panel of Figure 4.4, which is reproduced from Figure 3.6.

The accuracy of this approximate method can be controlled through the choice of $\varepsilon > 0$, which is an upper bound on the absolute sampling error $|C(\boldsymbol{u}) - z|$. The smaller ε is, the more accurate the results are, but, at the same time, the higher the computing time is. We will quantify this trade-off in the simulation study in Section 4.4.

Remark 4.11 (Improved acceptance ratio)**.** The acceptance ratio of this sampling approach is rather small, since $L_\varepsilon(z, C) := \{\boldsymbol{u} \in [0,1]^d : C(\boldsymbol{u}) \in (z - \varepsilon, z + \varepsilon)\}$ is a quite small subset of $[0,1]^d$. This is especially true in higher dimensions. One way to improve the acceptance ratio is not to sample from the copula C directly but from its restriction to the upper orthant $[z-\varepsilon, 1]^d$, since for $\boldsymbol{u} \in L_\varepsilon(z, C)$ it always holds that $\min\{u_1, ..., u_d\} > z-\varepsilon$ due to the upper Fréchet-Hoeffding bound (2.3). Whether this actually improves the computational efficiency of rejection-like sampling strongly depends on the efficiency of the method to generate observations in the upper orthant. For the multivariate Gaussian copula this can be carried out using a Gibbs sampler for the multivariate normal distribution truncated to $[\Phi^{-1}(z-\varepsilon), 1]^d$ (see Geweke (1991)). In high dimensions and if a large sample size is required for a given level z, this may be advantageous to unrestricted sampling from the copula. A similar approach is available for the Student's t copula. □

4.3 Bottom-up sampling

In this section, we propose two methods for bottom-up sampling of hierarchical Kendall copulas as alternatives to the top-down approach discussed above. Both approaches have in common that they start by drawing a sample of multivariate independent observations. These observations are then appropriately reordered (first method) or resampled (second

method) such that they represent an approximate sample of the hierarchical Kendall copula. This means that we essentially start at the bottom of Figure 3.2, which illustrates the definition of hierarchical Kendall copulas (see Definition 3.3), and then adjust the sample according to the hierarchical dependence structure of the model.

Here, the nature of the approximation to the true distribution is different to rejection-like sampling, which allows to set an explicit error bound. When using either one of the presented bottom-up sampling methods, convergence to the true distribution is only attained with increasing sample sizes. In other words, in small samples the exact underlying distribution is unknown, leading to potentially false conclusions. Such small-sample effects are therefore investigated in the simulation study in Section 4.4.

4.3.1 Sample reordering

In the context of hierarchical dependence models, the use of sample reordering, which was originally developed by Iman and Conover (1982), is proposed by Arbenz et al. (2012). The idea of sample reordering is to independently sample margins and the copula and then to reorder the independent margins according to the ranks of the dependent sample generated from the copula. For hierarchical copulas, this may be used to reorder aggregated samples from the different clusters according to a sample from the nesting copula. Bottom-up sampling using sample reordering proceeds as follows.

Algorithm 4.12 (Sample reordering of hierarchical Kendall copulas)**.** Let $C_{\mathcal{K}}$ be a hierarchical Kendall copula with cluster copulas $C_1, ..., C_{n_1}$ and nesting copula C_0.

(i) Generate a sample $(\boldsymbol{u}_k^I)_{k=1,...,N}$ of size N from the n-dimensional independence copula, where $\boldsymbol{u}_k^I := (u_{k1}^I, ..., u_{kn}^I)'$, $k = 1, ..., N$.

(ii) Generate samples $(u_{k,m_{i-1}+1}^0, ..., u_{k,m_i}^0)_{k=1,...,N}$, $i = 1, ..., n_1$, of size N from the cluster copulas $C_1, ..., C_{n_1}$.

(iii) Set $p_j(k) = \sum_{\ell=1}^N 1_{\{u_{\ell j}^0 \leq u_{kj}^0\}} \in \{1, ..., N\}$, the rank of u_{kj}^0 among $(u_{1j}^0, ..., u_{Nj}^0)'$, for $k = 1, ..., N$ and $j = 1, ..., n$.

(iv) Set $u_{p_j(k),j}^C = u_{(k),j}^I$ for $k = 1, ..., N$ and $j = 1, ..., n$, where $u_{(k),j}^I$ is the kth order statistic of $(u_{1j}^I, ..., u_{Nj}^I)'$ such that $u_{(1),j}^I \leq u_{(2),j}^I \leq ... \leq u_{(N),j}^I$.

(v) Set $v_{ki}^I = K_i(C_i(u_{k,m_{i-1}+1}^C, ..., u_{k,m_i}^C))$ for $k = 1, ..., N$ and $i = 1, ..., n_1$.

(vi) Generate a sample $(\boldsymbol{v}_k^0)_{k=1,...,N}$ of size N from the nesting copula C_0, where $\boldsymbol{v}_k^0 := (v_{k1}^0, ..., v_{kn_1}^0)'$, $k = 1, ..., N$.

(vii) Set $q_i(k) = \sum_{\ell=1}^N 1_{\{v_{\ell i}^0 \leq v_{ki}^0\}} \in \{1, ..., N\}$, the rank of v_{ki}^0 among $(v_{1i}^0, ..., v_{Ni}^0)'$, and $r_i(k) = \sum_{\ell=1}^N 1_{\{v_{\ell i}^I \leq v_{ki}^I\}} \in \{1, ..., N\}$, the rank of v_{ki}^I among $(v_{1i}^I, ..., v_{Ni}^I)'$, for $k = 1, ..., N$ and $i = 1, ..., n_1$.

(viii) Return the observations $u_{q_{i(j)}(k),j} = u_{r_{i(j)}(k),j}^I$, $k = 1, ..., N$, where $i(j)$ is the cluster of variable $j \in \{1, ..., n\}$.

In Steps (i)–(iv) a sample from each cluster copula $C_i,\ i = 1, ..., n_1$, is generated using the method by Iman and Conover (1982). Clearly, Steps (i), (iii) and (iv) are redundant given Step (ii), but we include them, since they show how other margins than uniform can be used instead. The final sample from the hierarchical Kendall copula is obtained by reordering the original independent sample according to the independent aggregated clusters in Step (v) and the dependent sample from the nesting copula C_0 in Step (vi). This means that the sample is reordered twice: first according to the cluster copulas, and second according to the nesting copula. We illustrate the approach in an example.

Example 4.13 (Sample reordering)**.** As in Example 3.6, let $C_\mathcal{K}$ be a four-dimensional hierarchical Kendall copula with $d_1 = d_2 = 2$, Clayton and Gumbel cluster copulas with parameters $\theta_1 = 1.33$ and $\theta_2 = 1.67$, respectively, and a Frank nesting copula with parameter $\theta_0 = 11.41$. To illustrate the sample reordering method of Algorithm 4.12, we set $N = 4$ and assume that our sample has already been reordered according to the cluster copulas C_1 and C_2, so that $(\boldsymbol{u}_k^I)_{k=1,...,4} = (\boldsymbol{u}_k^C)_{k=1,...,4}$. Let the sample be given by

$$\begin{aligned}
\boldsymbol{u}_1^I &= \boldsymbol{u}_1^C = (0.2, 0.5, 0.6, 0.2)', \\
\boldsymbol{u}_2^I &= \boldsymbol{u}_2^C = (0.4, 0.1, 0.2, 0.3)', \\
\boldsymbol{u}_3^I &= \boldsymbol{u}_3^C = (0.3, 0.6, 0.1, 0.6)', \\
\boldsymbol{u}_4^I &= \boldsymbol{u}_4^C = (0.9, 0.7, 0.8, 0.5)'.
\end{aligned}$$

This yields the aggregated values

$$\begin{aligned}
v_{11}^I &= K_1(C_1(0.2, 0.5)) = 0.30, & v_{12}^I &= K_2(C_2(0.6, 0.2)) = 0.36, \\
v_{21}^I &= K_1(C_1(0.4, 0.1)) = 0.16, & v_{22}^I &= K_2(C_2(0.2, 0.3)) = 0.27, \\
v_{31}^I &= K_1(C_1(0.3, 0.6)) = 0.43, & v_{32}^I &= K_2(C_2(0.1, 0.6)) = 0.22, \\
v_{41}^I &= K_1(C_1(0.9, 0.7)) = 0.87, & v_{42}^I &= K_2(C_2(0.8, 0.5)) = 0.68,
\end{aligned}$$

from which we obtain the ranks $r_i(k),\ k = 1, ..., 4,\ i = 1, 2$, as

$$\begin{aligned}
r_1(1) &= 2, & r_2(1) &= 3, \\
r_1(2) &= 1, & r_2(2) &= 2, \\
r_1(3) &= 3, & r_2(3) &= 1, \\
r_1(4) &= 4, & r_2(4) &= 4.
\end{aligned}$$

In the next step, we generate a sample $(\boldsymbol{v}_k^0)_{k=1,...,4}$ from the nesting copula C_0. Let this sample and the corresponding ranks $q_i(k),\ k = 1, ..., 4,\ i = 1, 2$, be given by

$$\begin{aligned}
\boldsymbol{v}_1^0 &= (0.1, 0.4)', & q_1(1) &= 1, & q_2(1) &= 2, \\
\boldsymbol{v}_2^0 &= (0.3, 0.7)', & q_1(2) &= 2, & q_2(2) &= 4, \\
\boldsymbol{v}_3^0 &= (0.7, 0.2)', & q_1(3) &= 3, & q_2(3) &= 1, \\
\boldsymbol{v}_4^0 &= (0.8, 0.5)', & q_1(4) &= 4, & q_2(4) &= 3.
\end{aligned}$$

These ranks are then used to reorder the sample $(\boldsymbol{u}_k^I)_{k=1,\dots,4}$, such that between-cluster dependence is respected. We calculate

$$\begin{aligned}
u_{11} &= u_{q_1(1),1} = u^I_{r_1(1),1} = u^I_{21} = 0.4,\\
u_{12} &= u_{q_1(1),2} = u^I_{r_1(1),2} = u^I_{22} = 0.1,\\
u_{13} &= u_{q_2(3),3} = u^I_{r_2(3),3} = u^I_{13} = 0.6,\\
u_{14} &= u_{q_2(3),4} = u^I_{r_2(3),4} = u^I_{14} = 0.2,
\end{aligned}$$

and similarly for $k = 2, 3, 4$. This yields

$$\begin{aligned}
\boldsymbol{u}_1 &= (0.4, 0.1, 0.6, 0.2)',\\
\boldsymbol{u}_2 &= (0.2, 0.5, 0.1, 0.6)',\\
\boldsymbol{u}_3 &= (0.3, 0.6, 0.8, 0.5)',\\
\boldsymbol{u}_4 &= (0.9, 0.7, 0.2, 0.3)',
\end{aligned}$$

which constitutes the final sample. □

Strong uniform consistency of the sample reordering method as $N \to \infty$ was recently shown under certain regularity conditions on the aggregation function and on the copulas by Mainik (2012), who also gives convergence rates. These conditions are satisfied by any component-wise non-decreasing aggregation function and by any copula with bounded density. Aggregation using the copula and the Kendall distribution function is component-wise non-decreasing, since copulas are multivariate distribution functions and Kendall distribution functions are non-decreasing. However, most common copulas such as the Gaussian, the Student's t or the Clayton have unbounded density. In Mainik (2012) convergence is shown for the Gaussian and the Clayton copula but remains an open question for other families. We complement the results by showing that convergence also holds for the bivariate Joe copula (see Example 2.18).

Remark 4.14 (Regularity conditions for the bivariate Joe copula)**.** To show convergence for the absolutely continuous Joe copula $C(\cdot,\cdot;\theta)$ with $\theta > 1$, we need to verify the regularity conditions (15) and (16) of Mainik (2012). Condition (16) requires that $\int_0^{1/2} \sqrt{\log K(\varepsilon^2)}\, d\varepsilon \le \infty$, where $K(\varepsilon) = \operatorname{esssup}\{c(u_1,u_2;\theta) :\ (u_1,u_2)' \in [\varepsilon, 1-\varepsilon]^2\}$ and $c(\cdot,\cdot;\theta)$ is the density of the bivariate Joe copula (see Equation (2.40)):

$$\begin{aligned}
c(u_1,u_2;\theta) = {} & \left((1-u_1)^\theta + (1-u_2)^\theta - (1-u_1)^\theta(1-u_2)^\theta\right)^{1/\theta-2} && (4.8)\\
& \times (1-u_1)^{\theta-1}(1-u_2)^{\theta-1} && (4.9)\\
& \times \left(\theta - 1 + (1-u_1)^\theta + (1-u_2)^\theta - (1-u_1)^\theta(1-u_2)^\theta\right). && (4.10)
\end{aligned}$$

According to Remark 4.1(a) of Mainik (2012), condition (16) is in particular satisfied if $K(\varepsilon)$ is polynomial. This holds, since the term in (4.10) is bounded by θ, $(1-u_1)^{\theta-1}$ and $(1-u_2)^{\theta-1}$ in (4.9) are bounded by 1, and the term in (4.8) can be rewritten as

$$\left(\left(\frac{1-u_1}{1-u_2}\right)^\theta + \left(\frac{1-u_2}{1-u_1}\right)^\theta - 1\right)^{1/\theta-2} ((1-u_1)(1-u_2))^{1-2\theta} \le ((1-u_1)(1-u_2))^{1-2\theta},$$

which is a polynomial.

Condition (15) is more technical and can be verified exactly along the lines of Proposition 4.2 in Mainik (2012). The proof requires the derivation of the maximum of $c(\cdot, u_2; \theta)$ for fixed $u_2 \in (0,1)$. It is reached at

$$u_1^* = u_1^*(u_2) = \max\left\{1 - \left(\frac{(1-u_2)^\theta(1-(1-u_2)^\theta-\theta)}{((1-u_2)^\theta-1)(\theta-(1-u_2)^\theta)}\right)^{1/\theta}, 0\right\}.$$

These results show that it is reasonable to use sample reordering for the bivariate Joe copula, as we do it in the simulation study in Section 4.4. □

4.3.2 Density resampling

In contrast to the bottom-up method presented in the previous section, this method does not modify the observations per variable but leaves them as they are. For copulas with available density, the density resampling approach has been proposed by Kurowicka and Cooke (2006, Section 6.4.3). The idea is to resample from a large number of independent uniform observations according to probabilities proportional to the density evaluated at the observations. Sampling efficiency can be increased if a good proposal distribution is known, which approximates the distribution of interest and from which it is easy to simulate. This has been named the sampling/importance resampling method by Rubin (1987, 1988). Here, we formulate the density resampling method for hierarchical Kendall copulas with density given in Equation (3.2) in terms of a general proposal distribution, which is not necessarily the multivariate independence copula.

Algorithm 4.15 (Density resampling of hierarchical Kendall copulas)**.** Let $C_{\mathcal{K}}$ be a hierarchical Kendall copula with density $c_{\mathcal{K}}$ and let $\widetilde{C}$ be the proposal copula with density $\widetilde{c}$. To generate a sample of size N, let $\widetilde{N} \gg N$.

(i) Generate a sample $(\boldsymbol{w}_k)_{k=1,\dots,\widetilde{N}}$ of size $\widetilde{N}$ from the proposal copula $\widetilde{C}$.

(ii) Resample N times from $(\boldsymbol{w}_k)_{k=1,\dots,\widetilde{N}}$ according to probabilities proportional to the importance ratios,

$$r_k := \frac{c_{\mathcal{K}}(\boldsymbol{w}_k)}{\widetilde{c}(\boldsymbol{w}_k)}, \quad k = 1, \dots, \widetilde{N}. \tag{4.11}$$

That is, draw a sample Λ of size N without replacement from $\{1, \dots, \widetilde{N}\}$ according to probabilities $p_k = r_k / \sum_{\ell=1}^{\widetilde{N}} r_\ell$, $k = 1, \dots, \widetilde{N}$.

(iii) Return the resulting sample $(\boldsymbol{u}_k)_{k=1,\dots,N} := (\boldsymbol{w}_k)_{k\in\Lambda}$.

If the proposal copula is not well-chosen, this is obviously rather inefficient, since $\widetilde{N}$ then should be chosen significantly larger than N to generate a reasonably good sample. In the simulation study in Section 4.4, we therefore compare density resampling with two different proposal copulas: the multivariate independence copula and the multivariate Gaussian copula with appropriately chosen correlation matrix (see Example 2.13). Sampling from a Gaussian copula as well as the evaluation of its density (2.20) are straightforward and

can be implemented efficiently, so that samples and the importance ratios (4.11) can be obtained conveniently. The Gaussian copula also allows for heterogeneous pairwise dependence, which is needed to appropriately capture within- and between-cluster dependence of a hierarchical Kendall copula.

But even if no reasonable proposal is available, density resampling with independence copula proposal may be considered as an alternative to rejection-like sampling (Section 4.2.4) and also to sample reordering (Section 4.3.1), which may also require large samples to generate an accurate sample. It is however important that the evaluation of the density $c_{\mathcal{K}}$ is computationally fast.

The finite sample performance and the computing time of all three approximate sampling approaches are considered in a numerical study in the next section. Particular focus is put on the choice of $\widetilde{N}$. We compare two different choices of $\widetilde{N}$ relative to N for the two proposal copulas, since no general rules how to adequately choose the ratio $\widetilde{N}/N$ are available for multivariate distributions. Clearly, the poorer the proposal copula $\widetilde{C}$ is chosen, the larger $\widetilde{N}/N$ should be.

4.4 Simulation study

Three of the methods discussed here are approximate: rejection-like sampling (Section 4.2.4), sample reordering (Section 4.3.1) and density resampling (Section 4.3.2). We therefore perform a simulation study in order to assess these approaches to sampling from hierarchical Kendall copulas. In particular, we investigate different choices of ε for rejection-like sampling (see Algorithm 4.10) and of the proposal copula $\widetilde{C}$ as well as of $\widetilde{N}$ for density resampling (see Algorithm 4.15); sample reordering does not require any choice of control parameters.

The sampling procedures are compared based on a four-dimensional hierarchical Kendall copula with bivariate Clayton and Joe cluster copulas and Gaussian nesting copula. The cluster copulas are reflection asymmetric with lower and upper tail dependence, respectively. The Gaussian nesting copula is tail independent. Parameters are chosen according to a Kendall's τ of $\tau_0 = 0.5$ for the Gaussian copula (medium dependence; copula parameter of $\theta_0 = 0.71$) and $\tau_1 = \tau_2 = 0.7$ for the cluster copulas (strong dependence; copula parameters of $\theta_1 = 4.67$ for the Clayton and $\theta_2 = 5.46$ for the Joe copula). In the case of sample reordering, convergence for the Clayton and the Gaussian copula is shown in Mainik (2012) and for the Joe copula in Remark 4.14.

As benchmark for the approximate methods, we generate a large sample $(\boldsymbol{u}_k^0)_{k=1,\dots,N^*}$ ($N^* = 1\,000\,000$) using closed-form conditional inverse sampling for Archimedean copulas (see Algorithms 4.5, 4.7 and 4.8), since both the Clayton and the Joe copulas are Archimedean. As an illustration, a sample of size 1000 of this benchmark data is shown in Figure 4.5.

To compare the different approaches, we generate samples $(\boldsymbol{u}_k)_{k=1,\dots,N}$ of different sizes $N \in \{100, 200, 500, 1000, 2000\}$ and compare different evaluation criteria. The following six criteria are considered based on $R = 1000$ repetitions:

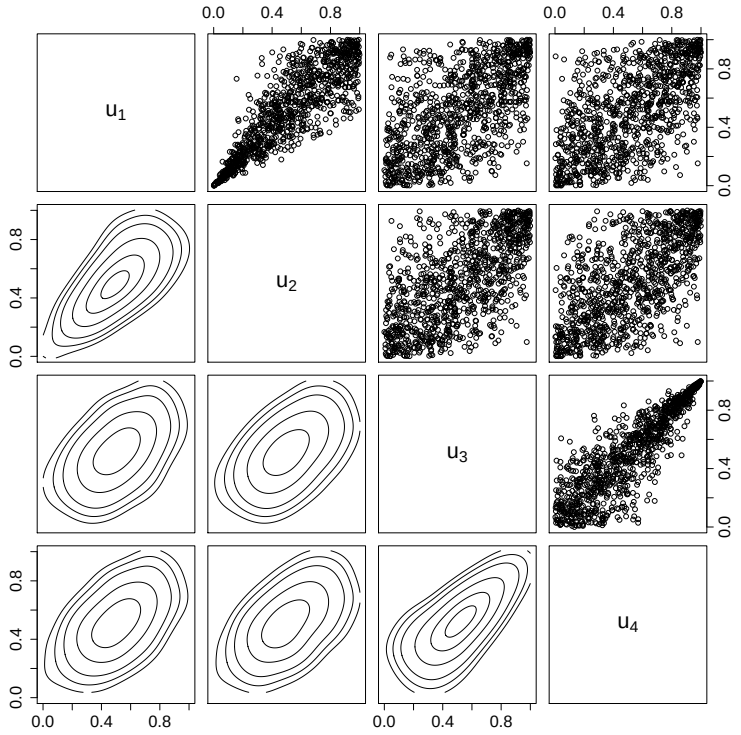


Figure 4.5: A sample of size 1000 from the simulated benchmark data. Variables 1 and 2 form the first cluster with Clayton copula, variables 3 and 4 the second with Joe copula. The lower triangle shows contour lines of the pairwise empirical densities with standard normal margins (the corresponding axes range from -3 to 3).

(i) Mean squared difference of pairwise Kendall's τ values:

$$\frac{1}{R}\sum_{r=1}^{R}\frac{1}{6}\sum_{1\le i<j\le 4}\left(\widehat{\tau}_{ij}^{(r)}-\widehat{\tau}_{ij}^{0}\right)^2,$$

where $\widehat{\tau}_{ij}^{(r)}$ is the empirical Kendall's τ of the variables i and j in repetition r and $\widehat{\tau}_{ij}^{0}$ that of the benchmark data.

(ii) Mean squared difference of pairwise lower-tail Kendall's τ values:

$$\frac{1}{R}\sum_{r=1}^{R}\frac{1}{6}\sum_{1\le i<j\le 4}\left(\widehat{\tau}_{ij,L}^{(r)}-\widehat{\tau}_{ij,L}^{0}\right)^2,$$

where $\widehat{\tau}_{ij,L}^{(r)}$ is the empirical 20% lower-tail Kendall's τ of the variables i and j in repetition r and $\widehat{\tau}_{ij,L}^{0}$ that of the benchmark data. The 20% lower-tail Kendall's τ is given for two uniform random variables U_1 and U_2 as the Kendall's τ of $(U_1, U_2 | U_1 < 0.2, U_2 < 0.2)$. It is a rank-based version of the exceedance correlation coefficient

by Longin and Solnik (2001) and Ang and Chen (2002) and serves as a measure of lower tail behavior.

(iii) Mean squared difference of pairwise upper-tail Kendall's τ values:

$$\frac{1}{R}\sum_{r=1}^{R}\frac{1}{6}\sum_{1\leq i<j\leq 4}\left(\widehat{\tau}_{ij,U}^{(r)}-\widehat{\tau}_{ij,U}^{0}\right)^2,$$

where $\widehat{\tau}_{ij,L}^{(r)}$ is the empirical 20% upper-tail Kendall's τ of the variables i and j in repetition r and $\widehat{\tau}_{ij,L}^{0}$ that of the benchmark data. The 20% upper-tail Kendall's τ is defined in an analogous way to the lower-tail Kendall's τ and measures upper tail behavior.

(iv) Mean squared difference of empirical copulas:

$$\frac{1}{R}\sum_{r=1}^{R}\frac{1}{|\Delta|^4}\sum_{\boldsymbol{v}\in\Delta^4}\left(\widehat{C}^{(r)}(\boldsymbol{v})-\widehat{C}^{0}(\boldsymbol{v})\right)^2,$$

where $\widehat{C}^{(r)}$ is the empirical copula of the sample $(\boldsymbol{u}_k^{(r)})_{k=1,\dots,N}$ in the rth repetition and $\widehat{C}^0$ that of the benchmark data. Further, Δ is an equispaced partition of $[0,1]$ of size $|\Delta| = 25$. Hence, in each repetition the empirical copulas are evaluated at $|\Delta|^4 = 25^4 = 390\,625$ points of $[0,1]^4$.

(v) Mean squared difference of log likelihoods:

$$\frac{1}{R}\sum_{r=1}^{R}\left(\frac{1}{N}\sum_{k=1}^{N}\log c_{\mathcal{K}}\big(\boldsymbol{u}_k^{(r)}\big)-\frac{1}{N^*}\sum_{k=1}^{N^*}\log c_{\mathcal{K}}\big(\boldsymbol{u}_k^{0}\big)\right)^2,$$

where $\boldsymbol{u}_k^{(r)}$ denotes the kth observations in the rth repetition and $c_{\mathcal{K}}$ is the density of the hierarchical Kendall copula (see Equations (3.2) and (3.15)) evaluated at the true parameters. The log likelihoods are standardized by the sample sizes to allow for comparison.

(vi) Mean squared difference of parameters:

$$\frac{1}{R}\sum_{r=1}^{R}\frac{1}{3}\sum_{i\in\{0,1,2\}}\left(\tau_i(\widehat{\theta}_i)-\tau_i\right)^2,$$

where $\widehat{\theta}_i$ are maximum likelihood estimates of $\theta_i,\ i=0,1,2$, which are transformed to Kendall's τ values using the respective relationships $\tau_i(\cdot)$ implied by the copulas (see Table 2.1).

While the first three criteria focus on data characteristics, namely the level of general dependence as well as the lower and upper tail behavior, the fourth criterion directly

compares the empirical copulas and the last two criteria investigate the effects on estimation and model selection.

Based on these six criteria we evaluate and compare the three approximate procedures. For rejection-like sampling we choose $\varepsilon \in \{10^{-2}, 10^{-3}, 10^{-4}\}$, while for density resampling we choose $\widetilde{N}$ as a multiple of N, namely $\widetilde{N}/N \in \{100, 1000\}$. The proposal copula is either chosen as the independence copula or as the Gaussian copula, where the entries of the correlation matrix are chosen according to the Spearman's ρ_S of the within- and between-cluster dependence parameters (see Equation (2.22)). Using Corollary 3.12, the Spearman's ρ_S of the variables 1 and 3, which are in different clusters, is given by

$$\begin{aligned}\rho_{S,13} &= 12 \int_{[0,1]^2} u_1 u_3 \, dC_{\mathcal{K},13}(u_1, u_3) - 3 \\ &= 12 \int_{[0,1]^4} u_1 u_3 \, c_0(K_1(C_1(u_1, u_2)), K_2(C_2(u_3, u_4))) \, c_1(u_1, u_2) \, c_2(u_3, u_4) \, d\boldsymbol{u} - 3,\end{aligned}$$

which can be solved numerically.

Simulation results for all three methods and all six criteria are shown in Table 4.1 and illustrated in Figure 4.6. In addition, we report the mean computing time in seconds (on a 2.6Ghz AMD Opteron). Standard errors of the evaluation criteria are shown in Table 4.2 and illustrated in Figure 4.7.

With respect to the data characteristics of Kendall's τ as well as lower- and upper-tail Kendall's τ, rejection-like sampling and sample reordering give very similar results. Most interestingly, the choice of the error ε in rejection-like sampling has only little influence on the results. The choice of $\widetilde{N}$ in relation to the sample size N for density resampling however strongly influences the performance. Only if $\widetilde{N}/N = 1000$, the results are similarly good as for the other two methods. The two proposal copulas lead to different results with respect to lower- and upper-tail Kendall's τ, where the Gaussian proposal copula yields much better results, which are even similar to those of the other methods. This is not the case for the independence proposal copula, which does not produce good results in the tails. On the whole, it should be noted that the general dependence in terms of Kendall's τ is better approximated than tail behavior as measured by the lower- and upper-tail Kendall's τ. For larger sample sizes, the differences are however quite small, so that we can conclude that simulation using rejection-like sampling, sample reordering and density resampling with Gaussian proposal copula appropriately reproduces data characteristics.

Results are similar for the empirical copula. The criterion and the corresponding standard error are minimized when sample reordering is used, but also rejection-like sampling and density resampling with Gaussian proposal copula work quite well. Density resampling with independence proposal only yields reasonable results if $\widetilde{N}/N = 1000$, while choosing $\widetilde{N}/N = 100$ is clearly too small. For rejection-like sampling, the performance does not depend on the choice of ε.

In terms of log likelihoods and copula parameters, the results provide additional insights which allow to better discriminate among the methods. While rejection-like sampling does very well and shows little dependence on ε (there is a notable difference between the choices of ε only for $N = 100$ and the log likelihood difference), sample reordering strongly suffers from small sample sizes (in particular $N \leq 1000$), where convergence of

	Rejection-like sampling			Reor-	Resamp. (indep.)		Resamp. (Gauss.)	
N	$\varepsilon=10^{-2}$	$\varepsilon=10^{-3}$	$\varepsilon=10^{-4}$	dering	$\frac{\widetilde{N}}{N}=10^2$	$\frac{\widetilde{N}}{N}=10^3$	$\frac{\widetilde{N}}{N}=10^2$	$\frac{\widetilde{N}}{N}=10^3$
Mean squared difference of pairwise Kendall's τ values ($\times 10^3$)								
100	2.362	2.263	2.424	2.395	6.963	2.699	6.963	2.699
200	1.159	1.160	1.126	1.124	6.103	1.609	6.103	1.609
500	0.430	0.431	0.421	0.430	5.248	0.922	5.248	0.922
1000	0.234	0.233	0.232	0.216	4.863	0.663	4.863	0.663
2000	0.112	0.113	0.112	0.107	4.735	0.586	4.735	0.586
Mean squared difference of pairwise lower-tail Kendall's τ values ($\times 10^3$)								
100	55.711	55.805	56.455	57.177	110.640	68.250	59.538	57.619
200	21.198	21.168	22.788	20.847	44.197	27.489	22.674	21.684
500	7.409	7.591	7.622	7.233	20.077	11.244	7.946	7.937
1000	3.712	3.766	3.772	3.536	14.252	6.491	4.038	3.776
2000	1.779	1.821	1.779	1.782	11.765	4.401	2.058	1.927
Mean squared difference of pairwise upper-tail Kendall's τ values ($\times 10^3$)								
100	56.577	56.047	56.638	55.059	136.183	70.978	55.553	55.870
200	21.272	21.743	21.258	21.643	51.750	29.722	20.526	21.352
500	7.445	7.718	7.616	7.652	24.475	13.084	7.787	7.749
1000	3.393	3.550	3.623	3.595	17.682	8.075	3.663	3.603
2000	1.706	1.770	1.783	1.694	14.680	6.172	1.873	1.778
Mean squared difference of empirical copulas ($\times 10^3$)								
100	1.067	1.062	1.069	0.838	1.380	1.057	1.018	1.053
200	0.524	0.521	0.524	0.427	0.987	0.539	0.547	0.541
500	0.212	0.212	0.212	0.171	0.692	0.263	0.230	0.205
1000	0.102	0.103	0.102	0.085	0.577	0.148	0.119	0.106
2000	0.051	0.051	0.051	0.042	0.541	0.105	0.068	0.056
Mean squared difference of log likelihoods ($\times 10^3$)								
100	36.527	36.780	37.230	528.955	184.999	51.703	33.689	38.533
200	18.918	19.884	17.612	194.076	181.721	42.193	17.949	17.895
500	7.319	7.089	7.217	45.092	177.304	32.627	6.968	7.627
1000	3.749	3.745	3.667	14.724	171.343	29.305	3.653	3.527
2000	1.832	1.852	1.802	4.570	170.900	28.416	1.985	1.972
Mean squared difference of parameters ($\times 10^3$)								
100	2.174	2.006	1.847	8.086	5.377	4.109	1.764	1.923
200	0.510	0.475	0.514	3.292	2.694	1.129	0.618	0.467
500	0.164	0.158	0.165	0.924	0.597	0.190	0.166	0.156
1000	0.086	0.081	0.083	0.356	0.339	0.105	0.085	0.077
2000	0.041	0.040	0.041	0.137	0.303	0.069	0.047	0.043
Mean computing time (in seconds)								
100	0.598	4.258	40.289	0.007	0.140	1.437	0.212	2.055
200	0.975	6.035	52.812	0.008	0.287	3.072	0.449	4.180
500	2.000	9.335	83.587	0.013	0.754	10.520	1.062	13.143
1000	4.073	15.002	129.981	0.023	1.680	27.464	2.255	32.301
2000	9.168	27.645	210.657	0.040	4.224	76.565	5.277	84.362

Table 4.1: Evaluation criteria.

	Rejection-like sampling			Reor-	Resamp. (indep.)		Resamp. (Gauss.)	
N	$\varepsilon = 10^{-2}$	$\varepsilon = 10^{-3}$	$\varepsilon = 10^{-4}$	dering	$\frac{\widetilde{N}}{N} = 10^2$	$\frac{\widetilde{N}}{N} = 10^3$	$\frac{\widetilde{N}}{N} = 10^2$	$\frac{\widetilde{N}}{N} = 10^3$
Mean squared difference of pairwise Kendall's τ values ($\times 10^3$)								
100	2.196	2.135	2.365	2.306	5.964	2.917	5.964	2.917
200	1.111	1.093	1.150	1.034	4.119	1.865	4.119	1.865
500	0.399	0.412	0.390	0.392	2.464	0.921	2.464	0.921
1000	0.238	0.235	0.239	0.189	1.663	0.510	1.663	0.510
2000	0.106	0.102	0.108	0.094	1.189	0.379	1.189	0.379
Mean squared difference of pairwise lower-tail Kendall's τ values ($\times 10^3$)								
100	49.193	50.027	49.114	45.366	101.992	61.745	52.074	54.149
200	17.180	17.548	20.126	15.309	32.135	22.174	17.867	17.293
500	5.559	5.579	5.869	5.364	13.039	8.165	5.848	6.379
1000	2.846	2.929	2.822	2.656	8.496	4.507	2.983	2.909
2000	1.327	1.324	1.342	1.314	5.162	2.826	1.503	1.413
Mean squared difference of pairwise upper-tail Kendall's τ values ($\times 10^3$)								
100	54.669	57.262	56.159	43.434	122.226	65.793	49.342	51.126
200	17.869	16.943	16.899	16.714	37.416	23.327	16.704	16.249
500	5.640	5.834	5.616	5.581	15.778	9.753	5.805	5.849
1000	2.624	2.783	2.814	2.636	9.635	5.544	2.710	2.531
2000	1.294	1.384	1.339	1.244	5.973	3.690	1.365	1.410
Mean squared difference of empirical copulas ($\times 10^3$)								
100	0.861	0.830	0.851	0.388	0.896	0.803	0.786	0.790
200	0.433	0.412	0.422	0.203	0.647	0.379	0.446	0.427
500	0.191	0.186	0.190	0.080	0.353	0.181	0.180	0.158
1000	0.075	0.074	0.074	0.039	0.235	0.103	0.093	0.079
2000	0.040	0.039	0.039	0.019	0.179	0.067	0.053	0.044
Mean squared difference of log likelihoods ($\times 10^3$)								
100	54.395	54.627	53.474	469.152	119.981	66.208	49.007	53.798
200	26.886	29.732	26.105	186.001	86.097	44.472	25.372	24.996
500	9.904	9.939	10.479	45.945	56.774	25.852	11.055	10.618
1000	5.309	5.772	5.294	15.609	39.484	17.332	5.106	5.204
2000	2.542	2.507	2.472	5.262	27.097	12.605	3.145	2.757
Mean squared difference of parameters ($\times 10^3$)								
100	6.104	5.822	5.442	6.146	10.188	9.042	5.143	5.522
200	1.605	1.300	1.792	2.759	7.259	4.406	2.386	1.317
500	0.157	0.145	0.149	0.780	2.070	0.196	0.164	0.140
1000	0.087	0.080	0.083	0.289	0.237	0.099	0.074	0.069
2000	0.038	0.038	0.039	0.112	0.163	0.065	0.040	0.038

Table 4.2: Standard errors of the evaluation criteria reported in Table 4.1.

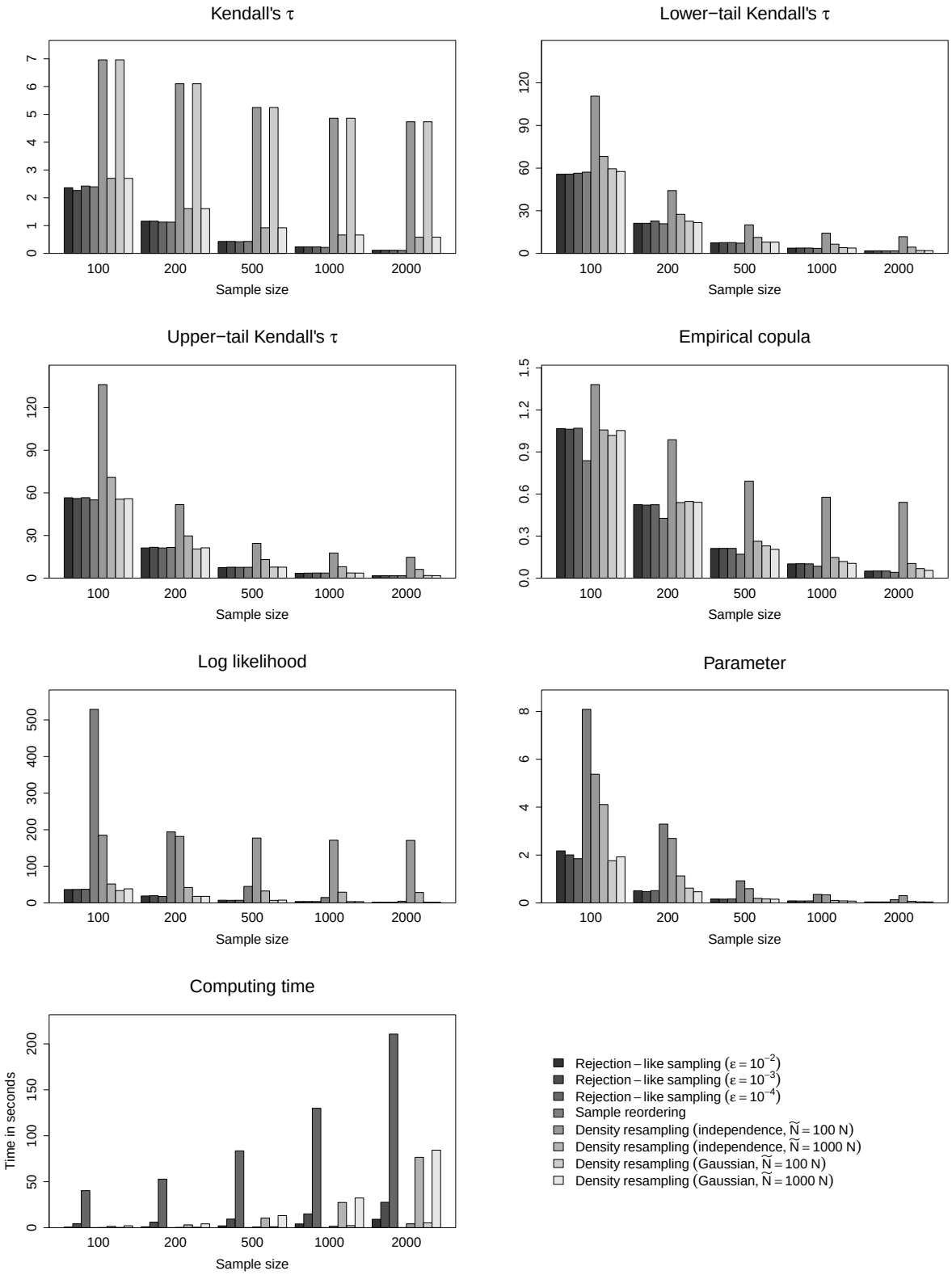


Figure 4.6: Illustration of the evaluation criteria reported in Table 4.1.

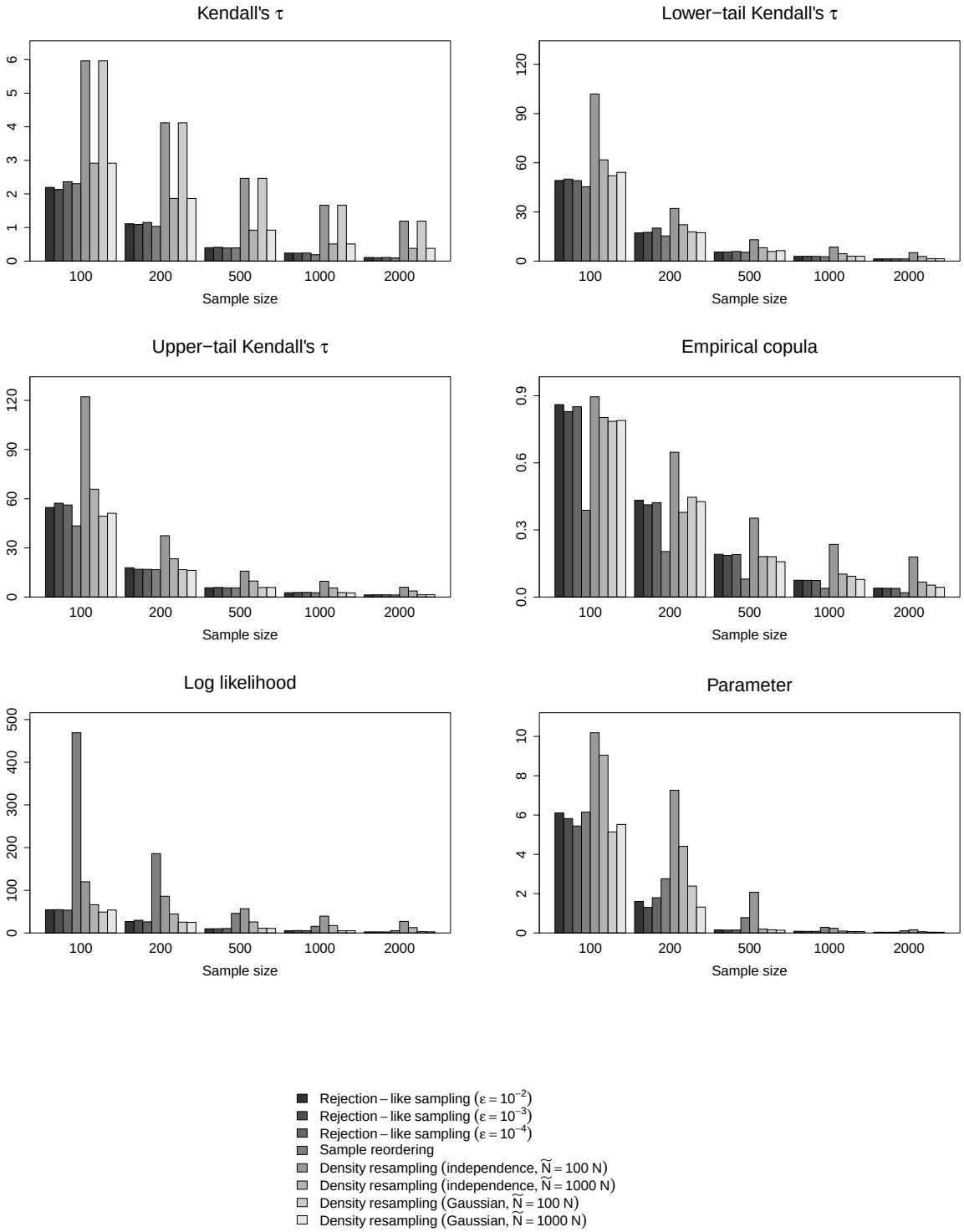


Figure 4.7: Illustration of the standard errors of the evaluation criteria reported in Table 4.1 (see Table 4.2).

the method cannot be assumed. This is reflected in large standard errors as well. Density resampling with independence proposal also performs rather poorly in terms of the log likelihood but does not suffer as heavily from small sample sizes as sample reordering. Once again, the results clearly show that $\widetilde{N}/N = 100$ is not appropriate. This is different for the Gaussian copula proposal, which also works well for $\widetilde{N}/N = 100$.

Finally, a look at the computing times reveals that rejection-like sampling with very small ε is inefficient and also density resampling with $\widetilde{N}/N = 1000$ is quite time-consuming. Sample reordering, on the other hand, is computationally very efficient for any sample size. This is due to the fact that no spare samples need to be generated, from which the final sample is selected, as in the other two methods. Rejection-like sampling with $\varepsilon = 10^{-2}$ also has reasonable computing times even for larger sample sizes.

In summary, rejection-like sampling appears to work best and already a choice of $\varepsilon = 10^{-2}$ seems to lead to a very good approximation. As the computing time strongly depends on ε, these results are very beneficial for the method of rejection-like sampling; a choice of $\varepsilon = 10^{-2}$ requires only little computing time. For density resampling we compared two choices of $\widetilde{N}/N$, which also determine the computing time. When an independence proposal is used, only $\widetilde{N}/N = 1000$ gave satisfactory results, yet indicating that choices of $\widetilde{N}/N > 1000$ may be necessary. But this would require an excessive computing time, even when the density of the hierarchical Kendall copula is efficient to evaluate, as it is the case here. Overall, based on the results of this simulation study, we cannot recommend the use of density resampling with independence proposal for hierarchical Kendall copulas. An appropriately chosen Gaussian copula proposal however showed a good performance and can also be used with $\widetilde{N}/N = 100$. Furthermore, sample reordering also proved to be a valid alternative to rejection-like sampling but only if sample sizes are sufficiently large (at least $N \geq 2000$), so that convergence of the method can be assumed. This is for example the case when risk capital figures need to be simulated in finance and insurance, as discussed in Arbenz et al. (2012) and in Section 4.5 below. The generation of such large samples using sample reordering is very time-efficient.

If a closed-form solution as for Archimedean (Section 4.2.1), Archimax (Section 4.2.2) or Plackett copulas (Section 4.2.3) is not available for the simulation of a hierarchical Kendall copula, we therefore recommend to use either top-down rejection-like sampling or, for sufficiently large sample sizes, bottom-up sample reordering. If the Gaussian copula can be calibrated adequately as proposal copula, also density resampling may be used.

4.5 Application: Value-at-Risk forecasting of stock portfolios

In Section 3.7 we analyzed the dependence among stock returns of the 30 constituents of the most important German stock market index DAX. It turned out that a hierarchical Kendall copula with Frank cluster copulas and Student's t nesting copula quite well fits the data of the training set of 1158 observations. 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 negative returns is continuous, the $(1-\alpha)$-VaR, $\alpha \in [0,1]$, is the $(1-\alpha)$-quantile of the distribution of negative returns or

the α-quantile of the distribution of returns, respectively. For market risk management, this value needs to be predicted on a daily basis. First, we describe how to carry this out. Then, the models of Section 3.7 are evaluated using so-called VaR backtests.

To start with, we select a rolling window size N (e.g., four years), a forecast period of length M and portfolio weights $\omega_j,\ j = 1, ..., 30$, with $\sum_{j=1}^{30} \omega_j = 1$ (the long-only constraint $\omega_j \geq 0$ is not required here). For each window $\{m, ..., N+m-1\},\ m = 1, ..., M$, we then proceed as follows to forecast the one day ahead Value-at-Risk at level $1-\alpha$ using a copula dependence model with GARCH(1,1) margins (see also Berg and Aas (2009), Nikoloulopoulos et al. (2012) and Brechmann and Czado (2013)).

(i) We specify GARCH(1,1) models with Student's t error distribution for the marginal time series $r_{tj},\ t = m, ..., N+m-1,\ j = 1, ..., 30$, (see Equation (3.19)) and compute standardized residuals $\widehat{z}_{tj}$, as described in Section 3.7.2.

(ii) Using the estimated parameters, we compute the ex-ante GARCH variance forecast for $j = 1, ..., 30$,

$$\widehat{\sigma}^2_{N+m,j} = \widehat{\omega}_j + \widehat{\alpha}_j \widehat{\sigma}^2_{N+m-1,j} \widehat{z}^2_{N+m-1,j} + \widehat{\beta}_j \widehat{\sigma}^2_{N+m-1,j}. \tag{4.12}$$

(iii) We fit a copula to the transformed standardized residuals $\widehat{u}_{tj} := \widetilde{T}_{\widehat{\nu}_j}(\widehat{z}_{tj}),\ t = m, ..., N+m-1,\ j = 1, ..., 30$, where $\widetilde{T}_\nu$ denotes the distribution function of the univariate standardized Student's t distribution with ν degrees of freedom (see Section 3.5 for the estimation of hierarchical Kendall copulas).

(iv) We repeat the following 10 000 times:

a) Observations $(\widehat{u}_{N+m,1}, ..., \widehat{u}_{N+m,30})'$ are generated from the fitted copula as described in this chapter.

b) We set $\widehat{z}_{N+m,j} = \widetilde{T}^{-1}_{\widehat{\nu}_j}(\widehat{u}_{N+m,j})$ for $j = 1, ..., 30$.

c) The sampled observations and the ex-ante GARCH variance forecasts (4.12) are used to compute the ex-ante return forecast $\widehat{r}_{N+m,j}$ for $j = 1, ..., 30$,

$$\widehat{r}_{N+m,j} = \widehat{\mu}_j + \widehat{\sigma}_{N+m,j} \widehat{z}_{N+m,j}.$$

d) The portfolio return forecast is then given by $\widehat{r}_{N+m,P} = \sum_{j=1}^{30} \omega_j \widehat{r}_{N+m,j}$.

(v) Finally, we compute the $(1-\alpha)$-VaR of the portfolio return, $\text{VaR}_{N+m,1-\alpha}(-\widehat{r}_{N+m,P})$, by taking the sample quantile at level α of the 10 000 portfolio return forecasts.

In Step (v) any other risk measure such as the expected shortfall could, of course, easily be used instead.

When repeating this procedure on a daily basis, the parameters of the copula and the GARCH margins are re-estimated for each window $\{m, ..., N+m-1\},\ m = 1, ..., M$. Here, we choose a moving window of length $N = 1158$, which is the length of the training set, and forecast the portfolio VaR for the $M = 500$ days of the testing set using the above procedure. In particular, we consider an equally weighted portfolio of the 30 DAX constituents, that is, $\omega_j = 1/30,\ j = 1, ..., 30$.

Resulting forecasts can then be used to assess the prediction accuracy of a model. For this, we consider the hit sequence of ex-post *exceedances* (Christoffersen, 1998),

$$I_t = \begin{cases} 1 & \text{if } -r_{t,P} > \text{VaR}_{t,1-\alpha}(-\widehat{r}_{t,P}), \\ 0 & \text{otherwise,} \end{cases} \qquad t = N+1, ..., N+M, \tag{4.13}$$

where $r_{t,P}$ denotes the ex-post observed portfolio return at time t. This sequence of Bernoulli variables should exhibit two properties if the forecasts are accurate. First, the exceedances should occur independently, that is, not in clusters, and second, the proportion of exceedances should approximately equal the VaR confidence level α ("unconditional coverage"). The term "conditional coverage" encompasses both properties.

In the literature, a wide range of tests, so-called *VaR backtests*, for these properties has been proposed (see, e.g., Campbell (2007) for a review). Since each test exhibits certain advantages and disadvantages and there are no general guidelines of which test to use, we recommend applying a battery of such tests to ensure that results are not biased in one or the other direction. For example, the following tests may be considered.

(i) The proportion of failures test of unconditional coverage by Kupiec (1995) (UC): Under the null hypothesis of unconditional coverage, it holds that the exceedances $I_{N+1}, ..., I_{N+M}$ have expectation α. Assuming independence of the exceedances, we therefore test the null hypothesis

$$H_0 : E(I_t) = \alpha \; \forall t = N+1, ..., N+M,$$

against the alternative that the expectation is different from α. For this, we use the likelihood-ratio test statistic proposed by Christoffersen (1998),

$$LR_{\text{uc}} := -2 \log \left(\frac{(1-\alpha)^{M-M_1} \, \alpha^{M_1}}{(1-\widehat{\alpha})^{M-M_1} \, \widehat{\alpha}^{M_1}} \right), \tag{4.14}$$

where $M_1 = \sum_{t=N+1}^{N+M} I_t$ is the observed number of exceedances and $\widehat{\alpha} = M_1/M$ is the empirical exceedance probability. The test statistic LR_{uc} asymptotically follows a χ^2 distribution with one degree of freedom, so that we reject H_0 at level γ if $LR_{\text{uc}} > F^{-1}_{\chi^2}(1-\gamma; 1)$, where $F^{-1}_{\chi^2}(\cdot; \nu)$ denotes the inverse χ^2 distribution function with ν degrees of freedom

(ii) The Markov test of independence by Christoffersen (1998) (IND1): If exceedances are independent, the probability of observing an exceedance should be independent of whether or not there was an exceedance on the previous day. Christoffersen (1998) proposes to investigate this by considering $I_{N+1}, ..., I_{N+M}$ as a first-order Markov chain with transition matrix

$$\begin{pmatrix} 1-p_{01} & p_{01} \\ 1-p_{11} & p_{11} \end{pmatrix},$$

where $p_{ij} = P(I_t = j | I_{t-1} = i), \; i, j \in \{0, 1\}$. We then test the null hypothesis

$$H_0 : p_{01} = p_{11},$$

against the alternative that the transition probabilities p_{01} and p_{11} are different. If M_{ij} denotes the number of observations with value i followed by j for $i, j \in \{0, 1\}$, the corresponding likelihood-ratio test statistic is given by

$$LR_{\text{ind}} := -2 \log \left(\frac{(1 - \widehat{p}_1)^{M_{00}+M_{10}} \, \widehat{p}_1^{M_{01}+M_{11}}}{(1 - \widehat{p}_{01})^{M_{00}} \, \widehat{p}_{01}^{M_{01}} \, (1 - \widehat{p}_{11})^{M_{10}} \, \widehat{p}_{11}^{M_{11}}} \right), \tag{4.15}$$

where $\widehat{p}_{ij} = M_{ij}/(M_{i0} + M_{i1}),\ i, j \in \{0, 1\}$, and $\widehat{p}_1 = (M_{01} + M_{11})/(M_{00} + M_{10} + M_{01} + M_{11})$ are the empirical transition probabilities. Similar to LR_{uc} defined in Equation (4.14), the test statistic LR_{ind} asymptotically follows a χ^2 distribution with one degree of freedom. We reject H_0 at level γ if $LR_{\text{ind}} > F^{-1}_{\chi^2}(1 - \gamma; 1)$.

(iii) The joint test of conditional coverage by Christoffersen (1998) (CC1): Combining the above two tests yields a test of conditional coverage, which tests the null hypothesis

$$H_0 : I_{N+1}, ..., I_{N+M} \text{ are independent and identically distributed and follow a Bernoulli distribution with probability } \alpha,$$

against the alternative that this is not the case. The appropriate likelihood-ratio statistic is given by

$$LR_{\text{cc}} := LR_{\text{uc}} + LR_{\text{ind}} = -2 \log \left(\frac{(1 - \alpha)^{M - M_1} \, \alpha^{M_1}}{(1 - \widehat{p}_{01})^{M_{00}} \, \widehat{p}_{01}^{M_{01}} \, (1 - \widehat{p}_{11})^{M_{10}} \, \widehat{p}_{11}^{M_{11}}} \right),$$

where LR_{uc} and LR_{ind} are defined in Equations (4.14) and (4.15), respectively. Christoffersen (1998) shows that this test statistic asymptotically follows a χ^2 distribution with two degrees of freedom and we therefore reject H_0 at level γ if $LR_{\text{cc}} > F^{-1}_{\chi^2}(1 - \gamma; 2)$.

(iv) The duration-based mixed Kupiec test of conditional coverage by Haas (2001) (CC2): This test and the following two tests are based on the time between two exceedances, the *duration*. Clearly, the durations should be independent of previous durations and have mean $1/\alpha$. For more details on the tests, we refer to the corresponding references.

(v) The duration-based Weibull test of independence by Christoffersen and Pelletier (2004) (IND2).

(vi) The duration-based GMM test of conditional coverage by Candelon et al. (2011) with orders 2 and 5 (CC3 and CC4, respectively).

For model comparison, we focus on the following models: the independence copula, the Gaussian and the Student's copula, the fully specified R-vine copula (without independence pair copulas) as well as the hierarchical Kendall copula with Frank cluster copulas and Student's t nesting copula, which provided the best fit among the examined hierarchical Kendall copulas in Section 3.7 (see Table 3.2). In addition, we consider a hierarchical Kendall copulas with Gaussian cluster copulas and Student's t nesting copula. In contrast

Copula	Level	Exceed.	UC	IND1	IND2	CC1	CC2	CC3	CC4
Independence	99%	20.6%	*0.00*	*0.04*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
	95%	27.0%	*0.00*	0.15	*0.03*	*0.00*	*0.00*	*0.00*	*0.00*
	90%	31.4%	*0.00*	0.17	*0.03*	*0.00*	*0.00*	*0.00*	*0.00*
Gaussian	99%	1.0%	1.00	0.75	0.79	0.95	0.35	0.93	1.00
	95%	5.2%	0.84	0.09	0.74	0.23	0.35	0.98	1.00
	90%	10.4%	0.77	0.46	0.87	0.73	*0.03*	0.29	0.73
Student's t	99%	0.8%	0.64	0.80	0.21	0.87	0.93	0.75	0.95
	95%	4.8%	0.84	0.12	0.73	0.29	0.39	0.89	1.00
	90%	10.6%	0.66	0.53	0.85	0.74	0.05	0.31	0.73
R-vine	99%	1.0%	1.00	0.75	0.81	0.95	0.35	0.93	1.00
	95%	5.4%	0.69	0.08	0.70	0.20	0.35	0.95	1.00
	90%	10.6%	0.66	0.53	0.87	0.74	*0.04*	0.24	0.66
Hier. Kendall	99%	0.4%	0.13	0.90	0.46	0.31	0.35	0.43	0.82
(Frank, Student's t)	95%	6.2%	0.23	*0.04*	0.76	0.06	0.21	0.55	0.92
	90%	11.4%	0.31	0.52	0.81	0.48	*0.03*	0.22	0.64
Hier. Kendall	99%	1.0%	1.00	0.75	0.77	0.95	0.35	0.93	1.00
(Gaussian, Student's t)	95%	5.4%	0.69	0.08	0.75	0.20	0.31	0.97	1.00
	90%	11.2%	0.38	0.46	0.83	0.51	*0.03*	0.23	0.65

Table 4.3: P-values of VaR backtests for the hypotheses of independence and (un)conditional coverage.

to Frank cluster copulas, this requires a lot more parameters and does not allow for sampling using a closed-form procedure, but this also adds flexibility, since Gaussian cluster copulas do not assume exchangeable within-sector dependence. For simulation, we use rejection-like sampling with an error of $\varepsilon = 10^{-2}$, which turned out to yield good results in the simulation study in Section 4.4. Sampling methods for other copulas can be found, e.g., in Mai and Scherer (2012). Backtesting results of the six models according to the above tests can be found in Table 4.3. The time series of VaR forecasts and exceedances are shown in Figure 4.8.

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, which is to be expected, since it completely ignores the dependence among the stocks. Between the two considered hierarchical Kendall copulas no significant difference can be detected. 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. Especially at the important 99% level, which is frequently used, the dependence models provide accurate forecasts.

This shows that hierarchical Kendall copulas are as good as the common Gaussian and Student's t copulas and also as the more flexible R-vine copula when it comes to out-of-sample validation. In particular, the hierarchical Kendall copula with Frank cluster copulas and Student's t nesting copula, which we consider here, is very parsimonious and allows for closed-form calculations and very efficient sampling due to its Archimedean clusters (see Section 4.2.1).

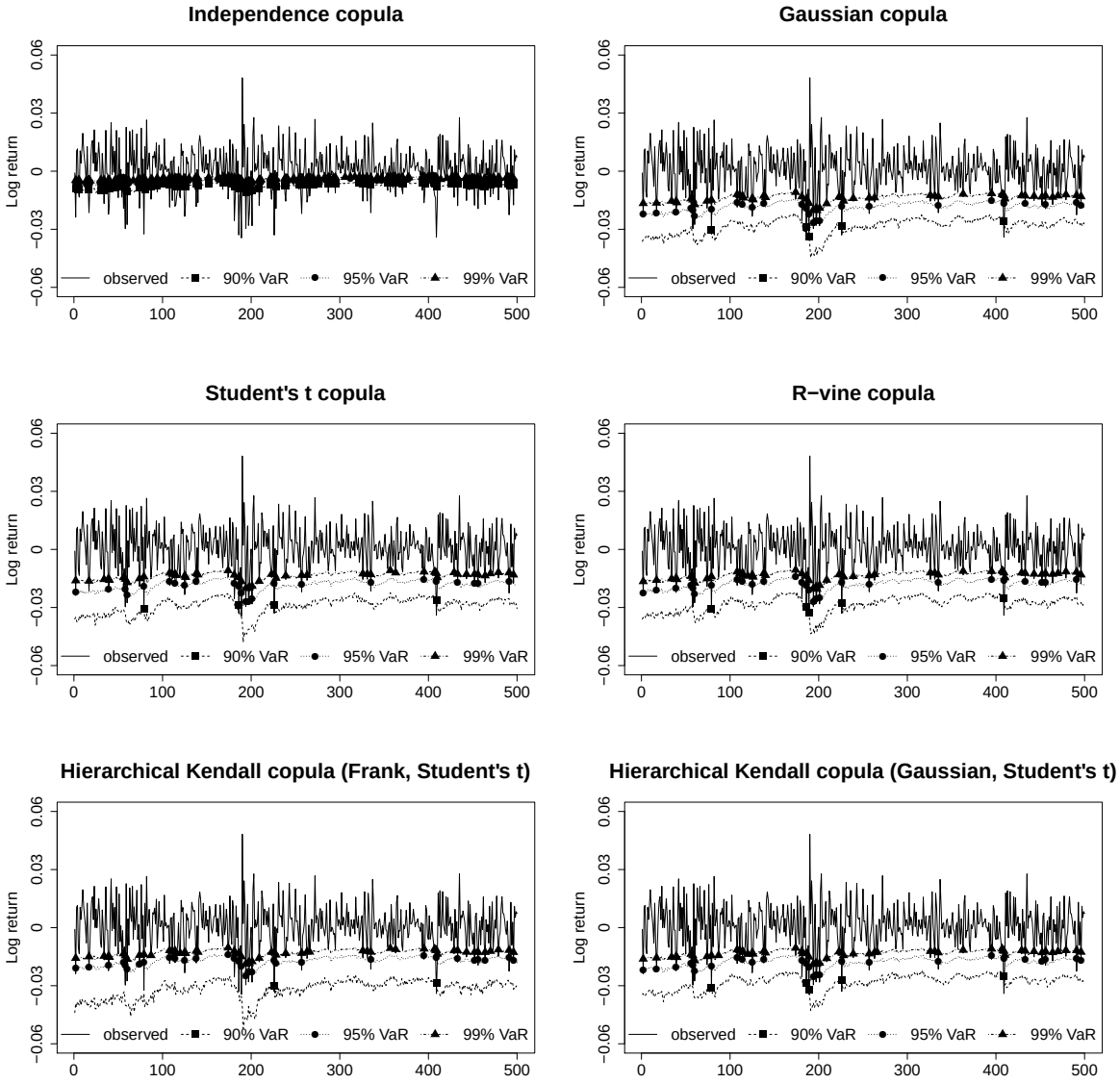


Figure 4.8: Time series of log returns of the DAX portfolio with 90%/95%/99%-VaR forecasts of the six considered dependence models. VaR exceedances (4.13) are marked by squares, circles and triangles, respectively.

4.6 Conclusion

The methods proposed in this chapter complement the results of Chapter 3, where the hierarchical Kendall copula is proposed as a new multivariate dependence model. We discuss the problem of sampling from a hierarchical Kendall copula and provide general guidelines how to solve it. While closed-form solutions can be derived for Archimedean, Archimax and Plackett copulas, approximate procedures have to be used otherwise. Three such approximate methods are proposed and evaluated in a simulation study. Especially top-down rejection-like sampling or, for sufficiently large sample sizes, bottom-up sample reordering can be recommended. The method of density resampling may be used if an adequate proposal copula can be calibrated.

An out-of-sample study of stock portfolio returns shows that hierarchical Kendall copulas are able to forecast the Value-at-Risk similarly well as common elliptical copulas and also as a flexible vine copula. These models can be considered as relevant benchmarks, since they constitute state-of-the-art dependence models for financial returns.

5 Systemic risk assessment

Since the financial crisis of 2007–2009 there is an active debate by regulators and academic researchers on systemic risk, with the aim of preventing similar crises in the future or at least reducing their impact. To investigate contagion effects among financial institutions, we develop methods for stress testing by exploiting the underlying dependence structure. New approaches for conditional copula simulation of the individual Student's t copula, of Archimedean copulas and of C-vine copulas are derived. The case of a hierarchical Kendall copula is also discussed. The chapter is mainly based on Brechmann, Hendrich, and Czado (2013).

5.1 Introduction

Dealing with the lessons learned from the financial crisis, the discussion about systemic risk has become more and more important. The collapse of Lehman Brothers in 2008 showed that the sudden and uncontrolled breakdown of a global financial company not only affected other financial institutions and seriously endangered the stability of the global financial sector but also had a great impact on the real economy of several countries around the world. As a result, the Financial Stability Board (FSB) developed guidelines to assess the systemic importance of financial institutions, markets, and instruments. The FSB defines systemic risk as "*the risk of disruption to financial services that is (i) caused by an impairment of all or parts of the financial system and (ii) has the potential to have serious negative consequences for the real economy*" (Financial Stability Board et al., 2009). Furthermore, an institution, market, or instrument is regarded as systemic if "*its failure or malfunction causes widespread distress, either as a direct impact or as a trigger for broader contagion*" on the financial system and/or the real economy.

The systemic relevance of an institution can be assessed based on several criteria that have been identified by the FSB. The three most important are size, lack of substitutability, and interconnectedness: Financial institutions whose "*distress or disorderly failure, because of their size, complexity, and systemic interconnectedness, would cause significant disruption to the wider financial system and economic activity*" (Financial Stability Board, 2011) are called systemically important. These institutions will face additional regulatory measures to reduce the systemic risk imposed by them. The Basel Committee on Banking Supervision (2011a) and the International Association of Insurance Supervisors (2012) have developed methodologies to determine globally systemically important banks and insurers, respectively. The assessment methodology for insurers differs to that used for banks, since it takes into account the fundamental differences in the business models of banks and insurance companies. While a systemic classification of insurers has not been published yet, a list of globally systemically important banks is released on a

yearly basis. In 2012, there were 28 banks on this list (Financial Stability Board, 2012).

Despite the popular expression "too big to fail", it has been argued in recent literature that the interconnectedness of an institution is much more important in the assessment of systemic risk: Cont and Moussa (2010) and Cont et al. (2013) find that the impact of the failure of an institution strongly depends on the interdependencies among institutions and less on its size. Similarly, Markose et al. (2012) observe in their analysis of interconnectedness in the US banking sector that only a few major institutions play a dominant role in terms of network centrality and connectivity. With respect to contagion in the US insurance sector, Park and Xie (2011) evaluate the impact of reinsurer downgradings on US property-casualty insurers and conclude that a systemic crisis caused by reinsurance transactions is rather unlikely. Billio et al. (2012) analyze the interdependencies among financial institutions from different sectors using principle component analysis and Granger-causality networks and detect an interesting asymmetry in the financial system, as banks are more likely to transmit shocks than insurers, hedge funds or brokers-dealers. Hence, in light of this research, it is more appropriate to speak of systemically important institutions as "too (inter-)connected to fail".

The exploration of contagion and interconnectedness is also the topic of this chapter. We propose to use copulas to analyze interdependencies in the global financial market, notably in the banking as well as in the insurance sector and not in both sectors in isolation, as it is often done. In doing so, we aim to find out whether there are significant differences in the dependence structure among banks and among (re-)insurers. As a statistical tool for dependence modeling, copulas allow for an accurate analysis beyond linear correlations and the common multivariate normal distribution. Therefore, we not only consider the popular classes of Archimedean and elliptical copulas, but also the grouped Student's t copula, vine copulas as well as hierarchical Kendall copulas. These copulas allow to take into account heterogeneous pairwise and tail dependencies more appropriately and therefore overcome limitations of the elliptical copulas, which are typically used in higher-dimensional dependence analysis. The grouped Student's t copula and hierarchical Kendall copulas also allow to explicitly respect groupings of variables, for instance, by sector and by geographical region.

Stress testing is an important tool for the assessment and classification of systemic risk. The systemic relevance of an institution is decisively determined by the potential impact of its failure on other institutions. It is therefore crucial to analyze such stress situations in the market by taking into account the interdependence among the institutions. Statistically speaking, we are interested in the following situation: Let $\boldsymbol{X} := (X_1, ..., X_d)'$ be a random vector of risk quantities. Then we want to analyze the case $\boldsymbol{X}_{-r}|X_r = x_r$, $r \in \{1, ..., d\}$, where $\boldsymbol{X}_{-r}$ denotes the random vector $\boldsymbol{X}$ without the rth component and the event $\{X_r = x_r\}$ corresponds to a stress situation. For instance, let X_r be the company value, then a stress situation occurs when x_r is very small.

Such an analysis requires the availability of the conditional distribution of $\boldsymbol{X}_{-r}|X_r = x_r$, given the specific underlying dependence model. As this distribution is typically not known in closed form, conditional simulation algorithms are needed for the scenario analysis. While these are straightforward and well-known in the case of elliptical copulas, we derive appropriate methods for Archimedean and vine copulas as well as for the individual Student's t copula, which includes the grouped Student's t copula as a special case. Hier-

archical Kendall copulas then exploit the methods developed for the other classes, which can form the building blocks of the hierarchical model specification.

The methodology developed in this chapter is used in a case study of 38 important financial institutions from all over the world, among them 20 insurers and 18 banks. Their credit default swap spreads, as market-based indicators of the credit worthiness, are statistically analyzed and appropriate multivariate dependence models are constructed. A stress testing exercise then provides insights into the systemic relevance of the different institutions. We detect differences among regional markets and, in addition, among the banking and the insurance sector. Interestingly, the official classification of globally systemically important banks is hardly reflected in the data. Furthermore, the analysis reveals new results regarding the classification of insurers, which, however, can not yet be compared to an official classification

The chapter is structured as follows. Conditional copula simulation procedures for the classes of elliptical, Archimedean, C-vine and hierarchical Kendall copulas as well as for the individual Student's t copula are derived in Section 5.2. The case study is then presented in Section 5.3, where the developed methods are used to conduct a systemic risk stress test in Section 5.3.4. Section 5.4 concludes.

5.2 Conditional copula simulation

As noted in the introduction, we are interested in the following situation for systemic risk stress testing. Let $\boldsymbol{X} = (X_1, ..., X_d)' \sim F$ be a continuous random vector and let $\boldsymbol{X}_{-r}$ denote the sub-vector of $\boldsymbol{X}$ having the rth component removed, $r \in \{1, ..., d\}$. Then what is the distribution $F_{-r|r}$ of $\boldsymbol{X}_{-r}|X_r = x_r$? If the event $\{X_r = x_r\}$ corresponds to an extreme situation, this distribution describes the impact of the rth variable being stressed.

Using the stressed distribution $F_{-r|r}$, we are then interested in calculating quantities like the conditional expected value $E(\boldsymbol{X}_{-r}|X_r = x_r)$ and the conditional variance $Var(\boldsymbol{X}_{-r}|X_r = x_r)$ to quantify size and variability of the impact. Similarly, quantiles of $\boldsymbol{X}_{-r}|X_r = x_r$ are of interest to assess extreme cases. Since the calculation of these quantities may not be feasible in closed form, it is often necessary to resort to statistical simulation from $F_{-r|r}$ to calculate Monte Carlo estimates of the quantities of interest.

A general approach can be formulated using the conditional inverse method for sampling from multivariate distributions (see also Section 4.2). Let $F_{j|1,...,j-1}$ denote the conditional distribution function of $X_j|(X_1 = x_1, ..., X_{j-1} = x_{j-1})$ for $j = 1, ..., d$ (for $j = 1$ the conditioning set is empty) and, without loss of generality, let $r = 1$. Hence, the aim is to sample $x_2, ..., x_d$ from $F_{-1|1}$, which is the conditional distribution of $\boldsymbol{X}_{-1}|X_1 = x_1$. We do this through iterative sampling from the distribution of $X_j|(X_1 = x_1, ..., X_{j-1} = x_{j-1})$ for $j = 2, ..., d$. For this, we sample $w_2, ..., w_d$ independently from the uniform distribution and set

$$x_j = F^{-1}_{j|1,...,j-1}(w_j|x_1, ..., x_{j-1}), \quad j = 2, ..., d, \tag{5.1}$$

which then define observations from $F_{-1|1}$.

This approach is very appealing if the conditional distribution functions $F_{j|1,...,j-1}$ can be determined in closed form. In the case of a C-vine copula as the underlying copula of

$\boldsymbol{X}$, this is, in fact, possible (see Section 5.2.4 below). Before that, we discuss the popular classes of elliptical and Archimedean copulas as well as the more recently proposed individual Student's t copula. While the conditional simulation of elliptical copulas is well-known, the procedures in the case of the individual Student's t copula and Archimedean copulas are more challenging and we derive new approaches here. Subsequently, conditional simulation of hierarchical Kendall copulas is discussed, which exploits the results for its building blocks as well as the sampling methods developed in Chapter 4.

Note that, if $U_j := F_j(X_j)$ for $j = 1, ..., d$, then it is equivalent to sample from the distribution of $\boldsymbol{X}_{-r}|X_r = x_r$ or that of $\boldsymbol{U}_{-r}|U_r = u_r$ where $u_r := F_r(x_r)$, since observations from the latter can be transformed back to the original level of the data by applying the inverse distribution functions $F_j^{-1},\ j \in \{1, ..., d\} \setminus \{r\}$. We therefore concentrate on the case $\boldsymbol{U}_{-r}|U_r = u_r$. As above, we let $r = 1$ unless otherwise stated.

Bivariate copulas such as extreme value copulas and the Plackett copula are not discussed here, since conditional simulation is straightforward in this case: The sampling approach using the conditional inverse method (5.1) then only requires the conditional copulas $C_{2|1}$ (2.14) (see also Equation (2.58)), which can be obtained by differentiation.

5.2.1 Elliptical copulas

For conditional simulation from elliptical copulas (see Section 2.2), it is advantageous to transform the random variables by the respective inverse distribution functions. That is, for the Gaussian copula set $Y_j := \Phi^{-1}(U_j),\ j = 1, ..., d$, and $y_1 := \Phi^{-1}(u_1)$ and for the Student's t copula set $Y_j := T_\nu^{-1}(U_j),\ j = 1, ..., d$, and $y_1 := T_\nu^{-1}(u_1)$. Then we sample $(y_2, ..., y_d)'$ from the corresponding conditional distribution function of $\boldsymbol{Y}_{-1}|Y_1 = y_1$ with appropriate parameters. These observations are finally transformed by $u_j = \Phi(y_j)$ or $u_j = T_\nu(y_j)$, respectively, for $j = 2, ..., d$.

For the multivariate normal case the conditional distribution of $\boldsymbol{Y}_{-1}|Y_1 = y_1$ is well-known (see, e.g., Kotz et al. (2000)). Let $\boldsymbol{Y} \sim \mathcal{N}_d(\boldsymbol{\mu}, \Sigma)$, where $\boldsymbol{\mu} = (\mu_1, \boldsymbol{\mu}_{-1}')'$ and

$$\Sigma = \begin{pmatrix} \sigma_{11} & \boldsymbol{\sigma}_1' \\ \boldsymbol{\sigma}_1 & \Sigma_{(-1,-1)} \end{pmatrix},$$

where $\boldsymbol{\sigma}_1 := (\sigma_{12}, ..., \sigma_{1d})'$ and $\Sigma_{(-1,-1)}$ denotes the covariance matrix Σ with first row and first column removed. Then $\boldsymbol{Y}_{-1}|Y_1 = y_1$ is again multivariate normal with modified mean vector and covariance matrix:

$$\boldsymbol{Y}_{-1}|Y_1 = y_1 \sim \mathcal{N}_{d-1}(\widetilde{\boldsymbol{\mu}}, \widetilde{\Sigma}),$$

where

$$\widetilde{\boldsymbol{\mu}} = \boldsymbol{\mu}_{-1} + \boldsymbol{\sigma}_1'(y_1 - \mu_1)/\sigma_{11} \quad \text{and} \quad \widetilde{\Sigma} = \Sigma_{(-1,-1)} - \boldsymbol{\sigma}_1\boldsymbol{\sigma}_1'/\sigma_{11}. \tag{5.2}$$

Since the conditional distribution is hence known in closed form, expectation and variance are given explicitly and simulation is only required if non-standard quantities need to be obtained. For the conditional copula simulation set $\boldsymbol{\mu} = \boldsymbol{0}$ and $\Sigma = R = (\rho_{jk})_{j,k=1,...,d}$, where R is the corresponding correlation matrix. Then, the conditional mean vector and the conditional correlation matrix simplify to

$$\widetilde{\boldsymbol{\mu}} = \boldsymbol{\rho}_1' y_1 \quad \text{and} \quad \widetilde{R} = R_{(-1,-1)} - \boldsymbol{\rho}_1\boldsymbol{\rho}_1', \tag{5.3}$$

where $\boldsymbol{\rho}_1 := (\rho_{12}, ..., \rho_{1d})'$ and $R_{(-1,-1)}$ denotes the correlation matrix R with the first row and the first column removed.

The conditional distribution in the case of a multivariate Student's t distribution is also known in closed form. Let $\boldsymbol{Y} \sim \mathcal{T}_d(\boldsymbol{\mu}, \Sigma, \nu)$. Then (see, e.g., Kotz and Nadarajah (2004))

$$F_{\boldsymbol{Y}_{-1}|Y_1}(\boldsymbol{y}_{-1}|y_1; \widetilde{\boldsymbol{\mu}}, \widetilde{\Sigma}, \nu+1) = T_{\widetilde{\boldsymbol{\mu}},\widetilde{\Sigma},\nu+1}\left(\sqrt{\frac{\nu+1}{\nu+(y_1-\mu_1)^2/\sigma_{11}}}\boldsymbol{y}_{-1}\right),$$

where $\widetilde{\boldsymbol{\mu}}$ and $\widetilde{\Sigma}$ are defined in Equation (5.2) and $T_{\widetilde{\boldsymbol{\mu}},\widetilde{\Sigma},\nu+1}$ is the distribution function of the $\mathcal{T}_{d-1}(\widetilde{\boldsymbol{\mu}}, \widetilde{\Sigma}, \nu+1)$ distribution. This means that observations from $F_{\boldsymbol{Y}_{-1}|Y_1}$ can be drawn by sampling $(\widetilde{y}_2, ..., \widetilde{y}_d)'$ from $\mathcal{T}_{d-1}(\widetilde{\boldsymbol{\mu}}, \widetilde{\Sigma}, \nu+1)$ and then setting $y_j = \widetilde{y}_j\sqrt{(\nu+(y_1-\mu_1)^2/\sigma_{11})/(\nu+1)}$ for $j = 2, ..., d$. To conditionally sample from a Student's t copula, set $\boldsymbol{\mu} = \boldsymbol{0}$ and $\Sigma = R$ as before.

5.2.2 Individual Student's t copula

Conditional simulation from an individual Student's t copula (see Section 2.3) is however not straightforward. We propose the following procedure: As before, transform the variables and the conditioning value such that $Y_j := T_{\nu_j}^{-1}(U_j),\ j = 1, ..., d$, and $y_1 := T_{\nu_1}^{-1}(u_1)$. Then $\boldsymbol{Y} \sim \mathcal{IT}_d(\boldsymbol{0}, R, \boldsymbol{\nu})$. From the representation in Equation (2.27), we know that $Y_j = W_j Z_j$, where $\boldsymbol{Z} \sim \mathcal{N}_d(\boldsymbol{0}, R)$ and $W_j = \sqrt{\nu_j/F_{\chi^2}^{-1}(Q;\nu_j)} \sim F_{W_j},\ j = 1, ..., d$, with $Q \sim U(0,1)$. We exploit this representation to sample from $\boldsymbol{Y}_{-1}|Y_1 = Y_1$ using the following steps.

(i) Sample a value w_1 from $W_1|Y_1 = y_1$.

(ii) Extract the value $q = F_{W_1}(w_1)$, which is common to all mixing variables, and set $w_j = F_{W_j}^{-1}(q)$ for $j = 2, ..., d$.

(iii) Compute the value of Z_1 given $Y_1 = y_1$ and $W_1 = w_1$ as $z_1 = y_1/w_1$ and use it to sample $(z_2, ..., z_d)'$ from $\boldsymbol{Z}_{-1}|Z_1 = z_1$ as described in Section 5.2.1.

(iv) Then $(y_2, ..., y_d)'$, where $y_j = w_j z_j,\ j = 2, ..., d$, are observations from $\boldsymbol{Y}_{-1}|Y_1 = y_1$.

Therefore, we need to find a way to sample from $W_1|Y_1 = y_1$ and we need to determine the distribution function $F_{W_j},\ j = 1, ..., d$. The latter is straightforward:

$$\begin{aligned} F_{W_j}(w_j;\nu_j) &= P\left(\sqrt{\nu_j/F_{\chi^2}^{-1}(Q;\nu_j)} \le w_j\right) = P\left(F_{\chi^2}^{-1}(Q;\nu_j) \ge \nu_j/w_j^2\right) \\ &= 1 - F_{\chi^2}(\nu_j/w_j^2;\nu_j), \quad w_j > 0. \end{aligned}$$

The corresponding density is therefore given by

$$f_{W_j}(w_j;\nu_j) = f_{\chi^2}\left(\frac{\nu_j}{w_j^2};\nu_j\right)\frac{2\nu_j}{w_j^3}, \quad w_j > 0. \tag{5.4}$$

To sample from $W_1|Y_1 = y_1$, we propose to use the conditional inverse method and derive the distribution $F_{W_1|Y_1}$ via its density for $w_1 > 0$,

$$f_{W_1|Y_1}(w_1|y_1;\nu_1) = \frac{f_{W_1,Y_1}(w_1,y_1;\nu_1)}{t_{\nu_1}(y_1)} = \frac{f_{W_1,Z_1}\left(w_1, \frac{y_1}{w_1};\nu_1\right)\frac{1}{w_1}}{t_{\nu_1}(y_1)} = \frac{f_{W_1}(w_1;\nu_1)\,\phi\left(\frac{y_1}{w_1}\right)\frac{1}{w_1}}{t_{\nu_1}(y_1)},$$

where we applied the change of variables $(W_1, Y_1)' \mapsto (W_1, Y_1/W_1)'$ and used that W_1 and Z_1 are independent as well as that $Y_1 \sim \mathcal{T}_1(0,1,\nu_1)$ and $Z_1 \sim \mathcal{N}_1(0,1)$. The density of W_1 is given in Equation (5.4). Hence,

$$F_{W_1|Y_1}(w_1|y_1;\nu_1) = \frac{1}{t_{\nu_1}(y_1)} \int_0^{w_1} f_{W_1}(x;\nu_1)\,\phi\left(\frac{y_1}{x}\right)\frac{1}{x}\,dx, \quad w_1 > 0.$$

As shown in Appendix B.5, we obtain after some calculations that

$$F_{W_1|Y_1}(w_1|y_1;\nu_1) = 1 - F_{\chi^2}\left(\frac{\nu_1 + y_1^2}{w_1^2};\nu_1 + 1\right), \quad w_1 > 0,$$

which can easily be used for conditional inverse sampling.

Resulting observations $(y_2, ..., y_d)'$ from $\boldsymbol{Y}_{-1}|Y_1 = Y_1$ are then transformed by $u_j = T_{\nu_j}(y_j)$, $j = 2, ..., d$. For the sake of clarity, we summarize the presented conditional simulation procedure in a compact algorithm.

Algorithm 5.1 (Conditional simulation from an individual Student's t copula)**.** Let $C(\cdot; R, \boldsymbol{\nu})$ be an individual Student's t copula with correlation matrix R and degrees of freedom $\boldsymbol{\nu}$. To generate observations given that $\{U_1 = u_1\}$, proceed as follows.

(i) Sample v from the uniform distribution.

(ii) Set $y_1 = T_{\nu_1}^{-1}(u_1)$.

(iii) Set $w_1 = \sqrt{(\nu_1 + y_1^2)/(F_{\chi^2}^{-1}(1 - v; \nu_1 + 1))}$.

(iv) Set $q = F_{W_1}(w_1)$ and $w_j = F_{W_j}^{-1}(q)$ for $j = 2, ..., d$.

(v) Set $z_1 = y_1/w_1$ and sample $(z_2, ..., z_d)'$ from $\mathcal{N}_{d-1}(\widetilde{\boldsymbol{\mu}}, \widetilde{R})$, where $\widetilde{\boldsymbol{\mu}} = \boldsymbol{\rho}_1' y_1$ and $\widetilde{R} = R_{(-1,-1)} - \boldsymbol{\rho}_1\boldsymbol{\rho}_1'$ (see Equation (5.3)).

(vi) Set $y_j = w_j z_j$ and $u_j = T_{\nu_j}(y_j)$ for $j = 2, ..., d$.

(vii) Return $(u_2, ..., u_d)'$.

Using this algorithm the individual Student's t copula as well as the grouped Student's t copula as a special case can be conveniently used for systemic risk stress testing.

5.2.3 Archimedean copulas

If $\boldsymbol{U} \sim C(\cdot;\varphi)$, where $C(\cdot;\varphi)$ is an Archimedean copula with generator φ (see Section 2.4), then Mesfioui and Quessy (2008) show that the conditional distribution of $U_j|(U_1 = u_1, ..., U_{j-1} = u_{j-1})$ for $j = 2, ..., d$ is given by

$$C_{j|1,...,j-1}(u_j|u_1, ..., u_{j-1};\varphi) = \frac{(\varphi^{-1})^{(j-1)}\left(\sum_{i=1}^{j}\varphi(u_i)\right)}{(\varphi^{-1})^{(j-1)}\left(\sum_{i=1}^{j-1}\varphi(u_i)\right)}, \quad u_j \in [0,1]. \tag{5.5}$$

The conditional inverse method (5.1) hence requires inversion of $(\varphi^{-1})^{(j-1)}$ for $j = 2, ..., d$, which may be numerically rather challenging, although explicit functional expressions of $(\varphi^{-1})^{(j-1)}$ for common Archimedean generators are provided in Hofert et al. (2012). We therefore derive an alternative conditional sampling strategy.

Here, we use a trick and introduce the copula level set variable $Z := C(\boldsymbol{U};\varphi) \in [0,1]$, which is distributed according to the Kendall distribution function $K(\cdot;\varphi)$ (see Equation (2.36)). Instead of directly sampling from the conditional distribution of $U_j|(U_1 = u_1, ..., U_{j-1} = u_{j-1})$ when using the conditional inverse method (5.1), the idea is to sample z from $Z|U_1 = u_1$ and use this to iteratively sample u_j from $U_j|(U_1 = u_1, ..., U_{j-1} = u_{j-1}, Z = z)$ for $j = 2, ..., d$. That is, first the level z of the copula level set $L(z;\varphi)$ is sampled given the event $\{U_1 = u_1\}$ and then the remaining variables are generated given this level set and the event $\{U_1 = u_1\}$. This approach is beneficial, since the distribution function $F_{U_j|U_1,...,U_{j-1},Z}$ of $U_j|(U_1 = u_1, ..., U_{j-1} = u_{j-1}, Z = z)$ is known explicitly (see Lemma 4.4) and can also be inverted in closed form. It is therefore numerically very efficient to use it for the sampling strategy using the conditional inverse method (5.1).

Hence, the open question is how to sample from $Z|U_1 = u_1 \sim F_{Z|U_1}$. For this, we begin with decomposing the density $f_{Z|U_1}$ corresponding to $F_{Z|U_1}$ as

$$f_{Z|U_1}(z|u_1;\varphi) = f_{U_1|Z}(u_1|z;\varphi)\, k(z;\varphi), \quad z \in (0,1), \tag{5.6}$$

which holds, since U_1 is uniform, so that $f_{U_1}(u_1) = 1,\ u_1 \in (0,1)$. The density $k(\cdot;\varphi)$ of the Kendall distribution function (2.36) is derived as

$$k(z;\varphi) = \frac{(-1)^{d-1}}{(d-1)!}\,\varphi(z)^{d-1}\,\varphi'(z)\,(\varphi^{-1})^{(d)}(\varphi(z)). \tag{5.7}$$

Further, since Equation (4.6) for $F_{U_j|U_1,...,U_{j-1},Z}$ also holds for $j = 1$ (with an empty conditioning set), the density $f_{U_1|Z}$ of $U_1|Z = z$ is given by

$$f_{U_1|Z}(u_1|z;\varphi) = -(d-1)\left(1 - \frac{\varphi(u_1)}{\varphi(z)}\right)^{d-2}\frac{\varphi'(u_1)}{\varphi(z)}. \tag{5.8}$$

Combining Equations (5.6), (5.7) and (5.8) then yields

$$f_{Z|U_1}(z|u_1;\varphi) = \frac{1}{(d-2)!}\,(\varphi(u_1) - \varphi(z))^{d-2}\,\varphi'(u_1)\,\varphi'(z)\,(\varphi^{-1})^{(d)}(\varphi(z)),$$

and

$$F_{Z|U_1}(z|u_1;\varphi) = \int_0^z f_{Z|U_1}(y|u_1;\varphi)\,dy \overset{x=\varphi(y)}{=} \frac{1}{(d-2)!}\,\varphi'(u_1)\int_{\varphi(0)}^{\varphi(z)} (\varphi(u_1)-x)^{d-2}\,(\varphi^{-1})^{(d)}(x)\,dx, \tag{5.9}$$

which we solve recursively using integration by parts (see Appendix B.2) to obtain

$$F_{Z|U_1}(z|u_1;\varphi) = \varphi'(u_1)\sum_{k=1}^{d-1}\frac{(\varphi(u_1)-\varphi(z))^{k-1}}{(k-1)!}\,(\varphi^{-1})^{(k)}(\varphi(z)), \quad z\in(0,1). \tag{5.10}$$

This closed-form expression is similar to that of the Kendall distribution function (see Equation 2.36). It can then be used for conditional inverse sampling from $Z|U_1 = u_1$. This means, in contrast to inversion of $(\varphi^{-1})^{(j-1)}$ for $j = 2,...,d$, as in direct conditional inverse sampling of Archimedean copulas using Equation (5.5), the only numerically challenging step of this newly proposed strategy is the inversion of $F_{Z|U_1}$, which is given in Equation (5.10). The procedure can be summarized in the following algorithm, which is similar to Algorithm 4.5 for sampling from an Archimedean copula on a given level set.

Algorithm 5.2 (Conditional simulation from an Archimedean copula)**.** Let $C(\cdot;\varphi)$ be an Archimedean copula with generator φ. To generate observations given that $\{U_1 = u_1\}$, proceed as follows.

(i) Sample $w_1,...,w_{d-1}$ independently from the uniform distribution.

(ii) Set $z = F^{-1}_{Z|U_1}(w_1|u_1;\varphi)$.

(iii) For $j = 2,...,d-1$: $u_j = \varphi^{-1}((1-w_j^{1/(d-j)})(\varphi(z) - \sum_{1\le i<j}\varphi(u_i)))$.

(iv) Set $u_d = \varphi^{-1}(\varphi(z) - \sum_{1\le i<d}\varphi(u_i))$.

(v) Return $(u_2,...,u_d)'$.

5.2.4 C-vine copulas

Among the large class of vine copulas (see Section 2.7) we focus here on the sub-class of C-vine copulas, which leads to a particularly appealing conditional simulation algorithm. Unconditional simulation from a C-vine copula can be carried out using the conditional inverse method (5.1) and the conditional distribution functions $C_{i|1,...,i-1}$ of $U_i|(U_1 = u_1,...,U_{i-1} = u_{i-1})$ given in Equation (2.70), which we here restate for convenience. For $i = 2,...,d-1$ and $k = 0,...,d-i$ it holds that

$$\begin{aligned} &C_{i+k|1,...,i-1}(u_{i+k}|u_1,...,u_{i-1}) \\ &\quad = C_{i+k|i-1;1,...,i-2}\left(C_{i+k|1,...,i-2}(u_{i+k}|u_1,...,u_{i-2})|C_{i-1|1,...,i-2}(u_{i-1}|u_1,...,u_{i-2})\right), \end{aligned} \tag{5.11}$$

where $u_{i+k}\in[0,1]$ and $C_{i+k|i-1;1,...,i-2}$ is defined in Equation (2.71).

The general sampling algorithm for C-vine copulas can be found in Aas et al. (2009). Since this sampling strategy makes use of the order of the C-vine root nodes, it is straightforward to conditionally sample from $\boldsymbol{U}_{-1}|U_1 = u_1$. However, in contrast to elliptical and Archimedean copulas, the case of $\boldsymbol{U}_{-r}|U_r = u_r$ for $r > 1$ needs to be considered explicitly, since the variables of a C-vine copula cannot simply be reordered.

Now, let $r > 1$. Clearly, for $i > r$ the sampling strategy of $U_i|(U_1 = u_1, ..., U_r = u_r, ..., U_{i-1} = u_{i-1})$ does not change. The question hence is how to sample from $U_1|(U_r = u_r)$, $U_2|(U_1 = u_1, U_r = u_r)$,..., $U_{r-1}|(U_1 = u_1, ..., U_{r-2} = u_{r-2}, U_r = u_r)$. This means that we need to compute the corresponding distribution functions $C_{i|1,...,i-1,r}$ for $1 \leq i < r$. It holds that

$$\begin{aligned} &C_{i|1,...,i-1,r}(u_i|u_1, ..., u_{i-1}, u_r) \\ &\quad = C_{i|r;1,...,i-1}\left(C_{i|1,...,i-1}(u_i|u_1, ..., u_{i-1})|C_{r|1,...,i-1}(u_r|u_1, ..., u_{i-1})\right), \end{aligned} \tag{5.12}$$

where $u_i \in [0, 1]$. Both arguments, $C_{i|1,...,i-1}$ and $C_{r|1,...,i-1}$, can be computed recursively according to Equation (5.11), since $r > i$.

To make things more concrete, we consider an illustrative example of a five-dimensional C-vine copula (the PCC of a four-dimensional C-vine can be found in Section 2.7.1).

Example 5.3 (Conditional simulation from a C-vine copula)**.** Let $d = 5$ and $r = 4$. Hence, we need to determine the conditional distribution functions $C_{1|4}$, $C_{2|1,4}$ and $C_{3|1,2,4}$, while $C_{5|1,2,3,4}$ is the same as in standard C-vine copula simulation. The first term, $C_{1|4}$, is straightforwardly given through the copula $C_{1,4}$, which is part of the C-vine PCC (2.69) and therefore known. According to Equation (5.12), we further have that

$$C_{2|1,4}(u_2|u_1, u_4) = C_{2|4;1}\left(C_{2|1}(u_2|u_1)|C_{4|1}(u_4|u_1)\right), \tag{5.13}$$

where the known copulas $C_{1,2}$, $C_{1,4}$ and $C_{2,4;1}$ are used to obtain $C_{2|1}$, $C_{4|1}$ and

$$C_{2|4;1}(u_2|u_4) = \frac{\partial C_{2,4;1}(u_2, u_4)}{\partial u_4},$$

respectively. By using Equation (5.12) again, we also compute $C_{3|1,2,4}$ as

$$C_{3|1,2,4}(u_3|u_1, u_2, u_4) = C_{3|4;1,2}\left(C_{3|1,2}(u_3|u_1, u_2)|C_{4|1,2}(u_4|u_1, u_2)\right), \tag{5.14}$$

where $C_{3|1,2}$ and $C_{4|1,2}$ are computed as in Equation (5.11). In particular, the copula $C_{3,4;1,2}$, which is part of the C-vine PCC (2.69), is used to derive

$$C_{3|4;1,2}(u_3|u_4) = \frac{\partial C_{3,4;1,2}(u_3, u_4)}{\partial u_4}.$$

Sampling using the conditional inverse method (5.1) is then feasible: Let w_1, w_2, w_3 and w_5 be independent observations from the uniform distribution. Then, we obtain

$$\begin{aligned} u_1 &= C^{-1}_{1|4}(w_1|u_4), \\ u_2 &= C^{-1}_{2|1}(C^{-1}_{2|4;1}(w_2|C_{4|1}(u_4|u_1))|u_1), \\ u_3 &= C^{-1}_{3|1}(C^{-1}_{3|2;1}(C^{-1}_{3|4;1,2}(w_3|C_{4|2;1}(C_{4|1}(u_4|u_1)|C_{2|1}(u_2|u_1)))|C_{2|1}(u_2|u_1))|u_1), \end{aligned}$$

by inverting the terms given in Equations (5.13) and (5.14), respectively, and finally u_5 by inverting Equation (5.11) for $i = 5$ and $k = 0$. □

The general sampling algorithm can be written down as outlined in the following (see Hendrich (2012)). This extends the standard C-vine simulation algorithm by Aas et al. (2009), from where notation is adopted.

Algorithm 5.4 (Conditional simulation from a C-vine copula)**.** Let $C(\cdot; \mathcal{V}, \mathcal{B}, \boldsymbol{\theta})$ be a C-vine copula with root node order $1, ..., d$. To generate observations given that $\{U_r = u_r\},\ r \geq 1$, proceed as follows.

(i) Let $V = (v_{j,k})_{j,k=1,...,d}$ be an auxiliary array.

(ii) Sample $w_1, ..., w_{r-1}, w_{r+1}, ..., w_d$ independently from the uniform distribution.

(iii) Set $v_{r,1} = u_r$.

(iv) For $i = 1, ..., r-1, r+1, ..., d$:

a) Set $v_{i,1} = w_i$.

b) If $i < r$, then set $v_{i,1} = C^{-1}_{i|r;1,...,i-1}(v_{i,1}|v_{r,i})$.

c) If $i > r$, then set $v_{i,1} = C^{-1}_{i|k;1,...,k-1}(v_{i,1}|v_{k,k})$ for $k = i-1, ..., 1$.

d) Set $u_i = v_{i,1}$.

e) If $i < d$, then set $v_{i,\ell+1} = C_{i|\ell;1,...,\ell-1}(v_{i,\ell}|v_{\ell,\ell})$ for $\ell = 1, ..., i-1$.

f) If $i < r$, then $v_{r,i+1} = C_{r|i;1,...,i-1}(v_{r,i}|v_{i,i})$.

(v) Return $(u_1, ..., u_{r-1}, u_{r+1}, ..., u_d)'$.

This algorithm hence enables us to simulate from a C-vine copula conditionally on a stress event to an arbitrary variable. It is of particularly convenient form for C-vine copulas, since C-vine trees are uniquely characterized by their root nodes. A similar algorithm for general R-vines has not been formulated yet.

5.2.5 Hierarchical Kendall copulas

Finally, we consider hierarchical Kendall copulas, which we discussed in Chapters 3 and 4. Let $\boldsymbol{U} := (U_1, ..., U_n)' \sim C_{\mathcal{K}}$, where $C_{\mathcal{K}}$ is a hierarchical Kendall copula as defined in Definition 3.3, and let $\{U_1 = u_1\}$ be the stress event as before. This means that the shock occurs in the first cluster. It then spreads to the other variables $\boldsymbol{U}_{-1}$ according to the following sampling scheme.

Algorithm 5.5 (Conditional simulation from a hierarchical Kendall copula)**.** Let $C_{\mathcal{K}}$ be a hierarchical Kendall copula with cluster copulas $C_1, ..., C_{n_1}$ and nesting copula C_0. To generate observations given that $\{U_1 = u_1\}$, proceed as follows.

(i) Sample $(u_2, ..., u_{m_1})'$ from $(U_2, ..., U_{m_1})'|U_1 = u_1$.

(ii) Set $v_1 = K_1(C_1(\boldsymbol{u}_1))$.

(iii) Sample $(v_2, ..., v_{n_1})'$ from $(V_2, ..., V_{n_1})'|V_1 = v_1$.

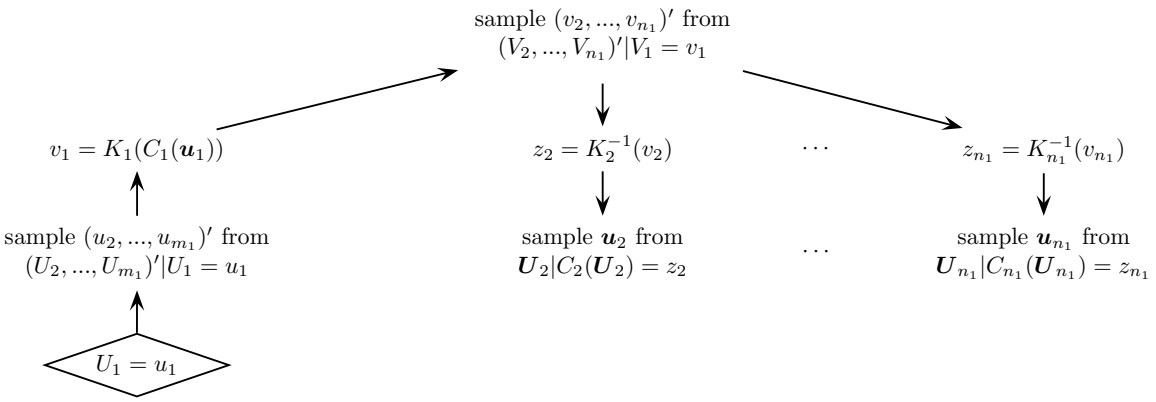


Figure 5.1: Illustration of the conditional simulation method for hierarchical Kendall copulas.

(iv) Set $z_i = K_i^{-1}(v_i)$ for all $i = 2, ..., n_1$.

(v) Sample $\boldsymbol{u}_i$ from $\boldsymbol{U}_i|C_i(\boldsymbol{U}_i) \equiv z_i$ for $i = 2, ..., n_1$.

(vi) Return $(u_2, ..., u_n)'$.

In Steps (i) and (iii), the methods discussed and developed in the previous sections can be used for conditional copula sampling. Step (v) can be carried out using the top-down procedures proposed in Section 4.2. The approach can also easily be extended to the case of k hierarchical levels (see Remark 3.7). For two levels it is illustrated in Figure 5.1.

5.3 Application: CDS spreads of financial institutions

The purpose of our case study on interconnectedness in the financial market is threefold. First, using appropriate statistical dependence models we carefully analyze the interdependencies among major financial institutions in the banking as well as the insurance sector and point out differences between these two sectors. Second, the developed statistical models are used to stress test the global financial market in order to obtain new insights with respect to the assessment and classification of systemically important institutions. Third, as we use credit default swap spreads for our analyses, we also investigate whether such data is actually useful to analyze systemic risk. The developed methodology is however independent of the data, which means that the first two questions can be investigated using the same tools but different data.

5.3.1 Data

Recently, there has been active research on the connection of credit default swaps and systemic risk. A credit default swap (CDS) is bilateral credit derivative contract, which allows the trading of default risks of an underlying corporate or government entity. Since the payoff of a CDS contract is caused by the default on debt, CDS spreads are a market-based

indicator of the credit worthiness of the reference entity. Rising CDS spreads indicate growing default expectations of the other market participants regarding the referenced entity. In fact, Hull et al. (2004) and Norden and Weber (2004) found that there is statistical evidence for the CDS market actually anticipating later rating announcements by the credit rating agencies.

The relationship between CDS and systemic risk seems obvious: If there is a systemic event in the market, default expectations of relevant institutions should rise, which is then reflected in increasing CDS spreads. Authors have therefore developed measures of systemic risk that are directly based on CDS spreads or the default probabilities derived from these (see for instance Acharya et al. (2011), Huang et al. (2009), and Giglio (2011)). CDS spreads have also been used to examine interdependencies among financial institutions: see Markose et al. (2012), Rahman (2013), Kaushik and Battiston (2013), and Chen et al. (2013). Nevertheless, none of the authors use copulas to account for non-standard interdependencies among the institutions. This is one aim of our study.

As data for our statistical analyses we use senior CDS spreads with a maturity of five years observed from January 2006 to October 2011 ($N = 1371$ daily observations), which are obtained from Bloomberg. In the attempt of a balanced selection of companies regarding their geographical region and sectoral belonging, we select 38 companies from the financial sector for the analysis of their interdependence structure. Among these are 18 banks and 20 (re-)insurers from different countries in three major geographical regions (abbreviations are shown in brackets):

(i) Systemically important banks according to the Financial Stability Board (2012):

- *Europe:* Banco Bilbao Vizcaya Argentaria (BBVA), Banco Santander (BS), Barclays, BNP Paribas, Deutsche Bank (DB), Royal Bank of Scotland (RBS), Société Générale (SG), Standard Chartered (SC), UBS, Unicredit
- *USA:* Citigroup, Goldman Sachs (GS), JP Morgan Chase (JPM)
- *Asia-Pacific:* Bank of China (BoC), Sumitomo Mitsui

Note that at the time of this analysis Banco Bilbao Vizcaya Argentaria and Standard Chartered had not yet been officially classified as systemic; see Financial Stability Board (2011).

(ii) Not systemically important banks:

- *Europe:* Intesa Sanpaolo
- *Asia-Pacific:* Kookmin Bank, Westpac Banking

(iii) (Re-)Insurers:

- *Europe:* Aegon, Allianz, Assicurazioni Generali, Aviva, AXA, Hannover Rück (HR), Legal & General (LG), Munich Re (MR), Prudential, SCOR, Swiss Re (SR), Zurich Insurance
- *USA:* ACE, Allstate, American International Group (AIG), Chubb, Hartford Financial Services, XL Group
- *Asia-Pacific:* QBE Insurance, Tokio Marine (TM)

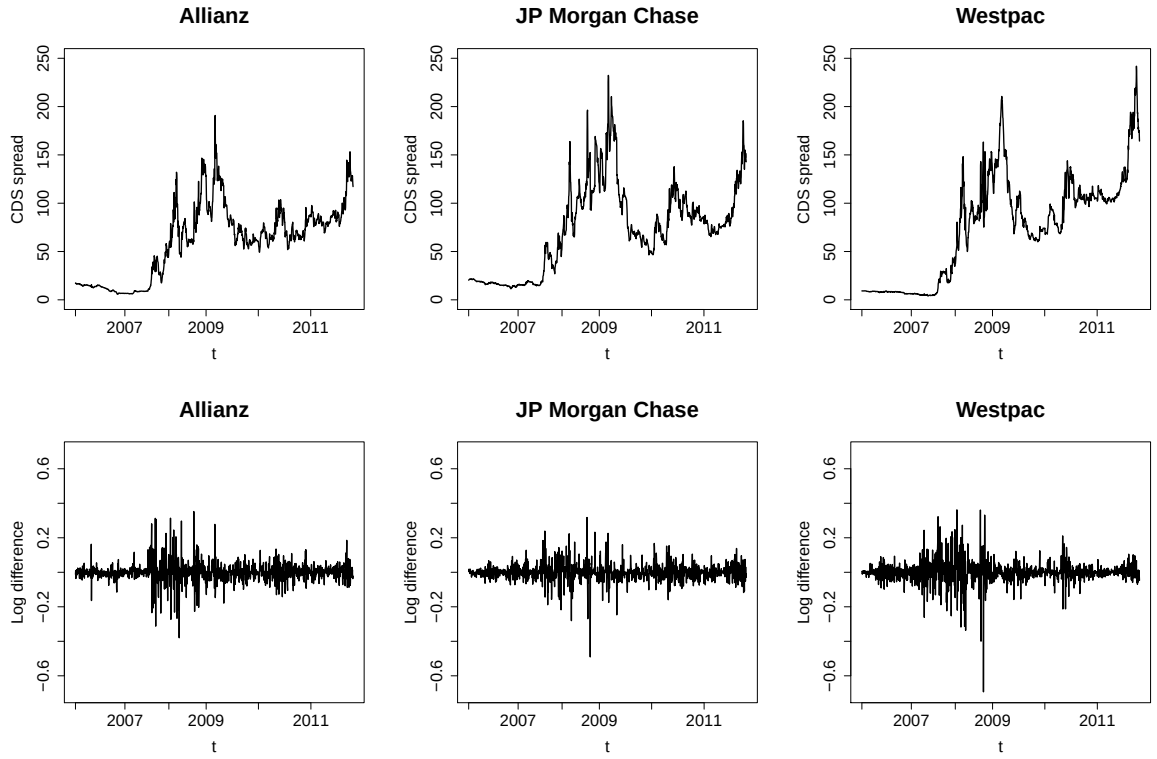


Figure 5.2: CDS spreads (top) and corresponding log differences (bottom) of Allianz, JP Morgan Chase and Westpac from January 2006 to October 2011.

For our analyses we use daily log differences of the CDS spreads of these 38 companies. Three time series of CDS spreads and their log differences are shown in Figure 5.2.

5.3.2 Marginal modeling

To deal with the serial dependence in the marginal time series as well as the between-series dependence, we employ the popular copula-GARCH approach as described in Section 3.7.2 and proceed according to the estimation method of inference functions for margins (IFM; see Section 3.5).

The time series of the log differences of the CDS spreads show common features of financial time series such as autocorrelation, leptokurtosis (heavy tails) and volatility clustering (see Figure 5.2 and Hendrich (2012, Table 4.2)). To remove these characteristics, we apply appropriate time series models. While often (ARMA-)GARCH models with (skewed) Student's t innovations provide good fits for financial time series (see, e.g., Chu et al. (2010) for an application to the iTraxx CDS index), this is not the case here. For each of the 38 time series we therefore separately consider extended GARCH models, such as the asymmetric exponential GARCH by Nelson (1991) or GARCH-in-mean by Engle et al. (1987), as well as non-standard innovations distributions like the generalized error, the generalized hyperbolic and the normal inverse Gaussian. All model fits are then carefully checked using a range of goodness-of-fit tests such as the Ljung-Box, the ARCH-LM or

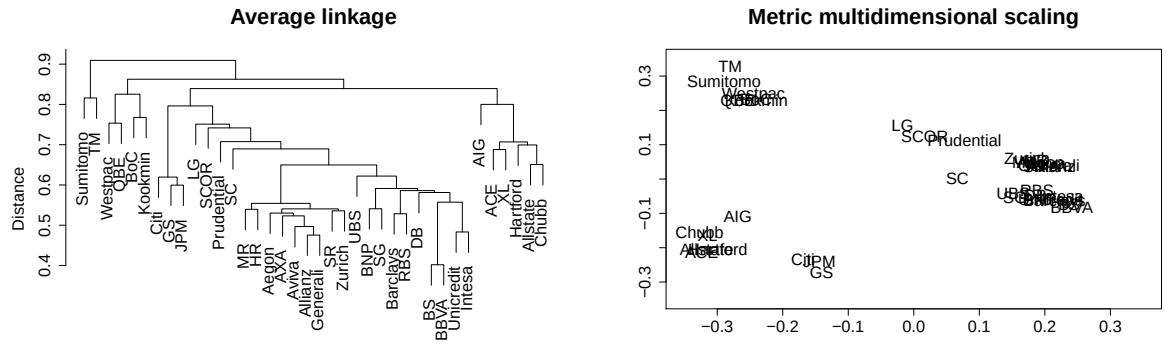


Figure 5.3: Hierarchical clustering with average linkage and metric multidimensional scaling of the institutions according to the metric $d(j,k) = \sqrt{1-\widehat{\rho}_{S,jk}}$ defined in Equation (3.17).

	EU banks	EU ins.	US banks	US ins.	AP banks	AP ins.
EU banks	0.51–0.84					
EU ins.	0.36–0.63	0.41–0.77				
US banks	0.33–0.41	0.28–0.41	0.60–0.64			
US ins.	0.24–0.35	0.22–0.34	0.27–0.39	0.31–0.58		
AP banks	0.12–0.30	0.14–0.32	0.10–0.24	0.15–0.28	0.18–0.41	
AP ins.	0.13–0.27	0.15–0.29	0.13–0.22	0.14–0.27	0.21–0.43	0.20–0.20

Table 5.1: Ranges of empirical Spearman's ρ_S values $\widehat{\rho}_{S,jk}$ within and between the sectors in the different regions.

the Nyblom stability test. More details on the fitting process can be found in Hendrich (2012, Section 5.1).

After adequately removing the serial dependence in each of the 38 univariate time series, we investigate the dependence among the residuals $\widehat{z}_{tj},\ t = 1,...,N,\ j = 1,...,38$. As we fix the estimated margins following the IFM method, we set $\widehat{u}_{tj} := \widehat{F}_j(\widehat{z}_{tj})$, where $\widehat{F}_j$ is the estimated innovations distribution of the jth time series.

5.3.3 Dependence modeling

To get a first impression of the interdependencies among the different institutions, we calculate empirical Spearman's ρ_S values, $\widehat{\rho}_{S,jk}$, for all pairs $j,k = 1,...,38,\ j < k$, and use metric multidimensional scaling to embed the institutions in the plane according to the metric $d(j,k) = \sqrt{1-\widehat{\rho}_{S,jk}}$ defined in Equation (3.17) (see, e.g., Hastie et al. (2009)). This means that the closer two institutions are to each other, the stronger is the dependence of their CDS spreads. In addition, we perform hierarchical clustering with the metric $d(j,k)$ and average linkage. The resulting plots are shown in Figure 5.3. Ranges of the empirical Spearman's ρ_S values per geographical region and sector are shown in Table 5.1.

The hierarchical clustering and the multidimensional scaling show that there is clear

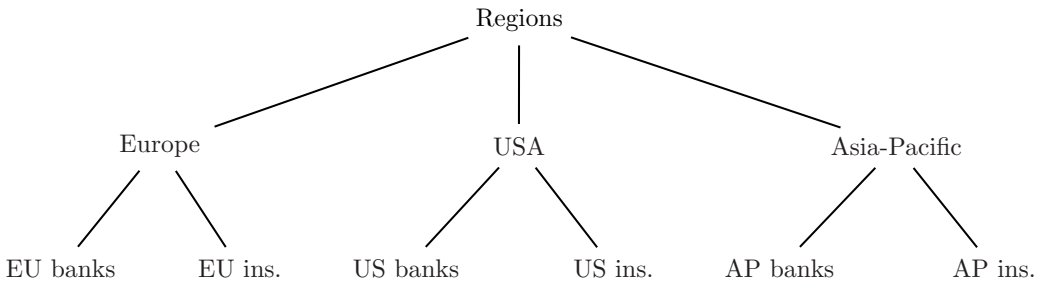


Figure 5.4: Sectoral clustering of CDS spreads.

geographical clustering present among the CDS spreads: European institutions can be found on the right of the multidimensional scaling plot, US institutions in the lower left corner and institutions from the Asia-Pacific region in the upper left corner. Within these regions there is also a clear separation of banks and insurers observable. Similar groupings of institutions can be observed in the dendogram. Hence, all pairs of companies within either one of the sectors show the strongest dependencies. Another interesting fact is that the banks that have not been officially classified as systemically important do not play a significantly different role than the other banks, as they cannot be explicitly distinguished based on the hierarchical clustering and the multidimensional scaling. The classification is not reflected here.

This exploratory look at the data illustrates that there are considerably different relationships among the institutions depending on the geographical region and the sector. Such heterogeneous dependencies cannot be appropriately captured using an Archimedean copula, which assumes exchangeability of all variables. While elliptical copulas are more appropriate for this purpose, they are still somewhat restrictive by imposing symmetric tail dependence. In the literature, it is however often observed that in times of crisis the dependence of joint negative events increases. For CDS spreads this means that we may expect the presence of upper tail dependence, which reflects the joint probability of extreme upward jumps in the expected default probabilities. In addition, the obvious sectoral clustering of variables, which is illustrated in Figure 5.4, should be taken into account.

We therefore consider a range of different models to account for these dependence characteristics:

(i) the Gumbel and the Frank copula (see Examples 2.16 and 2.17), where the Gumbel copula allows for upper tail dependence, but both copulas assume exchangeability;

(ii) the Gaussian and the Student's t copula (see Section 2.2), which are reflection symmetric and do not explicitly take into account the sectoral clustering of the institutions;

(iii) the grouped Student's t copula (see Section 2.3) with six sectors according to the lowest level shown in Figure 5.4 and thus with more flexibility in the modeling of the tails of the clusters;

(iv) a C-vine copula (see Section 2.7), which allows for a very flexible modeling of heterogeneous and pairwise dependencies;

(v) three-level hierarchical Kendall copulas (see Chapter 3, especially Remark 3.7) with clustering as illustrated in Figure 5.4 and with all cluster and nesting copulas chosen either as Gumbel, as Frank or as Gaussian.

The hierarchical Kendall copulas are the only dependence models, which explicitly take into account the three-level structure illustrated in Figure 5.4. The grouped Student's t copula can only model one level of grouping, so that we choose the most granular level here (the six sectors in the different regions). In addition, it does not explicitly allow to control the between-group dependence as hierarchical Kendall copulas do. All other copulas are not of hierarchical nature.

Due to the numerically more complex evaluation of the log likelihood compared to the standard Student's t copula, the grouped Student's t copula is only fitted sequentially as described in Section 3.7.3. Similarly, we concentrate on a sequential fit of the hierarchical Kendall copula with Gaussian cluster and nesting copulas (see Algorithm 3.19), since the Kendall distribution function has to be calculated using Monte Carlo simulation. The hierarchical Kendall copulas with Gumbel and Frank cluster and nesting copulas, for which the Kendall distribution functions are known in closed form (see Equation (2.36)), are estimated by joint maximum likelihood estimation of all dependence parameters.

The cluster and nesting copulas are fixed to be of the same type to limit the model complexity. The three considered models still cover a reasonable range of dependence patterns: Choosing only upper tail dependent Gumbel copulas as building blocks can be seen as a worst case model, while Gaussian building blocks are more flexible in the modeling of pairwise dependencies but assume tail independence. In fact, the sequential procedure of Algorithm 3.22 for the selection of an appropriate hierarchical Kendall copula almost yields the model with Gaussian cluster and nesting copulas. Only the fit of the bivariate within-region copulas, which specify the between-sector dependence, could be weakly improved by selecting a Gumbel copula for the US institutions and a Frank copula for the European institutions. The fits of respective bivariate Gaussian copulas are however very similar and actually yield a slightly higher log likelihood of the joint model in the end, so that we do not consider the sequentially selected model any further. Student's t copulas are not considered as building blocks, since it is difficult to efficiently calculate multivariate Student's t probabilities (see Genz and Bretz (2009)) and hence to efficiently evaluate the Student's t copula (2.23). It is therefore computationally very inefficient to obtain a Monte Carlo estimate of the Kendall distribution function.

The C-vine copula is selected according to the sequential selection algorithm by Czado et al. (2012), which is a special case of Algorithm 2.25 (see Dißmann et al. (2013)). Appropriate pair copulas are selected using the AIC from the following list: Gaussian, Student's t, Clayton, Gumbel and Frank as well as rotations by 90, 180 and 270 degrees of the reflection asymmetric copulas (see Table 2.1 and Figure 2.3). To obtain a more parsimonious model, the bivariate independence copula is also taken into account after performing an independence test of each pair. Subsequent to the sequential model selection, parameters are estimated jointly by maximum likelihood. We also fit an R-vine copula as described in Algorithm 2.25. The model however does not improve over the C-vine copula, so that we

Copula	Log lik.	# Par.	AIC	BIC
Gumbel	8640.22	1	-17278.45	-17273.22
Frank	8215.55	1	-16429.10	-16423.88
Gaussian	18326.53	703	-35247.07	-31575.09
Student's t	19915.88	704	-38423.76	-34746.56
Grouped Student's t*	20043.15	709	-38668.29	-34964.97
C-vine	20393.29	488	-39810.58	-37261.61
Hier. Kendall (Gumbel)	14206.19	10	-28392.38	-28340.15
Hier. Kendall (Frank)	13783.63	10	-27547.25	-27495.02
Hier. Kendall (Gaussian)*	17394.98	152	-34485.95	-33692.01

Table 5.2: Log likelihoods, numbers of parameters, AIC and BIC values of the copulas estimated by maximum likelihood (an asterisk indicates a sequential fit).

do not consider it any further here (see Hendrich (2012, Section 6.3.5) for more details). Note that especially in the first trees, where the strongest dependencies are modeled and therefore the simplifying assumption (see Section 2.7.2) has the strongest impact, mainly elliptical copulas are selected (see Hendrich (2012, Table 6.6)). We therefore expect that the simplifying assumption is approximately reasonable here.

Table 5.2 shows log likelihoods and AIC and BIC values of the copula fits for our 38-dimensional data set. The rather weak fits of the two considered Archimedean copulas illustrate the inappropriateness of an exchangeable Archimedean copula here. According to the AIC and the BIC, the C-vine copula can be regarded as the best model. In addition to a higher log likelihood compared to the elliptical copulas, it also benefits from a smaller number of parameters, which is achieved by using the independence copula for certain conditional pairs. As noted above, the C-vine copula may also better account for potential reflection asymmetry in the dependence structure, which is particularly important in the analysis of stress situations. The grouped Student's t copula also improves on the standard Student's t copula, which in turn provides a better fit than the Gaussian copula. In fact, the estimated degrees of freedom are 14.71 and therefore clearly indicate the presence of non-Gaussian dependence. The degrees of freedom per sector, as fitted for the grouped Student's t copula, range between 10.11 (European insurers) and 65.02 (insurers in the Asia-Pacific region). As observed in Section 3.7.3, we again find that the ad-hoc approach by Daul et al. (2003) would have underestimated the degrees of freedom of the different sectors, since it ignores the between-sector dependence.

The fits of the hierarchical Kendall copulas are not as good as that of the C-vine copula. Nevertheless, the models are much more parsimonious and provide a straightforward interpretation in terms of within- and between-sector dependence. In particular, the hierarchical Kendall copula with Gaussian cluster and nesting copulas is quite flexible but yet parsimonious, so that it provides a better fit than the Gaussian copula, when taking into account the number of model parameters. On the other hand, the two hierarchical Kendall copulas with Archimedean cluster and nesting copula show that, while standard Archimedean copulas are obviously inappropriate here, they can lead to a quite good

fit when used as building blocks of a hierarchical Kendall copula, which still constitutes an extremely parsimonious model. In both cases, the Gumbel copula is superior to the tail independent Frank copula. This underlines the importance of modeling upper tail dependence here, as it has been pointed out above.

Looking more closely at the fitted C-vine copula, we observe that the ordering of the institutions, which is implied by the selected root nodes, is less important here, since it is strongly driven by the number of institutions selected among certain regions and sectors and therefore does not directly provide an ordering of systemic importance. It is hence not surprising that the European institutions Allianz, BNP Paribas and Zurich Insurance are selected as the first three pivotal variables. Hartford Financial Services and JP Morgan Chase are then the first US institutions in the ordering, while QBE Insurance is the first institution from the Asia-Pacific region.

Interestingly, neither the exploratory analysis in Figure 5.3 nor the fitted models show that interdependencies involving systemically important banks are structurally different from those involving institutions that have not been classified as systemic. An additional finding is that the dependence of US banks and European institutions is determined to be higher than that of US insurers and European institutions. This indicates the, maybe not surprising, fact that especially the US banking sector plays a systemically important role in the financial market. This is in line with findings of Billio et al. (2012). To obtain a more differentiated view on the systemic importance of specific institutions and sectors, we conduct a stress testing exercise of the global CDS market.

5.3.4 Systemic risk stress test

According to the Financial Stability Board et al. (2009), a systemic crisis is defined as the distress of a whole system caused by the failure of one institution and the subsequent spreading of malfunction from one company to another. Hence, we now aim to further investigate the possibility of contagion among the institutions in our sample. We perform a simulation study to exploit the modeled dependence structure. More precisely, we assume a stress situation for one of the institutions and simulate the resulting impact on the remaining institutions as illustrated in Figure 5.5. In particular, we are interested to find out whether there are significant differences regarding the type of the institution that is stressed.

The fictitious stress situation that we analyze is a severe drop in the credit-worthiness of one particular institution. Assuming that the market works properly, this would result in a sharp increase of the CDS spreads for the company in question, since the market participants expect its default and require higher risk premia. Such an increase, in turn, would be reflected in large residuals of the fitted time series models for the log returns of the CDS spreads and thus in quantiles of the respective distributions that are close to one. This means that we are able to directly work on the copula level and not on the original level of the data. For our simulation study we assume that the variable of interest, U_r, $r \in \{1, ..., 38\}$, takes on the predefined quantile value of $u_r = 0.99$. Given this stress situation, we then use the methods developed in Section 5.2 to simulate the impact on the remaining institutions in terms of quantiles of their innovations distributions. That is, we draw observations from the distribution of $\boldsymbol{U}_{-r}|U_r = u_r$. This simulation is repeated

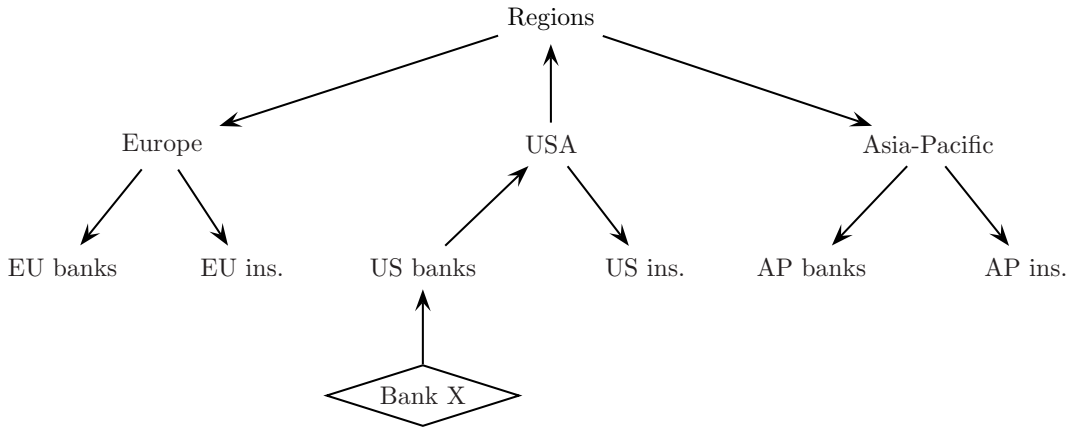


Figure 5.5: Illustration of contagion among the institutions if Bank X is in a stress situation.

$N' = 10\,000$ times for each institution and for the different copulas that have been fitted in the previous section, where we exclude the Frank copula and the hierarchical Kendall copula with Frank cluster and nesting copulas due to their weak fits. Especially conditional sampling from the hierarchical Kendall copulas works exactly as shown in Figure 5.5 (see also Figure 5.1), which illustrates the contagion effect among sectors and regions. This underlines the rationale of using a hierarchical dependence model here. For the hierarchical Kendall copula with Gaussian cluster and nesting copulas, we use rejection-like sampling as described in Section 4.2.4. The sampling error is chosen as $\varepsilon = 10^{-2}$, since this value was shown to already yield accurate sampling results in the simulation study in Section 4.4. We denote the observations sampled conditionally on institution r being stressed by $\widetilde{u}_{k,j|r},\ j \in \{1, ..., 38\} \setminus \{r\},\ k = 1, ..., N'$.

As an illustration, Figure 5.6 shows the mean impact per sector and region in the case of JP Morgan Chase and Hartford Financial Services being stressed: For sector s (within a specific region) with members M_s we define

$$\widetilde{\mu}_{s|r} := \frac{1}{N'} \sum_{k=1}^{N'} \widetilde{u}_{k,s|r}, \tag{5.15}$$

where

$$\widetilde{u}_{k,s|r} := \frac{1}{|M_s \setminus \{r\}|} \sum_{j \in M_s \setminus \{r\}} \widetilde{u}_{k,j|r}, \quad k = 1, ..., N'.$$

This gives an indication which sectors are most strongly influenced by stress to institution r. Of course, this is only informative if the underlying copula is non-exchangeable: The results obtained when using the (exchangeable) Gumbel copula are shown here only for comparison, since they imply that each sector is impacted in the same way, an obviously incorrect statement. As expected, JP Morgan Chase most strongly impacts the US banking sector and Hartford Financial Services the US insurance sector as shown by the non-exchangeable copulas. Interestingly, a stress to JP Morgan Chase influences the US banking sector as well as European banks and insurers similarly strongly, while the

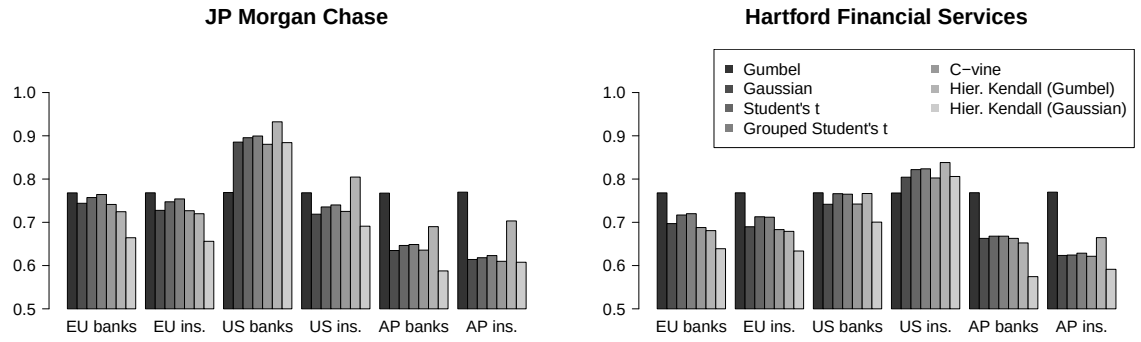


Figure 5.6: Mean impact $\widetilde{\mu}_{s|r}$ (5.15) per sector and region according to different copulas in the case of JP Morgan Chase (left panel) and Hartford Financial Services (right panel) being stressed. An impact of 0.5 corresponds to independence, since this is the mean of a uniform random variable.

impact of stress to Hartford Financial Services is stronger on US banks than on European institutions. This underlines the previous statement that US banks play a systemically important role in the global financial market.

Comparing the results of the non-exchangeable copulas, we observe that the tail dependence implied by the standard and by the grouped Student's t copula increases the mean values in comparison to the Gaussian case. The outcomes according to both variants of the Student's t copula are quite similar. The C-vine copula yields results that are about the same as for the Gaussian copula, although the C-vine copula also accounts for a reasonable degree of tail dependence. This indicates that the Student's t copulas possibly overestimate the tail dependence, as also the grouped Student's t copula still lacks some flexibility here, which is however provided by the C-vine copula. The hierarchical Kendall copula with Gumbel cluster and nesting copulas refines the inappropriate results of the exchangeable Gumbel copula. As noted above, it can be considered as a worst case model, which is confirmed by the rather large outcomes. On the other hand, the outcomes according to the hierarchical Kendall copula with Gaussian cluster and nesting copulas, which gives a quite good fit to the data, are mostly lower than those of the other non-exchangeable copulas. For further insights, we now concentrate on this hierarchical Kendall copula, which has a straightforward interpretation, and on the C-vine copula, which is the best-fitting non-hierarchical copula.

To investigate the question which of the sectors in the market is most systemic, we use both copulas to compute the mean impact of one sector s_1 on another s_2 as

$$\widetilde{\mu}_{s_2|s_1} := \frac{1}{|M_{s_1}|} \sum_{r \in M_{s_1}} \widetilde{\mu}_{s_2|r}, \tag{5.16}$$

which is shown in Table 5.3. Although the numbers according to the two copulas are somewhat different, the resulting picture is essentially the same. The values confirm the previous findings about the systemic role of US banks and also show that the impact of a stressed bank is, in general, often stronger than that of a stressed insurer. This is quite

		Stress situation in					
		EU banks	EU ins.	US banks	US ins.	AP banks	AP ins.
C-vine copula							
Impact on	EU banks	0.87	0.83	0.73	0.67	0.65	0.62
	EU ins.	0.83	0.87	0.72	0.68	0.66	0.64
	US banks	0.74	0.73	0.88	0.72	0.63	0.60
	US ins.	0.68	0.69	0.73	0.79	0.64	0.62
	AP banks	0.65	0.66	0.63	0.64	0.68	0.69
	AP ins.	0.63	0.65	0.61	0.62	0.69	0.62
Hierarchical Kendall copula (Gaussian)							
Impact on	EU banks	0.88	0.79	0.66	0.63	0.58	0.59
	EU ins.	0.79	0.87	0.66	0.63	0.58	0.59
	US banks	0.68	0.68	0.88	0.69	0.58	0.59
	US ins.	0.65	0.64	0.69	0.79	0.56	0.57
	AP banks	0.61	0.60	0.59	0.57	0.68	0.62
	AP ins.	0.63	0.63	0.61	0.59	0.63	0.63

Table 5.3: Mean impact $\widetilde{\mu}_{s_2|s_1}$ (5.16) per sector (rows) of another sector being stressed (columns) according to the C-vine copula and the hierarchical Kendall copula with Gaussian cluster and nesting copulas.

interesting in light of the argumentation of the Geneva Association (2010) claiming that insurers should not be treated as being similarly systemic as banks.

Finally, we move to the question of a possible classification of systemically important institutions. Here, we concentrate on the two largest sectors: European banks and insurers with eleven and twelve members, respectively. Among these institutions we not only consider the mean impacts, $\widetilde{\mu}_{\text{EU-banks}|r}$ and $\widetilde{\mu}_{\text{EU-ins.}|r}$, but also the corresponding confidence intervals to better assess the differences in the conditional simulations. For this we compute empirical quantiles from $\widetilde{u}_{\ell,s|r}$ (see Equation (5.15)). The results are shown in Figure 5.7. According to our analysis the systemically most important banks are Barclays, Banco Santander, BNP Paribas, Banco Bilbao Vizcaya Argentaria and Unicredit. The ranking of insurers is slightly different depending on the chosen dependence model, but the following six institutions can be identified as the most systemic ones: Allianz, Aviva, Assicurazioni Generali, Zurich Insurance, Aegon and Swiss Re. Nevertheless, the differences among the simulated values are quite small and the confidence intervals largely overlap. Also the two considered models yield quite similar results here.

It should be noted that, by the time of this analysis, Banco Bilbao Vizcaya Argentaria had not been officially classified as systemically important by the Financial Stability Board (2011). The 2012 classification (Financial Stability Board, 2012) included the bank and is therefore in line with our analysis. This however indicates that either a systemic risk analysis should not solely be based on CDS spreads or that the classification of the Financial Stability Board (2011, 2012) does not appropriately take into account the observed interdependence among default probabilities as reflected by CDS spreads.

This partly answers the question whether CDS spreads are actually useful for systemic

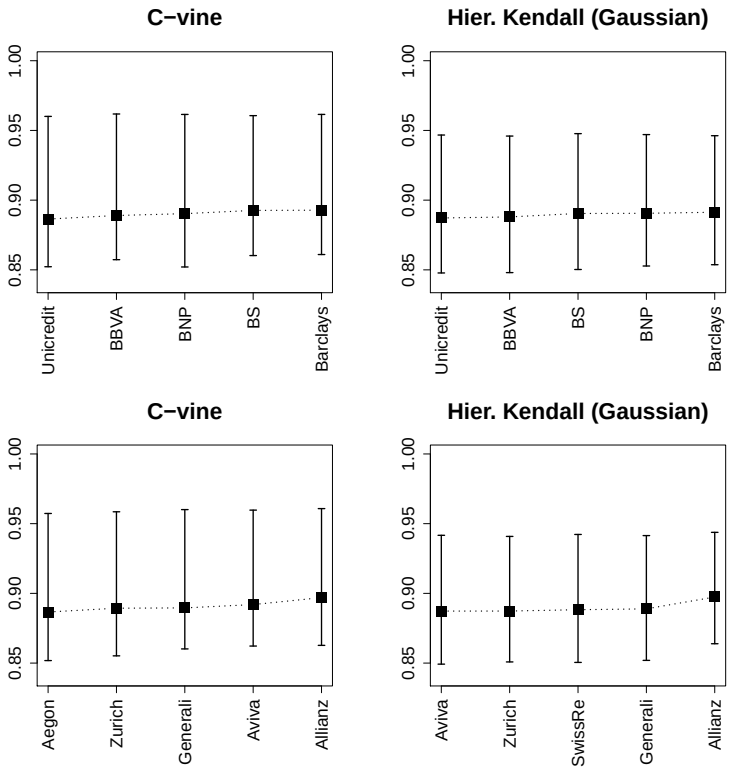


Figure 5.7: Top five banks (top row) and insurers (bottom row) impacting the European banking and insurance sectors, respectively, in case of a stress situation according to the C-vine copula and the hierarchical Kendall copula with Gaussian cluster and nesting copulas. Mean values and 50% confidence bounds are shown.

risk analysis. As a market-based indicator of the credit worthiness of an institution they certainly contain important information to be taken into account. However, we found that dependencies in the CDS market are strongly driven by geographical regions, which hinders a global classification of systemically important institutions. The removal of this geographical dependence in a copula framework is a prerequisite for further attempts to classify institutions using CDS spreads and subject of ongoing research (see Krupskii and Joe (2013a) for a recent copula-based approach to factor modeling, which exploits the C-vine structure; more details can be found in Chapter 7).

5.4 Conclusion

We propose a copula-based approach to the analysis of interdependencies among financial institutions for systemic risk measurement. For the purpose of stress testing the market, we develop necessary conditional simulation procedures. In particular, we derive new methods for Archimedean and for C-vine copulas as well as for the individual Student's t copula. The case of hierarchical Kendall copulas is also discussed and an appropriate conditional simulation algorithm is presented.

The application of these techniques in the analysis of the CDS spreads of 38 major international banks and insurers gives new insights into their interconnectedness and the closely related question of systemic relevance. In the dependence analysis we find evidence of non-elliptical structures, which are important to take into account in stress situations. We also find that banks are systemically more important than insurers. Particularly US banks strongly influence the international financial market. The question whether CDS spreads are actually useful for systemic risk analysis cannot be answered entirely: As a market-based indicator of the credit worthiness of an institution, they contain important information. However, they should not be the sole source of information for the assessment of systemic relevance.

We finally also take a first step towards a classification of institutions according to the performed stress test. It should nevertheless be kept in mind that the results also depend on the selected sample, although it already includes major institutions of the global financial market. The proposed methodology, especially the stress testing approach, is however not limited to the presented case study, but can easily be applied to other relevant data. A major purpose of such investigations certainly should be the further assessment and classification of systemically important institutions according to some appropriate systemic risk measure (see, e.g., Adrian and Brunnermeier (2011), Acharya et al. (2011), and Bernard et al. (2013)).

6 Operational risk measurement

Besides systemic risk assessment, a sound operational risk management is another important issue that financial institutions are currently facing. In this chapter, we develop and evaluate a flexible model for quantitative operational risk measurement, which explicitly deals with data scarcity and heterogeneous pairwise (tail) dependence of losses. By considering flexible families of copulas, we can easily move beyond modeling bivariate dependence among losses and estimate the total risk capital for the seven- and eight-dimensional distributions of event types and business lines. The chapter is based on Brechmann, Czado, and Paterlini (2013).

6.1 Introduction

The magnitude of operational losses observed in recent years and their potential systemic effects has pointed out the need for the development of realistic and therefore often more sophisticated quantitative risk management models (see Basel Committee on Banking Supervision (2009b)). Among the main challenges in operational risk modeling, we have the presence of very heterogeneous losses, usually classified in a matrix of 56 risk classes (eight business lines (BLs) $\times$ seven event types (ETs); see Table 6.1 and Basel Committee on Banking Supervision (2006)), scarcity of data and large numbers of zero losses for some classes, short time series with extreme tails and the need to estimate quantiles at very high confidence levels. In fact, banks are required to calculate the minimum capital requirement as the 99.9%-Value-at-Risk of the loss distribution such that

$$\text{MCR} = \text{VaR}_{99.9\%}\left(\sum_{j=1}^{56} L_j\right), \tag{6.1}$$

where L_j is the aggregate loss of one of the 56 BL-ET combinations. It is clear that this quantity is influenced by the dependencies among the different risk classes. The standard approach of the Basel Committee on Banking Supervision (2006) requires banks to marginally calculate the risk capital of all 56 BL-ET combinations and then determine the risk capital as the sum of these 56 figures, that is,

$$\text{MCR}^{\text{Basel}} = \sum_{j=1}^{56} \text{VaR}_{99.9\%}\left(L_j\right). \tag{6.2}$$

This corresponds to the assumption of comonotonicity (perfect positive dependence) among all 56 BL-ET combinations, which is often perceived by banks as a worst-case scenario for the MCR, assuming that $\text{MCR}^{\text{Basel}} \geq \text{MCR}$. The question if the standard

	No.	Description
BL	1	Corporate Finance
	2	Trading and Sales
	3	Retail Banking
	4	Commercial Banking
	5	Payment and Settlement
	6	Agency and Custody
	7	Asset Management
	8	Retail Brokerage
ET	1	Internal Fraud
	2	External Fraud
	3	Employment Practices & Workplace Safety
	4	Clients, Products & Business Practices
	5	Damage to Physical Assets
	6	Business Disruption & System Failures
	7	Execution, Delivery & Process Management

Table 6.1: Business lines (BLs) and event types (ETs) according to the Basel Committee on Banking Supervision (2006).

Basel approach is too conservative has therefore been raised many times by practitioners and researchers.

Regulators allow then, with due diligence, explicit dependence modeling under the Advanced Measurement Approach (AMA): "*Risk measures for different operational risk estimates must be added for purposes of calculating the regulatory minimum capital requirement. However, the bank may be permitted to use internally determined correlations in operational risk losses across individual operational risk estimates, provided it can demonstrate to the satisfaction of the national supervisor that its systems for determining correlations are sound, implemented with integrity, and take into account the uncertainty surrounding any such correlation estimates (particularly in periods of stress). The bank must validate its correlation assumptions using appropriate quantitative and qualitative techniques*" (Basel Committee on Banking Supervision, 2006, §669d). The supervisory guidelines for the AMA banks (Basel Committee on Banking Supervision, 2011b) underline that dependence modeling for operational risk is an evolving area, where many approaches are currently used by banks with potential impact on the capital requirements. Results from the 2008 Loss Data Collection Exercises (Basel Committee on Banking Supervision, 2009a) suggest that among the AMA banks only 17% use correlation coefficients, while most rely on copulas (43%), with a preference for Gaussian copulas, and 31% AMA banks use other methods.

Research is high then on the regulators agenda to avoid spurious differences in exposure estimates (see Basel Committee on Banking Supervision (2011b, §224)) and to provide sound guidelines for dependence modeling in operational risk, explicitly suggesting to move beyond Gaussian copulas and correlations coefficients. Theoretical and empirical evidence so far has mostly supported the idea that the assumption of perfect positive

dependence is unduly strong and using internally determined correlations, as Basel II allows, could lead to lower the risk capital requirements while still providing adequate coverage for future losses (see Artzner et al. (1999), Chapelle et al. (2008) and Frachot et al. (2004)). However, recently Mittnik et al. (2013) have also shown that, despite only for a small number of risk classes, modeling bivariate dependence could also lead to increase the risk capital and tail dependence should not be ignored. What is the total impact on risk capital of explicit dependence modeling among all BLs and ETs is still, to our knowledge, a question with no answer, as estimating realistic multivariate operational risk management models with more than two cells can be computationally challenging and data are often scarce for model validation.

In this work, we aim at analyzing how much the risk capital estimate may change when modeling dependencies in multivariate settings. That is, we consider the total impact of dependence modeling within the eight- or seven-dimensional BL and ET distributions by introducing a statistical model, which allows to explicitly consider the presence of extreme tails, heterogeneous pairwise dependence and large numbers of zero observations. In particular, we propose a flexible approach that, inspired by the work of Deb et al. (2013) on drug expenditures and Erhardt and Czado (2012) on dependent health insurance claims, directly models the dependence between the aggregate losses in BL-ET combinations using copulas. Since the non-occurrence of losses (zero events) also conveys information about the dependence characteristics, we explicitly incorporated it to allow a more accurate assessment of dependence. Finally, given that no excessive data aggregation is required, parameter estimation can be based on the maximum amount of available observations.

By using real-world data from the Italian Database of Operational Losses (DIPO) in the period from January 2003 to June 2011, we can provide new and much needed insights on the impact of different dependence modeling strategies on total capital requirements and their validation on real-world data. In fact, our results suggest that explicitly modeling dependence and zero inflation can lead to a reduction, as often expected, of the total regulatory capital, which might turn out to be up to 38% smaller than what the Basel comonotonicity approach would prescribe.

The chapter is organized as follows. Section 6.2 introduces the modeling framework by discussing separately the key components of our modeling strategy. Marginal modeling of operational losses is briefly treated in Section 6.3, while Section 6.4 is devoted to a discussion of relevant properties of major copula classes considered appropriate for operational losses: We consider elliptical, Archimedean and vine copulas as well as the individual Student's t copula. Modeling dependence among zero events as additional model component is described in Section 6.5, and the computation of risk measures to obtain operational risk capital figures is subsequently discussed in Section 6.6. Section 6.7 finally provides the results of the empirical investigation on real-world data from Italian banks, while Section 6.8 concludes.

6.2 Zero-inflated dependence model

Common characteristics of operational risk data can be summarized as follows: First, if losses are modeled on a weekly or monthly basis, it may frequently occur that there

are no losses observed for a particular BL or ET. An excessive number of zero losses is called *zero inflation*. Second, another important marginal property is that distributions of operational losses per BL or ET may be *heavy-tailed*. This means that there is a significant probability of extreme losses that has to be taken into account. Third, different BLs and ETs are not independent. Most importantly, the type of dependence, especially the *tail dependence*, may have a huge impact on risk capital estimates.

Since the heavy tails in the marginal distributions of operational losses have already been extensively discussed in the literature (see, e.g., Chavez-Demoulin et al. (2006) and Gourier et al. (2009)), we focus on the first and the third point and discuss appropriate zero-inflated dependence models for aggregate operational losses.

Suppose that we want to model the multivariate distribution of d BLs, ETs or cells of the 8×7 BL-ET matrix. For brevity, we henceforth always speak of d cells with $d \in \{7, 8, 56\}$. Let $L_j \geq 0,\ j = 1, ..., d$, denote the aggregate loss of the jth cell. Then, we explicitly model the presence of zero inflation in the aggregate loss by defining the following binary random variable $Y_j \sim P_{Y_j}$ for each cell $j \in \{1, ..., d\}$ as

$$Y_j := \begin{cases} 1 & \text{if there is a zero loss in cell } j, \\ 0 & \text{otherwise.} \end{cases}$$

This means that Y_j is the zero inflation component of the otherwise positive continuous part of L_j, which we denote by $L_j^+ > 0$. We obtain the following decomposition:

$$L_j = Y_j \times 0 + (1 - Y_j)L_j^+ = (1 - Y_j)L_j^+ \geq 0. \tag{6.3}$$

If $P_{Y_j}(0) = P(Y_j = 0) = 1$, then there is no zero inflation.

Inspired by the work of Deb et al. (2013) and Erhardt and Czado (2012), who build a three-dimensional model for dependent health insurance claims, we introduce the multivariate zero-inflated density of $\boldsymbol{L} := (L_1, ..., L_d)'$ and $\boldsymbol{Y} := (Y_1, ..., Y_d)'$ as

$$\begin{aligned} f_{\boldsymbol{Y},\boldsymbol{L}}(\boldsymbol{y}, \boldsymbol{\ell}) &= p_{\boldsymbol{Y}}(\boldsymbol{y})\, f_{\boldsymbol{L}|\boldsymbol{Y}}(\boldsymbol{\ell}|\boldsymbol{y}) \\ &= p_{\boldsymbol{Y}}(\boldsymbol{y})\, f_{\{L_j^+,\ j \in \mathcal{J}(\boldsymbol{y})\}}(\ell_j,\ j \in \mathcal{J}(\boldsymbol{y})), \end{aligned} \tag{6.4}$$

where $\boldsymbol{y} := (y_1, ..., y_d)' \in \{0, 1\}^d$, $\boldsymbol{\ell} := (\ell_1, ..., \ell_d)' \in \mathbb{R}_{\geq 0}^d$ and

$$\mathcal{J}(\boldsymbol{y}) = \{j \in \{1, ..., d\} :\ y_j = 0\}.$$

That is, $\mathcal{J}(\boldsymbol{y})$ contains all indices, for which the respective component of $\boldsymbol{y}$ is equal to 0. In other words, these are the cells with non-zero events. The $|\mathcal{J}(\boldsymbol{y})|$-dimensional density of $L_j^+,\ j \in \mathcal{J}(\boldsymbol{y})$, is denoted by $f_{\{L_j^+,\ j \in \mathcal{J}(\boldsymbol{y})\}}$; if $\mathcal{J}(\boldsymbol{y}) = \emptyset$, then $f_\emptyset := 1$. The multivariate binary probability mass function of $\boldsymbol{Y}$ is denoted by $p_{\boldsymbol{Y}}$. More explicitly, we can write Equation (6.4) also as

$$\begin{aligned} f_{\boldsymbol{Y},\boldsymbol{L}}(\boldsymbol{y}, \boldsymbol{\ell}) = p_{\boldsymbol{Y}}(\boldsymbol{y}) \Big(& 1_{\{\boldsymbol{y}=(1,...,1)'\}} + 1_{\{\boldsymbol{y}=(0,1,...,1)'\}}\, f_{L_1^+}(\ell_1) + 1_{\{\boldsymbol{y}=(1,0,1,...,1)'\}}\, f_{L_2^+}(\ell_2) \\ & + ... + 1_{\{\boldsymbol{y}=(1,...,1,0)'\}}\, f_{L_d^+}(\ell_d) + 1_{\{\boldsymbol{y}=(0,0,1,...,1)'\}}\, f_{L_1^+,L_2^+}(\ell_1, \ell_2) \\ & + 1_{\{\boldsymbol{y}=(0,1,0,1,...,1)'\}}\, f_{L_1^+,L_3^+}(\ell_1, \ell_3) + ... + 1_{\{\boldsymbol{y}=(1,...,1,0,0)'\}}\, f_{L_{d-1}^+,L_d^+}(\ell_{d-1}, \ell_d) \\ & + 1_{\{\boldsymbol{y}=(0,0,0,1,...,1)'\}}\, f_{L_1^+,L_2^+,L_3^+}(\ell_1, \ell_2, \ell_3) + ... + 1_{\{\boldsymbol{y}=(0,...,0)'\}}\, f_{L_1^+,...,L_d^+}(\boldsymbol{\ell}) \Big). \end{aligned}$$

Clearly, only exactly one of the indicator functions is different from zero.

In this way, we separate the dependence of the zero inflation component from the dependence of the positive losses. By applying Sklar's Theorem (2.2) to $f_{\{L_j^+,\ j \in \mathcal{J}(\boldsymbol{y})\}}$, we further obtain

$$f_{\{L_j^+,\ j \in \mathcal{J}(\boldsymbol{y})\}}(\ell_j,\ j \in \mathcal{J}(\boldsymbol{y})) = c_{\{j \in \mathcal{J}(\boldsymbol{y})\}}(F_{L_j^+}(\ell_j),\ j \in \mathcal{J}(\boldsymbol{y})) \prod_{j \in \mathcal{J}(\boldsymbol{y})} f_{L_j^+}(\ell_j),$$

where $c_{\{j \in \mathcal{J}(\boldsymbol{y})\}}$ is the copula density of $C_{\{j \in \mathcal{J}(\boldsymbol{y})\}}$, the $|\mathcal{J}(\boldsymbol{y})|$-dimensional margin of the d-dimensional copula C for all cells. More explicitly, it holds that

$$C_{\{j \in \mathcal{J}(\boldsymbol{y})\}}(u_j,\ j \in \mathcal{J}(\boldsymbol{y})) = C(\boldsymbol{v}),$$

where $u_j \in [0,1]$ for all $j \in \mathcal{J}(\boldsymbol{y})$ and $\boldsymbol{v} := (v_1, ..., v_d)'$ with

$$v_j = \begin{cases} u_j & \text{if } j \in \mathcal{J}(\boldsymbol{y}), \\ 1 & \text{otherwise,} \end{cases} \qquad j = 1, ..., d.$$

Then, it follows that we can state Equation (6.4) in terms of the copula C and its margins as

$$f_{\boldsymbol{Y},\boldsymbol{L}}(\boldsymbol{y}, \boldsymbol{\ell}) = p_{\boldsymbol{Y}}(\boldsymbol{y})\, c_{\{j \in \mathcal{J}(\boldsymbol{y})\}}(F_{L_j^+}(\ell_j),\ j \in \mathcal{J}(\boldsymbol{y})) \prod_{j \in \mathcal{J}(\boldsymbol{y})} f_{L_j^+}(\ell_j), \tag{6.5}$$

where $\boldsymbol{y} \in \{0,1\}^d$ and $\boldsymbol{\ell} \in \mathbb{R}^d_{\geq 0}$. This means that, as in the classical copula approach, we can separate the dependence modeling defined by the copula from the modeling of the marginal distributions. In addition, the multivariate binary distribution $p_{\boldsymbol{Y}}$ has to be modeled to account for the zero inflation component. Appropriate choices for these three model components are discussed in the following. We begin with the marginal distributions (Section 6.3), then turn to the copula of positive losses (Section 6.4) and also treat the zero loss distribution (Section 6.5).

6.3 Marginal modeling

The topic of marginal modeling of operational losses has already been extensively discussed in the literature (see, amongst others, Aue and Kalkbrener (2006)). We therefore only briefly mention the most relevant points.

The aggregate loss L_j is the sum of the individual operational losses within a given cell $j \in \{1, ..., d\}$, that is,

$$L_j = \sum_{m=1}^{N_j} X_{mj},$$

with the number of losses $N_j \sim F_{N_j}$ and the independent and identically distributed individual losses $X_{mj} \sim F_{X_j},\ m = 1, ..., N_j$. The severity distribution F_{X_j} is positive and continuous and the frequency distribution F_{N_j} is a count distribution with support on $\mathbb{N}_0 = \{0, 1, 2, ...\}$. The case $N_j = 0$ means that no losses are observed, that is, $L_j = 0$ (the empty sum is taken to be zero). As this case is explicitly taken into account by the

zero loss variable Y_j, we denote the positive number of losses by $N_j^+ \in \mathbb{N} = \{1, 2, ...\}$ with zero-truncated distribution $F_{N_j^+}$. Therefore, we have

$$L_j^+ = \sum_{m=1}^{N_j^+} X_{mj}.$$

Typical choices for $F_{N_j^+}$ are the zero-truncated Poisson and negative binomial distributions (see, e.g., Grogger and Carson (1991)). As is commonly known, the Poisson distribution assumes equidispersion, that is, it fixes mean and variance to be equal. Since this is a very strict assumption, the negative binomial distribution may often be a more realistic choice.

It is typically the case that operational losses below some threshold are not reported. Therefore, only losses above a certain threshold are observed. In this case, the severity distribution has to be translated to this threshold. Let F_0 be an arbitrary positive continuous distribution, then translation to the threshold $m > 0$ means that we define the severity distribution F_{X_j} for $x_j \in (m, \infty)$ as

$$F_{X_j}(x_j) = \frac{F_0(x_j) - F_0(m)}{1 - F_0(m)}. \tag{6.6}$$

An important issue of the distribution of individual operational losses is that it may be heavy-tailed as noted above. Examples of heavy-tailed (or subexponential) distributions are in particular the log normal, Pareto, Weibull (with shape parameter smaller than 1) and generalized Pareto, while the also commonly used gamma distribution is light-tailed (see Embrechts et al. (1997)). Furthermore, according to a classification based on moments, Pareto tails can be shown to be heavier than log normal tails. An alternative, which is proposed in extreme value theory, is to use the generalized Pareto as explicit tail distribution above a certain threshold, while a different distribution is used for the body of the distribution (see Chavez-Demoulin et al. (2006) for more details).

6.4 Dependence modeling of positive losses

The between-cell copula is central to the model in order to appropriately respect dependencies in operational risk data as stated above. In the simplest setting, the cells are simply assumed to be either independent or perfectly positively dependent (see Example 2.2 and Theorem 2.3). As this is not necessarily the case, the question is what properties a reasonable copula for operational losses should exhibit.

(i) *Pairwise dependence:* Dependence among different pairs of cells may be heterogeneous. Therefore the between-cell copula should be flexible enough to allow for different strengths of dependence for different pairs. In particular, it should be non-exchangeable (see Definition 2.5).

(ii) *Tail dependence:* The between-cell copula should allow for the modeling of tail dependence (see Remark 2.7). The presence of upper tail dependence means that very

large losses tend to occur jointly rather than independently. We therefore focus on upper tail dependence in the following.

(iii) *Interpretability:* Copulas may be specified in terms of many parameters with many different meanings. For reasons of internal and external communication, for example to the senior management or regulators, it is desirable to work with a model that has parameters, whose meaning can actually be interpreted, for example, in terms of Kendall's τ (2.4), Spearman's ρ_S (2.5) or the upper tail dependence coefficient λ_U (2.7) (see Table 2.1 for an overview).

(iv) *Computational tractability:* To work with model (6.5) requires the availability of all multivariate copula margins. This is particularly important for the statistical inference, where the density expression is needed for likelihood-based techniques.

While most of the literature on dependence modeling of operational losses (see Dalla Valle et al. (2008) and Giacometti et al. (2008)) and also practitioners (see Basel Committee on Banking Supervision (2009a)) focus on elliptical copulas, we here more generally evaluate a range of different multivariate copula models in light of the above four characteristics and discuss how useful they are for modeling operational risk data. A summary of the copulas characteristics is provided in Table 6.2 at the end of the discussion.

6.4.1 Elliptical copulas

The class of elliptical copulas is discussed in Section 2.2, while the extension of the standard Student's t copula to multiple degrees of freedom is treated in Section 2.3. This individual Student's t copula is considered in the next section. Here, we look at the properties of standard elliptical copulas.

(i) *Pairwise dependence:* Pairwise dependence of Gaussian and Student's t copulas can be different for each pair. In the correlation matrix $R \in [-1, 1]^{d \times d}$ each pair has a corresponding entry.

(ii) *Tail dependence:* While the Gaussian copula does not exhibit any tail dependence, the Student's t copula has symmetric upper and lower tail dependence (see Equation (2.25)).

(iii) *Interpretability:* As noted above, pairwise dependence is essentially determined by the correlation matrix, whose entries can be directly related to Kendall's τ (see Equations (2.21) and (2.24)). Correlation parameters of the Gaussian copula can also conveniently be interpreted in terms of Spearman's ρ_S (see Equation (2.22)). The tail dependence coefficient of the Student's t copula of cells j and k depends on the corresponding correlation parameter ρ_{jk} and the degrees of freedom $\nu > 2$ (see Equation (2.25)). In other words, parameters of elliptical copulas are well interpretable.

(iv) *Computational tractability:* The margins of elliptical copulas are again elliptical of the same class. The correlation matrix is a sub-matrix of the full correlation

matrix. Although easily tractable density expressions are available, the statistical inference is complicated by the fact that the correlation matrix has to be positive definite. Good starting values for numerical optimization are often provided by computing pairwise empirical Kendall's τ estimates and then inverting Equation (2.21) or (2.24), respectively.

The above properties present elliptical copulas as a rather appealing model for multivariate operational losses. Their major disadvantage is however certainly that the Gaussian copula does not exhibit any tail dependence, while the tail dependence of the Student's t copula is symmetric in both tails and governed by only one parameter for all pairs.

To illustrate the model defined in Equation (6.5) in terms of elliptical copulas, we consider a trivariate example. If $d = 3$, then Equation (6.5) can be written as

$$\begin{aligned} f_{\boldsymbol{Y},\boldsymbol{L}}(\boldsymbol{y},\boldsymbol{\ell}) = p_{\boldsymbol{Y}}(\boldsymbol{y}) \Big(& 1_{\{\boldsymbol{y}=(1,1,1)'\}} + 1_{\{\boldsymbol{y}=(0,1,1)'\}} f_{L_1^+}(\ell_1) \\ & + 1_{\{\boldsymbol{y}=(1,0,1)'\}} f_{L_2^+}(\ell_2) + 1_{\{\boldsymbol{y}=(1,1,0)'\}} f_{L_3^+}(\ell_3) \\ & + 1_{\{\boldsymbol{y}=(0,0,1)'\}} c_{1,2}(F_{L_1^+}(\ell_1), F_{L_2^+}(\ell_2)) f_{L_1^+}(\ell_1) f_{L_2^+}(\ell_2) \\ & + 1_{\{\boldsymbol{y}=(0,1,0)'\}} c_{1,3}(F_{L_1^+}(\ell_1), F_{L_3^+}(\ell_3)) f_{L_1^+}(\ell_1) f_{L_3^+}(\ell_3) \\ & + 1_{\{\boldsymbol{y}=(1,0,0)'\}} c_{2,3}(F_{L_2^+}(\ell_2), F_{L_3^+}(\ell_3)) f_{L_2^+}(\ell_2) f_{L_3^+}(\ell_3) \\ & + 1_{\{\boldsymbol{y}=(0,0,0)'\}} c(F_{L_1^+}(\ell_1), F_{L_2^+}(\ell_2), F_{L_3^+}(\ell_3)) f_{L_1^+}(\ell_1) f_{L_2^+}(\ell_2) f_{L_3^+}(\ell_3) \Big) . \end{aligned}$$

Hence, dependence is modeled by the trivariate copula C with bivariate margins $C_{1,2}$, $C_{1,3}$ and $C_{2,3}$. If C is an elliptical copula with correlation matrix $R = (\rho_{jk})_{j,k=1,2,3} \in [-1,1]^{3\times 3}$, then each bivariate margin $C_{j,k}$, $1 \leq j < k \leq 3$, is also elliptical of the same class and has a 2×2-correlation matrix with parameter ρ_{jk}. In case of the Student's t copula, each margin also has the same degrees of freedom as C.

To overcome the restrictions with respect to tail dependence even of the Student's t copula, we will consider two appropriate extensions: the individual Student's t copula and vine copulas. In addition, Archimedean copulas are also discussed for comparison.

6.4.2 Individual Student's t copula

The individual Student's t copula (see Section 2.3) generalizes the standard Student's t copula such that each variable has its own degrees of freedom parameter. This adds flexibility for appropriately modeling the tails, as discussed in the following.

(i) *Pairwise dependence:* As an extension of the standard Student's t copula, the pairwise dependence of the individual Student's t copula can also vary between different pairs. The dependence of a pair of variables is determined in terms of the corresponding entry in the correlation matrix and by the two parameters for the degrees of freedom of the variables.

(ii) *Tail dependence:* Due to the reflection symmetry, the tail dependence of the individual Student's t copula is also symmetric in both tails. In contrast to the standard Student's t copula, it is decisively determined by the degrees of freedom of the individual variables (see Equation (2.30)) and therefore more flexible.

(iii) *Interpretability:* Kendall's τ is approximately the same as for standard elliptical distributions (see Equation (2.29)) and therefore straightforward to interpret. The tail dependence coefficients however do not possess a simple closed-form expression (see Equation (2.30)), so that the interrelationship of the parameters and the strength of tail dependence is not obvious in the first place (see Figure 2.5).

(iv) *Computational tractability:* Despite its more sophisticated structure, the individual Student's t copula is still computationally tractable. The margins are simply individual Student's t copulas with corresponding degrees of freedom parameters and sub-matrix of the full correlation matrix. Furthermore, the multivariate density only involves a one-dimensional integration (see Equation (2.28)).

The statistical inference of the individual Student's t copula is therefore feasible also in higher dimensions. To obtain good starting values for numerical optimization, it is convenient to use inverted pairwise empirical Kendall's τ values as for elliptical copulas. For the individual degrees of freedom parameters, we propose to preliminarily fit bivariate individual Student's t copulas for each pair and then take the average estimated degrees of freedom of each variable as starting value. In our numerical examples, this typically provided reasonably good starting values, which sped up the numerical optimization.

The individual Student's t copula hence extends the standard Student's t copula at the critical point: Individual degrees of freedom parameters for each variable allow for a more flexible range of tail dependence of the different pairs, while the assumption of one common parameter for the degrees of freedom can be very restrictive if larger numbers of variables are considered. On the other hand, the individual Student's t copula is harder to interpret in terms of its parameters. The statistical inference is also more difficult due to a one-dimensional integration in the density expression.

6.4.3 Archimedean copulas

Archimedean copulas are introduced in Section 2.4. Their properties are:

(i) *Pairwise dependence:* Due to the exchangeability, each margin of an Archimedean copula is again of the same copula type with the same parameter. That is, pairwise dependence is fixed to be the same for all pairs.

(ii) *Tail dependence:* Depending on the chosen generator function, Archimedean copulas can exhibit different tail behavior (see Examples 2.15–2.18). For instance, the Gumbel copula is upper tail dependent, while the Frank does not have any tail dependence.

(iii) *Interpretability:* Parameters of Archimedean copulas can typically be expressed in terms of Kendall's τ. Closed-form expressions for the tail dependence coefficients in terms of the parameters are often also available (see Table 2.1).

(iv) *Computational tractability:* The major advantage of Archimedean copulas is that all margins are readily available and the copula often only depends on one or two

parameters. This strongly simplifies the statistical inference (see also Hofert et al. (2012)).

Clearly, because of the insufficient flexibility in modeling heterogeneous pairwise dependence, Archimedean copulas are ruled out as reasonable models for multivariate operational losses when moving beyond the bivariate case.

6.4.4 Vine copulas

Vine copulas (see Section 2.7) are a completely different approach to construct multivariate copulas and can also be used to generalize the Student's t copula, as we will discuss below. First, we go through the list of desirable properties.

(i) *Pairwise dependence:* Due to the flexibility in the choice of the pair copulas in the decomposition, different pairs of cells can have very different dependence structures such as asymmetry or tail dependence.

(ii) *Tail dependence:* Joe et al. (2010) show that for each pair of cells to have tail dependence, it is sufficient for the unconditional bivariate copulas to have tail dependence. That is, if for instance all $d-1$ copulas of unconditional pairs (pairs in the first vine tree) are specified as Student's t, then each pair of cells is tail dependent.

(iii) *Interpretability:* Interpretation of dependence patterns of vine copulas is complicated by the fact that most pairs are specified conditionally. Only for the dependence of the $d-1$ pairs in the first vine tree, which are specified unconditionally, interpretation is as simple as for the multivariate copulas discussed previously. However, simulation from vine copulas is very simple so that properties can be assessed empirically based on sufficiently large simulated data sets.

(iv) *Computational tractability:* The statistical inference of vine copulas is in principle rather straightforward, since the density of a vine copula is conveniently given in terms of a product of bivariate copulas (see Equation (2.66)). For log likelihood calculations, this nicely transforms into a sum of log copula densities.

To be useful for our multivariate operational loss model defined in Equation (6.5), we however also require the margins of vine copulas. The availability of the multivariate margins is unfortunately a major issue of vine copulas. Some margins are available in closed form, such as those of the $d-1$ unconditional pairs in the first vine tree T_1. Nevertheless, in most cases integration is needed to compute the margins. This is especially true for the bivariate margins of the $(d-1)(d-2)/2$ conditional pairs in vine trees $T_2, ..., T_{d-1}$. For example, in a three-dimensional vine copula with pair copulas $C_{1,2}$, $C_{1,3}$ and $C_{2,3;1}$ (see the example in Section 2.7.1, in particular Equation (2.60)), the bivariate margin of the variables 2 and 3 is given for $(u_2, u_3)' \in [0,1]^2$ by

$$\begin{aligned} c_{2,3}(u_2, u_3) &= \int_0^1 c(u_1, u_2, u_3)\, du_1 \\ &= \int_0^1 c_{1,2}(u_1, u_2)\, c_{1,3}(u_1, u_3)\, c_{2,3;1}(C_{2|1}(u_2|u_1), C_{3|1}(u_3|u_1))\, du_1. \end{aligned}$$

	Archim.	Gaussian	Student's t	Indiv. t	Vine
Pairwise dependence	–	+	+	+	+
Tail dependence	∘	–	∘	+	+
Interpretability	+	+	+	+	∘
Comput. tractability	+	+	+	∘	–

Table 6.2: Overview of copula characteristics: positive (+), neutral (∘), and negative (–).

In the worst case, for the pair of cells $\{j, k\}$ in the last vine tree T_{d-1}, which has the conditioning set $\{1, ..., d\} \setminus \{j, k\}$, $(d-2)$-dimensional integration may be needed to compute the corresponding bivariate margin. This renders the use of vine copulas in model (6.5) hardly feasible even if d is only as large as 5 or 6. Only if there are few zero events, margins may stay sufficiently well tractable. In our application in Section 6.7, we will show how to calibrate a seven-dimensional vine copula for operational losses per ET.

Although the building blocks of vine copulas can be of arbitrary type, we focus here on elliptical pair copulas. On the one hand, this narrows the wide range of possible constructions (see Section 2.7.4). On the other hand, this allows us to define an extension of the Student's t copula, which stays interpretable for the following reason: As noted in Section 2.7.2, Stöber et al. (2013) show that a Student's t copula can be represented as a vine copula, where the parameters of the pair copulas are obtained as partial correlations and degrees of freedom that are increased by one for each additional conditioning variable. The Student's t copula can therefore be generalized by a vine copula with Student's t pair copulas, where each bivariate Student's t copula is allowed to have different numbers of degrees of freedom. Since this results in a model with a large number of parameters (two parameters per pair copula), a simpler model can be constructed by only choosing Student's t copulas for the $d-1$ unconditional pairs and Gaussian copulas for all conditioned pairs. According to Joe et al. (2010), this construction also has tail dependence for all pairs. A further model simplification could be achieved by truncating the vine copula as discussed by Brechmann et al. (2012). The focus on elliptical pair copulas also means that the simplifying assumption of PCCs (see Section 2.7.2) can be regarded as not overly restrictive here.

6.4.5 Hierarchical copulas

The BLs and ETs shown in Table 6.1 can each be divided into sub-categories (see Basel Committee on Banking Supervision (2006)). This induces a natural hierarchy, which can be accounted for using a hierarchical dependence model as discussed in Chapter 3. Especially hierarchical Kendall copulas (see Definition 3.3) can provide the required flexibility in terms of the above discussed properties. However, as losses per (sub-)category are summed up for risk capital calculations, it would probably be more reasonable to work with a different aggregation function than the copula (see Section 3.2). One such approach for hierarchical risk capital aggregation was recently explored by Arbenz et al. (2012). Here, we do not further follow the approach of a hierarchical modeling, since data

availability today is typically a problem even at the coarser level of BLs and ETs, so that a detailed hierarchical dependence analysis is not feasible yet. Also in our application the information about the BL or ET sub-category of a loss is not available to us.

6.5 Dependence modeling of zero losses

The random vector $\boldsymbol{Y}$ of zero inflation components is multivariate binary, for which distributions are however rather non-standard and often require an excessive number of parameters (see, e.g., Johnson et al. (1997)). We propose to use a copula approach here, which we first illustrate in a bivariate example. For this, let $\boldsymbol{Y} = (Y_1, Y_2)'$ and recall that $P(Y_j = y_j) = P(Y_j \leq y_j) - P(Y_j \leq y_j - 1) = P_{Y_j}(y_j) - P_{Y_j}(y_j - 1),\ j = 1, 2$. Similarly, it holds for the bivariate probability mass function $p_{\boldsymbol{Y}}$ that

$$\begin{aligned}
p_{\boldsymbol{Y}}(\boldsymbol{y}) &= P(Y_1 = y_1, Y_2 = y_2) \\
&= P(Y_1 \leq y_1, Y_2 \leq y_2) - P(Y_1 \leq y_1 - 1, Y_2 \leq y_2) \\
&\quad - P(Y_1 \leq y_1, Y_2 \leq y_2 - 1) + P(Y_1 \leq y_1 - 1, Y_2 \leq y_2 - 1) \\
&= C_{\boldsymbol{Y}}(P_{Y_1}(y_1), P_{Y_2}(y_2)) - C_{\boldsymbol{Y}}(P_{Y_1}(y_1 - 1), P_{Y_2}(y_2)) \\
&\quad - C_{\boldsymbol{Y}}(P_{Y_1}(y_1), P_{Y_2}(y_2 - 1)) + C_{\boldsymbol{Y}}(P_{Y_1}(y_1 - 1), P_{Y_2}(y_2 - 1)),
\end{aligned}$$

where we used Sklar's Theorem (2.1) with an appropriate bivariate copula $C_{\boldsymbol{Y}}$. In general, $p_{\boldsymbol{Y}}$ can be represented as

$$p_{\boldsymbol{Y}}(\boldsymbol{y}) = \sum_{k_1=1}^{2} \ldots \sum_{k_d=1}^{2} (-1)^{k_1 + \ldots + k_d} C_{\boldsymbol{Y}}\left(u_1^{(k_1)}, \ldots, u_d^{(k_d)}\right), \tag{6.7}$$

where $u_j^{(1)} = P_{Y_j}(y_j)$ and $u_j^{(2)} = P_{Y_j}(y_j - 1)$ for $j = 1, \ldots, d$ (see Song (2007, Section 6.3.2)). For binary margins P_{Y_j}, it is either $y_j = 0$ or $y_j = 1$. If $y_j = 0$, then $u_j^{(1)} = P_{Y_j}(0)$ is the probability of a non-zero loss and $u_j^{(2)} = P_{Y_j}(-1) = 0$. Conversely, if $y_j = 1$, then $u_j^{(1)} = P_{Y_j}(1) = 1$ and $u_j^{(2)} = P_{Y_j}(0)$.

The copula $C_{\boldsymbol{Y}}$ can be any d-dimensional copula. As before, we recommend to use a copula that allows for heterogeneous pairwise dependence, which makes Archimedean copulas of no interest here. A vine copula, on the other hand, generally does not have a closed-form copula expression. This also rules out vine copulas (see Panagiotelis et al. (2012) for an alternative, PCC-based approach to modeling multivariate discrete data). Finally, to evaluate Equation (6.7), 2^d evaluations of the copula are needed. This may be very time-consuming so that also the individual Student's t copula, whose copula expression is of more complicated form than that of standard elliptical copulas, is typically not a good choice. The Gaussian copula may be used in moderate dimensions, since efficient algorithms for the numerical evaluation of $C_{\boldsymbol{Y}}$ are available (see Genz and Bretz (2009)). To evaluate a Student's t copula, the dimension should however be rather small. Parameters can be estimated using maximum likelihood techniques.

6.6 Operational risk capital

A major purpose of a multivariate model for operational losses is, of course, an accurate assessment of the regulatory risk capital to be held to cover future losses. The standard risk measure for computing the operational risk capital under Basel II is the Value-at-Risk (VaR) at the 99.9% level (see Equation (6.1)). In general, for a level $\alpha \in [0,1]$ the $(1-\alpha)$-VaR is defined as

$$\mathrm{VaR}_{1-\alpha}(L) = F_L^{-1}(1-\alpha), \tag{6.8}$$

where $L = \sum_{j=1}^{d} L_j$ is the total operational loss over d cells and F_L its continuous distribution function (see, e.g., McNeil et al. (2005)). Since F_L is not known in closed form, it has to be approximated by simulation. To simulate N' losses $\ell_{kj},\ k = 1, ..., N'$, for each cell $j \in \{1, ..., d\}$ using our multivariate model (6.5), we proceed as follows.

(i) Sample $u_{kj},\ k = 1, ..., N',\ j = 1, ..., d$, from the copula C for positive losses. For elliptical, Archimedean and vine copulas see Mai and Scherer (2012), for the individual Student's t copula see Luo and Shevchenko (2010).

(ii) Set $\ell_{kj}^{+} := F_{L_j^{+}}^{-1}(u_{kj}) > 0$ for $k = 1, ..., N'$ and $j = 1, ..., d$.

(iii) Sample $y_{kj} \in \{0,1\},\ k = 1, ..., N',\ j = 1, ..., d$, from the copula $C_{\boldsymbol{Y}}$ for zero losses and using the marginal distribution functions $P_{Y_j},\ j = 1, ..., d$.

(iv) Set $\ell_{kj} := (1 - y_{kj})\ell_{kj}^{+} \geq 0$ for $k = 1, ..., N'$ and $j = 1, ..., d$, according to Equation (6.3).

In this way, we generate a sample of size N' of the total operational loss, which is given by $\ell_k := \sum_{j=1}^{d} \ell_{kj},\ k = 1, ..., N'$. The $(1-\alpha)$-quantile $F_L^{-1}(1-\alpha)$, that is, the $(1-\alpha)$-VaR, can then be approximated by the corresponding empirical quantile of $\ell_1, ..., \ell_{N'}$.

6.7 Application: Operational losses of Italian banks

Having discussed our multivariate model for operational losses in detail, we now evaluate it based on a data base of losses of Italian banks.

6.7.1 Data

Our data set comprises operational losses reported from 33 Italian banking groups with about 180 entities to the Italian Database of Operational Losses (DIPO)[1]. The time period that we consider is from January 2003 to June 2011 for a total of 102 months or 451 weeks. The reporting threshold is 5000 Euro, below which no loss is reported. Further, it is known for each loss which ET and BL are affected.

Figure 6.1 displays pie charts for the proportion of aggregate losses and of numbers of losses per BL and per ET (see Table 6.3). It shows that the losses per cell are rather

[1]We are thankful to Claudia Pasquini, Claudia Capobianco and Vincenzo Buggè from DIPO and its Statistical Committee for their support. The views expressed in this chapter are those of the authors and do not necessarily reflect the viewpoints of DIPO or the DIPO Statistical Committee.

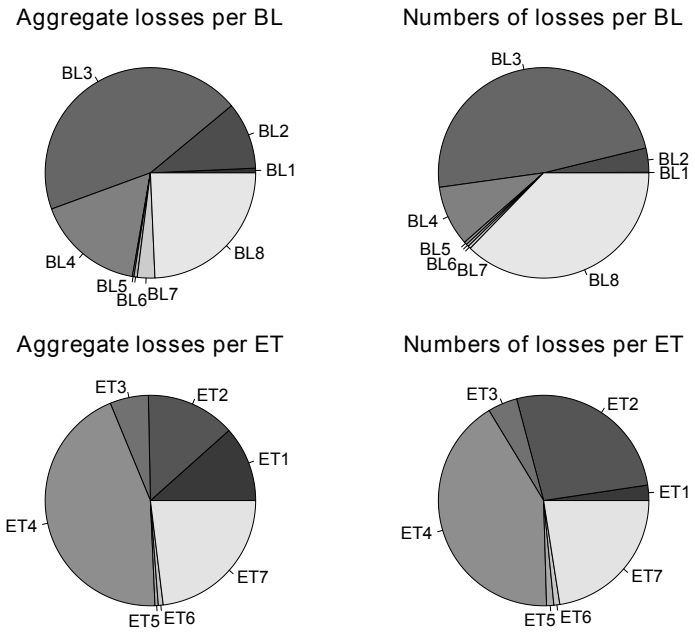


Figure 6.1: Proportion of aggregate losses L_j (left column) and numbers of losses N_j (right column) per BL (top row) and per ET (bottom row).

Agg.	ET1	ET2	ET3	ET4	ET5	ET6	ET7	Σ ET
BL1	0.00	0.01	0.01	0.49	0.00	0.01	0.20	0.72
BL2	0.70	1.19	0.05	5.22	0.00	0.19	2.96	10.30
BL3	5.53	10.73	5.41	11.22	0.41	0.32	10.95	44.57
BL4	1.63	1.65	0.20	9.27	0.08	0.07	3.78	16.67
BL5	0.00	0.02	0.01	0.02	0.00	0.03	0.16	0.24
BL6	0.06	0.01	0.03	0.02	0.00	0.01	0.39	0.52
BL7	0.01	0.01	0.07	2.32	0.00	0.03	0.24	2.69
BL8	3.68	0.10	0.14	15.91	0.01	0.09	4.37	24.31
Σ BL	11.60	13.71	5.92	44.46	0.51	0.75	23.05	

Freq.	ET1	ET2	ET3	ET4	ET5	ET6	ET7	Σ ET
BL1	0.00	0.00	0.01	0.02	0.00	0.00	0.06	0.10
BL2	0.03	0.01	0.02	2.13	0.01	0.12	1.40	3.71
BL3	1.45	23.74	4.16	6.74	0.82	0.41	11.07	48.39
BL4	0.06	2.75	0.15	2.35	0.27	0.10	3.59	9.27
BL5	0.01	0.02	0.02	0.03	0.00	0.05	0.30	0.44
BL6	0.01	0.04	0.02	0.04	0.00	0.01	0.26	0.38
BL7	0.01	0.02	0.02	0.07	0.01	0.03	0.27	0.44
BL8	0.79	0.19	0.15	30.38	0.00	0.19	5.57	37.26
Σ BL	2.36	26.77	4.54	41.77	1.11	0.91	22.52	

Table 6.3: Proportion (in %) of aggregate losses L_j (upper table) and numbers of losses N_j (lower table) for each BL-ET combination.

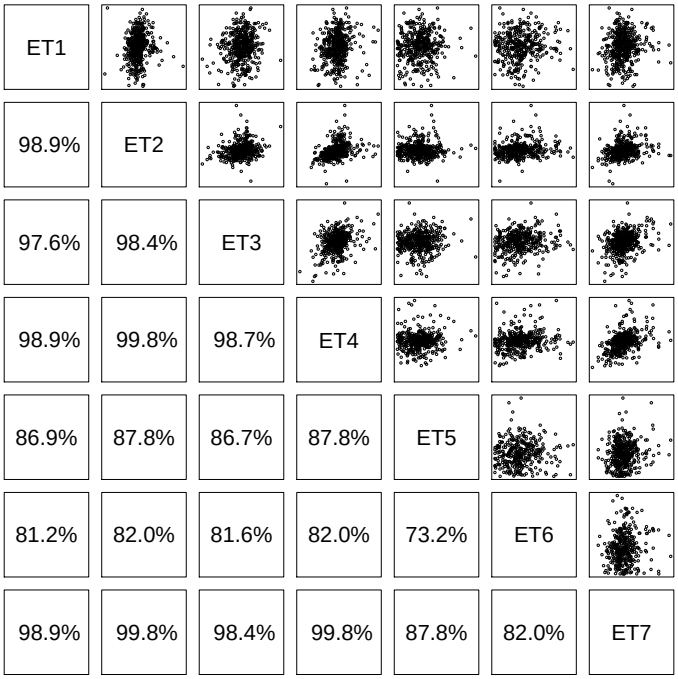


Figure 6.2: Pairwise scatter plots of aggregate losses L_j^+ per ET on the log scale. The lower triangle shows the percentage of pairwise complete observations of (L_j^+, L_k^+). Labels are omitted from the axes for confidentiality reasons.

heterogeneous. That is, operational losses in certain BLs and ETs occur more often and are of different magnitude. Especially Retail Banking (BL3) and Retail Brokerage (BL8) are often subject to operational losses, while the most frequent event types are External Fraud (ET2), Execution, Delivery & Process Management (ET7) and, in particular, Clients, Products & Business Practices (ET4).

We choose here to model on a weekly basis to balance the trade-off between having sufficiently many observations in each cell to avoid a large number of zero losses and also to being able of an accurate marginal and dependence modeling. In our investigation, we will both model the multivariate dependence between the eight BLs and between the seven ETs, to evaluate also the impact of such choice on the total risk capital estimate. Both approaches are followed in practice (see Basel Committee on Banking Supervision (2009a)), but so far no investigation showed the implications behind them. Figure 6.2 shows pairwise scatter plots of the aggregate losses of the seven ETs on the log scale. The general level of dependence appears to be rather weak, as the observations are rather dispersed, but there are also cases with seemingly stronger dependence such as between ETs 4 and 7. The dependence between positive losses is investigated in more detail in Section 6.7.3. Figure 6.2 also indicates that there is a considerable number of zero events that needs to be taken into account explicitly (see Section 6.7.4). First, we discuss an appropriate modeling of the margins.

6.7.2 Marginal modeling

For the estimation of weekly operational losses per BL and ET we consider the following frequency and severity distributions: zero-truncated Poisson, generalized Poisson (see Consul and Jain (1973)) and negative binomial for the loss frequency, and gamma, Weibull, log normal, Pareto and generalized Pareto for the loss severity. Each severity distribution is translated to the reporting threshold of 5000 Euro (see Equation (6.6)).

Using QQ-plots and goodness-of-fit tests, it turns out that negative binomial fits for the loss frequency are quite good and always superior to the respective Poisson fits postulating equidispersion, which cannot be observed with the available data. Generalized Poisson fits, which also allow for non-equidispersion, give similar results as the negative binomial, so that we decide to use the latter model.

In an exploratory pre-analysis, we also determine that significantly fewer losses are observed in three weeks in August. We account for this holiday season effect by fitting appropriate mean regressions with indicator variable for these three weeks. More precisely, let $f_{N_j^+}$ denote the probability mass function of the zero-truncated negative binomial distribution for the number of losses of a given cell j. Then,

$$f_{N_j^+}(n_j) = \frac{\Gamma(n_j + m_j)}{\Gamma(m_j)\, n_j!} \left(\frac{m_j}{m_j + \mu_{kj}}\right)^{m_j} \left(\frac{\mu_{kj}}{m_j + \mu_{kj}}\right)^{n_j} \left(1 - \left(\frac{m_j}{m_j + \mu_{kj}}\right)^{m_j}\right)^{-1}, \quad n_j \in \mathbb{N},$$

with season-dependent mean parameter $\mu_{kj} > 0$ and size $m_j > 0$. We model μ_{kj} as

$$\mu_{kj} = \exp\left(\beta_{0j} + \beta_{1j} 1_{\text{summer}}(k)\right),$$

where $\beta_{0j}, \beta_{1j} \in \mathbb{R}$ denote regression parameters, and the indicator variable 1_{summer} is defined as

$$1_{\text{summer}}(k) := \begin{cases} 1 & \text{if observation } k \text{ occurs in one of the three summer weeks,} \\ 0 & \text{otherwise.} \end{cases}$$

Hence, we have two different frequency distributions for each BL and ET, depending on whether a loss occurs in summer or not. Severities are not observed to be different in summer compared to the rest of the year.

For the loss severities, we also use QQ-plots and goodness-of-fit tests to determine the best fitting distributions per BL and ET. From the above list of distributions the log normal (translated to the minimum reported loss of 5000 Euro) gives the best fit for the individual positive losses. Figure 6.3 shows that the fitted log normal distribution functions, in fact, very closely follow the empirical distribution functions of the different BLs. A look at the very tail of the distributions (see Figure 6.4) underlines this. Although there may be a potential underestimation of the tails, the tails are generally only moderately heavy so that log normal fits are appropriate. For ETs similar results hold.

In the next step, we then compute the convolution of the chosen severity and frequency distributions for each BL and ET by Monte Carlo simulation with sample size 100 000 (see, e.g., Klugman et al. (2008)). For each BL and ET we obtain two convolutions: one for the three summer weeks with significantly fewer losses and one for the rest of the year. This provides us with estimates $\widehat{F}_{L_j^+}$ and $\widehat{F}^s_{L_j^+}$ of the marginal aggregate loss

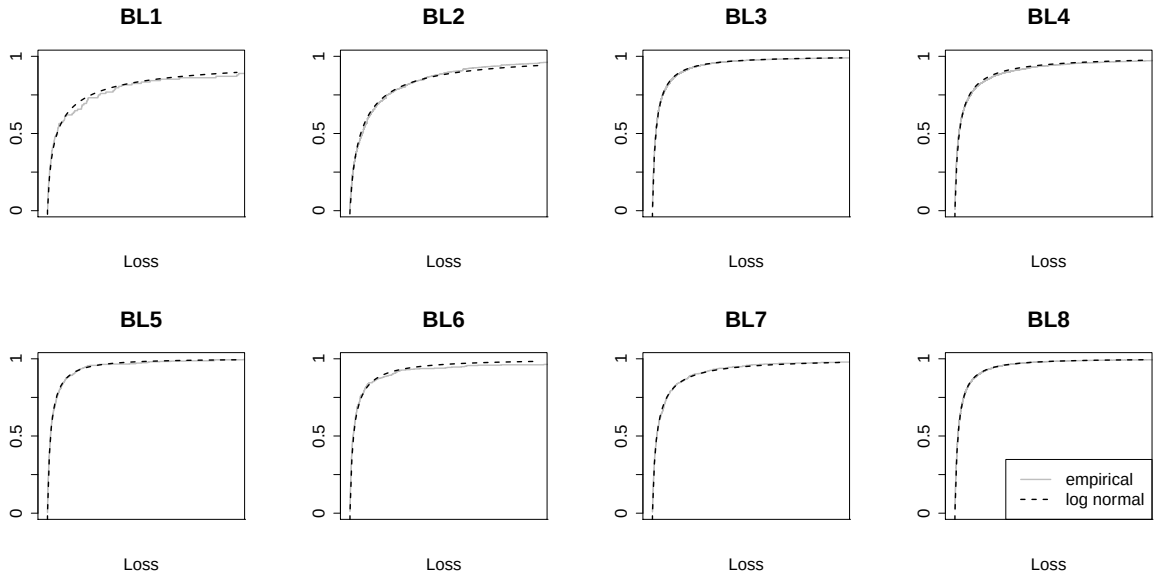


Figure 6.3: Comparison of empirical and fitted log normal distribution functions for all eight BLs and the main range of the individual losses X_{mj}. Labels are omitted from the horizontal axes to maintain confidentiality.

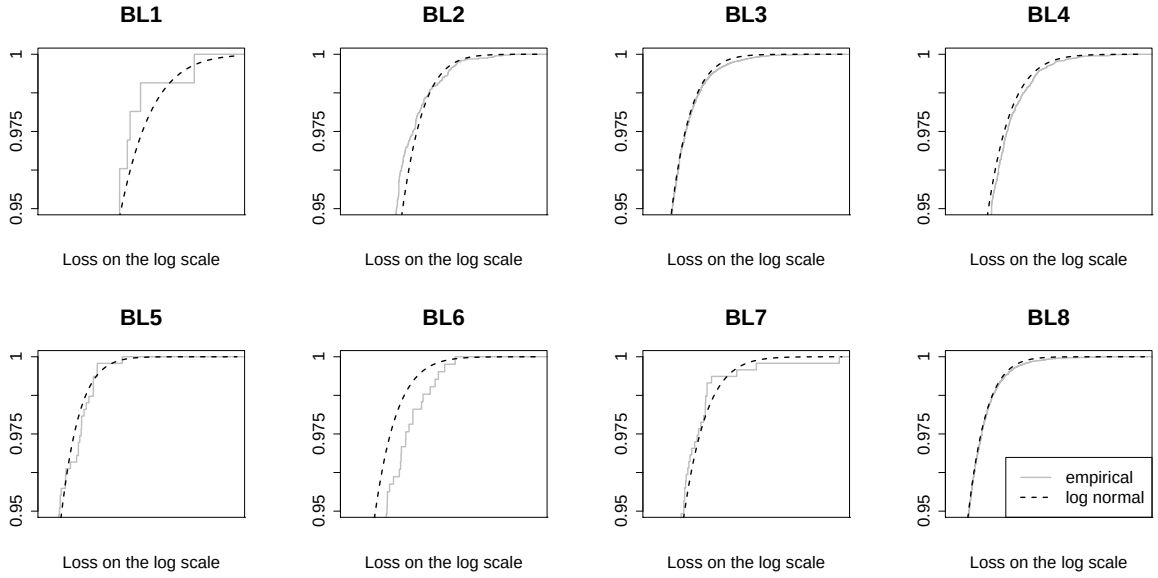


Figure 6.4: Comparison of empirical and fitted log normal distribution functions on the log scale for all eight BLs in the upper tail of the individual losses X_{mj}. Labels are omitted from the horizontal axes to maintain confidentiality.

ET\BL	1	2	3	4	5	6	7	8
1		-0.14	0.02	-0.07	0.08	0.18	0.05	-0.10
2	0.12		0.12	0.16	0.00	0.06	0.10	0.20
3	0.08	0.15		0.23	0.01	0.12	0.10	0.23
4	0.10	0.30	0.16		0.01	0.04	0.06	0.22
5	0.04	-0.01	0.06	0.04		-0.06	0.12	0.07
6	0.05	0.13	0.10	0.12	0.05		0.18	0.06
7	0.07	0.20	0.18	0.29	0.07	0.08		0.10

Table 6.4: Empirical Kendall's τ values of jointly observed pairs of positive aggregate losses (L_j^+, L_k^+) per BL (upper triangle) and per ET (lower triangle).

distribution functions $F_{L_j^+}$ and $F^s_{L_j^+}$, respectively, where the latter indicates the one for the three summer weeks. (Note that we have $j = 1, ..., 8$ for BLs and $j = 1, ..., 7$ for ETs.) These estimates are used to transform the observed aggregate losses $\ell_{kj},\ k = 1, ..., 451$, to approximately uniform data following the estimation method of inference functions for margins (IFM) described in Section 3.5. That is, we set

$$\widehat{u}_{kj} := \begin{cases} \widehat{F}^s_{L_j^+}(\ell_{kj}) & \text{if observation } k \text{ occurs in one of the three summer weeks,} \\ \widehat{F}_{L_j^+}(\ell_{kj}) & \text{otherwise.} \end{cases}$$

Following the IFM approach, these pseudo observations are then used in the dependence analysis.

6.7.3 Dependence modeling of positive losses

We then fit the copula models for positive losses of Section 6.4. As Archimedean copulas we choose the Frank, which does not exhibit any tail dependence (see Example 2.17), and the Gumbel, which has upper but no lower tail dependence (see Example 2.16). Upper tail dependence is of particular interest here, because it describes the joint probability of very large losses and therefore needs to be accounted for in risk capital calculations (see Section 6.7.5). General dependence is however rather weak between BLs and ETs, respectively: Pairwise empirical Kendall's τ values between BLs range from -0.14 to 0.23 and between ETs from -0.01 to 0.30 (see Table 6.4). This is in line with previous studies (see Dalla Valle et al. (2008), Cope and Antonini (2008), and Giacometti et al. (2008)) and clearly different from comonotonicity as postulated by the standard approach (see Equation (6.2)).

As described in Section 6.4, the use of vine copulas in our multivariate model for operational losses defined in Equation (6.5) is quite challenging, since multivariate margins are not available in closed form but involve possibly high-dimensional integration. We therefore check both for BLs and ETs which combinations of non-zero events are actually observed, that is, which margins need to be evaluated. For ETs these are the combinations {3,4}, {1,2,4,7}, {1,2,3,4,7}, {1,2,4,5,7}, {1,2,3,4,5,7}, {1,2,3,4,6,7}, {2,3,4,5,6,7},

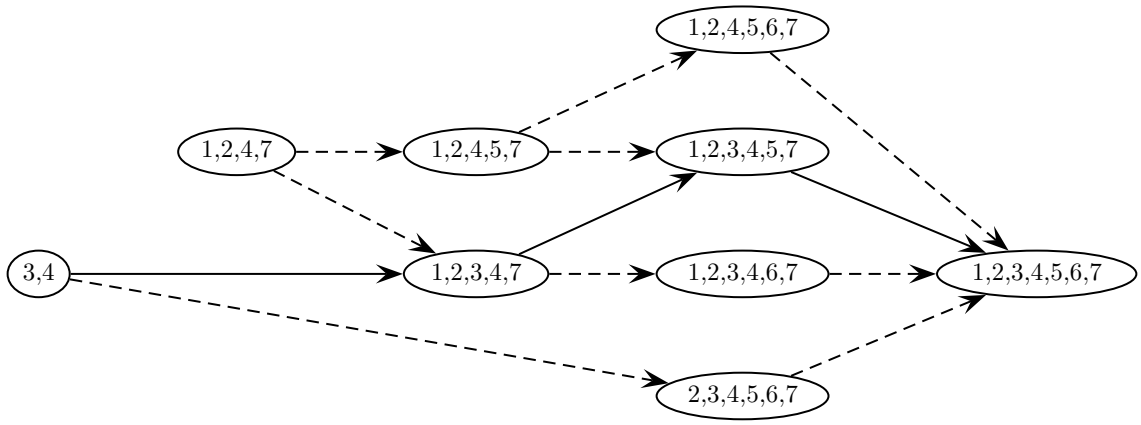


Figure 6.5: Combinations of observed non-zero events in ETs, where arrows indicate that a set is the subset of another.

{1,2,4,5,6,7} and {1,2,3,4,5,6,7}. For a vine copula to be tractable for the statistical inference, it is important to construct it such that only low-dimensional integrations are required for the margins. Such a vine copula can be determined according to the following tailor-made selection procedure, which is in contrast to the general selection approaches mentioned in Section 2.7.4 (see, in particular, Algorithm 2.25) and used in Sections 3.7.3 and 5.3.3.

From the conditional density decomposition (see Equations (2.55) and (2.61)) it is clear that a d-dimensional PCC can always be extended to a $(d+1)$-dimensional one by adding the term $f_{d+1|1,\dots,d}(x_{d+1}|x_1,\dots,x_d)$ to the existing PCC (see the Example in Section 2.7.1). This means that we can construct the vine copula by starting with a bivariate one and then iteratively extending it to the seven-dimensional one, which is needed to describe the dependence between all ETs. For this, we check which set of jointly observed ETs is a subset of another as displayed in Figure 6.5. One option then is to start with a PCC for {3,4}, which is simply a bivariate copula. It is reasonable to extend it to {1,2,3,4,7}, which is the smallest set of which {3,4} is a subset. Then, we choose {1,2,3,4,5,7} and finally {1,2,3,4,5,6,7} as indicated by the solid lines in Figure 6.5. The reason why we choose {1,2,3,4,5,7} instead of {1,2,3,4,6,7}, of which {1,2,3,4,7} is also a subset, is that {1,2,4,5,7}, which is not explicitly included in the PCC, is a subset of {1,2,3,4,5,7}. In this way, every multivariate margin that is not an explicit sub-model of the PCC is a subset of a set that is included and has only one element more. For the vine copula, this means that only one-dimensional integration is needed to integrate out this one additional element. For instance, the margin of {1,2,4,7} is obtained by integrating out variable 3 from the sub-model for {1,2,3,4,7}, since $\{1,2,4,7\} \subset \{1,2,3,4,7\}$ and $\{1,2,3,4,7\} \setminus \{1,2,4,7\} = \{3\}$.

To summarize, we found a PCC for ETs such that only one-dimensional integration is needed to evaluate the multivariate operational loss model (6.5). This is similar to the individual Student's t copula. Unfortunately, the best possible vine PCC for BLs still requires up to three-dimensional integration, which is numerically very demanding. We therefore do not consider a vine copula for BLs. For ETs, we however fit two different R-vine copulas: one with only Student's t pair copulas (model 'T') and one with Student's

	BL modeling			ET modeling		
Copula	Log lik.	# Par.	AIC	Log lik.	# Par.	AIC
Gumbel	29.14	1	-56.27	51.99	1	-101.99
Frank	36.62	1	-71.23	70.96	1	-139.92
Gaussian	92.92	28	-129.83	107.14	21	-172.28
Student's t	96.01	29	-134.02	111.32	22	-178.65
Indiv. Student's t	105.35	36	-138.70	119.50	28	-183.01
R-vine (T/N)	-	-	-	112.80	27	-171.61
R-vine (T)	-	-	-	121.39	42	-158.78

Table 6.5: Log likelihoods, numbers of parameters and AIC values of the copulas for dependence among BLs and ETs estimated by maximum likelihood.

t copulas for all unconditional variable pairs (first tree) and Gaussian copulas for all pairs that are specified conditionally in the remaining trees (model 'T/N').

Log likelihoods, numbers of parameters as well as AIC values of the considered models for BLs and ETs are shown in Table 6.5. While both Archimedean copulas are obviously not appropriate for the operational risk data per BL or ET, elliptical copulas provide a pretty good fit. Nevertheless, the individual Student's t copula provides an even better fit than the standard Student's t copula. This indicates that tail dependencies are not only present but also quite heterogeneous, implying that the standard Student's t copula with only one parameter for the degrees of freedom is too restrictive: The estimated numbers of degrees of freedom of the standard Student's t copula are 42.04 for BLs and 54.78 for ETs, while the estimated individual degrees of freedom of the individual Student's t copula vary between 6.30 and more than 300 for BLs and between 7.62 and also more than 300 for ETs. The standard Student's t copulas therefore create the false impression that dependence is almost Gaussian. The large estimated numbers of degrees of freedom can be seen as averages of the respective estimated individual degrees of freedom, which are rather heterogeneous. Entries of the correlation matrices range between -0.23 and 0.31 for BLs and between 0 and 0.33 for ETs, corresponding to a rather weak to medium level of dependence. Computing tail dependence coefficients for these parameters however results in values very close to 0. This is a very interesting result, since the use of the two Student's t copulas allows for a quantification of the strength of the tail dependence, which the non-tail dependent Gaussian copula does not allow for.

The two vine copulas for dependence between ETs as alternative extension of the standard Student's t copula also improve the log likelihood. This also stresses the need for a flexible tail dependence modeling, but, as before, estimated tail dependence coefficients are very small and close to 0. In contrast to the standard and the individual Student's t copulas, the vine copulas strongly suffer from the large number of model parameters, so that AIC values are better for the elliptical copulas and the individual Student's t copula. The number of parameters could be reduced by setting Student's t copulas with, for example, more than 100 degrees of freedom to Gaussian copulas. This would hardly impact the log likelihood value, but significantly reduce the number of parameters. If a

correlation parameter is very small and close to independence, even independence copulas could be used to reduce the number of parameters even further (see Sections 3.7.3 and 5.3.3). Also the individual Student's t copula could be simplified to some extent, either by grouping variables with similar numbers of degrees of freedom (the resulting model being the grouped Student's t copula by Daul et al. (2003); see Section 2.3) or by removing the mixing variables in the definition of the individual Student's t copula (see the representation in Equation (2.27)) for components with a large number of degrees of freedom.

6.7.4 Dependence modeling of zero losses

As discussed in Section 6.5, essentially any copula could be used to construct a flexible multivariate binary distribution for the zero losses, but due to computational and other limitations we concentrate here on a Gaussian copula with correlation matrix R^{BL} or R^{ET}, respectively. Empirical marginal probabilities of zero losses per BL range between 0 and 0.81 in the first BL, while for ETs the largest number of zero losses is observed for the sixth ET with about 20%. These empirical probabilities are used as parameters of the marginal Bernoulli distributions.

In a pre-analysis, it turns out that many of the empirical pairwise probabilities of zero losses are very small (below 0.01). We therefore restrict our copula analysis to those BLs and ETs with significant non-zero pairwise probabilities. These are the BLs 1, 5, 6 and 7 and the ETs 3, 5 and 6 (see Figure 6.2). Hence, we fit a four-dimensional Gaussian copula for zero losses of BLs and a three-dimensional one for the ETs. Estimated entries of the correlation matrix $R^{\mathrm{BL}}_{1,5,6,7}$ for BLs range between 0.02 and 0.28,

$$\widehat{R}^{\mathrm{BL}}_{1,5,6,7} = \begin{pmatrix} 1 & 0.07 & 0.07 & 0.07 \\ 0.07 & 1 & 0.02 & 0.16 \\ 0.07 & 0.02 & 1 & 0.28 \\ 0.07 & 0.16 & 0.28 & 1 \end{pmatrix},$$

and of the correlation matrix $R^{\mathrm{ET}}_{3,5,6}$ for ETs between 0.05 and 0.51,

$$\widehat{R}^{\mathrm{ET}}_{3,5,6} = \begin{pmatrix} 1 & 0.05 & 0.51 \\ 0.05 & 1 & 0.19 \\ 0.51 & 0.19 & 1 \end{pmatrix}.$$

Correlations to and between the other BLs and ETs are set to zero.

6.7.5 Operational risk capital

After carefully modeling the dependence between positive and zero losses for BLs and ETs, we evaluate the different models in terms of their risk capital estimates and assess the diversification benefit compared to comonotonicity. In particular, we concentrate on the Gaussian, the Student's t, the individual Student's t and the vine copula with only Student's t copulas, since they provide the best fit. They differ mainly in if and how tail dependencies are modeled. Those are however found to be very small for both BL and

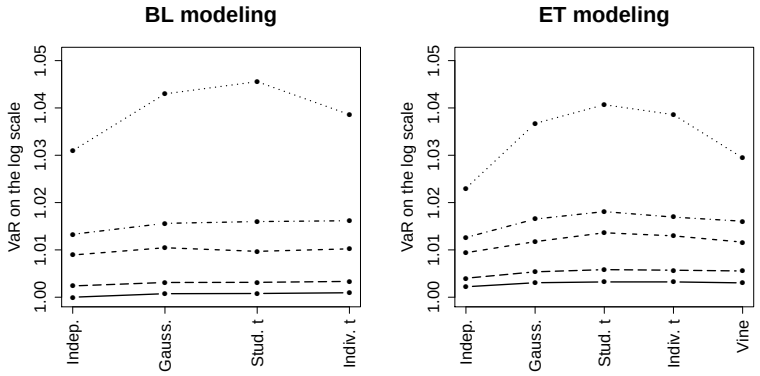


Figure 6.6: Annual $(1-\alpha)$-VaRs on the log scale for $\alpha = 10\%, 5\%, 9\%, 0.5\%, 0.1\%$ (line types: solid to dotted) and for the different copula models for positive losses. The left panel shows the results for BL modeling, the right for ET modeling. Vertical axes have been scaled to maintain confidentiality.

ET dependence. For comparison, a model with independence between positive and zero losses is also used.

We generate 100 000 annual observations from the different models (see Section 6.6), that is, we simulate each week of the year respecting the different marginal distributions in summer and then sum over the weeks. Based on these observations we compute annual VaRs (6.8) for different levels as shown in Figure 6.6. Obviously, risk measures are lowest when no dependence among BLs and ETs is assumed. Using a Student's t copula as underlying dependence model however results in the highest VaR estimates, while its extensions, the individual Student's t and the vine copula with only Student's t copulas, lead to smaller estimates. This is very interesting, since a standard Student's t copula is often chosen instead of a Gaussian copula to respect tail dependence, but it apparently overestimates the required risk capital. This is certainly due to the restrictive modeling approach of allowing only one parameter to control the overall level of tail dependence. Although tail dependence here is very small, its impact is obvious and a more accurate modeling of it, using extensions of the Student's t copula, yields refined estimates of risk measures. Comparing the BL and the ET modeling shows that the picture is essentially the same for both. This is reassuring given that the modeling is based on the same data, which only has been grouped differently.

The second question of interest is the diversification effect of considering the VaR of the total loss rather than the sum of the individual VaRs (see Equations (6.1) and (6.2)). That is, we are interested in the fraction

$$\mathrm{Div}_{1-\alpha} := \frac{\mathrm{VaR}_{1-\alpha}(\sum_{j=1}^{d} L_j) - \sum_{j=1}^{d} \mathrm{VaR}_{1-\alpha}(L_j)}{\sum_{j=1}^{d} \mathrm{VaR}_{1-\alpha}(L_j)}, \tag{6.9}$$

which gives the relative reduction in the VaR. If $\mathrm{Div}_{1-\alpha} < 0$, there is a diversification benefit.

The diversification effect for the copula models considered here is displayed in Figure 6.7. There is a clear diversification benefit of up to 38% for BLs and 32% for ETs when

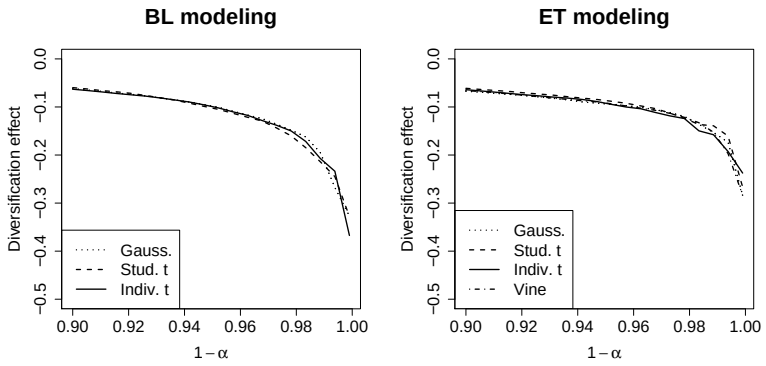


Figure 6.7: VaR diversification effect $\text{Div}_{1-\alpha}$ (6.9) for $\alpha \in [0.1\%, 10\%]$ and for the different copula models for positive losses. The left panel shows the results for BL modeling, the right for ET modeling.

explicitly modeling the dependence. This mirrors the fact that dependence between BLs and ETs, both in general and in the tails, is observed to be rather weak and therefore clearly differs from comonotonicity, where $\text{Div}_{1-\alpha} = 0$. Interestingly, there is no obvious difference between the different models. This indicates that the diversification effect is mainly driven by the general level of dependence modeled, which is about similar for the different models.

6.8 Conclusion

In this work, we introduce a rather general and flexible multivariate modeling approach for operational risk losses, which explicitly takes into account the multivariate dependence among losses and the presence of scarce data. Our aim is to introduce a more accurate model and then evaluate its implications on the estimation of the total risk capital compared to the Basel II comonotonicity assumption for the entire set of BLs and ETs using real-world data. Explicit dependence modeling is discussed critically by considering different copula classes and introducing, from both statistical and business perspectives, four key characteristics the ideal model should allow to take into account: heterogeneous pairwise dependence, tail dependence, interpretable parameters and computational tractability. It turns out that from a theoretical perspective, the individual Student's t copula is probably the most appealing model for operational losses, as it only poses a moderate computational challenge, while still being easily interpretable in its parameters and structure and allowing for the presence of different pairwise (tail) dependencies. Compared to the Student's t copula with only one degrees of freedom parameter, tail dependence is determined by the degrees of freedom of the individual variables and therefore flexible in considering potentially heterogeneous behaviors among losses. Vine copulas add even more flexibility, but at the expense of a more complex parameter interpretability and computational tractability, which make it less appealing in our zero-inflated model framework. Gaussian copulas are an option when it is reasonable to assume tail independence, while Archimedean copulas, despite used in practice, are of less interest for operational

risk modeling as they cannot account for heterogeneous pairwise dependence.

Empirical results on real-world data suggest that Gaussian and even better Student's t copulas can provide a good fit to positive losses grouped by BLs or ETs, despite the Gaussian implies tail independence and the Student's t copula results in a potential overestimation of tail dependence due to its inflexible degrees of freedom parameter. However, as expected, the individual Student's t copula shows the best fit in terms of AIC and log likelihood to the data at hand and allows to build an easily interpretable model for operational losses, from which to run Monte Carlo estimation to determine the overall effect on risk capital. In fact, while the Student's t copula results in the highest VaR estimates for both BL and ET modeling, with an increase with respect to the independence assumption of up to 35% and 43% for BLs and ETs, respectively, individual Student's t copula estimates result only in an increase of 17% and 37%. This reflects the more accurate assessment of tail dependence by the individual Student's t copula, which can account for heterogeneous tail dependence through the individual degrees of freedom per variable, of which the standard Student's t copula's degrees of freedom can be regarded as an average.

Finally, when considering the diversification ratio to evaluate the effect of a potential reduction of risk capital estimates compared to the standard Basel assumption of comonotonic losses, results suggest that a more realistic modeling of the multivariate distribution of BLs and ETs leads to a reduction in capital of up to 38% for BLs and 32% for ETs for high quantiles ($\alpha = 99.9\%$, with little differences between models). This confirms, as also previously reported in the literature for much simpler bivariate settings, that the comonotonicity assumption of the standard Basel framework is unduly strong when evaluating the overall effect. Banks might therefore have an incentive to move towards more sophisticated but then realistic and accurate risk management models. Nevertheless, further investigations on different real-world data and larger sample sizes for out-of-sample evaluation are still required to draw irrefutable conclusions.

With increasing availability of operational loss data, it will also be possible to investigate approaches to set up a 56-dimensional model for all BL-ET combination. In addition, hierarchical risk capital aggregation across BL and ET sub-categories will need to be considered in order to build a comprehensive dependence model for operational losses.

7 Conclusion and outlook

This thesis deals with approaches to hierarchical dependence modeling. The approach using aggregation via copulas and associated Kendall distribution functions is developed and analyzed in detail and we discuss why we think that it is a reasonable and useful approach. Nevertheless, our method to construct hierarchical copulas is only one possibility among many others to handle high-dimensional dependencies. Therefore, we would like to point out two potential avenues of future research. One is an alternative approach to define hierarchical copulas using a factor model. The other relates the developed methodology to the research on multivariate return periods. We begin with the latter.

Multivariate return periods: The notion of a return period is commonly used, for example, in hydrology or geophysics to quantify the riskiness of an event such as a flood or an earthquake. It is defined as the average time between two realizations of this specific event.

In the univariate case, let $X_1 \sim F_1$ and let $\mathcal{E}_1$ be the event of interest. Assuming that large values of X_1 are considered as critical, we define $\mathcal{E}_1 = (x_\mathcal{E}, \infty)$ with $x_\mathcal{E} \in \mathbb{R}$. That is, we analyze the return period of the exceedance of $x_\mathcal{E}$, which may be, for example, the height of a dike. The event $\mathcal{E}_1$ has probability $p_{\mathcal{E}_1} = P(X_1 > x_\mathcal{E}) = 1 - F_1(x_\mathcal{E})$. Therefore, the mean recurrence time of the event $\mathcal{E}_1$, which is called the *return period*, is given by

$$\mu_{\mathcal{E}_1} = \frac{1}{p_{\mathcal{E}_1}} = \frac{1}{1 - F_1(x_\mathcal{E})}.$$

Such return periods can also be considered in a multivariate context. For this, let $\boldsymbol{X} := (X_1, ..., X_d)' \sim F$, where $X_1, ..., X_d$ are dependent risk quantities such as sea levels along a coastline or flows of a river measured at different stations. Salvadori et al. (2011) then propose a notion of a multivariate return period that is consistent with the univariate case. They consider events

$$\mathcal{E}_d = \{\boldsymbol{x} \in \mathbb{R}^d : F(\boldsymbol{x}) > z_\mathcal{E}\}, \tag{7.1}$$

which depend on the multivariate quantile $z_\mathcal{E} \in [0, 1]$. To see how this generalizes the univariate case, we represent the event $\mathcal{E}_1$ as $\mathcal{E}_1 = \{x_1 \in \mathbb{R} : F_1(x_1) > F_1(x_\mathcal{E})\}$, which is the special case of $\mathcal{E}_d$ for $d = 1$ and $z_\mathcal{E} = F_1(x_\mathcal{E})$.

Using the notion of the Kendall distribution function (see Definition 2.9), we then calculate $p_{\mathcal{E}_d} = P(F(\boldsymbol{X}) > z_\mathcal{E}) = 1 - K(z_\mathcal{E}; C)$, where C is the copula of $\boldsymbol{X}$. Hence, the return period of $\mathcal{E}_d$ is

$$\mu_{\mathcal{E}_d} = \frac{1}{p_{\mathcal{E}_d}} = \frac{1}{1 - K(z_\mathcal{E}; C)},$$

and the corresponding critical level set is given by

$$L(z_{\mathcal{E}}; F) = \{\boldsymbol{x} \in \mathbb{R}^d : F(\boldsymbol{x}) = z_{\mathcal{E}}\}, \tag{7.2}$$

which has a one-to-one correspondence to the copula level set $L(z_{\mathcal{E}}; C)$ (see Equation (2.10)) through the probability integral transform: If $\boldsymbol{x} \in L(z_{\mathcal{E}}; F)$ and $u_j = F_j(x_j),\ j = 1, ..., d$, then $\boldsymbol{u} \in L(z_{\mathcal{E}}; C)$. Conversely, if $\boldsymbol{u} \in L(z_{\mathcal{E}}; C)$ and $x_j = F_j^{-1}(u_j),\ j = 1, ..., d$, then $\boldsymbol{x} \in L(z_{\mathcal{E}}; F)$.

This notion of a multivariate return period is consistent in the following sense. The larger $z_{\mathcal{E}}$ is, the smaller $p_{\mathcal{E}}$ and the larger $\mu_{\mathcal{E}}$ are. This is not the case for other notions of a multivariate return period, which typically depend on threshold values for each component of $\boldsymbol{X}$. One such notion is $\widetilde{\mathcal{E}}_d = \{\boldsymbol{x} \in \mathbb{R}^d : x_j > x_{\mathcal{E},j}\ \forall j = 1, ..., d\}$, which Salvadori et al. (2011) call the 'AND' case (see also Gräler et al. (2013) for a discussion of different notions). It is the event that all components of $\boldsymbol{X}$ exceed prescribed threshold levels $\boldsymbol{x}_{\mathcal{E}} := (x_{\mathcal{E},1}, ..., x_{\mathcal{E},d})' \in \mathbb{R}^d$ at the same time. Since $\boldsymbol{x}_{\mathcal{E}}$ is an arbitrary vector, no consistent ordering as for the return period $\mu_{\mathcal{E}_d}$ is possible.

There are scenarios where multivariate return periods of different sets of risk quantities are dependent and therefore should not be considered in isolation. For example, cities at the confluence of two (or more) rivers are particularly threatened by flooding if the flows of the rivers are positively dependent. This may be the case, for instance, due to similar catchment areas or because of the snowmelt in spring. Then, the return periods of extreme flows of the rivers need to be looked at jointly. Since return periods are characterized through their critical level sets (7.2), a reasonable approach to dependence modeling of return periods is by specifying the joint distribution of the critical level sets. This joint distribution corresponds to the nesting copula of a hierarchical Kendall copula (see Definition 3.3) with the respective Kendall distribution functions as marginal distribution functions. In other words, the sets of risk quantities can be identified with the clusters of a hierarchical Kendall copula and the joint distribution of the critical level sets, which are the level sets of the cluster copulas, is given in terms of the nesting copula. Since interest is mostly in extremal quantities, extreme value copulas (see Section 2.5) will typically be needed as cluster copulas. Especially the Tawn copula (see Example 2.19) provides a quite flexible dependence model for this purpose, since it allows for asymmetry, in contrast to most other popular copulas. In an analysis of annual maximum sea levels, Tawn (1988) detect such an asymmetry between measurements at two stations in England.

Having set up an appropriate statistical model for the return periods, the practitioner seeks to identify so-called design realizations, which characterize critical events in the best way. A reasonable choice for such a design realization is the most likely realization on the critical level set $L(z_{\mathcal{E}}; F)$ (see Equation (7.2)). Although the distribution on the level set is usually not known explicitly (see the discussion in Section 4.2), sampling on the level set is feasible using the methods developed in Chapter 4. In particular, for extreme value copulas a closed-form sampling procedure is available (see Section 4.2.2). Note that, if the underlying copula is asymmetric, then the most likely copula realization does not lie on the diagonal $u\mathbf{1},\ u \in [0, 1]$, as illustrated in Figure 4.3. Furthermore, conditional scenario analysis is then also possible using the methods introduced in Section 5.2.5 for systemic risk assessment.

Depending on the context, we could, of course, also define the event $\mathcal{E}_d$ (7.1) in terms of a different aggregation function, such as the maximum or the mean (see the discussion in Section 3.2). If the aggregation function is monotone (see Definition 3.2), then the resulting notion of a multivariate return period also satisfies the above consistency property.

Structured factor modeling: The general idea of factor modeling is to describe the behavior of a multivariate random vector in terms of a set of unobserved variables, which are called *factors*. If a small number of factors is sufficient for this purpose, this yields a parsimonious model formulation. Classically, factor models are based on multivariate normality of the random vector, but Krupskii and Joe (2013a) recently formulated a factor copula model, which can account for non-Gaussian dependence (see also Nikoloulopoulos et al. (2013) who develop a factor copula model for discrete data). Let $\boldsymbol{U} := (U_1, ..., U_n)' \sim C$. In the p-factor copula model, $U_1, ..., U_n$ are then assumed to be conditionally independent given p latent variables $V_1, ..., V_p$. Without loss of generality, we can assume that $V_1, ..., V_p$ are independent and identically distributed and $V_j \sim U(0,1)$ for all $j = 1, ..., p$. Hence, we have that

$$\begin{aligned} C(\boldsymbol{u}) &= \int_{[0,1]^p} P(\boldsymbol{U} \leq \boldsymbol{u} | V_1 = v_1, ..., V_p = v_p)\, dv_1 ... dv_p \\ &= \int_{[0,1]^p} \prod_{j=1}^{n} P(U_j \leq u_j | V_1 = v_1, ..., V_p = v_p)\, dv_1 ... dv_p \\ &= \int_{[0,1]^p} \prod_{j=1}^{n} C_{j|V_1,...,V_p}(u_j | v_1, ..., v_p)\, dv_1 ... dv_p, \quad \boldsymbol{u} \in [0,1]^n. \end{aligned} \tag{7.3}$$

Krupskii and Joe (2013a) discuss the cases of $p = 1$ and $p = 2$ factors in detail and propose to decompose $C_{j|V_1,...,V_p}$ as in a C-vine PCC (see Equation (2.70)).

Their approach can be used to specify a structured factor model, which respects groupings of variables. This is similar to a hierarchical Kendall copula as defined in Definition 3.3, of which we borrow here the notation. That is, let $d_1, ..., d_{n_1}$ denote the cluster sizes with $n = \sum_{i=1}^{n_1} d_i$ and define the cumulative sum $m_i = \sum_{j=1}^{i} d_j$ for $i = 1, ..., n_1$, and $m_0 = 0$. Further, let the set of latent variables be given by $V_0, V_1, ..., V_{n_1}$. Similar to the assumption $\mathcal{A}_2$ in Definition 3.3, we assume that

$\widetilde{\mathcal{A}}$: the conditional distribution of $\boldsymbol{U}_i | (V_0, V_1, ..., V_{n_1})'$ is the same as the conditional distribution of $\boldsymbol{U}_i | (V_0, V_i)'$ for all $i = 1, ..., n_1$.

This means that we assume that the dependence within cluster i is explained solely in terms of the cluster-specific latent variable V_i and in terms of V_0. The latent variable V_0 is an overall factor, which influences all variables in all clusters (see also the Gaussian factor copula proposed by Gregory and Laurent (2004)). Between- and within-cluster dependence is then implicitly given in terms of the dependence of the variables on the factors.

Under the assumption $\widetilde{\mathcal{A}}$ the $(n_1 + 1)$-factor copula of $\boldsymbol{U}$ (see Equation (7.3)) is given

by

$$\begin{aligned} C(\boldsymbol{u}) &= \int_{[0,1]^{n_1+1}} \prod_{j=1}^{n} C_{j|V_0,V_1,...,V_{n_1}}(u_j|v_0,v_1,...,v_{n_1})\, dv_0\, dv_1 ... dv_{n_1} \\ &= \int_{[0,1]^{n_1+1}} \prod_{i=1}^{n_1} \left(\prod_{j=m_{i-1}+1}^{m_i} C_{j|V_0,V_i}(u_j|v_0,v_i) \right) dv_0\, dv_1 ... dv_{n_1} \\ &= \int_0^1 \prod_{i=1}^{n_1} \left(\int_0^1 \prod_{j=m_{i-1}+1}^{m_i} C_{j|V_0,V_i}(u_j|v_0,v_i)\, dv_i \right) dv_0, \end{aligned}$$

where we can express $C_{j|V_0,V_i}$ as in a C-vine PCC (under the simplifying assumption (2.59)) as

$$C_{j|V_0,V_i}(u_j|v_0,v_i) = C_{j|V_i;V_0}(C_{j|V_0}(u_j|v_0)|C_{V_i|V_0}(v_i|v_0)) = C_{j|V_i;V_0}(C_{j|V_0}(u_j|v_0)|v_i),$$

since V_0 and V_i are assumed to be independent. The corresponding density of the (n_1+1)-factor copula can then be derived as

$$c(\boldsymbol{u}) = \int_0^1 \prod_{i=1}^{n_1} \left(\int_0^1 \prod_{j=m_{i-1}+1}^{m_i} c_{j,V_i;V_0}(C_{j|V_0}(u_j|v_0)|v_i)\, c_{j,V_0}(u_j,v_0)\, dv_i \right) dv_0. \tag{7.4}$$

Compared to the general (n_1+1)-factor copula, which requires (n_1+1)-dimensional integration to compute the density, this expression constitutes a significant simplification, since only one-dimensional integration nested within another one-dimensional integration is needed for density evaluations. This structured factor copula model is currently investigated in detail by Krupskii and Joe (2013b).

In contrast to hierarchical Kendall copulas, it is straightforward to show that the Gaussian copula is a special case of this factor copula model (see Krupskii and Joe (2013a)). Flexibility is however gained through different choices of the bivariate copulas as in a vine copula. This extends the classical bi-factor model by Holzinger and Swineford (1937), which is an important model in psychometrics and a special case of a more general model class called structural equation models (see, e.g., Bollen (1989)).

The presented bi-factor copula model can be extended, for example, by assuming dependence of the latent variables of the clusters. But then the density of the factor copula model no longer has the attractive form given in Equation (7.4) (see the proof of Theorem 3.8). An appealing density expression in terms of nested one-dimensional integrals can be kept if the cluster-specific latent variables $V_1, ..., V_{n_1}$ are assumed to be independent given the overall factor V_0. A density expression similar to Equation (7.4) can also be obtained when residual dependence within the clusters is assumed, that is, when the assumption that $U_{m_{i-1}+1}, ..., U_{m_i}$ are independent given V_0 and V_i is dropped for $i = 1, ..., n_1$.

A Bivariate copulas

For each bivariate copula discussed in Chapter 2 we show, from left to right, scatter plots of a sample, contour lines of the copula, and contour lines of the copula density combined with standard normal margins (expect for the countermonotonicity and the comonotonicity copula, which do not possess a density). The contour lines of the copulas are the level sets $L(z; C)$ as defined in Equation (2.10) (here: $z \in \{0.1, 0.2, ..., 0.9\}$). The level set of a copula density combined with standard normal margins is given by

$$L_\Phi(z; c) = \{(x_1, x_2)' \in \mathbb{R}^2 : c(\Phi(x_1), \Phi(x_2))\, \phi(x_1)\, \phi(x_2) = z\},$$

where we choose $z \in \{0.005, 0.01, 0.02, 0.05, 0.1, 0.15, 0.2\}$. The parameters of the copulas are chosen according to a Kendall's τ of 0.5.

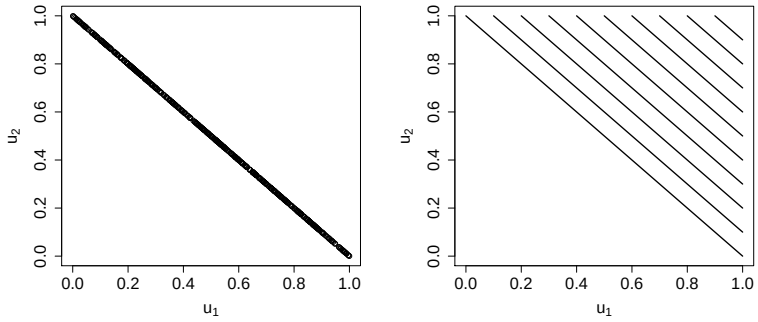


Figure A.1: Countermonotonicity copula.

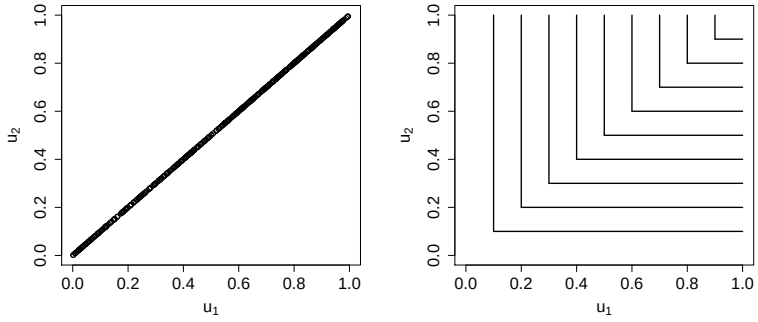


Figure A.2: Comonotonicity copula.

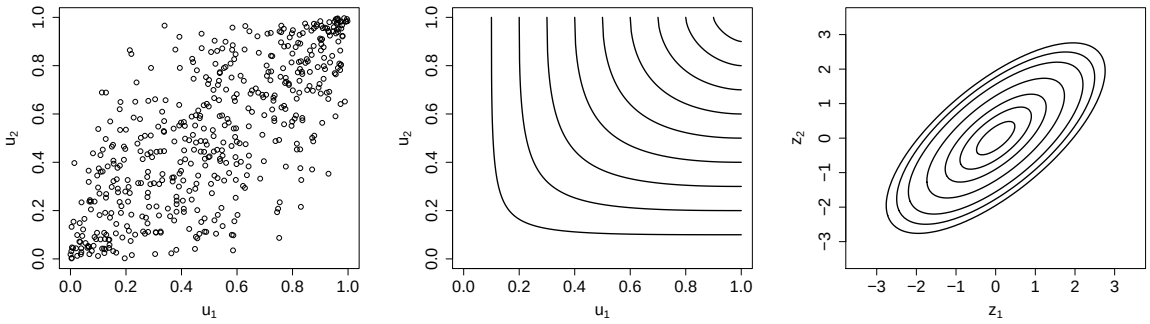


Figure A.3: Gaussian copula.

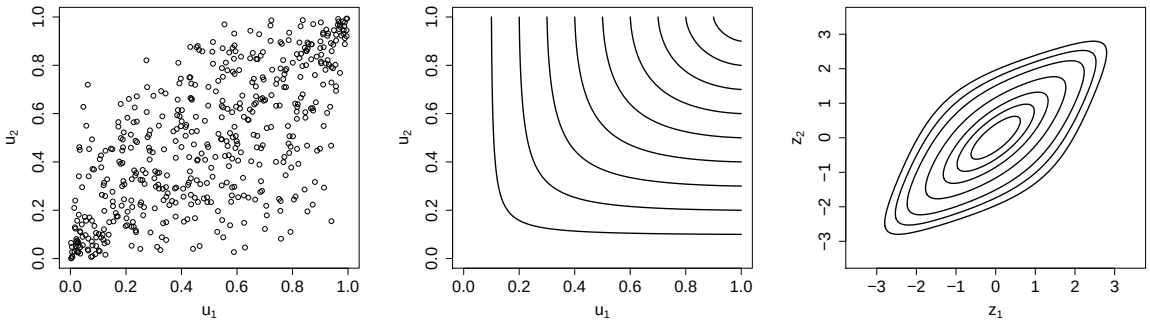


Figure A.4: Student's t copula with $\nu = 6$ degrees of freedom.

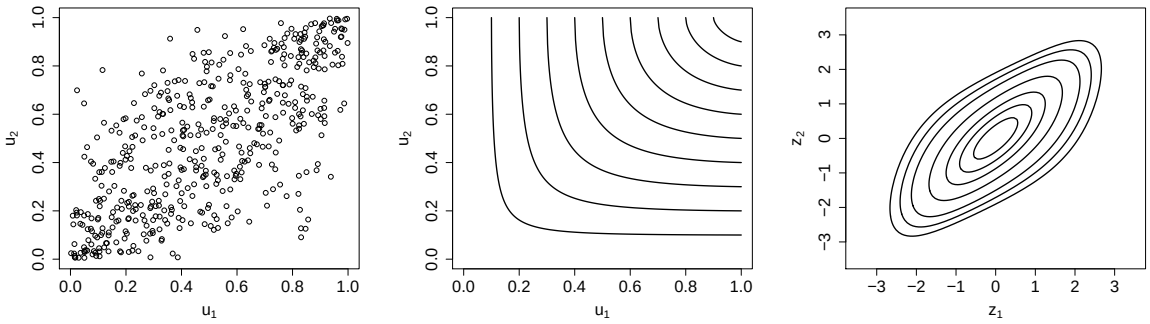


Figure A.5: Individual Student's t copula with $\nu_1 = 4$ and $\nu_2 = 20$ degrees of freedom.

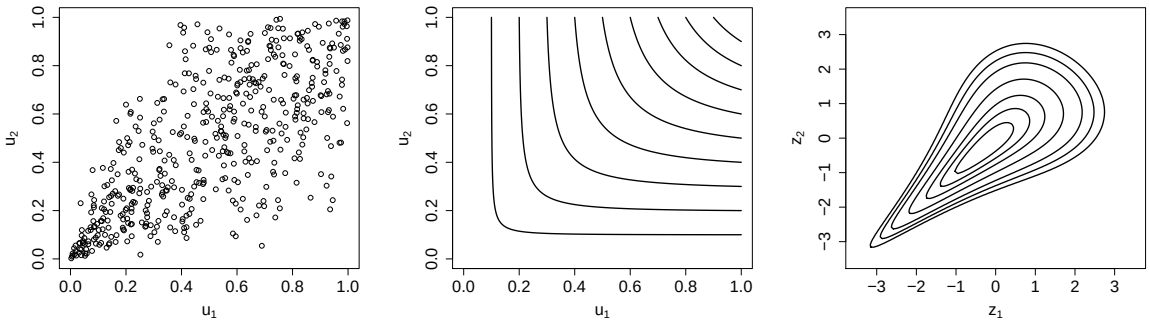


Figure A.6: Clayton copula.

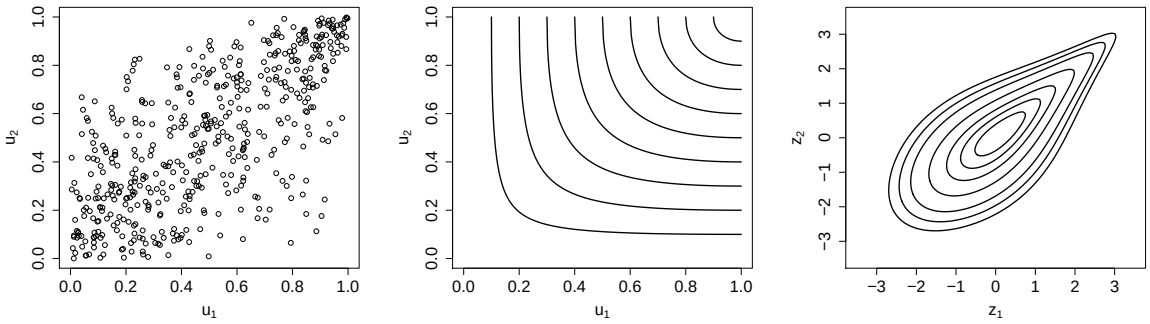


Figure A.7: Gumbel copula.

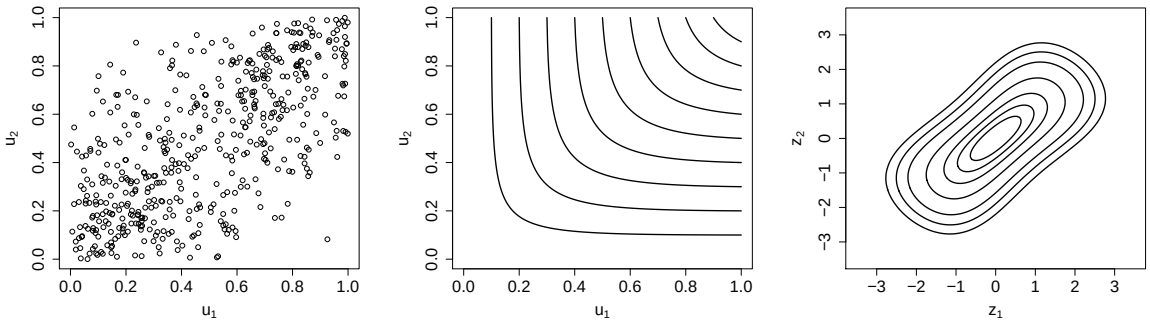


Figure A.8: Frank copula.

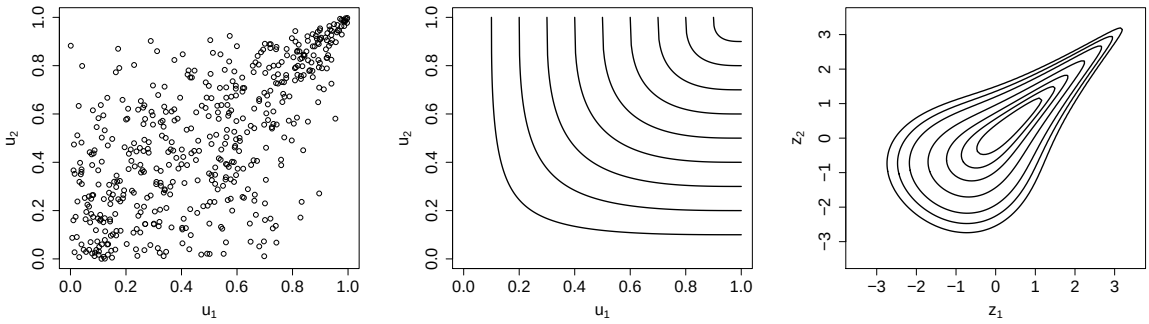


Figure A.9: Joe copula.

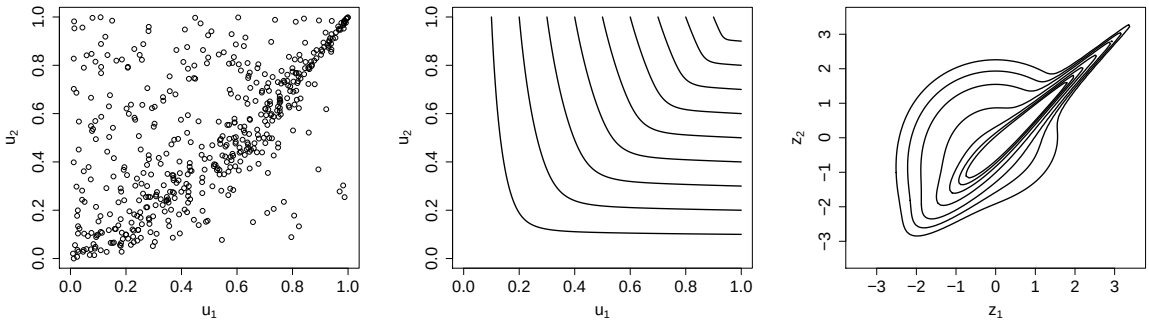


Figure A.10: Tawn copula with $\psi_1 = 0.6$ and $\psi_2 = 0.9$.

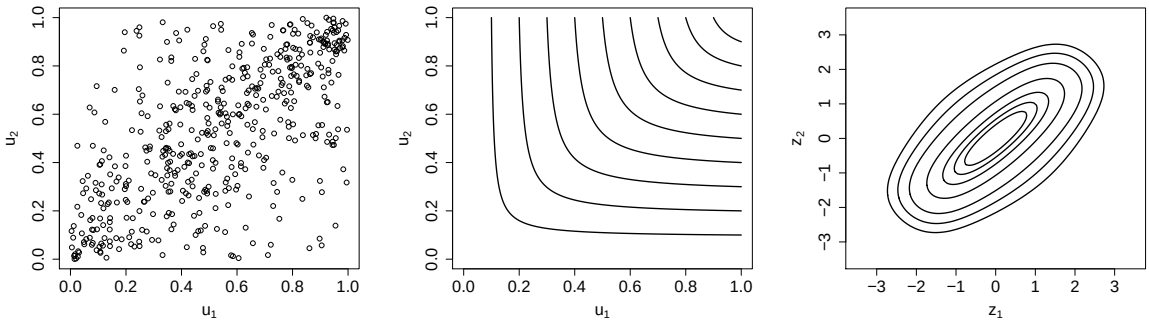


Figure A.11: Plackett copula.

B Technical derivations

We derive conditional distribution functions of Plackett and Archimedean copulas for a specific level set as well as of the level sets of Archimedean copulas and of the Student's t mixing variable. In addition, it is shown how to calculate the Kendall distribution function of the Plackett copula.

B.1 Conditional distribution function of Archimedean copulas

We prove Lemma 4.4 by showing for a d-dimensional Archimedean copula with generator φ that it holds for all $j = 1, ..., d-1$

$$F_{U_j|U_1,...,U_{j-1},C(\boldsymbol{U};\varphi)}(u|u_1, ..., u_{j-1}, z; \varphi) = \left(1 - \frac{\varphi(u)}{\varphi(z) - \sum_{1\leq i<j} \varphi(u_i)}\right)^{d-j},$$

where $u \in (C^{-1}(z|u_1, ..., u_{j-1}; \varphi), 1)$. For ease of notation, dependence of expressions such as the copula and its quantile function on the generator φ is dropped in the following.

We observe that the density (2.34) of an Archimedean copula C only depends on $u_1, ..., u_d$ through the first derivatives of φ and through $C(\boldsymbol{u})$:

$$c(\boldsymbol{u}) = (\varphi^{-1})^{(d)}(\varphi(u_1) + ... + \varphi(u_d)) \prod_{i=1}^{d} \varphi'(u_i) =: h(C(\boldsymbol{u})) \prod_{i=1}^{d} \varphi'(u_i). \tag{B.1}$$

To see this, note that any derivative of an inverse is a function of derivatives of the original function applied to the inverse, that is, $(f^{-1})^{(d)}(x) = \tilde{f}(f^{-1}(x))$, $d \in \mathbb{N}$, for an appropriately chosen $\tilde{f}$. The rest follows from the definition of Archimedean copulas (2.31).

Further, for Archimedean copulas the copula quantile function is given in closed form (see Equation (2.35)):

$$C^{-1}(z|u_1, ..., u_{d-1}) = \varphi^{-1}\left(\varphi(z) - \sum\nolimits_{1\leq i<d} \varphi(u_i)\right), \tag{B.2}$$

Its derivative with respect to z is

$$\frac{\partial}{\partial z} C^{-1}(z|u_1, ..., u_{d-1}) = \frac{\varphi'(z)}{\varphi'(\varphi^{-1}(\varphi(z) - \sum_{1\leq i<d} \varphi(u_i)))}. \tag{B.3}$$

Following Theorem 4.2, we then determine the function g_j given in Equation (4.2) by plugging in Equations (B.1)–(B.3):

$$\begin{aligned} g_j(u_1,...,u_j,z) &= \int_{C^{-1}(z|u_1,...,u_j)}^1 ... \int_{C^{-1}(z|u_1,...,u_{d-2})}^1 c(u_1,...,u_{d-1}, C^{-1}(z|u_1,...,u_{d-1})) \\ &\qquad \times \frac{\partial}{\partial z} C^{-1}(z|u_1,...,u_{d-1})\, du_{d-1}...du_{j+1} \\ &= \int_{C^{-1}(z|u_1,...,u_j)}^1 ... \int_{C^{-1}(z|u_1,...,u_{d-2})}^1 h(z)\, \varphi'(z) \prod_{i=1}^{d-1} \varphi'(u_i)\, du_{d-1}...du_{j+1} \\ &= h(z)\, \varphi'(z) \prod_{i=1}^{j} \varphi'(u_i) \\ &\quad \times \int_{C^{-1}(z|u_1,...,u_j)}^1 \varphi'(u_{j+1}) ... \int_{C^{-1}(z|u_1,...,u_{d-2})}^1 \varphi'(u_{d-1})\, du_{d-1}...du_{j+1}, \quad \text{(B.4)} \end{aligned}$$

since

$$\varphi'(C^{-1}(z|u_1,...,u_{d-1}))\, \frac{\partial}{\partial z} C^{-1}(z|u_1,...,u_{d-1}) = \varphi'(z),$$

and

$$h(C(u_1,...,u_{d-1}, C^{-1}(z|u_1,...,u_{d-1}))) = h(z).$$

Next, we iteratively solve the nested integrals in (B.4). First,

$$\begin{aligned} \int_{C^{-1}(z|u_1,...,u_{d-2})}^1 \varphi'(u_{d-1})\, du_{d-1} &= \varphi(1) - \varphi(C^{-1}(z|u_1,...,u_{d-2})) \\ &\overset{\varphi(1)=0}{\underset{(2.35)}{=}} - \left(\varphi(z) - \textstyle\sum_{1\le i<d-1} \varphi(u_i)\right) \\ &= \varphi(u_{d-2}) + \left(\textstyle\sum_{1\le i<d-2} \varphi(u_i) - \varphi(z)\right). \end{aligned}$$

The second integral is then given by

$$\begin{aligned} &\int_{C^{-1}(z|u_1,...,u_{d-3})}^1 \varphi'(u_{d-2}) \left(\varphi(u_{d-2}) + \left(\textstyle\sum_{1\le i<d-2} \varphi(u_i) - \varphi(z)\right)\right) du_{d-2} \\ &\qquad = \frac{1}{2} \left(\varphi(u_{d-2}) + \left(\textstyle\sum_{1\le i<d-2} \varphi(u_i) - \varphi(z)\right)\right)^2 \Big|_{u_{d-2}=C^{-1}(z|u_1,...,u_{d-3})}^{u_{d-2}=1} \\ &\qquad = \frac{1}{2} \left(\varphi(u_{d-3}) + \left(\textstyle\sum_{1\le i<d-3} \varphi(u_i) - \varphi(z)\right)\right)^2. \end{aligned}$$

Similarly, the third integral computes to

$$\begin{aligned} &\int_{C^{-1}(z|u_1,...,u_{d-4})}^1 \varphi'(u_{d-3}) \frac{1}{2} \left(\varphi(u_{d-3}) + \left(\textstyle\sum_{1\le i<d-3} \varphi(u_i) - \varphi(z)\right)\right)^2 du_{d-3} \\ &\qquad = ... = \frac{1}{2}\frac{1}{3} \left(\varphi(u_{d-4}) + \left(\textstyle\sum_{1\le i<d-4} \varphi(u_i) - \varphi(z)\right)\right)^3 \end{aligned}$$

By continuing iteratively, we finally arrive at

$$g_j(u_1,...,u_j,z) = h(z)\,\varphi'(z)\prod_{i=1}^{j}\varphi'(u_i)\,\frac{1}{(d-j-1)!}\left(\sum\nolimits_{1\le i<j+1}\varphi(u_i)-\varphi(z)\right)^{d-j-1},$$

and thus

$$\begin{aligned}
&\int_{C^{-1}(z|u_1,...,u_{j-1})}^{u} g_j(u_1,...,u_j,z)\,du_j \\
&\qquad = \frac{h(z)\,\varphi'(z)}{(d-j-1)!}\prod_{i=1}^{j-1}\varphi'(u_i) \\
&\qquad\quad \times \int_{C^{-1}_{u_1,...,u_{j-1}}(z)}^{u}\varphi'(u_j)\left(\varphi(u_j)+\left(\sum\nolimits_{1\le i<j}\varphi(u_i)-\varphi(z)\right)\right)^{d-j-1}du_j \\
&\qquad = \frac{h(z)\,\varphi'(z)}{(d-j-1)!}\prod_{i=1}^{j-1}\varphi'(u_i)\,\frac{1}{d-j}\left(\varphi(u)+\sum\nolimits_{1\le i<j}\varphi(u_i)-\varphi(z)\right)^{d-j} \\
&\qquad = \frac{h(z)\,\varphi'(z)}{(d-j)!}\prod_{i=1}^{j-1}\varphi'(u_i)\left(\varphi(u)+\sum\nolimits_{1\le i<j}\varphi(u_i)-\varphi(z)\right)^{d-j}. \qquad \text{(B.5)}
\end{aligned}$$

By plugging $u = 1$ into (B.5), we further obtain

$$\begin{aligned}
&\int_{C^{-1}(z|u_1,...,u_{j-1})}^{1} g_j(u_1,...,u_j,z)\,du_j \\
&\qquad = \frac{1}{(d-j)!}\,h(z)\,\varphi'(z)\prod_{i=1}^{j-1}\varphi'(u_i)\left(\sum\nolimits_{1\le i<j}\varphi(u_i)-\varphi(z)\right)^{d-j}. \qquad \text{(B.6)}
\end{aligned}$$

Combining Equations (B.5) and (B.6) as in Equation (4.1) then gives

$$\begin{aligned}
F_{U_j|U_1,...,U_{j-1},C(\boldsymbol{U})}(u|u_1,...,u_{j-1},z) &= \frac{\int_{C^{-1}(z|u_1,...,u_{j-1})}^{u} g_j(u_1,...,u_j,z)du_j}{\int_{C^{-1}(z|u_1,...,u_{j-1})}^{1} g_j(u_1,...,u_j,z)du_j} \\
&= \frac{\left(\varphi(u)+\sum_{1\le i<j}\varphi(u_i)-\varphi(z)\right)^{d-j}}{\left(\sum_{1\le i<j}\varphi(u_i)-\varphi(z)\right)^{d-j}} \\
&= \left(1-\frac{\varphi(u)}{\varphi(z)-\sum_{1\le i<j}\varphi(u_i)}\right)^{d-j},
\end{aligned}$$

which is the desired result.

As noted by an anonymous referee, this result can also be derived in a very elegant way by exploiting the representation (2.31) of Archimedean copulas, properties of the Dirichlet distribution (see Remark 4.3) and Proposition 4.6. Our proof of Lemma 4.4 however exploits the general formula provided in Theorem 4.2 and illustrates how it can be applied, as it is also the case for the Plackett copula (see Appendix B.4 below).

Remark B.1 (Alternative proof of Lemma 4.4)**.** According to Proposition 4.6, it holds that

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

where $\boldsymbol{S} = (S_1, ..., S_d)' \sim D(1, ..., 1)$. For the Dirichlet distribution, it holds that (see Fang et al. (1990, Theorem 1.6))

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

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

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

as claimed in Lemma 4.4. □

B.2 Conditional distribution function of the level sets of Archimedean copulas

Let $\boldsymbol{U} \sim C(\cdot; \varphi)$, where $C(\cdot; \varphi)$ is a d-dimensional Archimedean copula with generator φ, and define the copula level set variable $Z := C(\boldsymbol{U}; \varphi)$. We show that

$$F_{Z|U_1}(z|u_1; \varphi) = \varphi'(u_1) \sum_{k=1}^{d-1} \frac{(\varphi(u_1) - \varphi(z))^{k-1}}{(k-1)!} (\varphi^{-1})^{(k)}(\varphi(z)), \quad z \in (0,1). \tag{B.7}$$

According to Equation (5.9), it holds that

$$F_{Z|U_1}(z|u_1; \varphi) = \varphi'(u_1) \int_{\varphi(0)}^{\varphi(z)} \frac{(\varphi(u_1) - x)^{d-2}}{(d-2)!} (\varphi^{-1})^{(d)}(x)\, dx.$$

To derive Equation (B.7), we hence proof by induction that

$$\int_{\varphi(0)}^{\varphi(z)} \frac{(\varphi(u_1) - x)^{d-2}}{(d-2)!} (\varphi^{-1})^{(d)}(x)\, dx = \sum_{k=1}^{d-1} \frac{(\varphi(u_1) - \varphi(z))^{k-1}}{(k-1)!} (\varphi^{-1})^{(k)}(\varphi(z)), \tag{B.8}$$

for $d \geq 2$ and $z \in (0,1)$. In the bivariate case ($d = 2$), this is straightforward, since

$$\int_{\varphi(0)}^{\varphi(z)} (\varphi^{-1})^{(2)}(x)\, dx = (\varphi^{-1})^{(1)}(\varphi(z)).$$

For $d > 2$, we use integration by parts to obtain

$$\begin{aligned}\int_{\varphi(0)}^{\varphi(z)} &\frac{(\varphi(u_1) - x)^{d-2}}{(d-2)!} (\varphi^{-1})^{(d)}(x)\, dx \\ &= \frac{(\varphi(u_1) - \varphi(z))^{d-2}}{(d-2)!}(\varphi^{-1})^{(d-1)}(\varphi(z)) + \int_{\varphi(0)}^{\varphi(z)} \frac{(\varphi(u_1) - x)^{d-3}}{(d-3)!} (\varphi^{-1})^{(d-1)}(x)\, dx.\end{aligned}$$

Applying the induction hypothesis (B.8) for $d-1$ to the last term then proves the statement and hence Equation (B.7).

B.3 Kendall distribution function of the Plackett copula

According to Equation (2.15), it holds for the bivariate Plackett copula $C(\cdot, \cdot; \alpha)$ with parameter $\alpha \in (-1, \infty) \setminus \{0\}$ (see Equation (2.50)) that

$$K(z; \alpha) = z + \int_z^1 C_{2|1}(C^{-1}(z|u_1; \alpha)|u_1; \alpha)\, du_1, \quad z \in [0,1]. \tag{B.9}$$

The copula quantile function of the Plackett copula is given in closed form in Equation (2.52) and the first derivative of the Plackett copula with respect to the second argument is

$$C_{2|1}(u_2|u_1; \alpha) = \frac{1}{2}\left(1 - \frac{1 + \alpha(u_1 + u_2) - 2(\alpha+1)u_2}{\sqrt{(1 + \alpha(u_1+u_2))^2 - 4\alpha(\alpha+1)u_1u_2}}\right).$$

This yields for the integrand of Equation (B.9) (see also Genest and Rivest (2001)):

$$C_{2|1}(C^{-1}(z|u_1; \alpha)|u_1; \alpha) = \frac{(\alpha+1)z - \alpha z^2}{\alpha(\alpha+1)(u_1 - z)^2 + (\alpha+1)u_1 - \alpha z^2}.$$

Hence,

$$K(z; \alpha) = z + \frac{(\alpha+1)z - \alpha z^2}{\alpha(\alpha+1)} \int_z^1 \frac{1}{(u_1 + \frac{1}{2\alpha} - z)^2 + \mathcal{D}(z, \alpha)}\, du_1, \tag{B.10}$$

where

$$\mathcal{D}(z, \alpha) = \frac{4\alpha z(\alpha(1-z)+1) - \alpha - 1}{4\alpha^2(\alpha+1)}. \tag{B.11}$$

After substitution with $v = u_1 + 1/(2\alpha) - z$, the integral in Equation (B.10) is then given by

$$\int_z^1 \frac{1}{(u_1 + \frac{1}{2\alpha} - z)^2 + \mathcal{D}(z, \alpha)}\, du_1 = \int_{1/(2\alpha)}^{1+1/(2\alpha)-z} \frac{1}{v^2 + \mathcal{D}(z, \alpha)}\, dv =: \mathcal{I} \tag{B.12}$$

The solution of the integral $\mathcal{I}$ depends on whether $\mathcal{D}(z, \alpha) > 0$, $\mathcal{D}(z, \alpha) = 0$ or $\mathcal{D}(z, \alpha) < 0$. These three different cases are illustrated in Figure B.1.

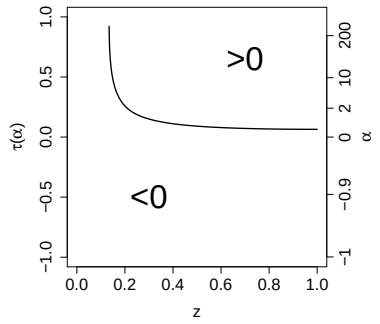


Figure B.1: Sets of parameters with $\mathcal{D}(z,\alpha) > 0$, $\mathcal{D}(z,\alpha) = 0$ (solid line) and $\mathcal{D}(z,\alpha) < 0$. The parameter α (right axis) has been transformed to the corresponding Kendall's τ value $\tau(\alpha)$ (left axis).

(i) If $\mathcal{D}(z,\alpha) > 0$, then we obtain

$$\mathcal{I} = \frac{1}{\sqrt{\mathcal{D}(z,\alpha)}} \left(\arctan\left(\frac{1 + \frac{1}{2\alpha} - z}{\sqrt{\mathcal{D}(z,\alpha)}} \right) - \arctan\left(\frac{1}{2\alpha\sqrt{\mathcal{D}(z,\alpha)}} \right) \right).$$

(ii) If $\mathcal{D}(z,\alpha) = 0$, then the integral is straightforward to solve as

$$\mathcal{I} = -\left(\frac{1}{1 + \frac{1}{2\alpha} - z} - 2\alpha \right).$$

(iii) If $\mathcal{D}(z,\alpha) < 0$, we have to distinguish two cases: $\alpha > 0$ and $\alpha < 0$. If $\alpha > 0$, we have

$$\mathcal{I} = -\frac{1}{\sqrt{-\mathcal{D}(z,\alpha)}} \left(\operatorname{arcoth}\left(\frac{1 + \frac{1}{2\alpha} - z}{\sqrt{-\mathcal{D}(z,\alpha)}} \right) - \operatorname{arcoth}\left(\frac{1}{2\alpha\sqrt{-\mathcal{D}(z,\alpha)}} \right) \right).$$

Conversely, if $\alpha < 0$, we need to use the inverse hyperbolic tangent instead of the cotangent:

$$\mathcal{I} = -\frac{1}{\sqrt{-\mathcal{D}(z,\alpha)}} \left(\operatorname{artanh}\left(\frac{1 + \frac{1}{2\alpha} - z}{\sqrt{-\mathcal{D}(z,\alpha)}} \right) - \operatorname{artanh}\left(\frac{1}{2\alpha\sqrt{-\mathcal{D}(z,\alpha)}} \right) \right).$$

These expressions can be plugged into Equations (B.10) and (B.12) to obtain the Kendall distribution function of the Plackett copula at an arbitrary point $z \in [0,1]$.

B.4 Conditional distribution function of the Plackett copula

Let $C(\cdot,\cdot;\alpha)$ be a Plackett copula with parameter $\alpha \in (-1,\infty) \setminus \{0\}$ (see Equation (2.50)) and let $z \in (0,1)$. As shown in Theorem 4.2, the conditional distribution function

$F_{U_1|C(U_1,U_2;\alpha)}(\cdot|z;\alpha)$ of $U_1|C(U_1,U_2;\alpha)=z$ can then be determined for $u\in(z,1)$ as

$$F_{U_1|C(U_1,U_2;\alpha)}(u|z;\alpha)=\frac{\int_z^u g_1(u_1,z;\alpha)\,du_1}{\int_z^1 g_1(u_1,z;\alpha)\,du_1}, \tag{B.13}$$

where

$$g_1(u_1,z;\alpha)=c\left(u_1,C^{-1}(z|u_1;\alpha);\alpha\right)\frac{\partial}{\partial z}C^{-1}(z|u_1;\alpha).$$

The density of the Plackett copula is given in Equation (2.51), the copula quantile function in Equation (2.52). We can therefore calculate g_1 and obtain after some simplifications

$$g_1(u_1,z;\alpha)=\frac{1}{\alpha^2(\alpha+1)}\frac{(\alpha+1)(1+\alpha u_1)u_1-2(1+\alpha u_1)\alpha u_1 z+(2u_1-1)\alpha^2z^2}{\left((u_1+\frac{1}{2\alpha}-z)^2+\mathcal{D}(z,\alpha)\right)^2},$$

where $\mathcal{D}$ is defined in Equation (B.11). To compute $F_{U_1|C(U_1,U_2;\alpha)}(\cdot|z;\alpha)$ (see Equation (B.13)), we hence have to solve the integral

$$\int_{1/(2\alpha)}^{u+1/(2\alpha)-z}\frac{\widetilde{a}v^2+\widetilde{b}v+\widetilde{c}}{\left(v^2+\mathcal{D}(z,\alpha)\right)^2}\,dv,$$

where $\widetilde{a}$, $\widetilde{b}$ and $\widetilde{c}$ are constants independent of v, which has been substituted for $u_1+1/(2\alpha)-z$ as in Equation (B.12). As for the Kendall distribution function, the solution depends on the value of $\mathcal{D}(z,\alpha)$ (see Figure B.1).

(i) If $\mathcal{D}(z,\alpha)>0$, we obtain, up to a multiplicative constant independent of u, that

$$\begin{aligned}\int_z^u & g_1(u_1,z;\alpha)\,du_1\\ &=\frac{\alpha^2z(1+\alpha)(u-z)\left((1+\alpha)(2\alpha(uz+z-u)-u-1)-2\alpha^2z^2\right)}{u(1+\alpha)(1+\alpha(u-2z))+\alpha^2z^2}\\ &\quad+h_1(u;z,\alpha,\sigma,\beta),\end{aligned}$$

where

$$h_1(w;z,\alpha,\sigma,\beta):=\sigma\beta(\arctan(\sigma)+\arctan(\sigma(2\alpha(z-w)-1)))$$

with

$$\begin{aligned}\sigma=\sigma(z,\alpha)&:=\sqrt{\frac{1+\alpha}{4\alpha z(1+\alpha(1-z))-\alpha-1}}=\sqrt{\frac{1}{4\alpha^2\mathcal{D}(z,\alpha)}}\\ \beta=\beta(z,\alpha)&:=-\left(1+\alpha(1-2z)+2\alpha^2z(z-1)\right)(1+\alpha(1-2z)).\end{aligned}$$

Setting $u=1$ further gives, up to the same multiplicative constant as above, that

$$\int_z^1 g_1(u_1,z;\alpha)\,du_1=2\alpha^2z(1+\alpha)(z-1)+h_1(1;z,\alpha,\sigma,\beta),$$

so that the conditional distribution function $F_{U_1|C(U_1,U_2;\alpha)}(\cdot|z;\alpha)$ can be obtained by plugging the above into Equation (B.13).

(ii) If $\mathcal{D}(z,\alpha) = 0$, the following closed-form expression can be derived:

$$\int_z^u g_1(u_1, z; \alpha)\, du_1 = \frac{4\alpha^2}{3}\left(1 - 3\alpha^2(z-1)z + \alpha(1+z) + h_2(u; z, \alpha)\right),$$

where

$$\begin{aligned} &h_2(w; z, \alpha) \\ &\quad = \frac{1}{(1+2\alpha(w-z))^3}(\alpha((5+3\alpha(1+2w(3+\alpha(1+2w))))z \\ &\qquad\qquad - 9\alpha(1+2\alpha w)z^2 + 6\alpha^2 z^3 - 1 - 6(1+\alpha)(1+\alpha w)w) - 1). \end{aligned}$$

The normalizing constant in Equation (B.13) is obtained by plugging in $u = 1$.

(iii) If $\mathcal{D}(z,\alpha) < 0$, we have to distinguish the cases $\alpha > 0$ and $\alpha < 0$, as in the derivation of the Kendall distribution function. If $\alpha > 0$, we have, similar to above and up to a multiplicative constant independent of u, that

$$\begin{aligned} &\int_z^u g_1(u_1, z; \alpha)\, du_1 \\ &\qquad = \frac{\alpha^2 z(1+\alpha)(u-z)\left((1+\alpha)(2\alpha(uz+z-u)-u-1) - 2\alpha^2 z^2\right)}{u(1+\alpha)(1+\alpha(u-2z)) + \alpha^2 z^2} \\ &\qquad\quad + h_3(u; z, \alpha, \sigma, \beta), \end{aligned}$$

where

$$h_3(w; z, \alpha, \sigma, \beta) := \sigma\beta(\operatorname{arcoth}(\sigma) + \operatorname{arcoth}(\sigma(2\alpha(z-w)-1))).$$

For $\alpha < 0$, we simply need to replace h_3 by $\widetilde{h}_3$, which is given by

$$\widetilde{h}_3(w; z, \alpha, \sigma, \beta) := \sigma\beta(\operatorname{artanh}(\sigma) + \operatorname{artanh}(\sigma(2\alpha(z-w)-1))).$$

In both cases, the conditional distribution function $F_{U_1|C(U_1,U_2;\alpha)}(\cdot|z;\alpha)$ is again obtained by plugging the respective expressions into Equation (B.13).

B.5 Conditional distribution function of the Student's t mixing variable

If we define $Y := WZ$, where $Z \sim \mathcal{N}_1(0,1)$ and $\nu/W^2 \sim \chi^2_\nu$, then $Y \sim \mathcal{T}_1(0,1,\nu)$ (see Section 2.3). We prove that the conditional distribution of the mixing variable W given $Y = y$ is

$$F_{W|Y}(w|y;\nu) = 1 - F_{\chi^2}\left(\frac{\nu + y^2}{w^2}; \nu + 1\right), \quad w > 0. \tag{B.14}$$

As shown in Section 5.2.2, it holds that

$$F_{W|Y}(w|y;\nu) = \frac{1}{t_\nu(y)} \int_0^w f_W(x;\nu)\, \phi\left(\frac{y}{x}\right) \frac{1}{x}\, dx, \tag{B.15}$$

where

$$t_\nu(y) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})}\left(1+\frac{y^2}{\nu}\right)^{-(\nu+1)/2},$$
$$f_W(x;\nu) = f_{\chi^2}\left(\frac{\nu}{x^2};\nu\right)\frac{2\nu}{x^3} = \frac{2^{-\nu/2}}{\Gamma(\frac{\nu}{2})}\left(\frac{\nu}{x^2}\right)^{\nu/2-1}\exp\left(-\frac{1}{2}\frac{\nu}{x^2}\right)\frac{2\nu}{x^3},$$
$$\phi\left(\frac{y}{x}\right) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\frac{y^2}{x^2}\right).$$

Plugging everything into Equation (B.15) yields

$$F_{W|Y}(w|y;\nu) = \frac{2^{-(\nu-1)/2}}{\Gamma(\frac{\nu+1}{2})}(\nu+y^2)^{(\nu+1)/2}\int_0^w x^{-\nu-2}\exp\left(-\frac{1}{2}\frac{\nu+y^2}{x^2}\right)dx.$$

After substitution with $t = (\nu+y^2)/x^2$ this reduces to

$$F_{W|Y}(w|y;\nu) = \frac{2^{-(\nu+1)/2}}{\Gamma(\frac{\nu+1}{2})}\int_{\frac{\nu+y^2}{w^2}}^\infty t^{(\nu+1)/2-1}\,e^{-t/2}\,dt,$$

which is the survival function of the χ^2 distribution with $\nu+1$ degrees of freedom evaluated at $(\nu+y^2)/w^2$, as claimed in Equation (B.14).

C Simulation results

For the simulation study of estimation methods of hierarchical Kendall copulas (see Section 3.5), the results of the remaining five cases are shown. The following five figures show the mean squared errors (MSEs) of the estimated nesting copula parameter θ_0 (transformed to Kendall's τ) for the three estimation procedures. The notation for the x-axis is (τ_0,τ_1,τ_2), where $L := 0.4$ and $H := 0.7$. The range of the y-axes is chosen such that the MSEs are comparable.

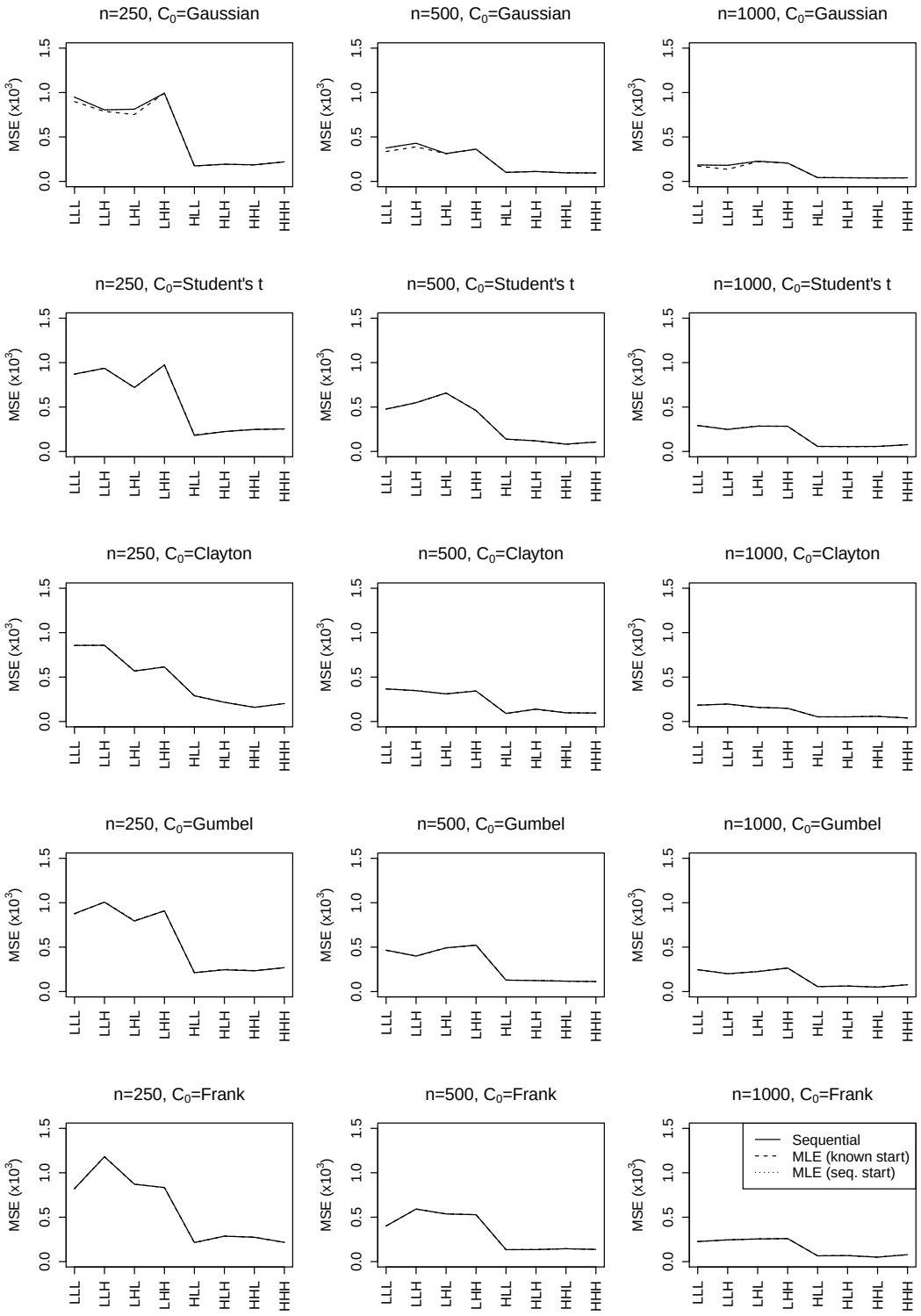


Figure C.1: Cluster copula 1: Clayton. Cluster copula 2: Clayton.

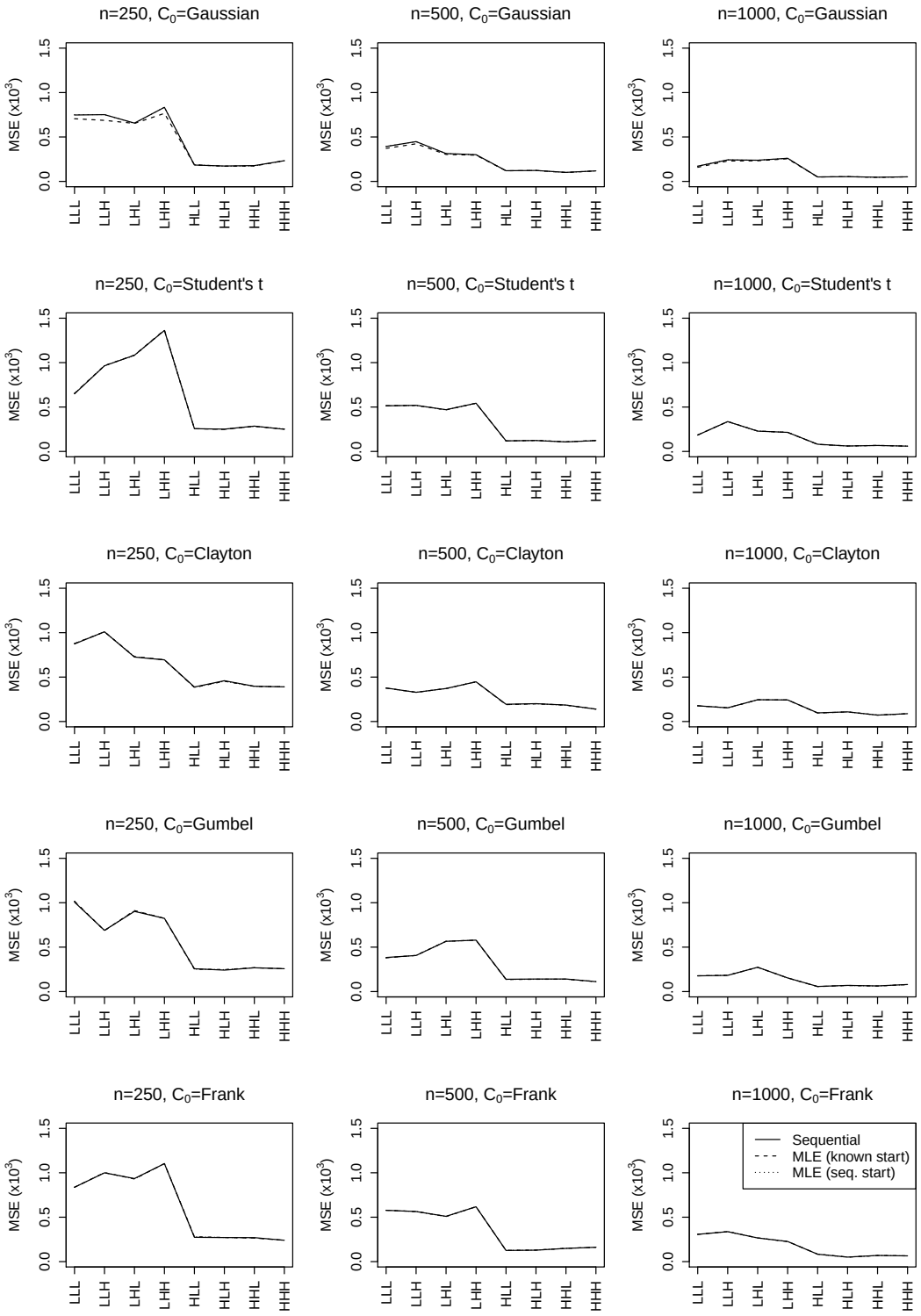


Figure C.2: Cluster copula 1: Clayton. Cluster copula 2: Frank.

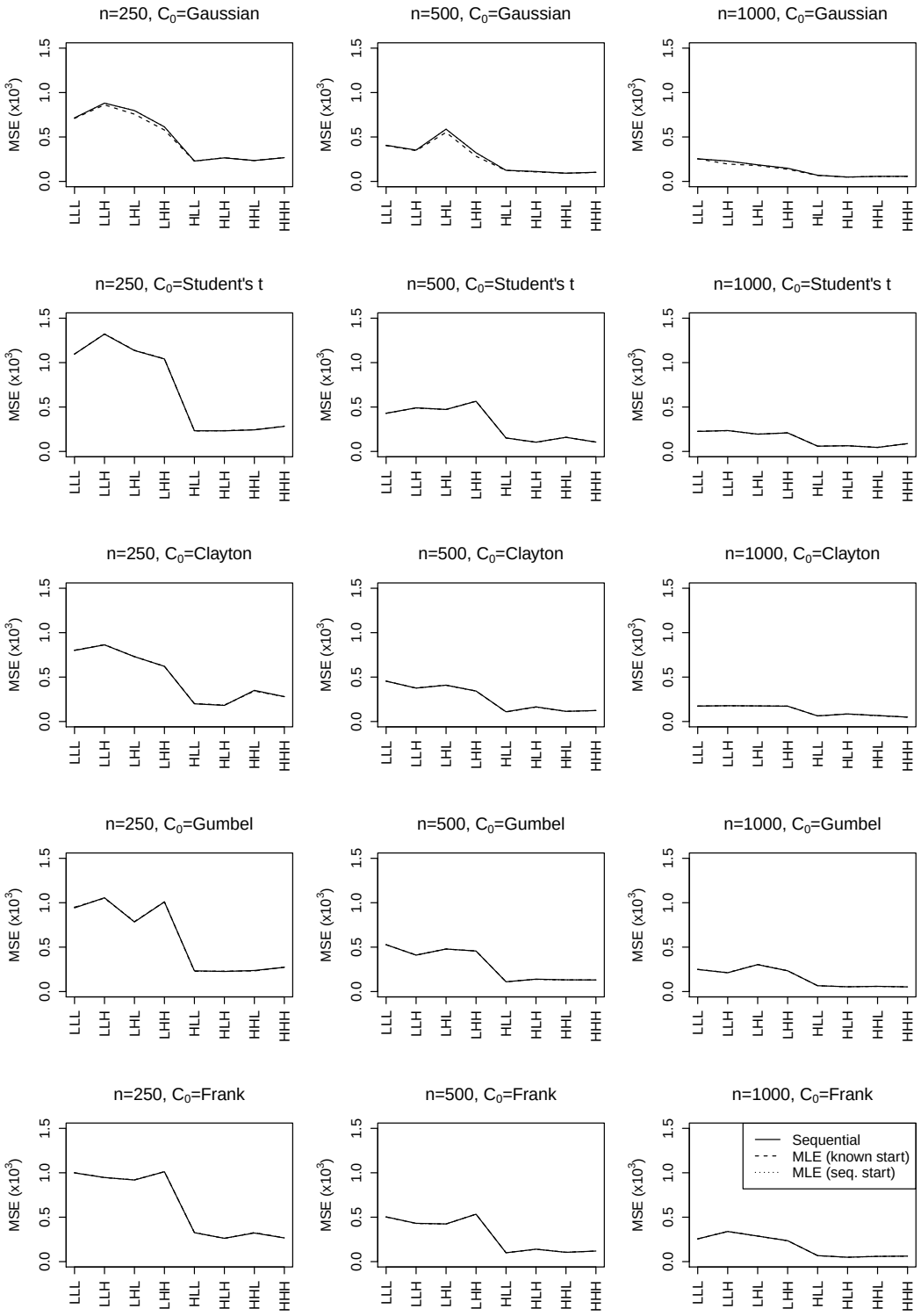


Figure C.3: Cluster copula 1: Gumbel. Cluster copula 2: Gumbel.

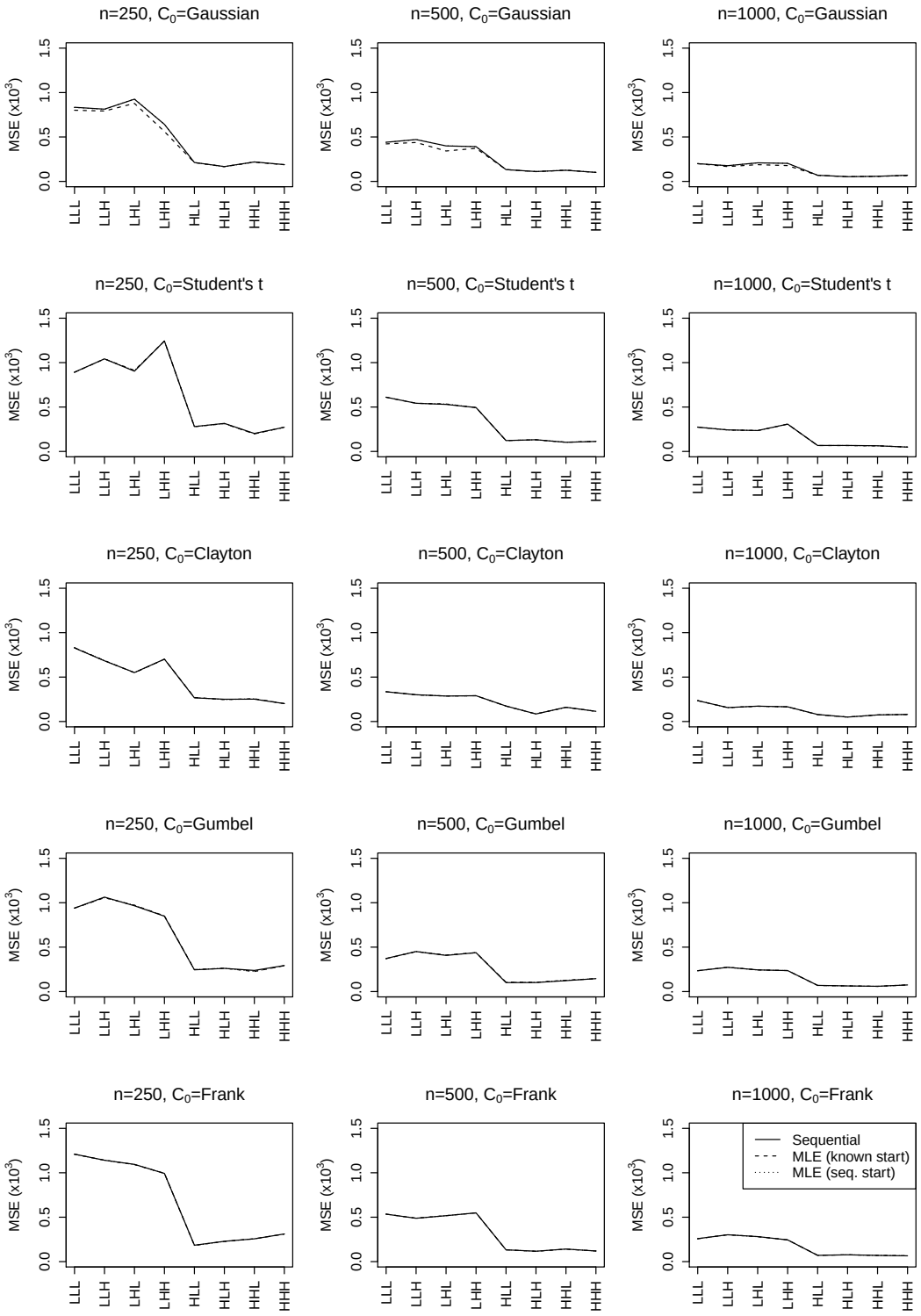


Figure C.4: Cluster copula 1: Gumbel. Cluster copula 2: Frank.

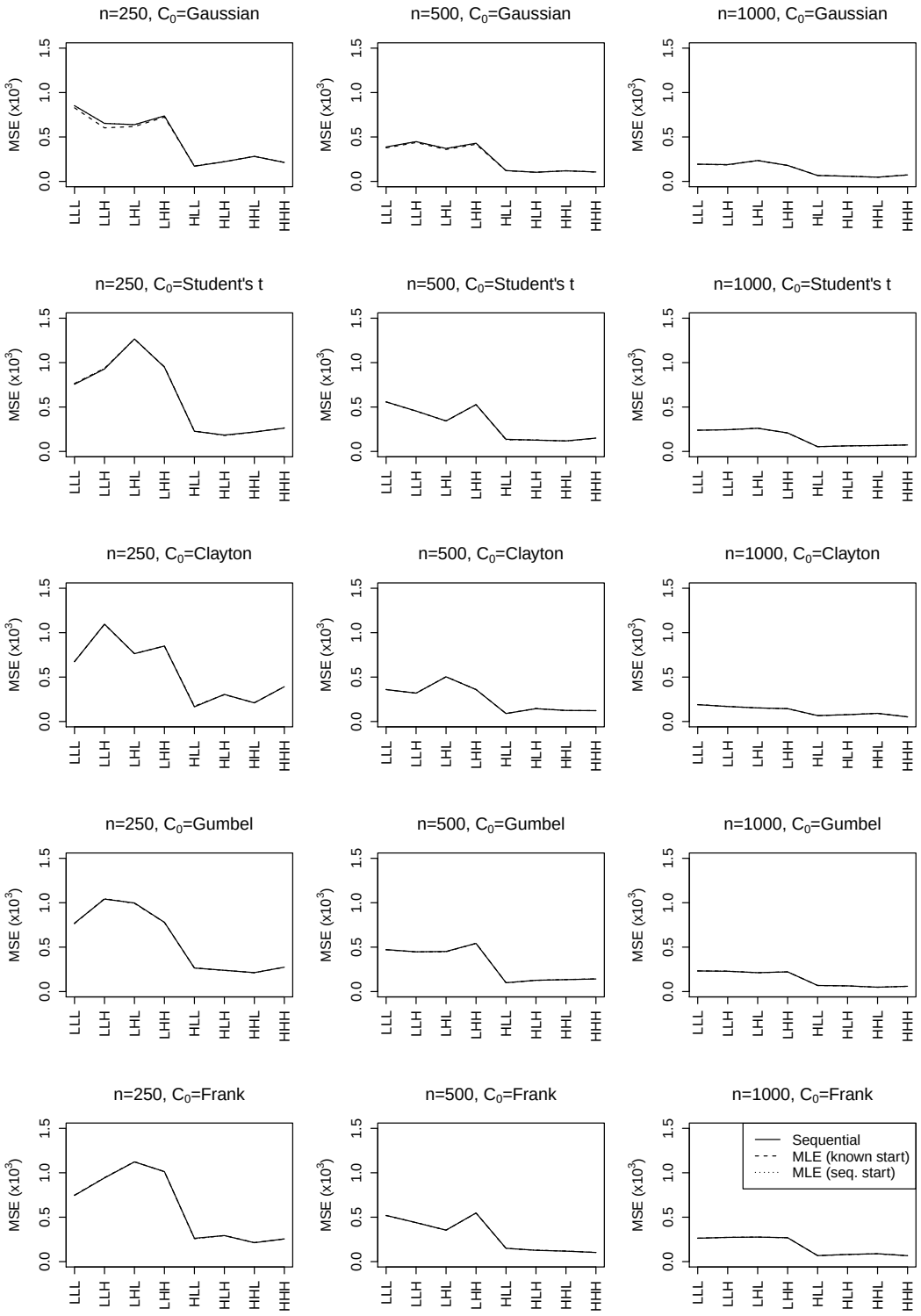


Figure C.5: Cluster copula 1: Frank. Cluster copula 2: Frank.

Bibliography

Aas, K., C. Czado, A. Frigessi, and H. Bakken (2009). Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics 44*(2), 182–198.

Acar, E. F., C. Genest, and J. Nešlehová (2012). Beyond simplified pair-copula constructions. *Journal of Multivariate Analysis 110*, 74–90.

Acharya, V., L. Pedersen, T. Philippon, and M. Richardson (2011). Measuring systemic risk. American Finance Association - Denver Meetings Paper.

Adrian, T. and M. Brunnermeier (2011). CoVaR. NBER Working Paper No. 17454.

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov and F. Csaki (Eds.), *Proceedings of the Second International Symposium on Information Theory Budapest*, Akademiai Kiado, pp. 267–281.

Ali, M. M., N. N. Mikhail, and M. S. Haq (1978). A class of bivariate distributions including the bivariate logistic. *Journal of Multivariate Analysis 8*(3), 405–412.

Ang, A. and J. Chen (2002). Asymmetric correlations of equity portfolios. *Journal of Financial Economics 63*(3), 443–494.

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.

Artzner, P., F. Delbaen, J. Eber, and D. Heath (1999). Coherent measures of risk. *Mathematical Finance 9*(3), 203–228.

Aue, F. and M. Kalkbrener (2006). LDA at work: Deutsche Bank's approach to quantifying operational risk. *Journal of Operational Risk 1*(4), 49–93.

Barbe, P., C. Genest, K. Ghoudi, and B. Rémillard (1996). On Kendall's process. *Journal of Multivariate Analysis 58*(2), 197–229.

Basel Committee on Banking Supervision (2006). International convergence of capital measurement and capital standards: A revised framework. Technical report, Bank for International Settlements.

Basel Committee on Banking Supervision (2009a). Observed range of practice in key elements of Advanced Measurement Approaches. Technical report, Bank for International Settlements.

Basel Committee on Banking Supervision (2009b). Results from the 2008 Loss Data Collection Exercise. Technical report, Bank for International Settlements.

Basel Committee on Banking Supervision (2011a). Global systemically important banks: Assessment methodology and the additional loss absorbency requirement. Technical report, Bank for International Settlements.

Basel Committee on Banking Supervision (2011b). Operational risk – Supervisory guidelines for the advanced measurement approaches. Technical report, Bank for International Settlements.

Bauer, A., C. Czado, and T. Klein (2012). Pair-copula constructions for non-Gaussian DAG models. *Canadian Journal of Statistics 40*(1), 86–109.

Bedford, T. and R. M. Cooke (2001). Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics and Artificial intelligence 32*(1–4), 245–268.

Bedford, T. and R. M. Cooke (2002). Vines - a new graphical model for dependent random variables. *Annals of Statistics 30*(4), 1031–1068.

Beliakov, G., A. Pradera, and T. Calvo (2008). *Aggregation Functions: A Guide for Practitioners*. Berlin: Springer.

Berg, D. (2009). Copula goodness-of-fit testing: An overview and power comparison. *European Journal of Finance 15*(7-8), 675–701.

Berg, D. and K. Aas (2009). Models for construction of higher-dimensional dependence: A comparison study. *European Journal of Finance 15*(7–8), 639–659.

Bernard, C., E. C. Brechmann, and C. Czado (2013). Statistical assessments of systemic risk measures. In J.-P. Fouque and J. Langsam (Eds.), *Handbook on Systemic Risk*. Cambridge: Cambridge University Press.

Billio, M., M. Getmansky, A. W. Lo, and L. Pelizzon (2012). Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *Journal of Financial Economics 104*(3), 535–559.

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

Brechmann, E. C. (2010). Truncated and simplified regular vines and their applications. Diploma Thesis, Faculty of Mathematics, Technische Universität München, Germany.

Brechmann, E. C. (2013a). Hierarchical Kendall copulas: Properties and inference. *Canadian Journal of Statistics*, forthcoming.

Brechmann, E. C. (2013b). Sampling from hierarchical Kendall copulas. *Journal de la Société Française de Statistique 154*(1), 192–209.

Brechmann, E. C. and C. Czado (2013). Risk management with high-dimensional vine copulas: An analysis of the Euro Stoxx 50. *Statistics & Risk Modeling*, forthcoming.

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.

Brechmann, E. C., C. Czado, and S. Paterlini (2013). Flexible dependence modeling of operational risk losses and its impact on total capital requirements. Submitted for publication.

Brechmann, E. C., K. Hendrich, and C. Czado (2013). Conditional copula simulation for systemic risk stress testing. *Insurance: Mathematics and Economics*, forthcoming.

Brechmann, E. C. and U. Schepsmeier (2013). Modeling dependence with C- and D-vine copulas: The R-package CDVine. *Journal of Statistical Software 52*(3), 1–27.

Calvo, T., G. Mayor, and R. Mesiar (2002). *Aggregation Operators: New Trends and Applications*. Heidelberg: Physica Verlag.

Campbell, S. D. (2007). A review of backtesting and backtesting procedures. *Journal of Risk 9*(2), 1–18.

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.

Capéraà, P., A.-L. Fougères, and C. Genest (1997). A nonparametric estimation procedure for bivariate extreme value copulas. *Biometrika 84*(3), 567–577.

Capéraà, P., A.-L. Fougères, and C. Genest (2000). Bivariate distributions with given extreme value attractor. *Journal of Multivariate Analysis 72*(1), 30–49.

Chakak, A. and M. Ezzerg (2000). Bivariate contours of copula. *Communications in Statistics - Simulation and Computation 29*(1), 175–185.

Chapelle, A., Y. Crama, G. Hübner, and J.-P. Peters (2008). Practical methods for measuring and managing operational risk in the financial sector: A clinical study. *Journal of Banking & Finance 32*(6), 1049–1061.

Chavez-Demoulin, V., P. Embrechts, and J. Nešlehová (2006). Quantitative models for operational risk: Extremes, dependence and aggregation. *Journal of Banking & Finance 30*(10), 2635–2658.

Chen, H., J. D. Cummins, K. S. Viswanathan, and M. A. Weiss (2013). Systemic risk and the inter-connectedness between banks and insurers: An econometric analysis. *Journal of Risk and Insurance*, forthcoming.

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.

Chu, Y.-S., N. Constantinou, and J. O'Hara (2010). An analysis of the determinants of the iTraxx CDS spreads using the skewed Student's t AR-GARCH model. University of Essex - Centre for Computational Finance and Aconomic Agents Working Paper Series 40.

Clayton, D. G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika 65*(1), 141–151.

Consul, P. C. and G. C. Jain (1973). A generalization of the Poisson distribution. *Technometrics 15*(4), 791–799.

Cont, R. and A. Moussa (2010). Too interconnected to fail: Contagion and systemic risk in financial networks. Columbia University - Financial Engineering Report 03.

Cont, R., A. Moussa, and E. B. Santos (2013). Network structure and systemic risk in banking systems. In J.-P. Fouque and J. Langsam (Eds.), *Handbook on Systemic Risk*. Cambridge: Cambridge University Press.

Cook, R. D. and M. E. Johnson (1981). A family of distributions for modeling non-elliptically symmetric multivariate data. *Journal of the Royal Statistical Society: Series B 43*(2), 210–218.

Cope, E. and G. Antonini (2008). Observed correlations and dependencies among operational losses in the ORX consortium database. *Journal of Operational Risk 3*(4), 47–74.

Czado, C. (2010). Pair-copula constructions of multivariate copulas. In P. Jaworski, F. Durante, W. Härdle, and T. Rychlik (Eds.), *Copula Theory and Its Applications*. Berlin: Springer.

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.

Czado, C., U. Schepsmeier, and A. Min (2012). Maximum likelihood estimation of mixed C-vines with application to exchange rates. *Statistical Modelling 12*(3), 229–255.

Dalla Valle, L., D. Fantazzini, and P. Giudici (2008). Copulae and operational risks. *International Journal of Risk Assessment and Management 9*(3), 238–257.

Daul, S., E. De Giorgi, F. Lindskog, and A. McNeil (2003). Using the grouped t-copula. *Risk 16*(11), 73–76.

Deb, P., P. K. Trivedi, and D. M. Zimmer (2013). Cost-offsets of prescription drug expenditures: Data analysis via a copula-based bivariate dynamic hurdle model. *Health Economics*, forthcoming.

Demarta, S. and A. J. McNeil (2005). The t copula and related copulas. *International Statistical Review 73*(1), 111–129.

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.

Embrechts, P., C. Klüppelberg, and T. Mikosch (1997). *Modelling Extremal Events for Insurance and Finance*. New York: Springer.

Embrechts, P., F. Lindskog, and A. McNeil (2003). Modelling dependence with copulas and applications to risk management. In S. T. Rachev (Ed.), *Handbook of Heavy Tailed Distributions in Finance*. Amsterdam: Elsevier.

Engle, R. F., D. M. Lilien, and R. P. Robins (1987). Estimating time varying risk premia in the term structure: The ARCH-M model. *Econometrica 55*(2), 391–407.

Erhardt, V. and C. Czado (2012). Modeling dependent claim totals including zero claims in private health care insurance. *Scandinavian Actuarial Journal 2012*(2), 106–299.

Eschenburg, P. (2013). Properties of extreme-value copulas. Diploma Thesis, Faculty of Mathematics, Technische Universität München, Germany.

Fang, K.-T. and B.-Q. Fang (1988). Generalized symmetrized Dirichlet distributions. *Acta Mathematicae Applicatae Sinica 4*(4), 316–323.

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

Financial Stability Board (2011). Policy measures to address systemically important financial institutions. Technical report.

Financial Stability Board (2012). Update of group of global systemically important banks. Technical report.

Financial Stability Board, International Monetary Fund, and Bank for International Settelments (2009). Report to the G-20 finance ministers and central bank governors: Guidance to assess the systemic importance of financial institutions, markets and instruments: Initial considerations. Technical report.

Frachot, A., T. Roncalli, and E. Salomon (2004). The correlation problem in operational risk. Working paper, Crédit Lyonnais.

Frahm, G., M. Junker, and A. Szimayer (2003). Elliptical copulas: applicability and limitations. *Statistics & Probability Letters 63*(3), 275–286.

Frank, M. J. (1979). On the simultaneous associativity of $f(x, y)$ and $x + y - f(x, y)$. *Aequationes Mathematicae 19*, 194–226.

Genest, C., I. Kojadinovic, J. Nešlehová, and J. Yan (2011). A goodness-of-fit test for bivariate extreme-value copulas. *Bernoulli 17*(1), 253–275.

Genest, C., J. Nešlehová, and J. Ziegel (2011). Inference in multivariate Archimedean copula models. *Test 20*(2), 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*(2), 337–366.

Genest, C., B. Rémillard, and D. Beaudoin (2009). Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and Economics 44*(2), 199–213.

Genest, C. and L.-P. Rivest (1989). A characterization of Gumbel's family of extreme value distributions. *Statistics & Probability Letters 8*(3), 207–211.

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.

Geneva Association (2010). Systemic risk in insurance - An analysis of insurance and financial stability. Technical report.

Genton, M. G. (2004). *Skew-Elliptical Distributions and Their Applications: A Journey Beyond Normality*. Boca Raton: Chapman & Hall/CRC.

Genz, A. and F. Bretz (2009). *Computation of Multivariate Normal and t Probabilities*. New York: Springer.

Geweke, J. F. (1991). Efficient simulation from the multivariate normal and Student-t distributions subject to linear constraints and the evaluation of constraint probabilities. In *Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface*, pp. 571–578.

Ghoudi, K., A. Khoudraji, and L.-P. Rivest (1998). Propriétés statistiques des copules de valeurs extrêmes bidimensionnelles. *Canadian Journal of Statistics 26*(1), 187–197.

Giacometti, R., S. Rachev, A. Chernobai, and M. Bertocchi (2008). Aggregation issues in operational risk. *Journal of Operational Risk 3*(3), 3–23.

Giglio, S. (2011). Credit default swap spreads and systemic financial risk. *Proceedings of the Federal Reserve Bank of Chicago 5*, 104–141.

Gonçalves, S. and H. White (2004). Maximum likelihood and the bootstrap for nonlinear dynamic models. *Journal of Econometrics 119*(1), 199–220.

Gourier, E., W. Farkas, and D. Abbate (2009). Operational risk quantification using extreme value theory and copulas: from theory to practice. *Journal of Operational Risk 4*(3), 3–26.

Gower, J. C. (1966). Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika 53*(3–4), 325–338.

Grabisch, M., J. L. Marichal, R. Mesiar, and E. Pap (2009). *Aggregation Functions*. Cambridge: Cambridge University Press.

Gräler, B., M. J. van den Berg, S. Vandenberghe, A. Petroselli, S. Grimaldi, B. De Baets, and N. E. C. Verhoest (2013). Multivariate return periods in hydrology: a critical and practical review focusing on synthetic design hydrograph estimation. *Hydrology and Earth System Sciences 17*(4), 1281–1296.

Gregory, J. and J.-P. Laurent (2004). In the core of correlation. *Risk 17*(10), 87–91.

Grogger, J. T. and R. T. Carson (1991). Models for truncated counts. *Journal of Applied Econometrics 6*(3), 225–238.

Grothe, O., F. Schmid, J. Schnieders, and J. Segers (2011). Measuring association between random vectors. Preprint available at http://arxiv.org/abs/1107.4381.

Gudendorf, G. and J. Segers (2010). Extreme-value copulas. In P. Jaworski, F. Durante, W. Härdle, and T. Rychlik (Eds.), *Copula Theory and Its Applications*. Berlin: Springer.

Gumbel, E. J. (1960). Distributions des valeurs extrêmes en plusiers dimensions. *Publications de l'Institut de Statistique de L'Université de Paris 9*, 171–173.

Haas, M. (2001). New methods in backtesting. Technical report, Research Center Caesar.

Hanea, A. M., D. Kurowicka, and R. M. Cooke (2006). Hybrid method for quantifying and analyzing Bayesian belief nets. *Quality and Reliability Engineering International 22*(6), 709–729.

Hashorva, E., S. Kotz, and A. Kume (2007). L_p-norm generalised symmetrised Dirichlet distributions. *Albanian Journal of Mathematics 1*(1), 31–56.

Hastie, T., R. Tibshirani, and J. H. Friedman (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2nd ed.). New York: Springer.

Hendrich, K. (2012). Copula-based analysis of interdependence among companies in the banking and insurance sector. Diploma Thesis, Faculty of Mathematics, Technische Universität München, Germany.

Hering, C. (2011). *Estimation techniques and goodness-of-fit tests for certain copula classes in large dimensions*. Ph. D. thesis, Universität Ulm.

Hobæk Haff, I. (2012). Comparison of estimators for pair-copula constructions. *Journal of Multivariate Analysis 110*, 91–105.

Hobæk Haff, I. (2013). Parameter estimation for pair-copula constructions. *Bernoulli 19*(2), 462–491.

Hobæk Haff, I., K. Aas, and A. Frigessi (2010). On the simplified pair-copula construction - simply useful or too simplistic? *Journal of Multivariate Analysis 101*(5), 1296–1310.

Hofert, M. (2008). Sampling Archimedean copulas. *Computational Statistics & Data Analysis 52*(12), 5163–5174.

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.

Holzinger, K. J. and F. Swineford (1937). The bi-factor method. *Psychometrika 2*(1), 41–54.

Hua, L. and H. Joe (2011). Tail order and intermediate tail dependence of multivariate copulas. *Journal of Multivariate Analysis 102*(10), 1454–1471.

Huang, X., H. Zhou, and H. Zhu (2009). A framework for assessing the systemic risk of major financial institutions. *Journal of Banking & Finance 33*(11), 2036–2049.

Hull, J., M. Predescu, and A. White (2004). The relationship between credit default swap spreads, bond yields, and credit rating announcements. *Journal of Banking & Finance 28*(11), 2789–2811.

Hürlimann, W. (2003). Hutchinson-Lai's conjecture for bivariate extreme value copulas. *Statistics & Probability Letters 61*(2), 191–198.

Iman, R. L. and W. J. Conover (1982). A distribution-free approach to inducing rank correlation among input variables. *Communcations in Statistics - Simulation and Computation 11*(3), 311–334.

Imlahi, L., M. Ezzerg, and A. Chakak (1999). Estimación de la curva mediana de una cópula $C(x_1, ..., x_m)$. *Revista de la Real Academia de Ciencias Exactas, Físicas y Naturales. Serie A: Matemáticas 93*(2), 241–250.

International Association of Insurance Supervisors (2012). Global systemically important insurers: Proposed assessment methodology. Technical report.

Joe, H. (1993). Parametric families of multivariate distributions with given margins. *Journal of Multivariate Analysis 46*(2), 262–282.

Joe, H. (1996). Families of m-variate distributions with given margins and $m(m-1)/2$ bivariate dependence parameters. In L. Rüschendorf, B. Schweizer, and M. D. Taylor (Eds.), *Distributions with fixed marginals and related topics*. Hayward: Institute of Mathematical Statistics.

Joe, H. (1997). *Multivariate Models and Dependence Concepts*. London: Chapman & Hall.

Joe, H. (2005). Asymptotic efficiency of the two-stage estimation method for copula-based models. *Journal of Multivariate Analysis 94*(2), 401–419.

Joe, H., H. Li, and A. K. Nikoloulopoulos (2010). Tail dependence functions and vine copulas. *Journal of Multivariate Analysis 101*(1), 252–270.

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.

Johnson, N., S. Kotz, and N. Balakrishnan (1997). *Discrete Multivariate Distributions*. New York: John Wiley.

Jondeau, E. and M. Rockinger (2006). The copula-GARCH model of conditional dependencies: An international stock market application. *Journal of International Money and Finance 25*(5), 827–853.

Kaushik, R. and S. Battiston (2013). Credit default swaps drawup networks: Too interconnected to be stable? *PLoS ONE 8*(7).

Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika 30*(1–2), 81–93.

Khoudraji, A. (1995). *Contributions à l'etude des copules et à la modélisation des valeurs extrêmes bivariées*. Ph. D. thesis, Université Laval Québec.

Kimberling, C. H. (1974). A probabilistic interpretation of complete monotonicity. *Aequationes Mathematicae 10*(2–3), 152–164.

Kimeldorf, G. and A. R. Sampson (1975). Uniform representations of bivariate distributions. *Communications in Statistics 4*, 617–627.

Klugman, S. A., H. H. Panjer, and G. E. Willmot (2008). *Loss Models: From Data to Decisions* (3rd ed.). New York: John Wiley.

Kotz, S., N. Balakrishnan, and N. L. Johnson (2000). *Continuous Multivariate Distributions. Volume 1: Models and Applications*. New York: Wiley.

Kotz, S. and S. Nadarajah (2004). *Multivariate t Distributions and Their Applications*. Cambridge: Cambridge University Press.

Krupskii, P. and H. Joe (2013a). Factor copula models for multivariate data. *Journal of Multivariate Analysis 120*, 85–101.

Krupskii, P. and H. Joe (2013b). Structured factor copula models: theory, inference and computation. Technical report, Department of Statistics, University of British Columbia.

Kullback, S. and R. A. Leibler (1951). On information and sufficiency. *Annals of Mathematical Statistics 22*(1), 79–86.

Kupiec, P. H. (1995). Techniques for verifying the accuracy of risk measurement models. *The Journal of Derivatives 3*(2), 73–84.

Kurowicka, D. (2011). Optimal truncation of vines. In D. Kurowicka and H. Joe (Eds.), *Dependence Modeling: Vine Copula Handbook*. Singapore: World Scientific Publishing Co.

Kurowicka, D. and R. M. Cooke (2003). A parametrization of positive definite matrices in terms of partial correlation vines. *Linear Algebra and its Applications 372*, 225–251.

Kurowicka, D. and R. M. Cooke (2006). *Uncertainty Analysis with High Dimensional Dependence Modelling*. Chichester: John Wiley.

Kurowicka, D. and H. Joe (2011). *Dependence Modeling: Vine Copula Handbook*. Singapore: World Scientific Publishing Co.

Lewandowski, D., D. Kurowicka, and H. Joe (2009). Generating random correlation matrices based on vines and extended onion method. *Journal of Multivariate Analysis 100*(9), 1989–2001.

Liu, Y. and R. Luger (2009). Efficient estimation of copula-GARCH models. *Computational Statistics & Data Analysis 53*(6), 2284–2297.

Longin, F. and B. Solnik (2001). Extreme correlation of international equity markets. *Journal of Finance 56*(2), 649–676.

Luo, X. and P. V. Shevchenko (2010). The t copula with multiple parameters of degrees of freedom: bivariate characteristics and application to risk management. *Quantitative Finance 10*(9), 1039–1054.

Mai, J.-F. and M. Scherer (2012). *Simulating Copulas: Stochastic Models, Sampling Algorithms, and Applications*. Singapore: World Scientific Publishing Co.

Mainik, G. (2012). Sampling dependence: empirical copulas, Latin hypercubes, and convergence of sums. Preprint available at http://www.math.ethz.ch/~gmainik/Papers/lhs-dep-sums.pdf.

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*(4), 259–265.

Mardia, K. V. (1970). *Families of Bivariate Distributions*. London: Griffin.

Markose, S., S. Giansante, and A. R. Shaghaghi (2012). "Too interconnected to fail" Financial network of US CDS market: Topological fragility and systemic risk. *Journal of Economic Behavior & Organization 83*(3), 627–646.

Marshall, A. W. and I. Olkin (1967). A multivariate exponential distribution. *Journal of the American Statistical Association 62*(317), 30–44.

McLeish, D. L. and C. G. Small (1988). *The Theory and Applications of Statistical Inference Functions*. New York: Springer.

McNeil, A. J. (2008). Sampling nested Archimedean copulas. *Journal of Statistical Computation and Simulation 78*(6), 567–581.

McNeil, A. J., R. Frey, and P. Embrechts (2005). *Quantitative Risk Management: Concepts Techniques and Tools*. Princeton: Princeton University Press.

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.

Mesfioui, M. and J.-F. Quessy (2008). Dependence structure of conditional Archimedean copulas. *Journal of Multivariate Analysis 99*(3), 372–385.

Mittnik, S., S. Paterlini, and T. Yener (2013). Operational-risk dependencies and the determination of risk capital. *Journal of Operational Risk*, forthcoming.

Morales-Nápoles, O. (2011). Counting vines. In D. Kurowicka and H. Joe (Eds.), *Dependence Modeling: Vine Copula Handbook*. Singapore: World Scientific Publishing Co.

Nelsen, R. B. (2006). *An Introduction to Copulas* (2nd ed.). Berlin: Springer.

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.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica 59*(2), 347–70.

Nikoloulopoulos, A. K., H. Joe, and H. Li (2009). Extreme value properties of multivariate t copulas. *Extremes 12*(2), 129–148.

Nikoloulopoulos, A. K., H. Joe, and H. Li (2012). Vine copulas with asymmetric tail dependence and applications to financial return data. *Computational Statistics & Data Analysis 56*(11), 3659–3673.

Nikoloulopoulos, A. K., H. Joe, and H. Li (2013). Factor copula models for item response data. *Psychometrika*, forthcoming.

Norden, L. and M. Weber (2004). Informational efficiency of credit default swap and stock markets: The impact of credit rating announcements. *Journal of Banking & Finance 28*(11), 2813–2843.

Okhrin, O., Y. Okhrin, and W. Schmid (2013). On the structure and estimation of hierarchical Archimedean copulas. *Journal of Econometrics 173*(2), 189–204.

Panagiotelis, A., C. Czado, and H. Joe (2012). Pair copula constructions for multivariate discrete data. *Journal of the American Statistical Association 107*(499), 1063–1072.

Park, S. C. and X. Xie (2011). Reinsurance and systemic risk: The impact of reinsurer downgrading on property-casualty insurers. Technical report, California State University.

Patton, A. J. (2006). Estimation of multivariate models for time series of possibly different lengths. *Journal of Applied Econometrics 21*(2), 147–173.

Pickands, J. (1981). Multivariate extreme value distributions. In *Proceedings of the 43rd Session of the International Statistical Institute, Buenos Aires*, pp. 859–878.

Plackett, R. L. (1965). A class of bivariate distributions. *Journal of the American Statistical Association 60*(310), 516–522.

Politis, D. N. and J. P. Romano (1994). The stationary bootstrap. *Journal of the American Statistical Association 89*(428), 1303–1313.

Rahman, D. (2013). Are banking systems increasingly fragile? Investigating financial institutions' CDS returns extreme co-movements. *Quantitative Finance*, forthcoming.

Rubin, D. B. (1987). A noniterative sampling/importance resampling alternative to the data augmentation algorithm for creating a few imputations when fractions of missing information are modest: the SIR algorithm. *Journal of the American Statistical Association 82*(398), 543–546.

Rubin, D. B. (1988). Using the SIR algorithm to simulate posterior distributions. In J. M. Bernardo, M. H. DeGroot, D. V. Lindley, and A. F. M. Smith (Eds.), *Bayesian Statistics 3*. Oxford: Oxford University Press.

Salmon, F. (2009). Recipe for Disaster: The Formula That Killed Wall Street. *Wired Magazine 17*(3).

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*(3), 295–304.

Schepsmeier, U. (2010). Maximum likelihood estimation of C-vine pair-copula constructions based on bivariate copulas from different families. Diploma Thesis, Faculty of Mathematics, Technische Universität München, Germany.

Schepsmeier, U. and E. C. Brechmann (2011). *CDVine: Statistical inference of C- and D-vine copulas*. R package.

Schepsmeier, U., J. Stöber, and E. C. Brechmann (2012). *VineCopula: Statistical inference of vine copulas*. R package.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics 6*(2), 461–464.

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.

Song, P. X.-K. (2007). *Correlated Data Analysis: Modeling, Analytics, and Applications*. New York: Springer.

Spearman, C. (1904). The proof and measurement of association between two things. *American Journal of Psychology 15*(1), 72–101.

Stöber, J., H. Joe, and C. Czado (2013). Simplified pair copula constructions - Limits and extensions. *Journal of Multivariate Analysis 119*, 101–118.

Stöber, J. and U. Schepsmeier (2013). Estimating standard errors in regular vine copula models. *Computational Statistics*, forthcoming.

Tawn, J. (1988). Bivariate extreme value theory: Models and estimation. *Biometrika 75*(3), 397–415.

Wang, W. and M. T. Wells (2000). Model selection and semiparametric inference for bivariate failure-time data. *Journal of the American Statistical Association 95*(449), 62–72.

Whelan, N. (2004). Sampling from Archimedean copulas. *Quantitative Finance 4*(3), 339–352.

Wu, F., E. Valdez, and M. Sherris (2007). Simulating from exchangeable Archimedean copulas. *Communications in Statistics - Simulation and Computation 36*(5), 1019–1034.