

# Copulae in High Dimensions: An Introduction

Ostap Okhrin<sup>1</sup>, Alexander Ristig<sup>\*2</sup>, and Yafei Xu<sup>1, 3</sup>

<sup>1</sup>*Institute of Transportation and Economics, Technische Universität Dresden, Germany*

<sup>2</sup>*Department of Statistics and Operations Research, Universität Wien, Austria*

<sup>3</sup>*School of Business and Economics, Humboldt-Universität zu Berlin, Germany*

## 1 Introduction

Researches of dependence modeling were burgeoning during the last decade. The traditional approaches that concentrate on the elliptical distributions such as Gaussian models are giving way to copula-based models. Albeit these Gaussian models sometimes own the convenience in model construction and computation, yet an abundant amount of empirical evidences do not support the underlying assumptions. De facto, shortcomings in the elliptical and especially Gaussian family are mainly in lack of asymmetrical and tail dependence which have been deeply discussed in numerous papers. Furthermore and of great importance, margins of elliptical distributions belong to the same elliptical family.

The seminal result of Sklar (1959) provides a partial solution to these problems. It allows to separate the marginal distributions from the dependency structure between the random variables. Since the theory on modeling and estimation of univariate distributions is well established compared to the multivariate case, the initial problem reduces to modeling the dependency by copulas. In particular, this approach dramatically widens the class of candidate distributions and allows a simple construction of distributions with less parameters than imposed by elliptical models.

In the beginning of the copula hype, researches were mainly focused on the bivariate dependence but as time passes problems raised by the financial, technological, biological

---

\*The support by the Vienna Science and Technologie Fund (WWTF) through project MA14-008 is gratefully acknowledged.

industries dictated the rules of further developments, namely moves to higher dimensions. Nonetheless, it has been realized as clearly stated in Mai and Scherer (2013), that “the step from one-dimensional modeling is clearly large. But, unfortunately, the step from two to three (or even more) dimensions is not a bit smaller.”.

Numerous steps are done in order to contribute to research on high-dimensional modeling approaches and these main branches have been established: pair copula construction, see Joe (1996), Bedford and Cooke (2001), Bedford and Cooke (2002) and Kurowicka and Cooke (2006), hierarchical Archimedean copula, see Savu and Trede (2010), Hofert (2011) and Okhrin, Okhrin and Schmid (2013a), and factor copula, see Krupskii and Joe (2013) and Oh and Patton (2015).

This chapter attempts at discussing such non-standard multivariate copula models and the subsequent sections are organized as follows. We introduce bivariate copulae and review modern multivariate copula families. Then, corresponding estimation methods and goodness of fit tests are presented. Last but not least, we study a risk management topic empirically.

## 2 Bivariate Copula

Modeling the dependence between only two random variables using copulae is the subject of this section. There are several equivalent definitions of the copula function. We define it as a bivariate distribution function and the simplest one is as follows:

**Definition 1.** *The copula  $C(u, v)$  is a bivariate distribution with margins being  $U[0, 1]$ .*

Term copula was mentioned for the first time in the seminal result of Sklar (1959). The separation of the bivariate distribution function into the copula function and margins is formally stated in the subsequent theorem. One possible proof is presented in Nelsen (2006, Section 2.3), for others we refer to Durante, Fernández-Sánchez and Sempi (2012), Durante, Fernández-Sánchez and Sempi (2013) and Durante and Sempi (2005, Chapter 2)

**Theorem 1.** *Let  $F$  be a bivariate distribution function with margins  $F_1$  and  $F_2$ , then there exists a copula  $C$  such that*

$$F(x_1, x_2) = C\{F_1(x_1), F_2(x_2)\}, \quad x_1, x_2 \in \overline{\mathbb{R}} = \mathbb{R} \cup \{\infty, -\infty\}. \quad (1)$$

If  $F_1$  and  $F_2$  are continuous than  $C$  is unique. Otherwise  $C$  is uniquely determined on  $F_1(\overline{\mathbb{R}}) \times F_2(\overline{\mathbb{R}})$ .

Conversely, if  $C$  is a copula and  $F_1$  and  $F_2$  are univariate distribution functions, then function  $F$  in (1) is a bivariate distribution function with margins  $F_1$  and  $F_2$ .

As indicated above, the theorem allows decomposing any continuous bivariate distribution into its marginal distributions and the dependency structure. Since by definition, the latter is the copula function with uniform margins, it follows that the copula density can be determined in the usual way

$$c(u_1, u_2) = \frac{\partial^2 C(u_1, u_2)}{\partial u_1 \partial u_2}, \quad u_1, u_2 \in [0, 1]. \quad (2)$$

Being armed with the Theorem 1 and (2), the density function  $f(\cdot)$  of the bivariate distribution  $F$  can be rewritten in terms of copula

$$f(x_1, x_2) = c\{F_1(x_1), F_2(x_2)\}f_1(x_1)f_2(x_2), \quad x_1, x_2 \in \overline{\mathbb{R}}.$$

A very important property of copulae is given in Nelsen (2006, Theorem 2.4.3) stating that copulae are invariant under strictly monotone transformations of margins. Seen from this angle, copulae capture only those features of the dependency which are invariant under increasing transformations.

## 2.1 Copula Families

Naturally, there is an infinite number of different copula functions satisfying the properties of Definition 1 and the number of them being deeply studied is expanding. In this section, we discuss three copula classes namely simple, elliptical and Archimedean copulae.

### 2.1.1 Simplest Copulae

To form basic intuition for copula functions, we first study some extreme special cases, like stochastically independent, perfect positive or negative dependent random variables. According to Theorem 1, the copula of two stochastically independent random variables  $X_1$  and  $X_2$  is given by the product (independence) copula defined as

$$\Pi(u_1, u_2) = u_1 u_2, \quad u_1, u_2 \in [0, 1].$$

The contour diagrams of the bivariate density function with product copula and either Gaussian or  $t_3$ -distributed margins are given in Figure 1. Two additional extremes are the lower and upper Fréchet-Hoeffding bounds. They represent the perfect negative and positive dependence of two random variables respectively

$$W(u_1, u_2) = \max(0, u_1 + u_2 - 1) \quad \text{and} \quad M(u_1, u_2) = \min(u_1, u_2), \quad u_1, u_2 \in [0, 1].$$

If  $C = W$  and  $(X_1, X_2) \sim C(F_1, F_2)$  then  $X_2$  is a decreasing function of  $X_1$ . Similarly, if  $C = M$ , then  $X_2$  is an increasing function of  $X_1$ . In general, we can argue that an arbitrary copula which represents some dependency structure lies between these two bounds, i.e.

$$W(u_1, u_2) \leq C(u_1, u_2) \leq M(u_1, u_2), \quad u_1, u_2 \in [0, 1].$$

The bounds serve as benchmarks for the evaluation of the dependency magnitude. There are numerous techniques for building new copulae by mixing at least two of the presented simplest copula. For example, copula families B11 and B12, see Joe (1997), arise as a combination of the upper Fréchet-Hoeffding bound and the product copula

$$\begin{aligned} C_{B11}(u_1, u_2, \theta) &= \theta M(u_1, u_2) + (1 - \theta)\Pi(u_1, u_2) = \theta \min\{u_1, u_2\} + (1 - \theta)u_1 u_2, \\ C_{B12}(u_1, u_2, \theta) &= M(u_1, u_2)^\theta \Pi(u_1, u_2)^{1-\theta} = (\min\{u_1, u_2\})^\theta (u_1 u_2)^{1-\theta}, \quad u_1, u_2, \theta \in [0, 1]. \end{aligned}$$

Family B11 builds on the fact that every convex combination of copulas is a copula as well. Family B12 is also known as Spearman or Cuadras-Augé copula, which is a weighted geometric mean of the upper Fréchet-Hoeffding bound and the product copula. Further generalization is done by using power mean over the upper Fréchet-Hoeffding bound and the product copula

$$\begin{aligned} C_p(u_1, u_2, \theta_1, \theta_2) &= \{\theta_1 M^{\theta_2}(u_1, u_2) + (1 - \theta_1)\Pi^{\theta_2}(u_1, u_2)\}^{1/\theta_2} \\ &= \{\theta_1 \min(u_1, u_2)^{\theta_2} + (1 - \theta_1)(u_1 u_2)^{\theta_2}\}^{1/\theta_2}, \end{aligned}$$

with  $\theta_1 \in [0, 1]$ ,  $\theta_2 \in \mathbb{R}$ . Last but not least, a convex combination of the Fréchet-Hoeffding lower bound, upper bound and product copula forms the Fréchet copula

$$C_F(u_1, u_2, \theta_1, \theta_2) = \theta_1 W(u_1, u_2) + (1 - \theta_1 - \theta_2)\Pi(u_1, u_2) + \theta_2 M(u_1, u_2),$$

subject to  $0 \leq \theta_1 + \theta_2 \leq 1$ . Note that any bivariate copula can be approximated by the Fréchet family and a bound of the resulting approximation error can be estimated. Nelsen (2006, Chapter 3) provides further methods for constructing multivariate copulas and discusses convex combination in more detail.

### 2.1.2 Elliptical Family

Due to the popularity of the Gaussian and  $t$ -distribution in several applications, elliptical copulae play an important role as well. The construction of this type of copulae is directly based on Sklar's Theorem showing how new bivariate distributions can be constructed. The copula-based modeling approach substantially widens the family of elliptical distributions by keeping the same elliptical copula function and varying the marginal distributions or vice versa.

To determine the copula function of a given bivariate distribution, we employ the transformation

$$C(u_1, u_2) = F\{F_1^{-1}(u_1), F_2^{-1}(u_2)\}, \quad u_1, u_2 \in [0, 1], \quad (3)$$

where  $F_i^{-1}$ ,  $i = 1, 2$ , are (generalized) inverses of the marginal distribution functions. Based on (3), arbitrary elliptical distributions can be derived. The problem, however, is that such copulae depend on the inverse distribution functions of the marginals which are rarely available in an explicit form.

For instance, from Formula 3 follows that the Gaussian copula and its density are given by

$$\begin{aligned} C_N(u_1, u_2, \delta) &= \Phi_\delta(\Phi^{-1}(u_1), \Phi^{-1}(u_2)), \\ c_N(u_1, u_2, \delta) &= (1 - \delta^2)^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2}(1 - \delta^2)^{-1}(u_1^2 + u_2^2 - 2\delta u_1 u_2) \right\} \\ &\quad \times \exp \left\{ \frac{1}{2}(u_1^2 + u_2^2) \right\}, \quad \text{for all } u_1, u_2 \in [0, 1], \delta \in [-1, 1], \end{aligned}$$

where  $\Phi$  is the distribution function of  $N(0, 1)$ ,  $\Phi^{-1}$  is the functional inverse of  $\Phi$  and  $\Phi_\delta$  denotes the bivariate standard normal distribution function with correlation coefficient  $\delta$ .

In the bivariate case, the  $t$ -copula and its density are given by

$$\begin{aligned} C_t(u_1, u_2, \nu, \delta) &= \int_{-\infty}^{t_\nu^{-1}(u_1)} \int_{-\infty}^{t_\nu^{-1}(u_2)} \frac{\Gamma(\frac{\nu+2}{2})}{\Gamma(\frac{\nu}{2})\pi\nu\sqrt{(1-\delta^2)}} \\ &\quad \times \left\{ 1 + \frac{x_1^2 - 2\delta x_1 x_2 + x_2^2}{(1-\delta^2)\nu} \right\}^{-\frac{\nu}{2}-1} dx_1 dx_2, \\ c_t(u_1, u_2, \nu, \delta) &= \frac{f_{\nu, \delta}\{t_\nu^{-1}(u_1), t_\nu^{-1}(u_2)\}}{f_\nu\{t^{-1}(u_1)\}f_\nu\{t^{-1}(u_2)\}}, \quad u_1, u_2, \delta \in [0, 1], \end{aligned}$$

where  $\delta$  denotes the correlation coefficient,  $\nu$  is the number of degrees of freedom.  $f_{\nu, \delta}$  and  $f_\nu$  are joint and marginal  $t$ -distributions respectively, while  $t_\nu^{-1}$  denotes the quantile

function of the  $t_\nu$  distribution. In-depth analysis of the  $t$ -copula is done in Rachev, Stoyanov and Fabozzi (2008) and Luo and Shevchenko (2010). Long-tailed distributed margins lead to more mass and variability in the tail areas of the corresponding bivariate distribution. However, the contour-curves of the  $t$ -copula are symmetric, which reflects the ellipticity of the underlying copula. This property is theoretically supported by Nelsen (2006, Theorem 2.3.7), stating that a bivariate copula is elliptical and thus, has reflection symmetry, if and only if

$$C(u_1, u_2, \theta) = u_1 + u_2 - 1 + C(1 - u_1, 1 - u_2, \theta), \quad u_1, u_2 \in [0, 1].$$

The next class of copulae and their generalizations provide an important flexible and rich family of alternatives to elliptical copulae.

### 2.1.3 Archimedean Family

In contrast to elliptical copulae, Archimedean copulae are not constructed via (3), but are related to Laplace transforms of bivariate distribution functions. The function  $C : [0, 1]^2 \rightarrow [0, 1]$  defined as

$$C(u_1, u_2) = \phi\{\phi^{-1}(u_1) + \phi^{-1}(u_2)\}, \quad u_1, u_2 \in [0, 1]$$

is a 2-dimensional Archimedean copula, where  $\phi \in \mathcal{L} = \{\phi : [0; \infty) \rightarrow [0, 1] \mid \phi(0) = 1, \phi(\infty) = 0; (-1)^j \phi^{(j)} \geq 0; j = 1, \dots, \infty\}$  is referred to as the generator of the copula. The generator usually depends on some parameters, however, mostly generators with a single parameter  $\theta$  are considered. Nelsen (2006) and Joe (2014) provide a thoroughly classified list of popular generators for Archimedean copulae and discuss their properties.

The useful applications in finance, see Patton (2012), appearing to be the Gumbel copula with the generator function  $\phi(x, \theta) = \exp\{-x^{1/\theta}\}$ ,  $1 \leq \theta < \infty$ ,  $x \in [0, 1]$ , leading to the copula function

$$C(u_1, u_2, \theta) = \exp\left\{-\left[(-\log u_1)^\theta + (-\log u_2)^\theta\right]^{1/\theta}\right\}, \quad u_1, u_2 \in [0, 1].$$

Genest and Rivest (1989) showed that a bivariate distribution based on the Gumbel copula with extreme valued marginal distributions is the only bivariate extreme value distribution belonging to the Archimedean family. Moreover, all distributions based on Archimedean copulae belong to its domain of attraction under common regularity conditions. In contrary to elliptical copulae, the Gumbel copula leads to asymmetric contour diagrams in

Figure 1. It exhibits a stronger linkage between positive values, however, more variability and more mass in the negative tail area. Opposite is observed for the Clayton copula with the generator  $\phi(x, \theta) = (\theta x + 1)^{-\frac{1}{\theta}}$  with  $-1 < \theta < \infty$ ,  $\theta \neq 0$ ,  $x \in [0, 1]$ , and copula function

$$C(u_1, u_2, \theta) = (u_1^{-\theta} + u_2^{-\theta} - 1)^{-\frac{1}{\theta}}, \quad u_1, u_2 \in [0, 1].$$

Also, the Frank generator  $\phi(x, \theta) = \theta^{-1} \log\{1 - (1 - e^{-\theta})e^{-x}\}$  with  $0 \leq \theta < \infty$ ,  $x \in [0, 1]$ , enjoys increased popularity and induces the copula function

$$C(u_1, u_2, \theta) = -\theta^{-1} \log \left\{ \frac{1 - e^{-\theta} - (1 - e^{-\theta u_1})(1 - e^{-\theta u_2})}{1 - e^{-\theta}} \right\}, \quad u_1, u_2 \in [0, 1].$$

The respective Frank copula is the only elliptical Archimedean copula.

## 2.2 Bivariate Copula and Dependence Measures

Since copulae define the dependence structure between random variables, there is a relationship between copulae and different dependency measures. The classical measures for continuous random variables are Kendall's  $\tau$  and Spearman's  $\rho$ . Similarly as copula functions, these measures are invariant under strictly increasing transformations. They are equal to 1 or  $-1$  under perfect positive or negative dependence respectively. In contrast to  $\tau$  and  $\rho$ , the Pearson correlation coefficient measures the linear dependence and, therefore, is not suitable for measuring non-linear relationships. Next, we discuss the relationship between  $\tau$ ,  $\rho$  and the underlying copula function.

**Definition 2.** Let  $F$  be a continuous bivariate cumulative distribution function with the copula  $C$ . Moreover, let  $(X_1, X_2) \sim F$  and  $(X'_1, X'_2) \sim F$  be independent random pairs. Then Kendall's  $\tau$  is given by

$$\begin{aligned} \tau_2 &= P\{(X_1 - X'_1)(X_2 - X'_2) > 0\} - P\{(X_1 - X'_1)(X_2 - X'_2) < 0\} \\ &= 2 P\{(X_1 - X'_1)(X_2 - X'_2) > 0\} - 1 = 4 \int_{[0,1]^2} C(u_1, u_2) dC(u_1, u_2) - 1. \end{aligned}$$

Kendall's  $\tau$  represents the difference between the probability of two random concordant pairs and the probability of two random discordant pairs. For most copula functions with

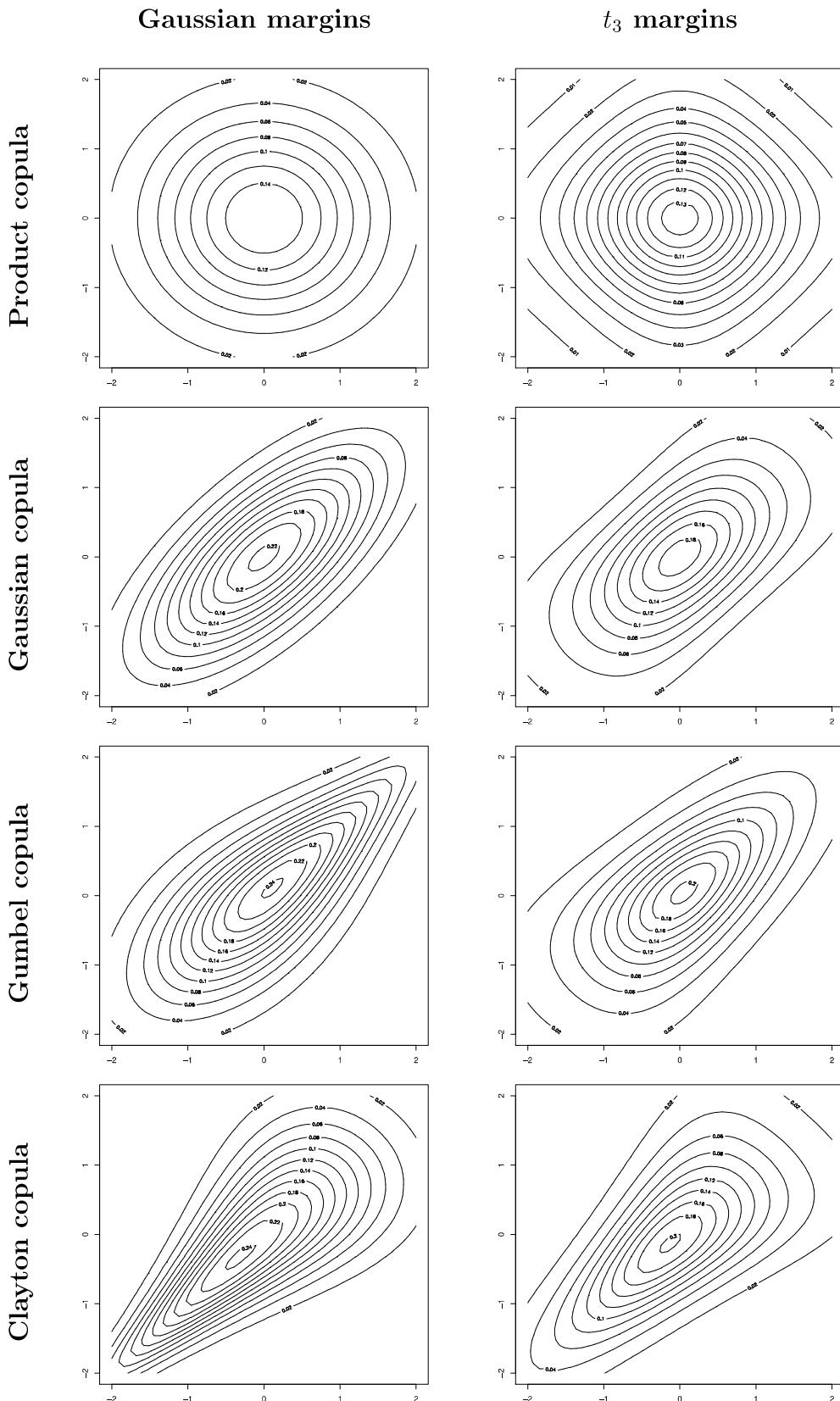


Figure 1: Contour diagrams for product, Gaussian, Gumbel and Clayton copulae with Gaussian (left column) and  $t_3$  distributed (right column) margins.

a single parameter  $\theta$  there is a one-to-one relationship between  $\theta$  and the Kendall's  $\tau_2$ . For example, it holds that

$$\begin{aligned}\tau_2(\text{Gaussian and } t) &= \frac{2}{\pi} \arcsin \delta, & \tau_2(\text{Archimedean}) &= 4 \int_0^1 \frac{\phi^{-1}(t)}{(\phi^{-1})'} dt + 1, \\ \tau_2(\Pi) &= 0, & \tau_2(W) &= 1, & \tau_2(M) &= -1.\end{aligned}$$

For instance, this implies that an unknown copula parameter  $\theta$  of the Gaussian,  $t$  and an arbitrary Archimedean copulae can be estimated using a type of method of moments procedure with a single moment condition. This requires, however, an estimator of  $\tau_2$ , c.f. Kendall (1970). Naturally, it is computed by

$$\tau_{2n} = \frac{4}{n(n-1)} P_n - 1,$$

where  $n$  stands for the sample size and  $P_n$  denotes the number of concordant pairs, e.g. such pairs  $(X_1, X_2)$  and  $(X'_1, X'_2)$  that  $(X_1 - X'_1)(X_2 - X'_2) > 0$ . Next we provide the definition and similar results for the Spearman's  $\rho$ .

**Definition 3.** Let  $F$  be a continuous bivariate distribution function with the copula  $C$  and the univariate margins  $F_1$  and  $F_2$  respectively. Assume that  $(X_1, X_2) \sim F$ . Then the Spearman's  $\rho$  is given by

$$\rho_2 = 12 \int_{\mathbb{R}^2} F_1(x_1) F_2(x_2) dF(x_1, x_2) = 12 \int_{[0,1]^2} u_1 u_2 dC(u_1, u_2) - 3.$$

Similarly as for Kendall's  $\tau$ , the relationship between Spearman's  $\rho$  and specific copulae is given through

$$\begin{aligned}\rho_2(\text{Gaussian and } t) &= \frac{6}{\pi} \arcsin \frac{\delta}{2}, \\ \rho_2(\Pi) &= 0, & \rho_2(W) &= 1, & \rho_2(M) &= -1.\end{aligned}$$

Unfortunately, there is no explicit representation of Spearman's  $\rho_2$  for Archimedean in terms of generator functions as by Kendall's  $\tau$ . The estimator of  $\rho$  is easily computed using

$$\rho_{2n} = \frac{12}{n(n+1)(n-1)} \sum_{i=1}^n R_i S_i - 3 \frac{n+1}{n-1},$$

where  $R_i$  and  $S_i$  denote the ranks of two samples. The exact regions determined by Kendall's  $\tau$  and Spearman's  $\rho$  has been recently given by Schreyer, Paulin and Trutschnig (2017).

### 3 Multivariate Copula: Primer and State-of-Art

As mentioned in the introduction, step from bivariate copulas to multivariate is large. Nevertheless, many works have been written properly different high-dimensional copulas. This section introduces simple multivariate models and most prominent families like hierarchical Archimedean copula (HAC), pair-copula construction and factor copula.

A  $d$ -dimensional copula is also the distribution function on  $[0, 1]^d$  having all marginal distributions uniform on  $[0, 1]$ . In Sklar's Theorem, the importance of copulas in the area of multivariate distributions is re-stated in an exquisite way.

**Theorem 2.** *Let  $F$  be a multivariate distribution function with margins  $F_1, \dots, F_d$ , then there exists the copula  $C$  such that*

$$F(x_1, \dots, x_d) = C\{F_1(x_1), \dots, F_d(x_d)\}, \quad x_1, \dots, x_d \in \overline{\mathbb{R}}.$$

*If  $F_i$  are continuous for  $i = 1, \dots, d$  then  $C$  is unique. Otherwise  $C$  is uniquely determined on  $F_1(\overline{\mathbb{R}}) \times \dots \times F_d(\overline{\mathbb{R}})$ .*

*Conversely, if  $C$  is a copula and  $F_1, \dots, F_d$  are univariate distribution functions, then function  $F$  defined above is a multivariate distribution function with margins  $F_1, \dots, F_d$ .*

As in the bivariate case, the representation in Sklar's Theorem can be used for constructing new multivariate distributions by changing either the copula function or marginal distributions. For an arbitrary continuous multivariate distribution we can determine its copula from the transformation

$$C(u_1, \dots, u_d) = F\{F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)\}, \quad u_1, \dots, u_d \in [0, 1], \quad (4)$$

where  $F_i^{-1}$  are inverse marginal distribution functions. Copula density and density of the multivariate distribution with respect to copula are

$$\begin{aligned} c(u_1, \dots, u_d) &= \frac{\partial^k C(u_1, \dots, u_d)}{\partial u_1 \dots \partial u_d}, \quad u_1, \dots, u_d \in [0, 1], \\ f(x_1, \dots, x_d) &= c\{F_1(x_1), \dots, F_d(x_d)\} \prod_{i=1}^d f_i(x_i), \quad x_1, \dots, x_d \in \overline{\mathbb{R}}. \end{aligned}$$

For the multivariate case as well as for the bivariate case copula functions are invariant under monotone transformations.

### 3.1 Extensions of Simple and Elliptical Bivariate Copulae

The independence copula and the upper and lower Fréchet-Hoeffding bounds can be straightforwardly generalized to the multivariate case. The independence copula is defined by the product  $\Pi(u_1, \dots, u_d) = \prod_{i=1}^d u_i$  and the bounds are given by

$$W(u_1, \dots, u_d) = \max \left( 0, \sum_{i=1}^d u_i + 1 - d \right),$$

$$M(u_1, \dots, u_d) = \min(u_1, \dots, u_d), \quad u_1, \dots, u_d \in [0, 1].$$

An arbitrary copula  $C(u_1, \dots, u_d)$  lies between the Fréchet-Hoeffdings bounds

$$W(u_1, \dots, u_d) \leq C(u_1, \dots, u_d) \leq M(u_1, \dots, u_d),$$

where the Fréchet-Hoeffding lower bound is not a copula function for  $d > 2$  though. The generalization of elliptical copulas to  $d > 2$  is straightforward as well. For example, the Gaussian case yields

$$C_N(u_1, \dots, u_d, \Sigma) = \Phi_\Sigma \{ \Phi^{-1}(u_1), \dots, \Phi^{-1}(u_d) \},$$

$$c_N(u_1, \dots, u_d, \Sigma) = |\Sigma|^{-1/2}$$

$$\exp \left[ -\frac{1}{2} \{ \Phi^{-1}(u_1), \dots, \Phi^{-1}(u_d) \}^\top (\Sigma^{-1} - I) \{ \Phi^{-1}(u_1), \dots, \Phi^{-1}(u_d) \} \right]$$

for all  $u_1, \dots, u_d \in [0, 1]$ , where  $\Phi_\Sigma$  is a  $d$ -dimensional Gaussian distribution with zero mean and correlation matrix  $\Sigma$ . Individual dispersion is imposed via the marginal distributions. Note that in the multivariate case the implementation of elliptical copulas can be involved due to technical difficulties with multivariate cdf's.

### 3.2 Hierarchical Archimedean Copula

A simple multivariate generalization of the Archimedean copulas is defined as

$$C(u_1, \dots, u_d) = \phi \{ \phi^{-1}(u_1) + \dots + \phi^{-1}(u_d) \}, \quad u_1, \dots, u_d \in [0, 1], \quad (5)$$

where  $\phi \in \mathcal{L}$ . This definition provides a simple, but rather limited technique for the construction of multivariate copulas, since a possibly complicated multivariate dependence structure is determined by a single copula parameter. Furthermore, multivariate Archimedean copulas imply that the variables are exchangeable. This means, that the distribution of  $(u_1, \dots, u_d)$  is the same as of  $(u_{j_1}, \dots, u_{j_d})$  for all  $j_\ell \neq j_v$ . This is certainly not an acceptable assumption in practical applications.

A more flexible method is provided by hierarchical Archimedean copulas (HAC) sometimes also called nested Archimedean copula which replace a uniform margin of a simple Archimedean copula by an additional Archimedean copula. The iterative substitution of margins by copulas widens the spectrum of attainable dependence structures. For example, the copula function for fully nested HAC is given by

$$\begin{aligned} C(u_1, \dots, u_d) &= \phi_{d-1}^{-1} \circ \phi_{d-2} \left( \dots [\phi_2^{-1} \circ \phi_1 \{ \phi_1^{-1}(u_1) + \phi_1^{-1}(u_2) \} \right. \\ &\quad \left. + \phi_2^{-1}(u_3)] + \dots + \phi_{d-2}^{-1}(u_{d-1}) \right) + \phi_{d-1}^{-1}(u_d) \} \\ &= \phi_{d-1}^{-1} [C(\{\phi_1, \dots, \phi_{d-2}\})(u_1, \dots, u_{d-1}) + \phi_{d-1}^{-1}(u_d)] \end{aligned} \quad (6)$$

for  $\phi_{d-i}^{-1} \circ \phi_{d-j} \in \mathcal{L}^*$ ,  $i < j$ , where

$$\mathcal{L}^* = \{ \omega : [0; \infty) \rightarrow [0, \infty) \mid \omega(0) = 0, \omega(\infty) = \infty; (-1)^{j-1} \omega^{(j)} \geq 0; j = 1, \dots, \infty \},$$

As indicated above, contrarily to the usual Archimedean copula (5), HAC define the dependency structure in a recursive way. At the lowest level of the so called HAC-tree, the dependency between the two variables is modeled by a copula function with the generator  $\phi_1$ , i.e.  $z_1 = C(u_1, u_2) = \phi_1 \{ \phi_1^{-1}(u_1) + \phi_1^{-1}(u_2) \}$ . At the second level, another copula function is used to model the dependency between  $z_1$  and  $u_3$ , etc. The generators  $\phi_i$  can come from the same family and differ only through the parameter or, to introduce more flexibility, come from different generator families, c.f. Hofert (2011). As an alternative to the fully nested model, so-called partially nested copulas combine arbitrarily many copula functions at each copula level. For example the following 4-dimensional copula, where the first and the last two variables are joined by individual copulas with generators  $\phi_{12}$  and  $\phi_{34}$ . Further, the resulted copulas are combined by a copula with the generator  $\phi$ .

$$C(u_1, u_2, u_3, u_4) = \phi \left( \phi^{-1} [\phi_{12} \{ \phi_{12}^{-1}(u_1) + \phi_{12}^{-1}(u_2) \}] + \phi^{-1} [\phi_{34} \{ \phi_{34}^{-1}(u_3) + \phi_{34}^{-1}(u_4) \}] \right).$$

The estimation of HAC is a challenging task, since both the copula structure and parameters of the generator functions have to be estimated. The variety of possible structures does not permit the enumeration of all possible structures and selecting that structure-parameter combination with the largest log-likelihood value.

Okhrin et al. (2013a) first propose methods for determining the optimal structure of HAC with (non-)parametrically estimated margins and provide asymptotic theory for the estimated parameters. The basic idea of the estimation procedure uses the fact that HAC are recursively defined and that dependencies decrease from the lowest to the highest hierarchical level for common parametric families. To sketch the procedure suppose margins are known: Parameters related to strongly dependent random variables are estimated first

and the variables grouped at the bottom of the HAC-tree. The determined HAC-tree is spanned by at least two random variables and the tree itself determines a univariate random variable. After removing all random variables spanning the tree from the set of variables and adding the univariate random variable determined by the tree, the parameter of the subsequent level is determined by selecting that pair of variables with the strongest dependency again. An additional level is added to the tree referring to the pair of variables with the strongest dependence and the set of variables is modified as explained above. The sketched steps are iteratively repeated until the HAC-tree is spanned by all random variables. This method is implemented in the `HAC` package for R, see Okhrin and Ristig (2014).

Segers and Uyttendaele (2014) introduce an algorithm for non-parametric structure determination by firstly decomposing the HAC's tree structure into four variants of trivariate structures. Then, the whole tree structure is subsequently determined based on testing the distance between trivariate copulas and Kendall's distribution function. Górecki, Hofert and Holeňa (2016) generalize the approach of Okhrin et al. (2013a) and propose an algorithm for simultaneous estimation of the structure and parameters based on the inversion of Kendall's  $\tau_2$ , i.e. based on the link between Kendall's  $\tau_2$  and Archimedean generators.

Properties and simulation procedures are comprehensively studied in Joe (1997), Whelan (2004), Savu and Trede (2010), Hofert (2011), Okhrin, Okhrin and Schmid (2013b), Rezapour (2015) and Górecki et al. (2016). Note that HAC became a standard tool for pricing credit derivatives in academia such as collateralized debt obligations, see Hering, Hofert, Mai and Scherer (2010), Hofert and Scherer (2011) and Choroś-Tomczyk, Härdle and Okhrin (2013).

Brechmann (2014) proposed hierarchical Kendall copula, which does not suffer from parameter restriction, but are slightly more complicated in estimation. Similar approach to avoid parameter restrictions and family limitations are proposed by using Lévy subordinated HAC, see Hering et al. (2010) and the corresponding application see Zhu, Wang and Tan (2016).

### 3.3 Factor Copula

In classical factor analysis, a function links the observed and latent variables under the assumption that the latent variables explain the observed variables, e.g., see Johnson

and Wichern (2013) and Härdle and Simar (2015). For example, a random variable  $X_i$ ,  $i = 1, \dots, d$ , is generated by an additive factor model, if

$$X_i = \sum_{j=1}^m \alpha_{ij} W_j + \varepsilon_i, \quad (7)$$

where  $W_j$ ,  $j = 1, \dots, m$ , are latent common factors and  $\varepsilon_i$ ,  $i = 1, \dots, d$ , are mutually independent idiosyncratic disturbances. The basic idea of factor models and their natural interpretation can be exported to the copula world in order to induce dependencies between independent idiosyncratic disturbances via common factors. Factor copula models, however, can be split in two complementary groups both having strength and weaknesses. On the one hand, there are (implicit) factor copula models inducing dependencies among random variables via a functional which links latent factors and idiosyncratic disturbances. Such models are a straightforward extension of factor models from multivariate analysis. On the other hand, factor copulas and dependencies also arise from integrating the product of conditionally independent distributions –given a latent factor– with respect to this factor. This approach benefits from the fact, that the copula collapses to the product copula in case of known factors.

Oh and Patton (2015) concentrate on (implicit) factor copulas for  $X = (X_1, \dots, X_d)^\top$  arising from a functional relation between the factor(s) and mutual independent idiosyncratic errors. In this sense, the dependence component of the joint distribution of  $X$  is implied from the factors' distribution, the distribution of the idiosyncratic disturbances and the link function. In particular,  $X$  follows a multivariate distribution specified via a copula, i.e.  $X \sim F(x_1, \dots, x_d) = C\{F_1(x_1), \dots, F_d(x_d)\}$ . For instance, the additive single factor copula model is represented as

$$\begin{aligned} X_i &= W + \varepsilon_i, \quad i = 1, \dots, d, \\ W &\sim F_W(\theta_W), \quad \varepsilon_i \stackrel{i.i.d.}{\sim} F_\varepsilon(\theta_\varepsilon), \quad W \perp \varepsilon_i, \quad \text{for all } i = 1, \dots, d, \end{aligned} \quad (8)$$

where  $W$  is the single common factor following the distribution of  $F_W(\theta_W)$  and  $\varepsilon_1, \dots, \varepsilon_d$  are mutually independent shock with distribution function  $F_\varepsilon(\theta_\varepsilon)$ . This model is extended to the non-linear factor copula based on the following representation,

$$\begin{aligned} Z_i &= h(W, \varepsilon_i), \quad i = 1, 2, \dots, d, \\ W &\sim F_W(\theta_W), \quad \varepsilon_i \stackrel{i.i.d.}{\sim} F_\varepsilon(\theta_\varepsilon), \quad W \perp \varepsilon_i, \quad \text{for all } i = 1, \dots, d, \end{aligned} \quad (9)$$

where  $h$  is a non necessarily linear link function. Thus, the dependence structure can be built in a more flexible way compared to the linear additive version. Model (8) implies

a joint Gaussian random vector  $X = (X_1, \dots, X_d)^\top$ , if the common factor and the idiosyncratic factor are both Gaussian. Therefore, a joint density function is available as well.

Nonetheless, a nice analytical expression of the joint density function for a factor copula with non-Gaussian margins and non-Gaussian factor is rarely available which makes parameter estimation demanding. Oh and Patton (2013) propose an estimation method for copula models without analytical form of the density function. This relies on a simulated method of moments approach building on the simplicity to draw random samples from a factor model. The proposed estimator for  $(\theta_W^\top, \theta_\varepsilon^\top)^\top$  is found numerically by minimizing the distance between scale free empirical dependence measures between  $X_k$  and  $X_\ell$ , such as  $\tau_{2n}^{k\ell}$ ,  $k = 1, \dots, d; \ell = k + 1, \dots, d$ , and those obtained from a drawn sample. Oh and Patton (2013) prove under weak regularity conditions that the simulated method of moment estimator is consistent and asymptotically normal. However, as argued by Genest, Ghoudi and Rivest (1995), method of moment estimators of copula parameters can be highly inefficient.

Another form of factor copulae relies on the assumption that the observed variables  $U_1, \dots, U_d$  are conditionally independent given latent factors  $V_1, \dots, V_m$ . Note that all random variables  $U_i$ ,  $i = 1, \dots, d$ , and  $V_j$ ,  $j = 1, \dots, m$ , are assumed to uniformly distributed. Then, the conditional distribution of  $U_i$  given  $m$  factors  $V_1, \dots, V_m$  is given by  $C_{U_i|V_1, \dots, V_m}$ . By using  $C_{U_i|V_1, \dots, V_m}$ , the dependence structure of the observed variables  $U_1, \dots, U_d$  can be specified by the following copula function, such that

$$C(u_1, \dots, u_d) = \int_{[0,1]^m} \prod_{i=1}^d C_{U_i|V_1, \dots, V_m}(u_i|v_1, \dots, v_m) dv_1 \cdots dv_m \quad \text{with } u_i \in (0, 1), \quad (10)$$

where the factors are out integrated. For the special case  $m = 1$ , the copula function (10) can be simplified to the form

$$C(u_1, \dots, u_d) = \int_{[0,1]} \prod_{i=1}^d C_{U_i|V_1}(u_i|v_1) dv_1 \quad \text{with } u_i \in (0, 1). \quad (11)$$

Let  $C_{U_i, V_1}$  and  $c_{U_i, V_1}$  be the joint cdf and density of the pairs of random variables  $(U_i, V_1)$ ,  $i = 1, \dots, d$ . Moreover, let the conditional distribution of  $U_i$  given  $V_1$  be denoted by  $C_{U_i|V_1}(u_i|v_1) = \partial C_{U_i, V_1}(u_i, v)/\partial v|_{v=v_1}$ . Then, the copula density of  $C(u_1, \dots, u_d)$  can be represented by

$$c(u_1, \dots, u_d) = \frac{\partial^d C(u_1, \dots, u_d)}{\partial u_1 \cdots \partial u_d} = \int_{[0,1]} \prod_{i=1}^d c_{U_i, V_1}(u_i, v_1) dv_1 \quad \text{with } u_i \in (0, 1), \quad (12)$$

where  $c_{U_i, V_1}(u_i, v_1) = \partial C(u_i | v_1) / \partial u_i$ . Seen from this angle, the dependencies between  $d$  observed variables is determined by  $d$  bivariate copulas  $C_{U_i, V_1}(u_i, v)$ . Based on a parametric copula density  $c(\cdot; \theta)$ , Krupskii and Joe (2013) separate the parameter estimation into two steps. In the first step, the margins are estimated parametrically or non-parametrically. In the second step, the maximum likelihood (ML) method is employed to estimate the parameter  $\theta$ .

Numerous literature about the factor copula's theory and applications can be referred to. Andersen, Sidenius and Basu (2003), Hull and White (2004) and Laurent and Gregory (2005) have contributed works on generalization of one factor copula models. A comprehensive review of the factor copula theory is given in Joe (2014). Some applications by using factor copula models can be referred to Li (2000) for credit derivative pricing, Krupskii and Joe (2013) for fitting stock returns and Oh and Patton (2015) for measuring systemic risk.

### 3.4 Vine Copula

Vine copula or pair-copula constructions are originally proposed in Joe (1996) and developed in depth by Bedford and Cooke (2001), Bedford and Cooke (2002), Kurowicka and Cooke (2006) and Aas, Czado, Frigessi and Bakken (2009). The catchy name is due to similarities of the graphical representation of vine copulae and botanical vines. The fundamental idea of the vine copula is to construct a  $d$ -dimensional copula by decomposing the dependence structure into  $d(d - 1)/2$  bivariate copulas.

Let  $S$  be the index subset of  $D = \{1, \dots, d\}$  referring to the index set of conditioning variables and  $T$  be the index set of conditioned variables with  $T \cup S = D$ . Let  $\#M$  denote the cardinality of set  $M$ . The cdf of variables with index in  $S$  is denoted by  $F_S$ , so that  $F(x) = F_D(x)$ . The conditional cdf of variables with index in  $T$  conditional on  $S$  is denoted  $F_{T|S}$ . A similar notation is used for the corresponding copulas. To derive a vine copula for a given  $x = (x_1, \dots, x_d)^\top$  in the spirit of Joe (2014), we start from a  $d$ -dimensional distribution function, i.e.

$$F(x) = \int_{(-\infty, x_S]} F_{T|S}(x_T | y_S) dF_S(y_S), \quad (13)$$

and replace the conditional distribution  $F_{T|S}(x_T | x_S)$  by the corresponding  $\#T$ -dimensional copula  $F_{T|S}(x_T | x_S) = C_{T;S}\{F_{j|S}(x_j | x_S) : j \in T\}$ . The copula  $C_{T;S}\{F_{j|S}(x_j | x_S) : j \in T\}$  is implied by Sklar's Theorem with margins  $F_{j|S}(x_j | x_S)$ ,  $j \in T$ . It is not a conditional

distribution although with conditional distribution as margins. This yields a copula-based representation of the joint  $d$ -dimensional distribution function from (13), which is given by

$$F(x) = \int_{(-\infty, x_S]} C_{T;S} \{ F_{j|S}(x_j | y_S) : j \in T \} dF_S(y_S). \quad (14)$$

Note that the support of the integral in (13) and (14) is a cube  $(-\infty, x_S] \in \mathbb{R}^{\#S}$ . Converting all univariate margins to uniform distributed random variables allows rewriting  $F(x)$  as a  $d$ -dimensional copula

$$C(u) = \int_{[0, u_S]} C_{T;S} \{ G_{j|S}(u_j | v_S) : j \in T \} dC_S(v_S), \quad (15)$$

where  $G_{j|S}(u_j | v_S)$  is a conditional distribution from copula  $C_{S \cup \{j\}}$ . If  $T = \{i_1, i_2\}$ , then

$$C_{S \cup \{i_1, i_2\}}(u_{S \cup \{i_1, i_2\}}) = \int_{[0, u_S]} C_{i_1, i_2; S} \{ G_{i_1|S}(u_{i_1} | v_S), G_{i_2|S}(u_{i_2} | v_S) \} dC_S(v_S). \quad (16)$$

Since the essential idea of vine copula is based on building a joint dependence structure by  $d(d - 1)/2$  bivariate copulae, (16) is an important building block in the construction of vines referring to a  $(\#S + 2)$ -dimensional copula built from a bivariate copula  $C_{i_1, i_2; S}$ .

In case of continuous random variables, the  $d$ -dimensional distribution function from (13) admits a density function  $f(x_1, \dots, x_d)$ , which can be decomposed and represented by bivariate copula densities in an analogue manner. Examples of density decompositions for the 6-dimensional case related to so called C-vine (canonical vine), D-vine (drawable vine) and R-vine (regular vine) copulas are given as follows. The C-vine structure is illustrated in the left column of Figure 2 and its density decomposition is

$$\begin{aligned} c\{F_1(x_1), \dots, F_6(x_6)\} &= c_{12}\{F_1(x_1), F_2(x_2)\} \cdot c_{13}\{F_1(x_1), F_3(x_3)\} \\ &\quad \cdot c_{14}\{F_1(x_1), F_4(x_4)\} \cdot c_{15}\{F_1(x_1), F_5(x_5)\} \cdot c_{16}\{F_1(x_1), F_6(x_6)\} \\ &\quad \cdot c_{23;1}\{F(x_2|x_1), F(x_3|x_1)\} \cdot c_{24;1}\{F(x_2|x_1), F(x_4|x_1)\} \\ &\quad \cdot c_{25;1}\{F(x_2|x_1), F(x_5|x_1)\} \cdot c_{26;1}\{F(x_2|x_1), F(x_6|x_1)\} \\ &\quad \cdot c_{34;12}\{F(x_3|x_{12}), F(x_4|x_{12})\} \cdot c_{35;12}\{F(x_3|x_{12}), F(x_5|x_{12})\} \\ &\quad \cdot c_{36;12}\{F(x_3|x_{12}), F(x_6|x_{12})\} \cdot c_{45;123}\{F(x_4|x_{123}), F(x_5|x_{123})\} \\ &\quad \cdot c_{46;123}\{F(x_4|x_{123}), F(x_6|x_{123})\} \cdot c_{56;1234}\{F(x_5|x_{1234}), F(x_6|x_{1234})\}. \end{aligned} \quad (17)$$

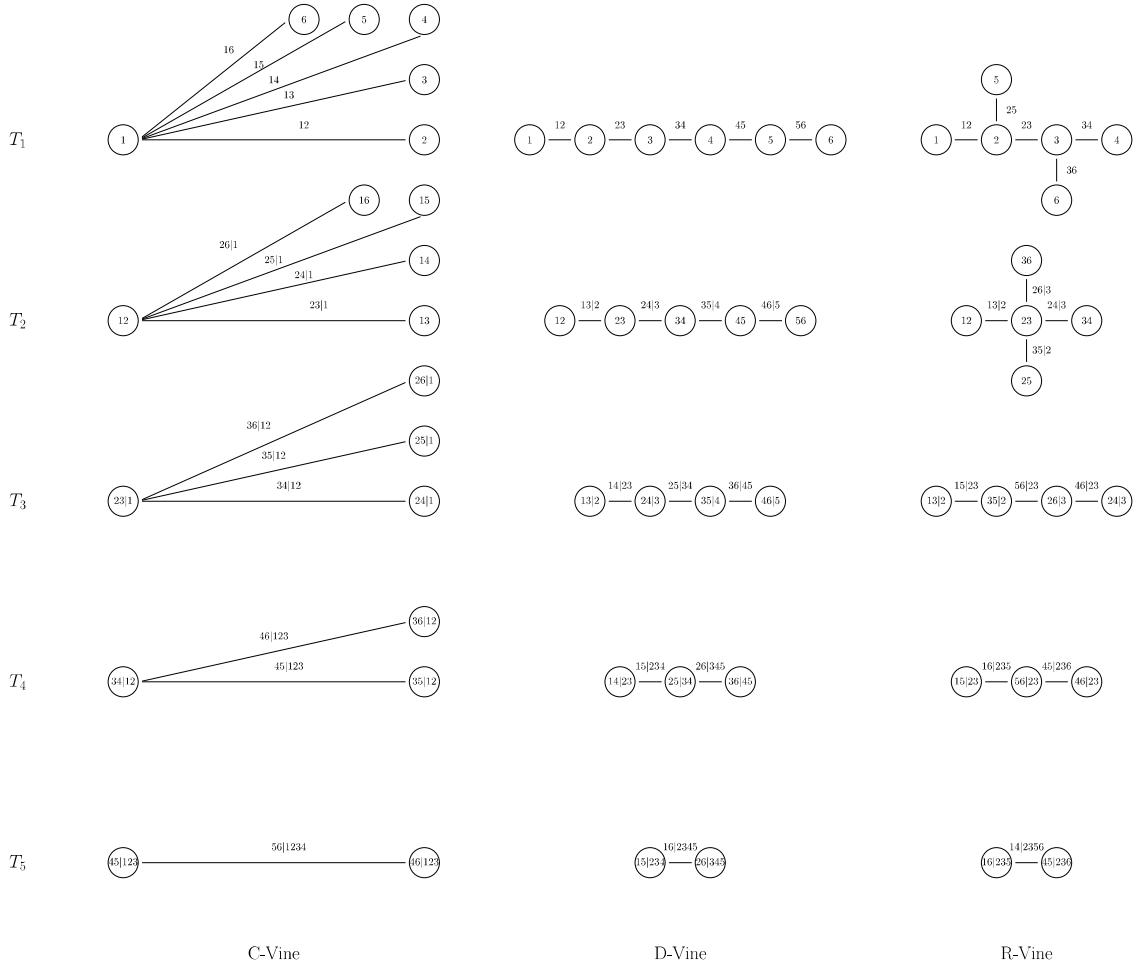


Figure 2: Vine tree structures of C-vine, D-vine and R-vine.

The density of the D-vine structure –given in the centred column of Figure 2– is

$$\begin{aligned}
c\{F_1(x_1), \dots, F_6(x_6)\} &= c_{12}\{F_1(x_1), F_2(x_2)\} \cdot c_{23}\{F_2(x_2), F_3(x_3)\} \\
&\quad \cdot c_{34}\{F_3(x_3), F_4(x_4)\} \cdot c_{45}\{F_4(x_4), F_5(x_5)\} \cdot c_{56}\{F_5(x_5), F_6(x_6)\} \\
&\quad \cdot c_{13;2}\{F(x_1|x_3), F(x_2|x_3)\} \cdot c_{24;3}\{F(x_2|x_3), F(x_4|x_3)\} \\
&\quad \cdot c_{35;4}\{F(x_3|x_4), F(x_5|x_4)\} \cdot c_{46;5}\{F(x_4|x_5), F(x_6|x_5)\} \\
&\quad \cdot c_{14;23}\{F(x_2|x_{23}), F(x_4|x_{23})\} \cdot c_{25;34}\{F(x_2|x_{34}), F(x_5|x_{34})\} \\
&\quad \cdot c_{36;45}\{F(x_3|x_{45}), F(x_6|x_{45})\} \cdot c_{15;234}\{F(x_1|x_{234}), F(x_5|x_{234})\} \\
&\quad \cdot c_{26;345}\{F(x_2|x_{345}), F(x_6|x_{345})\} \cdot c_{16;2345}\{F(x_1|x_{2345}), F(x_6|x_{2345})\}.
\end{aligned} \tag{18}$$

The density of the R-vine structure illustrated in the right column of Figure 2 is

$$\begin{aligned}
c\{F_1(x_1), \dots, F_6(x_6)\} &= c_{12}\{F_1(x_1), F_2(x_2)\} \cdot c_{23}\{F_2(x_2), F_3(x_3)\} \\
&\quad \cdot c_{34}\{F_3(x_3), F_4(x_4)\} \cdot c_{25}\{F_2(x_2), F_5(x_5)\} \cdot c_{36}\{F_3(x_3), F_6(x_6)\} \\
&\quad \cdot c_{13;2}\{F(x_1|x_2), F(x_3|x_2)\} \cdot c_{24;3}\{F(x_2|x_3), F(x_4|x_3)\} \\
&\quad \cdot c_{26;3}\{F(x_2|x_3), F(x_6|x_3)\} \cdot c_{35;2}\{F(x_3|x_2), F(x_5|x_2)\} \\
&\quad \cdot c_{15;23}\{F(x_1|x_{23}), F(x_5|x_{23})\} \cdot c_{56;23}\{F(x_5|x_{23}), F(x_6|x_{23})\} \\
&\quad \cdot c_{46;23}\{F(x_4|x_{23}), F(x_6|x_{23})\} \cdot c_{16;235}\{F(x_1|x_{235}), F(x_6|x_{235})\} \\
&\quad \cdot c_{45;236}\{F(x_4|x_{236}), F(x_5|x_{236})\} \cdot c_{14;2356}\{F(x_1|x_{2356}), F(x_4|x_{2356})\}.
\end{aligned} \tag{19}$$

In particular, the C-vine and D-vine have an intuitive graphical representation which can be immediately related to the decomposition of the copula density function into the product of bivariate copula densities. For example, the product of bivariate copula densities from the first two lines of the right hand side of Equation 17 refers to a C-vine represented in the upper left graphic of Figure 2. The formula and the corresponding graphic illustrate that the first variable  $X_1$  is pairwise coupled with the second, third ... and sixth random variable. The subsequent two lines (3-4) of Equation 17 are related to the second graphic of the left column of Figure 2. Conditional on  $X_1$ , random variable  $X_2$  is pairwise coupled with  $X_3, X_4, X_5$  and  $X_6$ . Connecting the remaining graphics with formulas is left to the reader. While the “formula-graphic” matching follows a similar scheme in case of the D-vine, the R-vine belongs to a more general vine copula class and contains the C-vine and D-vine as special cases. A rigorous definition of an R-vine copula can be found in Joe (2014).

In fact, vines can be estimated by either full or stage-wise ML such as the inference function for margins (IFM) method discussed below in Section 4. Nonetheless, the inference

approach derived in Haff (2013) namely the stepwise semi-parametric estimator deserves to be mentioned in more detail. Here, the marginal distributions are non-parametrically estimated by the empirical distribution function such as for factor copulae or HAC. In order to obtain a consistent and asymptotically Gaussian distributed estimator of a parametric vine copula, a so called simplifying assumption is required. The latter permits replacing “conditional” bivariate copula densities with unconditional densities. Then, it can be straightforwardly shown, that the log-likelihood can be maximized in a stage-wise manner. This is due to the decomposition of the density into the product of bivariate copula densities, so that the log-likelihood function is a sum of logarithmized copula densities. Coming back to the C-vine example from Figure 2. At the first stage, all parameters of bivariate copulas represented in the upper left graphic of Figure 2 are estimated, i.e. the parameters of the copulae for  $(X_1, X_2), \dots, (X_1, X_6)$ . Keeping the corresponding parameters fixed at estimated values, the four parameters of copulae referring to the pairs from the second graphic of the left column of Figure 2 are estimated. Holding these parameters fixed at estimated values again, all vine parameters of the remaining bivariate densities can be estimated iteratively. Literature on pair-copula construction is spreading steadily, and most recent information about it can be found on vine copula homepage <http://www.statistics.ma.tum.de/en/research/vine-copula-models/>.

## 4 Estimation Methods

The estimation of a copula-based multivariate distribution involves both the estimation of the copula parameters  $\boldsymbol{\theta}$  and the estimation of the margins  $F_j$ ,  $j = 1, \dots, d$ . The properties and goodness of the estimator of  $\boldsymbol{\theta}$  heavily depend on the estimators of  $F_j$ ,  $j = 1, \dots, d$ . We distinguish between a parametric and a non-parametric specification of the margins. If we are interested only in the dependency structure, the estimator of  $\boldsymbol{\theta}$  should be independent of any parametric models for the margins. However, Joe (1997) argues that complete distribution models and, therefore, parametric models for margins are actually more appropriate for applications.

In the bivariate case, a standard method of estimating the univariate parameter  $\theta$  is based on Kendall’s  $\tau_2$  statistic by Genest and Rivest (1993). The estimator of  $\tau_2$  complemented by the method of moments allows to estimate the parameters. However, as shown in Genest et al. (1995), the ML method leads to substantially more efficient estimators. For non-parametrically estimated margins, Genest et al. (1995) show the consistency

and asymptotic normality of ML estimators and derive the moments of the asymptotic distribution. The ML procedure can be performed simultaneously for the parameters of the margins and of the copula function. Alternatively, a two-stage procedure can be applied, where the parameters of margins are estimated at the first stage and the copula parameters at the second stage, see Joe (1997) and Joe (2005). Chen and Fan (2006) and Chen, Fan and Tsyrennikov (2006) analyze the case of non-parametrically estimated margins. Fermanian and Scaillet (2003) and Chen and Huang (2007) consider a fully non-parametric estimation of the copula. Next we provide details on both approaches. Note that estimation procedures for HAC, conditional-independence-based factor copulas and vines are in fact generalizations of the subsequent approaches taking specific needs of the copula into account, e.g., parameter restrictions.

## 4.1 Parametric Margins

Let  $\boldsymbol{\alpha} = (\boldsymbol{\alpha}_1^\top, \dots, \boldsymbol{\alpha}_d^\top)^\top$  denote the vector of parameters of marginal distributions and  $\boldsymbol{\theta}$  parameters of the copula. The classical full ML estimator  $\hat{\boldsymbol{\eta}}$  of  $\boldsymbol{\eta} = (\boldsymbol{\alpha}^\top, \boldsymbol{\theta}^\top)^\top$  solves the system of equations

$$\frac{\partial \mathcal{L}(\boldsymbol{\eta}, \mathbf{X})}{\partial \boldsymbol{\eta}} = \mathbf{0},$$

$$\begin{aligned} \text{where } \mathcal{L}(\boldsymbol{\eta}, \mathbf{X}) &= \sum_{i=1}^n \log \left\{ c(F_1(x_{1i}, \boldsymbol{\alpha}_1), \dots, F_d(x_{di}, \boldsymbol{\alpha}_d), \boldsymbol{\theta}) \prod_{j=1}^d f_j(x_{ji}, \boldsymbol{\alpha}_j) \right\} \\ &= \sum_{i=1}^n \left\{ \log c(F_1(x_{1i}, \boldsymbol{\alpha}_1), \dots, F_d(x_{di}, \boldsymbol{\alpha}_k), \boldsymbol{\theta}) + \sum_{j=1}^d \log f_j(x_{ji}, \boldsymbol{\alpha}_j) \right\}. \end{aligned}$$

Following the standard theory on ML estimation, the estimator  $\hat{\boldsymbol{\eta}}$  is efficient and asymptotically normal. However, it is often computationally demanding to solve the system simultaneously. Alternatively the multistage optimization proposed in Joe (1997, Chapter 10), also known as inference functions for margins, can be applied: Firstly, the parameters of the margins are separately estimated under the assumption that the copula is the product copula. Secondly, the parameters of the copula are estimated replacing the parameters of margins by estimates from the first step and treating them as known

quantities. The above optimization problem is then replaced by

$$\left( \frac{\partial \mathcal{L}_1}{\partial \boldsymbol{\alpha}_1^\top}, \dots, \frac{\partial \mathcal{L}_d}{\partial \boldsymbol{\alpha}_d^\top}, \frac{\partial \mathcal{L}_{d+1}}{\partial \boldsymbol{\theta}^\top} \right)^\top = \mathbf{0}, \quad (20)$$

$$\begin{aligned} \text{where} \quad & \mathcal{L}_j = \sum_{i=1}^n l_j(\mathbf{X}_i), \quad \text{for } j = 1, \dots, d+1, \\ & l_j(\mathbf{X}_i) = \log f_j(x_{ji}, \boldsymbol{\alpha}_j), \quad \text{for } j = 1, \dots, d, i = 1, \dots, n, \\ \text{and} \quad & l_{d+1}(\mathbf{X}_i) = \log c\{F_1(x_{1i}, \boldsymbol{\alpha}_1), \dots, F_d(x_{di}, \boldsymbol{\alpha}_d), \boldsymbol{\theta}\}, \quad \text{for } i = 1, \dots, n. \end{aligned}$$

The first  $d$  components in (20) correspond to the usual ML estimation of the parameters of the marginal distributions. The last component reflects the estimation of the copula parameters. Detailed discussion on this method can be found in Joe (1997, Chapter 10). Note, that this procedure does not lead to efficient estimators, however, as argued by Joe (1997) the loss in the efficiency is modest and mainly depends on the strength of dependencies. This method is a special case of the generalized method of moments with an identity weighting matrix, see Cherubini, Luciano and Vecchiato (2004, Section 4.5). The advantage of the two-stage procedure lies in the dramatic reduction of the numerical complexity.

## 4.2 Non-parametric Margins

In this section, we consider a non-parametric estimation of the marginal distributions also referred to as canonical ML. The asymptotic properties of the multistage estimator for  $\boldsymbol{\theta}$  do not depend explicitly on the type of the non-parametric estimator, but on its convergence properties. Here, we use the rectangular kernel (histogram) resulting in the estimator

$$\widehat{F}_j(x) = (n+1)^{-1} \sum_{i=1}^n \mathbf{1}(x_{ji} \leq x), \quad j = 1, \dots, d.$$

The factor  $n/(n+1)$  is used to restrict fitted values to the open unit interval. This is necessary as several copula densities are not bounded at zero and/or one. Let  $\widehat{F}_1, \dots, \widehat{F}_d$  denote the non-parametric estimators of  $F_1, \dots, F_d$ . The canonical ML estimator  $\hat{\boldsymbol{\theta}}$  of  $\boldsymbol{\theta}$  solves the system  $\partial \mathcal{L}/\partial \boldsymbol{\theta}^\top = \mathbf{0}$  by maximizing the pseudo log-likelihood with estimated margins  $\widehat{F}_1, \dots, \widehat{F}_d$ , i.e.

$$\begin{aligned} \mathcal{L} &= \sum_{i=1}^n l(\mathbf{X}_i) \quad \text{for } j = 1, \dots, p, \\ l(\mathbf{X}_i) &= \log c\{\widehat{F}_1(x_{1i}), \dots, \widehat{F}_d(x_{di}), \boldsymbol{\theta}\}, \quad \text{for } i = 1, \dots, n. \end{aligned}$$

As in the parametric case, the semi-parametric estimator  $\hat{\theta}$  is asymptotically normal under suitable regularity conditions. This method was first used in Oakes (1994) and then investigated by Genest et al. (1995) and Shih and Louis (1995). Additional properties of the estimator, such as the covariance matrix, are stated in these papers.

## 5 Goodness-of-Fit Tests for Copulae

Having a dataset and an estimated copula at hand, it arises the natural question whether the selected copula describes the data properly. For this purpose, a series of different goodness-of-fit tests has been developed in the last decade. Under the  $H_0$ -hypothesis one assumes that the true copula belongs to some parametric family  $H_0 : C \in C_0$ .

The most natural test approach is to measure the deviation of the parametric copula from the empirical one given through

$$C_n(u_1, \dots, u_d) = n^{-1} \sum_{i=1}^n \prod_{j=1}^d I\{\hat{F}_j(x_{ij}) \leq u_j\}.$$

Gaensler and Stute (1987) and Radulovic and Wegkamp (2004) show that  $C_n$  is a consistent estimation of the true underlying copula. Several tests are based on the empirical copula process, which is defined as follows

$$\mathbb{C}_n(u_1, \dots, u_d) = \sqrt{n}\{C_n(u_1, \dots, u_d) - C_{\hat{\theta}}(u_1, \dots, u_d)\}.$$

Fermanian (2005) and Genest and Rémillard (2008) propose to compute different measures to quantify the deviation of the assumed parametric copula from the empirical copula, one of those is Cramér-von Mises distance

$$S_n^E = \int_{[0,1]^d} \mathbb{C}_n(u_1, \dots, u_d)^2 dC_n(u_1, \dots, u_d)$$

or the weighted Cramér-von Mises distance, with tuning parameters  $m \geq 0$  and  $\zeta_m \geq 0$  given as

$$R_n^E = \int_{[0,1]^d} \left\{ \frac{\mathbb{C}_n(u_1, \dots, u_d)}{[C_{\hat{\theta}}(u_1, \dots, u_d)\{1 - C_{\hat{\theta}}(u_1, \dots, u_d)\} + \zeta_m]^m} \right\}^2 dC_n(u_1, \dots, u_d).$$

The usual Kolmogorov-Smirnov distance as for classical univariate tests is also applicable here

$$T_n^E = \sup_{\{u_1, \dots, u_d\} \in [0,1]^d} |\mathbb{C}_n(u_1, \dots, u_d)|.$$

The other group of tests developed and investigated by Genest and Rivest (1993), Wang and Wells (2000), Genest, Quessy and Rémillard (2006) are based on the probability integral transform and in particular on so called Kendall's transform. Having

$$(X_1, \dots, X_d) \sim F(x_1, \dots, x_d) = C_\theta\{F_1(x_1), \dots, F_d(x_d)\},$$

one concludes similar to  $F_i(X_i) \sim U(0, 1)$  that the copula-based random variable is

$$C_\theta\{F_1(X_1), \dots, F_d(X_d)\} \sim K_\theta(v)$$

where  $K_\theta(v)$  is the univariate Kendall's distribution (not necessarily uniform), see Barbe, Genest, Ghoudi and Rémillard (1996), Jouini and Clemen (1996). Empirically, the distribution function  $K$  can be estimated as

$$K_n(v) = n^{-1} \sum_{i=1}^n I\left[C_n\{\widehat{F}_1(x_{i1}), \dots, \widehat{F}_d(x_{id})\} \leq v\right], \quad v \in [0, 1].$$

Further usual test statistics for the univariate distributions like Cramér-von Mises or Kolmogorov-Smirnov, see Genest et al. (2006), can be applied

$$S_n^{(K)} = \int_0^1 \mathbb{K}_n(v)^2 dK_{\hat{\theta}}(v), \quad T_n^{(K)} = \sup_{v \in [0, 1]} |\mathbb{K}_n(v)|,$$

where  $\mathbb{K}_n = \sqrt{n}(K_n - K_{\hat{\theta}})$  is the Kendall's process. Here is, however, a little challenge in using this tests: as in testing for Kendall's distribution one tests in null hypothesis has  $H_0'': K \in \mathcal{K}_0 = \{K_\theta : \theta \in \Theta\}$ , and as  $H_0 \subset H_0''$ , the non-rejection of  $H_0''$  does not imply non rejection of  $H_0$ . For the bivariate Archimedean copulas  $H_0''$  and  $H_0$  are equivalent.

Another series of goodness-of-fit tests, is constructed via the other important integral transform, that dates back to Rosenblatt (1952). Based on the conditional distribution of  $U_i$  by

$$\begin{aligned} C_d(u_i | u_1, \dots, u_{i-1}) &= P\{U_i \leq u_i | U_1 = u_1 \dots U_{i-1} = u_{i-1}\} \\ &= \frac{\partial^{i-1} C(u_1, \dots, u_i, 1, \dots, 1) / \partial u_1 \dots \partial u_{i-1}}{\partial^{i-1} C(u_1, \dots, u_{i-1}, 1, \dots, 1) / \partial u_1 \dots \partial u_{i-1}}, \end{aligned}$$

the Rosenblatt transform is defined as follows.

**Definition 4.** *Rosenblatt's probability integral transform of a copula  $C$  is the mapping  $\mathfrak{R} : (0, 1)^d \rightarrow (0, 1)^d$ ,  $\mathfrak{R}(u_1, \dots, u_d) = (e_1, \dots, e_d)$  with  $e_1 = u_1$  and  $e_i = C_d(u_i | u_1, \dots, u_{i-1})$ ,  $\forall i = 2, \dots, d$ .*

Under this definition, the null hypothesis  $H_0 : C \in C_0$  can be rewritten as  $H_{0R} : (e_1, \dots, e_d)^\top \sim \Pi$ . The first test based on the Rosenblatt transform exploits information, that under  $H_0$  transformed observations should be exactly uniform distributed and independent, which is not the case, as those variables are not mutually independent and only approximately uniform. Nevertheless, two tests use Anderson-Darling test statistics, see Breymann, Dias and Embrechts (2003), and are constructed as

$$T_n = -n - \sum_{i=1}^n \frac{2i-1}{n} [\log G_{(i)} + \log\{1 - G_{(n+1-i)}\}]$$

where  $G_i$  might be constructed in two ways. In the first possibility

$$G_{i,\text{Gamma}} = \Gamma_d \left\{ \sum_{j=1}^d (-\log e_{ij}) \right\},$$

where  $\Gamma_d(\cdot)$  is the Gamma distribution with shape  $d$  and scale 1. The second way takes

$$G_{i,\chi^2} = \chi_d^2 \left[ \sum_{j=1}^d \{\Phi^{-1}(e_{ij})\}^2 \right],$$

where  $\chi_d^2$  refers to the Chi-squared distribution with  $d$  degrees of freedom and  $\Phi$  is standard normal distribution. Another possibility compares the variables not via the Anderson-Darling test statistics, but by purely deviations between estimated density functions, as in Patton, Fan and Chen (2004), where the test statistics is constructed by

$$C_n^{Ch} = \frac{n\sqrt{h}\hat{J}_n - c_n}{\sigma}$$

with  $c_n$  and  $\sigma$  are normalization factors and  $\hat{J}_n = \int_0^1 \{\frac{1}{n} \sum_{i=1}^n K_h(w, G_{i,\chi^2}) - 1\}^2 dw$ .

As discussed by Dobrić and Schmid (2007), the problem with those tests is that they have almost no power and even do not capture the type 1 error. Much better power have tests, that work directly on the copulas of the Rosenblatt transformed data, see Genest, Rémillard and Beaudoin (2009). The idea is to compute Cramer-von Mises statistics of the following form

$$\begin{aligned} S_n &= n \int_{[0,1]^d} \{D_n(u) - \Pi(u)\}^2 du \\ S_n^{(C)} &= n \int_{[0,1]^d} \{D_n(u) - \Pi(u)\}^2 dD_n(u) \end{aligned}$$

where the empirical distribution function

$$D_n(u) = D_n(u_1, \dots, u_d) = \frac{1}{n} \sum_{i=1}^n \prod_{j=1}^d I(e_{ij} \leq u_j)$$

should be “close” to product copula  $\Pi$  under  $H_0$ .

Different from previous test are those based on the kernel density estimators, and just to mention one, let us consider test developed by Scaillet (2007), where the test statistics is given through

$$J_n = \int_{[0,1]^d} \{\hat{c}(u) - K_H * c(u; \hat{\theta})\} w(u) du,$$

with “ $*$ ” being a convolution operator and  $w(u)$  a weight function. The kernel function  $K_H(y) = K(H^{-1}y) / \det(H)$  where  $K$  is the bivariate quadratic kernel with the bandwidth  $H = 2.6073n^{-1/6}\widehat{\Sigma}^{1/2}$  and  $\widehat{\Sigma}$  being a sample covariance matrix. The copula density is estimated non-parametrically as

$$\hat{c}(u) = n^{-1} \sum_{i=1}^n K_H[u - \{\widehat{F}_1(X_{i1}), \dots, \widehat{F}_d(X_{id})\}^\top],$$

where  $\widehat{F}_j$  refers to an estimated marginal distribution,  $j = 1, \dots, d$ . The most recent goodness of fit test for copulas have been proposed recently by Zhang, Okhrin, Zhou and Song (2016), where one compares the two-step pseudo maximum likelihood:

$$\hat{\theta} = \operatorname{argmax}_{\theta \in \Theta} \sum_{i=1}^n \mathcal{L}\{\widehat{F}_1(X_{i1}), \dots, \widehat{F}_d(X_{id}); \theta\}.$$

with the delete-one-block pseudo maximum likelihood  $\hat{\theta}_{-b}$ ,  $1 \leq b \leq B$ :

$$\hat{\theta}_{-b} = \operatorname{argmax}_{\theta \in \Theta} \sum_{b' \neq b}^B \sum_{i=1}^m \mathcal{L}\{\widehat{F}_1(X_{i1}), \dots, \widehat{F}_d(X_{id}); \theta\}, \quad b = 1, \dots, B.$$

Further, ”in-sample” and ”out-of-sample” pseudo-likelihoods are compared with the following test statistic:

$$T_n(m) = \sum_{b=1}^B \sum_{i=1}^m \left[ \mathcal{L}\{\widehat{F}_1(X_{i1}), \dots, \widehat{F}_d(X_{id}); \hat{\theta}\} - \mathcal{L}\{\widehat{F}_1(X_{i1}), \dots, \widehat{F}_d(X_{id}); \hat{\theta}_{-b}\} \right].$$

This leads to some challenges, like computation of  $\lceil \frac{n}{m} \rceil$  dependence parameters, but Zhang et al. (2016) proposed an asymptotically equivalent test statistics based on variability and sensitivity matrices. As most of the above mentioned tests, have complicated asymptotic distributions,  $p$ -values of the tests can be performed via the parametric bootstrap sketched in the subsequent procedure:

Step 1 Generate bootstrap sample  $\{\epsilon_i^{(k)}, i = 1, \dots, n\}$  from copula  $C(u; \hat{\theta})$  under  $H_0$  with  $\hat{\theta}$  and estimated marginal distribution  $\widehat{F}$  obtained from original data;

Step 2 Based on  $\{\epsilon_i^{(k)}, i = 1, \dots, n\}$  from Step 1, estimate  $\theta$  of the copula under  $H_0$ , and compute test statistics under consideration, say  $R_n^k$ ;

Step 3 Repeat Steps (1 – 2)  $N$ -times and obtain  $N$  statistics  $R_n^k, k = 1, \dots, N$ ;

Step 4 Compute an empirical  $p$ -value as  $p_e = N^{-1} \sum_{k=1}^N I(|R_n^k| \geq |R_n|)$  with  $R_n$  being the test statistic estimated from original data.

## 6 Empirical Study

Value-at-Risk (VaR) is an important measure in risk management. The traditional models for VaR estimation assume that the assets returns in a portfolio are jointly normally distributed. However, numerous empirical studies show that Gaussian based models are not sufficient to describe data characteristics, especially when extreme events happen such as financial crisis. The weak points of the Gaussian based models include the lack of asymmetry and tail dependence. Therefore copula methods come into the focus.

Twelve different copulas are used in this study to construct dependence structures. The employed families include the Gaussian copula,  $t$ -copula, Archimedean copulas (Clayton, Gumbel, Joe), HAC (Gumbel, Clayton, Frank), C- and D-vine structures and two factor copulas linked individually by a bivariate Gumbel and Clayton copula.

The data set utilized in this study includes five time series of stock close prices containing ADI (Analog Devices, Inc.), AVB (Avalonbay Communities Inc.), EQR (Equity Residential), LLY (Eli Lilly and Company) and TXN (Texas Instruments Inc.), from Yahoo finance. Here, ADI and TXN belong to high-tech industry, AVB and EQR to real estate industry and LLY to pharmacy industry. The time window spans from 20070113 to 20160116.

Let  $w = (w_1, \dots, w_d)^\top \in \mathbb{R}^d$  denote the long position vector of a  $d$ -dimensional portfolio,  $S_t = (S_{1,t}, \dots, S_{d,t})^\top$  stand for the vector of asset prices at time  $t \in \{1, \dots, T\}$  and  $X_{i,t} = \log(S_{i,t}/S_{i,t-1})$  for the one period log-return of the  $i$ -th asset at time  $t$ . Then,  $L_t = \sum_{i=1}^d w_i X_{i,t}$  denotes the portfolio return. The distribution function of the univariate random variable  $L_t$  is denoted by  $F_{L_t}(x) = P(L_t \leq x)$  and the Value-at-Risk at level  $\alpha$  for the portfolio is defined as the inverse of  $F_{L_t}(x)$ , namely  $\text{VaR}_t(\alpha) = F_{L_t}^{-1}(\alpha)$ .

	AVB	EQR	TXN	ADI	AVB	EQR	TXN	ADI	AVB	EQR	TXN	ADI
EQR	0.867				0.686				0.866			
TXN	0.359	0.375			0.260	0.264			0.376	0.381		
ADI	0.384	0.399	0.752		0.277	0.285	0.583		0.398	0.410	0.770	
LLY	0.358	0.370	0.358	0.362	0.268	0.260	0.272	0.270	0.390	0.376	0.393	0.391

Table 1: Pairwise dependence measures including Pearson’s correlation (left), Kendall’s correlation (center) and Spearman’s correlation (right).

## Copula Performance in Risk Management

From the above formulations can be concluded that the idiosyncratic dependence of the log-return process  $\{X_t\}_{t=1}^T$  is crucial for the appropriate estimation of the VaR. To remove temporal dependence from  $X_t$ , the single log-return processes are filtered through GARCH(1, 1) processes,

$$X_{i,t} = \mu_{i,t} + \sigma_{i,t}\epsilon_{i,t}, \quad (21)$$

$$\sigma_{i,t}^2 = a_i + \alpha_i(X_{i,t-1} - \mu_{i,t-1})^2 + \beta_i\sigma_{i,t-1}^2. \quad (22)$$

The GARCH(1, 1)-filtered log-returns are illustrated in Figure 3. Obviously, assets coming from the same sector have high correlation according to the GARCH residuals. For example, the AVB-EQR and TXN-ADI pairs have strong correlation coming from real estate industry and high technology industry respectively. The strong correlation is also observed in Table 1 presenting three dependence measures for pairs of AVB-EQR and TXN-ADI. LLY is from pharmacy industry and shows weak correlation with the other four companies according to the scatter-plots and the contours.

The performance of different copulas utilized for VaR estimation is evaluated via back-testing based on the exceeding ratio

$$ER^\alpha = (T - w)^{-1} \sum_{t=w}^T \mathbf{1}\{l_t < \widehat{\text{VaR}}_t(\alpha)\}, \quad (23)$$

where  $w$  is the sliding window size and  $l_t$  is the realization of  $L_t$ . For the twelve copulas, Table 2 presents the ERs which optimal if it equals  $\alpha$ . The Gaussian copula performs best for  $\alpha = 0.05$ , the HAC-Clayton copula has reached the most appropriate ER for  $\alpha \in \{0.01, 0.005\}$  and the Clayton copula for  $\alpha = 0.001$ . The Factor-Gumbel copula

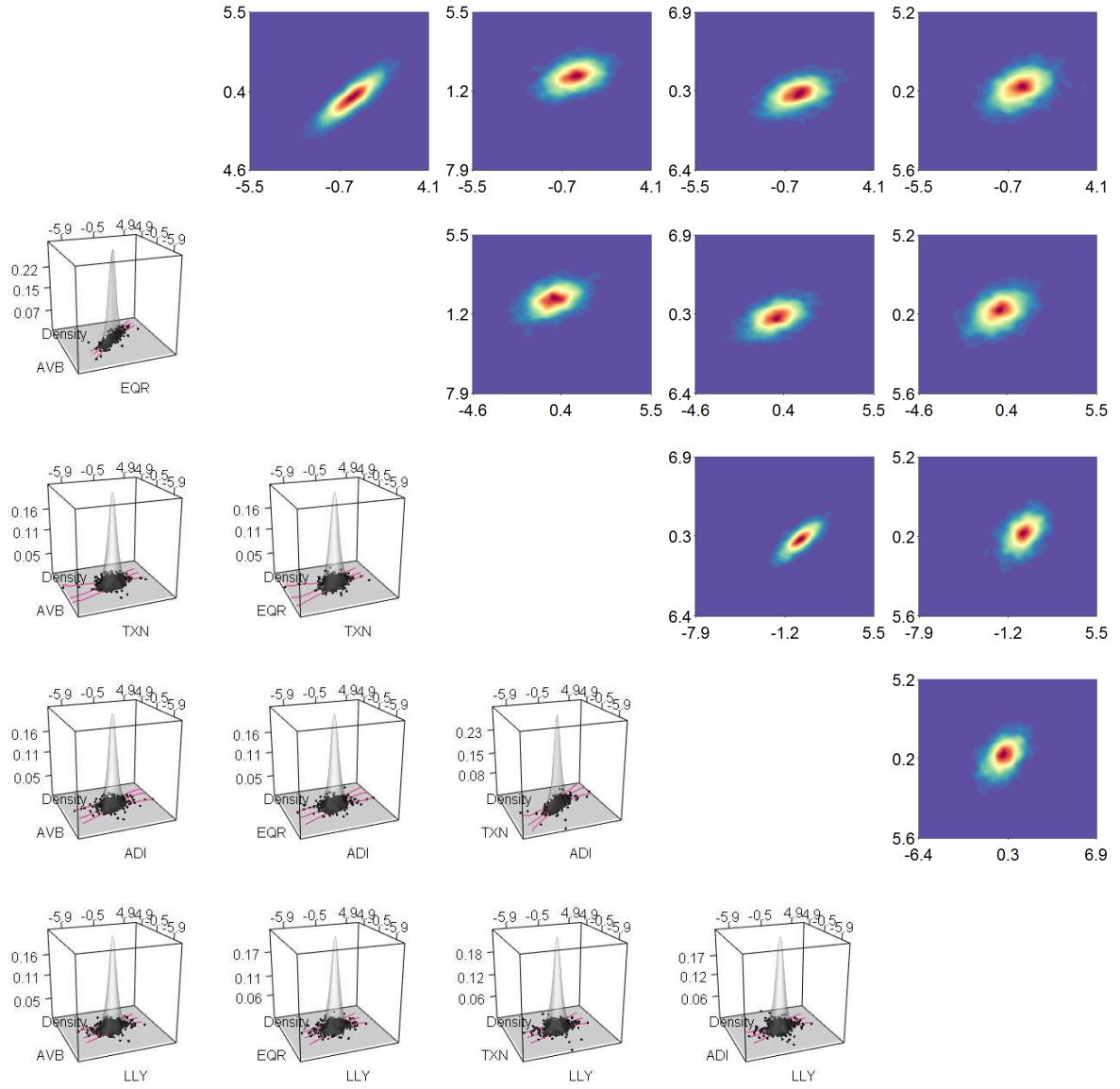


Figure 3: The lower triangular plots give 2-dimensional kernel density estimations containing scatter plots of pairwise GARCH(1, 1)-filtered log-returns with quantile regressions under 0.05, 0.5, 0.95 quantiles. The upper triangular plots give pairwise contours of five variables.

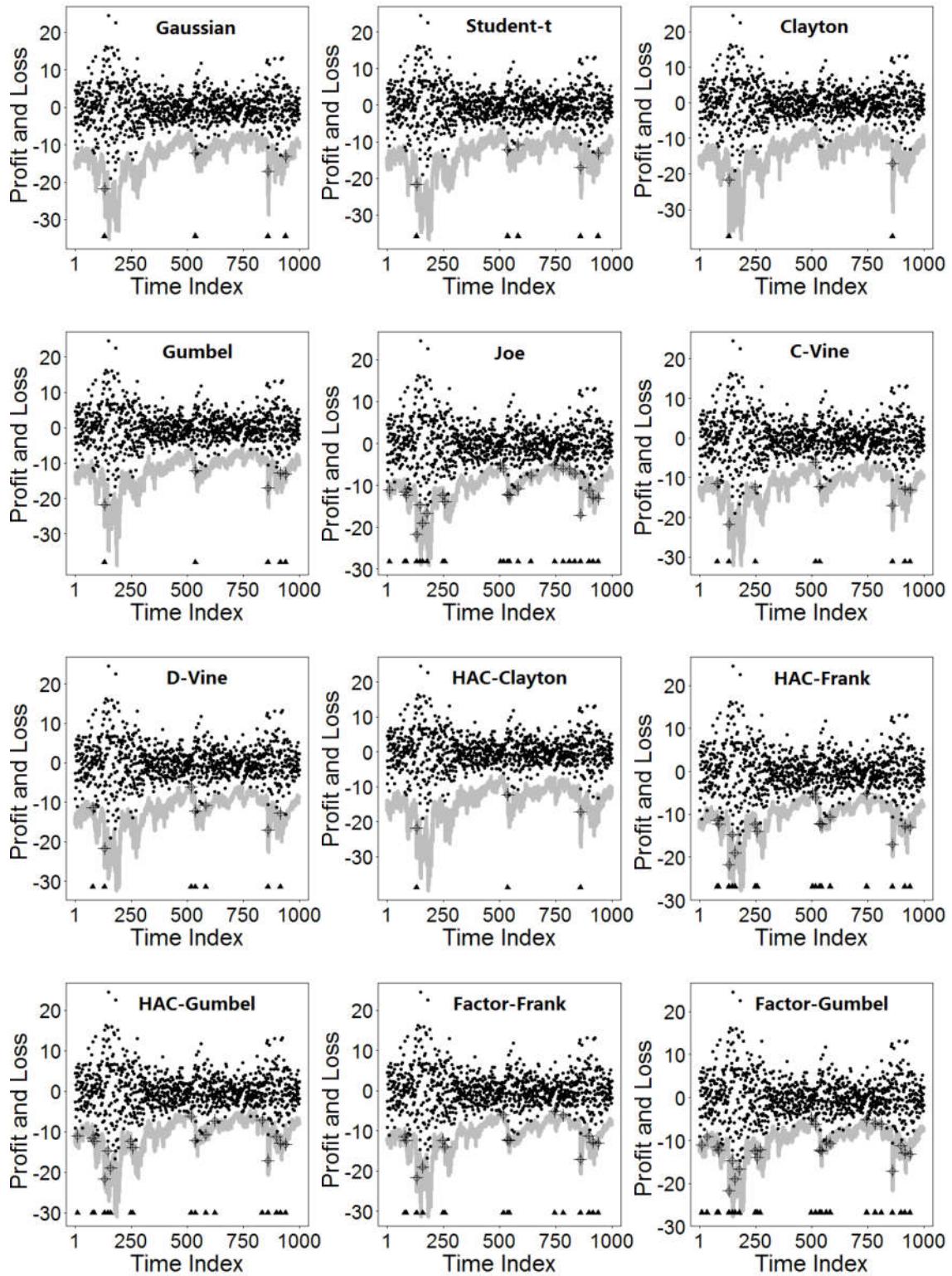


Figure 4: VaRs for  $\alpha = 0.001$  are constructed based on 1000 back-testing points with copulas of Gaussian,  $t$ , Clayton, Gumbel, Joe, C-Vine, D-Vine, HAC-Clayton, HAC-Frank, HAC-Gumbel, Factor-Frank, Factor-Gumbel, illustrated by row.

Copula	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.005$	$\alpha = 0.001$
Gaussian	<b>0.050</b>	0.018	0.009	0.004
t	0.048	0.014	0.011	0.005
Clayton	0.047	0.017	0.011	<b>0.002</b>
Gumbel	0.048	0.025	0.013	0.005
Joe	0.065	0.032	0.030	0.023
C-Vine	0.045	0.019	0.015	0.008
D-Vine	0.044	0.018	0.012	0.007
HAC-Clayton	0.044	<b>0.013</b>	<b>0.008</b>	0.003
HAC-Frank	0.055	0.033	0.026	0.016
HAC-Gumbel	0.070	0.036	0.028	0.017
Factor-Frank	0.046	0.026	0.017	0.015
Factor-Gumbel	0.086	0.042	0.032	0.024

Table 2: Exceeding ratios based on  $\alpha \in \{0.05, 0.01, 0.005, 0.001\}$ .

provides the worst ER values for all values of  $\alpha$ . Vines perform neither outstanding good nor bad. It deserves to be mentioned that copulas exhibiting upper-tail dependence show higher ER values, for instance, Joe copula, HAC-Gumbel copula and Factor-Gumbel copula. Even though some copulas are based on more parameters and thus, offer more flexibility, the increase of parameters does not essentially improve the ER.

## 7 Conclusion

This work discusses bivariate copula and focuses on three high dimensional copula models including the hierarchical Archimedean copula, the factor copula and the vine copula. The three models are developed in-depth with their advantages in modeling high dimensional data for diverse research fields. For the sake of comparison, an empirical study from risk management is presented. In this study, the estimation of Value-at-Risk is performed under 12 different copula models including the discussed state-of-art copulas as well as some classical benchmarks such as some of the elliptical and Archimedean family. Considered in toto, the hierarchical Archimedean copula with Clayton generator performs better than the alternatives in terms of the exceeding ratios measure.

## References

- Aas, K., Czado, C., Frigessi, A. and Bakken, H. (2009). Pair-copula constructions of multiple dependence, *Insurance: Mathematics and Economics* **44**(2): 182–198.
- Andersen, L., Sidenius, J. and Basu, S. (2003). Credit derivatives: All your hedges in one basket, *Risk* **16**: 67–72.
- Barbe, P., Genest, C., Ghoudi, K. and Rémillard, B. (1996). On Kendalls's process, *Journal of Multivariate Analysis* **58**: 197–229.
- Bedford, T. and Cooke, R. M. (2001). Probability density decomposition for conditionally dependent random variables modeled by vines, *Annals of Mathematical and Artificial Intelligence* **32**: 245–268.
- Bedford, T. and Cooke, R. M. (2002). Vines – a new graphical model for dependent random variables, *Annals of Statistics* **30**(4): 1031–1068.
- Brechmann, E. C. (2014). Hierarchical kendall copulas: Properties and inference, *Canadian Journal of Statistics* **42**(1): 78–108.
- Breymann, W., Dias, A. and Embrechts, P. (2003). Dependence structures for multivariate high-frequency data in finance, *Quantitative Finance* **1**: 1–14.
- Chen, S. X. and Huang, T. (2007). Nonparametric estimation of copula functions for dependence modeling, *The Canadian Journal of Statistics* **35**(2): 265–282.
- Chen, X. and Fan, Y. (2006). Estimation and model selection of semiparametric copula-based multivariate dynamic models under copula misspecification, *Journal of Econometrics* **135**(1–2): 125–154.
- Chen, X., Fan, Y. and Tsyrennikov, V. (2006). Efficient estimation of semiparametric multivariate copula models, *Journal of the American Statistical Association* **101**(475): 1228–1240.
- Cherubini, U., Luciano, E. and Vecchiato, W. (2004). *Copula Methods in Finance*, John Wiley & Sons, New York.
- Choroś-Tomczyk, B., Härdle, W. K. and Okhrin, O. (2013). Valuation of collateralized debt obligations with hierarchical Archimedean copulae, *Journal of Empirical Finance* **24**(C): 42–62.

- Dobrić, J. and Schmid, F. (2007). A goodness of fit test for copulas based on Rosenblatt's transformation, *Computational Statistics & Data Analysis* **51**(9): 4633–4642.
- Durante, F., Fernández-Sánchez, J. and Sempi, C. (2012). A topological proof of Sklar's theorem, *Applied Mathematical Letters* **26**: 945–948.
- Durante, F., Fernández-Sánchez, J. and Sempi, C. (2013). Sklar's theorem obtained via regularization techniques, *Nonlinear Analysis: Theory, Methods & Applications* **75**(2): 769–774.
- Durante, F. and Sempi, C. (2005). *Principles of Copula Theory*, Chapman and Hall/CRC.
- Fermanian, J.-D. (2005). Goodness-of-fit tests for copulas, *Journal of Multivariate Analysis* **95**(1): 119–152.
- Fermanian, J.-D. and Scailliet, O. (2003). Nonparametric estimation of copulas for time series, *Journal of Risk* **5**: 25–54.
- Gaensler, P. and Stute, W. (1987). *Seminar on empirical processes*, Springer Basel AG, Boca Raton.
- Genest, C., Ghoudi, K. and Rivest, L.-P. (1995). A semi-parametric estimation procedure of dependence parameters in multivariate families of distributions, *Biometrika* **82**(3): 543–552.
- Genest, C., Quessy, J.-F. and Rémillard, B. (2006). Goodness-of-fit procedures for copula models based on the probability integral transformation, *Scandinavian Journal of Statistics* **33**: 337–366.
- Genest, C. and Rémillard, B. (2008). Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models, *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques* **6**(44): 1096–1127.
- Genest, C., Rémillard, B. and Beaudoin, D. (2009). Goodness-of-fit tests for copulas: A review and a power study, *Insurance: Mathematics and Economics* **44**: 199–213.
- Genest, C. and Rivest, L.-P. (1989). A characterization of Gumbel family of extreme value distributions, *Statistics & Probability Letters* **8**(3): 207–211.
- Genest, C. and Rivest, L.-P. (1993). Statistical inference procedures for bivariate Archimedean copulas, *Journal of the American Statistical Association* **88**(3): 1034–1043.

- Górecki, J., Hofert, M. and Holeňa, M. (2016). An approach to structure determination and estimation of hierarchical Archimedean copulas and its application to bayesian classification, *Journal of Intelligent Information Systems* **46**(1): 21–59.
- Haff, I. H. (2013). Parameter estimation for pair-copula constructions, *Bernoulli* **19**(2): 462–491.
- Härdle, W. K. and Simar, L. (2015). *Applied Multivariate Statistical Analysis*, Springer, Berlin.
- Hering, C., Hofert, M., Mai, J.-F. and Scherer, M. (2010). Constructing hierarchical Archimedean copulas with Lévy subordinators, *Journal of Multivariate Analysis* **101**(6): 1428–1433.
- Hofert, M. (2011). Efficiently sampling nested Archimedean copulas, *Computational Statistics & Data Analysis* **55**(1): 57–70.
- Hofert, M. and Scherer, M. (2011). CDO pricing with nested Archimedean copulas, *Quantitative Finance* **11**(5): 775–87.
- Hull, J. and White, A. (2004). Valuation of a CDO and an  $n$ -th to default CDS without Monte Carlo simulation, *Journal of Derivatives* **12**(2): 8–23.
- 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. Taylor (eds), *Distribution with fixed marginals and related topics*, IMS Lecture Notes – Monograph Series, Institute of Mathematical Statistics.
- Joe, H. (1997). *Multivariate Models and Dependence Concepts*, Chapman & Hall, London.
- Joe, H. (2005). Asymptotic efficiency of the two-stage estimation method for copula-based models, *J. Multivariate Anal.* **94**(2): 401–419.
- Joe, H. (2014). *Dependence Modeling with Copulas*, Chapman and Hall/CRC, Boca Raton.
- Johnson, R. A. and Wichern, D. W. (2013). *Applied Multivariate Statistical Analysis*, 6 edn, Pearson, Harlow.
- Jouini, M. and Clemen, R. (1996). Copula models for aggregating expert opinions, *Operation Research* **3**(44): 444–457.
- Kendall, M. (1970). *Rank Correlation Methods*, Griffin, London.

Krupskii, P. and Joe, H. (2013). Factor copula models for multivariate data, *Journal of Multivariate Analysis* **120**: 85–101.

Kurowicka, M. and Cooke, R. M. (2006). *Uncertainty Analysis with High Dimensional Dependence Modelling*, John Wiley & Sons, New York.

Laurent, J.-P. and Gregory, J. (2005). Basket default swaps, CDO's and factor copulas, *Journal of Risk* **7**(4): 103–122.

Li, D. X. (2000). On default correlation: A copula function approach, *Journal of Fixed Income* **9**: 43–54.

Luo, X. and Shevchenko, P. V. (2010). The  $t$ -copula with multiple parameters of degrees of freedom: Bivariate characteristics and application to risk management, *Quantitative Finance* **10**: 1039–1054.

Mai, J.-F. and Scherer, M. (2013). What makes dependence modeling challenging? Pitfalls and ways to circumvent them, *Statistics & Risk Modeling* **30**(4): 287–306.

Nelsen, R. B. (2006). *An Introduction to Copulas*, Springer, New York.

Oakes, D. (1994). Multivariate survival distributions, *Journal of Nonparametric Statistics* **3**(3-4): 343–354.

Oh, D. H. and Patton, A. (2015). Modelling dependence in high dimensions with factor copulas, *Finance and Economics Discussion Series 2015-051*. Washington: Board of Governors of the Federal Reserve System .

Oh, D. H. and Patton, A. J. (2013). Simulated method of moments estimation for copula-based multivariate models, *Journal of the American Statistical Association* **108**(502): 689–700.

Okhrin, O., Okhrin, Y. and Schmid, W. (2013a). On the structure and estimation of hierarchical Archimedean copulas, *Journal of Econometrics* **173**(2): 189–204.

Okhrin, O., Okhrin, Y. and Schmid, W. (2013b). Properties of hierarchical Archimedean copulas, *Statistics & Risk Modeling* **30**(1): 21–54.

Okhrin, O. and Ristig, A. (2014). Hierarchical Archimedean copulae: The HAC package, *Journal of Statistical Software* **58**(4): 1–20.

Patton, A., Fan, Y. and Chen, X. (2004). Simple tests for models of dependence between

multiple financial time series, with applications to u.s. equity returns and exchange rates, *Working paper* .

Patton, A. J. (2012). A review of copula models for economic time series, *Journal of Multivariate Analysis* **110**: 4–18.

Rachev, S., Stoyanov, S. and Fabozzi, F. (2008). *Advanced stochastic models, risk assessment, and portfolio optimization: The ideal risk, uncertainty, and performance measures*, John Wiley & Sons.

Radulovic, J.-D. F. D. and Wegkamp, M. (2004). Weak convergence of empirical copula processes, *Bernoulli* **10**(5): 847–860.

Rezapour, M. (2015). On the construction of nested Archimedean copulas for  $d$ -monotone generators, *Statistics & Probability Letters* **101**(0): 21–32.

Rosenblatt, M. (1952). Remarks on a multivariate transformation, *Annals of Mathematical Statistics* **23**: 470–472.

Savu, C. and Trede, M. (2010). Hierarchies of Archimedean copulas, *Quantitative Finance* **10**(3): 295–304.

Scaillet, O. (2007). Kernel-based goodness-of-fit tests for copulas with fixed smoothing parameters, *Journal of Multivariate Analysis* **98**(3): 533–543.

Schreyer, M., Paulin, R. and Trutschnig, W. (2017). On the exact region determined by kendall's  $\tau$  and spearman's  $\rho$ , *to appear in: Journal of the Royal Statistical Society: Series B (Statistical Methodology)* .

Segers, J. and Uyttendaele, N. (2014). Nonparametric estimation of the tree structure of a nested Archimedean copula, *Computational Statistics & Data Analysis* **72**: 190–204.

Shih, J. H. and Louis, T. A. (1995). Inferences on the association parameter in copula models for bivariate survival data, *Biometrics* **51**(4): 1384–1399.

Sklar, A. (1959). Fonctions de répartition à  $n$  dimension et leurs marges, *Publications de l'Institut de Statistique de l'Université de Paris* **8**: 299–231.

Wang, W. and Wells, M. (2000). Model selection and semiparametric inference for bivariate failure-time data, *Journal of the American Statistical Association* **95**(449): 62–76.

Whelan, N. (2004). Sampling from Archimedean copulas, *Quantitative Finance* **4**(3): 339–352.

Zhang, S., Okhrin, O., Zhou, Q. M. and Song, P. X.-K. (2016). Goodness-of-fit test for specification of semiparametric copula dependence models, *Journal of Econometrics* **193**(1): 215 – 233.

Zhu, W., Wang, C.-W. and Tan, K. S. (2016). Structure and estimation of Lévy subordinated hierarchical Archimedean copulas (LSHAC): Theory and empirical tests, *Journal of Banking & Finance* .