SFB 649 Discussion Paper 2010-022

Fitting high-dimensional Copulae to Data

Ostap Okhrin*

* Humboldt-Universität zu Berlin, Germany

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

http://sfb649.wiwi.hu-berlin.de
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin
Spandauer Straße 1, D-10178 Berlin

SFB 649 ECONOMIC RISK BERLIN

Fitting high-dimensional Copulae to Data*

Ostap Okhrin†

May 13, 2010

Abstract: This paper make an overview of the copula theory from a practical side. We consider different methods of copula estimation and different Goodness-of-Fit tests for model selection. In the GoF section we apply Kolmogorov-Smirnov and Cramer-von-Mises type tests and calculate power of these tests under different assumptions. Novating in this paper is that all the procedures are done in dimensions higher than two, and in comparison to other papers we consider not only simple Archimedean and Gaussian copulae but also Hierarchical Archimedean Copulae. Afterwards we provide an empirical part to support the theory.

Keywords: copula; multivariate distribution; Archimedean copula; GoF.

JEL Classification: C13, C14, C50.

1 Introduction

Many practical problems arise from modelling high dimensional distributions. Precise modelling is important in fitting of asset returns, insurance payments, overflows from a dam and so on. Often practitioners stay ahead of potential problems by using assets backed up in huge portfolios, payments spatially distributed over land, and dams located on rivers where there are already other hydrological stations. This means that univariate problems are extended to multivariate ones in which all the univariate ones are dependent on each other. Until the late 1990s elliptical distribution, in particular the multivariate normal one, was the most desired distribution in practical applications. However the normal distribution does not, in practice, meet most applications. Some studies (see e.g Fama (1965), Mandelbrot (1965), etc.) show that daily returns are not normally distributed but follow stable distributions. This means that on one hand one cannot take the distribution in which margins are normal, and on the other hand, stable multivariate distributions are difficult to implement. In the hydrological problem, margins arise from extreme value distribution, while one is interested in the maximal value of the water collected after the winter season over a number of years, this value arises from the family

*The financial support from the Deutsche Forschungsgemeinschaft via SFB 649 "Okonomisches Risiko", Humboldt-Universität zu Berlin is gratefully acknowledged.

†C.A.S.E. - Center for Applied Statistics and Economics, Ladislaus von Bortkiewicz Chair of Statistics of Humboldt-Universität zu Berlin, Spandauer Straße 1, D-10178 Berlin, Germany. Email: ostap.okhrin@wiwi.hu-berlin.de

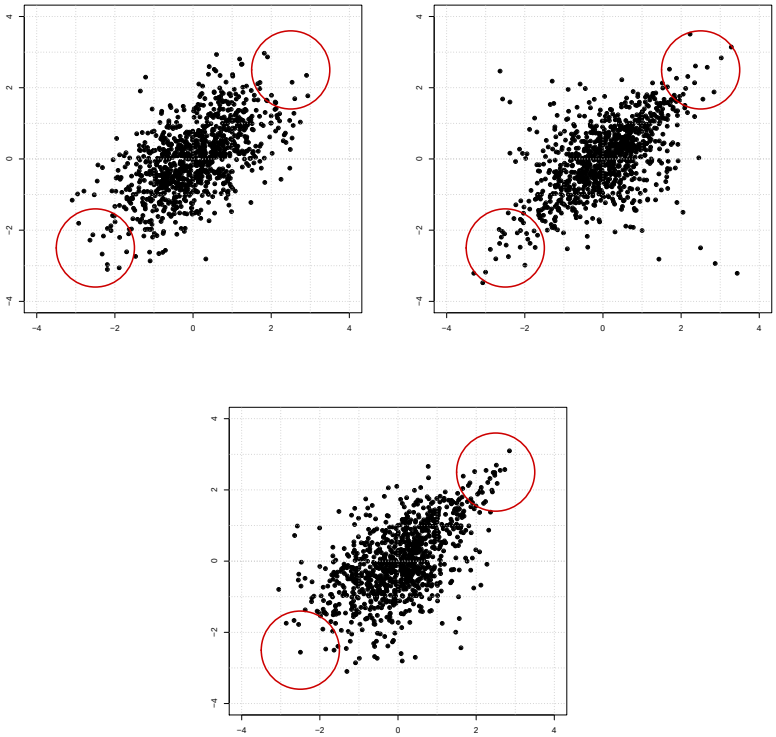

Figure 1: Scatter plots of bivariate samples with different dependency structures

of extreme distributions. As in the previous example, the multivariate extreme value distribution family is also somewhat restrictive.

Two further problems are illustrated in Figure 1. The scatter plot in the first figure shows the realisations of two Gaussian random variables. The points are symmetric and no extreme outliers can be observed. In contrary, the second picture exhibits numerous outliers. The outliers in the first and third quadrants show that extreme values often occur simultaneously for both variables. Such behaviour is observed in crisis periods, when strong negative movements on financial markets occur simultaneously. On the third figure we observe that the dependency between the negative values is different compared to the positive values. This type of non-symmetric dependency cannot be modeled by elliptical distributions, because they impose a very specific radially symmetric dependency structure.

Following these examples we need a solution to easily separate the modelling of the dependency structure and the margins. This is one of the tasks of copulae; to enable the modelling of marginals separately from the dependency. The above problem concerning assets could be solved by taking margins from the stable distribution and the dependency, as in the multivariate one. Similar solutions could be found for other problems. In finance, copulae are applied in different fields such as credit portfolio modelling and risk management.

Over the last 40 years, copula has only been attractive from a mathematical perspective, and only as late as 1999 were the different complicated properties of copula, such as the distribution (which made it more flexible), settled and solved. Nowadays dependency plays a key role in many financial models, starting from the basic portfolio theory of

Markowitz. Recent developments strongly support the joint non-Gaussianity of asset returns and exploit numerous alternative approaches to model the underlying distribution. The key role of dependency can be best illustrated by the famous quote "Given the prices of single-bet financial contracts, what is the price of multiple-bet contracts? There is no unique answer to that question...". The first application of copulae to financial data was carried out by Embrechts, McNeil and Straumann (1999). In this paper copulae were used in risk management framework which stimulated a series of ground breaking applied papers. Breymann, Dias and Embrechts (2003) model the dependencies of high-frequency data. An application to risk management is discussed in Junker and May (2005). Portfolio selection problems were considered in Hennessy and Lapan (2002) and in Patton (2004). Theoretical foundations of copula-based GARCH models and its application were proposed by Chen and Fan (2005). Lee and Long (2009), Giacomini, Härdle and Spokoiny (2009) and Härdle, Okhrin and Okhrin (2010) consider time varying copulae.

The new fields of application show the need for further theoretical developments. Each proposed model should be estimated with either parametric, semi- or nonparametric methods. The semiparametric estimation of the copula-based distribution, which is based on the nonparametric estimation of margins and estimation of the parameter for the fixed copula function, is discussed in Chen and Fan (2006), Chen, Fan and Tsyrennikov (2006), Genest, Ghoudi and Rivest (1995), Joe (2005), Wang and Wells (2000). Fully nonparametric estimation is discussed in Fermanian and Scaillet (2003), Chen and Huang (2007), Lejeune and Sarda (1992). To measure how well a copula-based statistical model fits the data, several goodness-of-fit tests were developed and discussed in the papers by Chen and Fan (2005), Chen, Fan and Patton (2004), Fermanian (2005) and Genest, Quessy and Rémillard (2006), Genest and Rémillard (2008), Breymann et al. (2003). In-depth discussion of simulation methodologies for Archimedean copulae can be found in Whelan (2004) and McNeil (2008). A detailed review and discussion of copula theory is given in Joe (1997) and Nelsen (2006).

In this chapter we describe the attractive features of copulae from the statistical perspective, with examples and applications in real data. We consider the most important copula classes with different methods of estimation and goodness-of-fit tests. We compare different goodness-of-fit tests by their rejection rates, for which a profound simulation study has been devised. In the empirical part of the chapter we apply different copula models to the normalised residuals and test the quality of the fit by discussed goodness-of-fit tests. We found that for the selected datasets hierarchical Archimedean copula outperform the simple Archimedean copula and the Gaussian copula by all goodness-of-fit tests.

2 Theoretical Background

From the early days of the multivariate probability theory it is well known, that given the d-variate distribution function $F : \mathbb{R} \to [0; 1]$ of a d-variate random vector $(X_1, \ldots, X_d)$ the distribution function, called marginal distribution function of each of the d components

$X_1, \ldots, X_d$ is easily computed:

$$\begin{aligned} F_1(x) &= F(x, +\infty, \ldots, +\infty), \\ F_2(x) &= F(+\infty, x, +\infty, \ldots, +\infty), \\ &\cdots \\ F_d(x) &= F(+\infty, \ldots, +\infty, x). \end{aligned}$$

The converse problem was studied by Fréchet (1951), Hoeffding (1940), Hoeffding (1941), where having the distribution functions $F_1, \ldots, F_d$ of d random variables $X_1, \ldots, X_d$ defined on the same probability space $(\Omega, \mathcal{F}, \mathcal{P})$ they wanted to make a conclusions about the set $\Gamma(F_1 \ldots, F_d)$ of the d-variate distribution functions whose marginals are $F_1, \ldots, F_d$

$$F \in \Gamma(F_1, \ldots, F_d) \Leftrightarrow \left\{ \begin{array}{rcl} F_1(x) &=& F(x, +\infty, \ldots, +\infty), \\ F_2(x) &=& F(+\infty, x, +\infty, \ldots, +\infty), \\ & \cdots & \\ F_d(x) &=& F(+\infty, \ldots, +\infty, x). \end{array} \right.$$

Nowadays the set $\Gamma(F_1, \ldots, F_d)$ is called the Fréchet class of $F_1, \ldots, F_d$. Γ is not empty, because it always contains the independence case in which $F(x_1, \ldots, x_d) = F_1(x_1) \cdot \cdots \cdot F_d(x_d),\ \forall x_1, \ldots, x_d \in \mathbb{R}$. Dealing with Fréchet classes, one often interests in the bounds and members of the Γ. Dall'Aglio (1972) studies conditions under which there is only one distribution function which belongs to $\Gamma(F_1, \ldots, F_d)$. A nice and short review of the Fréchet classes can be found in Joe (1997).

In 1959 Sklar found the partial solution to the above mentioned problem by introducing copulae. Because there are a variety of copula definitions we will first look at the most general one. For this we will need to define the C-volume with the d-box that is a cartesian product $[\mathbf{a}, \mathbf{b}] = \prod_{j=1}^{d} [a_j, b_j]$, where, for every index $j \in \{1, 2, \ldots, d\},\ 0 \leq a_j \leq b_j \leq 1$.

Definition 1 *For a function $C : [0;1]^d \to [0;1]$, the C-volume V_c of the box $[\mathbf{a}, \mathbf{b}]$ is defined via*

$$V_c([\mathbf{a}, \mathbf{b}]) \stackrel{def}{=} \sum_{\mathbf{v}} sign(\mathbf{v}) C(\mathbf{v}),$$

where the sum is carried over all the 2^d vertices $\mathbf{v}$ of the box $[\mathbf{a}, \mathbf{b}]$. Here also

$$sing(\mathbf{v}) = \left\{ \begin{array}{ll} 1, & \text{if } v_j = a_j \text{ for an even number of vertices,} \\ -1, & \text{if } v_j = a_j \text{ for an odd number of vertices.} \end{array} \right.$$

Here is the definition of a copula, see Härdle and Simar (2007):

Definition 2 *A function $C : [0,1]^d \to [0,1]$ is a d-dimensional copula if*

1. *$C(x_1, \ldots, x_d) = 0$, when $x_j = 0$ for at least one index $j \in \{1, \ldots, d\}$;*

2. $C(1, 1, \ldots, x_j, 1 \ldots, 1) = x_j$;

3. *the V_c-volume of every d-box* $[\mathbf{a}, \mathbf{b}]$ *is positive:* $V_c([\mathbf{a}, \mathbf{b}]) \geq 0$.

The set of all the d-dimensional copulae ($d \geq 3$) in the rest of the chapter is denoted as $\mathcal{C}_d$, while the set of all bivariate ($d = 2$) copulae is denoted by $\mathcal{C}$. As already mentioned above, this simple family of functions has been extremely popular because of its property given in the Sklar (1959) theorem

Theorem 1 *Given a d-dimensional distribution function F, a copula $C \in \mathcal{C}_d$ exists such that for all $(x_1, \ldots, x_d) \in \overline{\mathbb{R}}^d$:*

$$F(x_1, \ldots, x_d) = C\{F_1(x_1), \ldots, F_d(x_d)\}. \tag{1}$$

The copula C is uniquely defined on $\prod_{j=1}^{d} F_j(\overline{\mathbb{R}})$ and therefore unique if all margins are continuous, thus

$$C(u_1, \ldots, u_d) = F\{F_1^{-1}(u_1), \ldots, F_d^{-1}(u_d)\}. \tag{2}$$

Conversely, if $F_1, \ldots, F_d$ are d one-dimensional distribution functions, then the function F defined in (1) is a d-dimensional distribution function.

Sklar's theorem also answers the question of the uniqueness of the copula C. However, if, for example, in the two dimensional case at least one of the two distribution functions has a discrete component, there may be more than one copula extending C from $F_1(\overline{\mathbb{R}}) \times F_2(\overline{\mathbb{R}})$ to the whole unit square $[0, 1]^2$. This is due to a fact that C is uniquely defined only on the product of the ranges $F_1(\overline{\mathbb{R}}) \times F_2(\overline{\mathbb{R}})$. In this case it is good to have a procedure of bilinear interpolation in order to single out a unique copula. In the variety of papers where copulae are applied in different fields, authors usually do not consider the assumption that the random variables are continuous. This assumption is necessary to avoid problems with non-uniqueness. The second part of the Sklar's theorem is based on the construction of the multivariate distribution from the margins and the copula function. It is extremely popular in practice, where, for example, in risk management, analysts may have a better idea about the marginal behaviour of individual risk factors, than about their dependency structure. This approach allows them to combine marginal models and to investigate the sensitivity of risk to the dependence specification.

New multivariate distributions are created in two steps. At first, all univariate random variables $X_1, \ldots, X_d$ are separately described by their marginal distributions $F_1, \ldots, F_d$. Then secondly, the copula $C \in \mathcal{C}_d$ which contains all the information about the relationship between the original variables $X_1, \ldots, X_d$ – not taking into account the information provided by $F_1, \ldots, F_d$ – is introduced.

Being armed with the remarks written above, one can write the following copula definition

Definition 3 *A d-dimensional copula is a cumulative distribution function on $[0, 1]^d$ with standard uniform marginal cumulative distribution functions.*

As in the case of the multivariate distribution, mentioned at the beginning, setting all of the arguments equal to $+\infty$ one gets an univariate marginal distribution. A univariate marginal of copula C is obtained by setting some of its arguments equal to 1. Similarly the m-marginal of C, $m < d$ is given by setting all $d - m$ arguments equal to 1, from the simple combinatoric problem, we see that there are $\binom{d}{m}$ different m-margins of the copula C.

A copula C satisfies a set of different important conditions, one of which is the Lipschitz condition which says that:

$$|C(u_1, \ldots, u_d) - C(v_1, \ldots, v_d)| \leq \sum_{j=1}^{d} |v_j - u_j|.$$

Another property says, that $\forall j \in \{1, \ldots, d\}$, $\{u_1, \ldots, u_{j-1}, t, u_{j+1}, \ldots, u_d\}$, $\forall t \in [0, 1]$, the functions $t \mapsto C(u_1, \ldots, u_{j-1}, t, u_{j+1}, \ldots, u_d)$ are increasing as functions of t.

To get a better impression of what a copula is from a definition, let us consider a special bivariate case. Explicitly, a bivariate copula is a function $C : [0, 1]^2 \to [0, 1]$ such that:

1. $\forall u \in [0, 1] \quad C(u, 0) = C(0, u) = 0$;
2. $\forall u \in [0, 1] \quad C(u, 1) = C(1, u) = u$;
3. $\forall u, u', v, v' \in [0, 1]$ with $u \leq u'$ and $v \leq v'$

$$C(u', v') - C(u', v) - C(u, v') + C(u, v) \geq 0.$$

The last inequality is referred to as the rectangular inequality and the function that satisfies it is said to be 2-increasing. The bivariate copula is always of special interest, because of the properties that are difficult to derive in higher dimensions.

The property of increasingness with respect to each argument could be profound for the bivariate copula in the following way. As we know from above, if C is a bivariate copula, then functions $[0, 1] \ni t \mapsto C(t, v)$ and $[0, 1] \ni t \mapsto C(v, t)$ are increasing with respect to t. The increasingness with respect to each argument means that derivatives with respect to Lebegue measure exist almost everywhere, and those derivatives are positive where they exist. From the Lipschitz conditions they are also bound above

$$0 \leq \frac{\partial C(s, t)}{\partial s} \leq 1, \quad 0 \leq \frac{\partial C(s, t)}{\partial t} \leq 1.$$

Every copula can be expressed in the form of the sum of absolutely continuous and singular part and an absolutely continuous copula C has a density c such that

$$C(u_1, \ldots, u_d) = \int\limits_{[0,1]^d} c(s_1, \ldots, s_d) ds_1 \ldots ds_d = \int\limits_0^1 ds_1 \ldots \int\limits_0^1 c(s_1, \ldots, s_d) ds_d$$

from which the copula density is found by differentiation

$$c(u_1, \ldots, u_d) = \frac{\partial^d C(u_1, \ldots, u_d)}{\partial u_1 \ldots \partial u_d}.$$

Following the Sklar theorem, the multivariate distribution F with margins $F_1, \ldots, F_d$ has multivariate density f with marginal densities $f_1, \ldots, f_d$ respectively. If, from the Sklar theorem copula C exists such that $F(x_1, \ldots, x_d) = C\{F_1(x_1), \ldots, F_d(x_d)\}$ then the d-variate density is

$$f(x_1, \ldots, x_d) = c\{F_1(x_1), \ldots, F_d(x_d)\} \cdot f_1(x_1) \ldots f_d(x_d). \tag{3}$$

Notice, however, that, as a consequence of the Lipschitz condition, for every bivariate copula C and for every $v \in [0, 1]$, both functions $t \mapsto C(t, v)$ and $t \mapsto C(v, t)$ are absolutely continuous so that

$$C(t, v) = \int_0^t c_{1v}(s)ds \text{ and } C(v, t) = \int_0^t c_{2v}(s)ds.$$

Unfortunately, this representation has no application so far.

3 Copula Classes

Naturally, there are an infinite number of different copula functions satisfying the assumptions of definition. In this section we discuss in details three important classes of simple, elliptical and Archimedean copulae.

3.1 Simple Copulae

Often we are interested in some extreme, special cases, like independence and perfect positive or negative dependence. If d-random variables $X_1, \ldots, X_d$ are stochastically independent from the Sklar Theorem the structure of such a relationship is given by the product (independence) copula defined as

$$\Pi(u_1, \ldots, u_d) = \prod_{j=1}^{d} u_j, \ u_1, \ldots, u_d \in [0, 1].$$

Another two extremes are the lower and upper Fréchet-Hoeffding bounds. They represent the perfect negative and positive dependencies respectively

$$\begin{aligned} W(u_1, \ldots, u_d) &= \max\big(0, \sum_{j=1}^{d} u_j + 1 - d\big), \\ M(u_1, \ldots, u_d) &= \min(u_1, \ldots, u_d), \quad u_1, \ldots, u_d \in [0, 1]. \end{aligned}$$

If, in a two dimensional case $C = W$ 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 other words both M and W are singular, where M uniformly spreads the probability mass on the diagonal $X_1 = X_2$ and W uniformly spreads the probability mass on the opposite

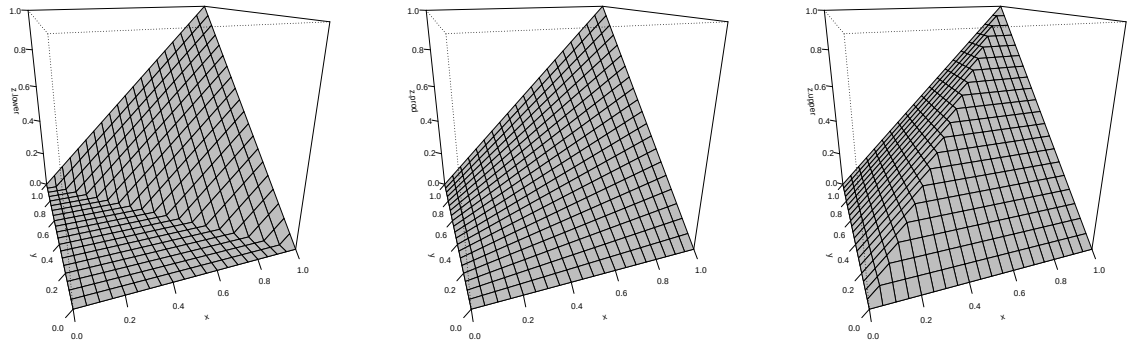

Figure 2: Lower Frechet Hoeffdings bound, Product copula and upper Frechet Hoeffdings bound in two-dimensional case (from left to right).

diagonal $X_1 = -X_2$. In general we can argue that an arbitrary copula which represents some dependency structure lies between these two bounds, i.e.

$$W(u_1, \ldots, u_d) \leq C(u_1, \ldots, u_d) \leq M(u_1, \ldots, u_d).$$

The bounds serve as benchmarks for the evaluation of the dependency magnitude. Note, however, that the lower Fréchet-Hoeffding bound is not a proper copula function for $d > 2$ but is a proper quasi-copula. Both upper and lower bounds are sharp, because there are copulae, that are either equal, at some points, to one of the two bounds.

The simple copulae for the two dimensional case are plotted in Figure 2.

3.2 Elliptical Copulae

Due to the popularity of Gaussian and t-distributions in financial applications, elliptical copulae also play an important role. For example, in the modelling of collateralized debt obligations, where the assumption of the Gaussian one-factor dependency between joint default of the obligors, proposed by Li (2000), is seen as a standard approach. The construction of this type of copulae is based directly on the Sklar Theorem. The Gaussian copula and its copula density are given by:

$$\begin{aligned} C_N(u_1, \ldots, u_d, \boldsymbol{\Sigma}) &= \boldsymbol{\Phi}_{\boldsymbol{\Sigma}}\{\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_d)\}, \\ c_N(u_1, \ldots, u_d, \boldsymbol{\Sigma}) &= \\ &= |\boldsymbol{\Sigma}|^{-1/2} \exp\left\{-\frac{[\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_d)]'(\boldsymbol{\Sigma}^{-1} - \boldsymbol{I})[\boldsymbol{\Phi}^{-1}(u_1), \ldots, \boldsymbol{\Phi}^{-1}(u_d)]}{2}\right\}, \\ &\qquad \text{for all } u_1, \ldots, u_d \in [0, 1], \end{aligned}$$

where $\boldsymbol{\Phi}_{\boldsymbol{\Sigma}}$ is a d-dimensional normal distribution with a zero mean and the correlation matrix $\boldsymbol{\Sigma}$. The variances of the variables are imposed by the marginal distributions. Note, that in the multivariate case the implementation of elliptical copulae is very involved due to technical difficulties with multivariate cdf's. The level plots of the two-dimensional respective densities with different margins are given in Figure 3.

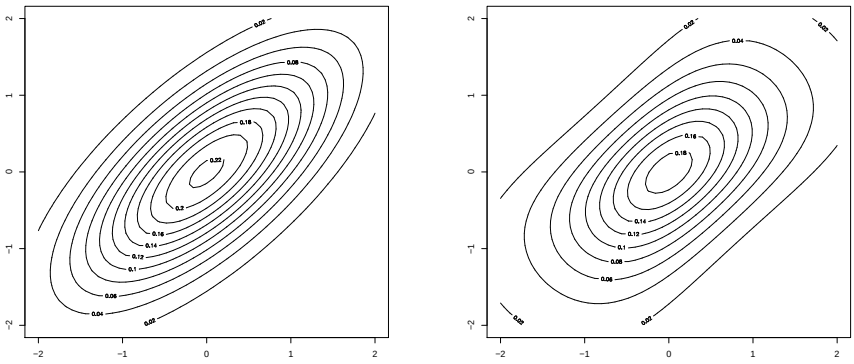

Figure 3: Contour diagrams for Gaussian copula with Gaussian (left column) and t_3 distributed (right column) margins.

Using (2) one can derive the copula function for an arbitrary elliptical distribution. The problem is, however, that such copulae depend on the inverse distribution functions and these are rarely available in an explicit form. Therefore, the next class of copulae with its generalisations provides an important flexible and rich family of alternatives to the elliptical copulae.

3.3 Archimedean Copulae

In contrast to elliptical copulae, Archimedean copulae have a special method of construction which does not use (2), but fulfills all the conditions of the copula. Having M as an univariate distribution function of the positive random variable let ϕ be the Laplace transform of M, $\phi = \mathcal{LS}(M)$

$$\phi(s) = \int_0^\infty e^{-sw}\, dM(w), \quad s \geq 0. \tag{4}$$

Thus, M is said to be the inverse Laplace transform of ϕ, $M = \mathcal{LS}^{-1}(\phi)$. We denote as $\mathcal{L}$ the class of Laplace transforms which contain strictly decreasing differentiable functions, see Joe (1997):

$$\mathcal{L} = \{\phi : [0;\infty) \to [0,1] \,|\, \phi(0) = 1,\, \phi(\infty) = 0;\, (-1)^j \phi^{(j)} \geq 0;\, j = 1, \ldots, \infty\}.$$

It is known, that for an arbitrary univariate distribution function F, a unique distribution function G exists such that

$$F(x) = \int_0^\infty G^\alpha(x)\, dM(\alpha) = \phi\{-\log G(x)\}.$$

This leads to $G = \exp\{-\phi^{[-1]}(F)\}$, where $\phi^{[-1]}$ is the generalised inverse

$$\phi^{[-1]}(x) = \begin{cases} \phi^{-1}(x) & \text{for } 0 \leq x < \phi(0); \\ 0 & \text{else.} \end{cases}$$

Taking d univariate distributions $F_1, \ldots, F_d$, a simple extension leads to the multivariate distribution function that belongs to $\Gamma(F_1, \ldots, F_d)$

$$\begin{aligned} F &= \int G_1^{\alpha} \ldots G_d^{\alpha} \, dM(\alpha) = \phi(-\log G_1 - \cdots - \log G_d) \\ &= \phi\left\{\sum_{j=1}^{d} \phi^{[-1]}(F_j)\right\}, \end{aligned}$$

with Archimedean copula given by

$$C(u_1, \ldots, u_d) = \phi\left\{\sum_{j=1}^{d} \phi^{[-1]}(u_j)\right\}. \tag{5}$$

The function ϕ is called the *generator* of the Archimedean copula. Throughout the chapter the notation ϕ^{-1} is understood as the generalised inverse $\phi^{[-1]}$. Usually generator function depends on the parameter θ which is set to be the parameter of the copula. It is easy to see, that Archimedean copulae are exchangeable. In two-dimensional cases they are symmetric in the sense that $C(u, v) = C(v, u), \ \forall u, v \in [0, 1]$. Joe (1997) and Nelsen (2006) provide a classified list of the typical Archimedean generators. Here we discuss the three most commonly used ones in financial applications, Archimedean copulae.

The first, widely used (in practice) copula is the Gumbel (1960) copula, which gained its popularity from the extreme value theory. The multivariate distribution based on the Gumbel copula with univariate extreme value marginal distributions is the only extreme value distribution based on an Archimedean copula, see Genest and Rivest (1989). Moreover, all distributions based on Archimedean copulae belong to its domain of attraction under common regularity conditions. Direct and inverse generators of the Gumbel copula with the copula function are given by

$$\begin{aligned} \phi(x, \theta) &= \exp\{-x^{1/\theta}\}, \quad 1 \le \theta < \infty, \ x \in [0, \infty), \\ \phi^{-1}(x, \theta) &= (-\log x)^{\theta}, \quad 1 \le \theta < \infty, \ x \in [0, 1], \\ C_\theta(u_1, \ldots, u_d) &= \exp\left[-\left\{\sum_{j=1}^{d} (-\log u_j)^{\theta}\right\}^{\theta^{-1}}\right], \ u_1, \ldots, u_d \in [0, 1]. \end{aligned}$$

The Gumbel copula leads to asymmetric contour diagrams and shows stronger linkage between positive values, however, is also shows more variability and more mass in the negative tail.

For $\theta = 1$, the Gumbel copula reduces to the product copula and for $\theta \to \infty$ we obtain the Fréchet-Hoeffding upper bound. This copula does not have an extension to the negative dependence. The Gumbel copula is one of a few Archimedean copulae for which we have an explicit form of the distribution function M from (4). In the case of Gumbel copula M is the stable distribution, see Renyi (1970). This information is very useful in the simulation techniques, especially for the Marshall and Olkin (1988) method, see Section Simulations 4.

Another example is the Clayton (1978) copula which, in contrary to the Gumbel, has more mass on the lower tail, and less on the upper. This copula is often used in the

modelling of the losses, which is of interest, for example, in insurance and finance. The necessary functions for this example are

$$\begin{aligned}
\phi(x,\theta) &= (\theta x+1)^{-\frac{1}{\theta}}, \quad -1/(d-1) \le \theta < \infty,\ \theta \neq 0,\ x \in [0,\infty), \\
\phi^{-1}(x,\theta) &= \frac{1}{\theta}(u^{-\theta}-1),\ -1/(d-1) \le \theta < \infty,\ \theta \neq 0,\ x \in [0,1], \\
C_\theta(u_1,\ldots,u_d) &= \left\{ \left(\sum_{j=1}^{d} u_j^{-\theta} \right) - d + 1 \right\}^{-\theta^{-1}}, \ u_1,\ldots,u_d \in [0,1].
\end{aligned}$$

The Clayton copula is one of few copulae that has a truncation property and has a simple explicit form of density for any dimension

$$c_\theta(u_1,\ldots,u_d) = \prod_{j=1}^{d} \{1+(j-1)\theta\} u_j^{-(\theta+1)} \left(\sum_{j=1}^{d} u_j^{-\theta} - d + 1 \right)^{-(\theta^{-1}+d)}.$$

As the parameter θ tends to infinity, dependence becomes maximal and the copula gives the upper Frechet-Hoeffding bound. As θ tends to zero, we have independence. As $\theta \to -1/(d-1)$, the distribution tends to the lower Fréchet bound.

Another interesting Archimedean copula is the so called Frank (1979) copula, which, in the bivariate case, is the only elliptical Archimedean copula in the sense that $C(u,v) = u+v-1+C(1-u,1-v) = \overline{C}(u,v)$, where $\overline{C}(u,v)$ is called the survival or associative copula. $\overline{C}(u,v)$ is also a copula for a survival bivariate distribution. Direct and inverse generator of the Frank copula with the copula functions are

$$\begin{aligned}
\phi(x,\theta) &= -\frac{1}{\theta}\log\{1+e^u(e^{-\theta}-1)\}, \quad 0 \le \theta < \infty,\ x \in [0,\infty), \\
\phi^{-1}(x,\theta) &= \log\left\{ \frac{e^{-\theta x}-1}{e^{-\theta}-1} \right\},\ 0 \le \theta < \infty,\ x \in [0,1], \\
C_\theta(u_1,\ldots,u_d) &= -\frac{1}{\theta}\log\left[1 + \frac{\prod_{j=1}^{d}\{\exp(-\theta u_j)-1\}}{\{\exp(-\theta)-1\}^{d-1}} \right], \ u_1,\ldots,u_d \in [0,1].
\end{aligned}$$

The dependence becomes maximal when θ tends to infinity and independence is achieved when $\theta = 0$.

The level plots of the bivariate copula-based densities with t_3 and normal margins are given in Figure 4.

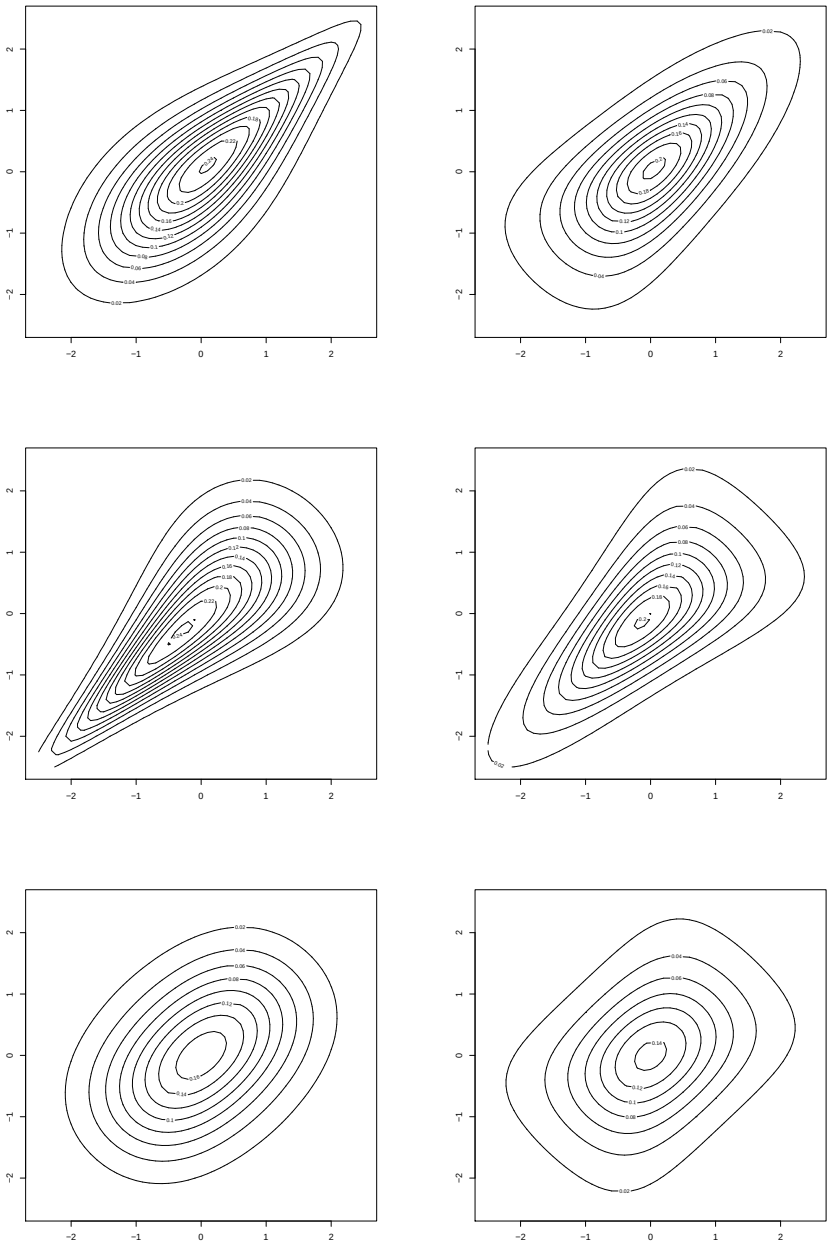

Figure 4: Contour diagrams for (from top to bottom) Gumbel, Clayton and Frank copula with Normal (left column) and t_3 distributed (right column) margins.

3.4 Hierarchical Archimedean Copulae

A recently developed flexible method is provided by hierarchical Archimedean copulae (HAC). The special, so called partially nested, case of HAC:

$$\begin{aligned} C(u_1,\ldots,u_d) &= C_0\{C_1(u_1,\ldots,u_{k_1}),\ldots,C_m(u_{k_{m-1}+1},\ldots,u_d)\} \\ &= \phi_0\left[\sum_{p=1}^{m}\phi_0^{-1}\circ\phi_i\left\{\sum_{j=k_{p-1}+1}^{k_p}\phi_p^{-1}(u_j)\right\}\right] \end{aligned} \tag{6}$$

for $\phi_0^{-1}\circ\phi_p \in \{w : [0;\infty) \to [0;\infty) | w(0) = 0;\, w(\infty) = \infty;\, (-1)^{j-1}w^{(j)} \geq 0;\, j = 1,\ldots,\infty\}$, $p = 1,\ldots,m$, with $k_0 = 1$. In contrast to the Archimedean copula, HAC defines the whole dependency structure in a recursive way. At the lowest level the dependency between the first two variables is modelled by a copula function with the generator ϕ_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 an another copula function is used to model the dependency between z_1 and u_3, etc. Note, that the generators ϕ_i can come from the same family and differ only through the parameter or, to introduce more flexibility, come from different generator families. As an alternative to the fully nested model, we can consider copula functions, with arbitrarily chosen combinations at each copula level. Okhrin, Okhrin and Schmid (2008) provide several methodologies of determining the structure of the HAC from the data, Okhrin, Okhrin and Schmid (2009) provide necessary theoretical properties of HAC, there are also several empirical papers on the application HAC to CDO (see Choros, Härdle and Okhrin (2009)) and to weather data (see Filler, Odening, Okhrin and Xu (2010)).

4 Simulation Techniques

To investigate the properties of some multivariate distributions, one needs the algorithms of the simulations because many of those properties are to be checked by Monte Carlo techniques. In this section we provide different methods of sampling from copula.

4.1 Conditional Inverse Method

The conditional inverse method is a general approach for the simulation of random variables from an arbitrary multivariate distribution. This method can be also used to simulate from copulae. The idea is to generate random variables recursively from the conditional distributions. To sample $U_1,\ldots,U_d$ from copula C we proceed with the following steps

1. sample $V_1,\ldots,V_d$ from $U(0,1)$;
2. $U_1 = V_1$;

3. $U_j = C_j^{-1}(V_j|U_1, \ldots, U_{j-1})$ for $j = 2, \ldots, d$ where the conditional distribution of U_j is given by

$$\begin{aligned} C_j(u_j|u_1, \ldots, u_{j-1}) &= P(U_j \leq u_j | U_1 = u_1 \ldots U_{j-1} = u_{j-1}) \\ &= \frac{\frac{\partial^{j-1} C_j(u_1, \ldots, u_j)}{\partial u_1 \ldots \partial u_{j-1}}}{\frac{\partial^{j-1} C_{j-1}(u_1, \ldots, u_{j-1})}{\partial u_1 \ldots \partial u_{j-1}}} \end{aligned} \tag{7}$$

with $C_j = C(u_1, \ldots, u_j, 1, \ldots, 1) = C(u_1, \ldots, u_j)$.

The approach is numerically expensive, due to high order derivatives of C and the calculation of the inverse of the conditional distribution function.

4.2 Marshall and Olkin (1988) Method

To simulate from Archimedean copulae a simpler method was introduced in Marshall and Olkin (1988). The idea of the method is based on the fact that Archimedean copulae are derived from Laplace transforms (4). Following Marshall and Olkin (1988) we proceed with the following three steps procedure:

1. sample U from $M = \mathcal{LS}^{-1}(\phi)$;
2. sample independent $(V_1, \ldots, V_d) \sim U[0, 1]$;
3. $U_j = \phi\{-\ln(V_j)/U\}$ for $j = 1, \ldots, d$.

This method works much faster than the classic conditional inverse technique. The drawback is that the distribution M can only be determined explicitly for a few generator functions ϕ. For example for Gumbel copula $M(\theta) = St(1/\theta, 1, [cos\{\pi/(2\theta)\}]^\theta)$ and for Clayton copula $M(\theta) = \Gamma(1/\theta, 1)$.

4.3 McNeil (2008) Method

Methods of simulation from the different HAC structures were proposed in McNeil (2008); this is an extension of the Marshall and Olkin (1988) method. Below is the algorithm for partially nested copulae (6)

1. sample U from $M = LS^{-1}(\phi_0)$;
2. for $i = 1, \ldots, m$ sample

   $$V_{k_{p-1}+1}, \ldots, V_{k_p} \text{ from } C[u_{k_{p-1}+1}, \ldots, u_{k_p}; \exp\{-U\phi_0^{-1} \circ \phi_p(\cdot)\}]$$

   using Marshall and Olkin (1988) method where

   $$C[u_{k_{p-1}+1}, \ldots, u_{k_p}; \exp\{-U\phi_0^{-1} \circ \phi_p(\cdot)\}]$$

   is the simple Archimedean copula with the generator function given by $\exp\{-U\phi_0^{-1} \circ \phi_p(\cdot)\}$;

3. $(U_{k_{p-1}+1}, \ldots, U_{k_p})^\top = \phi_0[-\log\{(V_{k_{p-1}+1}, \ldots, V_{k_p})^\top\}/U],\ p = 1, \ldots, m.$

This method, however also has some drawbacks because the inverse Laplace transform of the composition of the generator function does not always have an explicit form. Nevertheless, McNeil (2008) provides a list of combinations, which enable this.

5 Estimation

For a given data-set one needs to find an appropriate model, and to estimate the parameter when the model is fixed. In this section we describe different methods of the estimation of the copula from the data. All methods are similar and are based on the equation (2). Having the sample $X_{ij},\ i = 1, \ldots, n,\ j = 1, \ldots, d$ one needs to estimate the copula. To estimate the marginal distributions $\hat{F}_j(\cdot),\ j = 1, \ldots, d$ at least three possible methods are available. The most simple one is to use the empirical distribution function

$$\hat{F}_j(x) = \frac{1}{n+1} \sum_{i=1}^{n} \mathbf{I}\{X_{ij} \leq x\}.$$

The change of the fraction before the sum from the classical $\frac{1}{n}$ to $\frac{1}{n+1}$ is made to bound the empirical distribution from 1; otherwise this causes problems in the maximum likelihood (ML) calculation. The inverse function of $\hat{F}_j(x)$ is then an empirical quantile. Instead of this simplest empirical estimation one can smooth the distribution function by using a kernel method, see Härdle and Linton (1994). Using kernel function $\kappa : \mathbb{R} \to \mathbb{R},\ \int \kappa = 1$ with the bandwidth $h > 0$ one gets following estimator

$$\tilde{F}_j(x) = \frac{1}{n+1} \sum_{i=1}^{n} K\left(\frac{x - X_{ij}}{h}\right),$$

with $K(x) = \int_{-\infty}^{x} \kappa(t)dt$. Apart from nonparametric methods, there is also a parametric method that is based on the assumption of a parametric form of the marginal distribution $F_j(x, \hat{\alpha}_j)$, where α_j is the parameter of the distribution, and $\hat{\alpha}_j$ is its estimator based on the ML method or method of moments. The last case considers the full knowledge of the true marginal distribution $F_j(x)$, which is rare in practice.

In the same way, there are four possible choices of the copula function. Let us first determine general margins $\breve{F}_j(x)$ that could be one of $\hat{F}_j(x),\ \tilde{F}_j(x),\ F_j(x, \hat{\alpha})$ or $F_j(x)$. The empirical copula is then defined as

$$\widehat{C}(u_1, \ldots, u_d) = \frac{1}{n} \sum_{i=1}^{n} \prod_{j=1}^{d} \mathbf{I}\{\breve{F}_j(X_{ij}) \leq u_j\}. \tag{8}$$

Let $K_j,\ j = 1, \ldots, d$ be the same symmetric kernel for each direction as in the estimation of marginal distributions, and let $h_j,\ j = 1, \ldots, d$ be the set of bandwidths, then the kernel based copula estimation considered in Fermanian and Scaillet (2003) is

$$\widetilde{C}(u_1, \ldots, u_d) = \frac{1}{n} \sum_{i=1}^{n} \prod_{j=1}^{d} K_j\left\{\frac{u_j - \breve{F}_j(X_{ij})}{h_j}\right\}. \tag{9}$$

In the bivariate case ($d = 2$) to avoid boundary bias, one uses (Chen and Huang, 2007) local linear kernel to smooth at $u \in [0, 1]$

$$K_{uh} = \frac{K(x)\{a_2(u,h) - a_1(u,h)x\}}{a_0(u,h)a_2(u,h) - a_1^2(u,h)},$$

with $a_\ell(u,h) = \int_{(u-1)/h}^{u/h} t^\ell K(t)dt,\ \ell = 0, 1, 2$ and $h > 0$ (see Lejeune and Sarda (1992), Jones (1993)). Let $G_{uh}(t) = \int_{-\infty}^{t} K_{uh}(x)dx$ and $T_{uh} = G_{uh}\{(u-1)/h\}$, then an unbiased kernel based estimator of the bivariate copula is given by

$$\begin{aligned} \widetilde{C}(u_1, u_2) &= \frac{1}{n} G_{u_1 h}\left\{\frac{u_1 - \breve{F}_1(X_{i1})}{h}\right\} G_{u_2 h}\left\{\frac{u_2 - \breve{F}_2(X_{i2})}{h}\right\} \\ &- (u_1 T_{u_2 h} + u_2 T_{u_1 h} + T_{u_1 h} T_{u_2 h}). \end{aligned} \tag{10}$$

The last situation is the parametric copula $C(\mathbf{u}, \theta)$, where the copula comes from some fixed family. In this case the parameter of the copula function is estimated using the ML method. From (3) the likelihood function for the case $\breve{F}_j(x) = F_j(x, \alpha_j),\ j = 1, \ldots, d$ is

$$L(\theta, \alpha_1, \ldots, \alpha_d) = \prod_{i=1}^{n} f(X_{i1}, \ldots, X_{id}; \alpha_1, \ldots, \alpha_d, \theta)$$

and the log-likelihood function is given by

$$\begin{aligned} \ell(\theta, \alpha_1, \ldots, \alpha_d) &= \sum_{i=1}^{n} \log c\{F_1(X_{i1}; \alpha_1), \ldots, F_d(X_{id}; \alpha_d); \theta\} \\ &+ \sum_{i=1}^{n} \sum_{j=1}^{d} \log f_j(X_{ij}; \alpha_j), \end{aligned}$$

where $f_j(\cdot)$ are marginal densities. All parameters $\{\theta, \alpha_1, \ldots, \alpha_d\}$ can be estimated in one or two steps. For practical applications, however, a two step estimation procedure is more efficient. A one step procedure, also called *full maximum likelihood*, is carried out by maximising likelihood function simultaneously over all parameters, thus by solving

$$(\partial\ell/\partial\alpha_1, \ldots, \partial\ell/\partial\alpha_d, \partial\ell/\partial\theta) = \mathbf{0},$$

with respect to $(\theta, \alpha_1, \ldots, \alpha_d)$. Following the standard theory on ML estimation estimators are efficient and asymptotically normal. However, it is often computationally demanding to solve the system simultaneously.

The two step procedure can be done for any kind of marginal distribution $\breve{F}_j(x) \in \{\widehat{F}_j(x), \widetilde{F}_j(x), F_j(x, \hat{\alpha})\}$. Firstly, we estimate the marginal distribution by using any of the above methods and secondly, we estimate the copula parameter by the pseudo log-likelihood function

$$\ell_p(\theta) = \sum_{i=1}^{n} \log c\{\breve{F}_1(X_{i1}), \ldots, \breve{F}_d(X_{id}); \theta\}.$$

The solution is then

$$\hat{\theta} = \arg\max_{\theta} \ell_p(\theta).$$

If the marginal distributions are from parametric families $\breve{F}_j(x) = F_j(x, \hat{\alpha}_j),\ j = 1, \ldots, d$, then the method is called *inference for margins*. Otherwise, if margins, are nonparametrically estimated $\breve{F}_j(x) \in \{\widehat{F}_j(x), \widetilde{F}_j(x)\},\ j = 1, \ldots, d$, then the method is called *canonical maximum likelihood* method.

6 Goodness-of-Fit (GoF) Tests

After the copula is estimated, one needs to test how well the estimated copula describes the sample. Nonparametric copula is certainly the best choice for this, and is usually considered the benchmark in many tests. With the GoF tests one checks whether the underlying copula belongs to any copula family. The test problem could be written as a composite null hypothesis

$$H_0: \ C \in \mathcal{C}_0, \quad \text{against} \quad H_1: \ C \notin \mathcal{C}_0,$$

where $\mathcal{C}_0 = \{C_\theta : \theta \in \Theta\}$ is a known parametric family of copulae. In some cases we restrict ourselves to the one element family $\mathcal{C}_0 = C_0$, thus the hypothesis in this case in the simple one. The test problem is, in general, equivalent to the GoF tests for multivariate distributions. However, since the margins are estimated we cannot apply the standard test procedures directly.

Here we consider several methodologies recently introduced in the literature. We can categorised them into three classes: tests based on the empirical copula, tests based on the Kendall's process and tests based on Rosenblatt's transform.

6.1 Tests based on the empirical copula

These tests are based directly on the distance between C and C_0. Naturally, as C is unknown one takes the empirical copula which is fully nonparametric $\hat{C}$ or $\widetilde{C}$ instead. The estimated copula C_0, that should be tested, is the parametric one $C(\cdot, \hat{\theta})$. Two statistics considered in the literature (see e.g Fermanian (2005), Genest and Rémillard (2008), etc.) are similar to Crámer-von Mises and Kolmogorov-Smirnov test statistics

$$\begin{aligned} S &= n \int_{[0,1]^d} \{\widehat{C}(u_1, \ldots, u_d) - C(u_1, \ldots, u_d, \widehat{\theta})\}^2 \, d\widehat{C}(u_1, \ldots, u_d), \\ T &= \sup_{u_1, \ldots, u_d \in [0,1]} \sqrt{n} |\widehat{C}(u_1, \ldots, u_d) - C(u_1, \ldots, u_d, \widehat{\theta})|. \end{aligned}$$

Genest and Rémillard (2008) show the convergence of $\sqrt{n}\{\widehat{C}(u_1, \ldots, u_d) - C(u_1, \ldots, u_d, \widehat{\theta})\}$ in distribution, they also show that tests based on S and T are consistent. In actual fact, the p-values of the test statistics depends on this limiting distribution and in practice p-values are calculated using the bootstrap methods described in Genest and Rémillard (2008). This is quite expensive numerically, but leads to proper results.

6.2 Tests based on Kendall's process

Genest and Rivest (1993), Wang and Wells (2000) and Barbe, Genest, Ghoudi and Rémillard (1996) consider a test based on the true and empirical distributions of the pseudo random variable $V = C(U_1, \ldots, U_d) \sim K$. The expectation of v is the transformation of the multivariate extension of Kendall's τ, hence the deviation of the true K and empirical $\hat{K}$ as a univariate function is called Kendall's process. The most natural empirical estimation of K is

$$\hat{K}(v) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{I}\{V_i \leq v\}.$$

The theoretical form of the K was discussed in Barbe et al. (1996), Okhrin et al. (2009) for different copula functions. In the bivariate case of the Archimedean copulae it is related to the generator function as

$$K(v, \theta) = v - \frac{\phi_\theta^{-1}(v)}{\{\phi_\theta^{-1}(v)\}'}.$$

As in the tests based on the empirical copulae Wang and Wells (2000) and Genest et al. (2006) propose to compute a Kolmogorov-Smirnov and Crámer-von-Mises statistics for the K

$$\begin{aligned} S_K &= n \int_0^1 \{\hat{K}(v) - K(v, \theta)\}^2 \, dv, \\ T_K &= \sup_{v \in [0,1]} |\hat{K}(v) - K(v, \theta)|, \end{aligned}$$

where $\hat{K}(v)$ and $K(v, \theta)$ are empirical and theoretical K-distributions of the variable $v = C(u_1, \ldots, u_d)$. However, as in the previous tests, exact p-values for this statistic cannot be computed explicitly. Savu and Trede (2004) propose a χ^2-test based on the K-distribution. Unfortunately, in most cases the distribution of the test statistic does not follow a standard distribution and either a bootstrap or another computationally intensive methods should be used.

6.3 Tests based on Rosenblatt's process

An alternative global approach is based on the probability integral transform introduced in Rosenblatt (1952) and applied in Breymann et al. (2003), Chen et al. (2004) and Dobrić and Schmid (2007). The idea of the transformation is to construct the variables

$$\begin{aligned} Y_{i1} &= \breve{F}_1(X_{i1}), \\ Y_{ij} &= C\{\breve{F}_j(X_{ij}) | \breve{F}_1(X_{i1}), \ldots, \breve{F}_{j-1}(X_{i,j-1})\}, \quad \text{for} \quad j = 2, \ldots, d, \end{aligned} \tag{11}$$

where the conditional copula is defined in (7). Under H_0 the variables Y_{ij}, for $j = 1, \ldots, d$ are independently and uniformly distributed on the interval $[0, 1]$. Here we discuss the

second test based on Y_{ij} proposed in Chen et al. (2004). Consider the variable $W_i = \sum_{j=1}^{d}[\Phi^{-1}(Y_{ij})]^2$. Under H_0 it holds that $W_i \sim \chi^2_d$. Breymann et al. (2003) assume that estimating margins and copula parameters does not significantly affect the distribution of $\hat{W}_i$ and apply a standard χ^2 test directly to the pseudo-observations. Chen et al. (2004) developed a kernel-based test for the distribution of W and, thus, an account for estimation errors. Let $\tilde{g}_W(w)$ denote the kernel estimator of the density of W. Under H_0 the density $g_W(w)$ is equal to one, as the density of the uniform distribution. As a measure of divergency Chen et al. (2004) used $\hat{J}_n = \int_0^1 \{\tilde{g}_W(w) - 1\}^2 dw$. Assuming non-parametric estimator of the marginal distributions Chen et al. (2004) prove under regularity conditions that

$$T_n = (n\sqrt{h}\hat{J}_n - c_n)/\sigma \to N(0,1),$$

where the normalisation parameters h, c_n and σ are defined in Chen et al. (2004). The proof of this statement does not depend explicitly on the type of the non-parametric estimator of the marginals $\breve{F}_j$, but uses the order of $\breve{F}_j(X_{ij}) - F_j(X_{ij})$ as a function of n. It can be shown that if the parametric families of marginal distributions are correctly specified and their parameters are consistently estimated, then the statement also holds if we use parametric estimators for marginal distributions.

7 Simulation Study

A Monte Carlo experiment has been provided to discuss the finite sample properties of the goodness-of-fit tests based on the empirical copula and different estimation techniques on the simulated data. We restrict ourselves to the three dimensional case of three copula families, namely Gaussian, simple AC with Gumbel generator and HAC with Gumbel generator. For the simulation from the AC we use the Marshall and Olkin (1988) method and for simulation from HAC the McNeil (2008) method. To simulate from the Gaussian copula we simulate first from normal distribution and then apply the Sklar's theorem (1).

The main characteristic of interest in this study is to see whether the tests are able to maintain their nominal level fixed at $\alpha = 0.1$ and to see the power of the tests under the variety of alternatives. This is the only study that discusses the power of goodness-of-fit tests for copula in dimensions higher that $d = 2$. We consider all possible copulae with parameters $\tau \in \{0.25, 0.5, 0.75\}$. This means that under consideration were three AC: $C_{\theta(0.25)}(\cdot)$, $C_{\theta(0.5)}(\cdot)$, $C_{\theta(0.75)}(\cdot)$, three HAC: $C_{\theta(0.25)}\{C_{\theta(0.50)}(u_1, u_2), u_3\}$, $C_{\theta(0.25)}\{C_{\theta(0.75)}(u_1, u_2), u_3\}$, $C_{\theta(0.75)}\{C_{\theta(0.50)}(u_1, u_2), u_3\}$, and 15 Gaussian copulae with all possible positive definite correlation matrices containing values $\rho \in \{0.25, 0.5, 0.75\}$. Here $\theta(\tau)$ converts Kendall's τ correlation coefficient into a corresponding copula parameter.

The results are provided in Table 1 for AC, in Table 2 for HAC and in Table 4 for Gaussian copulae. To save the workspace we provide results for only 3 Gaussian copulae out of 15 with the largest difference between parameters. For HAC, a vector function $\boldsymbol{\theta}(\tau_1, \tau_2)$ converts two Kendall's τ into HAC copula parameters. If $\tau_1 < \tau_2$ then copula $C_{\theta(\tau_1)}\{C_{\theta(\tau_2)}(u_1, u_2), u_3\}$ is considered. For Gaussian copula

$$\boldsymbol{\Sigma}(\tau_1, \tau_2, \tau_3) = \begin{pmatrix} 1 & \tau_1 & \tau_2 \\ \tau_1 & 1 & \tau_3 \\ \tau_2 & \tau_3 & 1 \end{pmatrix}.$$

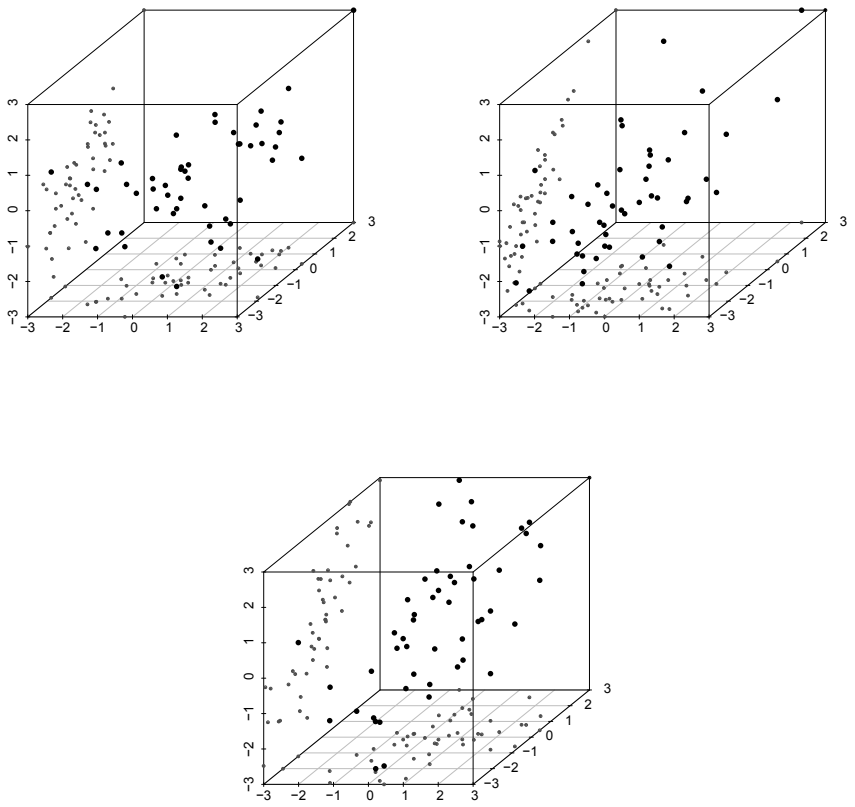

Figure 5: Samples of size $n = 50$ from $C_{0.25}(\cdot)$, $C_{\theta(0.75)}\{C_{\theta(0.50)}(u_1, u_2), u_3\}$ and Gaussian copula with upper diagonal elements of the correlation matrix given by $\rho = (0.25, 0.25, 0.75)^\top$

From each copula we simulate a sample of $n = 50$ or $n = 150$ observations with standard normal margins. The margins are then estimated parametrically (normal distribution with estimated mean and variance) or nonparametrically. Respective columns in the tables are marked by "*par.*" and "*emp.*". For each sample we estimate the AC using inference for the margins method, HAC using Okhrin et al. (2008) and the Gaussian copula using the generalised method of moments. Then we test how good these distributions fit the sample. The empirical copula for both tests has been calculated as in (8). Number of bootstrap steps provided for the tests is equal to $N = 1000$. To sum up the simulation procedure, we used

1. F : two methods of estimation of margins (parametric and nonparametric);
2. C_0 : hypothesised copula models under H_0 (three models);
3. C : copula model from which the data were generated (three models with 3, 3 and 15 levels of dependence respectively);
4. n : size of each sample drawn from C (two possibilities, $n = 50$ and $n = 150$).

Thus, for all these $2 \times 3 \times (3 + 3 + 15) \times 2 = 252$ situations we perform 100 repetitions in order to calculate the power of both tests. This study is hardly comparable to other similar studies, because, as far as we know, this is the only one that considers the three dimensional case, and the only one that considers a hierarchical Archimedean copulae.

To understand the numbers in the tables more deeply let us consider first the value in Table 1. The number 0.88 says, that testing using Kolmogorov-Smirnov type statistic T_n for the AC with $\tau = 0.25$ from the sample of a size $n = 50$, with nonparametrically estimated margins, rejects the null hypotheses H_0, assuming that the data are from HAC, in $100\% - 88\% = 12\%$ of chances. It is very natural that the rejection rate for the AC, that have HAC under H_0, is very close to the case, where AC is under H_0. In general AC is a special case of HAC. If the true distribution is AC, the rejection rates should be equal, or close to each other, and the difference based only on the estimation error.

Figure 6 represents the level of both goodness-of-fit tests for different sizes in terms of three quartiles; the outliers are marked with closed dots. In general, values lies below 0.1, which implies that the bootstrap performs well. Increasing the number of runs improves this graph. We see that if the sample size has enlarged three times, then the tests have approximately doubled their power in S statistics, and a slightly smaller coefficient is given for the T statistics. In general, small size samples from different models look very similar (see Figure 5), this makes detection of the model that best fits the data hardly applicable, this also explains a lot of outliers in Figure 6.

From the tables we see, that S_n performs, on average, better than T_n statistics, this can be also seen from the Figure 6. In the tables, rejection rates for S_n under false H_0 are in general higher, than for T_n statistics. We can also conclude that the larger the difference between parameters of the model is the faster AC is rejected. This can be expressed by the only parameter in AC that does not covers the whole dependency.

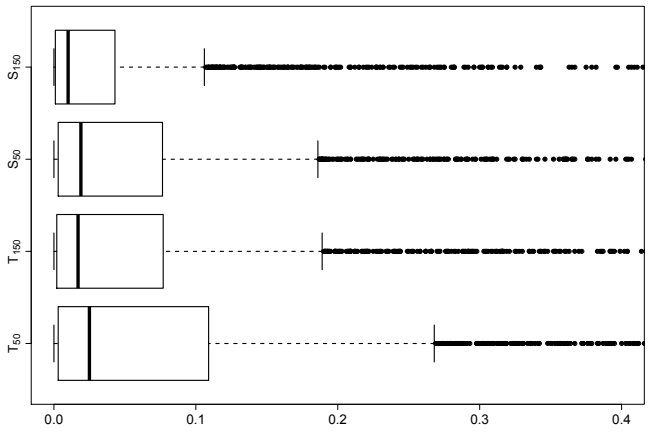

Figure 6: Levels of goodness-of-fit tests for different sample size, for parametric margins.

AC								
θ		$n = 50$				$n = 150$			
		T		S		T		S	
		emp.	par.	emp.	par.	emp.	par.	emp.	par.
$\theta(0.25)$	HAC	0.88	0.51	0.83	0.38	0.93	0.36	0.90	0.35
	AC	0.88	0.51	0.89	0.50	0.95	0.32	0.90	0.34
	Gauss	0.71	0.29	0.56	0.22	0.69	0.11	0.43	0.08
$\theta(0.5)$	HAC	0.90	0.38	0.94	0.30	0.87	0.35	0.88	0.27
	AC	0.96	0.55	0.95	0.45	0.90	0.45	0.92	0.35
	Gauss	0.76	0.30	0.65	0.19	0.47	0.13	0.31	0.02
$\theta(0.75)$	HAC	0.93	0.29	0.93	0.15	0.89	0.27	0.89	0.10
	AC	0.93	0.29	0.93	0.22	0.90	0.25	0.91	0.13
	Gauss	0.77	0.19	0.65	0.10	0.57	0.11	0.24	0.05

Table 1: Non-rejection rate of the different models, where the sample is drawn from the simple AC

HAC								
$\boldsymbol{\theta}$		$n = 50$				$n = 150$			
		T		S		T		S	
		emp.	par.	emp.	par.	emp.	par.	emp.	par.
$\theta(0.25, 0.5)$	HAC	0.88	0.29	0.90	0.24	0.96	0.31	0.92	0.26
	AC	0.91	0.26	0.93	0.36	0.54	0.13	0.53	0.07
	Gauss	0.82	0.20	0.69	0.19	0.57	0.14	0.37	0.04
$\theta(0.25, 0.75)$	HAC	0.93	0.21	0.92	0.13	0.88	0.18	0.88	0.09
	AC	0.46	0.14	0.54	0.07	0.00	0.00	0.00	0.00
	Gauss	0.84	0.19	0.71	0.13	0.52	0.10	0.42	0.01
$\theta(0.5, 0.75)$	HAC	0.86	0.31	0.87	0.18	0.91	0.20	0.94	0.08
	AC	0.89	0.36	0.92	0.28	0.44	0.04	0.47	0.02
	Gauss	0.70	0.19	0.55	0.12	0.50	0.11	0.30	0.05

Table 2: Non-rejection rate of the different models, where the sample is drawn from the HAC

Gauss								
$\boldsymbol{\Sigma}$	$n = 50$				$n = 150$			
	T		S		T		S	
	emp.	par.	emp.	par.	emp.	par.	emp.	par.
HAC	0.89	0.20	0.93	0.11	0.78	0.08	0.81	0.02
$\boldsymbol{\Sigma}(0.25, 0.25, 0.75)$ AC	0.43	0.13	0.47	0.09	0.00	0.00	0.00	0.00
Gauss	0.88	0.22	0.89	0.12	0.87	0.11	0.86	0.03
HAC	0.92	0.20	0.91	0.14	0.76	0.07	0.69	0.04
$\boldsymbol{\Sigma}(0.25, 0.75, 0.25)$ AC	0.39	0.12	0.39	0.04	0.00	0.00	0.00	0.00
Gauss	0.90	0.18	0.87	0.13	0.92	0.12	0.94	0.10
HAC	0.89	0.30	0.93	0.16	0.78	0.10	0.75	0.04
$\boldsymbol{\Sigma}(0.75, 0.25, 0.25)$ AC	0.51	0.16	0.46	0.07	0.00	0.00	0.00	0.00
Gauss	0.91	0.28	0.90	0.17	0.88	0.13	0.86	0.06

Table 4: Non-rejection rate of the different models, where the sample is drawn from the Gaussian copula

8 Empirical Results

The empirical part of this study is based on the calculation of the Value-at-Risk for the Profit and Loss function of the portfolio containing three assets. Asset returns follow some GARCH-type process with residuals from copula based models. We consider the daily stock prices of three American banks, namely Bank of America, Citigroup and Santander from 29.09.2000 to 16.02.2001. This results in $T = 100$ observations being consistent with the simulation study provided above. We take this time interval because several U.S. banks have recorded strong earnings in the fourth quarter of 2000. Rising profits were reported by U.S. industry leaders, namely Citigroup and Bank of America. At the same time bad forecasts for technology companies were reported; these influence the financial sector as well. Prices $\{X_{tj}\}$, $j = 1, 2, 3$ behave (over the chosen period) as in Figure 7. Assuming the log-returns $R_{tj} = \log(X_{tj}/X_{t-1,j})$, $j = 1, 2, 3$, $t = 1, \ldots, T$ (see Figure 8) follow an ARMA(1,1)-GARCH(1,1) process, we have

$$R_{tj} = \mu_j + \gamma_j R_{t-1,j} + \zeta_j \sigma_{t-1,j} \varepsilon_{t-1,j} + \sigma_{tj} \varepsilon_{tj},$$

where

$$\sigma_{tj}^2 = \omega_j + \alpha_j \sigma_{t-1,j}^2 + \beta_j \sigma_{t-1,j}^2 \varepsilon_{t-1,j}^2$$

and $\omega > 0$, $\alpha_j \geq 0$, $\beta_j \geq 0$, $\alpha_j + \beta_j < 1$, $|\zeta| < 1$.

The fit of an ARMA(1,1)-GARCH(1,1) model to the log returns $\mathbf{R}_t = (R_{t1}, R_{t2}, R_{t3})^\top$, $T = 100$, gives the estimates $\hat{\omega}_j$, $\hat{\alpha}_j$, $\hat{\beta}_j$, $\hat{\zeta}_j$ and $\hat{\gamma}_j$, as in Table 5. Empirical residuals $\{\hat{\varepsilon}_t\}_{t=1}^T$, where $\hat{\varepsilon}_t = (\hat{\varepsilon}_{t1}, \hat{\varepsilon}_{t2}, \hat{\varepsilon}_{t3})^\top$ are assumed to be normally distributed; this is not rejected by the Kolmogorov-Smirnov test at the high level of significance for all three banks. Residuals are also assumed to be independent, because of the Box-Ljung auto-correlation test with lag 12. Thus, in the estimation of copula we use an inference for margins method, where margins are normal, thus, estimated parametrically.

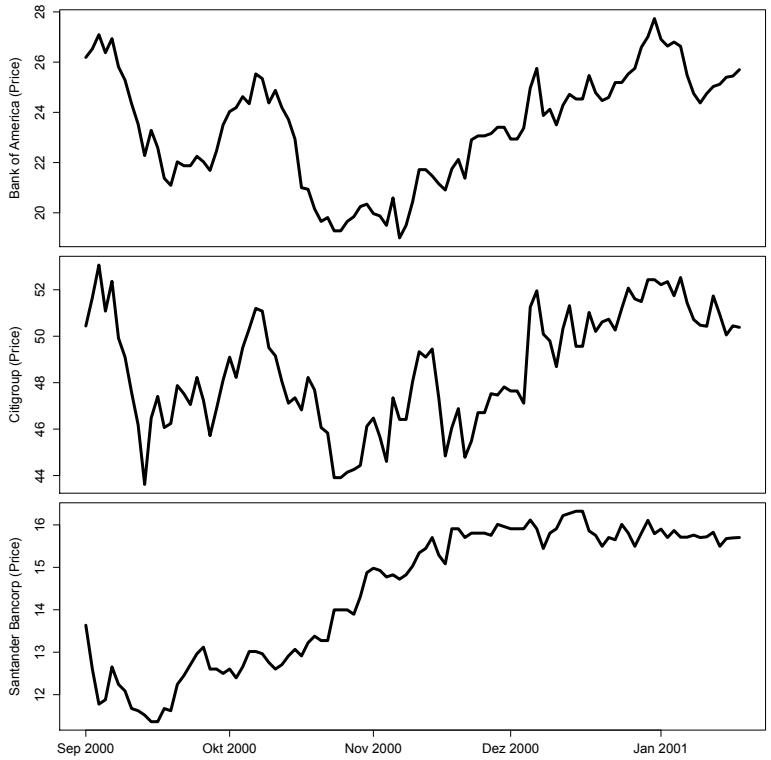

Figure 7: Stock prices for Bank of America, Citigroup and Santander (from top to bottom).

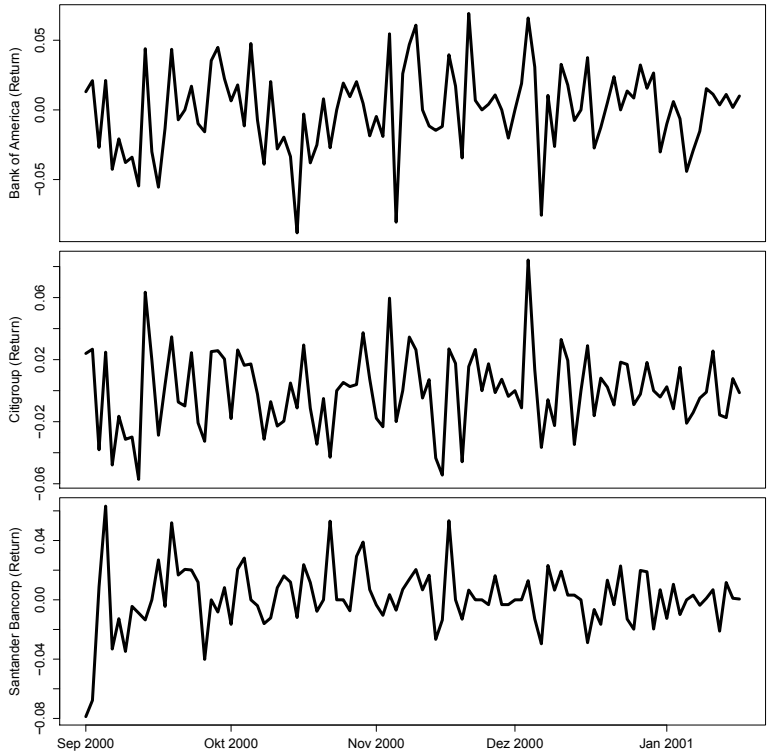

Figure 8: Log-returns for Bank of America, Citigroup and Santander (from top to bottom).

	$\hat{\mu}_j$	$\hat{\gamma}_j$	$\hat{\zeta}_j$	$\hat{\omega}_j$	$\hat{\alpha}_j$	$\hat{\beta}_j$	BL	KS
Bank of America	1.879e-03	0.226	-0.232	3.465e-04	0.551	0.170	0.567	0.829
	(2.598e-03)	(0.642)	(0.654)	(1.369e-04)	(0.284)	(0.155)		
Citigroup	0.116e-03	0.305	-0.455	2.669e-04	0.096	0.471	0.569	0.786
	(1.487e-03)	(0.296)	(0.288)	(5.533e-04)	(0.165)	(1.008)		
Santander	1.359e-03	0.430	-0.566	4.512e-10	0.012	0.979	0.914	0.781
	(0.908e-03)	(0.149)	(0.174)	(1.376e-05)	(0.018)	(0.049)		

Table 5: Fitting of univariate ARMA(1,1)-GARCH(1,1) to asset returns. The standard deviation of the parameters, which are quiet big because of the small sample size, are given in parentheses. The last two columns provide the p-values of the Box-Ljung test (BL) for autocorrelations and Kolmogorov-Smirnov test (KS) for testing of normality of the residuals.

	T_{100}	S_{100}	estimates
HAC	0.3191	0.1237	$C\{C(u_1, u_2; 1.996), u_3; 1.256\}$
AC	0.0012	0.0002	$C(u_1, u_2, u_3; 1.276)$
Gauss	0.0160	0.0078	$C_N\{u_1, u_2, u_3; \mathbf{\Sigma}(0.697, 0.215, 0.312)\}$

Table 6: p-values of both GoFs and estimates of the models under different H_0 hypotheses.

Upper diagonal cells of Figure 9 represent pair wise scatterplots of ARMA-GARCH residuals. In the lower diagonal cells of the same figure we show the scatterplots of the residuals mapped on the unit square by the estimated marginal cdf, $\hat{F}(\hat{\varepsilon})$.

We estimated three different models, namely simple AC, HAC and Gaussian copula. Afterwards two tests, used in the simulation study, were applied to see how good these models describe data. In this case the number of bootstrap runs has been increased to $N = 10000$ to make the test results more precise. Estimated models, and p-values are represented in Table 6. We see that parameters in the HAC model deviate from each other, we may conclude therefore, that a simple AC is not a proper model that fits the data. On the other hand, from Figure 9 we see that the points are not elliptical; this convinces us to expect a low p-value of the test where the Gaussian copula is under H_0. In the first two columns of Table 6 we put p-values for all tests. We conclude that HAC is the most appropriate model for this particular dataset, because it has the largest p-value. Based on two tests only HAC can not be rejected under significance level $\alpha = 0.05$. This means that our data may not be described by the simple three-dimensional normal distribution, but the margins are still normal.

To see if knowledge of preferable distribution is worth knowing in a financial problem, we estimate the Vale-at-Risk from a Profit and Loss of a linear portfolio using copulae. The portfolio is composed of the stocks discussed above. We also perform an evaluation of the estimators through backtesting. Let w be the portfolio, which is represented by the number of assets for a specified stock in the portfolio, $w = \{w_1, \ldots, w_d\},\ w_i \in \mathbb{Z}$. The value V_t of the portfolio w is given by

$$V_t = \sum_{j=1}^{d} w_j X_{tj} \tag{12}$$

Figure 9: Scatterplots from ARMA-GARCH residuals (upper triangular) and from residuals mapped on unit square by the cdf (lower triangular).

and the random variable defined as the absolute change in the portfolio

$$L_{t+1} = (V_{t+1} - V_t) = \sum_{j=1}^{d} w_j X_{tj} \left\{\exp(R_{t+1,j}) - 1\right\} \tag{13}$$

also called profit and loss (P&L) function, expresses the absolute change in the portfolio value in one period. The distribution function of L, dropping the time index, is given by

$$F_L(x) = P(L \leq x). \tag{14}$$

As usual the Value-at-Risk at level α from a portfolio w is defined as the α-quantile from F_L:

$$\text{VaR}(\alpha) = F_L^{-1}(\alpha). \tag{15}$$

It follows from (14) that F_L depends on the d-dimensional distribution of log-returns F_X. In general, the loss distribution F_L depends on a random process representing the risk factors influencing the P&L from a portfolio. In the present case log-returns modelled by an ARMA(1,1)-GARCH(1,1) model are a suitable risk factor choice. Thus, modelling their distribution is essential to obtain the quantiles from F_L. To estimate the VaR we simulate samples of residuals ε_t from HAC, AC and Gaussian copula with normal margins, then apply simulated residuals to the estimated ARMA(1,1)-GARCH(1,1) model and calculate it based on the values of the Profit and Loss $\hat{L}$ with $w = (1, 1, 1)^\top$. The $\widehat{VaR}(\alpha)$ is then an empirical α-quantile from the $\hat{L}$. In Figure 10 we represent the series of estimated Value-at-Risk with $\alpha = 0.1$ and the P&L function. Afterwards backtesting is used to

α	$\hat{\alpha}_{HAC}$	$\hat{\alpha}_{AC}$	$\hat{\alpha}_{Gauss}$
0.10	0.091	0.122	0.081
0.05	0.040	0.061	0.031
0.01	0.000	0.010	0.000

Table 7: Backtesting for the estimation of VaR under different alternatives.

evaluate the performance of the specified copula family $\mathcal{C}$. The estimated values for the VaR are compared with the true realisations $\{L_t\}$ of the P&L function, an exceedance occuring for each L_t smaller than $\widehat{VaR}_t(\alpha)$. The ratio of the number of exceedances to the number of observations gives the exceedances ratio $\hat{\alpha}$:

$$\hat{\alpha} = \frac{1}{T}\sum_{t=1}^{T} \mathbf{I}\{L_t < \widehat{VaR}_t(\alpha)\}.$$

The backtesting results are provided in Table 7. From them we see that the Gaussian copula usually underestimates the VaR. This is natural because this copula does not have nor upper nor a lower tail dependence. The simple Archimedean copula overestimates the VaR. Results provided by HAC are the closest to the true ones, but this copula underestimates the true VaR in all levels of significance. This is also natural because Gumbel copula describes wins rather than losses best. In general these results were expected due to the fact, that HAC is the only copula that was accepted by both tests under a high level of significance.

9 Conclusions

In this chapter we gave a short survey on copulae. We discussed different copula classes, methods of simulation and estimation and several goodness-of-fit tests. We provided an extensive simulation study in which two goodness-of-fit tests and two estimation techniques were considered. Afterwards, copulae were applied to de-GARCHed real world time-series. From the empirical study we conclude that, in some cases, even if margins are normal, the dependency is certainly not linearly normal, and more flexible dependency models are asked for.

References

Barbe, P., Genest, C., Ghoudi, K. and Rémillard, B. (1996). On Kendalls's process, *Journal of Multivariate Analysis* **58**: 197–229.

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**: 265–282.

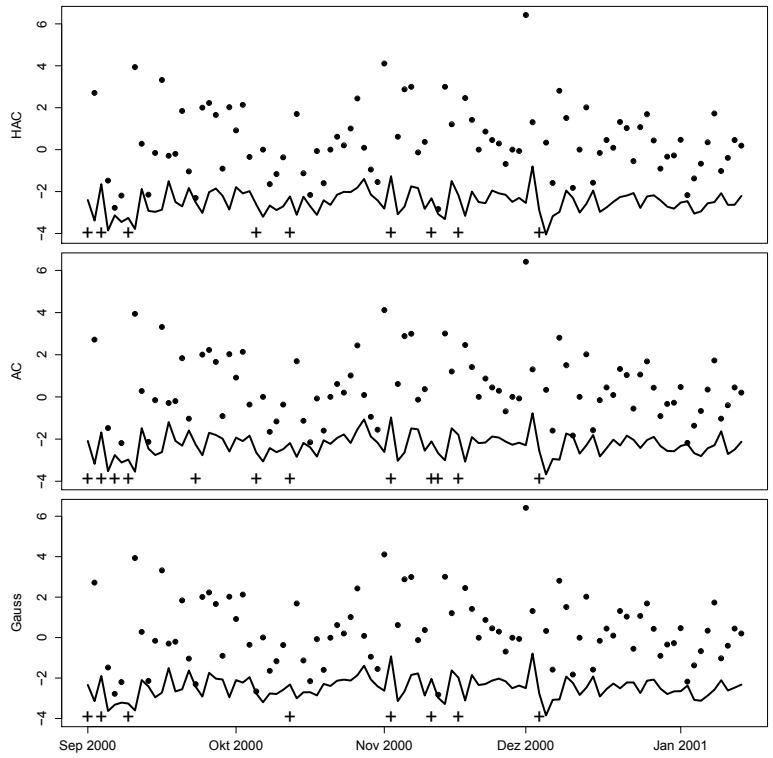

Figure 10: $\widehat{VaR}(\alpha)$, P&L (dots) and exceedances (crosses), estimated with 3-dimensional HAC with Gumbel generator (top), simple Gumbel copula (middle) and Gaussian copula (bottom) with $\alpha = 0.1$.

Chen, X. and Fan, Y. (2005). Pseudo-likelihood ratio tests for model selection in semiparametric multivariate copula models, *The Canadian Journal of Statistics* **33**(2): 389–414.

Chen, X. and Fan, Y. (2006). Estimation and model selection of semiparametric copula-based multivariate dynamic models under copula misspesification, *Journal of Econometrics* **135**: 125–154.

Chen, X., Fan, Y. and Patton, A. (2004). Simple tests for models of dependence between multiple financial time series, with applications to U.S. equity returns and exchange rates, *Discussion paper 483*, Financial Markets Group, London School of Economics.

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.

Choros, B., Härdle, W. and Okhrin, O. (2009). CDO and HAC, *SFB 649 Discussion Paper 2009-038*, Sonderforschungsbereich 649, Humboldt Universität zu Berlin, Germany. available at http://sfb649.wiwi.hu-berlin.de/papers/pdf/SFB649DP2009-038.pdf.

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**: 141–151.

Dall'Aglio, G. (1972). Fréchet classes and compatibility of distribution functions, *Symp. Math.* **9**: 131–150.

Dobrić, J. and Schmid, F. (2007). A goodness of fit test for copulas based on Rosenblatt's transformation, *Computational Statistics and Data Analysis* **51**: 4633 – 4642.

Embrechts, P., McNeil, A. J. and Straumann, D. (1999). Correlation and dependence in risk management: Properties and pitfalls, *RISK* pp. 69–71.

Fama, E. F. (1965). The behavior of stock market prices, *Journal of Business* pp. 34–105.

Fermanian, J.-D. (2005). Goodness-of-fit tests for copulas, *Journal of Multivariate Analysis* **95**(1): 119–152.

Fermanian, J.-D. and Scaillet, O. (2003). Nonparametric estimation of copulas for time series, *Journal of Risk* **5**: 25–54.

Filler, G., Odening, M., Okhrin, O. and Xu, W. (2010). On the systemic nature of weather risk, *Agricultural Finance Review* . forthcoming.

Frank, M. J. (1979). On the simultaneous associativity of $f(x, y)$ and $x + y - f(x, y)$, *Aequationes Mathematicae* **19**: 194–226.

Fréchet, M. (1951). Sur les tableaux de corrélation dont les marges sont donnés, *Annales de l 'Université de Lyon* **4**.

Genest, C., Ghoudi, K. and Rivest, L.-P. (1995). A semi-parametric estimation procedure of dependence parameters in multivariate families of distributions, *Biometrika* **82**: 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* **44**. in press.

Genest, C. and Rivest, L.-P. (1989). A characterization of Gumbel family of extreme value distributions, *Statistics and Probability Letters* **8**: 207–211.

Genest, C. and Rivest, L.-P. (1993). Statistical inference procedures for bivariate Archimedean copulas, *Journal of the American Statistical Association* **88**: 1034–1043.

Giacomini, E., Härdle, W. K. and Spokoiny, V. (2009). Inhomogeneous dependence modeling with time-varying copulae, *Journal of Business and Economic Statistics* **27**(2): 224–234.

Gumbel, E. J. (1960). Distributions des valeurs extrêmes en plusieurs dimensions, *Publ. Inst. Statist. Univ. Paris* **9**: 171–173.

Härdle, W. K. and Linton, O. (1994). Applied nonparametric methods, *in* R. Engle and D. McFadden (eds), *Handbook of Econometrics*, Elsevier, North-Holland.

Härdle, W., Okhrin, O. and Okhrin, Y. (2010). Time varying hierarchical Archimedean copulae, *submitted for publication* .

Härdle, W. and Simar, L. (2007). *Applied Multivariate Statistical Analysis*, 2 edn, Springer Verlag, Heidelberg.

Hennessy, D. A. and Lapan, H. E. (2002). The use of Archimedean copulas to model portfolio allocations, *Mathematical Finance* **12**: 143–154.

Hoeffding, W. (1940). Masstabinvariante Korrelationstheorie, *Schriften des Mathematischen Instituts und des Instituts für Angewandte Mathematik der Universität Berlin* **5**(3): 179–233.

Hoeffding, W. (1941). Masstabinvariante Korrelationsmasse für diskontinuierliche Verteilungen, *Arkiv für matematischen Wirtschaften und Sozial forschung* **7**: 49–70.

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, *Journal of Multivariate Analisys* **94**: 401–419.

Jones, M. C. (1993). Simple boundary corrections for kernel density estimation, *Statistic Computing* **3**: 135 – 146.

Junker, M. and May, A. (2005). Measurement of aggregate risk with copulas, *Econometrics Journal* **8**: 428–454.

Lee, T.-H. and Long, X. (2009). Copula-based multivariate garch model with uncorrelated dependent errors, *Journal of Econometrics* **150**(2): 207–218.

Lejeune, M. and Sarda, P. (1992). Smooth estimators of distribution and density functions, *Computational Statistics & Data Analysis* **14**: 457–471.

Li, D. X. (2000). On default correlation: A copula function approach, *The Journal of Fixed Income* **6**: 43–54.

Mandelbrot, B. (1965). The variation of certain speculative prices, *Journal of Business* **36**(4): 34–105.

Marshall, A. W. and Olkin, J. (1988). Families of multivariate distributions, *Journal of the American Statistical Association* **83**: 834–841.

McNeil, A. J. (2008). Sampling nested Archimedean copulas, *Journal Statistical Computation and Simulation* **78**(6): 567581.

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

Okhrin, O., Okhrin, Y. and Schmid, W. (2008). On the structure and estimation of hierarchical Archimedean copulas, *under revision* .

Okhrin, O., Okhrin, Y. and Schmid, W. (2009). Properties of Hierarchical Archimedean Copulas, *SFB 649 Discussion Paper 2009-014*, Sonderforschungsbereich 649, Humboldt-Universität zu Berlin, Germany. available at http://sfb649.wiwi.hu-berlin.de/papers/pdf/SFB649DP2009-014.pdf.

Patton, A. J. (2004). On the out-of-sample importance of skewness and asymmetric dependence for asset allocation, *Journal of Financial Econometrics* **2**: 130–168.

Renyi, A. (1970). *Probability Theory*, North-Holland, Amsterdam.

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

Savu, C. and Trede, M. (2004). Goodness-of-fit tests for parametric families of Archimedean copulas, *Discussion paper*, University of Muenster.

Sklar, A. (1959). Fonctions dé repartition á n dimension et leurs marges, *Publ. Inst. Stat. Univ. 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**: 62–76.

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

SFB 649 Discussion Paper Series 2010

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

001 "Volatility Investing with Variance Swaps" by Wolfgang Karl Härdle and Elena Silyakova, January 2010.
002 "Partial Linear Quantile Regression and Bootstrap Confidence Bands" by Wolfgang Karl Härdle, Ya'acov Ritov and Song Song, January 2010.
003 "Uniform confidence bands for pricing kernels" by Wolfgang Karl Härdle, Yarema Okhrin and Weining Wang, January 2010.
004 "Bayesian Inference in a Stochastic Volatility Nelson-Siegel Model" by Nikolaus Hautsch and Fuyu Yang, January 2010.
005 "The Impact of Macroeconomic News on Quote Adjustments, Noise, and Informational Volatility" by Nikolaus Hautsch, Dieter Hess and David Veredas, January 2010.
006 "Bayesian Estimation and Model Selection in the Generalised Stochastic Unit Root Model" by Fuyu Yang and Roberto Leon-Gonzalez, January 2010.
007 "Two-sided Certification: The market for Rating Agencies" by Erik R. Fasten and Dirk Hofmann, January 2010.
008 "Characterising Equilibrium Selection in Global Games with Strategic Complementarities" by Christian Basteck, Tijmen R. Daniels and Frank Heinemann, January 2010.
009 "Predicting extreme VaR: Nonparametric quantile regression with refinements from extreme value theory" by Julia Schaumburg, February 2010.
010 "On Securitization, Market Completion and Equilibrium Risk Transfer" by Ulrich Horst, Traian A. Pirvu and Gonçalo Dos Reis, February 2010.
011 "Illiquidity and Derivative Valuation" by Ulrich Horst and Felix Naujokat, February 2010.
012 "Dynamic Systems of Social Interactions" by Ulrich Horst, February 2010.
013 "The dynamics of hourly electricity prices" by Wolfgang Karl Härdle and Stefan Trück, February 2010.
014 "Crisis? What Crisis? Currency vs. Banking in the Financial Crisis of 1931" by Albrecht Ritschl and Samad Sarferaz, February 2010.
015 "Estimation of the characteristics of a Lévy process observed at arbitrary frequency" by Johanna Kappusl and Markus Reiß, February 2010.
016 "Honey, I'll Be Working Late Tonight. The Effect of Individual Work Routines on Leisure Time Synchronization of Couples" by Juliane Scheffel, February 2010.
017 "The Impact of ICT Investments on the Relative Demand for High-, Medium-, and Low-Skilled Workers: Industry versus Country Analysis" by Dorothee Schneider, February 2010.
018 "Time varying Hierarchical Archimedean Copulae" by Wolfgang Karl Härdle, Ostap Okhrin and Yarema Okhrin, February 2010.
019 "Monetary Transmission Right from the Start: The (Dis)Connection Between the Money Market and the ECB's Main Refinancing Rates" by Puriya Abbassi and Dieter Nautz, March 2010.
020 "Aggregate Hazard Function in Price-Setting: A Bayesian Analysis Using Macro Data" by Fang Yao, March 2010.
021 "Nonparametric Estimation of Risk-Neutral Densities" by Maria Grith, Wolfgang Karl Härdle and Melanie Schienle, March 2010.

SFB 649 Discussion Paper Series 2010

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

022 "Fitting high-dimensional Copulae to Data" by Ostap Okhrin, April 2010.