Technische Universität München
Fakultät für Mathematik
Lehrstuhl für Finanzmathematik

Goodness-of-fit tests for elliptical copulas

Miriam Angelika Jaser

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

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzende:		Prof. Claudia Czado, Ph.D.
Prüfer der Dissertation:	1.	Priv.-Doz. Dr. Aleksey Min
	2.	Prof. Dr. Yarema Okhrin (Universität Augsburg)
	3.	Prof. Dr. Hajo Holzmann (Philipps-Universität Marburg)

Die Dissertation wurde am 25.11.2020 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik am 08.03.2021 angenommen.

Abstract

This thesis explores statistical testing procedures to detect whether the underlying dependence structure between the components of a given multidimensional random sample may be captured by an elliptical copula. First, we disregard that the univariate marginal distribution functions of the components are usually unknown and assume given copula data. For this scenario, we develop a simple non-parametric goodness-of-fit test for multivariate elliptical copulas exploiting the equality of Kendall's tau and Blomqvist's beta for all bivariate margins. In a next step, we derive simple non-parametric tests for two further intrinsic properties of bivariate elliptical copulas, namely symmetry and radial symmetry. The combination of the three proposed tests yields a powerful testing procedure for bivariate elliptical copulas. Finally, we expand the simple goodness-of-fit test for elliptical copulas by taking the estimation of unknown marginal distribution functions into account. In addition, we resolve some limitation of the simple goodness-of-fit test for copula data in higher dimensions. The finite sample performances of all presented statistical procedures are analyzed in extensive simulation studies and, if applicable, compared to the performance of competing ones. The practical relevance of the proposed tests is illustrated in empirical analyses.

Zusammenfassung

In dieser Arbeit werden statistische Testverfahren untersucht, um festzustellen, ob die zugrunde liegende Abhängigkeitsstruktur zwischen den Komponenten einer gegebenen mehrdimensionalen Zufallsstichprobe durch eine elliptische Copula beschrieben werden kann. Dabei vernachlässigen wir zunächst, dass die eindimensionalen Randverteilungsfunktionen der Komponenten in der Regel unbekannt sind, und gehen von gegebenen Copuladaten aus. Für dieses Szenario entwickeln wir einen einfachen nichtparametrischen Anpassungstest für multivariate elliptische Copulas, der die Gleichheit von Kendalls Tau und Blomqvists Beta für alle bivariaten Ränder ausnutzt. Im nächsten Schritt leiten wir einfache nichtparametrische Tests für zwei weitere intrinsische Eigenschaften bivariater elliptischer Copulas her, nämlich Symmetrie und Radialsymmetrie. Die Kombination der drei vorgeschlagenen Tests ergibt ein trennscharfes Testverfahren für zweidimensionale elliptische Copulas. Zu guter Letzt erweitern wir den einfachen Anpassungstest für elliptische Copulas, indem wir die Schätzung der unbekannten Randverteilungsfunktionen berücksichtigen. Darüber hinaus beheben wir einige Einschränkungen des einfachen Anpassungsgütetests für Copuladaten in höheren Dimensionen. Das Verhalten für endliche Stichproben (also das empirische Niveau und die empirische Güte) aller vorgestellten statistischen Verfahren wird in umfangreichen Simulationsstudien analysiert und gegebenenfalls mit dem Verhalten konkurrierender Verfahren verglichen. Die praktische Relevanz der vorgeschlagenen Tests wird in empirischen Analysen veranschaulicht.

Acknowledgements

First of all, I would like to thank my supervisor PD Dr. Aleksey Min for providing me the opportunity to write this thesis. Throughout my doctoral studies, he was always available for discussions and never hesitated to invest his time into my development as a researcher, teaching assistant, and as a person. Without his guidance and support this thesis would not have been possible. Moreover, I am very grateful to Prof. Dr. Rudi Zagst for the teaching position that enabled me to pursue my doctoral studies at the Chair of Mathematical Finance.

My sincere thanks go to my co-author Dr. Stephan Haug from the Research group Mathematical Statistics for his excellent support, the many fruitful discussions, and the entirely successful collaboration. In this context, I would also like to thank Prof. Dr. Nina Gantert from the Chair of Probability who served as my mentor and Prof. Dr. Matthias Scherer for his valuable feedback in numerous discussions.

Moreover, I would like to express my sincere gratitude to Prof. Dr. Axel Bücher from the Heinrich-Heine-Universität Düsseldorf for his visit in Munich and even more for the kind invitation to Düsseldorf. I am very thankful for the extraordinarily constructive time we spent together, the fruitful discussions before and afterwards and the valuable feedback which has lead to a significantly improved presentation of this thesis. Working with him was not only very inspiring, but also a great pleasure.

I also would like to acknowledge the financial support by the German Research Foundation (DFG) and the Technical University of Munich (TUM) within the funding program Open Access Publishing as well es by the TUM Graduate School's International School of Applied Mathematics, which enabled me to participate in the CFE 2019 conference in London.

Furthermore, I would like to thank my former and current colleagues at the Chair of Financial Mathematics for the supportive working atmosphere and the many interesting discussions on mathematics and beyond.

Last but not least, my special gratitude goes to my friends and my family, especially my parents Reinhold and Angelika, and my sister Jasmin for their love, understanding and constant support and for making my life outside my doctoral studies worth living. Most importantly, I thank my wife Caro not only for her endless love, but also for helping me to set my priorities right. It is her, who guides me to find out what really matters to me in life and who gives me the reassurance to not be afraid of what I will find.

Miriam Jaser
November 23, 2020

Contents

1 Introduction

The International Data Corporation (IDC) predicts that the “Global Datasphere” will grow from 45 Zettabytes in 2019 to 175 Zettabytes by 2025 (see Reinsel et al. (2018)). The world from business through society to the everyday life of each individual is and will be increasingly driven by this inconceivable amount of data. In view of the currently spreading SARS-CoV-2 pandemic, we become more than ever and painfully aware of another major phenomenon of our time: the globalization which makes our world an interconnected and interdependent place. In this light, statistical methods for an accurate data analysis of huge amounts of data and especially for the analysis of dependencies become indispensable.

In scientific research, the modeling of dependence structures between multivariate random quantities has been an important field in probability theory and statistics for many years. The monograph by Joe (1997) gives an overview of various dependence concepts. Originally, the multivariate normal distribution and with it the correlation coefficient were widely used to model and measure multivariate dependence. However, this approach reveals some inadequacies as for example the correlation coefficient can only measure the linear relationship between two random variables resulting in the necessity for alternatives (see, e.g., Embrechts et al. (1999)).

A more general concept to model dependence structures is given by the so-called copulas. It is based on the famous Sklar's theorem dating back to Sklar (1959), which claims that any multivariate distribution function F on $\mathbb{R}^d$ can be separated into its marginal distribution functions $F_1, \ldots, F_d$ and a copula $C : [0,1]^d \to [0,1]$ in the following way

$$F(\boldsymbol{x}) = C\left(F_1(x_1), \ldots, F_d(x_d)\right) , \quad \boldsymbol{x} \in \mathbb{R}^d .$$

Thus, the modeling of the joint behavior of random variables can be split into two simpler sub-problems, namely the consideration of the individual behavior of the random variables as well as their interaction that is the dependence structure, which is grasped by the copula.

Due to their mathematical elegance, copulas have attracted an increasing attention since the late 1990's and are nowadays a common tool to model dependencies (see, e.g., Nelsen (1999), Joe (2015)). The theory of copulas has been applied in various fields such as actuarial sciences (see, e.g., Frees and Valdez (1998)), finance (see, e.g., Cherubini et al. (2004), Genest et al. (2009a)), hydrology (see, e.g., Genest and Favre

(2007), Salvadori and De Michele (2007)), machine learning (see, e.g., Elidan (2013)), and risk management (see, e.g., McNeil et al. (2005)), to name just a few.

A simple transformation of Sklar's theorem allows to construct a copula C from a given multivariate distribution function F and its marginal distribution functions $F_1, \ldots, F_d$ with generalized inverses $F_1^-, \ldots, F_d^-$ in the following way

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

This gives rise to the popular class of elliptical copulas, which are simply the copulas of elliptical distributions. Its most prominent representatives are the Gaussian and the t copula. Elliptical copulas are fully specified through an association matrix and the generator function of the corresponding elliptical distribution. Note that distributions whose dependence structure is captured by an elliptical copula are called meta-elliptical distributions and were introduced by Fang et al. (2002), Fang et al. (2005). Elliptical copulas and meta-elliptical distributions are well investigated (see, e.g., Frahm et al. (2003), Abdous et al. (2005), Demarta and McNeil (2005)) and applied in several areas such as actuarial sciences (see, e.g., Oh et al. (2020)), finance (see, e.g., Fischer et al. (2009)), hydrology (see, e.g., Song and Singh (2010)), or risk management (see, e.g., Li (2000), Embrechts et al. (2003)).

By definition, every mathematical model suffers from limitations and so do elliptical copulas. Not considering the limitations of the Gaussian copula approach to assess credit risk by Li (2000) paved the way for the financial crisis of 2007 to 2009 (see Salmon (2009)). Not only because of this example but in general it is very important to check whether the dependence structure of given data can be represented by the very specific dependence structure of an elliptical copula. This is where goodness-of-fit tests for copulas come into play. For an overview of goodness-of-fit tests for copulas, see Genest et al. (2009b), Berg (2009), or Fermanian (2013). To the best of our knowledge, regarding elliptical copulas, so far, there only exist goodness-of-fit tests for previously specified parametric families, like the Gaussian or the t copula. The procedures presented in Quessy and Bellerive (2013) are based on a fixed or at least parametric generator function. Therefore, these are also only suitable for specified parametric families. The test by Li and Peng (2009) can be regarded as a test for elliptical copulas as well. However, it is based on the concept of tail-dependence and thereby excludes for example the Gaussian copula from the null hypothesis. Thus, our search for a procedure to test the goodness-of-fit for the entire class of elliptical copulas was not crowned by success.

The main motivation of this thesis is to develop non-parametric formal statistical testing procedures which enable us to determine whether the dependence structure of a multivariate random sample is well-represented by an elliptical copula. To this end, we consider the test problem

$$H_0 : C \in \mathcal{C}^{ellipt} \quad \text{vs.} \quad H_1 : C \notin \mathcal{C}^{ellipt} ,$$

where C denotes the unknown copula and $\mathcal{C}^{ellipt}$ the class of elliptical copulas.

Outline of the thesis

Parts of this thesis are based on the following two research papers:

- Chapter 3:

  Jaser et al. (2017):
  Jaser, M., Haug, S., and Min, A. (2017). A simple non-parametric goodness-of-fit test for elliptical copulas. *Depend. Model.*, 5(1):330–353

- Chapter 4:

  Jaser and Min (2020):
  Jaser, M. and Min, A. (2020). On tests for symmetry and radial symmetry of bivariate copulas towards testing for ellipticity. *Comput. Stat.*

For this thesis, the content of the papers has been revised and where appropriate extended by additional arguments or illustrations. Especially Chapter 4 contains some new materials. Moreover, note that some parts of Chapter 2 are very similar to parts of the above listed research papers as well.

In the following, we provide an outline of the thesis structure and present the main contributions of the thesis.

In **Chapter 2**, we introduce the mathematical fundamentals for this thesis which revolve around the concept of copulas.

For the next two chapters, we assume given copula data and neglect that in real-life situations, marginal distributions would be needed to be estimated first. In **Chapter 3**, which is based on the research paper Jaser et al. (2017), we build a simple non-parametric goodness-of-fit test for multivariate elliptical copulas of any dimension. For this, we use a property which is common to all elliptical copulas: the equality of Kendall's tau and Blomqvist's beta (Fang et al. (2002), Schmid and Schmidt (2007)). We derive a Wald-type test statistic based on the equality of Kendall's tau and Blomqvist's beta for all bivariate margins and establish its asymptotic chi-square distribution. After analyzing the empirical level and the empirical power, the presented simple goodness-of-fit test is applied to the dependence structure underlying a financial data set.

In the next step, we take two further properties of bivariate elliptical copulas into consideration: symmetry and radial symmetry. In **Chapter 4**, which is based on the research paper Jaser and Min (2020), we build simple non-parametric tests for symmetry and radial symmetry of bivariate copulas and incorporate them together with the test for equality of Kendall's tau and Blomqvist's beta from Chapter 3 into a powerful testing procedure for bivariate ellipticity. We establish the asymptotic normality of the corresponding test statistics. In an extensive simulation study, we compare the tests to existing more advanced tests for symmetry and radial symmetry by Genest et al. (2012) and Genest and Nešlehová (2014), respectively. In a further simulation study, we analyze the finite-sample performance of the testing procedure for ellipticity. We illustrate the testing procedure in practice with applications to financial and insurance data. In the supplementary material to this chapter, we first show how the tests for symmetry and radial symmetry can be refined utilizing variance reduction techniques (see, e.g., Korn et al. (2010)). Secondly, we validate the choice of Kendall's tau for the tests for symmetry and radial symmetry by analyzing the performance of competing alternative tests based on Spearman's rho.

Ultimately, we waive the requirement of given copula data. In **Chapter 5**, we further develop the simple non-parametric goodness-of-fit test for elliptical copulas of any dimension to take the estimation of unknown marginals into account and to resolve the problems of the simple test to hold its nominal level in higher dimensions. The latter is achieved by utilizing an L_2-type instead of a Wald-type test statistic. This avoids the estimation of the covariance matrix, which becomes more and more burdensome with increasing dimension. With the help of empirical copula theory, we derive the limiting Gaussian field which depends on the unknown copula. To perform the test, we make use of the subsampling approximation by Kojadinovic and Stemikovskaya (2019). It should be mentioned that Quessy (2020) also deals with the two described limitations of the simple test from Chapter 3 published in Jaser et al. (2017). However, our proofs differ from Quessy (2020) and we utilize an alternative bootstrap procedure. All results presented in Chapter 5 have been developed independently since Jaser et al. (2017). In an extensive simulation study, we compare our advanced test to the simple test for pseudo-observations as well as to the competing test by Quessy (2020).

Finally, note that conclusions and directions for future research are given in the respective chapters.

2
Preliminaries

This chapter introduces the central terms and concepts forming the basis of the thesis. Note that some parts of this chapter are taken from Jaser et al. (2017) and Jaser and Min (2020). In Section 2.1, we introduce copulas. Section 2.2 presents the concept of ordinal measures of dependence and gives two concrete examples. Moreover, two different notions of symmetry are discussed in Section 2.3. Elliptical copulas are presented in Section 2.4 and Archimedean copulas in Section 2.5. Finally, we close the chapter with the definition of pseudo-observations and different versions of the empirical copula in Section 2.6.

2.1 Copulas

A d-dimensional copula is a cumulative distribution function over the unit hypercube $[0,1]^d$ with uniformly distributed margins. By providing a link between copulas and multivariate distribution functions, the following fundamental result by Sklar (1959) makes copulas a mathematical tool for modeling the dependence structure of multivariate distributions.

Theorem 2.1. ***(Sklar's theorem)***
Let F be a cumulative distribution function on $\mathbb{R}^d$ with continuous margins $F_1, \ldots, F_d$. Then there exists a unique copula $C : [0,1]^d \to [0,1]$ such that for all $\boldsymbol{x} \in \mathbb{R}^d$ it holds that

$$F(\boldsymbol{x}) = C\left(F_1(x_1), \ldots, F_d(x_d)\right) .$$

In particular, Sklar's theorem allows to treat margins and the copula separately resulting in two independent and simpler problems. Further, Sklar's theorem provides a universal construction framework for copulas. Without loss of generality, let F_i^- be the generalized inverses of F_i, $i \in \{1, \ldots, d\}$. Then, the copula $C(\boldsymbol{u})$ of F for any $\boldsymbol{u} \in [0,1]^d$ is given by

$$C(\boldsymbol{u}) = F\left(F_1^-(u_1), \ldots, F_d^-(u_d)\right) .$$

Note that for any distribution function H on $[0,1]^d$, $H_{k\ell}$ denotes the bivariate marginal distribution function of the k-th and ℓ-th component with $k, \ell \in \{1, \ldots, d\}$.

More precisely, it holds that $H_{k\ell}(u_k, u_\ell) = H(\boldsymbol{u}^{(k\ell)})$, where, for any vector $\boldsymbol{u} \in [0,1]^d$ and $A \subset \{1, \ldots, d\}$, the vector $\boldsymbol{u}^{(A)}$ denotes the vector where all components of $\boldsymbol{u}$ except the components of the index set A are replaced by 1. Furthermore, for $k, \ell \in \{1, \ldots, d\}$ with $k < \ell$, the corresponding bivariate margin of C is a bivariate copula, which is called the marginal copula of the k-th and ℓ-th component and denoted by $C_{k\ell}$ (see Nelsen (1999)).

Finally, an analog to Sklar's theorem links multivariate survival functions to a copula and its marginal survival functions. For any random vector $\boldsymbol{X} \in \mathbb{R}^d$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with cumulative distribution function F, the survival function $\bar{F}$ is defined as $\bar{F}(\boldsymbol{x}) = \mathbb{P}(X_1 > x_1, \ldots, X_d > x_d)$, $\boldsymbol{x} \in \mathbb{R}^d$.

Theorem 2.2. ***(Sklar's theorem for survival functions)***
Let $\bar{F}$ be a d-dimensional survival function with continuous marginal survival functions $\bar{F}_1, \ldots, \bar{F}_d$. Then there exists a unique copula $\hat{C} : [0,1]^d \to [0,1]$ such that for all $\boldsymbol{x} \in \mathbb{R}^d$ it holds that

$$\bar{F}(\boldsymbol{x}) = \hat{C}\left(\bar{F}_1(x_1), \ldots, \bar{F}_d(x_d)\right) .$$

The copula $\hat{C}$ from Theorem 2.2 is called the survival copula of a random vector $\boldsymbol{X} \in \mathbb{R}^d$ with survival function $\bar{F}$. Note that the survival function $\bar{C}$ of a copula C is not a copula itself, however, the survival copula $\hat{C}$ is indeed a copula. Furthermore, it holds that the random vector $\boldsymbol{U}$ is distributed according to the copula C if and only if the random vector $(1 - U_1, \ldots, 1 - U_d)$ is distributed according to the survival copula $\hat{C}$. Moreover, the copula C and the survival copula $\hat{C}$ can be analytically connected via the inclusion exclusion principle.

For a more extensive treatment of copula theory, we refer to the standard textbooks Joe (2015) and Nelsen (1999).

2.2 Ordinal measures of dependence

In this section, we consider ordinal or concordance measures of dependence, which are invariant with respect to monotone increasing, not necessarily linear transformations and can also be expressed in terms of the underlying copula. Concordance measures allow to summarize the strength of dependence of a random vector and, therefore, inherent in its copula, by a single number. In the sequel, we introduce Kendall's tau and Blomqvist's beta, which are fundamental for the tests derived in the following chapters. These tests will be based on the dependence between all bivariate pairs of the components of the random vector $\boldsymbol{X} \in \mathbb{R}^d$. Therefore, we will introduce these measures in a bivariate setting. For multivariate extensions of Kendall's tau, we refer to Kendall and Smith (1940) and Joe (1990). A multivariate extension of Blomqvist's beta was introduced in Schmid and Schmidt (2007).

2.2.1 Kendall's tau

We start with the concordance measure Kendall's tau, which belongs to the most popular dependence measures and is defined as follows.

Definition 2.3. ***(Kendall's tau)***
Let (X'_k, X'_ℓ) be an independent copy of the random vector (X_k, X_ℓ) of continuous random variables X_k and X_ℓ. Then, Kendall's tau is defined by

$$\begin{aligned}\tau_{k\ell} :&= \mathbb{E}[\operatorname{sgn}(X_k - X'_k)\operatorname{sgn}(X_\ell - X'_\ell)]\\ &= \mathbb{P}((X_k - X'_k)(X_\ell - X'_\ell) > 0) - \mathbb{P}((X_k - X'_k)(X_\ell - X'_\ell) < 0)\,,\end{aligned}$$

where sgn *denotes the sign function.*

Hence, Kendall's tau equals the probability of concordance minus the probability of discordance. Furthermore, for continuous random variables X_k and X_l with copula $C_{k\ell}$, Kendall's tau is completely determined by their copula $C_{k\ell}$ (see Theorem 5.1.3 in Nelsen (1999)) and can be expressed as

$$\tau_{k\ell} = \tau_{C_{k\ell}} = 4\int_0^1\int_0^1 C_{k\ell}(u_k, u_\ell)\, dC_{k\ell}(u_k, u_\ell) - 1\,. \tag{2.1}$$

For the sample version of Kendall's tau, we look at a random sample of n observations $(X_{k1}, X_{\ell 1}), \ldots, (X_{kn}, X_{\ell n})$ from the random vector (X_k, X_ℓ). In total, there are $\binom{n}{2} = \frac{n(n-1)}{2}$ different pairs of observations $(X_{ki}, X_{\ell i})$ and $(X_{kj}, X_{\ell j})$ and we get

$$\tau_{k\ell,n} := \frac{2}{n(n-1)} \sum_{1\le i<j\le n} \operatorname{sgn}(X_{ki} - X_{kj})\operatorname{sgn}(X_{\ell i} - X_{\ell j}) \tag{2.2}$$

as the minimum variance unbiased estimator for Kendall's tau (see Denker (1985)).

2.2.2 Blomqvist's beta

The second concordance measure, we want to consider, is Blomqvist's beta, also referred to as the medial correlation coefficient. The intention of Blomqvist (1950) was to design a simple rank correlation coefficient which can be easily applied in practice. Blomqvist's beta is defined as follows.

Definition 2.4. ***(Blomqvist's beta)***
Let X_k and X_ℓ be continuous random variables. Then, Blomqvist's beta is defined by

$$\begin{aligned}\beta_{k\ell} :&= \mathbb{E}[\operatorname{sgn}(X_k - \tilde{x}_k)\operatorname{sgn}(X_\ell - \tilde{x}_\ell)]\\ &= \mathbb{P}((X_k - \tilde{x}_k)(X_\ell - \tilde{x}_\ell) > 0) - \mathbb{P}((X_k - \tilde{x}_k)(X_\ell - \tilde{x}_\ell) < 0)\,,\end{aligned}$$

where $\tilde{x}_k$ and $\tilde{x}_\ell$ denote the population medians of X_k and X_ℓ, respectively.

Hence, Blomqvist's beta equals the probability of X_1 and X_2 being both either smaller or greater than their respective medians minus the probability of one being smaller and the other one being greater than its median. Blomqvist's beta can easily be expressed in terms of the copula C_{12} of the distribution of (X_1, X_2) and is given by

$$\beta_{k\ell} = \beta_{C_{k\ell}} = 4C_{k\ell}\left(\frac{1}{2}, \frac{1}{2}\right) - 1\,. \tag{2.3}$$

Consequently, for copulas with a closed-form analytical expression, Blomqvist's beta can be explicitly derived. This displays one advantage of Blomqvist's beta over other more complicated dependence measures.

Now, let $(X_{k1}, X_{\ell 1}), \ldots, (X_{kn}, X_{\ell n})$ be again a random sample of n observations from the random vector (X_k, X_ℓ) and let $\widetilde{X}_{k,n}$ and $\widetilde{X}_{\ell,n}$ be the sample medians of the components of the sample. Definition 2.4 trivially leads to the following sample version of Blomqvist's beta given by

$$\beta_{k\ell,n} := \frac{1}{n}\sum_{i=1}^{n} \operatorname{sgn}\left(X_{ki} - \widetilde{X}_{k,n}\right) \operatorname{sgn}\left(X_{\ell i} - \widetilde{X}_{\ell,n}\right)\,. \tag{2.4}$$

2.3 Symmetries of copulas

Copulas can be classified with respect to their symmetry properties. In this section, we introduce two notions of symmetry, namely exchangeability and radial symmetry.

Definition 2.5. ***(Exchangeability)***
A copula C is called exchangeable if,

$$C(u_1, \ldots, u_d) = C(u_{\pi(1)}, \ldots, u_{\pi(d)})\,, \quad u_1, \ldots, u_d \in [0,1],$$

for all permutations $(\pi(1), \ldots, \pi(d))$ of $\{1, \ldots, d\}$.

We follow Nelsen (1999), and simply call a bivariate copula C symmetric if $C(u,v) = C(v,u)$, for all $(u,v) \in [0,1]^2$. If the bivariate copula C is symmetric and the distribution function of a random vector (U,V), then the dependence structure between U and V is symmetric and, hence, we have

$$(U,V) \stackrel{d}{=} (V,U)\,. \tag{2.5}$$

The scatter plots of random samples of symmetric bivariate copulas therefore show symmetry with respect to the main diagonal $\{u = v\}$.

Definition 2.6. ***(Radial symmetry)***
A copula C is called radially symmetric if it coincides with its survival copula, that is if

$$C = \hat{C}\,.$$

Let C be a bivariate copula and the distribution function of a random vector (U, V). Then, the corresponding survival copula $\hat{C}$ is the distribution function of $(1-U, 1-V)$. Thus, C is radially symmetric if $(U, V) \stackrel{d}{=} (1-U, 1-V)$ or, equivalently,

$$(U - 0.5, V - 0.5) \stackrel{d}{=} (0.5 - U, 0.5 - V)\,.$$

The scatter plots of random samples of radially symmetric bivariate copulas therefore show symmetry with respect to the point $(0.5, 0.5)$.

2.4 Elliptical copulas

One of the most prominent parametric classes of copulas are elliptical copulas. They are implicit copulas, which do not possess a simple closed-form analytic expression. More precisely, elliptical copulas are derived from multivariate elliptical distribution functions with the help of Sklar's theorem. Therefore, we first introduce elliptical distributions. Our exposition follows Chapter 2 in Fang et al. (1990) and is based on spherical distributions that stay invariant under orthogonal transformations of the underlying random vectors. Spherical distributions are an important sub-class of elliptical distributions.

Definition 2.7. ***(Elliptical distribution)***
Let $\mathbb{S}_d$ denote the space of all symmetric $d \times d$ matrices. A random vector $\boldsymbol{X} \in \mathbb{R}^d$ is said to have an (non-degenerate) elliptical distribution with parameters $\boldsymbol{\mu} \in \mathbb{R}^d$ and $\boldsymbol{\Sigma} = (\sigma_{k\ell})_{k,\ell \in \{1,\dots,d\}} \in \mathbb{S}_d$, if

$$\boldsymbol{X} = \boldsymbol{\mu} + \mathcal{A}\boldsymbol{Y},$$

where $\boldsymbol{Y}$ has a m-dimensional spherical distribution and $\mathcal{A}$ is a $d \times m$ matrix such that $\mathcal{A}\mathcal{A}^\top = \boldsymbol{\Sigma}$ with $\operatorname{rank}(\boldsymbol{\Sigma}) = m$.

Thus, elliptical distributions are defined as the class of affine transformations of spherical distributions. A bivariate elliptically distributed random vector $\boldsymbol{X}$ resulting from the application of the linear transformation $\mathcal{A}$ to the spherically distributed random vector $\boldsymbol{Y}$ has elliptically contoured density level surfaces. This explains the name of elliptical distributions. Definition 2.7 is the stochastic representation of elliptical distributions. Note that elliptical distributions can alternatively be defined through their generator function. For further details about elliptical distributions and the definition of spherical distributions we refer to Fang et al. (1990).

Since Sklar's theorem (see Theorem 2.1) determines the copula of multivariate distributions with continuous margins in an unique way, elliptical copulas are defined as follows.

Definition 2.8. ***(Elliptical copula)***
Elliptical copulas are the copulas of elliptical distributions.

Consequently, an elliptical copula C is defined as the copula of the underlying elliptical distribution F and is typically not available in closed form. Distributions with an elliptical copula are called (meta)-elliptical distributions (see Fang et al. (2002)). These distributions are fully specified through the matrix

$$\mathcal{R} = (\rho_{k\ell})_{k,\ell\in\{1,\dots,d\}} := (\sigma_{k\ell}/\sqrt{\sigma_{kk}\sigma_{\ell\ell}})_{k,\ell\in\{1,\dots,d\}}$$

the generator function and the marginal distributions.

Note that elliptical copulas are radially symmetric (see, e.g., Lemma 4.6 in Mai and Scherer (2012)). In addition, bivariate elliptical copulas are symmetric. The two most popular elliptical copulas are the Gaussian and the t copula.

2.4.1 Bivariate Gaussian copula

The Gaussian (also normal or just Gauss) copula C_P^{Gauss} is the copula of $(X, Y) \sim N_2(\mathbf{0}, P)$, where we denote with $N_2(\mathbf{0}, P)$ a bivariate normal distribution with mean $\mathbf{0}$ and correlation matrix P. In the present bivariate case, we write C_ρ^{Gauss}, where $\rho = \rho_{XY}$ stands for the linear correlation of X and Y. The implicit form of the Gaussian copula is given by

$$C_\rho^{Gauss}(u, v) = \Phi_\rho\left(\Phi^{-1}(u), \Phi^{-1}(v)\right), \quad u, v \in [0, 1], \tag{2.6}$$

where Φ_ρ denotes the distribution function of $N_2(\mathbf{0}, P)$ and Φ^{-1} represents the quantile function of the univariate standard normal distribution. For $\rho \in (0, 1)$, Equation (2.6) implies the following expression for the bivariate Gaussian copula:

$$C_\rho^{Gauss}(u, v) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi(1-\rho^2)^{1/2}} \exp\left(\frac{-(s^2 - 2\rho st + t^2)}{2(1-\rho^2)}\right) ds\, dt\,.$$

2.4.2 Bivariate t copula

The t copula is the copula of $(X, Y) \sim t_2(\nu, \mathbf{0}, P)$, where we denote with $t_2(\nu, \mathbf{0}, P)$ a bivariate t distribution with degrees of freedom parameter $\nu > 0$, location parameter $\mathbf{0}$ and association matrix P having the structure of correlation matrices (see Kotz and Nadarajah (2004)). Without loss of generality, we assume $\nu > 2$. In this case, the correlation matrix of (X, Y) exists and coincides with the association matrix P. Similarly to the Gaussian copula, $C_{\nu,\rho}^t$ denotes the bivariate t copula with linear correlation parameter $\rho = \rho_{XY}$. The implicit form of the t copula is given by

$$C_{\nu,\rho}^t(u, v) = t_{\nu,\rho}\left(t_\nu^{-1}(u), t_\nu^{-1}(v)\right), \quad u, v \in [0, 1], \tag{2.7}$$

where $t_{\nu,\rho}$ denotes the distribution function of $t_2(\nu, \mathbf{0}, P)$ and t_ν^{-1} represents the quantile function of the univariate t distribution with ν degrees of freedom. For $\rho \in (0, 1)$, Equation (2.7) implies the following expression for the bivariate t copula:

$$C^t_{\nu,\rho}(u,v) = \int_{-\infty}^{t_\nu^{-1}(u)} \int_{-\infty}^{t_\nu^{-1}(v)} \frac{1}{2\pi(1-\rho^2)^{1/2}} \left(1 + \frac{s^2 - 2\rho st + t^2}{\nu(1-\rho^2)}\right)^{-(\nu+2)/2} ds\, dt\,.$$

2.5 Archimedean copulas

Here, we outline bivariate Archimedean copulas and follow Nelsen (1999). For d-dimensional Archimedean copulas with $d > 2$, we refer to Chapter 2 of Mai and Scherer (2012). Thus, we consider the simplest construction of multivariate Archimedean copulas, which are exchangeable and have only one parameter.

Definition 2.9. ***(Bivariate Archimedean copula)***
Let $\varphi : [0,1] \to [0,\infty]$ be a continuous, strictly decreasing, convex function with $\varphi(1) = 0$. Then, the function $C_\varphi : [0,1]^2 \to [0,1]$ given by

$$C_\varphi(u,v) = \varphi^{[-1]}(\varphi(u) + \varphi(v)) \tag{2.8}$$

is a copula, where $\varphi^{[-1]}$ is a pseudo-inverse of φ. Copulas of this form are called Archimedean copulas and φ is called a generator. If $\varphi(0) = \infty$, the generator is called strict, $\varphi^{[-1]} = \varphi^{-1}$ and $C_\varphi(u,v) = \varphi^{-1}(\varphi(u) + \varphi(v))$ is said to be a strict Archimedean copula.

Table 2.1 summarizes generators with parameter ranges and the resulting explicit expression for the bivariate Archimedean copulas from the Frank, Clayton and Gumbel family.

Copula family	$\varphi_\theta(t)$	$\theta \in$	$C_\theta(u,v)$
Frank	$-\ln \frac{e^{-\theta t} - 1}{e^{-\theta} - 1}$	$\mathbb{R} \setminus \{0\}$	$-\frac{1}{\theta} \ln \left(1 + \frac{\left(e^{-\theta u} - 1\right)\left(e^{-\theta v} - 1\right)}{e^{-\theta} - 1}\right)$
Clayton	$\frac{1}{\theta}\left(t^{-\theta} - 1\right)$	$(0,\infty)$	$\left(u^{-\theta} + v^{-\theta} - 1\right)^{-1/\theta}$
Gumbel	$(-\ln t)^\theta$	$[1,\infty)$	$\exp\left(-\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{1/\theta}\right)$

Table 2.1: Summary of generators, parameter ranges and explicit expressions for the bivariate Frank, Clayton and Gumbel copula.

For a bivariate Archimedean copula C, one can compute Kendall‘s tau using its generator φ. More precisely, the following relation (see Genest and MacKay (1986)) holds,

$$\tau = 1 + 4 \int_0^1 \frac{\varphi(t)}{\varphi'(t)} \, dt \,.$$

Further, Equation (2.3) and Definition 2.9 imply for Blomqvist's beta that

$$\beta = 4 \, \varphi^{[-1]} \left(2 \, \varphi \left(\frac{1}{2} \right) \right) - 1 \,.$$

Archimedean copulas are exchangeable by construction. Moreover, the bivariate Frank copula is even radially symmetric, that is the survival copula coincides with the copula itself. Being exchangeable and radially symmetric, bivariate Frank copulas possess the same symmetry properties as bivariate elliptical copulas. Therefore, it is very important to distinguish between them when modeling the dependence of bivariate data. Table 2.2 reports Kendall's tau, Blomqvist's beta and the symmetry properties for the bivariate Frank, Clayton and Gumbel copula.

Copula family	τ_θ	β_θ	exch.	rad. sym.
Frank	$1 + \frac{4}{\theta} \left(D_1(\theta) - 1 \right)$	$1 + \frac{4}{\theta} \ln \left(\frac{1}{2} \left(e^{-\theta/2} + 1 \right) \right)$	✓	✓
Clayton	$\frac{\theta}{\theta + 2}$	$4 \left(2^{\theta+1} - 1 \right)^{-1/\theta} - 1$	✓	✗
Gumbel	$\frac{\theta - 1}{\theta}$	$2^{2-2^{1/\theta}} - 1$	✓	✗

Table 2.2: Summary of Kendall's tau, Blomqvist's beta, and the symmetry properties for the bivariate Frank, Clayton and Gumbel copula. Note: $D_k(x)$ is the Debye function, which is defined for any $k \in \mathbb{N}$ by $D_k(x) = \frac{k}{x^k} \int_0^x \frac{t^k}{e^t - 1} \, dt$.

2.6 Pseudo-observations and the empirical copula

Let $\boldsymbol{X} \in \mathbb{R}^d$ be a d-dimensional random vector with cumulative distribution function F and continuous univariate marginal distribution functions $F_1, \ldots, F_d$ and copula C. Further, let $\boldsymbol{X}_1 = (X_{11}, \ldots, X_{d1}), \ldots, \boldsymbol{X}_n = (X_{1n}, \ldots, X_{dn}) \in \mathbb{R}^d$ be a random sample of n independent observations from the random vector $\boldsymbol{X}$. If the marginal distribution functions $F_1, \ldots, F_d$ are known, copula data can be derived by computing $U_{ki} = F_k(X_{ki})$ for all $k \in \{1, \ldots, d\}$ and $i \in \{1, \ldots, n\}$. The corresponding empirical copula can be defined by

$$\widetilde{C}_n(\boldsymbol{u}) = \frac{1}{n} \sum_{i=1}^{n} I \left\{ U_{1i} \leq u_1, \ldots, U_{di} \leq u_d \right\} \,, \tag{2.9}$$

where $I\{\cdot, \dots, \cdot\}$ denotes the indicator function.

In practice, however, the marginal distribution functions $F_1, \dots, F_d$ are usually unknown. Dropping the assumption of known marginal distribution functions, it is a natural approach to estimate the unknown margins by their corresponding empirical counterparts $F_{1,n}, \dots, F_{d,n}$, where

$$F_{k,n}(x) = \frac{1}{n} \sum_{i=1}^{n} I\{X_{ki} \leq x\}, \quad \text{for } k \in \{1, \dots, d\} \text{ and } x \in \mathbb{R}. \tag{2.10}$$

Furthermore, let F_n denote the empirical joint cumulative distribution function of the sample. Inspired by Sklar's theorem (see Theorem 2.1), the empirical copula C_n of the sample is then defined by

$$C_n(\boldsymbol{u}) = F_n\left(F_{1,n}^{-}(u_1), \dots, F_{d,n}^{-}(u_d)\right), \quad \text{for } \boldsymbol{u} \in [0,1]^d. \tag{2.11}$$

In the case of unknown marginal distributions, pseudo-observations

$$\widehat{\boldsymbol{U}}_1 = (\widehat{U}_{11}, \dots, \widehat{U}_{d1}), \dots, \widehat{\boldsymbol{U}}_n = (\widehat{U}_{1n}, \dots, \widehat{U}_{dn}),$$

can be derived from the sample by defining

$$\widehat{U}_{ki} := \frac{n}{n+1} F_{k,n}(X_{ki}), \quad \text{for all } k \in \{1, \dots, d\} \text{ and } i \in \{1, \dots, n\}. \tag{2.12}$$

The empirical copula can then be defined as the empirical cumulative distribution function of the sample of pseudo-observations. Hence, another definition of the empirical copula is given by

$$\widehat{C}_n(\boldsymbol{u}) = \frac{1}{n} \sum_{i=1}^{n} I\{\widehat{U}_{1i} \leq u_1, \dots, \widehat{U}_{di} \leq u_d\}. \tag{2.13}$$

Note that the difference between C_n and $\widehat{C}_n$ is small and vanishes asymptotically. More precisely, according to Kojadinovic and Stemikovskaya (2019) it holds that

$$\sup_{\boldsymbol{u} \in [0,1]^d} |C_n(\boldsymbol{u}) - \widehat{C}_n(\boldsymbol{u})| \leq \frac{d}{n}.$$

Thus, for statistical testing the two definitions can be used equivalently since using C_n or $\widehat{C}_n$ does not affect the asymptotic behavior.

3

A simple non-parametric goodness-of-fit test for elliptical copulas

This chapter is similar to Jaser et al. (2017).

3.1 Introduction

Nowadays, copulas are a standard tool for modeling multivariate dependence. There exist many copula classes such as Archimedean, elliptical and Marshall-Olkin copulas (see, e.g., Mai and Scherer (2012)) and the choice of the right copula class is crucial for an accurate multivariate data analysis. Therefore, goodness-of-fit tests for copulas have been an objective of active research in recent years, see, e.g., Genest et al. (2009b), Berg (2009) and Fermanian (2013). In financial applications, elliptical copulas are commonly used to capture the dependence structure.

This chapter is concerned with the construction of a simple non-parametric goodness-of-fit test to examine whether the underlying dependence structure follows some elliptical copula of any dimension. Therefore, the null hypothesis that the unknown copula C of the given data belongs to the class of elliptical copulas $\mathcal{C}^{ellipt}$,

$$H_0 : C \in \mathcal{C}^{ellipt} ,$$

is tested against the alternative

$$H_1 : C \notin \mathcal{C}^{ellipt} .$$

In case of bivariate elliptical copulas, which are symmetric and radially symmetric, one could first use our simple non-parametric tests for symmetry and radial symmetry presented in Chapter 4 to statistically confirm both symmetry properties. If at least one of these statistical tests is rejected, then the bivariate copula of the underlying data cannot be elliptical. Therefore,we propose a multiple testing procedure for ellipticity of copula data in Chapter 4. Otherwise, a new statistical test is needed to identify bivariate elliptical copulas within symmetric and radially symmetric copulas. In case of multivariate elliptical copulas, one could test only for radial symmetry.

Li and Peng (2009) construct a goodness-of-fit test for the tail copula of a d-dimensional distribution, whose dependence structure is expressed by an elliptical copula. Klüppelberg et al. (2008) derive, in Lemma 1, the parametric form of the tail copula of elliptical distributions and argue, in Section 2, that it depends only on the underlying elliptical copula and is independent of the marginal distributions. Hence, the test in Li and Peng (2009) can also be seen as a goodness-of-fit test for elliptical copulas. To the best of our knowledge, this was the only goodness-of-fit test for elliptical copulas of any dimension d prior to our research. However, this test utilizes the tail dependence concept and therefore, the class of copulas for the null hypothesis has to be restricted to elliptical copulas with positive tail dependence. This tail dependence assumption discards for example the Gaussian copula from the null hypothesis and consequently shrinks the class of elliptical copulas under consideration. Furthermore, the test is based on the upper order statistics of the data and therefore has to deal with the difficulties of extreme value statistics. We propose a new simple non-parametric goodness-of-fit test, which takes into account the dependence structure of the whole data set. In particular, it is based on the equality of Kendall's tau and Blomqvist's beta for all bivariate margins of meta-elliptical distributions resulting from Fang et al. (2002) and Schmid and Schmidt (2007).

Elliptical copulas are specified by their generator function and parameters. If the choice of the generator function is fixed, many general goodness-of-fit tests can be used to test whether an underlying copula belongs to this specified subclass of elliptical copulas. However, the choice of the generator function is not an obvious and simple task. Our goodness-of-fit test does not require the knowledge of the generator function and in this sense, it is general. Moreover, it is simple since its critical values are directly computed from an asymptotic χ^2-distribution of a test statistic.

The remainder of this chapter is organized as follows. Section 3.2 discusses the relation of Kendall's tau and Blomqvist's beta under the null hypothesis. Section 3.3 presents our test statistic and its limiting χ^2-distribution under the assumption that the copula data comes from an elliptical family. Non-elliptical copula classes for the power study are given in Section 3.4. Section 3.5 provides the simulation study with numerical results on the nominal level of the test as well as on its power. Section 3.6 deals with an application of our goodness-of-fit test to real data, before Section 3.7 concludes. One technical proof is deferred to Section 3.8.

3.2 Relation between Kendall's tau and Blomqvist's beta for elliptical distributions and copulas

Kendall's tau and Blomqvist's beta are fundamental for our test statistic. The test will be based on the dependence between all bivariate pairs of the components of the random vector $\boldsymbol{X} \in \mathbb{R}^d$. In Theorem 3.1 of Fang et al. (2002), it is proven that the classical relation between Kendall's tau and the linear correlation coefficient known for bivariate normal distributions is valid within the more general class of meta-elliptical distributions. In particular, let (X_k, X_ℓ) be a meta-elliptically distributed random vector with association $\rho_{k\ell}$, which coincides with the correlation between X_k and X_ℓ in case of finite second moments of the latter two. Then, the following relation between Kendall's tau $\tau_{k\ell}$ and $\rho_{k\ell}$ holds:

$$\tau_{k\ell} = \frac{2}{\pi} \arcsin(\rho_{k\ell}) \,. \tag{3.1}$$

Furthermore, Proposition 8 in Schmid and Schmidt (2007) implies a similar result for Blomqvist's beta $\beta_{k\ell}$ and $\rho_{k\ell}$:

$$\beta_{k\ell} = \frac{2}{\pi} \arcsin(\rho_{k\ell}) \,. \tag{3.2}$$

Equations (3.1) and (3.2) show that Kendall's tau $\tau_{k\ell}$ and Blomqvist's beta $\beta_{k\ell}$ are uniquely determined by the association $\rho_{k\ell}$ for bivariate meta-elliptical distributions. Second, they coincide. The equality of Kendall's tau and Blomqvist's beta is an intrinsic property of meta-elliptical distributions and therefore of elliptical copulas. Hence, we build our goodness-of-fit test on this characteristic of elliptical copulas. To the best of our knowledge, such a simple goodness-of-fit test has not been considered in the literature so far.

Note that the equality of Kendall's tau and Blomqvist's beta is a necessary but not sufficient condition of elliptical copulas and, therefore, does not completely characterize them. More precisely, given an elliptical copula Kendall's tau and Blomqvist's beta are equal but the converse is not true. This is shown by the following example. Let U, V be independent random variables both uniformly distributed on $[0, 1]$. Then, set

$$(U_1, U_2) = \begin{cases} (U, V/4)\,, & \text{if } U \leq 1/4 \\ (U, V/4 + 3/4), & \text{if } 1/4 < U \leq 1/2 \\ (U, V/4 + 1/2), & \text{if } 1/2 < U \leq 3/4 \\ (U, V/4 + 1/4), & \text{if } 3/4 < U \leq 1\,. \end{cases} \tag{3.3}$$

Now, let the copula C be the distribution function of (U_1, U_2) defined in Equation (3.3). A scatter plot of a random sample of size 1000 from the copula C is illustrated in Figure 3.1. It is clear that C is non-elliptical. Furthermore, Kendall's tau and Blomqvist's beta can be computed and we get

$$\tau = \beta = 0\,. \tag{3.4}$$

The proof of Equation (3.4) can be found in Section 3.8.

3.3 Goodness-of-fit test for elliptical copulas

In financial applications, it is often assumed that a copula C belongs to the class of elliptical copulas. Therefore, our aim is to provide a statistical test to verify this

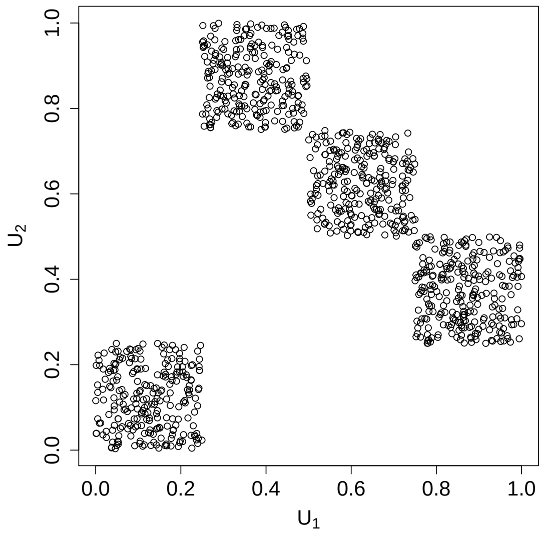

Figure 3.1: Scatter plot of a random sample of size 1000 from the copula C corresponding to the random vector defined in Equation (3.3).

assumption. From now on, we assume that we are given a copula sample and neglect unknown marginal distribution functions and their estimation. In practice, marginal distribution functions can be estimated parametrically and non-parametrically, which will affect the statistical inference of the test statistic. This subject is addressed in Chapter 5.

Let $\boldsymbol{U}_1, \ldots, \boldsymbol{U_n} \in [0,1]^d$ be a sample from the statistical model

$$\left(([0,1]^d)^n, \mathcal{B}([0,1]^d)^{\otimes n}, P^{\otimes n}\right),$$

where P is a distribution with copula C and uniform margins. Under the hypothesis of an elliptical copula C, also all marginal copulas have to be elliptical. We construct our test on the equality of Kendall's tau $\tau_{C_{k\ell}}$ and Blomqvist's beta $\beta_{C_{k\ell}}$ given by

$$\tau_{C_{k\ell}} = \beta_{C_{k\ell}}\,, \tag{3.5}$$

for all pairs $k, \ell \in \{1, \ldots, d\}$ with $k < \ell$. By virtue of (3.5), our test statistic will be constructed using the difference between the empirically estimated Blomqvist's beta and Kendall's tau. Asymptotic distributions of the empirical estimators for Kendall's tau and Blomqvist's beta are well known and reviewed below.

First, we outline the derivation of the asymptotic distribution of the Kendall's tau estimator. According to (2.2), an unbiased estimator of $\tau_{k\ell}$ is given by

$$\tau_{k\ell,n} = \frac{2}{n(n-1)} \sum_{1\le i<j\le n} \operatorname{sgn}(U_{ki} - U_{kj})\operatorname{sgn}(U_{\ell i} - U_{\ell j})\,. \tag{3.6}$$

The estimator $\tau_{k\ell,n}$ is a U-statistic and Hoeffding (1948) showed that $\sqrt{n}(\tau_{k\ell,n} - \tau_{k\ell})$ converges weakly to a centered Gaussian random variable with variance

$$\sigma^2_{\tau_{k\ell}} := \mathbb{V}\mathrm{ar}(2\tilde{h}_{k\ell,1}((U_{k1}, U_{\ell 1}))) ,$$

where

$$\tilde{h}_{k\ell,1}((U_{k1}, U_{\ell 1})) := \mathbb{E}\left[\mathrm{sgn}(U_{k1} - U_{k2})\mathrm{sgn}(U_{\ell 1} - U_{\ell 2})|U_{k1}, U_{\ell 1}\right] .$$

If the copula $C_{k\ell}$ is assumed to be known, then $\tilde{h}_{k\ell,1}$ has the following representation

$$\tilde{h}_{k\ell,1}((U_{k1}, U_{\ell 1})) = 1 - 2U_{k1} - 2U_{\ell 1} + 4C_{k\ell}(U_{k1}, U_{\ell 1}) \tag{3.7}$$

and $\sigma^2_{\tau_{k\ell}}$ can be represented through the copula $C_{k\ell}$ (see Theorem 4.3 in Dengler (2010)) as

$$\begin{aligned}\sigma^2_{\tau_{k\ell}} &= 64\mathbb{E}\left[C^2_{k\ell}(U_{k1}, U_{\ell 1})\right] - 64\mathbb{E}\left[U_{k1}C_{k\ell}(U_{k1}, U_{\ell 1})\right] - 64\mathbb{E}\left[U_{\ell 1}C_{k\ell}(U_{k1}, U_{\ell 1})\right] \\ &\quad + 32\mathbb{E}\left[C_{k\ell}(U_{k1}, U_{\ell 1})\right] + 16\mathbb{E}\left[U^2_{k1}\right] + 16\mathbb{E}\left[U^2_{\ell 1}\right] - 16\mathbb{E}\left[U_{k1}\right] - 16\mathbb{E}\left[U_{\ell 1}\right] \\ &\quad + 32\mathbb{E}\left[U_{k1}U_{\ell 1}\right] + 1 - 4\tau^2_{k\ell} .\end{aligned} \tag{3.8}$$

The variance $\sigma^2_{\tau_{k\ell}}$ can be further simplified using the theoretical moments of uniformly distributed random variables and Equation (2.1) for Kendall's tau. We get

$$\begin{aligned}\sigma^2_{\tau_{k\ell}} &= 64\mathbb{E}\left[C^2_{k\ell}(U_{k1}, U_{\ell 1})\right] - 64\mathbb{E}\left[U_{k1}C_{k\ell}(U_{k1}, U_{\ell 1})\right] - 64\mathbb{E}\left[U_{\ell 1}C_{k\ell}(U_{k1}, U_{\ell 1})\right] \\ &\quad + 32\mathbb{E}\left[U_{k1}U_{\ell 1}\right] + \frac{20}{3} + 8\tau_{k\ell} - 4\tau^2_{k\ell} .\end{aligned} \tag{3.9}$$

If we do not impose any parametric assumption on the copula $C_{k\ell}$, the asymptotic variance from (3.9) needs to be estimated non-parametrically. For this, each expectation involving $C_{k\ell}$ can be consistently estimated with the corresponding V-statistic (see Denker (1985) or von Mises (1947)) by employing the empirical copula $\widetilde{C}_{k\ell,n}$ defined in Equation (2.9). The remaining mixed moment can be consistently estimated by the corresponding empirical moment and $\tau_{k\ell}$ can be estimated by $\tau_{k\ell,n}$ from (3.6). However, this framework cannot ensure a positive variance estimate, since $\sigma^2_{\tau_{k\ell}}$ from (3.9) has been computed using theoretical moments of the uniform distribution as well as Equation (2.1). If we additionally estimate the moments of the uniform distribution in Equation (3.8) empirically, then the resulting variance estimate can still be negative due to the direct estimation of $\tau_{k\ell}$.

Below, we describe our estimation framework for $\sigma^2_{\tau_{k\ell}}$, which is the variance of

$$2\tilde{h}_{k\ell,1}((U_{k1}, U_{\ell 1})) .$$

For a sample $(U_{k1}, U_{\ell 1})$, …, $(U_{kn}, U_{\ell n})$, we propose to estimate $\tilde{h}_{k\ell,1}((U_{ki}, U_{\ell i}))$ non-parametrically by

$$\widehat{h}_{k\ell,1}((U_{ki}, U_{\ell i})) = 1 - 2U_{ki} - 2U_{\ell i} + 4C_{k\ell,n}(U_{ki}, U_{\ell i}) , \quad i \in \{1, \ldots, n\} . \tag{3.10}$$

Now, $\sigma^2_{\tau_{k\ell}}$ is estimated by the sample variance of

$$2\widehat{h}_{k\ell,1}((U_{k1}, U_{\ell 1})) \, , \dots \, , 2\widehat{h}_{k\ell,1}((U_{kn}, U_{\ell n})) \, .$$

This leads to a consistent and positive estimation of $\sigma^2_{\tau_{k\ell}}$. Consistency follows again from the consistency of the corresponding V-statistics resulting from the empirical copula $\widetilde{C}_{k\ell,n}$ combined with the estimation of moments. Note that our variance estimate is equivalent to the estimate based on Equation (3.8), when $\tau_{k\ell}$ is estimated using the empirical copula $\widetilde{C}_{k\ell,n}$.

For copula data $(U_{k1}, U_{\ell 1}), \dots, (U_{kn}, U_{\ell n})$, the marginal medians are known to be equal to 0.5. Resulting from (2.4), an empirical estimator for Blomqvist's beta $\beta_{k\ell}$ is given by

$$\beta^\star_{k\ell,n} = \frac{1}{n} \sum_{i=1}^{n} \operatorname{sgn}\left(U_{ki} - 0.5\right) \operatorname{sgn}\left(U_{\ell i} - 0.5\right) \, .$$

The asymptotic normality of the estimator $\beta^\star_{k\ell,n}$ of Blomqvist's beta follows in the case of known marginal distributions trivially from the central limit theorem and was already stated in Blomqvist (1950). Thus, we have the following result

$$\sqrt{n}\left(\beta^\star_{k\ell,n} - \beta_{k\ell}\right) \xrightarrow{d} N(0, \sigma^2_{\beta_{k\ell}}) \, ,$$

where

$$\sigma^2_{\beta_{k\ell}} = \mathbb{V}\mathrm{ar}\left[\operatorname{sgn}\left(U_{k1} - 0.5\right) \operatorname{sgn}\left(U_{\ell 1} - 0.5\right)\right] = 1 - \beta^2_{k\ell}$$

and $\xrightarrow{d}$ denotes convergence in distribution.

Now, we know how to estimate Kendall's tau and Blomqvist's beta for each pair (k, ℓ) of coordinates. The test statistic will be based on all $d(d-1)/2$ differences between the corresponding estimators for Kendall's tau and Blomqvist's beta. Hence, we define the statistic $\boldsymbol{D}_n$

$$\boldsymbol{D}_n := \operatorname{vec}_u(\boldsymbol{\beta}_n) - \operatorname{vec}_u(\boldsymbol{\tau}_n) \, , \tag{3.11}$$

in terms of the matrices $\boldsymbol{\beta}^\star_n := (\beta^\star_{k\ell,n})_{k,\ell \in \{1,\dots,d\}}$ and $\boldsymbol{\tau}_n := (\tau_{k\ell,n})_{k,\ell \in \{1,\dots,d\}}$, where $\beta^\star_{kk,n} = \tau_{kk,n} := 1$ and $\operatorname{vec}_u(\mathcal{A})$ is the vectorization operator that extracts the elements strictly above the main diagonal of a matrix $\mathcal{A} \in \mathbb{S}_d$ in a row-wise manner, that is

$$\operatorname{vec}_u(\mathcal{A}) := (a_{12}, a_{13}, \dots, a_{1d}, a_{23}, a_{24}, \dots, a_{2d}, \dots, a_{d-1,d}).$$

The following theorem contains the asymptotic distribution of $\boldsymbol{D}_n$ for a sample from an elliptical copula. Moreover, it states our test statistic T_n and its limiting distribution under the null hypothesis $C \in \mathcal{C}^{ellipt}$.

Theorem 3.1.
Let $\boldsymbol{U}_1, \dots, \boldsymbol{U}_n \in [0,1]^d$ *be a sample from the statistical model* $\Big(([0,1]^d)^n, \ \mathcal{B}([0,1]^d)^{\otimes n},$

$P^{\otimes n}\Big)$, *where P is a distribution with elliptical copula C and uniform margins. Then, the statistic $\boldsymbol{D}_n$ defined in (3.11) has the following asymptotic distribution*

$$\sqrt{n} \cdot \boldsymbol{D}_n \xrightarrow{d} N\left(0, \mathcal{V}\right)$$

with

$$\mathcal{V} = \begin{pmatrix} \mathcal{I}_{d(d-1)/2} & -\mathcal{I}_{d(d-1)/2} \end{pmatrix} \boldsymbol{\Sigma} \begin{pmatrix} \mathcal{I}_{d(d-1)/2} \\ -\mathcal{I}_{d(d-1)/2} \end{pmatrix},$$

where $\boldsymbol{\Sigma}$ is defined in Equation (3.15) and $\mathcal{I}_{d(d-1)/2}$ is the unit matrix of dimension $d(d-1)/2$.

Now, let $\mathcal{V}_n$ be a consistent estimator of $\mathcal{V}$ and consider the Wald-type statistic

$$T_n := n\boldsymbol{D}_n^\top \mathcal{V}_n^{-1} \boldsymbol{D}_n \,. \tag{3.12}$$

Then, it holds that

$$T_n \xrightarrow{d} \chi^2_{d(d-1)/2} \,,$$

where χ^2_m denotes the χ^2-distribution with m degrees of freedom.

Proof. Let $d \geq 2$ be the dimension of the sample $\boldsymbol{U}_1, \ldots, \boldsymbol{U_n} \in [0,1]^d$ from the statistical model

$$\Big(([0,1]^d)^n, \mathcal{B}([0,1]^d)^{\otimes n}, P^{\otimes n}\Big) \,,$$

where P is a distribution with copula C and uniform marginals. Next, we define the matrices

$$\mathcal{U}_i := (\operatorname{sgn}(U_{ki} - 0.5)\operatorname{sgn}(U_{\ell i} - 0.5))_{k,\ell \in \{1,\ldots,d\}} \in \mathbb{S}_d$$

and

$$\mathcal{H}_i := \Big(2\tilde{h}_{k\ell,1}((U_{ki}, U_{\ell i}))\Big)_{k,\ell \in \{1,\ldots,d\}} \in \mathbb{S}_d$$

as well as

$$\boldsymbol{\beta} := (\beta_{k\ell})_{k,\ell \in \{1,\ldots,d\}} \in \mathbb{S}_d$$

and

$$\boldsymbol{\tau} := (\tau_{k\ell})_{k,\ell \in \{1,\ldots,d\}} \in \mathbb{S}_d \,,$$

where $\beta_{kk} := 1$ and $\tau_{kk} := 1$. Using the matrices $\mathcal{U}_i$ and $\mathcal{H}_i$, we define the vectors $\boldsymbol{Z}_i^\beta := \operatorname{vec}_u(\mathcal{U}_i)$ and $\boldsymbol{Z}_i^\tau := \operatorname{vec}_u(\mathcal{H}_i)$.

Now, we consider the empirical estimator

$$\tau_{k\ell,n} = \frac{2}{n(n-1)} \sum_{1 \leq i < j \leq n} \operatorname{sgn}(U_{ki} - U_{kj})\operatorname{sgn}(U_{\ell i} - U_{\ell j}) \tag{3.13}$$

of $\tau_{k\ell}$, which is a U-statistic of degree two with kernel function $h\left(\boldsymbol{w}_1, \boldsymbol{w}_2\right) := \operatorname{sgn}(w_{11} - w_{12})\operatorname{sgn}(w_{21} - w_{22})$, where $\boldsymbol{w}_i = (w_{1i}, w_{2i})$, for $i = 1, 2$. Hoeffding's decomposition for U-statistics implies (see Theorem 1.2.1 in Denker (1985)) that $\tau_{k\ell,n} - \tau_{k\ell}$ can be represented as $2U_{k\ell,n1} + U_{k\ell,n2}$, where

$$U_{k\ell,n1} := \frac{1}{n}\sum_{i=1}^{n}\left(\tilde{h}_{k\ell,1}((U_{ki}, U_{\ell i})) - \tau_{k\ell}\right)$$

and $U_{k\ell,n2} := (\tau_{k\ell,n} - \tau_{k\ell}) - 2U_{k\ell,n1}$. Note that $U_{k\ell,n2}$ is a U-statistic of degree two with the degenerate kernel $h_{k\ell,2}(\boldsymbol{w}_1, \boldsymbol{w}_2) = h(\boldsymbol{w}_1, \boldsymbol{w}_2) - \tilde{h}_{k\ell,1}(\boldsymbol{w}_1) - \tilde{h}_{k\ell,1}(\boldsymbol{w}_2) + \tau_{k\ell}$, that is

$$\mathbb{E}\left[h_{k\ell,2}\left((U_{k1}, U_{\ell 1}), (U_{k2}, U_{\ell 2})\right) | (U_{k1}, U_{\ell 1})\right] = \mathbb{E}\left[h_{k\ell,2}\left(\boldsymbol{W}_1, \boldsymbol{W}_2\right) | \boldsymbol{W}_1\right] = 0$$

almost surely with independent and identically distributed (i.i.d.) $\boldsymbol{W}_1, \boldsymbol{W}_2 \sim C_{k\ell}$. From Theorem 1.2.4 in Denker (1985) it follows that

$$\mathbb{E}\left[(\sqrt{n}\cdot U_{k\ell,n2})^2\right] \leq \frac{A_{k\ell,h}}{n}, \tag{3.14}$$

where $A_{k\ell,h}$ is a constant depending only on the kernel $h(\cdot,\cdot)$. Therefore, $\sqrt{n}\cdot U_{k\ell,n2} \xrightarrow{L_2} 0$ as $n \longrightarrow \infty$, and

$$\sqrt{n}\cdot(\tau_{k\ell,n} - \tau_{k\ell}) \quad \text{and} \quad \sqrt{n}\cdot 2U_{k\ell,n1} = \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} 2\tilde{h}_{k\ell,1}((U_{ki}, U_{\ell i})) - 2\tau_{k\ell}\right)$$

have the same limiting normal distribution. The multivariate central limit theorem implies

$$\sqrt{n}\left(\overline{\boldsymbol{Z}}_n - \begin{pmatrix}\mathrm{vec}_u(\boldsymbol{\beta})\\ \mathrm{vec}_u(\boldsymbol{\tau})\end{pmatrix}\right) \xrightarrow{d} N(\boldsymbol{0}, \boldsymbol{\Sigma}),$$

where $\boldsymbol{Z}_i := \begin{pmatrix}\boldsymbol{Z}_i^{\beta}\\ \boldsymbol{Z}_i^{\tau}\end{pmatrix}$, $\overline{\boldsymbol{Z}}_n = \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{Z}_i$ and

$$\boldsymbol{\Sigma} = \mathbb{C}\mathrm{ov}(\boldsymbol{Z}_1). \tag{3.15}$$

For example, the covariance matrix $\boldsymbol{\Sigma}$ for $d = 2$, $k = 1$ and $l = 2$ has the following form

$$\begin{pmatrix}\mathbb{V}\mathrm{ar}\left(\mathrm{sgn}(U_{11}-0.5)\mathrm{sgn}(U_{21}-0.5)\right) & \mathbb{C}\mathrm{ov}\left(\mathrm{sgn}(U_{11}-0.5)\mathrm{sgn}(U_{21}-0.5), 2\tilde{h}_1((U_{11},U_{21}))\right)\\ \mathbb{C}\mathrm{ov}\left(\mathrm{sgn}(U_{11}-0.5)\mathrm{sgn}(U_{21}-0.5), 2\tilde{h}_1((U_{11},U_{21}))\right) & \mathbb{V}\mathrm{ar}\left(2\tilde{h}_1((U_{11},U_{21}))\right)\end{pmatrix}.$$

For an elliptical copula C, the multivariate statistic $\boldsymbol{D}_n$ is then equal to

$$\overline{\boldsymbol{Z}^{\beta}}_n - \mathrm{vec}_u\left((\tau_{k\ell,n})_{k,\ell\in\{1,\ldots,d\}}\right),$$

which has the same limiting distribution as

$$\overline{\boldsymbol{Z}^{\beta}}_n - \overline{\boldsymbol{Z}^{\tau}}_n.$$

By applying the Delta method (see, e.g., Proposition 6.2 in Bilodeau and Brenner (1999)) with

$$\phi : \mathbb{R}^{d(d-1)} \to \mathbb{R}^{d(d-1)/2}, \boldsymbol{x} \mapsto (x_1, \ldots, x_{d(d-1)/2}) - (x_{d(d-1)/2+1}, \ldots, x_{d(d-1)}),$$

we obtain

$$\sqrt{n}\left(\overline{\boldsymbol{Z}^{\beta}}_n - \overline{\boldsymbol{Z}^{\tau}}_n\right) \xrightarrow{d} N(\boldsymbol{0}, \mathcal{V})$$

under the null hypothesis $C \in \mathcal{C}^{ellipt}$, where

$$\mathcal{V} = \phi'\left(\begin{pmatrix}\text{vec}_u(\boldsymbol{\beta})\\ \text{vec}_u(\boldsymbol{\tau})\end{pmatrix}\right)\boldsymbol{\Sigma}\phi'\left(\begin{pmatrix}\text{vec}_u(\boldsymbol{\beta})\\ \text{vec}_u(\boldsymbol{\tau})\end{pmatrix}\right)^{\top}$$

and ϕ' denotes the Jacobian matrix of ϕ. Since ϕ is a linear map, ϕ' is independent of $\boldsymbol{\beta}$ and $\boldsymbol{\tau}$. Moreover, it is given by

$$\begin{pmatrix}\mathcal{I}_{d(d-1)/2} & -\mathcal{I}_{d(d-1)/2}\end{pmatrix},$$

where $\mathcal{I}_{d(d-1)/2}$ is the unit matrix of dimension $d(d-1)/2$.

The second statement of the theorem is obvious. The asymptotic distribution of T_n defined in (3.12) follows from the asymptotic normality of $\boldsymbol{D}_n$, the multivariate Slutsky Theorem (see, e.g., Lemma 6.3 in Bilodeau and Brenner (1999)) and the continuous mapping theorem.

□

The second result of Theorem 3.1 depends on a consistent estimator of the covariance matrix $\mathcal{V}$ since $\boldsymbol{\Sigma}$ is unknown. In the following remark, we indicate the construction of such a consistent estimator $\mathcal{V}_n$.

Remark 3.2. *The asymptotic covariance matrix $\boldsymbol{\Sigma}$ depends on the unobserved*

$$\tilde{h}_{k\ell,1}((U_{k1}, U_{\ell 1})),$$

for $k, \ell \in \{1, \ldots, d\}$ and $k \neq \ell$. However, $\boldsymbol{\Sigma}$ can be consistently estimated using

$$\widehat{h}_{k\ell,1}((U_{ki}, U_{\ell i})),$$

$i = 1, \ldots, n$, defined in (3.10). This results in the consistent estimator $\mathcal{V}_n$ of the covariance matrix $\mathcal{V}$.

Based on Theorem 3.1, we propose the test function

$$\delta(\boldsymbol{U}_1, \ldots, \boldsymbol{U}_n) = I\{T_n > \chi^2_{d(d-1)/2, 1-\alpha}\}$$

to test

$$H_0 : C \in \mathcal{C}^{ellipt} \qquad \text{against} \qquad H_1 : C \notin \mathcal{C}^{ellipt},$$

where $\chi^2_{m,\alpha}$ denotes the α-quantile of the χ^2-distribution with m degrees of freedom.

3.4 Non-elliptical copula classes for the power study

In the following, we briefly overview copulas based on special mixtures of elliptical copulas or elliptical distributions, which constitute two non-elliptical copula classes used for the power study.

3.4.1 Mixture of bivariate elliptical copulas

The aim of this section is to consider another class of bivariate non-elliptical copulas, which are symmetric (or exchangeable) and radially symmetric. For this, we mix two bivariate elliptical copulas with different parameters. More precisely, a bivariate Gaussian copula with correlation ρ_G and a bivariate t copula with ν degrees of freedom and association parameter ρ_t, where $\rho_t \neq \rho_G$, are mixed with probabilities $p \in [0,1]$ and $1-p$, respectively. The resulting bivariate mixture copula is given by

$$C^{mixt,cop}(u,v) = p\, C^{Gauss}_{\rho_G}(u,v) + (1-p)\, C^{t}_{\nu,\rho_t}(u,v), \quad (u,v) \in [0,1]^2 .$$

By choosing $\rho_G \neq \rho_t$, we expect this bivariate mixture copula to be non-elliptical. However, this is not trivial to show since elliptical copulas are only implicitly defined as the copulas of elliptical distributions. To the best of our knowledge, such mixtures of elliptical copulas have not been investigated so far.

It should be noted that the proposed construction of such mixture copulas is general and can be extended to any dimension. Further, it is very easy to draw a random sample from the mixture copula. For this, the random sample is drawn from the Gaussian copula $C^{Gauss}_{\rho_G}$ with probability p and with probability $(1-p)$ from the t copula C^{t}_{ν,ρ_t}. In our simulation study, we set $p = 0.5$, $\nu = 5$ and varied the association parameters ρ_G and ρ_t. By virtue of the one-to-one correspondence between Kendall's tau and the association parameter ρ (correlation coefficient for $\nu \geq 2$) given in (3.1), this is equivalent to varying Kendall's tau.

3.4.2 Copulas derived from the mixture of bivariate elliptical distributions

Here, we have a closer look on bivariate copulas derived from the mixture of bivariate elliptical distributions. Again, the framework presented below is general and can be extended to any dimension. The idea is to mix two bivariate elliptical distributions in such a way that the resulting bivariate distribution is not elliptical any more. We expect that its copula is then also non-elliptical, but we have no theoretical justification. Without loss of generality, we set $\boldsymbol{\mu} = \mathbf{0}$ in Definition 2.7. Now, one can easily argue that the mixture of two bivariate elliptical distributions with different parameters $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ is not elliptical.

In the following, a bivariate Gaussian distribution $N_2(\mathbf{0}, \mathcal{P}_G)$ with correlation ρ_G and a bivariate t distribution $t_2(\nu, \mathbf{0}, \mathcal{P}_t)$ with correlation ρ_t, where $\rho_t \neq \rho_G$, are mixed with probabilities $p \in [0,1]$ and $1-p$, respectively. The cumulative distribution function F^{mixt} of the resulting bivariate mixture distribution is given by

$$F^{mixt}(x,y) = p\,\Phi_{\rho_G}(x,y) + (1-p)\,t_{\nu,\rho_t}(x,y), \quad (x,y) \in \mathbb{R}^2\,,$$

where Φ_{ρ_G} and t_{ν,ρ_t} are the cumulative distribution functions of $N_2(\mathbf{0}, \mathcal{P}_G)$ and $t_2(\nu, \mathbf{0}, \mathcal{P}_t)$, respectively. The margins F_1^{mixt} and F_2^{mixt} of this mixture distribution can be determined using the margins of the underlying Gaussian and t distribution. Then, according to Sklar's theorem (see Theorem 2.1), the bivariate copula $C^{mixt,distr}(u,v)$ of the mixture distribution F^{mixt}, for any $u, v \in [0,1]$, is given by

$$C^{mixt,distr}(u,v) = F^{mixt}((F_1^{mixt})^{-}(u), (F_2^{mixt})^{-}(v))\,,$$

where $(F_1^{mixt})^{-}$ and $(F_2^{mixt})^{-}$ denote the generalized inverses of F_1^{mixt} and F_2^{mixt}, respectively. Since we chose $\rho_G \neq \rho_t$, the resulting bivariate mixture distribution F^{mixt} is non-elliptical.

Just like for the mixture of elliptical copulas, it is easy to draw a random sample from the mixture distribution. First, the random sample is drawn with probability p from the bivariate Gaussian distribution $N_2(\mathbf{0}, \mathcal{P}_G)$ and with probability $(1-p)$ from the bivariate t distribution $t_2(\nu, \mathbf{0}, \mathcal{P}_t)$.Then, the random sample is transformed using the margins F_1^{mixt} and F_2^{mixt} to get copula data. For the simulation study, we set again $p = 0.5$, $\nu = 5$ and varied the association parameters ρ_G and ρ_t. With the same argument as before, this is equivalent to varying Kendall's tau.

3.5 Simulation study

In order to assess the finite-sample performance of the proposed test for ellipticity based on the test statistic T_n, a Monte Carlo study was conducted. We are interested in the ability of the test to hold its nominal level as well as the power of the test to detect alternatives. For ease of notation we skip all indices in the bivariate examples and just use them when they are needed.

3.5.1 Setup

First of all, we fixed a significance level of $\alpha = 0.05$ for the test throughout the study. Furthermore, the number of Monte Carlo replications was set to $N = 1000$. The simulation study was then carried out for different dimensions d, copula families, levels of dependence (measured in terms of Kendall's tau) and sample sizes. In particular, we have considered samples of dimension $d = 2, 3$ and 6.

To investigate the level of the test, random samples from two elliptical copula families were considered, namely the Gaussian copula and the t copula with 5 degrees of freedom ($t_{\nu=5}$). To study the power of the test, random samples from non-elliptical copula families were examined (see Section 3.4). Here, we looked at random samples from the Frank, Clayton and Gumbel family as well as from a mixture of two elliptical copulas and a copula derived from the mixture of two elliptical distributions with different association parameters, respectively. For the mixtures, we chose a Gaussian and a t copula as well as a Gaussian and a t distribution, respectively.

In order to assess the effect of the strength of dependence, five different levels of dependence were chosen, according to $\tau \in \{0.1, 0.25, 0.5, 0.75, 0.9\}$. Each value of τ was converted to a unique association or dependence parameter of a multivariate copula. As a consequence, all bivariate marginal copulas of the resulting multivariate copula are then identical. For the copulas based on mixtures, four different levels of dependence were considered. The different levels are given by a combination of τ_G for the Gaussian copula/distribution and τ_t for the t copula/distribution. These parameters (τ_G, τ_t) had values in $\{(0.25, 0.75), (0.75, 0.25), (0.5, 0.25), (0.5, 0.75)\}$. Finally, for every choice of copula family and fixed level of dependence, random samples of size $n \in \{100, 250, 500, 1000, 5000\}$ were considered.

To get an impression of the common copula families used in the simulation study, Figure 3.2 displays scatter plots of bivariate random samples of size $n = 1000$ for the levels of dependence corresponding to $\tau \in \{0.25, 0.5, 0.75\}$. Further, scatter plots of the bivariate mixture copula and of the copula derived from the mixture of bivariate elliptical distributions are illustrated for the different combinations of τ_G and τ_t in Figure 3.3 and Figure 3.4, respectively. First, we would like to point out that the scatter plots for the Gaussian and the Frank copula in Figure 3.2 are quite difficult to distinguish. Moreover, the scatter plots for the mixtures in Figures 3.3 and 3.4 could easily be assigned erroneously to data from elliptical copulas.

3.5.2 Level

Tables 3.1, 3.2 and 3.3 display the empirical level of the test for ellipticity with significance level $\alpha = 0.05$ as observed in 1000 random samples for dimension $d = 2, 3$ and 6, respectively, and all possible scenarios from the simulation setup. Note that for $d = 3$ and $d = 6$, all off-diagonal elements of the correlation matrix $\mathcal{R}$ of the Gaussian and $t_{\nu=5}$ copula are identical and related to the level of dependence τ.

For dimensions $d = 2$ and $d = 3$, the test seems to hold its nominal level (see Tables 3.1 and 3.2). Only for large values of Kendall's tau in combination with a small sample size of $n = 100$ or $n = 250$, the test turns out to be too liberal. As the distributional result for the test statistic holds only asymptotically, this explains why there might occur some problems especially for small sample sizes.

Table 3.3 shows that the proposed test requires large sample sizes to hold its level for higher dimensions. For $d = 6$ and medium level of dependence τ, a sample size of at least $n = 1000$ is required. This can be explained by the asymptotic nature of our test. The accuracy of the distributional approximation with the limiting χ^2-distribution is very poor for small sample sizes and gets improved significantly for larger sample

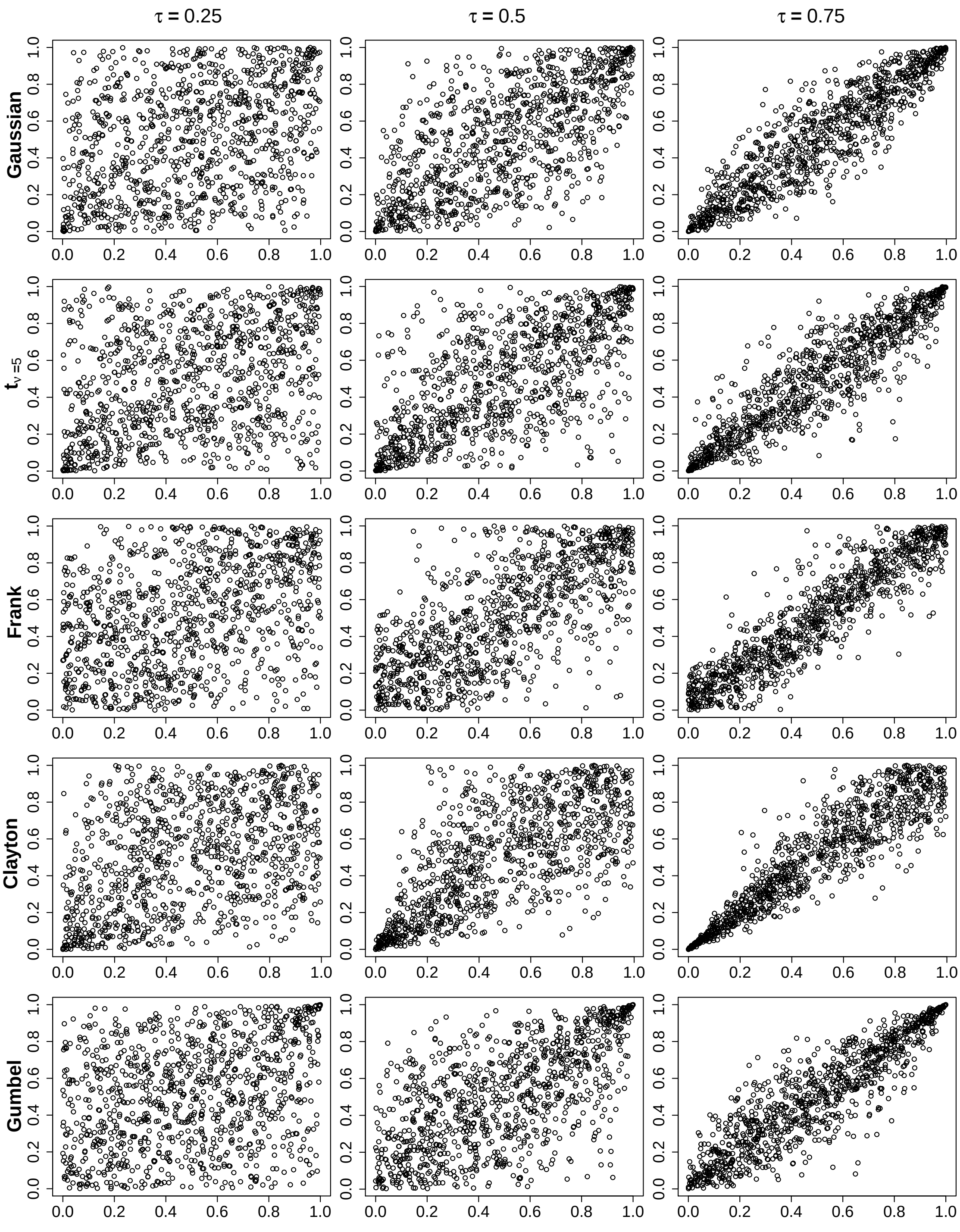

Figure 3.2: Scatter plots of random samples of size 1000 from the bivariate Gaussian, $t_{\nu=5}$, Frank, Clayton, and Gumbel copula (from top to bottom) with $\tau = 0.25$ (left), 0.5 (middle), 0.75 (right).

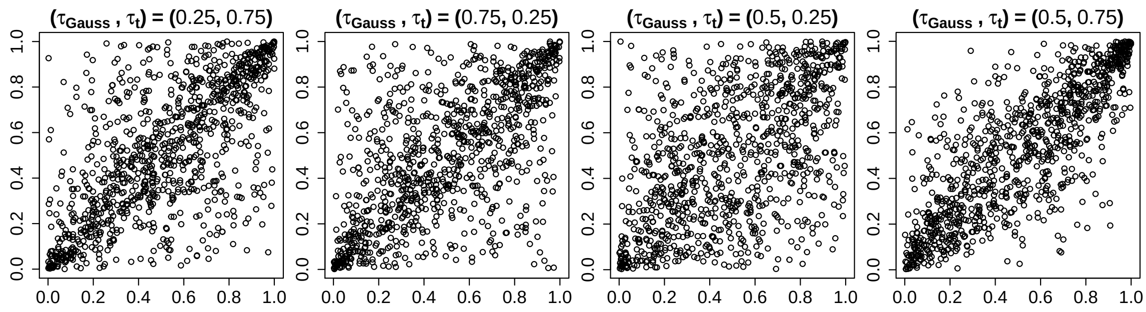

Figure 3.3: Scatter plots of random samples of size 1000 from the copula based on the mixture of two bivariate elliptical copulas with $(\tau_G, \tau_t) = (0.25, 0.75)$, $(0.75, 0.25)$, $(0.25, 0.5)$, and $(0.75, 0.5)$ (from left to right).

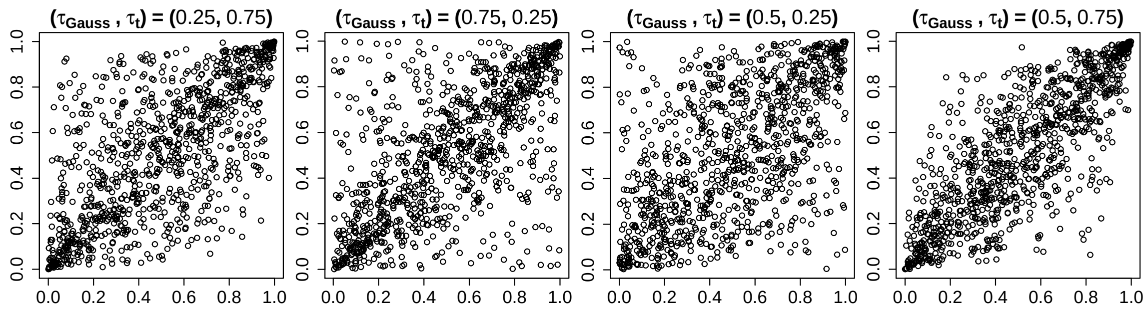

Figure 3.4: Scatter plots of random samples of size 1000 from the copula based on the mixture of two bivariate elliptical distributions with $(\tau_G, \tau_t) = (0.25, 0.75)$, $(0.75, 0.25)$, $(0.25, 0.5)$, and $(0.75, 0.5)$ (from left to right).

C	τ	$n = 100$	$n = 250$	$n = 500$	$n = 1000$	$n = 5000$
	0.10	0.046	0.052	0.046	0.061	0.044
	0.25	0.044	0.045	0.060	0.053	0.053
Gauss	0.50	0.063	0.044	0.044	0.055	0.048
	0.75	0.054	0.040	0.049	0.054	0.051
	0.90	0.090	0.047	0.053	0.053	0.064
	0.10	0.048	0.048	0.047	0.049	0.053
	0.25	0.049	0.051	0.042	0.043	0.053
$t_{\nu=5}$	0.50	0.047	0.042	0.034	0.051	0.047
	0.75	0.072	0.052	0.049	0.050	0.046
	0.90	0.077	0.046	0.035	0.054	0.059

Table 3.1: Dimension d=2: Empirical level of the test for ellipticity with significance level $\alpha = 0.05$ based on the test statistic T_n: rate of rejecting H_0 as observed in 1000 random samples of size n from copula family C with Kendall's tau τ.

C	τ	$n = 100$	$n = 250$	$n = 500$	$n = 1000$	$n = 5000$
	0.10	0.054	0.046	0.045	0.049	0.041
	0.25	0.054	0.044	0.062	0.055	0.045
Gauss	0.50	0.048	0.066	0.062	0.062	0.059
	0.75	0.081	0.057	0.055	0.053	0.045
	0.90	0.213	0.111	0.077	0.063	0.043
	0.10	0.047	0.048	0.042	0.039	0.048
	0.25	0.067	0.039	0.051	0.058	0.061
$t_{\nu=5}$	0.50	0.055	0.052	0.054	0.038	0.053
	0.75	0.071	0.058	0.045	0.054	0.042
	0.90	0.173	0.080	0.066	0.072	0.048

Table 3.2: Dimension d=3: Empirical level of the test for ellipticity with significance level $\alpha = 0.05$ based on the test statistic T_n: rate of rejecting H_0 as observed in 1000 random samples of size n from copula family C with Kendall's tau τ.

C	τ	$n = 100$	$n = 250$	$n = 500$	$n = 1000$	$n = 5000$
Gauss	0.10	0.155	0.092	0.069	0.063	0.043
	0.25	0.195	0.091	0.075	0.044	0.059
	0.50	0.266	0.126	0.074	0.060	0.067
	0.75	0.484	0.243	0.153	0.104	0.057
	0.90	0.642	0.591	0.403	0.193	0.075
$t_{\nu=5}$	0.10	0.167	0.069	0.070	0.054	0.059
	0.25	0.196	0.092	0.090	0.054	0.046
	0.50	0.244	0.122	0.067	0.054	0.037
	0.75	0.479	0.220	0.127	0.069	0.040
	0.90	0.501	0.528	0.285	0.154	0.056

Table 3.3: Dimension d=6: Empirical level of the test for ellipticity with significance level $\alpha = 0.05$ based on the test statistic T_n: rate of rejecting H_0 as observed in 1000 random samples of size n from copula family C with Kendall's tau τ.

sizes. This is illustrated by the QQ-plots for the $t_{\nu=5}$ copula in Figure 3.5. Hence, the results of our simulation study for dimension $d = 6$ are reliable only for large sample sizes. Therefore, we consider only samples of size $n = 1000$ and $n = 5000$ in the power study for dimension $d = 6$.

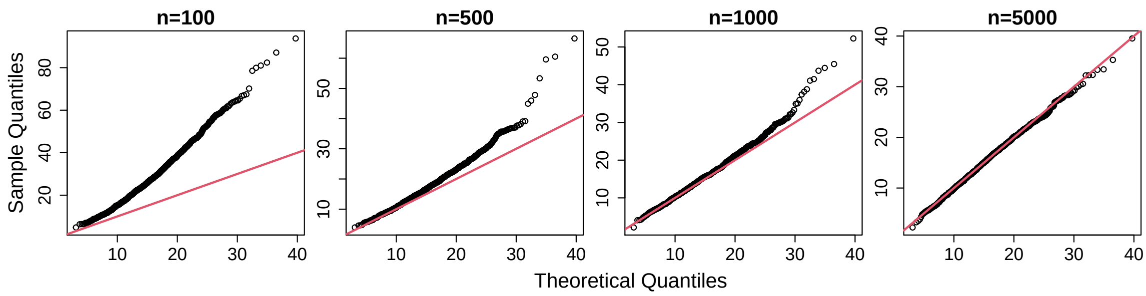

Figure 3.5: QQ-plots of T_n for $t_{\nu=5}$ copula, $d = 6$, $\tau = 0.75$, and $n = 100, 500, 1000,$ and 5000.

3.5.3 Power

The results for the empirical power of the test for ellipticity with significance level $\alpha = 0.05$ based on 1000 random samples from the Frank, Clayton and Gumbel family are presented in Tables 3.4, 3.5 and 3.6 for the different dimensions $d = 2, 3$ and 6. For the random samples from the mixture copula and the copula derived from elliptical distributions, we report the results only for $d = 2$ in Tables 3.7 and 3.8, respectively. This is due to the fact that huge sample sizes are generally needed to achieve satisfactory empirical power for the bivariate mixture copula constructions. This lacks in practical relevance and, therefore, we do not consider these mixture copulas in higher dimensions.

C	τ	$n = 100$	$n = 250$	$n = 500$	$n = 1000$	$n = 5000$
	0.10	0.063	0.072	0.063	0.074	0.243
	0.25	0.074	0.097	0.179	0.298	0.903
Frank	0.50	0.145	0.256	0.492	0.743	1.000
	0.75	0.157	0.341	0.567	0.854	1.000
	0.90	0.181	0.232	0.379	0.620	1.000
	0.10	0.049	0.052	0.052	0.056	0.058
	0.25	0.062	0.061	0.047	0.053	0.066
Clayton	0.50	0.053	0.050	0.075	0.102	0.228
	0.75	0.121	0.136	0.266	0.452	0.981
	0.90	0.169	0.194	0.288	0.514	0.986
	0.10	0.050	0.050	0.051	0.062	0.049
	0.25	0.047	0.052	0.037	0.067	0.067
Gumbel	0.50	0.056	0.053	0.044	0.029	0.055
	0.75	0.059	0.044	0.073	0.053	0.080
	0.90	0.088	0.046	0.065	0.074	0.086

Table 3.4: Dimension d=2: Empirical power of the test for ellipticity with significance level $\alpha = 0.05$ based on the test statistic T_n: rate of rejecting H_0 as observed in 1000 random samples of size n from copula family C with Kendall's tau τ.

First of all, when we look at Tables 3.4 - 3.8, we notice that the observed power varies enormously across the level of dependence and the sample size as well as across the copula families. In general, the rejection rate increases with the sample size, as expected. In addition to that, the rejection rate increases with the level of dependence. Since the non-ellipticity becomes more apparent for higher values of Kendall's tau, which can also be observed in Figure 3.2, this is also expected. The empirical power also increases with increasing dimension as soon as the distributional approximation with the χ^2 distribution is sufficiently accurate. Some exceptions occur in connection with the Gumbel family, which we discuss later on.

Power for Archimedean copula families

For the Frank copula, the test appears to perform well for all considered dimensions. If Kendall's tau has a value of at least 0.5, a sample size of $n = 1000$ suffices to achieve a good power. For dimension $d = 2$ and $d = 3$ and small levels of dependence, a larger

C	τ	$n = 100$	$n = 250$	$n = 500$	$n = 1000$	$n = 5000$
Frank	0.10	0.070	0.065	0.077	0.114	0.376
	0.25	0.096	0.145	0.234	0.430	0.992
	0.50	0.175	0.323	0.603	0.883	1.000
	0.75	0.232	0.379	0.651	0.918	1.000
	0.90	0.411	0.348	0.457	0.710	1.000
Clayton	0.10	0.056	0.056	0.052	0.054	0.059
	0.25	0.059	0.044	0.057	0.049	0.065
	0.50	0.075	0.065	0.083	0.091	0.272
	0.75	0.142	0.177	0.272	0.504	1.000
	0.90	0.336	0.284	0.365	0.549	0.997
Gumbel	0.10	0.065	0.050	0.053	0.048	0.068
	0.25	0.044	0.047	0.060	0.050	0.076
	0.50	0.064	0.050	0.063	0.056	0.051
	0.75	0.099	0.057	0.073	0.068	0.082
	0.90	0.217	0.120	0.105	0.076	0.104

Table 3.5: Dimension d=3: Empirical power of the test for ellipticity with significance level $\alpha = 0.05$ based on the test statistic T_n: rate of rejecting H_0 as observed in 1000 random samples of size n from copula family C with Kendall's tau τ.

sample size is needed. Table 3.6 shows that the empirical power for dimension $d = 6$ is larger than the corresponding power for lower dimensions, such that here a sample size of $n = 1000$ is sufficient for small levels of dependence.

The bivariate Frank copula is the only Archimedean copula which is not only symmetric but also radially symmetric. Since radial symmetry is an important necessary condition for a copula to be elliptical, the fact that the test performs quite well for this family is a very promising feature. Note that elliptical copulas of dimension $d > 3$ can but do not have to be exchangeable.

In case of the Clayton family, quite similar observations can be made, though with slightly lower rejection rates. Still, we can say that the test seems to be good in detecting the lack of ellipticity if the level of dependence is not too close to independence.

In contrast to the previous results, the rejection rates for the Gumbel family appear to be very low. Since the test statistic T_n is based on the difference between Kendall's tau and Blomqvist's beta, we have to take a closer look at those two measures in order to find some explanation. Figure 3.6 illustrates Kendall's tau and Blomqvist's beta as functions of the copula family parameter θ for the bivariate Frank, Clayton and Gumbel copulas. Here, the reason for the low rejection rates becomes apparent: Kendall's tau and Blomqvist's beta are very close and almost not distinguishable for the Gumbel family. Nevertheless, even in this case, the test is able to provide some

C	τ	$n = 1000$	$n = 5000$
Frank	0.10	0.176	0.794
	0.25	0.678	1.000
	0.50	0.974	1.000
	0.75	0.959	1.000
	0.90	0.862	1.000
Clayton	0.10	0.059	0.077
	0.25	0.056	0.061
	0.50	0.122	0.260
	0.75	0.532	0.996
	0.90	0.745	1.000
Gumbel	0.10	0.062	0.080
	0.25	0.055	0.057
	0.50	0.074	0.062
	0.75	0.101	0.104
	0.90	0.238	0.139

Table 3.6: Dimension d=6: Empirical power of the test for ellipticity with significance level $\alpha = 0.05$ based on the test statistic T_n: rate of rejecting H_0 as observed in 1000 random samples of size n from copula family C with Kendall's tau τ.

C	τ_G	τ_t	$n = 100$	$n = 250$	$n = 500$	$n = 1000$	$n = 5000$
Mixture ($p = 0.5$)	0.25	0.75	0.077	0.080	0.134	0.249	0.776
	0.75	0.25	0.079	0.070	0.118	0.176	0.604
	0.50	0.25	0.041	0.043	0.061	0.046	0.054
	0.50	0.75	0.048	0.062	0.069	0.075	0.202

Table 3.7: Dimension d=2: Empirical power of the test for ellipticity with significance level $\alpha = 0.05$ based on the test statistic T_n: rate of rejecting H_0 as observed in 1000 random samples of size n from a mixture C of bivariate elliptical copulas with Kendall's tau combinations (τ_G, τ_t).

C	τ_G	τ_t	$n = 100$	$n = 250$	$n = 500$	$n = 1000$	$n = 5000$
	0.25	0.75	0.053	0.064	0.068	0.099	0.232
Mixture	0.75	0.25	0.092	0.128	0.236	0.366	0.948
$(p = 0.5)$	0.50	0.25	0.047	0.042	0.056	0.061	0.150
	0.50	0.75	0.041	0.061	0.039	0.051	0.065

Table 3.8: Dimension d=2: Empirical power of the test for ellipticity with significance level $\alpha = 0.05$ based on the test statistic T_n: rate of rejecting H_0 as observed in 1000 random samples of size n from the copula C of a mixture of bivariate elliptical distributions with Kendall's tau combinations (τ_G, τ_t).

indication against the null hypothesis for huge sample sizes if the level of dependence is high enough, meaning Kendall's tau being equal to 0.75 or higher. To confirm this presumption, we carried out the simulation study for the bivariate Gumbel copula with a Kendall's tau of 0.75 and chose a sample size of $n = 10^5$, which delivered a quite acceptable rejection rate of 0.648.

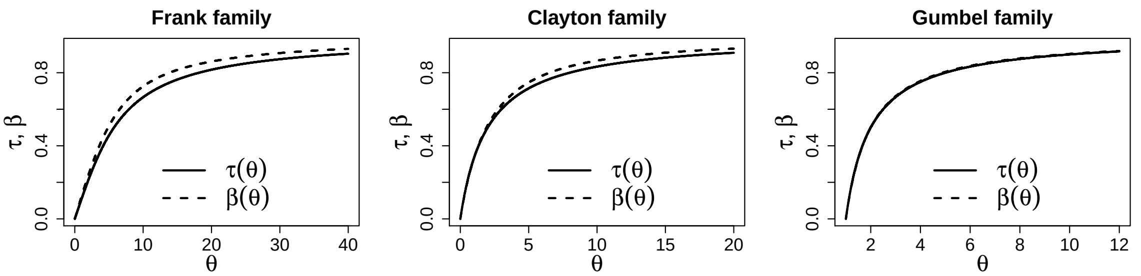

Figure 3.6: Comparison of Kendall's tau and Blomqvist's beta as functions of the copula family parameter θ for different copula families.

Power for the bivariate mixture copula constructions

For the mixture of bivariate elliptical copulas, the test generally achieves good power only for huge sample sizes ($n = 10^5$), which we do not consider in our simulation study. If the absolute difference of the values of Kendall's tau for the Gaussian and the t copula is large enough then an acceptable empirical power can be observed already for a sample size of $n = 5000$.

Similar observations on the empirical power can be made for the copula derived from the mixture of bivariate elliptical distributions. There is only one exception. It turns out that the empirical power depends not only on the absolute difference but also on the sign of the difference. Thus, the empirical power of 0.948 for the combination of

$\tau_G = 0.75$ and $\tau_t = 0.25$ is observed. Whereas, we get the empirical power of 0.232 if we switch the values of the dependence levels.

Since it is not easy to graphically detect the non-ellipticity for the samples of the mixture copulas used in the simulation study, our test is still useful.

3.5.4 Level and power for pseudo-observations

In this section, empirical level and power of the proposed test are investigated for the more realistic situation with unknown marginal distributions. For this, we simulated copula data from the considered copula families and transformed the uniform marginal distributions to exponential distributions with unit rate to get observations $\boldsymbol{X}_i \in \mathbb{R}_+^d$, $i = 1, \ldots, n$. The test was then applied to the corresponding pseudo-observations resulting from Equation (2.12) and denoted by $\widehat{\boldsymbol{U}}_i$, $i = 1, \ldots, n$. Thus, we do not make any assumptions on the marginal distributions, which is in accordance with practical applications. Below, we present our results for the bivariate case.

Table 3.9 shows the empirical level of our test for $d = 2$ and unknown margins. As one can observe, the test keeps its nominal level across all sample sizes and dependence levels for the considered copula families. Compared to Table 3.1, the empirical levels are similar for both situations: known and unknown margins. This supports our testing procedure for copula data in real applications.

C	τ	$n = 100$	$n = 250$	$n = 500$	$n = 1000$	$n = 5000$
Gauss	0.10	0.063	0.049	0.054	0.042	0.043
	0.25	0.053	0.048	0.063	0.048	0.049
	0.50	0.046	0.060	0.058	0.044	0.041
	0.75	0.069	0.055	0.036	0.053	0.045
	0.90	0.082	0.060	0.051	0.050	0.049
$t_{\nu=5}$	0.10	0.063	0.049	0.059	0.056	0.053
	0.25	0.055	0.047	0.050	0.043	0.046
	0.50	0.058	0.055	0.043	0.060	0.048
	0.75	0.058	0.053	0.044	0.058	0.057
	0.90	0.075	0.067	0.048	0.045	0.050

Table 3.9: Dimension d=2: Empirical level of the test for ellipticity with significance level $\alpha = 0.05$ based on the test statistic T_n: rate of rejecting H_0 as observed in 1000 random samples of pseudo-observations of size n from copula family C with Kendall's tau τ.

Table 3.10 now shows the empirical power of our test for $d = 2$ and unknown margins. We do not observe any significant differences in comparison to the empirical power results from Table 3.4. Thus, it seems that the test is equally powerful for known as well as unknown marginal distributions. Summarizing the empirical findings of this section, we can recommend our test also in the case of unknown marginal distributions, although the observations are now dependent and therefore the limit results do not hold as stated in Theorem 3.1.

C	τ	$n = 100$	$n = 250$	$n = 500$	$n = 1000$	$n = 5000$
	0.10	0.052	0.082	0.060	0.093	0.251
	0.25	0.092	0.101	0.176	0.310	0.910
Frank	0.50	0.125	0.258	0.475	0.753	1.000
	0.75	0.169	0.336	0.534	0.851	1.000
	0.90	0.166	0.224	0.372	0.609	1.000
	0.10	0.047	0.055	0.051	0.063	0.049
	0.25	0.041	0.050	0.038	0.072	0.065
Clayton	0.50	0.054	0.043	0.061	0.087	0.218
	0.75	0.098	0.138	0.272	0.445	0.982
	0.90	0.168	0.176	0.292	0.506	0.990
	0.10	0.050	0.055	0.057	0.052	0.056
	0.25	0.049	0.040	0.051	0.049	0.061
Gumbel	0.50	0.039	0.056	0.057	0.033	0.046
	0.75	0.046	0.047	0.047	0.057	0.077
	0.90	0.078	0.071	0.058	0.070	0.086

Table 3.10: Dimension d=2: Empirical power of the test for ellipticity with significance level $\alpha = 0.05$ based on the test statistic T_n: rate of rejecting H_0 as observed in 1000 random samples of pseudo-observations of size n from copula family C with Kendall's tau τ.

3.5.5 Power under the local alternatives

The simple functional form of the test statistic allows to investigate the power of the proposed goodness-of-fit test under local alternatives. Since the accuracy of the distributional approximation for our test statistic T_n is not satisfactory for small sample sizes and large dimensions, we restrict ourselves to dimension $d = 2$. For the null hypothesis $H_0 : \beta = \tau$, local alternatives of the form $\beta = \tau + \Delta/\sqrt{n}$ are considered for varying Δ. It follows in lines of the proof of Theorem 3.1 that the asymptotic distribution of the test statistic under the local alternatives is the non-central χ^2-distribution with one degree of freedom and non-centrality parameter Δ^2/v^2, where v^2 is the asymptotic variance of $\boldsymbol{D}_n$ for $d = 2$. In applications, the asymptotic variance v^2 should be consistently estimated and depends on the underlying data.

For varying Δ, Figure 3.7 shows the theoretical asymptotic power of our test under the sequence of local alternatives when the data comes from a Frank copula with $\tau = 0.75$ and $\beta = 0.804$. The asymptotic variance is estimated using a sample of size

10000. This estimate is then used instead of the unknown asymptotic variance v^2. Further, the five circles in Figure 3.7 indicate the empirical power of our test from Table 3.4 for the Frank copula and the five sample sizes $n = 100, 250, 500, 1000$ and 5000. For each sample size n, the position of the circles on the x-axis is computed by $\sqrt{n}\,(\beta - \tau)$. Thus, the circles are located further to the right with increasing sample size n. We see that the asymptotic local power is in good agreement with our empirical results. Moreover, the five triangles in Figure 3.7 similarly display the empirical power of our test applied to pseudo-observations from Section 3.5.4. For the considered simulation scenario, Figure 3.7 shows that the empirical power of our asymptotic test does not significantly fall in quality and agrees well with the asymptotic local power even if marginal distributions are unknown.

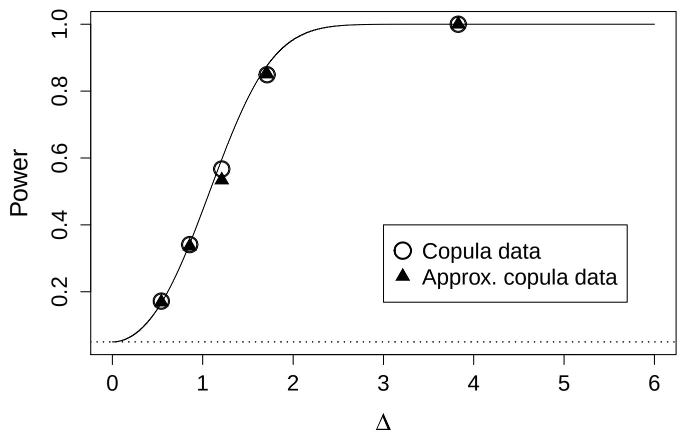

Figure 3.7: Asymptotic local power curve for the bivariate Frank copula with $\tau = 0.75$ and $\beta = 0.804$. Circles and triangles correspond to the empirical power for copula data and pseudo-observations, respectively, of sample sizes $n = 100, 250, 500, 1000, 5000$.

3.6 Empirical analysis

We consider the daily log-returns of the DAX, the Dow Jones Industrial Average and the Euro Stoxx 50 indices for 11 years starting from January 1, 2006 till December 31, 2016. For our test, we need i.i.d. data. Therefore, we fit a time series model to each series of log-returns and then use the standardized residuals of these models. More precisely, we choose ARMA$(1,1)$ - GARCH$(1,1)$ models with Student's t innovations to capture autocorrelation and volatility clustering in the daily log-returns. The model fits have been validated with QQ-plots of the standardized residuals.

To get the copula data, the standardized residuals have to be transformed to achieve approximate i.i.d. uniform margins. This can be done non-parametrically by using the empirical cumulative distribution functions. Apart from that, one can use a Student's t distribution to parametrically transform the residuals, which is due to the fact that the considered ARMA$(1,1)$ - GARCH$(1,1)$ models have Student's t innovations. Figure 3.8 displays the scatter plots of the standardized residuals after the non-parametric (above the diagonal) as well as the parametric transformation with the fitted t distribution (below the diagonal). Here, we can visually observe a high dependence between the margins as well as symmetry and radial symmetry of the underlying data. Therefore, an elliptical copula would be a natural choice to model the dependence structure of the standardized residuals of the three indices.

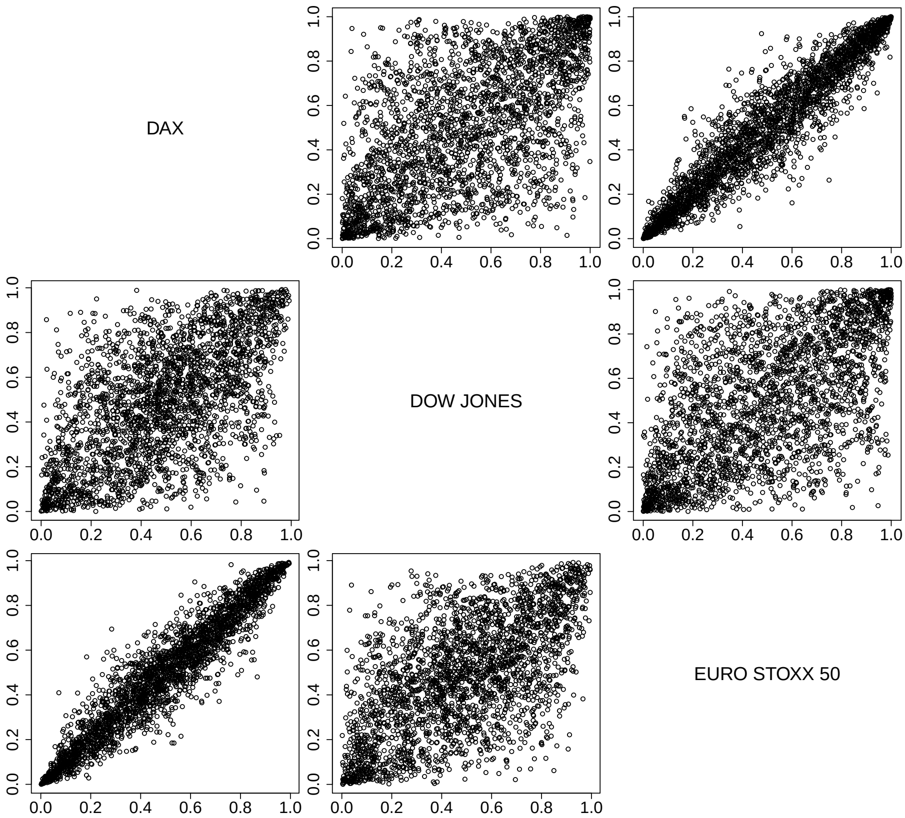

Figure 3.8: Pairwise scatter plots of the non-parametrically (above the diagonal) and parametrically (below the diagonal) transformed residuals of the ARMA-GARCH models for the DAX, Dow Jones and Euro Stoxx 50 indices.

Now, we apply our goodness-of-fit test to the underlying copula data. We get p-values of 0.037 and 0.045 for the non-parametrically and the parametrically transformed residuals, respectively. Hence, our test rejects the null hypothesis that the dependence structure of the considered data can be captured by a three-dimensional elliptical copula at the significance level of 5%. This is a surprising statistical result and indicates that one should be careful when choosing elliptical copulas in financial applications.

Further, we get p-values between 0.018 and 0.063, when we apply our test to the bivariate margins of the non-parametrically and the parametrically transformed residuals. Even if we cannot reject the null hypothesis of ellipticity for some bivariate margins, we would not favour elliptical copulas for modeling the two-dimensional dependence structures of the given data.

3.7 Conclusion

In this chapter, we derive a simple non-parametric goodness-of-fit test for elliptical copulas of any dimension. It is based on the equality of Kendall's tau and Blomqvist's beta for all bivariate margins. The distinguishing property of our test is its ability to differentiate between elliptical and non-elliptical copulas of any dimension even if the underlying copulas are radially symmetric. In the bivariate case, our test can even detect symmetric non-elliptical copulas. For bivariate copula data, we propose to use the test in combination with the tests for symmetry and radial symmetry presented in Chapter 4. This leads to a powerful non-parametric testing procedure, which is also developed in Chapter 4.

In an intensive Monte Carlo study, we investigate the nominal level and the power of the proposed test. Unfortunately, our test is not powerful enough to reject samples of moderate and large sizes from the Gumbel copula. For the considered Archimedean copulas, except the Gumbel family, our test has sufficient power for moderate dependent data starting from sample size 1000. When considering bivariate copulas derived from mixture constructions, the power depends on the values of the association parameters and the distance between them. In some cases, sufficient power can already be achieved using samples of size 5000.

Our test requires copula data, which is usually not available in empirical applications due to unknown marginal distributions. It seems that the performance of our asymptotic test is not significantly influenced by non-parametric estimation of unknown marginal distributions. In Chapter 5, we extend our test to the case of unknown margins. Furthermore, the problem of the test to hold its nominal level in higher dimensions gets addressed. In case of given bivariate data with unknown marginal distributions, we propose to use the test in combination with the advanced tests for symmetry and radial symmetry from Genest et al. (2012) and Genest and Nešlehová (2014), respectively.

3.8 Proof

Recall that the copula C is defined as the distribution function of (U_1, U_2) given according to Equation (3.3) as

$$(U_1, U_2) = \begin{cases} (U, V/4)\,, & \text{if } U \leq 1/4 \\ (U, V/4 + 3/4), & \text{if } 1/4 < U \leq 1/2 \\ (U, V/4 + 1/2), & \text{if } 1/2 < U \leq 3/4 \\ (U, V/4 + 1/4), & \text{if } 3/4 < U \leq 1\,. \end{cases}$$

In the following, we prove that Equation (3.4) holds, that is

$$\tau = \beta = 0\,.$$

Furthermore, note that

$$\mathbb{P}(U \leq 1/4) = \mathbb{P}(1/4 < U \leq 1/2) = \mathbb{P}(1/2 < U \leq 3/4) = \mathbb{P}(3/4 < U \leq 1) = 1/4\,.$$

For the computation of Blomqvist's beta β, let $\tilde{u}_1$ and $\tilde{u}_2$ denote the population medians of U_1 and U_2, respectively. Using Definition 2.4 together with the law of total probability, it follows

$$\begin{aligned}
\beta =& \mathbb{P}((U_1 - \tilde{u}_1)(U_2 - \tilde{u}_2) > 0) - \mathbb{P}((U_1 - \tilde{u}_1)(U_2 - \tilde{u}_2) < 0) \\
=& 2\mathbb{P}((U_1 - \tilde{u}_1)(U_2 - \tilde{u}_2) > 0) - 1 \\
=& 2\Big[\mathbb{P}((U_1 - \tilde{u}_1)(U_2 - \tilde{u}_2) > 0 \mid U \leq 1/4) \cdot \mathbb{P}(U \leq 1/4) \\
& + \mathbb{P}((U_1 - \tilde{u}_1)(U_2 - \tilde{u}_2) > 0 \mid 1/4 < U \leq 1/2) \cdot \mathbb{P}(1/4 < U \leq 1/2) \\
& + \mathbb{P}((U_1 - \tilde{u}_1)(U_2 - \tilde{u}_2) > 0 \mid 1/2 < U \leq 3/4) \cdot \mathbb{P}(1/2 < U \leq 3/4) \\
& + \mathbb{P}((U_1 - \tilde{u}_1)(U_2 - \tilde{u}_2) > 0 \mid 3/4 < U \leq 1) \cdot \mathbb{P}(3/4 < U \leq 1)\Big] - 1 \\
=& 2\left[1 \cdot \frac{1}{4} + 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{4} + 0 \cdot \frac{1}{4}\right] - 1 \\
=& 0\,.
\end{aligned}$$

For the computation of Kendall's tau, note that the random vector (U_1, U_2) given that $U \leq 1/4$ is uniformly distributed on the square $[0, 1/4]^2$ and its conditional density is given by

$$c_{U_1,U_2|U\leq 1/4}(u_1, u_2) = 16 \cdot I\{0 \leq u_1 \leq 1/4, 0 \leq u_2 \leq 1/4\}\,.$$

The conditional density of (U_1, U_2) given that $1/4 < U \leq 1/2$, $1/2 < U \leq 3/4$, or $3/4 < U \leq 1$ is given similarly. Now, let (U_1', U_2') be an independent copy of the random vector (U_1, U_2). According to Definition 2.3, Kendall's tau can be computed as

$$\begin{aligned}
\tau &= \mathbb{E}[\operatorname{sgn}(U_1 - U_1')\operatorname{sgn}(U_2 - U_2')] \\
&= \int_{[0,1]^2}\int_{[0,1]^2} \operatorname{sgn}(u_1 - u_1')\operatorname{sgn}(u_2 - u_2') dC(u_1,u_2) dC(u_1',u_2') \\
&= \int_0^{1/4}\int_0^{1/4}\int_0^{1/4}\int_0^{1/4} \operatorname{sgn}(u_1 - u_1')\operatorname{sgn}(u_2 - u_2') \cdot 16 \cdot \frac{1}{4} \cdot 16 \cdot \frac{1}{4}\, du_1 du_2 du_1' du_2' \\
&\quad + 16 \cdot \int_{3/4}^{1}\int_{1/4}^{1/2}\int_0^{1/4}\int_0^{1/4} \operatorname{sgn}(u_1 - u_1')\operatorname{sgn}(u_2 - u_2')\, du_1 du_2 du_1' du_2' \\
&\quad + 16 \cdot \int_{1/2}^{3/4}\int_{1/2}^{3/4}\int_0^{1/4}\int_0^{1/4} \operatorname{sgn}(u_1 - u_1')\operatorname{sgn}(u_2 - u_2')\, du_1 du_2 du_1' du_2' \\
&\quad + 16 \cdot \int_{1/4}^{1/2}\int_{3/4}^{1}\int_0^{1/4}\int_0^{1/4} \operatorname{sgn}(u_1 - u_1')\operatorname{sgn}(u_2 - u_2')\, du_1 du_2 du_1' du_2' \\
&\quad + 16 \cdot \int_0^{1/4}\int_0^{1/4}\int_{3/4}^{1}\int_{1/4}^{1/2} \operatorname{sgn}(u_1 - u_1')\operatorname{sgn}(u_2 - u_2')\, du_1 du_2 du_1' du_2' \\
&\quad + 16 \cdot \int_{3/4}^{1}\int_{1/4}^{1/2}\int_{3/4}^{1}\int_{1/4}^{1/2} \operatorname{sgn}(u_1 - u_1')\operatorname{sgn}(u_2 - u_2')\, du_1 du_2 du_1' du_2' \\
&\quad + 16 \cdot \int_{1/2}^{3/4}\int_{1/2}^{3/4}\int_{3/4}^{1}\int_{1/4}^{1/2} \operatorname{sgn}(u_1 - u_1')\operatorname{sgn}(u_2 - u_2')\, du_1 du_2 du_1' du_2' \\
&\quad + 16 \cdot \int_{1/4}^{1/2}\int_{3/4}^{1}\int_{3/4}^{1}\int_{1/4}^{1/2} \operatorname{sgn}(u_1 - u_1')\operatorname{sgn}(u_2 - u_2')\, du_1 du_2 du_1' du_2' \\
&\quad + 16 \cdot \int_0^{1/4}\int_0^{1/4}\int_{1/2}^{3/4}\int_{1/2}^{3/4} \operatorname{sgn}(u_1 - u_1')\operatorname{sgn}(u_2 - u_2')\, du_1 du_2 du_1' du_2' \\
&\quad + 16 \cdot \int_{3/4}^{1}\int_{1/4}^{1/2}\int_{1/2}^{3/4}\int_{1/2}^{3/4} \operatorname{sgn}(u_1 - u_1')\operatorname{sgn}(u_2 - u_2')\, du_1 du_2 du_1' du_2' \\
&\quad + 16 \cdot \int_{1/2}^{3/4}\int_{1/2}^{3/4}\int_{1/2}^{3/4}\int_{1/2}^{3/4} \operatorname{sgn}(u_1 - u_1')\operatorname{sgn}(u_2 - u_2')\, du_1 du_2 du_1' du_2' \\
&\quad + 16 \cdot \int_{1/4}^{1/2}\int_{3/4}^{1}\int_{1/2}^{3/4}\int_{1/2}^{3/4} \operatorname{sgn}(u_1 - u_1')\operatorname{sgn}(u_2 - u_2')\, du_1 du_2 du_1' du_2' \\
&\quad + 16 \cdot \int_0^{1/4}\int_0^{1/4}\int_{1/4}^{1/2}\int_{3/4}^{1} \operatorname{sgn}(u_1 - u_1')\operatorname{sgn}(u_2 - u_2')\, du_1 du_2 du_1' du_2' \\
&\quad + 16 \cdot \int_{3/4}^{1}\int_{1/4}^{1/2}\int_{1/4}^{1/2}\int_{3/4}^{1} \operatorname{sgn}(u_1 - u_1')\operatorname{sgn}(u_2 - u_2')\, du_1 du_2 du_1' du_2' \\
&\quad + 16 \cdot \int_{1/2}^{3/4}\int_{1/2}^{3/4}\int_{1/4}^{1/2}\int_{3/4}^{1} \operatorname{sgn}(u_1 - u_1')\operatorname{sgn}(u_2 - u_2')\, du_1 du_2 du_1' du_2' \\
&\quad + 16 \cdot \int_{1/4}^{1/2}\int_{3/4}^{1}\int_{1/4}^{1/2}\int_{3/4}^{1} \operatorname{sgn}(u_1 - u_1')\operatorname{sgn}(u_2 - u_2')\, du_1 du_2 du_1' du_2' .
\end{aligned}$$

In the following, we compute the first and the second summand. For the first summand, we get

$$
\begin{aligned}
&16 \cdot \int_0^{1/4}\int_0^{1/4}\int_0^{1/4}\int_0^{1/4} \operatorname{sgn}(u_1-u_1')\operatorname{sgn}(u_2-u_2')\, du_1du_2du_1'du_2' \\
&= 16 \cdot \int_0^{1/4}\int_0^{1/4}\int_0^{1/4} \operatorname{sgn}(u_2-u_2') \int_0^{1/4} \operatorname{sgn}(u_1-u_1')\, du_1du_2du_1'du_2' \\
&= 16 \cdot \int_0^{1/4}\int_0^{1/4}\int_0^{1/4} \operatorname{sgn}(u_2-u_2') \left[\int_{u_1'}^{1/4} 1du_1 + \int_0^{u_1'} -1du_1\right] du_2du_1'du_2' \\
&= 16 \cdot \int_0^{1/4}\int_0^{1/4} \left(\frac{1}{4}-2u_1'\right) \left[\int_{u_2'}^{1/4} 1du_2 + \int_0^{u_2'} -1du_2\right] du_1'du_2' \\
&= 16 \cdot \int_0^{1/4}\int_0^{1/4} \left(\frac{1}{4}-2u_1'\right)\left(\frac{1}{4}-2u_2'\right) du_1'du_2' \\
&= \int_0^{1/4}\int_0^{1/4} 1-8u_2'-8u_1'+64u_1'u_2'\, du_1'du_2' \\
&= \int_0^{1/4} \left[u_1'-8u_2'u_1'-4u_1'^2+32u_2'u_1'^2\right]_0^{1/4} du_2' \\
&= \int_0^{1/4} \frac{1}{4}-2u_2'-\frac{1}{4}+2u_2'\, du_2' \\
&= 0\,.
\end{aligned}
$$

For the second summand, we get

$$
\begin{aligned}
&16 \cdot \int_{3/4}^{1}\int_{1/4}^{1/2}\int_0^{1/4}\int_0^{1/4} \operatorname{sgn}(u_1-u_1')\operatorname{sgn}(u_2-u_2')\, du_1du_2du_1'du_2' \\
&= 16 \cdot \int_{3/4}^{1}\int_{1/4}^{1/2}\int_0^{1/4} \operatorname{sgn}(u_2-u_2') \int_0^{1/4} \operatorname{sgn}(u_1-u_1')\, du_1du_2du_1'du_2' \\
&= 16 \cdot \int_{3/4}^{1}\int_{1/4}^{1/2}\int_0^{1/4} \operatorname{sgn}(u_2-u_2') \int_0^{1/4} -1du_1\, du_2du_1'du_2' \\
&= (-4)\int_{3/4}^{1}\int_{1/4}^{1/2}\int_0^{1/4} -1du_2\, du_1'du_2' \\
&= \int_{3/4}^{1}\int_{1/4}^{1/2} 1\, du_1'du_2' \\
&= \frac{1}{16}\,.
\end{aligned}
$$

Inserting these results together with the results for the remaining summands yields

$$\tau = 0+\frac{1}{16}+\frac{1}{16}+\frac{1}{16}+\frac{1}{16}+0-\frac{1}{16}-\frac{1}{16}+\frac{1}{16}-\frac{1}{16}+0-\frac{1}{16}+\frac{1}{16}-\frac{1}{16}-\frac{1}{16}+0=0\,.$$

Altogether, this proves Equation (3.4), that is $\tau=\beta=0$.

4

On tests for symmetry and radial symmetry of bivariate copulas towards testing for ellipticity

This chapter is similar to Jaser and Min (2020).

4.1 Introduction

Since Embrechts et al. (2003), Frees and Valdez (1998), and Li (2000), copulas were widely used in economics, finance, and risk management to capture the dependence of multivariate data. Bivariate parametric copulas are usually the basis of many multivariate copula constructions (see, e.g., Aas et al. (2009) or Fischer et al. (2009)). Therefore, the choice of a parametric bivariate copula family is very crucial to accurately capture the multivariate dependence. For large and huge sample sizes, carrying out known goodness-of-fit tests is very time consuming. Graphical tools like scatter plots can significantly reduce the amount of copulas to be considered but may lead to erroneous decisions. In this chapter, we fill this existing gap and propose simple statistical tests to detect symmetry or radial symmetry of the underlying bivariate copula data. More precisely, a test for the hypothesis that the unknown copula C is symmetric, that is

$$H_0^s : C(u, v) = C(v, u)\,, \text{ for all } (u, v) \in [0, 1]^2\,,$$

against the alternative

$$H_1^s : \exists (u, v) \in [0, 1]^2\,, \text{ such that } C(u, v) \neq C(v, u)\,,$$

is proposed. The null hypothesis and the alternative to test whether the unknown copula C is radially symmetric are given by

$$H_0^r : C = \hat{C} \text{ versus } H_1^r : C \neq \hat{C}\,,$$

where $\hat{C}$ denotes the survival copula of C.

The existing tests for symmetry and radial symmetry of bivariate copulas by Genest et al. (2012), Genest and Nešlehová (2014), Li and Genton (2013), and Quessy (2016) assume unknown marginal distributions and take into account their non-parametric estimation. Therefore, the asymptotic distribution of their test statistics is of complex nature and derived using the weak convergence of empirical copula processes. In applications, bootstrap techniques are needed for the computation of p-values, and this is computationally expensive for huge sample sizes.

Assuming given copula data, we propose simpler non-parametric tests for symmetry and radial symmetry of bivariate copulas. We manipulate the underlying copula data without changing its dependence structure to create two bivariate samples. Our test statistics are then based on the difference between the empirical Kendall's tau of both samples. The limiting distributions of the test statistics can be derived using the classical theory of U-statistics. Therefore, our non-parametric tests are related to asymptotic normal distributions and are very simple at work. Our tests are based only on a sample characteristic of the bivariate copula data. Therefore, they are easy to implement and computationally very fast. In times of Big Data, this nice feature of our tests is very useful in the analysis of data sets with huge sample sizes.

In Chapter 3, we proposed a goodness-of-fit test for elliptical copulas under the assumption of given copula data. It utilizes the known equality of Kendall's tau and Blomqvist's beta for elliptical copulas, that is the null hypothesis

$$H_0^e : \tau = \beta \text{ is tested against the alternative } H_1^e : \tau \neq \beta\,.$$

This test may illustrate poor performance in finite samples if Kendall's tau and Blomqvist's beta are very close for a particular copula family. In this chapter, we propose a multiple testing procedure for ellipticity of copula data, which combines our simple non-parametric tests for symmetry, radial symmetry, and the equality of Kendall's tau and Blomqvist's beta. Thus, the proposed multiple testing procedure utilizes the most common properties of elliptical copulas, which should make it powerful to detect a non-elliptical dependence structure in bivariate copula data. Let C be the unknown bivariate copula and $\mathcal{C}^{ellipt}$ the class of elliptical copulas, then the null hypothesis and the alternative of the testing procedure are given by

$$H_0 : C \in \mathcal{C}^{ellipt} \text{ versus } H_1 : C \notin \mathcal{C}^{ellipt}\,.$$

This chapter is organized as follows. In Section 4.2, we propose simple non-parametric tests for symmetry and radial symmetry. Section 4.3 presents a Monte Carlo simulation study to evaluate the finite-sample performance and compare it to the one of existing competitors. In Section 4.4, a simple and powerful non-parametric testing procedure is proposed to decide whether the dependence structure of underlying bivariate copula data may be captured by an elliptical copula. Applications to financial and insurance data are reported in Section 4.5 to illustrate the testing procedure at work. Finally, Section 4.6 concludes. Some supplementary material is deferred to Section 4.7.

4.2 Simple non-parametric tests for symmetry and radial symmetry

In this section, we derive our two statistical tests for symmetry and radial symmetry for bivariate copulas. We assume that we are given a copula sample and neglect unknown marginal distributions and their estimation. In practical applications, one usually estimates marginal distribution functions non-parametrically to avoid misspecification. For the following subsections, let $(U_1, V_1), \ldots, (U_n, V_n) \in [0,1]^2$ be a sample from the statistical model $(([0,1]^2)^n,\ \mathcal{B}([0,1]^2)^{\otimes n},\ \ P^{\otimes n})$, where P is a distribution with copula C and uniform margins.

4.2.1 Test for symmetry

Let (U,V) be distributed according to the symmetric copula C, that is $(U,V) \overset{d}{=} (V,U)$. Further, we assume that $\mathbb{P}(U{=}V) = 0$. For a given sample realization from C, the scatter plot displays symmetry with respect to the main diagonal. By interchanging the coordinates, any two observations, one below and one above the diagonal, can be mirrored to the opposite side of the diagonal. The modified data set can still be considered as a realization from the given copula C. Therefore, a sample realization from the copula C can be generated just using all observations either above or below the diagonal.

The complementary events that (U,V) is below or above the diagonal, that is

$$B^s := \{\omega : U - V > 0\} \quad \text{and} \quad \overline{B^s} := \{\omega : U - V < 0\}\,, \tag{4.1}$$

have equal probabilities of 0.5. Using (2.5), it follows for the events in (4.1) that

$$\mathbb{P}(B^s) = \mathbb{P}(U - V > 0) = \mathbb{P}(U - V < 0) = \mathbb{P}(\overline{B^s}) = 0.5\,. \tag{4.2}$$

Using the law of total probability as well as (2.5) and (4.2), it follows that

$$\begin{aligned} C(u,v) &= \mathbb{P}(U \le u, V \le v) \\ &= \mathbb{P}(U \le u, V \le v \mid B^s) \cdot \mathbb{P}(B^s) + \mathbb{P}(U \le u, V \le v \mid \overline{B^s}) \cdot \mathbb{P}(\overline{B^s}) \\ &= 0.5 \cdot \mathbb{P}(U \le u, V \le v \mid U - V > 0) + 0.5 \cdot \mathbb{P}(V \le u, U \le v \mid V - U < 0) \\ &= 0.5 \cdot F_{U,V|B^s}(u,v) + 0.5 \cdot F_{V,U|B^s}(u,v)\,. \end{aligned} \tag{4.3}$$

Here, $F_{X,Y|A}$ denotes the conditional distribution function of (X,Y) given $(X,Y) \in A$. Thus, the symmetric copula C can be represented as a mixture of two conditional distribution functions. Similarly, it holds that

$$\begin{aligned} C(u,v) &= 0.5 \cdot \mathbb{P}(V \le u, U \le v \mid V - U > 0) + 0.5 \cdot \mathbb{P}(U \le u, V \le v \mid U - V < 0) \\ &= 0.5 \cdot F_{V,U|\overline{B^s}}(u,v) + 0.5 \cdot F_{U,V|\overline{B^s}}(u,v)\,. \end{aligned} \tag{4.4}$$

According to Equation (4.3) and (4.4), the symmetric copula C can be represented either as a mixture of two conditional distribution functions given the event that (U, V) is below the diagonal or as a mixture of two conditional distribution functions given the event that (U, V) is above the diagonal. This constitutes the key idea of our testing procedure for symmetric copulas pursued to produce two i.i.d. random samples out of a given i.i.d. random sample from C.

Let $(U_1, V_1), \ldots, (U_n, V_n)$ be an i.i.d. random sample from the symmetric copula C. First, we consider the sub-sample $(U_1^{B^s}, V_1^{B^s}), \ldots, (U_{N_{B^s}}^{B^s}, V_{N_{B^s}}^{B^s})$ for which $U_\cdot^{B^s} - V_\cdot^{B^s} > 0$ holds, that is, whose realizations are below the diagonal. By virtue of Equation (4.3), a new sample from C can be obtained by choosing either $(U_i^{B^s}, V_i^{B^s})$ with probability 0.5 or $(V_i^{B^s}, U_i^{B^s})$ also with probability 0.5, for $i \in \{1, \ldots, N_{B^s}\}$. The resulting random sample is denoted by

$$(\tilde{U}_1^{B^s}, \tilde{V}_1^{B^s}), \ldots, (\tilde{U}_{N_{B^s}}^{B^s}, \tilde{V}_{N_{B^s}}^{B^s}) \,. \tag{4.5}$$

Similarly, we proceed with the sub-sample $(U_1^{\overline{B^s}}, V_1^{\overline{B^s}}), \ldots, (U_{N_{\overline{B^s}}}^{\overline{B^s}}, V_{N_{\overline{B^s}}}^{\overline{B^s}})$ for which $U_\cdot^{\overline{B^s}} - V_\cdot^{\overline{B^s}} < 0$ holds, that is, whose realizations are above the diagonal, and create a second random sample

$$(\tilde{U}_1^{\overline{B^s}}, \tilde{V}_1^{\overline{B^s}}), \ldots, (\tilde{U}_{N_{\overline{B^s}}}^{\overline{B^s}}, \tilde{V}_{N_{\overline{B^s}}}^{\overline{B^s}}) \,. \tag{4.6}$$

It should be mentioned that the sampling algorithm can be generalized for $0 < \mathbb{P}(U{=}V) < 1$ by discarding observations with $U_i = V_i$.

Note that the sample size N_{B^s} is a binomially distributed random variable with size n and success probability 0.5. From the law of large numbers, it follows that N_{B^s}/n converges to 0.5 in probability as n tends to infinity. The same conclusions can be drawn for the sample size $N_{\overline{B^s}}$ since the relation $N_{\overline{B^s}} = n - N_{B^s}$ holds. Defining the sequence of random variables $N_n^s := \min\left(N_{B^s}, N_{\overline{B^s}}\right)$, it follows that N_n^s/n similarly converges to 0.5 in probability as n tends to infinity. Choosing the first N_n^s realizations from (4.5) and (4.6) yields random samples of equal sample size N_n^s given by

$$(\tilde{U}_1^{B^s}, \tilde{V}_1^{B^s}), \ldots, (\tilde{U}_{N_n^s}^{B^s}, \tilde{V}_{N_n^s}^{B^s}) \quad \text{and} \quad (\tilde{U}_1^{\overline{B^s}}, \tilde{V}_1^{\overline{B^s}}), \ldots, (\tilde{U}_{N_n^s}^{\overline{B^s}}, \tilde{V}_{N_{N_n^s}}^{\overline{B^s}}) \,. \tag{4.7}$$

Under the null hypothesis H_0^s of C being symmetric, the two newly generated random samples have the same underlying copula C and, hence, Kendall's tau. Therefore, the empirically estimated Kendall's tau for both random samples should be of the same magnitude. Now, we base our test on the difference

$$S_{N_n^s} := \tau_{N_n^s}^{B^s} - \tau_{N_n^s}^{\overline{B^s}} \,,$$

where $\tau_{N_n^s}^{B^s}$ and $\tau_{N_n^s}^{\overline{B^s}}$ denote the empirically estimated Kendall's taus based on the two samples from (4.7). Note that the indices for the identification of the margins can be skipped since we are dealing with bivariate data in this chapter.

It is clear that

$$\frac{N_{B^s}}{n} \xrightarrow{\mathbb{P}} 0.5 \quad \text{and} \quad \frac{N_{\overline{B^s}}}{n} \xrightarrow{\mathbb{P}} 0.5 \,.$$

For $n \geq 2$, the above sampling algorithm can be slightly modified to ensure that N_{B^s} and $N_{\overline{B^s}}$ are positive random variables. Therefore, N_n^s is a sequence of positive integer-valued random variables with

$$\frac{N_n^s}{n} \xrightarrow{\mathbb{P}} 0.5\,. \tag{4.8}$$

To state the asymptotic distribution of the test statistic $S_{N_n^s}$ in Theorem 4.1, we define $\tilde{h}_1\big((U_1, V_1)\big) := \mathbb{E}\left[\operatorname{sgn}(U_1 - U_2)\operatorname{sgn}(V_1 - V_2) \mid U_1, V_1\right]$.

Theorem 4.1. *Let $(U_1, V_1), \ldots, (U_n, V_n)$ be an i.i.d. random sample from a bivariate random vector (U, V) with $\mathbb{P}(U{=}V) = 0$, whose distribution function is a symmetric copula C. Further, let (4.8) hold. Then,*

$$\sqrt{n^\star} \cdot S_{N_n^s} \xrightarrow{d} N\left(0, 2\sigma^2\right) ,$$

where $n^\star = n/2$ and $\sigma^2 = \mathbb{V}\mathrm{ar}\left(2\tilde{h}_1\big((U_1, V_1)\big)\right)$.

Proof. Let $(U_1, V_1), \ldots, (U_n, V_n) \in [0, 1]^2$ be a sample from the statistical model

$$\Big(([0, 1]^2)^n, \mathcal{B}([0, 1]^2)^{\otimes n}, P^{\otimes n}\Big) ,$$

where P is a distribution with symmetric copula C and uniform margins. The samples given in (4.7) can then be derived and the test statistic $S_{N_n^s}$ is given by the difference of the corresponding empirical Kendall's tau estimators $\tau^{B^s}_{N_n^s}$ and $\tau^{\overline{B^s}}_{N_n^s}$.

For the random sample size N_n^s, it holds that N_n^s/n converges to 0.5 in probability as n tends to infinity. It follows for $n \to \infty$ that

$$\frac{N_n^s}{\lfloor n/2 \rfloor} \xrightarrow{\mathbb{P}} 1\,,$$

where $\lfloor x \rfloor$, $x \in \mathbb{R}$, denotes the integer part of x. Thus, the assumption of Theorem 1 from Anscombe (1952) is satisfied, and it is sufficient to show that the difference $\tau_n^{B^s} - \tau_n^{\overline{B^s}}$ satisfies the conditions (C1) and (C2) of Anscombe (1952).

To make it easier to follow this and remaining similar proofs, we state conditions (C1) and (C2) here. A sequence of random variables $\{Y_n\}$ satisfies condition (C1) of Anscombe (1952) if there exist a real number θ, a sequence of positive numbers $\{w_n\}$, and a distribution function F, such that for any x with $F(x)$ continuous it holds that

$$\mathbb{P}\left(Y_n - \theta \leq x w_n\right) \to F(x)\,, \qquad \text{as } n \to \infty\,. \tag{4.9}$$

Moreover, a sequence of random variables $\{Y_n\}$ satisfies condition (C2) of Anscombe (1952) if, given $\epsilon > 0$ and $\eta > 0$, there exists a large $\nu_{\epsilon,\eta}$ and a small $c > 0$ such that for all $n > \nu_{\epsilon,\eta}$ it holds that

$$\mathbb{P}\left(\sup_{n\prime : |n\prime - n| < cn} |Y_{n\prime} - Y_n| \geq \epsilon w_n\right) < \eta\,. \tag{4.10}$$

Finally, note that given two sequences of random variables that satisfy (C2), it can be shown that the difference also satisfies (C2).

From the theory of U-statistics (see Hoeffding (1947)), it holds that $\sqrt{n}\,(\tau_n - \tau)$ converges in distribution to a centered normal distribution with variance $\sigma^2 = \mathbb{V}\text{ar}\left(2\tilde{h}_1\left((U_1, V_1)\right)\right)$. The independence of $\tau_n^{B^s}$ and $\tau_n^{\overline{B^s}}$, and the Delta method imply

$$\sqrt{n}\left(\tau_n^{B^s} - \tau_n^{\overline{B^s}}\right) \xrightarrow{d} N\left(0, 2\sigma^2\right).$$

Thus, the difference $\tau_n^{B^s} - \tau_n^{\overline{B^s}}$ satisfies condition (C1) with $w_n = 1/\sqrt{n}$.

Further, the proof of Theorem 6 in Sproule (1974) yields that $\tau_n^{B^s}$ and $\tau_n^{\overline{B^s}}$ satisfy condition (C2). Therefore, the difference $\tau_n^{B^s} - \tau_n^{\overline{B^s}}$ also satisfies condition (C2). Finally, Theorem 1 of Anscombe (1952) implies the desired asymptotic convergence

$$\sqrt{n^\star}\left(\tau_{N_n^s}^{B^s} - \tau_{N_n^s}^{\overline{B^s}}\right) \xrightarrow{d} N\left(0, 2\sigma^2\right).$$

□

In practical applications, the unknown variance σ^2 in Theorem 4.1 should be consistently estimated. The following remark describes a possible consistent estimation procedure for σ^2.

Remark 4.2. *The function $\tilde{h}_1$ has the representation (see, e.g., Theorem 4.3 in Dengler (2010))*

$$\tilde{h}_1\left((U, V)\right) = 1 - 2U - 2V + 4C(U, V).$$

Subsequently, the asymptotic variance of $S_{N_n^s}$ can be consistently estimated in the framework of Chapter 3. Using the whole random sample (U_1, V_1), …, (U_n, V_n), $\tilde{h}_1\left((U_i, V_i)\right)$ is estimated non-parametrically by

$$\widehat{h}_1\left((U_i, V_i)\right) = 1 - 2U_i - 2V_i + 4C_n(U_i, V_i), i \in \{1, \ldots, n\},$$

where $\widetilde{C}_n$ denotes the empirical copula defined in Equation (2.9) Now, σ^2 is consistently estimated by the sample variance $\widehat{\sigma}_n^2$ of

$$2\widehat{h}_1\left((U_1, V_1)\right), \ldots, 2\widehat{h}_1\left((U_n, V_n)\right).$$

For details see Chapter 3.

Based on Theorem 4.1, we propose the test function

$$\delta^s(U_1, \ldots, U_n) = I\left\{\left|\sqrt{n^\star} \cdot S_{N_n^s} \,/\, \widehat{\sigma}_n\right| > z_{1-\alpha/2}\right\}$$

to test H_0^s against H_1^s at the significance level α, where z_α denotes the α-quantile of the standard normal distribution.

4.2.2 Test for radial symmetry

Let (U, V) be distributed according to the radially symmetric copula C. Hence, C coincides with its survival copula $\hat{C}$, and it holds that $(U, V) \stackrel{d}{=} (1-U, 1-V)$. Further, we assume that $\mathbb{P}(U{+}V{=}1) = 0$. For sample realizations from C, scatter plots show symmetry with respect to the the point $(0.5, 0.5)$. Now, we split a given data set with respect to the counter-diagonal into two sub-sets: one below and the other above the counter-diagonal. By reflecting any two observations from different sub-sets with respect to the point $(0.5, 0.5)$, the copula of the resulting sample is not changed. Therefore, a sample from the copula C can be generated just using all observations either below or above the counter-diagonal.

More precisely, note that the complementary events

$$B^r := \{\omega : U + V < 1\} \quad \text{and} \quad \overline{B^r} := \{\omega : U + V > 1\}$$

have equal probabilities of 0.5. We follow the idea of our test for symmetry and use two mixture representations conditioned on the events that (U, V) is below and above the counter-diagonal, respectively, in order to generate two i.i.d. random samples of size N_{B^r} and $N_{\overline{B^r}}$ out of one given i.i.d. random sample from C.

Similarly to Section 4.2.1, the corresponding test statistic is given by

$$R_{N_n^r} := \tau_{N_n^r}^{B^r} - \tau_{N_n^r}^{\overline{B^r}},$$

where $\tau_{N_n^r}^{B^r}$ and $\tau_{N_n^r}^{\overline{B^r}}$ denote the empirically estimated Kendall's taus based on the two samples, and $N_n^r := \min\left(N_{B^r}, N_{\overline{B^r}}\right)$. As before, N_n^r can be assumed to be a sequence of positive integer-valued random variables with

$$\frac{N_n^r}{n} \xrightarrow{\mathbb{P}} 0.5\,. \tag{4.11}$$

The asymptotic distribution of the test statistic $R_{N_n^r}$ is given in the following theorem.

Theorem 4.3. *Let $(U_1, V_1), \ldots, (U_n, V_n)$ be an i.i.d. random sample from a bivariate random vector (U, V) with $\mathbb{P}(U{+}V{=}1) = 0$, whose distribution function is a radially symmetric copula C. Further, let (4.11) hold. Then,*

$$\sqrt{n^\star} \cdot R_{N_n^r} \xrightarrow{d} N\left(0, 2\sigma^2\right),$$

where $n^\star = n/2$ and $\sigma^2 = \mathbb{V}\mathrm{ar}\left(2\tilde{h}_1\Big((U_1, V_1)\Big)\right)$.

The proof of Theorem 4.3 is similar to the proof of Theorem 4.1 and, therefore, omitted. Note that the asymptotic variance σ^2 is the same as in Theorem 4.1. Hence, Remark 4.2 yields a consistent estimation procedure for the asymptotic variance of $R_{N_n^r}$ and the test function δ^r is constructed similarly to δ^s.

4.3 Simulation study

In order to assess the finite-sample performance of our proposed tests for symmetry and radial symmetry, a Monte Carlo study was conducted for the test problems H_0^s and H_0^r. First, we would like to point out that the tests are based on a random sampling algorithm. Therefore, the value of the test statistic inherits some variability. The upcoming simulation study shows that the randomness of the test statistic does not affect the empirical level of the tests and the tests still provide good empirical power.

As a benchmark, we use the more advanced tests from Genest et al. (2012) and Genest and Nešlehová (2014), respectively, which are available in the R-package `copula` (see `exchTest` and `radSymTest` in Hofert et al. (2018)). Note that our proposed tests rely on the assumption of known marginal distributions, while the tests from Genest et al. (2012) and Genest and Nešlehová (2014) take into account their non-parametric estimation. Further, their tests compare the whole copulas while our proposed tests are based on two sample characteristics of the bivariate copula. We assume that this fact is mainly responsible for the differences between our and their numerical results.

The mixture representations for symmetric or radial symmetric copulas may not hold if marginal distributions are estimated. Therefore, it is not straightforward for us to extend the proposed tests for unknown margins. Further, if marginal distributions are estimated non-parametrically, the two newly generated samples may contain ties. Our Monte Carlo study empirically assesses the influence of non-parametrically estimated marginal distributions on the level and power of our proposed tests. For this, each copula sample $(U_1, V_1), \ldots, (U_n, V_n)$ is replaced by the corresponding bivariate pseudo-observations $(\widehat{U}_1, \widehat{V}_1)$, ..., $(\widehat{U}_n, \widehat{V}_n)$ defined via Equation (2.12).

4.3.1 Setup

First of all, the number of Monte Carlo replications was set to $N = 1000$, and all tests were performed at a significance level of $\alpha = 0.05$. To determine the empirical level and power of the tests, the simulation study was carried out for different sample sizes, levels of dependence measured in terms of Kendall's tau and types of dependence expressed in terms of copula families.

More precisely, random samples of size $n \in \{100, 250, 500, 1000\}$ were considered for all tests throughout the study. In addition, the influence of the strength of dependence was investigated by choosing five different levels of dependence in terms of Kendall's tau given by $\tau \in \{0.1, 0.25, 0.5, 0.75, 0.9\}$. Finally, the type of dependence is determined through the choice of a specific copula family. For this, some of the most popular copula families and some derived special cases were considered in the simulation study. The performance of all tests was studied for samples from the Gaussian, t (with 5 degrees of freedom, $t_{\nu=5}$), Frank, Clayton, and Gumbel copula families. The Gaussian and the t copula are elliptical copulas and, thus, also symmetric and radially symmetric. Further, the Frank, Clayton, and Gumbel copula are symmetric Archimedean copulas.

In addition, the Frank copula is also radially symmetric.

Since all listed copulas are symmetric, asymmetrized versions of the Gaussian, Clayton, and Gumbel copula families were additionally used to assess the power of the test for symmetry. Regarding the asymmetrization, we followed the procedure in Genest et al. (2012) and used Khoudraji's device (see Khoudraji (1995)). The asymmetric copulas are given in terms of an asymmetrization parameter $\delta \in (0, 1)$. According to Genest et al. (2012), maximum asymmetry is observed for $\delta = 0.5$ and, hence, we also chose $\delta \in \{0.25, 0.5, 0.75\}$. Since there is only little asymmetry for small values of τ, we analyzed the performance of the test for symmetry for $\tau \in \{0.5, 0.75, 0.9\}$ in this context. To illustrate the asymmetrized versions of the considered copula families used in the simulation study, Figure 4.1 displays scatter plots of random samples of size $n = 1000$ for the asymmetrization parameter $\delta \in \{0.25, 0.5, 0.75\}$. Following Genest and Nešlehová (2014), a Skewed-t copula with 4 degrees of freedom and skewness parameter $\gamma = (1, 1)$ (Skewed-$t_{\nu=4}$) was chosen to study the power of the test for radial symmetry. Figure 4.2 illustrates random samples of size $n = 1000$ from this Skewed-$t_{\nu=4}$ copula for the levels of dependence corresponding to $\tau \in \{0.25, 0.5, 0.75\}$. Note that the skewness can be observed to decrease with increasing level of dependence.

4.3.2 Test for symmetry

In this section, the finite-sample performance of the test of H_0^s for symmetry based on the test statistics $S_{N_n^s}$ is analyzed. To study the level of the test, random samples from the Gaussian, t, Frank, Clayton, and Gumbel copula were considered. Table 4.1 reports the empirical level of our test (in Column JMS), of our test for pseudo-observations (in Column JMSP), and of the test from Genest et al. (2012) (in Column GNQ).

First, note that our test holds its nominal level across all copula models, sample sizes, and values of Kendall's tau. Compared to the more advanced test from Genest et al. (2012), our test seems to hold its nominal level a little better. For pseudo-observations, our test is generally rather conservative and its empirical level is decreasing with increasing sample size. Surprisingly, this does not influence the empirical power negatively.

Random samples from the asymmetrized versions of the Gaussian, Clayton, and Gumbel copula families were used to investigate the power of the test for symmetry. Table 4.2 displays the empirical power of our test (in Column JMS), of our test for pseudo-observations (in Column JMSP), and of the test from Genest et al. (2012) (in Column GNQ). Even if the results vary noticeably across the different combinations of factors, our test generally achieves sufficient power. As expected, the rejection rates increase with the sample size as well as with the strength of dependence. In terms of the asymmetrization parameter δ, the largest power is mostly observed for $\delta = 0.5$. Since maximum asymmetry occurs near $\delta = 0.5$, this is also expected.

Compared to the test from Genest et al. (2012), our test has slightly lower power and needs higher sample sizes to achieve similar power. The empirical power of our test

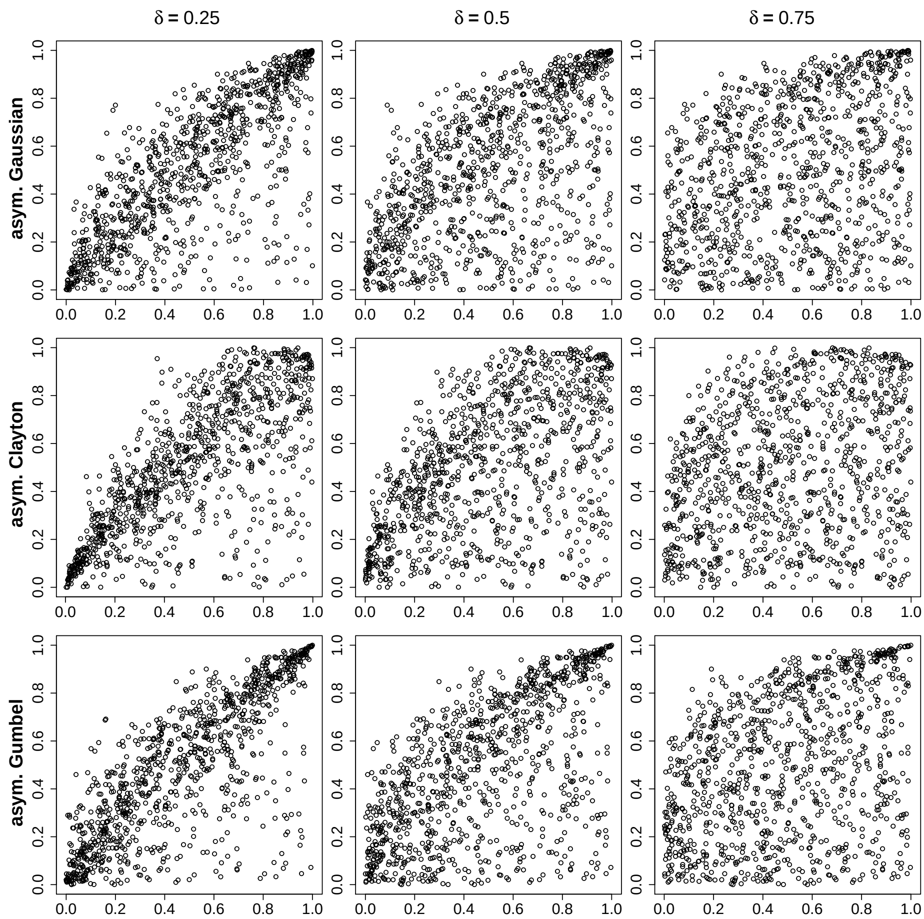

Figure 4.1: Scatter plots of random samples of size 1000 from the asymmetrized versions of the Gaussian, Clayton, and Gumbel copula (from top to bottom) with asymmetrization parameter $\delta = 0.25$ (left), 0.5 (middle), 0.75 (right).

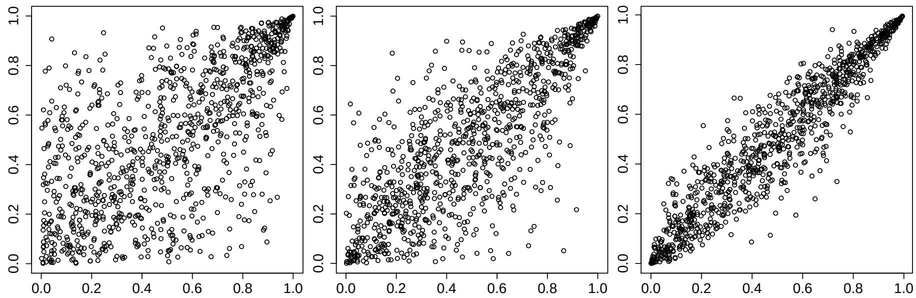

Figure 4.2: Scatter plots of random samples of size 1000 from the Skewed-$t_{\nu=4}$ copula with $\tau = 0.25$ (left), 0.5 (middle), 0.75 (right).

C	$n = 100$			$n = 250$			$n = 500$			$n = 1000$		
τ	JMS	JMSP	GNQ	JMS	JMSP	GNQ	JMS	JMSP	GNQ	JMS	JMSP	GNQ
Gauss												
0.25	0.066	0.012	0.022	0.055	0.016	0.039	0.055	0.012	0.048	0.054	0.007	0.044
0.50	0.060	0.012	0.015	0.061	0.006	0.027	0.045	0.010	0.020	0.051	0.006	0.032
0.75	0.041	0.024	0.011	0.052	0.021	0.011	0.046	0.016	0.004	0.050	0.006	0.013
$t_{\nu=5}$												
0.25	0.059	0.014	0.033	0.043	0.016	0.046	0.052	0.023	0.035	0.059	0.013	0.035
0.50	0.047	0.018	0.014	0.058	0.016	0.035	0.049	0.009	0.031	0.062	0.006	0.046
0.75	0.034	0.027	0.022	0.055	0.019	0.014	0.057	0.010	0.013	0.051	0.008	0.019
Frank												
0.25	0.054	0.014	0.031	0.052	0.014	0.038	0.051	0.010	0.043	0.045	0.009	0.032
0.50	0.057	0.024	0.015	0.060	0.013	0.025	0.052	0.005	0.038	0.061	0.007	0.035
0.75	0.036	0.017	0.016	0.032	0.009	0.011	0.042	0.010	0.006	0.048	0.009	0.016
Clayton												
0.25	0.060	0.024	0.033	0.062	0.018	0.040	0.052	0.010	0.032	0.051	0.009	0.043
0.50	0.063	0.029	0.031	0.050	0.013	0.029	0.059	0.003	0.029	0.045	0.004	0.035
0.75	0.059	0.051	0.021	0.059	0.028	0.015	0.056	0.015	0.021	0.049	0.009	0.027
Gumbel												
0.25	0.063	0.022	0.036	0.049	0.020	0.038	0.061	0.011	0.035	0.049	0.013	0.042
0.50	0.057	0.015	0.027	0.050	0.010	0.026	0.053	0.004	0.024	0.049	0.006	0.039
0.75	0.053	0.034	0.017	0.051	0.021	0.013	0.050	0.019	0.008	0.049	0.003	0.028

Table 4.1: Empirical level of our test for symmetry (JMS), our test for pseudo-observations (JMSP), and the test from Genest et al. (2012) (GNQ) with significance level $\alpha = 0.05$: rate of rejecting H_0 as observed in 1000 random samples of size n from copula family C with Kendall's tau τ.

C	$n = 100$			$n = 250$			$n = 500$			$n = 1000$		
τ	JMS	JMSP	GNQ	JMS	JMSP	GNQ	JMS	JMSP	GNQ	JMS	JMSP	GNQ
Gauss												
$\delta = 0.25$												
0.50	0.095	0.067	0.082	0.157	0.140	0.233	0.285	0.242	0.466	0.480	0.530	0.803
0.75	0.434	0.553	0.618	0.849	0.954	0.995	0.991	1.000	1.000	1.000	1.000	1.000
0.90	0.963	0.983	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\delta = 0.5$												
0.50	0.155	0.145	0.199	0.296	0.312	0.499	0.564	0.645	0.851	0.846	0.936	0.989
0.75	0.672	0.798	0.907	0.976	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.90	0.973	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\delta = 0.75$												
0.50	0.158	0.134	0.168	0.298	0.289	0.393	0.542	0.604	0.764	0.846	0.889	0.968
0.75	0.521	0.576	0.626	0.896	0.961	0.988	0.996	1.000	1.000	1.000	1.000	1.000
0.90	0.657	0.744	0.844	0.979	0.996	0.999	1.000	1.000	1.000	1.000	1.000	1.000
Clayton												
$\delta = 0.25$												
0.50	0.177	0.149	0.093	0.361	0.343	0.260	0.586	0.633	0.548	0.885	0.954	0.909
0.75	0.678	0.763	0.779	0.958	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.90	0.892	0.981	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\delta = 0.5$												
0.50	0.148	0.115	0.111	0.307	0.264	0.339	0.509	0.578	0.715	0.813	0.901	0.965
0.75	0.463	0.550	0.834	0.871	0.953	1.000	0.991	1.000	1.000	1.000	1.000	1.000
0.90	0.817	0.920	0.999	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\delta = 0.75$												
0.50	0.092	0.058	0.072	0.132	0.110	0.169	0.197	0.173	0.295	0.370	0.365	0.586
0.75	0.190	0.189	0.366	0.440	0.469	0.814	0.751	0.832	0.988	0.961	0.990	1.000
0.90	0.460	0.515	0.764	0.827	0.919	0.997	0.989	0.998	1.000	1.000	1.000	1.000
Gumbel												
$\delta = 0.25$												
0.50	0.142	0.149	0.110	0.263	0.268	0.275	0.515	0.573	0.637	0.744	0.855	0.916
0.75	0.592	0.743	0.679	0.944	0.982	0.997	0.998	1.000	1.000	1.000	1.000	1.000
0.90	0.990	0.991	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\delta = 0.5$												
0.50	0.285	0.305	0.272	0.599	0.669	0.704	0.895	0.963	0.974	0.992	0.999	1.000
0.75	0.862	0.950	0.970	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.90	0.987	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\delta = 0.75$												
0.50	0.273	0.284	0.284	0.638	0.690	0.690	0.888	0.966	0.963	0.990	1.000	1.000
0.75	0.619	0.693	0.752	0.951	0.985	0.993	0.999	1.000	1.000	1.000	1.000	1.000
0.90	0.722	0.799	0.893	0.987	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 4.2: Empirical power of our test for symmetry (JMS), our test for pseudo-observations (JMSP), and the test from Genest et al. (2012) (GNQ) with significance level $\alpha = 0.05$: rate of rejecting H_0 as observed in 1000 random samples of size n from copula family C asymmetrized with parameter δ and with Kendall's tau τ.

for pseudo-observations is in most cases comparable to the one for the copula samples. Moreover, across all different combinations of factors, there are several scenarios with higher empirical power for the pseudo-observations even though the empirical level for them is lower than for copula data.

Our test for symmetry is computationally less intensive than the more advanced test from Genest et al. (2012), where bootstrap methods are applied. Table 4.3 illustrates the running times of the tests (in Row JMS and GNQ, respectively) for samples of size $n = 10^3, 10^4$, and 10^5. For one sample of size $n = 10^4$, the running time of our test is about 2 seconds in comparison to more than 2 minutes for the corresponding test from Genest et al. (2012). For $n = 10^5$, it was not possible to conduct the test for symmetry of Genest et al. (2012) using the R-package `copula`, while our test runs in a bit more than 3 minutes. Thus, our test for symmetry is up to 75 times faster and can especially be recommended for huge samples.

	$n = 10^3$	$n = 10^4$	$n = 10^5$
JMS	0.02	1.77	197.70
GNQ	1.34	134.03	–
JMR	0.04	3.55	399.91
GN	7.60	655.05	63982.67 (17.77 h)

Table 4.3: Running times in seconds for our tests (JMS/JMR) and the tests from Genest et al. (2012)/Genest and Nešlehová (2014) (GNQ/GN) for samples of size $n = 10^3, 10^4$, and 10^5.

4.3.3 Test for radial symmetry

In this section, the finite-sample performance of the test of H_0^r for radial symmetry based on the test statistic $R_{N_n^r}$ is analyzed. Random samples from the Gaussian, t, and Frank copula were considered in order to examine the empirical level. Table 4.4 presents the empirical level of our test (in Column JMR), of our test for pseudo-observations (in Column JMRP), and of the test from Genest and Nešlehová (2014) (in Column GN). In general, our test and the test from Genest and Nešlehová (2014) hold their nominal level. For pseudo-observations, our test also holds its nominal level in most cases. One exception is the Frank copula for $\tau = 0.75$. Further analysis showed that increasing the sample size does not reduce the problem of inflated rejection rates as the empirical levels oscillate around 0.119. Hence, our test for radial symmetry is systematically too liberal in this setting.

To assess the empirical power, random samples from the Clayton, Gumbel, and Skewed-$t_{\nu=4}$ copula were used. Table 4.5 reports the empirical power of our test (in

C	$n = 100$			$n = 250$			$n = 500$			$n = 1000$		
τ	JMR	JMRP	GN	JMR	JMRP	GN	JMR	JMRP	GN	JMR	JMRP	GN
Gauss												
0.25	0.049	0.049	0.041	0.058	0.029	0.047	0.054	0.030	0.042	0.042	0.030	0.049
0.50	0.055	0.059	0.037	0.053	0.053	0.044	0.060	0.055	0.059	0.047	0.045	0.044
0.75	0.052	0.073	0.042	0.046	0.077	0.051	0.039	0.060	0.052	0.056	0.079	0.051
$t_{\nu=5}$												
0.25	0.063	0.038	0.050	0.065	0.040	0.052	0.054	0.034	0.039	0.056	0.036	0.049
0.50	0.041	0.040	0.031	0.053	0.042	0.043	0.056	0.056	0.057	0.049	0.039	0.040
0.75	0.052	0.054	0.029	0.053	0.064	0.051	0.047	0.049	0.036	0.051	0.052	0.052
Frank												
0.25	0.061	0.033	0.039	0.053	0.034	0.045	0.052	0.032	0.044	0.042	0.036	0.043
0.50	0.044	0.067	0.052	0.046	0.067	0.049	0.045	0.066	0.052	0.061	0.068	0.048
0.75	0.032	0.111	0.037	0.044	0.116	0.040	0.050	0.119	0.036	0.037	0.125	0.052

Table 4.4: Empirical level of our test for radial symmetry (JMR), our test for pseudo-observations (JMRP), and the test from Genest and Nešlehová (2014) (GN) with significance level $\alpha = 0.05$: rate of rejecting H_0 as observed in 1000 random samples of size n from copula family C with Kendall's tau τ.

Column JMR), of our test for pseudo-observations (in Column JMRP), and of the test from Genest and Nešlehová (2014) (in Column GN). First, note that the results differ considerably for the various combinations of factors. For all copulas, the power increases with the sample size, which is expected. Further, for the Clayton and the Gumbel copula, the power also increases with the degree of dependence, whereas for the Skewed-$t_{\nu=4}$ copula, the power decreases with increasing τ. This is in accordance with our observations from Figure 4.2 in which the skewness can be observed to decrease with increasing level of dependence in terms of Kendall's tau. Lastly, note that the rejection rates are slightly lower for the Gumbel copula.

Our test overall achieves satisfactory empirical power against the various alternatives. Compared to the test from Genest and Nešlehová (2014), it is in many cases somewhat less powerful. However, it achieves equal or even slightly higher power especially in scenarios where the more advanced test has difficulties to detect the radial asymmetry. Examples are given by the Gumbel copula and the Skewed-$t_{\nu=4}$ copula for $n = 100$ and $n = 250$ in combination with $\tau = 0.75$. The empirical power of our test for pseudo-observations is overall slightly higher than the one for copula samples, which might be caused by possible high empirical levels.

Table 4.3 illustrates the running times for our test (in Row JMR) and the test from Genest and Nešlehová (2014) (in Row GN) for samples of size $n = 10^3, 10^4$, and 10^5. For one sample of size $n = 10^4$, the running time of our test is less than 4 seconds in comparison to almost 11 minutes for the corresponding test from Genest and Nešlehová (2014). For one sample of size $n = 10^5$, it runs in less than 7 minutes, while the test

C	$n = 100$			$n = 250$			$n = 500$			$n = 1000$		
τ	JMR	JMRP	GN	JMR	JMRP	GN	JMR	JMRP	GN	JMR	JMRP	GN
Clayton												
0.25	0.256	0.229	0.377	0.517	0.506	0.730	0.851	0.858	0.959	0.983	0.993	1.000
0.50	0.625	0.640	0.811	0.955	0.965	0.997	1.000	1.000	1.000	1.000	1.000	1.000
0.75	0.775	0.884	0.921	0.997	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Gumbel												
0.25	0.119	0.123	0.092	0.207	0.215	0.246	0.343	0.342	0.491	0.615	0.638	0.800
0.50	0.166	0.193	0.161	0.413	0.447	0.458	0.703	0.722	0.814	0.934	0.948	0.987
0.75	0.166	0.234	0.132	0.516	0.575	0.495	0.814	0.823	0.828	0.981	0.985	0.992
S-$t_{\nu=4}$												
0.25	0.470	0.493	0.514	0.885	0.905	0.951	0.996	0.998	1.000	1.000	1.000	1.000
0.50	0.331	0.395	0.336	0.713	0.734	0.770	0.965	0.963	0.991	1.000	1.000	1.000
0.75	0.152	0.230	0.113	0.497	0.575	0.436	0.834	0.878	0.843	0.991	0.993	0.997

Table 4.5: Empirical power of our test for radial symmetry (JMR), our test for pseudo-observations (JMRP), and the test from Genest and Nešlehová (2014) (GN) with significance level $\alpha = 0.05$: rate of rejecting H_0 as observed in 1000 random samples of size n from copula family C with Kendall's tau τ.

from Genest and Nešlehová (2014) requires almost 18 hours. Thus, it is up to 190 times faster and, similarly to our test for symmetry, it can especially be recommended for huge samples.

4.4 Testing procedure for ellipticity

This section presents a powerful and simple non-parametric statistical procedure to test whether the dependence structure of a bivariate random vector with uniform margins is captured by an elliptical copula.

4.4.1 The testing procedure

The testing procedure consists of the following three steps. First, the hypothesis that the unknown copula C is symmetric, that is H_0^s is tested against the alternative H_1^s. If the hypothesis H_0^s cannot be rejected, we test the hypothesis that the unknown copula C is radially symmetric, that is H_0^r against the alternative H_1^r. In the third step of our testing procedure, the equality of Kendall's tau and Blomqvist's beta is tested, that is H_0^e is tested against the alternative H_1^e. If any of the three hypotheses is rejected, we also reject our original null hypothesis H_0 that C belongs to the class of elliptical copulas. If none of the three hypotheses can be rejected, we cannot reject the null hypothesis H_0 of C being elliptical. To assess the effect of non-parametrically estimated marginal distribution functions on the proposed testing procedure, the following simulation study is also conducted for pseudo-observations.

4.4.2 Simulation study

In this section, the finite-sample performance of the proposed testing procedure is analyzed. The corresponding Monte Carlo study was set up similarly to Section 4.3. Note that our testing procedure for ellipticity consists of a multiple test problem with three sub-hypotheses. In order to maintain the global level $\alpha = 0.05$, we made use of the standard Bonferroni procedure (see, e.g., Miller (1981)). For this, the three null hypotheses H_0^s, H_0^r, and H_0^e were tested sequentially and separately at the significance level $\alpha/3$. Finally, the null hypothesis $H_0 : C \in \mathcal{C}^{ellipt}$ was rejected if any of the considered sub-hypotheses was rejected.

Table 4.6 reports the empirical level of the testing procedure (in Column JMT) and of the testing procedure for pseudo-observations (in Column JMTP) based on random samples from the Gaussian and the t copula. The testing procedure appears to hold its nominal level for copula data as well as for pseudo-observations across all combinations of factors.

C	$n = 100$		$n = 250$		$n = 500$		$n = 1000$	
τ	JMT	JMTP	JMT	JMTP	JMT	JMTP	JMT	JMTP
Gauss								
0.25	0.039	0.034	0.050	0.035	0.053	0.026	0.047	0.023
0.50	0.061	0.041	0.071	0.038	0.053	0.037	0.052	0.032
0.75	0.049	0.044	0.057	0.046	0.054	0.038	0.057	0.048
$t_{\nu=5}$								
0.25	0.062	0.036	0.054	0.032	0.051	0.035	0.047	0.031
0.50	0.046	0.037	0.052	0.037	0.051	0.034	0.052	0.029
0.75	0.039	0.038	0.056	0.046	0.062	0.043	0.055	0.040

Table 4.6: Empirical level of our testing procedure for ellipticity (JMT) and of our testing procedure for pseudo-observations (JMTP) with significance level $\alpha = 0.05$: rate of rejecting H_0 as observed in 1000 random samples of size n from copula family C with Kendall's tau τ.

To study the power of the testing procedure, random samples of the Frank, Clayton, and Gumbel copula were considered. Table 4.7 shows the empirical power of the testing procedure (in Column JMT) and of the testing procedure for pseudo-observations (in Column JMTP). As already observed for all individual tests, the rejection rates vary clearly across copula families, levels of dependence, and sample sizes. As expected, the power increases with the sample size and with the level of dependence. The lowest rejection rates are observed for the Frank copula. However, it is still sufficiently good in detecting the lack of ellipticity if the sample size is large enough and the level of dependence is not too close to independence. For the Clayton copula, the testing procedure performs best in detecting the non-ellipticity, even in very small samples of size $n = 100$. The results for the Gumbel copula are only slightly worse than for the Clayton copula and, hence, the testing procedure is still powerful. Similar observations can be made for the empirical power of the testing procedure for pseudo-observations. As for the individual tests, there are scenarios with higher empirical power for the pseudo-observations than for copula data.

C	$n = 100$		$n = 250$		$n = 500$		$n = 1000$	
τ	JMT	JMTP	JMT	JMTP	JMT	JMTP	JMT	JMTP
Frank								
0.25	0.074	0.045	0.071	0.055	0.112	0.099	0.190	0.169
0.50	0.085	0.087	0.175	0.165	0.331	0.316	0.619	0.597
0.75	0.098	0.167	0.226	0.238	0.459	0.444	0.747	0.749
Clayton								
0.25	0.177	0.138	0.375	0.338	0.729	0.729	0.959	0.977
0.50	0.455	0.462	0.896	0.915	0.998	1.000	1.000	1.000
0.75	0.591	0.779	0.993	0.999	1.000	1.000	1.000	1.000
Gumbel								
0.25	0.096	0.091	0.140	0.126	0.219	0.218	0.467	0.441
0.50	0.107	0.095	0.270	0.275	0.557	0.551	0.863	0.874
0.75	0.112	0.156	0.340	0.415	0.680	0.714	0.949	0.947

Table 4.7: Empirical power of our testing procedure for ellipticity (JMT) and of our testing procedure for pseudo-observations (JMTP) with significance level $\alpha = 0.05$: rate of rejecting H_0 as observed in 1000 random samples of size n from copula family C with Kendall's tau τ.

Compared to the results of our test based on the equality of Kendall's tau and Blomqvist's beta in Chapter 3, the testing procedure performs much better for the Clayton and the Gumbel copula. It is now possible to distinguish non-elliptical copulas with very close Kendall's tau and Blomqvist's beta if they are not symmetric or not radially symmetric. Furthermore, the testing procedure is still able to detect the non-ellipticity of the Frank copula, which is symmetric and radially symmetric. All in all, the idea of using the most known properties of elliptical copulas in the testing procedure shows clear advantages.

4.5 Empirical analysis

The main aim of this section is to illustrate our testing procedure for ellipticity in practice using financial and insurance data. For this, the results of the three building tests are reported. Since our tests for symmetry and radial symmetry are based on a random sampling algorithm (see Section 4.2.1 and 4.2.2), we performed the tests within the testing procedure for ellipticity 1000 times and consider the averages of the resulting p-values.

In the sequel, the testing procedure is applied to six different data sets in total. For the majority of these data sets, the decision resulting from the testing procedure is the same for all 1000 replications. For two datasets, we get a different decision from 2 and 1 out of 1000 replications, respectively, than we get from the testing procedure using the average of the p-values. Hence, the number of cases with a different decision seems to be negligible. We recommend to perform the testing procedure more than 2 times, if the decision to accept or reject the null hypothesis is very close.

4.5.1 Financial data

As a first illustration, our testing procedure is applied to financial data from the US stock market. Two major US stock price indices are selected: the Standard & Poor's 500 (S&P 500), as one of the most popular indices of large-cap US equities, and the Russell 2000, as one of the most popular small-cap US indices. In order to get data sets of large sample sizes, daily returns of the two indices over different periods of three years are considered. It is well known that the dependence structure of financial data for crisis and non-crisis periods differs. Therefore, the following analysis is based on the daily log-returns of the S&P 500 and the Russell 2000 indices for the crisis periods from 1999 to 2001 and 2007 to 2009, as well as for the non-crisis periods from 2003 to 2005 and from 2011 to 2013. Furthermore, we are also interested in the dependence structure between monthly returns of the two indices, which is of more interest from a macroeconomic point of view. For this, monthly returns are considered for the period of the last 30 years from 1988 to 2017.

To remove temporal dependencies, ARMA-GARCH time series models are fitted to each series of log-returns. The choice of the final model is done using the BIC (see, e.g., Schwarz (1978)). The resulting standardized residuals are transformed non-parametrically by using the empirical cumulative distribution functions to achieve approximate i.i.d. uniform margins. Figures 4.3 and 4.4 display the scatter plots of the underlying copula data for the different data sets comprised of daily and monthly returns of the S&P 500 and the Russell 2000 for the selected time periods, respectively. In Figure 4.3, an elliptical shape is visually observable for the non-crisis periods from 2003 to 2005 and from 2011 to 2013, whereas the shape of the data for the crisis periods from 1999 to 2001 and from 2007 to 2009 might be non-elliptical from the visual impression. The shape of the copula data in Figure 4.4 is clearly non-elliptical.

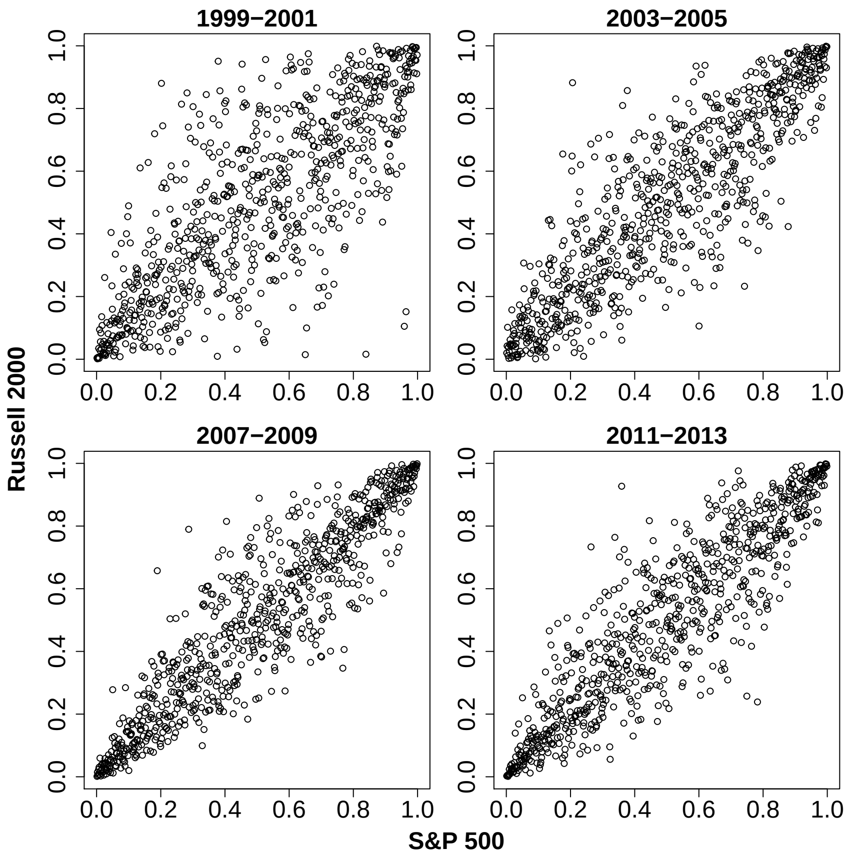

Figure 4.3: Daily data: Scatter plots of the non-parametrically transformed standardized residuals of the ARMA-GARCH models for the log-returns of the S&P 500 and the Russell 2000 indices for different time periods.

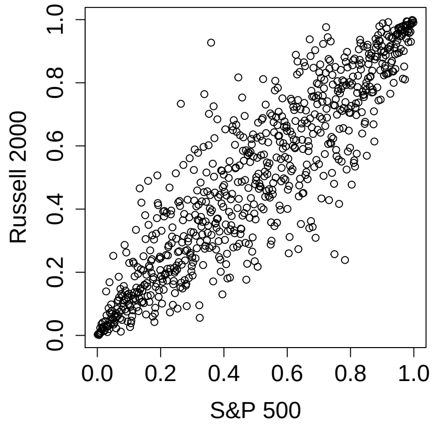

Figure 4.4: Monthly data: Scatter plot of the non-parametrically transformed standardized residuals of the ARMA-GARCH model for the log-returns of the S&P 500 and the Russell 2000 indices for the time period from 1988 to 2017.

Table 4.8 presents p-values for the above discussed data sets. For the two non-crisis periods, the null hypothesis H_0 of the elliptical dependence structure cannot be rejected at the considered significance level of 5%. In contrast, the testing procedure rejects H_0 for the crisis period from 1999 to 2001 due to a very low p-value of the test for equality. For the crisis period from 2007 to 2009, H_0 cannot be rejected at the considered significance level of 5%. However, the p-values of 0.042 and 0.030 for the test for radial symmetry and the test for equality, respectively, are quite low and provide some indication against H_0. Note that the test for radial symmetry leads to a rejection of H_0 for 2 out of the 1000 replications. Hence, also for the crisis period from 2007 to 2009, elliptical copulas cannot be recommended to model the dependence structure of the underlying data. The same applies for the data set comprised of the monthly log-returns. Due to the very low p-value of the test for radial symmetry, the null hypothesis of ellipticity is rejected at the considered significance level of 5%. All in all, the results are in accordance with our expectations and the visual observations from Figures 4.3 and 4.4.

4.5.2 Insurance data

One famous example for a bivariate data set from the insurance sector is given by losses and corresponding allocated loss adjustment expenses (short ALAE) of insurance claims. The US Insurance Services Office has collected data on 1500 general liability claims randomly chosen from late settlement lags. Each claim contains an indemnity payment (loss) and an allocated loss adjustment expense (ALAE). A detailed description of

data	time period	symmetry	radial symmetry	equality
Daily data (crisis)	1999-2001	0.154	0.124	0.004
Daily data (non-crisis)	2003-2005	0.943	0.352	0.705
Daily data (crisis)	2007-2009	0.857	0.043	0.030
Daily data (non-crisis)	2011-2013	0.346	0.345	0.965
Monthly data	1988-2017	0.417	0.002	0.380

Table 4.8: p-values of our tests for symmetry, radial symmetry, and equality of Kendall's tau and Blomqvist's beta for the dependence structure of the financial data (S&P 500 and Russell 2000) for different time periods.

the data set can be found in Frees and Valdez (1998). The modeling of the joint distribution of losses and ALAEs has also been analyzed in Genest et al. (1998), Klugman and Parsa (1999), Denuit et al. (2006), Chen and Fan (2005), and Zhang et al. (2016), among others. In Figure 4.5, scatter plots of the observations (left) and of the logarithm of the observations (middle) are displayed.

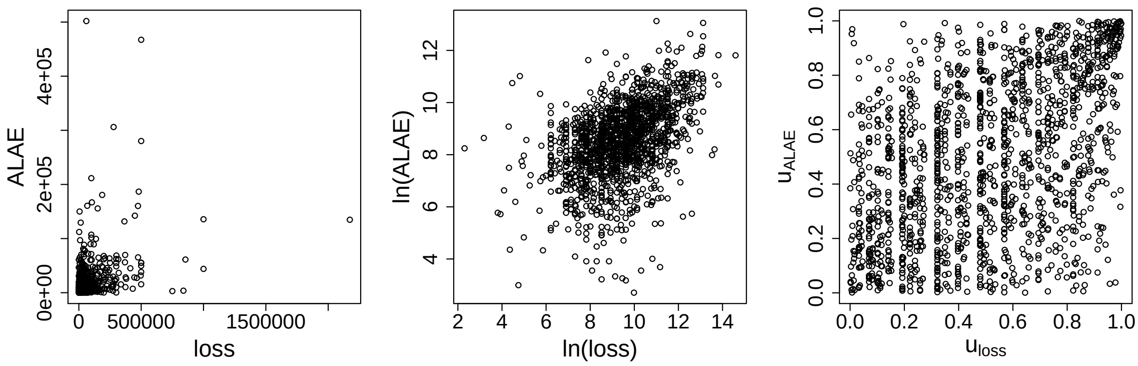

Figure 4.5: Scatter plots for data of loss and ALAE (left), logarithms of loss and ALAE (middle), non-parametrically transformed data of loss and ALAE (right). Sample size is $n = 1500$.

To achieve the approximate i.i.d. uniform margins, data is transformed non-parametrically by using the marginal empirical cumulative distribution functions. A scatter plot of the resulting transformed loss and ALAE is presented in Figure 4.5 (right). Applying our testing procedure for ellipticity then leads to p-values of 0.391 for symmetry, 0.034 for radial symmetry, and 0.118 for the equality of Blomqvist's beta and Kendall's tau. At the considered significance level of 5%, our testing procedure cannot reject the null hypothesis H_0. However, the low p-value of 0.034 for the radial symmetry provides some indication against H_0. Note that the test for radial symmetry leads to a rejection of H_0 for 1 out of the 1000 replications. Hence, we would not recommend elliptical copulas to model the dependence structure of the underlying loss ALAE data.

The different parametric and semi-parametric model selection procedures in Frees and Valdez (1998), Genest et al. (1998), Denuit et al. (2006), Chen and Fan (2005), and Zhang et al. (2016) all resulted in the Gumbel copula as the preferred model for the given loss ALAE data set. In the scatter plot of the copula data (Figure 4.5, right), positive upper-tail dependence but no lower-tail dependence can be observed between the two variables. This is expected by actuaries, since large losses are often accompanied by large ALAEs, and in line with the tail dependence properties of the Gumbel copula, which exhibits only upper-tail dependence. The choice of the Gumbel copula is therefore not surprising.

4.6 Conclusion

In this chapter, we derive very simple non-parametric tests for symmetry and radial symmetry for bivariate copula data, which are computationally very fast. An extensive simulation study is conducted to investigate the finite-sample performance and to compare the proposed tests to the already existing more advanced tests for symmetry and radial symmetry from Genest et al. (2012) and Genest and Nešlehová (2014), respectively, which do not require copula data and are applicable on the original scale of the observations. The results of the Monte Carlo simulation show that the proposed tests for symmetry and radial symmetry overall achieve sufficient empirical power against the various alternatives. In comparison to the more advanced tests with non-parametrically estimated margins, they are slightly less powerful and equally powerful starting from a sample size of 1000. It should be mentioned that the proposed tests are simpler and computationally less expensive and, hence, attractive for huge samples. However, the proposed tests are not consistent and may fail to detect the asymmetry if the two samples resulting from our algorithm have similar Kendall's taus.

Our next contribution is the construction of a powerful non-parametric goodness-of-fit testing procedure for elliptical copulas by combining our proposed tests for symmetry and radial symmetry with our test for copula data in Chapter 3. Hence, the most common intrinsic properties of bivariate elliptical copulas, namely symmetry, radial symmetry, and the equality of Kendall's tau and Blomqvist's beta are utilized. The corresponding Monte Carlo simulation study shows that the proposed testing procedure is more powerful than the test in Chapter 3 for samples from non-symmetric or non-radially symmetric copula families.

Elliptical copulas are very popular in applied sciences. However, their application should be treated with caution. To illustrate the testing procedure for ellipticity in practice, it is applied to financial and insurance data. The first empirical application to data from the US stock market highlights that the dependence structure of two major US stock price indices is not always captured by an elliptical copula. The second application to the loss and ALAE insurance data set indicates that an elliptical copula might not be the right choice to model the corresponding dependence structure.

Our tests for symmetry and radial symmetry can be combined with variance reduction techniques (see, e.g., Korn et al. (2010)). The derivation of the adapted statistical tests for symmetry and radial symmetry is carried out in Section 4.7.1. Here, we also investigate the asymptotic distributions of the corresponding test statistics. Moreover, note that the proposed tests can be based on any bivariate non-parametric measure of ordinal association. Our tests based on Kendall's tau outperform the tests based on Blomqvist's beta while they are comparable with the tests based on Spearman's rho. As an illustration, the results of the simulation study to compare our tests based on Kendall's tau and Spearman's rho are portrayed in Section 4.7.2.

Finally, note that this chapter is devoted to the bivariate case. The development of the testing procedure for ellipticity in higher dimensions is left open for future research.

4.7 Supplementary material

In the following two sections, we give the theoretical derivation of the reflection approach for the test for symmetry and radial symmetry, respectively, and present the empirical results of the tests for symmetry and radial symmetry based on Spearman's rho.

4.7.1 Reflection approach

The proposed tests for symmetry and radial symmetry can be refined by utilizing variance reduction techniques (see, e.g., Korn et al. (2010)). In contrast to the previous approach (see Section 4.2), the considered sub-samples are completely reflected with respect to the main diagonal or the point $(0.5, 0.5)$, respectively. The empirical estimator of Kendall's tau is then based on the enlarged random samples with dependent sample points. As a result, the precision of the statistical estimation is expected to be improved. The theoretical derivations are given in the following two subsections.

Test for symmetry

Consider the sub-sample $(U_1^{B^s}, V_1^{B^s}), \ldots, (U_{N_n^s}^{B^s}, V_{N_n^s}^{B^s})$, whose realizations are below the diagonal and expand it by the reflected counterparts $(V_1^{B^s}, U_1^{B^s}), \ldots, (V_{N_n^s}^{B^s}, U_{N_n^s}^{B^s})$. Thus, instead of choosing either $(U_i^{B^s}, V_i^{B^s})$ or $(V_i^{B^s}, U_i^{B^s})$, both sample points are in the new random sample

$$(U_1^{B^s}, V_1^{B^s}), \ldots, (U_{N_n^s}^{B^s}, V_{N_n^s}^{B^s}), (V_1^{B^s}, U_1^{B^s}), \ldots, (V_{N_n^s}^{B^s}, U_{N_n^s}^{B^s}). \tag{4.12}$$

For the second random sample, we proceed in a similar way with the sub-sample $(U_1^{\overline{B^s}}, V_1^{\overline{B^s}})$, ..., $(U_{N_n^s}^{\overline{B^s}}, V_{N_n^s}^{\overline{B^s}})$, whose realizations are above the diagonal. The resulting random sample is given by

$$(U_1^{\overline{B^s}}, V_1^{\overline{B^s}}), \ldots, (U_{N_n^s}^{\overline{B^s}}, V_{N_n^s}^{\overline{B^s}}), (V_1^{\overline{B^s}}, U_1^{\overline{B^s}}), \ldots, (V_{N_n^s}^{\overline{B^s}}, U_{N_n^s}^{\overline{B^s}}). \tag{4.13}$$

Note that the sample points within both random samples are not independent any more, while the random samples themselves are still independent.

Again, we base our test on the difference of the empirically estimated Kendall's tau $\tau_{2N_n^s}^{B^s,ref}$ and $\tau_{2N_n^s}^{\overline{B^s},ref}$ of the two newly generated random samples in (4.12) and (4.13), respectively, and get

$$S_{N_n^s}^{ref} := \tau_{2N_n^s}^{B^s,ref} - \tau_{2N_n^s}^{\overline{B^s},ref}.$$

Similarly to Section 4.2.1, N_n^s can be assumed to be a sequence of positive integer-valued random variables for which the convergence result in (4.8) holds. The asymptotic distribution of the test statistic $S_{N_n^s}^{ref}$ coincides with the one of the test statistic $S_{N_n^s}$ and is stated in the following theorem.

Theorem 4.4. *Let $(U_1, V_1), \ldots, (U_n, V_n)$ be an i.i.d. random sample from a bivariate random vector (U, V), whose distribution function is a symmetric copula C. Further, let (4.8) hold. Then,*

$$\sqrt{n^\star} \cdot S_{N_n^s}^{ref} \xrightarrow{d} N\left(0, 2\sigma^2\right),$$

where $n^\star = n/2$ and $\sigma^2 = \mathbb{V}\mathrm{ar}\left(2\tilde{h}_1\left((U_1, V_1)\right)\right)$.

Proof of Theorem 4.4. Note that due to the present dependence in the underlying random samples, the proof cannot rely on Sproule (1974) as the proof of Theorem 4.1. Let $(U_1, V_1), \ldots, (U_n, V_n) \in [0,1]^2$ be a sample from the statistical model

$$\left(([0,1]^2)^n, \mathcal{B}([0,1]^2)^{\otimes n}, P^{\otimes n}\right),$$

where P is a distribution with symmetric copula C and uniform marginals. For the samples given in (4.12) and (4.13), the test statistic $S_{N_n^s}^{ref}$ is given by the difference of the corresponding empirical Kendall's tau estimators $\tau_{2N_n^s}^{B^s,ref}$ and $\tau_{2N_n^s}^{\overline{B^s},ref}$.

As a direct consequence from Equation (4.8), $2N_n^s/n$ converges to 1 in probability for $n \to \infty$. By virtue of Theorem 1 from Anscombe (1952), the stated asymptotic convergence holds if the difference $\tau_{2n}^{B^s,ref} - \tau_{2n}^{\overline{B^s},ref}$ satisfies the conditions (C1) and (C2) of Anscombe (1952). Similarly to the proof of Theorem 4.1, it is sufficient to show that $\tau_{2n}^{B^s,ref}$ satisfies (C1) and (C2) (see Equations (4.9) and (4.10) and the statement below). Using the sample given in (4.5) and $N_n^s = n$, the sample given in (4.12) can be alternatively represented by

$$\boldsymbol{W}_i = (W_{1i}, W_{2i}) = \begin{cases} (\tilde{U}_i^{B^s}, \tilde{V}_i^{B^s}), & \text{for } i \in \{1, \dots, n\} \\ (\tilde{V}_i^{B^s}, \tilde{U}_i^{B^s}), & \text{for } i \in \{n+1, \dots, 2n\} . \end{cases}$$

Thus, $\boldsymbol{W}_i$ is for all $i \in \{1, \dots, 2n\}$ a random vector distributed according to C. Additionally, $\boldsymbol{W}_{i_1}$ and $\boldsymbol{W}_{i_2}$ are dependent for $i_1 < i_2$ only if $i_2 = i_1 + n$. Now, we consider the empirical estimator of Kendall's tau

$$\tau_{2n}^{B^s,ref} = \frac{2}{2n(2n-1)} \sum_{1 \le i < j \le 2n} h(\boldsymbol{W}_i, \boldsymbol{W}_j) , \tag{4.14}$$

where

$$h(\boldsymbol{W}_i, \boldsymbol{W}_j) := \operatorname{sgn}(W_{1i} - W_{1j}) \operatorname{sgn}(W_{2i} - W_{2j}) .$$

Noting the possible dependence between $\boldsymbol{W}_i$ and $\boldsymbol{W}_j$, the empirical estimator in (4.14) is split up as follows

$$\tau_{2n}^{B^s,ref} = \frac{2}{2n(2n-1)} \sum_{\substack{1 \le i < j \le 2n \\ j \neq i+n}} h(\boldsymbol{W}_i, \boldsymbol{W}_j) + \frac{2}{2n(2n-1)} \sum_{i=1}^{n} h(\boldsymbol{W}_i, \boldsymbol{W}_{i+n}) . \tag{4.15}$$

For the second summand, it is obvious that

$$\sqrt{n} \cdot \frac{2}{2n(2n-1)} \sum_{i=1}^{n} h(\boldsymbol{W}_i, \boldsymbol{W}_{i+n}) \xrightarrow{a.s.} 0 , \qquad \text{as } n \to \infty . \tag{4.16}$$

Furthermore, the components of the function h in the first summand are independent now and we define

$$U_{2n}^{ind}(h) := \frac{2}{2n(2n-1)} \sum_{\substack{1 \le i < j \le 2n \\ j \neq i+n}} h(\boldsymbol{W}_i, \boldsymbol{W}_j) .$$

Let

$$h_1(\boldsymbol{W}_1) := \tilde{h}_1(\boldsymbol{W}_1) - \tau$$

and

$$h_2(\boldsymbol{W}_1, \boldsymbol{W}_2) := h(\boldsymbol{W}_1, \boldsymbol{W}_2) - \tilde{h}_1(\boldsymbol{W}_1) - \tilde{h}_1(\boldsymbol{W}_2) + \tau .$$

Note that h_2 is degenerate for independent $\boldsymbol{W}_i$ and $\boldsymbol{W}_j$, that is

$$\mathbb{E}\left[h_2(\boldsymbol{W}_i, \boldsymbol{W}_j) | \boldsymbol{W}_i\right] = 0 .$$

Similarly to the Hoeffding decomposition (see Theorem 1.2.1 in Denker (1985)) for U-statistics, we obtain

$$U_{2n}^{ind}(h) = U_{2n,1}^{ind} + U_{2n,2}^{ind} + \frac{2n-2}{2n-1} \tau , \tag{4.17}$$

where

$$U_{2n,1}^{ind} := \frac{2n-2}{n(2n-1)} \sum_{i=1}^{2n} h_1(\boldsymbol{W}_i)$$

and

$$U_{2n,2}^{ind} := \frac{2}{2n(2n-1)} \sum_{\substack{1 \le i < j \le 2n \\ j \ne i+n}} h_2(\boldsymbol{W}_i, \boldsymbol{W}_j) \,.$$

To show that $\sqrt{n}\, U_{2n,2}$ converges to 0 almost surely for $n \to \infty$, consider the fourth moment $\mathbb{E}\left[\left(\sqrt{n} \cdot U_{2n,2}^{ind}\right)^4\right]$. From the degeneracy of h_2, the boundedness of h and straightforward computations, it follows that

$$\begin{aligned}
&\mathbb{E}\left[\left(\sqrt{n} \cdot U_{2n,2}^{ind}\right)^4\right] \\
&= n^2 \left(\frac{2}{2n(2n-1)}\right)^4 \sum_{\substack{1 \le i_1 < j_1 \le 2n \\ j \ne i+n}} \sum_{\substack{1 \le i_2 < j_2 \le 2n \\ j \ne i+n}} \sum_{\substack{1 \le i_3 < j_3 \le 2n \\ j \ne i+n}} \sum_{\substack{1 \le i_4 < j_4 \le 2n \\ j \ne i+n}} \\
&\qquad \mathbb{E}\left[h_2(\boldsymbol{W}_{i_1}, \boldsymbol{W}_{j_1}) h_2(\boldsymbol{W}_{i_2}, \boldsymbol{W}_{j_2}) h_2(\boldsymbol{W}_{i_3}, \boldsymbol{W}_{j_3}) h_2(\boldsymbol{W}_{i_4}, \boldsymbol{W}_{j_4})\right] \\
&\le \frac{Kn^4}{n^2(2n-1)^4} \le \frac{K}{n^2} \,,
\end{aligned}$$

where $K > 0$ is an absolute constant and Kn^4 constitutes an upper limit for the number of non-zero summands. The Markov inequality implies

$$\sum_{n=1}^{\infty} \mathbb{P}\left(|\sqrt{n}\, U_{2n,2}^{ind}| > \epsilon\right) \le \sum_{n=1}^{\infty} \frac{1}{\epsilon^4} \mathbb{E}\left[|\sqrt{n} \cdot U_{2n,2}^{ind}|^4\right] \le \frac{K}{\epsilon^4} \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty$$

and, hence, $\sqrt{n}\, U_{2n,2}$ converges to 0 almost surely for $n \to \infty$.

From the above considerations, it follows that

$$\sqrt{n}\, \left(\tau_{2n}^{B^s,ref} - \tau\right) \qquad \text{and} \qquad \sqrt{n} \cdot U_{2n,1}^{ind}$$

have the same asymptotic distribution. Note that h_1 is symmetric with respect to the components of $\boldsymbol{W}_i$ under the null hypothesis and, hence,

$$U_{2n,1}^{ind} = \frac{2n-2}{2n-1} \cdot \frac{2}{n} \sum_{i=1}^{n} h_1(\boldsymbol{W}_i) \,.$$

The central limit theorem now implies that $\sqrt{n}\, U_{2n,1}^{ind}$ converges in distribution to a centered normal distribution with variance $\sigma^2 = \mathbb{V}\mathrm{ar}\left(2\tilde{h}_1\left((U_1, V_1)\right)\right)$. Thus, $U_{2n,1}^{ind}$ and therefore also $\tau_{2n}^{B^s,ref}$ satisfy (C1) with $w_n = 1/\sqrt{n}$.

By Theorem 3 of Anscombe (1952), $U_{2n,1}^{ind}$ satisfies (C2). Therefore, (4.15), (4.16), (4.17) and the almost sure convergence of $\sqrt{n}\, U_{2n,2}^{ind}$ imply that $\tau_{2n}^{B^s,ref}$ satisfies (C2).

Similarly, it follows that $\tau_{2n}^{\overline{B^s},ref}$ satisfies (C1) and (C2). Following the steps in the proof of Theorem 4.1 finally yields

$$\sqrt{n^\star}\left(\tau_{2N_n^s}^{B^s,ref} - \tau_{2N_n^s}^{\overline{B^s},ref}\right) \xrightarrow{d} N\left(0, 2\sigma^2\right).$$

□

The asymptotic variance $2\sigma^2$ is identical to the asymptotic variance in Theorem 4.1 and, therefore, Remark 4.2 provides a consistent estimation procedure for the unknown variance of $S_{N_n^s}^{ref}$.

Test for radial symmetry

The two sub-samples with realizations below and above the counter-diagonal are expanded by their reflected counterparts resulting in the new random samples

$$(U_1^{B^r}, V_1^{B^r}), \dots, (U_{N_n^r}^{B^r}, V_{N_n^r}^{B^r}), (1-U_1^{B^r}, 1-V_1^{B^r}), \dots, (1-U_{N_n^r}^{B^r}, 1-V_{N_n^r}^{B^r}) \tag{4.18}$$

and

$$(U_1^{\overline{B^r}}, V_1^{\overline{B^r}}), \dots, (U_{N_n^r}^{\overline{B^r}}, V_{N_n^r}^{\overline{B^r}}), (1-U_1^{\overline{B^r}}, 1-V_1^{\overline{B^r}}), \dots, (1-U_{N_n^r}^{\overline{B^r}}, 1-V_{N_n^r}^{\overline{B^r}}). \tag{4.19}$$

Again, the sample points within both random samples are not independent any more, while the random samples themselves are still independent.

The test is based on the difference

$$R_{N_n^r}^{ref} := \tau_{2N_n^r}^{B^r,ref} - \tau_{2N_n^r}^{\overline{B^r},ref},$$

where $\tau_{2N_n^r}^{B^r,ref}$ and $\tau_{2N_n^r}^{\overline{B^r},ref}$ denote the empirically estimated Kendall's tau based on the random samples in (4.18) and (4.19), respectively. Similarly to Section 4.2.2, N_n^r can be assumed to be a sequence of positive integer-valued random variables for which the convergence result in (4.11) holds. Note that the asymptotic distribution of the test statistic $R_{N_n^r}^{ref}$ is stated in the following theorem does not coincide with the one of the test statistic $R_{N_n^r}$.

Theorem 4.5. *Let $(U_1, V_1), \dots, (U_n, V_n)$ be an i.i.d. random sample from a bivariate random vector (U, V), whose distribution function is a radially symmetric copula C. Further, let (4.11) hold. Then,*

$$\sqrt{n^\star} \cdot R_{N_n^r}^{ref} \xrightarrow{d} N\left(0, 2\sigma_{r,ref}^2\right),$$

where $n^\star = n/2$ and

$$\sigma_{r,ref}^2 = \mathbb{V}\mathrm{ar}\left(\tilde{h}_1\big((U_1, V_1)\big)\right) + \mathbb{V}\mathrm{ar}\left(\tilde{h}_1\big((1-U_1, 1-V_1)\big)\right) + 2\mathbb{C}\mathrm{ov}\left(\tilde{h}_1\big((U_1, V_1)\big), \tilde{h}_1\big((1-U_1, 1-V_1)\big)\right).$$

Proof of Theorem 4.5. The proof is in line with the proof of Theorem 4.4. The difference in the asymptotic variance emerges from the fact that the function $\tilde{h}_1$ is not radially symmetric with respect to the components of its argument. Note that, for $N_n^r = n$, the sample given in (4.18) can be alternatively represented by

$$\boldsymbol{W}_i = (W_{1i}, W_{2i}) = \begin{cases} (\tilde{U}_i^{B^r}, \tilde{V}_i^{B^r}), & \text{for } i \in \{1, \ldots, n\} \\ (1 - \tilde{U}_i^{B^r}, 1 - \tilde{V}_i^{B^r}), & \text{for } i \in \{n+1, \ldots, 2n\}, \end{cases}$$

where $(\tilde{U}_i^{B^r}, \tilde{V}_i^{B^r})$ is defined similarly to (4.5) but for the case of radial symmetry.

Since $\tilde{h}_1$ is not radially symmetric with respect to the components of $\boldsymbol{W}_i$ under the null hypothesis, $U_{2n,1}^{ind}$ from the proof of Theorem 4.4 becomes

$$U_{2n,1}^{ind} = \frac{2n-2}{2n-1} \cdot \frac{1}{n} \sum_{i=1}^{n} \left(h_1(\boldsymbol{W}_i) + h_1(\boldsymbol{1} - \boldsymbol{W}_i) \right) ,$$

where $\boldsymbol{1} - \boldsymbol{W}_i = (1 - W_{1i}, 1 - W_{2i})$. The central limit theorem now implies that $\sqrt{n} \cdot U_{2n,1}^{ind}$ converges in distribution to a centered normal distribution with variance

$$\begin{aligned} \sigma_{r,ref}^2 = \mathbb{V}\mathrm{ar}\left(\tilde{h}_1\Big((U_1, V_1)\Big)\right) + \mathbb{V}\mathrm{ar}\left(\tilde{h}_1\Big((1 - U_1, 1 - V_1)\Big)\right) \\ + 2\mathbb{C}\mathrm{ov}\left(\tilde{h}_1\Big((U_1, V_1)\Big), \tilde{h}_1\Big((1 - U_1, 1 - V_1)\Big)\right) . \end{aligned}$$

Altogether, this results in

$$\sqrt{n^\star} \left(\tau_{2N_n^r}^{B^r, ref} - \tau_{2N_n^r}^{\overline{B^r}, ref} \right) \xrightarrow{d} N\left(0, 2\sigma_{r,ref}^2\right) .$$

□

The components of the asymptotic variance $2\sigma_{r,ref}^2$ can be estimated consistently in the framework of Remark 4.2.

Additional remarks

Note that the tests, as presented in Section 4.2, are based on a random sampling algorithm. The reflection of the entire sample leads to an elimination of the inherited variability of the test statistic. Hence, the resulting tests for symmetry and radial symmetry based on the reflection approach do not require a repeated execution as done in the empirical analysis in Section 4.5. This contemplates a computational advantage due to reduced processing times for the tests.

We conducted a simulation study to compare the tests for symmetry and radial symmetry from Section 4.2 to the ones based on the reflection approach presented above. In accordance with the theoretical results, we got similar empirical results for the test for symmetry with and without the reflection approach and better empirical results for the test for radial symmetry using the reflection approach than without. However, note that the stated reduced computation time also makes the reflection approach beneficial for the test for symmetry.

4.7.2 Performance of the simple tests for symmetry and radial symmetry based on Spearman's rho

Recall that the tests for symmetry and radial symmetry are originally based on Kendall's tau. However, they can be based on any bivariate non-parametric measure of bivariate rank correlation. In this section, we illustrate the performance of the tests for symmetry and radial symmetry when they are based on Spearman's rho instead. Note that these tests and the corresponding asymptotic distributions can be derived in a similar manner to that in Section 4.2.

Spearman's rho is an alternative rank correlation coefficient, which is defined as follows.

Definition 4.6. ***(Spearman's rho)***
Let (X', Y') and (X'', Y'') be independent copies of the random vector (X, Y) of continuous random variables X and Y. Then, Spearman's rho is defined by

$$\rho_{12} := 3\left(\mathbb{P}((X - X')(Y - Y'') > 0) - \mathbb{P}((X - X')(Y - Y'') < 0)\right) .$$

Hence, Spearman's rho equals the probability of concordance minus the probability of discordance for the two vectors (X, Y) and (X', Y''). Furthermore, for continuous random variables X and Y with copula C, Spearman's rho is completely determined by their copula C (see Theorem 5.1.6 in Nelsen (1999)) and can be expressed as

$$\rho_{XY} = \rho_C = 12 \int_0^1 \int_0^1 uv \, dC(u, v) - 3 .$$

It easily follows that for copula data Spearman's rho is equal to the linear correlation. More, precisely, let (U, V) be distributed according to the bivariate copula C, then it holds that

$$\rho_C = \mathbb{C}\mathrm{or}(U, V) .$$

To compare the performance of the tests for symmetry and radial symmetry based on Kendall's tau and Spearman's rho, respectivley, we have conducted a simulation study. with a similar setup to the one described in Section 4.3. Tables 4.9 and 4.10 present the empirical level and power of our test for symmetry based on Kendall's tau (in Column JMS), the test for symmetry based on Spearman's rho (in Column SpS), the corresponding tests for pseudo-observations (in Column JMSP and SpSP), and the test from Genest et al. (2012) (in Column GNQ). Tables 4.11 and 4.12 display the empirical level and power of our test for radial symmetry based on Kendall's tau (in Column JMR), the test for radial symmetry based on Spearman's rho (in Column SpR), the corresponding tests for pseudo-observations (in Column JMRP and SpRP), and the test from Genest and Nešlehová (2014) (in Column GN).

Regarding the empirical level in Tables 4.9 and 4.11, we observe that the test for symmetry as well as the test for radial symmetry based on Spearman's rho are slightly too liberal for small sample sizes, especially for $n = 100$. As a consequence, the higher empirical power for the tests based on Spearman's rho compared to the ones based on Kendall's tau for small sample sizes (see Tables 4.10 and 4.12) cannot be unequivocally classified as a superior behavior of the former. For larger sample sizes, all tests hold their nominal level well independently of the underlying rank correlation coefficient (see Tables 4.9 and 4.11). For pseudo-observations, the tests based on Spearman's rho and Kendall's tau show very similar behavior with respect to the empirical level.

Looking at the corresponding empirical power results in Tables 4.10 and 4.12 for larger sample sizes, that is $n = 500$ and $n = 1000$, contradictory trends can be observed for the tests for symmetry compared to the tests for radial symmetry. More precisely, for the symmetry tests, the test based on Spearman's rho performs better than the one based on Kendall's tau, whereas for the radial symmetry tests, the rejection rates for the test based on Kendall's tau tend to be higher than the ones for the test based on Spearman's rho. For pseudo-observations, the observations are similar to the ones described above for copula data.

Since there is no clear winner regarding the empirical power results and the tests based on Spearman's rho for copula data had some problems in holding their nominal level, we decided to base our tests for symmetry and radial symmetry on Kendall's tau.

C τ	JMS	SpS	JMSP	SpSP	GNQ	JMS	SpS	JMSP	SpSP	GNQ
	$n = 100$					$n = 250$				
Gauss										
0.25	0.066	0.077	0.012	0.016	0.022	0.055	0.054	0.016	0.019	0.039
0.50	0.060	0.072	0.012	0.015	0.015	0.061	0.064	0.006	0.009	0.027
0.75	0.041	0.048	0.024	0.026	0.011	0.052	0.049	0.021	0.014	0.011
$t_{\nu=5}$										
0.25	0.059	0.070	0.014	0.026	0.033	0.043	0.051	0.016	0.020	0.046
0.50	0.047	0.061	0.018	0.028	0.014	0.058	0.066	0.016	0.017	0.035
0.75	0.034	0.034	0.027	0.031	0.022	0.055	0.063	0.019	0.019	0.014
Frank										
0.25	0.054	0.080	0.014	0.024	0.031	0.052	0.061	0.014	0.016	0.038
0.50	0.057	0.070	0.024	0.025	0.015	0.060	0.063	0.013	0.014	0.025
0.75	0.036	0.054	0.017	0.013	0.016	0.032	0.050	0.009	0.007	0.011

(To be continued)

Table 4.9: Empirical level of our test for symmetry (JMS), the test for symmetry based on Spearman's rho (SpS), the corresponding tests for pseudo-observations (JMSP and SpSP), and the test from Genest et al. (2012) (GNQ) with significance level $\alpha = 0.05$: rate of rejecting H_0 as observed in 1000 random samples of size n from copula family C with Kendall's tau τ.

C										
τ	JMS	SpS	JMSP	SpSP	GNQ	JMS	SpS	JMSP	SpSP	GNQ
Clayton										
0.25	0.060	0.088	0.024	0.033	0.033	0.062	0.073	0.018	0.018	0.040
0.50	0.063	0.070	0.029	0.030	0.031	0.050	0.060	0.013	0.008	0.029
0.75	0.059	0.058	0.051	0.043	0.021	0.059	0.059	0.028	0.031	0.015
Gumbel										
0.25	0.063	0.075	0.022	0.033	0.036	0.049	0.054	0.020	0.020	0.038
0.50	0.057	0.054	0.015	0.022	0.027	0.050	0.054	0.010	0.013	0.026
0.75	0.053	0.049	0.034	0.029	0.017	0.051	0.054	0.021	0.022	0.013
	$n = 500$					$n = 1000$				
Gauss										
0.25	0.055	0.052	0.012	0.014	0.048	0.054	0.058	0.007	0.009	0.044
0.50	0.045	0.050	0.010	0.005	0.020	0.051	0.044	0.006	0.007	0.032
0.75	0.046	0.049	0.016	0.008	0.004	0.050	0.050	0.006	0.002	0.013
$t_{\nu=5}$										
0.25	0.052	0.055	0.023	0.030	0.035	0.059	0.063	0.013	0.016	0.035
0.50	0.049	0.051	0.009	0.009	0.031	0.062	0.057	0.006	0.010	0.046
0.75	0.057	0.053	0.010	0.013	0.013	0.051	0.052	0.008	0.009	0.019
Frank										
0.25	0.051	0.046	0.010	0.011	0.043	0.045	0.051	0.009	0.008	0.032
0.50	0.052	0.052	0.005	0.005	0.038	0.061	0.069	0.007	0.006	0.035
0.75	0.042	0.044	0.010	0.010	0.006	0.048	0.040	0.009	0.007	0.016
Clayton										
0.25	0.052	0.059	0.010	0.009	0.032	0.051	0.046	0.009	0.008	0.043
0.50	0.059	0.042	0.003	0.005	0.029	0.045	0.039	0.004	0.006	0.035
0.75	0.056	0.054	0.015	0.023	0.021	0.049	0.051	0.009	0.008	0.027
Gumbel										
0.25	0.061	0.066	0.011	0.014	0.035	0.049	0.050	0.013	0.011	0.042
0.50	0.053	0.054	0.004	0.006	0.024	0.049	0.051	0.006	0.007	0.039
0.75	0.050	0.052	0.019	0.013	0.008	0.049	0.043	0.003	0.011	0.028

Table 4.9: (continued)

C τ	JMS	SpS	JMSP	SpSP	GNQ	JMS	SpS	JMSP	SpSP	GNQ
	$n = 100$					$n = 250$				
Gauss										
$\delta = 0.25$										
0.50	0.095	0.133	0.067	0.084	0.082	0.157	0.199	0.140	0.167	0.233
0.75	0.434	0.571	0.553	0.668	0.618	0.849	0.942	0.954	0.992	0.995
0.90	0.963	0.987	0.983	0.996	0.996	1.000	1.000	1.000	1.000	1.000
$\delta = 0.5$										
0.50	0.155	0.203	0.145	0.186	0.199	0.296	0.365	0.312	0.378	0.499
0.75	0.672	0.783	0.798	0.883	0.907	0.976	0.989	0.997	1.000	1.000
0.90	0.973	0.988	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\delta = 0.75$										
0.50	0.158	0.201	0.134	0.167	0.168	0.298	0.339	0.289	0.338	0.393
0.75	0.521	0.598	0.576	0.647	0.626	0.896	0.919	0.961	0.969	0.988
0.90	0.657	0.736	0.744	0.783	0.844	0.979	0.976	0.996	0.996	0.999
Clayton										
$\delta = 0.25$										
0.50	0.177	0.203	0.149	0.158	0.093	0.361	0.348	0.343	0.341	0.260
0.75	0.678	0.722	0.763	0.771	0.779	0.958	0.966	0.998	0.996	1.000
0.90	0.892	0.974	0.981	0.998	0.999	1.000	1.000	1.000	1.000	1.000
$\delta = 0.5$										
0.50	0.148	0.194	0.115	0.141	0.111	0.307	0.313	0.264	0.265	0.339
0.75	0.463	0.569	0.550	0.651	0.834	0.871	0.918	0.953	0.972	1.000
0.90	0.817	0.907	0.920	0.975	0.999	0.998	1.000	1.000	1.000	1.000
$\delta = 0.75$										
0.50	0.092	0.121	0.058	0.072	0.072	0.132	0.147	0.110	0.111	0.169
0.75	0.190	0.276	0.189	0.250	0.366	0.440	0.491	0.469	0.533	0.814
0.90	0.460	0.541	0.515	0.560	0.764	0.827	0.867	0.919	0.944	0.997
Gumbel										
$\delta = 0.25$										
0.50	0.142	0.193	0.149	0.172	0.110	0.263	0.301	0.268	0.308	0.275
0.75	0.592	0.702	0.743	0.793	0.679	0.944	0.983	0.982	0.996	0.997
0.90	0.990	0.997	0.991	0.999	1.000	1.000	1.000	1.000	1.000	1.000
$\delta = 0.5$										
0.50	0.285	0.350	0.305	0.357	0.272	0.599	0.651	0.669	0.714	0.704
0.75	0.862	0.920	0.950	0.963	0.970	0.997	1.000	1.000	1.000	1.000
0.90	0.987	0.998	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000

(To be continued)

Table 4.10: Empirical power of our test for symmetry (JMS), the test for symmetry based on Spearman's rho (SpS), the corresponding tests for pseudo-observations (JMSP and SpSP), and the test from Genest et al. (2012) (GNQ) with significance level $\alpha = 0.05$: rate of rejecting H_0 as observed in 1000 random samples of size n from copula family C asymmetrized with parameter δ and with Kendall's tau τ.

C τ	JMS	SpS	JMSP	SpSP	GNQ	JMS	SpS	JMSP	SpSP	GNQ
$\delta = 0.75$										
0.50	0.273	0.348	0.284	0.337	0.284	0.638	0.670	0.690	0.725	0.690
0.75	0.619	0.702	0.693	0.734	0.752	0.951	0.963	0.985	0.987	0.993
0.90	0.722	0.792	0.799	0.842	0.893	0.987	0.991	0.999	0.997	1.000
	$n = 500$					$n = 1000$				
Gauss										
$\delta = 0.25$										
0.50	0.285	0.371	0.242	0.340	0.466	0.480	0.589	0.530	0.665	0.803
0.75	0.991	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.90	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\delta = 0.5$										
0.50	0.564	0.633	0.645	0.727	0.851	0.846	0.895	0.936	0.965	0.989
0.75	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.90	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\delta = 0.75$										
0.50	0.542	0.581	0.604	0.645	0.764	0.846	0.859	0.889	0.917	0.968
0.75	0.996	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.90	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Clayton										
$\delta = 0.25$										
0.50	0.586	0.579	0.633	0.583	0.548	0.885	0.873	0.954	0.925	0.909
0.75	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.90	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\delta = 0.5$										
0.25	0.509	0.537	0.578	0.586	0.715	0.813	0.822	0.901	0.898	0.965
0.50	0.991	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.75	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\delta = 0.75$										
0.50	0.197	0.213	0.173	0.190	0.295	0.370	0.375	0.365	0.374	0.586
0.75	0.751	0.788	0.832	0.877	0.988	0.961	0.974	0.990	0.996	1.000
0.90	0.989	0.990	0.998	0.999	1.000	1.000	1.000	1.000	1.000	1.000
Gumbel										
$\delta = 0.25$										
0.50	0.515	0.576	0.573	0.632	0.637	0.744	0.796	0.855	0.890	0.916
0.75	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.90	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\delta = 0.5$										
0.50	0.895	0.911	0.963	0.974	0.974	0.992	0.998	0.999	0.999	1.000
0.75	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.90	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\delta = 0.75$										
0.50	0.888	0.900	0.966	0.971	0.963	0.990	0.993	1.000	1.000	1.000
0.75	0.999	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.90	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 4.10: (continued)

C τ	JMR	SpR	JMRP	SpRP	GN	JMR	SpR	JMRP	SpRP	GN
	$n = 100$					$n = 250$				
Gauss										
0.25	0.049	0.071	0.049	0.052	0.041	0.058	0.062	0.029	0.039	0.047
0.50	0.055	0.062	0.059	0.065	0.037	0.053	0.053	0.053	0.060	0.044
0.75	0.052	0.050	0.073	0.076	0.042	0.046	0.048	0.077	0.082	0.051
$t_{\nu=5}$										
0.25	0.063	0.076	0.038	0.057	0.050	0.065	0.071	0.040	0.043	0.052
0.50	0.041	0.058	0.040	0.045	0.031	0.053	0.054	0.042	0.052	0.043
0.75	0.052	0.059	0.054	0.056	0.029	0.053	0.054	0.064	0.067	0.051
Frank										
0.25	0.061	0.084	0.033	0.043	0.039	0.053	0.061	0.034	0.037	0.045
0.50	0.044	0.068	0.067	0.071	0.052	0.046	0.058	0.067	0.066	0.049
0.75	0.032	0.062	0.111	0.095	0.037	0.044	0.047	0.116	0.098	0.040
	$n = 500$					$n = 1000$				
Gauss										
0.25	0.054	0.057	0.030	0.029	0.042	0.042	0.049	0.030	0.025	0.049
0.50	0.060	0.060	0.055	0.056	0.059	0.047	0.053	0.045	0.044	0.044
0.75	0.039	0.041	0.060	0.066	0.052	0.056	0.057	0.079	0.079	0.051
$t_{\nu=5}$										
0.25	0.052	0.055	0.023	0.030	0.035	0.059	0.063	0.013	0.016	0.035
0.50	0.049	0.051	0.009	0.009	0.031	0.062	0.057	0.006	0.010	0.046
0.75	0.057	0.053	0.010	0.013	0.013	0.051	0.052	0.008	0.009	0.019
Frank										
0.25	0.054	0.054	0.034	0.029	0.039	0.056	0.059	0.036	0.038	0.049
0.50	0.056	0.061	0.056	0.054	0.057	0.049	0.049	0.039	0.044	0.040
0.75	0.047	0.048	0.049	0.052	0.036	0.051	0.050	0.052	0.054	0.052

Table 4.11: Empirical level of our test for radial symmetry (JMR), the test for radial symmetry based on Spearman's rho (SpR), the corresponding tests for pseudo-observations (JMRP and SpRP), and the test from Genest and Nešlehová (2014) (GN) with significance level $\alpha = 0.05$: rate of rejecting H_0 as observed in 1000 random samples of size n from copula family C with Kendall's tau τ.

C τ	JMR	SpR	JMRP	SpRP	GN	JMR	SpR	JMRP	SpRP	GN
	$n = 100$					$n = 250$				
Clayton										
0.25	0.256	0.295	0.229	0.245	0.377	0.517	0.526	0.506	0.483	0.730
0.50	0.625	0.628	0.640	0.600	0.811	0.955	0.938	0.965	0.945	0.997
0.75	0.775	0.846	0.884	0.800	0.921	0.997	0.998	0.999	0.994	1.000
Gumbel										
0.25	0.119	0.151	0.123	0.145	0.092	0.207	0.226	0.215	0.217	0.246
0.50	0.166	0.177	0.193	0.198	0.161	0.413	0.365	0.447	0.373	0.458
0.75	0.166	0.218	0.234	0.222	0.132	0.516	0.450	0.575	0.457	0.495
S-$t_{\nu=4}$										
0.25	0.470	0.494	0.493	0.501	0.514	0.885	0.853	0.905	0.867	0.951
0.50	0.331	0.338	0.395	0.358	0.336	0.713	0.661	0.734	0.640	0.770
0.75	0.152	0.218	0.230	0.231	0.113	0.497	0.472	0.575	0.479	0.436
	$n = 500$					$n = 1000$				
Clayton										
0.25	0.851	0.826	0.858	0.824	0.959	0.983	0.978	0.993	0.984	1.000
0.50	1.000	0.998	1.000	0.998	1.000	1.000	1.000	1.000	1.000	1.000
0.75	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Gumbel										
0.25	0.343	0.334	0.342	0.315	0.491	0.615	0.555	0.638	0.565	0.800
0.50	0.703	0.595	0.722	0.607	0.814	0.934	0.859	0.948	0.871	0.987
0.75	0.814	0.668	0.823	0.688	0.828	0.981	0.917	0.985	0.909	0.992
S-$t_{\nu=4}$										
0.25	0.996	0.987	0.998	0.987	1.000	1.000	1.000	1.000	1.000	1.000
0.50	0.965	0.912	0.963	0.894	0.991	1.000	0.997	1.000	0.998	1.000
0.75	0.834	0.750	0.878	0.722	0.843	0.991	0.953	0.993	0.937	0.997

Table 4.12: Empirical power of our test for radial symmetry (JMR), the test for radial symmetry based on Spearman's rho (SpR), the corresponding tests for pseudo-observations (JMRP and SpRP), and the test from Genest and Nešlehová (2014) (GN) with significance level $\alpha = 0.05$: rate of rejecting H_0 as observed in 1000 random samples of size n from copula family C with Kendall's tau τ.

5

Further development of the simple non-parametric goodness-of-fit test for elliptical copulas

5.1 Introduction

For a classification of this chapters topic within the research context, we refer to the introduction of Chapter 3 (see Section 3.1). Therein, the relevance of a statistical testing procedure to assess whether the underlying dependence structure of a given random sample is captured by an elliptical copula is outlined. For given copula data, Chapter 3 then provides us with a simple goodness-of fit test for the null hypothesis that the unknown copula C of the given data belongs to the class of elliptical copulas $\mathcal{C}^{ellipt}$, that is for

$$H_0 : C \in \mathcal{C}^{ellipt} \,,$$

against the alternative

$$H_1 : C \notin \mathcal{C}^{ellipt} \,.$$

The core element of the simple test is the known equality of Kendall's tau and Blomqvist's beta for all bivariate margins under the null hypothesis H_0 of an elliptical copula. The test is then based on a Wald-type test statistic which, under H_0, asymptotically follows a χ^2 distribution (for Details see Chapter 3).

The simple goodness-of-fit test from Chapter 3 is restrictive with respect to the following two issues. On the one hand, it assumes given copula data and, therefore, neglects unknown marginal distribution functions and their estimation. In practical applications, marginal distribution functions are usually unknown and are estimated parametrically or non-parametrically. The empirical results in Section 3.5 indicate that the finite sample performance of the simple test is not influenced by the non-parametric estimation of unknown marginal distributions. However, the resulting dependence of the pseudo-observations affects the statistical inference of the test statistic. On the other hand, from the simulation study in Section 3.5 it becomes apparent that the simple test has problems to hold its nominal level in higher dimensions.

The main objective of this chapter is the advanced development of the simple test to resolve the two stated restrictions. Note that these also get addressed in Quessy (2020).

However, the results presented in the following have been developed independently over the past few years since the publication of Jaser et al. (2017). For bivariate data with unknown marginal distributions, the asymptotic distribution of the statistic defined by the difference between Blomqvist's beta and Kendall's tau is outlined in Jaser et al. (2017) by considering the empirical copula process and applying the functional Delta method (see Theorem 3.9.4 in van der Vaart and Wellner (1996)). One objective of this chapter is to formalize this derivation and generalize it for data of any dimension. This enables us to drop the first restriction of given copula data and constitutes an alternative to the proofs in Quessy (2020).

The second restriction is a consequence of the fact that the design of the test statistic T_n defined in Equation (3.12) involves a consistent estimator of the covariance matrix ν. As the dimension d increases, the number of elements to be estimated in this $d(d-1)/2 \times d(d-1)/2$-dimensional matrix gets large and huge sample sizes are needed to reasonably estimate the covariance matrix. The slow rate of the distributional approximation with the asymptotic χ^2-distribution for increasing sample sizes in dimension $d = 6$ becomes apparent from Figure 3.5. To overcome this second restriction of the simple test, we propose an L_2-type test statistic in this chapter, which does not require an estimation of the covariance matrix. The limiting distribution of the new test statistic still depends on the unknown copula C and, hence, some bootstrap procedure is needed. The most popular approximation is the multiplier bootstrap, which is also used in Quessy (2020). In the following, we present an alternative approach by making use of the subsampling approximation introduced in Kojadinovic and Stemikovskaya (2019). Regarding the finite sample performance, the subsampling approximation shows significantly superior behavior compared to the empirical bootstrap and equivalent performance compared to the multiplier bootstrap.

The remainder of this chapter is organized as follows. In Section 5.2 the empirical copula process and its weak convergence are introduced. The test statistic and its asymptotic behavior as well as a subsampling approximation are presented in Section 5.3. Section 5.4 describes the setup of the simulation study to analyze the finite sample performance. Finally, Section 5.5 concludes. Some auxiliary results are deferred to Section 5.6.

5.2 Preliminaries

Let $\boldsymbol{X} \in \mathbb{R}^d$ be a d-dimensional random vector with cumulative distribution function F and continuous univariate marginal distribution functions $F_1, \ldots, F_d$ and copula C. Now, let $\boldsymbol{X}_1 = (X_{11}, \ldots, X_{d1}), \ldots, \boldsymbol{X}_n = (X_{1n}, \ldots, X_{dn}) \in \mathbb{R}^d$ be a random sample of n independent observations from the random vector $\boldsymbol{X}$. If the marginal distribution functions $F_1, \ldots, F_d$ are known, copula data can be easily derived. Dropping the assumption of known marginal distribution functions, it is a natural approach to estimate the unknown margins by their corresponding empirical counterparts $F_{1,n}, \dots, F_{d,n}$ defined in Equation (2.10) to get pseudo-observations $\widehat{\boldsymbol{U}}_1 = (\widehat{U}_{11}, \dots, \widehat{U}_{d1}), \dots, \widehat{\boldsymbol{U}}_n = (\widehat{U}_{1n}, \dots, \widehat{U}_{dn})$ defined in Equation (2.12).

Inspired by Sklar's theorem (see Theorem 2.1), the empirical copula C_n of the sample is defined in Equation (2.11). Alternatively, the empirical copula can also be defined as the empirical cumulative distribution function of the sample of pseudo-observations resulting in the definition of $\widehat{C}_n$ in Equation (2.13). These definitions of the empirical copula lead to the following empirical copula processes

$$\mathbb{C}_n(\boldsymbol{u}) = \sqrt{n}(C_n(\boldsymbol{u}) - C(\boldsymbol{u}))\,,$$

and

$$\widehat{\mathbb{C}}_n(\boldsymbol{u}) = \sqrt{n}(\widehat{C}_n(\boldsymbol{u}) - C(\boldsymbol{u}))\,.$$

In the sequel, $\rightsquigarrow$ denotes weak convergence in the metric space $\ell^\infty([0,1]^d)$, which is the space of all uniformly bounded functions on the unit hypercube $[0,1]^d$ equipped with the metric induced by the supremum norm. For details on weak convergence and empirical processes, we refer to van der Vaart and Wellner (1996) and Kosorok (2008). In the following condition, a non restrictive smoothness assumption on the copula C is stated.

Condition 5.1. *For $k = 1, \dots, d$, the first-order partial derivatives $\partial_k C(\boldsymbol{u})$ exist and are continuous on the set $\{\boldsymbol{u} \in [0,1]^d : u_k \in (0,1)\}$.*

For completeness, we define $\partial_k C(\boldsymbol{u}) = \limsup_{h \downarrow 0} \{C(\boldsymbol{u} + h\boldsymbol{e}_k) - C(\boldsymbol{u})\}/h$ for $\boldsymbol{u} \in [0,1]^d \setminus \{\boldsymbol{u} \in [0,1]^d : u_k \in (0,1)\}$, where $\boldsymbol{e}_k$ denotes the kth unit vector. Based on this condition, Segers (2012) derived the asymptotics of the empirical copula process $\mathbb{C}_n$.

Proposition 5.2. (Segers (2012))
If the copula C satisfies Condition 5.1, the empirical copula process $\mathbb{C}_n$ converges weakly towards the Gaussian field $\mathbb{G}_C$, that is

$$\mathbb{C}_n = \sqrt{n}(C_n - C) \rightsquigarrow \mathbb{G}_C\,, \qquad \textit{in} \quad \ell^\infty([0,1]^d)\,.$$

The limiting Gaussian field $\mathbb{G}_C$ is defined, for all $\boldsymbol{u} \in [0,1]^d$, by

$$\mathbb{G}_C(\boldsymbol{u}) = \mathbb{B}_C(\boldsymbol{u}) - \sum_{j=1}^{d} \partial_j C(\boldsymbol{u}) \mathbb{B}_C(\boldsymbol{u}^{(j)})\,, \tag{5.1}$$

Further, $\mathbb{B}_C$ is a d-dimensional C-Brownian bridge on $[0,1]^d$, that is a tight centered Gaussian process with covariance function given, for all $\boldsymbol{u}, \boldsymbol{v} \in [0,1]^d$, by

$$\mathbb{C}\mathrm{ov}(\mathbb{B}_C(\boldsymbol{u}), \mathbb{B}_C(\boldsymbol{v})) = C(\boldsymbol{u} \wedge \boldsymbol{v}) - C(\boldsymbol{u})C(\boldsymbol{v})\,,$$

where $\boldsymbol{u} \wedge \boldsymbol{v} = (\min\{u_1, v_1\}, \dots, \min\{u_d, v_d\})$.

The main result of this paper is based on the weak convergence of the empirical copula process $\widehat{\mathbb{C}}_n$. Kojadinovic and Stemikovskaya (2019) outline that Proposition 5.2 also holds for $\widehat{\mathbb{C}}_n$, that is

$$\widehat{\mathbb{C}}_n = \sqrt{n}(\widehat{C}_n - C) \rightsquigarrow \mathbb{G}_C\,, \qquad in \quad \ell^\infty([0,1]^d)\,. \tag{5.2}$$

Furthermore for $k, \ell \in \{1, \dots, d\}$ with $k < \ell$, the corresponding bivariate margin $\mathbb{G}_{C_{k\ell}}$ of $\mathbb{G}_C$ can be derived in the following way. Let $(u_k, u_\ell) \in [0,1]^2$, then we have

$$\begin{aligned}\mathbb{G}_{C_{k\ell}}(u_k, u_\ell) = \mathbb{G}_C(\boldsymbol{u}^{(k,\ell)}) &= \mathbb{B}_C(\boldsymbol{u}^{(k,\ell)}) - \sum_{j=1}^{d} \partial_j C(\boldsymbol{u}^{(k,\ell)})\mathbb{B}_C(\boldsymbol{u}^{(k,\ell)^{(j)}}) \\ &= \mathbb{B}_{C_{k\ell}}(u_k, u_\ell) - \partial_k C_{k\ell}(u_k, u_\ell)\mathbb{B}_{C_{k\ell}}(u_k, 1) - \partial_\ell C_{k\ell}(u_k, u_\ell)\mathbb{B}_{C_{k\ell}}(1, u_\ell)\,,\end{aligned}$$

where $\mathbb{B}_{C_{k\ell}}$ is a bivariate $C_{k\ell}$-Brownian bridge on $[0,1]^2$. Note that $\mathbb{G}_{C_{k\ell}}$ is completely tucked, that is $\mathbb{G}_{C_{k\ell}}$ is pinned down to 0 on the entire boundary of the unit square $[0,1]^2$ (see, e.g., Example A.2.12 in van der Vaart and Wellner (1996)).

5.3 Testing ellipticity

5.3.1 The test statistic and its asymptotic behavior

Under the hypothesis of an elliptical copula C, also all marginal copulas have to be elliptical. The test in Chapter 3 is based on the equality of Kendall's tau $\tau_{C_{k\ell}}$ and Blomqvist's beta $\beta_{C_{k\ell}}$ for all pairs $k, \ell \in \{1, \dots, d\}$ with $k < \ell$. By virtue of this equality (see Equation (3.5)), the test statistic is constructed using all $d(d-1)/2$ pairwise differences $\beta^\star_{k\ell,n} - \tau_{k\ell,n}$, $k, \ell \in \{1, \dots, d\}$ with $k < \ell$, between the empirically estimated Blomqvist's beta $\beta^\star_{k\ell,n}$ and Kendall's tau $\tau_{k\ell,n}$. Since the utilized estimator of β depends on the marginal medians, the test in Chapter 3 requires copula data. In this chapter, we discard the assumption of known marginal distribution functions.

Recall from Equations (2.1) and (2.3) that Kendall's tau and Blomqvist's beta are completely determined by the bivariate marginal copula $C_{k\ell}$ and can be expressed as

$$\tau_{k\ell} = \tau_{C_{k\ell}} = 4\int_0^1\int_0^1 C_{k\ell}(u_k, u_\ell)\, dC_{k\ell}(u_k, u_\ell) - 1$$

and

$$\beta_{k\ell} = \beta_{C_{k\ell}} = 4C_{k\ell}(0.5, 0.5) - 1\,.$$

A suitable non-parametric estimator of $\beta_{k\ell}$ can be obtained by replacing the copula $C_{k\ell}$ in Equation (2.3) with the empirical counterpart given by $\widehat{C}_{k\ell,n}(u_k, u_\ell) = \widehat{C}_n(\boldsymbol{u}^{(k\ell)})$, for all $(u_k, u_\ell) \in [0,1]^2$. This yields

$$\widehat{\beta}_{k\ell,n} = 4\widehat{C}_{k\ell,n}(0.5, 0.5) - 1\,. \tag{5.3}$$

Similarly, a non-parametric estimator of $\tau_{k\ell}$ is given by

$$\widehat{\tau}_{k\ell,n} = 4\int_0^1 \int_0^1 \widehat{C}_{k\ell,n}(u_k, u_\ell)\, d\widehat{C}_{k\ell,n}(u_k, u_l) - 1\,. \tag{5.4}$$

This results in the statistic $\widehat{\boldsymbol{D}}_n$ defined by

$$\widehat{\boldsymbol{D}}_n = \widehat{\boldsymbol{\beta}}_n - \widehat{\boldsymbol{\tau}}_n\,, \tag{5.5}$$

where $\widehat{\boldsymbol{\beta}}_n = \left(\widehat{\beta}_{12,n}, \widehat{\beta}_{13,n}, \ldots, \widehat{\beta}_{d-1\,d,n}\right)$ and $\widehat{\boldsymbol{\tau}}_n = (\widehat{\tau}_{12,n}, \widehat{\tau}_{13,n}, \ldots, \widehat{\tau}_{d-1\,d,n})$.

Note that the estimators $\widehat{\beta}_{k\ell,n}$ and $\widehat{\tau}_{k\ell,n}$ are asymptotically equivalent to the estimators $\beta_{k\ell,n}$ and $\tau_{k\ell,n}$ defined in (2.4) and (2.2), respectively. We get the following lemma.

Lemma 5.3. *Let $\beta_{k\ell,n}$ and $\tau_{k\ell,n}$ as well as $\widehat{\beta}_{k\ell,n}$ and $\widehat{\tau}_{k\ell,n}$ be defined as in* (2.4) *and* (2.2) *as well as* (5.3) *and* (5.4)*, respectively. Then, it holds that*

(i) $\beta_{k\ell,n} = \widehat{\beta}_{k\ell,n} + o_{\mathbb{P}}\left(\frac{1}{\sqrt{n}}\right)$,

(ii) $\tau_{k\ell,n} = \widehat{\tau}_{k\ell,n} + o_{\mathbb{P}}\left(\frac{1}{\sqrt{n}}\right)$.

Proof. The proof of (i) for Blomqvist's beta is a straightforward consequence of Lemma 1 in Genest et al. (2013). The proof of (ii) for Kendall's tau is given in Section 5.6.1. □

Now, let us consider the map

$$\Phi : \mathbb{D}_\Phi \to \mathbb{R}^{\frac{d(d-1)}{2}},\ H \mapsto \begin{pmatrix} 4H_{12}(0.5, 0.5) - 4\int_{[0,1]^2} H_{12}(u_1, u_2) dH_{12}(u_1, u_2) \\ 4H_{13}(0.5, 0.5) - 4\int_{[0,1]^2} H_{13}(u_1, u_2) dH_{13}(u_1, u_2) \\ \cdots \\ 4H_{d-1,d}(0.5, 0.5) - 4\int_{[0,1]^2} H_{d-1,d}(u_1, u_2) dH_{d-1,d}(u_1, u_2) \end{pmatrix},$$

where $\mathbb{D}_\Phi$ is the space of all distribution functions H on $[0,1]^d$ with $H_{k\ell}(u, 0) = H_{k\ell}(0, u) = 0$ for all $u \in [0,1]$ and all $k, \ell \in \{1, \ldots, d\}$ and $k < \ell$. Note that under the null hypothesis,

$$\sqrt{n}\,\widehat{\boldsymbol{D}}_n = \sqrt{n}(\widehat{\boldsymbol{\beta}}_n - \widehat{\boldsymbol{\tau}}_n) = \sqrt{n}\left(\Phi(\widehat{C}_n) - \Phi(C)\right)\,.$$

Thus, in order to derive the weak convergence of $\widehat{\boldsymbol{D}}_n$ for a sample whose dependence structure is captured by an elliptical copula, it is sufficient to show Hadamard-differentiability of Φ tangentially to suitable subspaces and to apply the functional delta method (see Theorem 3.9.4 in van der Vaart and Wellner (1996)).

Theorem 5.4. *Suppose Condition 5.1 holds. Then Φ is Hadamard-differentiable at C tangentially to $\mathbb{D}_0$, the set of continuous functions on $[0,1]^d$. Its derivative at C in $\alpha \in \mathbb{D}_0$ is given by*

$$(\Phi'_C(\alpha))(\boldsymbol{u}) = \begin{pmatrix} 4\alpha_{12}(0.5,0.5) - 4\int_{[0,1]^2} C_{12}\, d\alpha_{12} - 4\int_{[0,1]^2} \alpha_{12}\, dC_{12} \\ 4\alpha_{13}(0.5,0.5) - 4\int_{[0,1]^2} C_{13}\, d\alpha_{13} - 4\int_{[0,1]^2} \alpha_{13}\, dC_{13} \\ \dots \\ 4\alpha_{d-1,d}(0.5,0.5) - 4\int_{[0,1]^2} C_{d-1,d}\, d\alpha_{d-1,d} - 4\int_{[0,1]^2} \alpha_{d-1,d}\, dC_{d-1,d} \end{pmatrix},$$

where $\alpha_{k\ell}(u_k, u_l) = \alpha(\boldsymbol{u}^{(k\ell)})$, for $k, \ell \in \{1,\dots,d\}$ with $k < \ell$.

Proof. Let $\mathbb{E}$ be the space of all distribution functions η on $[0,1]^2$ with $\eta(u,0) = \eta(0,u) = 0$ for all $u \in [0,1]$ and decompose $\Phi = \Phi_2 \circ \Phi_1$, where

$$\Phi_1 : \begin{cases} \mathbb{D}_\Phi \to \mathbb{E}^{\frac{d(d-1)}{2}}, \\ H \mapsto (H_{12}, H_{13}, \dots, H_{d-1,d}), \end{cases}$$

$$\Phi_2 : \begin{cases} \mathbb{E}^{\frac{d(d-1)}{2}} \to \mathbb{R}^{\frac{d(d-1)}{2}}, \\ (H_{12}, H_{13}, \dots, H_{d-1,d}) \mapsto \\ \qquad \begin{pmatrix} 4H_{12}(0.5,0.5) - 4\int_{[0,1]^2} H_{12}(u_1,u_2) dH_{12}(u_1,u_2) \\ 4H_{13}(0.5,0.5) - 4\int_{[0,1]^2} H_{13}(u_1,u_2) dH_{13}(u_1,u_2) \\ \dots \\ 4H_{d-1,d}(0.5,0.5) - 4\int_{[0,1]^2} H_{d-1,d}(u_1,u_2) dH_{d-1,d}(u_1,u_2) \end{pmatrix}. \end{cases}$$

The first map Φ_1 is Hadamard differentiable at C since it is linear and continuous. Its derivative at C is given by $\Phi'_{1,C} = \Phi_1$.

For the second map Φ_2, Hadamard differentiability in every component is shown first. Therefore, consider the mapping

$$\Phi_2^{(k\ell)} : \mathbb{E} \to \mathbb{R}, H_{k\ell} \mapsto 4H_{k\ell}(0.5,0.5) - 4\int_{[0,1]^2} H_{k\ell}(u_1,u_2) dH_{k\ell}(u_1,u_2).$$

With Lemma 1 in Veraverbeke et al. (2011), it easily follows that $\Phi_2^{(k\ell)}$ is Hadamard differentiable at $C_{k\ell}$ tangentially to the set $\mathbb{E}_0$ of functions that are continuous on $[0,1]^2$. The derivative is given by

$$\begin{aligned}\left(\Phi_2^{(k\ell)}\right)'_{C_{k\ell}}(\alpha_{k\ell}) =& 4\alpha_{k\ell}(0.5,0.5) - 4\int_{[0,1]^2} C_{k\ell}(u_1,u_2)\, d\alpha_{k\ell}(u_1,u_2) \\ &- 4\int_{[0,1]^2} \alpha_{k\ell}(u_1,u_2)\, dC_{k\ell}(u_1,u_2),\end{aligned}$$

where the integral $\int C_{k\ell} d\alpha_{k\ell}$ is defined by Equation (5.7) if α is not of bounded variation. For the sake of completeness, we give a more detailed proof of the result of Lemma 1 in Veraverbeke et al. (2011) in the Appendix 5.6.2.

As a consequence, the map Φ_2 is Hadamard differentiable at $(C_{12}, C_{23}, \dots, C_{d-1,d})$ tangentially to $\mathbb{E}_0^{\frac{d(d-1)}{2}}$ and its derivative is given by

$$\Phi'_{2,(C_{12},C_{13},\dots,C_{d-1,d})}(\alpha_{12}, \alpha_{13}, \dots, \alpha_{d-1,d}) = \begin{pmatrix} 4\alpha_{12}(0.5, 0.5) - 4\int_{[0,1]^2} C_{12}\, d\alpha_{12} - 4\int_{[0,1]^2} \alpha_{12}\, dC_{12} \\ 4\alpha_{13}(0.5, 0.5) - 4\int_{[0,1]^2} C_{13}\, d\alpha_{13} - 4\int_{[0,1]^2} \alpha_{13}\, dC_{13} \\ \dots \\ 4\alpha_{d-1,d}(0.5, 0.5) - 4\int_{[0,1]^2} C_{d-1,d}\, d\alpha_{d-1,d} - 4\int_{[0,1]^2} \alpha_{d-1,d}\, dC_{d-1,d} \end{pmatrix}.$$

Finally, applying the chain rule (see Lemma 3.9.3 in van der Vaart and Wellner (1996)) to $\Phi = \Phi_2 \circ \Phi_1$ yields the Hadamard differentiability of Φ at C tangentially to the set $\mathbb{D}_0$. Its derivative is then given by $\Phi'_C = \Phi'_{2,\Phi_1(C)} \circ \Phi'_{1,C}$. □

The following weak convergence result for $\widehat{\boldsymbol{D}}_n$ now follows directly with the functional delta method.

Theorem 5.5. *Let $\mathbf{X}_1, \dots, \mathbf{X}_n \in \mathbb{R}^d$ be i.i.d. d-dimensional random vectors with common cumulative distribution function F, continuous univariate marginal distribution functions $F_1, \dots, F_d$ and elliptical copula C. Under the assumption that the copula C satisfies the first order property of Condition 5.1, it follows that*

$$\sqrt{n}\, \widehat{\boldsymbol{D}}_n \rightsquigarrow N(0, \Sigma),$$

where $\Sigma_{pq} = \mathbb{E}\left[(\Phi'_C(\mathbb{G}_C))_p \cdot (\Phi'_C(\mathbb{G}_C))_q\right]$, for $p, q \in \{1, \dots, \frac{d(d-1)}{2}\}$. The map Φ'_C and the Gaussian process $\mathbb{G}_C$ are defined in (5.6) and (5.1), respectively.

Proof. Given the weak convergence of the empirical copula process $\widehat{\mathbb{C}}_n$ from (5.2) together with the Hadamard differentiability of Φ at C tangentially to the set $\mathbb{D}_0$ from Theorem 5.4, the Delta method (see Theorem 3.9.4 in van der Vaart and Wellner (1996)) immediately yields

$$\sqrt{n}\, \widehat{\boldsymbol{D}}_n = \sqrt{n}\left(\Phi(\widehat{C}_n) - \Phi(C)\right) \rightsquigarrow \Phi'_C(\mathbb{G}_C).$$

With the definition of the integral $\int C_{k\ell}\, d\alpha_{k\ell}$ given in Equation (5.7) and the characteristics of the process $\mathbb{G}_{C_{k\ell}}$ from (5.1), it finally follows that

$$\Phi'_C(\mathbb{G}_C) = \begin{pmatrix} 4\mathbb{G}_{C_{12}}(0.5, 0.5) - 8\int_{[0,1]^2} \mathbb{G}_{C_{12}}(u_1, u_2)\, dC_{12}(u_1, u_2) \\ 4\mathbb{G}_{C_{13}}(0.5, 0.5) - 8\int_{[0,1]^2} \mathbb{G}_{C_{13}}(u_1, u_2)\, dC_{13}(u_1, u_2) \\ \dots \\ 4\mathbb{G}_{C_{d-1,d}}(0.5, 0.5) - 8\int_{[0,1]^2} \mathbb{G}_{C_{d-1,d}}(u_1, u_2)\, dC_{d-1,d}(u_1, u_2) \end{pmatrix}. \tag{5.6}$$

This proves the stated weak convergence result for the statistic $\widehat{\boldsymbol{D}}_n$. □

The limiting Gaussian field $\Phi'_C(\mathbb{G}_C)$ of the empirical copula process and, hence, the covariance matrix Σ_{Φ_C} depend on the unknown copula C. Furthermore, a Wald-type test statistic involving the covariance matrix causes the corresponding test problems to hold its nominal level in higher dimensions. To construct an asymptotic test for the hypothesis of ellipticity, we propose the L_2-type statistic

$$\widehat{T}_n := n \cdot \widehat{\boldsymbol{D}}_n^\top \widehat{\boldsymbol{D}}_n \,.$$

Theorem 5.5 and the Continuous Mapping Theorem (see Theorem 1.3.6 in van der Vaart and Wellner (1996)) yield

$$\widehat{T}_n \rightsquigarrow \Phi'_C(\mathbb{G}_C)^\top \Phi'_C(\mathbb{G}_C) \,.$$

5.3.2 A subsampling approximation

The limiting distribution of the test statistic as well as the limiting covariance matrix depend on the unknown copula C. Therefore, we make use of the subsampling procedure from Kojadinovic and Stemikovskaya (2019) to perform the test. For this, subsamples of size $b < n$ are drawn by sampling without replacement from $\boldsymbol{X}_1, \dots, \boldsymbol{X}_n$. The number of all possible subsamples is then given by $N_{b,n} = \binom{n}{b}$ and the subsamples are denoted by

$$\mathcal{X}_b^{[m]} = (\boldsymbol{X}_1^{[m]}, \dots, \boldsymbol{X}_b^{[m]}), \quad m \in \{1, \dots, N_{b,n}\}.$$

Let $M \in \mathbb{N}$ denote the number of bootstrap replications. The following algorithm defines the test:

1. Compute pseudo-observations $\widehat{\boldsymbol{U}}_1, \dots, \widehat{\boldsymbol{U}}_n$.
2. Compute the statistic $\widehat{T}_n$ from the pseudo-observations.
3. For $m = 1, \dots, M$:
   - Randomly select a subsample $\mathcal{X}_b^{[I_{m,n}]} = (\boldsymbol{X}_1^{[I_{m,n}]}, \dots, \boldsymbol{X}_b^{[I_{m,n}]})$ of size b by choosing $I_{m,n}$ with replacement from $\{1, \dots, N_{b,n}\}$.
   - Compute pseudo-observations $\widehat{\boldsymbol{U}}_1^{[I_{m,n}]}, \dots, \widehat{\boldsymbol{U}}_b^{[I_{m,n}]}$ of the subsample.
   - Compute the statistic $\widehat{\boldsymbol{\beta}}_b^{[I_{m,n}]}$ and $\widehat{\boldsymbol{\tau}}_b^{[I_{m,n}]}$ from the pseudo-observations of the subsample.
   - Compute the bootstrap statistic

   $$\widehat{T}_{b,n}^{[I_{m,n}]} = b \cdot \left(\widehat{\boldsymbol{\beta}}_b^{[I_{m,n}]} - \widehat{\boldsymbol{\tau}}_b^{[I_{m,n}]} - (\widehat{\boldsymbol{\beta}}_n - \widehat{\boldsymbol{\tau}}_n)\right)^\top \left(\widehat{\boldsymbol{\beta}}_b^{[I_{m,n}]} - \widehat{\boldsymbol{\tau}}_b^{[I_{m,n}]} - (\widehat{\boldsymbol{\beta}}_n - \widehat{\boldsymbol{\tau}}_n)\right)$$

4. An approximate p-value for the test based on $\widehat{T}_n$ is then given by

$$\widehat{p}_{M,b,n} = \frac{1}{M} \sum_{m=1}^{M} I\{\widehat{T}_{b,n}^{[I_{m,n}]} > \widehat{T}_n\} \,.$$

Kojadinovic and Stemikovskaya (2019) recommend to generally multiply the subsample replicates by the so-called finite population correction, which is given by $(1 - b/n)^{-1/2}$. For our choice of b in the following simulation study (see Section 5.4), it holds that $b/n \to 0$. It follows that $(1 - b/n)^{-1/2}$ tends to 1. Hence, the weak convergence result from Kojadinovic and Stemikovskaya (2019) holds without the correction factor and we compute our bootstrap statistic $\widehat{T}_{b,n}^{[I_{m,n}]}$ without it.

5.4 Simulation study

In order to assess the finite-sample performance of the proposed test for ellipticity based on the L_2-type statistic $\widehat{T}_n$ in combination with the subsampling approximation described in Section 5.3.2, in the following shortly L_2S test, a Monte Carlo study was conducted.

In Chapter 3, it was shown that the non-parametric estimation of unknown marginal distribution functions does not affect the finite sample performance of the simple test based on the Wald-type statistic T_n, in the following shortly simple test. We use the following simulation study to compare the results for the L_2S test to the results for the simple test applied to pseudo-observations.

Furthermore, we compare the L_2S test to the competing one from Quessy (2020) based on the similar L_2-type test statistic $S_n^{L_2}$; in the following shortly Q test. In contrast to our subsampling approximation, in Quessy (2020) a multiplier bootstrap method is utilized.

5.4.1 Setup

The simulation study was carried out for different copula families C, sample sizes n, dimensions d, and levels of dependence in terms of Kendall's tau τ. To investigate the empirical level of the test, random samples from the Gaussian copula and the t copula with 5 degrees of freedom ($t_{\nu=5}$) were considered. Random samples from the Frank and the Clayton copula were examined, to study the empirical power of the test. In Chapter 3, also samples of the Gumbel copula were included into the power study. However, the rejection rates for the Gumbel family were very low. Closer analysis showed that Kendall's tau and Blomqvist's beta are very close for the Gumbel family and huge sample sizes are needed for the test to provide some indication against the null hypothesis. Since the test statistic $\widehat{T}_n$ is also based on the equality of Kendall's tau and Blomqvist's beta, the rejection rates for samples from the Gumbel copula stayed quite low and the results are not included in the following power study.

To assess the effect of sample size, dimension and strength of dependence, the values of n, d, and τ are chosen to vary in the sets $\{100, 250, 500, 1000\}$, $\{2, 3, 6\}$, and $\{0.1, 0.25, 0.5, 0.75, 0.9\}$, respectively. Finally, the number of Monte replications was set to $N = 1000$, the number of subsampling replications to $M = 300$, and all tests were performed at a significance level of $\alpha = 0.05$.

Following the approach in Kojadinovic and Stemikovskaya (2019), we also chose the size of the subsamples b depending on the sample size n. For samples of size $n \leq 200$, the best results in Kojadinovic and Stemikovskaya (2019) were derived for $b \in \{\lfloor 0.1 \rfloor n, \lfloor 0.28 \rfloor n\}$. Our simulations for larger samples showed that with increasing sample size, a smaller percentage of n is needed for our test to hold its nominal level across different dimensions. Based on the results in Kojadinovic and Stemikovskaya (2019) as well as our own simulations, we suggest to choose the subsample size $b = \lfloor 2.4 \cdot n^{0.54} \rfloor$, where $\lfloor x \rfloor$, with $x \in \mathbb{R}$, denotes the integer part of x. This leads to subsample sizes $b \in \{28, 47, 68, 100\}$ for the considered sample sizes $n \in \{100, 250, 500, 1000\}$.

5.4.2 Finite-sample performance

In this section, the finite-sample performance is analyzed. Tables 5.1 to 5.6 report the empirical level and the empirical power of our L_2S test based on the statistic $\widehat{T}_n$ and the subsampling approach (in Column L_2S), the simple test for ellipticity from Chapter 3 for pseudo-observations (in Column S), and the Q test from Quessy (2020) based on a multiplier bootstrap (in Column Q) for dimension $d = 2, 3$ and 6, respectively.

C	$n = 100$			$n = 250$			$n = 500$			$n = 1000$		
τ	L_2S	S	Q	L_2S	S	Q	L_2S	S	Q	L_2S	S	Q
Gauss												
0.10	0.048	0.063	0.061	0.037	0.041	0.048	0.049	0.050	0.054	0.051	0.046	0.057
0.25	0.048	0.059	0.057	0.040	0.060	0.062	0.049	0.046	0.049	0.059	0.050	0.052
0.50	0.035	0.053	0.051	0.044	0.058	0.061	0.056	0.061	0.065	0.047	0.041	0.050
0.75	0.020	0.051	0.056	0.026	0.040	0.043	0.028	0.040	0.041	0.046	0.052	0.054
0.90	0.007	0.060	0.079	0.020	0.060	0.067	0.034	0.057	0.061	0.030	0.045	0.053
$t_{\nu=5}$												
0.10	0.046	0.058	0.062	0.041	0.052	0.049	0.046	0.045	0.045	0.057	0.052	0.056
0.25	0.039	0.046	0.056	0.036	0.041	0.046	0.045	0.043	0.043	0.060	0.053	0.059
0.50	0.027	0.052	0.054	0.035	0.061	0.062	0.055	0.049	0.055	0.062	0.062	0.062
0.75	0.025	0.066	0.069	0.028	0.051	0.061	0.040	0.045	0.050	0.052	0.066	0.069
0.90	0.013	0.087	0.098	0.014	0.048	0.056	0.029	0.048	0.055	0.044	0.064	0.065

Table 5.1: Dimension d=2: Empirical level of our test for ellipticity based on the L_2-type statistic $\widehat{T}_n$ and the subsampling approach (L_2S), the simple test for ellipticity from Chapter 3 for pseudo-observations (S), and the test from Quessy (2020) based on the statistic S^{L_2} and a multiplier bootstrap (Q) with significance level $\alpha = 0.05$: rate of rejecting H_0 as observed in 1000 random samples of size n from copula family C with Kendall's tau τ.

C	$n = 100$			$n = 250$			$n = 500$			$n = 1000$		
τ	L_2S	S	Q	L_2S	S	Q	L_2S	S	Q	L_2S	S	Q
Gauss												
0.10	0.039	0.061	0.045	0.034	0.062	0.054	0.058	0.055	0.054	0.060	0.054	0.058
0.25	0.026	0.063	0.039	0.025	0.045	0.046	0.043	0.042	0.041	0.066	0.064	0.063
0.50	0.029	0.080	0.042	0.030	0.070	0.057	0.048	0.051	0.043	0.044	0.048	0.047
0.75	0.027	0.101	0.056	0.031	0.083	0.060	0.046	0.066	0.049	0.056	0.063	0.062
0.90	0.010	0.210	0.051	0.012	0.105	0.050	0.027	0.083	0.048	0.033	0.062	0.051
$t_{\nu=5}$												
0.10	0.032	0.061	0.041	0.034	0.067	0.062	0.052	0.051	0.051	0.047	0.046	0.052
0.25	0.040	0.061	0.047	0.034	0.060	0.050	0.052	0.051	0.050	0.066	0.063	0.064
0.50	0.030	0.061	0.039	0.029	0.058	0.058	0.042	0.048	0.049	0.044	0.052	0.045
0.75	0.025	0.093	0.054	0.024	0.061	0.046	0.043	0.066	0.048	0.047	0.055	0.047
0.90	0.012	0.201	0.047	0.009	0.115	0.042	0.025	0.091	0.059	0.030	0.067	0.053

Table 5.2: Dimension d=3: Empirical level of our test for ellipticity based on the L_2-type statistic $\widehat{T}_n$ and the subsampling approach (L_2S), the simple test for ellipticity from Chapter 3 for pseudo-observations (S), and the test from Quessy (2020) based on the statistic S^{L_2} and a multiplier bootstrap (Q) with significance level $\alpha = 0.05$: rate of rejecting H_0 as observed in 1000 random samples of size n from copula family C with Kendall's tau τ.

To study the empirical level of the tests, random samples from the Gaussian and the $t_{\nu=5}$ copula were considered. Tables 5.1, 5.2, and 5.3 present the empirical level for dimension $d = 2, 3$ and 6, respectively. Our L_2S test and the Q test seem to hold their nominal level across the two copula models, all sample sizes, and all values of Kendall's tau for the considered dimensions. Only for small sample sizes in combination with large values of Kendall's tau, our L_2S test turns out to be slightly too conservative. As already discussed in Section 3.5, the simple test requires large sample sizes to hold its level in higher dimensions (see also Tables 5.2 and 5.3).

To assess the empirical power, random samples of the Frank and the Clayton copula were used. The corresponding empirical power results for dimension $d = 2, 3$ and 6 are stated in Tables 5.4, 5.5, and 5.6, respectively. To take into account that the simple test has problems to keep its nominal level in higher dimensions, the empirical power results corresponding to these cases are displayed in gray. First of all, note that the rejection rates vary clearly across the different combinations of factors. As expected, the empirical power generally increases with increasing sample size. In addition, the empirical power also increases with increasing level of dependence, at least up to $\tau = 0.75$. This is reasonable, since the deviation from the null hypothesis of ellipticity becomes more apparent with increasing values of Kendall's tau (see Figure 3.2). For extremely high levels of dependence as for $\tau = 0.9$, however, it might be harder to distinguish the non-ellipticity of the sample. To illustrate this scenario, Figure 5.1

C	$n=100$			$n=250$			$n=500$			$n=1000$		
τ	L_2S	S	Q	L_2S	S	Q	L_2S	S	Q	L_2S	S	Q
Gauss												
0.10	0.032	0.198	0.018	0.011	0.093	0.035	0.060	0.069	0.035	0.053	0.054	0.040
0.25	0.043	0.228	0.033	0.009	0.084	0.024	0.047	0.075	0.038	0.054	0.062	0.050
0.50	0.044	0.283	0.034	0.021	0.132	0.032	0.056	0.080	0.042	0.066	0.088	0.061
0.75	0.037	0.567	0.040	0.017	0.256	0.037	0.043	0.132	0.039	0.051	0.096	0.052
0.90	0.011	0.933	0.036	0.006	0.681	0.045	0.022	0.385	0.042	0.026	0.195	0.047
$t_{\nu=5}$												
0.10	0.044	0.209	0.025	0.012	0.098	0.040	0.061	0.071	0.043	0.062	0.064	0.048
0.25	0.040	0.223	0.029	0.022	0.107	0.040	0.071	0.071	0.053	0.063	0.057	0.047
0.50	0.032	0.293	0.024	0.021	0.142	0.042	0.054	0.091	0.046	0.066	0.081	0.051
0.75	0.037	0.601	0.041	0.018	0.277	0.048	0.042	0.143	0.042	0.036	0.068	0.035
0.90	0.016	0.919	0.038	0.007	0.666	0.040	0.028	0.379	0.038	0.031	0.193	0.051

Table 5.3: Dimension d=6: Empirical level of our test for ellipticity based on the L_2-type statistic $\widehat{T}_n$ and the subsampling approach (L_2S), the simple test for ellipticity from Chapter 3 for pseudo-observations (S), and the test from Quessy (2020) based on the statistic S^{L_2} and a multiplier bootstrap (Q) with significance level $\alpha = 0.05$: rate of rejecting H_0 as observed in 1000 random samples of size n from copula family C with Kendall's tau τ.

presents scatter plots of bivariate random samples of size $n = 500$ from the Gaussian, t_{ν_5}, Frank, and Clayton copula for the very high level of dependence corresponding to $\tau = 0.9$. Furthermore, the empirical power increases with the dimension. Since the L_2S test and the Q test, in contrast to the simple test, also hold their nominal level in higher dimensions, this is an especially appealing feature of the L_2S test and the Q test when applied to high-dimensional data.

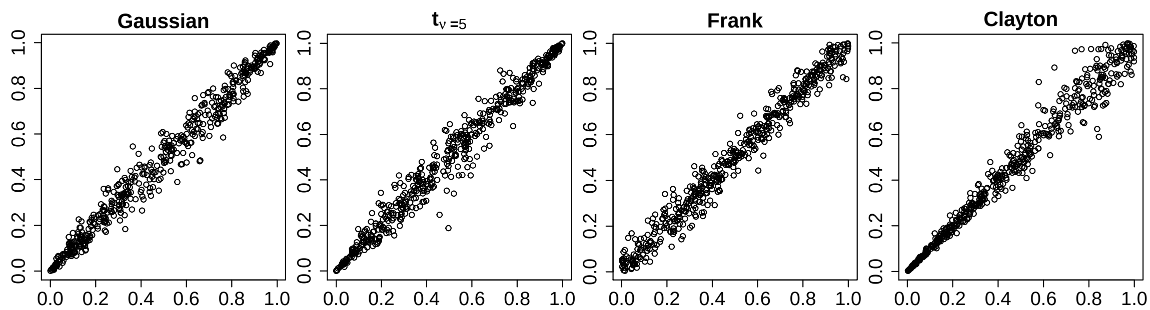

Figure 5.1: Scatter plots of random samples of size 500 from the bivariate Gaussian, $t_{\nu=5}$, Frank, and Clayton copula (from left to right) with $\tau = 0.9$.

C	$n=100$			$n=250$			$n=500$			$n=1000$		
τ	L_2S	S	Q	L_2S	S	Q	L_2S	S	Q	L_2S	S	Q
Frank												
0.10	0.039	0.050	0.051	0.052	0.054	0.055	0.063	0.065	0.063	0.101	0.094	0.102
0.25	0.042	0.057	0.063	0.139	0.135	0.145	0.213	0.197	0.204	0.324	0.310	0.321
0.50	0.099	0.123	0.122	0.262	0.264	0.263	0.496	0.481	0.483	0.754	0.730	0.733
0.75	0.100	0.176	0.179	0.290	0.316	0.318	0.593	0.587	0.594	0.863	0.853	0.852
0.90	0.020	0.174	0.188	0.185	0.228	0.236	0.298	0.393	0.406	0.589	0.627	0.622
Clayton												
0.10	0.032	0.053	0.053	0.038	0.054	0.058	0.066	0.063	0.065	0.063	0.060	0.063
0.25	0.043	0.053	0.057	0.036	0.052	0.053	0.054	0.051	0.050	0.053	0.051	0.056
0.50	0.044	0.061	0.064	0.052	0.056	0.048	0.073	0.055	0.064	0.105	0.085	0.095
0.75	0.039	0.091	0.099	0.132	0.124	0.126	0.255	0.246	0.253	0.476	0.436	0.444
0.90	0.018	0.131	0.138	0.128	0.199	0.198	0.233	0.308	0.307	0.483	0.499	0.515

Table 5.4: Dimension d=2: Empirical power of our test for ellipticity based on the L_2-type statistic $\widehat{T}_n$ and the subsampling approach (L_2S), the simple test for ellipticity from Chapter 3 for pseudo-observations (S), and the test from Quessy (2020) based on the statistic S^{L_2} and a multiplier bootstrap (Q) with significance level $\alpha = 0.05$: rate of rejecting H_0 as observed in 1000 random samples of size n from copula family C with Kendall's tau τ.

C	$n=100$			$n=250$			$n=500$			$n=1000$		
τ	L_2S	S	Q	L_2S	S	Q	L_2S	S	Q	L_2S	S	Q
Frank												
0.10	0.039	0.078	0.052	0.050	0.075	0.061	0.097	0.091	0.090	0.147	0.124	0.138
0.25	0.059	0.090	0.069	0.147	0.144	0.156	0.279	0.226	0.258	0.539	0.448	0.519
0.50	0.159	0.190	0.169	0.360	0.338	0.391	0.728	0.606	0.691	0.949	0.886	0.950
0.75	0.116	0.271	0.159	0.380	0.396	0.421	0.776	0.661	0.768	0.960	0.910	0.958
0.90	0.010	0.385	0.159	0.152	0.364	0.280	0.416	0.479	0.489	0.733	0.671	0.734
Clayton												
0.10	0.033	0.058	0.035	0.026	0.049	0.052	0.070	0.071	0.064	0.062	0.060	0.065
0.25	0.038	0.064	0.046	0.027	0.049	0.047	0.052	0.056	0.051	0.050	0.044	0.046
0.50	0.034	0.071	0.045	0.058	0.053	0.051	0.088	0.070	0.075	0.120	0.079	0.097
0.75	0.048	0.165	0.089	0.199	0.188	0.176	0.386	0.270	0.333	0.665	0.491	0.614
0.90	0.013	0.349	0.123	0.134	0.292	0.209	0.303	0.363	0.371	0.658	0.561	0.665

Table 5.5: Dimension d=3: Empirical power of our test for ellipticity based on the L_2-type statistic $\widehat{T}_n$ and the subsampling approach (L_2S), the simple test for ellipticity from Chapter 3 for pseudo-observations (S), and the test from Quessy (2020) based on the statistic S^{L_2} and a multiplier bootstrap (Q) with significance level $\alpha = 0.05$: rate of rejecting H_0 as observed in 1000 random samples of size n from copula family C with Kendall's tau τ.

C	$n = 100$			$n = 250$			$n = 500$			$n = 1000$		
τ	L_2S	S	Q	L_2S	S	Q	L_2S	S	Q	L_2S	S	Q
Frank												
0.10	0.042	0.211	0.027	0.045	0.129	0.075	0.128	0.118	0.099	0.272	0.197	0.229
0.25	0.115	0.321	0.082	0.211	0.301	0.251	0.624	0.409	0.571	0.910	0.717	0.894
0.50	0.254	0.579	0.215	0.563	0.573	0.634	0.954	0.754	0.934	0.999	0.973	1.000
0.75	0.186	0.851	0.214	0.529	0.738	0.592	0.914	0.813	0.911	0.996	0.959	0.996
0.90	0.009	0.992	0.125	0.181	0.945	0.310	0.561	0.879	0.633	0.909	0.893	0.917
Clayton												
0.10	0.032	0.191	0.015	0.022	0.106	0.044	0.069	0.072	0.048	0.092	0.071	0.067
0.25	0.048	0.210	0.026	0.017	0.085	0.027	0.052	0.071	0.033	0.077	0.063	0.053
0.50	0.032	0.302	0.016	0.056	0.138	0.047	0.126	0.106	0.077	0.228	0.104	0.135
0.75	0.077	0.706	0.082	0.278	0.444	0.218	0.560	0.405	0.450	0.882	0.546	0.817
0.90	0.014	0.970	0.110	0.153	0.876	0.239	0.425	0.787	0.477	0.818	0.773	0.802

Table 5.6: Dimension d=6: Empirical power of our test for ellipticity based on the L_2-type statistic $\widehat{T}_n$ and the subsampling approach (L_2S), the simple test for ellipticity from Chapter 3 for pseudo-observations (S), and the test from Quessy (2020) based on the statistic S^{L_2} and a multiplier bootstrap (Q) with significance level $\alpha = 0.05$: rate of rejecting H_0 as observed in 1000 random samples of size n from copula family C with Kendall's tau τ.

In general, all three tests seem to be good in detecting deviations from the null hypothesis of ellipticity if the level of dependence is not too close to independence. The L_2S test shows an inferior performance compared to the other two tests for the very high level of dependence given by $\tau = 0.9$. However, the disadvantage gets smaller with increasing sample size and increasing dimension resulting in the L_2S test being as powerful as the Q test as of $n = 1000$ in dimensions $d = 3$ and 6 and even outperforming the simple test in the latter scenarios. Having a closer look on the power properties in dimension $d = 2$ (see Table 5.4), similar rejection rates can be observed for all three tests. In higher dimensions (see Tables 5.5 and 5.6), the simple test is less efficient in detecting the lack of ellipticity compared to the other two tests, with some rare exceptions as for example for $n = 100$ in $d = 3$. Comparing the performance of the L_2S and the Q test in dimensions $d = 3$ and 6, mixed results can be observed for the samples from the Frank copula. For the sample size $n = 100$ in combination with lower dependence, the L_2S test performs better, whereas for higher dependence the Q test is better. For the sample size $n = 250$, the Q test is more powerful than the L_2S test. However, with increasing sample size the L_2S outperforms the Q test and the superiority becomes larger with increasing dimension. For samples of the Clayton copula in dimensions $d = 3$ and 6, the Clayton copula is systematically more powerful than the Q test and the advantage increases again with the dimension.

Overall, the finite-sample performance of the simple test in dimension $d = 2$ and the L_2S test as well as the Q test in all dimensions is very convincing. They hold their nominal level and are reasonably powerful in detecting deviations from the null hypothesis of ellipticity. Although the results are not completely unambiguous, the L_2S test is not only a solid competitor to the Q test but might be even preferable. Note that further optimization of the choice of the subsample size b might also leave space for improvement of the L_2S test. Furthermore, note that the bootstrap statistic in the subsampling approximation (see Section 5.3.2) is computed from subsamples having a size of only 10 to at most 30 percent of the size of the original sample. Thus, the L_2S test is also computationally very attractive.

5.5 Conclusion

In this chapter, the key idea of the simple non-parametric goodness-of-fit test for elliptical copulas from Chapter 3 is further developed to get an advanced statistical test that resolves the two stated restrictions of the simple test. The first restriction is the fact that the simple test assumes given copula data. Taking the estimation of unknown marginal distribution functions into account, a weak convergence result is derived for the statistic containing the differences between Kendall's tau and Blomqvist's beta for all bivariate margins. The second restriction is the fact that the simple test has problems to keep its level in higher dimensions. This is due to slow convergence rates resulting from the estimation of the covariance matrix, which gets large in higher dimensions. We propose a L_2-type test statistic, which does not suffer from these restrictions since there is no covariance matrix involved. The limiting distribution of the test statistic depends on the unknown underlying copula. Thus, in order to perform the test, we apply the subsampling procedure from Kojadinovic and Stemikovskaya (2019).

Note that the objective of this chapter also gets adressed in Quessy (2020). However, the results in this chapter have been developed independently. We present a different proof for the main result and, with the subsampling approach, we provide an alternative bootstrap procedure.

In the simulation study, we compare the finite sample performance of the advanced test based on the L_2 test statistic and a subsampling approximation to the one of the competing test from Quessy (2020) based on a similar statistic and a multiplier bootstrap. Furthermore, the empirical results of the simple test from Chapter 3 are included into the analysis. Note that the simple test can also be recommended in the case of unknown marginal distributions since their estimation does not affect the finite sample performance of the test (see Chapter 3). The contradicting statement in Quessy (2020) is based on a mistake in his Matlab code, which he generously provided us. Thus, our empirical findings and the deduced recommendations are correct. The simulations illustrate that the advanced tests keep their nominal level in higher dimensions already beginning from the small sample size of $n = 100$. At the same time, the empirical power of the advanced tests increases with the dimension and they generally outperform the simple test for dimensions larger than 2. Overall, the advanced test developed in this chapter is able to compete with the test from Quessy (2020) and it even narrowly beats it in many scenarios.

The derivations in this chapter are based on empirical process theory for i.i.d. data. Using the results in Bücher and Volgushev (2013), the results can easily be extended to strictly stationary time series. The subsampling approximation from Kojadinovic and Stemikovskaya (2019) can also be applied in this scenario. However, note that the choice of the subsample size b in the time series case is still subject of current research.

5.6 Auxiliary results

This section covers the proof of Lemma 5.3 and a revision of the Hadamard differentiability of Kendall's tau.

5.6.1 Proof of Lemma 5.3(ii)

Let $\tau_{k\ell,n}$ and $\widehat{\tau}_{k\ell,n}$ be defined as in (2.2) and (5.4), respectively. The total number of different pairs of observations in the sample of size n is given by $\binom{n}{2} = \frac{n(n-1)}{2}$ and is equal to the number of concordant pairs plus the number of discordant pairs (given that the sample does not contain any ties, which is the case here). This yields,

$$\begin{aligned}
\tau_{k\ell,n} &= \frac{2}{n(n-1)} \sum_{1\le i<j\le n} \operatorname{sgn}(X_{ki}-X_{kj})\operatorname{sgn}(X_{li}-X_{lj}) \\
&= \frac{2}{n(n-1)} \sum_{1\le i<j\le n} \Big(I\{(X_{ki}-X_{kj})(X_{li}-X_{lj})>0\} \\
&\qquad\qquad - I\{(X_{ki}-X_{kj})(X_{li}-X_{lj})<0\}\Big) \\
&= \frac{2}{n(n-1)} \left[2\left(\sum_{1\le i<j\le n} I\{(X_{ki}-X_{kj})(X_{li}-X_{lj})>0\}\right) - \binom{n}{2}\right] \\
&= \frac{4}{n(n-1)} \left(\sum_{1\le i<j\le n} I\{(X_{ki}-X_{kj})(X_{li}-X_{lj})>0\}\right) - 1 \\
&= \frac{4}{n(n-1)} \sum_{i,j=1}^{n} I\{X_{ki}>X_{kj}, X_{li}>X_{lj}\} - 1
\end{aligned}$$

Now, using the simple equality

$$\frac{4}{n(n-1)} = \frac{4}{n^2} + \frac{4}{n^2(n-1)}$$

as well as the definition of pseudo-observations in Equation (2.12) and the definition of the empirical copula from Equation (2.13), we further get

$$\begin{aligned}
\tau_{k\ell,n} &= \frac{4}{n^2}\sum_{i,j=1}^{n} I\{X_{ki} > X_{kj}, X_{li} > X_{lj}\} - 1 + o_{\mathbb{P}}\left(\frac{1}{\sqrt{n}}\right) \\
&= \frac{4}{n^2}\left[\left(\sum_{i,j=1}^{n} I\{X_{ki} \geq X_{kj}, X_{li} \geq X_{lj}\}\right) - n\right] - 1 + o_{\mathbb{P}}\left(\frac{1}{\sqrt{n}}\right) \\
&= \frac{4}{n^2}\sum_{i,j=1}^{n} I\{X_{ki} \geq X_{kj}, X_{li} \geq X_{lj}\} - 1 + o_{\mathbb{P}}\left(\frac{1}{\sqrt{n}}\right) \\
&= \frac{4}{n^2}\sum_{i,j=1}^{n} I\{X_{kj} \leq X_{ki}, X_{lj} \leq X_{li}\} - 1 + o_{\mathbb{P}}\left(\frac{1}{\sqrt{n}}\right) \\
&= \frac{4}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} I\{\widehat{U}_{kj} \leq \widehat{U}_{ki}, \widehat{U}_{lj} \leq \widehat{U}_{li}\} - 1 + o_{\mathbb{P}}\left(\frac{1}{\sqrt{n}}\right) \\
&= 4 \cdot \frac{1}{n}\sum_{i=1}^{n} \widehat{C}_{kl,n}(\widehat{U}_{ki}, \widehat{U}_{li}\} - 1 + o_{\mathbb{P}}\left(\frac{1}{\sqrt{n}}\right) \\
&= 4\int_{[0,1]^2} \widehat{C}_{kl,n}(u_k, u_l)\, d\widehat{C}_{kl,n}(u_k, u_l) - 1 + o_{\mathbb{P}}\left(\frac{1}{\sqrt{n}}\right) .
\end{aligned}$$

5.6.2 On the Hadamard differentiability of Kendall's tau

In this section, we review Lemma 1 in Veraverbeke et al. (2011) on the Hadamard differentiability of Kendall's tau and its proof. Therein, a definition of the integral $\int C\, d\alpha$ is needed if α is not of bounded variation. First, note that there exist several definitions of bounded variation for multivariate functions (see, e.g., Clarkson and Adams (1933) for the bivariate case or Owen (2005) for more details). Veraverbeke et al. (2011) do not state which kind of variation they consider. In the sequel, we choose one of these definitions in a way such that a valid bivariate integration by parts formula is obtained. Multivariate integration by parts formulas have been proven recently by e.g. Berghaus et al. (2017) and Radulović et al. (2017). Veraverbeke et al. (2011) do not give any proof or reference for their integration by parts formula. We will use the results from Berghaus et al. (2017) in the sequel.

Let $f(\boldsymbol{x})$ be a real-valued function on $[0,1]^2$. Consider the set of two partitions $u_1^{(0)}$, $u_1^{(1)}$, …, $u_1^{(m(1))}$ and $u_2^{(0)}$, $u_2^{(1)}$, …, $u_2^{(m(2))}$ of $[0,1]$ with

$$0 = u_1^{(0)} < u_1^{(1)} < \ldots < u_1^{(m(1))} = 1, \quad \text{and} \quad 0 = u_2^{(0)} < u_2^{(1)} < \ldots < u_2^{(m(2))} = 1 .$$

For $u \in [0,1]$, define the operators

$$\Delta_1 f(u_1^{(i)}, u) = f(u_1^{(i+1)}, u) - f(u_1^{(i)}, u) .$$

and

$$\Delta_2 f(u, u_2^{(j)}) = f(u, u_2^{(j+1)}) - f(u, u_2^{(j)}) \, .$$

Further, Δ_{12} stands for $\Delta_1\Delta_2$ and, hence,

$$\begin{aligned}\Delta_{12} f(u_1^{(i)}, u_2^{(j)}) &= \Delta_1(\Delta_2 f(u_1^{(i)}, u_2^{(j)})) = \Delta_1(f(u_1^{(i)}, u_2^{(j+1)}) - f(u_1^{(i)}, u_2^{(j)})) \\ &= f(u_1^{(i+1)}, u_2^{(j+1)}) - f(u_1^{(i+1)}, u_2^{(j)}) - f(u_1^{(i)}, u_2^{(j+1)}) + f(u_1^{(i)}, u_2^{(j)}) \, .\end{aligned}$$

Given these operators, the variation in the sense of Hardy and Krause (HK-variation) is defined as follows.

Definition 5.6. *(HK-variation in 2 dimensions)*
Let $f : [0,1]^2 \mapsto \mathbb{R}$. The HK-variation of f on the unit hypercube $[0,1]^2$ anchored at $(1,1)$ is given by

$$\begin{aligned}V_{HK}(f) = &\sup \sum_{i=0}^{m(1)-1} \sum_{j=0}^{m(2)-1} |\Delta_{12} f(u_1^{(i)}, u_2^{(j)})| + \sup \sum_{i=0}^{m(1)-1} |\Delta_1 f(u_1^{(i)}, 1)| \\ &+ \sup \sum_{j=0}^{m(2)-1} |\Delta_2 f(1, u_2^{(j)})| \, ,\end{aligned}$$

where the supremum is taken over the corresponding set of partitions $u_1^{(0)}, u_1^{(1)}, \ldots, u_1^{(m(1))}$ and $u_2^{(0)}, u_2^{(1)}, \ldots, u_2^{(m(2))}$ with

$$0 = u_1^{(0)} < u_1^{(1)} < \ldots < u_1^{(m(1))} = 1 \, ,$$

and

$$0 = u_2^{(0)} < u_2^{(1)} < \ldots < u_2^{(m(2))} = 1 \, ,$$

respectively.
If $V_{HK}(f) < \infty$, f is said to be of bounded HK-variation.

For bivariate distribution functions on the unit square, we can prove the following result.

Lemma 5.7. *Let $H : [0,1]^2 \mapsto \mathbb{R}$ be a distribution function on $[0,1]^2$. Then, it holds that*

$$V_{HK}(H) = 3 \, .$$

Proof. Let $H : [0,1]^2 \mapsto \mathbb{R}$ be a distribution function on $[0,1]^2$. The result then easily follows from the definition of the HK-variation. In particular, we get

$$\begin{aligned}
V_{HK}(H) &= \sup \sum_{i=0}^{m(1)-1} \sum_{j=0}^{m(2)-1} |\Delta_{12} H(u_1^{(i)}, u_2^{(j)})| + \sup \sum_{i=0}^{m(1)-1} |\Delta_1 H(u_1^{(i)}, 1)| \\
&\quad + \sup \sum_{j=0}^{m(2)-1} |\Delta_2 H(1, u_2^{(j)})| \\
&= \sup \sum_{i=0}^{m(1)-1} \sum_{j=0}^{m(2)-1} \Big(H(u_1^{(i+1)}, u_2^{(j+1)}) - H(u_1^{(i+1)}, u_2^{(j)}) - H(u_1^{(i)}, u_2^{(j+1)}) \\
&\qquad + H(u_1^{(i)}, u_2^{(j)}) \Big) \\
&\quad + \sup \sum_{i=0}^{m(1)-1} \Big(H(u_1^{(i+1)}, 1) - H(u_1^{(i)}, 1) \Big) \\
&\quad + \sup \sum_{j=0}^{m(2)-1} \Big(H(1, u_2^{(j+1)}) - H(1, u_2^{(j)}) \Big) \\
&\quad = 1 + 1 + 1 = 3 \,.
\end{aligned}$$

□

Aistleitner and Dick (2015) prove that any function that is of bounded HK-variation and in addition right-continuous defines a finite signed measure on $[0,1]^2$ (see Theorem 3 in Aistleitner and Dick (2015)). Berghaus et al. (2017) define the corresponding Lebesgue-Stieltjes integral (see Definition A.5 in Berghaus et al. (2017)) and then prove the following integration by parts formula.

Proposition 5.8. *(Corollary A.7 in Berghaus et al. (2017))*
Let $f, g : [0,1]^2 \mapsto \mathbb{R}$ be of bounded HK-variation and right-continuous with either f or g continuous. Then, for any $(\boldsymbol{a}, \boldsymbol{b}] \subset [0,1]^2$ with $\boldsymbol{a} < \boldsymbol{b}$,

$$\begin{aligned}
\int_{(\boldsymbol{a},\boldsymbol{b}]} f dg &= \int_{(\boldsymbol{a},\boldsymbol{b}]} g df + fg(b_1, b_2) - fg(a_1, b_2) - fg(b_1, a_2) + fg(a_1, a_2) \\
&\quad - \int_{(a_1,b_1]} g(u, b_2) d_1 f(u, b_2) + \int_{(a_1,b_1]} g(u, a_2) d_1 f(u, a_2) \\
&\quad - \int_{(a_2,b_2]} g(b_1, u) d_2 f(b_1, u) + \int_{(a_2,b_2]} g(a_1, u) d_2 f(a_1, u) \,.
\end{aligned}$$

Let f and g be given as in Proposition 5.8. For Lebesgue-Stieltjes integrals one can integrate over arbitrary measurable sets to get the following useful identity

$$\int_{[0,1]^2} f dg = \int_{(0,1]^2} f dg + \int_{\partial_1 [0,1]^2} f dg \,,$$

where $\partial_1 [0,1]^2 = [0,1]^2 \backslash (0,1]^2$. Since g is of bounded HK-variation and right-continuous it defines a signed measure (see Aistleitner and Dick (2015)). If $\partial_1 [0,1]^2$ has measure zero under the corresponding signed measure, then the second integral vanishes and we get the equality

$$\int_{[0,1]^2} f\,dg = \int_{(0,1]^2} f\,dg\,.$$

This transition will be needed in the proof of Lemma 5.9 before the integration by parts formula from Proposition 5.8 can be applied.

Now, we come back to the integral $\int C\,d\alpha$. Let $\alpha : [0,1]^2 \mapsto \mathbb{R}$ be continuous and let C be a copula on $[0,1]^2$. If α is not of bounded variation, the integral $\int C\,d\alpha$ is defined by

$$\begin{aligned}\int_{[0,1]^2} C(u_1,u_2)\,d\alpha(u_1,u_2) = \int_{[0,1]^2} \alpha(u_1,u_2)\,dC(u_1,u_2) + \alpha(1,1)\\ - \int_0^1 \alpha(u,1)\,d_1u - \int_0^1 \alpha(1,u)\,d_2u\,.\end{aligned} \tag{5.7}$$

After clarifying the basics, we have a closer look on Lemma 1 in Veraverbeke et al. (2011) and adjust their proof.

Lemma 5.9. *Let $\mathbb{E}$ be the space of all distribution functions η on $[0,1]^2$ with $\eta(u,0) = \eta(0,u) = 0$ for all $u \in [0,1]$ and consider the map*

$$\Psi : \mathbb{E} \to \mathbb{R},\ \eta \mapsto \int_{[0,1]^2} \eta(u_1,u_2)d\eta(u_1,u_2)\,.$$

The map Ψ is Hadamard-differentiable at every copula C, tangentially to the set $\mathbb{E}_0$ of functions that are continuous on $[0,1]^2$. The derivative is given by

$$\Psi'_C : \mathbb{E}_0 \to \mathbb{R},\ \alpha \mapsto \int_{[0,1]^2} C(u_1,u_2)\,d\alpha(u_1,u_2) + \int_{[0,1]^2} \alpha(u_1,u_2)\,dC(u_1,u_2)\,,$$

where $\int C\,d\alpha$ is defined by Equation (5.7) if α is not of bounded HK-variation.

Proof. Let $\alpha_t \in \ell^\infty([0,1]^2)$ and $\alpha \in \mathbb{E}_0$ with $\alpha_t \to \alpha$ uniformly on $[0,1]^2$ such that $C_t := C + t\alpha_t \in \mathbb{E}$. In the following expression, plugging in the definition of C_t yields

$$\int_{[0,1]^2} C_t\,dC_t = \int_{[0,1]^2} C\,dC + t\int_{[0,1]^2} \alpha_t\,dC + t\int_{[0,1]^2} C\,d\alpha_t + t\int_{[0,1]^2} \alpha_t\,d(C_t - C)\,.$$

After transposing and extending the equation by $-\Psi'_C(\alpha) = -\int_{[0,1]^2} C\,d\alpha - \int_{[0,1]^2} \alpha\,dC$, we get

$$\begin{aligned}&\frac{\int_{[0,1]^2} C_t\,dC_t - \int_{[0,1]^2} C\,dC}{t} - \Psi'_C(\alpha)\\ &= \int_{[0,1]^2} (\alpha_t - \alpha)\,dC + \left(\int_{[0,1]^2} C\,d\alpha_t - \int_{[0,1]^2} C\,d\alpha\right) + \int_{[0,1]^2} \alpha_t\,d(C_t - C)\,.\end{aligned} \tag{5.8}$$

Now, the three terms on the right-hand side of Equation (5.8) are considered separately. Since $\alpha_t \to \alpha$ uniformly and C is of bounded HK-variation (see Lemma 5.7), the first term converges to zero.

For the second term of Equation (5.8), first note that the signed measure corresponding to $\alpha_t = (C_t - C)/t$ puts no mass in the axes through zero. Thus, the integration by parts formula from Proposition 5.8 with integration area $(0,1]^2$ can be applied. Further, we have $C(0,u) = C(u,0) = 0$, $C(u,1) = u$, and $C(1,u) = u$, for all $u \in [0,1]$, as well as $C(1,1) = 1$. Together, this yields

$$\begin{aligned}
\int_{[0,1]^2} C\,d\alpha_t = \int_{(0,1]^2} C\,d\alpha_t = \\
&= \int_{(0,1]^2} \alpha_t dC + C\alpha_t(1,1) - C\alpha_t(0,1) - C\alpha_t(1,0) + C\alpha_t(0,0) \\
&\quad - \int_{(0,1]} \alpha_t(u,1) d_1 C(u,1) + \int_{(0,1]} \alpha_t(u,0) d_1 C(u,0) \\
&\quad - \int_{(0,1]} \alpha_t(1,u) d_2 C(1,u) + \int_{(0,1]} \alpha_t(0,u) d_2 C(0,u) \\
&= \int_{(0,1]^2} \alpha_t dC + \alpha_t(1,1) - \int_{(0,1]} \alpha_t(u,1) d_1 u - \int_{(0,1]} \alpha_t(1,u) d_2 u \\
&= \int_{[0,1]^2} \alpha_t dC + \alpha_t(1,1) - \int_{[0,1]} \alpha_t(u,1) d_1 u - \int_{[0,1]} \alpha_t(1,u) d_2 u\,.
\end{aligned}$$

Using the definition of the integral $\int C d\alpha$ in Equation (5.7), we get

$$\begin{aligned}
\int_{[0,1]^2} C\,d\alpha_t - \int_{[0,1]^2} C\,d\alpha = &\int_{[0,1]^2} (\alpha_t - \alpha)\,dC + \alpha_t(1,1) - \alpha(1,1) \\
&- \int_{[0,1]} (\alpha_t(u,1) - \alpha(u,1))\,d_1 u \\
&- \int_{[0,1]} (\alpha_t(1,u) - \alpha(1,u))\,d_2 u\,.
\end{aligned}$$

With similar arguments as for the first term of Equation (5.8), we conclude that the second term of Equation (5.8) also converges to zero.

Expanding the third term of Equation (5.8) yields

$$\int_{[0,1]^2} \alpha_t\,d(C_t - C) = \int_{[0,1]^2} (\alpha_t - \alpha)\,d(C_t - C) + \int_{[0,1]^2} \alpha\,d(C_t - C)\,. \tag{5.9}$$

The class of functions of bounded HK-variation is closed under summation (see, e.g., Owen (2005)). Thus $C_t - C$ is of bounded HK-variation and, using the same reasons one more time, the first term on the right-hand side of Equation (5.9) converges to zero. For the last term of Equation (5.9), the continuity of α is exploited. For a given $\epsilon > 0$, there exist partitions $0 = t_0 < t_1 < \ldots < t_{m_1} = 1$ and $0 = s_0 < s_1 < \ldots < s_{m_2} = 1$, such that α varies less than ϵ on each rectangle $[t_{i-1}, t_i] \times [s_{j-1}, s_j]$, for $i = 1, \ldots, m_1$

and $j = 1, \ldots, m_2$. Now, denote the function that is equal to $\alpha(t_{i-1}, s_{j-1})$, and therefore constant, on the rectangle $[t_{i-1}, t_i] \times [s_{j-1}, s_j]$, for all $i = 1, \ldots, m_1$ and all $j = 1, \ldots, m_2$, by $\tilde{\alpha}$. For the last term of Equation (5.9), the following upper bound can then be derived

$$\begin{aligned}\left|\int \alpha\, d(C_t - C)\right| &\leq \left|\int (\alpha - \tilde{\alpha})\, d(C_t - C)\right| + \left|\int \tilde{\alpha}\, d(C_t - C)\right| \\ &\leq 6\|\alpha - \tilde{\alpha}\|_\infty + \sum_{i=1}^{m_1}\sum_{j=1}^{m_2} |\alpha(t_{i-1}, s_{j-1})| \cdot \left|\int_{[t_{i-1},t_i]\times[s_{j-1},s_j]} d(C_t - C)\right| \\ &\leq 6\epsilon + \|\alpha\|_\infty \sum_{i=1}^{m_1}\sum_{j=1}^{m_2} 4\|C_t - C\|_\infty \\ &\leq 6\epsilon + 4m_1 m_2 \|\alpha\|_\infty \|C_t - C\|_\infty \\ &\to 6\epsilon\,.\end{aligned}$$

Since ϵ can be chosen to be arbitrarily small, $\int \alpha\, d(C_t - C)$ converges to zero. Having shown that all three terms on the right-hand side of equation (5.8) converge to zero concludes the proof of Lemma 5.9. □

Bibliography

Aas, K., Czado, C., Frigessi, A., and Bakken, H. (2009). Pair-copula constructions of multiple dependence. *Insurance Math. Econom.*, 44(2):182–198.

Abdous, B., Genest, C., and Rémillard, B. (2005). Dependence properties of meta-elliptical distributions. In *Statistical modeling and analysis for complex data problems*, volume 1 of *GERAD 25th Anniv. Ser.*, pages 1–15. Springer, New York.

Aistleitner, C. and Dick, J. (2015). Functions of bounded variation, signed measures, and a general Koksma-Hlawka inequality. *Acta Arith.*, 167(2):143–171.

Anscombe, F. J. (1952). Large-sample theory of sequential estimation. *Proc. Cambridge Philos. Soc.*, 48:600–607.

Berg, D. (2009). Copula goodness-of-fit testing: an overview and power comparison. *Eur. J. Finance*, 15(7-8):675–701.

Berghaus, B., Bücher, A., and Volgushev, S. (2017). Supplement to “weak convergence of the empirical copula process with respect to weighted metrics”.

Bilodeau, M. and Brenner, D. (1999). *Theory of multivariate statistics*. Springer Texts in Statistics. Springer-Verlag, New York.

Blomqvist, N. (1950). On a measure of dependence between two random variables. *Ann. Math. Statistics*, 21:593–600.

Bücher, A. and Volgushev, S. (2013). Empirical and sequential empirical copula processes under serial dependence. *J. Multivariate Anal.*, 119:61–70.

Chen, X. and Fan, Y. (2005). Pseudo-likelihood ratio tests for semiparametric multivariate copula model selection. *Canad. J. Statist.*, 33(3):389–414.

Cherubini, U., Luciano, E., and Vecchiato, W. (2004). *Copula methods in finance*. Wiley Finance Series. John Wiley & Sons, Ltd., Chichester.

Clarkson, J. A. and Adams, C. R. (1933). On definitions of bounded variation for functions of two variables. *Transactions of the American Mathematical Society*, 35(4):824–854.

Demarta, S. and McNeil, A. J. (2005). The t copula and related copulas. *International statistical review*, 73(1):111–129.

Dengler, B. (2010). *On the asymptotic behaviour of the estimator of Kendall’s Tau*. PhD thesis, Vienna University of Technology (Austria).

Denker, M. (1985). *Asymptotic distribution theory in nonparametric statistics.* Advanced Lectures in Mathematics. Friedr. Vieweg & Sohn, Braunschweig.

Denuit, M., Purcaru, O., and Keilegom, I. V. (2006). Bivariate archimedean copula models for censored data in non-life insurance. *Journal of Actuarial Practice 1993-2006*, 13:5–32.

Elidan, G. (2013). Copulas in machine learning. In *Copulae in mathematical and quantitative finance*, volume 213 of *Lect. Notes Stat.*, pages 39–60. Springer, Heidelberg.

Embrechts, P., Lindskog, F., and McNeil, A. (2003). Chapter 8 - modelling dependence with copulas and applications to risk management. In *Handbook of heavy tailed distributions in finance*, volume 1 of *Handbooks in Finance*, pages 329–384. North-Holland, Amsterdam.

Embrechts, P., Mcneil, E., and Straumann, D. (1999). Correlation: Pitfalls and alternatives. *RISK Magazine*, pages 69–71.

Fang, H.-B., Fang, K.-T., and Kotz, S. (2002). The meta-elliptical distributions with given marginals. *J. Multivariate Anal.*, 82(1):1–16.

Fang, H.-B., Fang, K.-T., and Kotz, S. (2005). Corrigendum to: "The meta-elliptical distributions with given marginals" [J. Multivariate Anal. **8**2 (2002), no. 1, 1–16; mr1918612]. *J. Multivariate Anal.*, 94(1):222–223.

Fang, K. T., Kotz, S., and Ng, K. W. (1990). *Symmetric multivariate and related distributions*, volume 36 of *Monographs on Statistics and Applied Probability.* Chapman and Hall, Ltd., London.

Fermanian, J.-D. (2013). An overview of the goodness-of-fit test problem for copulas. In *Copulae in mathematical and quantitative finance*, volume 213 of *Lect. Notes Stat.*, pages 61–89. Springer, Heidelberg.

Fischer, M., Köck, C., Schlüter, S., and Weigert, F. (2009). An empirical analysis of multivariate copula models. *Quant. Finance*, 9(7):839–854.

Frahm, G., Junker, M., and Szimayer, A. (2003). Elliptical copulas: applicability and limitations. *Statist. Probab. Lett.*, 63(3):275–286.

Frees, E. W. and Valdez, E. A. (1998). Understanding relationships using copulas. *N. Am. Actuar. J.*, 2(1):1–25.

Genest, C., Carabarí n Aguirre, A., and Harvey, F. (2013). Copula parameter estimation using Blomqvist's beta. *J. SFdS*, 154(1):5–24.

Genest, C. and Favre, A.-C. (2007). Everything you always wanted to know about copula modeling but were afraid to ask. *J. Hydrol. Eng.*, 12(4):347–368.

Genest, C., Gendron, M., and Bourdeau-Brien, M. (2009a). The advent of copulas in finance. *The European Journal of Finance*, 15(7-8):609–618.

Genest, C., Ghoudi, K., and Rivest, L.-P. (1998). "Understanding relationships using copulas," by Edward Frees and Emiliano Valdez, January 1998. *N. Am. Actuar. J.*, 2(3):143–149.

Genest, C. and MacKay, J. (1986). The joy of copulas: bivariate distributions with uniform marginals. *Amer. Statist.*, 40(4):280–283.

Genest, C., Nešlehová, J., and Quessy, J.-F. (2012). Tests of symmetry for bivariate copulas. *Ann. Inst. Statist. Math.*, 64(4):811–834.

Genest, C. and Nešlehová, J. G. (2014). On tests of radial symmetry for bivariate copulas. *Statist. Papers*, 55(4):1107–1119.

Genest, C., Rémillard, B., and Beaudoin, D. (2009b). Goodness-of-fit tests for copulas: a review and a power study. *Insurance Math. Econom.*, 44(2):199–213.

Hoeffding, W. (1947). On the distribution of the rank correlation coefficient τ when the variates are not independent. *Biometrika*, 34:183–196.

Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statistics*, 19:293–325.

Hofert, M., Kojadinovic, I., Maechler, M., and Yan, J. (2018). *copula: Multivariate Dependence with Copulas.* R package version 0.999-19.1.

Jaser, M., Haug, S., and Min, A. (2017). A simple non-parametric goodness-of-fit test for elliptical copulas. *Depend. Model.*, 5(1):330–353.

Jaser, M. and Min, A. (2020). On tests for symmetry and radial symmetry of bivariate copulas towards testing for ellipticity. *Comput. Stat.*

Joe, H. (1990). Multivariate concordance. *J. Multivariate Anal.*, 35(1):12–30.

Joe, H. (1997). *Multivariate models and dependence concepts*, volume 73 of *Monographs on Statistics and Applied Probability.* Chapman & Hall, London.

Joe, H. (2015). *Dependence modeling with copulas*, volume 134 of *Monographs on Statistics and Applied Probability.* CRC Press, Boca Raton, FL.

Kendall, M. G. and Smith, B. B. (1940). On the method of paired comparisons. *Biometrika*, 31:324–345.

Khoudraji, A. (1995). *Contributions à l'étude des copules et à la modélisation de valeurs extrêmes bivariées*. PhD thesis, Université Laval (Canada).

Klugman, S. A. and Parsa, R. (1999). Fitting bivariate loss distributions with copulas. *Insurance Math. Econom.*, 24(1-2):139–148.

Klüppelberg, C., Kuhn, G., and Peng, L. (2008). Semi-parametric models for the multivariate tail dependence function—the asymptotically dependent case. *Scand. J. Statist.*, 35(4):701–718.

Kojadinovic, I. and Stemikovskaya, K. (2019). Subsampling (weighted smooth) empirical copula processes. *J. Multivariate Anal.*, 173:704–723.

Korn, R., Korn, E., and Kroisandt, G. (2010). *Monte Carlo methods and models in finance and insurance*. Chapman & Hall/CRC Financial Mathematics Series. CRC Press, Boca Raton, FL.

Kosorok, M. R. (2008). *Introduction to empirical processes and semiparametric inference*. Springer Series in Statistics. Springer, New York.

Kotz, S. and Nadarajah, S. (2004). *Multivariate t distributions and their applications*. Cambridge University Press, Cambridge.

Li, B. and Genton, M. G. (2013). Nonparametric identification of copula structures. *J. Amer. Statist. Assoc.*, 108(502):666–675.

Li, D. and Peng, L. (2009). Goodness-of-fit test for tail copulas modeled by elliptical copulas. *Statist. Probab. Lett.*, 79(8):1097–1104.

Li, D. X. (2000). On default correlation: A copula function approach. *Journal of Fixed Income*, 9(4):43–54.

Mai, J.-F. and Scherer, M. (2012). *Simulating copulas*, volume 4 of *Series in Quantitative Finance*. Imperial College Press, London. Stochastic models, sampling algorithms, and applications.

McNeil, A. J., Frey, R., and Embrechts, P. (2005). *Quantitative risk management*. Princeton Series in Finance. Princeton University Press, Princeton, NJ. Concepts, techniques and tools.

Miller, Rupert G., J. (1981). *Simultaneous statistical inference*. Springer-Verlag, New York-Berlin, second edition. Springer Series in Statistics.

Nelsen, R. B. (1999). *An introduction to copulas*, volume 139 of *Lecture Notes in Statistics*. Springer-Verlag, New York.

Oh, R., Ahn, J. Y., and Lee, W. (2020). On copula-based collective risk models: from elliptical copulas to vine copulas. *Scandinavian Actuarial Journal*.

Owen, A. B. (2005). Multidimensional variation for quasi-Monte Carlo. In *Contemporary multivariate analysis and design of experiments*, volume 2 of *Ser. Biostat.*, pages 49–74. World Sci. Publ., Hackensack, NJ.

Quessy, J.-F. (2016). A general framework for testing homogeneity hypotheses about copulas. *Electron. J. Stat.*, 10(1):1064–1097.

Quessy, J.-F. (2020). On nonparametric tests of multivariate meta-ellipticity. *Stat Papers*.

Quessy, J.-F. and Bellerive, R. (2013). Statistical procedures for the selection of a multidimensional meta-elliptical distribution. *J. SFdS*, 154(1):78–101.

Radulović, D., Wegkamp, M., and Zhao, Y. (2017). Weak convergence of empirical copula processes indexed by functions. *Bernoulli*, 23(4B):3346–3384.

Reinsel, D., Gantz, J., and Rydning, J. (2018). Data age 2025: The digitization of the world from edge to core. (November 2018, data refreshed May 2020). An IDC White Paper, International Data Corporation (IDC), Framingham, MA, USA. https://www.seagate.com/files/www-content/our-story/trends/files/dataage-idc-report-final.pdf [Accessed 2020-11-18].

Salmon, F. (2009). Recipe for disaster: The formula that killed wall street. *Wired magazine*. 2009-02-23, 12:00 PM. https://www.wired.com/2009/02/wp-quant/ [Accessed 2020-11-20].

Salvadori, G. and De Michele, C. (2007). On the use of copulas in hydrology: theory and practice. *Journal of Hydrologic Engineering*, 12(4):369–380.

Schmid, F. and Schmidt, R. (2007). Nonparametric inference on multivariate versions of Blomqvist's beta and related measures of tail dependence. *Metrika*, 66(3):323–354.

Schwarz, G. (1978). Estimating the dimension of a model. *Ann. Statist.*, 6(2):461–464.

Segers, J. (2012). Asymptotics of empirical copula processes under non-restrictive smoothness assumptions. *Bernoulli*, 18(3):764–782.

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, 8:229–231.

Song, S. and Singh, V. P. (2010). Meta-elliptical copulas for drought frequency analysis of periodic hydrologic data. *Stochastic Environmental Research and Risk Assessment*, 24(3):425–444.

Sproule, R. N. (1974). Asymptotic properties of U-statistics. *Trans. Amer. Math. Soc.*, 199:55–64.

van der Vaart, A. W. and Wellner, J. A. (1996). *Weak convergence and empirical processes*. Springer Series in Statistics. Springer-Verlag, New York. With applications to statistics.

Veraverbeke, N., Omelka, M., and Gijbels, I. (2011). Estimation of a conditional copula and association measures. *Scand. J. Stat.*, 38(4):766–780.

von Mises, R. (1947). On the asymptotic distribution of differentiable statistical functions. *Ann. Math. Statistics*, 18(3):309–348.

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