Estimating standard errors and efficient goodness-of-fit tests for regular vine copula models

ULF SCHEPSMEIER

Fakultät für Mathematik
Technische Universität München
85748 Garching

TECHNISCHE UNIVERSITÄT MÜNCHEN
Lehrstuhl für Mathematische Statistik

Estimating standard errors and efficient goodness-of-fit tests for regular vine copula models

Ulf Schepsmeier

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:		Univ.-Prof. Donna Pauler Ankerst, Ph.D.
Prüfer der Dissertation:	1.	Univ.-Prof. Claudia Czado, Ph.D.
	2.	Prof. Peter X.K. Song, Ph.D. University of Michigan, USA
	3.	Prof. Dr. Kjersti Aas University of Bergen, Norway

Die Dissertation wurde am 16.10.2013 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik am 27.12.2013 angenommen.

Zusammenfassung

In der vorliegenden Arbeit betrachten wir die sehr flexible Klasse der regulären Vine-Copula Modelle (R-vines). Diese multivariaten Copulas werden hierarchisch konstruiert, wobei lediglich zwei-dimensionale Copulas als Bausteine benutzt werden. Die Zerlegung selber wird Paar-Copula-Konstruktion (PCC) genannt. Für diese R-vine Copula Modelle führen wir die konzeptionelle wie auch algorithmisch umgesetzte Berechnung der Scorefunktion und der beobachteten Informationsmatrix ein. Die damit einhergehende Abschätzung der Fisher-Information erlaubt es uns die Standardfehler der Parameterschätzer routinemäßig zu bestimmen. Außerdem beseitigen die analytischen Ausdrücke des R-vine log-likelihood Gradienten und der Hessematrix Genauigkeitsdefizite bei den bisher gebräuchlichen numerischen Ableitungen. Die hierfür benötigten bivariaten Ableitungen bezüglich der Copulaparameter und -argumente werden ebenfalls berechnet. Hier sind insbesondere die schwierigen Ableitungen der Student's t-copula hervorzuheben.

Um R-vine Copulas mittels statistischer Tests zu validieren führen wir einige Tests für die Güte der Anpassung (Goodness-of fit Tests) ein. Speziell betrachten wir zwei neue Gütetests, die aus dem Informations- und Ausprägungstest von White (1982) und dem Informationsverhältnistest von Zhang et al. (2013) hervorgehen. Wir berechnen die entsprechenden Teststatistiken und beweisen deren asymptotische Verteilung. Für Vergleichszwecke führen wir noch 13 weitere Gütetests ein, die wir aus dem zwei-dimensionalen Fall adaptieren und für den R-vine Fall erweitern. Eine intensive Vergleichsstudie, die die Power untersucht, zeigt die Überlegenheit der informationsmatrixbasierten Tests. Berechnet man die size und power basierend auf simulierten gebootstrapten p-Werten, so können exzellente Resultate erzielt werden, während Tests, die auf asymptotischen p-Werten basieren ungenau sind. Insbesondere in höheren Dimensionen.

Eine Anwendung der gezeigten Algorithmen auf US-Wechselkurs Daten zeigt die asymptotische Effizienz unserer Methoden bei der Berechnung von Standardfehlern für reguläre Vine-Copulas. Ferner wenden wir die besten Gütetests auf die genannten US-Wechselkurse, sowie auf ein Portfolio von Aktienindizes und deren Volatilitätsindizes an, um die beste Vine-Copula zu wählen.

Abstract

In this thesis we consider the flexible class of regular vine (R-vine) copula models. R-vine copulas are multivariate copulas based on a pair-copula construction (PCC) which is constructed hierarchically from only bivariate copulas as building blocks. We introduce theory and algorithms for the computation of the score function and the observed information matrix in R-vine models. The corresponding approximation of the Fisher information allows to routinely estimate parameter standard errors. Furthermore, the analytical expression of the R-vine log-likelihood gradient and Hessian matrix overcomes reliability and accuracy issues associated with numerical differentiation. Needed bivariate derivatives with respect to copula parameters and arguments are derived, in particular for the Student's t-copula.

To validate R-vine copula models based on statistical tests we introduce several goodness-of fit tests. In particular we propose two new goodness-of-fit tests arising from the information matrix and specification test proposed by White (1982) and the information ratio test by Zhang et al. (2013). The test statistics are derived and their asymptotic distribution proven. Further 13 goodness-of-fit tests are adapted from the bivariate case and compared in an extensive power study, which shows the superiority of the information matrix based tests. The bootstrapped simulation based tests show excellent performance with respect to size and power, while the asymptotic theory based tests are inaccurate in higher dimensions.

To illustrate the need for estimated standard errors and the asymptotic efficiency of our algorithms we apply our methods in a rolling window analysis of US exchange rates. The best performing goodness-of-fit tests are applied to the US exchange rates data set as well as to a portfolio of stock indices and their related volatility indices selecting among different R-vine specifications.

Acknowledgments

This thesis would not be possible without the help, support and encouragement of many people. All of them I want to thank here.

First of all, I want to thank my supervisor Prof. Claudia Czado for giving me the opportunity for this thesis, her helpful advice and support throughout the last years. I am very grateful for the position in her BMBF project “SIMDATA-NL: Nichtlineare Charakterisierung und Analyse von FEM-Simulationsergebnissen für Autobauteile und Crash-Tests” giving me the financial funding and the opportunity to learn new scientific areas and application areas. Furthermore, it enabled me to attend international scientific conferences and meet renowned researchers of different fields and parts of the world.

I also would like to thank Prof. Kjersti Aas and Prof. Peter Song for acting as referees of this thesis.

A special thanks goes to Jakob Stöber and Eike Brechamnn for their fruitful collaborations, inspiring discussions and support with the implementation of some of the algorithms developed in this thesis. Many thanks to Harry Joe and in particular Peter Song for their inspiring ideas, great discussions and encouraging feedback during their stay in Munich. It was a great pleasure to work with you. I also want to thank Aleksey Min and my colleagues at the Chair of Mathematical Statistics for many stimulating exchanges of ideas, and for the great time we spent together.

Also, I gratefully acknowledge the financial support of the TUM Graduate School's Faculty Graduate Center ISAM (International School of Applied Mathematics).

Lastly, and most importantly, I am grateful to my family and friends for their support, time, patience and steady encouragement.

Contents

Introduction

With more data becoming accessible the modeling of high-dimensional data is nowadays an often discussed problem and received considerable attention in the last decades. Beside the classical multivariate distribution functions as for example the multivariate Gaussian distribution the concept of copulas has been developed to model dependencies between random variables. Copulas allow to model marginal distributions and the dependence structure separately. In this thesis we discuss in particular the flexible class of regular vine (R-vine) copula models, a pair copula construction (PCC). They decompose the d-dimensional density into unconditional and conditional bivariate copulas, so called pair-copulas. The high flexibility is gained on the one hand from the independently chosen building block copula families and secondly from the choice of the decomposition itself. Starting with Aas et al. (2009), statistical inference such as maximum likelihood estimation of the pair-copula parameters or simulation algorithms for vine copula models are developed in the existing literature.

Although there exists an asymptotic theory for the maximum likelihood parameter estimates, there is a surprising scarcity in the literature considering the uncertainty in point parameter estimates for R-vines. One main contribution of this thesis is the development and implementation of R-vine copula standard error estimation by deriving the observed Hessian matrix as approximation for the Fisher information matrix.

The second theme is model verification by goodness-of-fit testing, which was not investigated for R-vine copula models although several goodness-of-fit tests are available for bivariate copulas. We developed two new goodness-of-fit tests for vine copula models based on the Information matrix equality and specification test of White (1982), and extended several goodness-of-fit tests, considered so far for bivariate copulas, to the vine copula case.

Copulas, the central concept in this thesis, have their origin in the Fréchet classes (Fréchet 1951), but its statistical break through came with the Theorem of Sklar (1959). It states that for each set of continuous random variables $X_1, \ldots, X_d$ with marginal distribution functions $F_1, \ldots, F_d$ there exists an unique copula C giving the joint distribution function of $X_1, \ldots, X_d$. A different naming of the same concept is used in the studies of Kimeldorf and Sampson (1975) or Galambos (1987). More recently the books of Joe (1997), Nelsen (2006) or Mai and Scherer (2012) give a good overview of the concept of copulas, their properties and inference methods.

Therefore bivariate copulas are well studied and many different copula families are proposed and investigated having specific properties. Especially in economics and finance copulas are wildly applied (see for example Patton 2012) because of their high flexibility and easy computation. For model verification based on statistical tests several goodness-of-fit approaches have been developed and applied. For example tests based on the empirical copula process (Genest et al. 2009), the probability integral transform (Breymann et al. 2003, Berg and Bakken 2007) or other approaches (Genest et al. 2012, Huang and Prokhorov 2013 and many more) are available. Further, model comparison can be achieved by tests of Vuong (1989) or Clarke (2007), or distances like the Kullback and Leibler (1951) information criterion (KLIC) can be derived.

But copulas of classical classes such as the Archimedean ones are rather limited in higher dimensions. They loose their flexibility, easy and fast computation and model verification by goodness-of-fit tests or the calculation of the KLIC became unstable, difficult or even impossible. Pair-copula constructions overcome most of these difficulties. Build from bivariate building blocks only they inherit several beneficial properties of the bivariate case. The hierarchical structure of the decomposition of the d-dimensional density into $d(d-1)/2$ unconditional and conditional pair-copulas allow for a high flexibility in the construction. Further, each pair-copula can be selected independently from set of (parametric) copula families. This is still tractable in moderate dimensions. While most applications stay in dimension up to 50, Heinen and Valdesogo (2009) propose a vine copula model with even 100 variables.

The graphical representation of such a construction in a nested set of trees was called a regular vine (R-vine) copula by Bedford and Cooke (2001, 2002). Later the statistical inference for PCCs was developed by Aas et al. (2009). Since then vine copulas are on the rise (see for example Kurowicka and Cooke 2006, Dißmann et al. 2013, Joe et al. 2010 or Gräler and Pebesma 2011 just to name a few). They have been applied to model dependence in various areas including agricultural science and electricity loads (Smith et al. 2010), exchange rates (Czado et al. 2012, Stöber and Czado 2012), order books and headache data (Panagiotelis et al. 2012).

In this thesis we extend the existing literature about vines by two important statistical issues. First, while it is a standard exercise in multivariate statistics to compute the uncertainty incorporated in parameter point estimates this was not possible so far for vine copula models. We will develop algorithms to calculate the log-likelihood gradient and Hessian matrix to approximate the Fisher information for standard error estimation. The needed first and second derivatives with respect to the parameter(s) and copula arguments of bivariate copulas will be derived. Secondly, goodness-of-fit tests for vine copula model verification will be introduced and investigated with respect to their size and power. In particular, two new goodness-of-fit tests based on the information matrix equality will be introduced. Another 13 goodness-of-fit tests are adapted from the bivariate case for comparison. The general outline of this thesis is as follows.

Chapter 1, which is taken from Schepsmeier and Stöber (2012) and Schepsmeier and Stöber (2013), introduces the concept of copulas. It gives the most important definitions and the main notation concerning copulas used in the subsequent chapters. In particular, the class of Archimedean copulas and elliptical copulas are presented. The second part of this chapter derives expressions for the observed and expected Fisher information for the proposed bivariate copula families. In particular for the Student's t-copula the first and second derivative with respect to both copula parameters are derived as well as the Fisher information. It corrects several flaws in the existing literature. Numerical issues and our implementation is discussed. Further, a practical example computing standard errors in a rolling window analysis shows the usefulness of the derived quantities.

In **Chapter 2**, which is partly based on material from Stöber and Schepsmeier (2013), pair-copula constructions (PCCs) and regular vine (R-vine) copula models are introduced. They are the models of consideration in the Chapters 3-4. We will define PCCs and the class of R-vine copula models, and give the general form of its density function. The last part extensively attends to the implemented software regarding bivariate copulas and R-vine copulas. It will lay down the main functionality of the R-package **VineCopula** of Schepsmeier et al. (2012) enabling statisticians and practitioners to facilitate inference for vine copulas.

The estimation of standard errors for an R-vine copula model is proposed in **Chapter 3**. The content is based on material from Stöber and Schepsmeier (2013). The first and second derivative of the regular vine copula log-likelihood is derived with respect to its parameters in a new algorithmic manner. Thus we can estimate the observed Fisher information matrix of R-vines copula models. The algorithms make use of the hierarchical nature for subsequent calculation of the log-likelihood. The routinely calculation of the observed Fisher information overcomes reliability and accuracy issues associated with numerical differentiation in multi-dimensional models. In particular, for statistical estimation methods based on numerical optimization closed form expressions for the gradient are computationally advantageous. For example this allows to perform maximum likelihood estimation in a multi-dimensional setup. Here optimization based on numerical differentiation can be highly unreliable.

A simulation study in Section 3.3 confirms that the standard errors we estimate are appropriate. Confidence intervals are estimated using the observed information gained from our algorithms and by sample estimates of the sequential approach proposed by Hobæk Haff (2013). An example of application to a financial data set illustrates their computation and gives proof of its usefulness.

Chapter 4, which is based on material from Schepsmeier (2013a) and Schepsmeier (2013b), covers the examination of goodness-of-fit (GOF) tests for vine copula models. In particular, two new GOF tests based on the information matrix equation and specification test of White (1982) are introduced. The first one directly applies White's theorem to the vine copula case. It extends the approach of Huang and Prokhorov (2013), who proposed a GOF test for bivariate copulas based on White (1982). The test statistic derivation is mainly facilitated by the point-wise calculation of the sum of the Hessian matrix and outer product of gradient using the algorithms of Chapter 3. The corresponding critical value can be approximated by asymptotic theory or simulation using bootstrap. The simulation based tests show excellent performance with regard to observed size and power in extensive simulation studies, while the asymptotic theory based test is inaccurate for $n \leq 10000$ for a 5-dimensional model. In the applied 8-dimensional case even 20000 observations are not enough.

The second new introduced GOF test is inspired by Presnell and Boos (2004), Zhou et al. (2012) and Zhang et al. (2013) and arises from the information matrix ratio. In contrast to the White test, which is based on the difference of the negative Hessian matrix and the outer product of the score function, the information ratio test considers the ratio of these two quantities. The corresponding test statistic is derived and its asymptotic normality proven. The test's power is again investigated in a simulation study.

Both new GOF tests are compared in a high dimensional setting to 13 other GOF tests, adapted from the bivariate copula case. In particular, we will compare to GOF tests based on the empirical copula process as suggested by Genest et al. (2009) and based on the multivariate probability integral transform (see Breymann et al. 2003 or Berg and Bakken 2007). A combination of these two approaches (see Genest et al. 2009) is also considered in the comparison setting. An extensive simulation study shows the excellent performance of the introduced information based GOF tests with respect to size and power. Furthermore, the superiority of these two tests against most other goodness-of-fit tests is illustrated.

Finally, the best performing tests are applied in two examples validating different R-vine specifications. The first example is an application to a portfolio of stock indices and their related volatility indices, while the second one selects among different R-vine specifications to model dependency among exchange rates.

An outlook and discussion of the previous chapters is given in **Chapter 5**. Some further ideas on goodness-of-fit tests for regular vine copula models are proposed. Possible pros and cons of possible extension of bivariate copula goodness-of-fit tests are considered. In particular the very interesting hybrid approach for copula goodness-of-fit testing suggested by Zhang et al. (2013) will be briefly discussed.

The appendices finally contain additional material for the previous chapters. Appendix A gives details for the calculation of the second derivative of the R-vine log-likelihood, the calculation of the covariance matrix in the Gaussian case and the R-vine copula model specification for the exchange rate data set used in the example. Some technical details for the goodness-of-fit tests considered in Chapter 4 are given in Appendix B, while Appendix C gives the model specifications considered in the power studies of Section 4.1.3 and 4.4.

Chapter 1

Copulas[1]

A copula is a multivariate distribution function C: $[0,1]^d \mapsto [0,1]$ on the unit hypercube with uniform univariate marginal distributions. Let $F_1, \ldots, F_d$ be the marginal distribution functions of the continuous random variables $X_1, \ldots, X_d$. Now, for an arbitrary multivariate continuous distribution function F we know (Sklar 1959) that there exists a unique copula C such that

$$F(x_1, \ldots, x_d) = C(F_1(x_1), \ldots, F_d(x_d)).$$

Thus, copulas allow to model univariate effects of a random vector $\boldsymbol{X} = (X_1, \ldots, X_d)$ and their joint dependence structures, separately. Standard references on copula theory include the books by Joe (1997) and Nelsen (2006). In this chapter we concentrate on one- or two-parametric copulas such as copulas from the elliptical class (Section 1.2) or Archimedean ones (Section 1.1) in dimension $d = 2$, so called bivariate copulas. These will form the building blocks for the investigated vine copula models in Chapters 2-4. If C is two-times partial differentiable the bivariate copula density is

$$c(u_1, u_2) = \frac{\partial^2 C(u_1, u_2)}{\partial u_1 \partial u_2},$$

where $u_1 := F_1(x_1)$ and $u_2 := F_2(x_2)$ are so called copula data on $[0, 1]$.

Beside the parametric copulas on which we will focus here non-parametric approaches are available as well. The use of non-parametric copulas avoids the need for the error-prone selection from pre-specified sets of parametric copulas. But they require other conditions, for example the choice of a bandwidth parameter, needing expert knowledge. Non-parametric approaches such as the Bernstein copula are for example treated in Sancetta and Satchell (2004) or the empirical copula in Rüschendorf (1976).

In the last decades we have seen a rising interest in the concept of copulas both by statisticians and practitioners (see e.g. Joe 1997, Embrechts et al. 2003, Nelsen 2006, Chan et al. 2009, Mai and Scherer 2012) and various applications in particular of bivariate copulas have been considered (see for example Patton 2006, Acar et al. 2012). Thus

[1] The contents of this chapter is based on Schepsmeier, U. and Stöber, J. (2013), Derivatives and Fisher information of bivariate copulas, forthcoming in Statistical Papers, and Schepsmeier, U. and Stöber, J. (2012), Web supplement: Derivatives and Fisher information of bivariate copulas, TU München

many copula families were analyzed and new ones are added to the literature. Statistical inference is facilitated by maximum likelihood or Bayesian approaches, but estimation of standard errors of copula parameter estimates are treated rather poorly. One reason may be that so far closed form expressions for the Fisher information matrix are missing for important copula families. We will fill this gap in the forthcoming sections.

To assess the uncertainty of parameter estimates, which is at the core of statistical analysis, the observed information matrix based on the Hessian matrix of the log-likelihood or the asymptotic Fisher information can be considered. This however requires the calculation of derivatives with respect to the copula parameters which is not straightforward when the copula is not available in closed form as for the popular class of elliptical copulas. Also for an efficient numerical treatment and algorithms such as the maximization by parts (Song et al. 2005), the derivatives will be required and cannot be approximated by finite differences. These can be numerically unstable, especially for higher-order derivatives (see McCullough 1999 and references therein, in particular Donaldson and Schnabel 1987). Also, for p parameters, approximating the score function by finite differences amounts to at least $2p$ evaluations of the likelihood function which is infeasible in higher dimensions.

Next, we will introduce some of the most important and well known copulas, starting with the Archimedean class. These and the Gaussian and Student's t-copula of the elliptical class of Section 1.2 we will treat in more detail with respect to their derivatives and Fisher information matrix in Section 1.3. Our particular focus lies on the bivariate Student's t-distribution and its copula, the bivariate t-copula. These are perhaps the most important distributions in financial applications due to their tail behavior which is considered to be more realistic than the behavior of the Gaussian distribution (see Demarta and McNeil 2005 and references therein). Our main contributions are: We obtain all derivatives of the bivariate t-copula which are required in statistical applications and calculate the Fisher information of the related Student's t-distribution. Further, we provide a numerical implementation for all parametric families investigated in this chapter and demonstrate its accuracy.

1.1 Archimedean copulas

A very popular class of parametric copulas are the Archimedean copulas because of their simple construction and nice properties such as the easy calculation of the density or the closed form expression for the rank dependence coefficient Kendall's τ. An Archimedean copula is a function C,

$$C(u_1, u_2) = \varphi(\varphi^{[-1]}(u_1) + \varphi^{[-1]}(u_2)),$$

where φ is the so called generator function and $\varphi^{[-1]}$ is the pseudo-inverse of φ which is defined as follows:

$$\begin{aligned}&\varphi^{[-1]} : [0, 1] \to [0, \infty)\\&\varphi^{[-1]}(x) := \inf\{u : \varphi(u) \leq x\}.\end{aligned}$$

For necessary and sufficient conditions on the generator function φ, we refer to McNeil and Nešlehová (2009). In order to ensure the existence of continuous derivatives, we will assume throughout this thesis that $\varphi : [0, \infty) \mapsto [0, 1]$, with $(-1)^j \varphi^{(j)} \geq 0$, $0 \leq j \leq 2$, where $\varphi^{(2)}$ is continuous, $\varphi(0) = 1$, and $\lim_{x\to\infty} \varphi(x) = 0$. A good overview discussing Archimedean copulas and extensions is given in Joe (1997) together with a list of common parametric families. Here, we concentrate on four of the most popular members of this class, namely the Clayton/MTCJ, Gumbel, Frank and Joe copula. The following examples defining the copula cdf and its density. Furthermore, the conditional cdf function of U_1 given $U_2 = u_2$ where $(U_1, U_2) \sim C$ is given. It is needed for example for the calculation of the pair-copula construction which we will introduce in Chapter 2. Further, the inverse function of the conditional cdf function can be used for the sampling algorithm of Genest and Favre (2007).

Thus we denote the conditional cdf as

$$h(u_1, u_2; \theta) := \partial_2 C(u_1, u_2; \theta) = \frac{\partial C(u_1, u_2; \theta)}{\partial u_2},$$

and call it a h-function as it is standard in the literature. For all the forthcoming definitions we refer to Joe (1997).

Example 1.1 (Clayton/MTCJ copula)
The first Archimedean copula we consider is the Clayton/MTCJ copula. The generator function φ of this copula is $\varphi(t) = (1 + t)^{-\frac{1}{\theta}}$ and the copula is given by

$$C(u_1, u_2; \theta) = (u_1^{-\theta} + u_2^{-\theta} - 1)^{-\frac{1}{\theta}} = A(u_1, u_2, \theta)^{-\frac{1}{\theta}},$$

with $A(u_1, u_2, \theta) := u_1^{-\theta} + u_2^{-\theta} - 1$, and corresponding density

$$c(u_1, u_2; \theta) = (1 + \theta)(u_1 u_2)^{-1-\theta}(u_1^{-\theta} + u_2^{-\theta} - 1)^{-\frac{1}{\theta}-2} = \frac{(1 + \theta)(u_1 u_2)^{-1-\theta}}{A(u_1, u_2, \theta)^{\frac{1}{\theta}+2}},$$

where $0 < \theta < \infty$ controls the degree of dependence. For $\theta \to \infty$ the Clayton/MTCJ copula converges to the monotonicity copula with perfect positive dependence, $\theta \to 0$ corresponds to independence.

The h-function of the Clayton copula is

$$h(u_1, u_2; \theta) = u_2^{-\theta-1} \cdot A(u_1, u_2, \theta)^{-1-\frac{1}{\theta}}.$$

Example 1.2 (Gumbel copula)
The Gumbel copula is given by

$$C(u_1, u_2; \theta) = \exp[-\{(-\ln(u_1))^\theta + (-\ln(u_2))^\theta\}^{\frac{1}{\theta}}] = \exp[-(t_1 + t_2)^{\frac{1}{\theta}}],$$

where $t_i := (-\ln(u_i))^\theta, i = 1, 2$. Here, $\theta \geq 1$ is the dependence parameter. For $\theta \to \infty$ the Gumbel copula converges to the comonotonic copula with perfect positive dependence, in contrast $\theta = 1$ corresponds to independence. The h-function and the density are as follows:

$$\begin{aligned}
h(u_1, u_2; \theta) &= -\frac{\mathrm{e}^{-(t_1+t_2)^{\frac{1}{\theta}}} (t_1 + t_2)^{\frac{1}{\theta}-1} t_2}{u_2 \ln(u_2)}, \\
c(u_1, u_2; \theta) &= C(u_1, u_2; \theta)(u_1 u_2)^{-1}\{(-\ln(u_1))^\theta + (-\ln(u_2))^\theta\}^{-2+\frac{2}{\theta}} \\
&\quad \times (\ln(u_1)\ln(u_2))^{\theta-1}\{1 + (\theta - 1)((-\ln(u_1))^\theta + (-\ln(u_2))^\theta)^{-\frac{1}{\theta}}\} \\
&= C(u_1, u_2; \theta)\frac{1}{u_1 u_2}(t_1 + t_2)^{-2+\frac{2}{\theta}}(\ln(u_1)\ln(u_2))^{\theta-1}\{1 + (\theta - 1)(t_1 + t_2)^{-\frac{1}{\theta}}\}.
\end{aligned}$$

Further, we consider the Frank and Joe copula. While the Frank copula has no tail dependence and parameter support $\theta \in (-\infty, \infty)\backslash\{0\}$ the Joe copula has upper tail dependence $2 - 2^{1/\theta}$ with $\theta \geq 1$.

$$C_{Frank}(u_1, u_2; \theta) = -\frac{1}{\theta} \ln \left(\frac{1}{1 - e^{-\theta}} [(1 - e^{-\theta}) - (1 - e^{-\theta u_1})(1 - e^{-\theta u_2})] \right),$$
$$C_{Joe}(u_1, u_2; \theta) = 1 - \left((1 - u_1)^\theta + (1 - u_2)^\theta - (1 - u_1)^\theta (1 - u_2)^\theta \right)^{\frac{1}{\theta}}.$$

Figure 1.1 illustrates the bivariate contour plots corresponding to a bivariate meta distribution with standard normal margins and specified bivariate copula, namely Clayton, Gumbel, Frank and Joe, with a Kendall's τ value of 0.5. For a further discussion including densities and h-functions, we refer to the web supplement Schepsmeier and Stöber (2012) of Schepsmeier and Stöber (2013).

1.2 Elliptical copulas

The far more interesting and wildly used copulas are of the elliptical class. These consist of the copulas corresponding to elliptical distributions by Sklar's theorem. The Gaussian and Student's t-distributions are the most prominent members. For the following definitions and calculations we refer to Schepsmeier and Stöber (2013).

Gaussian copula

The Gaussian copula is given by

$$C(u_1, u_2; \rho) = \Phi_2(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \rho),$$

where $\Phi_2(\cdot, \cdot, \rho)$ is the joint distribution function of two standard normally distributed random variables with correlation $\rho \in (-1, 1)$, Φ is the cumulative distribution function (cdf) of $N(0, 1)$ (the standard normal distribution) and Φ^{-1} (the quantile function) is its functional inverse. The density of the bivariate Gaussian copula is

$$c(u_1, u_2; \rho) = \frac{1}{\sqrt{1 - \rho^2}} \exp \left\{ -\frac{\rho^2 (x_1^2 + x_2^2) - 2\rho x_1 x_2}{2(1 - \rho^2)} \right\},$$

where $x_1 := \Phi^{-1}(u_1)$ and $x_2 := \Phi^{-1}(u_2)$.
The conditional distribution function of the first variable U_1 given $U_2 = u_2$ is

$$h(u_1, u_2; \rho) = \frac{\partial}{\partial u_2} C(u_1, u_2; \rho) = \Phi_2 \left(\frac{\Phi^{-1}(u_1) - \rho \Phi^{-1}(u_2)}{\sqrt{1 - \rho^2}} \right),$$

see Aas et al. (2009). The top left panel of Figure 1.2 illustrates the contour plot of the bivariate Gauss copula considering standard normal margins and Kendall's $\tau = 0.5$.

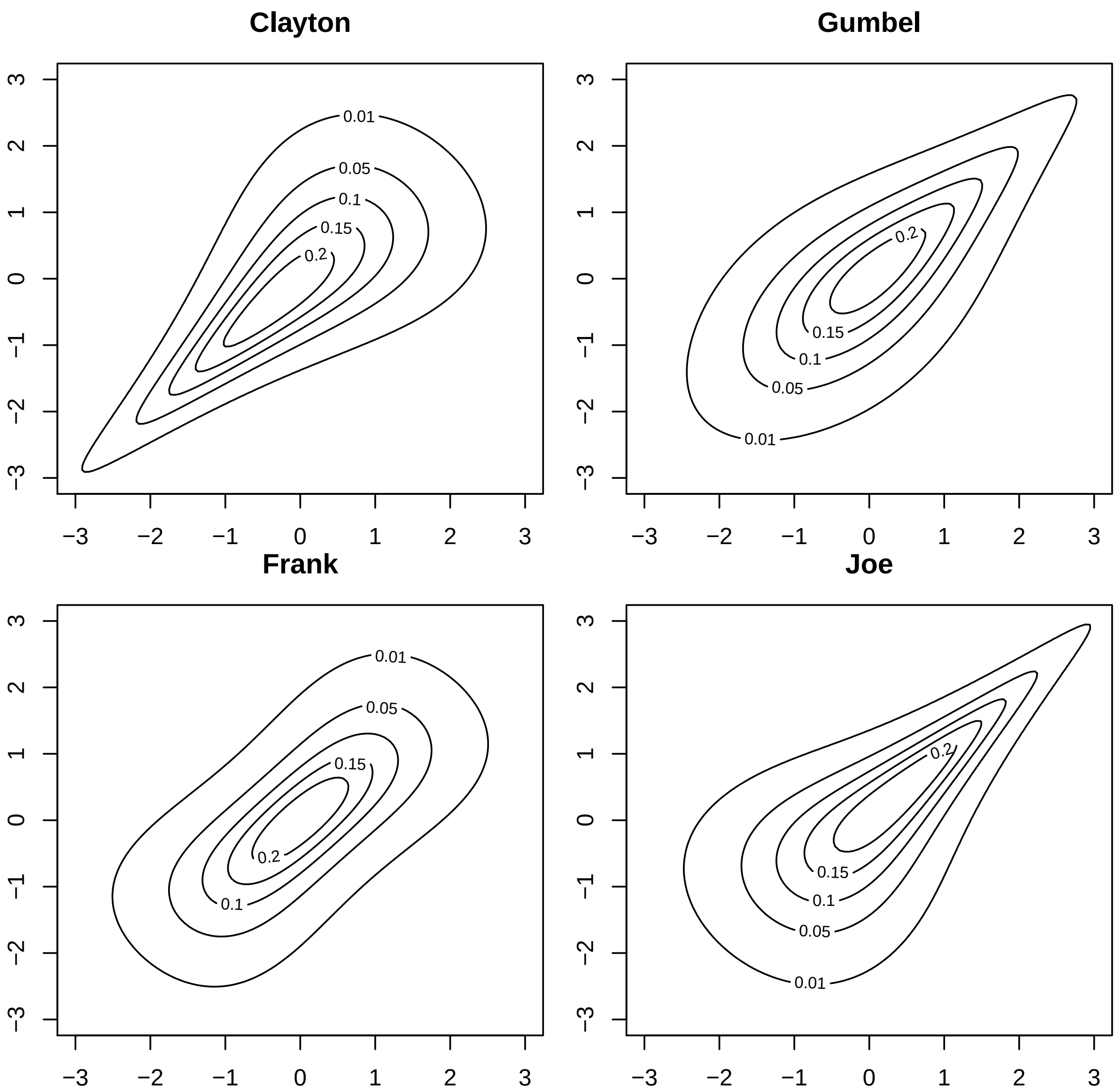


Figure 1.1: Bivariate contour plots of Clayton (top left), Gumbel (top right), Frank (bottom left) and Joe (bottom right) with standard normal margins and a Kendall's τ value of 0.5.

Student's t-copula

We denote the cumulative distribution function of the univariate Student's t-distribution with $\nu > 0$ degrees of freedom by t_ν and the corresponding quantile function by t_ν^{-1}. The corresponding density is given by

$$dt(x_i;\nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\pi\nu}}\left(1+\frac{x_i^2}{\nu}\right)^{-\frac{\nu+1}{2}}, \quad i = 1, 2,$$

where $\Gamma(\cdot)$ is the Gamma function. Let us further write the quantile of the univariate Student's t-distribution as

$$x_i := t_\nu^{-1}(u_i), \quad u_i \in (0,1), \quad i = 1,2,$$

then the bivariate t-copula is defined by

$$C(u_1, u_2; \rho, \nu) = t_{2;\rho,\nu}(x_1, x_2),$$

where $t_{2;\rho,\nu}$ is the cumulative distribution function of the bivariate Student's t-distribution. The density of the bivariate t-copula with association $\rho \in (-1,1)$ and degrees of freedom $\nu > 0$ is given by

$$c(u_1, u_2; \rho, \nu) = \frac{1}{2\pi\sqrt{1-\rho^2}} \frac{1}{dt(x_1;\nu)dt(x_2;\nu)} \left(1 + \frac{x_1^2 + x_2^2 - 2\rho x_1 x_2}{\nu(1-\rho^2)}\right)^{-\frac{\nu+2}{2}}.$$

Again, the h-function corresponding to the t-copula has been derived in Aas et al. (2009):

$$h(u_1, u_2; \rho, \nu) = t_{\nu+1}\left(\frac{t_\nu^{-1}(u_1) - \rho t_\nu^{-1}(u_2)}{\sqrt{\frac{\left(\nu + \left(t_\nu^{-1}(u_2)\right)^2\right)(1-\rho^2)}{\nu+1}}}\right) = t_{\nu+1}\left(\frac{x_1 - \rho x_2}{\sqrt{\frac{\left(\nu + x_2^2\right)(1-\rho^2)}{\nu+1}}}\right).$$

Bivariate contour plots of the Student's t copula with Kendall's $\tau = 0.5$ and different degrees-of-freedom parameter ν are plotted in Figure 1.2 considering standard normal margins.

1.3 Derivatives and Fisher Information

Considering asymptotic properties of ML estimation, it is well known that under certain regularity conditions (c.f. Bickel and Doksum 2007, p. 386 or Lehmann and Casella 1998, p. 449), the ML estimator $\hat{\boldsymbol{\theta}}_n \in \mathbb{R}^p$ obtained from n observations is strongly consistent and asymptotically normal:

$$\sqrt{n}\, \mathcal{I}(\boldsymbol{\theta})^{1/2}\left(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\right) \xrightarrow{d} N(0, I_p) \text{ as } n \to \infty,$$

where $\boldsymbol{\theta} \in \mathbb{R}^p$ is the true parameter and I_p is the $p \times p$ identity matrix. Here, the (expected) Fisher information matrix $\mathcal{I}(\boldsymbol{\theta})$ can be obtained as

$$\mathcal{I}(\boldsymbol{\theta}) = -\mathbb{E}_{\boldsymbol{\theta}}\left[\left(\frac{\partial^2}{\partial\theta_i \partial\theta_j} l(\boldsymbol{\theta}|\boldsymbol{X})\right)_{i,j=1,\ldots,p}\right] = \mathbb{E}_{\boldsymbol{\theta}}\left[\left(\frac{\partial}{\partial\theta_i} l(\boldsymbol{\theta}|\boldsymbol{X}) \cdot \frac{\partial}{\partial\theta_j} l(\boldsymbol{\theta}|\boldsymbol{X})\right)_{i,j=1,\ldots,p}\right],$$

where $l(\boldsymbol{\theta}|\boldsymbol{x})$ is the log-likelihood of $\boldsymbol{\theta}$ given an observation of $\boldsymbol{X} = \boldsymbol{x}$. In a finite sample of n independent observation $(x_1, \ldots, x_n)$, it has been argued (Efron and Hinkley 1978) that the Fisher information should be replaced by the observed information $\mathcal{I}_n(\hat{\boldsymbol{\theta}}_n)$ at the ML estimate $\hat{\boldsymbol{\theta}}_n$

$$\mathcal{I}_n(\hat{\boldsymbol{\theta}}_n) = \left[\left(\sum_{k=1}^{n} \frac{\partial^2}{\partial\theta_i \partial\theta_j} l(\boldsymbol{\theta}|x_k)\right)_{i,j=1,\ldots,p}\right]_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_n}.$$

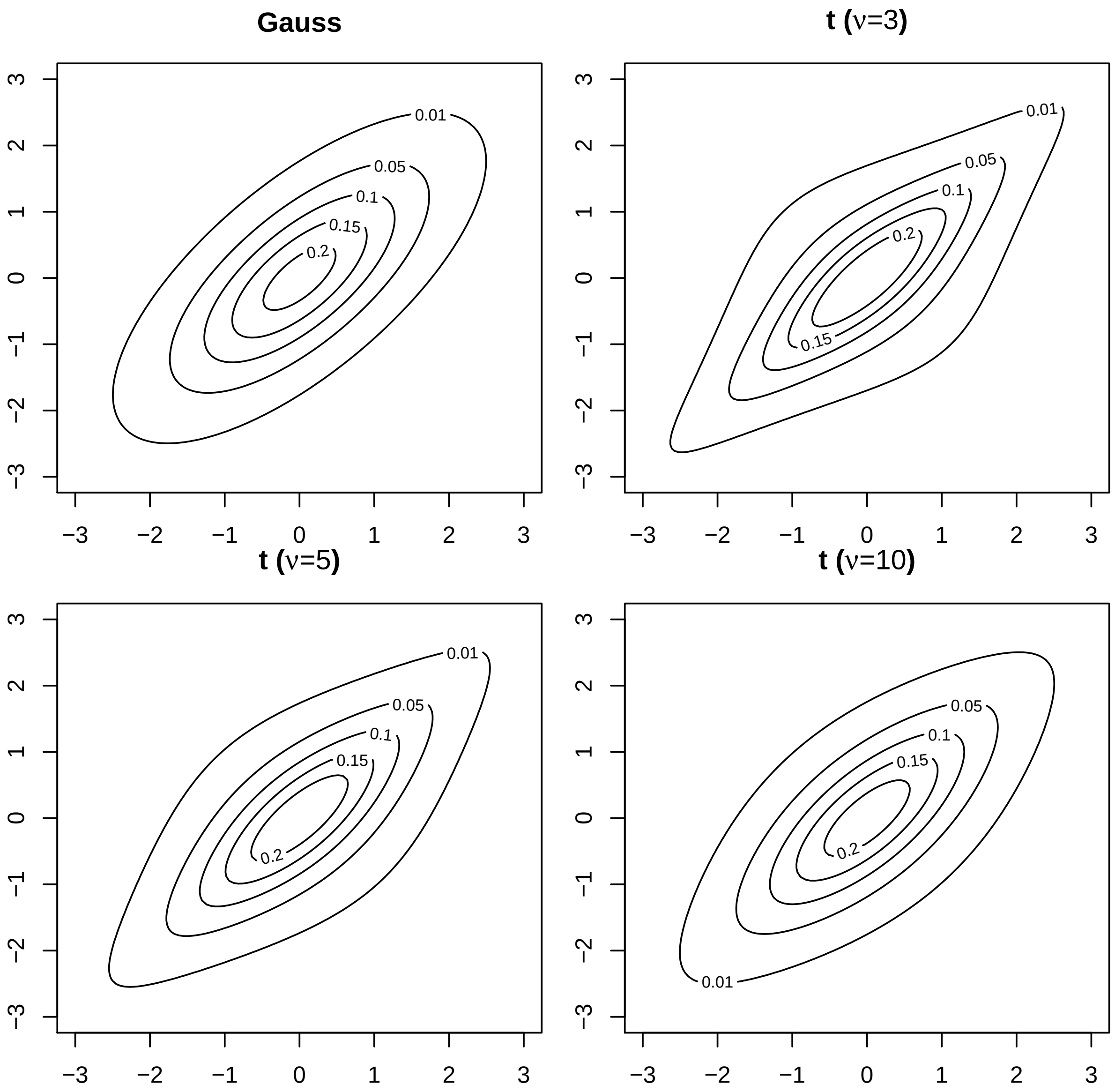


Figure 1.2: Bivariate contour plots of Gauss (top left), Student's t with degrees-of-freedom parameter $\nu = 3$ (top right), Student's t with $\nu = 5$ (bottom left) and Student's t with $\nu = 10$ (bottom right) with standard normal margins and a Kendall's τ value of 0.5.

Thus, we require second derivatives of copula log-likelihoods in order to study the covariance structure of ML estimators when copulas are involved. In particular, they are required for bivariate copulas because of their prominent role in PCCs and applications.

1.3.1 Archimedean copulas

In the Archimedean case, all functions of interest are given in closed form such that also computer algebra systems as Maple[2] or Mathematica[3] can be used for the calculation of derivatives. If available we will also state the expected Fisher information. For the sake of notational shortness, we will often suppress function arguments in the following.

Since we will need the first and second derivative of the density as well as of h-function in the forthcoming chapters, we state in the first example the derivatives with respect to the copula parameter and the copula arguments, too.

Example 1.3 (Clayton/MTCJ copula; continue of Example 1.1)
For the Clayton copula we can state the Fisher information with respect to θ as

$$\mathcal{I}(\theta) = \frac{1}{\theta^2} + \frac{2}{\theta((\theta-1)(2\theta-1)} + \frac{4\theta}{3\theta-2} - \frac{2(2\theta-1)}{\theta-1}\rho(\theta),$$

with

$$\begin{aligned}\rho(\theta) = &\frac{1}{(3\theta-2)(2\theta-1)} \\ &+ \frac{\theta}{2(3\theta-2)(2\theta-1)(\theta-1)}\left[\Psi_1\left(\frac{1}{2(\theta-1)}\right) - \Psi_1\left(\frac{\theta}{2(\theta-20}\right)\right] \\ &+ \frac{1}{2(3\theta-2)(2\theta-1)(\theta-1)}\left[\Psi_1\left(\frac{\theta}{2(\theta-1)}\right) - \Psi_1\left(\frac{2\theta-1}{2(\theta-20}\right)\right],\end{aligned}$$

where $\Psi(\cdot)$ is the trigamma function (Oakes 1982).

Derivatives of the density function
The partial derivative of the density c with respect to the parameter θ is

$$\begin{aligned}\frac{\partial c}{\partial\theta} &= (u_1u_2)^{-\theta-1}(u_1^{-\theta}+u_2^{-\theta}-1)^{-2-\frac{1}{\theta}} - (1+\theta)(u_1u_2)^{-\theta-1}\ln(u_1u_2)(u_1^{-\theta}+u_2^{-\theta}-1)^{-2-\frac{1}{\theta}} \\ &+ (1+\theta)(u_1u_2)^{-\theta-1}(u_1^{-\theta}+u_2^{-\theta}-1)^{-2-\frac{1}{\theta}} \\ &\left(\frac{\ln(u_1^{-\theta}+u_2^{-\theta}-1)}{\theta^2} + \frac{(-2-\frac{1}{\theta})(-u_1^{-\theta}\ln(u_1)-u_2^{-\theta}\ln(u_2))}{u_1^{-\theta}+u_2^{-\theta}-1}\right) \\ &= -c(u_1,u_2)\left(\ln(u_1u_2) - \left(\frac{\ln(A(u_1,u_2,\theta))}{\theta^2} + \frac{(-2-\frac{1}{\theta})(-u_1^{-\theta}\ln(u_1)-u_2^{-\theta}\ln(u_2))}{A(u_1,u_2,\theta)}\right)\right),\end{aligned}$$

and the derivative with respect to u_1 is

$$\begin{aligned}\frac{\partial c}{\partial u_1} &= (1+\theta)(u_1u_2)^{-\theta-1}(-\theta-1)\left({u_1}^{-\theta}+{u_2}^{-\theta}-1\right)^{-2-\theta^{-1}}{u_1}^{-1} - (1+\theta)(u_1u_2)^{-\theta-1} \\ &\left({u_1}^{-\theta}+{u_2}^{-\theta}-1\right)^{-2-\theta^{-1}}\left(-2-\theta^{-1}\right){u_1}^{-\theta}\theta{u_1}^{-1}\left({u_1}^{-\theta}+{u_2}^{-\theta}-1\right)^{-1} \\ &= -\frac{c(u_1,u_2)\cdot(\theta+1)}{u_1} + \frac{c(u_1,u_2)\cdot\left(2+\frac{1}{\theta}\right)\theta}{u_1^{\theta+1}\cdot A(u_1,u_2,\theta)}.\end{aligned}$$

[2]Maple 13.0. Maplesoft, a division of Waterloo Maple Inc., Waterloo, Ontario.
[3]Wolfram Research, Inc., Mathematica, Version 8.0, Champaign, IL (2012).

Derivatives of the h-function
We calculate the derivatives of the h-function with respect to the copula parameter θ and u_2, respectively:

$$\begin{aligned}
\frac{\partial h}{\partial \theta} &= -u_2^{-\theta-2}(u_1^{-\theta}+u_2^{-\theta}-1)^{-\frac{\theta+1}{\theta}}\theta - u_2^{-\theta-2}(u_1^{-\theta}+u_2^{-\theta}-1)^{-\frac{\theta+1}{\theta}} \\
&+ u_2^{-2\theta-2}(u_1^{-\theta}+v^{-\theta}-1)^{-\frac{2\theta+1}{\theta}}\theta + u_2^{-2\theta-2}(u_1^{-\theta}+u_2^{-\theta}-1)^{-\frac{2\theta+1}{\theta}} \\
&= \left(-\frac{h}{u_2} + u_2^{-2\theta-2}A(u_1,u_2,\theta)^{-\frac{2\theta+1}{\theta}}\right)(\theta+1), \\
\frac{\partial h}{\partial u_2} &= (-\theta-1)u_2^{-\theta-2}A(u_1,u_2,\theta)^{-1-\frac{1}{\theta}} + (-1-\frac{1}{\theta})A(u_1,u_2,\theta)^{-2-\frac{1}{\theta}}\frac{\partial A}{\partial u_2},
\end{aligned}$$

where

$$\frac{\partial A}{\partial u_1} = -\theta u_1^{-\theta-1}.$$

Second derivatives of the density function
For the following, we use the partial derivative of $A(u_1,u_2,\theta)$ to shorten the notation.

$$\begin{aligned}
\frac{\partial A}{\partial \theta} &= -u_1^{-\theta}\ln(u_1) - u_2^{-\theta}\ln(u_2), &\qquad \frac{\partial^2 A}{\partial^2 \theta} &= -u_1^{-\theta}\ln(u_1)^2 - u_2^{-\theta}\ln(u_2)^2, \\
\frac{\partial A}{\partial u_1} &= -\theta u_1^{-\theta-1}, & \frac{\partial^2 A}{\partial^2 u_1} &= \theta(\theta+1)u_1^{-\theta-2}
\end{aligned}$$

The second partial derivative of the Clayton copula density with respect to θ is

$$\begin{aligned}
\frac{\partial^2 c}{\partial^2 \theta} &= \frac{\partial c}{\partial \theta}\cdot\left(-\ln(u_2) + \frac{\ln(A(u_1,u_2,\theta))}{\theta^2} + \frac{\left(-2-\frac{1}{\theta}\right)\frac{\partial A}{\partial \theta}}{A(u_1,u_2,\theta)}\right) \\
&+ c(u_1,u_2)\cdot\left(\frac{\frac{\frac{\partial A}{\partial \theta}}{A(u_1,u_2,\theta)}\theta^2 - 2\ln(A(u_1,u_2,\theta))\theta}{\theta^4}\right. \\
&+ \left.\frac{\left(\frac{1}{\theta^2}\frac{\partial A}{\partial \theta} + (-2-\frac{1}{\theta})\frac{\partial^2 A}{\partial^2 \theta}\right)\cdot A(u_1,u_2,\theta) - (-2-\frac{1}{\theta})\left(\frac{\partial A}{\partial \theta}\right)^2}{A(u_1,u_2,\theta)^2}\right).
\end{aligned}$$

Further, the partial derivative with respect to u_1 is

$$\begin{aligned}
\frac{\partial^2 c}{\partial^2 u_1} &= -\frac{\frac{\partial c}{\partial u_1}(\theta+1)u_1 - (\theta+1)c(u_1,u_2)}{u_1^2} \\
&+ \frac{(2+\frac{1}{\theta})\left(\frac{\partial c}{\partial u_1}\frac{\partial A}{\partial u_1} + c(u_1,u_2)\frac{\partial^2 A}{\partial^2 u_1}\right) - c(u_1,u_2)(2+\frac{1}{\theta})\left(\frac{\partial^2 A}{\partial^2 u_1}\right)^2}{A(u_1,u_2,\theta)^2},
\end{aligned}$$

and finally, we obtain the derivatives with respect to u_1, and θ and u_2, respectively.

$$\frac{\partial^2 c}{\partial u_1 \partial \theta} = -\frac{\frac{\partial c}{\partial \theta}(\theta+1) + c(u_1,u_2)}{u_1} + \frac{\left(u_1^{\theta+1} A(u_1,u_2,\theta)\right)\left[\frac{\partial c}{\partial \theta}(2\theta+1) + 2c(u_1,u_2)\right]}{u_1^{2\theta+2} A(u_1,u_2,\theta)^2}$$
$$- \frac{c(u_1,u_2)(2\theta+1)\left[u_1^{\theta+1}\ln(u_1)A(u_1,u_2,\theta) + u_1^{\theta+1}\frac{\partial A}{\partial \theta}\right]}{u_1^{2\theta+2}A(u_1,u_2,\theta)^2},$$
$$\frac{\partial^2 c}{\partial u_1 \partial u_2} = -\frac{\frac{\partial c}{\partial u_2}(\theta+1)}{u_1} + \frac{\frac{\partial c}{\partial u_2}(2\theta+1)}{u_1^{\theta+1}A(u_1,u_2,\theta)} - \frac{c(u_1,u_2)(2\theta+1)}{u_1^{2\theta+2}A(u_1,u_2,\theta)^2} \cdot \frac{\partial A}{\partial u_2}.$$

Second derivatives of the h-function
The second partial derivatives of the conditional distribution function with respect to θ and u_2 are given as follows,

$$\frac{\partial^2 h}{\partial^2 \theta} = \frac{\partial h}{\partial \theta} \cdot \left(-\ln(u_2) + \frac{\ln(A(u_1,u_2,\theta))}{\theta^2} + \frac{\left(-1-\frac{1}{\theta}\right)\frac{\partial A}{\partial \theta}}{A(u_1,u_2,\theta)}\right)$$
$$+ h(u_1,u_2) \cdot \left(\frac{\frac{\frac{\partial A}{\partial \theta}}{A(u_1,u_2,\theta)}\theta^2 - 2\ln(A(u_1,u_2,\theta))\theta}{\theta^4}\right.$$
$$\left. + \frac{\left(\frac{1}{\theta^2}\frac{\partial A}{\partial \theta} + \left(-1-\frac{1}{\theta}\right)\frac{\partial^2 A}{\partial^2 \theta}\right) \cdot A(u_1,u_2,\theta) - \left(-1-\frac{1}{\theta}\right)\left(\frac{\partial A}{\partial \theta}\right)^2}{A(u_1,u_2,\theta)^2}\right),$$

$$\frac{\partial^2 h}{\partial^2 u_2} = (-\theta-1)(-\theta-2)u_2^{-\theta-3}A(u_1,u_2,\theta)^{-1-\frac{1}{\theta}}$$
$$+ (-\theta-1)u_2^{-\theta-1}\left(-1-\frac{1}{\theta}\right)A(u_1,u_2,\theta)^{-2-\frac{1}{\theta}}\frac{\partial A}{\partial u_2}$$
$$+ \left(-1-\frac{1}{\theta}\right)\left(-2-\frac{1}{\theta}\right)A(u_1,u_2,\theta)^{-3-\frac{1}{\theta}}\left(\frac{\partial A}{\partial u_2}\right)^2$$
$$+ \left(-1-\frac{1}{\theta}\right)A(u_1,u_2,\theta)^{-2-\frac{1}{\theta}}\frac{\partial^2 A}{\partial^2 u_2},$$
$$\frac{\partial^2 h}{\partial \theta \partial u_2} = (\theta+1)\left(-\frac{\frac{\partial h}{\partial u_2}u_2 - h(u_1,u_2)}{u_2^2} + (-2\theta-2)u_2^{-2\theta-3}A(u_1,u_2,\theta)^{-2-\frac{1}{\theta}}\right.$$
$$\left. + u_2^{-2\theta-2}\left(-2-\frac{1}{\theta}\right)A(u_1,u_2,\theta)^{-3-\frac{1}{\theta}}\frac{\partial A}{\partial u_2}\right).$$

Example 1.4 (Gumbel copula; continue of Example 1.2)
Using that the Fisher information with respect to dependence parameters does not depend on the marginal distributions (Smith 2007), the Fisher information can be determined

using the corresponding distribution with Weibull marginals (Oakes and Manatunga 1992),

$$\mathcal{I}(\theta) = \frac{1}{\theta^4}\left[\theta^2\left(-\frac{2}{3}+\frac{\pi^2}{9}\right) - \theta + \frac{2K_0}{\theta} + \left(\theta^3+\theta^2+(K_0-1)\theta - 2K_0 + \frac{K_0}{\theta}\right)E_1(\theta-1)e^{\theta-1}\right],$$

where $K_0 = \left(\frac{5}{6} - \frac{\pi^2}{18}\right)$ *and* $E_1(\theta) = \int_\theta^\infty \frac{1}{u}e^{-u}du$ *(Abramowitz and Stegun 1992, Exponential Integral, p. 228).*

1.3.2 Elliptical copulas

Since the information with respect to the dependence parameters is independent of the marginal distributions (Smith 2007), the Fisher information of the bivariate Gauss copula with respect to ρ is the same as for the bivariate Gaussian distribution (Berger and Sun 2008), i.e.

$$\mathcal{I}(\rho) = \frac{1+\rho^2}{(1-\rho^2)^2}.$$

But for the Student's t-copula the calculations for the Fisher information is much more involved and can only be given in closed form with respect to the correlation parameter, i.e. only $\mathcal{I}_\rho$ of

$$\mathcal{I}(\rho,\nu) = \begin{pmatrix} \mathcal{I}_\rho & \mathcal{I}_{\rho\nu} \\ \mathcal{I}_{\rho\nu} & \mathcal{I}_\nu \end{pmatrix}$$

can be calculated. To do so we will consider the corresponding elliptical distribution. The bivariate Student's t-distribution is a special case of the bivariate Pearson type VII distribution for which the following theorem holds.

Theorem 1.1 (Fisher Information of the Pearson VII distribution)
Let X and Y be two random variables with joint Pearson VII distribution with parameters $N > 1, m > 0, \rho \in (-1, 1)$ and density

$$f(x,y;\rho,N,m) = \frac{N-1}{\pi m\sqrt{1-\rho^2}}\left(1+\frac{x^2+y^2-2\rho xy}{m(1-\rho^2)}\right)^{-N}.$$

Then the elements of the Fisher information matrix with respect to (ρ,ν) are

$$\mathcal{I}(\rho,\nu) = \begin{pmatrix} \mathcal{I}_\rho & \mathcal{I}_{\rho m} \\ \mathcal{I}_{\rho m} & \mathcal{I}_m \end{pmatrix}$$

$$1)\ \mathcal{I}_\rho = \frac{1+\rho^2}{(1-\rho^2)^2} + \frac{N(N-1)\rho^2}{(1-\rho^2)^2} B(3, N-1)$$
$$+ \frac{N(N-1)(2-3\rho^2+\rho^6)}{4(1-\rho^2)^4} B(3, N-1) - \frac{2N(N-1)(1+\rho^2)}{(1-\rho^2)^2} B(2, N-1)$$
$$2)\ \mathcal{I}_m = \frac{2N(N-1)B(2, N-1) + N(1-N)B(3, N-1) - 1}{m^2}$$
$$3)\ \mathcal{I}_{\rho m} = \frac{N(N-1)[B(3, N-1) - B(2, N-1)]\rho}{m(1-\rho^2)},$$

where $B(\alpha, \beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1}dt$ *is the Beta-function.*

Proof
This proof is based on an earlier attempt by Nadarajah (2006).
1) Fisher information with respect to ρ*:*
For the transformed variables U *and* V*, corresponding to* $\rho = 0$*,*

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \sqrt{1+\rho} + \sqrt{1-\rho} & \sqrt{1+\rho} - \sqrt{1-\rho} \\ \sqrt{1+\rho} - \sqrt{1-\rho} & \sqrt{1+\rho} + \sqrt{1-\rho} \end{pmatrix} \begin{pmatrix} U \\ V \end{pmatrix},$$

the moments are given as

$$E_N[U^p V^q] = \frac{m^{(p+q)/2}(N-1)}{\pi} B\left(\frac{p+q}{2} + 1, N - \frac{p+q}{2} - 1\right) B\left(\frac{p+1}{2}, \frac{q+1}{2}\right), \quad (1.1)$$

if both p *and* q *are even integers (Nadarajah 2006). If either* p *or* q *is odd, Expression (1.1) is equal to zero.*

For the calculation of the Fisher Information of the Pearson VII distribution with respect to ρ *we need the second partial derivative of the log-density with respect to* ρ*.*

$$\frac{\partial \log L}{\partial \rho} = \frac{\rho}{1-\rho^2} - \frac{N}{\left(1 + \frac{x^2+y^2-2\rho xy}{m(1-\rho^2)}\right)} \cdot \frac{-2xy(1-\rho^2) + 2\rho(x^2+y^2-2\rho xy)}{m(1-\rho^2)^2},$$
$$\frac{\partial^2 \log L}{\partial^2 \rho} = \frac{1+\rho^2}{(1-\rho^2)^2} - \left(1 + \frac{x^2+y^2-2\rho xy}{m(1-\rho^2)}\right)^{-1}$$
$$\cdot \left[\frac{2N(x^2+y^2-2\rho xy)}{m(1-\rho^2)^2} + \frac{8N(\rho x^2+\rho y^2-(1-\rho^2)xy)\rho}{m(1-\rho^2)^3}\right]$$
$$+ \left(1 + \frac{x^2+y^2-2\rho xy}{m(1-\rho^2)}\right)^{-2} \frac{4N(\rho x^2+\rho y^2-(1-\rho^2)xy)^2}{m^2(1-\rho^2)^4}$$

Using the equations in the Appendix of Nadarajah (2006) for $E[X^pY^q]$ *in terms of* $E[U^pV^q]$ *and Expression (1.1), we can determine the Fisher Information of Pearson VII distribu-*

tion with respect to ρ:

$$\begin{aligned}
\mathcal{I}_\rho &= E_N\left[-\frac{\partial^2 \log L}{\partial^2\rho}\right] = \frac{1+\rho^2}{(1-\rho^2)^2} - E_N\left[\left(1+\frac{x^2+y^2-2\rho xy}{m(1-\rho^2)}\right)^{-1}\right. \\
&\cdot \frac{2N((x^2+y^2-2\rho xy)(1-\rho^2)+4\rho(\rho x^2+\rho y^2-(1+\rho^2)xy)}{m(1-\rho^2)^3} \\
&\left. + \left(1+\frac{x^2+y^2-2\rho xy}{m(1-\rho^2)}\right)^{-2} \frac{4N\left(\rho x^2+\rho y^2-(1-\rho^2)xy\right)}{m^2(1-\rho^2)^4}\right] \\
&= \frac{1+\rho^2}{(1-\rho^2)^2} + E_{N+2}\left[\frac{4N(N-1)\left(\rho x^2+\rho y^2-(1-\rho^2)xy\right)}{(N+1)m^2(1-\rho^2)^4}\right] \\
&- E_{N+1}\left[\frac{2(N-1)((x^2+y^2-2\rho xy)(1-\rho^2)+4\rho(\rho x^2+\rho y^2-(1+\rho^2)xy)}{m(1-\rho^2)^3}\right] \\
&= \frac{1+\rho^2}{(1-\rho^2)^2} + \frac{N(N-1)\rho^2}{(1-\rho^2)^2}B(3,N-1) + \frac{N(N-1)(2-3\rho^2+\rho^6)}{4(1-\rho^2)^4}B(3,N-1) \\
&- \frac{2N(N-1)(1+\rho^2)}{(1-\rho^2)^2}B(2,N-1),
\end{aligned}$$

where $B(a,b)$ is the Beta-function.

2) and 3) of Theorem 1.1: Nadarajah (2006, p.198-200) □

Corollary 1.2 (Fisher Information of the Student's t-distribution)
Let X and Y be two random variables with density

$$f(x,y;\rho,\nu) = \frac{1}{2\pi\sqrt{1-\rho^2}}\left(1+\frac{x^2+y^2-2\rho xy}{\nu(1-\rho^2)}\right)^{-\frac{\nu+2}{2}}$$

the bivariate Student's t-distribution. The elements of the Fisher information matrix are

$$\begin{aligned}
&\text{1) } \mathcal{I}_\rho = \frac{1+\rho^2}{(1-\rho^2)^2} + \frac{(\nu^2+2\nu)\rho^2}{4(1-\rho^2)^2}B\left(3,\frac{\nu}{2}\right) \\
&\qquad + \frac{(\nu^2+2\nu)(2-3\rho^2+\rho^6)}{16(1-\rho^2)^4}B\left(3,\frac{\nu}{2}\right) - \frac{(\nu^2+2\nu)(1+\rho^2)}{2(1-\rho^2)^2}B\left(2,\frac{\nu}{2}\right) \\
&\text{2) } \mathcal{I}_\nu = \frac{1}{\nu}B\left(2,\frac{\nu}{2}\right) - \frac{\nu+2}{4\nu}B\left(3,\frac{\nu}{2}\right) \\
&\text{3) } \mathcal{I}_{\rho\nu} = -\frac{\rho}{2(1-\rho^2)}\left(B\left(2,\frac{\nu}{2}\right) - \frac{\nu+2}{2}B\left(3,\frac{\nu}{2}\right)\right).
\end{aligned}$$

Proof
The t-distribution with parameters $\rho \in (-1,1)$ and $\nu > 0$ results from the Pearson VII distribution for $N = \frac{m+2}{2}$ and $m = \nu$. The log density is

$$l(x,y;\rho,\nu) = -\log(2\pi) - \frac{1}{2}\log(1-\rho^2) - \frac{\nu+2}{2}\log\left(1+\frac{x^2+y^2-2\rho xy}{\nu(1-\rho^2)}\right).$$

Thus, 1) follows directly from Theorem 1.1 1).

Since the Fisher Information for the Pearson VII distribution with respect to N, m *and* ρ *is known from Theorem 1.1, we can compute the elements for 2) and 3) of the Fisher Information matrix easily. Denoting the Fisher Information of Pearson VII by* $\mathcal{I}^{P7}$ *and the Fisher Information of the t-distribution by* $\mathcal{I}^t$ *we obtain*

$$\mathcal{I}^t_m = \mathcal{I}^{P7}_m + 2\mathcal{I}^{P7}_{mN} \underbrace{\left(\frac{\partial N}{\partial m}\right)}_{1/2} + \mathcal{I}^{P7}_N \underbrace{\left(\frac{\partial N}{\partial m}\right)^2}_{1/4} = \frac{1}{m} B\left(2, \frac{m}{2}\right) - \frac{m+2}{4m} B\left(3, \frac{m}{2}\right)$$

and

$$\mathcal{I}^t_{\rho m} = \mathcal{I}^{P7}_{\rho m} + \mathcal{I}^{P7}_{\rho N} \underbrace{\left(\frac{\partial N}{\partial m}\right)}_{1/2} = -\frac{\rho}{2(1-\rho^2)} \left(B\left(2, \frac{m}{2}\right) - \frac{m+2}{2} B\left(3, \frac{m}{2}\right)\right). \qquad \square$$

Here, $\mathcal{I}_m = E[-\frac{\partial^2 \log l}{\partial^2 m}], \mathcal{I}_{mN} = E[-\frac{\partial^2 \log l}{\partial N \partial m}], \mathcal{I}_N = E[-\frac{\partial^2 \log l}{\partial^2 N}], \mathcal{I}_{\rho m} = E[-\frac{\partial^2 \log l}{\partial \rho \partial m}]$ *and* $\mathcal{I}_{\rho N} = E[-\frac{\partial^2 \log l}{\partial \rho \partial N}]$.

Since ρ is a pure dependence parameter, the Fisher information of the t-copula is a the same as for the corresponding distribution.

Corollary 1.3 (Fisher information of the bivariate t-copula)
Let U_1 *and* U_2 *be two uniform distributed random variables distributed according to the bivariate t-copula* $C(u_1, u_2; \rho, \nu)$ *. Then the Fisher Information with respect to the association parameter* ρ *is*

$$\mathcal{I}_\rho = \frac{1+\rho^2}{(1-\rho^2)^2} + \frac{(\nu^2+2\nu)\rho^2}{4(1-\rho^2)^2} B\left(3, \frac{\nu}{2}\right) + \frac{(\nu^2+2\nu)(2-3\rho^2+\rho^6)}{16(1-\rho^2)^4} B\left(3, \frac{\nu}{2}\right) - \frac{(\nu^2+2\nu)(1+\rho^2)}{2(1-\rho^2)^2} B\left(2, \frac{\nu}{2}\right).$$

Proof
The result follows directly from Corollary 1.2 and Smith (2007). $\square$

For the Fisher Information with respect to the degrees-of-freedom ν and with respect to the mixed term ρ and ν no closed form expression is available. Thus, we provide a numerical solution (see Subsection 1.3.3).

1.3.3 Numerical issues and implementation in C

The derivatives of the bivariate copula models we have discussed are included in the R-package **VineCopula** (Schepsmeier et al. 2012, functions `BiCopDeriv`, `BiCopDeriv2`, `BiCopHfuncDeriv` and `BiCopHfuncDeriv2`). For details on input and output arguments of these functions we refer to the manual of the **VineCopula** package. In order to speed up the calculations, all implementations were done in C, using **R** (R Development Core Team 2013) for a convenient front end. Since the derivatives for the bivariate t-copula include derivatives of the (regularized) incomplete beta function (see details in the web supplement Schepsmeier and Stöber 2012), an efficient calculation of these is a key issue to achieve accuracy as well as fast computation times. For this, we employ the algorithm of Boik and Robinson-Cox (1998). Where this was possible, the C-code was optimized using the Computer Algebra Software Maple(TM).

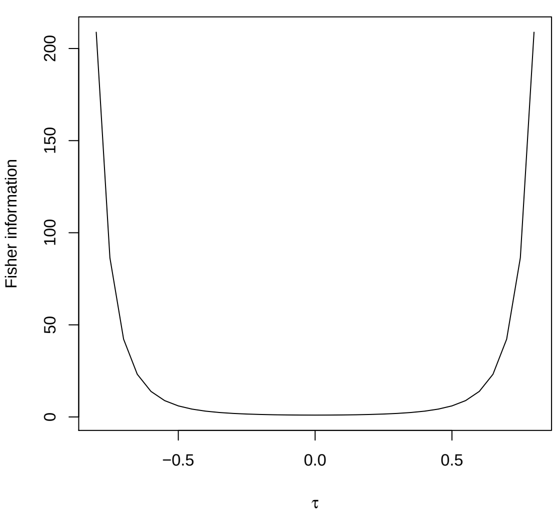
To obtain a benchmark for evaluating the numerical accuracy of our implementation, we consider the Fisher information which is analytically available for several families. We compare the known values with numerical results using the obtained derivatives of copula densities. For this, we employ the adaptive integration routines supplied in the **cubature** package (C code by Steven G. Johnson and R by Balasubramanian Narasimhan 2011), available on CRAN (based on Genz and Malik 1980 and Berntsen et al. 1991), with a maximum relative tolerance of $1e^{-5}$. The results show a maximum relative error in the order of $1e^{-6}$, confirming the accuracy of our implementation. The Fisher information with respect to the standard parametrization of different copula families is illustrated in Figure 1.3 (with corresponding values in Table 1.1) and Figure 1.4 (Table 1.2), respectively. For better comparison, the parameter values on the x-axis are transformed to the respective values of Kendall's τ. Note, that while the Fisher information is increasing with the absolute value of Kendall's τ for the Gaussian and Student's t-copula, the same is not true for the Archimedean families. This, however, is a mere consequence of the standard parametrization for these families. If we consider the Fisher information with respect to a parametrization in the form of Kendall's τ, the shapes look similar as for the Gaussian copula. This re-parametrization implies

$$\begin{aligned}
\mathcal{I}_\tau(\boldsymbol{\tau}) &= \mathbb{E}_{\boldsymbol{\tau}}\left[\Big(\frac{\partial}{\partial\tau_i} l(\boldsymbol{\tau}|\boldsymbol{X}) \cdot \frac{\partial}{\partial\tau_j} l(\boldsymbol{\tau}|\boldsymbol{X})\Big)_{i,j=1,\ldots,p}\right] \\
&= \mathbb{E}_{\boldsymbol{\theta}(\boldsymbol{\tau})}\left[\Big(\frac{\partial}{\partial\tau_i} l(\boldsymbol{\theta}(\boldsymbol{\tau})|\boldsymbol{X}) \cdot \frac{\partial}{\partial\tau_j} l(\boldsymbol{\theta}(\boldsymbol{\tau})|\boldsymbol{X})\Big)_{i,j=1,\ldots,p}\right] \\
&= \mathbb{E}_{\boldsymbol{\theta}(\boldsymbol{\tau})}\left[\Big(\frac{\partial}{\partial\theta_i} l(\boldsymbol{\theta}(\boldsymbol{\tau})|\boldsymbol{X}) \cdot \frac{\partial\theta_i}{\partial\tau_i} \cdot \frac{\partial}{\partial\theta_j} l(\boldsymbol{\theta}(\boldsymbol{\tau})|\boldsymbol{X}) \cdot \frac{\partial\theta_j}{\partial\tau_j}\Big)_{i,j=1,\ldots,p}\right] \\
&= \Big(\mathcal{I}_{\theta,i,j}(\boldsymbol{\theta}(\boldsymbol{\tau}))\frac{\partial\theta_i}{\partial\tau_i}\frac{\partial\theta_j}{\partial\tau_j}\Big)_{i,j=1,\ldots,p},
\end{aligned}$$

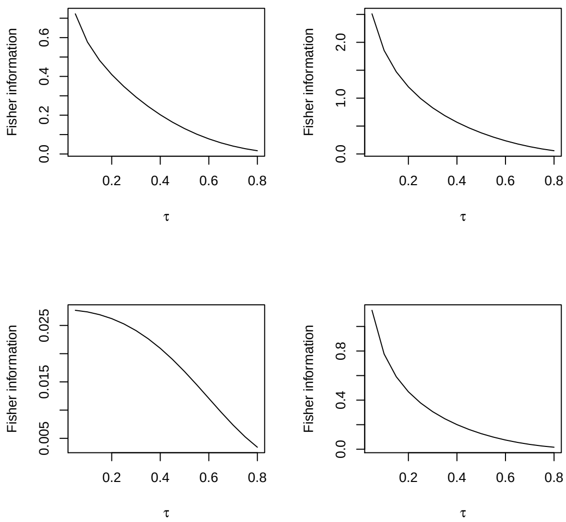
where $\mathcal{I}_{\theta,i,j}(\boldsymbol{\theta}(\boldsymbol{\tau}))$ is the i,j element of the Fisher information $\mathcal{I}_\theta(\boldsymbol{\theta}(\boldsymbol{\tau}))$ with respect to $\boldsymbol{\theta}$.

1.3.4 Example: Stock returns

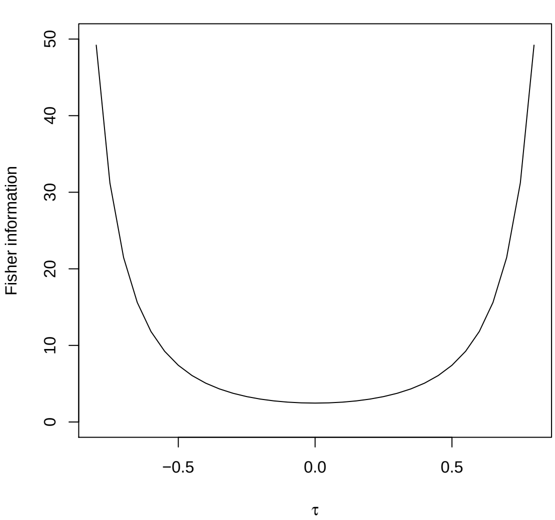
In this section we apply the developed methods to financial data available in the R-package **VineCopula** (Schepsmeier et al. 2012). We consider stock returns of Allianz AG (ALV) and Deutsche Bank (DBK) and conduct a rolling window analysis assuming a bivariate Student's t-copula for the dependence structure. We estimate the correlation as well as



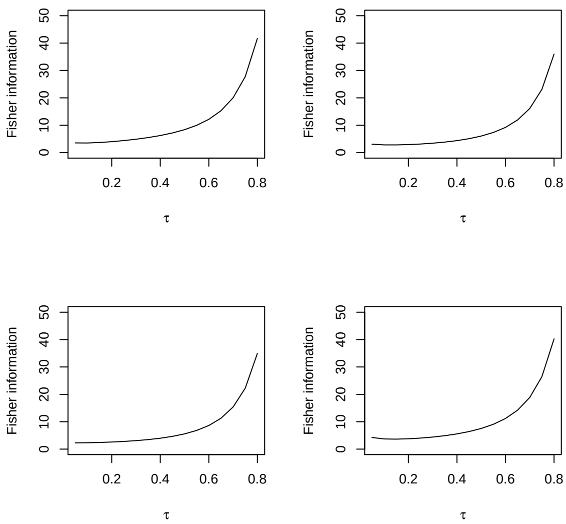
(a) Fisher information for the bivariate Gauss copula over a Kendall's τ range of $(-0.8, 0.8)$.



(b) Fisher information for the bivariate Archimedean copulas over a Kendall's τ range of $(0.05, 0.8)$. Top left: Clayton, top right: Gumbel, bottom left: Frank, bottom right: Joe.



(c) see above



(d) see above

Figure 1.3: Fisher information with respect to the standard parametrization (upper panel, (a)+(b)) and with respect to Kendall's τ (lower panel, (c)+(d)).

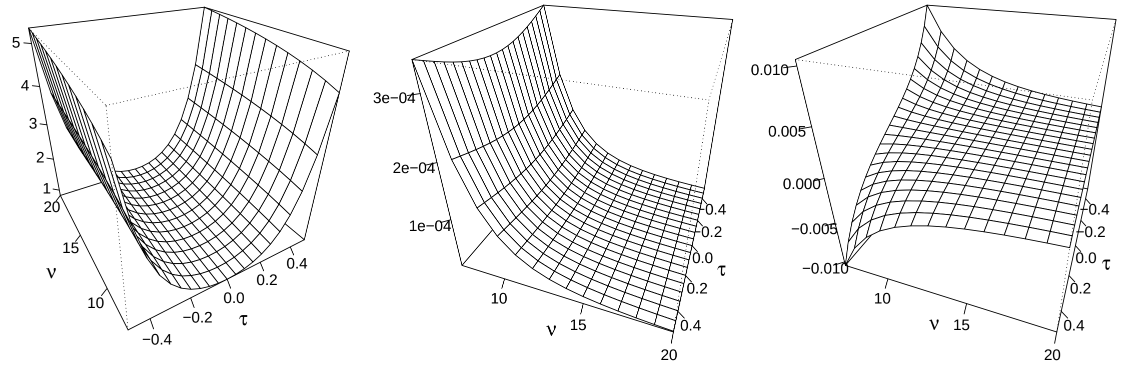


Figure 1.4: Fisher Information of the Student's t-copula over a Kendall's τ range of $(-0.5, 0.5)$ and degrees of freedom $\nu \in (7, 20)$ with respect to the correlation parameter ρ (left) the degrees of freedom (middle) and both (right).

τ	0.05	0.10	0.20	0.25	0.30	0.40	0.50	0.60	0.70	0.80
Gauss	2.50	2.59	2.99	3.31	3.75	5.07	7.40	11.82	21.48	49.21
Clayton	3.55	3.52	4.00	4.40	4.89	6.24	8.38	12.16	20.02	41.70
Gumbel	3.08	2.83	2.93	3.14	3.44	4.38	6.05	9.22	16.13	35.97
Frank	2.27	2.33	2.59	2.80	3.09	3.98	5.56	8.61	15.38	34.93
Joe	4.25	3.71	3.77	4.03	4.40	5.55	7.54	11.18	18.87	40.26

Table 1.1: Fisher information with respect to Kendall's τ for selected values of τ.

the degrees-of-freedom parameter and compare standard error estimates obtained from the observed and expected information.

Considering the time horizon from January 4th, 2005 to August 7th, 2009 we obtain 1158 daily log returns. Each time series is filtered using a GARCH(1,1) model with Student's t-innovations. A non-parametric rank transformation is applied to transform the residuals to uniformly distributed copula data. For the remainder, we assume the marginal distributions to be known and study the time variability only in the dependence model, i.e. the copula parameters, and not in the margins. This is done in a rolling windows analysis as follows: We move a window of size 100, 200 and 400 data points over the data set, respectively, and estimate the copula parameters for each subsample. Furthermore, the expected (see Corollary 1.3) and the observed Fisher Information are computed to estimate the standard errors $\sigma_{expexted}$ and $\sigma_{observed}$, respectively. The result is presented in Figure 1.5 where the solid line corresponds to the parameter estimates for each subsample and the dashed line to the overall maximum likelihood estimate (MLE) using all observations. For each estimate, the point-wise 95% confidence interval approximated by $[\hat{\theta}_{MLE} - 2\hat{\sigma}, \hat{\theta}_{MLE} + 2\hat{\sigma}]$ is given in dark grey for $\sigma_{observed}$ and light grey for $\sigma_{expected}$. For

ν	τ	0.05	0.10	0.20	0.25	0.30	0.40	0.50	0.60	0.70	0.80
3	$\partial^2\rho$	1.78	1.83	2.06	2.25	2.49	3.25	4.58	7.10	12.62	28.47
	$100 \cdot \partial^2\nu$	0.73	0.73	0.73	0.73	0.72	0.72	0.71	0.70	0.68	0.67
	$100 \cdot \partial\rho\partial\nu$	-0.38	-0.78	-1.66	-2.17	-2.77	-4.35	-6.85	-11.35	-20.95	-48.27
5	$\partial^2\rho$	1.94	2.00	2.27	2.48	2.77	3.66	5.21	8.15	14.59	33.08
	$100 \cdot \partial^2\nu$	0.11	0.11	0.11	0.11	0.11	0.11	0.12	0.12	0.12	0.12
	$100 \cdot \partial\rho\partial\nu$	-0.13	-0.26	-0.55	-0.73	-0.93	-1.46	-2.29	-3.80	-7.01	-16.16
10	$\partial^2\rho$	2.14	2.21	2.52	2.78	3.12	4.16	5.99	9.46	17.05	38.84
	$100 \cdot \partial^2\nu$	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	$100 \cdot \partial\rho\partial\nu$	-0.02	-0.05	-0.11	-0.14	-0.18	-0.28	-0.44	-0.73	-1.35	-3.12

Table 1.2: t-copula: Fisher information with respect to Kendall's τ for selected values of τ and degrees of freedom ν.

comparison, non-parametric bootstrap estimates are given (dotted).

Dependence between the bank and the insurer was not constant over the observation period. However, the variability is less significant than one might guess from an initial analysis since the MLE for the overall dataset is still within the 95% confidence band for most windows. In particular, our analysis does not support the assumption of time-varying degrees of freedom. The estimates obtained from the observed Fisher information yield results close to those obtained from the expected information. This supports their use for the routinely calculation of standard errors in practical applications. With increasing window size, also the estimates obtained from the non-parametric bootstrap coincide with those from the information matrix for large parts of the data. The deviation at the beginning and the end of the observation period, where the expected and observed information yield sharper bounds than the bootstrap, could be an indicator that the pure t-copula is misspecified. While we did not find strong evidence for time-varying parameters in our analysis, this suggest to investigate models where the copula family is allowed to switch (e.g. Stöber and Czado 2012) in further research.

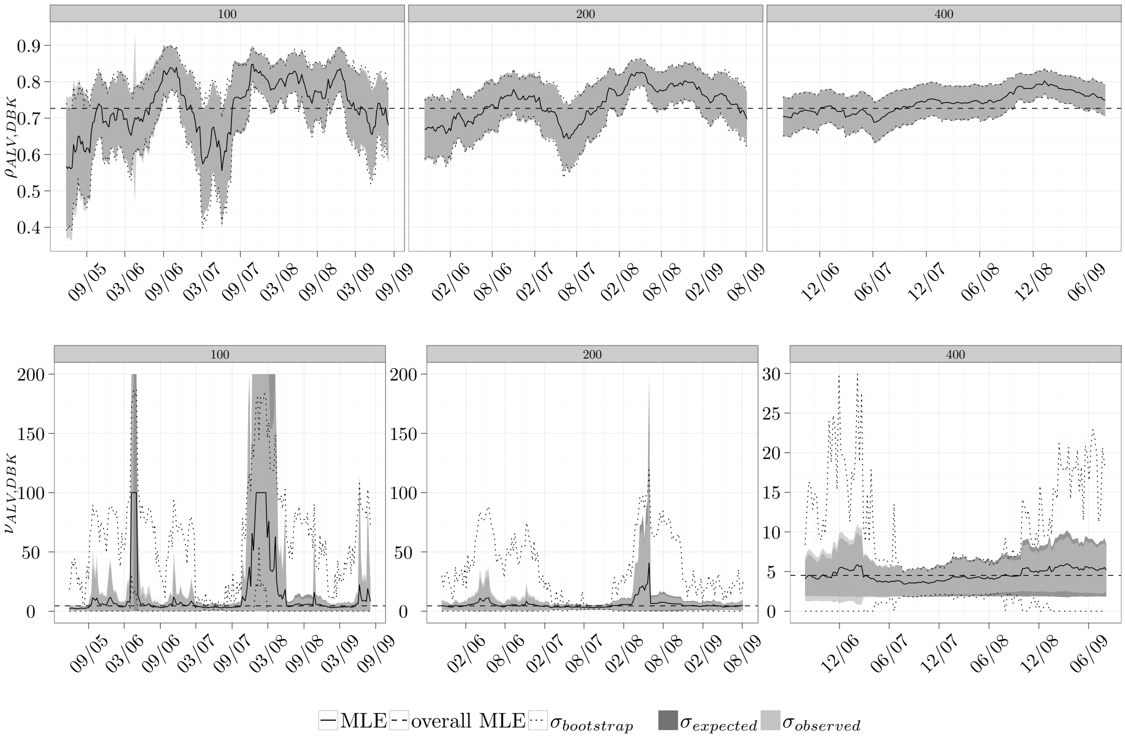


Figure 1.5: Rolling window analysis with window size 100 (left), 200 (middle) and 400 (right). The x-axis indicates the endpoint of each window (format: MM/YY), with the corresponding parameter estimate on the y-axis, i.e. the dependence parameter ρ (top panel) and the degrees-of-freedom parameter ν. The dashed horizontal line in each plot is the MLE corresponding to the whole data set (overall MLE).

Chapter 2

Regular vine copula models[1]

Since high dimensional problems and modeling issues become more and more relevant for practical purposes the concept of (multivariate) copulas becomes highly attractive for statisticians. The possibility to separate the modeling of the margins and the joint dependence structure, as discussed in Chapter 1, are very advantageous and allows for more flexibility.

But common multivariate parametric copulas such as the multivariate Gauss or Clayton copula are limited in their flexibility in higher dimensions. Different bivariate dependence properties such as tail dependence can not be captured. The pair-copula construction (PCC) circumvents these problems by making use of the highly flexible bivariate copulas as building blocks in a decomposition of the multivariate copula distribution into bivariate unconditional and conditional copulas. This main idea goes back to the paper of Joe (1996). He developed "a class of m-variate distributions with given [univariate] margins and $m(m-1)/2$ dependence parameters". Although he did not call it a PCC or a vine it was what we call a D-vine.

The second great advantage of Joe's construction was the integral free expression of the density corresponding to his distribution construction. This is possible since the calculation of the required conditional distribution functions as arguments of the conditional copula can be derived as derivative of a copula with respect to the second argument. This is that we will call later h-function. This property makes PCCs computational attractive compared to other (graphical) ways of dependence modeling such as (non-Gaussian) directed acyclic graphs (DAGs) (see for example Bauer et al. 2012).

The graphical illustration of the construction as a set of nested trees and the term "vine" goes back to the papers of Bedford and Cooke (2001, 2002). They developed the regular vine (R-vine) as a set of connected trees specifying the construction of the distribution. Their figure of a D-vine had the shape of a vine, giving the name of the class. Further, they derived a general expression for the density of a PCC distribution. In particular, they discussed the special case of a multivariate normal distribution formulated as a vine. Here, the partial correlations derived from the correlations matrix specify the vine copula parameters demanding bivariate Gaussian copulas as building blocks.

The statistical inference for PCCs was developed in Aas et al. (2009). Primarily the

[1]The contents of this chapter is based on J. Stoeber and U. Schepsmeier (2013), Estimating standard errors in regular vine copula models, forthcoming in Computational Statistics.

arbitrary selection of bivariate copula families as building blocks allows for high flexibility in modeling dependence structures in multivariate data sets. This is illustrated by them in an application to financial data preparing the basis for further applications (see for example Brechmann and Czado 2013). Since then the theory of vine copulas arising from the PCC were studied in the literature. Inference, selection methods and algorithms are for example investigated in Czado (2010), Min and Czado (2010, 2012), Brechmann et al. (2012) and Czado et al. (2012). The asymptotic theory we will use in Chapter 3 was developed and applied in Hobæk Haff (2013), Stöber and Schepsmeier (2013) and Dißmann et al. (2013), while extensions to discrete data or discrete-continuous data are proposed in Panagiotelis et al. (2012) and Stöber et al. (2012).

In the next section we will define PCCs and regular vines in detail and give in Section 2.2 the computation of the likelihood forming the basis for the subsequent chapters. In Section 2.3 the developed R-packages are presented containing the main functionality of R-vine inference, estimation, simulation and model selection.

2.1 Pair-copula constructions

Pair-copula constructions are a very flexible way to model multivariate distribution functions with bivariate copulas. The model is based on the decomposition of the d-dimensional density into $d(d-1)/2$ bivariate copula densities, including $d-1$ unconditional and $(d-2)(d-1)/2$ conditional ones. Bedford and Cooke (2001, 2002) introduced for the organization of the construction a set of nested trees $T_i = (V_i, E_i)$. Here V_i denotes the nodes while E_i represents the set of edges. This process was called by Aas et al. (2009) a pair-copula construction.

Let $F_1, \ldots, F_d$ be the marginal distribution functions of the random variables $X_1, \ldots, X_d$. The corresponding marginal density functions are denoted as $f_1, \ldots, f_d$. Following the notation of Czado (2010) with a set of bivariate copula densities $\mathcal{B} = \left\{c_{j(e),k(e)|D(e)} | e \in E_i, 1 \le i \le d-1\right\}$ corresponding to edges $j(e), k(e)|D(e)$ in E_i, for $1 \le i \le d-1$ a valid d-dimensional density can be constructed by setting

$$\begin{aligned} &f_{1,\ldots,d}(x_1, \ldots, x_d) = \\ &\prod_{k=1}^{d} f_k(x_k) \prod_{i=1}^{d-1} \prod_{e \in E_i} c_{j(e),k(e)|D(e)}(F_{j(e)|D(e)}(x_{j(e)}|\boldsymbol{x}_{D(e)}), F_{k(e)|D(e)}(x_{k(e)}|\boldsymbol{x}_{D(e)})). \end{aligned} \tag{2.1}$$

Here, $\boldsymbol{x}_{D(e)}$ is the subvector of $\boldsymbol{x}$ determined by the set of indices in $D(e)$, which is called *conditioning set* while the indices $j(e)$ and $k(e)$ form the *conditioned set*. The required conditional cumulative distribution functions (cdf) $F_{j(e)|D(e)}(x_{j(e)}|\boldsymbol{x}_{D(e)})$ and $F_{k(e)|D(e)}(x_{k(e)}|\boldsymbol{x}_{D(e)}))$ can be calculated as the first derivative of the corresponding copula cdf with respect to the second copula argument (see Joe 1996).

$$\begin{aligned} F(x_{j(e)}|\boldsymbol{x}_{D(e)}) &= \frac{\partial C_{j(e),j'(e)|D(e)\setminus j'(e)}(F(x_{j(e)}|\boldsymbol{x}_{D(e)\setminus j'(e)}), F(x_{j'(e)}|\boldsymbol{x}_{D(e)\setminus j'(e)}))}{\partial F(x_{j'(e)}|\boldsymbol{x}_{D(e)\setminus j'(e)})} \\ &=: h_{j(e),j'(e)|D(e)\setminus j'(e)}(F(x_{j(e)}|\boldsymbol{x}_{D(e)\setminus j'(e)}), F(x_{j'(e)}|\boldsymbol{x}_{D(e)\setminus j'(e)})) \end{aligned} \tag{2.2}$$

For regular vines there is an index $j'(e)$ in the conditioning set of indices given by edge e, such that the copula $C_{j(e),j'(e);D(e)\setminus j'(e)}$ is in the set of pair-copulas $\mathcal{B}$ (Dißmann et al.

2013). Doing so, we assume that the copula $C_{j(e),j'(e);D(e)\backslash j'(e)}$ does not depend on the values $\boldsymbol{x}_{D(e)\backslash j'(e)}$, i.e. on the conditioning set without the chosen variable $x_{j'}(e)$. This is called the *simplifying assumption*. In the literature Expression (2.4) is often denoted as a h-function. A PCC is called an R-vine copula if all marginal densities are uniform.

But it is often more convenient to define and work with copula data. Therefore, assume known margins $F_1, \ldots, F_d$ for the random variables $X_1, \ldots, X_d$ and base the notation on copula data $\boldsymbol{u} = (u_1, \ldots, u_d)$, with $u_i := F_i(x_i), i = 1, \ldots, d$. Here x_i is the observed value of X_i. If the margins $F_i, i = 1, \ldots, d$ are unknown, they can be estimated empirically. As proposed by Genest et al. (1995) these estimates can be used to transform the data $\boldsymbol{x}$ to an approximate sample $\boldsymbol{u} := (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$ in the copula space.

On the copula scale the vine copula density is defined as

$$c_{1,\ldots,d}(\boldsymbol{u}) = \prod_{i=1}^{d-1} \prod_{e \in E_i} c_{j(e),k(e);D(e)}(C_{j(e)|D(e)}(u_{j(e)}|\boldsymbol{u}_{D(e)}), C_{k(e)|D(e)}(u_{k(e)}|\boldsymbol{u}_{D(e)})). \tag{2.3}$$

and the corresponding h-function as

$$\begin{aligned} C(u_{j(e)}|\boldsymbol{u}_{D(e)}) &= \frac{\partial C_{j(e),j'(e);D(e)\backslash j'(e)}(C(u_{j(e)}|\boldsymbol{u}_{D(e)\backslash j'(e)}), C(u_{j'(e)}|\boldsymbol{u}_{D(e)\backslash j'(e)}))}{\partial C(u_{j'(e)}|\boldsymbol{u}_{D(e)\backslash j'(e)})} \\ &=: h_{j(e),j'(e);D(e)\backslash j'(e)}(C(u_{j(e)}|\boldsymbol{u}_{D(e)\backslash j'(e)}), C(u_{j'(e)}|\boldsymbol{u}_{D(e)\backslash j'(e)})). \end{aligned} \tag{2.4}$$

Sometimes the conditional copula $C(u|\boldsymbol{v})$ is also denoted as $F(u|\boldsymbol{v})$ since C is a distribution function. In the next section as well as in Chapter 3 we will use the F-notation for better distinction between the copula and its arguments. For example in the evaluation of the copula density $c_{2,3;1}(F(u_2|u_1), F(u_3|u_1))$ small c denotes the copula density and its arguments are denoted in the F-notation. Throughout this thesis we will work on the copula scale, i.e. assuming known margins.

Example 2.1 (3-dim pair-copula construction)
Let $x_1 \sim F_1, x_2 \sim F_2$ and $x_3 \sim F_3$. Then we have on the original scale

$$\begin{aligned} f(x_1,x_2,x_3) &= f_3(x_3)f_2(x_2)f_1(x_1) && \textit{(marginals)} \\ &\times c_{12}(F_1(x_1),F_2(x_2))c_{23}(F_2(x_2),F_3(x_3)) && \textit{(unconditional pairs)} \\ &\times c_{13;2}(F_{1|2}(x_1|x_2),F_{3|2}(x_3|x_2)|x_2) && \textit{(conditional pair)} \\ &\overset{\textit{simpli.}}{\underset{\textit{assump.}}{=}} f_3(x_3)f_2(x_2)f_1(x_1)c_{12}(F_1(x_1),F_2(x_2))c_{23}(F_2(x_2),F_3(x_3)) \\ &\times c_{13;2}(F_{1|2}(x_1|x_2),F_{3|2}(x_3|x_2)) \end{aligned}$$

The conditional cdfs such as $F_{1|2}(x_1|x_2)$ and $F_{3|2}(x_3|x_2)$ in the example can be recursively calculated using the h-function (see Expression (2.2)).

On the copula scale, i.e. let $u_1 = F_1(x_1), u_2 = F_2(x_2)$ and $u_3 = F_3(x_3)$, and call u_1, u_2 and u_3 copula data, the corresponding pair-copula construction under the simplifying assumption is

$$c_{123}(u_1,u_2,u_3) = c_{12}(u_1,u_2)c_{23}(u_2,u_3)c_{13;2}(C_{1|2}(u_1|u_2), C_{3|2}(u_3|u_2)).$$

where $C_{i,j}$ denotes the conditional distribution function of U_i given U_j.

Morales-Nápoles (2010) showed that there is a huge number of possible constructions. A set of nested trees is used to illustrate and order all these possible constructions. Each edge in a tree corresponds to a pair-copula in the PCC, while the nodes identify the pair-copula arguments. Bedford and Cooke (2001) formulated the following conditions, which a sequence of trees $\mathcal{V} = (T_1, \ldots, T_{d-1})$ has to fulfill to form an R-vine.

1. T_1 is a tree with nodes $V_1 = \{1, \ldots, d\}$ and edges E_1.
2. For $i \geq 2$, T_i is a tree with nodes $V_i = E_{i-1}$ and edges E_i.
3. If two nodes in T_{i+1} are joint by an edge, the corresponding edges in T_i must share a common node *(proximity condition)*.

In our notation we follow Dißmann et al. (2013) by denoting the vine structure with $\mathcal{V}$, the set of bivariate copulas with $\mathcal{B}(\mathcal{V})$ and the corresponding copula parameter with $\boldsymbol{\theta}(\mathcal{B}(\mathcal{V}))$. A specified regular vine copula we denote by $RV(\mathcal{V}, \mathcal{B}(\mathcal{V}), \boldsymbol{\theta}(\mathcal{B}(\mathcal{V})))$.

The special cases of an R-vine tree structure $\mathcal{V}$ are line like and star structures of the trees. The first one is called drawable vine (D-vine) in which each node has a maximum degree of 2, while the second one is a canonical vine (C-vine) with a root node of degree $d-1$. All other nodes, so called leafs, have degree 1. D-vines are for example studied in Aas et al. (2009) or Min and Czado (2010). An introduction to the statistical inference and model selection for C-vines are given for example in Czado et al. (2012). Examples of a five-dimensional C-vine and D-vine are illustrated in Figure 2.1.

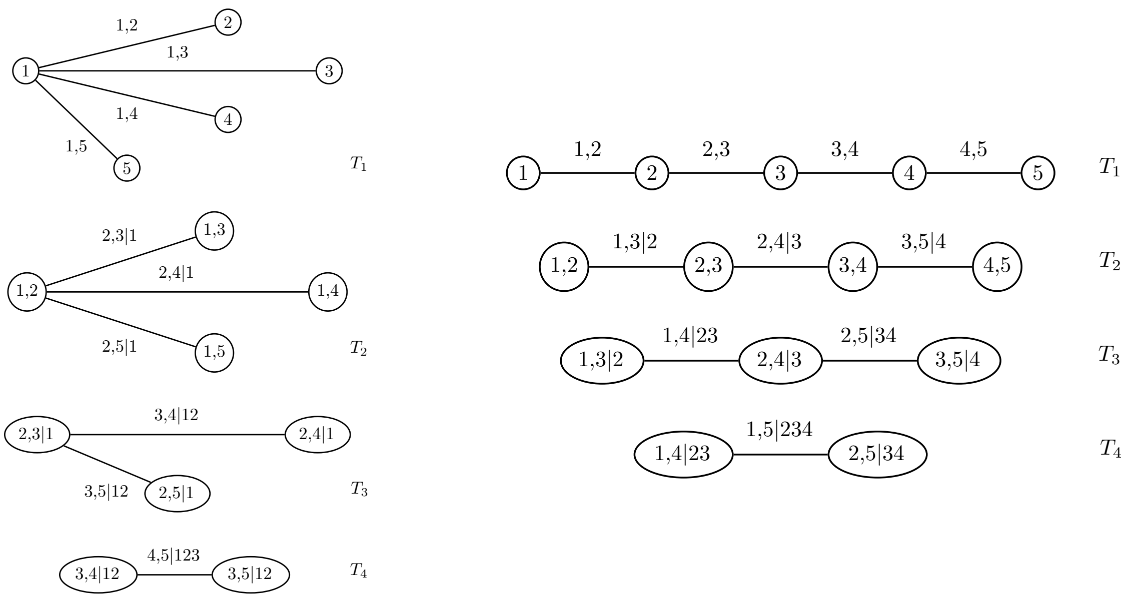


Figure 2.1: Examples of five-dimensional C- (left panel) and D-vine trees (right panel) with edge indices.

The pair-copula selection for an R-vine copula can be done by Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) or bivariate goodness-of-fit tests, while for the structure selection at least two algorithms are available in the literature.

Dißmann et al. (2013) favor a maximum spanning tree (MST) algorithm maximizing absolute Kendall's τ tree-wise, whereas Gruber and Czado (2012) follow a Bayesian approach. They propose an algorithm to select the tree structure as well as the pair-copula families in a regular vine copula model jointly in a tree-by-tree reversible jump MCMC approach. Figure 2.2 illustrates the tree structure of a 5-dimensional R-vine.

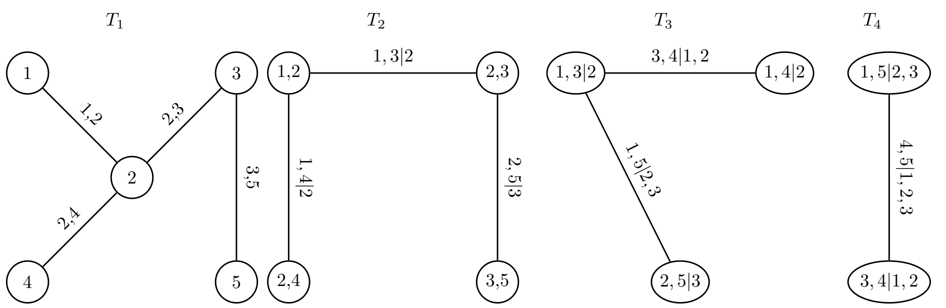


Figure 2.2: An R-vine tree sequence in 5 dimensions with edge indices corresponding to the pair-copulas in an R-vine copula model.

2.2 Computation of the R-vine likelihood

In order to calculate the (log-) likelihood function of an R-vine model, we have to develop an algorithmic way to evaluate the copula terms in the decomposition (2.3) with respect to the appropriate arguments. To simplify notation, we will assume that all copulas under investigation are symmetric in their arguments, such that we do not have to distinguish between the h-functions $h_{j(e),j'(e)|D(e)\setminus j'(e)}$ and $h_{j'(e),j(e)|D(e)\setminus j'(e)}$ given in Equation (2.4). While this is valid for most common parametric copula families, we can easily drop this assumption later. For calculation purposes it is convenient to use matrix notation which has been introduced by Morales-Nápoles et al. (2010), Dißmann (2010) and Dißmann et al. (2013). It stores the edges of an R-vine tree sequence in the following way: Consider the R-vine in Figure 2.2 which can be described in matrix notation as follows:

$$M = \begin{pmatrix} 5 & & & & \\ 4 & 4 & & & \\ 1 & 3 & 3 & & \\ 2 & 1 & 1 & 2 & \\ 3 & 2 & 2 & 1 & 1 \end{pmatrix}. \tag{2.5}$$

As an illustration for how the R-vine matrix is derived from the tree sequence in Figure 2.2 and vice versa, let us consider the second column of the matrix. Here we have 4 on the diagonal, and 3 as a second entry. The set of remaining entries below 3 is $\{1, 2\}$. This corresponds to the edge $4, 3|1, 2$ in T_3 of Figure 2.2. Similarly, the edge $4, 1|2$ corresponds to the third entry 1 in the third row and $4, 2$ to the fourth and last entry. Note that the diagonal of M is sorted in descending order which can always be achieved by reordering the node labels. From now on, we will assume that all matrices are "normalized" in this way as this allows to simplify notation. Therefore we have $m_{i,i} = d-i+1$. For the special cases of C-Vines and D-Vines, the matrices in 5 dimensions are

$$M_{C-Vine} = \begin{pmatrix} 5 & & & & \\ 4 & 4 & & & \\ 3 & 3 & 3 & & \\ 2 & 2 & 2 & 2 & \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \quad \text{and} \quad M_{D-Vine} = \begin{pmatrix} 5 & & & & \\ 1 & 4 & & & \\ 2 & 1 & 3 & & \\ 3 & 2 & 1 & 2 & \\ 4 & 3 & 2 & 1 & 1 \end{pmatrix}.$$

In the following, we will introduce some additional notation required to develop programmable algorithms for R-vine models. For simplicity, we consider the 5-dimensional example. Similar to the way the R-vine tree sequence is given in M, we can store the copula families $\mathcal{B}$ and the corresponding parameters $\boldsymbol{\theta}$.

$$\begin{aligned}\mathcal{B} &= \begin{pmatrix} \mathcal{B}_{5,m_{2,1}|m_{3,1},m_{4,1},m_{5,1}} & & & \\ \mathcal{B}_{5,m_{3,1}|m_{4,1},m_{5,1}} & \mathcal{B}_{4,m_{3,2}|m_{4,2},m_{5,2}} & & \\ \mathcal{B}_{5,m_{4,1}|m_{5,1}} & \mathcal{B}_{4,m_{4,2}|m_{5,2}} & \mathcal{B}_{3,m_{4,3}|m_{5,3}} & \\ \mathcal{B}_{5,m_{5,1}} & \mathcal{B}_{4,m_{5,2}} & \mathcal{B}_{3,m_{5,3}} & \mathcal{B}_{2,m_{5,4}} \end{pmatrix} \\ &= \begin{pmatrix} \mathcal{B}_{5,4|1,2,3} & & & \\ \mathcal{B}_{5,1|2,3} & \mathcal{B}_{4,3|1,2} & & \\ \mathcal{B}_{5,2|3} & \mathcal{B}_{4,1|2} & \mathcal{B}_{3,1|2} & \\ \mathcal{B}_{5,3} & \mathcal{B}_{4,2} & \mathcal{B}_{3,2} & \mathcal{B}_{2,1} \end{pmatrix} \in \mathbb{R}^{5\times 5}\end{aligned}$$

$$\begin{aligned}\boldsymbol{\theta} &= \begin{pmatrix} \theta_{5,m_{2,1}|m_{3,1},m_{4,1},m_{5,1}} & & & \\ \theta_{5,m_{3,1}|m_{4,1},m_{5,1}} & \theta_{4,m_{3,2}|m_{4,2},m_{5,2}} & & \\ \theta_{5,m_{4,1}|m_{5,1}} & \theta_{4,m_{4,2}|m_{5,2}} & \theta_{3,m_{4,3}|m_{5,3}} & \\ \theta_{5,m_{5,1}} & \theta_{4,m_{5,2}} & \theta_{3,m_{5,3}} & \theta_{2,m_{5,4}} \end{pmatrix} \\ &= \begin{pmatrix} \theta_{5,4|1,2,3} & & & \\ \theta_{5,1|2,3} & \theta_{4,3|1,2} & & \\ \theta_{5,2|3} & \theta_{4,1|2} & \theta_{3,1|2} & \\ \theta_{5,3} & \theta_{4,2} & \theta_{3,2} & \theta_{2,1} \end{pmatrix} \in \mathbb{R}^{5\times 5}\end{aligned}$$

To evaluate the (log-) likelihood function (Equation (2.3)) of an R-vine model, we require the conditional distributions $F_{j(e)|D(e)}$ and $F_{k(e)|D(e)}$, evaluated at a d-dimensional vector of observations $(u_1, \ldots, u_d)$, as arguments of the copula density $c_{j(e),k(e);D(e)}$ corresponding to edge e. We will store also these values in two matrices. In particular, we calculate

$$V^{direct} = \begin{pmatrix} \dots & & & & \\ \dots & F(u_4|u_{m_{3,2}}, u_{m_{4,2}}, u_{m_{5,2}}) & & & \\ \dots & F(u_4|u_{m_{4,2}}, u_{m_{5,2}}) & F(u_3|u_{m_{4,3}}, u_{m_{5,3}}) & & \\ \dots & F(u_4|u_{m_{5,2}}) & F(u_3|u_{m_{5,3}}) & F(u_2|u_{m_{5,4}}) & \\ \dots & u_4 & u_3 & u_2 & u_1 \end{pmatrix} \quad (2.6)$$

$$V^{indirect} = \begin{pmatrix} \dots & & & \\ \dots & F(u_{m_{3,2}}|u_{m_{4,2}}, u_{m_{5,2}}, u_4) & & \\ \dots & F(u_{m_{4,2}}|u_{m_{5,2}}, u_4) & F(u_{m_{4,3}}|u_{m_{5,3}}, u_3) & \\ \dots & F(u_{m_{5,2}}|u_4) & F(u_{m_{5,3}}|u_3) & F(u_{m_{5,4}}|u_2) \\ \dots & u_{m_{5,2}} & u_{m_{5,3}} & u_{m_{5,4}} \end{pmatrix}. \quad (2.7)$$

Note that for each pair-copula term $c_{j(e),k(e);D(e)}$ in (2.3) the corresponding terms of V^{direct} and $V^{indirect}$ involving $F_{j(e)|D(e)}$ and $F_{k(e)|D(e)}$ can be easily determined by applying (2.4). When being able to evaluate

$$c_{4,m_{3,2};m_{4,2},m_{5,2}}(F(u_4|u_{m_{4,2}}, u_{m_{5,2}}), F(u_{m_{3,2}}|u_{m_{4,2}}, u_{m_{5,2}}))$$

we do also obtain

$$\begin{aligned} F(u_4|u_{m_{6,5}}, u_{m_{7,5}}, u_{m_{8,5}}) &= (\partial_1 C)_{4,m_{3,2}|m_{4,2},m_{5,2}}(F(u_4|u_{m_{4,2}}, u_{m_{5,2}}), F(u_{m_{3,2}}|u_{m_{4,2}}, u_{m_{5,2}})), \\ F(u_{m_{3,2}}|u_4, u_{m_{4,2}}, u_{m_{5,2}}) &= (\partial_2 C)_{4,m_{3,2}|m_{4,2},m_{5,2}}(F(u_4|u_{m_{4,2}}, u_{m_{5,2}}), F(u_{m_{3,2}}|u_{m_{4,2}}, u_{m_{5,2}})), \end{aligned}$$

where $(\partial_1 C)$ and $(\partial_2 C)$ denote partial derivatives with respect to the first and second argument, respectively. With all such conditional distribution functions being available for e.g. T_3, the copula terms in (2.3) corresponding to the next tree T_4 can be evaluated. Following the notation in (2.4), we write $h(\cdot, \cdot|\mathcal{B}^{k,i}, \theta^{k,i})$ for the conditional distribution function corresponding to a parametric family $\mathcal{B}^{k,i}$ with parameter $\theta^{k,i}$, where $\mathcal{B}^{k,i}$ and $\theta^{k,i}$ denote the (k,i)th element of the matrices $\mathcal{B}$ and $\boldsymbol{\theta}$, respectively. Exempli gratia,

$$\begin{aligned} F(u_4|u_3, u_1, u_2) &= h(F(u_4|u_1, u_2), F(u_3|u_1, u_2)|\mathcal{B}_{4,3|1,2}, \theta_{4,3|1,2}) \\ &= h(v^{direct}_{3,2}, v^{direct}_{3,3}|\mathcal{B}^{6,5}, \theta^{6,5}). \end{aligned}$$

The only question which is left to solve for the computation of the likelihood is whether the arguments in each step (i.e. $F(u_4|u_1, u_2), F(u_3|u_1, u_2)$ in the example) have to be picked from the matrix V^{direct} or $V^{indirect}$. For this, we exploit the descending order of the diagonal of M. From the structure of V^{direct}, we see that the first argument of the copula term with family $\mathcal{B}^{k,i}$ and parameter $\theta^{k,i}$ is stored as the (k,i)th element $v^{\text{direct}}_{k,i}$ of V^{direct}. To locate the second entry, let us denote $\tilde{M} = (\tilde{m}_{k,i}|i = 1, \dots, d; k = i, \dots, d)$, where $\tilde{m}_{k,i} := \max\{m_{k,i}, \dots, m_{d,i}\}$ for all $i = 1, \dots, d$ and $k = i, \dots, d$. The second argument, which is $F(u_{m_{k,i}}|m_{i,i}, m_{k+1,i}, \dots, m_{d,i})$ must be in column $(d - \tilde{m}_{k,i} + 1)$ of V^{direct} or $V^{indirect}$ by the ordering of variables. If $\tilde{m}_{k,i} = m_{k,i}$, the conditioning variable $u_{m_{k,i}}$ has the biggest index and thus the entry we are looking for must be in V^{direct}. Similarly, if $\tilde{m}_{k,i} > m_{k,i}$, the variable with the biggest index is in the conditioning set and we must choose from $V^{indirect}$. In particular, for a C-Vine it always holds that $\tilde{m}_{k,i} = m_{k,i}$ and we choose from V^{direct} while for the D-vine $\tilde{m}_{k,i} > m_{k,i}$ and we must choose from $V^{indirect}$.

Example 2.2 (Selection of arguments for $c_{4,3|1,2}$)
As an example for how this procedure selects the correct arguments for copula terms in R-vine let us consider the copula $c_{4,3|1,2}$ in our example distribution. The corresponding parameter $\theta_{4,3|1,2}$ is stored as $\theta^{3,2}$, thus we are in the case where $i = 2$ and $k = 3$. Since $\tilde{m}_{3,2} = \max\{m_{3,2}, m_{4,2}, m_{5,2}\} = \max\{3, 1, 2\} = 3$ and $\tilde{m}_{3,2} = 3 = m_{3,2}$ we select as second argument the entry $v^{direct}_{k,(d-\tilde{m}_{k,i}+1)} = v^{direct}_{3,3} = F(u_3|u_1, u_2)$. Together with $v^{direct}_{3,2} = F(u_4|u_1, u_2)$, which we have already selected, this is the required argument.

These sequential selections and calculations are performed in Algorithm 2.2.1, which was developed in Dißmann (2010) and Dißmann et al. (2013).

Algorithm 2.2.1 Log-likelihood of an R-vine specification.

Require: d-dimensional R-vine specification in matrix form, i.e., M, $\mathcal{B}$, $\boldsymbol{\theta}$, set of observations $(u_1, \dots, u_d)$.

Set $L = 0$.
Let $V^{\text{direct}} = (v^{\text{direct}}_{k,i} | i = 1, \dots, d; k = i, \dots, d)$.
Let $V^{\text{indirect}} = (v^{\text{indirect}}_{k,i} | i = 1, \dots, d; k = i, \dots, d)$.
Set $(v^{\text{direct}}_{d,1}, v^{\text{direct}}_{d,2}, \dots, v^{\text{direct}}_{d,d}) = (u_d, u_{d-1}, \dots u_1)$.
Let $\tilde{M} = (\tilde{m}_{k,i} | i = 1, \dots, d; k = i, \dots, d)$ where $\tilde{m}_{k,i} = \max\{m_{k,i}, \dots, m_{d,i}\}$ for all $i = 1, \dots, d$ and $k = i, \dots, d$.
for $i = d-1, \dots, 1$ **do** {Iteration over the columns of M}
  for $k = d, \dots, i+1$ **do** {Iteration over the rows of M}
    Set $z_1 = v^{\text{direct}}_{k,i}$
    if $\tilde{m}_{k,i} = m_{k,i}$ **then**
      Set $z_2 = v^{\text{direct}}_{k,(d-\tilde{m}_{k,i}+1)}$.
    else
      Set $z_2 = v^{\text{indirect}}_{k,(d-\tilde{m}_{k,i}+1)}$.
    end if
    Set $L = L + c(z_1, z_2 | \mathcal{B}^{k,i}, \theta^{k,i})$.
    Set $v^{\text{direct}}_{k-1,i} = h(z_1, z_2 | \mathcal{B}^{k,i}, \theta^{k,i})$ and $v^{\text{indirect}}_{k-1,i} = h(z_2, z_1 | \mathcal{B}^{k,i}, \theta^{k,i})$.
  end for
end for
return L

2.3 Software

There are two packages **CDVine** and **VineCopula** publicly available for the selection, estimation and validation of vine copula models based on the statistical software **R** (R Development Core Team 2013) given on the Comprehensive **R** Archive Network (CRAN, http://cran.r-project.org/). While the R-package **CDVine** of Brechmann and Schepsmeier (2013) provides the functionality for the special classes of C- and D-vines, the R-package **VineCopula** of Schepsmeier et al. (2012) is available for R-vines.

Since vine copula models are based on bivariate copulas as building blocks, both packages contain tools for bivariate exploratory data analysis and inference of a wide range of bivariate copula families. In particular, we provide bivariate copula families from the two major classes of elliptical and Archimedean copulas as introduced in Chapter 1. Graphical tools as contour plots, λ-functions, K-plots or Chi-plots are included for copula selection. They assist the automatized Akaike information criterium (AIC) or Bayesian information criterium (BIC) based methods. Furthermore, copula selection or verification are possible with several implemented goodness-of-fit tests returning the corresponding test statistic and a (bootstrapped) p-value. Dependence measures such as Kendall's τ or tail dependence can be calculated, as well, to analyze and characterize bivariate copulas. It is important to know whether a particular parametric copula might be suitable for a given data set (Joe 1997).

But the main functionality is focused on vine copula models. The packages enable the user to facilitate vine copula inference. Parameter estimation and model selection as well as sampling algorithms are available. Therefore, the recursive algorithm of the log-likelihood is implemented in a fast and optimized manner for maximum likelihood estimation (MLE). A much faster but only approximating approach is sequential estimation (Aas et al. 2009), which is also included. It is typically used to generate starting values for the MLE. Furthermore, there is an extensive functionality for model verification and model comparison. The tests of Vuong (1989) and Clarke (2007), introduced for non-nested models, can be applied to vines as well as classical Akaike or Bayesian information criteria.

Recently, the possibility to calculate the Hessian matrix or the gradient are implemented thus providing estimated standard errors of the vine copula parameters as we will discuss in Chapter 3. Verification of a given or estimated vine copula model by goodness-of fit tests will also be available in the near future (see Chapter 4).

In the next two subsection we will discuss methods for bivariate data analysis and for statistical inference of R-vine copula models in more detail. For explicit documentation of each function we refer to the extensive manuals of **CDVine** and **VineCopula**, and the vignette of Brechmann and Schepsmeier (2013).

Bivariate data analysis

In both packages several bivariate copula families are included for bivariate analysis as well as for multivariate analysis using vine copulas. Beside the four mentioned Archimedean copulas, Clayton, Gumbel, Frank and Joe, the packages provide functionality for four Archimedean copula families with two parameters, namely the Clayton-Gumbel, the Joe-Gumbel, the Joe-Clayton and the Joe-Frank. Following Joe (1997) we simply refer to them as BB1, BB6, BB7 and BB8, respectively. Their more flexible structure allow for different non-zero lower and upper tail dependence coefficients. As boundary cases they include the Clayton and Gumbel, the Joe and Gumbel, the Joe and Clayton as well as the Joe and Frank copula, respectively.

In addition to these families, we also implemented rotated versions of the Clayton, Gumbel, Joe and the BB families. By rotating them by 180 degrees, one obtains the corresponding survival copulas. In contrast, rotation by 90 and 270 degrees allows for the modeling of negative dependence which is not possible with the standard non-rotated versions. In particular, the distribution functions C_{90}, C_{180} and C_{270} of a copula C rotated

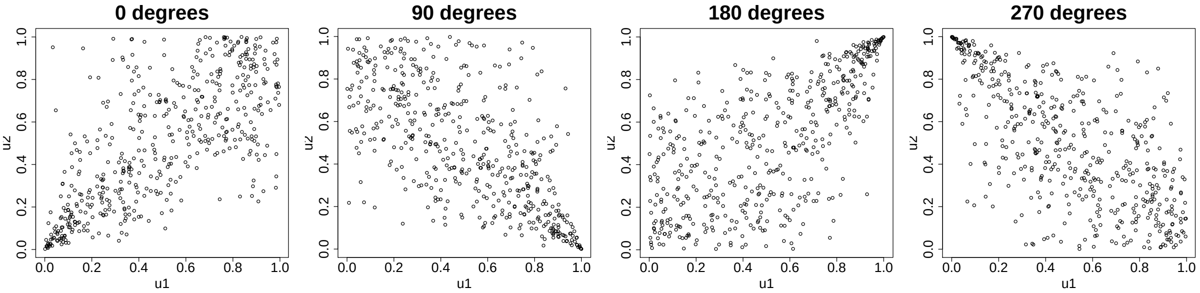


Figure 2.3: Samples from Clayton copulas rotated by 0, 90, 180 and 270 degrees with parameters corresponding to Kendall's τ values of 0.5 for positive dependence and -0.5 for negative dependence.

by 90, 180 and 270 degrees, respectively, are given as follows:

$$\begin{aligned} C_{90}(u_1,u_2) &= u_2 - C(1-u_1,u_2),\\ C_{180}(u_1,u_2) &= u_1+u_2-1+C(1-u_1,1-u_2),\\ C_{270}(u_1,u_2) &= u_1 - C(u_1,1-u_2). \end{aligned}$$

To illustrate rotated bivariate Archimedean copulas, we simulate samples of size $N = 500$ from Clayton copulas rotated by 0, 90, 180 and 270 degrees, respectively. Parameters are chosen according to Kendall's τ values of 0.5 for positive dependence and -0.5 for negative dependence. Corresponding scatter plots are shown in Figure 2.3. The first plot corresponds to the standard Clayton copula modeling positive dependence and lower tail dependence. The survival copula in the third plot also covers the positive Kendall's τ range put has upper tail dependence. The rotated versions of the Clayton copula in the second and forth plot model negative dependence. Again different tail behavior can be captured by rotation.

The main functions for exploratory data analysis, simulation, selection and estimation of bivariate copulas are:

- `BiCopMetaContour` gives a bivariate contour plot corresponding to a bivariate meta distribution with different margins and specified bivariate copula and parameter values or creates corresponding empirical contour plots based on bivariate copula data.

- `BiCopSim` simulates from a given parametric bivariate copula.

- `BiCopGofTest` calculates goodness-of fit tests for bivariate copulas, either based on White's information matrix equality (White 1982) as introduced by Huang and Prokhorov (2013) or based on Kendall's process (Genest et al. 2009). It computes the test statistics and calculates bootstrapped p-values.

- `BiCopSelect` selects an appropriate bivariate copula family for given bivariate copula data using one of a range of methods. As selection criterion either AIC or BIC can be used.

- `BiCopEst` estimates the parameter(s) for a bivariate copula using either inversion of empirical Kendall's τ for single parameter copula families or maximum likelihood estimation for one and two parameter copula families.

To calculate the maximum likelihood estimates of the copula parameters in a fast and efficient way the partial derivatives with respect to the copula parameter(s) or arguments of several copula families are implemented. In particular, the derivatives of the elliptical bivariate copulas Gauss and Student's t as well as of the one-parametric Archimedean copulas calculated in Chapter 1 are available in the functions `BiCopDeriv` and `BiCopDeriv2`. Since the bivariate derivatives of the copula density as well as of the h-function form the building blocks of the R-vine score function and Hessian matrix these are also implemented in `BiCopHfuncDeriv` and `BiCopHfuncDeriv2`. They provide the first and second derivatives of a given conditional parametric bivariate copula (h-function; `BiCopHfunc`) with respect to its parameter(s) or one of its arguments, respectively. See therefor also Section 1.3.3.

Statistical inference for vine copulas

The statistical inference for regular vines is provided in the package **VineCopula** on which we will concentrate here. Most of the functions explained in the following are also available for C- and D-vine in the package **CDVine**. Brechmann and Schepsmeier (2013) gave a detailed introduction for the handling and the functionality of the **CDVine** package, which follows a slightly different notation than the **VineCopula** package.

As mentioned above it is convenient for calculation purposes to use the matrix notation for R-vines as introduced by Morales-Nápoles et al. (2010), Dißmann (2010) and Dißmann et al. (2013). The function `RVineMatrix` creates an R-vine matrix object which encodes an R-vine copula model. It contains the R-vine structure matrix M, denoted as `Matrix`, and three further matrices identifying the pair-copula families $\mathcal{B}$ (`family`) utilized and their parameter matrices $\boldsymbol{\theta}_1$ (`par`) and $\boldsymbol{\theta}_2$ (`par2`). The RVineMatrix-object forms the basis for all R-vine related functions, for example the simulation from a given R-vine copula model (`RVineSim`) or the calculation of the log-likelihood (`RVineLogLik`).

Having decided the structure of the R-vine to be used, one has to select pair-copula families for each (conditional) pair of variables or using the function `RVineCopSelect`. Based on `BiCopSelect`, this function selects for a given copula data set appropriate bivariate copula families from a set of possible copula families according to the AIC (default) or the BIC criterion. This copula selection proceeds tree by tree, since the conditional pairs in trees $2, ..., d-1$ depend on the specification of the previous trees through the h-functions. Hence, initially R-vine copula models are typically fitted sequentially by proceeding iteratively tree by tree and thus only involving bivariate estimation for each individual pair-copula term. This can be established using the function `RVineSeqEst` which internally calls the function `BiCopEst`.

Even though these sequential estimates often provide a good fit, one typically is interested in maximizing the (log-)likelihood of a vine copula specification, which is facilitated in `RVineMLE` using the sequential estimates as initial values. The log-likelihood of a vine copula for given copula data, pair-copula families and parameters can be obtained using the function `RVineLogLik` which implements the algorithm given in Algorithm 2.2.1.

If the vine structure has to be selected as well given a d-dimensional copula data set, one can apply the function `RVineStructureSelect`. It performs the maximum spanning tree algorithm suggested by Dißmann et al. (2013) maximizing the absolute sum of Kendall's τ tree by tree. In particular, the following optimization problem is solved for each tree

$$\max \sum_{\textit{edges } e_{ij} \textit{ in spanning tree}} |\hat{\tau}_{ij}|,$$

where $\hat{\tau}_{ij}$ denote the pairwise empirical Kendall's taus and a spanning tree is a tree on all nodes. The setting of the first tree selection step is always a complete graph. For subsequent trees, the setting depends on the R-vine construction principles, in particular on the proximity condition. The special case of a C-vine can be selected too, using the algorithm of Czado et al. (2012). Tree structures are determined and appropriate pair-copula families are selected using `BiCopSelect` and estimated sequentially.

The vine tree structure can be plotted with `RVineTreePlot`. This function plots one or all trees of a given R-vine copula model. The edges can be labeled with the pair-copula families, its parameters, the theoretical or empirical Kendall's τ values, or the indices of (conditioned) pair of variables identified by the edges. For an example see Figure 4.12 in Chapter 4 illustrating the first example.

In the next chapter we will introduce the estimation of standard errors in regular vine copula models and therefore the first and second derivatives of the R-vine log-likelihood are needed. Of cause this functionality is also included in the package by the R-functions `RVineStdError`, `RVineGradient` and `RVineHessian`, respectively. Note that the ordering of the gradient and Hessian matrix follows the ordering of the R-vine matrix. The gradient starts at the lower right corner of the R-vine matrix and goes column by column to the left and up, i.e. the first entry of the gradient is the last entry of the second last column of the parameter matrix followed by the last entry of the third last column and the second last entry of this column. If there is a copula family with two parameters, i.e. the t-copula, the derivative with respect to the second parameter is at the end of the gradient vector in order of their occurrence. Since derivatives of the bivariate two parameter Archimedean copulas, i.e. BB1, BB6, BB7, BB8 and their rotated versions are not implemented the gradient and Hessian matrix is not available for R-vine copula models with BB copulas.

The function `RVineStdError` returns two matrices of standard errors corresponding to the parameter matrices given the Hessian matrix. Note that the negative Hessian matrix has to be positive semidefinite to calculate the covariance matrix as inversion of the observed Fisher information, i.e. the Hessian matrix.

Furthermore, the R-vine goodness-of-fit tests proposed in Chapter 4 can be calculated too, using the function `RVineGofTest`. Its function argument `method` determines which goodness-of-fit method is used. Further function arguments like `statistic` or `B` manage the different sub-methods and the calculation of bootstrapped p-values, respectively.

The main functionality is coded in C to speed up the calculation and R is used as convenient user interface. Both packages depend on the packages **igraph** (Csardi 2010; Csardi and Nepusz 2006) and **mvtnorm** (Genz et al. 2011; Genz and Bretz 2009). For the calculation of the Anderson-Darling test statistics for the goodness-of-fit tests of Chapter 4 the R-package **ADGofTest** (Bellosta 2011) is required.

For reproducibility we want to mention that all the calculations given in this thesis are executed either on R 2.13.1 or R 2.15.0. The results are, to our knowledge, not affected by the different versions or operating systems (UNIX or Windows). Most computations have been performed using VineCopula 1.1. But unpublished pre-releases are used too. The numerical calculations in Section 1.3.3, the simulation studies in Section 3.3, 4.1.3 and 4.4, and the examples in Section 1.3.4, 3.4 and 4.5 were performed on a Linux cluster supported by DFG grant INST 95/919-1 FUGG.

Chapter 3

Estimating standard errors in regular vine copula models[1]

In the last chapter we introduced the very flexible class of regular vine copula models for complex multivariate data. However, while it is a standard exercise in multivariate statistics to compute the uncertainty incorporated in parameter point estimates for classes like the multivariate normal distribution, this is often not possible for the more complex models. But these advanced models are required to accurately capture the dependence in real-world multivariate data. Despite the wide range of applications of R-vine copula based models in practice, there is a surprising scarcity in the literature considering the uncertainty in point estimates of the copula parameters. It is well known that maximum likelihood estimates $\hat{\boldsymbol{\theta}}_n$ of the p-dimensional vine parameter vector $\boldsymbol{\theta}$ will be strongly consistent and asymptotically normal under regularity conditions on the bivariate building blocks (see Hobæk Haff 2013), i.e.

$$\sqrt{n}\, \mathcal{I}(\boldsymbol{\theta})^{1/2}\Big(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\Big) \xrightarrow{d} N(0, I_p) \text{ as } n \to \infty. \tag{3.1}$$

Here, n is the number of (i.i.d.) observations, I_p the $p \times p$ identity matrix, and

$$\mathcal{I}(\boldsymbol{\theta}) = -\mathbb{E}_{\boldsymbol{\theta}}\left[\Big(\frac{\partial^2}{\partial\theta_i\partial\theta_j} l(\boldsymbol{\theta}|\boldsymbol{X})\Big)_{i,j=1,\ldots,p}\right] = \mathbb{E}_{\boldsymbol{\theta}}\left[\Big(\frac{\partial}{\partial\theta_i} l(\boldsymbol{\theta}|\boldsymbol{X}) \cdot \frac{\partial}{\partial\theta_j} l(\boldsymbol{\theta}|\boldsymbol{X})\Big)_{i,j=1,\ldots,p}\right], \tag{3.2}$$

denotes the *Fisher Information Matrix* with $l(\boldsymbol{\theta}|\boldsymbol{x})$ being the log-likelihood of parameter $\boldsymbol{\theta}$ for one multivariate observation $\boldsymbol{x}$. For a given data set, we can replace the expectation in Equation (3.2) by taking the sample mean to obtain the *observed information*. Given the fact that full maximum likelihood inference is numerically difficult when non-uniform marginal distributions are involved, also two-step procedures have been developed.

In particular, Joe and Xu (1996) employ the probability integral transform using parametric marginal distributions which are fitted in a first step to obtain uniform (copula) data on which the copula is estimated using ML in a second step. As an alternative, Genest et al. (1995) proposed to use non-parametric rank transformations in the first step. While these methods are computationally more tractable, they are asymptotically less efficient.

[1]The contents of this chapter is based on J. Stoeber and U. Schepsmeier (2013), Estimating standard errors in regular vine copula models, forthcoming in Computational Statistics.

Following Hobæk Haff (2013), we can decompose the asymptotic covariance matrix for the estimates of dependence parameters $\boldsymbol{\theta}$ in a marginal and a dependence part:

$$\boldsymbol{V}^{\boldsymbol{\theta},two-step} = \boldsymbol{V}^{dependence} + \boldsymbol{V}^{margins},$$

where $\boldsymbol{V}^{dependence,ML} = \mathcal{I}(\boldsymbol{\theta})^{-1}$ for ML estimation and the second part is zero only when there is no uncertainty about the margins. Further, also the estimation of copula parameters can be performed in a tree by tree fashion. Denote the log-likelihood arising from parameters in tree i by

$$l_i(\boldsymbol{\theta}_i|\boldsymbol{x}) := \sum_{e\in E_i} \log\left(c_{j(e),k(e)|D(e)}(F_{j(e)|D(e)}(x_{j(e)}|\boldsymbol{x}_{D(e)}), F_{k(e)|D(e)}(x_{k(e)}|\boldsymbol{x}_{D(e)})|\theta_e)\right),$$

where the set of components of $\boldsymbol{\theta}_i$ is $\{\theta_e|e \in E_i\}$. In particular $\boldsymbol{\theta}_1$ is estimated by maximizing $l_1(\boldsymbol{\theta}_1|\boldsymbol{x})$ and the obtained estimates $\hat{\boldsymbol{\theta}}_1$ are used to calculate the arguments of the copula functions in $l_2(\boldsymbol{\theta}_2|\boldsymbol{x})$. Now, we maximize $l_2(\boldsymbol{\theta}_2|\boldsymbol{x})$ in $\boldsymbol{\theta}_2$ to obtain $\hat{\boldsymbol{\theta}}_2$ and proceed until all parameters are estimated (for an algorithm see Stöber and Czado 2012). This implies that l_i also implicitly depends on the parameters $\boldsymbol{\theta}_j$ for $j < i$ through its arguments, i.e. $l_i(\boldsymbol{\theta}_i|\boldsymbol{x}) = l_i(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_{i-1}, \ldots, \hat{\boldsymbol{\theta}}_1|\boldsymbol{x})$. The parameter estimates obtained from this sequential procedure have asymptotical covariance

$$\boldsymbol{V}^{dependence,seq.} = \boldsymbol{\mathcal{J}}_{\boldsymbol{\theta}}^{-1}\boldsymbol{\mathcal{K}}_{\boldsymbol{\theta}}\left(\boldsymbol{\mathcal{J}}_{\boldsymbol{\theta}}^{-1}\right)^T, \tag{3.3}$$

where $\boldsymbol{\mathcal{J}}_{\boldsymbol{\theta}}$ involves second derivatives of the R-vine log-likelihood function and $\boldsymbol{\mathcal{K}}_{\boldsymbol{\theta}}$ involves elements of the score function, see Hobæk Haff (2013). To be more precise, $\boldsymbol{\mathcal{K}}_{\boldsymbol{\theta}}$ and $\boldsymbol{\mathcal{J}}_{\boldsymbol{\theta}}$ are defined as

$$\boldsymbol{\mathcal{K}}_{\boldsymbol{\theta}} = \begin{pmatrix} \boldsymbol{\mathcal{K}}_{\boldsymbol{\theta},1,1} & & & \\ \vdots & \ddots & & \\ \boldsymbol{0}^T & \cdots & \boldsymbol{\mathcal{K}}_{\boldsymbol{\theta},d-2,d-2} & \\ \boldsymbol{0}^T & \cdots & \boldsymbol{0}^T & \boldsymbol{\mathcal{K}}_{\boldsymbol{\theta},d-1,d-1} \end{pmatrix}, \tag{3.4}$$

$$\boldsymbol{\mathcal{J}}_{\boldsymbol{\theta}} = \begin{pmatrix} \boldsymbol{\mathcal{J}}_{\boldsymbol{\theta},1,1} & & & \\ \vdots & \ddots & & \\ \boldsymbol{\mathcal{J}}_{\boldsymbol{\theta},d-2,1} & \cdots & \boldsymbol{\mathcal{J}}_{\boldsymbol{\theta},d-2,d-2} & \\ \boldsymbol{\mathcal{J}}_{\boldsymbol{\theta},d-1,1} & \cdots & \boldsymbol{\mathcal{J}}_{\boldsymbol{\theta},d-1,d-2} & \boldsymbol{\mathcal{J}}_{\boldsymbol{\theta},d-1,d-1} \end{pmatrix}, \tag{3.5}$$

with $\boldsymbol{\mathcal{K}}_{\boldsymbol{\theta},i,j} = E\left[\left(\frac{\partial l_i(\boldsymbol{\theta}_i,\ldots,\boldsymbol{\theta}_1|\boldsymbol{X})}{\partial\boldsymbol{\theta}_i}\right)\left(\frac{\partial l_j(\boldsymbol{\theta}_j,\ldots,\boldsymbol{\theta}_1|\boldsymbol{X})}{\partial\boldsymbol{\theta}_j}\right)^T\right]$ and $\boldsymbol{\mathcal{J}}_{\boldsymbol{\theta},i,j} = -E\left[\frac{\partial^2 l_i(\boldsymbol{\theta}_i,\ldots,\boldsymbol{\theta}_1|\boldsymbol{X})}{\partial\boldsymbol{\theta}_i\partial\boldsymbol{\theta}_j}\right], i,j = 1,\ldots,d-1$.

Again, for a given data set, the expectations can be replaced by taking the sample mean to obtain an equivalent of the observed information. While this asymptotic theory is well known, it is almost never applied in practice since the estimation of the asymptotic covariance matrix will involve the Hessian matrix, i.e. the second derivatives of the R-vine likelihood function. For these derivatives, no analytical expressions have been available, which caused a gap between theoretical knowledge about estimation errors and practical applicability which we will fill in this chapter. In particular, our algorithms overcome the reliability and accuracy issues associated with statistical estimation methods based on numerical derivatives (see McCullough 1998, 1999 for a discussion and examples). This allows to perform maximum likelihood estimation and compute the observed information also in a multidimensional setup where algorithms based on numerical differentiation can be highly unreliable.

3.1 Computation of the score function

In this section we develop an algorithm to calculate the derivatives of the R-vine log-likelihood with respect to copula parameters and thus the score function of the model. Throughout the remainder, we will assume that all occurring copula densities are continuously differentiable with respect to their arguments and parameters. Further, we assume that the copula parameters are all in $\mathbb{R}$, the extension to two or higher dimensional parameter spaces is straightforward but makes the notation unnecessarily complex.

To determine the log-likelihood derivatives, we will again exploit the hierarchical structure of the R-vine copula model and proceed similarly as for the likelihood calculation. The first challenge in developing an algorithm for the score function is to determine which of the copula terms in Expression (2.3) depend on which parameter directly or indirectly through one of their arguments. Following the steps of the log-likelihood computation and exploiting the structure of the R-vine matrix M, this is decided in Algorithm 3.1.1.

Algorithm 3.1.1 Determine copula terms which depend on a specific parameter.

The input of the algorithm is a d-dimensional R-vine matrix M with elements $(m_{l,j})_{l,j=1,\dots,d}$ and the row number k and column number i corresponding to the position of the parameter of interest in the corresponding parameter matrix $\boldsymbol{\theta}$. The output will be a matrix C (with elements $(c_{l,j})_{l,j=1,\dots,d}$) of zeros and ones, a one indicating that the copula term corresponding to this position in the matrix will depend on the parameter under consideration.

Set $g := (m_{i,i}, m_{k,i}, m_{k+1,i}, \dots, m_{d,i})$
Set $c_{l,j} := 0 \quad l, j = 1, \dots, d$
for $a = i, \dots, 1$ **do**
  for $b = k, \dots, a+1$ **do**
    Set $h := (m_{a,a}, m_{b,a}, m_{b+1,a}, \dots, m_{d,a})$
    if $\#(g \cap h) == \#g$ **then**
      Set $c_{b,a} := 1$
    end if
  end for
end for
return C

Knowing how a specific copula term depends on a given parameter, we can proceed with calculating the corresponding derivatives. Before we explain the derivatives in detail let us consider an example in 3 dimensions, where two of the three possible cases of dependence on a given parameter are illustrated.

Example 3.1 (3-dim, continuation of Example 2.1)
Let $x_1 \sim F_1, x_2 \sim F_2, x_3 \sim F_3$ and $u_1 = F(x_1), u_2 = F(x_2)$, $u_3 = F(x_3)$, then the joint density can be decomposed as

$$f_{123}(x_1, x_2, x_3) = f_1(x_1)f_2(x_2)f_3(x_3) \cdot c_{1,2}(u_1, u_2|\theta_{1,2}) \cdot c_{2,3}(u_2, u_3|\theta_{2,3}) \\ \cdot c_{1,3|2}(h_{1,2}(u_2, u_1|\theta_{1,2}), h_{2,3}(u_3, u_2|\theta_{2,3})|\theta_{1,3|2}).$$

The first derivatives of $\ln f_{123}$ *with respect to the copula parameters are*

$$\begin{aligned}
\frac{\partial(\ln f_{123}(x_1, x_2, x_3))}{\partial\theta_{1,2}} &= \frac{\partial_{\theta_{1,2}} c_{1,2}(u_1, u_2|\theta_{1,2})}{c_{1,2}(u_1, u_2|\theta_{1,2})} \\
&+ \frac{\partial_1 c_{1,3|2}(h_{1,2}(u_2, u_1|\theta_{1,2}), h_{2,3}(u_3, u_2|\theta_{2,3})|\theta_{1,3|2})}{c_{1,3|2}(h_{1,2}(u_2, u_1|\theta_{1,2}), h_{2,3}(u_3, u_2|\theta_{2,3})|\theta_{1,3|2})} \cdot \partial_{\theta_{1,2}} h_{1,2}(u_1, u_2|\theta_{1,2}), \\
\frac{\partial(\ln f_{123}(x_1, x_2, x_3))}{\partial\theta_{2,3}} &= \frac{\partial_{\theta_{2,3}} c_{2,3}(u_2, u_3|\theta_{2,3})}{c_{2,3}(u_2, u_3|\theta_{2,3})} \\
&+ \frac{\partial_2 c_{1,3|2}(h_{1,2}(u_2, u_1|\theta_{1,2}), h_{2,3}(u_3, u_2|\theta_{2,3})|\theta_{1,3|2})}{c_{1,3|2}(h_{1,2}(u_2, u_1|\theta_{1,2}), h_{2,3}(u_3, u_2|\theta_{2,3})|\theta_{1,3|2})} \cdot \partial_{\theta_{2,3}} h_{2,3}(u_2, u_3|\theta_{2,3}), \\
\frac{\partial(\ln f_{123}(x_1, x_2, x_3))}{\partial\theta_{1,3|2}} &= \frac{\partial_{\theta_{1,3|2}} c_{1,3|2}(h_{1,2}(u_2, u_1|\theta_{1,2}), h_{2,3}(u_3, u_2|\theta_{2,3})|\theta_{1,3|2})}{c_{1,3|2}(h_{1,2}(u_2, u_1|\theta_{1,2}), h_{2,3}(u_3, u_2|\theta_{2,3})|\theta_{1,3|2})}.
\end{aligned}$$

The first case which occurs in our example is that the copula densities $c_{1,2}$ and $c_{2,3}$ depend on their respective parameters directly. For a general term involving a copula $c_{U,V|\mathbf{Z}}$ with parameter θ, the derivative is

$$\begin{aligned}
\frac{\partial}{\partial\theta} \ln \left(c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z}), F_{V|\mathbf{Z}}(v|\mathbf{z})|\theta\right)\right) &= \frac{\frac{\partial}{\partial\theta}\left(c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z}), F_{V|\mathbf{Z}}(v|\mathbf{z})|\theta\right)\right)}{c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z}), F_{V|\mathbf{Z}}(v|\mathbf{z})|\theta\right)} \\
&= \frac{\partial_\theta c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z}), F_{V|\mathbf{Z}}(v|\mathbf{z})|\theta\right)}{c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z}), F_{V|\mathbf{Z}}(v|\mathbf{z})|\theta\right)}.
\end{aligned} \tag{3.6}$$

Further, like for $c_{1,3|2}$, a $c_{U,V|\mathbf{Z}}$ term can depend on a parameter θ through one of its arguments, say $F_{U|\mathbf{Z}}(u|\mathbf{z}, \theta)$:

$$\begin{aligned}
&\frac{\partial}{\partial\theta} \ln \left(c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z}, \theta), F_{V|\mathbf{Z}}(v|\mathbf{z})\right)\right) = \\
&\qquad = \frac{\frac{\partial c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z},\theta), F_{V|\mathbf{Z}}(v|\mathbf{z})\right)}{\partial F_{U|\mathbf{Z}}(u|\mathbf{z},\theta)}}{c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z}, \theta), F_{V|\mathbf{Z}}(v|\mathbf{z})\right)} \cdot \frac{\partial}{\partial\theta} F_{U|\mathbf{Z}}(u|\mathbf{z}, \theta) \\
&\qquad = \frac{\partial_1 c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z}, \theta), F_{V|\mathbf{Z}}(v|\mathbf{z})\right)}{c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z}, \theta), F_{V|\mathbf{Z}}(v|\mathbf{z})\right)} \cdot \frac{\partial}{\partial\theta} F_{U|\mathbf{Z}}(u|\mathbf{z}, \theta).
\end{aligned} \tag{3.7}$$

Finally, in dimension $d \geq 4$, both arguments of a $c_{U,V|\mathbf{Z}}$ copula term can depend on a parameter θ. In this case,

$$\begin{aligned}
&\frac{\partial}{\partial\theta}\ln\left(c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z},\theta),F_{V|\mathbf{Z}}(v|\mathbf{z},\theta)\right)\right)\\
&\qquad=\frac{\partial_1 c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z},\theta),F_{V|\mathbf{Z}}(v|\mathbf{z},\theta)\right)}{c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z},\theta),F_{V|\mathbf{Z}}(v|\mathbf{z},\theta)\right)}\cdot\frac{\partial}{\partial\theta}F_{U|\mathbf{Z}}(u|\mathbf{z},\theta)\\
&\qquad+\frac{\partial_2 c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z},\theta),F_{V|\mathbf{Z}}(v|\mathbf{z},\theta)\right)}{c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z},\theta),F_{V|\mathbf{Z}}(v|\mathbf{z},\theta)\right)}\cdot\frac{\partial}{\partial\theta}F_{V|\mathbf{Z}}(v|\mathbf{z},\theta).
\end{aligned}\tag{3.8}$$

We see that the derivatives of copula terms corresponding to tree T_i in the vine will involve derivatives of conditional distribution functions which are determined in tree T_{i-1}. Thus, it will be convenient to store their derivatives in matrices $S1^{direct,\theta}$ and $S1^{indirect,\theta}$ related to the matrices V^{direct} and $V^{indirect}$ which have been determined during the calculation of the log-likelihood together with the terms

$$\ln\left(c_{j(e),k(e)|D(e)}\left(F_{j(e)|D(e)}(x_{j(e)}|\boldsymbol{x}_{D(e)}),F_{k(e)|D(e)}(x_{k(e)}|\boldsymbol{x}_{D(e)})\right)\right)=:\varrho_{j(e),k(e)|D(e)},$$

for each edge e in the R-vine $\mathcal{V}$, which can also be stored in a matrix V^{values}:

$$V^{values}=\begin{pmatrix}\cdots & & & \\ \cdots & \varrho_{4,m_{3,2}|m_{4,2},m_{5,2}} & & \\ \cdots & \varrho_{4,m_{4,2}|m_{5,2}} & \varrho_{3,m_{4,3}|m_{5,3}} & \\ \cdots & \varrho_{4,m_{5,2}} & \varrho_{3,m_{5,3}} & \varrho_{2,1} \\ \cdots & & & \end{pmatrix}\tag{3.9}$$

In particular, we will determine the following matrices:

$$S1^{direct,\theta}=\begin{pmatrix}\cdots & & & \\ \cdots & \frac{\partial}{\partial\theta}F(u_4|u_{m_{3,2}},u_{m_{4,2}},u_{m_{5,2}}) & & \\ \cdots & \frac{\partial}{\partial\theta}F(u_4|u_{m_{4,2}},u_{m_{5,2}}) & \frac{\partial}{\partial\theta}F(u_3|u_{m_{4,3}},u_{m_{5,3}}) & \\ \cdots & \frac{\partial}{\partial\theta}F(u_4|u_{m_{5,2}}) & \frac{\partial}{\partial\theta}F(u_3|u_{m_{5,3}}) & \frac{\partial}{\partial\theta}F(u_2|u_1) \\ \cdots & & & \end{pmatrix},\tag{3.10}$$

$$S1^{indirect,\theta}=\begin{pmatrix}\cdots & & & \\ \cdots & \frac{\partial}{\partial\theta}F(u_{m_{3,2}}|u_{m_{4,2}},u_{m_{5,2}},u_4) & & \\ \cdots & \frac{\partial}{\partial\theta}F(u_{m_{4,2}}|u_{m_{5,2}},u_4) & \frac{\partial}{\partial\theta}F(u_{m_{4,3}}|u_{m_{5,3}},u_3) & \\ \cdots & \frac{\partial}{\partial\theta}F(u_{m_{5,2}}|u_4) & \frac{\partial}{\partial\theta}F(u_{m_{5,3}}|u_3) & \frac{\partial}{\partial\theta}F(u_1|u_2) \\ \cdots & & & \end{pmatrix},\tag{3.11}$$

$$S1^{values,\theta}=\begin{pmatrix}\cdots & & & \\ \cdots & \frac{\partial}{\partial\theta}\varrho_{4,m_{3,2}|m_{4,2},m_{5,2}} & & \\ \cdots & \frac{\partial}{\partial\theta}\varrho_{4,m_{4,2}|m_{5,2}} & \frac{\partial}{\partial\theta}\varrho_{3,m_{4,3}|m_{5,3}} & \\ \cdots & \frac{\partial}{\partial\theta}\varrho_{4,m_{5,2}} & \frac{\partial}{\partial\theta}\varrho_{3,m_{5,3}} & \frac{\partial}{\partial\theta}\varrho_{2,1} \\ \cdots & & & \end{pmatrix}.\tag{3.12}$$

Here, the terms in $S1^{direct,\theta}$ and $S1^{indirect,\theta}$ can be determined by differentiating (2.4) similarly as we did for the copula terms in (3.6) - (3.8). For instance, we have

$$\begin{aligned}\frac{\partial}{\partial\theta}F_{U|V,\mathbf{z}}(u|v,\mathbf{z},\theta) &= \frac{\partial}{\partial\theta}\left(h_{U|V,\mathbf{z}}\left(F_{U|\mathbf{z}}(u|\mathbf{z},\theta),F_{V|\mathbf{z}}(v|\mathbf{z})\right)\right)\\ &= \partial_1 h_{U|V,\mathbf{z}}\left(F_{U|\mathbf{z}}(u|\mathbf{z},\theta),F_{V|\mathbf{z}}(v|\mathbf{z})\right)\cdot\frac{\partial}{\partial\theta}F_{U|\mathbf{z}}(u|\mathbf{z},\theta)\\ &= c_{U|V,\mathbf{z}}\left(F_{U|\mathbf{z}}(u|\mathbf{z},\theta),F_{V|\mathbf{z}}(v|\mathbf{z})\right)\cdot\frac{\partial}{\partial\theta}F_{U|\mathbf{z}}(u|\mathbf{z},\theta)\end{aligned}$$

and

$$\frac{\partial}{\partial\theta}h_{U|V,\mathbf{z}}\big(F_{U|\mathbf{z}}(u|\mathbf{z}),F_{V|\mathbf{z}}(v|\mathbf{z},\theta)\big) = \partial_2 h_{U|V,\mathbf{z}}\left(F_{U|\mathbf{z}}(u|\mathbf{z}),F_{V|\mathbf{z}}(v|\mathbf{z},\theta)\right)\cdot\frac{\partial}{\partial\theta}F_{V|\mathbf{z}}(v|\mathbf{z},\theta).$$

The complete calculations required to obtain the derivative of the log-likelihood with respect to one copula parameter θ are performed in Algorithm 3.1.2.

Algorithm 3.1.2 Log-likelihood derivative with respect to the parameter $\theta^{\tilde{k},\tilde{i}}$.

The input of the algorithm is a d-dimensional R-vine matrix M with maximum matrix $\tilde{M}$ and parameter matrix $\boldsymbol{\theta}$, and a matrix C determined using Algorithm 3.1.1 for a parameter $\theta^{\tilde{k},\tilde{i}}$ positioned at row $\tilde{k}$ and $\tilde{i}$ in the R-vine parameter matrix $\boldsymbol{\theta}$. Further, we assume the matrices V^{direct}, $V^{indirect}$ and V^{values} corresponding to one observation from the R-vine copula distribution, which have been determined during the calculation of the log-likelihood, to be given. The output will be the value of the first derivative of the copula log-likelihood for the given observation with respect to the parameter $\theta^{\tilde{k},\tilde{i}}$.

```
Set z_1 = v^{direct}_{k̃,ĩ}
Set s1^{direct}_{k,i} := 0, s1^{indirect}_{k,i} := 0, s1^{values}_{k,i} := 0, i = 1,...,d; k = i,...,d
if m_{k̃,ĩ} == m̃_{k̃,ĩ} then
    Set z_2 = v^{direct}_{k̃,d−m̃_{k̃,ĩ}+1}
else
    Set z_2 = v^{indirect}_{k̃,d−m̃_{k̃,ĩ}+1}
end if
Set s1^{direct}_{k̃−1,ĩ} = ∂_{θ^{k̃,ĩ}} h(z_1, z_2 | B^{k̃,ĩ}, θ^{k̃,ĩ})
Set s1^{indirect}_{k̃−1,ĩ} = ∂_{θ^{k̃,ĩ}} h(z_2, z_1 | B^{k̃,ĩ}, θ^{k̃,ĩ})
Set s1^{values}_{k̃,ĩ} = ∂_{θ^{k̃,ĩ}} c(z_1, z_2 | B^{k̃,ĩ}, θ^{k̃,ĩ}) / exp(v^{values}_{k̃,ĩ})
for i = ĩ,...,1 do
    for k = k̃ − 1,...,i + 1 do
        if c_{k,i} == 1 then
            Set z_1 = v^{direct}_{k,i}, z̃_1 = s1^{direct}_{k,i}
            if m_{k,i} == m̃_{k,i} then
                Set z_2 = v^{direct}_{k,d−m̃_{k,i}+1}, z̃_2 = s1^{direct}_{k,d−m̃_{k,i}+1}
            else
                Set z_2 = v^{indirect}_{k,d−m̃_{k,i}+1}, z̃_2 = s1^{indirect}_{k,d−m̃_{k,i}+1}
```

19: **end if**
20: **if** $c_{k+1,i} == 1$ **then**
21: Set $s1^{values}_{k,i} = s1^{values}_{k,i} + \frac{\partial_1 c(z_1,z_2|\mathcal{B}^{k,i},\theta^{k,i})}{v^{values}_{k,i}} \cdot \tilde{z}_1$
22: Set $s1^{direct}_{k-1,i} = s1^{direct}_{k-1,i} + \partial_1 h(z_1, z_2|\mathcal{B}^{k,i}, \theta^{k,i}) \cdot \tilde{z}_1$
23: Set $s1^{indirect}_{k-1,i} = s1^{indirect}_{k-1,i} + \partial_2 h(z_2, z_1|\mathcal{B}^{k,i}, \theta^{k,i}) \cdot \tilde{z}_1$
24: **end if**
25: **if** $c_{k+1,d-m+1} == 1$ **then**
26: Set $s1^{values}_{k,i} = s1^{values}_{k,i} + \frac{\partial_2 c(z_1,z_2|\mathcal{B}^{k,i},\theta^{k,i})}{exp(v^{values}_{k,i})} \cdot \tilde{z}_2$
27: Set $s1^{direct}_{k-1,i} = s1^{direct}_{k-1,i} + \partial_2 h(z_1, z_2|\mathcal{B}^{k,i}, \theta^{k,i}) \cdot \tilde{z}_2$
28: Set $s1^{indirect}_{k-1,i} = s1^{indirect}_{k-1,i} + \partial_1 h(z_2, z_1|\mathcal{B}^{k,i}, \theta^{k,i}) \cdot \tilde{z}_2$
29: **end if**
30: **end if**
31: **end for**
32: **end for**
33: **return** $\sum_{k,i=1,\ldots,d} s1^{values}_{k,i}$

In particular, this algorithm allows to replace finite-differences based numerical maximization of R-vine likelihood functions with maximization based on the analytical gradient. In a numerical comparison study across different R-vine models in 5-8 dimensions this resulted in a decrease in computation time by a factor of 4-8. Note that Algorithm 3.1.2 also allows to determine the terms $\partial l_j(\boldsymbol{\theta}_j, \ldots, \boldsymbol{\theta}_1)/\partial\boldsymbol{\theta}_j$ which are required to estimate the matrix $\boldsymbol{\mathcal{K}_\theta}$ needed for the calculation of standard errors of the sequential estimator. Here the sum in line 33 of Algorithm 3.1.2 has to be replaced by $\sum_{i=1,\ldots,d} s1^{values}_{d-j,i}$ for derivatives with respect to parameters $\theta^{d-j+1,\tilde{i}}$ in $\boldsymbol{\theta}_j$.

3.2 Computation of the observed information

Based on the calculation of the score function performed in the previous section, we will present an algorithm to determine the Hessian matrix corresponding to the R-vine log-likelihood function in this section. Considering a derivative with respect to bivariate copula parameters θ and γ associated with the vine, it is clear that the expressions for the derivatives of the log-densities in this case will contain second derivatives of the occurring h-functions. Thus, our algorithm will determine the following matrices:

$$S2^{direct,\theta,\gamma} =$$
$$\begin{pmatrix} \cdots & & & \\ \cdots & \frac{\partial}{\partial\theta\partial\gamma}F(u_4|u_{m_{3,2}},u_{m_{4,2}},u_{m_{5,2}}) & & \\ \cdots & \frac{\partial}{\partial\theta\partial\gamma}F(u_4|u_{m_{4,2}},u_{m_{5,2}}) & \frac{\partial}{\partial\theta\partial\gamma}F(u_3|u_{m_{4,3}},u_{m_{5,3}}) & \\ \cdots & \frac{\partial}{\partial\theta\partial\gamma}F(u_4|u_{m_{5,2}}) & \frac{\partial}{\partial\theta\partial\gamma}F(u_3|u_{m_{5,3}}) & \frac{\partial}{\partial\theta\partial\gamma}F(u_2|u_1) \\ \cdots & & & \end{pmatrix}, \quad (3.13)$$

$$S2^{indirect,\theta,\gamma} = \begin{pmatrix} \cdots & & & \\ \cdots & \frac{\partial}{\partial\theta\partial\gamma}F(u_{m_{3,2}}|u_{m_{4,2}},u_{m_{5,2}},u_4) & & \\ \cdots & \frac{\partial}{\partial\theta\partial\gamma}F(u_{m_{4,2}}|u_{m_{5,2}},u_4) & \frac{\partial}{\partial\theta\partial\gamma}F(u_{m_{4,3}}|u_{m_{5,3}},u_3) & \\ \cdots & \frac{\partial}{\partial\theta\partial\gamma}F(u_{m_{5,2}}|u_4) & \frac{\partial}{\partial\theta\partial\gamma}F(u_{m_{5,3}}|u_3) & \frac{\partial}{\partial\theta\partial\gamma}F(u_1|u_2) \\ \cdots & & & \end{pmatrix}, \tag{3.14}$$

$$S2^{values,\theta,\gamma} = \begin{pmatrix} \cdots & & & \\ \cdots & \frac{\partial}{\partial\theta\partial\gamma}\varrho_{4,m_{3,2}|m_{4,2},m_{5,2}} & & \\ \cdots & \frac{\partial}{\partial\theta\partial\gamma}\varrho_{4,m_{4,2}|m_{5,2}} & \frac{\partial}{\partial\theta\partial\gamma}\varrho_{3,m_{4,3}|m_{5,3}} & \\ \cdots & \frac{\partial}{\partial\theta\partial\gamma}\varrho_{4,m_{5,2}} & \frac{\partial}{\partial\theta\partial\gamma}\varrho_{3,m_{4,2}} & \frac{\partial}{\partial\theta\partial\gamma}\varrho_{2,1} \\ \cdots & & & \end{pmatrix}. \tag{3.15}$$

Since not all entries in (2.6), (2.7) and (3.9) depend on both θ and γ, not all entries in (3.13) - (3.15) will be non-zero and required in the algorithm. Employing Algorithm 3.1.1 to obtain matrices C^θ and C^γ corresponding to the parameters θ and γ, respectively, we see that the second derivatives of all elements where the corresponding matrix entry of either C^θ or C^γ is zero vanish.
To derive an algorithm similar to Algorithm 3.1.2 which recursively determines all terms of the second derivatives of the log-likelihood with respect to parameters θ, γ, we need to distinguish 7 basic cases of dependence on the two parameters (Table 3.1) which can occur for a term $c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z}), F_{V|\mathbf{Z}}(v|\mathbf{z})\right)$.

	dependence on	θ	γ
case 1		$F_{U\|\mathbf{Z}}$	$F_{V\|\mathbf{Z}}$
case 2		$F_{U\|\mathbf{Z}}, F_{V\|\mathbf{Z}}$	$F_{V\|\mathbf{Z}}$
case 3		$F_{U\|\mathbf{Z}}, F_{V\|\mathbf{Z}}$	$F_{U\|\mathbf{Z}}, F_{V\|\mathbf{Z}}$
case 4	through	$F_{U\|\mathbf{Z}}$	$c_{U,V\|\mathbf{Z}}$
case 5		$F_{U\|\mathbf{Z}}, F_{V\|\mathbf{Z}}$	$c_{U,V\|\mathbf{Z}}$
case 6		$F_{U\|\mathbf{Z}}$	$F_{U\|\mathbf{Z}}$
case 7		$c_{U,V\|\mathbf{Z}}$	$c_{U,V\|\mathbf{Z}}$

Table 3.1: The 7 cases of how a copula terms can depend on θ and γ.

Here, case 7 is relevant only for derivatives where $\theta = \gamma$, since we assume that all bivariate copulas occurring in the vine density have one parameter in $\mathbb{R}$. Because of symmetry in the parameters, all other possible combinations are already included in these cases, we only have to exchange θ and γ. A more detailed description of the occurring derivatives is given in Appendix A.1.
As before, the terms in $S2^{direct,\theta,\gamma}$ and $S2^{indirect,\theta,\gamma}$ can be determined by differentiating (2.4) similarly as we did for the copula terms in (A.1) - (A.7) of Appendix A.1. Thus, the second derivatives can again be calculated recursively (see Algorithm A.1.1). Again, the algorithm allows to determine also the terms $\partial^2 l_j(\boldsymbol{\theta}_j, \ldots, \boldsymbol{\theta}_1)/\partial\boldsymbol{\theta}_j\partial\boldsymbol{\theta}_k$ required to estimate

$\boldsymbol{\mathcal{J}_\theta}$. For this, the summation over $S2^{values,\theta,\gamma}$ is replaced by summation over line $d-j$, i.e.

$$\frac{\partial^2 l_j(\boldsymbol{\theta}_j,\ldots,\boldsymbol{\theta}_1)}{\partial\theta\partial\gamma} = \sum_{i=1,\ldots,d} S2_{d-j,i}^{values,\theta,\gamma}.$$

Combining Algorithm A.1.1 and numerical integration techniques[2] we can also calculate the Fisher information (see Equation (3.2)) matrix of R-vine copula models and determine asymptotical standard errors for ML estimates.

Example 3.2 (3-dim. Gaussian and Student's t copula vine models)
Let us consider a 3-dimensional vine copula model with Gaussian pair-copulas (structure matrix M_3 and family matrix $\mathcal{B}^{Gauss}$). The parameter matrix $\boldsymbol{\theta}^{Gauss}$ gives the correlation parameters of the bivariate Gaussian copulas.

$$M_3 = \begin{pmatrix} 3 & & \\ 1 & 2 & \\ 2 & 1 & 1 \end{pmatrix}, \quad \mathcal{B}^{Gauss} = \begin{pmatrix} & & \\ Gauss & & \\ Gauss & Gauss & \end{pmatrix}, \quad \boldsymbol{\theta}^{Gauss} = \begin{pmatrix} & & \\ 0.34 & & \\ 0.79 & 0.35 & \end{pmatrix}.$$

Then, we can calculate the asymptotic standard errors for each parameter based on the expected information matrix $\mathcal{I}(\boldsymbol{\theta}^{Gauss})$ or rather $\boldsymbol{V}^{dep}$ for the MLE case (see Equation (3.2)) and for the sequential estimation case (Equation (3.3) and Hobæk Haff 2013), respectively. The order of the entries in the according asymptotic standard error matrices $\boldsymbol{ASE}^{MLE}$ and $\boldsymbol{ASE}^{seq}$ are the same as in the parameter matrix $\boldsymbol{\theta}^{Gauss}$. For $\mathcal{I}(\boldsymbol{\theta}^{Gauss})$ and $\boldsymbol{V}^{dep}$, the order of the parameters is $(\rho_{12}, \rho_{23}, \rho_{13|2})$.

$$\mathcal{I}^{MLE}(\boldsymbol{\theta}^{Gauss}) = \mathbb{E}_{\boldsymbol{\theta}}\left[-\left(\frac{\partial^2}{\partial\theta_i^{Gauss}\partial\theta_j^{Gauss}} l(\boldsymbol{\theta}^{Gauss})\right)_{i,j=1,\ldots,3}\right] = \begin{pmatrix} 1.62 & -0.77 & 0.15 \\ -0.77 & 12.40 & 0.80 \\ 0.15 & 0.80 & 1.42 \end{pmatrix},$$

$$\boldsymbol{V}^{dep,MLE} = \left(\mathcal{I}^{MLE}\right)^{-1}(\boldsymbol{\theta}^{Gauss}) = \begin{pmatrix} 0.65 & 0.05 & -0.10 \\ 0.05 & 0.09 & -0.05 \\ -0.10 & -0.05 & 0.75 \end{pmatrix}.$$

This implies that the asymptotic standard errors are given by the square roots of the diagonal elements as

$$\boldsymbol{ASE}^{MLE} = \begin{pmatrix} & \\ 0.86 & \\ 0.29 & 0.80 \end{pmatrix}, \quad \boldsymbol{ASE}^{seq} = \begin{pmatrix} & \\ 0.89 & \\ 0.31 & 0.83 \end{pmatrix}.$$

For the Gaussian distribution the occurring integrals can be computed both numerically and analytically (see Appendix A.2). The results indicate that regardless of the applied estimation method approximately 100 observations are required to estimate the parameters of the 3-dimensional Gaussian copula up to $\sigma = 0.01$.
In the second setting we change the bivariate copula families to Student's t copulas. The

[2]We use the adaptive integration routines supplied by Steven G. Johnson and Balasubramanian Narasimhan in the **cubature** package available on CRAN which are based on Genz and Malik (1980) and Berntsen, Espelid, and Genz (1991).

corresponding copula parameters are stored in $\boldsymbol{\theta}^{Student}$*, where the lower triangle gives the association parameter (parameter 1) and the upper triangle gives the degrees of freedom (parameter 2). As before* $\boldsymbol{ASE}^{MLE}$ *and* $\boldsymbol{ASE}^{seq}$ *denote the corresponding standard errors based on Equation (3.1) for the full ML estimation and Equation (3.3) for sequential estimation, respectively.*

$$\boldsymbol{\theta}^{Student} = \begin{pmatrix} & 3 & 3 \\ 0.34 & & 3 \\ 0.79 & 0.35 & \end{pmatrix},$$

$$\boldsymbol{ASE}^{MLE} = \begin{pmatrix} & 12 & 12 \\ 1.04 & & 11 \\ 0.39 & 0.97 & \end{pmatrix}, \quad \boldsymbol{ASE}^{seq} = \begin{pmatrix} & 12 & 14 \\ 1.04 & & 12 \\ 0.48 & 1.15 & \end{pmatrix}.$$

Note that our results show that for uniform $[0,1]$ *marginal distributions, the sequential procedure is less efficient than full MLE. This is interesting in combination with (Hobæk Haff 2013, Theorem 2) which states that together with non-parametric estimation of the marginal distributions, the sequential estimation procedure for the Gaussian distribution is asymptotically as efficient as the full MLE.*

3.3 Simulation study

This section presents the results of a simulation study which shows that the standard errors we compute for sequential and ML estimation are appropriate. We will estimate confidence intervals using the observed information and the sample estimate of (3.3) and compute their coverage rate.

The simulation setup is as follows: We use the 5 dimensional R-vine structure given in Equation 2.5 (Figure 2.2), with pair-copula families $\mathcal{B}$,

$$\mathcal{B} = \begin{pmatrix} \text{Indep.} & & & \\ \text{Clayton} & \text{Clayton} & & \\ \text{Gumbel} & \text{Gumbel} & \text{Frank} & \\ \text{Gumbel} & \text{Gauss} & t_5 & \text{Clayton} \end{pmatrix}.$$

Here, t_5 is a Student's t copula with degrees of freedom parameter $\nu = 5$. For the corresponding parameters, we consider two setups: all copula parameters are chosen to correspond to a Kendall's τ rank correlation of $\tau = 0.2$ (Setup 1) and $\tau = 0.6$ (Setup 2). Further, we consider four different sample sizes with $N = 100, 200, 400$ and 1000 observations and simulate 10000 data sets for each scenario. The coverage rate of estimated 90% confidence intervals, computed as the percentage of simulations for which the true parameter was within the $[\hat{\theta} - 1.645\hat{\sigma}, \hat{\theta} + 1.645\hat{\sigma}]$ interval, is given in Table 3.2.

As our results show, the coverage rate is close to the expected 90% for all parameters. The only parameter where we observe coverage rates closer to 95% for low sample sizes is the degrees of freedom parameter of the Student's t copula. With increasing sample size, however, the coverage rate decreases to 90% also for this two-parametric family. These results are independent of the strength of dependence (as measured in terms of Kendall's τs which is present in the data).

					$\tau = 0.2$						
n	est	$\theta_{1,2}$	$\rho_{2,4}$	$\nu_{2,4}$	$\rho_{2,3}$	$\theta_{3,5}$	$\theta_{1,4\|2}$	$\theta_{1,3\|2}$	$\theta_{2,5\|3}$	$\theta_{3,4\|1,2}$	$\theta_{1,5\|2,3}$
100	MLE	89.1	88.5	94.8	88.8	88.9	89.7	89.3	89.3	89.3	89.4
	seq.	88.2	88.8	95.0	88.0	88.7	89.7	88.8	88.9	88.9	88.9
200	MLE	89.5	89.3	90.2	89.6	89.7	89.8	89.7	89.3	89.4	89.4
	seq.	89.4	89.1	90.8	89.3	89.6	89.7	89.5	89.6	89.0	89.4
400	MLE	89.8	89.9	91.2	90.1	90.2	90.1	89.8	89.7	89.4	89.8
	seq.	89.3	89.8	91.4	89.5	90.1	90.1	89.7	89.6	89.4	89.5
1000	MLE	90.2	90.1	90.4	90.3	89.9	90.0	90.2	89.5	89.9	89.8
	seq.	89.3	89.9	90.4	90.4	89.9	90.0	90.2	89.5	89.9	89.5
					$\tau = 0.6$						
n	est	$\theta_{1,2}$	$\rho_{2,4}$	$\nu_{2,4}$	$\rho_{2,3}$	$\theta_{3,5}$	$\theta_{1,4\|2}$	$\theta_{1,3\|2}$	$\theta_{2,5\|3}$	$\theta_{3,4\|1,2}$	$\theta_{1,5\|2,3}$
100	MLE	89.0	88.1	89.9	88.6	88.9	89.5	89.0	89.1	89.0	89.3
	seq.	88.9	87.4	96.9	87.7	88.8	89.8	89.4	88.9	86.6	88.8
200	MLE	89.6	88.4	90.2	88.7	89.2	89.5	89.2	89.9	89.5	90.1
	seq.	89.4	88.4	93.6	89.3	89.6	89.8	89.5	89.5	88.4	90.0
400	MLE	89.8	89.2	89.7	89.8	89.9	89.8	90.1	89.9	90.0	90.0
	seq.	89.6	89.6	91.2	89.5	89.7	90.1	89.7	89.8	88.6	89.4
1000	MLE	90.2	90.2	89.7	90.4	90.1	90.2	89.7	89.8	89.9	89.5
	seq.	89.9	89.7	90.3	90.4	89.9	90.6	90.1	90.0	89.8	89.1

Table 3.2: Coverage rate of the 90% confidence interval.

Like many complex models, our implementation shows some numerical instabilities for low sample sizes. While the numerically stable computation of derivatives has been discussed in Chapter 1 and Schepsmeier and Stöber (2013), the recursive structure of an R-vine can amplify numerical inaccuracies. For low sample sizes and weak dependence, this resulted in some of the computed Hessian matrices not being negative (semi-) definite at the estimated parameters. In the simulation setup with Kendall's $\tau = 0.2$ and $N = 100$, this occurred for 263 of the 10000 data sets. For $N = 400$ only one exception was observed and no exceptions for bigger sample sizes and $\tau = 0.6$.

If the Hessian matrix is not negative semidefinite, a "close" negative definite matrix can be considered instead. We use the function `nearPD` of the package **Matrix** provided by Bates and Maechler (2012). Their implementation is based on an algorithm of Higham (2002).

Finally we apply our algorithms for the computation of the score function and observed information to a financial data set. It illustrates the applicability of our methods allowing to compute standard errors in a routinely manner for R-vines. Here, as in the simulation study before, we used the implemented algorithms of the functions `RVineStdError` of the R-package **VineCopula** to compute the standard errors out of the inverse Hessian matrix given by `RVineHessian`. The matrices $\boldsymbol{\mathcal{K}}_{\boldsymbol{\theta}}$ and $\boldsymbol{\mathcal{J}}_{\boldsymbol{\theta}}$ for the covariance matrix in the sequential approach of Hobæk Haff (2013) can be computed by sub-routines of `RVineHessian`.

3.4 Example: Rolling window analysis of exchange rate data

In this section, we illustrate the computation of standard errors in a rolling window analysis of the exchange rate data analyzed by Czado et al. (2012) and Stöber and Czado (2012). The data consists of 8 daily exchange rates quoted with respect to the US dollar during the period from July 22, 2005 to July 17, 2009, resulting in 1007 data points in total: Australian dollar (AUD), Japanese yen (JPY), Brazilian real (BRL), Canadian dollar (CAD), Euro (EUR), Swiss frank (CHF), Indian rupee (INR) and British pound (GBP). The marginal time dependence has been removed in a pre-analysis as described by Schepsmeier (2010, Chapter 5) (see Appendix A.3 for details) and the data is given as copula data on the unit hypercube. We employ the sequential model selection procedure of Dißmann et al. (2013) to select an adequate model for the data set. The selected R-vine structure, pair copula families, and estimated copula parameters for the whole data set are given in Appendix A.3.

To study possible inhomogeneities in dependence over time, we apply a rolling window analysis as follows: A window with a window-size of 100, 200 and 400 data points, respectively, is run over the data with step-size five, i.e. the window is moved by five trading days in each step. For each window dataset under investigation we estimate the R-vine parameters using ML (and sequential estimation) while keeping the R-vine structure $\mathcal{V}$ given by (A.17) and the copula families $\mathcal{B}$ given by (A.19) fixed. Additionally, we compute the observed information (and the equivalent for the sequential estimator) for each window and use it to obtain standard errors for the parameter estimates. In some windows the estimated degrees of freedom parameters are very high leading to numerical instabilities and indicating that the Student's t copulas which are associated with some bivariate marginal distributions might not be appropriate for the whole dataset. In Figure 3.1 we illustrate the estimated copula parameters and pointwise confidence intervals given by $[\hat{\theta} - 2\hat{\sigma}, \hat{\theta} + 2\hat{\sigma}]$ for some selected pair-copulas.

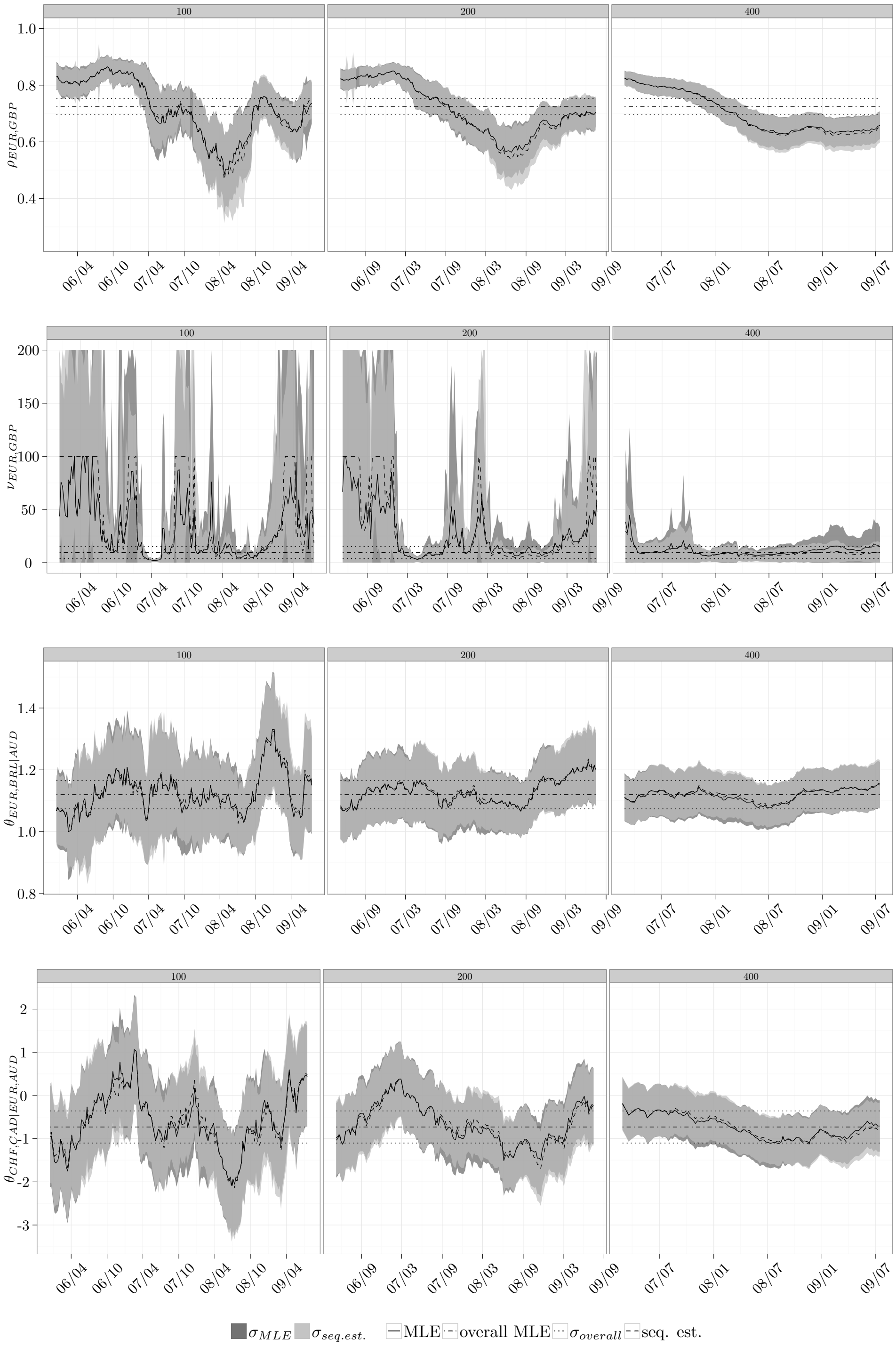


Figure 3.1: Rolling window analysis for the exchange rate data with window size 100 (left), 200 (middle) and 400 (right) for some selected par-copulas with different copula families (t-copula (row 1-2, degrees of freedom are cut off at 100), Gumbel (row 3) and Frank (row 4)). The x-axis indicates the endpoint of each window, with the corresponding parameter estimate (ML: solid, sequential estimation: dashed) on the y-axis. The dash-dotted horizontal line in each plot is the MLE corresponding to the whole dataset. The confidence intervals for ML and sequential estimation are indicated by the dark and light grey areas, respectively.

Chapter 4

Efficient goodness-of-fit tests in multi-dimensional vine copula models[1]

Although vine copula model inference and model selection is covered in several books and papers in the last decade the important statistical tool of goodness-of-fit testing is poorly addressed. It is unfortunate that there is little progress known in the theory and method concerning a goodness-of-fit (GOF) test for vine copulas, an important aspect of statistical model diagnostics. In fact, most of the published work has been only focused on bivariate copula models (see for example Genest et al. 2006, Dobrić and Schmid 2005, Dobrić and Schmid 2007, Huang and Prokhorov 2013, Genest et al. 2012 and may more). Comparison studies of Genest et al. 2009 or Berg 2009 investigated the most promising ones. But very little is provided for the validation of vine copula models. First GOF approaches are for example proposed in Aas et al. (2009) or Berg and Aas (2009) but not further studied or tested.

But model diagnosis becomes ever so imperative in the application of multi-dimensional vine copulas. Developing efficient GOF tests is now a timely task as already noted in Fermanian (2012), and an important addition to the current literature of vine copulas. In addition, comprehensive comparisons for many of the classical GOF tests are lacking in terms of their relative merits when they are applied to multi-dimensional copulas. So far model verification methods for vine copulas are usually based on the likelihood, or on the Akaike Information Criterion (AIC) or Bayesian Information Criterion (BIC) as classical comparison measures, which take the model complexity into account. The tests proposed by Vuong (1989) and Clarke (2007), suitable for non-nested models, may be applicable for vine copula models (see for example Brechmann et al. 2012). Note that these tests are not GOF tests, since they only compare two given (estimated) vine copula models based on their likelihood ratio.

[1]The contents of this chapter is based on Schepsmeier (2013), A goodness-of-fit test for regular vine copula models, submitted for publication, and Schepsmeier (2013), Efficient goodness-of-fit tests in multi-dimensional vine copula models, Submitted for publication

In our goodness-of-fit tests we would like to test

$$H_0 : C \in \mathcal{C}_0 = \{C_{\boldsymbol{\theta}} : \boldsymbol{\theta} \in \Theta\} \quad \text{against} \quad H_1 : C \notin \mathcal{C}_0 = \{C_{\boldsymbol{\theta}} : \boldsymbol{\theta} \in \Theta\}, \tag{4.1}$$

where C denotes the (vine) copula distribution function and $\mathcal{C}_0$ is a class of parametric (vine) copulas with $\Theta \subseteq \mathbb{R}^p$ being the parameter space of dimension p. Here we assume a known tree structure in the case of parametric vine copulas.

In this chapter we will propose two new rank based, "blanket" GOF tests for the vine copulas, based on the information matrix equality and specification test proposed by White (1982). The equality is sometimes called the Bartlett identity and is defined as

$$-\mathbb{H}(\boldsymbol{\theta}) = \mathbb{C}(\boldsymbol{\theta}). \tag{4.2}$$

Here $\mathbb{H}(\boldsymbol{\theta})$ is the expected Hessian or variability matrix, and $\mathbb{C}(\boldsymbol{\theta})$ is the expected outer product of the gradient or sensitivity matrix, which will be defined in more detail in Section 4.1.

Thus, the copula/vine misspecification test will be

$$H_0 : \mathbb{H}(\boldsymbol{\theta}_0) + \mathbb{C}(\boldsymbol{\theta}_0) = 0 \text{ against } H_1 : \mathbb{H}(\boldsymbol{\theta}_0) + \mathbb{C}(\boldsymbol{\theta}_0) \neq 0, \tag{4.3}$$

with $\boldsymbol{\theta}_0$ being the true value of the vine parameter vector. Consequently, the corresponding test statistic will be based on the difference between $-\mathbb{H}(\boldsymbol{\theta})$ and $\mathbb{C}(\boldsymbol{\theta})$. Such a test statistic was already proposed by Huang and Prokhorov (2013) for the bivariate copula case specializing White's information matrix equality and specification test. They investigated the size and power behavior for several bivariate parametric copulas given different alternatives of the same set of copula families. But multivariate extensions and generalizations were not investigated in their simulation study. Especially the asymptotic distribution function of the test statistic and the derived p-values were not taken into account in higher dimensions.

Here, we will derive the test statistic for the vine copula case and prove its asymptotic distribution. An extensive simulation study will reveal that the simulation based test shows excellent performance with regard to observed size and power, while the asymptotic theory based test is inaccurate for $n \leq 10000$ for a 5-dimensional model (in $d = 8$ even 20000 are not enough). We will denote this GOF test as White test.

In contrast to the White test, which relies on the difference between $-\mathbb{H}(\boldsymbol{\theta})$ and $\mathbb{C}(\boldsymbol{\theta})$, the second new test is based on the information matrix ratio (IMR) of Zhou et al. (2012):

$$\Psi(\boldsymbol{\theta}) := -\mathbb{C}(\boldsymbol{\theta})^{-1}\mathbb{H}(\boldsymbol{\theta}).$$

Here, our test problem is the reformulated general test problem of White (1982):

$$H_0 : \Psi(\boldsymbol{\theta}) = I_p \quad \text{against} \quad H_1 : \Psi(\boldsymbol{\theta}) \neq I_p,$$

where I_p is the p-dimensional identity matrix.

In Section 4.2 the IMR based test statistic for vine models will be derived and its asymptotic normality under the Bartlett identity will be proven. Secondly, the small sample performance for size and power will be investigated. Furthermore, the two new

GOF tests are compared to 13 other GOF tests for vines in a high dimensional setting ($d = 5$ and $d = 8$). These test are simple extensions of good performing (Genest et al. 2009) GOF tests for bivariate copulas. In particular, we will compare to GOF tests based on the following quantities

- empirical copula process $\hat{C}_n(\boldsymbol{u}) - C_{\hat{\boldsymbol{\theta}}_n}(\boldsymbol{u})$, with $\boldsymbol{u} = (u_1, \ldots, u_d) \in [0,1]^d$,

$$\hat{C}_n(\boldsymbol{u}) = \frac{1}{n+1} \sum_{t=1}^{n} \mathbf{1}_{\{U_{t1} \le u_1, \ldots, U_{td} \le u_d\}}, \tag{4.4}$$

  and $C_{\hat{\boldsymbol{\theta}}_n}(\boldsymbol{u})$ being the copula with estimated parameter(s) $\hat{\boldsymbol{\theta}}_n$, and/or

- multivariate probability integral transform (PIT).

The multivariate PIT aggregation to univariate test data is facilitated using different aggregation functions. For the univariate test data then standard univariate GOF test statistics such as Anderson-Darling (AD), Cramér-von Mises (CvM) and Kolmogorov-Smirnov (KS) are used. In contrast, the empirical copula process (ECP) based test use the multivariate Cramér-von Mises (mCvM) and multivariate Kolmogorov-Smirnov (mKS) test statistics. The different GOF tests are given in the appendix for the convenience of the reader.

The power study will expose that the information based GOF tests outperform the other GOF tests in terms of size and power. The PIT based GOF tests reveal little to no power against the considered alternatives. But applying the PIT transformed data to the empirical copula process, as first suggested by Genest et al. (2009), is more promising. Here $C_{\hat{\boldsymbol{\theta}}_n}(\boldsymbol{u})$ is replaced by the independence copula $C_\perp$ in the ECP.

A first overview of the considered GOF tests is given in Figure 4.1. The two new GOF test are highlighted.

4.1 White's information matrix test

The first goodness-of-fit test we introduce for regular vine copula models arises directly from the information matrix equality and specification test proposed by White (1982) and extends the goodness-of-fit test for copulas introduced by Huang and Prokhorov (2013). Therefore, let us define first the misspecification test problem as already shortly touched in Equation (4.3).

4.1.1 The misspecification test problem

Let $\boldsymbol{U} = (U_1, \ldots, U_d)^T \in [0,1]^d$ be a d dimensional random vector with distribution function $G(\boldsymbol{u}) = C_{\boldsymbol{\theta}}(u_1, \ldots, u_d)$, where $C_{\boldsymbol{\theta}}$ is a d dimensional copula with parameter $\boldsymbol{\theta}$ and $U_i \sim \mathcal{U}(0,1)$ for $i = 1, \ldots, d$. Let $c_{\boldsymbol{\theta}}$ be the corresponding copula density, then

$$\begin{aligned} \mathbb{H}(\boldsymbol{\theta}) &= E\left[\partial_{\boldsymbol{\theta}}^2 ln(c_{\boldsymbol{\theta}}(U_1, \ldots, U_d))\right], \\ \mathbb{C}(\boldsymbol{\theta}) &= E\left[\partial_{\boldsymbol{\theta}} ln(c_{\boldsymbol{\theta}}(U_1, \ldots, U_d)) \big(\partial_{\boldsymbol{\theta}} ln(c_{\boldsymbol{\theta}}(U_1, \ldots, U_d))\big)^T\right] \end{aligned} \tag{4.5}$$

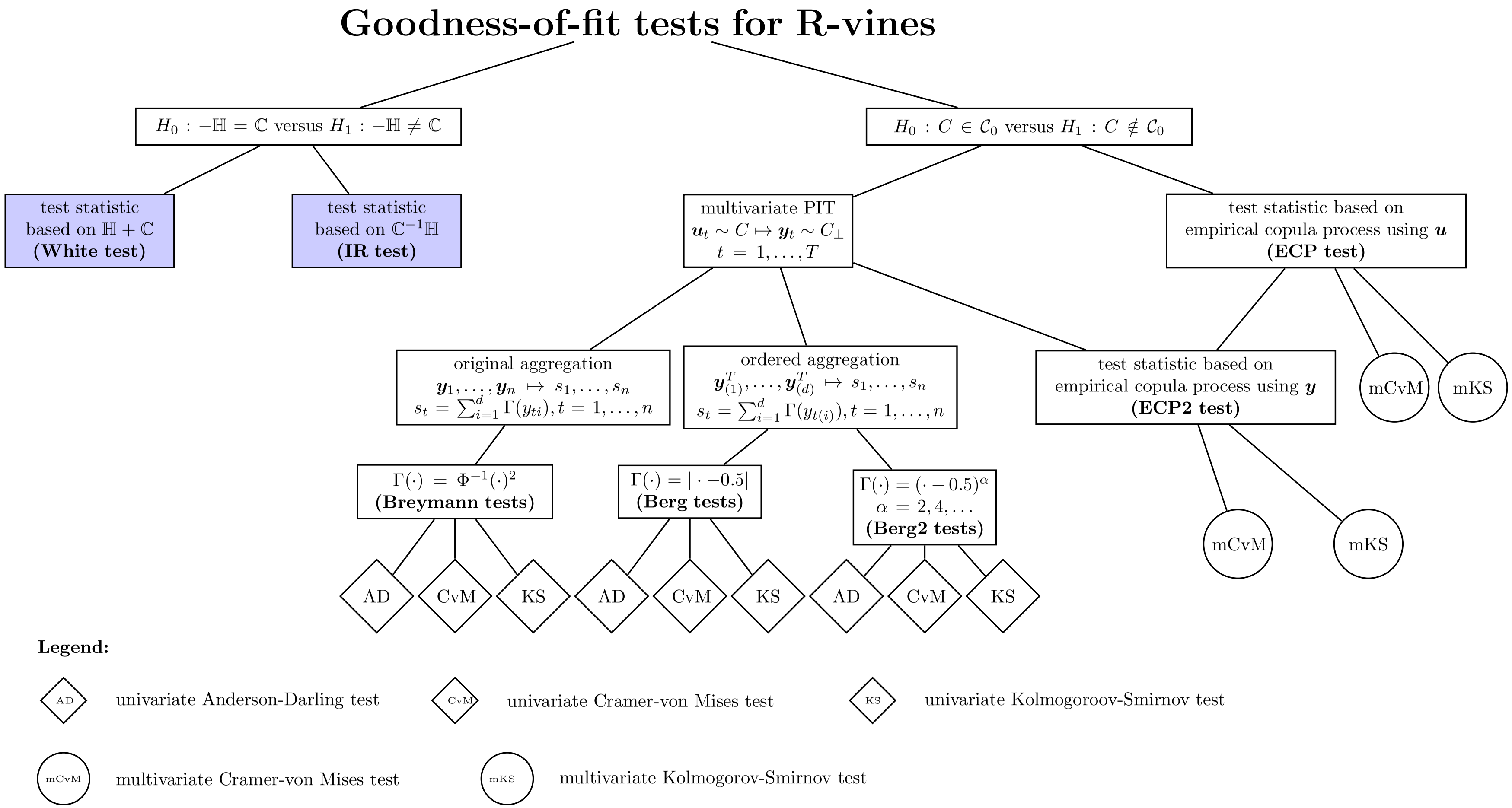


Figure 4.1: Structured overview of the suggested goodness-of-fit test hypotheses and their test statistics.

are the expected Hessian matrix of $ln(c_{\boldsymbol{\theta}}(u_1, \ldots, u_d))$ and the expected outer product of the corresponding score function, respectively, where $\partial_{\boldsymbol{\theta}}$ denotes the gradient with respect to the copula parameter $\boldsymbol{\theta}$. Now, the theorem of White (1982) shows that under correct model specification ($\boldsymbol{\theta} = \boldsymbol{\theta}_0$) the negative expected Hessian matrix $\mathbb{H}(\boldsymbol{\theta}_0)$ and the expected outer product of the corresponding score function $\mathbb{C}(\boldsymbol{\theta}_0)$ are equal, see Equation (4.2). Thus, the corresponding vine copula misspecification test problem is therefore given by

$$H_0 : \mathbb{H}(\boldsymbol{\theta}_0) + \mathbb{C}(\boldsymbol{\theta}_0) = 0 \text{ against } H_1 : \mathbb{H}(\boldsymbol{\theta}_0) + \mathbb{C}(\boldsymbol{\theta}_0) \neq 0,$$

as already stated in Equation (4.3). Here we will expand the GOF test of Huang and Prokhorov (2013) who used White's information matrix equality (4.2) to the vine copula. In our setting we replace the d-dimensional density $c_{\boldsymbol{\theta}}(u_1, \ldots, u_d)$ in (4.5) with the vine density $c_{1,\ldots,d}(u_1, \ldots, u_d)$ given in (2.3), with a parametric vine copula, i.e. the parameter vector $\boldsymbol{\theta}$ is given by $\boldsymbol{\theta} = (\theta_{j(e),k(e);D(e)}, e \in E_i, i = 1, \ldots, d)$. Remember we assume copula data, i.e. known margins.

4.1.2 Goodness-of-fit test for R-vine copulas

The first step in the development of a GOF test for the testing problem given in (4.3) will be the estimation of the Hessian matrix $\mathbb{H}(\boldsymbol{\theta})$ and the outer product of gradient $\mathbb{C}(\boldsymbol{\theta})$. For this we assume the availability of an i.i.d. sample in d dimensions of size n, denoted by $\boldsymbol{u} := (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$ in the copula space. This pseudo data $\boldsymbol{u}$ is used to estimate the unknown true parameter $\boldsymbol{\theta}_0$ of the vine copula. Algorithms for maximum likelihood estimation for vine copula parameters are given in e.g. Czado et al. (2012) for C-vine copulas or Dißmann et al. (2013) for R-vine copulas. We call the estimate derived from these algorithms applied to $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ pseudo maximum likelihood (ML) estimate and denote it by $\hat{\boldsymbol{\theta}}_n = \hat{\boldsymbol{\theta}}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$. Assuming that $\boldsymbol{U}$ follows an R-vine copula model, i.e. $\boldsymbol{U} \sim RV(\mathcal{V}, \mathcal{B}(\mathcal{V}), \boldsymbol{\theta}(\mathcal{B}(\mathcal{V})))$ we define the random matrices

$$\mathbb{H}(\boldsymbol{\theta}|\boldsymbol{U}) := \frac{\partial^2}{\partial^2 \boldsymbol{\theta}} l(\boldsymbol{\theta}|\boldsymbol{U}) \qquad \text{and} \qquad \mathbb{C}(\boldsymbol{\theta}|\boldsymbol{U}) := \frac{\partial}{\partial \boldsymbol{\theta}} l(\boldsymbol{\theta}|\boldsymbol{U}) \left(\frac{\partial}{\partial \boldsymbol{\theta}} l(\boldsymbol{\theta}|\boldsymbol{U}) \right)^T \tag{4.6}$$

for the second derivative of the log-likelihood function $l(\boldsymbol{\theta}|\boldsymbol{U})$ and the outer product of the score function, respectively. Further, we denote the sample counter parts of the Hessian matrix and the outer product of the score function for the copula data $\boldsymbol{u}_t = (u_{1t}, \ldots, u_{dt})^T, 1 \leq t \leq n$ at $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}_n$ of a vine copula with density given in (2.3) (in (2.3) the margins are assumed to be uniform) by

$$\hat{\mathbb{H}}_t(\hat{\boldsymbol{\theta}}_n) := \mathbb{H}(\hat{\boldsymbol{\theta}}_n|\boldsymbol{u}_t) \in \mathbb{R}^{p \times p} \qquad \text{and} \qquad \hat{\mathbb{C}}_t(\hat{\boldsymbol{\theta}}_n) := \mathbb{C}(\hat{\boldsymbol{\theta}}_n|\boldsymbol{u}_t) \in \mathbb{R}^{p \times p}.$$

Here p is the length of the parameter vector $\boldsymbol{\theta}$. Thus, the sample equivalents to $\mathbb{H}(\boldsymbol{\theta})$ (expected Hessian) and $\mathbb{C}(\boldsymbol{\theta})$ (expected outer product of gradient) for the pseudo ML estimate $\hat{\boldsymbol{\theta}}_n$ are

$$\bar{\mathbb{H}}(\hat{\boldsymbol{\theta}}_n) := \frac{1}{n} \sum_{t=1}^{n} \hat{\mathbb{H}}_t(\hat{\boldsymbol{\theta}}_n) \qquad \text{and} \qquad \bar{\mathbb{C}}(\hat{\boldsymbol{\theta}}_n) := \frac{1}{n} \sum_{t=1}^{n} \hat{\mathbb{C}}_t(\hat{\boldsymbol{\theta}}_n), \tag{4.7}$$

given n observations. In Chapter 3 we provided algorithms for the calculation of the gradient as well as of the Hessian matrix for R-vines. Alternatively numerical versions based on finite differences can be used as well, unless they are not precise enough.

Note that the matrices $\hat{\mathbb{H}}_t(\hat{\boldsymbol{\theta}}_n)$ and $\hat{\mathbb{C}}_t(\hat{\boldsymbol{\theta}}_n)$ are of size $d(d-1)/2 \times d(d-1)/2$ if all pair-copulas are one-parametric. In this case we have $p = d(d-1)/2$. For each two or higher parametrized bivariate copula in the vine the dimension of the matrices increases by 1 or the number of additional parameters. An example for a two-parameter bivariate copula is the bivariate Student's t-copula. But the dimension of the information matrices decrease if independence copulas are used in the pair-copula construction. This often appears in higher trees since the dependence usually decreases as number of trees increases. For higher dimensional vines a truncation, i.e. setting all pair-copula families of higher order trees to the independence copula, may be helpful to reduce the number of parameters significantly (Brechmann et al. 2012).

To formulate the test statistic we vectorize the sum of the expected Hessian matrix $\mathbb{H}(\boldsymbol{\theta}_0)$ and the expected outer product of gradient $\mathbb{C}(\boldsymbol{\theta}_0)$. Therefore we define the random vector

$$\boldsymbol{d}(\boldsymbol{\theta}|\boldsymbol{U}) := vech(\mathbb{H}(\boldsymbol{\theta}|\boldsymbol{U}) + \mathbb{C}(\boldsymbol{\theta}|\boldsymbol{U})) \in \mathbb{R}^{\frac{p(p+1)}{2}}$$

and its empirical version by

$$\hat{\boldsymbol{d}}_t(\hat{\boldsymbol{\theta}}_n) := \boldsymbol{d}(\hat{\boldsymbol{\theta}}_n|\boldsymbol{u}_t) \qquad \text{and} \qquad \bar{\boldsymbol{d}}(\hat{\boldsymbol{\theta}}_n) := \frac{1}{n}\sum_{t=1}^{n} \hat{\boldsymbol{d}}_t(\hat{\boldsymbol{\theta}}_n) \in \mathbb{R}^{\frac{p(p+1)}{2}}.$$

Note that because of symmetry only the lower triangle (including the diagonal) of the matrices has to be vectorized. Further, assuming existence of derivative and finite expectation we define the expected gradient matrix of the random vector $\boldsymbol{d}(\boldsymbol{\theta}|\boldsymbol{U})$ as

$$\nabla D_{\boldsymbol{\theta}} := E\left[\partial_{\boldsymbol{\theta}_k} \boldsymbol{d}_l(\boldsymbol{\theta}|\boldsymbol{U})\right]_{l=1,\dots,\frac{p(p+1)}{2},k=1,\dots,p} \in \mathbb{R}^{\frac{p(p+1)}{2}\times p}, \quad \text{and}$$
$$\widehat{\nabla D_{\boldsymbol{\theta}}} := \frac{1}{n}\sum_{t=1}^{n}\left[\partial_{\boldsymbol{\theta}_k} \hat{\boldsymbol{d}}_l(\hat{\boldsymbol{\theta}}_n|\boldsymbol{u}_t)\right]_{l=1,\dots,\frac{p(p+1)}{2},k=1,\dots,p} \in \mathbb{R}^{\frac{p(p+1)}{2}\times p}$$

as its estimate. White (1982, Appendix) derived the corresponding asymptotic covariance matrix of $\sqrt{n}\bar{\boldsymbol{d}}(\hat{\boldsymbol{\theta}}_n)$, which is given by

$$V_{\boldsymbol{\theta}_0} = \mathbb{E}\left[\left(\boldsymbol{d}(\boldsymbol{\theta}_0|\boldsymbol{U}) - \nabla D_{\boldsymbol{\theta}_0}\mathbb{H}(\boldsymbol{\theta}_0)^{-1}\partial_{\boldsymbol{\theta}_0} l(\boldsymbol{\theta}_0|\boldsymbol{U})\right)\left(\boldsymbol{d}(\boldsymbol{\theta}_0|\boldsymbol{U}) - \nabla D_{\boldsymbol{\theta}_0}\mathbb{H}(\boldsymbol{\theta}_0)^{-1}\partial_{\boldsymbol{\theta}_0} l(\boldsymbol{\theta}_0|\boldsymbol{U})\right)^T\right] \tag{4.8}$$

In particular $\sqrt{n}\bar{\boldsymbol{d}}(\hat{\boldsymbol{\theta}}_n) \xrightarrow{d} N(0, V_{\boldsymbol{\theta}_0})$, as $n \to \infty$.
The following proposition of Whites theorem is valid under the assumptions A1-A10 of White (1982). This assures that $l(\hat{\boldsymbol{\theta}}_n|\boldsymbol{u}_t)$ is a continuous measurable function and its derivatives exist; A10 assumes that $V_{\boldsymbol{\theta}_0}$ is nonsingular.

Proposition 4.1
Under the correct vine copula specification and suitable regularity conditions (A1-A10 in White 1982) the information matrix test statistic is defined as

$$\mathcal{T}_n = n\left(\bar{\boldsymbol{d}}(\hat{\boldsymbol{\theta}}_n)\right)^T \hat{V}_{\hat{\boldsymbol{\theta}}_n}^{-1}\bar{\boldsymbol{d}}(\hat{\boldsymbol{\theta}}_n), \tag{4.9}$$

where $\hat{V}_{\hat{\boldsymbol{\theta}}_n}^{-1}$ *is an consistent estimate for the inverse asymptotic covariance matrix* $V_{\boldsymbol{\theta}_0}$*. It follows that* $\mathcal{T}_n$ *is asymptotically* $\chi^2_{p(p+1)/2}$ *distributed.*

The proof is an extension of the proof of White (1982) since the maximum likelihood estimator $\hat{\boldsymbol{\theta}}_n$ in a vine copula is also to be shown normally distributed (see Hobæk Haff 2013 and Equation (3.1) in Chapter 3). Since the asymptotic distribution is independent of model parameters the test is asymptotically pivotal. The χ^2-distribution only depends on the parameter vector dimension p, which is known aforehand given the pair-copula families, and not on $\boldsymbol{\theta}_0$. Furthermore, the test is a so called "blanket" test in the sense of Genest et al. (2009).

Note that all calculations are performed using copula data, thus ignoring the uncertainty in the margins. In Section 4.1.4 we extend our GOF test adjusting $V_{\boldsymbol{\theta}_0}$ for the estimation of the margins. For general multivariate copulas this is already done by Huang and Prokhorov (2013). Given the test statistic $\mathcal{T}_n$ of Proposition 4.1 we can define an α-level test.

Corollary 4.2
Let $\alpha \in (0,1)$ *and* $\mathcal{T}_n$ *as in Proposition 4.1. Then the test*

$$\begin{aligned}&\textit{Reject } H_0 : \mathbb{H}(\boldsymbol{\theta}_0) + \mathbb{C}(\boldsymbol{\theta}_0) = 0 \quad \textit{versus} \quad H_1 : \mathbb{H}(\boldsymbol{\theta}_0) + \mathbb{C}(\boldsymbol{\theta}_0) \neq 0 \\ &\Leftrightarrow \mathcal{T} > \left(\chi^2_{p(p+1)/2}\right)^{-1}(1-\alpha)\end{aligned}$$

is an asymptotic α*-level test. Here* $\left(\chi^2_{df}\right)^{-1}(\beta)$ *denotes the* β *quantile of a* χ^2_{df} *distribution with df degrees of freedom.*

4.1.3 Power studies

A GOF test's performance is usually measured by its power, which is often unknown. In this case, it can only be investigated by simulations. Given a significance level α a high power at a specified alternative indicates a good discrimination against the alternative. In this section we will investigate the introduced test for a suitable large n and under a variety of alternatives. In particular we determine the power function under the true data generating process (DGP) and under the alternative DGP. In the first case the null hypothesis H_0 holds, while in the second case the null hypothesis H_0 does not hold. The power at the true DGP assesses the ability of the test in Corollary 4.2 to maintain its nominal level.

Performance measures

As performance measures we use p-value plots and size-power curves, introduced by Davidson and MacKinnon (1998) and explained in the following.

Given an observed test statistic $\mathcal{T}_n = t$ the p-value at t is

$$p(t) := P(\mathcal{T}_n \geq t),$$

i.e. the smallest α level at which the test can reject H_0 when $\mathcal{T}_n = t$ is observed. The distribution of $\mathcal{T}_n$ has to be known at least asymptotically such as the χ^2-distribution in our case or has to be estimated empirically using simulated data.

Let $Z_{M_1} := p(\mathcal{T}_n(M_1))$ be the random variable on $[0,1]$ with value $p(t)$ when $\mathcal{T}_n = t$ is observed and data is generated from model M_1. The distribution function for Z_{M_1} we denote by $F_{M_1}(\cdot)$, i.e. $F_{M_1}(\alpha) := P(Z_{M_1} \leq \alpha)$, being the size of the test. To estimate Z_{M_1} and F_{M_1} we assume B realizations of the test statistic $\mathcal{T}_n(M_1)$ when n observations are generated from model M_1, denoted as $t_n^j(M_1), j = 1, \dots, B$, and estimate the p-values $p_{M_1}^j$ as

$$\hat{p}_{M_1}^j := \hat{p}(t_n^j(M_1)) := \frac{1}{B} \sum_{r=1}^{R} \mathbf{1}_{\{t_n^r(M_1) \geq t_n^j(M_1)\}}$$

and consider the empirical distribution function of them. Thus,

$$\hat{F}_{M_1}(\alpha) := \frac{1}{B} \sum_{r=1}^{B} \mathbf{1}_{\{\hat{p}_{M_1}^r \leq \alpha\}} \quad \alpha \in (0,1),$$

forms an estimate for the size of the test at level α. $\hat{F}_{M_1}(\alpha)$ is called *actual* size or *actual* alpha (probability of the outcome under the null hypothesis), while α is known as *nominal* size.

p-value plot: The p-value plot graphs α versus $\hat{F}_{M_1}(\alpha)$, i.e. *nominal* size against *actual* size. The plot indicates if the test reaches its *nominal* size, i.e. if the assumed asymptotic holds.

Size-power curve: We are not only interested in the size of the test but also in its power. For data generated under model M_2 we want to determine $F_{M_2}(\alpha)$, which gives the power of the test when H_1 is true, i.e. M_2 holds, and a level α is used for the test. Generate B i.i.d. data sets from M_2 and use them to estimate $F_{M_2}(\alpha)$ by $\hat{F}_{M_2}(\alpha)$. The plot of $\hat{F}_{M_1}(\alpha)$ versus $\hat{F}_{M_2}(\alpha)$ is called the size-power curve. A test with good power should have small power when the size is small and approach power one rapidly as the size increases.

Remark: Note that these curves correspond to the better known Receiver-operating-characteristic curves (ROC), which plot the fraction of true positives (TP) out of the positives (TPR = true positive rate) versus the fraction of false positives (FP) out of the negatives (FPR = false positive rate), at various threshold settings. TPR is also known as sensitivity, and FPR is one minus the specificity or true negative rate. False positive (FP) is better known as Type I error. Since α (*nominal* size) is the false positive rate and power is one minus the false negative rate (FNR; prob. of of a Type II error occurring), ROC plots *nominal* size versus power (Fawcett 2006).

General remarks on size and power, and the implementation of the test

It is shown for other statistical models that the size behavior of the information matrix test (IMT) is very poor. E.g. in the regression model context the IMT has poor size properties even in samples of size 1000 and more (see Hall 1989 and references within, especially Taylor 1987). Similar observations are made by e.g. Chesher and Spady (1991).

In the bivariate copula case the asymptotic approximation holds even for relative small number of observations. But this is not investigated or documented in Huang and Prokhorov (2013).

In this simulation study three possible errors can occur, which may influence the asymptotic behavior: simulation error, estimation and model error, and numerical errors. Simulation errors are always involved since only pseudo random variables can be generated on a computer. For the parameter estimation maximum likelihood is used based on Newton-Raphson algorithms for maximization. In higher dimensions this can be quite challenging, even given the analytical gradient (and the analytical Hessian matrix). A local maximum may be returned. Further, numerical instabilities can occur, especially in the calculation of the score function and the Hessian matrix as discussed in Chapter 3.

Even the estimator of $V_{\boldsymbol{\theta}_0}$ may not be positive definite, though this becomes increasingly unlikely as the sample size increases.

Furthermore, the normal asymptotic theory only holds for full maximum likelihood, but a sequential maximum likelihood, i.e. a tree-wise estimation, is performed due to resource and time limits. Usually sequential estimates are close to full ML estimates, except of the degree-of-freedom parameter ν of the Student's t copula. There exists even an asymptotic theory for sequential estimates similar to Equation (3.1), see Hobæk Haff (2013).

General simulation setup

We test if our goodness-of-fit test $\mathcal{T}_n$ based on the vine copula model $M_1 = RV(\mathcal{V}_1, \mathcal{B}_1(\mathcal{V}_1), \boldsymbol{\theta}_1(\mathcal{B}_1(\mathcal{V}_1)))$ has suitable power against an alternative vine copula model $M_2 = RV(\mathcal{V}_2, \mathcal{B}_2(\mathcal{V}_2), \boldsymbol{\theta}_2(\mathcal{B}_2(\mathcal{V}_2)))$, where $M_2 \neq M_1$. To produce the corresponding p-value plots for M_1 and the size-power curves we proceed as follows:

1. Set vine copula model M_1.

2. Generate a copula data sample of size $n = 1000$ from model M_1 (pre-run).

3. Given the data of the pre-run select and estimate M_2 using e.g. the step-wise selection algorithm of Dißmann et al. (2013).

4. For $r = 1, \ldots, B$

    - Generate copula data $\boldsymbol{u}^r_{M_1} = (\boldsymbol{u}^{1r}_{M_1}, \ldots, \boldsymbol{u}^{dr}_{M_1})$ from M_1 of size n.
    - Estimate $\boldsymbol{\theta}_1(\mathcal{B}_1(\mathcal{V}_1))$ of model M_1 given data $\boldsymbol{u}^r_{M_1}$ and denote it by $\hat{\boldsymbol{\theta}}_1(\mathcal{B}_1(\mathcal{V}_1); \boldsymbol{u}^r_{M_1})$.
    - Calculate test statistic $t^r_n(M_1) := t^r_n(\hat{\boldsymbol{\theta}}_1(\mathcal{B}_1(\mathcal{V}_1); \boldsymbol{u}^r_{M_1}))$ based on data $\boldsymbol{u}^r_{M_1}$ assuming the vine copula model $M_1 = RV(\mathcal{V}_1, \mathcal{B}_1(\mathcal{V}_1), \hat{\boldsymbol{\theta}}_1(\mathcal{B}_1(\mathcal{V}_1)))$.

- Calculate asymptotic p-values $p(t_n^r(M_1)) = (\chi^2_{p(p+1)/2})^{-1}(t_n^r(M_1))$, where p is the number of parameters of $\boldsymbol{\theta}_1(\mathcal{B}_1(\mathcal{V}_1))$.
- Generate copula data $\boldsymbol{u}_{M_2}^r = (\boldsymbol{u}_{M_2}^{1r}, \ldots, \boldsymbol{u}_{M_2}^{dr})$ from M_2 of size n.
- Estimate $\boldsymbol{\theta}_1(\mathcal{B}_1(\mathcal{V}_1))$ of model M_1 given data $\boldsymbol{u}_{M_2}^r$ and denote it by $\hat{\boldsymbol{\theta}}_1(\mathcal{B}_1(\mathcal{V}_1); \boldsymbol{u}_{M_2})$.
- Calculate test statistic $t_n^r(M_2) := t_n^r(\hat{\boldsymbol{\theta}}_1(\mathcal{B}_1(\mathcal{V}_1); \boldsymbol{u}_{M_2}^r))$ based on data $\boldsymbol{u}_{M_2}^r$ assuming vine copula model M_1.
- Calculate asymptotic p-values $p(t_n^r(M_2)) = (\chi^2_{p(p+1)/2})^{-1}(t_n^r(M_2))$.

end for

5. Estimate p-values $p_{M_1}^j$ and $p_{M_2}^j$ by

$$\hat{p}_{M_1}^j = \hat{p}(t_n^j(M_1)) := \frac{1}{B}\sum_{r=1}^{B} \mathbf{1}_{\{t_n^r(M_1) \geq t_n^j(M_1)\}} \quad \text{and}$$

$$\hat{p}_{M_2}^j = \hat{p}(t_n^j(M_2)) := \frac{1}{B}\sum_{r=1}^{B} \mathbf{1}_{\{t_n^r(M_2) \geq t_n^j(M_2)\}},$$

respectively, for $j = 1, \ldots, B$.

6. Estimate the distribution function of Z_{M_1} and Z_{M_2} by

$$\hat{F}_{M_1}(\alpha) = \frac{1}{B}\sum_{r=1}^{B} \mathbf{1}_{\{\hat{p}_{M_1}^r \leq \alpha\}} \quad \text{and} \quad \hat{F}_{M_2}(\alpha) = \frac{1}{B}\sum_{r=1}^{B} \mathbf{1}_{\{\hat{p}_{M_2}^r \leq \alpha\}},$$

respectively.

The following simulation results are based on $B = 10000$ replications and the number of observations n are chosen to be 300, 500, 750 or 1000. The dimension of the vine copula models $M_i, i = 1, 2$ is 5. Possible pair-copula families are the elliptical Gauss and Student's t-copula, the Archimedean Clayton, Gumbel, Frank and Joe copula, and the rotated Archimedean copulas. A p-value plot to assess the nominal size of the test is achieved by plotting α versus $\hat{F}_{M_1}(\alpha)$. Evaluating $\hat{F}_{M_1}(\alpha)$ and $\hat{F}_{M_2}(\alpha)$ on the grid

$$\alpha = 0.001, 0.002, \ldots, 0.010, 0.015, \ldots, 0.990, 0.991, \ldots, 0.999$$

with smaller grid size near 0 and 1 we can plot a size-power curve $\hat{F}_{M_1}(\alpha)$ versus $\hat{F}_{M_2}(\alpha)$.

All calculations are performed with **R** (R Development Core Team 2013), the R-package **VineCopula** of Schepsmeier et al. (2012) (see Section 2.3) and the **copula**-package (Yan 2007; Kojadinovic and Yan 2010; Hofert et al. 2013).

Specific setting

In the following three power studies we investigate the properties of the introduced test with respect to its size and power. In the first power study we determine the power of the test assuming an R-vine copula as true model (M_1 in the notation from above) under three alternatives of simpler copula models such as

- the multivariate Gauss copula,
- the C-vine copula and
- the D-vine copula,

which are special cases of the R-vine. Every multivariate Gaussian copula can be written as a vine copula with Gaussian pair-copulas and vice versa (Czado 2010). Only in the Gaussian case the conditional correlation parameters, forming the pair-copula parameters, are equal to the partial correlation parameter, which can be calculated recursively using the entries of the multivariate Gauss copula variance-covariance matrix.

The second power study investigates the power of the test between two R-vines, which are chosen with two different selection methods - a maximum spanning tree approach introduced by Dißmann et al. (2013) versus a Bayesian approach (MCMC) investigated by Gruber and Czado (2012), based on a generated data set given a specified R-vine copula model.

In a third simulation study we compare the often used multivariate t-copula under the alternative of an R-vine with only bivariate t-copulas, and vise versa. The difference is the common degree-of-freedom parameter ν in the multivariate t-copula versus variable, separately estimated νs in the R-vine model. The correlation parameters ρ can be set/estimated such as in the Gaussian case described above.

Table 4.1 gives an overview of all three power studies, their true model and their alternatives.

Study	True model (M_1)	Alternative (M_2)
I	R-vine $(\mathcal{V}_R, \mathcal{B}_R(\mathcal{V}_R), \boldsymbol{\theta}_R(\mathcal{B}_R(\mathcal{V}_R)))$	multivariate Gauss
		C-vine $(\mathcal{V}_C, \mathcal{B}_C(\mathcal{V}_C), \boldsymbol{\theta}_C(\mathcal{B}_C(\mathcal{V}_C)))$
		D-vine $(\mathcal{V}_D, \mathcal{B}_D(\mathcal{V}_D), \boldsymbol{\theta}_D(\mathcal{B}_D(\mathcal{V}_D)))$
II	R-vine $(\mathcal{V}_R, \mathcal{B}_R(\mathcal{V}_R), \boldsymbol{\theta}_R(\mathcal{B}_R(\mathcal{V}_R)))$	R-vine estimated by Dißmann et al. (2013)
		$(\mathcal{V}_{MST}, \mathcal{B}_{MST}(\mathcal{V}_{MST}), \boldsymbol{\theta}_{MST}(\mathcal{B}_{MST}(\mathcal{V}_{MST})))$
		R-vine estimated by Gruber and Czado (2012)
		$(\mathcal{V}_B, \mathcal{B}_B(\mathcal{V}_B), \boldsymbol{\theta}_B(\mathcal{B}_B(\mathcal{V}_B)))$
III	multivariate t-copula	R-vine estimated by Dißmann et al. (2013)
		with only Student's t-copulas
	R-vine with only t-copulas	multivariate t-copula

Table 4.1: Overview of the studied test settings

Power study I

We investigated three variants of the dependence:

- M_1 with mixed Kendall's τ values,
- M_1 with constant low ($\tau = 0.25$) Kendall's τ values and
- M_1 with constant medium ($\tau = 0.5$) Kendall's τ values

for the $d(d-1)/2$ pair-copulas. An R-vine with constant high dependencies is omitted since the power in the medium case are already very high and allow to draw conclusions for the high dependency case. The structure of the chosen R-vine is given in Figure C.1 and Equation (C.1) of Appendix C.1. The chosen bivariate copula families and Kendall's τ values can be found in Table C.1 of Appendix C.1.

The selected D-vine structure (Step 3 in the test procedure) is already defined by the ordering of its variables in the first tree. Here the ordering is 3-4-5-1-2, see Equation (C.3) of Appendix C.1. Similarly, the C-vine structure is defined by its root nodes. The root in the first tree is variable 2 while in the second tree variable 1 is added to the root, i.e. the root in Tree 2 is 1,2. Variable 4, 5 and 3 are added in the next trees, respectively, see Equation (C.2) of Appendix C.1. The selected copula structure and pair-copula parameters in Step 3 are quite stable given more than one data set in Step 2, i.e. no changes in the vine copula structure and minor changes in the pair-copula choice (e.g. the algorithm selects a rotated Gumbel copula instead of a Clayton copula, which are quite similar given a low Kendall's τ). Neglecting possible small variations in the pair-copula selection given R different data sets we fix the C- and D-vine structure as well as the pair-copula family selection after one run of Step 2.

An overview of all investigated models is given in Table 4.2.

Model	$\mathcal{V}$	$\mathcal{B}(\mathcal{V})$	$\boldsymbol{\theta}(\mathcal{B}(\mathcal{V}))_1$	$\boldsymbol{\theta}(\mathcal{B}(\mathcal{V}))_2$	$\boldsymbol{\theta}(\mathcal{B}(\mathcal{V}))_3$
R-vine	Eq. (C.1), Fig. C.1	Tab. C.1	Tab. C.1	$\tau = 0.25$	$\tau = 0.5$
Gauss	-	Gauss	$\hat{\boldsymbol{\theta}}_1$	$\hat{\boldsymbol{\theta}}_2$	$\hat{\boldsymbol{\theta}}_3$
C-vine	$\hat{\mathcal{V}}_C$, Eq. (C.2)	$\hat{\mathcal{B}}_C(\hat{\mathcal{V}}_C)$	$\hat{\boldsymbol{\theta}}_C(\hat{\mathcal{B}}_C(\hat{\mathcal{V}}_C))_1$	$\hat{\boldsymbol{\theta}}_C(\hat{\mathcal{B}}_C(\hat{\mathcal{V}}_C))_2$	$\hat{\boldsymbol{\theta}}_C(\hat{\mathcal{B}}_C(\hat{\mathcal{V}}_C))_3$
D-vine	$\hat{\mathcal{V}}_D$, Eq. (C.3)	$\hat{\mathcal{B}}_D(\hat{\mathcal{V}}_D)$	$\hat{\boldsymbol{\theta}}_D(\hat{\mathcal{B}}_D(\hat{\mathcal{V}}_D))_1$	$\hat{\boldsymbol{\theta}}_D(\hat{\mathcal{B}}_D(\hat{\mathcal{V}}_D))_2$	$\hat{\boldsymbol{\theta}}_D(\hat{\mathcal{B}}_D(\hat{\mathcal{V}}_D))_3$

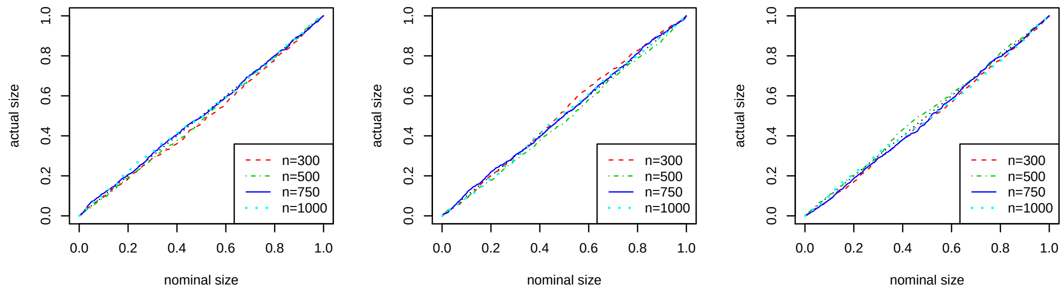
Table 4.2: Model specifications

Results with regard to nominal size: A p-value plot (α versus $\hat{F}_{M_1}(\alpha)$) for the simulated results shows that the test works perfectly under the null independent of the number of observations, since the p-value plot fits the 45 degree line nearly perfect (see Figure 4.2a). That means that the test reaches its nominal level in the case of simulated p-values.

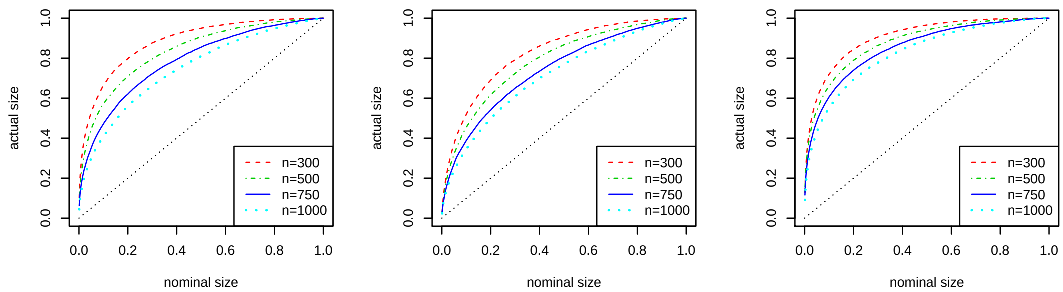
Given the asymptotic p-values $p(t_n^r(M_1)) = (\chi^2_{p(p+1)/2})^{-1}(t_n^r(M_1))$ and their corresponding empirical distribution function $\hat{F}_{M_1}^{asy}(\alpha)$ we have a different picture. The *actual* size is much greater than the *nominal* size (see Figure 4.2b). Also the test does not hold its nominal level in the case of asymptotic p-values given a small/medium data set. The test over-rejects quite too often based on asymptotic p-values.

Comparing the finite sample distribution of the test statistic with the theoretical χ^2 distribution in Figure 4.3a (left panel) we can clearly see, that even in the case of 1000 observation points the χ^2 distribution does not fit the empirical distribution given the observed test statistics $t_n^r(M_1)$. For the investigated 5-dimensional case the empirical distribution fits the theoretical one not until $n = 10000$. In this case the actual size is the nominal size (see Figure 4.3a, right panel).

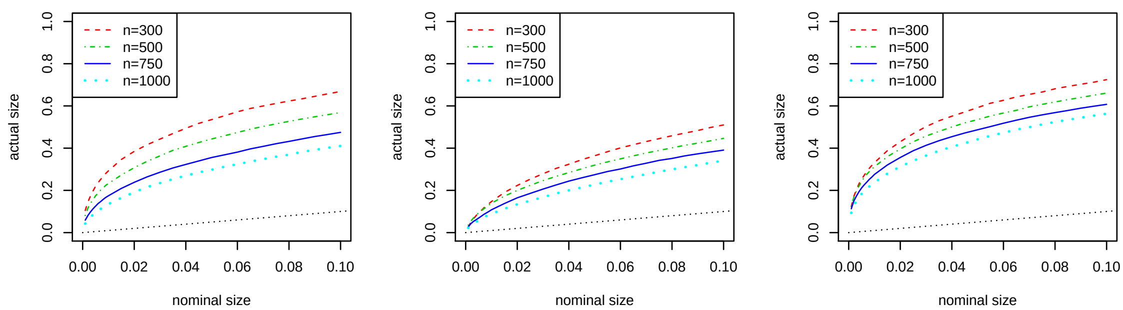
Additionally, we investigated the size behavior in a 8-dimensional vine copula model, whose details are not provided here. Figure 4.3b clearly illustrates that the asymptotic theory based test is too conservative, while in the 5-dimensional case the test is too liberal. Even for a sample size of $n = 20000$ for $d = 8$ the the actual size does not reach the nominal size.



(a) p-value plots using simulated p-values

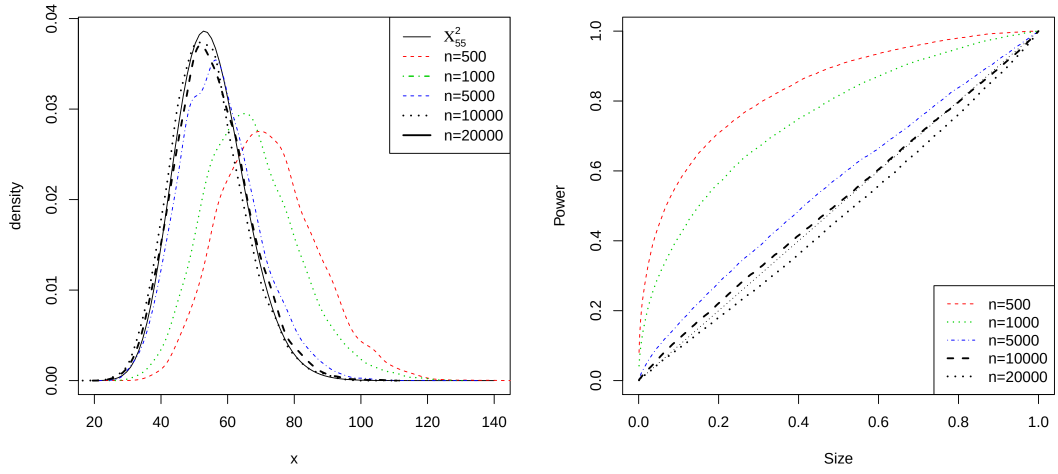


(b) p-value plots using asymptotic p-values

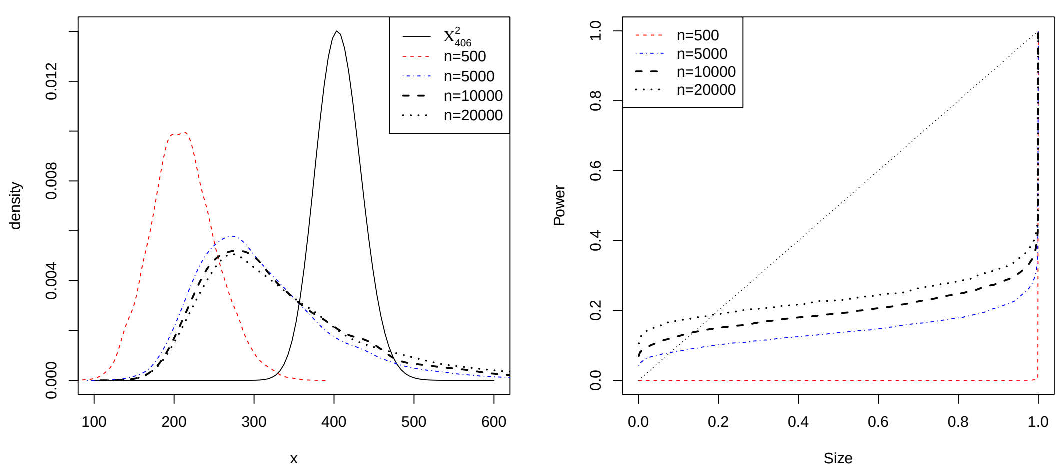


(c) detail plots of (b)

Figure 4.2: p-value plots for the three different scenarios; left: mixed Kendall's τ, center: constant Kendall's $\tau = 0.25$, right: constant Kendall's $\tau = 0.5$.



(a) $d = 5$



(b) $d = 8$

Figure 4.3: Empirical density plot (left panel) and p-value plot for the asymptotic p-values (right panel)

Results with regard to power: In Figure 4.4 we show the behavior of the size-power curve for varying number of observations n in each scenario. The number of observations increase from $n = 300$ in the upper left panel over $n = 500$ (upper right panel) and $n = 750$ (lower left panel) to $n = 1000$ in the lower right panel. Due to the results of the p-value plots we only consider the results of the simulated p-values in the following. The dotted diagonal line represents the case where size (x-axes) equals power (y-axes). In addition, we list in Table 4.3 the power for $n = 500$ at *nominal* size 5% in several senarios. Given a true vine model different vine models are tested. E.g. given a C-vine with Kendall's τ value of 0.25 the test for an R-vine returned on average a power of 18.4%

using simulated p-values (Simul.) and 59.9% using asymptotic p-values (Asy.).

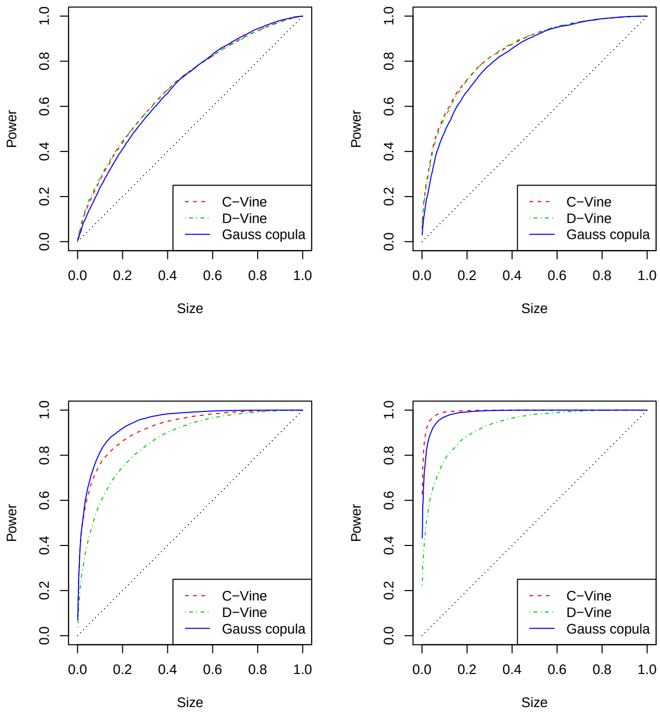
The first observation evaluating the plots is that the power is always greater than the size. This indicates a good performance of the test in mean. Further, an increasing number of observations increases the power of the test. In the medium dependence scenarios the tests are consistent since the power reaches one at quite low size.

Further conclusions are:

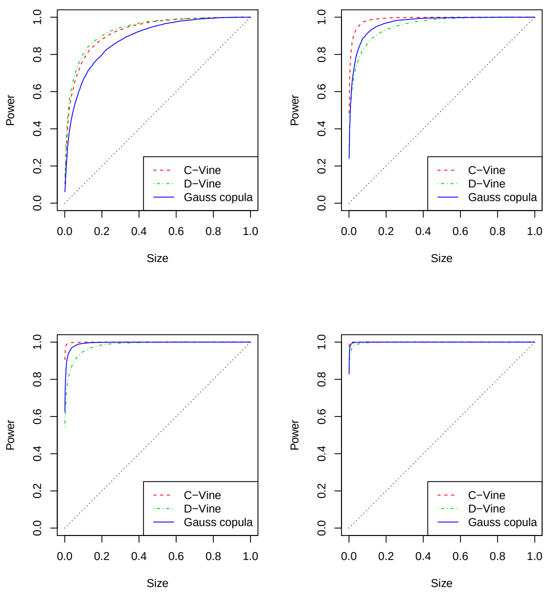
- The size-power curves of the C-vine and the D-vine are close to each other in each scenario. This changes in higher dimensional data sets (a 8-dimensional scenario was performed but is not documented here in detail). In the 5-dimensional case all vine structures are very similar, i.e. a change of one edge can change an R-vine into a C- or D-vine.
- The Gauss model is the first detected model, i.e. its size-power curve is the steepest and outperforms the other two. The C- and D-vine are more flexible in their choice of pair-copula families and thus can fit the data better.
- If the number of observations is too low, e.g. $n = 300$, conclusions are less robust. This weakness can be often observed in the inferential context and is e.g. documented for other goodness-of-fit tests for copulas in the comparison study of Genest et al. (2009).
- Very low dependencies yield to flatter size-power curves. If Kendall's τ is very small (in absolute terms) all copulas are close to the product copula and the choice of the vine structure as well as of the pair-copula families are less significant. Increasing power by increasing strength of dependence are already observed in other goodness-of-fit tests for copulas, see e.g. Genest et al. (2009) or Genest et al. (2012).
- Additional simulation studies with 8-dimensional vine copula models show that the power decreases (slightly) for increasing dimension. With increasing dimension the number of pair-copulas and thus the number of copula parameters increases, e.g. in scenario 3 (Kendall's $\tau = 0.5$) the power of an R-vine copula against a C-vine copula decreases at size 5% from 92% to 57%, and against a Gauss copula from 78% to 20%, but against a D-vine copula it increases from 82% to 89%. For bigger size the gap shrinks. Our assumption is that numerical instabilities in the calculation of the gradient and especially of the Hessian matrix increase. Nevertheless, in 8 dimensions the power is still substantive.

Power study II

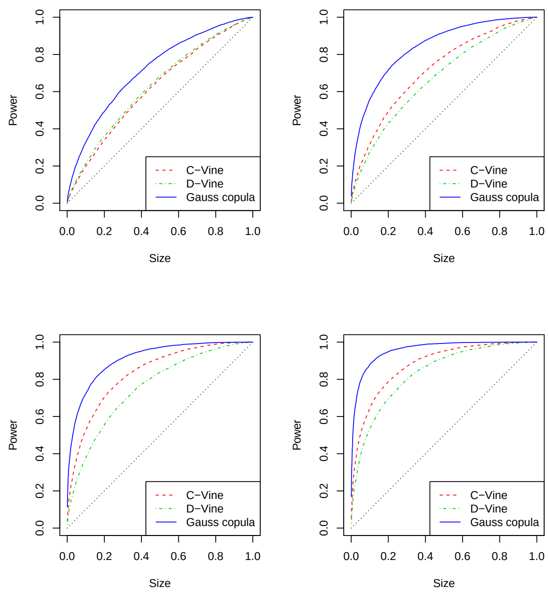
In the second scenario we investigate if the test can distinguish between two different 5-dimensional R-vines. One is selected and estimated with a maximum spanning tree (MST) algorithm and sequential estimation based on bivariate maximum likelihood (MLE), which is quite close to the full MLE (Dißmann et al. 2013). The comparison candidate is an R-vine chosen via MCMC in a Bayesian approach with the highest probability among the



(a) empirical with mixed Kendall's τ



(b) empirical with constant Kendall's $\tau = 0.5$



(c) empirical with constant Kendall's $\tau = 0.25$

Figure 4.4: Size-power curves for different number of observations. In each panel: Upper left: $n = 300$, upper right: $n = 500$, lower left: $n = 750$ and lower right: $n = 1000$.

True model in the alternative H_1	Vine under H_0	$\tau = 0.25$ Simul.	$\tau = 0.25$ Asy.	$\tau = 0.5$ Simul.	$\tau = 0.5$ Asy.	mixed τ Simul.	mixed τ Asy.
R-vine	R-vine	*5.0*	*31.9*	*5.0*	*53.5*	*5.0*	*44.4*
	C-vine	18.4	59.9	92.8	100.0	42.5	89.8
	D-vine	15.6	56.8	82.7	98.2	40.8	90.1
	Gauss	38.7	81.9	78.1	99.6	33.2	88.5
C-vine	R-vine	15.8	54.6	59.8	97.7	30.8	88.0
	C-vine	*5.0*	*29.3*	*5.0*	*47.8*	*5.0*	*46.3*
	D-vine	14.1	50.5	67.6	98.2	51.8	94.1
	Gauss	17.3	57.6	58.4	97.3	36.5	92.1
D-vine	R-vine	6.8	36.1	35.8	93.8	54.3	95.5
	C-vine	9.4	41.6	62.3	98.8	35.8	90.0
	D-vine	*5.0*	*29.0*	*5.0*	*51.2*	*5.0*	*44.4*
	Gauss	17.5	56.5	37.9	94.3	60.4	96.5
Gauss	R-Vine	6.1	9.1	7.3	28.3	7.7	17.9
	C-vine	6.1	9.3	6.3	25.4	6.6	16.0
	D-vine	5.8	8.4	8.0	30.3	7.4	17.5
	Gauss	*5.0*	*7.6*	*5.0*	*23.1*	*5.0*	*13.2*

Table 4.3: Estimated power (in %) for $n = 500$ at *nominal* size 5% (Values in italic give the actual size of the test)

visited variants in the MCMC (Gruber and Czado 2012). The true data generating model and the specification of the two estimated models are given in Table C.2 of Appendix C.2. The model selected by the Bayesian approach differs from the original R-vine just in two copulas, i.e. $c_{2,4}$ instead of $c_{3,4}$ and consequently $c_{3,4|2}$ instead of $c_{2,4|3}$. In contrast, the MST model differs in its fitted structure $\hat{\mathcal{V}}$ and pair-copula family selection $\hat{\mathcal{B}}(\hat{\mathcal{V}})$ to the other two more pronounced. Figure 4.5 shows the results and the conclusions are:

- The MST model is clearly detected as different from the true R-vine model for $n > 500$, while the MCMC model, which is much "closer" to the true model, since the size-power curves are close to 45 degree for all sample sizes n.

- As in power study I the power for the MST model is increasing with increasing number of observations.

- The observed log-likelihood is a first indicator for the misspecification in the MST case, since $l_{MST} = 3360 < l_{MCMC} = 3731 < l_{true} = 3757$ (n=1000)[2]. The log-likelihood of the MCMC model is much closer to the true one.

[2]Results from Gruber and Czado (2012)

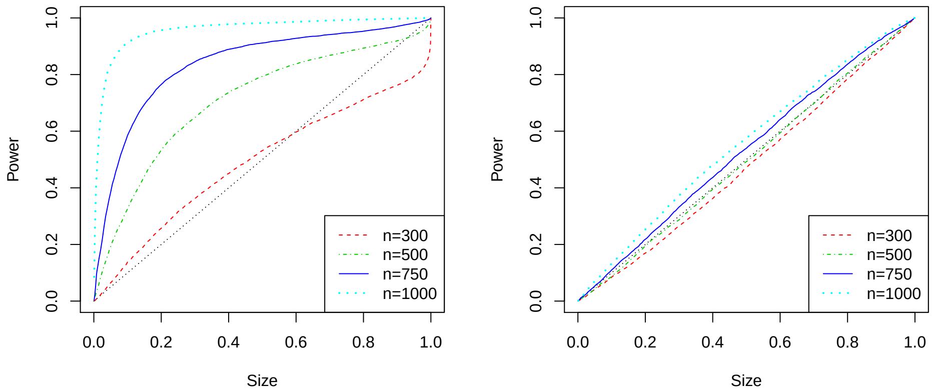


Figure 4.5: Simulated size-power curves for different sample sizes in power study II. Left panel: R-vine model fitted with maximum spanning tree (MST) and sequential estimation, right panel: R-vine model fitted with Bayesian approach.

Power study III

In financial applications often the multivariate Student's t-distribution or multivariate t-copula is used. In the last simulation study we investigate the difference of the multivariate t-copula with common degree-of-freedom versus an R-vine copula based on bivariate t-copulas with variable and separately estimated degrees-of-freedom in terms of our goodness-of-fit test. Again we choose the mixed R-vine model from power study I but change all copula families to t-copulas with the degree-of-freedom parameters in the range $\nu \in [4, 20]$.

Using the convergence of the t-copula for large ν to the Gauss copula we replace t-copulas with estimated degrees-of-freedom greater or equal 30 with the Gaussian copula in the R-vine copula model. This is done for stability and accuracy reasons in the calculations of the derivatives.

In Figure 4.6 we give the size-power curves for the simulated p-values for the two scenarios. The right panel indicates that given an R-vine with t-copulas the multivariate t-copula is not a good fit, since the test has power to discriminate against the multivariate t-copula. In contrast, the test has less power to discriminate an R-vine as alternative to a multivariate t-copula regardless of the sample size. For $n = 300$ the test has no detection power since the size-power curves are close to the 45 degree line.

Further observations are: The shape of the size-power curves depends heavily on the sample size. The numerical instability in estimation of the degrees-of-freedom may be a reason for this behavior. Compared to the other scenarios we have less power for small size.

4.1.4 Extension to unknown margins

Goodness-of-fit tests such as the tests discussed Berg (2009) do not account for uncertainties in the margins. He uses rank transformed data, i.e. so called pseudo-samples

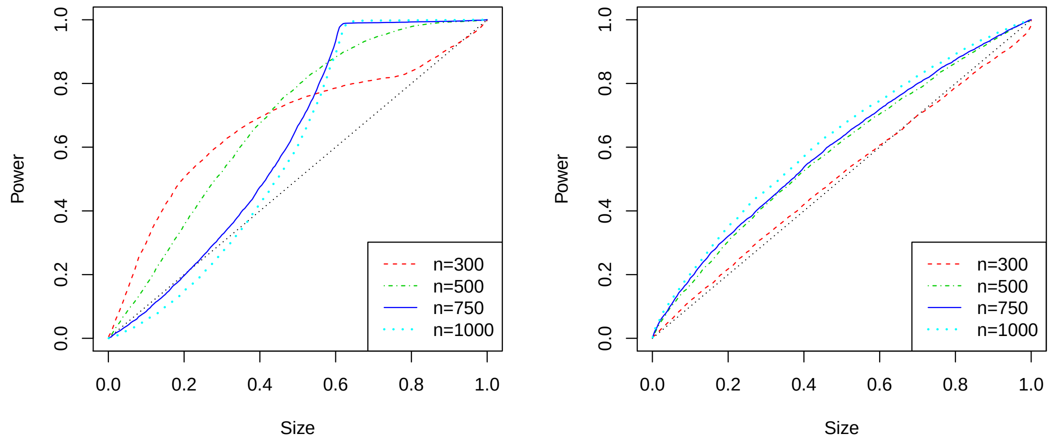


Figure 4.6: Size-power curves based on simulated p-values for different sample sizes. Left: R-vine (vine in H_1) versus multivariate t-copula (vine under H_0); right: multivariate t-copula (vine in H_1) versus R-vine (vine under H_0).

$\boldsymbol{u}_1 = (u_{11}, \ldots, u_{1d}), \ldots, \boldsymbol{u}_n = (u_{n1}, \ldots, u_{nd})$, where

$$\boldsymbol{u}_j = (u_{j1}, \ldots, u_{jd}) := \left(\frac{R_{j1}}{n+1}, \ldots, \frac{R_{jd}}{n+1} \right),$$

and R_{ji} is the rank of $x_{ij}, i = 1, \ldots, n$ and $j = 1, \ldots, d$, which are i.i.d. observations of the random vector $\boldsymbol{X} = (X_1, \ldots, X_d)$. The denominator $n+1$ instead of n avoids numerical problems at the boundaries of the unit hypercube.

A second method to handle unknown margins is the inference functions for margins (IFM) approach by Joe (1997), which uses parametric estimates $F_{\hat{\gamma}_i}, i = 1, \ldots, d$ of the margins based on marginal parameter estimates $\hat{\gamma}_i$ and use them to transform to copula data.

In the case of unknown margins one has to adjust the computation of the test statistic. Huang and Prokhorov (2013) did this for their copula goodness-of fit test. Similarly we can adjust our proposed R-vine copula goodness-of-fit test.

The asymptotic variance matrix $V_{\boldsymbol{\theta}_0}$ (Expression (4.8)) for the test statistic $\mathcal{T}_n$ (Expression (4.9)) is extended using expected derivatives with respect to the margins of the log-likelihood and the expected derivatives of the vectorized sum of the Hessian matrix and the outer product of gradient, respectively. More precisely define

$$W_i(F_i) := \int_{[0,1]^d} \left[I_{\{F_i \le u_i\}} - u_i \right] \partial^2_{\boldsymbol{\theta}, u_i} \ln(c_{\boldsymbol{\theta}_0}(u_1, \ldots, u_d)) c_{\boldsymbol{\theta}_0}(u_1, \ldots, u_d) du_1 \ldots du_d,$$

$$M_i(F_i) := \int_{[0,1]^d} \left[I_{\{F_i \le u_i\}} - u_i \right] \partial_{u_i} vech\Big(\partial^2_{\boldsymbol{\theta}} \ln(c_{\boldsymbol{\theta}_0}(u_1, \ldots, u_d)) + \partial_{\boldsymbol{\theta}} \ln(c_{\boldsymbol{\theta}_0}(u_1, \ldots, u_d)) \big(\partial_{\boldsymbol{\theta}} \ln(c_{\boldsymbol{\theta}_0}(u_1, \ldots, u_d)) \big)^T \Big) du_1 \ldots du_d,$$

with $F_i := F_i(x_i), i = 1, \ldots, d$. Furthermore, $\boldsymbol{d}(\boldsymbol{\theta}_0)$ is now defined in terms of the random vector $\boldsymbol{X}$:

$$\boldsymbol{d}(\boldsymbol{\theta}_0|\boldsymbol{X}) := vech(\mathbb{H}(\boldsymbol{\theta}_0) + \mathbb{C}(\boldsymbol{\theta}_0)),$$

where $\mathbb{H}(\boldsymbol{\theta}_0)$ and $\mathbb{C}(\boldsymbol{\theta}_0)$ are defined in (4.5). With $l(\boldsymbol{\theta}|\boldsymbol{X}) := \ln(c_{\boldsymbol{\theta}_0}(F_1, \ldots, F_d))$ the adjusted variance matrix is

$$V_{\boldsymbol{\theta}_0} = E\Bigg[\left(\boldsymbol{d}(\boldsymbol{\theta}_0|\boldsymbol{X}) - \nabla D_{\boldsymbol{\theta}_0}\mathbb{H}^{-1}(\boldsymbol{\theta}_0)\left(\partial_{\boldsymbol{\theta}} l(\boldsymbol{\theta}_0|\boldsymbol{X}) + \sum_{i=1}^{d} W_i(F_i)\right) + \sum_{i=1}^{d} M_i(F_i)\right)$$
$$\left(\boldsymbol{d}(\boldsymbol{\theta}_0|\boldsymbol{X}) - \nabla D_{\boldsymbol{\theta}_0}\mathbb{H}^{-1}(\boldsymbol{\theta}_0)\left(\partial_{\boldsymbol{\theta}} l(\boldsymbol{\theta}_0|\boldsymbol{X}) + \sum_{i=1}^{d} W_i(F_i)\right) + \sum_{i=1}^{d} M_i(F_i)\right)^T\Bigg].$$

But, this correction of White's original formula would involve multidimensional integrals, not computational tractable in appropriate time. An adjusted gradient and Hesse matrix may avoid this problem. Since the goodness-of-fit test calculation does not depend directly on the density function f but on the product of pair-copulas and marginal density functions, the derivatives should not only be with respect to the parameters but to the marginals too. This approach may be a topic for further research. Thus (4.8) will be used as an approximation in the case of unknown margins, e.g. in our application in the next section.

To justify the good approximation behavior of (4.8) we run power study I of Section 4.1.3 with unknown margins. We limit ourselves here to the mixed copula case with $n = 500$ observations. The chosen data generating process uses the standard normal distribution for all 5 margins in a first scenario, and centered normal distributions with different standard deviations $\sigma \in \{1, 2, 3, 4, 5\}$ in a second scenario. The margins are estimated via moments in an IFM approach.

In Figure 4.7 we illustrate the resulting size-power curves. Comparing the middle panel to the right panel or to the left panel we cannot detect significant differences in the power if the margins are unknown. Further, the choice of the marginal distribution F_γ has no significant influence on the size-power curves. This is confirmed by further power studies not presented in this manuscript, e.g. marginal Student's t-distributions. This is no longer true if the choice of the marginal distribution fits badly the (generated) data, e.g. data generated from a Student's t-distribution fitted with an exponential distribution. An estimation of the margins with the rank based approach returned similar results as the IFM approach above.

4.2 Information matrix ratio test

As second goodness-of-fit test for vine copula models we introduce the information ratio (IR) test. It is inspired by the paper of Zhou et al. (2012), who propose an IR test for general model misspecification of the variance or covariance structures. Their test is related to the "in-and-out-sample" (IOS) test of Presnell and Boos (2004), which is a likelihood ratio test. Additionally Presnell and Boos (2004) showed that the IOS test statistic can

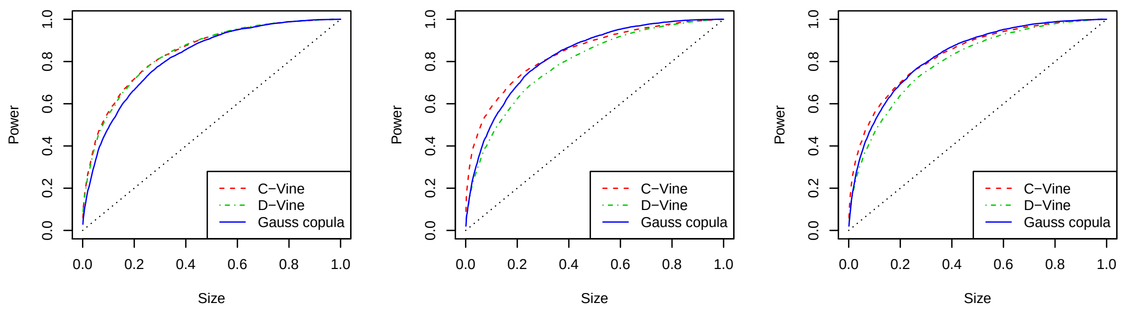


Figure 4.7: Simulated size-power curves for the mixed Kendall's τ R-vine copula model considering $n = 500$ observations. left panel: no uncertainty in the margins; middle panel: data generating process with standard normal margins for all 5 dimensions, right panel: data generating process with centered normal margins with different standard deviations.

be expressed as a ratio of the expected Hessian and the expected outer product of the gradient. Following the notation of Section 4.1, $\mathbb{H}(\boldsymbol{\theta})$ and $\mathbb{C}(\boldsymbol{\theta})$ denote the expected Hessian matrix and the expected outer product of the score function, respectively. See Equation (4.5) for definition of $\mathbb{H}(\boldsymbol{\theta})$ and $\mathbb{C}(\boldsymbol{\theta})$. Now the information matrix ratio (IMR) is defined as

$$\Psi(\boldsymbol{\theta}) := -\mathbb{H}(\boldsymbol{\theta})^{-1}\mathbb{C}(\boldsymbol{\theta}) \tag{4.10}$$

and the test problem is

$$H_0 : \Psi(\boldsymbol{\theta}) = I_p \quad \text{against} \quad H_1 : \Psi(\boldsymbol{\theta}) \neq I_p,$$

where I_p is the p-dimensional identity matrix.

Given again the sample equivalents to $\mathbb{H}(\boldsymbol{\theta})$ and $\mathbb{C}(\boldsymbol{\theta})$ denoted as $\bar{\mathbb{H}}(\hat{\boldsymbol{\theta}}_n)$ and $\bar{\mathbb{C}}(\hat{\boldsymbol{\theta}}_n)$ as defined in Equation (4.7) we get as empirical version of (4.10):

$$\bar{\Psi}(\hat{\boldsymbol{\theta}}_n) := -\bar{\mathbb{H}}(\hat{\boldsymbol{\theta}}_n)^{-1}\bar{\mathbb{C}}(\hat{\boldsymbol{\theta}}_n).$$

As in Zhou et al. (2012) we define the *information ratio* **(IR)** statistic as

$$IR_n := tr(\bar{\Psi}(\hat{\boldsymbol{\theta}}_n))/p, \tag{4.11}$$

where $tr(A)$ denotes the trace of matrix A. To derive the asymptotic normality of the test statistic IR_n some conditions have to be set. The first two conditions $\mathcal{C}_1$ and $\mathcal{C}_2$ guarantee the existence of the gradient and the Hessian matrix.

$\mathcal{C}_1$: The density function (2.3) is twice continuous differentiable with respect to $\boldsymbol{\theta}$.

$\mathcal{C}_2$: -$\bar{\mathbb{H}}(\hat{\boldsymbol{\theta}}_n)$ and $\bar{\mathbb{C}}(\hat{\boldsymbol{\theta}}_n)$ are positive definite.

Condition $\mathcal{C}_3 - \mathcal{C}_5$ are more technical and are the same as in Presnell and Boos (2004).

$\mathcal{C}_3$: There exist $\boldsymbol{\theta}_0$ such that $\hat{\boldsymbol{\theta}}_n \xrightarrow{P} \boldsymbol{\theta}_0$ as $n \to \infty$.

$\mathcal{C}_4$: The estimator $\hat{\boldsymbol{\theta}}_n \in \mathbb{R}^p$ has an approximating influence curve function $h(\boldsymbol{\theta}|\boldsymbol{u})$ such that

$$\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta} = \frac{1}{n}\sum_{i=1}^{n} h(\boldsymbol{\theta}_0|U_i) + R_{n1},$$

where $\sqrt{n}R_{n1} \xrightarrow{P} 0$ as $n \to \infty$, $E[h(\boldsymbol{\theta}_0|U_1] = 0$, and $cov(h(\boldsymbol{\theta}_0|U_1)$ is finite.

$\mathcal{C}_5$: The real-valued function $q(\boldsymbol{\theta}|\boldsymbol{u})$ possesses second order partial derivatives with respect to $\boldsymbol{\theta}$, and

(a) $Var(q(\boldsymbol{\theta}_0|U_1))$ and $E\left[\frac{\partial}{\partial\boldsymbol{\theta}} q(\boldsymbol{\theta}_0|U_1)\right]$ are finite.

(b) There exists a function $M(\boldsymbol{u})$ such that for all $\boldsymbol{\theta}$ in a neighborhood of $\boldsymbol{\theta}_0$ and all $j, k \in \{1, \ldots, p\}$, $\left|\frac{\partial^2}{\partial^2\boldsymbol{\theta}} q(\boldsymbol{\theta}|\boldsymbol{u})_{jk}\right| \leq M(\boldsymbol{u})$, where $E[M(U_1)] < \infty$.

In the following $vech(A) \in \mathbb{R}^{p(p+1)/2}$ represents the vectorization of the symmetric matrix $A \in \mathbb{R}^{p\times p}$. Let $\boldsymbol{W} := (W_1, \ldots, W_{p(p+1)})^T = (vech(\bar{\mathbb{C}}(\hat{\boldsymbol{\theta}}_n)), vech(\bar{\mathbb{H}}(\hat{\boldsymbol{\theta}}_n)))^T \in \mathbb{R}^{p(p+1)}$, then Presnell and Boos (2004) showed that

$$\Sigma_W^{-1/2}\sqrt{n}\boldsymbol{W} - \boldsymbol{\mu}_W \xrightarrow{d} N_{p(p+1)}(\mathbf{0}_{p(p+1)}, I_{p(p+1)}),$$

where $\boldsymbol{\mu}_W$ is the mean vector and Σ_W is the asymptotic covariance matrix of $\boldsymbol{W}$. Here $\mathbf{0}_{p(p+1)} := (0, \ldots, 0)^T$ is the $p(p+1)$-dimensional zero vector and $I_{p(p+1)}$ is the $p(p+1)$-dimensional identity matrix. Furthermore, let $D(\hat{\boldsymbol{\theta}}_n)$ define the partial derivatives of IR_n taken with respect to the components of $\boldsymbol{W}$, i.e.

$$D(\hat{\boldsymbol{\theta}}_n) := \left(\frac{\partial IR_n}{\partial W_i}\right)_{i=1,\ldots,p(p+1)} \in \mathbb{R}^{p(p+1)}.$$

Theorem 4.3
Let $\boldsymbol{U} \sim RV(\mathcal{V}, \mathcal{B}(\mathcal{V}), \boldsymbol{\theta}(\mathcal{B}(\mathcal{V})))$ satisfy the conditions $\mathcal{C}_1 - \mathcal{C}_3$. Further, let $\mathcal{C}_4$ hold for the maximum likelihood estimator $\hat{\boldsymbol{\theta}}_n$ with $h(\boldsymbol{\theta}_0|\boldsymbol{u}) := \mathbb{C}(\boldsymbol{\theta}_0)^{-1}\frac{\partial}{\partial\boldsymbol{\theta}} l(\boldsymbol{\theta}_0|\boldsymbol{u})$. Additionally, the condition $\mathcal{C}_5$ has to be satisfied for both $q(\boldsymbol{\theta}|\boldsymbol{u}) := -\frac{\partial^2}{\partial^2\boldsymbol{\theta}} l(\boldsymbol{\theta}|\boldsymbol{u})_{jk}$ and $q(\boldsymbol{\theta}|\boldsymbol{u}) := \left(\frac{\partial}{\partial\boldsymbol{\theta}} l(\boldsymbol{\theta}|\boldsymbol{u}) \left(\frac{\partial}{\partial\boldsymbol{\theta}} l(\boldsymbol{\theta}|\boldsymbol{u})\right)^T\right)_{jk}$ for each $j, k \in \{1, \ldots, p\}$. Then the IR test statistic

$$Z_n := \frac{IR_n - 1}{\sigma_{IR}} \xrightarrow{d} N(0, 1) \text{ as } n \to \infty,$$

where σ_{IR} is the standard error of the IR test statistic, defined as

$$\sigma_{IR}^2 := \frac{1}{n} D^T \Sigma_W D.$$

Here Σ_W is the asymptotic covariance matrix arising from the joint asymptotic normality of $vech(\bar{\mathbb{C}}(\hat{\boldsymbol{\theta}}_n))$ and $vech(\bar{\mathbb{H}}(\hat{\boldsymbol{\theta}}_n))$ defined above. By D we denote the $p(p+1)$-dimensional vector of partial derivatives of IR_n taken with respect to the components of $\boldsymbol{W}$ and evaluated at their limits in probability, i.e. $D := D(\hat{\boldsymbol{\theta}}_n)|_{\hat{\boldsymbol{\theta}}_n \xrightarrow{P} \boldsymbol{\theta}_0}$.

Proof
The proof follows directly from the proof of Theorem 3 in Presnell and Boos (2004), since we have a fully specified likelihood and the conditions of Theorem 3 are assumed to be satisfied for vine copulas considered in Theorem 4.3. □

Since the theoretical asymptotic variance σ_{IR}^2 is quite difficult to compute, an empirical version is used in practice. To evaluate the standard error σ_{IR} numerically, Zhou et al. (2012) suggest a perturbation resampling approach. Furthermore, Presnell and Boos (2004) state that the convergence to normality is slow and thus they suggest obtaining p-values using a parametric bootstrap under the null hypothesis.

The condition $\mathcal{C}_4$ for $q(\boldsymbol{\theta}|\boldsymbol{u}) := -\frac{\partial^2}{\partial^2 \boldsymbol{\theta}} l(\boldsymbol{\theta}|\boldsymbol{u})_{jk}$ implies, that the copula density function (2.3) is four times differentiable with respect to $\boldsymbol{\theta}$. Furthermore, the first and second moment of the second derivative has to be finite. The vine copula density is four times differentiable if all selected pair-copulas are four times differentiable. These assumptions are satisfied for the elliptical Gauss and Student's t-copula as well as for the parametric Archimedean copulas in all dimensions.

Note that in Zhou et al. (2012) the information matrix ratio is defined inversely, i.e. $\Psi(\boldsymbol{\theta}) = -\mathbb{C}(\boldsymbol{\theta})^{-1}\mathbb{H}(\boldsymbol{\theta})$. This does not change the asymptotic normality of Theorem 4.3, since $\bar{\mathbb{H}}(\hat{\boldsymbol{\theta}}_n) + \bar{\mathbb{C}}(\hat{\boldsymbol{\theta}}_n) = 0$ as $n \to \infty$ (under the correct model). The asymptotic mean matrix of $\Psi(\boldsymbol{\theta})$ is the identity matrix I_p and therefore the mean of the IR statistic is $tr(I_p)/p = 1$.

Let $\alpha \in (0,1)$ and Z_n as in Theorem 4.3. Then the test

$$\text{Reject } H_0 : \Psi(\boldsymbol{\theta}) = I_p \quad \text{against} \quad \Psi(\boldsymbol{\theta}) \neq I_p \qquad \Leftrightarrow \qquad Z_n > \Phi^{-1}(1-\alpha)$$

is an asymptotic α-level test. Here Φ^{-1} denotes the quantile of a $N(0,1)$-distribution.

4.3 Further goodness-of-fit tests for vine copulas

In the recent years many GOF test were suggested for copulas. The most promising ones were investigated in Genest et al. (2009) and Berg (2009). However only the size and power of the elliptical and one-parametric Archimedean copulas for $d \in \{2,4,8\}$ were analyzed. The multivariate case is therefore poorly addressed. For vine copulas little is done. A first test for vine copulas was suggested but not investigated in Aas et al. (2009). Their GOF is based on the multivariate PIT and an aggregation introduced by Breymann et al. (2003). After aggregation standard univariate GOF tests such as the Anderson-Darling (AD), the Cramér-von Mises (CvM) or the Kolmogorov-Smirnov (KS) tests are applied. They are described in more detail in B.2. We will denote the resulting tests as **Breymann**.

Similar approaches based on the multivariate PIT are proposed by Berg and Bakken (2007). Beside new aggregation functions forming univariate test data, they perform the aggregation step on the ordered PIT output data $\boldsymbol{y}_{(1)}^T, \ldots, \boldsymbol{y}_{(d)}^T$ instead of $\boldsymbol{y}_1^T, \ldots, \boldsymbol{y}_d^T$. Again standard univariate GOF tests are applied. These approaches will be called **Berg** and **Berg2**, respectively.

Berg and Aas (2009) applied a test for $H_0 : C \in \mathcal{C}_0$ against $H_1 : C \notin \mathcal{C}_0$ based on the empirical copula process (ECP) to a 4-dimensional vine copula. As the Breymann test, their GOF test is not described in detail or investigated with respect to its power. We will denote this test as **ECP**. An extension of the ECP-test is the combination of the multivariate PIT approach with the ECP. The general idea is that the transformed data of a multivariate PIT should be "close" to the independence copula $C_{\perp}$ Genest et al. (2009). Thus a distance of CvM or KS type between them is considered. This approach is called **ECP2**.

In the forthcoming sections we will introduce the multivariate PIT based GOF such as the ones of Breymann et al. (2003) and Berg and Bakken (2007), and the two ECP based GOF tests. For an overview see the diagram in Figure 4.1 at the beginning of this Chapter.

4.3.1 Rosenbatt's transform test

The vine copula GOF test suggested by Aas et al. (2009) is based on the multivariate probability integral transform (PIT) of Rosenblatt (1952) applied to copula data $\boldsymbol{u} = (\boldsymbol{u}_1^T, \ldots, \boldsymbol{u}_d^T), \boldsymbol{u}_i = (u_{1i}, \ldots, u_{ni})^T$, $i = 1, \ldots, d$ and a given estimated vine copula model $(\mathcal{V}, \mathcal{B}(\mathcal{V}), \hat{\boldsymbol{\theta}}(\mathcal{B}(\mathcal{V})))$. The general multivariate PIT definition and the explicit algorithm for the R-vine copula model is given in Appendix B.1. The PIT output data $\boldsymbol{y} = (\boldsymbol{y}_1^T, \ldots, \boldsymbol{y}_d^T), y_i = (y_{1i}, \ldots, y_{ni})^T, i = 1 \ldots, d$ is assumed to be i.i.d. with $y_{it} \sim U[0, 1]$ for $t = 1, \ldots, n$. Now, a common approach in multivariate GOF testing is dimension reduction. Here the aggregation is performed by

$$s_t := \sum_{i=1}^{d} \Gamma(y_{ti}), \quad t = \{1, \ldots, n\}, \tag{4.12}$$

with a weighting function $\Gamma(\cdot)$. Breymann et al. (2003) suggest as weight function the

$$\begin{pmatrix} u_{11} & \ldots & u_{1d} \\ \vdots & & \vdots \\ u_{n1} & \ldots & u_{nd} \end{pmatrix} \xrightarrow[(PIT)]{Rosenblatt} \begin{pmatrix} y_{11} & \ldots & y_{1d} \\ \vdots & & \vdots \\ y_{n1} & \ldots & y_{nd} \end{pmatrix} \xrightarrow[\Gamma(y_{ti})]{Aggregation} \begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix} \xrightarrow[GOFtests]{univariate}$$

Figure 4.8: Schematic procedure of the PIT based goodness-of-fit tests.

squared quantile of the standard normal distribution, i.e. $\Gamma(y_{ti}) = \Phi^{-1}(y_{ti})^2$, with $\Phi(\cdot)$ denoting the $N(0, 1)$ cdf. Finally, they apply a univariate Anderson-Darling test to the univariate test data s_t. The three step procedure is summarized in Figure 4.8.

Berg and Bakken (2007) point out that the approach of Breymann et al. (2003) has some weaknesses and limitations. The weighting function $\Phi^{-1}(y_{ti})^2$ strongly weights data along the boundaries of the d-dimensional unit hypercube. They suggest a generalization and extension of the PIT approach. First, they propose two new weighting functions for the aggregation in (4.12):

$$\Gamma(y_{ti}) = |y_{ti} - 0.5| \quad \text{and} \quad \Gamma(y_{ti}) = (y_{ti} - 0.5)^{\alpha}, \alpha = (2, 4, \ldots).$$

Further, they use the order statistics of the random vector $\boldsymbol{Y} = (Y_1, \ldots, Y_d)$, denoted by $Y_{(1)} \leq Y_{(2)} \leq \ldots \leq Y_{(d)}$ with observed values $y_{(1)} < y_{(2)} \leq \ldots \leq y_{(d)}$. The calculation of the order statistics PIT can be simplified by using the fact that $Y_{(1)} \leq Y_{(2)} \leq \ldots \leq Y_{(d)}$ are i.i.d. $U(0,1)$ random variables and $\{Y_{(i)}, 1 \leq i \leq d\}$ is a Markov chain (David 1981, Theorem 2.7). Now Theorem 1 of Deheuvels (1984) can be applied and the calculation of the PIT ease to

$$v_i := F_{Y_{(i)}|Y_{(i-1)}}(y_{(i)}) = 1 - \left(\frac{1 - y_{(i)}}{1 - y_{(i-1)}}\right)^{d-(i-1)}, \quad i = 1, \ldots, d, y_{(0)} = 0. \tag{4.13}$$

Now, Berg and Bakken (2007) construct the aggregation as the sum of a product of two weighting functions applied to $\boldsymbol{y}$ and $\boldsymbol{v} = (v_1, \ldots, v_d)$, respectively, i.e.

$$s_t := \sum_{i=1}^{d} \Gamma_{\boldsymbol{y}}(y_{ti}) \cdot \Gamma_{\boldsymbol{v}}(v_{ti}), \quad t = \{1, \ldots, n\}.$$

Here $\Gamma_{\boldsymbol{y}}(\cdot)$ and $\Gamma_{\boldsymbol{v}}(\cdot)$ are chosen from the suggested weighting functions including the one of Breymann et al. (2003). Let S_t be the corresponding random aggregation of s_t. If $\Gamma_{\boldsymbol{y}}(\cdot) = 1$ and $\Gamma_{\boldsymbol{v}}(\cdot) = \Phi^{-1}(\cdot)^2$ or vise versa, the asymptotic distribution of S_t follows a χ^2_d distributed random variable (Breymann et al. 2003). In all other cases the asymptotic distribution of S_t is unknown.

The combinations with $\Gamma_{\boldsymbol{y}}(y_{ti}) = |y_{ti} - 0.5|$ and $\Gamma_{\boldsymbol{y}}(y_{ti}) = (y_{ti} - 0.5)^\alpha$ for $\alpha = 2, 4, \ldots$ performed very poorly in the simulation setup considered later. Thus we will not include them in the forthcoming power study. Only the weighting functions listed in Table 4.4 will be investigated. As final test statistics to the test data s_t we apply the univariate Cramér-von Mises (CvM) or Kolmogorov-Smirnov (KS) test, as well as the mentioned univariate Anderson-Darling (AD) test. All three test statistics are given in Appendix B.2 for the convenience of the reader.

Short	Description	
Breymann	$\Gamma_{\boldsymbol{y}}(y_{ti}) = \Phi^{-1}(y_{ti})$	$\Gamma_{\boldsymbol{v}}(v_{ti}) = 1$
Berg	$\Gamma_{\boldsymbol{y}}(y_{ti}) = 1$	$\Gamma_{\boldsymbol{v}}(v_{ti}) = \|v_{ti} - 0.5\|$
Berg2	$\Gamma_{\boldsymbol{y}}(y_{ti}) = 1$	$\Gamma_{\boldsymbol{v}}(v_{ti}) = (v_{ti} - 0.5)^2$

Table 4.4: Specifications of the PIT based goodness-of-fit tests.

Let $s^{AD}_{1-\alpha}, s^{CvM}_{1-\alpha}$ and $s^{KS}_{1-\alpha}$ denote the $1-\alpha$ quantile of the univariate AD, CvM or KS test statistic, respectively. Then the test rejects the null hypothesis (4.1) if $W^2_n > s^{AD}_{1-\alpha}, \omega^2 > s^{CvM}_{1-\alpha}$ or $D_n > s^{KS}_{1-\alpha}$, respectively.

4.3.2 Empirical copula process tests

A rather different approach is suggested by Genest and Rémillard (2008) for copula GOF testing. They propose to use the difference of the copula distribution function $C_{\hat{\boldsymbol{\theta}}_n}(\boldsymbol{u})$ with estimated parameter $\hat{\boldsymbol{\theta}}_n$ and the empirical copula $\hat{C}_n(\boldsymbol{u})$ (see Equation (4.4)) given the copula data $\boldsymbol{u}$. This stochastic process is known as the empirical copula process (ECP) and will be used to test (4.1). For a vine copula model the copula distribution function $C_{\hat{\boldsymbol{\theta}}_n}(\boldsymbol{u})$ is not given in closed form. Thus a bootstrapped version has to be used.

Now, the ECP $\hat{C}_n(\boldsymbol{u}) - C_{\hat{\boldsymbol{\theta}}_n}(\boldsymbol{u})$ is utilized in a multivariate Cramér-von Mises (mCvM) or multivariate Kolmogorov-Smirnov (mKS) based test statistic. The multivariate distribution functions $\hat{F}_n(\boldsymbol{y})$ and $F(\boldsymbol{y})$ in Equation (B.1) and (B.2) of Appendix B.2.1 are replaced by their (vine) copula equivalents $\hat{C}_n(\boldsymbol{u})$ and $C_{\hat{\boldsymbol{\theta}}_n}(\boldsymbol{u})$, respectively. Thus we consider

$$\textbf{ECP-mCvM:} \quad n\omega^2_{ECP} := n \int_{[0,1]^d} (\hat{C}_n(\boldsymbol{u}) - C_{\hat{\boldsymbol{\theta}}_n}(\boldsymbol{u}))^2 d\hat{C}_n(\boldsymbol{u}) \quad \text{and}$$

$$\textbf{ECP-mKS:} \quad D_{n,ECP} := \sup_{\boldsymbol{u}\in[0,1]^d} |\hat{C}_n(\boldsymbol{u}) - C_{\hat{\boldsymbol{\theta}}_n}(\boldsymbol{u})|.$$

To avoid the calculation/approximation of $C_{\hat{\boldsymbol{\theta}}_n}(\boldsymbol{u})$ Genest et al. (2009) and other authors propose to use the transformed data $\boldsymbol{y} = (y_1, \ldots, y_d)$ of the PIT approach and plug them into the ECP. The idea is to calculate the distance between the empirical copula $\hat{C}_n(\boldsymbol{y})$ of the transformed data $\boldsymbol{y}$ and the independence copula $C_\perp(\boldsymbol{y})$. Thus, the considered multivariate CvM and KS test statistics are

$$\textbf{ECP2-mCvM:} \quad n\omega^2_{ECP2} := n \int_{[0,1]^d} (\hat{C}_n(\boldsymbol{y}) - C_\perp(\boldsymbol{y}))^2 d\hat{C}_n(\boldsymbol{y}) \quad \text{and}$$

$$\textbf{ECP2-mKS:} \quad D_{n,ECP2} := \sup_{\boldsymbol{y}\in[0,1]^d} |\hat{C}_n(\boldsymbol{y}) - C_\perp(\boldsymbol{y})|,$$

respectively. Since neither the mCvM nor the mKS test statistic has a known asymptotic distribution function a parametric bootstrap procedure has to be applied to estimate p-values. Thus a computer intensive double bootstrap procedure has to be implemented. As before the test rejects the null hypothesis (4.1) if $n\omega^2_{ECP} > s^{mCvM}_{1-\alpha}$ or $D_{n,ECP} > s^{mKS}_{1-\alpha}$, respectively. Here $s^{mCvM}_{1-\alpha}$ and $s^{mKS}_{1-\alpha}$ are the $1-\alpha$ quantiles of the mCvM and mKS test statistic's empirical distribution function, respectively. Similar rejection regions are defined for the ECP2 test statistics.

4.4 Comparison in a power study

To investigate the power behavior of the proposed GOF tests and to compare them to each other we conduct several Monte Carlo studies of different dimension. The second property of interest is the ability of the test to maintain the nominal level or size, usually chosen at 5%.

As in Section 4.1.3 we use the estimated size (actual size) and power as performance measures. Furthermore, the general simulation setup stated in Section 4.1.3 is also applied to the different test statistics defined in Section 4.2 and 4.3, namely $\mathcal{T}_n$ (White), Z_n (IR), W_n^2 (AD), $n\omega^2$ (CvM or mCvM) and D_n (KS or mKS).

In all of the forthcoming simulation studies we used $B = 2500$ replications and the number of observations were chosen to be $n = 500, n = 750, n = 1000$ or $n = 2000$. As model dimension we chose $d = 5$ and $d = 8$ and the critical level α is 0.05. As before all calculations are performed using the statistical software **R** and the R-package **VineCopula** of Schepsmeier et al. (2012) introduced in Section 2.3.

As test specification we consider again an R-vine as true model (M_1) and the alternatives

- multivariate Gauss copula,
- C-vine copula and
- D-vine copula,

as defined in Power study I of Section 4.1.3 and Appendix C.1. For the 8 dimensional example we refer to Table C.3 and Figure C.2 of Appendix C.3.

Although all three stated alternatives have different vine structures and pair-copula families we do not know which vine copula model is "closer" to the true R-vine model. A often proposed approach for model comparison is the Kullback and Leibler (1951) information criterion (KLIC). It measures the distance between a true unknown distribution and a specified, but estimated model. In the following definition we follow Vuong (1989). Let $c_0(\cdot)$ be the true (vine) copula density function of a d-dimensional random vector $\boldsymbol{U}$. Further, E_0 denotes the expected value with respect to this true distribution. The estimated (vine) copula density of $\boldsymbol{U}$ is denoted as $c(\hat{\boldsymbol{\theta}}_n|\boldsymbol{U})$, where $\hat{\boldsymbol{\theta}}_n$ is the estimated model parameter (vector) given n samples of $\boldsymbol{U}$. Then, the KLIC between c_0 and c is defined as

$$KLIC(c_0, c) := \int_{(0,1)^d} c_0(\boldsymbol{u}) \ln \left(\frac{c_0(\boldsymbol{u})}{c(\hat{\boldsymbol{\theta}}_n|\boldsymbol{u})} \right) d\boldsymbol{u} = E_0[\ln c_0(\boldsymbol{U})] - E_0[\ln c(\hat{\boldsymbol{\theta}}_n|\boldsymbol{U})].$$

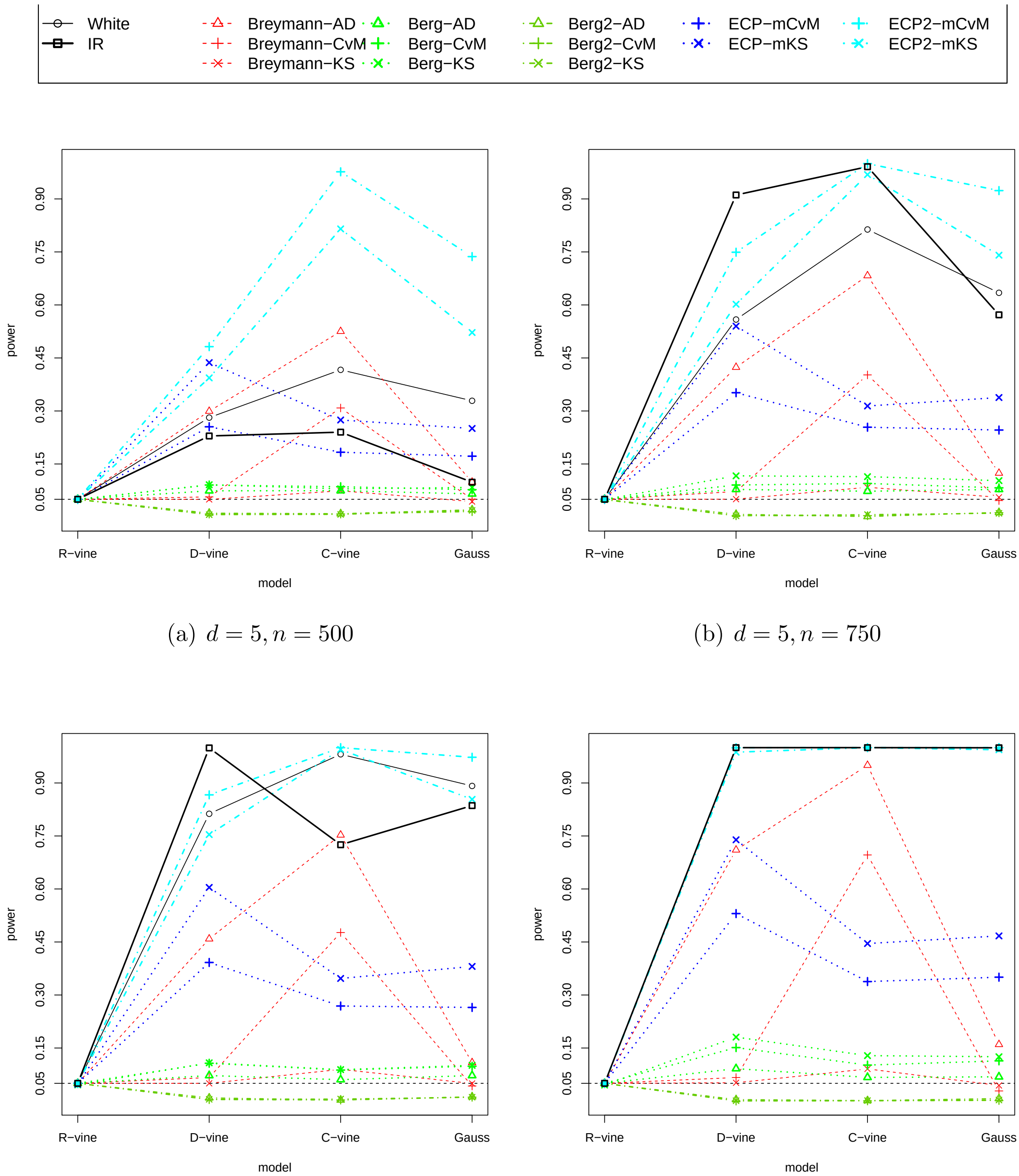
The model with the smallest KLIC is "closest" to the true model. In the plots of the following power study we ordered the alternatives on the x-axis by their KLIC as listed in Table 4.5, e.g. for $d = 5$ we have the order D-vine, C-vine, Gauss.

The approximation of the multidimensional integral is facilitated by Monte Carlo or a numerical integration based on the R-package **cubature** (C code by Steven G. Johnson and R by Balasubramanian Narasimhan 2011). In the numerical integration copula data, i.e. $\boldsymbol{u} \in (0, 1)^d$, or standard normal transformed data, i.e. $\boldsymbol{x} = \Phi(\boldsymbol{u}) \in \mathbb{R}^d$, are used. We see that it is quite challenging to estimate the KLIC distance in high dimensions.

Results

Since all proposed GOF tests have either no asymptotic distribution at all or face substantial numerical problems estimating the asymptotic variance or have shown to have low power in small samples, we only investigate the bootstrapped version of the tests. In the Figures 4.9 and 4.10 we illustrate the estimated power of all 15 proposed GOF tests for $d = 5$ and $d = 8$, respectively. On the x-axis we have the R-vine as true model and the three alternatives ordered by their KLIC. For the true model the actual size is plotted. A horizontal black dashed line indicates the 5% α-level.

Size: All proposed GOF tests maintain their given size independently of the number of sample points for $d = 5$. In the 8-dimensional case the GOF tests based on the Berg



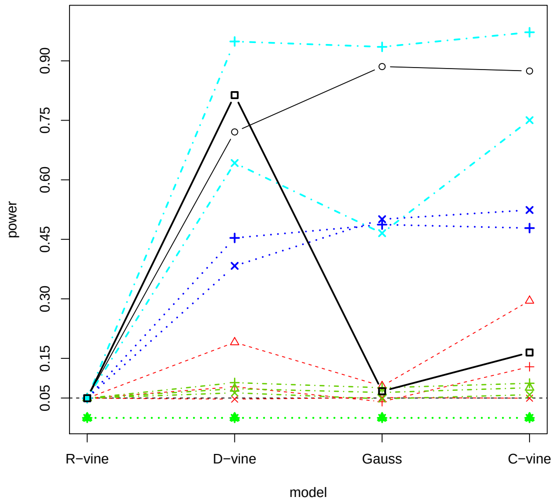
(a) $d = 5, n = 500$

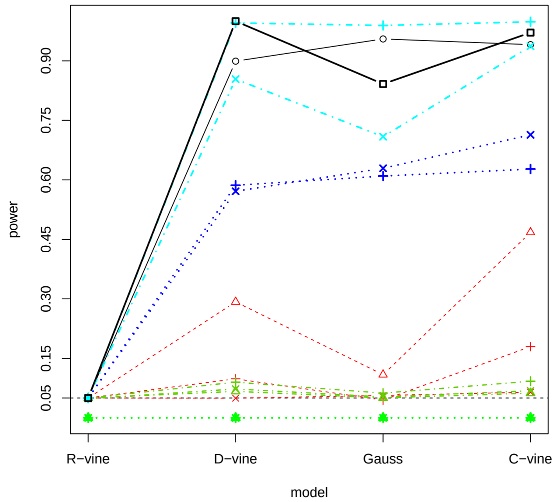
(b) $d = 5, n = 750$

(c) $d = 5, n = 1000$

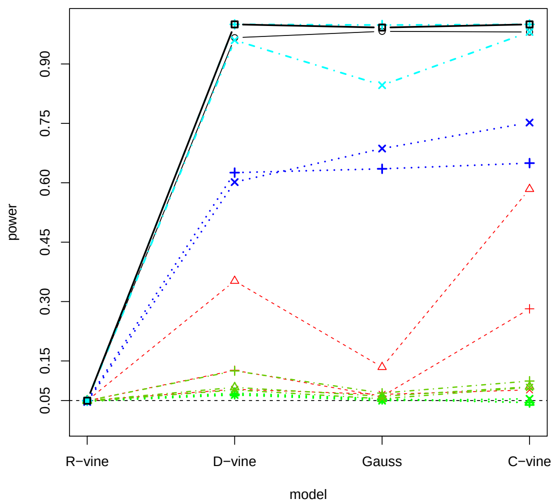
(d) $d = 5, n = 2000$

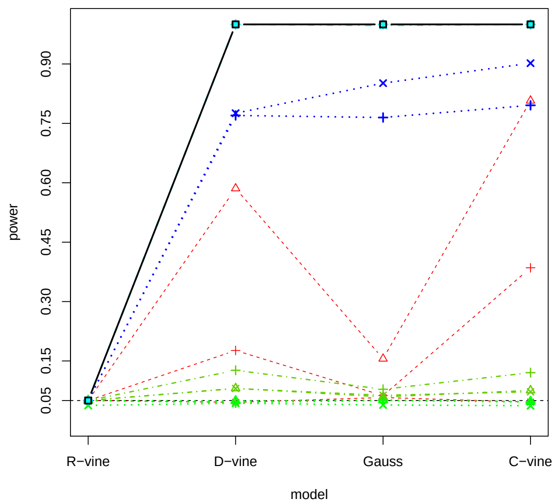
Figure 4.9: Power comparison of the proposed goodness-of-fit tests in 5 dimensions with different number of sample points. The alternatives are ordered on the x-axis by the rank of their KLIC value with respect to the true R-vine.





(a) $d = 8, n = 500$ (b) $d = 8, n = 750$





(c) $d = 8, n = 1000$ (d) $d = 8, n = 2000$

Figure 4.10: Power comparison of the proposed goodness-of-fit tests in 8 dimensions with different number of sample points. The alternatives are ordered on the x-axis by the rank of their KLIC value with respect to the true R-vine.

d	method	C-vine	D-vine	Gauss
5	Monte Carlo	0.65	0.64	0.72
	numerical integration based on copula margins	0.62[a]	0.45[a]	0.71[a]
	numerical integration based on normal margins	0.48[a]	0.51[a]	0.50[a]
8	Monte Carlo	1.66	0.13	0.73
	numerical integration based on copula margins	1.46[b]	1.29[b]	1.91[b]
	numerical integration based on normal margins	2.15[c]	3.20[c]	2.14[c]

Table 4.5: Kullback-Leibler distances of the proposed vine copula models with respect to the true R-vine copula model ([a]estimated relative error < 0.01, [b]estimated relative error ≈ 1.4, [c]estimated relative error ≈ 3.5).

approaches do not maintain their nominal size in case of $n = 500$ and $n = 750$. All other GOF tests do hold the 5% level and thus control the type I error.

Sample size effects on power: We have increasing power with increasing sample size for the White, IR, ECP, ECP2 and Breymann (in combination with the AD test statistic) GOF test. The tests based on Berg and Berg2 have no or very low power independently of the number of observations. This is also true for the Breymann GOF test in combination with the univariate CvM and KS test statistics. In eight dimensions the number of sample points are important for the IR test since the tests has very small power considering only 500 data points. In five dimensions the effect is not that eye-catching but can be found too. Almost independent from the the number of sample points is the ordering of the test by their power. In all test scenarios the ECP2 test with mCvM test statistic outperforms the others, followed by the IR test, the test based on White and the ECP2 test based on the mKS test statistic. The next GOF tests are the tests based on the ECP and the Breymann transformation with AD test statistic.

Dimension effect on the power: The power of the top four GOF tests (IR, White, ECP and ECP2) are almost independent of the dimension. Only in the case of $n = 500$ sample points a clearly increase of power can be observed from $d = 5$ to $d = 8$ dimensions. For the weaker tests the reverse is true. With increasing dimension the Breymann GOF test decreases in power. The Berg and Berg2 tests are independently of the dimension.

Effect of alternatives on the power: The results with respect to the KLIC are two-fold. For $d = 5$ the power increases with increasing KLIC for most the GOF tests except for the Gauss copula in H_1. For $d = 8$ it is again the multivariate Gauss copula which is out of line for many of the tests. The exceptions are the ECP tests. For $n \geq 1000$ the power of the four “good” tests mentioned before increases with KLIC. Some of them have even a power of 100%. The Breymann test is conspicuous, since the test is working quite well for the C- and D-vine alternative but is relatively poor for the multivariate Gaussian copula independent of the dimension or sample size. While the Breymann tests have much lower power than the four best GOF tests, they still have power to distinguish between the null and alternative models.

Effects of the test functionals on power: For ECP, ECP2 or Breymann tests it appears that CvM based test statistics are more powerful than the KS type test statistics. This is in line with Genest et al. (2009) for bivariate copula GOF tests.

The poor performance of the Breymann, Berg and Berg2 approach was also recognized in the comparison studies of Genest et al. (2009) in the bivariate case and in Berg (2009) for copulas of dimension 2, 4 and 8. The analyzed copulas in Berg (2009) were the Gauss, Student's t, Clayton, Gumbel and Frank copula. But there the test statistics maintained their nominal level and had some explanatory power.

The bootstrapped p-values or power values stabilize fast for increasing bootstrap replications, for all GOF tests. This happens for 1000-1500 replications, irrespective of sample size or alternative. In many cases the stabilization is even faster.

Beside these last points, no clear hierarchy among the best performing proposed test statistics is recognizable. But some tests perform rather well while others do not even maintain their nominal level. In particular, our new IR test performs quite well in terms of power against false alternatives.

Of cause the computation time for the different proposed GOF tests is also a point of interest for practical applications. Therefore, in Table 4.6 the computation times in seconds for the different methods run on a Intel(R) Core(TM) i5-2450M CPU @ 2.50GHz computer for $n = 1000$ are given alongside with a summary of our findings. The computing time of the information matrix based methods White and IR are clearly higher than the other test statistics. Given the complex calculation of the R-vine gradient and Hessian matrix (see Chapter 3) this is not very surprising.

4.5 Examples

4.5.1 Indices and volatility indices

As first application we consider a financial data set of four indices and their corresponding volatility indices, namely the German DAX and VDAX-NEW, the European EuroSTOXX50 and VSTOXX, the US S&P500 and VIX, and the Swiss SMI and VSMI. The daily data cover the time horizon of the current financial crisis starting at August, 9th, 2007 when a sharp increase of inter bank interest rates was noticed, until April 30th, 2013, resulting in 1405 data points. For each marginal time series we calculated the log-returns and modeled them with an AR(1)-GARCH(1,1) model using Student's t innovations. The resulting standardized residuals are transformed using the non-parametric rank transformation (see Genest et al. 1995) to obtain $[0, 1]^8$ copula data.

The contour and pair plots in Figure 4.11 reveal the expected elliptical positive dependence behavior among the indices and among the volatility indices. But between the indices and the volatilities a negative dependence can be observed. Furthermore, a slight asymmetric tail dependence is recognizable.

To model the dependence structure we investigated four models. In particular, an R-vine copula model, selected using the maximum spanning tree algorithm by Dißmann et al. (2013), a C-vine copula, selected by the heuristic proposed by Czado et al. (2012), a D-vine copula, selected using a traveling sales man algorithm, and a multivariate Gaussian copula. The corresponding first trees of the vine models are illustrated in Figure 4.12. For the R-vine copula as well as in the D-vine copula we can see that the indices and the volatilities cluster except for the US ones. The C-vine copula is too restrictive to recognize such groupings. Another interesting point is that the first tree structure of the R-vine is

	White	Breymann	Berg	Berg2	ECP	ECP2	IR
main idea	$\mathbb{H}(\boldsymbol{\theta}) + \mathbb{C}(\boldsymbol{\theta}) = 0$	PIT+Aggregation+uniform test			$\hat{C}_n - C_{\hat{\boldsymbol{\theta}}_n}$	PIT+ECP	$-\mathbb{H}(\boldsymbol{\theta})^{-1}\mathbb{C}(\boldsymbol{\theta}) = I_p$
hold nominal level	+	+	−	−	+	+	+
power against alternatives	+	0 (partly)	−	−	0 (partly)	+	+
consistentency	+	−	−	−	−	+	+
asymptotic distribution	+ (only for high n)	0 (proved to be incorrect)	−	−	−	−	+ (not tested)
complexity	0 (difficult cov. matrix)	− (3 step procedure)	−	−	+	0 (2 step procedure)	+
computation time							
$d = 5$	3.41	0.06	0.06	0.07	0.08	0.06	1.66
$d = 8$	58.79	0.14	0.14	0.18	0.17	0.13	30.30

Table 4.6: Overview of the performance of the proposed GOF tests.

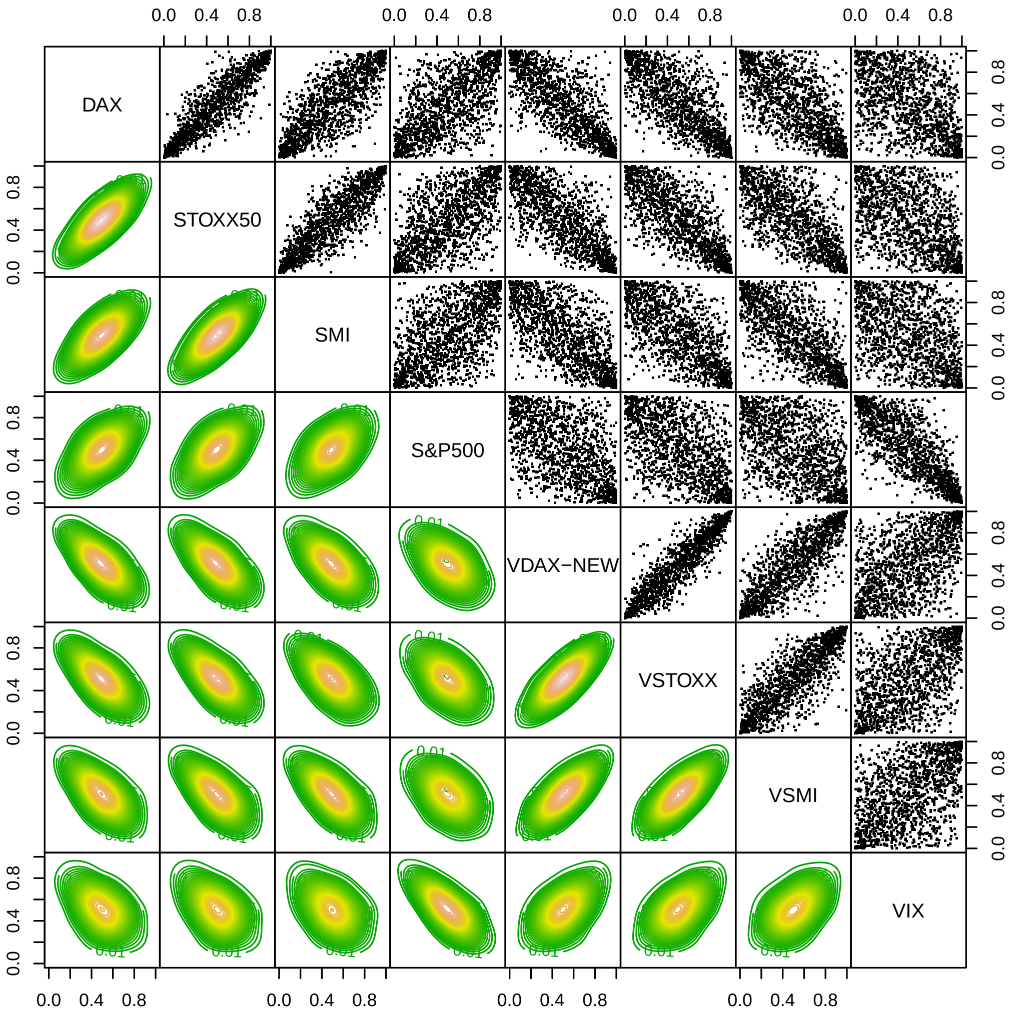


Figure 4.11: Lower left: Contour plots with standard normal margins; upper right: pairs plots of the transformed data set.

very close to the first tree structure of the D-vine. If we delete the edge "DAX-VDAX-NEW" and add a new edge "VSMI-SMI" in the R-vine we get the D-vine tree structure. Further, we see evidence of asymmetric tail dependence since (rotated) Gumbel copulas are selected.

Performing a parametric bootstrap with $B = 2500$ most of the good performing proposed GOF tests, namely White, IR, ECP (with CvM) and ECP2 (with CvM), confirm that a vine copula model can not be rejected at a 5% significance level (see Table 4.7). Only the ECP2 approach returns a p-value of 0.01 below the chosen significance level of 0.05 for the estimated C-vine copula, and the White based test a pvalue < 0.01 for the estimated R-vine copula model. The multivariate Gauss copula is rejected by the White, the IR and ECP2 GOF test, while the ECP based test returns a p-value of 0.6. In 3 of 4 GOF tests the highest returned p-value is for the D-vine copula. But note that the size of the p-value or the ordering of the p-values do not give an ordering of the considered models.

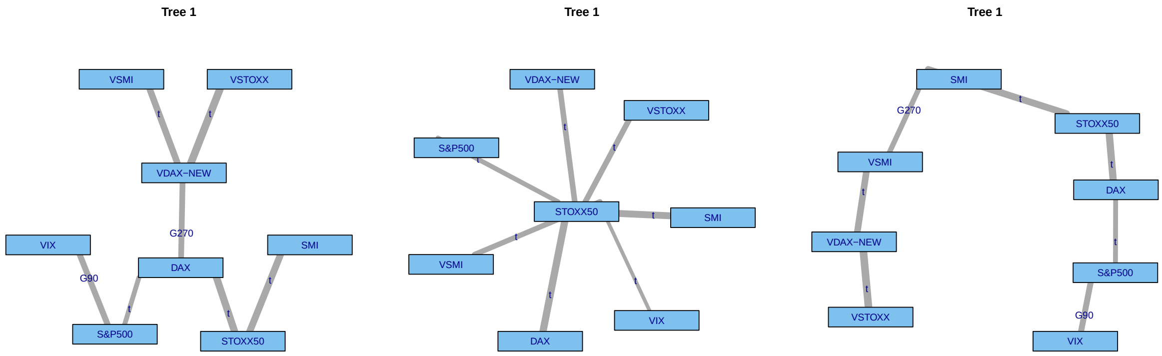


Figure 4.12: First tree structure of the selected R-vine (left), C-vine (center) and D-vine (right). The edge label denote the corresponding pair-copula family ($t \hat{=}$Student's t; $G90, G270 \hat{=}$rotated Gumbel).

As in the simulation study the GOF tests differ in their rejection decision and several GOF tests are needed to get a better picture of the better fitting model. The discrimination between the estimated vine copula models is even harder than in the power study. The MC-estimated KLIC of the R-vine to the C-vine is only 0.15, while the KLIC of the R-vine to the D-vine is even smaller (0.11). Even the multivariate Gauss copula has an estimated small KLIC with 0.31. Additional simulation studies based on the estimated vine copula models for $n = 1000$ show that the simulated power is quite small for all proposed GOF tests.

In terms of log-likelihood the D-vine is also the best fitting vine copula to the data unless the R-vine has a better AIC and BIC. The significant smaller number of parameters favors the R-vine compared to the D- or C-vine.

The economical interpretation of these findings is, that the assumption of multivariate Gaussian distributed random vectors is not fulfilled in times of financial and economic crises. More flexible models are needed to capture the asymmetric behaviors and tail dependencies. R-vines are able to model these properties as already shown in Brechmann and Czado (2013), Almeida and Czado (2012) or Min and Czado (2012).

	log-lik	#par	AIC	BIC	White	IR	ECP		ECP2	
							CvM	KS	CvM	KS
R-vine	7652	33	**-15238**	**-15065**	0.002	0.75	0.18	0.98	0.30	0.67
C-vine	7585	42	-15086	-14865	0.14	0.74	0.51	0.36	0.01	< 0.01
D-vine	**7654**	41	-15226	-15011	0.41	0.52	0.82	0.24	0.55	0.67
Gauss	7320	28	-14584	-14445	< 0.01	< 0.01	0.60	0.28	< 0.01	< 0.01

Table 4.7: Likelihood based validation quantities and bootstrapped p-values of the White, ECP, ECP2 and IR goodness-of-fit test for the 4 considered (vine) copula models

4.5.2 Exchange rate data

The second example to illustrate the introduced goodness-of-fit tests is the exchange rate data set of Section 3.4. We investigate three different fitted vine models. First, we consider a C-vine as already discussed in Schepsmeier (2010, Chapter 5) and Czado et al. (2012). Further, the R-vine of Section 3.4 and Appendix A.3, which was also already applied in Stöber and Czado (2012) in a regime switching model, is taken into consideration. Finally, a multivariate Gauss copula is considered, as commonly often applied in finance.

Performing a parametric bootstrap with repetition rate $B = 2500$ and sample size $N = 5000$ our goodness-of-fit test results (see Table 4.8) confirm that the C-vine model of Czado et al. (2012) can not be rejected at a 5% significance level given the White, ECP and ECP2 test. The R-vine model of Section 3.4 has only a bootstrapped p-value of 2% in the White test, but is greater than the significance level of 5% in all other considered tests. As in the example before the ECP based tests can not reject the assumption of a multivariate Gauss copula. All other GOF test reject the multivariate Gauss copula at a significance level of 5%.

The log-likelihood, AIC and BIC show a similar picture since the C-vine is preferred in terms of log-likelihood and AIC. Looking at the BIC criterion the R-vine is favorable due to its relative small number of parameters and a quite high log-likelihood. Also, our goodness-of-fit tests are consistent with the previous findings that the more flexible R-vine and C-vine are appropriate to model financial data.

	log-lik	#par	AIC	BIC	White	IR	ECP		ECP2	
							CvM	KS	CvM	KS
C-vine	**2213**	34	**-4358**	-4191	0.51	0.04	1	0.72	0.11	0.56
R-vine	2199	28	-4343	**-4205**	0.02	0.06	1	0.92	0.08	0.57
Gauss	2089	29	-4121	-3984	0.04	< 0.01	0.99	0.78	0.03	0.03

Table 4.8: Likelihood based validation quantities and bootstrapped p-values of the White, ECP, ECP2 and IR goodness-of-fit test for the 4 considered (vine) copula models

Chapter 5

Outlook and discussion

In the recent chapters we have introduced two important statistical inference tools for the very flexible class of regular vine copula models. The possibility to estimate standard errors of the estimated vine copula parameters closes a gap in the corresponding literature. Therefore, algorithms for the R-vine log-likelihood gradient and Hessian matrix are developed. The necessary first and second derivatives with respect to the copula parameters of bivariate copulas are derived and implemented. Now statisticians are able to quantify the uncertainty of their vine copula parameter estimates. Secondly, we discussed several goodness-of-fit tests for R-vine copulas. In particular, two new approaches arising from the Information matrix equality and specification test of White (1982) were introduced and their small sample performance studied. For comparison other GOF tests were extended from the bivariate copula case.

The algorithm for the gradient allows us to replace finite-differences of R-vine log-likelihood functions with the analytical gradient. Thus a great numerical improvement in the computation of the MLE can be achieved. As mentioned in Section 3.1 a decrease of computation time by a factor of 4-8 is possible. Other methods like maximization by parts of Song et al. (2005) will also gain from the analytical derivatives of the (copula) log-likelihood.

Although the calculation of copula log-likelihood derivatives seem to be a rather technical issue, it is quite challenging and was derived with flaws in the past. In particular the derivatives of the Student's t-copula with respect to the degrees-of-freedom parameter was wrong in the existing literature, for example Nadarajah (2006).

Despite the fast and optimized implementation in C the algorithms for the calculation of the log-likelihood gradient and Hessian matrix of an R-vine are very time consuming. Assuming only one-parametric copula families for all pair-copulas of the PCC $d(d-1)/2$ parameters have to be estimated. Thus the dimension of the parameter vector $\boldsymbol{\theta}(\mathcal{B}(\mathcal{V}))$ grows quadratically in d and the dimension of the Hessian matrix roughly in the order of d^4. It follows that the estimation of the Fisher information by the negative Hessian matrix and thereby the estimation of standard errors is limited to small to medium dimensions of data sets.

Therefore truncation methods for the vine copulas have to be considered to reduce the number of model parameters. First approaches of parameter cutback are already proposed by Brechmann et al. (2012) truncating the vine tree structure after a certain tree level. Here, truncation means that all (conditional) pair-copulas are set to the independence copula after the selected tree level. It is the authors opinion that further approaches have to be considered to stay computationally trackable. One idea is to include empirical copulas as building blocks. Especially in higher order trees. A first non-parametric method is suggested by Haff and Segers (2012) using an idea similar to the empirical copula, called empirical pair-copula. Here only one bandwidth parameter $h > 0$ has to be selected.

But current practice is to model the vine parametrically. New approaches of non-parametric vine copulas try to obviate the possible false copula selection form a set of parametric copula families. In some cases parametric copulas are too limited to catch for example the specific tail dependence of a pair of variables. Non-parametric copulas estimation can be more flexible since they refrain from any strong parametric assumption of the data structure. Vine copula models with exclusive non-parametric pair-copulas are for example proposed by Weiß and Scheffer (2012) using non-parametric Bernstein copulas, or by Kauermann and Schellhase (2013) using bivariate penalized splines. Since so called non-parametric copulas need parameters as well, the term "non-parametric" may be misleading. Here the parameters are smoothing, bandwidth (Weiß and Scheffer 2012), or spline and penalization parameters (Kauermann and Schellhase 2013), which are not connected anymore to copula moments such as Kendall's τ. If at all the number of model parameters is not reduced significantly.

Another important issue, which is broadly discussed in this thesis, is the validation of R-vine copula models using goodness-of-fit tests. We introduced two new goodness-of-fit tests for vine copulas based on White's information matrix test. The calculation of the test statistics as well as their asymptotic distribution functions showed up to be challenging. But good empirical approximations have been found as shown in several extensive power studies. The studies revealed good performance of the tests in terms of power against false alternatives given simulated p-values. Given sufficient data points the tests are even empirical consistent. Furthermore, the new GOF tests maintained always their nominal level, controlling the type I error, independently of sample size, dimension or alternative.

Further GOF tests are extended from the (bivariate) copula case to facilitate a wider comparison. The small sample performance for size and power were investigated for GOF tests based on the empirical copula process and the multivariate PIT. The application of the PIT data in the empirical copula process was also considered, which showed good performance results.

We discuss now some additional approaches for copula GOF tests, which might yield further useful GOF tests for regular vines.

1. Of cause further known GOF tests for copulas can be extended to the vine copula case. But most of them will have major problems in higher dimensions. For example the **likelihood ratio based GOF test** or the **Chi-squared type GOF test**, both introduced for copulas by Dobrić and Schmid (2005), have to partition the unit hypercube. This will probably result in long computation time in high dimensions as well as the need of sufficient large number of observations.

2. Berg and Aas (2009) developed a further GOF test based on the **Kendall's process**. To define the Kendall's process we need $U_1, \ldots, U_d$ uniform distributed random variables with joint distribution (copula) C. Then, Kendall's transform K is the (univariate) distribution of the random variable $V = C(U_1, \ldots, U_d)$. Now we consider $\boldsymbol{u}_j = (u_{j1}, \ldots, u_{jd})^T$ as $U(0,1)^d$ pseudo-observations, defined as normalized ranks. Further, let $\hat{\boldsymbol{\theta}}_n$ the maximum likelihood estimator corresponding to the parametric copula $C_{\boldsymbol{\theta}}$ given n observations. Now, Berg and Aas (2009) apply a GOF test based on Kendall's process $\mathcal{K}_n = \sqrt{n}\{K_n - K_{\hat{\boldsymbol{\theta}}_n}\}$ to a four dimensional vine copula. Here,

$$K_n(t) = \frac{1}{n+1} \sum_{j=1}^{n} 1_{\{C_n(\boldsymbol{u}_j) \leq t\}} \qquad t \in (0,1)$$

is the empirical distribution function of $C_n(\boldsymbol{u})$. Further, $K_{\hat{\boldsymbol{\theta}}_n}(t) = P(C_{\hat{\boldsymbol{\theta}}_n} \leq t)$ is the parametric estimate of Kendall's dependence function $K(t)$. Remember C_n is the empirical copula introduced by Deheuvels (1979),

$$C_n(\boldsymbol{u}) = \frac{1}{n+1} \sum_{t=1}^{n} \mathbf{1}_{\{u_{t1} \leq u_1, \ldots, u_{td} \leq u_d\}}.$$

$\mathcal{K}_n$ is then used in the Cramér-von Mises statistic, i.e.:

$$T_n = \int_{[0,1]^d} \{K_n(\boldsymbol{u}) - K_{\hat{\boldsymbol{\theta}}_n}(\boldsymbol{u})\}^2 dK_n(\boldsymbol{u}) = \sum_{j=1}^{n} \left\{K_n(\boldsymbol{u}_j) - K_{\hat{\boldsymbol{\theta}}_n}(\boldsymbol{u}_j)\right\}^2.$$

This approach revealed good results in the comparison study of Genest et al. (2009) for bivariate copulas, where $K_{\hat{\boldsymbol{\theta}}_n}$ has a closed form. For the vine copulas the Kendall's transform is not trackable. Thus the Kendall's process based GOF test needs like the ECP based GOF tests a double bootstrap procedure. Unfortunately Berg and Aas (2009) did not test or investigate this GOF test in detail for the vine copula case.

3. Further suggestions for copula GOF tests are for example presented in Fermanian (2012). Some special **GOF test designed for Archimedean or extreme value copulas** only are presented. Since the multivariate Clayton copula is the only Archimedean copula which allows a vine representation (Stöber et al. 2013), this approach is limited for vine copula GOF tests. Further, a **density based GOF** is proposed to avoid the calculation of the copula cdf like in the ECP or Kendall's process approach. Testing the closeness between the true copula density and one of its estimates out of a set of possible copula families is equivalent to study the identity $C = C_0$ (Fermanian 2005).

4. A very interesting hybrid approach was suggested by Zhang et al. (2013). Since no GOF test outperforms in all cases a **hybrid test** is introduced. In our simulation studies we noted, that for different scenarios or even just for different alternatives diverse GOF tests performed best. Although the Information based tests performed in most of the investigated scenarios best in terms of power against false alternatives

the ECP based GOF tests outperformed them in some settings. Thus a hybrid test for vine copula goodness-of-fit testing may be a good extension to the proposed tests. Given the complex structure of vine copula models the strategy of a hybrid test is particularly appealing. We hope that a hybrid test approach will lead to more consistent test decisions in those cases where different GOF tests come to conflicting test decisions (see for example the two applications in Section 4.5). As possible members of a hybrid test we suggest the Information based test, namely the White test and the IR test, as well as the ECP and ECP2 test showing good performing results in our studies. The latter ones in the Cramér-von Mises type. But other GOF tests may be considered as well.

We discuss now in more detail the hybrid approach. Given m test statistics $t_n^{(i)}, i = 1, \ldots, m$ with sample size n and controlling type I error for any given significance level α under the null hypothesis the hybrid p-value is defined as

$$p_n^{hybrid} := m * \min\{p_n^{(1)}, \ldots, p_n^{(m)}\}.$$

Here $p_n^{(i)}, i = 1, \ldots, m$ denote the p-values of the test statistics $t_n^{(i)}$. Zhang et al. (2013) showed that the power function is bounded from below and if there is at least one test which is consistent, then the hybrid test is consistent. The rejection rule of the hybrid test is

$$\text{Reject } H_0 : C \in \mathcal{C}_0 = \{C_{\boldsymbol{\theta}} : \boldsymbol{\theta} \in \Theta\} \quad \text{versus} \quad H_1 : C \notin \mathcal{C}_0 \quad \Leftrightarrow \quad p_n^{hybrid} \leq \alpha,$$

where C denotes the (vine) copula distribution function and $\mathcal{C}_0$ is a class of parametric (vine) copulas with $\Theta \subseteq \mathbb{R}^p$ being the parameter space of dimension p. This is equivalent to the situation where there is at least one test rejecting the null at the level of α/m (Zhang et al. 2013).

Beside the choice of the best GOF test there is another issue to discuss. By testing the validity of the null hypothesis $H_0 : C \in \mathcal{C}_0$ one has to take the margins into account if the margins are unknown. As pointed out by Genest et al. (2009) the marginal distribution functions $F_1, \ldots, F_d$ of the random variables $X_1, \ldots, X_d$ can be considered as nuisance parameters. So far we always considered known margins. Only for the White test we extended the GOF test statistic to unknown margins so far (see Section 4.1.4). Thus an extension of the proposed GOF tests to unknown margins has to be considered in the future. But this will be a tough question resulting possibly in numerical difficult solutions as in the White case. A more practical approach is to validate the margins first and then the dependence structure resulting from using the validated margins.

A more interesting question is the extension of the GOF tests to variable vine structures. Remember that we assume a known tree structure for the considered tests so far. Thus GOF tests over structures of vines would be an interesting point. The question arises "Is the underlying dependence structure of a given data set an R-vine?" Or, how much influence does a specific vine structure have on the goodness of fit? For example different vine copula structures can have a similar good fit in terms of likelihood, AIC or BIC.

Different selection heuristics may estimate different vine copulas, including different vine tree structures or just different pair-copula families, and return a comparable fit. Such heuristics are for example the maximum spanning tree approach of Dißmann et al. (2013) or the Bayesian approach of Gruber and Czado (2012).

Model comparison in terms of distances such as the Kullback and Leibler (1951) information criterion (KLIC) are computational difficult to handle and not trackable in higher dimension. As shown in the examples of Section 4.4 and Section 4.5 the results of the numerical integration algorithms can be quite unstable. Furthermore, the number of possible vine copula models grows exponentially in the dimension, making direct comparisons between all members of a specified vine copula class impossible. Thus a goodness-of-fit test checking a general R-vine class would be helpful.

Note that in an extended case the R-vine does not have to be specified in terms of structure $\mathcal{V}$, pair-copula families $\mathcal{B}(\mathcal{V})$ or $\boldsymbol{\theta}(\mathcal{B}(\mathcal{V}))$. Consider for example an arbitrary C-, D- or R-vine with only Student's t-copulas or any other fixed copula family. Or the class of truncated R-vines with arbitrary copula families on the first trees and independence pair-copulas in higher trees. The test problem would be then

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

where $\mathcal{C}_0$ is a subclass of arbitrary R-vine copula models. The GOF tests proposed in Chapter 4 or in the previous paragraphs would be special cases defining $\mathcal{C}_0 = \{C_{\boldsymbol{\theta}} : \boldsymbol{\theta} \in \Theta \text{ and } \mathcal{V}_{C_{\boldsymbol{\theta}}} \text{ fix}\}$. To the authors knowledge there is no such GOF test available.

Appendix A

Appendix for Estimating standard errors

A.1 Algorithm for the calculation of second derivatives

In Section 3.2 we introduced the seven possible cases of dependence which can occur during the calculation of the second log-likelihood derivative. In the following, we illustrate these cases in detail. In case 1 we determine

$$\begin{aligned}&\frac{\partial^2}{\partial\theta\partial\gamma}\ln\left(c_{U,V|\mathbf{z}}\left(F_{U|\mathbf{z}}(u|\mathbf{z},\theta),F_{V|\mathbf{z}}(v|\mathbf{z},\gamma)\right)\right)\\&=\frac{\partial_1\partial_2 c_{U,V|\mathbf{z}}\left(F_{U|\mathbf{z}}(u|\mathbf{z},\theta),F_{V|\mathbf{z}}(v|\mathbf{z},\gamma)\right)}{c_{U,V|\mathbf{z}}\left(F_{U|\mathbf{z}}(u|\mathbf{z},\theta),F_{V|\mathbf{z}}(v|\mathbf{z},\gamma)\right)}\cdot\left(\frac{\partial}{\partial\theta}F_{U|\mathbf{z}}(u|\mathbf{z},\theta)\right)\cdot\left(\frac{\partial}{\partial\gamma}F_{V|\mathbf{z}}(v|\mathbf{z},\gamma)\right)\\&-\left(\frac{\partial}{\partial\theta}\ln\left(c_{U,V|\mathbf{z}}\left(F_{U|\mathbf{z}}(u|\mathbf{z},\theta),F_{V|\mathbf{z}}(v|\mathbf{z},\gamma)\right)\right)\right)\\&\cdot\left(\frac{\partial}{\partial\gamma}\ln\left(c_{U,V|\mathbf{z}}\left(F_{U|\mathbf{z}}(u|\mathbf{z},\theta),F_{V|\mathbf{z}}(v|\mathbf{z},\gamma)\right)\right)\right),\end{aligned} \tag{A.1}$$

for case 2

$$\begin{aligned}&\frac{\partial^2}{\partial\theta\partial\gamma}\ln\left(c_{U,V|\mathbf{z}}\left(F_{U|\mathbf{z}}(u|\mathbf{z},\theta),F_{V|\mathbf{z}}(v|\mathbf{z},\theta,\gamma)\right)\right)\\&=\frac{\partial_1\partial_2 c_{U,V|\mathbf{z}}\left(F_{U|\mathbf{z}}(u|\mathbf{z},\theta),F_{V|\mathbf{z}}(v|\mathbf{z},\theta,\gamma)\right)}{c_{U,V|\mathbf{z}}\left(F_{U|\mathbf{z}}(u|\mathbf{z},\theta),F_{V|\mathbf{z}}(v|\mathbf{z},\theta,\gamma)\right)}\cdot\left(\frac{\partial}{\partial\theta}F_{U|\mathbf{z}}(u|\mathbf{z},\theta)\right)\cdot\left(\frac{\partial}{\partial\gamma}F_{V|\mathbf{z}}(v|\mathbf{z},\theta,\gamma)\right)\\&-\left(\frac{\partial}{\partial\theta}\ln\left(c_{U,V|\mathbf{z}}\left(F_{U|\mathbf{z}}(u|\mathbf{z},\theta),F_{V|\mathbf{z}}(v|\mathbf{z},\theta,\gamma)\right)\right)\right)\\&\cdot\left(\frac{\partial}{\partial\gamma}\ln(\left(c_{U,V|\mathbf{z}}\left(F_{U|\mathbf{z}}(u|\mathbf{z},\theta),F_{V|\mathbf{z}}(v|\mathbf{z},\theta,\gamma)\right)\right)\right)\end{aligned} \tag{A.2}$$

$$+\frac{\partial_2\partial_2 c_{U,V|\mathbf{z}}\left(F_{U|\mathbf{z}}(u|\mathbf{z},\theta),F_{V|\mathbf{z}}(v|\mathbf{z},\theta,\gamma)\right)}{c_{U,V|\mathbf{z}}\left(F_{U|\mathbf{z}}(u|\mathbf{z},\theta),F_{V|\mathbf{z}}(v|\mathbf{z},\theta,\gamma)\right)}\cdot\left(\frac{\partial}{\partial\theta}F_{V|\mathbf{z}}(v|\mathbf{z},\theta,\gamma)\right)\cdot\left(\frac{\partial}{\partial\gamma}F_{V|\mathbf{z}}(v|\mathbf{z},\theta,\gamma)\right)$$
$$+\frac{\partial_2 c_{U,V|\mathbf{z}}\left(F_{U|\mathbf{z}}(u|\mathbf{z},\theta),F_{V|\mathbf{z}}(v|\mathbf{z},\theta,\gamma)\right)}{c_{U,V|\mathbf{z}}\left(F_{U|\mathbf{z}}(u|\mathbf{z},\theta),F_{V|\mathbf{z}}(v|\mathbf{z},\theta,\gamma)\right)}\cdot\left(\frac{\partial^2}{\partial\theta\partial\gamma}F_{V|\mathbf{z}}(v|\mathbf{z},\theta,\gamma)\right), \tag{A.3}$$

and case 3 yields

$$\frac{\partial^2}{\partial\theta\partial\gamma}\ln(c_{U,V|\mathbf{z}}(F_{U|\mathbf{z}}(u|\mathbf{z},\theta,\gamma),F_{V|\mathbf{z}}(v|\mathbf{z},\theta,\gamma)))$$
$$=\frac{\partial_1\partial_2 c_{U,V|\mathbf{z}}\left(F_{U|\mathbf{z}}(u|\mathbf{z},\theta,\gamma),F_{V|\mathbf{z}}(v|\mathbf{z},\theta,\gamma)\right)}{c_{U,V|\mathbf{z}}\left(F_{U|\mathbf{z}}(u|\mathbf{z},\theta,\gamma),F_{V|\mathbf{z}}(v|\mathbf{z},\theta,\gamma)\right)}\cdot\left(\frac{\partial}{\partial\theta}F_{U|\mathbf{z}}(u|\mathbf{z},\theta,\gamma)\right)\left(\frac{\partial}{\partial\gamma}F_{V|\mathbf{z}}(v|\mathbf{z},\theta,\gamma)\right)$$
$$-\left(\frac{\partial}{\partial\theta}\ln\left(c_{U,V|\mathbf{z}}\left(F_{U|\mathbf{z}}(u|\mathbf{z},\theta,\gamma),F_{V|\mathbf{z}}(v|\mathbf{z},\theta,\gamma)\right)\right)\right)$$
$$\cdot\left(\frac{\partial}{\partial\gamma}\ln(\left(c_{U,V|\mathbf{z}}\left(F_{U|\mathbf{z}}(u|\mathbf{z},\theta,\gamma),F_{V|\mathbf{z}}(v|\mathbf{z},\theta,\gamma)\right)\right)\right)$$
$$+\frac{\partial_1\partial_1 c_{U,V|\mathbf{z}}\left(F_{U|\mathbf{z}}(u|\mathbf{z},\theta,\gamma),F_{V|\mathbf{z}}(v|\mathbf{z},\theta,\gamma)\right)}{c_{U,V|\mathbf{z}}\left(F_{U|\mathbf{z}}(u|\mathbf{z},\theta,\gamma),F_{V|\mathbf{z}}(v|\mathbf{z},\theta,\gamma)\right)}\cdot\left(\frac{\partial}{\partial\theta}F_{U|\mathbf{z}}(u|\mathbf{z},\theta,\gamma)\right)\left(\frac{\partial}{\partial\gamma}F_{U|\mathbf{z}}(u|\mathbf{z},\theta,\gamma)\right)$$
$$+\frac{\partial_1\partial_2 c_{U,V|\mathbf{z}}\left(F_{U|\mathbf{z}}(u|\mathbf{z},\theta,\gamma),F_{V|\mathbf{z}}(v|\mathbf{z},\theta,\gamma)\right)}{c_{U,V|\mathbf{z}}\left(F_{U|\mathbf{z}}(u|\mathbf{z},\theta,\gamma),F_{V|\mathbf{z}}(v|\mathbf{z},\theta,\gamma)\right)}\cdot\left(\frac{\partial}{\partial\theta}F_{V|\mathbf{z}}(v|\mathbf{z},\theta,\gamma)\right)\left(\frac{\partial}{\partial\gamma}F_{U|\mathbf{z}}(u|\mathbf{z},\theta,\gamma)\right)$$
$$+\frac{\partial_1\partial_2 c_{U,V|\mathbf{z}}\left(F_{U|\mathbf{z}}(u|\mathbf{z},\theta,\gamma),F_{V|\mathbf{z}}(v|\mathbf{z},\theta,\gamma)\right)}{c_{U,V|\mathbf{z}}\left(F_{U|\mathbf{z}}(u|\mathbf{z},\theta,\gamma),F_{V|\mathbf{z}}(v|\mathbf{z},\theta,\gamma)\right)}\cdot\left(\frac{\partial}{\partial\theta}F_{U|\mathbf{z}}(u|\mathbf{z},\theta,\gamma)\right)\left(\frac{\partial}{\partial\gamma}F_{V|\mathbf{z}}(v|\mathbf{z},\theta,\gamma)\right)$$
$$+\frac{\partial_2\partial_2 c_{U,V|\mathbf{z}}\left(F_{U|\mathbf{z}}(u|\mathbf{z},\theta,\gamma),F_{V|\mathbf{z}}(v|\mathbf{z},\theta,\gamma)\right)}{c_{U,V|\mathbf{z}}\left(F_{U|\mathbf{z}}(u|\mathbf{z},\theta,\gamma),F_{V|\mathbf{z}}(v|\mathbf{z},\theta,\gamma)\right)}\cdot\left(\frac{\partial}{\partial\theta}F_{V|\mathbf{z}}(v|\mathbf{z},\theta,\gamma)\right)\left(\frac{\partial}{\partial\gamma}F_{V|\mathbf{z}}(v|\mathbf{z},\theta,\gamma)\right)$$
$$+\frac{\partial_1 c_{U,V|\mathbf{z}}\left(F_{U|\mathbf{z}}(u|\mathbf{z},\theta,\gamma),F_{V|\mathbf{z}}(v|\mathbf{z},\theta,\gamma)\right)}{c_{U,V|\mathbf{z}}\left(F_{U|\mathbf{z}}(u|\mathbf{z},\theta,\gamma),F_{V|\mathbf{z}}(v|\mathbf{z},\theta,\gamma)\right)}\cdot\left(\frac{\partial^2}{\partial\theta\partial\gamma}F_{V|\mathbf{z}}(v|\mathbf{z},\theta,\gamma)\right)$$
$$+\frac{\partial_2 c_{U,V|\mathbf{z}}\left(F_{U|\mathbf{z}}(u|\mathbf{z},\theta,\gamma),F_{V|\mathbf{z}}(v|\mathbf{z},\theta,\gamma)\right)}{c_{U,V|\mathbf{z}}\left(F_{U|\mathbf{z}}(u|\mathbf{z},\theta,\gamma),F_{V|\mathbf{z}}(v|\mathbf{z},\theta,\gamma)\right)}\cdot\left(\frac{\partial^2}{\partial\theta\partial\gamma}F_{V|\mathbf{z}}(v|\mathbf{z},\theta,\gamma)\right). \tag{A.4}$$

Similarly, we have for case 4 that

$$\frac{\partial^2}{\partial\theta\partial\gamma}\ln\left(c_{U,V|\mathbf{z}}\left(F_{U|\mathbf{z}}(u|\mathbf{z},\theta),F_{V|\mathbf{z}}(v|\mathbf{z})|\gamma\right)\right)$$
$$=\left(\frac{\partial}{\partial\theta}\ln\left(c_{U,V|\mathbf{z}}\left(F_{U|\mathbf{z}}(u|\mathbf{z},\theta),F_{V|\mathbf{z}}(v|\mathbf{z})|\gamma\right)\right)\right)\cdot\frac{-\partial_\gamma c_{U,V|\mathbf{z}}\left(F_{U|\mathbf{z}}(u|\mathbf{z},\theta),F_{V|\mathbf{z}}(v|\mathbf{z})|\gamma\right)}{c_{U,V|\mathbf{z}}\left(F_{U|\mathbf{z}}(u|\mathbf{z},\theta),F_{V|\mathbf{z}}(v|\mathbf{z})|\gamma\right)}$$
$$+\frac{\partial_\gamma\partial_1 c_{U,V|\mathbf{z}}\left(F_{U|\mathbf{z}}(u|\mathbf{z},\theta),F_{V|\mathbf{z}}(v|\mathbf{z})|\gamma\right)}{c_{U,V|\mathbf{z}}\left(F_{U|\mathbf{z}}(u|\mathbf{z},\theta),F_{V|\mathbf{z}}(v|\mathbf{z})|\gamma\right)}\cdot\left(\frac{\partial}{\partial\theta}F_{U|\mathbf{z}}(u|\mathbf{z},\theta)\right), \tag{A.5}$$

and

$$\begin{aligned}
&\frac{\partial^2}{\partial\theta\partial\gamma}\ln\left(c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z},\theta),F_{V|\mathbf{Z}}(v|\mathbf{z},\theta)|\gamma\right)\right)\\
&=\left(\frac{\partial}{\partial\theta}\ln\left(c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z},\theta),F_{V|\mathbf{Z}}(v|\mathbf{z},\theta)|\gamma\right)\right)\right)\\
&\cdot\frac{-\partial_\gamma c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z},\theta),F_{V|\mathbf{Z}}(v|\mathbf{z},\theta)|\gamma\right)}{c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z},\theta),F_{V|\mathbf{Z}}(v|\mathbf{z},\theta)|\gamma\right)}\\
&+\frac{\partial_\gamma\partial_1 c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z},\theta),F_{V|\mathbf{Z}}(v|\mathbf{z},\theta)|\gamma\right)}{c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z},\theta),F_{V|\mathbf{Z}}(v|\mathbf{z},\theta)|\gamma\right)}\cdot\left(\frac{\partial}{\partial\theta}F_{U|\mathbf{Z}}(u|\mathbf{z},\theta)\right)\\
&+\frac{\partial_\gamma\partial_2 c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z},\theta),F_{V|\mathbf{Z}}(v|\mathbf{z},\theta)|\gamma\right)}{c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z},\theta),F_{V|\mathbf{Z}}(v|\mathbf{z},\theta)|\gamma\right)}\cdot\left(\frac{\partial}{\partial\theta}F_{V|\mathbf{Z}}(v|\mathbf{z},\theta)\right),
\end{aligned}\tag{A.6}$$

for the fifth case. Finally,

$$\begin{aligned}
&\frac{\partial^2}{\partial\theta\partial\gamma}\ln\left(c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z},\theta,\gamma),F_{V|\mathbf{Z}}(v|\mathbf{z})\right)\right)\\
&=\frac{\partial_1\partial_1 c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z},\theta,\gamma),F_{V|\mathbf{Z}}(v|\mathbf{z})\right)}{c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z},\theta,\gamma),F_{V|\mathbf{Z}}(v|\mathbf{z})\right)}\cdot\left(\frac{\partial}{\partial\theta}F_{U|\mathbf{Z}}(u|\mathbf{z},\theta,\gamma)\right)\cdot\left(\frac{\partial}{\partial\gamma}F_{U|\mathbf{Z}}(u|\mathbf{z},\theta,\gamma)\right)\\
&-\left(\frac{\partial}{\partial\gamma}\ln\left(c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z},\theta,\gamma),F_{V|\mathbf{Z}}(v|\mathbf{z})\right)\right)\right)\\
&\cdot\left(\frac{\partial}{\partial\theta}\ln\left(c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z},\theta,\gamma),F_{V|\mathbf{Z}}(v|\mathbf{z})\right)\right)\right)\\
&+\frac{\partial_1 c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z},\theta,\gamma),F_{V|\mathbf{Z}}(v|\mathbf{z})\right)}{c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z},\theta,\gamma),F_{V|\mathbf{Z}}(v|\mathbf{z})\right)}\cdot\left(\frac{\partial^2}{\partial\gamma\partial\theta}F_{U|\mathbf{Z}}(u|\mathbf{z},\theta,\gamma)\right),
\end{aligned}\tag{A.7}$$

$$\begin{aligned}
&\frac{\partial^2}{\partial\theta\partial\gamma}\ln\left(c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z}),F_{V|\mathbf{Z}}(v|\mathbf{z})|\theta,\gamma\right)\right)\\
&=\frac{\partial_\theta\partial_\gamma c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u),F_{V|\mathbf{Z}}(v|\mathbf{z})|\theta,\gamma\right)}{c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z}),F_{V|\mathbf{Z}}(v|\mathbf{z})|\theta,\gamma\right)}\\
&-\frac{\partial_\theta c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u),F_{V|\mathbf{Z}}(v|\mathbf{z})|\theta,\gamma\right)\cdot\partial_\gamma c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z}),F_{V|\mathbf{Z}}(v)|\theta,\gamma\right)}{c_{U,V|\mathbf{Z}}\left(F_{U|\mathbf{Z}}(u|\mathbf{z}),F_{V|\mathbf{Z}}(v|\mathbf{z})|\theta,\gamma\right)^2}.
\end{aligned}\tag{A.8}$$

Algorithm A.1.1 Second derivative with respect to the parameters $\theta^{\tilde{k},\tilde{i}}$ and $\theta^{\hat{k},\hat{i}}$.

The input of the algorithm is a d-dimensional R-vine matrix M with maximum matrix $\tilde{M}$ and parameter matrix $\boldsymbol{\theta}$, and matrices $C^{\tilde{k},\tilde{i}}$, $C^{\hat{k},\hat{i}}$ determined using Algorithm 3.1.1 for parameters $\theta^{\tilde{k},\tilde{i}}$ and $\theta^{\hat{k},\hat{i}}$ of the R-vine parameter matrix. Further, we assume the matrices V^{direct}, $V^{indirect}$ and V^{values}, the matrices $S1^{direct,\tilde{k},\tilde{i}}$, $S1^{indirect,\tilde{k},\tilde{i}}$ and $S1^{values,\tilde{k},\tilde{i}}$ and $S1^{direct,\hat{k},\hat{i}}$, $S1^{indirect,\hat{k},\hat{i}}$ and $S1^{values,\hat{k},\hat{i}}$ to be given. The output will be the value of the second derivative of the copula log-likelihood for the given observation with respect to parameters $\theta^{\tilde{k},\tilde{i}}$ and $\theta^{\hat{k},\hat{i}}$. Without loss of generality, we assume that $\hat{i}\geq\tilde{i}$, and $\hat{k}\geq\tilde{k}$ if $\hat{i}=\tilde{i}$.

if $c^{\hat{k},\hat{i}}_{\tilde{k},\tilde{i}} == 1$ **then**
  Set $m = \tilde{m}_{\tilde{k},\tilde{i}}$
  Set $z_1 = v^{direct}_{\tilde{k},\tilde{i}}, \tilde{z}_1 = s1^{direct,\hat{k},\hat{i}}_{\tilde{k},\tilde{i}}$
  if $m == m_{\tilde{k},\tilde{i}}$ **then**
    Set $z_2 = v^{direct}_{\tilde{k},d-m+1}, \tilde{z}_2 = s1^{direct,\hat{k},\hat{i}}_{\tilde{k},d-m+1}$
  else
    Set $z_2 = v^{indirect}_{\tilde{k},d-m+1}, \tilde{z}_2 = s1^{indirect,\hat{k},\hat{i}}_{\tilde{k},d-m+1}$
  end if
  Set $s2^{direct}_{\tilde{k}-1,\tilde{i}} = 0, s2^{indirect}_{\tilde{k}-1,\tilde{i}} = 0, s2^{values}_{\tilde{k},\tilde{i}} = 0$
  if $\tilde{k} == \hat{k}$ & $\tilde{i} == \hat{i}$ **then**
    Set $s2^{direct}_{\tilde{k}-1,\tilde{i}} = \partial_{\theta^{\tilde{k},\tilde{i}}}\partial_{\theta^{\tilde{k},\tilde{i}}} h(z_1, z_2|\mathcal{B}^{\tilde{k},\tilde{i}}, \theta^{\tilde{k},\tilde{i}})$
    Set $s2^{indirect}_{\tilde{k}-1,\tilde{i}} = \partial_{\theta^{\tilde{k},\tilde{i}}}\partial_{\theta^{\tilde{k},\tilde{i}}} h(z_2, z_1|\mathcal{B}^{\tilde{k},\tilde{i}}, \theta^{\tilde{k},\tilde{i}})$
    Set $s2^{values}_{\tilde{k},\tilde{i}} = \frac{\partial_{\theta^{\tilde{k},\tilde{i}}}\partial_{\theta^{\tilde{k},\tilde{i}}} c(z_1,z_2|\mathcal{B}^{\tilde{k},\tilde{i}},\theta^{\tilde{k},\tilde{i}})}{exp(v^{values}_{\tilde{k},\tilde{i}})} - (s1^{values,\hat{k},\hat{i}}_{\tilde{k},\tilde{i}})^2$
  end if
  if $c^{\hat{k},\hat{i}}_{\tilde{k}+1,\tilde{i}} == 1$ **then**
    Set $s2^{values}_{\tilde{k},\tilde{i}} = s1^{values,\hat{k},\hat{i}}_{\tilde{k},\tilde{i}} \cdot \frac{-\partial_{\theta^{\tilde{k},\tilde{i}}} c(z_1,z_2|\mathcal{B}^{\tilde{k},\tilde{i}},\theta^{\tilde{k},\tilde{i}})}{exp(v^{values}_{\tilde{k},\tilde{i}})} + \frac{\partial_1\partial_{\theta^{\tilde{k},\tilde{i}}} c(z_1,z_2|\mathcal{B}^{\tilde{k},\tilde{i}},\theta^{\tilde{k},\tilde{i}})}{exp(v^{values}_{\tilde{k},\tilde{i}})} \cdot \tilde{z}_1$
    Set $s2^{direct}_{\tilde{k}-1,\tilde{i}} = \partial_1\partial_{\theta^{\tilde{k},\tilde{i}}} h(z_1, z_2|\mathcal{B}^{\tilde{k},\tilde{i}}, \theta^{\tilde{k},\tilde{i}}) \cdot \tilde{z}_1$
    Set $s2^{indirect}_{\tilde{k}-1,\tilde{i}} = \partial_2\partial_{\theta^{\tilde{k},\tilde{i}}} h(z_2, z_1|\mathcal{B}^{\tilde{k},\tilde{i}}, \theta^{\tilde{k},\tilde{i}}) \cdot \tilde{z}_1$
  end if
  if $c^{\hat{k},\hat{i}}_{\tilde{k}+1,d-m+1} == 1$ **then**
    Set $s2^{values}_{\tilde{k},\tilde{i}} = s2^{values}_{\tilde{k},\tilde{i}} + \frac{\partial_2\partial_{\theta^{\tilde{k},\tilde{i}}} c(z_1,z_2|\mathcal{B}^{\tilde{k},\tilde{i}},\theta^{\tilde{k},\tilde{i}})}{exp(v^{values}_{\tilde{k},\tilde{i}})} \cdot \tilde{z}_2$
    if $c^{\hat{k},\hat{i}}_{k+1,i} == 0$ **then**
      Set $s2^{values}_{\tilde{k},\tilde{i}} = s2^{values}_{\tilde{k},\tilde{i}} + s1^{values,\hat{k},\hat{i}}_{\tilde{k},\tilde{i}} \cdot \frac{-\partial_{\theta^{\tilde{k},\tilde{i}}} c(z_1,z_2|\mathcal{B}^{\tilde{k},\tilde{i}},\theta^{\tilde{k},\tilde{i}})}{exp(v^{values}_{\tilde{k},\tilde{i}})}$
    end if
    Set $s2^{direct}_{\tilde{k}-1,\tilde{i}} = s2^{direct}_{\tilde{k}-1,\tilde{i}} + \partial_2\partial_{\theta^{\tilde{k},\tilde{i}}} h(z_1, z_2|\mathcal{B}^{\tilde{k},\tilde{i}}, \theta^{\tilde{k},\tilde{i}}) \cdot \tilde{z}_2$
    Set $s2^{indirect}_{\tilde{k}-1,\tilde{i}} = s2^{indirect}_{\tilde{k}-1,\tilde{i}} + \partial_1\partial_{\theta^{\tilde{k},\tilde{i}}} h(z_2, z_1|\mathcal{B}^{\tilde{k},\tilde{i}}, \theta^{\tilde{k},\tilde{i}}) \cdot \tilde{z}_2$
  end if
end if
for $i = \tilde{i}, \ldots, 1$ **do**
  for $k = \tilde{k} - 1, \ldots, i + 1$ **do**
    Set $m = \tilde{m}_{k,i}$
    Set $z_1 = v^{direct}_{k,i}, \tilde{z}^{\hat{k},\hat{i}}_1 = s1^{direct,\hat{k},\hat{i}}_{k,i}, \tilde{z}^{\tilde{k},\tilde{i}}_1 = s1^{direct,\tilde{k},\tilde{i}}_{k,i}, \bar{z}_1 = s2^{direct}_{k,i}$
    if $m == m_{k,i}$ **then**
      Set $z_2 = v^{direct}_{k,d-m+1}, \tilde{z}^{\hat{k},\hat{i}}_2 = s1^{direct,\hat{k},\hat{i}}_{k,d-m+1}, \tilde{z}^{\tilde{k},\tilde{i}}_2 = s1^{direct,\tilde{k},\tilde{i}}_{k,d-m+1}, \bar{z}_2 = s2^{direct}_{k,d-m+1}$
    else
      Set $z_2 = v^{indirect}_{k,d-m+1}, \tilde{z}^{\hat{k},\hat{i}}_2 = s1^{indirect,\hat{k},\hat{i}}_{k,d-m+1}, \tilde{z}^{\tilde{k},\tilde{i}}_2 = s1^{indirect,\tilde{k},\tilde{i}}_{k,d-m+1}, \bar{z}_2 = s2^{indirect}_{k,d-m+1}$
    end if

Set $s2^{values}_{k,i} = -s1^{values,\hat{k},\hat{i}}_{k,i} \cdot s1^{values,\tilde{k},\tilde{i}}_{k,i}, s2^{direct}_{k-1,i} = 0, s2^{indirect}_{k-1,i} = 0$
if $c^{\hat{k},\hat{i}}_{k+1,i} == 1 \;\&\; c^{\tilde{k},\tilde{i}}_{k+1,i} == 1$ **then**
Set $s2^{values}_{k,i} = s2^{values}_{k,i} + \frac{\partial_1\partial_1 c(z_1,z_2|\mathcal{B}^{k,i},\theta^{k,i})}{exp(v^{values}_{k,i})} \cdot \tilde{z}_1^{\hat{k},\hat{i}} \cdot \tilde{z}_1^{\tilde{k},\tilde{i}} + \frac{\partial_1 c(z_1,z_2|\mathcal{B}^{k,i},\theta^{k,i})}{cop} \cdot \bar{z}_1$
Set $s2^{direct}_{k-1,i} = s2^{direct}_{k-1,i} + \partial_1 h(z_1,z_2|\mathcal{B}^{k,i},\theta^{k,i}) \cdot \bar{z}_1 + \partial_1\partial_1 h(z_1,z_2|\mathcal{B}^{k,i},\theta^{k,i}) \cdot \tilde{z}_1^{\hat{k},\hat{i}} \cdot \tilde{z}_1^{\tilde{k},\tilde{i}}$
Set $s2^{indirect}_{k-1,i} = s2^{indirect}_{k-1,i} + \partial_2 h(z_2,z_1|\mathcal{B}^{k,i},\theta^{k,i}) \cdot \bar{z}_1 + \partial_2\partial_2 h(z_2,z_1|\mathcal{B}^{k,i},\theta^{k,i}) \cdot \tilde{z}_1^{\hat{k},\hat{i}} \cdot \tilde{z}_1^{\tilde{k},\tilde{i}}$
end if
if $c^{\hat{k},\hat{i}}_{k+1,d-m+1} == 1 \;\&\; c^{\tilde{k},\tilde{i}}_{k+1,d-m+1} == 1$ **then**
Set $s2^{values}_{k,i} = s2^{values}_{k,i} + \frac{\partial_2\partial_2 c(z_1,z_2|\mathcal{B}^{k,i},\theta^{k,i})}{exp(v^{values}_{k,i})} \cdot \tilde{z}_2^{\hat{k},\hat{i}} \cdot \tilde{z}_2^{\tilde{k},\tilde{i}} + \frac{\partial_2 c(z_1,z_2|\theta^{k,i})}{exp(v^{values}_{k,i})} \cdot \bar{z}_2$
Set $s2^{direct}_{k-1,i} = s2^{direct}_{k-1,i} + \partial_2 h(z_1,z_2|\mathcal{B}^{k,i},\theta^{k,i}) \cdot \bar{z}_2 + \partial_2\partial_2 h(z_1,z_2|\mathcal{B}^{k,i},\theta^{k,i}) \cdot \tilde{z}_2^{\hat{k},\hat{i}} \cdot \tilde{z}_2^{\tilde{k},\tilde{i}}$
Set $s2^{indirect}_{k-1,i} = s2^{indirect}_{k-1,i} + \partial_1 h(z_2,z_1|\mathcal{B}^{k,i},\theta^{k,i}) \cdot \bar{z}_2 + \partial_1\partial_1 h(z_2,z_1|\mathcal{B}^{k,i},\theta^{k,i}) \cdot \tilde{z}_2^{\hat{k},\hat{i}} \cdot \tilde{z}_2^{\tilde{k},\tilde{i}}$
end if
if $c^{\hat{k},\hat{i}}_{k+1,i} == 1 \;\&\; c^{\tilde{k},\tilde{i}}_{k+1,d-m+1} == 1$ **then**
Set $s2^{values}_{k,i} = s2^{values}_{k,i} + \frac{\partial_1\partial_2 c(z_1,z_2|\mathcal{B}^{k,i},\theta^{k,i})}{exp(v^{values}_{k,i})} \cdot \tilde{z}_1^{\hat{k},\hat{i}} \cdot \tilde{z}_2^{\tilde{k},\tilde{i}}$
Set $s2^{direct}_{k-1,i} = s2^{direct}_{k-1,i} + \partial_1\partial_2 h(z_1,z_2|\mathcal{B}^{k,i},\theta^{k,i}) \cdot \tilde{z}_1^{\hat{k},\hat{i}} \cdot \tilde{z}_2^{\tilde{k},\tilde{i}}$
Set $s2^{indirect}_{k-1,i} = s2^{direct}_{k-1,i} + \partial_1\partial_2 h(z_2,z_1|\mathcal{B}^{k,i},\theta^{k,i}) \cdot \tilde{z}_1^{\hat{k},\hat{i}} \cdot \tilde{z}_2^{\tilde{k},\tilde{i}}$
end if
if $c^{\hat{k},\hat{i}}_{k+1,d-m+1} == 1 \;\&\; c^{\tilde{k},\tilde{i}}_{k+1,i} == 1$ **then**
Set $s2^{values}_{k,i} = s2^{values}_{k,i} + \frac{\partial_2\partial_1 c(z_1,z_2|\mathcal{B}^{k,i},\theta^{k,i})}{exp(v^{values}_{k,i})} \cdot \tilde{z}_2^{\hat{k},\hat{i}} \cdot \tilde{z}_1^{\tilde{k},\tilde{i}}$
Set $s2^{direct}_{k-1,i} = s2^{direct}_{k-1,i} + \partial_1\partial_2 h(z_1,z_2|\mathcal{B}^{k,i},\theta^{k,i}) \cdot \tilde{z}_2^{\hat{k},\hat{i}} \cdot \tilde{z}_1^{\tilde{k},\tilde{i}}$
Set $s2^{indirect}_{k-1,i} = s2^{direct}_{k-1,i} + \partial_1\partial_2 h(z_2,z_1|\mathcal{B}^{k,i},\theta^{k,i}) \cdot \tilde{z}_2^{\hat{k},\hat{i}} \cdot \tilde{z}_1^{\tilde{k},\tilde{i}}$
end if
end for
end for
return $\sum_{k,i=1,\dots,d} s2^{values}_{k,i}$

A.2 Calculation of the covariance matrix in the Gaussian case

While analytical results on the Fisher information for the multivariate normal distribution are well known (Mardia and Marshall 1984) we will now illustrate how the matrices $\boldsymbol{\mathcal{K}}_\theta$ and $\boldsymbol{\mathcal{J}}_\theta$ (Equation (3.4) and (3.5)) can be calculated. We consider a 3-dimensional Gaussian distribution

$$\begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} \sim N_3(\mathbf{0}, \mathbf{\Sigma}), \quad \mathbf{\Sigma} = \begin{pmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{12} & 1 & \rho_{23} \\ \rho_{13} & \rho_{23} & 1 \end{pmatrix},$$

with density f_{123} and corresponding copula c_{123}. Exampli gratia, we show the computation for the entry $(2,1)$ in $\boldsymbol{\mathcal{K}}_\theta$ in detail. The other entries in $\boldsymbol{\mathcal{K}}_\theta$ and $\boldsymbol{\mathcal{J}}_\theta$ are obtained similarly.

The first step is to calculate the following integral:

$$\int_{[0,1]^3} \left(\frac{\partial}{\partial\rho_{12}} \ln(c_{12}(u_1, u_2|\rho_{12}))\right) \left(\frac{\partial}{\partial\rho_{23}} \ln(c_{23}(u_2, u_3|\rho_{23}))\right) c_{123}(u_1, u_2, u_3) du_1 du_2 du_3, \tag{A.9}$$

where c_{12} and c_{23} are the corresponding copulas to the bivariate marginal distributions f_{12} and f_{23}, respectively. Since the integral is independent of the univariate marginal distributions, we can compute it using standard normal margins (see Smith 2007):

$$\int_{\mathbb{R}^3} \left(\frac{\partial}{\partial\rho_{12}} \ln(f_{12}(x_1, x_2|\rho_{12}))\right) \left(\frac{\partial}{\partial\rho_{23}} \ln(f_{23}(x_2, x_3|\rho_{23}))\right) f_{123}(x_1, x_2, x_3) dx_1 dx_2 dx_3, \tag{A.10}$$

where f_{12} and f_{23} are the according bivariate normal distributions. The 3-dimensional and bivariate normal densities in (A.9) and (A.10) can be expressed as

$$\begin{aligned} f_{123}(x_1, x_2, x_3) &= \frac{\sqrt{2}}{\pi^{3/2}\sqrt{2\,\rho_{13}\rho_{12}\rho_{23} - {\rho_{13}}^2 - {\rho_{12}}^2 + 1 - {\rho_{23}}^2}} \\ &\cdot \exp\left\{-\frac{1}{2}\frac{-{x_1}^2 - {x_2}^2 - {x_3}^2 + {x_1}^2{\rho_{23}}^2 + {x_2}^2{\rho_{13}}^2 + {x_3}^2{\rho_{12}}^2 + 2\,x_1x_2\rho_{12} + 2\,x_1x_3\rho_{13}}{-2\,\rho_{13}\rho_{12}\rho_{23} + {\rho_{13}}^2 + {\rho_{12}}^2 - 1 + {\rho_{23}}^2}\right\} \\ &\cdot \exp\left\{\frac{+2\,x_2x_3\rho_{23} - 2\,x_1x_2\rho_{13}\rho_{23} - 2\,x_1x_3\rho_{12}\rho_{23} - 2\,x_2x_3\rho_{13}\rho_{12}}{-2\,\rho_{13}\rho_{12}\rho_{23} + {\rho_{13}}^2 + {\rho_{12}}^2 - 1 + {\rho_{23}}^2}\right\}, \end{aligned} \tag{A.11}$$

and

$$f_{12}(x_1, x_2) = \frac{1}{2\pi}\frac{1}{\sqrt{1 - {\rho_{12}}^2}} \exp\left\{-\frac{1}{2}\frac{-{x_1}^2 + 2\,x_1x_2\rho_{12} - {x_2}^2}{(-1 + \rho_{12})(\rho_{12} + 1)}\right\}. \tag{A.12}$$

Further, the derivatives needed in Equation (A.10) are

$$\frac{\partial}{\partial\rho_{12}} \ln(f_{12}(x_1, x_2|\rho_{12})) = -\frac{{\rho_{12}}^3 - x_1x_2{\rho_{12}}^2 + {x_2}^2\rho_{12} - \rho_{12} + {x_1}^2\rho_{12} - x_1x_2}{(-1 + \rho_{12})^2(\rho_{12} + 1)^2}, \tag{A.13}$$

and

$$\frac{\partial}{\partial\rho_{23}} \ln(f_{23}(x_2, x_3|\rho_{23})) = -\frac{{\rho_{23}}^3 - x_2x_3{\rho_{23}}^2 + {x_3}^2\rho_{23} - \rho_{23} + {x_2}^2\rho_{23} - x_2x_3}{(-1 + \rho_{23})^2(\rho_{23} + 1)^2}. \tag{A.14}$$

Using (A.11), (A.13) and (A.14) in (A.10) we get

$$\begin{aligned} (A.10) = \int_{\mathbb{R}^3} &\frac{{\rho_{12}}^3 - x_1x_2{\rho_{12}}^2 + {x_2}^2\rho_{12} - \rho_{12} + {x_1}^2\rho_{12} - x_1x_2}{(-1 + \rho_{12})^2(\rho_{12} + 1)^2} \\ &\cdot \frac{{\rho_{23}}^3 - x_2x_3{\rho_{23}}^2 + {x_3}^2\rho_{23} - \rho_{23} + {x_2}^2\rho_{23} - x_2x_3}{(-1 + \rho_{23})^2(\rho_{23} + 1)^2} \cdot f_{123}(x_1, x_2, x_3) dx_1 dx_2 dx_3. \end{aligned} \tag{A.15}$$

The integral (A.15) can be solved using well known results on product moments of multivariate normal distributions (see Isserlis 1918).

$$\begin{aligned}(A.15) = & \frac{\rho_{23}{\rho_{12}}^3 + {\rho_{12}}^3{\rho_{23}}^3 - 3\,{\rho_{23}}^2{\rho_{12}}^2\rho_{13} - {\rho_{12}}^2\rho_{13} + 2\,\rho_{23}{\rho_{13}}^2\rho_{12}}{(\rho_{23}+1)^2(-1+\rho_{23})^2(-1+{\rho_{12}}^2)^2} \\ & + \frac{-\rho_{12}\rho_{23} + {\rho_{23}}^3\rho_{12} + \rho_{13} - {\rho_{23}}^2\rho_{13}}{(\rho_{23}+1)^2(-1+\rho_{23})^2(-1+{\rho_{12}}^2)^2}\end{aligned} \quad (A.16)$$

Since $(\rho_{23}+1)^2(-1+\rho_{23})^2 = (1-\rho_{23}^2)^2$ we can simplify Equation (A.16) to

$$\begin{aligned}(A.9) = (A.16) &= \frac{(\rho_{13}-\rho_{12}\rho_{23})(1-\rho_{12}^2)(1-\rho_{23}^2) + 2\rho_{12}\rho_{23}(\rho_{13}-\rho_{12}\rho_{23})^2}{(1-\rho_{12}^2)^2(1-\rho_{23}^2)^2} \\ &= \frac{\rho_{13}-\rho_{12}\rho_{23}}{(1-\rho_{12}^2)(1-\rho_{23}^2)} + 2\rho_{12}\rho_{23}\frac{(\rho_{13}-\rho_{12}\rho_{23})^2}{(1-\rho_{12}^2)^2(1-\rho_{23}^2)^2} \\ &= \frac{k_{12}}{(1-\rho_{12}^2)(1-\rho_{23}^2)},\end{aligned}$$

with

$$k_{12} = (\rho_{13}-\rho_{12}\rho_{23})\left(1 + 2\rho_{12}\rho_{23}\frac{\rho_{13}-\rho_{12}\rho_{23}}{(1-\rho_{12}^2)(1-\rho_{23}^2)}\right).$$

For the computation of terms corresponding to parameter $\rho_{13|2}$, note that

$$\rho_{13|2} = \frac{\rho_{13}-\rho_{12}\rho_{23}}{\sqrt{(1-\rho_{12}^2)(1-\rho_{23}^2)}},$$

and

$$\rho_{13} = \rho_{13|2}\sqrt{(1-\rho_{12}^2)(1-\rho_{23}^2)} + \rho_{12}\rho_{23},$$

which means that (A.11), (A.13) and (A.14) can easily be re-parametrized.
The final matrices are

$$\boldsymbol{\mathcal{K}}_\theta = \begin{pmatrix} \frac{1+\rho_{12}^2}{(1-\rho_{12}^2)^2} & \frac{k_{12}}{(1-\rho_{12}^2)(1-\rho_{23}^2)} & 0 \\ \frac{k_{12}}{(1-\rho_{12}^2)(1-\rho_{23}^2)} & \frac{1+\rho_{23}^2}{(1-\rho_{23}^2)^2} & 0 \\ 0 & 0 & \frac{1+\rho_{13|2}}{(\rho_{13|2}^2-1)^2} \end{pmatrix},$$

$$\boldsymbol{\mathcal{J}}_\theta = \begin{pmatrix} \frac{1+\rho_{12}^2}{(1-\rho_{12}^2)^2} & 0 & 0 \\ 0 & \frac{1+\rho_{23}^2}{(1-\rho_{23}^2)^2} & 0 \\ \frac{\rho_{13|2}\rho_{12}}{(\rho_{12}^2-1)(\rho_{13|2}^2-1)} & \frac{\rho_{13|2}\rho_{23}}{(\rho_{23}^2-1)(\rho_{13|2}^2-1)} & \frac{1+\rho_{13|2}}{(\rho_{13|2}^2-1)^2} \end{pmatrix}.$$

A.3 Selected model for the exchange rate data

To obtain marginally uniformly distributed copula data on $[0,1]^8$, we conduct a pre-analysis as described in (Schepsmeier 2010, Chapter 5). AR(1)-GARCH(1,1) models (Table A.1) are selected for the marginal time series, and the resulting standardized residuals are transformed using the non-parametric rank transformation (see Genest et al. 1995). We could also employ the probability integral transformation based on the parametric error distributions (IFM, Joe and Xu 1996) but since we are only interested in dependence properties here, we choose the non-parametric alternative which is more robust with respect to misspecification of marginal error distributions.

	μ	a_1	ω	α_1	β_1	skew	shape
EUR	0.0011	0.9982	0.0000	0.0422	0.9554	0.9844	9.2961
GBP	0.0014	0.9972	0.0000	0.0465	0.9511	1.0533	8.1926
CAD	0.0049	0.9954	0.0000	0.0576	0.9363	0.9590	9.5090
AUD	0.0025	0.9978	0.0000	0.0755	0.9170	1.2211	7.4130
BRL	0.0023	0.9982	0.0000	0.1860	0.8137	1.1349	10.0000
JPY	0.1660	0.9986	0.0068	0.0367	0.9514	0.8823	7.1649
CHF	0.0042	0.9963	0.0000	0.0342	0.9619	0.8839	9.2343
INR	0.0305	0.9994	0.0011	0.2588	0.8169	1.0625	2.9743

Table A.1: Parameters of the AR(1)-GARCH(1,1) models. μ and a_1 correspond to the AR-process, while ω, α_1, and β_1 define the GARCH(1,1)-process. skew and shape describe the Skew-t error distribution.

The R-vine describing the exchange rate data set is specified by the structure matrix M, the copula family matrix $\mathcal{B}$ and the estimated copula parameter matrix $\hat{\boldsymbol{\theta}}^{MLE}$. For simplicity, we use the following abbreviations: 1=AUD (Australian dollar), 2=JPY (Japanese yen), 3=BRL (Brazilian real), 4=CAD (Canadian dollar), 5=EUR (Euro), 6=CHF (Swiss frank), 7=INR (Indian rupee) and 8=GBP (British pound).

The pair-copula families in the application were chosen from the elliptical copulas Gauss and Student's t copula, the Archimedean Clayton, Gumbel, Frank and Joe copula, and their rotated versions. The selection is done using to AIC/BIC.

$$M = \begin{pmatrix} 8 & & & & & & & \\ 7 & 7 & & & & & & \\ 2 & 2 & 6 & & & & & \\ 3 & 3 & 2 & 5 & & & & \\ 6 & 4 & 3 & 2 & 4 & & & \\ 4 & 1 & 4 & 3 & 2 & 3 & & \\ 1 & 5 & 1 & 4 & 3 & 2 & 2 & \\ 5 & 6 & 5 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \tag{A.17}$$

$$\hat{\boldsymbol{\theta}}^{MLE} = \begin{pmatrix} & & & & & & 11.85 & 8.96 \\ & & & & & & & 7.74 \\ & & & & & & & 3.76 \\ & -0.70 & -0.09 & & & & & 9.97 \\ & -0.88 & -1.44 & & & & & 8.41 \\ 0.07 & -1.10 & -0.73 & & & & & 7.46 \\ 0.26 & -1.23 & -1.17 & 1.13 & 1.08 & 0.63 & & \\ 0.72 & 0.55 & 0.88 & 0.63 & 0.54 & 0.48 & 0.3 & \end{pmatrix} \tag{A.18}$$

$$\mathcal{B} = \begin{pmatrix} & & & & & & \\ \text{Indep.} & & & & & & \\ \text{Indep.} & \text{Indep.} & & & & & \\ \text{Indep.} & \text{Frank} & \text{Gauss} & & & & \\ \text{Indep.} & \text{Frank} & \text{Frank} & \text{Indep.} & & & \\ \text{Gauss} & \text{r. Joe} & \text{Frank} & \text{Indep.} & \text{Indep.} & & \\ \text{Student's t} & \text{r. Gumbel} & \text{r. Gumbel} & \text{Gumbel} & \text{Frank} & \text{Frank} & \\ \text{Student's t} & \text{Student's t} & \text{Student's t} & \text{Student's t} & \text{Student's t} & \text{Student's t} & \text{Gauss} \end{pmatrix} \tag{A.19}$$

The standard errors corresponding to the parameters in (A.18) are

$$\boldsymbol{SE}_n^{MLE} = \begin{pmatrix} & & & & & & 5.00 & 2.54 \\ & & & & & & & 1.95 \\ & & & & & & & 0.59 \\ & 0.19 & 0.03 & & & & & 3.04 \\ & 0.19 & 0.20 & & & & & 2.40 \\ 0.03 & 0.03 & 0.19 & & & & & 2.01 \\ 0.03 & 0.03 & 0.03 & 0.02 & 0.19 & 0.19 & & \\ 0.01 & 0.02 & 0.01 & 0.02 & 0.02 & 0.02 & 0.03 & \end{pmatrix}. \tag{A.20}$$

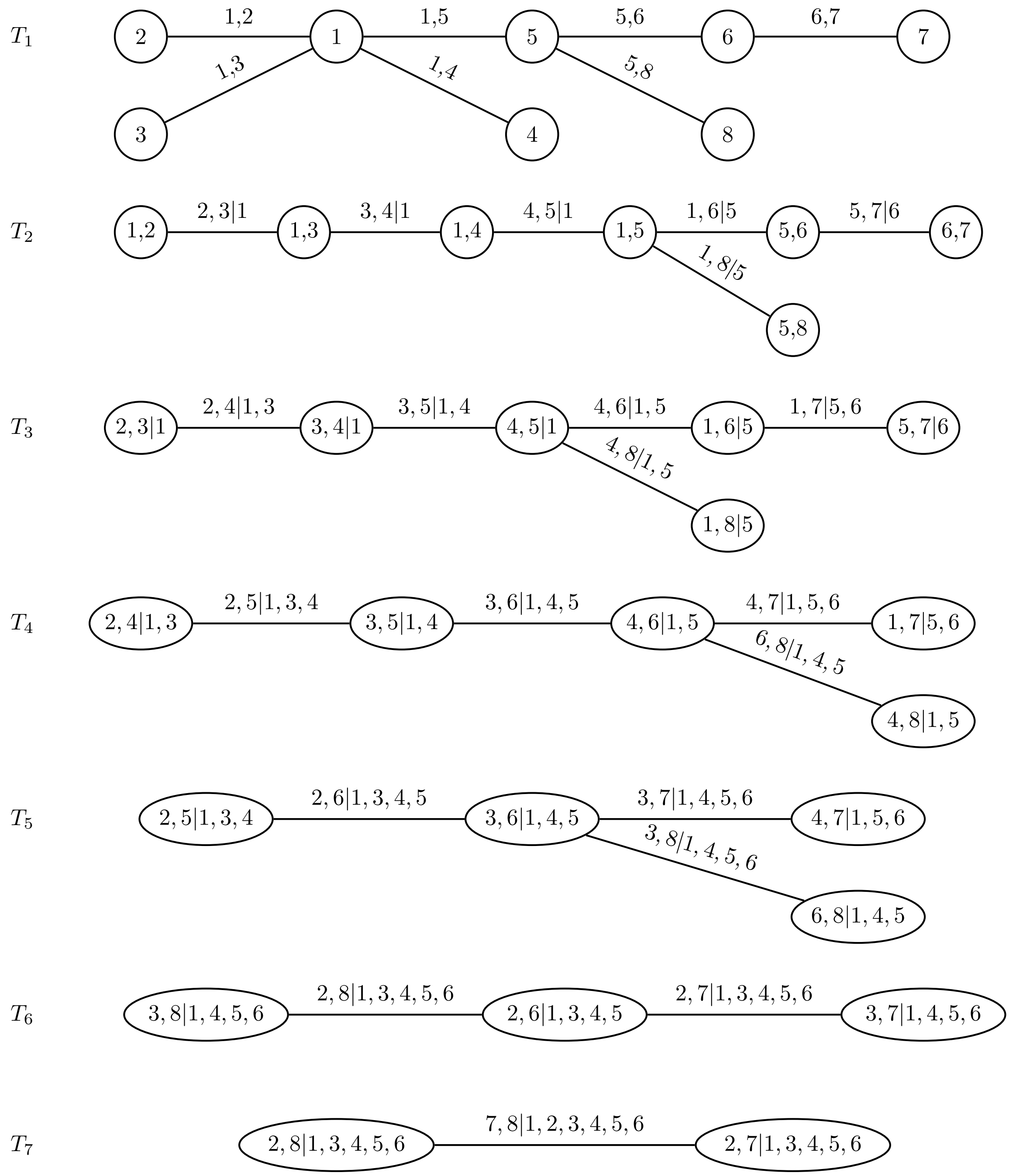


Figure A.1: The R-vine tree sequence of an 8-dimensional R-vine fitted to our exchange rate data set used in Section 3.4 and Section 4.5.2 applying the MST approach.

Appendix B

Technical details for the goodness-of-fit tests

B.1 Rosenblatt's transform

The multivariate probability integral transformation (PIT) of Rosenblatt (1952) transforms the copula data $\boldsymbol{u} = (u_1, \ldots, u_d)$ with a given multivariate copula C into independent data in $[0, 1]^d$, where d is the dimension of the data set.

Definition B.1 (Rosenblatt's transform)
Let $\boldsymbol{u} = (u_1, \ldots, u_d)$ denote copula data of dimension d. Further let C be the joint cdf of $\boldsymbol{u}$. Then Rosenblatt's transformation of $\boldsymbol{u}$, denoted as $\boldsymbol{y} = (y_1, \ldots, y_d)$, is defined as

$$y_1 := u_1, \qquad y_2 := C(u_2|u_1), \quad \ldots \quad y_d := C(u_d|u_1, \ldots, u_{d-1}),$$

where $C(u_k|u_1, \ldots, u_{k-1})$ is the conditional copula of U_k given $U_1 = u_1, \ldots, U_{k-1} = u_{k-1}, k = 2, \ldots, d$.

The data vector $\boldsymbol{y} = (y_1, \ldots, y_d)$ is now i.i.d. with $y_i \sim U[0, 1]$. In the context of vine copulas the multivariate PIT is given for the special classes of C- and D-vine in Aas et al. (2009, Algorithm 5 and 6). It is a straight forward application of the Rosenblatt transformation of Definition B.1 to the recursive structure of a C- or D-vine copula. Similar, an algorithm for the R-vine can be stated, see Algorithm B.1.1. Here we make use of a similar structured algorithm of Dißmann et al. (2013) for calculating the log-likelihood of an R-vine copula given in Algorithm 2.2.1.

Algorithm B.1.1 now calculates the PIT of an R-vine copula model. The vector $\boldsymbol{y} = (y_1, \ldots, y_d)$ stores at the end the transformed PIT variables.

Algorithm B.1.1 Probability integral transform (PIT) of an R-vine

Require: d-dimensional R-vine specification in matrix form, i.e., M, $\mathcal{B}$, $\boldsymbol{\theta}$, where $m_{k,k} = d - k + 1, k = 1, \ldots, d$, and a set of observations $(u_1, \ldots, u_d)$.

1: Let $V^{\text{direct}} = (v_{k,i}^{\text{direct}}|i = 1, \ldots, d; k = i, \ldots, d)$.

2: Let $V^{\text{indirect}} = (v_{k,i}^{\text{indirect}}|i = 1, \ldots, d; k = i, \ldots, d)$.

3: Set $(v_{d,1}^{\text{direct}}, v_{d,2}^{\text{direct}}, \ldots, v_{d,d}^{\text{direct}}) = (u_d, u_{d-1}, \ldots u_1)$.

Let $\tilde{M} = (\tilde{m}_{k,i}|i = 1, \dots, d; k = i, \dots, d)$ where $\tilde{m}_{k,i} = \max\{m_{k,i}, \dots, m_{d,i}\}$ for all $i = 1, \dots, d$ and $k = i, \dots, d$.
Set $y_1 = u_1$
for $i = d-1, \dots, 1$ **do** {Iteration over the columns of M}
  for $k = d, \dots, i+1$ **do** {Iteration over the rows of M}
    Set $z_1 = v_{k,i}^{\text{direct}}$
    if $\tilde{m}_{k,i} = m_{k,i}$ **then**
      Set $z_2 = v_{k,(d-\tilde{m}_{k,i}+1)}^{\text{direct}}$.
    else
      Set $z_2 = v_{k,(d-\tilde{m}_{k,i}+1)}^{\text{indirect}}$.
    end if
    Set $v_{k-1,i}^{\text{direct}} = h(z_1, z_2|\mathcal{B}^{k,i}, \theta^{k,i})$ and $v_{k-1,i}^{\text{indirect}} = h(z_2, z_1|\mathcal{B}^{k,i}, \theta^{k,i})$.
    Set $y_{d-k+1} = v_{i-1,k}^{direct}$
  end for
end for
return $\boldsymbol{y} = (y_1, \dots, y_d)$

B.2 Cramér-von Mises, Kolmogorov-Smirnov and Anderson Darling goodness-of-fit test

B.2.1 Multivariate and univariate Cramér-von Mises and Kolmogorov-Smirnov test

Already in the third century of 1900 two model specification tests were developed by Cramér and von Mises, and by Kolmogorov and Smirnov. Both tests treat the hypothesis that n i.i.d. samples $\boldsymbol{y}_1, \dots, \boldsymbol{y}_n$ of the random vector $\boldsymbol{Y} = (Y_1, \dots, Y_d)$ follow a specified continuous distribution function F, i.e.

$$H_0 : \boldsymbol{Y} \sim F \quad \text{versus} \quad H_1 : \boldsymbol{Y} \not\sim F.$$

The general **multivariate Cramér-von Mises (mCvM)** test statistic for a d-dimensional random vector $\boldsymbol{Y}$ is defined as

$$\textbf{mCvM:} \quad \omega^2 = \int_{\mathbb{R}^d} \left[\hat{F}_n(\boldsymbol{y}) - F(\boldsymbol{y})\right]^2 dF(\boldsymbol{y}), \tag{B.1}$$

while the **multivariate Kolmogorov-Smirnov (mKS)** test statistic is

$$\textbf{mKS:} \quad D_n = \sup_{\boldsymbol{y}} |\hat{F}_n(\boldsymbol{y}) - F(\boldsymbol{y})|. \tag{B.2}$$

Here $\hat{F}_n(\boldsymbol{y}) = \frac{1}{n+1} \sum_{j=1}^{n} \mathbf{1}_{\{\boldsymbol{y}_j \leq \boldsymbol{y}\}}$ denotes the empirical distribution function corresponds to the i.i.d. sample $(\boldsymbol{y}_1, \dots, \boldsymbol{y}_n)$ of $\boldsymbol{Y}$.
The **univariate cases** for the random variable Y are then denoted by

$$\textbf{CvM:} \quad \omega^2 = \int_{-\infty}^{\infty} \left[\hat{F}_n(y) - F(y)\right]^2 dF(y) \quad \text{and}$$

$$\textbf{KS:} \quad D_n = \sup_{y} |\hat{F}_n(y) - F(y)|.$$

B.2.2 Univariate Anderson-Darling test

The Anderson and Darling (1954) test, is a statistical test of whether a given probability distribution fits a given set of data samples. It extends the Cramér-von Mises test statistics by adding more weight in the tails of the distribution in consideration. Although it has a general multivariate definition we introduce only the univariate case, since only the univariate case is needed in Section 4.3.1. Let Y be a random variable then the null hypothesis of the Anderson-Darling test is again $H_0 : Y \sim F(y)$ against the alternative $H_1 : Y \nsim F(y)$. The general **univariate Anderson-Darling (AD)** test statistic is defined as

$$W_n^2 = n \int_{-\infty}^{\infty} \left[\hat{F}_n(y) - F(y)\right]^2 \psi(F(y)) dF(y), \tag{B.3}$$

where $\psi(F(y))$ is a non-negative weighting function. With the weighting function $\psi(u) = \frac{1}{u(1-u)}$ Anderson and Darling (1954) put more weight in the tails since this function is large near $u = 0$ and $u = 1$. Setting the weight function to $\psi(u) = 1$ one gets as a special case the Cramér-von Mises test statistic. In the case of uniform margins (B.3) simplifies to

$$\textbf{AD:} \quad W_n^2 = n \int_0^1 \frac{\left[\hat{F}_n(y) - y\right]^2}{y(1-y)} dy, \qquad y \in [0, 1]. \tag{B.4}$$

Appendix C

Model specifications in the power studies

C.1 Model specification in power study I

For the vine copula density (see Equation (2.3)) often a short hand notation is used. For this the pair-copula arguments are omitted and denotes only the conditioned and conditioning set. Thus, for the R-vine considered in the first power study and given in Figure C.1 we can write

$$c_{12345} = c_{1,2} \cdot c_{1,3} \cdot c_{1,4} \cdot c_{4,5} \cdot c_{2,4;1} \cdot c_{1,5;4} \cdot c_{2,3;1,4} \cdot c_{3,5;1,4} \cdot c_{2,5;1,3,4}. \tag{C.1}$$

Similarly the considered C- and D-vine copula can be expressed as

$$c_{12345} = c_{1,2} \cdot c_{2,3} \cdot c_{2,4} \cdot c_{2,5} \cdot c_{1,3|2} \cdot c_{1,4|2} \cdot c_{1,5|2} \cdot c_{3,4|1,2} \cdot c_{4,5|1,2} \cdot c_{3,5|1,2,4} \tag{C.2}$$
$$c_{12345} = c_{1,2} \cdot c_{1,5} \cdot c_{4,5} \cdot c_{3,4} \cdot c_{2,5|1} \cdot c_{1,4|5} \cdot c_{3,5|4} \cdot c_{2,4|1,5} \cdot c_{1,3|4,5} \cdot c_{2,3|1,4,5} \tag{C.3}$$

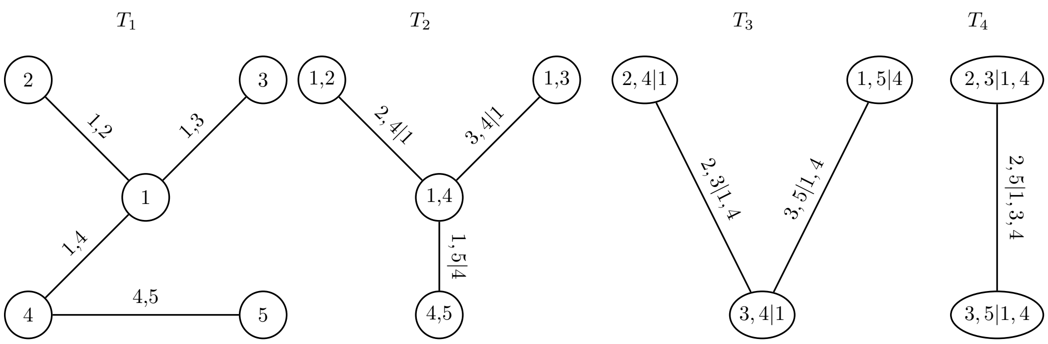


Figure C.1: Tree structure of the 5 dimensional R-vine copula used in the power studies.

Tree	$\mathcal{V}_R$	$\mathcal{B}_R(\mathcal{V}_R)$	τ
1	$c_{1,2}$	Gauss	0.71
	$c_{1,3}$	Gauss	0.33
	$c_{1,4}$	Clayton	0.71
	$c_{4,5}$	Gumbel	0.74
2	$c_{2,4\|1}$	Gumbel	0.38
	$c_{3,4\|1}$	Gumbel	0.47
	$c_{1,5\|4}$	Gumbel	0.33
3	$c_{2,3\|1,4}$	Clayton	0.35
	$c_{3,5\|1,4}$	Clayton	0.31
4	$c_{2,5\|1,3,4}$	Gauss	0.13

Table C.1: Copula families and Kendall's τ values of the investigated (mixed) R-vine copula model defined by (C.1).

C.2 Model specification in power study II

True model (M_1)			M_2^{MST}		M_2^{MCMC}	
$\mathcal{V}$	$\mathcal{B}(\mathcal{V})$	τ	$\hat{\mathcal{V}}$	$\hat{\mathcal{B}}(\hat{\mathcal{V}})$	$\hat{\mathcal{V}}$	$\hat{\mathcal{B}}(\hat{\mathcal{V}})$
$c_{1,2}$	Gauss	0.10	$c_{1,3}$	t_ν	$c_{1,2}$	Gauss
$c_{2,3}$	t_3	-0.15	$c_{1,5}$	Gauss	$c_{2,3}$	t_ν
$c_{3,4}$	t_3	-0.10	$c_{2,5}$	t_ν	$c_{2,4}$	Gumbel 90
$c_{3,5}$	t_3	0.15	$c_{4,5}$	Gumbel 270	$c_{3,5}$	t_ν
$c_{1,3\|2}$	N	0.70	$c_{1,2\|5}$	t_ν	$c_{1,3\|2}$	Gauss
$c_{2,4\|3}$	Gumbel 90	-0.60	$c_{1,4\|5}$	t_ν	$c_{3,4\|2}$	Gumbel
$c_{2,5\|3}$	Gumbel	0.85	$c_{3,5\|1}$	t_ν	$c_{2,5\|3}$	t_ν
$c_{1,4\|2,3}$	Gauss	0.45	$c_{2,3\|1,5}$	t_ν	$c_{1,4\|2,3}$	Gauss
$c_{1,5\|2,3}$	Gauss	-0.50	$c_{3,4\|1,5}$	t_ν	$c_{1,5\|2,3}$	Gauss
$c_{4,5\|1,2,3}$	Gauss	0.10	$c_{2,4\|1,3,5}$	Gauss	$c_{4,5\|1,2,3}$	Gauss

Table C.2: Copula families and Kendall's τ values of the investigated R-vine models in power study 2 ($t_\nu \hat{=}$ t-copula with ν degrees-of-freedom, Gumbel 90 $\hat{=}$ 90 degree rotated Gumbel copula, Gumbel 270 $\hat{=}$ 270 degree rotated Gumbel copula).

C.3 Model specification in power comparison study ($d = 8$)

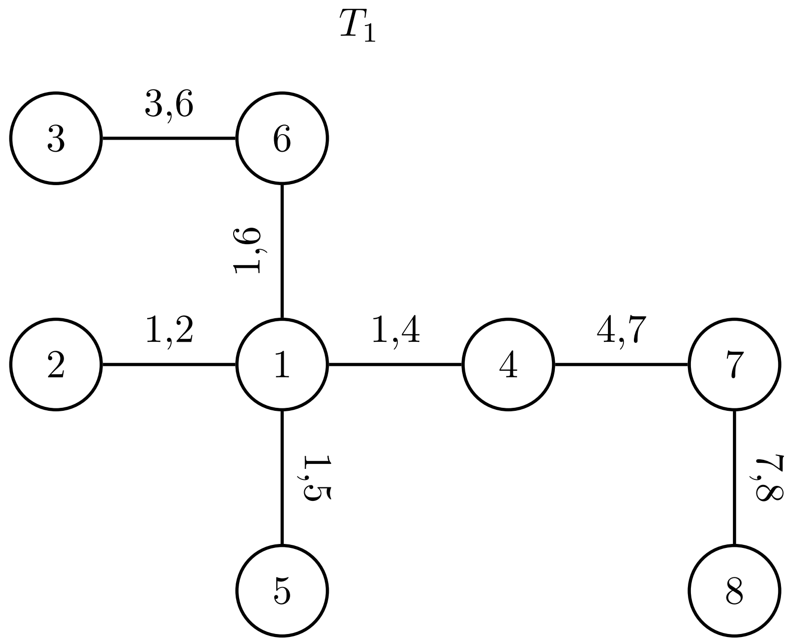


Figure C.2: First tree of the 8 dimensional R-vine copula used in the power study.

Tree	$\mathcal{V}_R^8$	$\mathcal{B}_R^8(\mathcal{V}_R^8)$	τ	Tree	$\mathcal{V}_R^8$	$\mathcal{B}_R^8(\mathcal{V}_R^8)$	τ
1	$c_{1,2}$	Joe	0.41	3	$c_{6,7\|1,4}$	Frank 0.03	7
	$c_{1,4}$	Gauss	0.59		$c_{1,8\|4,7}$	Gumbel	0.22
	$c_{1,5}$	Gauss	0.59		$c_{3,4\|1,6}$	Gauss	0.41
	$c_{1,6}$	Frank	0.23		$c_{2,3\|1,6}$	Gumbel	0.68
	$c_{3,6}$	Frank	0.19	4	$c_{6,8\|1,4,7}$	Clayton	0.17
	$c_{4,7}$	Clayton	0.44		$c_{5,7\|1,4,6}$	Gauss	0.09
	$c_{7,8}$	Gumbel	0.64		$c_{3,5\|1,4,6}$	Frank	0.21
2	$c_{2,6\|1}$	Clayton	0.58		$c_{2,4\|1,3,6}$	Gumbel	0.57
	$c_{1,3\|6}$	Gumbel	0.44	5	$c_{2,5\|1,3,4,6}$	Joe	0.25
	$c_{4,6\|1}$	Frank	0.11		$c_{3,7\|1,4,5,6}$	Gumbel	0.17
	$c_{4,5\|1}$	Clayton	0.53		$c_{5,8\|1,4,6,7}$	Frank	0.02
	$c_{1,7\|4}$	Clayton	0.29	6	$c_{2,7\|1,3,4,5,6}$	Gumbel	0.31
	$c_{4,8\|7}$	Gauss	0.53		$c_{3,8\|1,4,5,6,7}$	Clayton	0.20
3	$c_{5,6\|1,4}$	Gauss	0.19	7	$c_{2,8\|1,3,4,5,6,7}$	Frank	0.03

Table C.3: Copula families and Kendall's τ values of the investigated (mixed) R-vine copula model in the 8-dimensional case.

Bibliography

Aas, K., C. Czado, A. Frigessi, and H. Bakken (2009). Pair-copula construction of multiple dependence. *Insurance: Mathematics and Economics 44*, 182–198.

Abramowitz, M. and I. Stegun (1992). *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. New York: Dover Publishing.

Acar, E. F., C. Genest, and J. Nešlehová (2012). Beyond simplified pair-copula constructions. *Journal of Multivariate Analysis 110*, 74–90.

Almeida, C. and C. Czado (2012). Efficient Bayesian inference for stochastic time-varying copula models. *Computational Statistics and Data Analysis 56*(6), 1511–1527.

Anderson, T. W. and D. A. Darling (1954). A test of goodness of fit. *Journal of the American Statistical Association 49*(268), 765–769.

Bates, D. and M. Maechler (2012). *Matrix: Sparse and Dense Matrix Classes and Methods*. R package version 1.0-6.

Bauer, A., C. Czado, and T. Klein (2012). Pair-copula constructions for non-Gaussian DAG models. *Canadian Journal of Statistics 40*(1), 86–109.

Bedford, T. and R. Cooke (2001). Probability density decomposition for conditionally dependent random variables modeled by vines. *Ann. Math. Artif. Intell. 32*, 245–268.

Bedford, T. and R. Cooke (2002). Vines - a new graphical model for dependent random variables. *Annals of Statistics 30*, 1031–1068.

Bellosta, C. J. G. (2011). *ADGofTest: Anderson-Darling GoF test*. R package version 0.3.

Berg, D. (2009). Copula goodness-of-fit testing: An overview and power comparison. *The European Journal of Finance 15*, 1466–4364.

Berg, D. and K. Aas (2009). Models for construction of multivariate dependence: A comparison study. *The European Journal of Finance 15*(17-18), 639–659.

Berg, D. and H. Bakken (2007). A copula goodness-of-fit approach based on the conditional probability integral transformation. http://www.danielberg.no/publications/Btest.pdf.

Berger, J. O. and D. Sun (2008). Objective priors for the bivariate normal model. *The Annals of Statistics 36*(2), 963–82.

Berntsen, J., T. O. Espelid, and A. Genz (1991). An adaptive algorithm for the approximate calculation of multiple integrals. *ACM Trans. Math. Soft. 17*(4), 437–451.

Bickel, P. J. and K. A. Doksum (2007). *Mathematical Statistics: Basic Ideas and selected Topics* (second ed.), Volume 1. Pearson Prentice Hall, Upper Saddle River.

Boik, R. J. and J. F. Robinson-Cox (1998). Derivatives of the incomplete beta function. *Journal of Statistical Software 3*(1).

Brechmann, E., C. Czado, and K. Aas (2012). Truncated regular vines in high dimensions with applications to financial data. *Canadian Journal of Statistics 40*(1), 68–85.

Brechmann, E. C. and C. Czado (2013). Risk management with high-dimensional vine copulas: An analysis of the Euro Stoxx 50. *Statistics & Risk Modeling*, forthcoming.

Brechmann, E. C. and U. Schepsmeier (2013). Dependence modeling with C- and D-vine copulas: The R-package CDVine. *Journal of Statistical Software 52*(3), 1–27.

Breymann, W., A. Dias, and P. Embrechts (2003). Dependence structures for multivariate high-frequency data in finance. *Quantitative Finance 3*(1), 1–14.

C code by Steven G. Johnson and R by Balasubramanian Narasimhan (2011). *cubature: Adaptive multivariate integration over hypercubes.* R package version 1.1-1.

Chan, N.-H., J. Chen, X. Chen, Y. Fan, and L. Peng (2009). Statistical inference for multivariate residual copula of GARCH models. *Statist. Sinica 19*, 53–70.

Chesher, A. and R. Spady (1991). Asymptotic expansions of the information matrix test statistic. *Econometrica 59*(3), 787–815.

Clarke, K. (2007). A Simple Distribution-Free Test for Nonnested Model Selection. *Political Analysis 15*, 347–363.

Csardi, G. (2010). *igraph: Network analysis and visualization.* R package version 0.5.5-1.

Csardi, G. and T. Nepusz (2006). The igraph software package for complex network research. *InterJournal Complex Systems*, 1695.

Czado, C. (2010). Pair-Copula Constructions of Multivariate Copulas. In Jaworski, P. and Durante, F. and Härdle, W.K. and Rychlik, T (Ed.), *Copula Theory and Its Applications, Lecture Notes in Statistics*, Volume 198, Berlin Heidelberg, pp. 93–109. Springer-Verlag.

Czado, C., U. Schepsmeier, and A. Min (2012). Maximum likelihood estimation of mixed C-vines with application to exchange rates. *Statistical Modelling 12*(3), 229–255.

David, H. (1981). *Order statistics.* Wiley series in probability and mathematical statistics. Applied probability and statistics. Wiley.

Davidson, R. and J. MacKinnon (1998). Graphical methods for investigating the size and power of hypothesis tests. *The Manchester School 66*(1), 1–26.

Deheuvels, P. (1979). La fonction de dépendance empirique et ses propriétés. un test non paramétrique d'indépendance. *Acad. Roy. Belg. Bull. Cl. Sci.(5) 65*(6), 274–292.

Deheuvels, P. (1984). The characterization of distributions by order statistics and record values: A unified approach. *Journal of Applied Probability 21*(2), 326–334.

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

Dißmann, J. (2010). Statistical inference for regular vines and application. Diploma Thesis, Faculty of Mathematics, Technische Universität München, Germany.

Dißmann, J., E. Brechmann, C. Czado, and D. Kurowicka (2013). Selecting and estimating regular vine copulae and application to financial returns. *Computational Statistics and Data Analysis 59*(1), 52 – 69.

Dobrić, J. and F. Schmid (2005). Testing goodness of fit for parametric families of copula - application to financial data. *Communications in Statistic - Simulation and Computation® 34*(4), 1053–1068.

Dobrić, J. and F. Schmid (2007). A goodness of fit test for copulas based on rosenblatt's transformation. *Computational Statistics & Data Analysis 51*(9), 4633 – 4642.

Donaldson, J. R. and R. B. Schnabel (1987). Computational Experience with Confidence Regions and Confidence Intervals for Nonlinear Least Squares. *Technometrics 29*(1), 67–82.

Efron, B. and D. V. Hinkley (1978). Assessing the accuracy of the maximum likelihood estimator: Observed versus expected fisher information. *Biometrika 65*(3), 457–482.

Embrechts, P., F. Lindskog, and A. McNail (2003). Modelling dependence with copulas and applications to risk management. In S. Rachev (Ed.), *Handbook of heavy tailed distributions in finance*, pp. 329–384. Amsterdam: Elsevier/North-Holland.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters 27*(8), 861–874.

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

Fermanian, J.-D. (2012). An overview of the goodness-of-fit test problem for copulas. *ArXiv e-prints*.

Fréchet, M. (1951). Sur les tableaux de corrélation dont les marges sont données. *Ann. Univ. Lyon Sect. A 9*, 53–77.

Galambos, J. (1987). *The asymptotic theory of extreme order statistics*. Krieger Malabar.

Genest, C. and A. Favre (2007). Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of Hydrologic Engineering 12*, 347–368.

Genest, C., K. Ghoudi, and L.-P. Rivest (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika 82*, 543–552.

Genest, C., W. Huang, and J.-M. Dufour (2012). A regularized goodness-of-fit test for copulas. *Journal de la Société Française de Statistique & revue de statistique appliquée 154*(1), 64–77.

Genest, C., J.-F. Quessy, and B. Rémillard (2006). Goodness-of-fit Procedures for Copula Model Based on the Probability Integral Transformation. *Scandinavian Journal of Statistics 33*, 337–366.

Genest, C. and B. Rémillard (2008). Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models. *Annales de l'Institut Henri Poincaré - Probabilités et Statistiques 44*(6), 1096–1127.

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

Genz, A. et al. (2011). *mvtnorm: Multivariate Normal and t Distributions*. R package version 0.9-96.

Genz, A. and F. Bretz (2009). *Computation of Multivariate Normal and t Probabilities*. Lecture notes in statistics. Springer.

Genz, A. C. and A. A. Malik (1980). An adaptive algorithm for numeric integration over an n-dimensional rectangular region. *J. Comput. Appl. Math. 6*(4), 295–302.

Gräler, B. and E. Pebesma (2011). The pair-copula construction for spatial data: a new approach to model spatial dependency. *Procedia Environmental Sciences 7*, 206–211.

Gruber, L. and C. Czado (2012). Bayesian Model Selection for Regular Vine Copulas. Preprint.

Haff, I. H. and J. Segers (2012). Nonparametric estimation of pair-copula constructions with the empirical pair-copula. *arXiv preprint arXiv:1201.5133*.

Hall, A. (1989). On the calculation of the information matrix test in the normal linear regression model. *Economics Letters 29*(1), 31–35.

Heinen, A. and A. Valdesogo (2009). *Asymmetric CAPM dependence for large dimensions: the canonical vine autoregressive model*. CORE.

Higham, N. J. (2002). Computing the nearest correlation matrix - a problem from finance. *IMA Journal of Numerical Analysis 22*, 329–343.

Hobæk Haff, I. (2013). Parameter estimation for pair-copula constructions. *Bernoulli 19*(2), 462–491.

Hofert, M., I. Kojadinovic, M. Maechler, and J. Yan (2013). *copula: Multivariate Dependence with Copulas*. R package version 0.999-7.

Huang, W. and A. Prokhorov (2013). A goodness-of-fit test for copulas. *Economic Reviews to appear*.

Isserlis, L. (1918). On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables. *Biometrika 12*(1/2), 134–139.

Joe, H. (1996). Families of m-variate distributions with given margins and m(m-1)/2 bivariate dependence parameters. In L. Rüschendorf and B. Schweizer and M. D. Taylor (Ed.), *Distributions with Fixed Marginals and Related Topics*, Volume 28, Hayward, CA, pp. 120–141. Inst. Math. Statist.

Joe, H. (1997). *Multivariate Models and Dependence Concepts.* Chapman und Hall, London.

Joe, H., H. Li, and A. K. Nikoloulopoulos (2010). Tail dependence functions and vine copulas. *Journal of Multivariate Analysis 101*(1), 252–270.

Joe, H. and J. J. Xu (1996). The estimation method of inference functions of margins for multivariate models. *Technical report 166, Department of Statistics, University of British Columbia.*

Kauermann, G. and C. Schellhase (2013). Flexible pair-copula estimation in d-vines using bivariate penalized splines. *Statistics and Computing*, 1–20.

Kimeldorf, G. and A. Sampson (1975). Uniform Representations of Bivariate Distributions. *Communications in Statistics-simulation and Computation 4*, 617–627.

Kojadinovic, I. and J. Yan (2010). Modeling multivariate distributions with continuous margins using the copula R package. *Journal of Statistical Software 34*(9), 1–20.

Kullback, S. and R. A. Leibler (1951). On Information and Sufficiency. *The Annals of Mathematical Statistics 22*(1), 79–86.

Kurowicka, D. and R. Cooke (2006). *Uncertainty analysis with high dimensional dependence modeling.* Wiley Series in Probability and Statistics. Chichester: John Wiley & Sons Ltd.

Lehmann, E. L. and G. Casella (1998). *Theory of Point Estimation* (second ed.). Springer, New York.

Mai, J.-F. and M. Scherer (2012). *Simulating Copulas: Stochastic Models, Sampling Algorithms, and Applications.* Singapore: World Scientific Publishing Co.

Mardia, K. V. and R. J. Marshall (1984). Maximum likelihood estimation of models for residual covariance in spatial regression. *Biometrika 71*(1), 135–46.

McCullough, B. D. (1998). Assessing the reliability of statistical software: Part I. *The American Statistician 52*(4), 358–366.

McCullough, B. D. (1999). Assessing the reliability of statistical software: Part II. *The American Statistician 53*(2), 149–159.

McNeil, A. J. and J. Nešlehová (2009). Multivariate archimedean copulas, d-monotone functions and l1-norm symmetric distributions. *The Annals of Statistics 37*, 3059–3097.

Min, A. and C. Czado (2010). Bayesian inference for multivariate copulas using pair-copula constructions. *Journal of Financial Econometrics 8*(4), 511–546.

Min, A. and C. Czado (2012). SCOMDY models based on pair-copula constructions with application to exchange rates. *Computational Statistics and Data Analysis online first: http://dx.doi.org/10.1016/j.csda.2012.08.003.*

Morales-Nápoles, O. (2010). Counting Vines. In H. Kurowicka, D. Joe (Ed.), *Dependence Modeling-Handbook on Vine Copulas*, Singapore, pp. 189–218. World Scientific Publishing.

Morales-Nápoles, O., R. M. Cooke, and D. Kurowicka (2010). About the number of vines and regular vines on n nodes. *preprint, personal communication.*

Nadarajah, S. (2006). Fisher information for the elliptically symmetric pearson distributions. *Applied Mathematics and Computation 178*(2), 195 – 206.

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

Oakes, D. (1982). A model for association in bivariate survival data. *Journal of the Royal Statistical Society, Series B 44*(3), 414–422.

Oakes, D. and A. K. Manatunga (1992). Fisher information for a bivariate extreme value distribution. *Biometrika 79*(4), 827–832.

Panagiotelis, A., C. Czado, and H. Joe (2012). Pair copula constructions for multivariate discrete data. *Journal of the American Statistical Association, 107*(499), 1063–1072.

Patton, A. J. (2006). Estimation of multivariate models for time series of possibly different lengths. *Journal of Applied Econometrics 21*(2), 147–173.

Patton, A. J. (2012). Copula methods for forecasting multivariate time series. *Handbook of Economic Forecasting 2.*

Presnell, B. and D. D. Boos (2004). The ios test for model misspecification. *Journal of the American Statistical Association 99*(465), 216–227.

R Development Core Team (2013). *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0.

Rosenblatt, M. (1952). Remarks on a Multivariate Transformation. *The Annals of Mathematical Statistics 23*(3), 470–472.

Rüschendorf, L. (1976). Asymptotic distributions of multivariate rank order statistics. *The Annals of Statistics 4*(5), 912–923.

Sancetta, A. and S. Satchell (2004). The bernstein copula and its applications to modeling and approximations of multivariate distributions. *Econometric Theory 20*(3), 535–562.

Schepsmeier, U. (2010). Maximum likelihood estimation of C-vine pair-copula constructions on bivariate copulas from different families. Diploma thesis, Center of Mathematical Sciences, Munich University of Technology, Garching bei München.

Schepsmeier, U. (2013a). Efficient goodness-of-fit tests in multi-dimensional vine copula models. preprint, available at http://arxiv.org/abs/1309.5808.

Schepsmeier, U. (2013b). A goodness-of-fit test for regular vine copula models. preprint, available at: http://arxiv.org/abs/1306.0818.

Schepsmeier, U. and J. Stöber (2012). Web supplement: Derivatives and Fisher information of bivariate copulas. Technical report, TU München. available online: https://mediatum.ub.tum.de/node?id=1119201.

Schepsmeier, U. and J. Stöber (2013). Derivatives and fisher information of bivariate copulas. *Statistical Papers online first: http://link.springer.com/article/10.1007/s00362-013-0498-x.*

Schepsmeier, U., J. Stoeber, and E. C. Brechmann (2012). *VineCopula: Statistical inference of vine copulas*. R package version 1.0.

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

Smith, M., A. Min, C. Almeida, and C. Czado (2010). Modelling longitudinal data using a pair-copula decomposition of serial dependence. *The Journal of the American Statistical Association 105*(492), 1467–1479.

Smith, M. D. (2007). Invariance theorems for Fisher information. *Communications in Statistics - Theory and Methods 36*(12), 2213–2222.

Song, P. X.-K., Y. Fan, and J. D. Kalbfleisch (2005). Maximization by parts in likelihood inference. *Journal of the American Statistical Association 100*(472), 1145–1158.

Stöber, J., H. Hong, C. Czado, and P.Ghosh (2012). Comorbidity of chronic diseases in the elderly: longitudinal patterns identifed by a copula design for mixed responses. Preprint.

Stöber, J. and C. Czado (2012). Detecting regime switches in the dependence structure of high dimensional financial data. *arXiv preprint arXiv:1202.2009*. forthcoming in Computational Statistics and Data Analysis.

Stöber, J., C. Czado, and H. Joe (2013). Simplified pair copula constructions - limitations and extensions. *Journal of Multivariate Analysis. 119*, 101–118.

Stöber, J. and U. Schepsmeier (2013). Estimating standard errors in regular vine copula models. *Computational Statistics online first: http://link.springer.com/article/10.1007/s00180-013-0423-8*.

Taylor, L. W. (1987). The size bias of white's information matrix test. *Economics Letters 24*(1), 63–67.

Vuong, Q. (1989). Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses. *Econometrica 57*(2), 307–333.

Weiß, G. and M. Scheffer (2012, October). Smooth Nonparametric Bernstein Vine Copulas. *ArXiv e-prints*.

White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica 50*, 1–26.

Yan, J. (2007). Enjoy the joy of copulas: With a package copula. *Journal of Statistical Software 21*(4), 1–21.

Zhang, S., O. Okhrin, Q. M. Zhou, and P. X.-K. Song (2013). Goodness-of-fit Test For Specification of Semiparametric Copula Dependence Models. personal communication http://sfb649.wiwi.hu-berlin.de/papers/pdf/SFB649DP2013-041.pdf.

Zhou, Q. M., P. X.-K. Song, and M. E. Thompson (2012). Information ratio test for model misspecification in quasi-likelihood inference. *Journal of the American Statistical Association 107*(497), 205–213.