The University of Sydney Business School
The University of Sydney

BUSINESS ANALYTICS WORKING PAPER SERIES

Generalized Information Matrix Tests for Copulas

Artem Prokhorov

Ulf Schepsmeier

Yajing Zhu

June 2015

We propose a family of goodness-of-fit tests for copulas. The tests use generalizations of the information matrix (IM) equality of White (1982) and so relate to the copula test proposed by Huang and Prokhorov (2014). The idea is that eigenspectrum-based statements of the IM equality reduce the degrees of freedom of the test's asymptotic distribution and lead to better size-power properties, even in high dimensions. The gains are especially pronounced for vine copulas, where additional benefits come from simplifications of score functions and the Hessian. We derive the asymptotic distribution of the generalized tests, accounting for the non-parametric estimation of the marginals and apply a parametric bootstrap procedure, valid when asymptotic critical values are inaccurate. In Monte Carlo simulations, we study the behavior of the new tests, compare them with several Cramer-von Mises type tests and confirm the desired properties of the new tests in high dimensions.

BA Working Paper No: BAWP-2015-05

http://sydney.edu.au/business/business_analytics/research/working_papers

Generalized Information Matrix Tests for Copulas[*]

Artem Prokhorov[†]   **Ulf Schepsmeier**[‡]   **Yajing Zhu**[§]

June 2015

Abstract

We propose a family of goodness-of-fit tests for copulas. The tests use generalizations of the information matrix (IM) equality of White (1982) and so relate to the copula test proposed by Huang and Prokhorov (2014). The idea is that eigenspectrum-based statements of the IM equality reduce the degrees of freedom of the test's asymptotic distribution and lead to better size-power properties, even in high dimensions. The gains are especially pronounced for vine copulas, where additional benefits come from simplifications of score functions and the Hessian. We derive the asymptotic distribution of the generalized tests, accounting for the non-parametric estimation of the marginals and apply a parametric bootstrap procedure, valid when asymptotic critical values are inaccurate. In Monte Carlo simulations, we study the behavior of the new tests, compare them with several Cramer-von Mises type tests and confirm the desired properties of the new tests in high dimensions.

JEL Codes: C13
Key Words: information matrix equality, copula, goodness-of-fit, vine copulas, R-vines

---

[†]University of Sydney Business School, Sydney; email: artem.prokhorov@sydney.edu.au
[‡]Lehrstuhl für Mathematische Statistik, Technische Universität München, München; email: ulf.schepsmeier@tum.de
[§]Department of Economics, Concordia University, Montreal; email: yajing.s.zhu@gmail.com

1 Introduction

Consider a continuous random vector $\mathbf{X} = (X_1, \ldots, X_d)$ with a joint cumulative distribution function H and marginals $F_1, \ldots, F_d$. By Sklar's theorem, H has the following copula representation

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

where C is a unique cumulative distribution function, whose marginals are uniform on $[0, 1]^d$. Copulas represent the dependence structure between elements of $\mathbf{X}$ and this allows one to model and estimate distributions of random vectors by estimating the marginals and the copula separately. In economics, finance and insurance, this ability is very important because it facilitates accurate pricing of risk (see, e.g., Zimmer, 2012). In such problems d is often quite high – tens or hundreds – and this has spurred a lot of interest to high dimensional copula modeling and testing in the recent years (see, e.g., Patton, 2012).

In such high dimensions, classical multivariate parametric copulas such as the elliptical or Archimedean copulas are often insufficiently flexible in modeling different correlations or tail dependencies. On the other hand, they are very flexible and powerful in bivariate modeling. This advantage was used by Joe (1996) and later by Bedford and Cooke (2001, 2002) to construct multivariate densities using hierarchically bivariate copulas as building blocks. This process – known as a pair-copula construction (PCC, Aas et al., 2009) – results in a very flexible class of regular vine (R-vine) copula models, which can have a relative large dimension, yet remain computationally tractable (see, e.g., Czado, 2010; Kurowicka and Cooke, 2006, for introductions to vine copulas).

A copula model for $\mathbf{X}$ arises when C is unknown but belongs to a parametric family $\mathcal{C}_0 = \{C_{\boldsymbol{\theta}} : \boldsymbol{\theta} \in \mathcal{O}\}$, where $\mathcal{O}$ is an open subset of $\mathbb{R}^p$ for some integer $p \geq 1$, and $\boldsymbol{\theta}$ denotes the copula parameter vector. There is a wide literature on estimation of $\boldsymbol{\theta}$ under the assumption $\mathcal{H}_0 : C \in \mathcal{C}_0 = \{C_{\boldsymbol{\theta}} : \boldsymbol{\theta} \in \mathcal{O}\}$ given independent copies $\mathbf{X}_1 = (X_{11}, \ldots, X_{1d}), \ldots, \mathbf{X}_n = (X_{n1}, \ldots, X_{nd})$ of $\mathbf{X}$; see, e.g., Genest et al. (1995), Joe (2005), Prokhorov and Schmidt (2009). The complementary issue of testing

$$\mathcal{H}_0 : C \in \mathcal{C}_0 = \{C_{\boldsymbol{\theta}} : \boldsymbol{\theta} \in \mathcal{O}\} \quad \text{vs.} \quad \mathcal{H}_1 : C \notin \mathcal{C}_0 = \{C_{\boldsymbol{\theta}} : \boldsymbol{\theta} \in \mathcal{O}\}$$

is more recent – surveys of available tests can be found in Berg (2009) and Genest et al. (2009).

Currently, the main problem in testing is to develop operational "blanket" tests, powerful in high dimensions. This means we need tests which remain computationally feasible and powerful against a wide class of high-dimensional alternatives, rather

than against specific low-dimensional families, and which do not require ad hoc choices, such as a bandwidth, a kernel, or a data categorization (see, e.g., Klugman and Parsa, 1999; Genest and Rivest, 1993; Junker and May, 2005; Fermanian, 2005; Scaillet, 2007). Genest et al. (2009) discuss five testing procedures that qualify as "blanket" tests. We will use some of them in our simulations.

Recently, Huang and Prokhorov (2014) proposed a "blanket" test based on the information matrix equality for copulas and Schepsmeier (2013, 2015) extended that test to vine copulas. The point of this test is to compare the expected Hessian for $\boldsymbol{\theta}$ with the expected outer-product-of-the-gradient (OPG) form of the covariance matrix – under $\mathcal{H}_0$, their sum should be zero. This is the so called Bartlett identity. So in multi-parameter cases, the statistic is based on a random vector whose dimension – being equal to the number of distinct elements in the Hessian – grows as the square of the number of parameters. Even though the statistic has a standard asymptotic distribution, simulations suggest that using analytical critical values leads to severe oversize distortions, especially when dimension is high.

The tests we propose in this paper are motivated by recent developments in information matrix equality testing (see, e.g., Golden et al., 2013). Specifically, we use alternative, eigenspectrum-based statements of the information matrix equality. This means we use functions of the eigenvalues of the two matrices, instead of the distinct elements of the matrices. This leads to a noticeable reduction in dimension of the random vector underlying the test statistic, which permits significant size and power improvements. The improvements are more pronounced for high dimensional dependence structures. Regular vine copulas are effective in this setting because of a further dimension reduction they permit. We argue that R-vines offer additional computational benefits for our tests. Compared to available alternatives, our tests applied to vine copula constructions remain operational and powerful in fairly high dimensions and seem to be the only tests allowing for copula specification testing in high dimensions.

The paper is organized as follows. In Section 2, we introduce seven new goodness-of-fit tests for copulas and discuss their asymptotic properties. Section 3 describes the computational benefits that result from applying our tests to vine copulas. In Section 4 we use the new tests in a Monte Carlo study where we first study the new copula tests in terms of their size and power performance and then examine the effect of dimensionality, sample size and dependence strength on size and power of these tests, as compared with three popular "blanket" tests that perform well in simulations. Section 5 concludes.

2 Generalized Information Matrix Test for Copulas

In the setting of general specification testing, Golden et al. (2013) introduced an extension to the original information equality test of White (1982), which they call Generalized Information Matrix Test (GIMT). Unlike the original test which is based on the negative expected Hessian and OPG, GIMT is based on functions of the eigenspectrum of the two matrices. In this section we develop a series of copula goodness-of-fit tests which draw on GIMT and we study their properties.

2.1 Basic Asymptotic Result

Let $\mathbf{X}_i = (X_{i1}, \ldots, X_{id})$, $i = 1, \ldots, n$, denote realizations of a random vector $\mathbf{X} = (X_1, \ldots, X_d) \in \mathbb{R}^d$. All tests we consider are based on a pseudo-sample $\mathbf{U}_1 = (U_{11}, \ldots, U_{1d}), \ldots, \mathbf{U}_n = (U_{n1}, \ldots, U_{nd})$, where $\mathbf{U}_i = (U_{i1}, \ldots, U_{id}) = (\frac{R_{i1}}{n+1}, \ldots, \frac{R_{id}}{n+1})$, and R_{ij} is the rank of X_{ij} amongst $X_{1j}, \ldots, X_{nj}$. This transformation of each X_{ij} to its normalized rank can be viewed as the empirical marginal distribution of X_j, $j = 1, \ldots, d$. The denominator $n + 1$ is used instead of n to avoid numerical problems at the boundaries of $[0, 1]^d$. Given an independent sample $\{\mathbf{X}_1, \ldots, \mathbf{X}_n\}$, the pseudo-sample $\{\mathbf{U}_1, \ldots, \mathbf{U}_n\}$ – no longer independent due to the rank transformation – can be viewed as a sample from the underlying copula C.

Assume that the copula density $c_{\boldsymbol{\theta}}$ exists. Let $\mathbb{H}(\boldsymbol{\theta})$ denote the expected Hessian matrix of $\ln c_{\boldsymbol{\theta}}$ and let $\mathbb{C}(\boldsymbol{\theta})$ denote the expected outer product of the corresponding score function (OPG), i.e.,

$$\mathbb{H}(\boldsymbol{\theta}) = \mathbb{E}\nabla_{\boldsymbol{\theta}}^2 \ln c_{\boldsymbol{\theta}}(\mathbf{U}) \qquad \text{and} \qquad \mathbb{C}(\boldsymbol{\theta}) = \mathbb{E}\nabla_{\boldsymbol{\theta}} \ln c_{\boldsymbol{\theta}}(\mathbf{U}) \, \nabla_{\boldsymbol{\theta}}^{'} \ln \, c_{\boldsymbol{\theta}}(\mathbf{U}),$$

where "$\nabla_{\boldsymbol{\theta}}$" denotes derivatives with respect to $\boldsymbol{\theta}$ and expectations are with respect to the true distribution H. Let $\boldsymbol{\theta}_0$ denote the true value of $\boldsymbol{\theta}$ and assume $\mathbb{H}(\boldsymbol{\theta}_0)$ and $\mathbb{C}(\boldsymbol{\theta}_0)$ are in the interior of a compact set $S^{p \times p} \subseteq \mathbb{R}^{p \times p}$. For $i = 1, \ldots, n$, let

$$\mathbb{H}_i(\boldsymbol{\theta}) = \nabla_{\boldsymbol{\theta}}^2 \ln c_{\boldsymbol{\theta}}(\mathbf{U}_i) \qquad \text{and} \qquad \mathbb{C}_i(\boldsymbol{\theta}) = \nabla_{\boldsymbol{\theta}} \ln c_{\boldsymbol{\theta}}(\mathbf{U}_i) \, \nabla_{\boldsymbol{\theta}}^{'} \ln c_{\boldsymbol{\theta}}(\mathbf{U}_i).$$

For any $\boldsymbol{\theta} \in \mathcal{O}$, define the sample analogues of $\mathbb{H}(\boldsymbol{\theta})$ and $\mathbb{C}(\boldsymbol{\theta})$:

$$\bar{\mathbb{H}}(\boldsymbol{\theta}) := n^{-1} \sum_{i=1}^{n} \mathbb{H}_i(\boldsymbol{\theta}) \qquad \text{and} \qquad \bar{\mathbb{C}}(\boldsymbol{\theta}) := n^{-1} \sum_{i=1}^{n} \mathbb{C}_i(\boldsymbol{\theta}).$$

Then, given an estimate $\hat{\boldsymbol{\theta}}_n$ of $\boldsymbol{\theta}_0$, we can denote estimates of $\mathbb{H}(\boldsymbol{\theta}_0)$ and $\mathbb{C}(\boldsymbol{\theta}_0)$ by

$$\bar{\mathbb{H}}_n := \bar{\mathbb{H}}(\hat{\boldsymbol{\theta}}_n) \qquad \text{and} \qquad \bar{\mathbb{C}}_n := \bar{\mathbb{C}}(\hat{\boldsymbol{\theta}}_n).$$

Definition 1 (Hypothesis Function) *Let $s : S^{p \times p} \times S^{p \times p} \to \mathbb{R}^r$ be a continuous differentiable function in both of its matrix arguments. s is called a hypothesis function if for every $A, B \in S^{p \times p}$ it follows:*

$$\text{If } A = -B, \text{then } s(A, B) = 0_r,$$

where 0_r is the zero vector of dimension r.

Definition 2 (GIMT) *A test statistic $\hat{s}_n := s(\bar{\mathbb{H}}_n, \bar{\mathbb{C}}_n)$ is a GIMT for copula $C_{\boldsymbol{\theta}}$ if it tests the null hypothesis:*

$$H_0 : s(\mathbb{H}(\boldsymbol{\theta}_0), \mathbb{C}(\boldsymbol{\theta}_0)) = 0_r.$$

We can now look at the properties of the GIMT for copulas.

Lemma 1 (Asymptotic Normality of $\sqrt{n}\hat{s}_n$) *Let $s : S^{p \times p} \times S^{p \times p} \to \mathbb{R}^r$ be a GIMT hypothesis function with $\nabla_{\boldsymbol{\theta}} s(\mathbb{H}(\boldsymbol{\theta}), \mathbb{C}(\boldsymbol{\theta}))$ evaluated at $\boldsymbol{\theta}_0$ having full row rank r. Then, under $\mathcal{H}_0$ and suitable regularity conditions,*

$$\sqrt{n}\hat{s}_n \xrightarrow{d} N(0, \Sigma_s(\boldsymbol{\theta}_0)),$$

where the asymptotic covariance matrix is given by

$$\Sigma_s(\boldsymbol{\theta}_0) := (\nabla s_{\boldsymbol{\theta}_0}) V_{\boldsymbol{\theta}_0} (\nabla s_{\boldsymbol{\theta}_0})', \tag{1}$$

where $\nabla s_{\boldsymbol{\theta}_0}$ and $V_{\boldsymbol{\theta}_0}$ are given in Eqs.(6)-(7) of Appendix A.

Proof: see Appendix A for all proofs.

The regularity conditions used in Lemma 1 are standard assumptions of continuity and differentiability of the likelihood and rank conditions on information (see, e.g., White, 1982, Assumptions A1-A10). In the copula context, they translate into equivalent assumptions on the copula density (see, e.g., Genest et al., 1995).

The main difference between Lemma 1 and the specification tests of White (1982) and Golden et al. (2013) is in the form of $V_{\boldsymbol{\theta}_0}$. The complication arises from the rank transformation which requires a non-trivial adjustment to the variance of $\hat{s}_n$, accounting for the estimation error (see Huang and Prokhorov, 2014).

Theorem 1 (Asymptotic Theory) *Let $\hat{\Sigma}_{n,s}$ denote a consistent estimate of the asymptotic covariance matrix $\Sigma_s(\boldsymbol{\theta}_0)$. Then, under $\mathcal{H}_0$ and suitable regularity conditions, the GIMT statistic for copulas*

$$\mathcal{W}_n := n\,\hat{s}_n'\hat{\Sigma}_{n,s}^{-1}\hat{s}_n \tag{2}$$

is asymptotically χ_r^2 distributed.

These results suggest that we can use any function of $\mathbb{H}(\boldsymbol{\theta}_0)$ and $\mathbb{C}(\boldsymbol{\theta}_0)$ with a known probability limit for testing copula validity. One of the main insights of Golden et al. (2013) is that different hypothesis functions permit misspecification testing in different directions. For example, a test comparing the determinants of $\mathbb{H}$ and $\mathbb{C}$ will detect small variations in eigenvalues of the two matrices, while a test comparing traces will focus on differences in the major principal components of the two matrices.

In multivariate settings, the dimension of $\boldsymbol{\theta}$ often grows faster than the dimension of $\mathbf{U}$. For example, a d-variate t-copula has $O(d^2)$ parameters. The eigenspectrum-based hypothesis functions allow to reduce the dimension of the test statistic (and thus the degrees of freedom of the test) from $p(p+1)/2$, where p is the number of copula parameters, to the number of values of the hypothesis function, r.

A consistent estimator $\hat{\Sigma}_{n,s}$ would require estimation of $\nabla s_{\boldsymbol{\theta}_0}$ and V_{θ_0}. Some aspects of consistent estimation of V_{θ_0} are discussed by Huang and Prokhorov (2014) so in the propositions that follow we focus on the additional complexity introduced by the various hypothesis functions through ∇s_{θ_0}.

Table 1 lists the hypothesis functions we consider. The original *White* and *IR (Information Ratio) Tests* are special cases. We introduce the *Trace White Test* to focus on the sum of the eigenvalues of $\mathbb{H} + \mathbb{C}$ and the *Determinant White Test* to focus on the product of the eigenvalues of $\mathbb{H} + \mathbb{C}$. The focused testing allows for directional power which we discuss later.

Two more tests are log-versions of the last two. The *(Log) Determinant IR Test* focuses on the determinant of the information matrix ratio, and the *Log Trace Test* looks at whether the sum of the eigenvalues is the same for the negative Hessian and the OPG form. We use logarithms here as variance stabilizing transformations. In contrast to the White (or IR) version, the *Log Trace Test* does not use the eigenvalues of the sum (or the ratio) of $\mathbb{H}$ and $\mathbb{C}$, rather it looks at the eigenvalues of each matrix separately.

The *Log GAIC* (Generalized Akaike Information Criterion) *Test* picks on the idea of the *IR Test* that the negative Hessian multiplied by the inverse of the OPG (or vice versa) equals the identity matrix. The new feature is that we focus on the

Table 1: Summary of eigenspectrum tests

Name	short	$s(\mathbb{H}, \mathbb{C})$
White Test	$\mathcal{T}_n$	$vech(\mathbb{H}) + vech(\mathbb{C}) = 0$
Determinant White Test	$\mathcal{T}_n^{(D)}$	$det(\mathbb{H} + \mathbb{C}) = 0$
Trace White Test	$\mathcal{T}_n^{(T)}$	$tr(\mathbb{H} + \mathbb{C}) = 0$
IR Test	$\mathcal{Z}_n$	$tr(-\mathbb{H}^{-1}\mathbb{C}) - p = 0$
Determinant IR Test	$\mathcal{Z}_n^{(D)}$	$det(-\mathbb{H}^{-1}\mathbb{C}) - 1 = 0$
Log Trace IMT	Tr_n	$\log(tr(-\mathbb{H}^{-1})) - \log(tr(\mathbb{C})) = 0$
Log GAIC IMT	$\mathcal{G}_n$	$\log[\frac{1}{p}(1_p)'(\Lambda(-\mathbb{H}^{-1}) \odot \Lambda(\mathbb{C}))] = 0$
Log Eigenspectrum IMT	$\mathcal{P}_n$	$\log(\Lambda(-\mathbb{H}^{-1})) - \log(\Lambda(\mathbb{C}^{-1})) = 0_p$
Eigenvalue Test	$\mathcal{Q}_n$	$\Lambda(-\mathbb{H}^{-1}\mathbb{C}) = 1_p$

average product of the Hessian-based eigenvalues and OPG-based eigenvalues. The last two tests are explicitly based on the full eigenspectrum. The *Eigenspectrum Test* compares the eigenvalues of $\mathbb{H}$ and $\mathbb{C}$ separately, the *Eigenvalue Test* uses the eigenvalues of the information matrix ratio.

All these hypothesis functions are identical under the null, yet the behavior of these tests varies widely. We first look at the asymptotic approximations of the behavior.

2.2 White Test for Copulas

In the case of the original White (1982) test, the asymptotic covariance matrix in Lemma 1 simplifies. Huang and Prokhorov (2014, Proposition 1) provide the asymptotic variance matrix for this case. It can be obtained by rearranging the building blocks used in construction of the test statistic (elements of $\boldsymbol{d}(\boldsymbol{\theta})$ in Appendix A), and by setting $\nabla s_{\boldsymbol{\theta}_0} = \left(\boldsymbol{I}_{p(p+1)/2} \ \boldsymbol{I}_{p(p+1)/2} \right)$, where $\boldsymbol{I}_k$ is a $k \times k$ identity matrix.

One of the most important criticisms of this test is its slow convergence to the asymptotic distribution. One cause of this problem is its high degrees of freedom. For example, in the setting of a vine copula estimation, Schepsmeier (2013) shows that for a five-dimensional vine ($df = p(p+1)/2 = 55$), the number of observations needed to show acceptable size and power behavior using asymptotic critical values is at least 10,000; for an eight-dimensional vine ($df = 406$) that number is greater than 20,000. The alternatives we propose are determinant- and trace-based.

Proposition 1 (Determinant White Test) *Let $\hat{\Sigma}_{s,n}$ be as defined in Theorem 1,*

7

with an estimator of $\nabla s_{\boldsymbol{\theta}_0}$ given by

$$\widehat{\nabla s}_{\boldsymbol{\theta}_0} = det(\bar{\mathbb{H}}_n + \bar{\mathbb{C}}_n) vech[(\bar{\mathbb{H}}_n + \bar{\mathbb{C}}_n)^{-1}]' \left[\boldsymbol{I}_{p(p+1)/2}, \; \boldsymbol{I}_{p(p+1)/2} \right].$$

Then, under $\mathcal{H}_0$, the asymptotic distribution of the test statistic

$$\mathcal{T}_n^{(D)} := n \frac{[det(\bar{\mathbb{H}}_n + \bar{\mathbb{C}}_n)]^2}{\hat{\Sigma}_{s,n}}$$

is χ_1^2.

Proposition 2 (Trace White Test) *Let $\hat{\Sigma}_{s,n}$ be as defined in Theorem 1 with $\nabla s_{\boldsymbol{\theta}_0}$ defined as follows*

$$\widehat{\nabla s}_{\boldsymbol{\theta}_0} = (vech(\boldsymbol{I}_p)', \quad vech(\boldsymbol{I}_p)').$$

Then, under $\mathcal{H}_0$, the asymptotic distribution of the test statistic

$$\mathcal{T}_n^{(T)} := n \frac{tr(\bar{\mathbb{H}}_n + \bar{\mathbb{C}}_n)^2}{\hat{\Sigma}_{s,n}}$$

is χ_1^2.

The two tests are chi-square with one degree of freedom, rather than $p(p+1)/2$, and have important differences allowing for what can be called directional testing. Because larger eigenvalues have larger effect on determinant than on trace, the *Trace White Test* will be less sensitive to changes in eigenvalues, especially small ones, and thus less powerful than the *Determinant White Test*.

2.3 Information Ratio Test for Copulas

As extensions of the original White test, Zhou et al. (2012) and Presnell and Boos (2004) consider using a ratio of the Hessian and OPG. Under correct specification, the matrix $\mathbb{H}(\boldsymbol{\theta})^{-1}\mathbb{C}(\boldsymbol{\theta})$ is equal to a p-dimensional identity matrix. We propose two versions of this test for copulas.

Proposition 3 (IR Test) *Let $\hat{\Sigma}_{s,n}$ be as defined in Theorem 1 with an estimator of $\nabla s_{\boldsymbol{\theta}_0}$ given by*

$$\widehat{\nabla s}_{\boldsymbol{\theta}_0} = \left[vech \left(\bar{\mathbb{H}}_n^{-1} \bar{\mathbb{C}}_n \bar{\mathbb{H}}_n^{-1} \right)', \quad vech \left(-\bar{\mathbb{H}}_n^{-1} \right)' \right].$$

Then, under $\mathcal{H}_0$, the asymptotic distribution of the test statistic

$$\mathcal{Z}_n := n \frac{\left[tr(-\bar{\mathbb{H}}_n^{-1}\bar{\mathbb{C}}_n) - p \right]^2}{\hat{\Sigma}_{s,n}}$$

is χ_1^2.

Proposition 4 (Log-Determinant IR Test) *Let $\hat{\Sigma}_{s,n}$ be as defined in Theorem 1, with an estimator of $\nabla s_{\boldsymbol{\theta}_0}$ given by*

$$\widehat{\nabla s}_{\boldsymbol{\theta}_0} = \det(\bar{\mathbb{H}}_n^{-1}\bar{\mathbb{C}}_n) \left(vech\left(-\bar{\mathbb{C}}_n\bar{\mathbb{H}}_n^{-1}\bar{\mathbb{C}}_n \right)', \quad vech\left(\bar{\mathbb{C}}_n^{-1} \right)' \right).$$

Then, under $\mathcal{H}_0$, the asymptotic distribution of the test statistic

$$\mathcal{Z}_n^{(D)} := n \frac{\left(\log(\det(-\bar{\mathbb{H}}_n^{-1}\bar{\mathbb{C}}_n)) \right)^2}{\hat{\Sigma}_{s,n}}$$

is χ_1^2.

2.4 Log Trace Test for Copulas

Similar to the Log-Determinant IR Test we can construct a test using the log of traces of $-\mathbb{H}$ and $\mathbb{C}$, which should be identical under the null.

Proposition 5 (Log Trace Test) *Let $\hat{\Sigma}_{s,n}$ be as defined in Theorem 1, with an estimator of $\nabla s_{\boldsymbol{\theta}_0}$ given by*

$$\widehat{\nabla s}_{\boldsymbol{\theta}_0} = \left(-\frac{1}{tr(-\bar{\mathbb{H}}_n^{-1})} vech(\bar{\mathbb{H}}_n^{-2})', \ -\frac{1}{tr(-\bar{\mathbb{C}}_n^{-1})} vech(I_p)' \right).$$

Then, under $\mathcal{H}_0$, the asymptotic distribution of the test statistic

$$Tr_n := n \frac{\left[\log(tr(-\bar{\mathbb{H}}_n^{-1})) - \log(tr(\bar{\mathbb{C}}_n)) \right]^2}{\hat{\Sigma}_{s,n}}$$

is χ_1^2.

As mentioned earlier, trace-based tests pick up changes in larger eigenvalues easier than in smaller – a property desirable for some alternatives.

2.5 Log GAIC Test for Copulas

It is well known (see, e.g., Takeuchi, 1976) that under model misspecification the Generalized Akaike Information Criterion defined as follows

$$GAIC := -2\log\prod_{i=1}^{n} f(U_i; \hat{\boldsymbol{\theta}}_n) + 2tr(-\bar{\mathbb{H}}^{-1}(\hat{\boldsymbol{\theta}}_n)\bar{\mathbb{C}}(\hat{\boldsymbol{\theta}}_n))$$

is an unbiased estimator of the expected value of $-2\log\prod_{i=1}^{n} f(U_i; \hat{\boldsymbol{\theta}}_n)$. Under correct model specification $2tr(-\bar{\mathbb{H}}^{-1}(\hat{\boldsymbol{\theta}}_n)\bar{\mathbb{C}}(\hat{\boldsymbol{\theta}}_n)) \to 2p$, since $-\bar{\mathbb{H}}^{-1}(\hat{\boldsymbol{\theta}}_n)\bar{\mathbb{C}}(\hat{\boldsymbol{\theta}}_n) \to I_p$ a.s., and so GAIC becomes AIC. This motivates the use of the IR Test but also of the following form of the GIMT.

Let $\Lambda(A) = (\lambda_1, \ldots, \lambda_p)'$ denote the vector of sorted eigenvalues of $A \in \mathbb{R}^{p\times p}$. Further, let $\Lambda^{-1}(A) := 1/\Lambda(A)$ denote component-wise $\{1/\lambda_j\}_{j=1}^{p}$ and $\Lambda(A^{-1}) = \Lambda^{-1}(A)$. Then, under the null, $tr(-\mathbb{H}^{-1}\mathbb{C}) = (1_p)'(\Lambda(-\mathbb{H}^{-1}) \odot \Lambda(\mathbb{C}))$, where $\odot$ denotes the Hadamard product, i.e. component-wise multiplication; however, generally, eigenvalues of the product matrix are not equal to the product of eigenvalues of the components.

Proposition 6 (GAIC Test) *Let $\hat{\Sigma}_{s,n}$ be as defined in Theorem 1, with an estimator of $\nabla s_{\boldsymbol{\theta}_0}$ given by*

$$\widehat{\nabla s}_{\boldsymbol{\theta}_0} = \frac{1}{\text{tr}(\bar{\mathbb{H}}_n^{-1}\bar{\mathbb{C}}_n)} \left[vech\left(\bar{\mathbb{H}}_n^{-1}\bar{\mathbb{C}}_n\bar{\mathbb{H}}_n^{-1}\right)', \quad vech\left(-\bar{\mathbb{H}}_n^{-1}\right)' \right].$$

Then, under $\mathcal{H}_0$, the asymptotic distribution of the test statistic

$$\mathcal{G}_n := n\frac{\left\{\log\left[\frac{1}{p}(1_p)'\left(\Lambda(-\bar{\mathbb{H}}_n^{-1}) \odot \Lambda(\bar{\mathbb{C}}_n)\right)\right]\right\}^2}{\hat{\Sigma}_{s,n}}$$

is χ_1^2.

In contrast to the *IR Test* the eigenvalues of the Hessian and the OPG are calculated separately. Thus, similar to the *Log Determinant IR Test*, the *Log GAIC Test* is more sensitive to changes in the entire eigenspectrum than the *IR Test* (see Golden et al., 2013, for a more detailed discussion).

2.6 Eigenvalue Test for Copulas

The form of the *Log Eigenspectrum IMT* was initially proposed by Golden et al. (2013). It is a p-degrees-of-freedom test so the reduction in the degrees-of-freedom from $p(p+1)/2$ is more noticeable for larger p, which would typically mean a higher dimensional copula.

In order to derive its asymptotic distribution we need additional notation. For a real symmetric matrix A, let $y_j(A)$ denote the normalized eigenvector corresponding to eigenvalue $\lambda_j(A)$, $j = 1, \ldots, p$. Let D denote the duplication matrix, i.e. such a matrix that $Dvech(A) = vec(A)$ (see, e.g. Magnus and Neudecker, 1999).

Proposition 7 (Log Eigenspectrum Test) *Let $\hat{\Sigma}_{s,n}$ be as defined in Theorem 1, with an estimator of $\nabla s_{\boldsymbol{\theta}_0}$ given by*

$$
\widehat{\nabla s}_{\boldsymbol{\theta}_0} = \begin{bmatrix}
-\frac{1}{\lambda_1(\bar{\mathbb{H}}_n)}[y_1(\bar{\mathbb{H}}_n)' \otimes y_1(\bar{\mathbb{H}}_n)']D & \frac{1}{\lambda_1(\bar{\mathbb{C}}_n)}[y_1(\bar{\mathbb{C}}_n)' \otimes y_1(\bar{\mathbb{C}}_n)']D \\
\vdots & \vdots \\
-\frac{1}{\lambda_p(\bar{\mathbb{H}}_n)}[y_p(\bar{\mathbb{H}}_n)' \otimes y_p(\bar{\mathbb{H}}_n)']D & \frac{1}{\lambda_p(\bar{\mathbb{C}}_n)}[y_p(\bar{\mathbb{C}}_n)' \otimes y_p(\bar{\mathbb{C}}_n)']D
\end{bmatrix}.
$$

Then, under $\mathcal{H}_0$, the asymptotic distribution of the test statistic

$$
\mathcal{P}_n := n\big[\log(\Lambda(-\bar{\mathbb{H}}_n^{-1})) - \log(\Lambda(\bar{\mathbb{C}}_n^{-1}))\big]'\hat{\Sigma}_{s,n}^{-1}\big[\log(\Lambda(-\bar{\mathbb{H}}_n^{-1})) - \log(\Lambda(\bar{\mathbb{C}}_n^{-1}))\big]
$$

is χ_p^2.

A similar approach is based on the eigenspectrum of the information matrix ratio $\Lambda(-\mathbb{H}^{-1}(\boldsymbol{\theta}_0)\mathbb{C}(\boldsymbol{\theta}_0))$. We will call this test the *Eigenvalue Test*.

Proposition 8 (Eigenvalue Test) *Let $\hat{\Sigma}_{s,n}$ be as defined in Theorem 1 with an estimator of $\nabla s_{\boldsymbol{\theta}_0}$ given by*

$$
\widehat{\nabla s}_{\boldsymbol{\theta}_0} = \begin{bmatrix}
\frac{1}{\lambda_1(\bar{\mathbb{H}}_n)}[y_1(\bar{\mathbb{C}}_n)' \otimes y_1(\bar{\mathbb{C}}_n)']D & -\frac{\lambda_1(\bar{\mathbb{C}}_n)}{\lambda_1(\bar{\mathbb{H}}_n)^2}[y_1(\bar{\mathbb{H}}_n)' \otimes y_1(\bar{\mathbb{H}}_n)']D \\
\vdots & \vdots \\
\frac{1}{\lambda_p(\bar{\mathbb{H}}_n)}[y_p(\bar{\mathbb{C}}_n)' \otimes y_p(\bar{\mathbb{C}}_n)']D & -\frac{\lambda_p(\bar{\mathbb{C}}_n)}{\lambda_p(\bar{\mathbb{H}}_n)^2}[y_p(\bar{\mathbb{H}}_n)' \otimes y_p(\bar{\mathbb{H}}_n)']D
\end{bmatrix}.
$$

Then, under $\mathcal{H}_0$, the asymptotic distribution of the test statistic

$$
\mathcal{Q}_n := n\big[\Lambda(-\bar{\mathbb{H}}_n^{-1}\bar{\mathbb{C}}_n) - 1_p\big]'\hat{\Sigma}_{s,n}^{-1}\big[\Lambda(-\bar{\mathbb{H}}_n^{-1}\bar{\mathbb{C}}_n) - 1_p\big]
$$

is χ_p^2.

3   GIMTs for Vine Copulas

A regular vine (R-vine) copula is a nested set of bivariate copulas representing uncon-
ditional and conditional dependence between elements of the initial random vector
(see, e.g., Joe, 1996; Bedford and Cooke, 2001, 2002). Any d-variate copula can be
expressed as a product of such (conditional) bivariate copulas and there are many
ways of writing this product. Graphically, R-vine copulas can be illustrated by a
set of connected trees $\mathcal{V} = \{T_1, \ldots, T_{d-1}\}$, where each edge represents a bivariate
conditional copula. The nodes illustrate the arguments of the associated copula.
The edges of tree T_i form the nodes of tree T_{i+1}, $i \in \{1, \ldots, d-2\}$. The proximity
condition of Bedford and Cooke (2001) then defines which possible edges are allowed
between the nodes to form an R-vine. If we denote the set of bivariate copulas used
in trees $\mathcal{V}$ by $\mathcal{B}(\mathcal{V})$ and the corresponding set of parameters by $\boldsymbol{\theta}(\mathcal{B}(\mathcal{V}))$, then we can
specify an R-vine copula by $(\mathcal{V}, \mathcal{B}(\mathcal{V}), \boldsymbol{\theta}(\mathcal{B}(\mathcal{V})))$.

Let $U_1, \ldots, U_d$ denote a pseudo-sample as introduced in Section 2.1. The edges
$j(e), k(e)|D(e)$ in E_i, for $1 \leq i \leq d-1$ correspond the set of bivariate copula
densities $\mathcal{B} = \{c_{j(e),k(e)|D(e)}|e \in E_i, 1 \leq i \leq d-1\}$. The indices $j(e)$ and $k(e)$ form
the *conditioned set* while $D(e)$ is called *conditioning set*. Then a regular vine copula
density is given by the product

$$c_{1,\ldots,d}(\boldsymbol{u}) = \prod_{i=1}^{d-1} \prod_{e \in E_i} c_{j(e),k(e);D(e)}(C_{j(e)|D(e)}(u_{j(e)}|\boldsymbol{u}_{D(e)}), C_{k(e)|D(e)}(u_{k(e)}|\boldsymbol{u}_{D(e)})). \quad (3)$$

The copula arguments $C_{j(e)|D(e)}(u_{j(e)}|\boldsymbol{u}_{D(e)})$ and $C_{k(e)|D(e)}(u_{k(e)}|\boldsymbol{u}_{D(e)})$ can be derived
integral-free by the formula derived from the first derivative of the corresponding cdf
with respect to the second copula argument (Joe, 1996):

$$C_{j(e)|D(e)}(u_{j(e)}|\boldsymbol{u}_{D(e)}) = \frac{\partial C_{j(e),j'(e);D(e)\backslash j'(e)}(C(u_{j(e)}|\boldsymbol{u}_{D(e)\backslash j'(e)}), C(u_{j'(e)}|\boldsymbol{u}_{D(e)\backslash j'(e)}))}{\partial C(u_{j'(e)}|\boldsymbol{u}_{D(e)\backslash j'(e)})}.$$

An example of a 5-dimensional R-vine is given in Figure 1.

The canonical vine (C-vine) and the drawable vine (D-vine) are two special R-
vines. The C-vine has in each tree a root node which is connected to all other nodes
in this tree. In the D-vine each node is connected to two other nodes at most.

The copula parameter vector $\boldsymbol{\theta}(\mathcal{B}(\mathcal{V}))$ can be estimated either in a tree-by-tree
approach called sequential estimation, or in a full maximum likelihood estimation
(MLE) procedure (Aas et al., 2009). The sequential procedure uses the hierarchical
structure of R-vines and is quick – its results are often used as starting values for
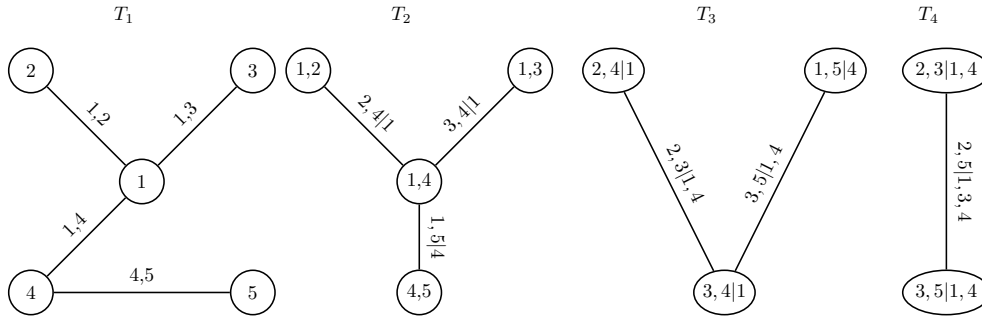the MLE approach. Both are consistent estimators.

Figure 1: Tree structure of a 5-dimensional R-vine copula.

Vine copulas have gained popularity because of the benefits they offer when dimension d is high. First, they permit a decomposition of a d-variate copula with $O(d^2)$ or more parameters into $d(d-1)/2$ bivariate (one-parameter) copulas, which reduces computational burden. Second, they offer a natural way to impose conditional independence by dropping selected higher-order edges in $\mathcal{V}$. Finally, the integral free expressions for the conditional copulas offer an additional computational benefit.

Such a reduction of parameters using the conditional independence copula can be achieved in two ways. First, single conditional copulas can be assumed independent, especially if some pre-testing procedure confirms this (see, e.g., Genest and Favre, 2007). Further, by setting all pair-copula families above a certain tree order to the independence copula, the number of parameters can be reduced significantly. This involves no testing and is often done heuristically; Brechmann et al. (2012) call this approach *truncation*.

In our settings, vine copulas offer an additional advantage over conventional copulas. As an example, consider testing goodness-of-fit of a d-variate Eyraud-Farlie-Gumbel-Morgenstern (EFGM) copula. This copula has $2^d - d - 1$ parameters so the number of degrees-of-freedom for the *White Test* is of order $O(2^{2d})$, while for the eigenspectrum-based tests that number is as low as one. Regardless of the GIMT, the calculation of the test statistic involves evaluating, analytically or numerically, the score function and the Hessian. The score $\nabla_{\boldsymbol{\theta}} \ln c_{\boldsymbol{\theta}}$ is a vector-valued function with $2^d - d - 1$ elements, each a function of all $2^d - d - 1$ elements of $\boldsymbol{\theta}$. The Hessian is a large matrix-valued function, in which each component is a function of the entire vector $\boldsymbol{\theta}$. Now what changes if we replace that copula with a d-variate vine?

Consider the case of $d = 3$. Suppose we use the following R-vine representation

$$c_{123}(u_1, u_2, u_3; \boldsymbol{\theta}) = c_{12}(u_1, u_2; \theta_1) c_{23}(u_2, u_3; \theta_2) c_{13;2}(C_{1|2}(u_1|u_2; \theta_1), C_{3|2}(u_3|u_2; \theta_2); \theta_3),$$

where each bivariate copula is EFGM and $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)$. Then, it is easy to see that $\nabla_{\boldsymbol{\theta}} \ln c_{\boldsymbol{\theta}}$ has the form

$$\begin{pmatrix} \nabla_{\theta_1} \ln c_{12} + \nabla_{\theta_1} \ln c_{13;2} \\ \nabla_{\theta_2} \ln c_{23} + \nabla_{\theta_2} \ln c_{13;2} \\ \nabla_{\theta_3} \ln c_{13;2} \end{pmatrix},$$

where each element is a score function for the corresponding element of $\boldsymbol{\theta}$ – a simpler function with fewer argument (see, e.g., Stöber and Schepsmeier, 2013). The term $\nabla_{\theta_1} \ln c_{13;2}$ is the only term that has all three parameters but if a sequential procedure is used, estimates of θ_1 and θ_2 come from previous steps and are treated as known so only θ_3 is effectively unknown in $c_{13;2}$. Regardless of the estimation method, only derivatives of bivariate copulas are needed, which are much simpler than in higher dimensions. Closed form expressions for the first two derivatives of several bivariate copulas are given in Schepsmeier and Stöber (2014, 2012). The Hessian $\mathbb{H}$ will simplify accordingly – some cross derivatives will be zero (Stöber and Schepsmeier, 2013). The same is true for the third-order derivatives used to obtain $\hat{\Sigma}_{s,n}$.

These are sizable simplifications when dealing with high dimensional copulas. The problem is that multivariate dependence requires sufficiently rich parametrization which affects tests' properties. It also imposes heavy computational burdens as most available "blanket" tests use parametric bootstrap, which is harder to implement in high dimensions. Our simulations suggest that goodness-of-fit tests including GIMTs deteriorate quickly for copulas with dimension $d > 2$ unless the copulas are vines.

4   Power study

In this section we analyze the size and power properties of the new copula goodness-of-fit tests. We start by comparing performance of the various versions of GIMT for vine copulas. This is the case where we believe our tests are paticularly useful in high dimensions. Then, for classical (non-vine) copula specifications, we compare the best performing tests with "blanket" non-GIMT alternatives favored in an extensive simulation study by Genest et al. (2009).

4.1   Comparison Between GIMTs for Vine Copulas

4.1.1   Simulation Setup

We follow the simulation procedure of Schepsmeier (2013) and consider testing the null that the vine copula model is $M_0 = RV(\mathcal{V}_0, \mathcal{B}_0(\mathcal{V}_0), \boldsymbol{\theta}_0(\mathcal{B}_0(\mathcal{V}_0)))$ against the alternative $M_1 = RV(\mathcal{V}_1, \mathcal{B}_1(\mathcal{V}_1), \boldsymbol{\theta}_1(\mathcal{B}_1(\mathcal{V}_1))), M_1 \neq M_0$. In each Monte Carlo simulation

r, we generate n observations on $\boldsymbol{u}^r_{M_0} = (\boldsymbol{u}^{1r}_{M_0}, \ldots, \boldsymbol{u}^{dr}_{M_0})$ from model M_0, estimate the vine copula parameters $\boldsymbol{\theta}_0(\mathcal{B}_0(\mathcal{V}_0))$ and $\boldsymbol{\theta}_1(\mathcal{B}_1(\mathcal{V}_1))$ and calculate the test statistic under the null, $t^r_n(M_0)$, and under the alternative, $t^r_n(M_1)$, for all the tests considered in Section 2. The number of simulations is $B = 5000$.

Then we obtain approximate p-values $\hat{p}_r$ for each test statistic as $\hat{p}_j := \hat{p}(t_j) := 1/B \sum_{r=1}^{B} \mathbf{1}_{\{t_r \geq t_j\}}, j = 1, \ldots, B$ and the actual size $\hat{F}_{M_0}(\alpha)$ and (size-adjusted) power $\hat{F}_{M_1}(\alpha)$ using the formula

$$\hat{F}(\alpha) = \frac{1}{B} \sum_{r=1}^{B} \mathbf{1}_{\{\hat{p}_r \leq \alpha\}}, \qquad \alpha \in (0, 1) \tag{4}$$

We use an R-vine copula with $d = 5$ and $d = 8$ as M_0. As M_1 we use (a) a multivariate Gaussian copula, which can also be represented as a vine, (b) a C-vine copula and (c) a D-vine copula. The details on the copulas under the null and alternatives, as well as on the method used for choosing the specific bivariate components, are provided in Appendix B. All calculations in this section were performed with **R** (R Development Core Team, 2013) and the R-package VineCopula of Schepsmeier et al. (2013).[1]

4.1.2 Simulation Results

We start by assessing the asymptotic approximation of the tests. Figures 2-3 show empirical distributions of the test statistics for two sample sizes, $n = 500$ and 1000. Several observations seem important here. First, overall we observe convergence to the asymptotic distribution even for the fairly high dimensional copulas we consider but asymptotics serve as a very poor approximator in all, except for a few, cases. Second, the sequential approach performs better than the MLE approach – an observation for which we do not have an explanation. Third, the sampling distributions of the Trace White and Determinant IR Tests – one-degree-of-freedom tests – are much closer to their asymptotic limits, regardless of the dimension, than tests with other functional forms and tests with greater degrees of freedom. Fourth, the Determinant White, Log Trace, and Eigenvalue Tests deteriorate quickly as dimension increases. The Trace White and Determinant IR Tests dominate other tests in terms of asymptotic approximation.

Now we look at size-power behavior. Since some of the proposed tests face substantial numerical problems with the asymptotic variance estimation and many exhibit large deviations from the χ^2_r distribution in small samples, especially when

---

[1]The R code used in this section, as well as the Matlab codes used in the next section are available from the authors upon request.
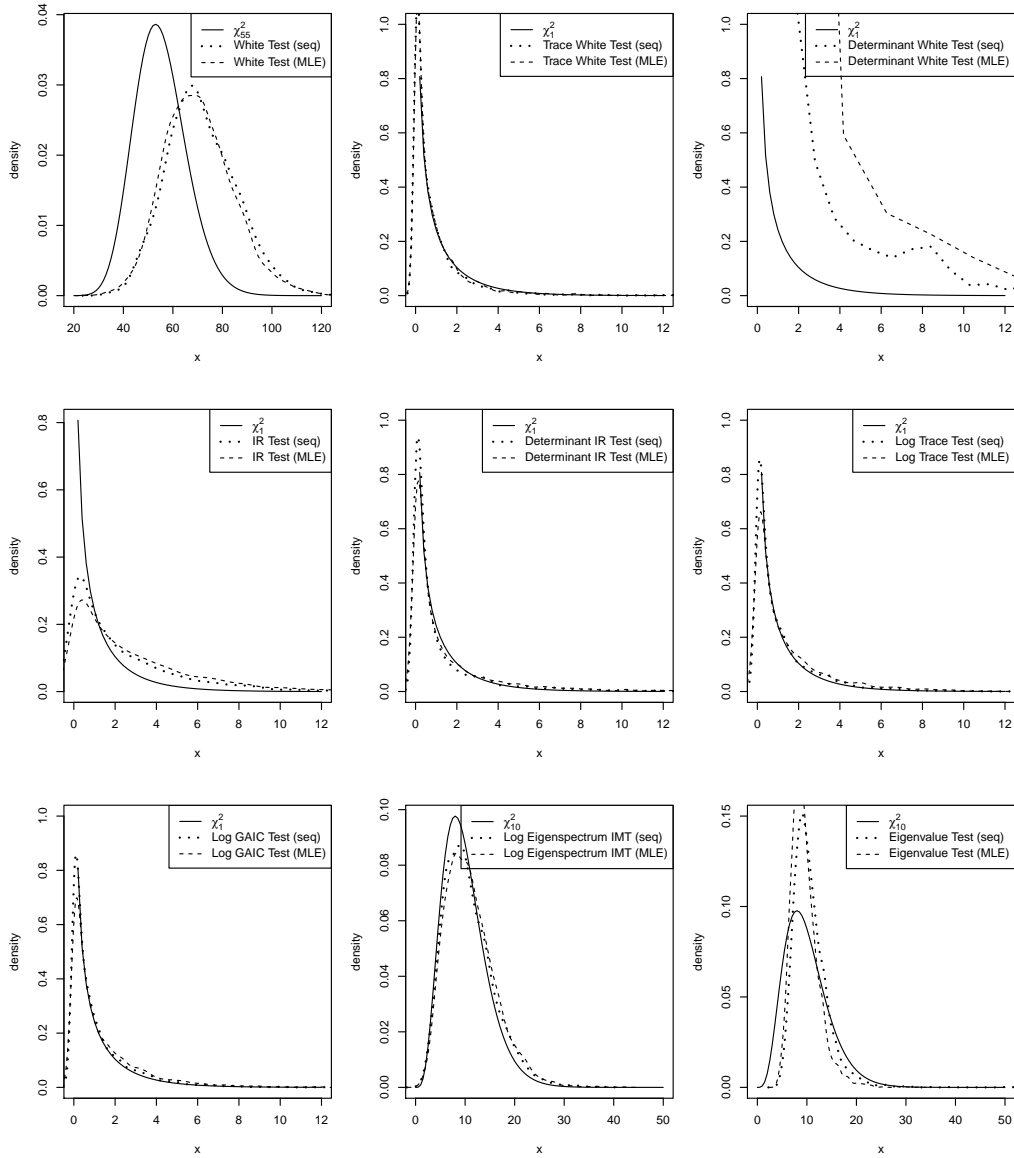
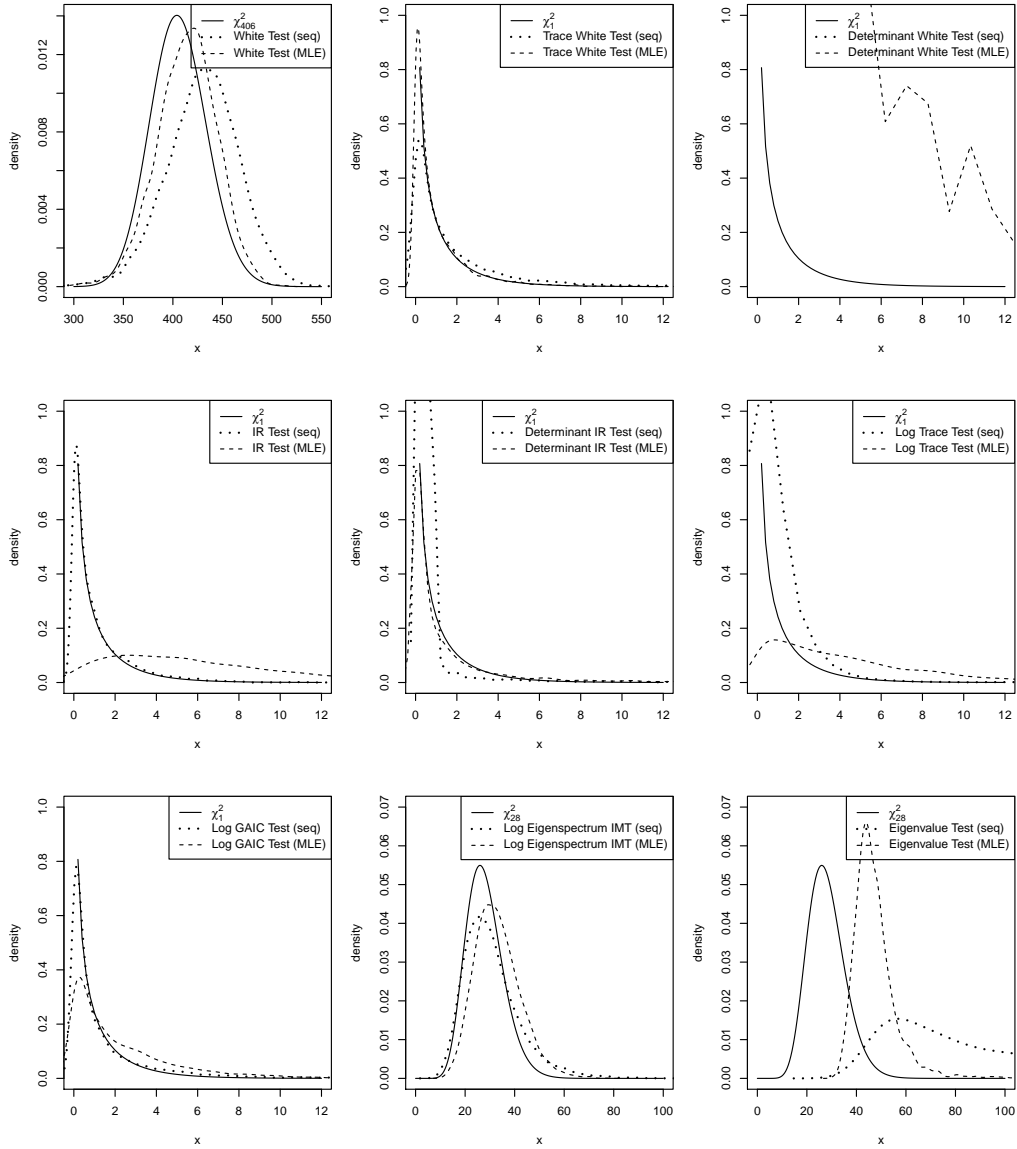Figure 2: Empirical densities of GIMT for R-vine copulas: $d = 5, n = 500$

Figure 3: Empirical densities for GIMT for R-vine copulas: $d = 8, n = 1000$

dimension is high, we only investigate the bootstrap version of the tests. Figures 4-5 illustrate the estimated power of all 9 proposed tests. We consider three dimensions, $d = 5, 8$ and 16; and two versions, sequential (dotted lines) and MLE (solid lines). The two sample sizes we consider are $n = 500$ and 1000 for $d = 5$ and 8; and $n = 1000$ and 5000 for $d = 16$. Percentage of rejections of $\mathcal{H}_0$ is on the y-axis, while the truth (R-vine) and the alternatives are on the x-axis. Obviously, the power is equal to the actual size for the true model. A horizontal black dashed line indicates the 5% nominal size.

All proposed tests maintain their given size independently of the number of sample points, dimension or estimation method. For $d = 5$ we can observe increasing power as sample size increases for all tests except the Determinant White Test. If $d = 8$ the behavoir of the tests, especially the MLE versions, is more erratic. The Determinant White Test seems to be the only test that continues to perform poorly in terms of power when sample size increases. Other tests show improvement in power for either the mle or sequential version or both. Interestingly, the Trace White, Eigenvalue and IR Tests at times show very strong power in one of the two versions (mle or sequential) and no power in the other. Overall, all tests except the Determinant White show power against each alternative, showing that they are consistent.

For $d = 16$ we report only sequential estimates as they were most time efficient. The Log Eigenspectrum, Eigenvalue, IR and Determinant IR tests show consistently good behavior in terms of power against the two alternatives. The power of the Determinant IR and Log Eigenspectrum Tests remains high independent of the dimension or the sample size.

4.2   Comparison with Non-GIMT Tests

4.2.1   Simulation Setup

In this section we compare selected GIMTs for copulas with the original White test $\mathcal{T}_n$ and three "blanket" copula goodness-of-fit tests analyzed by Genest et al. (2009). The GIMTs we select are the Log GAIC Test $\mathcal{G}_n$ and the Eigenvalue Test $\mathcal{Q}_n$ – they showed favorable size and power properties in the simulations of previous sections. The selected non-GIMTs are based on the empirical copula process and the Rosenblatt and Kendall transformation – they showed a favorable size and power behavior in an extensive Monte Carlo study by Genest et al. (2009). We provide details on the three tests in Appendix D and we summarize them in Table 2. For vine copulas such comparisons are provided by Schepsmeier (2015) so in this section we focus on classical multivariate (non-vine) copulas.

Again, since the limiting approximation is poor and depends on an unknown

18
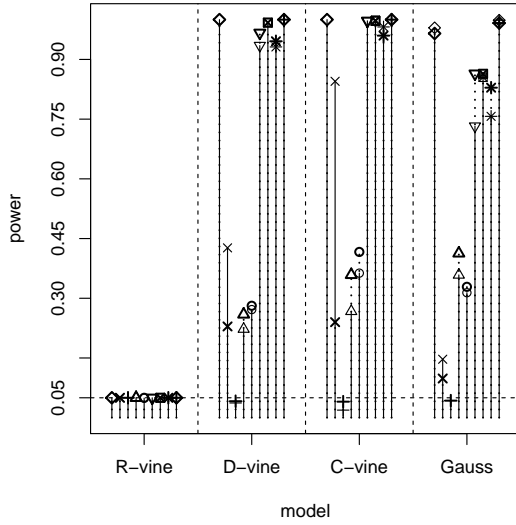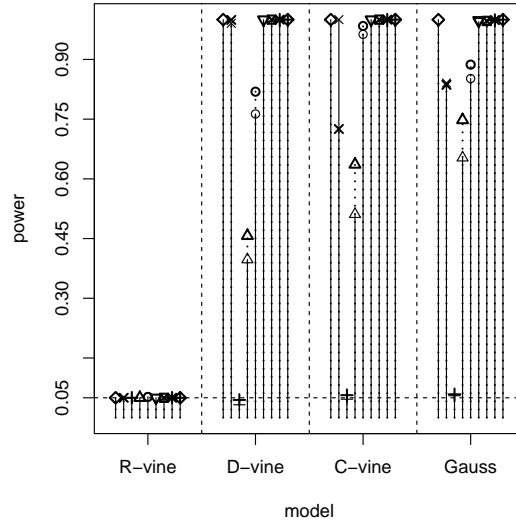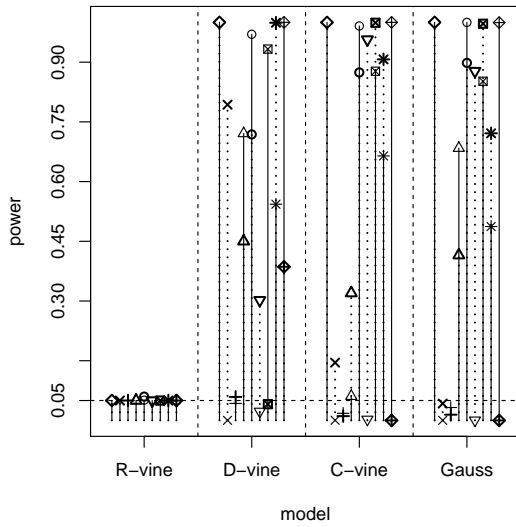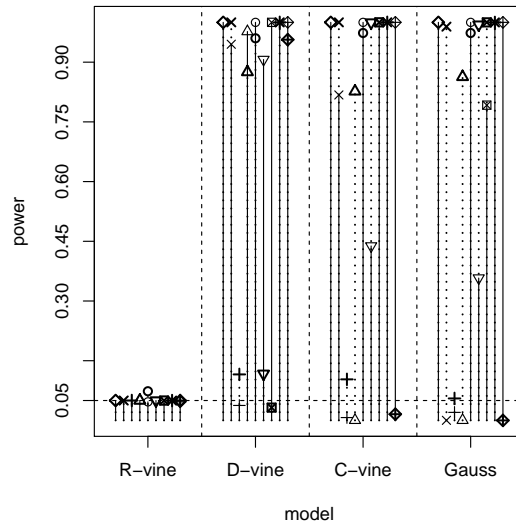
Figure 4: Size and power comparison for bootstrap versions of proposed tests in 5 and 8 dimensions with different sample sizes.
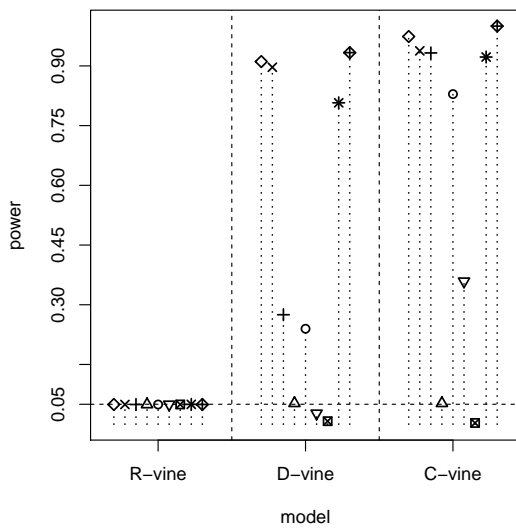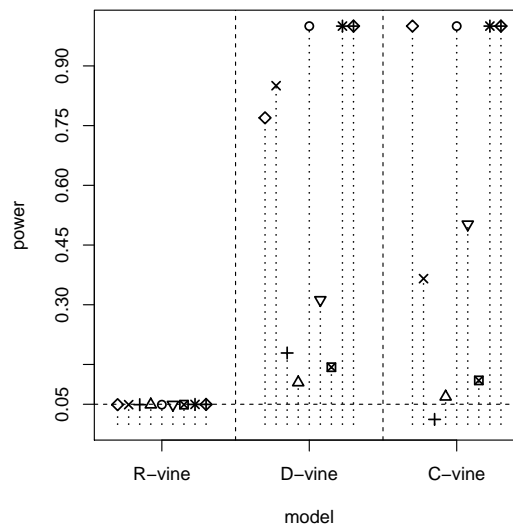
(a) $d = 16, n = 1000$ (b) $d = 16, n = 5000$

Figure 5: Size and power comparison for boostrap versions of proposed tests in 16 dimensions and different sample sizes (only sequential estimates are reported).

Table 2: Summary of non-GIMTs.

Empirical copula process	$\mathcal{S}_n$	$n \int_{[0,1]^d} (C_n(u) - C_{\hat{\boldsymbol{\theta}}_n}(u))^2 \, dC_n(u)$
		$= \sum_{j=1}^n \{C_n(\mathbf{U}_j) - C_{\hat{\boldsymbol{\theta}}_n}(\mathbf{U}_j)\}^2$
Rosenblatt's transform	$\mathcal{S}_n^R$	$\{\mathbf{V}_j = \mathcal{R}_{C_{\hat{\boldsymbol{\theta}}_n}}(\mathbf{U}_j)\}_{j=1}^n$
		$\sum_{j=1}^n \{C_n(\mathbf{V}_j) - C_{\perp}(\mathbf{V}_j)\}^2$
Kendall's transform	$\mathcal{S}_n^K$	$C_{\boldsymbol{\theta}}(\mathbf{U}) \sim K_{\boldsymbol{\theta}}$
		$n \int_{[0,1]} (K_n(v) - K_{\hat{\boldsymbol{\theta}}_n}(v))^2 \, dK\hat{\boldsymbol{\theta}}_n(v)$

parameter $\boldsymbol{\theta}$, we resort to parametric bootstrap to obtain valid p-values. Furthermore, $\boldsymbol{\theta}_0$ and $F_1, \ldots, F_d$ are unknown as before. Therefore we use the pseudo-sample $\{U_{i1}, \ldots, U_{id}\}_{i=1}^n$ to approximate $F_1(X_{i1}), \ldots, F_d(X_{id})$, where $U_{ij} = \frac{R_{ij}}{n+1}$, $i = 1, \ldots, n, j = 1, \ldots, d$, and R_{ij} is the rank of X_{ij} amongst $X_{1j}, \ldots, X_{nj}$. We can use any consistent estimator of $\boldsymbol{\theta}_0$, e.g., the estimator based on Kendall's τ, or the canonical maximum likelihood estimator (CMLE), which maximizes the pseudo-likelihood $\ell(\boldsymbol{\theta}) = \sum_{i=1}^n \ln c_{\boldsymbol{\theta}}(\mathbf{U}_i)$, where $\mathbf{U}_i = (U_{i1}, \ldots, U_{id})$. In this section, we use the estimator based on Kendall's τ in all bivariate and multivariate cases except for tests involving the Outer Power Clayton copula, for which the estimator based on Kendall's τ is not feasible. For details see Appendix C.

4.2.2 Simulation Results

We report selected size and power results in tables similar to those reported by Genest et al. (2009) and Huang and Prokhorov (2014). The point of the tables is to examine the effect of the sample size, degree of dependence and dimension on size and power of the seven tests. The nominal level is fixed at 5% as before.

We first report bivariate results for selected values of Kendall's τ and four one-parameter copula families, where we obtain an estimate of the parameter by inverting the sample version of Kendall's τ. The results are based on 1,000 random samples of size $n = 150$ and 500. Tables 3 and 4 report the size and power results for $n = 150$ and Kendall's τ equal 0.5 and 0.75, respectively. Table 5 reports the results for $n = 500$ and Kendall's $\tau = 0.5$. In each row we report the percentage of rejections of $\mathcal{H}_0$ associated with $\mathcal{S}_n$, $\mathcal{S}_n^R$, $\mathcal{S}_n^K$, $\mathcal{T}_n$ and $\mathcal{Q}_n$. As an example, Table 3 shows that when testing the null of the Gaussian copula using $\mathcal{Q}_n$ and $n = 150$, we reject the null about 42% of the time when the true copula was Gumbel with Kendall's $\tau = 0.5$. For all tests, except $\mathcal{T}_n$, we bootstrap critical values. We use analytical values for $\mathcal{T}_n$ to show that the conventional version of IMT is badly oversized (more comparisons including bootstrap $\mathcal{T}_n$ can be found in Huang and Prokhorov (2014)).

Table 3: Percentage of rejections of $\mathcal{H}_0$ by various tests for sample size $n = 150$ arising from different copula models with $d = 2$ and Kendall's $\tau = 0.50$

Copula under $\mathcal{H}_0$	True copula	Test based on				
		$\mathcal{S}_n$	$\mathcal{S}_n^R$	$\mathcal{S}_n^K$	$\mathcal{T}_n$	$\mathcal{Q}_n$
Gaussian	Gaussian	**4.9**	**5.0**	**4.9**	**7.5**	**4.0**
	Frank	20.2	13.4	17.4	6.8	36.0
	Clayton	80.0	90.8	90.3	30.8	70.5
	Gumbel	38.3	42.0	16.1	15.4	42.0
Frank	Gaussian	19.9	8.9	22.6	14.6	2.0
	Frank	**4.8**	**4.8**	**4.8**	**9.4**	**4.0**
	Clayton	89.1	86.9	98.6	5.7	10.1
	Gumbel	63.0	44.1	28.3	5.3	12.0
Clayton	Gaussian	93.7	89.0	75.1	80.6	34.2
	Frank	95.7	94.4	89.5	90.2	70.0
	Clayton	**5.3**	**5.1**	**4.5**	**12.0**	**4.9**
	Gumbel	99.9	99.7	98.5	90.5	54.2
Gumbel	Gaussian	18.3	33.7	37.7	5.2	8.0
	Frank	39.8	52.1	42.4	29.3	37.6
	Clayton	99.6	99.7	99.9	75.5	78.8
	Gumbel	**4.6**	**4.5**	**4.6**	**10.0**	**4.4**

The results indicate that all the tests maintain the nominal size and generally have power against the alternatives. We note that in the bivariate case we use only one indicator in constructing $\mathcal{T}_n$ and so $\mathcal{Q}_n$ provides no dimension reduction. The analytical p-values used for $\mathcal{T}_n$ lead to noticeable oversize distortions, while $\mathcal{Q}_n$ retains size close to nominal and is often conservative compared with $\mathcal{S}_n$, $\mathcal{S}_n^R$, and $\mathcal{S}_n^K$. The tables also show that a higher dependence or a larger sample size give higher power, which is true for all the tests we consider. The increase in power resulting from the sample size increase is an indication of $\mathcal{Q}_n$ being consistent.

Table 6 presents selected results for $d = 4$. Here we focus on $\mathcal{S}_n$, $\mathcal{T}_n$ and $\mathcal{Q}_n$ but report two versions of $\mathcal{T}_n$, one based on bootstrapped critical values ($\mathcal{T}_n^b$) and the other based on the analytical asymptotic critical values ($\mathcal{T}_n^a$) – this high dimensional comparison was not considered by Huang and Prokhorov (2014). We do not include

Table 4: Percentage of rejections of $\mathcal{H}_0$ by various tests for sample size $n = 150$ arising from different copula models with $d = 2$ and Kendall's $\tau = 0.75$

Copula under $\mathcal{H}_0$	True copula	Test based on				
		$\mathcal{S}_n$	$\mathcal{S}_n^R$	$\mathcal{S}_n^K$	$\mathcal{T}_n$	$\mathcal{Q}_n$
Gaussian	Gaussian	**4.9**	**4.9**	**4.4**	**10.4**	**5.3**
	Frank	42.2	32.8	41.4	40.0	86.6
	Clayton	91.8	99.9	97.3	60.5	99.2
	Gumbel	38.5	55.5	17.9	23.7	71.2
Frank	Gaussian	40.9	18.4	40.2	8.0	3.6
	Frank	**4.7**	**5.0**	**4.5**	**11.0**	**5.3**
	Clayton	96.6	99.7	99.6	20.4	7.2
	Gumbel	81.9	59.9	53.2	8.7	2.2
Clayton	Gaussian	99.8	99.5	94.9	90.2	66.4
	Frank	99.1	99.9	97.0	91.6	99.8
	Clayton	**5.4**	**5.1**	**4.9**	**11.0**	**4.2**
	Gumbel	99.9	99.9	99.9	96.2	97.2
Gumbel	Gaussian	12.3	60.7	29.4	9.6	4.6
	Frank	51.7	83.8	61.6	76.4	89.2
	Clayton	99.9	99.9	99.9	90.4	100.0
	Gumbel	**4.5**	**5.2**	**4.4**	**10.9**	**5.0**

$\mathcal{S}_n^R$ and $\mathcal{S}_n^K$ because their behavior appears similar to that of $\mathcal{S}_n$. Under the null, we have three one-parameter Archimedean copulas, the Gaussian copula with six distinct parameters in the correlation matrix and the Outer Power Clayton copula with two parameters. The alternatives are six four-dimensional copula families. We did not include the Student-t copula under the null (but include it under the alternative) due to the heavy computational burden associated with generating from this copula. All the other true copulas are also considered under the null.

Several observations are unique to the multivariate simulations because they involve more than one parameter and more than two marginals. To simulate from the Outer Power Clayton copula, which has two parameters, we set $(\beta, \theta) = (4/3, 1)$, which corresponds to Kendall's τ equal 0.5. For the Gaussian copula, after estimating the pairwise Kendall τ's, we invert them to obtain the corresponding elements of the correlation matrix. For the Archimedean copulas, we follow Berg (2009) and

Table 5: Percentage of rejections of $\mathcal{H}_0$ by various tests for sample size $n = 500$ arising from different copula models with $d = 2$ and Kendall's $\tau = 0.50$

Copula under $\mathcal{H}_0$	True copula	Test based on				
		$\mathcal{S}_n$	$\mathcal{S}_n^R$	$\mathcal{S}_n^K$	$\mathcal{T}_n$	$\mathcal{Q}_n$
Gaussian	Gaussian	**4.6**	**5.4**	**4.9**	**7.5**	**4.0**
	Frank	36.9	60.7	33.4	60.7	66.5
	Clayton	99.8	100.0	99.6	90.4	99.5
	Gumbel	65.3	18.9	62.9	62.3	71.1
Frank	Gaussian	42.5	35.1	32.7	20.6	15.2
	Frank	**4.2**	**6.4**	**4.7**	**7.1**	**4.8**
	Clayton	100.0	99.9	100.0	10.6	15.1
	Gumbel	95.2	47.5	85.8	13.3	14.9
Clayton	Gaussian	100.0	99.5	99.7	100.0	99.0
	Frank	100.0	99.4	99.9	99.2	99.9
	Clayton	**5.0**	**5.2**	**4.7**	**12.0**	**4.9**
	Gumbel	100.0	100.0	100.0	99.5	100.0
Gumbel	Gaussian	74.1	38.4	61.6	20.7	21.2
	Frank	95.5	47.8	85.1	89.3	99.2
	Clayton	100.0	100.0	100.0	100.0	100.0
	Gumbel	**5.2**	**5.5**	**5.0**	**7.2**	**4.4**

obtain the dependence parameter by inverting the average of six pairwise Kendall τ's. For the Outer Power Clayton copula, we can only estimate the parameters by CMLE. Details on simulating from and estimation of the Outer Power Clayton copula can be found in Hofert et al. (2012). For a given value of τ and each combination of copulas under the null and under the alternative, the results we report are based on 1,000 random samples of size $n = 150$. Each of these samples is then used to test goodness-of-fit. Table 6 reports size and power for (average) Kendall's τ equal 0.5.

The key observation from Table 6 is that $\mathcal{Q}_n$ dominates both versions of $\mathcal{T}_n$ in terms of power. We attribute this to the dimension reduction permitted by $\mathcal{Q}_n$. The table also shows that our test maintains the nominal size of 5% in the multivariate cases. Overall, the behavior of $\mathcal{Q}_n$ is as good if not better than that of $\mathcal{S}_n$. A remarkable case of the better performance of $\mathcal{Q}_n$ is the tests involving the Student-t alternative, where $\mathcal{S}_n$ does worse, regardless of the copula under the null.

Table 6: Percentage of rejections of $\mathcal{H}_0$ for $n = 150$ and $d = 4$ with Kendall's $\tau = 0.50$.

Copula under $\mathcal{H}_0$	True copula	Test based on			
		$\mathcal{S}_d$	$\mathcal{T}_n^a$	$\mathcal{T}_n^b$	$\mathcal{Q}_n$
Gaussian	Gaussian	**5.0**	**4.9**	**5.0**	**4.9**
	Frank	15.4	4.7	6.5	56.1
	Clayton	88.5	14.4	10.2	72.5
	Gumbel	52.1	12.1	13.6	75.5
	Student	11.3	14.6	7.0	90.4
	Outer Power Clayton	60.2	13.9	11.4	72.4
Frank	Gaussian	43.4	16.3	19.6	47.8
	Frank	**4.2**	**7.3**	**5.3**	**4.9**
	Clayton	97.0	14.5	7.1	27.3
	Gumbel	67.3	7.0	4.5	25.6
	Student	56.7	77.3	50.5	80.9
	Outer Power Clayton	77.6	8.2	13.1	42.7
Clayton	Gaussian	92.2	99.4	42.6	98.8
	Frank	94.1	99.9	38.1	99.9
	Clayton	**5.1**	**10.3**	**4.2**	**4.7**
	Gumbel	99.3	99.9	55.4	99.8
	Student	96.7	98.5	50.8	96.9
	Outer Power Clayton	70.3	50.6	12.5	75.8
Gumbel	Gaussian	76.3	49.8	20.2	83.4
	Frank	60.1	33.8	16.9	76.1
	Clayton	99.4	99.6	82.6	99.9
	Gumbel	**5.0**	**6.5**	**5.2**	**5.1**
	Student	77.5	79.0	30.3	93.2
	Outer Power Clayton	89.7	50.9	22.3	78.5
Outer Power Clayton	Gaussian	62.8	14.6	6.7	18.4
	Frank	60.1	20.2	9.1	45.1
	Clayton	9.4	8.9	9.0	11.1
	Gumbel	25.4	13.5	8.1	20.9
	Student	19.5	8.4	7.9	75.7
	Outer Power Clayton	**5.3**	**7.7**	**5.0**	**4.8**

Table 7: Percentage of rejections of $\mathcal{H}_0$ for $n = 150$ and $d = 5$ with Kendall's $\tau = 0.50$.

Copula under $\mathcal{H}_0$	True copula	Test based on			
		$\mathcal{S}_d$	$\mathcal{Q}_n$	$\mathcal{T}_n^b$	$\mathcal{G}_n$
Gaussian	Gaussian	**5.1**	**4.8**	**5.0**	**5.0**
	Frank	15.2	63.4	7.1	50.6
	Clayton	93.8	76.9	17.7	71.2
	Gumbel	52.3	74.6	12.4	62.5
	Student	9.1	92.6	7.6	90.1
	Outer Power Clayton	61.7	74.7	13.5	57.5
Frank	Gaussian	60.4	61.4	21.3	51.7
	Frank	**5.0**	**4.9**	**5.1**	**4.9**
	Clayton	98.3	34.6	8.3	30.5
	Gumbel	69.7	20.1	4.1	19.2
	Student	64.2	51.8	60.4	56.4
	Outer Power Clayton	75.4	77.3	13.9	80.1
Clayton	Gaussian	91.4	98.1	50.4	92.0
	Frank	89.9	99.2	38.9	99.4
	Clayton	**4.9**	**4.9**	**5.0**	**4.9**
	Gumbel	97.5	99.9	59.5	99.8
	Student	97.1	98.1	55.4	98.9
	Outer Power Clayton	72.6	74.1	17.6	64.3
Gumbel	Gaussian	81.0	86.5	24.9	85.4
	Frank	67.5	77.4	20.7	82.0
	Clayton	99.3	99.9	83.4	99.9
	Gumbel	**5.1**	**5.0**	**5.1**	**5.1**
	Student	74.2	90.4	40.2	76.5
	Outer Power Clayton	91.1	80.5	30.5	62.1
Outer Power Clayton	Gaussian	60.2	17.3	8.2	12.8
	Frank	60.6	51.6	17.4	41.3
	Clayton	7.5	11.3	10.2	15.9
	Gumbel	26.7	21.7	13.1	17.8
	Student	5.2	76.4	10.4	63.7
	Outer Power Clayton	**5.3**	**5.0**	**4.9**	**5.0**

26

Table 8: Percentage of rejections of $\mathcal{H}_0$ for $n = 150$ and $d = 8$ with Kendall's $\tau = 0.50$.

Copula under $\mathcal{H}_0$	True copula	Test based on			
		$\mathcal{S}_n$	$\mathcal{Q}_n$	$\mathcal{T}_n^b$	$\mathcal{G}_n$
Gaussian	Gaussian	**5.0**	**4.8**	**5.0**	**5.0**
	Frank	25.6	86.3	22.5	81.5
	Clayton	98.7	91.2	29.6	93.8
	Gumbel	75.5	87.2	36.1	90.5
	Student	12.2	99.9	18.9	99.9
	Outer Power Clayton	75.4	95.6	39.2	82.7
Frank	Gaussian	97.8	87.9	32.3	82.2
	Frank	**4.9**	**4.9**	**5.0**	**4.9**
	Clayton	99.5	60.2	19.4	42.2
	Gumbel	85.6	32.4	9.8	29.3
	Student	99.5	79.8	64.4	82.3
	Outer Power Clayton	91.4	93.7	42.3	96.7
Clayton	Gaussian	99.7	99.9	75.4	99.9
	Frank	97.9	100	62.2	99.9
	Clayton	**4.9**	**4.9**	**5.0**	**5.0**
	Gumbel	99.9	99.9	82.3	99.9
	Student	99.9	99.9	65.2	99.9
	Outer Power Clayton	81.1	95.8	34.6	81.6
Gumbel	Gaussian	99.5	98.9	42.1	97.5
	Frank	63.4	81.9	40.3	85.1
	Clayton	100	99.9	99.0	99.9
	Gumbel	**5.2**	**5.0**	**5.1**	**5.1**
	Student	99.5	99.5	54.2	90.1
	Outer Power Clayton	99.9	99.9	42.2	82.1
Outer Power Clayton	Gaussian	67.6	38.2	33.4	20.7
	Frank	71.4	54.1	16.2	42.9
	Clayton	14.2	12.5	11.7	16.6
	Gumbel	45.3	28.4	32.3	35.8
	Student	18.6	97.6	52.4	67.9
	Outer Power Clayton	**5.0**	**5.1**	**5.3**	**5.0**

An interesting observation is how the power of $\mathcal{Q}_n$ changes between Table 3 and Table 6. Consider, for example, the test of the null of the Frank copula. Regardless of the alternative, $\mathcal{Q}_n$ performs poorly in the bivariate case. However, with the increased dimension the behavior of $\mathcal{Q}_n$ improves substantially. This is especially pronounced in comparison with $\mathcal{T}_n$, whose power remains particularly low against the Archimedean alternatives. At the same time, for the Student-t and Gaussian alternatives, the performance of $\mathcal{Q}_n$ stands out even compared with $\mathcal{S}_n$.

Table 7 and Table 8 present selected results for $d = 5$ and $d = 8$, respectively. Here we focus on $\mathcal{S}_d$, $\mathcal{Q}_n$, $\mathcal{T}_n$ and $\mathcal{G}_n$. We use $\mathcal{T}_n$ (bootstrap) as a benchmark. The Log GAIC Test $\mathcal{G}_n$ is another GIMT that performed well in Section 4.1 – we use it to further illustrate the dimension reduction permitted by GIMTs. In Tables 7 and 8, under the null we have three one-parameter Achimedean copulas, the Outer Power Clayton copula with two parameters, and the Gaussian copula with $\frac{d(d-1)}{2}$ distinct parameters in the correlation matrix. The alternatives are Frank, Clayton, Gumbel, Outer Power Clayton, Gaussian, and t copulas. Samples are in every scenario are simulated from a copula with Kendall's τ equal to 0.5. The parameter estimation here is done by CMLE, rather than by conversion of Kendall's τ used for $d = 4$ in Table 7. The explicit expressions of the score functions of the selected Archimedean copulas can be found in Hofert et al. (2012).

The results in Tables 7-8 show that, as expected, $\mathcal{Q}_n$, $\mathcal{G}_n$ and $\mathcal{T}_n$ all maintain the nominal size and show power. More interestingly, the power of the three GIMT tests increases as the dimension increases. In particular, $\mathcal{Q}_n$ and $\mathcal{G}_n$ behave similarly under all null hypotheses and both show significant increases in power in almost all scenarios as the dimension grows. We also see that $\mathcal{Q}_n$ and $\mathcal{G}_n$ dominate $\mathcal{T}_n$ in all scenarios. Note that for the Frank, Clayton, and Gumbel copulas, both Hessian and OPG matrices degenerate to scalars; therefore there is no dimension reduction in $\mathcal{Q}_n$ and $\mathcal{G}_n$ compared to $\mathcal{T}_n$. Yet, we observe that $\mathcal{Q}_n$ and $\mathcal{G}_n$ are more powerful than $\mathcal{T}_n$, which may be due to the fact that the eigenvalues of $-\mathbb{H}^{-1}\mathbb{C}$ are more sensitive to changes in $\mathbb{H}$ and $\mathbb{C}$ than the eigenvalues of $\mathbb{H} + \mathbb{C}$. When testing multi-parameter copulas, e.g., multivariate Gaussian, due to the additional dimension reduction, $\mathcal{Q}_n$ and $\mathcal{G}_n$ perform much better than $\mathcal{T}_n$.

5   Conclusion

We consider a battery of tests resulting from eigenspectrum-based versions of the information matrix equality applied to copulas. The benefit of this generalization is due to a reduction in degrees of freedom of the tests and to the focused hypothesis function used to construct them. For example, in testing the validity of

high-dimension, multi-parameter copulas we manage to reduce the information matrix based test statistic to an asymptotically χ^2 with one degree of freedom, and we can focus on the effect of larger or smaller eigenvalues by using specific functions of the eigenspectrum such as *det* or *trace*. However, only a few of the proposed tests can be well approximated by their asymptotic distributiions in realistic sample sizes so we have also looked at the boostrap version of the tests.

The main argument of the paper is that the bootstrap versions of GIMTs dominate other available tests of copula validity when copulas are high-dimensional and multi-parameter. We use this argument to motivate the use of GIMTs on vine copulas, where additional simplifications result from the functional form of the Hessian and the score.

References

AAS, K., C. CZADO, A. FRIGESSI, AND H. BAKKEN (2009): "Pair-copula construction of multiple dependence," *Insurance: Mathematics and Economics*, 44, 182–198.

BEDFORD, T. AND R. M. COOKE (2001): "Probability density decomposition for conditionally dependent random variables modeled by vines." *Ann. Math. Artif. Intell.*, 32, 245–268.

——— (2002): "Vines–a new graphical model for dependent random variables," *The Annals of Statistics*, 30, 1031–1068.

BERG, D. (2009): "Copula goodness-of-fit testing: an overview and power comparison," *The European Journal of Finance*, 15, 675–701.

BRECHMANN, E., C. CZADO, AND K. AAS (2012): "Truncated Regular Vines in High Dimensions with Applications to Financial Data," *Canadian Journal of Statistics*, 40, 68–85.

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

CHEN, X. AND Y. FAN (2006): "Estimation of copula-based semiparametric time series models," *Journal of Econometrics*, 130, 307–335.

CZADO, C. (2010): "Pair-Copula Constructions of Multivariate Copulas," in *Copula Theory and Its Applications, Lecture Notes in Statistics*, ed. by Jaworski, P. and Durante, F. and Härdle, W.K. and Rychlik, T, Berlin Heidelberg: Springer-Verlag, vol. 198, 93–109.

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

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 estimatiion procedure of dependence parameters in multivariate families of distributions," *Biometrika*, 82, 543–552.

GENEST, C., J.-F. QUESSY, AND B. REMILLARD (2006): "Goodness-of-fit Procedures for Copula Models 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 semi-parametric models," *Annales de l'Institut Henri Poincare - Probabilites et Statistiques*, 44, 1096–1127.

GENEST, C., B. RÉMILLARD, AND D. BEAUDOIN (2009): "Goodness-of-fit tests for copulas: A review and a power study," *Insurance: Mathematics and Economics*, 44, 199–213.

GENEST, C. AND L.-P. RIVEST (1993): "Statistical inference procedures for bivariate Archimedean copulas," *Journal of the American Statistical Association*, 88, 1034–1043.

GOLDEN, R., S. HENLEY, H. WHITE, AND T. M. KASHNER (2013): "New Directions in Information Matrix Testing: Eigenspectrum Tests," in *Recent Advances and Future Directions in Causality, Prediction, and Specification Analysis*, ed. by X. Chen and N. R. Swanson, Springer New York, 145–177.

HOFERT, M., M. MACHLER, AND A. J. MCNEIL (2012): "Likelihood inference for Archimedean copulas in high dimensions under known margins," *Journal of Multivariate Analysis*, 110, 133–150.

HUANG, W. AND A. PROKHOROV (2014): "A goodness-of-fit test for copulas," *Econometric Reviews*, 98, 533–543.

JOE, H. (1996): "Families of m-variate distributions with given margins and m(m-1)/2 bivariate dependence parameters," in *Distributions with Fixed Marginals and Related Topics*, ed. by L. Rüschendorf and B. Schweizer and M. D. Taylor, Hayward, CA: Inst. Math. Statist., vol. 28, 120–141.

——— (2005): "Asymptotic efficiency of the two-stage estimation method for copula-based models," *Journal of Multivariate Analysis*, 94, 401–419.

JUNKER, M. AND A. MAY (2005): "Measurement of aggregate risk with copulas," *The Econometrics Journal*, 8, 428–454.

KLUGMAN, S. AND R. PARSA (1999): "Fitting bivariate loss distributions with copulas," *Insurance: Mathematics and Economics*, 24, 139–148.

KOLLO, T. AND D. VON ROSEN (2006): *Advanced Multivariate Statistics with Ma- trices. Mathematics and Its Applications*, Springer.

KUROWICKA, D. AND R. M. COOKE (2006): *Uncertainty Analysis with High Dimensional Dependence Modelling*, John Wiley & Sons Ltd, Chichester.

LEEUW, J. D. (2007): "Derivatives of generalized eigen systems with applications," Tech. rep., Center for environmental statistics, Department of Statistics, University of California, Los Angeles, CA.

MAGNUS, J. (1985): "On differentiating eigenvalues and eigenvectors," *Econometric Theory*, 1, 179–191.

MAGNUS, J. AND H. NEUDECKER (1999): *Matrix Differential Calculus with Applications in Statistics and Econometrics*, John Wiley & Sons Ltd, Chichester.

PATTON, A. J. (2012): "A Review of Copula Models for Economic Time Series," *J. Multivar. Anal.*, 110, 4–18.

PRESNELL, B. AND D. D. BOOS (2004): "The IOS Test for Model Misspecification," *Journal of the American Statistical Association*, 99, 216–227.

PROKHOROV, A. AND P. SCHMIDT (2009): "Likelihood-based estimation in a panel setting: Robustness, redundancy and validity of copulas," *Journal of Econometrics*, 153, 93–104.

R DEVELOPMENT CORE TEAM (2013): *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0.

ROSENBLATT, M. (1952): "Remarks on a Multivariate Transformation," *The Annals of Mathematical Statistics*, 23, 470–472.

SCAILLET, O. (2007): "Kernel Based Goodness-of-Fit Tests for Copulas with Fixed Smoothing Parameters," *Journal of Multivariate Analysis*, 98, 533–543.

SCHEPSMEIER, U. (2013): "A goodness-of-fit test for regular vine copula models," *Arxiv working paper*.

——— (2015): "Efficient goodness-of-fit tests in multi-dimensional vine copula models," *Journal of Multivariate Analysis*, 138, 35–52, high-Dimensional Dependence and Copulas.

SCHEPSMEIER, U. AND J. STÖBER (2012): "Web supplement: Derivatives and Fisher information of bivariate copulas," Tech. rep., TU München, available online: https://mediatum.ub.tum.de/node?id=1119201.

——— (2014): "Derivatives and Fisher information of bivariate copulas," *Statistical Papers*, 55, 525–542.

SCHEPSMEIER, U., J. STOEBER, AND E. C. BRECHMANN (2013): *VineCopula: Statistical inference of vine copulas*, r package version 1.2.

SERFLING, R. (1980): *Approximation Theorems of Mathematical Statistics*, New York: John Wiley & Sons.

STÖBER, J. AND U. SCHEPSMEIER (2013): "Estimating standard errors in regular vine copula models," *Computational Statistics*, 28, 2679–2707.

TAKEUCHI, K. (1976): "Distribution of Information statistics and a criterion of model fitting for adequacy of models," *Mathematical Sciences*, 153, 12–18.

WHITE, H. (1982): "Maximum Likelihood Estimation of Misspecified Models," *Econometrica*, 50, 1–25.

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, 205–213.

ZIMMER, D. (2012): "The role of copulas in the housing crisis," *Review of Economics and Statistics*, 94, 607–620.

A Proofs

Proof of Lemma 1: The proof is based on combining the results of Golden et al. (2013) and Huang and Prokhorov (2014). It also relates to the work of Presnell and Boos (2004) on information ratio test.

We start with $d = 2$ for simplicity and later give the formulas for any d. Let

$$\boldsymbol{d}_i(\boldsymbol{\theta}) := \begin{pmatrix} vech(\mathbb{H}_i(\boldsymbol{\theta})) \\ vech(\mathbb{C}_i(\boldsymbol{\theta})) \end{pmatrix} \in \mathbb{R}^{p(p+1)}$$

denote the lower triangle vectorizations of $-\mathbb{H}_i(\boldsymbol{\theta})$ and $\mathbb{C}_i(\boldsymbol{\theta})$ and let $\nabla D_{\boldsymbol{\theta}} := \mathbb{E}\nabla_{\boldsymbol{\theta}}\boldsymbol{d}_i(\boldsymbol{\theta}) \in \mathbb{R}^{p(p+1)\times p}$ denote the expected Jacobian matrix of the random vector $\boldsymbol{d}_i(\boldsymbol{\theta})$. We can estimate $\mathbb{E}\boldsymbol{d}_i(\boldsymbol{\theta}_0)$ by $\bar{\boldsymbol{d}}(\hat{\boldsymbol{\theta}}_n)$, where $\bar{\boldsymbol{d}}(\boldsymbol{\theta}) := \frac{1}{n}\sum_{i=1}^n \boldsymbol{d}_i(\boldsymbol{\theta})$.

Let $\hat{F}_{ji} = \hat{F}_j(x_{ji})$, $j = 1, 2$, $i = 1, \ldots, n$, be the empirical cdf's. Note that $\boldsymbol{d}_i(\boldsymbol{\theta})$ implicitely depends on the nonparametric estimates of the marginals, $\hat{F}_1(x_1)$ and $\hat{F}_2(x_2)$. Then,

$$\boldsymbol{d}_i(\boldsymbol{\theta}) = \left\{ vech[\nabla_\theta^2 \ln c(\hat{F}_{1i}, \hat{F}_{2i}; \boldsymbol{\theta})]', vech[\nabla_\theta \ln c(\hat{F}_{1i}, \hat{F}_{2i}; \theta)\nabla_\theta' \ln c(\hat{F}_{1i}, \hat{F}_{2i}; \boldsymbol{\theta})]' \right\}'.$$

Provided that the derivatives and expectation exist, let

$$\nabla D_\theta = \mathbb{E}\nabla_\theta \boldsymbol{d}_i(\boldsymbol{\theta})$$

and

$$\nabla \bar{D}_\theta = n^{-1}\sum_{i=1}^n \nabla_\theta \boldsymbol{d}_i(\boldsymbol{\theta}).$$

First, expand $\sqrt{n}\bar{\boldsymbol{d}}_{\hat{\theta}}$ with respect to $\boldsymbol{\theta}$:

$$\sqrt{n}\bar{\boldsymbol{d}}(\hat{\boldsymbol{\theta}}) = \sqrt{n}\bar{\boldsymbol{d}}(\boldsymbol{\theta}_0) + \nabla D_{\theta_o}\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) + o_p(1).$$

Chen and Fan (2006) show that

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_o) \to N(0, B^{-1}GB^{-1}),$$

where

$$B = -\mathbb{H}(\boldsymbol{\theta}_0),$$
$$G = \lim_{n\to\infty} Var(\sqrt{n}A_n^*),$$
$$A_n^* = \frac{1}{n}\sum_{i=1}^n (\nabla_\theta \ln c(F_{1i}, F_{2i}; \boldsymbol{\theta}_0) + W_1(F_{1i}) + W_2(F_{2i})).$$

Here terms $W_1(F_{1i})$ and $W_2(F_{2i})$ are the adjustments needed to account for the empirical distributions used in place of the true distributions. These terms are calculated as follows:

$$W_1(F_{1i}) = \int_0^1 \int_0^1 [I\{F_{1i} \leq u\} - u] \nabla_{\theta,u}^2 \ln c(u,v;\boldsymbol{\theta}_0)\, c(u,v;\boldsymbol{\theta}_0) \mathrm{d}u\mathrm{d}v,$$

$$W_2(F_{2i}) = \int_0^1 \int_0^1 [I\{F_{2i} \leq v\} - v] \nabla_{\theta,v}^2 \ln c(u,v;\boldsymbol{\theta}_0)\, c(u,v;\boldsymbol{\theta}_0) \mathrm{d}u\mathrm{d}v.$$

So, rewriting the consistency result from Chen and Fan (2006) we have

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) = B^{-1}\sqrt{n}A_n^* + o_p(1).$$

Second, expand $\sqrt{n}\bar{\boldsymbol{d}}(\boldsymbol{\theta}_0)$ with respect to F_1 and F_1:

$$\sqrt{n}\bar{\boldsymbol{d}}(\boldsymbol{\theta}_0) \simeq \frac{1}{\sqrt{n}}\sum_{i=1}^n \boldsymbol{d}_i(\boldsymbol{\theta}_0) + \frac{1}{n}\sum_{i=1}^n \nabla_{F_1}\boldsymbol{d}_i(\boldsymbol{\theta}_0)\sqrt{n}(\hat{F}_{1i}-F_{1i}) + \frac{1}{n}\sum_{i=1}^n \nabla_{F_2}\boldsymbol{d}_i(\boldsymbol{\theta}_0)\sqrt{n}(\hat{F}_{2i}-F_{2i}).$$

(5)

Under suitable regularity conditions (see, e.g., Genest et al., 1995; Chen and Fan, 2006),

$$\frac{1}{n}\sum_{i=1}^n \nabla_{F_{1i}}\boldsymbol{d}_i(\boldsymbol{\theta}_0)\sqrt{n}(\hat{F}_{1i} - F_{1i})$$

$$\simeq \int_0^1 \int_0^1 \nabla_u \{vech[\nabla_\theta^2 \ln c(u,v;\boldsymbol{\theta}_0)]', vech[\nabla_\theta \ln c(u,v;\boldsymbol{\theta}_0)\nabla_\theta' \ln c(u,v;\boldsymbol{\theta}_0)]'\}'$$

$$\sqrt{n}(\hat{F}_1(F_1^{-1}(u)) - u)c(u,v;\boldsymbol{\theta}_0)\mathrm{d}u\mathrm{d}v$$

$$= \frac{1}{\sqrt{n}}\sum_{i=1}^n \int_0^1 \int_0^1 [I\{F_{1i} \leq u\} - u]$$

$$\nabla_u\{vech[\nabla_\theta^2 \ln c(u,v;\boldsymbol{\theta}_0)]', vech[\nabla_\theta \ln c(u,v;\boldsymbol{\theta}_0)\nabla_\theta' \ln c(u,v;\boldsymbol{\theta}_0)]'\}'c(u,v;\boldsymbol{\theta}_0)\mathrm{d}u\mathrm{d}v.$$

Denote

$$M_1(F_1) = \int_0^1 \int_0^1 [I\{F_{1i} \leq u\} - u]$$

$$\nabla_u\{vech[\nabla_\theta^2 \ln c(u,v;\boldsymbol{\theta}_0)]', vech[\nabla_\theta \ln c(u,v;\boldsymbol{\theta}_0)\nabla_\theta' \ln c(u,v;\boldsymbol{\theta}_0)]'\}'c(u,v;\boldsymbol{\theta}_0)\mathrm{d}u\mathrm{d}v,$$

then

$$\frac{1}{n}\sum_{i=1}^n \nabla_{F_{1i}}\boldsymbol{d}_i(\boldsymbol{\theta}_0)\sqrt{n}(\hat{F}_{1i} - F_{1i}) = \frac{1}{\sqrt{n}}\sum_{i=1}^n M_1(F_{1i}).$$

33

Similarly, denote

$$M_2(F_{2i}) = \int_0^1 \int_0^1 [I\{F_{2i} \le v\} - v]$$
$$\nabla_v\{vech[\nabla_\theta^2 \ln c(u, v; \boldsymbol{\theta}_0)]', vech[\nabla_\theta \ln c(u, v; \boldsymbol{\theta}_0)\nabla_\theta' \ln c(u, v; \boldsymbol{\theta}_0)]'\}'c(u, v; \boldsymbol{\theta}_0)\mathrm{d}u\mathrm{d}v,$$

then

$$\frac{1}{n}\sum_{i=1}^n \nabla_{F_{2i}}\boldsymbol{d}_i(\boldsymbol{\theta}_0)\sqrt{n}(\hat{F}_{2i} - F_{2i}) = \frac{1}{\sqrt{n}}\sum_{t=i}^n M_2(F_{2i}).$$

Therefore, equation (5) can be rewritten as

$$\sqrt{n}\bar{\boldsymbol{d}}(\boldsymbol{\theta}_0) = \frac{1}{\sqrt{n}}\sum_{i=1}^n \boldsymbol{d}_i(\boldsymbol{\theta}_0) + \sqrt{n}B_n^* + o_p(1),$$

where

$$B_n^* = \frac{1}{n}\sum_{i=1}^n [M_1(F_{1i}) + M_2(F_{2i})].$$

Finally, combining the expansions gives

$$\sqrt{n}\bar{\boldsymbol{d}}(\hat{\boldsymbol{\theta}}) = \frac{1}{\sqrt{n}}\sum_{i=1}^n \mathbf{d}_i(\boldsymbol{\theta}_0) + \sqrt{n}B_n^* + \nabla D_{\boldsymbol{\theta}_0}B^{-1}\sqrt{n}A_n^* + o_p(1).$$

So $\bar{\boldsymbol{d}}(\hat{\boldsymbol{\theta}})$ converges in distribution to a multivariate normal with variance matrix $V_{\boldsymbol{\theta}_0}$:

$$\sqrt{n}\bar{\boldsymbol{d}}(\hat{\boldsymbol{\theta}}) \to N(0, V_{\boldsymbol{\theta}_0}),$$

where

$$\begin{aligned}
V_{\boldsymbol{\theta}_0} = \mathbb{E}\,\{&\boldsymbol{d}_i(\boldsymbol{\theta}_0) + M_1(F_{1t}) + M_2(F_{2t}) \\
&+\nabla D_{\boldsymbol{\theta}_0}B^{-1}\left[\nabla_\theta \ln c(F_{1t}, F_{2t}; \boldsymbol{\theta}_0) + W_1(F_{1t}) + W_2(F_{2t})\right]\} \\
\times \{&\boldsymbol{d}_i(\boldsymbol{\theta}_0) + M_1(F_{1t}) + M_2(F_{2t}) \\
&+\nabla D_{\boldsymbol{\theta}_0}B^{-1}\left[\nabla_\theta \ln c(F_{1t}, F_{2t}; \boldsymbol{\theta}_0) + W_1(F_{1t}) + W_2(F_{2t}))\right]\}'.
\end{aligned}$$

Extension to $d \ge 2$ is straightforward. Now

$$\boldsymbol{d}_i(\boldsymbol{\theta}) = \begin{pmatrix} vech(\nabla_\theta^2 \ln c(F_{1i}, F_{2i}, \dots, F_{di}; \boldsymbol{\theta})) \\ vech(\nabla_\theta \ln c(F_{1i}, F_{2i}, \dots, F_{di}; \boldsymbol{\theta})\nabla_\theta' \ln c(F_{1i}, F_{2i}, \dots, F_{di}; \boldsymbol{\theta})) \end{pmatrix}$$

and the asymptotic variance matrix becomes

$$
V_{\boldsymbol{\theta}_0} = \mathbb{E} \left\{ \boldsymbol{d}_i(\theta_0) - \nabla D_{\theta_0} \mathbb{H}^{-1} \left[\nabla_{\boldsymbol{\theta}} \ln c(F_{1i}, F_{2i}, \dots, F_{di}; \boldsymbol{\theta}_0) + \sum_{j=1}^{d} W_j(F_{ji}) \right] + \sum_{j=1}^{d} M_j(F_{ji}) \right\}
$$

$$
\times \left\{ \boldsymbol{d}_i(\boldsymbol{\theta}_0) - \nabla D_{\theta_0} \mathbb{H}^{-1} \left[\nabla_{\boldsymbol{\theta}} \ln c(F_{1i}, F_{2i}, \dots, F_{di}; \boldsymbol{\theta}_0) + \sum_{j=1}^{d} W_j(F_{ji}) \right] + \sum_{j=1}^{d} M_j(F_{ji}) \right\}',
$$

(6)

where, for $j = 1, 2, \dots, d$,

$$
W_j(F_{ji}) = \int_0^1 \int_0^1 \cdots \int_0^1 [I\{F_{ji} \leq u_n\} - u_j] \nabla^2_{\boldsymbol{\theta}, u_j} \ln c(u_1, u_2, \dots, u_d; \boldsymbol{\theta}_0)
$$

$$
c(u_1, u_2, \dots, u_d; \boldsymbol{\theta}_0) \mathrm{d}u_1 \mathrm{d}u_2 \cdots \mathrm{d}u_d,
$$

and

$$
M_j(F_{ji}) = \int_0^1 \int_0^1 \cdots \int_0^1 [I\{F_{ji} \leq u_j\} - u_j] \nabla_{u_j} vech[\nabla^2_{\boldsymbol{\theta}} \ln c(u_1, u_2, \dots, u_d; \boldsymbol{\theta}_0)
$$

$$
+ \nabla_{\boldsymbol{\theta}} \ln c(u_1, u_2, \dots, u_d; \boldsymbol{\theta}_0) \nabla'_{\boldsymbol{\theta}} \ln c(u_1, u_2, \dots, u_d; \boldsymbol{\theta}_0)]
$$

$$
c(u_1, u_2, \dots, u_d; \boldsymbol{\theta}_0) \mathrm{d}u_1 \mathrm{d}u_2 \cdots \mathrm{d}u_d.
$$

Now, since $\hat{s}_n$ is a function of $\bar{\boldsymbol{d}}(\hat{\boldsymbol{\theta}})$, its asymptotic distribution can be easily obtained using the delta method. Define

$$
\nabla s_{\boldsymbol{\theta}_0} := \left(\left. \frac{\partial s}{\partial vech(\mathbb{H})'} \right|_{\boldsymbol{\theta}_0}, \left. \frac{\partial s}{\partial vech(\mathbb{C})'} \right|_{\boldsymbol{\theta}_0} \right)
$$

(7)

Then,

$$
\sqrt{n} \hat{s}_n \xrightarrow{d} N(0, \Sigma_s(\boldsymbol{\theta}_0)),
$$

where

$$
\Sigma_s(\boldsymbol{\theta}_0) := (\nabla s_{\boldsymbol{\theta}_0}) V_{\boldsymbol{\theta}_0} (\nabla s_{\boldsymbol{\theta}_0})'.
$$

Proof of Theorem 1: Follows trivially from Lemma 1 and consistency of $\hat{\Sigma}_{n,s}$ for Σ.

Lemma A1: For any real-valued square matrices A and B, let the elements of $B \in \mathbb{R}^{r \times r}$ be functions of $A \in \mathbb{R}^{p \times p}$. Let the matrix $\frac{dB}{dA} \in \mathbb{R}^{p^2 \times r^2}$ be called matrix derivative of B by A if

$$
\frac{dB}{dA} = \frac{\partial}{\partial vec(A)} vec(B)',
$$

35

where *vec* denotes the vectorization operator. Let D denote the transition matrix, i.e. such a matrix that for, any A, $vech(A) = Dvec(A)$ and $D^+vech(A) = vec(A)$, where D^+ is the Moore-Penrose inverse of D. Then, the following results hold (see, e.g., Kollo and von Rosen, 2006):

$$\frac{dA}{dA} = I_{p^2}$$

$$\frac{dC'A}{dA} = I_p \otimes C, \text{ where C is a matrix of proper size with constant elements}$$

$$\frac{d(C'B)}{dA} = \frac{dB}{dA}(I \otimes C)$$

$$\frac{dBC}{dA} = \frac{dB}{dA}(C \otimes I)$$

$$\frac{dA^{-1}}{dA} = -A^{-1} \otimes (A')^{-1}$$

$$\frac{d\,\mathrm{tr}(B)}{dA} = \frac{dB}{dA}vec(I_r)$$

$$\frac{d\,\mathrm{tr}(C'A)}{dA} = vec(C), \text{ where C is a matrix of proper size with constant elements}$$

$$\frac{d\det(A)}{dA} = \det(A)vec(A^{-1})'$$

$$\frac{dA(B(C))}{dC} = \frac{dB}{dC}\frac{dA}{dB}$$

Lemma A2: Let λ denote an eigenvalue of a symmetric matrix A and let y denote the corresponding normalized eigenvector, i.e. the solution of the equation system $Ay = \lambda y$, such that $y'y = 1$. Let D denote the duplication matrix. Then, the following result holds (see Magnus, 1985):

$$\frac{\partial \lambda}{\partial vech(A)} = [y' \otimes y']D$$

Proof of Proposition 1: First use Lemma A1 on determinant differentiation, as well as properties of *vec* and *vech* operators, to obtain

$$\nabla s_{\boldsymbol{\theta}_0} = det(\mathbb{H}(\boldsymbol{\theta}_0) + \mathbb{C}(\boldsymbol{\theta}_0))vech((\mathbb{H}(\boldsymbol{\theta}_0) + \mathbb{C}(\boldsymbol{\theta}_0))^{-1})' \left[\mathbf{I}_{p(p+1)/2}, \mathbf{I}_{p(p+1)/2}\right]$$

Now use $\hat{\boldsymbol{\theta}}_n$, which is consistent for $\boldsymbol{\theta}_0$, and the sample equivalents $\bar{\mathbb{H}}$ and $\bar{\mathbb{C}}$, which are consistent for $\mathbb{H}$ and $\mathbb{C}$, to obtain the consistent estimator $\widehat{\nabla s}_{\boldsymbol{\theta}_0}$ given in the proposition.

The asymptotic distribution of $\mathcal{T}_n^{(D)}$ then follows from Theorem 1.

Proof of Proposition 2: First use Lemma A1 on trace differentiation to obtain the form of $\nabla s_{\boldsymbol{\theta}_0}$, then the result follows trivially from Theorem 1.

Proof of Proposition 3: First use Lemma A1 on trace and inverse differentiation as well as the fact that $[C' \otimes A] vec(B) = vec(ABC)$, to obtain

$$\nabla s_{\boldsymbol{\theta}_0} = \left(vech \left(\mathbb{H}(\boldsymbol{\theta}_0)^{-1} \mathbb{C}(\boldsymbol{\theta}_0) \mathbb{H}(\boldsymbol{\theta}_0)^{-1} \right)', \quad vech \left(-\mathbb{H}(\boldsymbol{\theta}_0)^{-1} \right)' \right)$$

then replace the population values with consistent estimates as before, and apply Theorem 1 to obtain the result.

Proof of Proposition 4: Similar to previous propositions, using Lemma A1 on determinant differentiation to obtain

$$\nabla s_{\boldsymbol{\theta}_0} = \det(\mathbb{H}(\boldsymbol{\theta}_0)^{-1} \mathbb{C}(\boldsymbol{\theta}_0)) \left(vech \left(-\mathbb{C}(\boldsymbol{\theta}_0)^{-1} \mathbb{H}(\boldsymbol{\theta}_0)^{-1} \mathbb{C}(\boldsymbol{\theta}_0) \right)', \quad vech \left(\mathbb{C}(\boldsymbol{\theta}_0)^{-1} \right)' \right).$$

Proof of Proposition 5: Similar to previous propositions, using Lemma A1 on trace differentiation to obtain

$$\nabla s_{\boldsymbol{\theta}_0} = \left(-\frac{1}{\text{tr}(\mathbb{H}(\boldsymbol{\theta}_0)^{-1})} vech \left(\mathbb{H}(\boldsymbol{\theta}_0)^{-2} \right)', \quad -\frac{1}{\text{tr}(\mathbb{C}(\boldsymbol{\theta}_0))} vech \left(I_p \right)' \right).$$

Proof of Proposition 6: Under the null, this is a log version of the IR test, so

$$\nabla s_{\boldsymbol{\theta}_0} = \frac{1}{\text{tr}(\mathbb{H}(\boldsymbol{\theta}_0)^{-1} \mathbb{C}(\boldsymbol{\theta}_0))} \left(vech \left(\mathbb{H}(\boldsymbol{\theta}_0)^{-1} \mathbb{C}(\boldsymbol{\theta}_0) \mathbb{H}(\boldsymbol{\theta}_0)^{-1} \right)', \quad vech \left(-\mathbb{H}(\boldsymbol{\theta}_0)^{-1} \right)' \right)$$

The rest of the proof is the same as in previous propositions.

Proof of Proposition 7: Similar to above, using Lemma A2 to obtain

$$\nabla s_{\boldsymbol{\theta}_0} = \begin{bmatrix} -\frac{1}{\lambda_1(\mathbb{H}(\boldsymbol{\theta}_0))}[y_1(\mathbb{H}(\boldsymbol{\theta}_0))' \otimes y_1(\mathbb{H}(\boldsymbol{\theta}_0))']D & \frac{1}{\lambda_1(\mathbb{C}(\boldsymbol{\theta}_0)))}[y_1(\mathbb{C}(\boldsymbol{\theta}_0))' \otimes y_1(\mathbb{C}(\boldsymbol{\theta}_0))']D \\ \vdots & \vdots \\ -\frac{1}{\lambda_p(\mathbb{H}(\boldsymbol{\theta}_0))}[y_p(\mathbb{H}(\boldsymbol{\theta}_0))' \otimes y_p(\mathbb{H}(\boldsymbol{\theta}_0))']D & \frac{1}{\lambda_p(\mathbb{C}(\boldsymbol{\theta}_0)))}[y_p(\mathbb{C}(\boldsymbol{\theta}_0))' \otimes y_p(\mathbb{C}(\boldsymbol{\theta}_0))']D \end{bmatrix}.$$

Proof of Proposition 8: Similar to above, using Lemma A2 to obtain

$$\nabla s_{\boldsymbol{\theta}_0} = \begin{bmatrix} \frac{1}{\lambda_1(\mathbb{H}(\boldsymbol{\theta}_0))}[y_1(\mathbb{C}(\boldsymbol{\theta}_0))' \otimes y_1(\mathbb{C}(\boldsymbol{\theta}_0))']D & -\frac{\lambda_1(\mathbb{C}(\boldsymbol{\theta}_0))}{\lambda_1(\mathbb{H}(\boldsymbol{\theta}_0))^2}[y_1(\mathbb{H}(\boldsymbol{\theta}_0))' \otimes y_1(\mathbb{H}(\boldsymbol{\theta}_0))']D \\ \vdots & \vdots \\ \frac{1}{\lambda_p(\mathbb{H}(\boldsymbol{\theta}_0))}[y_p(\mathbb{C}(\boldsymbol{\theta}_0))' \otimes y_p(\mathbb{C}(\boldsymbol{\theta}_0))']D & -\frac{\lambda_p(\mathbb{C}(\boldsymbol{\theta}_0))}{\lambda_p(\mathbb{H}(\boldsymbol{\theta}_0))^2}[y_p(\mathbb{H}(\boldsymbol{\theta}_0))' \otimes y_p(\mathbb{H}(\boldsymbol{\theta}_0))']D \end{bmatrix}.$$

B Vines Used in Simulations

In Section 4.1.2 we used the following vine copula for our simulation study. Table 9 for $d = 5$ and Table 10 for $d = 8$ give details about the vine copula decomposition (structure) $\mathcal{V}$, their selected pair-copula families $\mathcal{B}$ and Kendall's τ for the vine copula under the null hypothesis. For the C-vine and D-vine, $\mathcal{V}$ as well as $\mathcal{B}$ are selected by the algorithms provided in the VineCopula package (Schepsmeier et al., 2013). $\hat{\tau}$ denotes the estimated Kendall's τ in the pre-run step of the simulation procedure of Schepsmeier (2013). Note that the vine copula density is written in a short hand notation omitting the pair-copula arguments. The notation of the pair-copula families follows Brechmann and Schepsmeier (2013).

For the C- and D-vine the calculation of the vine copula density (3) simplifies. For the five-dimensional example used in the simulation study, (3) can be expressed as

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

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

Similar representations used for $d = 8$ and 16 as well as a similar table for $d = 16$ are available from the authors upon request.

C Outer Power Clayton Copula

The Outer Power Clayton copula is defined as follows:

$$C(u) = \psi(\psi^{-1}(u_1) + \cdots + \psi^{-1}(u_d)),$$

where $\psi(t) = \widetilde{\psi}(t^{1/\beta})$ for some $\beta \in [1, \infty)$ and $\widetilde{\psi}(t)$ is the Clayton copula generator $\widetilde{\psi}(t) = (1 + t)^{-1/\theta}$ for some $\theta \in (0, \infty)$. The inversion of Kendall's τ is not feasible here because $\tau = \tau(\theta, \beta) = 1 - \frac{2}{\beta(\theta+2)}$ and so (β, θ) are not identifiable individually. Our simulations using the CMLE instead of the inversion of Kendall's τ for other copulas (not reported here) suggest that the CMLE leads to a substantial power improvement of some GIMT, e.g., of $\mathcal{Q}_n$. We do not have an explanation for this phenomenon and so only report the least favorable results. The power reported in Section 4.2.2 for tests that do not involve the Outer Power Clayton copula is therefore conservative.

	R-vine			C-vine			D-vine		
T	$\mathcal{V}_R^5$	$\mathcal{B}_R^5(\mathcal{V}_R^5)$	τ	$\mathcal{V}_C^5$	$\mathcal{B}_C^5(\mathcal{V}_C^5)$	$\hat{\tau}$	$\mathcal{V}_D^5$	$\mathcal{B}_D^5(\mathcal{V}_D^5)$	$\hat{\tau}$
1	$c_{1,2}$	N	0.71	$c_{1,2}$	N	0.71	$c_{1,2}$	N	0.71
	$c_{1,3}$	N	0.33	$c_{2,3}$	N	0.51	$c_{1,5}$	F	0.70
	$c_{1,4}$	C	0.71	$c_{2,4}$	$G180$	0.70	$c_{4,5}$	G	0.75
	$c_{4,5}$	G	0.74	$c_{2,5}$	F	0.73	$c_{3,4}$	G	0.48
2	$c_{2,4;1}$	G	0.38	$c_{1,3;2}$	$G90$	-0.33	$c_{2,5;1}$	N	0.37
	$c_{3,4;1}$	G	0.47	$c_{1,4;2}$	$G180$	0.29	$c_{1,4;5}$	$G180$	0.22
	$c_{1,5;4}$	G	0.33	$c_{1,5;2}$	$G180$	0.25	$c_{3,5;4}$	C	0.15
3	$c_{2,3;1,4}$	C	0.35	$c_{3,4;1,2}$	N	0.27	$c_{2,4;1,5}$	F	0.18
	$c_{3,5;1,4}$	C	0.31	$c_{3,5;1,2}$	N	0.25	$c_{1,3;4,5}$	F	-0.26
4	$c_{2,5;1,3,4}$	N	0.13	$c_{4,5;1,2,3}$	G	0.20	$c_{2,3;1,4,5}$	$G180$	0.31

Table 9: Chosen vine copula structures, copula families and Kendall's τ values for the R-vine copula model and the C- and D-vine alternatives in the five-dimensional case (N:=Normal, C:=Clayton, G:=Gumbel, F:=Frank, J:=Joe; 90, 180, 270:= degrees of rotation).

D Non-GIMTs for Copulas

Here we provide details on the non-GIMTs used in Section 4.2. We start with a few definitions.

Given a multivariate distribution, the Rosenblatt transformation (Rosenblatt, 1952) yields a set of independent uniforms on $[0, 1]$ from possibly dependent realizations obtained using that multivariate distribution. The Rosenblatt transform can be specialized to copulas as follows:

Definition 3 *Rosenblatt's probability integral transformation (PIT) of a copula C is the mapping $\mathcal{R} : (0,1)^d \to (0,1)^d$ which to every $\boldsymbol{u} = (u_1, \ldots, u_d) \in (0,1)^d$ assigns a vector $\mathcal{R}(u) = (e_1, \ldots, e_d)$ with $e_1 = u_1$ and, for $i \in \{2, \ldots, d\}$,*

$$e_i = \frac{\partial^{i-1} C(u_1, \ldots, u_i, 1, \ldots, 1)}{\partial u_1 \cdots \partial u_{i-1}} \Big/ \frac{\partial^{i-1} C(u_1, \ldots, u_{i-1}, 1, \ldots, 1)}{\partial u_1 \cdots \partial u_{i-1}}. \tag{8}$$

As noted by Genest et al. (2009), the initial random vector $\mathbf{U}$ has distribution C, denoted $\mathbf{U} \sim C$, if and only if the distribution of the Rosenblatt transform $\mathcal{R}(\mathbf{U})$ is the d-variate independence copula defined as $C_\perp(e_1, \ldots, e_d) = \prod_{j=1}^d e_j$. Thus $\mathcal{H}_0 : \mathbf{U} \sim C \in \mathcal{C}_0$ is equivalent to $\mathcal{H}_0^* : \mathcal{R}_{\boldsymbol{\theta}}(\mathbf{U}) \sim C_\perp$.

T	$\mathcal{V}_R^8$	$\mathcal{B}_R^8(\mathcal{V}_R^8)$	τ	$\mathcal{V}_C^8$	$\mathcal{B}_C^8(\mathcal{V}_C^8)$	$\hat{\tau}$	$\mathcal{V}_D^8$	$\mathcal{B}_D^8(\mathcal{V}_D^8)$	$\hat{\tau}$
1	$c_{1,2}$	J	0.41	$c_{1,8}$	F	0.59	$c_{1,4}$	N	0.61
	$c_{1,4}$	N	0.59	$c_{2,8}$	F	0.51	$c_{4,5}$	G180	0.71
	$c_{1,5}$	N	0.59	$c_{3,8}$	N	0.55	$c_{5,8}$	F	0.60
	$c_{1,6}$	F	0.23	$c_{4,8}$	G180	0.59	$c_{7,8}$	G	0.65
	$c_{3,6}$	F	0.19	$c_{5,8}$	F	0.60	$c_{3,7}$	G180	0.41
	$c_{4,7}$	C	0.44	$c_{6,8}$	F	0.27	$c_{2,3}$	G	0.52
	$c_{7,8}$	G	0.64	$c_{7,8}$	G	0.65	$c_{2,6}$	J180	0.57
2	$c_{2,6;1}$	C	0.58	$c_{1,2;8}$	J	0.10	$c_{1,5;4}$	C	0.22
	$c_{1,3;6}$	G	0.44	$c_{2,3;8}$	J	0.29	$c_{4,8;5}$	C	0.22
	$c_{4,6;1}$	F	0.11	$c_{2,4;8}$	G	0.24	$c_{5,7;8}$	J90	-0.05
	$c_{4,5;1}$	C	0.53	$c_{2,5;8}$	G	0.29	$c_{3,8;7}$	G	0.41
	$c_{1,7;4}$	C	0.29	$c_{2,6;8}$	J180	0.52	$c_{2,7;3}$	J	0.10
	$c_{4,8;7}$	N	0.53	$c_{2,7;8}$	N	-0.17	$c_{3,6;2}$	G270	-0.48
3	$c_{5,6;1,4}$	N	0.19	$c_{1,4;2,8}$	N	0.28	$c_{1,8;4,5}$	N	0.20
	$c_{6,7;1,4}$	F	0.03	$c_{3,4;2,8}$	N	0.22	$c_{4,7;5,8}$	N	-0.13
	$c_{1,8;4,7}$	G	0.22	$c_{4,5;2,8}$	G180	0.41	$c_{3,5;7,8}$	G	0.18
	$c_{3,4;1,6}$	N	0.41	$c_{4,6;2,8}$	G270	-0.20	$c_{2,8;3,7}$	G	0.25
	$c_{2,3;1,6}$	G	0.68	$c_{4,7;2,8}$	I	0	$c_{6,7;2,3}$	C	0.08
4	$c_{6,8;1,4,7}$	C	0.17	$c_{1,6;2,4,8}$	J180	0.09	$c_{6,8;2,3,7}$	C	0.05
	$c_{5,7;1,4,6}$	N	0.09	$c_{3,6;2,4,8}$	N	-0.33	$c_{2,5;3,7,8}$	G	0.19
	$c_{3,5;1,4,6}$	F	0.21	$c_{5,6;2,4,8}$	F	-0.04	$c_{3,4;5,7,8}$	C180	0.09
	$c_{2,4;1,3,6}$	G	0.57	$c_{6,7;2,4,8}$	I	0	$c_{1,7;4,5,8}$	J180	0.06
5	$c_{2,5;1,3,4,6}$	J	0.25	$c_{1,5;2,4,6,8}$	C	0.23	$c_{5,6;2,3,7,8}$	C90	-0.04
	$c_{3,7;1,4,5,6}$	G	0.17	$c_{3,5;2,4,6,8}$	F	0.10	$c_{2,4;3,5,7,8}$	C90	-0.02
	$c_{5,8;1,4,6,7}$	F	0.02	$c_{5,7;2,4,6,8}$	F	0.05	$c_{1,3;4,5,7,8}$	G90	-0.09
6	$c_{2,7;1,3,4,5,6}$	G	0.31	$c_{1,3;2,4,5,6,8}$	F	0.07	$c_{4,6;2,3,5,7,8}$	C90	-0.14
	$c_{3,8;1,4,5,6,7}$	C	0.20	$c_{3,7;2,4,5,6,8}$	I	0	$c_{1,2;3,4,5,7,8}$	G90	-0.13
7	$c_{2,8;1,3,4,5,6,7}$	F	0.03	$c_{1,7;2,3,4,5,6,8}$	I	0	$c_{1,6;2,3,4,5,7,8}$	G180	0.24

Table 10: Chosen vine copula structures, copula families and Kendall's τ values for R-vine copula model and the C- and D-vine alternatives in the eight-dimensional case (I:=indep., N:=Normal, C:=Clayton, G:=Gumbel, F:=Frank, J:=Joe; 90, 180, 270:= degrees of rotation).

The PIT algorithm for R-vine copulas is given in the Appendix of Schepsmeier (2015). It makes use of the hierarchical structure of the R-vine, which simplifies the calculation of (8).

Definition 4 *Kendall's transformation is the mapping* $\boldsymbol{X} \mapsto V = C(U_1, \ldots, U_d)$, *where* $U_i = F_i(X_i)$ *for* $i = 1, \ldots, d$ *and* C *denotes the joint distribution of* $\boldsymbol{U} = (U_1, \ldots, U_d)$.

Let K denote the (univariate) distribution function of Kendall's transform V and let K_n denote the empirical analogue of K defined by

$$K_n(v) = \frac{1}{n} \sum_{j=1}^{n} \mathbf{1}(V_j \leq v), \ v \in [0, 1], \tag{9}$$

where $\mathbf{1}(\cdot)$ is the indicator function. Then, under standard regularity conditions, K_n is a consistent estimator of K. Also, under $\mathcal{H}_0$, the vector $\mathbf{U} = (U_1, \ldots, U_d)$ is distributed as $C_{\boldsymbol{\theta}}$ for some $\boldsymbol{\theta} \in \mathcal{O}$, and hence Kendall's transformation $C_{\boldsymbol{\theta}}(\mathbf{U})$ has distribution $K_{\boldsymbol{\theta}}$.

Note that K is not available for all parametric copula families in closed form, especially not for vine copulas. Thus Genest et al. (2009) use a bootstrap procedure to approximate K in such cases.

We now describe the non-GIMTs used in the simulation study.

D.1 Empirical copula process test

This test is based on the empirical copula defined as follows:

$$C_n(\mathbf{u}) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}(U_{i1} \leq u_1, \ldots, U_{id} \leq u_d). \tag{10}$$

It is a well-known result that, under regularity conditions, C_n is a consistent estimator of the true underlying copula C, whether or not $\mathcal{H}_0$ is true. Note that $C_n(\mathbf{u})$ is different from $K_n(v)$, which is a univariate empirical distribution function.

A natural goodness-of-fit test would be based on a "distance" between C_n and an estimated copula C_{θ_n} obtained under $\mathcal{H}_0$. In this paper, $\hat{\boldsymbol{\theta}}_n = \Gamma_n(\mathbf{U}_1, \ldots, \mathbf{U}_n)$ stands for an estimator of $\boldsymbol{\theta}$ obtained using the pseudo-observations.

Thus the test relies on the empirical copula process (ECP) $\sqrt{n}(C_n - C_{\hat{\boldsymbol{\theta}}_n})$. In particular, it has the following rank-based Cramér-von Mises form:

$$\mathcal{S}_n = \int_{[0,1]^d} (C_n - C_{\hat{\boldsymbol{\theta}}_n})^2 \, dC_n(u) = \sum_{j=1}^n \{C_n(\mathbf{U}_j) - C_{\hat{\boldsymbol{\theta}}_n}(\mathbf{U}_j)\}^2, \qquad (11)$$

where large values of $\mathcal{S}_n$ would lead to a rejection of $\mathcal{H}_0$. Genest et al. (2009) demonstrate that the test is consistent, that is, that if $C \notin \mathcal{C}_0$ then $\mathcal{H}_0$ is rejected with probability one as $n \to \infty$.

In the vine copula case we have to perform a double bootstrap procedure to obtain p-values since $C_{\hat{\boldsymbol{\theta}}_n}$ is not available in closed form.

D.2 Rosenblatt's transformation test

As an alternative to $\mathcal{S}_n$, Genest and Rémillard (2008) proposed using $\{\mathbf{V}_j = \mathcal{R}_{C_{\hat{\boldsymbol{\theta}}_n}}(\mathbf{U}_j)\}_{j=1}^n$ instead of $\mathbf{U}_j$, where $\mathcal{R}_{C_{\hat{\boldsymbol{\theta}}}}$ represents Rosenblatt's transformation with respect to the copula $C_{\hat{\boldsymbol{\theta}}_n} \in \mathcal{C}_0$ and $\hat{\boldsymbol{\theta}}_n$ is a consistent estimator of the true value $\boldsymbol{\theta}_0$, under $\mathcal{H}_0 : C \in \mathcal{C}_0 = \{C_{\boldsymbol{\theta}} : \boldsymbol{\theta} \in \mathcal{O}\}$.

The idea is then to compare $C_n(\mathbf{V}_j)$ with the independence copula $C_\perp(\mathbf{V}_j)$ and the corresponding Cramér-von Mises type statistic can be written as follows:

$$\mathcal{S}_n^R = \sum_{j=1}^n \{C_n(\mathbf{V}_j) - C_\perp(\mathbf{V}_j)\}^2. \qquad (12)$$

In the vine copula context Schepsmeier (2015) called this GOF test ECP2 test addressing its close relation to the ECP.

D.3 Kendall's transformation test

Since under $\mathcal{H}_0$, the Kendall's transformation $C_{\boldsymbol{\theta}}(\mathbf{U})$ has distribution $K_{\boldsymbol{\theta}}$, the distance between K_n and a parametric estimator $K_{\hat{\boldsymbol{\theta}}_n}$ of K is another natural testing criterion. We are testing the null $\mathcal{H}_0^{**} : K \in \mathcal{K}_0 = \{K_{\boldsymbol{\theta}} : \boldsymbol{\theta} \in \mathcal{O}\}$ using the empirical process $\mathbb{K} = \sqrt{n}(K_n - K_{\hat{\boldsymbol{\theta}}_n})$. The specific statistic considered by Genest et al. (2006) is the following rank-based analogue of the Cramér-von Mises statistic

$$S_n^K = \int_0^1 \mathbb{K}_n(v)^2 dK_{\hat{\boldsymbol{\theta}}_n}(v)$$