

# A test for Archimedeanity in bivariate copula models

Axel Bücher, Holger Dette and Stanislav Volgushev

Ruhr-Universität Bochum

Fakultät für Mathematik

44780 Bochum, Germany

e-mail: axel.buecher@ruhr-uni-bochum.de

e-mail: holger.dette@ruhr-uni-bochum.de

e-mail: stanislav.volgushev@ruhr-uni-bochum.de

September 9, 2011

## Abstract

We propose a new test for the hypothesis that a bivariate copula is an Archimedean copula. The test statistic is based on a combination of two measures resulting from the characterization of Archimedean copulas by the property of associativity and by a strict upper bound on the diagonal by the Fréchet-upper bound. We prove weak convergence of this statistic and show that the critical values of the corresponding test can be determined by the multiplier bootstrap method. The test is shown to be consistent against all departures from Archimedeanity if the copula satisfies weak smoothness assumptions. A simulation study is presented which illustrates the finite sample properties of the new test.

Keywords and Phrases: Archimedean Copula, associativity, functional delta method, multiplier bootstrap

AMS Subject Classification: Primary 62G10 ; secondary 62G20

## 1 Introduction

Let  $F$  be a bivariate continuous distribution function with marginal distribution functions  $F_1$  and  $F_2$ . By Sklar's Theorem [see Sklar (1959)] we can decompose  $F$  as follows

$$F(\mathbf{x}) = C(F_1(x_1), F_2(x_2)), \quad \mathbf{x} = (x_1, x_2) \in \mathbb{R}^2, \quad (1.1)$$

where  $C$  is the unique copula associated to  $F$ . By definition,  $C$  is a bivariate distribution function on the unit square  $[0, 1]^2$  whose univariate marginals are standard uniform distributions on the interval  $[0, 1]$ . Equation (1.1) is usually interpreted in the way that the copula  $C$  completely

characterizes the information about the stochastic dependence contained in  $F$ . For an extensive exposition on the theory of copulas we refer the reader to the monograph Nelsen (2006).

In the last decades, various parametric models for copulas have been developed, among which the class of Archimedean copulas forms one the most famous and largest class, see Genest and MacKay (1986); Nelsen (2006); McNeil and Nešlehová (2009) among many others. Many widely used copulas, such as Clayton-, Gumbel- and Frank-copulas are in fact Archimedean copulas. The elements of this class may be characterized by a continuous, strictly decreasing and convex function  $\Phi : [0, 1] \rightarrow [0, \infty]$  satisfying  $\Phi(1) = 0$  such that

$$C(\mathbf{u}) = \Phi^{[-1]}[\Phi(u_1) + \Phi(u_2)] \quad \text{for all } \mathbf{u} = (u_1, u_2) \in [0, 1]^2.$$

The function  $\Phi$  is called the *generator* of  $C$  and its *pseudo-inverse*  $\Phi^{[-1]}(t)$  is defined as the usual inverse  $\Phi^{-1}(t)$  for  $t \in [0, \Phi(0)]$  and is set to 0 for  $t \geq \Phi(0)$ . The prominence of the class of Archimedean copulas basically stems from the fact that they are easy to handle and to simulate, see Genest et al. (2011). While the estimation of Archimedean copulas has been investigated in Genest and Rivest (1993) and recently more thoroughly in Genest et al. (2011), the issue of testing for the hypothesis that the copula is an Archimedean one has found much less interest in the literature. The present paper fills this gap by developing a consistent test for this hypothesis.

Our interest in this problem stems from recent work of Genest and Rivest (1993), Wang and Wells (2000) and Naifar (2011) who proposed Archimedean copulas for modeling dependencies between bivariate observations (among many others). We also refer to the work of Rivest and Wells (2001) who used Archimedean copulas for modeling the dependence in the context of censored data.

To the best of our knowledge, the only available test hitherto has been discussed in Jaworski (2010). This author proposed a procedure which is based on a characterization of Archimedean copulas similar to the one stated in Theorem 4.1.6 in Nelsen (2006) [which dates back to Ling (1965)]. To be precise recall that a bivariate copula  $C$  is called associative if and only if the identity

$$C(x, C(y, z)) = C(C(x, y), z) \tag{1.2}$$

holds for all  $(x, y, z) \in [0, 1]^3$ . Theorem 4.1.6 in Nelsen (2006) shows that a bivariate copula  $C$  is an Archimedean copula if and only if  $C$  is associative and the inequality  $C(u, u) < u$  holds for all  $u \in (0, 1)$ , i.e. on the diagonal  $C$  is strictly dominated by the Fréchet-upper bound  $M(\mathbf{u}) = \min(u_1, u_2)$ . The procedure suggested in Jaworski (2010) is in fact to test for associativity in order to check the validity of an Archimedean copula model. The corresponding test statistic is defined as

$$\mathcal{T}_n(x, y) = \sqrt{n}(C_n(x, C_n(y, y)) - C_n(C_n(x, y), y)),$$

where  $(x, y)$  is some fixed point in the open cube  $(0, 1)^2$  and  $C_n$  denotes the empirical copula, see Section 2 for details. The main advantage of this approach is its simplicity, in particular the simple limit distribution of the resulting test statistic, which is in fact normal. On the other hand this simplicity has its price in terms of consistency. In our opinion, the method proposed by Jaworski (2010) has at least three mayor drawbacks. First of all, it is clearly not consistent against a large

class of alternatives since it only tests for equation (1.2) with  $y = z$ . Second, Jaworski (2010) uses a pointwise approach in order to test for a global hypothesis as in (1.2). This means that the test may not reject the hypothesis because (1.2) is satisfied at the particular point  $(x, y, y)$  under investigation, although there may exist many other points where (1.2) is not satisfied. Third, there exist copulas which are in fact associative but not Archimedean. These problems also have strong implications for the practical applicability of the test as demonstrated by results in a simulation study in Jaworski (2010), where the sample size has to be chosen extremely large in order to get reasonable rejection probabilities.

To the best of our knowledge there exists no test for an Archimedean copula, which is consistent against general alternatives and it is the primary purpose of this paper to develop such a procedure and to investigate its statistical properties. We propose a test statistic which is based on a combination of two measures resulting from the characterization of Archimedean copulas, namely the property of associativity as described in (1.2) and the strict upper bound on the diagonal  $C(u, u) < u$  for all  $u \in (0, 1)$ . In Section 2 we define a new process which is based on an estimate of the difference of the left and right hand side of the defining equation (1.2) for associativity. We prove weak convergence of this process in the space of all uniformly bounded functions on the cube  $[0, 1]^3$ . As a consequence, we also obtain weak convergence of a corresponding Cramér-von-Mises and a Kolmogorov-Smirnov type statistic. Because the asymptotic distribution depends in a complicated manner on the underlying copula we propose a multiplier bootstrap procedure to obtain the critical values and show its validity. As a first main result we obtain a test for associativity, which is consistent against all alternatives satisfying weak smoothness assumptions on  $C$ . In Section 3 we utilize these findings to develop an asymptotic test for the hypothesis of Archimedeanity. Finally in Section 4 we investigate the finite sample performance of the new test by means of a simulation study.

## 2 Testing Associativity

### 2.1 The test statistic and its asymptotic behavior

In the following let  $\mathbf{X}_1, \dots, \mathbf{X}_n$ ,  $\mathbf{X}_i = (X_{i1}, X_{i2})$  denote independent identically distributed bivariate random vectors with continuous distribution function  $F$ , marginal distribution functions  $F_1$  and  $F_2$  and copula  $C = F(F_1^-, F_2^-)$ . In this paragraph we will introduce a test statistic for the null hypothesis that the underlying copula is associative, i.e.  $C$  satisfies condition (1.2) for all  $(x, y, z) \in [0, 1]^3$ .

For this purpose we briefly summarize relevant notations and results on the empirical copula, which is the simplest and most popular nonparametric estimator of the copula. In particular we define the empirical copula by

$$C_n(\mathbf{u}) = F_n(F_{n1}^-(u_1), F_{n2}^-(u_2)),$$

where  $F_n(\mathbf{x}) = n^{-1} \sum_{i=1}^n \mathbb{I}\{\mathbf{X}_i \leq \mathbf{x}\}$  and  $F_{np}(x_p) = n^{-1} \sum_{i=1}^n \mathbb{I}\{X_{ip} \leq x_p\}$ ,  $p = 1, 2$  are the joint and marginal empirical distribution functions of the sample  $\mathbf{X}_1, \dots, \mathbf{X}_n$ , respectively. It is a well

known result that under the assumptions of continuous partial derivatives of  $C$  the corresponding empirical copula process

$$\mathbb{C}_n = \sqrt{n}(C_n - C) \quad (2.1)$$

converges weakly towards a Gaussian limit field  $\mathbb{G}_C$  in  $l^\infty([0, 1]^2)$ , see Rüschendorf (1976); Fermanian et al. (2004); Tsukahara (2005) among others. Defining  $\dot{C}_p$  as the  $p$ -th partial derivative of  $C$  ( $p = 1, 2$ ) the process  $\mathbb{G}_C$  can be expressed as

$$\mathbb{G}_C(\mathbf{x}) = \mathbb{B}_C(\mathbf{x}) - \dot{C}_1(\mathbf{x})\mathbb{B}_C(x_1, 1) - \dot{C}_2(\mathbf{x})\mathbb{B}_C(1, x_2) \quad (2.2)$$

with the copula-brownian bridge  $\mathbb{B}_C$ , i.e.  $\mathbb{B}_C$  is a centered Gaussian field with  $\text{Cov}(\mathbb{B}_C(\mathbf{x}), \mathbb{B}_C(\mathbf{y})) = C(\mathbf{x} \wedge \mathbf{y}) - C(\mathbf{x})C(\mathbf{y})$ , where the minimum of two vectors is defined component-wise. As explained in Segers (2011) the assumption of continuity of the partial derivatives of  $C$  on the whole unit square does not hold for many (even most) commonly used copula models and as a consequence Segers provides the result that the following nonrestrictive smoothness condition is sufficient in order to obtain weak convergence of the empirical copula process defined in (2.1).

**Condition 2.1.** *For  $p = 1, 2$  the first order partial derivative  $\dot{C}_p$  of the copula  $C$  with respect to  $x_p$  exists and is continuous on the set  $V_p = \{\mathbf{u} \in [0, 1]^2 : 0 < u_p < 1\}$ .*

Now, in order to test for associativity we consider the process

$$\mathbb{H}_n(x, y, z) = \sqrt{n} \{C_n(x, C_n(y, z)) - C_n(C_n(x, y), z)\},$$

where  $(x, y, z) \in [0, 1]^3$ . The asymptotic properties of the process  $\{\mathbb{H}_n(x, y, z)\}_{(x, y, z) \in [0, 1]^3}$  are summarized in the following Theorem. Throughout this paper  $l^\infty(T)$  denotes the set of all uniformly bounded functions on  $T$ , and the symbol  $\rightsquigarrow$  denotes uniform convergence in a metric space (which will be specified in the corresponding statements).

**Theorem 2.2.** *If the copula  $C$  is associative and satisfies Condition 2.1, then it holds*

$$\mathbb{H}_n \rightsquigarrow \mathbb{H}_C \quad \text{in } l^\infty([0, 1]^3),$$

where the limit field  $\mathbb{H}_C$  can be expressed as

$$\mathbb{H}_C(x, y, z) = \mathbb{G}_C(x, C(y, z)) - \mathbb{G}_C(C(x, y), z) + \dot{C}_2(x, C(y, z))\mathbb{G}_C(y, z) - \dot{C}_1(C(x, y), z)\mathbb{G}(x, y).$$

**Proof.** If the copula  $C$  is associative we can write the process  $\mathbb{H}_n$  as

$$\mathbb{H}_n = \sqrt{n} \{\Phi(C_n) - \Phi(C)\},$$

where the functional  $\Phi : \mathbb{D}_\Phi \rightarrow l^\infty([0, 1]^3)$  is defined for

$$\alpha \in \mathbb{D}_\Phi = \{F : F \text{ cdf on } [0, 1]^2\}$$

by

$$\Phi(\alpha)(x, y, z) = \alpha(x, \alpha(y, z)) - \alpha(\alpha(x, y), z).$$

We will show later that under Condition 2.1 the mapping  $\Phi$  is Hadamard-differentiable at  $C$  tangentially to the space

$$\mathbb{D}_0 = \{\gamma \in C[0, 1]^2 \mid \gamma(\mathbf{u}) = 0 \text{ for all } \mathbf{u} \in [0, 1]^2 \text{ s.t. } C(\mathbf{u}) \in \{0, 1\}\},$$

with derivative given by

$$\Phi'_C(\alpha)(x, y, z) = \alpha(x, C(y, z)) - \alpha(C(x, y), z) + \dot{C}_2(x, C(y, z))\alpha(y, z) - \dot{C}_1(C(x, y), z)\alpha(x, y).$$

Observing that  $\mathbb{B}_C \in \mathbb{D}_0$  a.s., the functional delta method, see Theorem 3.9.4 in Van der Vaart and Wellner (1996), yields the assertion.

We now briefly sketch how to see the Hadamard-differentiability of the mapping  $\Phi$ : let  $t_n \rightarrow 0$  and  $\alpha_n \in l^\infty([0, 1]^2)$  with  $\alpha_n \rightarrow \alpha \in \mathbb{D}_0$  such that  $C + t_n\alpha_n \in \mathbb{D}_\Phi$ . Then

$$t_n^{-1}\{\Phi(C + t_n\alpha_n) - \Phi(C)\} = L_{n1} + L_{n2} - L_{n3}$$

where

$$\begin{aligned} L_{n1}(x, y, z) &= \alpha_n(x, (C + t_n\alpha_n)(y, z)) - \alpha_n((C + t_n\alpha_n)(x, y), z) \\ L_{n2}(x, y, z) &= t_n^{-1}\{C(x, (C + t_n\alpha_n)(y, z)) - C(x, C(y, z))\} \\ L_{n3}(x, y, z) &= t_n^{-1}\{C((C + t_n\alpha_n)(x, y), z) - C(C(x, y), z)\}. \end{aligned}$$

Exploiting the fact that  $\alpha_n$  converges uniformly to a bounded function and that  $\alpha$  is uniformly continuous one can conclude that  $L_{n1}(x, y, z) = \alpha(x, C(y, z)) - \alpha(C(x, y), z) + o(1)$  uniformly in  $(x, y, z) \in [0, 1]^3$ . Regarding the summand  $L_{n2}$  we have to split the investigation in two cases. First, we consider all those  $(x, y, z) \in [0, 1]^3$  for which  $C(y, z) \in (0, 1)$ . A Taylor expansion of  $C(x, \cdot)$  at  $C(y, z)$  yields

$$L_{n2}(x, y, z) = \dot{C}_2(x, C(y, z))\alpha_n(y, z) + r_n(x, y, z),$$

where the error term can be written as

$$r_n(x, y, z) = (\dot{C}_2(x, u_n) - \dot{C}_2(x, C(y, z)))\alpha_n(y, z)$$

with some intermediate point  $u_n$  between  $C(y, z)$  and  $(C + t_n\alpha_n)(y, z)$ . The main term uniformly converges to  $\dot{C}_2(x, C(y, z))\alpha(y, z)$  [note that partial derivatives of copulas are uniformly bounded by 1] and it remains to show that  $r_n(x, y, z) = o(1)$  uniformly in  $(x, y, z)$  with  $C(y, z) \in (0, 1)$ .

To see this, we will show at the end of this proof that for any  $\varepsilon > 0$  there exists a  $\delta > 0$ , such that

$$\limsup_{n \rightarrow \infty} \sup_{\mathbf{v} \in A_\delta} |\alpha_n(\mathbf{v})| \leq \varepsilon. \quad (2.3)$$

where  $\mathbf{v} = (y, z)$ ,  $A_\delta = \{\mathbf{v} \in [0, 1]^2 \mid C(\mathbf{v}) \in [0, \delta] \cup (1 - \delta, 1]\}$ . Then, since partial derivatives of copulas are bounded by 1, we can conclude that

$$\limsup_{n \rightarrow \infty} \sup_{x \in [0, 1], (y, z) \in A_\delta} |r_n(x, y, z)| \leq \varepsilon.$$

Due to Condition 2.1 the partial derivative  $\dot{C}_2$  is uniformly continuous on the quadrangle  $[0, 1] \times [\delta, 1 - \delta]$ . Thus, since  $\alpha$  is uniformly bounded and since  $u_n \rightarrow C(y, z)$ , we obtain uniform convergence of  $r_n(x, y, z)$  to 0 for all  $(y, z)$  s.t.  $C(y, z) \in [\delta, 1 - \delta]$ , i.e. for  $(y, z) \in [0, 1]^2 \setminus A_\delta$ . Combining the two facts derived above, it follows that

$$\limsup_{n \rightarrow \infty} \sup_{x \in [0, 1], C(y, z) \in (0, 1)} |r_n(x, y, z)| \leq \varepsilon.$$

Since  $\varepsilon > 0$  was arbitrary, this lim sup must be zero. Summarizing, the case  $(x, y, z) \in [0, 1]^3$  such that  $C(y, z) \in (0, 1)$  is finished.

In the remaining case  $C(y, z) \in \{0, 1\}$ , i.e.  $(y, z) \in A_0$ , Lipschitz-continuity of  $C$  entails that

$$|L_{n2}(x, y, z)| = t_n^{-1} |C(x, C(y, z) + t_n \alpha_n(y, z)) - C(x, C(y, z))| \leq \alpha_n(y, z) = \alpha(y, z) + o(1) = o(1)$$

uniformly in  $(x, y, z)$  since in this case  $\alpha(y, z) = 0$ . Finally, the summand  $L_{n3}$  may be treated analogously.

To complete the proof it remains to show (2.3). Exploiting uniform convergence of  $\alpha_n$ , uniform continuity of  $\alpha$  and the fact that  $\alpha(\mathbf{v}) = 0$  for all  $\mathbf{v} \in A_0 = \{\mathbf{v} \mid C(\mathbf{v}) \in \{0, 1\}\}$ , we can conclude that there exists a  $\kappa > 0$  such that  $|\alpha_n(\mathbf{v})| \leq \varepsilon$  for all  $\mathbf{v} \in A_0^\kappa = \{\mathbf{v} \mid \exists \mathbf{u} \in A_0 \text{ s.t. } \|\mathbf{u} - \mathbf{v}\| \leq \kappa\}$  and sufficiently large  $n$ . For  $v_1 \in [\kappa, 1]$  let  $\delta(v_1) = \sup\{C(v_1, z) \mid (v_1, z) \in A_0^\kappa\}$  [which equals  $C(v_1, z(v_1))$  for some  $z(v_1)$  such that  $(v_1, z(v_1)) \in \partial A_0^\kappa \cap (0, 1)^2$  since for fixed any  $v_1$  the function  $u \mapsto C(v_1, u)$  is increasing] and set  $\delta = \inf_{v_1 \in [\kappa, 1]} \delta(v_1)$ , which is strictly positive due to compactness of  $\partial A_0^\kappa \cap (0, 1)^2$  and continuity of  $C$ . We will now show that this choice of  $\delta$  yields (2.3). Now, if  $C(\mathbf{v}) \leq \delta$ , we have either  $v_1 < \kappa$  [then  $\mathbf{v} \in A_0^\kappa$  since  $C(0, v_2) = 0$ ] or  $v_1 \geq \kappa$ . In the latter case,  $C(\mathbf{v}) \leq \delta(v_1)$  and monotonicity of  $C$  imply  $\mathbf{v} \in A_0^\kappa$ . This proves (2.3) and completes the proof of Theorem 2.2.  $\square$

As a consequence of Theorem 2.2 and the continuous mapping Theorem [see e.g. Theorem 1.3.6 in Van der Vaart and Wellner (1996)], we obtain the weak convergence of a corresponding Cramér-von-Mises and Kolmogorov-Smirnov type test statistic, i.e.

$$\mathbb{T}_{n, L_2} = \int_{[0, 1]^3} \{\mathbb{H}_n(x, y, z)\}^2 d(x, y, z) \rightsquigarrow \mathbb{T}_{C, L_2} = \int_{[0, 1]^3} \{\mathbb{H}_C(x, y, z)\}^2 d(x, y, z), \quad (2.4)$$

$$\mathbb{T}_{n, \text{KS}} = \sup_{[0, 1]^3} |\mathbb{H}_n(x, y, z)| \rightsquigarrow \mathbb{T}_{C, \text{KS}} = \sup_{[0, 1]^3} |\mathbb{H}_C(x, y, z)|, \quad (2.5)$$

which can be used to construct an asymptotic test for the hypothesis of associativity. Since  $\mathbb{T}_{n, M} \xrightarrow{\mathbb{P}} \infty$  [ $M \in \{L_2, \text{KS}\}$ ] if the copula is not associative the null hypothesis should be rejected for unlikely large values of  $\mathbb{T}_{n, M}$ . This gives rise to the demand for critical values of  $\mathbb{T}_{C, M}$  which can be obtained by multiplier bootstrap methods as described in the subsequent paragraph.

## 2.2 A multiplier bootstrap approximation

It is the purpose of this Section to provide a bootstrap approximation for the distribution of the limiting variables  $\mathbb{T}_{C,M}$  whose variances depend on the unknown copula in a complicated manner. We begin with an approximation of the distribution of the limiting process  $\mathbb{H}_C$ . For this purpose we rewrite the decomposition of the process  $\mathbb{G}_C$  defined in (2.2) as

$$\begin{aligned} \mathbb{H}_C(x, y, z) &= \mathbb{B}_C(x, C(y, z)) - \dot{C}_1(x, C(y, z))\mathbb{B}_C(x, 1) - \dot{C}_2(x, C(y, z))\mathbb{B}_C(1, C(y, z)) \\ &\quad - \left\{ \mathbb{B}_C(C(x, y), z) - \dot{C}_1(C(x, y), z)\mathbb{B}_C(C(x, y), 1) - \dot{C}_2(C(x, y), z)\mathbb{B}_C(1, z) \right\} \\ &\quad + \dot{C}_2(x, C(y, z)) \left\{ \mathbb{B}_C(y, z) - \dot{C}_1(y, z)\mathbb{B}_C(y, 1) - \dot{C}_2(y, z)\mathbb{B}_C(1, z) \right\} \\ &\quad + \dot{C}_1(C(x, y), z) \left\{ \mathbb{B}_C(x, y) - \dot{C}_1(x, y)\mathbb{B}_C(x, 1) - \dot{C}_2(x, y)\mathbb{B}_C(1, y) \right\}. \end{aligned} \quad (2.6)$$

In the following discussion the symbol

$$G_n \xrightarrow[\xi]{\mathbb{P}} G \quad (2.7)$$

denotes weak convergence in some metric space  $\mathbb{D}$  conditionally on the data in probability [see Kosorok (2008)]. More precisely, (2.7) holds for random variables  $G_n = G_n(\mathbf{X}_1, \dots, \mathbf{X}_n, \xi_1, \dots, \xi_n)$ ,  $G \in \mathbb{D}$  if and only if

$$\sup_{h \in BL_1(\mathbb{D})} |\mathbb{E}_\xi h(G_n) - \mathbb{E}h(G)| \xrightarrow{\mathbb{P}} 0 \quad (2.8)$$

and

$$\mathbb{E}_\xi h(G_n)^* - \mathbb{E}_\xi h(G_n)_* \xrightarrow{\mathbb{P}} 0 \quad \text{for every } h \in BL_1(\mathbb{D}), \quad (2.9)$$

where

$$BL_1(\mathbb{D}) = \{f : \mathbb{D} \rightarrow \mathbb{R} \mid \|f\|_\infty \leq 1, |f(\beta) - f(\gamma)| \leq d(\beta, \gamma) \forall \gamma, \beta \in \mathbb{D}\}$$

denotes the set of all Lipschitz-continuous functions bounded by 1. The subscript  $\xi$  in the expectations in (2.8) and (2.9) indicates the conditional expectation with respect to the weights  $\xi = (\xi_1, \dots, \xi_n)$  given the data and  $h(G_n)^*$  and  $h(G_n)_*$  denote measurable majorants and minorants with respect to the joint data, including the weights  $\xi$ . Note also that condition (2.8) is motivated by the metrization of weak convergence by the bounded Lipschitz-metric, see e.g. Theorem 1.12.4 in Van der Vaart and Wellner (1996).

The process  $\mathbb{B}_C$  can be approximated by multiplier bootstrap methods, see Bücher (2011); Bücher and Dette (2010); Rémillard and Scaillet (2009); Segers (2011). More precisely, let  $\xi_1, \dots, \xi_n$  denote independent identically distributed random variables with mean 0 and variance 1 such that

$$\|\xi_i\|_{2,1} = \int_0^\infty \sqrt{\mathbb{P}(|\xi_i| > x)} dx < \infty, \quad (2.10)$$

and consider the process

$$\alpha_n^\xi = \sqrt{n}(C_n^\xi - C_n), \quad (2.11)$$

where

$$C_n^\xi(\mathbf{x}) = n^{-1} \sum_{i=1}^n \frac{\xi_i}{\xi_n} \mathbb{I}\{X_{i1} \leq F_{n1}^-(x_1), X_{i2} \leq F_{n2}^-(x_2)\}$$

denotes a multiplier bootstrap version of the estimator. It was shown in Bücher and Dette (2010) and in more detail in Bücher (2011) that

$$\alpha_n^\xi \xrightarrow[\xi]{\mathbb{P}} \mathbb{B}_C$$

i.e. the process  $\alpha_n^\xi$  defined in (2.11) converges weakly to  $\mathbb{B}_C$  in  $l^\infty([0, 1]^2)$  conditionally on the data in probability in the sense of Kosorok (2008).

For the approximation of the partial derivatives in (2.6) let  $\widehat{C}_p$  be some estimator of  $\dot{C}_p$ ; for instance an estimator based on the differential quotient as in Rémillard and Scaillet (2009) defined by

$$\widehat{C}_1(\mathbf{u}) := \begin{cases} \frac{C_n(u_1+h, u_2) - C_n(u_1-h, u_2)}{2h} & \text{if } u_1 \in [h, 1-h] \\ \frac{C_n(2h, u_2)}{2h} & \text{if } u_1 \in [0, h) \\ \frac{u_2 - C_n(1-2h, u_2)}{2h} & \text{if } u_1 \in (1-h, 1] \end{cases} \quad (2.12)$$

$$\widehat{C}_2(\mathbf{u}) := \begin{cases} \frac{C_n(u_1, u_2+h) - C_n(u_1, u_2-h)}{2h} & \text{if } u_2 \in [h, 1-h] \\ \frac{C_n(u_1, 2h)}{2h} & \text{if } u_2 \in [0, h) \\ \frac{u_1 - C_n(u_1, 1-2h)}{2h} & \text{if } u_2 \in (1-h, 1] \end{cases} \quad (2.13)$$

where  $h = h_n \rightarrow 0$  such that  $\inf_n h_n \sqrt{n} > 0$  [for a smooth version of these estimators see Scaillet (2005)].

**Theorem 2.3.** *Assume that there exists a constant  $K$  such that  $\|\widehat{C}_p\|_\infty \leq K$  for all  $n \in \mathbb{N}, p = 1, 2$  and that*

$$\sup_{\mathbf{u} \in [0, 1]^2 : u_p \in [\delta, 1-\delta]} \left| \widehat{C}_p(\mathbf{u}) - \dot{C}_p(\mathbf{u}) \right| \xrightarrow{\mathbb{P}} 0$$

for all  $\delta \in (0, 1/2)$ . If moreover the copula  $C$  satisfies Condition 2.1 and if the multipliers  $\xi_i$  satisfy (2.10), then the multiplier process  $\mathbb{H}_n^\xi$  defined as

$$\begin{aligned} \mathbb{H}_n^\xi(x, y, z) &= \alpha_n^\xi(x, C_n(y, z)) - \widehat{C}_1(x, C_n(y, z))\alpha_n^\xi(x_1, 1) - \widehat{C}_2(x, C_n(y, z))\alpha_n^\xi(1, C_n(y, z)) \\ &\quad - \left\{ \alpha_n^\xi(C_n(x, y), z) - \widehat{C}_1(C_n(x, y), z)\alpha_n^\xi(C_n(x, y), 1) - \widehat{C}_2(C_n(x, y), z)\alpha_n^\xi(1, z) \right\} \\ &\quad + \widehat{C}_2(x, C_n(y, z)) \left\{ \alpha_n^\xi(y, z) - \widehat{C}_1(y, z)\alpha_n^\xi(y, 1) - \widehat{C}_2(y, z)\alpha_n^\xi(1, z) \right\} \\ &\quad + \widehat{C}_1(C_n(x, y), z) \left\{ \alpha_n^\xi(x, y) - \widehat{C}_1(x, y)\alpha_n^\xi(x, 1) - \widehat{C}_2(x, y)\alpha_n^\xi(1, y) \right\} \end{aligned}$$

converges weakly to the process  $\mathbb{H}_C$  conditional on the data in probability, i.e.  $\mathbb{H}_n^\xi \xrightarrow[\xi]{\mathbb{P}} \mathbb{H}_C$ .



**Proof.** Define the process  $\tilde{\mathbb{H}}_n^\xi$  by substituting the estimators  $\widehat{C}_1, \widehat{C}_2$  and  $C_n$  in the definition of  $\mathbb{H}_n^\xi$  by the true but unknown objects  $\dot{C}_1, \dot{C}_2$  and  $C$ . By Lemma A.1 in Bücher (2011) it suffices to show that

$$\|\mathbb{H}_n^\xi - \tilde{\mathbb{H}}_n^\xi\|_\infty = \sup_{(x,y,z) \in [0,1]^3} |\mathbb{H}_n^\xi(x,y,z) - \tilde{\mathbb{H}}_n^\xi(x,y,z)| \xrightarrow{\mathbb{P}} 0.$$

Using the triangle inequality we have to estimate the following 12 summands

$$\begin{aligned} \|\mathbb{H}_n^\xi - \tilde{\mathbb{H}}_n^\xi\|_\infty &\leq \|\alpha_n^\xi(x, C_n(y, z)) - \alpha_n^\xi(x, C(y, z))\|_\infty \\ &+ \|\widehat{\dot{C}}_1(x, C_n(y, z))\alpha_n^\xi(x, 1) - \dot{C}_1(x, C(y, z))\alpha_n^\xi(x, 1)\|_\infty \\ &+ \|\widehat{\dot{C}}_2(x, C_n(y, z))\alpha_n^\xi(1, C_n(y, z)) - \dot{C}_2(x, C(y, z))\alpha_n^\xi(1, C(y, z))\|_\infty \\ &+ \|\alpha_n^\xi(C_n(x, y), z) - \alpha_n^\xi(C(x, y), z)\|_\infty \\ &+ \|\widehat{\dot{C}}_1(C_n(x, y), z)\alpha_n^\xi(C_n(x, y), 1) - \dot{C}_1(C(x, y), z)\alpha_n^\xi(C(x, y), 1)\|_\infty \\ &+ \|\widehat{\dot{C}}_2(C_n(x, y), z)\alpha_n^\xi(1, z) - \dot{C}_2(C(x, y), z)\alpha_n^\xi(1, z)\|_\infty \\ &+ \|\widehat{\dot{C}}_2(x, C_n(y, z))\alpha_n^\xi(y, z) - \dot{C}_2(x, C(y, z))\alpha_n^\xi(y, z)\|_\infty \\ &+ \|\widehat{\dot{C}}_2(x, C_n(y, z))\widehat{\dot{C}}_1(y, z)\alpha_n^\xi(y, 1) - \dot{C}_2(x, C(y, z))\dot{C}_1(y, z)\alpha_n^\xi(y, 1)\|_\infty \\ &+ \|\widehat{\dot{C}}_2(x, C_n(y, z))\widehat{\dot{C}}_2(y, z)\alpha_n^\xi(1, z) - \dot{C}_2(x, C(y, z))\dot{C}_2(y, z)\alpha_n^\xi(1, z)\|_\infty \\ &+ \|\widehat{\dot{C}}_1(C_n(x, y), z)\alpha_n^\xi(x, y) - \dot{C}_1(C(x, y), z)\alpha_n^\xi(x, y)\|_\infty \\ &+ \|\widehat{\dot{C}}_1(C_n(x, y), z)\widehat{\dot{C}}_1(x, y)\alpha_n^\xi(x, 1) - \dot{C}_1(C(x, y), z)\dot{C}_1(x, y)\alpha_n^\xi(x, 1)\|_\infty \\ &+ \|\widehat{\dot{C}}_1(C_n(x, y), z)\widehat{\dot{C}}_2(x, y)\alpha_n^\xi(1, y) - \dot{C}_1(C(x, y), z)\dot{C}_2(x, y)\alpha_n^\xi(1, y)\|_\infty, \end{aligned}$$

of which one of the hardest cases will be considered exemplarily in the following, namely the third summand

$$\sup_{(x,y,z) \in [0,1]^3} \left| \widehat{\dot{C}}_2(x, C_n(y, z))\alpha_n^\xi(1, C_n(y, z)) - \dot{C}_2(x, C(y, z))\alpha_n^\xi(1, C(y, z)) \right|.$$

The treatment of the other summands is similar and is omitted for the sake of brevity. We estimate

$$\begin{aligned} &\left| \widehat{\dot{C}}_2(x, C_n(y, z))\alpha_n^\xi(1, C_n(y, z)) - \dot{C}_2(x, C(y, z))\alpha_n^\xi(1, C(y, z)) \right| \\ &\leq \left| \widehat{\dot{C}}_2(x, C_n(y, z)) - \dot{C}_2(x, C_n(y, z)) \right| \times |\alpha_n^\xi(1, C_n(y, z))| \\ &\quad + \left| \dot{C}_2(x, C_n(y, z)) - \dot{C}_2(x, C(y, z)) \right| \times |\alpha_n^\xi(1, C_n(y, z))| \\ &\quad + \left| \dot{C}_2(x, C(y, z)) \right| \times |\alpha_n^\xi(1, C_n(y, z)) - \alpha_n^\xi(1, C(y, z))| \\ &=: A_1(x, y, z) + A_2(x, y, z) + A_3(x, y, z) \end{aligned}$$

and consider each term separately. For arbitrary  $\varepsilon > 0$  and  $\delta \in (0, 1/2)$  we estimate

$$\mathbb{P}(\sup A_1(x, y, z) > \varepsilon) \leq \mathbb{P}\left(\sup_{C_n(y, z) \in [\delta, 1-\delta]} A_1(x, y, z) > \varepsilon\right) + \mathbb{P}\left(\sup_{C_n(y, z) \notin [\delta, 1-\delta]} A_1(x, y, z) > \varepsilon\right)$$

where we suppressed the index  $(x, y, z) \in [0, 1]^3$  at the suprema. The first probability can be made arbitrary small by the assumptions on  $\widehat{C}_2$  and by the asymptotic tightness of the process  $\alpha_n^\xi$ , see Theorem 2.3 in Bücher (2011). For the second summand use uniform boundedness of  $\widehat{C}_2$  and the fact that the (unconditional) limit process  $\mathbb{B}_C(1, \cdot)$  of  $\alpha_n^\xi(1, \cdot)$  is a standard Brownian bridge having continuous trajectories which vanish at 0 and 1. By decreasing  $\delta$  the probability can be made arbitrary small, see Segers (2011) for an rigorous treatment of this argument.

Since  $\widehat{C}_2$  is uniformly continuous if the second coordinate is bounded away from zero and one the second summand  $A_2(x, y, z)$  can be treated similarly. Regarding  $A_3(x, y, z)$  note that  $\alpha_n^\xi$  is asymptotically uniformly equicontinuous [Theorem 2.3 in Bücher (2011)] and that  $\sup_{(y, z) \in [0, 1]^2} |C_n(y, z) - C(y, z)| \xrightarrow{\mathbb{P}} 0$  which yields

$$\sup_{(y, z) \in [0, 1]^2} |\alpha_n^\xi(1, C_n(y, z)) - \alpha_n^\xi(1, C(y, z))| \xrightarrow{\mathbb{P}} 0.$$

By boundedness of  $\widehat{C}_2$  this yields the assertion  $\sup_{(x, y, z) \in [0, 1]^3} A_3(x, y, z) \xrightarrow{\mathbb{P}} 0$ .  $\square$

**Remark 2.4.**

(a) Note that the assumptions on the estimator  $\widehat{C}_p$  for the partial derivatives  $\dot{C}_p$  are e.g. satisfied for the estimators defined in (2.12) and (2.13), see Lemma 4.1 in Segers (2010).

(b) Note that Theorem 2.3 holds independently of the hypothesis of associativity. As a consequence of the continuous mapping theorem for the bootstrap, see Proposition 10.7 in Kosorok (2008), we can conclude that

$$\mathbb{T}_{n, L_2}^\xi = \int_{[0, 1]^3} \{\mathbb{H}_n^\xi(x, y, z)\}^2 d(x, y, z) \xrightarrow[\xi]{\mathbb{P}} \mathbb{T}_{C, L_2}, \quad \mathbb{T}_{n, KS}^\xi = \sup_{[0, 1]^3} |\mathbb{H}_n^\xi(x, y, z)| \xrightarrow[\xi]{\mathbb{P}} \mathbb{T}_{C, KS} \quad (2.14)$$

and the latter convergence suggests to use the following approach in order to obtain an asymptotic level- $\alpha$  test for the hypothesis of associativity.

1. Compute the statistic  $\mathbb{T}_{n, M}$  [ $M \in \{L_2, KS\}$ ].
2. Choose the number of bootstrap replications  $B \in \mathbb{N}$ . For  $b = 1, \dots, B$  simulate independent replications of the random variables  $\xi_1, \dots, \xi_n$  and denote the result from the  $b$ -th iteration by  $\xi_{1, b}, \dots, \xi_{n, b}$ .
3. For  $b = 1, \dots, B$  compute the statistics  $\mathbb{T}_{n, M}^{(\xi, b)}$  defined in (2.14) from the data  $\mathbf{X}_1, \dots, \mathbf{X}_n$  and the multipliers  $\xi_{1, b}, \dots, \xi_{n, b}$  and determine the  $(1 - \alpha)$ -quantile  $q_{1-\alpha, M}^\xi$  of the empirical distribution of the sample  $\{\mathbb{T}_{n, M}^{(\xi, b)}\}_{b=1, \dots, B}$ .

4. Reject the null hypothesis of associativity whenever  $\mathbb{T}_{n,M} > q_{1-\alpha,M}^\xi$

Since  $\mathbb{T}_{n,M} \xrightarrow{\mathbb{P}} \infty$  and  $\mathbb{T}_{n,M}^\xi \xrightarrow[\xi]{\mathbb{P}} \mathbb{T}_{C,M}$  if the copula is not associative the test is consistent against all alternatives satisfying the conditions of Theorem 2.3.

### 3 Testing Archimedeanity

As stated in the Introduction a bivariate copula  $C$  is Archimedean if and only if  $C$  is an associative copula satisfying  $C(u, u) < u$  for all  $u \in (0, 1)$ . Associativity has been dealt with in the preceding paragraph and it remains to handle non-Archimedean copulas which may be associative but satisfy  $C(q, q) = q$  for some  $q \in (0, 1)$ . Due to Theorem 1 in Jaworski (2010) or by the results in Section 2.4 of Alsina et al. (2006) all those copulas may be expressed as an ordinal sum of Archimedean copulas. An ordinal sum copula is defined as following [cf. Section 3.2.2 in Nelsen (2006)]: let  $\{J_i\}_{i \in I}$  be a countable partition of non-overlapping closed intervals  $J_i = [a_i, b_i]$  whose union is  $[0, 1]$ . If moreover  $\{C_i\}_{i \in I}$  is a collection of copulas, then the ordinal sum of  $\{C_i\}_{i \in I}$  with respect to  $\{J_i\}_{i \in I}$  is the copula  $C$  defined by

$$C(\mathbf{u}) = \begin{cases} a_i + (b_i - a_i)C_i\left(\frac{u_1 - a_i}{b_i - a_i}, \frac{u_2 - a_i}{b_i - a_i}\right) & \text{if } \mathbf{u} \in J_i \times J_i \\ \min\{u_1, u_2\} & \text{otherwise.} \end{cases}$$

Note that  $C(b_i, b_i) = b_i$  for all  $b_i$  and that ordinal sum copulas put no mass on  $[0, 1]^2 \setminus \bigcup_{i \in I} J_i \times J_i$ . In Figure 1 we illustrate the ordinal sum of a Gumbel copula  $C_1$  with parameter  $\theta_1 = 1.5$  and a Clayton Copula  $C_2$  with parameter  $\theta_2 = 1$ , where  $J_1 = [0, 1/2]$ ,  $J_2 = [1/2, 1]$ . Note that Kendall's  $\tau$  of  $C$  is equal to  $2/3$ , while it equals  $1/3$  for both  $C_1$  and  $C_2$ .

In order to check for  $C(q, q) = q$  for some  $q \in (0, 1)$  we propose the following modification of the statistic  $\mathbb{T}_{n,M}$

$$\mathbb{S}_{n,M} = \mathbb{T}_{n,M} + k_n \phi(A_n(C_n)),$$

where  $\mathbb{T}_{n,M}$  is defined in (2.4) and (2.5),  $k_n \sim n^\alpha$ ,  $\alpha \in (0, 1/2)$  is some constant chosen by the statistician,  $\phi$  is some increasing function with  $\phi(0) = 0$  and

$$A_n(C_n) = \max \left\{ \frac{i}{n} \left( 1 - \frac{i}{n} \right) : C_n \left( \frac{i}{n}, \frac{i}{n} \right) = \frac{i}{n} \right\}.$$

Intuitively,  $A_n(C_n)$  should be “large” for copulas which satisfy  $C(q, q) = q$  for some  $q \in (0, 1)$ . For a decent choice of  $k_n$  and  $\phi$  we refer the reader to Section 4.

In Figure 2 we illustrate the points  $i/n$  at which  $C_n(i/n, i/n) = i/n$  for two specific examples, the Clayton copula with  $\theta = 1$  and the ordinal sum copula depicted in Figure 1. The solid and dashed lines correspond to the true copula and the empirical copula calculated for a set of  $n = 100$  observations, respectively. For the ordinal sum copula there always exist some points  $i/n$  in a neighbourhood of  $1/2$  such that  $C_n(i/n, i/n) = i/n$ , see the proof of the following Proposition, which is sufficient for the derivation of the asymptotic properties of the statistic  $\mathbb{S}_{n,M}$ .

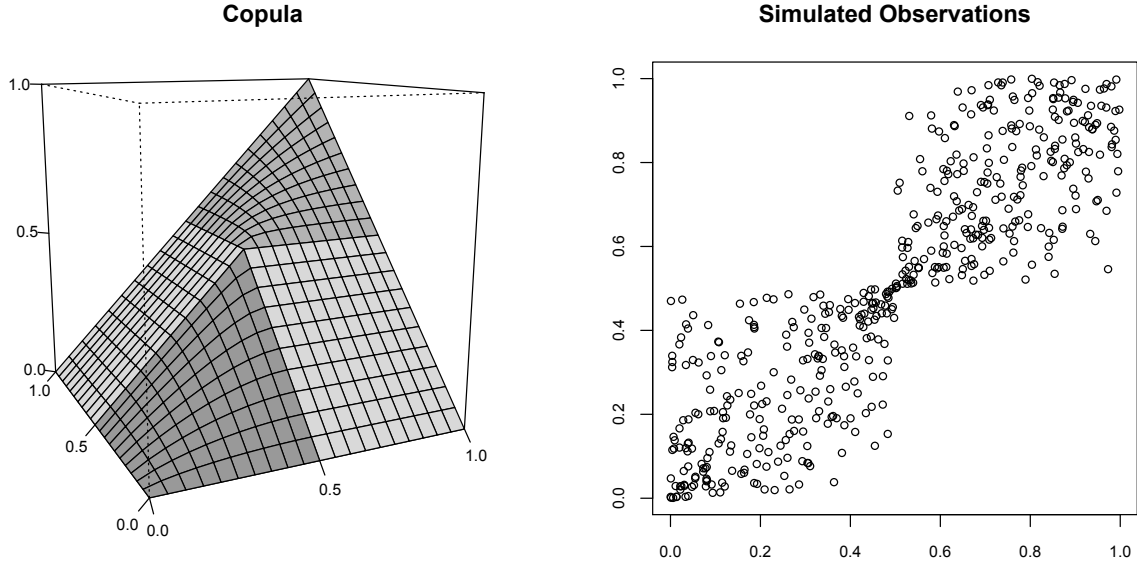


Figure 1: *Left picture: Ordinal sum copula. Right picture: 500 corresponding simulated observations.*

**Proposition 3.1.**

a) Suppose  $C$  is an Archimedean copula satisfying Condition 2.1 and that the coefficients of tail dependence

$$\lambda_L = \lim_{u \rightarrow 0} \frac{C(u, u)}{u} \quad \text{and} \quad \lambda_U = \lim_{u \rightarrow 1} \frac{1 - 2u + C(u, u)}{1 - u}$$

exist and are smaller than 1. Then it holds

$$A_n(C_n) = o_{\mathbb{P}}(n^{-\alpha})$$

for any  $\alpha < 1/2$ .

b) If there exists a  $q \in (0, 1)$  such that  $C(q, q) = q$  then it holds

$$A_n(C_n) \geq q(1 - q) + o_{\mathbb{P}}(1).$$

**Proof.** We start with the proof of a). First choose  $\delta > 0$  and  $\lambda < 1$  such that

$$\frac{C(u, u)}{u} \vee \frac{1 - 2u + C(u, u)}{1 - u} \leq \lambda$$

for all  $u \in [0, \delta] \cup [1 - \delta, 1]$  and use the decomposition

$$A_n(C_n) = A_n(C_n, [0, \delta]) + A_n(C_n, [\delta, 1 - \delta]) + A_n(C_n, (1 - \delta, 1]), \quad (3.1)$$

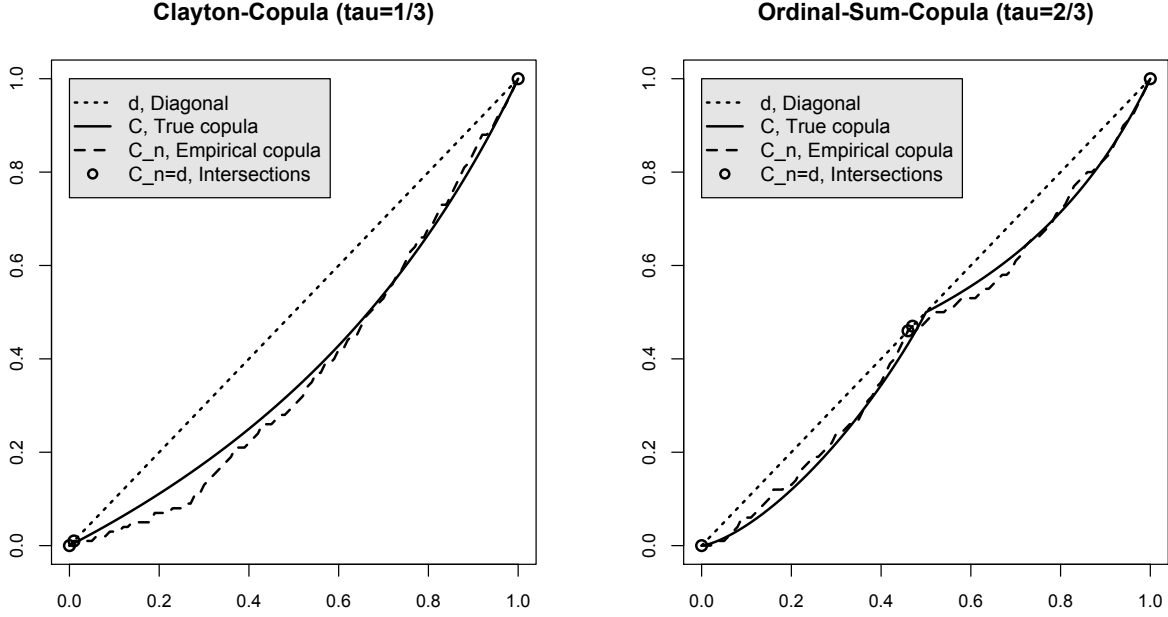


Figure 2: The solid lines show the diagonal section of a Clayton copula (left) and an ordinal sum copula (right), while the dashed line show one realization of the corresponding empirical copula. The circled points mark the locations where  $C_n(i/n, i/n) = i/n$ .

where

$$A_n(C_n, B) = \max \{i/n(1 - i/n) : C_n(i/n, i/n) = i/n \text{ and } i/n \in B\}$$

for some set  $B \subset [0, 1]$  (with the convention that  $\max \emptyset = 0$ ). Consider each term separately and define  $M_n = \sup_{u \in [0, 1]} |C_n(u, u) - C(u, u)|$ , which is of order  $O_{\mathbb{P}}(n^{-1/2})$  under Condition 2.1. Now let  $i/n \in (0, \delta)$  be such that  $C_n(i/n, i/n) = i/n$ . Due to the estimate

$$i/n(1 - \lambda) \leq i/n \left(1 - \frac{C(i/n, i/n)}{i/n}\right) = i/n - C(i/n, i/n) = C_n(i/n, i/n) - C(i/n, i/n) \leq M_n$$

we have  $i/n(1 - i/n) \leq i/n \leq M_n/(1 - \lambda)$  and we can conclude that

$$A_n(C_n, [0, \delta]) \leq \frac{M_n}{1 - \lambda} = O_{\mathbb{P}}(n^{-1/2}). \quad (3.2)$$

A similar calculation shows that for  $i/n \in (1 - \delta, 1]$  with  $C_n(i/n, i/n) = i/n$  we have  $(1 - i/n)(1 - \lambda) \leq M_n$  which in turn implies

$$A_n(C_n, (1 - \delta, 1]) \leq \frac{M_n}{1 - \lambda} = O_{\mathbb{P}}(n^{-1/2}). \quad (3.3)$$

It remains to estimate the second summand  $A_n(C_n, [\delta, 1 - \delta])$  of the decomposition (3.1). For continuity reasons we can choose a  $\kappa > 0$  such that  $u - C(u, u) \geq \kappa$  for all  $u \in [\delta, 1 - \delta]$ . If there

was a  $q \in [\delta, 1 - \delta]$  such that  $C_n(q, q) = q$ , it would follow that  $M_n \geq C_n(q, q) - C(q, q) \geq \kappa$  and therefore we have for any  $\varepsilon > 0$

$$\mathbb{P}(n^\alpha A_n(C_n, [\delta, 1 - \delta]) > \varepsilon) \leq \mathbb{P}(\exists q \in [\delta, 1 - \delta] : C_n(q, q) = q) \leq \mathbb{P}(M_n \geq \kappa) \xrightarrow{\mathbb{P}} 0.$$

A combination of (3.2) and (3.3) with this result proves part *a*) of the proposition.

For the proof of part *b*) let  $n_1 = \#\{1 \leq i \leq n : (F_1(X_{i1}), F_2(X_{i2})) \in [0, q]^2\}$ . Since  $C(q, q) = q$  implies that the mass of  $C$  is concentrated on  $(0, q)^2 \cup (q, 1)^2$  we have  $(F_1(X_{i1}), F_2(X_{i2})) \in [0, q]^2$  if and only if  $X_{i1} \leq X_{n_1:n,1}$  and  $X_{i2} \leq X_{n_1:n,2}$ , where  $X_{j:n,p} = F_{np}^-(j/n)$  denotes the  $j$ -th order statistic of  $X_{1p}, \dots, X_{np}$  (for  $p = 1, 2$ ). This yields  $C_n(n_1/n, n_1/n) = n_1/n$ , which entails the assertion by

$$A_n(C_n) \geq \frac{n_1}{n} \left(1 - \frac{n_1}{n}\right) \xrightarrow{\mathbb{P}} q(1 - q) > 0.$$

□

### Remark 3.2.

- a) The conditions on the coefficients of tail dependence in part *a*) of Proposition 3.1 can be equivalently expressed by conditions on the regular variation of the Archimedean generator of  $C$ . For a thorough discussion of these issues the reader is referred to the work of Charpentier and Segers (2009).
- b) Exploiting Theorem G.1 in Genest and Segers (2009) and Proposition 4.2 in Segers (2011) one can improve the rate of convergence in part *a*) of Proposition 3.1 to any  $\alpha < 3/4$ . It is our conjecture that the term is in fact of order  $O_{\mathbb{P}}(1/n)$ , but we were not able to derive the asymptotic distribution of  $nA_n(C_n)$ . Since we do not need a refined rate for our purposes here we omit a deeper discussion and defer these issues to future research.

From now on, suppose that the conditions of Theorem 2.3 and Proposition 3.1 hold. We can conclude that  $\mathbb{S}_{n,M}$  weakly converges to  $\mathbb{T}_{C,M}$  if the copula  $C$  is Archimedean, while  $\mathbb{S}_{n,M}$  converges to  $+\infty$  in probability if  $C$  is non-Archimedean, i.e. if it is either non-associative (by the results of Section 2) or if there exists a  $q \in (0, 1)$  such that  $C(q, q) = q$  (by Proposition 3.1). The quantiles of  $\mathbb{T}_{C,M}$  can be approximated by the multiplier method described in Section 2.2. Analogously to the discussion at the end of the Section 2.2 we can use the multiplier bootstrap to obtain an asymptotic level- $\alpha$  test for the hypothesis of Archimedeanity, which is consistent against all alternatives satisfying the Condition 2.1. Its finite sample properties will be investigated in the following section.

## 4 Finite sample properties

We conclude this paper with a simulation study regarding the finite sample performance of the proposed tests for Archimedeanity and Associativity. We consider the following six copula models:

- The Gumbel copula, which is Archimedean.
- The Clayton copula, which is Archimedean.
- The  $t$ -copula with fixed degree of freedom  $df = 1$ , which is non-associative.
- The asymmetric negative logistic model [see Joe (1990)] with fixed parameters  $\psi_1 = 2/3, \psi_2 = 1$ , which is non-associative.
- An ordinal sum model based on the partition  $J_1 = [0, 1/2], J_2 = [1/2, 1]$  and the Gumbel ( $C_1$ ) and Clayton ( $C_2$ ) copula, denoted by  $\text{Ordinal}_A$ . The model is associative.
- An ordinal sum model based on the partition  $J_1 = [0, 1/2], J_2 = [1/2, 1]$  and the two Clayton ( $C_1 = C_2$ ) copulas, denoted by  $\text{Ordinal}_B$ . The model is associative.

The parameters of the models are chosen in such a way that the coefficient of upper tail dependence  $\lambda_U$  is either  $1/4$  or  $1/2$  [for the asymmetric negative logistic model] or that Kendall's- $\tau$  is either  $1/3$  or  $2/3$  [for the remaining five models]. For  $\tau_{\text{Ordinal}_A} = 1/3$  (or  $2/3$ , resp.) we chose  $\tau_{\text{Gumbel}} = 0$  ( $1/3$ ) and  $\tau_{\text{Clayton}} = -2/3$  ( $1/3$ ), while  $\tau_{\text{Clayton}} = -1/3$  ( $1/3$ ) for  $\tau_{\text{Ordinal}_B} = 1/3$  ( $2/3$ ).

We generated 1000 random samples of sample sizes  $n = 200$  and  $n = 500$  and calculated the empirical probability of rejecting the null hypotheses of Archimedeanity or Associativity for  $M \in \{\text{L}^2, \text{KS}\}$ . For each sample of size  $n = 200, 500$  we carried out  $B = 200$  Bootstrap replications based on the multiplier method, where we chose a  $\mathcal{U}(\{0, 2\})$ -distribution for the multipliers [i.e.  $\mathbb{P}(\xi = 0) = \mathbb{P}(\xi = 2) = 1/2$ , s.t.  $\mu = \tau = 1$ ] and used  $h = n^{-1/4}$  to estimate the partial derivatives. The critical values of the tests are determined by the method described in Section 2. The penalty term  $\mathbb{S}_{n,M} - \mathbb{T}_{n,M} = k_n \phi(A_n(C_n))$  is chosen in the following data-adaptive way: first of all, we set  $\phi(x) = (4x)^2$  in order to give more emphasis to values around the maximal value of  $A_n(C_n)$  [which equals  $1/4$ ]. The constant  $k_n$  is chosen according to the distribution of the bootstrap approximation: if  $q_{0.05,M}^\xi$  denotes the 0.05-quantile of the sample  $\{\mathbb{T}_{n,M}^{\xi,b}\}_{b=1,\dots,B}$  we set  $k_n = q_{0.05,M}^\xi n^{1/4}$ . The latter choice guarantees that under  $H_0$  the error term is small compared to the distribution of  $\mathbb{T}_{C,M}$ .

The results are stated in Table 1. The entries of the table represent the empirical probabilities of rejecting the null hypothesis of Archimedeanity and of Associativity [in brackets] for both the  $\text{L}^2$ -test [first two columns] and the KS-test [last two columns]. We observe that the nominal level of the four tests are accurately approximated for the four Archimedean copulas under investigation. The  $\text{L}^2$ -test tends to be more conservative than the KS-test. Also note that the values for  $\mathbb{S}_{n,M}$  and  $\mathbb{T}_{n,M}$  differ only by a very small amount meaning that the penalty term  $k_n \phi(A_n(C_n))$  is of negligible magnitude under the null hypothesis.

The  $t$ -copula and the asymmetric negative logistic models are non-associative and the results in Table 1 reveal that these deviations are detected by both tests for Associativity, with better results for the  $t$ -copula and for stronger dependence [measured by either  $\tau$  or  $\lambda_U$ ]. The power properties of the  $\text{L}^2$ -test outclass the properties of the KS-test for all four non-associative models under investigation, such that the former test seems to be generally preferable.

Regarding the (associative) ordinal sum models both tests for associativity are very conservative. Note that the asymptotic theory developed in Section 2 does not apply for these models since the partial derivatives of the corresponding copulas are not continuous on  $\{1/2\} \times [0, 1]$  and  $[0, 1] \times \{1/2\}$ . Regarding the power properties the KS-test for Archimedeanity performs slightly better for the ordinal sum alternatives.

**Acknowledgements** This work has been supported by the Collaborative Research Center “Statistical modeling of nonlinear dynamic processes” (SFB 823, Teilprojekt A1, C1) of the German Research Foundation (DFG).

## References

- Alsina, C., Frank, M. J., and Schweizer, B. (2006). *Associative functions*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ. Triangular norms and copulas.
- Bücher, A. (2011). *Statistical inference for copulas and extremes*. PhD thesis, Ruhr-University Bochum, Germany.
- Bücher, A. and Dette, H. (2010). A note on bootstrap approximations for the empirical copula process. *Statist. Probab. Lett.*, 80:1925–1932.
- Charpentier, A. and Segers, J. (2009). Tails of multivariate Archimedean copulas. *J. Multivariate Anal.*, 100(7):1521–1537.
- Fermanian, J.-D., Radulović, D., and Wegkamp, M. (2004). Weak convergence of empirical copula processes. *Bernoulli*, 10(5):847–860.
- Genest, C. and MacKay, R. J. (1986). Copules archimédiennes et familles de lois bidimensionnelles dont les marges sont données. *Canad. J. Statist.*, 14(2):145–159.
- Genest, C., Nešlehová, J., and Ziegel, J. (2011). Inference in multivariate archimedean copula models. *Test*. to appear.
- Genest, C. and Rivest, L.-P. (1993). Statistical inference procedures for bivariate Archimedean copulas. *J. Amer. Statist. Assoc.*, 88(423):1034–1043.
- Genest, C. and Segers, J. (2009). Rank-based inference for bivariate extreme-value copulas. *Ann. Statist.*, 37(5B):2990–3022.
- Jaworski, P. (2010). Testing archimedeanity. In Borgelt, C., González-Rodríguez, G., Trutschnig, W., Lubiano, M., Gil, M., Grzegorzewski, P., and Hryniewicz, O., editors, *Combining Soft Computing and Statistical Methods in Data Analysis*, volume 77 of *Advances in Intelligent and Soft Computing*, pages 353–360. Springer, Berlin.



	L 2-Test		KS-Test	
	0.1	0.05	0.1	0.05
<b>n=200:</b>				
Clayton( $\tau = 1/3$ )	0.071 (0.071)	0.038 (0.037)	0.088 (0.088)	0.050 (0.050)
Clayton( $\tau = 2/3$ )	0.030 (0.016)	0.011 (0.009)	0.124 (0.078)	0.068 (0.036)
Gumbel( $\tau = 1/3$ )	0.079 (0.079)	0.043 (0.043)	0.082 (0.082)	0.046 (0.045)
Gumbel( $\tau = 2/3$ )	0.034 (0.032)	0.015 (0.013)	0.108 (0.098)	0.065 (0.057)
$t(\tau = 1/3, df = 1)$	0.953 (0.953)	0.886 (0.884)	0.562 (0.558)	0.380 (0.376)
$t(\tau = 2/3, df = 1)$	0.748 (0.726)	0.592 (0.564)	0.392 (0.355)	0.258 (0.231)
Aneglog( $\lambda_U = 0.25$ )	0.112 (0.112)	0.061 (0.061)	0.105 (0.105)	0.059 (0.059)
Aneglog( $\lambda_U = 0.5$ )	0.641 (0.641)	0.536 (0.536)	0.363 (0.356)	0.225 (0.222)
Ordinal <sub>A</sub> ( $\tau = 1/3$ )	0.996 (0.000)	0.827 (0.000)	1.000 (0.012)	1.000 (0.005)
Ordinal <sub>A</sub> ( $\tau = 2/3$ )	1.000 (0.004)	1.000 (0.001)	1.000 (0.079)	1.000 (0.041)
Ordinal <sub>B</sub> ( $\tau = 1/3$ )	1.000 (0.000)	1.000 (0.000)	1.000 (0.045)	1.000 (0.021)
Ordinal <sub>B</sub> ( $\tau = 2/3$ )	1.000 (0.006)	1.000 (0.001)	1.000 (0.057)	1.000 (0.030)
<b>n=500:</b>				
Clayton( $\tau = 1/3$ )	0.082 (0.082)	0.051 (0.051)	0.088 (0.088)	0.036 (0.036)
Clayton( $\tau = 2/3$ )	0.062 (0.059)	0.032 (0.027)	0.090 (0.082)	0.046 (0.039)
Gumbel( $\tau = 1/3$ )	0.091 (0.091)	0.045 (0.045)	0.105 (0.105)	0.050 (0.050)
Gumbel( $\tau = 2/3$ )	0.072 (0.070)	0.033 (0.032)	0.124 (0.121)	0.068 (0.066)
$t(\tau = 1/3, df = 1)$	1.000 (1.000)	1.000 (1.000)	0.954 (0.953)	0.871 (0.871)
$t(\tau = 2/3, df = 1)$	0.998 (0.990)	0.998 (0.990)	0.818 (0.811)	0.655 (0.650)
Aneglog( $\lambda_U = 0.25$ )	0.237 (0.237)	0.173 (0.173)	0.124 (0.124)	0.069 (0.069)
Aneglog( $\lambda_U = 0.5$ )	0.979 (0.979)	0.947 (0.947)	0.716 (0.716)	0.584 (0.584)
Ordinal <sub>A</sub> ( $\tau = 1/3$ )	1.000 (0.000)	0.980 (0.000)	1.000 (0.022)	1.000 (0.007)
Ordinal <sub>A</sub> ( $\tau = 2/3$ )	1.000 (0.021)	1.000 (0.009)	1.000 (0.093)	1.000 (0.038)
Ordinal <sub>B</sub> ( $\tau = 1/3$ )	1.000 (0.000)	1.000 (0.000)	1.000 (0.005)	1.000 (0.023)
Ordinal <sub>B</sub> ( $\tau = 2/3$ )	1.000 (0.013)	1.000 (0.004)	1.000 (0.082)	1.000 (0.037)

Table 1: *Simulated rejection probabilities for the multiplier bootstrap-tests for Archimedeanity and for Associativity (in brackets). The sample size is  $n = 200$  or  $n = 500$ ,  $B = 200$  Bootstrap-replicates and 1000 simulation runs have been performed. The first four lines are Archimedean copula models, the  $t$ -models are not associative and the ordinal sum-models are associative, but not Archimedean.*

- Joe, H. (1990). Families of min-stable multivariate exponential and multivariate extreme value distributions. *Statist. Probab. Lett.*, 9(1):75–81.
- Kosorok, M. R. (2008). *Introduction to Empirical Processes and Semiparametric Inference*. Springer, New York.
- Ling, C.-h. (1965). Representation of associative functions. *Publ. Math. Debrecen*, 12:189–212.
- McNeil, A. J. and Nešlehová, J. (2009). Multivariate Archimedean copulas,  $d$ -monotone functions and  $l_1$ -norm symmetric distributions. *Ann. Statist.*, 37(5B):3059–3097.
- Naifar, N. (2011). Modelling dependence structure with Archimedean copulas and applications to the iTraxx CDS index. *Journal of Computational and Applied Mathematics*, 235:2459–2466.
- Nelsen, R. B. (2006). *An introduction to copulas*. Springer Series in Statistics. Springer, New York, second edition.
- Rémillard, B. and Scaillet, O. (2009). Testing for equality between two copulas. *Journal of Multivariate Analysis*, 100:377–386.
- Rivest, L.-P. and Wells (2001). A martingale approach to the copula-graphic estimator for the survival function under dependent censoring. *Journal of Multivariate Analysis*, 79:138–.
- Rüschendorf, L. (1976). Asymptotic distributions of multivariate rank order statistics. *Annals of Statistics*, 4:912–923.
- Scaillet, O. (2005). A Kolmogorov-Smirnov type test for positive quadrant dependence. *Canad. J. Statist.*, 33(3):415–427.
- Segers, J. (2010). Weak convergence of empirical copula processes under nonrestrictive smoothness assumptions. *arXiv:1012.2133v1*.
- Segers, J. (2011). Asymptotics of empirical copula processes under nonrestrictive smoothness assumptions. *arXiv:1012.2133v2*.
- Sklar, A. (1959). Fonctions de répartition à  $n$  dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, 8:229–231.
- Tsukahara, H. (2005). Semiparametric estimation in copula models. *Canad. J. Statist.*, 33(3):357–375.
- Van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes - Springer Series in Statistics*. Springer, New York.
- Wang, W. and Wells, M. T. (2000). Model selection and semiparametric inference for bivariate failure-time data. *Journal of the American Statistical Association*, 95(449):62–72.