

# ON WEAK CONDITIONAL CONVERGENCE OF BIVARIATE ARCHIMEDEAN AND EXTREME VALUE COPULAS, AND CONSEQUENCES TO NONPARAMETRIC ESTIMATION

THIMO M. KASPER, SEBASTIAN FUCHS, WOLFGANG TRUTSCHNIG

## KEYWORDS

Archimedean copula, Extreme Value copula, Checkerboard copula, weak convergence, estimation, dependence measure

**ABSTRACT.** Looking at bivariate copulas from the perspective of conditional distributions and considering weak convergence of almost all conditional distributions yields the notion of weak conditional convergence. At first glance, this notion of convergence for copulas might seem far too restrictive to be of any practical importance - in fact, given samples of a copula  $C$  the corresponding empirical copulas do not converge weakly conditional to  $C$  with probability one in general. Within the class of Archimedean copulas and the class of Extreme Value copulas, however, standard pointwise convergence and weak conditional convergence can even be proved to be equivalent. Moreover, it can be shown that every copula  $C$  is the weak conditional limit of a sequence of checkerboard copulas. After proving these three main results and pointing out some consequences we sketch some implications for two recently introduced dependence measures and for the nonparametric estimation of Archimedean and Extreme Value copulas.

## 1. INTRODUCTION

Suppose that  $\{C_\theta : \theta \in \Theta\}$  is a parametric class of bivariate copulas with  $\Theta \subseteq \mathbb{R}^d$  for some  $d \in \mathbb{N}$  and let  $\{K_\theta : \theta \in \Theta\}$  denote the corresponding conditional distributions (Markov kernels), i.e., if  $X, Y$  are uniformly distributed on  $[0, 1]$  and  $(X, Y)$  has distribution function  $C_\theta$  then  $K_\theta(x, E) = \mathbb{P}(Y \in E | X = x)$ . Many standard classes of copulas are not only continuous in the parameter with respect to pointwise/uniform convergence (see [10, 30]) but exhibit the even stronger property that if  $(\theta_n)_{n \in \mathbb{N}}$  converges to  $\theta$  then almost all conditional distributions  $(K_{\theta_n}(x, \cdot))_{n \in \mathbb{N}}$  converge weakly to  $K_\theta(x, \cdot)$ . In the sequel we will refer to weak convergence of almost all conditional distributions as weak conditional convergence. It is straightforward to verify that (among many others) the family of Gaussian copulas and the family of  $t$ -copulas exhibit the just mentioned continuity with respect to the parameter. Moreover, leaving the absolutely continuous setting, the same is true, e.g., for the Marshall-Olkin family.

Despite the afore-mentioned examples, at first glance, weak conditional convergence might seem as a concept far too restrictive to be of any practical importance outside the purely parametric setting. This impression is reinforced by the fact that given samples  $(X_1, Y_1), (X_2, Y_2), \dots$  from a copula  $C$  and letting  $\hat{E}_n$  denote the corresponding empirical copula (bivariate interpolation of the induced subcopula, see [30]) we do not have weak conditional convergence of  $(\hat{E}_n)_{n \in \mathbb{N}}$  to  $C$  unless  $C$  is completely dependent in the sense that random variable  $Y$  is a measurable function of random variable  $X$  (see [26]).

As we will demonstrate in this contribution, however, within the class of Archimedean copulas and the class of Extreme Value copulas (neither of them being a parametric class of the afore-mentioned type) standard pointwise/uniform convergence and weak conditional convergence are even equivalent, a result having direct implications for the dependence

measures  $\zeta_1$  and  $r$  introduced in [36] and [9], respectively, as well as for the nonparametric estimation of Archimedean and Extreme Value copulas (see [1, 13, 14] and [15, 19]). We will show that convexity of the univariate ‘generating’ functions (the normalized generator in the Archimedean and the Pickands dependence function in the Extreme Value case) is the key property entailing weak conditional convergence. Additionally, building upon the theorems in [27] we will derive a universal approximation result with respect to weak conditional convergence and show that for every bivariate copula  $C$  we can find a sequence  $(C_n)_{n \in \mathbb{N}}$  of checkerboard copulas that converges weakly conditional to the copula  $C$ .

The authors’ interest in studying convergence of Archimedean copulas was triggered by [4] where the authors among other things showed that pointwise/uniform convergence of a sequence of Archimedean copulas to an Archimedean copula is equivalent to convergence of the corresponding sequence of Kendall distribution functions. In our contribution we first derive a slightly modified version of this result (including the fact that we can have convergence of the copulas without having convergence of the corresponding generators in 0, see Theorem 4.1) and then go one step further (see Theorem 4.2) and prove the equivalence of six different notions of convergence (some involving the copulas, some the generators), weak conditional convergence being one of them.

The rest of this contribution is organized as follows: Section 2 gathers preliminaries and notations that will be used throughout the paper. In Section 3 we formally define weak conditional convergence, prove that checkerboard copulas are dense with respect to weak conditional convergence, and show that weak conditional convergence implies convergence with respect to the metric  $D_1$  introduced in [36] but in general not vice versa. Section 4 derives the afore-mentioned equivalence of pointwise/uniform and weak conditional convergence within the family of Archimedean copulas in several steps. In Section 5 we prove an analogous characterization of convergence within the class of Extreme Value copulas. Direct consequences of these two main results to the estimation of Archimedean and Extreme Value copulas are sketched and illustrated via simulations in Section 6. Finally, we use the obtained results for estimating the recently introduced coefficient of correlation (see [5]) and compare the performance of the estimators incorporating or ignoring the Extreme Value/Archimedean information.

## 2. NOTATION AND PRELIMINARIES

In the sequel we will let  $\mathcal{C}$  denote the family of all bivariate copulas. For each copula  $C$  the corresponding doubly stochastic measure will be denoted by  $\mu_C$ , i.e.  $\mu_C([0, x] \times [0, y]) = C(x, y)$  for all  $x, y \in [0, 1]$ ,  $\mathcal{P}_{\mathcal{C}}$  will denote the family of all doubly stochastic measures. For more background on copulas and doubly stochastic measures we refer to [10, 30]. For every metric space  $(S, d)$  the Borel  $\sigma$ -field on  $S$  will be denoted by  $\mathcal{B}(S)$ .

In what follows Markov kernels will play a prominent role. A *Markov kernel* from  $\mathbb{R}$  to  $\mathbb{R}$  is a mapping  $K : \mathbb{R} \times \mathcal{B}(\mathbb{R}) \rightarrow [0, 1]$  such that for every fixed  $E \in \mathcal{B}(\mathbb{R})$  the mapping  $x \mapsto K(x, E)$  is (Borel-)measurable and for every fixed  $x \in \mathbb{R}$  the mapping  $E \mapsto K(x, E)$  is a probability measure. Given two real-valued random variables  $X, Y$  on a probability space  $(\Omega, \mathcal{A}, \mathbb{P})$  we say that a Markov kernel  $K$  is a *regular conditional distribution* of  $Y$  given  $X$  if  $K(X(\omega), E) = \mathbb{E}(\mathbb{1}_E \circ Y | X)(\omega)$  holds  $\mathbb{P}$ -almost surely for every  $E \in \mathcal{B}(\mathbb{R})$ . It is well-known (see, e.g., [22, 25]) that for  $X, Y$  as above, a regular conditional distribution of  $Y$  given  $X$  always exists and is unique for  $\mathbb{P}^X$ -a.e.  $x \in \mathbb{R}$ . If  $(X, Y)$  has distribution function  $H$  (in which case we will also write  $(X, Y) \sim H$  and let  $\mu_H$  denote the corresponding probability measure on  $\mathcal{B}(\mathbb{R}^2)$ ) we will let  $K_H$  denote (a version of) the regular conditional distribution of  $Y$  given  $X$  and simply refer to it as *Markov kernel of  $H$* . If  $C$  is a copula then we will consider the Markov kernel of  $C$

automatically as mapping  $K_C : [0, 1] \times \mathcal{B}([0, 1]) \rightarrow [0, 1]$ . Defining the  $x$ -section of a set  $G \in \mathcal{B}(\mathbb{R}^2)$  as  $G_x := \{y \in \mathbb{R} : (x, y) \in G\}$  the so-called disintegration theorem (see [22, 25]) yields

$$(2.1) \quad \int_{\mathbb{R}} K_H(x, G_x) \, d\mathbb{P}^X(x) = \mu_H(G).$$

As a direct consequence, for every  $C \in \mathcal{C}$  we get

$$\int_{[0,1]} K_C(x, E) \, d\mathbb{P}^X(x) = \int_{[0,1]} K_C(x, E) \, d\lambda(x) = \lambda(E)$$

for every  $E \in \mathcal{B}([0, 1])$ , whereby  $\lambda$  denotes the Lebesgue measure on  $\mathbb{R}$ . For more background on conditional expectation and general disintegration we refer to [22, 25].

We call a copula  $C$  *completely dependent* if there exists a  $\lambda$ -preserving transformation  $h : [0, 1] \rightarrow [0, 1]$  (i.e., a transformation fulfilling  $\lambda(h^{-1}(E)) = \lambda(E)$  for every  $E \in \mathcal{B}([0, 1])$ ) such that  $K(x, E) := \mathbf{1}_E(h(x))$  is a Markov kernel of  $C$ . For more properties of complete dependence we refer to [26] as well as to [36] and the references therein.

Markov kernels can be used to define metrics stronger than the standard *uniform metric*  $d_\infty$ , defined by

$$(2.2) \quad d_\infty(C_1, C_2) := \max_{(x,y) \in [0,1]^2} |C_1(x, y) - C_2(x, y)|,$$

on  $\mathcal{C}$ . It is well known that the metric space  $(\mathcal{C}, d_\infty)$  is compact and that pointwise and uniform convergence of a sequence of copulas  $(C_n)_{n \in \mathbb{N}}$  are equivalent (see [10, 37]). Following [36] and defining

$$(2.3) \quad \begin{aligned} D_1(C_1, C_2) &:= \int_{[0,1]} \int_{[0,1]} |K_{C_1}(x, [0, y]) - K_{C_2}(x, [0, y])| \, d\lambda(x) d\lambda(y), \\ D_2^2(C_1, C_2) &:= \int_{[0,1]} \int_{[0,1]} (K_{C_1}(x, [0, y]) - K_{C_2}(x, [0, y]))^2 \, d\lambda(x) d\lambda(y), \\ D_\infty(C_1, C_2) &:= \sup_{y \in [0,1]} \int_{[0,1]} |K_{C_1}(x, [0, y]) - K_{C_2}(x, [0, y])| \, d\lambda(x) \end{aligned}$$

it can be shown that  $D_1, D_2, D_\infty$  are metrics generating the same topology on  $\mathcal{C}$ . In the sequel we will mainly work with  $D_1$  and refer to [11, 36] for more information on  $D_2$  and  $D_\infty$ . The metric space  $(\mathcal{C}, D_1)$  is complete and separable but not compact. Moreover, we have  $D_1(C, \Pi) \in [0, \frac{1}{3}]$  for every  $C \in \mathcal{C}$  and  $D_1(C, \Pi)$  is maximal if and only if  $C$  is completely dependent. The metric  $D_1$  was originally introduced in order to construct a dependence measure which, contrary to  $d_\infty$ , is capable of separating independence and complete dependence. The resulting  $D_1$ -based dependence measure  $\zeta_1$  introduced in [36] is defined as

$$(2.4) \quad \zeta_1(C) := 3 \cdot D_1(C, \Pi)$$

for every  $C \in \mathcal{C}$ . In the sequel we will also consider the dependence measure  $r(X, Y) = r(C)$  introduced in [9] as

$$(2.5) \quad r(C) := 6 \cdot \int_{[0,1]} \int_{[0,1]} K_C(x, [0, y])^2 d\lambda(x) d\lambda(y) - 2.$$

It is straightforward to verify that  $r(C)$  can be expressed in terms of  $D_2$  and that  $r(C) = 6 \cdot D_2^2(C, \Pi)$  holds. Both  $\zeta_1$  and  $r$  attain values in  $[0, 1]$ , are 0 if, and only if  $C = \Pi$ , and 1 if, and only if  $C$  is completely dependent.

A symmetric version of  $D_1$ -convergence (or, equivalently,  $D_2$ -convergence) was introduced by Mikusiński and Taylor in [29] under the name  $\partial$ -convergence: A sequence  $(C_n)_{n \in \mathbb{N}}$   $\partial$ -converges to a copula  $C \in \mathcal{C}$  ( $C_n \xrightarrow{\partial} C$  for short) if and only if

$$\lim_{n \rightarrow \infty} D_1(C_n, C) + D_1(C_n^t, C^t) = 0.$$

We will see in the next section that in general weak conditional convergence does not imply  $\partial$ -convergence (Example 3.4) nor vice versa. For a thorough survey of different notions on convergence of copulas we refer to [35].

### 3. WEAK CONDITIONAL CONVERGENCE AND CHECKERBOARDS

Sticking to the idea of viewing bivariate copulas in terms of their conditional distributions and considering weak convergence gives rise to what we refer to as weak conditional convergence in the sequel:

**Definition 3.1.** Suppose that  $C, C_1, C_2, \dots$  are copulas and let  $K_C, K_{C_1}, K_{C_2}, \dots$  be (versions of) the corresponding Markov kernels. We will say that  $(C_n)_{n \in \mathbb{N}}$  converges *weakly conditional* to  $C$  if and only if for  $\lambda$ -almost every  $x \in [0, 1]$  we have that the sequence  $(K_{C_n}(x, \cdot))_{n \in \mathbb{N}}$  of probability measures on  $\mathcal{B}([0, 1])$  converges weakly to the probability measure  $K_C(x, \cdot)$ . In the latter case we will write  $C_n \xrightarrow{\text{wcc}} C$  (where ‘wcc’ stands for ‘weak conditional convergence’).

As already mentioned in the Introduction, many standard parametric classes  $\{C_\theta : \theta \in \Theta\}$  of copulas depend on the parameter  $\theta$  weakly conditional in the sense that if  $(\theta_n)_{n \in \mathbb{N}}$  converges to  $\theta$  then the corresponding sequence  $(C_{\theta_n})_{n \in \mathbb{N}}$  converges weakly conditional to  $C_\theta$ . This is obviously true for parametric classes  $\{C_\theta : \theta \in \Theta\}$  of absolutely continuous copulas whose corresponding densities  $\{k_\theta : \theta \in \Theta\}$  have the property that if  $(\theta_n)_{n \in \mathbb{N}}$  converges to  $\theta \in \Theta$  then  $(k_{\theta_n})_{n \in \mathbb{N}}$  converges to  $k_\theta$   $\lambda_2$ -almost everywhere (whereby  $\lambda_2$  denotes the two-dimensional Lebesgue measure on  $\mathcal{B}(\mathbb{R}^2)$ ). In fact, in this case the corresponding Markov kernels  $K_{C_{\theta_n}}$  are given by

$$(3.1) \quad K_{C_{\theta_n}}(x, [0, y]) = \int_{[0, y]} k_{\theta_n}(x, s) d\lambda(s)$$

and if we let  $\Lambda \in \mathcal{B}([0, 1]^2)$  denote the set of all points  $(x, y)$  fulfilling  $\lim_{n \rightarrow \infty} k_{\theta_n}(x, y) = k_\theta(x, y)$  then disintegration yields  $\lambda(\Lambda_x) = 1$  for  $\lambda$ -almost every  $x \in [0, 1]$ , so the property follows immediately. It is straightforward to verify that (among many others) the family of Gaussian copulas and the family of  $t$ -copulas fulfill this property.

The same is true for other, not necessarily absolutely continuous classes like the Marshall-Olkin family  $(M_{\alpha, \beta})_{(\alpha, \beta) \in [0, 1]^2}$  (see [7, 10, 30]) given by

$$(3.2) \quad M_{\alpha, \beta}(x, y) = \begin{cases} x^{1-\alpha} y & \text{if } x^\alpha \geq y^\beta \\ x y^{1-\beta} & \text{if } x^\alpha < y^\beta. \end{cases}$$

According to [36] the corresponding Markov kernel  $K_{M_{\alpha, \beta}}$  is given by

$$(3.3) \quad K_{M_{\alpha, \beta}}(x, [0, y]) = \begin{cases} (1 - \alpha)x^{-\alpha} y & \text{if } y^\beta < x^\alpha \\ y^{1-\beta} & \text{if } y^\beta \geq x^\alpha \end{cases}$$

and it is straightforward to verify that if the parameter vector  $(\alpha_n, \beta_n)_{n \in \mathbb{N}}$  converges to  $(\alpha, \beta)$  then we also have  $M_{\alpha_n, \beta_n} \xrightarrow{\text{wcc}} M_{\alpha, \beta}$ .

Before focusing on the Archimedean and the Extreme Value setting we prove a general approximation result saying that the class of checkerboard copulas is dense in  $\mathcal{C}$  w.r.t. weak conditional convergence (see Theorem 3.2). Recall that a copula  $C$  is called a *checkerboard copula* with resolution  $N \in \mathbb{N}$  if and only if  $\mu_C$  distributes its mass uniformly on each rectangle  $R_{ij}^N = [\frac{i-1}{N}, \frac{i}{N}] \times [\frac{j-1}{N}, \frac{j}{N}]$  with  $i, j \in \{1, \dots, N\}$ . We will refer to  $\mathcal{S}_N$  as the family of all checkerboard copulas with resolution  $N$ , the set  $\mathcal{S} = \bigcup_{N=1}^{\infty} \mathcal{S}_N$  is called the class of all checkerboard copulas (checkerboards for short). For every  $N \in \mathbb{N}$  the (unique) checkerboard copula  $\mathfrak{CB}_N(C) \in \mathcal{S}_N$  fulfilling

$$\mu_C(R_{ij}^N) = \mu_{\mathfrak{CB}_N(C)}(R_{ij}^N)$$

for all  $i, j \in \{1, \dots, N\}$  will be referred to as *N-checkerboard approximation of the copula C*.

It is well-known that the class of checkerboard copulas  $\mathcal{S}$  is dense in  $(\mathcal{C}, d_{\infty})$  and in  $(\mathcal{C}, D_1)$ , see [10, 27, 36]. The following theorem implies these two interrelations:

**Theorem 3.2.** *Given a copula  $C$ , there is a sequence of elements of  $\mathcal{S}$  that converges to  $C$  with respect to weak conditional convergence. In other words,  $\mathcal{S}$  is dense in  $\mathcal{C}$  with respect to weak conditional convergence.*

*Proof.* Fix  $C \in \mathcal{C}$  and suppose that  $K_C$  is a Markov kernel of  $C$ . We are going to show that  $(\mathfrak{CB}_{2^n}(C))_{n \in \mathbb{N}}$  converges weakly conditional to  $C$ . Using Lipschitz continuity of copulas in each coordinate, for every  $y$  of the form  $y = \frac{j}{2^m}$  with  $m \in \mathbb{N}$  and  $j \in \{0, \dots, 2^m\}$  there exists a set  $\Lambda_y \in \mathcal{B}([0, 1])$  with the following three properties:

- (1) For every  $x \in \Lambda_y$  the function  $t \mapsto C(t, y)$  is differentiable at  $x$  and fulfills  $\frac{\partial C}{\partial x}(x, y) = K_C(x, [0, y])$ ,
- (2)  $\lambda(\Lambda_y) = 1$ ,
- (3)  $\Lambda_y \subseteq \left( \bigcup_{l=1}^{\infty} \{0, \frac{1}{2^l}, \frac{2}{2^l}, \dots, \frac{2^l-1}{2^l}, 1\} \right)^c$ .

Letting  $\mathcal{W}$  denote the set of all points  $y$  of the form  $y = \frac{j}{2^m}$  with  $m \in \mathbb{N}$  and  $j \in \{0, \dots, 2^m\}$  it follows that  $\mathcal{W}$  is countably infinite, hence setting  $\Lambda = \bigcap_{y \in \mathcal{W}} \Lambda_y$  yields  $\lambda(\Lambda) = 1$ .

Consider  $y = \frac{j}{2^m} \in \mathcal{W}$  and  $x \in \Lambda$ . For every  $n \geq m$  there exists exactly one index  $i_n(x) \in \{0, \dots, 2^n\}$  with

$$x \in \left( \frac{i_n(x)-1}{2^n}, \frac{i_n(x)}{2^n} \right).$$

Considering that  $t \mapsto K_{\mathfrak{CB}_{2^n}(C)}(t, \cdot)$  is constant on the interval  $\left( \frac{i_n(x)-1}{2^n}, \frac{i_n(x)}{2^n} \right)$  disintegration yields

$$\begin{aligned} C\left(\frac{i_n(x)}{2^n}, y\right) - C\left(\frac{i_n(x)-1}{2^n}, y\right) &= \int_{\left(\frac{i_n(x)-1}{2^n}, \frac{i_n(x)}{2^n}\right]} K_{\mathfrak{CB}_{2^n}(C)}(t, [0, y]) d\lambda(t) \\ &= \frac{1}{2^n} K_{\mathfrak{CB}_{2^n}(C)}(x, [0, y]), \end{aligned}$$

from which we directly get

$$K_{\mathfrak{CB}_{2^n}(C)}(x, [0, y]) = \frac{C\left(\frac{i_n(x)}{2^n}, y\right) - C\left(\frac{i_n(x)-1}{2^n}, y\right)}{\frac{1}{2^n}} \xrightarrow{n \rightarrow \infty} \frac{\partial C}{\partial x}(x, y) = K_C(x, [0, y]).$$

Since  $(x, y) \in \Lambda \times \mathcal{W}$  was arbitrary we have shown that for each  $x \in \Lambda$  the conditional distribution functions  $y \mapsto K_{\mathcal{CB}_{2^n}(C)}(x, [0, y])$  converge to  $y \mapsto K_C(x, [0, y])$  for every  $y \in \mathcal{W}$ , i.e., on a dense set. Having this, weak conditional convergence of  $(\mathcal{CB}_{2^n}(C))_{n \in \mathbb{N}}$  to  $C$  follows immediately.  $\square$

We conclude this section with an example clarifying the interrelation between the aforementioned notions of convergence: According to Lemma 7 in [36] weak conditional convergence of  $(C_n)_{n \in \mathbb{N}}$  to  $C$  implies convergence w.r.t.  $D_1$ . Additionally, convergence w.r.t.  $D_1$  implies convergence in  $d_\infty$  but not vice versa. The following simple example shows that we can have  $D_1$ -convergence without having weak conditional convergence.

**Example 3.3.** Let  $N \in \mathbb{N}, i \in \{1, \dots, N\}$  and  $n = 2^N + i - 2$ . Define the copula  $C_n$  via its Markov kernel by

$$K_{C_n}(x, [0, y]) = \begin{cases} \mathbb{1}_{[0, y]}(2^N \cdot x + 1 - i) & \text{if } x \in [\frac{i-1}{2^N}, \frac{i}{2^N}] \\ y & \text{if } x \in [0, 1] \setminus [\frac{i-1}{2^N}, \frac{i}{2^N}]. \end{cases}$$

Then  $C_n$  does not converge weakly conditional to  $\Pi$  since for every  $x \in [0, 1]$  the probability measure  $K_{C_n}(x, \cdot)$  is, on the one hand, degenerated for infinitely many  $n \in \mathbb{N}$  and, on the other hand, coincides with  $\lambda$  restricted to  $[0, 1]$  infinitely many times too. Nevertheless  $C_n$  converges to  $\Pi$  w.r.t.  $D_1$  since

$$D_1(C_n, \Pi) \leq \frac{1}{2^N}$$

holds and if  $n$  goes to infinity then so does  $N$ .

We conclude this section with an example showing that, in contrast to convergence w.r.t. to  $d_\infty$ , neither convergence w.r.t.  $D_1$  nor weak conditional convergence is a symmetric concept, i.e., considering Markov kernels  $K_{C_n^t}(x, \cdot)$  instead of  $K_{C_n}(x, \cdot)$ ,  $n \in \mathbb{N}$ , may yield a different notion of convergence (as usual  $C^t$  denotes the transpose of  $C$ , i.e.  $C^t(x, y) = C(y, x)$ ).

**Example 3.4.** For every  $n \in \mathbb{N}$  define the  $\lambda$ -preserving transformation  $h_n : [0, 1] \rightarrow [0, 1]$  by

$$h_n(x) := 2^n x \pmod{1}$$

and let  $C_n$  denote the corresponding completely dependent copula. Since  $K_{C_n^t}(x, \cdot)$  is the (discrete) uniform distribution on  $\{\frac{x}{2^n} + \frac{i}{2^n} : i \in \{0, \dots, 2^n - 1\}\}$ ,  $K_{C_n^t}(x, \cdot)$  converges weakly to  $K_\Pi(x, \cdot)$  for  $\lambda$ -almost every  $x \in [0, 1]$  and we have

$$C_n^t \xrightarrow{wcc} \Pi,$$

implying that  $(C_n^t)_{n \in \mathbb{N}}$  converges to  $\Pi$  w.r.t.  $D_1$ . Additionally (see [36, Theorem 14])  $D_1(C_n, \Pi) = 1/3$  holds for every  $n \in \mathbb{N}$  from which it follows immediately that  $(C_n)_{n \in \mathbb{N}}$  does not converge to  $\Pi$  neither weakly conditional nor w.r.t.  $D_1$ . It can even be shown that  $(C_n)_{n \in \mathbb{N}}$  does not converge in  $(\mathcal{C}, D_1)$  at all: If  $(C_n)_{n \in \mathbb{N}}$  would converge to some copula  $C$  w.r.t.  $D_1$  then according to [36, Proposition 15])  $C$  would be completely dependent, i.e., there would be some  $\lambda$ -preserving transformation  $h$  such that  $C = C_h$  holds. Considering that  $D_1$ -convergence implies  $d_\infty$ -convergence  $C_h^t = \Pi$  would follow, a contradiction since  $C_h^t$  is singular whereas  $\Pi$  is absolutely continuous.

**Remark 3.5.** The preceding examples imply that weak conditional convergence does not imply  $\partial$ -convergence (Example 3.4) nor vice versa (Example 3.3).

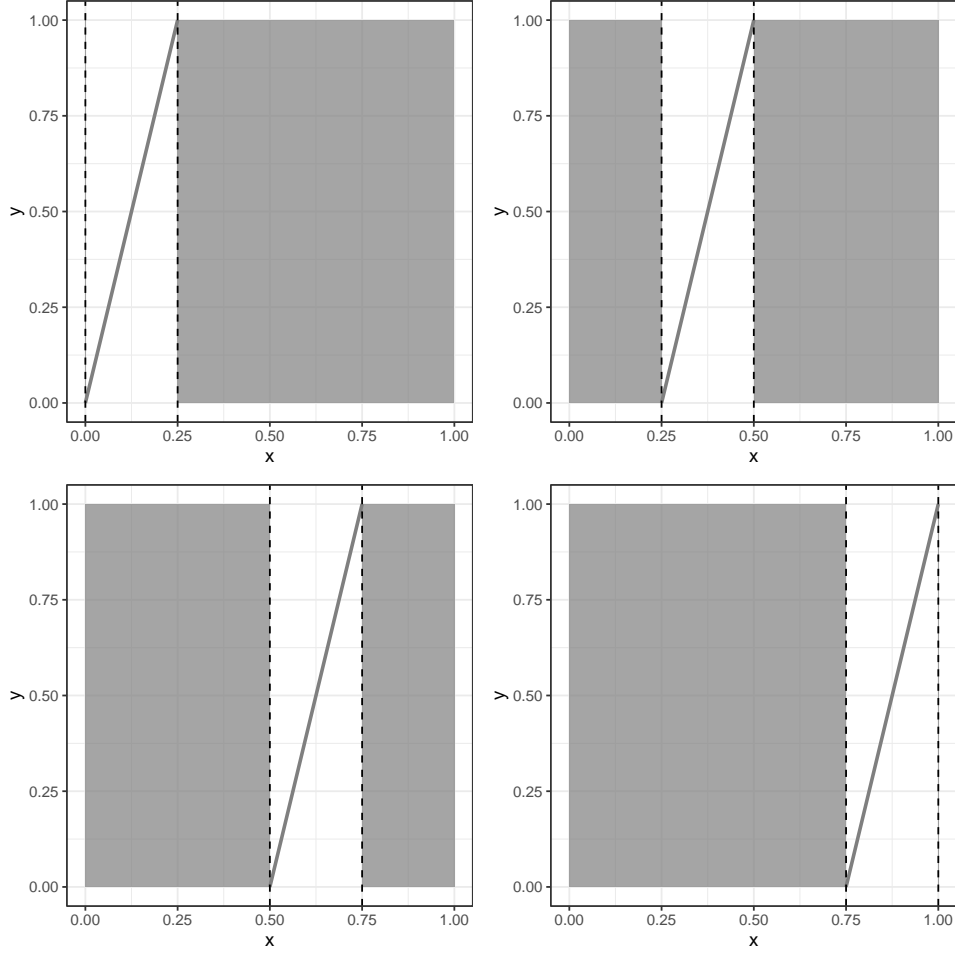


FIGURE 1. Supports of the copulas  $C_3, C_4, C_5, C_6$  considered in Example 3.3.

#### 4. ARCHIMEDEAN COPULAS

Recall that a generator of a bivariate Archimedean copula (see [30]) is a convex and strictly decreasing function  $\varphi : [0, 1] \rightarrow [0, \infty]$  with  $\varphi(1) = 0$ . Every generator induces a copula  $C$  via

$$(4.1) \quad C(x, y) = \varphi^{-}(\varphi(x) + \varphi(y)), \quad x, y \in [0, 1]$$

where  $\varphi^{-} : [0, \infty] \rightarrow [0, 1]$  denotes the pseudoinverse of  $\varphi$  defined by

$$(4.2) \quad \varphi^{-}(x) := \begin{cases} \varphi^{-1}(x) & \text{if } x \in [0, \varphi(0+)) \\ 0 & \text{if } x \geq \varphi(0+) \end{cases}$$

where  $\varphi(x\pm) := \lim_{t \rightarrow x\pm} \varphi(t)$  denotes the respective one-sided limit. We refer to  $C$  as the Archimedean copula induced by  $\varphi$  and call  $C$  *strict* if  $\varphi(0+) = \infty$  and *non-strict* otherwise. In what follows  $\mathcal{C}_{ar}$  will denote the family of all bivariate Archimedean copulas.

Since  $\varphi$  is convex obviously  $\varphi(0) \geq \varphi(0+)$  holds. Defining the right-continuous version  $\psi$  of  $\varphi$  by

$$\psi(t) := \begin{cases} \varphi(0+) & \text{if } t = 0 \\ \varphi(t) & \text{if } t \in (0, 1] \end{cases}$$

it is straightforward to verify that  $\psi$  and  $\varphi$  generate the same copula. In other words, the value of  $\varphi$  at 0 is irrelevant and we may, without loss of generality, from now on assume

that all generators are right-continuous at 0. Additionally, since for every generator  $\varphi$  and every constant  $a > 0$  we have that  $a \cdot \varphi$  generates the same copula we will from now on also assume (without explicit reference) that the generator is normalized in the sense that  $\varphi(\frac{1}{2}) = 1$  holds.

Following [12, 30] we define the  $t$ -level set  $L_t$  of the Archimedean copula  $C$  by

$$L_t := \{(x, y) \in [0, 1]^2 : C(x, y) = t\}$$

and the  $t$ -level function  $f^t : [t, 1] \rightarrow [0, 1]$  by  $f^t(x) := \varphi^{-1}(\varphi(t) - \varphi(x))$  implying that

$$\text{graph}(f^t) = \{(x, f^t(x)) : x \in [t, 1]\} = \{(x, y) \in [0, 1]^2 : C(x, y) = t\} = L_t$$

holds for every  $t \in (0, 1]$ .

For every generator  $\varphi : [0, 1] \rightarrow [0, \infty]$  we will let  $D^+\varphi(x)$  ( $D^-\varphi(x)$ ) denote the right-hand (left-hand) derivative of  $\varphi$  at  $x \in (0, 1)$ . Convexity of  $\varphi$  implies that  $D^+\varphi(x) = D^-\varphi(x)$  holds for all but at most countably many  $x \in (0, 1)$ , i.e.  $\varphi$  is differentiable outside a countable subset of  $(0, 1)$ , and that  $D^+\varphi$  is non-decreasing and right-continuous (see [23, 32]). In the sequel we will let  $\text{Cont}(D^+\varphi) \subseteq (0, 1)$  denote the set of all continuity points of  $D^+\varphi$  in  $(0, 1)$  (by definition, 0 and 1 are not contained in  $\text{Cont}(D^+\varphi)$ ) and make use of the fact that  $[0, 1] \setminus \text{Cont}(D^+\varphi)$  is at most countably infinite and has Lebesgue measure 0. Setting  $D^+\varphi(0) = -\infty$  in case of strict  $\varphi$  as well as  $D^+\varphi(1) = 0$  (for strict and non-strict  $\varphi$ ) allows to view  $D^+\varphi$  as non-decreasing and right-continuous function on the full unit interval  $[0, 1]$ . Additionally (again see [23, 32]) we have  $D^-\varphi(x) = D^+\varphi(x-)$  for every  $x \in (0, 1)$ .

If  $\varphi$  is strict then according to [12]

$$(4.3) \quad K_C(x, [0, y]) = \begin{cases} \frac{D^+\varphi(x)}{D^+\varphi(C(x, y))} & \text{if } x \in (0, 1) \\ 1 & \text{if } x \in \{0, 1\} \end{cases}$$

is (a version of) the Markov kernel of  $C$ , if  $\varphi$  is non-strict, then

$$(4.4) \quad K_C(x, [0, y]) = \begin{cases} 0 & \text{if } x \in (0, 1), y < f^0(x) \\ \frac{D^+\varphi(x)}{D^+\varphi(C(x, y))} & \text{if } x \in (0, 1), y \geq f^0(x) \\ 1 & \text{if } x \in \{0, 1\} \end{cases}$$

is a (version of a) Markov kernel of  $C$ . Recall that for every Archimedean copula  $C$  with generator  $\varphi$  the Kendall distribution function is given by (see, e.g., [14])

$$(4.5) \quad F^{\text{Kendall}}(x) = x - \frac{\varphi(x)}{D^+\varphi(x)}.$$

We now prove in several steps that in  $\mathcal{C}_{ar}$  weak conditional convergence and pointwise convergence coincide. The following theorem serves as starting point for this result. Up to a slight modification this intermediate theorem was already established by Charpentier and Segers in [4], however, the slight modification will turn out to be crucial in the sequel.

**Theorem 4.1.** *Let  $C, C_1, C_2, \dots$  be Archimedean copulas with generators  $\varphi, \varphi_1, \varphi_2, \dots$ , respectively. Then the following conditions are equivalent:*

- (a)  $\lim_{n \rightarrow \infty} C_n(x, y) = C(x, y)$  for all  $x, y \in [0, 1]$ ,
- (b)  $\lim_{n \rightarrow \infty} F_n^{\text{Kendall}}(x) = F^{\text{Kendall}}(x)$  for all  $x \in \text{Cont}(D^+\varphi)$ ,
- (c)  $\lim_{n \rightarrow \infty} \varphi_n(x) = \varphi(x)$  for all  $x \in (0, 1]$ ,
- (d)  $\lim_{n \rightarrow \infty} D^+\varphi_n(x) = D^+\varphi(x)$  for all  $x \in \text{Cont}(D^+\varphi)$ .



*Proof.* The equivalence of (a) and (b) can be found in [4, Proposition 2] and the equivalence of (a) and (c) is contained in [24, Theorem 8.14] where the authors prove equivalence of pointwise convergence of the sequence of multiplicative generators and the pointwise convergence of the induced continuous Archimedean  $t$ -norms which readily translates to the copula setting. The fact that (c) implies (d) is a direct consequence of the convexity of the generators, see [33].

Finally, suppose that (d) holds and consider  $x \in \text{Cont}(D^+\varphi)$ . For every  $\varepsilon > 0$  we can find an index  $n_0 \in \mathbb{N}$  such that for all  $n \geq n_0$  we have  $D^+\varphi(x) - \varepsilon < D^+\varphi_n(x)$ . Using monotonicity of  $D^+\varphi_n$  we get

$$D^+\varphi(x) - \varepsilon \leq D^+\varphi_n(x) \leq D^+\varphi_n(t) \leq 0$$

for  $n \geq n_0$  and every  $t \in [x, 1]$ . Having this, Dominated convergence yields

$$-\varphi(y) = \int_{[y,1]} D^+\varphi(t) dt = \int_{[y,1]} \lim_{n \rightarrow \infty} D^+\varphi_n(t) dt = \lim_{n \rightarrow \infty} \int_{[y,1]} D^+\varphi_n(t) dt = - \lim_{n \rightarrow \infty} \varphi_n(y)$$

for every  $y \in [x, 1]$ . Considering that  $\text{Cont}(D^+\varphi)$  is dense in  $(0, 1]$  condition (c) now follows immediately.  $\square$

The afore-mentioned modification of the result in [4] is that it may happen that a sequence of Archimedean copulas  $(C_n)_{n \in \mathbb{N}}$  converges to an Archimedean copula  $C$  although the corresponding generators  $(\varphi_n)_{n \in \mathbb{N}}$  do not converge to  $\varphi$  in the point 0.

We now state the main result of this section saying that pointwise convergence and weak conditional convergence coincide in  $\mathcal{C}_{ar}$ :

**Theorem 4.2.** *Let  $C, C_1, C_2, \dots$  be Archimedean copulas with generators  $\varphi, \varphi_1, \varphi_2, \dots$ , respectively. Then the following assertions are equivalent:*

- (a)  $\lim_{n \rightarrow \infty} C_n(x, y) = C(x, y)$  for all  $x, y \in [0, 1]$ ,
- (b)  $\lim_{n \rightarrow \infty} F_n^{\text{Kendall}}(x) = F^{\text{Kendall}}(x)$  for all  $x \in \text{Cont}(D^+\varphi)$ ,
- (c)  $\lim_{n \rightarrow \infty} \varphi_n(x) = \varphi(x)$  for all  $x \in (0, 1]$ ,
- (d)  $\lim_{n \rightarrow \infty} D^+\varphi_n(x) = D^+\varphi(x)$  for all  $x \in \text{Cont}(D^+\varphi)$ ,
- (e)  $\lim_{n \rightarrow \infty} D_1(C_n, C) = 0$ ,
- (f)  $C_n \xrightarrow{wcc} C$  for  $n \rightarrow \infty$ .

**Remark 4.3.** Since Archimedean copulas are symmetric Theorem 4.2 implies that within the class of all Archimedean copulas weak conditional convergence and  $\partial$ -convergence are equivalent (compare with Remark 3.5).

We are now going to prove Theorem 4.2 by showing that each of the conditions (a) to (d) from Theorem 4.1 implies weak conditional convergence. Doing so we will work with level curves and distinguish two cases concerning the 0-level curve  $f^0$  of the limit copula. For the first case, we present two different proofs since they use different ideas - the first one builds upon convexity of the generators and direct consequences to the sequence of derivatives, the second one uses some additional information about the behaviour of the corresponding sequence of level curves as described in the next lemma:

**Lemma 4.4.** *Let  $C, C_1, C_2, \dots$  be non-strict Archimedean copulas with generators  $\varphi, \varphi_1, \varphi_2, \dots$ , respectively. If  $(\varphi_n)_{n \in \mathbb{N}}$  converges pointwise to  $\varphi$  on  $(0, 1]$  then the following two assertions hold:*

- (i)  $\liminf_{n \rightarrow \infty} \varphi_n(0) \geq \varphi(0)$ ,
- (ii)  $f^t(x) \geq \limsup_{n \rightarrow \infty} f_n^t(x)$  for every  $t \in [0, 1)$  and every  $x \in [t, 1]$ .

*Proof.* To prove assertion (i) we proceed as follows: Considering that for every generator  $\psi$  and every  $z \in (0, 1)$  by convexity we have  $\psi(0) \geq \psi(z) - z \cdot D^+\psi(z)$  and applying Theorem 4.1 to the case  $z \in \text{Cont}(D^+\varphi)$  yields

$$\liminf_{n \rightarrow \infty} \varphi_n(0) \geq \liminf_{n \rightarrow \infty} (\varphi_n(z) - z \cdot D^+\varphi_n(z)) = \varphi(z) - z \cdot D^+\varphi(z) > \varphi(z).$$

Since for every  $\varepsilon > 0$  we can choose  $z \in \text{Cont}(D^+\varphi)$  in such a way that  $\varphi(z) > \varphi(0) - \varepsilon$  holds assertion (i) now follows. To prove assertion (ii) fix  $t \in [0, 1)$  and consider  $y > f^t(x)$ . In this case we have  $\varphi(x) + \varphi(y) < \varphi(t)$  implying that there exists an index  $n_0 \in \mathbb{N}$  such that  $\varphi_n(x) + \varphi_n(y) < \varphi_n(t)$ , hence  $y > f_n^t(x)$ , holds for every  $n \geq n_0$ . It follows that  $y \geq \limsup_{n \rightarrow \infty} f_n^t(x)$  from which the assertion follows immediately since  $y > f^t(x)$  was arbitrary.  $\square$

We now use the previous result to show level curve convergence:

**Lemma 4.5.** *Let  $C, C_1, C_2, \dots$  be Archimedean copulas with generators  $\varphi, \varphi_1, \varphi_2, \dots$  converging pointwise on  $(0, 1]$ . Then for every  $t > 0$  the  $t$ -level curves converge pointwise, i.e.,*

$$(4.6) \quad \lim_{n \rightarrow \infty} f_n^t(x) = f^t(x)$$

*holds for all  $x \in [t, 1]$ . If, in addition,  $\lim_{n \rightarrow \infty} \varphi_n(0) = \varphi(0)$  holds then eq. (4.6) is also true for  $t = 0$  and  $x \in [0, 1]$ .*

*Proof.* Suppose that  $t > 0$ . As a by-product of [24, Theorem 8.14] we obtain uniform convergence of the sequence  $(\varphi_n^-)_{n \in \mathbb{N}}$  to  $\varphi^-$  on each compact interval of the form  $[0, s]$  with  $s \in [0, \infty)$ . Fix  $x \in [t, 1]$ , set  $s := 2 \sup_{n \in \mathbb{N}} \varphi_n(t) < \infty$  and define  $z_n := \varphi_n(t) - \varphi_n(x)$ . Then  $(z_n)_{n \in \mathbb{N}}$  converges to  $z := \varphi(t) - \varphi(x)$  for  $n \rightarrow \infty$  and, using the afore-mentioned uniform convergence of  $(\varphi_n^-)_{n \in \mathbb{N}}$  on  $[0, s]$ , the equality  $\lim_{n \rightarrow \infty} f_n^t(x) = f^t(x)$  follows.

Notice that the second assertion is trivial for strict  $\varphi$ , so it remains to prove the assertion for  $\varphi(0) < \infty$ . Fix  $x \in (0, 1)$ . If  $f^0(x) = 0$  then the result follows directly from Lemma 4.4 part (ii). Suppose therefore that  $f^0(x) > 0$ . Then for every  $y \in (0, f^0(x))$  we have  $\varphi(x) + \varphi(y) > \varphi(0)$  and we can find an index  $n_0 \in \mathbb{N}$  such that  $\varphi_n(x) + \varphi_n(y) > \varphi_n(0)$ , hence  $y < f_n^0(x)$ , holds for every  $n \geq n_0$ . As direct consequence we get  $f^0(x) \leq \liminf_{n \rightarrow \infty} f_n^0(x)$ , which in combination with Lemma 4.4 assertion (ii) yields

$$f^0(x) \leq \liminf_{n \rightarrow \infty} f_n^0(x) \leq \limsup_{n \rightarrow \infty} f_n^0(x) \leq f^0(x).$$

This completes the proof.  $\square$

Recall that for univariate distribution functions  $F, F_1, F_2, \dots$  weak convergence of  $(F_n)_{n \in \mathbb{N}}$  to  $F$  is equivalent to pointwise convergence on a dense subset (see, e.g., [2]). In the following two lemmata we prove convergence on a dense set above and below the zero level curve  $f^0$  of the limit copula  $C$ . Notice that the first lemma is sufficient within the family of strict Archimedean copulas since in this case  $f^0(x) = 0$  for every  $x \in (0, 1]$ .

**Lemma 4.6.** *Let  $C, C_1, C_2, \dots$  be Archimedean copulas with generators  $\varphi, \varphi_1, \varphi_2, \dots$  and assume that one of the conditions of Theorem 4.1 holds. Then there exists a set  $\Lambda \in \mathcal{B}([0, 1])$  fulfilling  $\lambda(\Lambda) = 1$  such that for every  $x \in \Lambda$  we have that*

$$(4.7) \quad \lim_{n \rightarrow \infty} K_{C_n}(x, [0, y]) = K_C(x, [0, y])$$

holds for every  $y \in U_x \subseteq [f^0(x), 1]$ , where  $U_x$  is dense in  $[f^0(x), 1]$ .

*Proof (1).* Setting  $\Lambda := \text{Cont}(D^+\varphi)$  we obviously have  $\lambda(\Lambda) = 1$ . We are going to prove the even stronger property that for every  $x \in \Lambda$  the identity

$$(4.8) \quad \lim_{n \rightarrow \infty} |K_{C_n}(x, [0, y]) - K_C(x, [0, y])| = \lim_{n \rightarrow \infty} \left| \frac{D^+\varphi_n(x)}{D^+\varphi_n(C_n(x, y))} - \frac{D^+\varphi(x)}{D^+\varphi(C(x, y))} \right| = 0$$

holds for  $\lambda$ -almost all  $y \in [f^0(x), 1]$ . First of all notice that the set

$$U_x := \{y \in [f^0(x), 1] : C(x, y) \in \text{Cont}(D^+\varphi)\}$$

is of full measure in  $[f^0(x), 1]$ . In fact, for  $x \in (0, 1)$  the function  $h_x : [f^0(x), 1] \rightarrow [0, x]$ , defined by  $h_x(y) := C(x, y)$  is an increasing homeomorphism (see [30]) and therefore the set  $h_x^{-1}(\text{Cont}(D^+\varphi)^c)$  is at most countably infinite. Convexity of the generators implies that the sequence  $(D^+\varphi_n)_{n \in \mathbb{N}}$  of derivatives converges continuously to  $D^+\varphi$  on  $\text{Cont}(D^+\varphi)$  (see [33]), hence we obtain

$$\lim_{n \rightarrow \infty} D^+\varphi_n(C_n(x, y)) = D^+\varphi(C(x, y))$$

from which the desired property follows.  $\square$

*Proof (2).* First of all notice that for  $t, x \in \text{Cont}(D^+\varphi)$  we have

$$(4.9) \quad K_{C_n}(x, [0, f_n^t(x)]) = \frac{D^+\varphi_n(x)}{D^+\varphi_n(C(x, f_n^t(x)))} = \frac{D^+\varphi_n(x)}{D^+\varphi_n(t)}$$

and the right-hand side converges to  $K_C(x, [0, f^t(x)])$  by Theorem 4.1. Exploiting this fact we consider  $x \in \text{Cont}(D^+\varphi)$  and proceed as follows: Fix  $\varepsilon > 0$  again and let  $y \in [f^0(x), 1]$  denote a continuity point of the map  $v \mapsto K_C(x, [0, v])$ . Furthermore choose  $t, s \in \text{Cont}(D^+\varphi)$  with  $t < s$  in such a way that  $y \in (f^t(x), f^s(x))$  and  $K_C(x, [f^t(x), f^s(x)]) < \varepsilon$  holds. According to Lemma 4.5 there exists some index  $n_0 \in \mathbb{N}$  such that for all  $n \geq n_0$  we have  $y < f_n^s(x)$ , which using eq. (4.6) implies

$$K_{C_n}(x, [0, y]) \leq K_{C_n}(x, [0, f_n^s(x)]) \xrightarrow{n \rightarrow \infty} K_C(x, [0, f^s(x)]) \leq K_C(x, [0, y]) + \varepsilon.$$

As a direct consequence we get

$$\limsup_{n \rightarrow \infty} K_{C_n}(x, [0, y]) \leq K_C(x, [0, y]) + \varepsilon.$$

Replacing  $s$  by  $t$  and proceeding analogously yields

$$\liminf_{n \rightarrow \infty} K_{C_n}(x, [0, y]) \geq K_C(x, [0, y]) - \varepsilon,$$

which completes the proof.  $\square$

As second step we consider the case  $0 < y < f^0(x)$  (implying that  $C$  is non-strict). The following simple lemma will be crucial for the proof of Lemma 4.8:

**Lemma 4.7.** *Suppose that  $(X, d_1)$  is a compact metric space and that  $(Y, d_2)$  is another (not necessarily compact) metric space. Furthermore let  $f : X \rightarrow Y$  be an arbitrary function and  $(x_n)_{n \in \mathbb{N}}$  a sequence in  $(X, d_1)$ . If there exists some  $y \in Y$  such that for every convergent subsequence  $(x_{n_j})_{j \in \mathbb{N}}$  we have  $\lim_{j \rightarrow \infty} d_2(f(x_{n_j}), y) = 0$  then  $\lim_{n \rightarrow \infty} d_2(f(x_n), y) = 0$  follows.*

*Proof.* Suppose that the assumptions of the lemma are fulfilled but the sequence  $(y_n)_{n \in \mathbb{N}}$  with  $y_n := f(x_n)$  does not converge to  $y$  for  $n \rightarrow \infty$ . In this case there exists some  $\varepsilon > 0$  and a subsequence  $(y_{n_j})_{j \in \mathbb{N}}$  with  $d_2(y_{n_j}, y) \geq \varepsilon$  for every  $j \in \mathbb{N}$ . Compactness of  $(X, d_1)$  implies the existence of a subsequence  $(x_{n_{j_l}})_{l \in \mathbb{N}}$  of  $(x_{n_j})_{j \in \mathbb{N}}$ , and, by assumption, this sequence fulfills

$$\lim_{l \rightarrow \infty} d_2(y_{n_{j_l}}, y) = \lim_{l \rightarrow \infty} d_2(f(x_{n_{j_l}}), y) = 0,$$

a contradiction.  $\square$

To simplify notation we say that  $\lim_{n \rightarrow \infty} \varphi_n(0) = \infty$  if, and only if for every  $N \in \mathbb{N}$  there exists some index  $n_0 \in \mathbb{N}$  such that for all  $n \geq n_0$  we have  $\varphi_n(0) > N$ . Lemma 4.4 part (i) together with Lemma 4.7 allow us to distinguish the following three types of convergent subsequences  $(\varphi_{n_j})_{j \in \mathbb{N}}$  of  $(\varphi_n)_{n \in \mathbb{N}}$ : (a)  $\lim_{j \rightarrow \infty} \varphi_{n_j}(0) = \varphi(0)$ , (b)  $\lim_{j \rightarrow \infty} \varphi_{n_j}(0) = \infty$  or (c)  $\lim_{j \rightarrow \infty} \varphi_{n_j}(0) = \alpha \in (\varphi(0), \infty)$ .

**Lemma 4.8.** *Let  $C, C_1, C_2, \dots$  be Archimedean copulas with generators  $\varphi, \varphi_1, \varphi_2, \dots$  and assume that one of the conditions of Theorem 4.1 holds. Then there exists a set  $\Lambda \in \mathcal{B}([0, 1])$  fulfilling  $\lambda(\Lambda) = 1$  such that for every  $x \in \Lambda$  we have that*

$$(4.10) \quad \lim_{n \rightarrow \infty} K_{C_n}(x, [0, y]) = 0 = K_C(x, [0, y])$$

*holds for every  $y < f^0(x)$ .*

*Proof.* As in the previous proof we set  $\Lambda = \text{Cont}(D^+\varphi)$ . Fix  $x \in \text{Cont}(D^+\varphi)$  and  $y \in (0, f^0(x))$  and distinguish the following two different situations:

(a) Suppose that  $(\varphi_{n_j})_{j \in \mathbb{N}}$  is a subsequence of  $(\varphi_n)_{n \in \mathbb{N}}$  fulfilling  $\lim_{j \rightarrow \infty} \varphi_{n_j}(0) = \varphi(0)$ . According to Lemma 4.5 we have  $\lim_{j \rightarrow \infty} f_{n_j}^0(x) = f^0(x)$ , so there exists an index  $j_0 \in \mathbb{N}$  such that  $y < f_{n_j}^0(x)$ , hence  $K_{C_{n_j}}(x, [0, y]) = 0 = K_C(x, [0, y])$ , holds for every  $j \geq j_0$ , from which  $\lim_{j \rightarrow \infty} K_{C_{n_j}}(x, [0, y]) = K_C(x, [0, y])$  follows immediately.

(b) & (c) Suppose that  $(\varphi_{n_j})_{j \in \mathbb{N}}$  is a subsequence of  $(\varphi_n)_{n \in \mathbb{N}}$  fulfilling  $\lim_{j \rightarrow \infty} \varphi_{n_j}(0) = \infty$  or  $\lim_{j \rightarrow \infty} \varphi_{n_j}(0) =: \alpha \in (\varphi(0), \infty)$ . Choose  $\varepsilon > 0$  in such a way that  $y \leq f^0(x) - \varepsilon$  holds and define the set  $M_\varepsilon$  (see Figure 2) by

$$M_\varepsilon = \{(a, b) \in [0, 1]^2 : b \leq f^0(a) - \varepsilon\}.$$

Then  $\mu_C(M_\varepsilon) = 0$  and  $M_\varepsilon$  is a  $\mu_C$ -continuity set, so applying Portmanteau's theorem (see [2]) yields

$$(4.11) \quad \lim_{j \rightarrow \infty} \mu_{C_{n_j}}(M_\varepsilon) = 0.$$

Assume now that there exists some  $\delta > 0$  such that  $K_{C_{n_j}}(x, [0, y]) \geq \delta > 0$  would hold for infinitely many  $j \in \mathbb{N}$  and denote the corresponding subsequence by  $(C_{n_{j_l}})_{l \in \mathbb{N}}$ . It follows from eq. (4.4) that  $y \geq f_{n_{j_l}}^0(x)$  holds for every  $l \in \mathbb{N}$ . Set  $t_{n_{j_l}} := C_{n_{j_l}}(x, y)$  for every  $l \in \mathbb{N}$  and let  $l^*$  denote the smallest index fulfilling that  $t_{n_{j_l}} < x$  holds for all  $l \geq l^*$ . Using the fact that for every Archimedean copula  $A$  with generator  $\psi$  and for every  $t \in [0, 1)$  the mapping

$$x \mapsto K_A(x, [0, f^t(x)]) = \frac{D^+\psi(x)}{D^+\psi(t)}$$

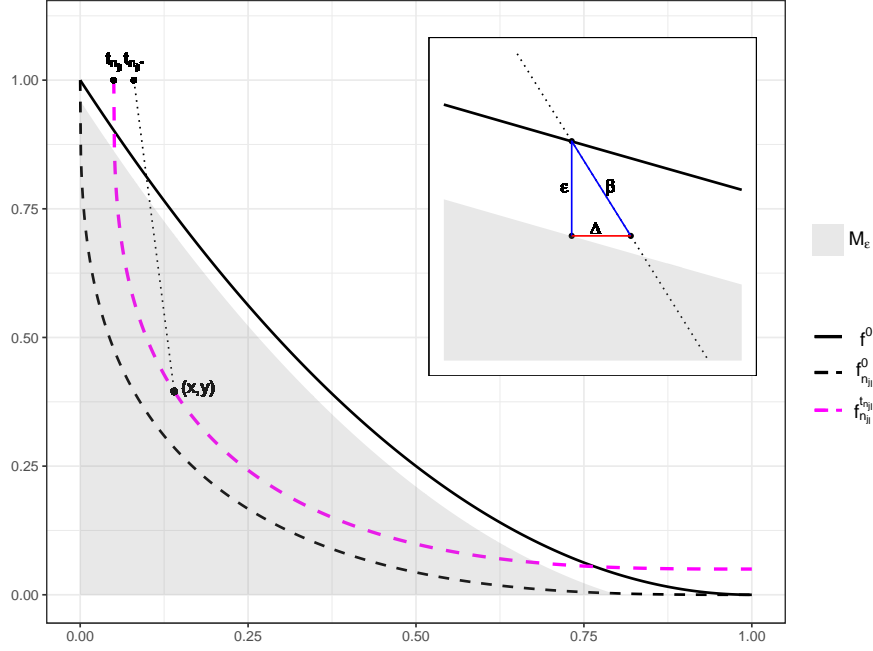


FIGURE 2. The  $\mu_C$ -null set  $M_\varepsilon$ , the level curves  $f_{n_{j_l}}^{t_{n_{j_l}}}$ ,  $f_{n_{j_l}}^0$ ,  $f^0$  considered in the proof of Lemma 4.8 and a zoomed-in illustration of  $\Delta$ , where  $\beta = \frac{1-y}{x-t_{n_{j_l}*}}$ .

is decreasing in  $x$  it follows that for every  $l \in \mathbb{N}$  we have

$$K_{C_{n_{j_l}}}(u, [0, f_{n_{j_l}}^{t_{n_{j_l}}}(u)]) \geq K_{C_{n_{j_l}}}(x, [0, \underbrace{f_{n_{j_l}}^{t_{n_{j_l}}}(x)}_{=y}]) \geq \delta > 0$$

for every  $u \in (0, x]$ . The proof idea now is to use this monotonicity in combination with convexity of the level curves (see [30]) to construct a contradiction to eq. (4.11): In fact, convexity implies that the graph of each  $f_{n_{j_l}}^{t_{n_{j_l}}}$  restricted to  $[t_{n_{j_l}*}, x]$  lies below the straight line connecting the points  $(x, y)$  and  $(t_{n_{j_l}*}, 1)$  (again see Figure 2). Hence, defining  $\Delta > 0$  by

$$(4.12) \quad \Delta = \frac{\varepsilon}{\frac{1-y}{x-t_{n_{j_l}*}}} = \varepsilon \frac{x-t_{n_{j_l}*}}{1-y}$$

it follows that for every  $u \in [x-\Delta, x]$  we have  $(u, f_{n_{j_l}}^{t_{n_{j_l}}}(u)) \in M_\varepsilon$ . As direct consequence it follows that

$$\mu_{C_{n_{j_l}}}(M_\varepsilon) \geq \int_{[x-\Delta, x]} K_{C_{n_{j_l}}}(u, [0, f_{n_{j_l}}^{t_{n_{j_l}}}(u)]) \, d\lambda(u) \geq \delta \cdot \Delta > 0$$

holds for every  $l > l^*$  which contradicts eq. (4.11). Altogether in case (b) & (c) we have also shown now that  $\lim_{j \rightarrow \infty} K_{C_{n_j}}(x, [0, y]) = 0 = K_C(x, [0, y])$  holds.

Taking (a) and (b) & (c) together we have proved that for each convergent subsequence  $(\varphi_{n_j}(0))_{j \in \mathbb{N}}$  of  $(\varphi_n(0))_{n \in \mathbb{N}}$  we have

$$\lim_{j \rightarrow \infty} K_{C_{n_j}}(x, [0, y]) = 0.$$

The result now follows from Lemma 4.7.  $\square$

Considering that weak conditional convergence of the copulas implies convergence in  $D_1$  the proof of this section's main result Theorem 4.2 is complete. In Section 6 we will use the following interesting consequence:

**Corollary 4.9.** *Let  $C, C_1, C_2, \dots$  be Archimedean copulas with generators  $\varphi, \varphi_1, \varphi_2, \dots$ , respectively and suppose that  $(\varphi_n)_{n \in \mathbb{N}}$  converges to  $\varphi$  on  $(0, 1]$ . Then the following identities holds:*

$$(4.13) \quad \lim_{n \rightarrow \infty} \zeta_1(C_n) = \zeta_1(C), \quad \lim_{n \rightarrow \infty} r(C_n) = r(C)$$

*In other words: Within  $\mathcal{C}_{ar}$  both  $\zeta_1$  and  $r$  are continuous w.r.t. pointwise convergence of the generators on  $(0, 1]$ .*

We conclude this section by recalling the fact that the class of Archimedean copulas is not closed w.r.t. uniform convergence, i.e., the limit of a sequence of Archimedean copulas may fail to be Archimedean, see [4] (however, we necessarily have associativity). As easily verified, the same is true if we consider weak conditional convergence or convergence w.r.t.  $D_1$ .

## 5. EXTREME VALUE COPULAS

We are now going to prove a result similar to Theorem 4.2 for bivariate Extreme Value copulas. Remember that  $C \in \mathcal{C}$  is called bivariate *Extreme Value copula* if one of the following three equivalent conditions is fulfilled (see [8, 10, 18, 31]):

(a) There is a copula  $B \in \mathcal{C}$  such that for all  $x, y \in [0, 1]$  we have

$$(5.1) \quad C(x, y) = \lim_{n \rightarrow \infty} B^n(x^{\frac{1}{n}}, y^{\frac{1}{n}}).$$

(b)  $C(x, y) = C^n(x^{\frac{1}{n}}, y^{\frac{1}{n}})$  holds for all  $n \in \mathbb{N}$  and all  $x, y \in [0, 1]$ .

(c) There exists a convex map  $A : [0, 1] \rightarrow [0, 1]$  satisfying  $A(0) = A(1) = 1$  and  $\max(1 - x, x) \leq A(x) \leq 1$  for all  $x \in [0, 1]$  such that for all  $x, y \in (0, 1)$  the copula  $C$  can be expressed in terms of  $A$  as

$$(5.2) \quad C(x, y) = C_A(x, y) := (xy)^{A\left(\frac{\ln(x)}{\ln(xy)}\right)}.$$

In the following we will let  $\mathcal{C}_{ev}$  denote the class of all bivariate Extreme Value copulas,  $\mathcal{A}$  the family of all *Pickands dependence functions*, i.e., the family of all functions  $A$  fulfilling assertion (c). Using either max-stability or Arzela-Ascoli theorem [34] it is straightforward to verify that  $\mathcal{C}_{ev}$  is a compact subset of  $(\mathcal{C}, d_\infty)$ . Furthermore, letting  $\|\cdot\|_\infty$  denote the uniform norm on  $\mathcal{A}$ , obviously the mapping  $\Phi : (\mathcal{A}, \|\cdot\|_\infty) \rightarrow (\mathcal{C}_{ev}, d_\infty)$ , defined by  $\Phi(A) = C_A$ , is continuous and it is straightforward to verify that a sequence of Extreme Value copulas  $(C_{A_n})_{n \in \mathbb{N}}$  converges pointwise (hence uniformly) to an Extreme Value copula  $C_A$  if, and only if  $(A_n)_{n \in \mathbb{N}}$  converges uniformly to  $A$ .

Following [38] we will let  $D^+A$  denote the right-hand derivative of the Pickands dependence function  $A$  on  $[0, 1]$  and  $D^-A$  the left-hand derivative on  $(0, 1]$ . Furthermore, convexity implies that  $D^-A(x) = D^+A(x)$  holds for all but at most countably infinitely many  $x \in (0, 1)$ . In the sequel we will let  $\text{Cont}(D^+A)$  denote the set of all continuity points of  $D^+A$  in  $(0, 1)$ . Setting  $D^+A(1) := D^-A(1)$  we can view  $D^+A$  as a function on the whole unit interval that attains values in  $[-1, 1]$ . Furthermore  $D^+A : [0, 1] \rightarrow [-1, 1]$  is a non-decreasing, right-continuous function and it is straightforward to verify that  $\mathcal{A}$  can be identified with  $\mathcal{D}_A$ , defined by

$$\mathcal{D}_A = \left\{ f : [0, 1] \rightarrow [-1, 1] : f \text{ non-decreasing, right-continuous, } \int_{[0,1]} f \, d\lambda = 0 \right\},$$

in the sense that for every  $A \in \mathcal{A}$  we have  $D^+A \in \mathcal{D}_\mathcal{A}$  and, given  $f \in \mathcal{D}_\mathcal{A}$  setting  $A(x) := 1 + \int_{[0,x]} f d\lambda$  yields  $A \in \mathcal{A}$  as well as  $D^+A = f$  on  $[0, 1)$  (see [38]). For more information on Pickands dependence functions and the approach via right-hand derivatives we refer to [3, 16].

Returning to weak conditional convergence first notice that according to [38]

(5.3)

$$K_C(x, [0, y]) = \begin{cases} 1 & \text{if } x \in \{0, 1\} \\ C(x, y) [D^+A(\frac{\log(x)}{\log(xy)}) \frac{\log(y)}{x \log(xy)} + \frac{1}{x} A(\frac{\log(x)}{\log(xy)})] & \text{if } x, y \in (0, 1) \\ y & \text{if } x \in (0, 1), y \in \{0, 1\} \end{cases}$$

is a Markov kernel of the Extreme Value copula  $C$  with Pickands dependence function  $A$ .

We now state the main result of this section saying that in  $\mathcal{C}_{ev}$  pointwise convergence and weak conditional convergence are equivalent:

**Theorem 5.1.** *Let  $C, C_1, C_2, \dots$  be Extreme Value copulas with Pickands dependence functions  $A, A_1, A_2, \dots$ , respectively. Then the following assertions are equivalent:*

- (a)  $\lim_{n \rightarrow \infty} C_n(x, y) = C(x, y)$  for all  $x, y \in [0, 1]$ ,
- (b)  $\lim_{n \rightarrow \infty} A_n(x) = A(x)$  for all  $x \in [0, 1]$ ,
- (c)  $\lim_{n \rightarrow \infty} D^+A_n(x) = D^+A(x)$  for all  $x \in \text{Cont}(D^+A)$ ,
- (d)  $\lim_{n \rightarrow \infty} D_1(C_n, C) = 0$ ,
- (e)  $C_n \xrightarrow{wcc} C$  for  $n \rightarrow \infty$ .

**Remark 5.2.** Since for every Extreme Value copula  $C_A$  its transpose  $C_{A^t} := (C_A)^t$  is an Extreme Value copula with Pickands dependence function  $A^t$  given by  $A^t(x) = A(1 - x)$  for every  $x \in [0, 1]$  it follows that in the class of bivariate Extreme Value copulas the properties  $C_n \xrightarrow{wcc} C$  and  $C_n^t \xrightarrow{wcc} C^t$  are equivalent. As a consequence of Theorem 5.1 (and in contrast to Remark 3.5) weak conditional convergence and  $\partial$ -convergence are therefore equivalent in  $\mathcal{C}_{ev}$ .

Theorem 5.1 is a direct consequence of the following analogue of Lemma 4.6 and Lemma 4.8 (notice that the result implies weak conditional convergence for ANY choice of the Markov kernels):

**Lemma 5.3.** *Let  $C, C_1, C_2, \dots$  be bivariate Extreme Value copulas with Pickands dependence functions  $A, A_1, A_2, \dots$ , respectively. Suppose that  $(C_n)_{n \in \mathbb{N}}$  converges pointwise to  $C$  and choose the corresponding kernels according to (5.3). Then for every  $x \in (0, 1)$  there exists a set  $U_x \subset [0, 1]$  that is dense in  $[0, 1]$  such that*

$$(5.4) \quad \lim_{n \rightarrow \infty} K_{C_n}(x, [0, y]) = K_C(x, [0, y])$$

holds for every  $y \in U_x$ .

*Proof.* If  $(C_n)_{n \in \mathbb{N}}$  converges pointwise to  $C$  then, as mentioned before, it follows that  $\lim_{n \rightarrow \infty} \|A_n - A\|_\infty = 0$  holds. Thus (as in the case of Archimedean generators) convexity yields

$$\lim_{n \rightarrow \infty} D^+A_n(x) = D^+A(x)$$

for every  $x \in \text{Cont}(D^+A)$ . Defining  $h_x : (0, 1) \rightarrow (0, 1)$  for every  $x \in (0, 1)$  by

$$(5.5) \quad h_x(y) = \frac{\log(x)}{\log(xy)}$$

yields a strictly increasing homeomorphism of  $(0, 1)$ . Since  $(0, 1) \setminus \text{Cont}(D^+A)$  is at most countably infinite  $h_x^{-1}((0, 1) \setminus \text{Cont}(D^+A))$  is as well and

$$\lambda(h_x^{-1}(\text{Cont}(D^+A))) = 1$$

follows. Being a set of full measure  $U_x := h_x^{-1}(\text{Cont}(D^+A))$  is dense in  $[0, 1]$  and the result follows.  $\square$

Altogether we have proved Theorem 5.1 which, in turn, has the following corollary:

**Corollary 5.4.** *Let  $C, C_1, C_2, \dots$  be Extreme Value copulas with Pickands dependence functions  $A, A_1, A_2, \dots$ , respectively and suppose that  $(A_n)_{n \in \mathbb{N}}$  converges to  $A$  on  $[0, 1]$ . Then the following identities holds:*

$$(5.6) \quad \lim_{n \rightarrow \infty} \zeta_1(C_n) = \zeta_1(C), \quad \lim_{n \rightarrow \infty} r(C_n) = r(C)$$

*In other words: Within  $\mathcal{C}_{ev}$  both  $\zeta_1$  and  $r$  are continuous w.r.t. pointwise convergence of the Pickands dependence functions.*

We conclude this section with the following remark:

**Remark 5.5.** Suppose that  $H, H_1, H_2, \dots$  are the continuous bivariate distribution functions of the pairs  $(X, Y), (X_1, Y_1), (X_2, Y_2), \dots$ , with corresponding marginal distribution functions  $F_X, F_{X_1}, F_{X_2}, \dots$  and  $G_Y, G_{Y_1}, G_{Y_2}, \dots$  and corresponding copulas  $C, C_1, C_2, \dots$ . Defining  $H_n \xrightarrow{\text{wcc}} H$  analogously to Definition 3.1 (notice that in this case  $\lambda$  is replaced by  $\mathbb{P}^X$ ) it is straightforward to verify that  $H_n \xrightarrow{\text{wcc}} H$  implies  $C_n \xrightarrow{\text{wcc}} C$  but not necessarily vice versa. In the case that all  $C, C_1, C_2, \dots$  are Extreme Value copulas and the Pickands function  $A$  of the limit copula  $C$  is twice differentiable, however, the reverse implication also holds. In the class  $\mathcal{C}_{ar}$  the two concepts are equivalent too if, for instance, all generators are 3-monotone.

## 6. CONSEQUENCES FOR THE ESTIMATION OF ARCHIMEDEAN AND EXTREME VALUE COPULAS

**6.1. Extreme Value copulas.** Suppose that  $C_A$  is an Extreme Value copula with Pickands function  $A$  and suppose that  $(X_1, Y_1), (X_2, Y_2), \dots$  is a random sample from  $(X, Y) \sim C_A$ . Letting  $\hat{A}_n$  denote the CFG estimator according to [3, 15] (for an estimator in the multivariate setting see [19]) it can be shown that if  $A$  is twice continuously differentiable then the corresponding process  $\sqrt{n}(\hat{A}_n - A)$  (in the space of  $C([0, 1], \|\cdot\|_\infty)$  of all continuous functions on the unit interval) has a weak limit, and that, for suitable weight functions,  $(\hat{A}_n)_{n \in \mathbb{N}}$  is uniformly, strongly consistent (see [3, Proposition 4.1]). Although the estimator  $\hat{A}_n$  may fail to be convex in general, following an idea from [15] it can be used to construct a convex estimator  $\hat{A}_n^*$  given by

$$\hat{A}_n^* := \text{greatest convex minorant of } \max\{\min\{\hat{A}_n, 1\}, \text{id}, 1 - \text{id}\}$$

where  $\text{id}$  denotes the identity function on  $[0, 1]$ .  $\hat{A}_n^*$  is a Pickands dependence function (see [15, Section 3.3]) and the estimator  $\hat{A}_n^*$  is uniformly, strongly consistent (the latter follows from [28]). Hence Theorem 5.1 directly yields weak conditional convergence of the sequence of corresponding Extreme Value copulas  $(C_{\hat{A}_n^*})_{n \in \mathbb{N}}$  to  $C_A$ . Moreover, according to Corollary 5.4

$$\lim_{n \rightarrow \infty} |\zeta_1(C_{\hat{A}_n^*}) - \zeta_1(C_A)| = 0$$



holds and the same is true for the dependence measure  $r$  studied in [9], i.e., for estimating  $\zeta_1(C_A)$  it suffices to have a good estimator of the Pickands dependence function  $A$  (we refer to [21] for more details concerning the estimation of  $\zeta_1(X, Y)$ ).

**Example 6.1.** Consider the Galambos copula  $C_A$  with parameter  $\theta = 3$ , i.e., the Extreme Value copula whose Pickands dependence function is given by  $A(x) = 1 - (x^{-3} + (1 - x)^{-3})^{-1/3}$ . Figure 3 depicts a sample (and corresponding histograms) for the case  $n = 10000$ .

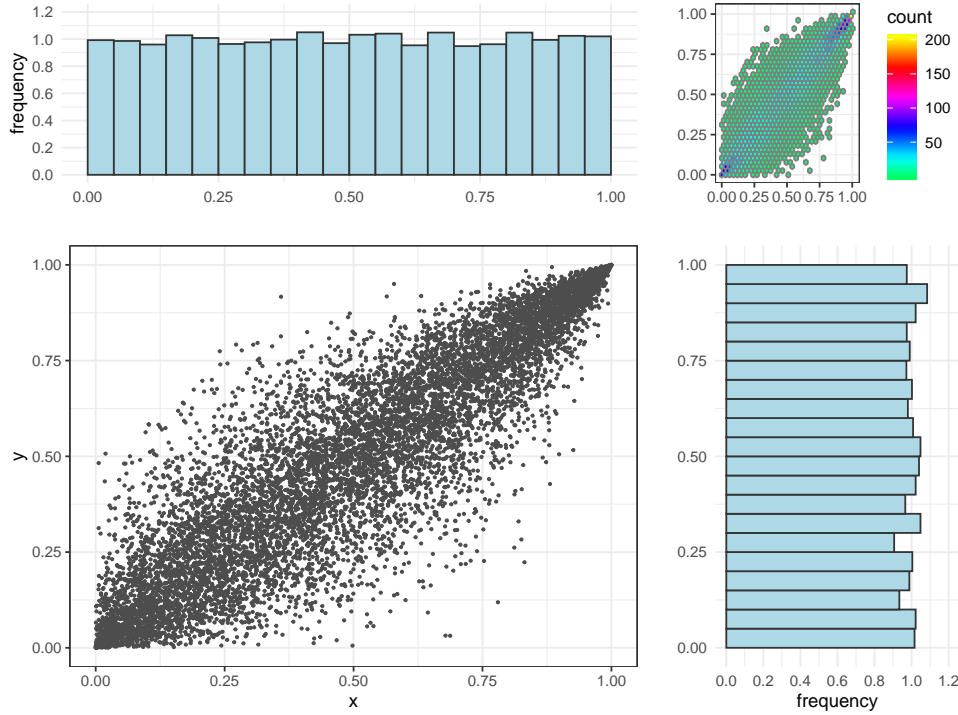


FIGURE 3. Sample of size  $n = 10000$  from the Galambos copula with parameter  $\theta = 3$  as considered in Example 6.1 (lower left panel); two-dimensional histogram (upper right panel) and marginal histograms (upper left and lower right panel).

Using the R-package ‘copula’ we calculate the estimator  $\hat{A}_n^*$  of  $A$  as described above. Figure 4 depicts the obtained generators together with the true Pickands function  $A$ . For the dependence measure  $\zeta_1$  (again using the R-package ‘qad’) we obtained the following values:  $\zeta_1(C_{\hat{A}_{500}^*}) = 0.7746$ ,  $\zeta_1(C_{\hat{A}_{10000}^*}) = 0.7594$ ,  $\zeta_1(C_A) = 0.7513$ .

We now focus on the estimation of  $\zeta_2 := r$  introduced in [9], denote the estimator of  $r$  developed by Chatterjee in [5] by  $\hat{r}_n$  (for an implementation see the R-package ‘XICOR’ [6]) and proceed with a small simulation study comparing the afore-mentioned plugin approach (using the Extreme Value information) with  $\hat{r}_n$  (not taking into account the Extreme Value information). In other words: Given a random sample  $(X_1, Y_1), (X_2, Y_2), \dots$  from  $(X, Y) \sim C_A \in \mathcal{C}_{ev}$  we calculate  $\hat{r}_n$  and  $r(C_{\hat{A}_n^*}) = 6 \cdot D_2^2(C_{\hat{A}_n^*}, \Pi)$  for different sample sizes a total of  $R = 5000$  times. Doing so we consider two cases of  $C_A$ : the Galambos copula with parameter  $\theta = 3$  and the Extreme Value copula with piecewise linear Pickands dependence function  $A$  given by

$$(6.1) \quad A(x) = \mathbf{1}_{[0, \frac{1}{4}]}(x) \cdot (1 - x) + \mathbf{1}_{(\frac{1}{4}, \frac{7}{10}]} \left( -\frac{1}{9}(x - 7) \right) + \mathbf{1}_{(\frac{7}{10}, 1]}(x) \cdot x.$$

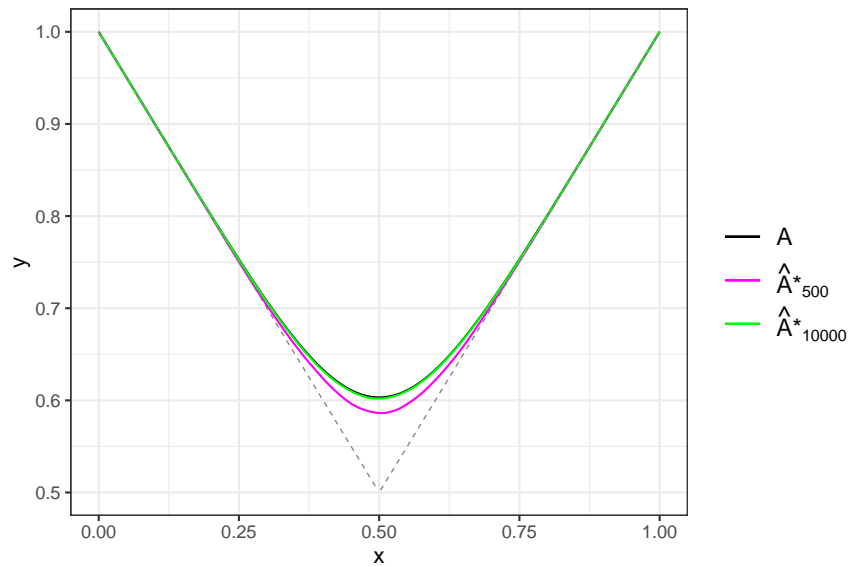


FIGURE 4. The Pickands dependence function  $A$  (black) as well as the estimate  $\hat{A}_n^*$  for  $n = 500$  and  $n = 10000$  as considered in Example 6.1.

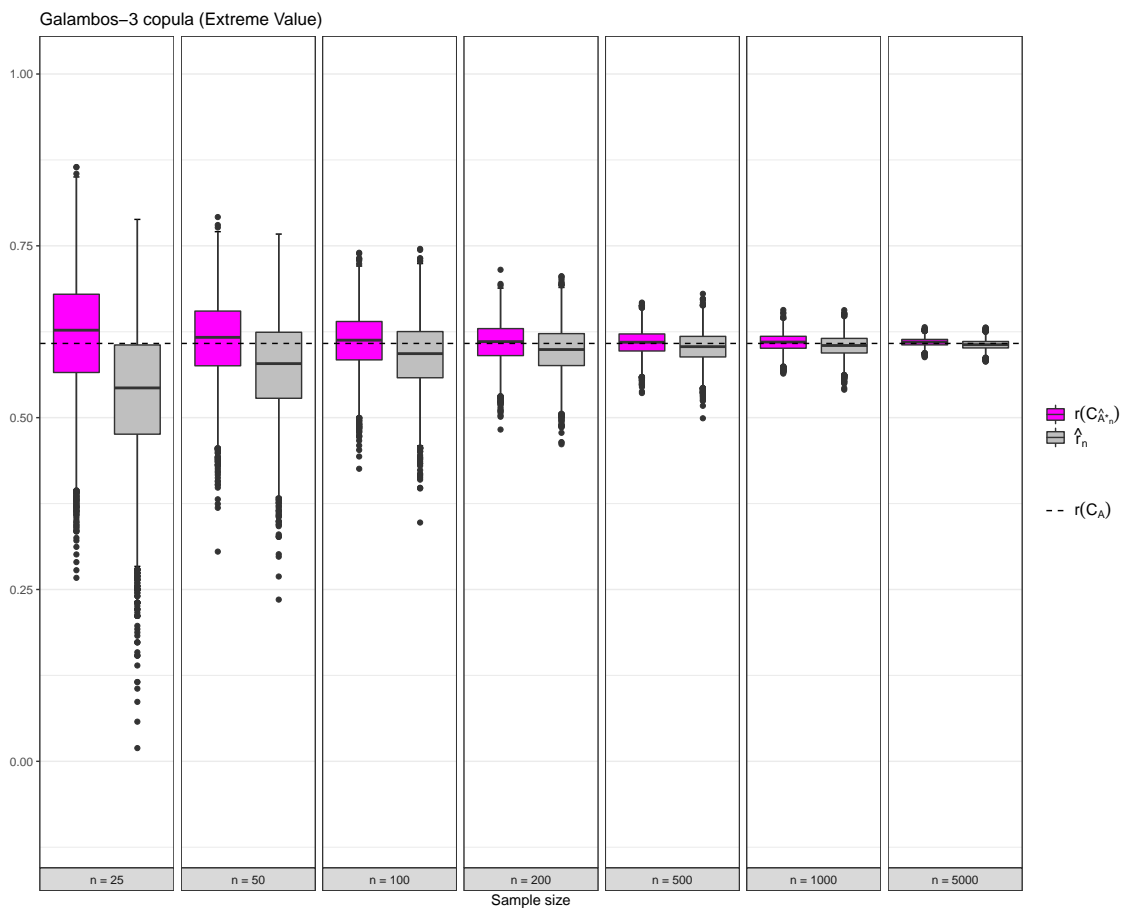


FIGURE 5. Boxplots of the obtained value of  $\hat{r}_n$  and  $r(C_{\hat{A}_n^*})$  based on samples from the Galambos copula with parameter  $\theta = 3$ .

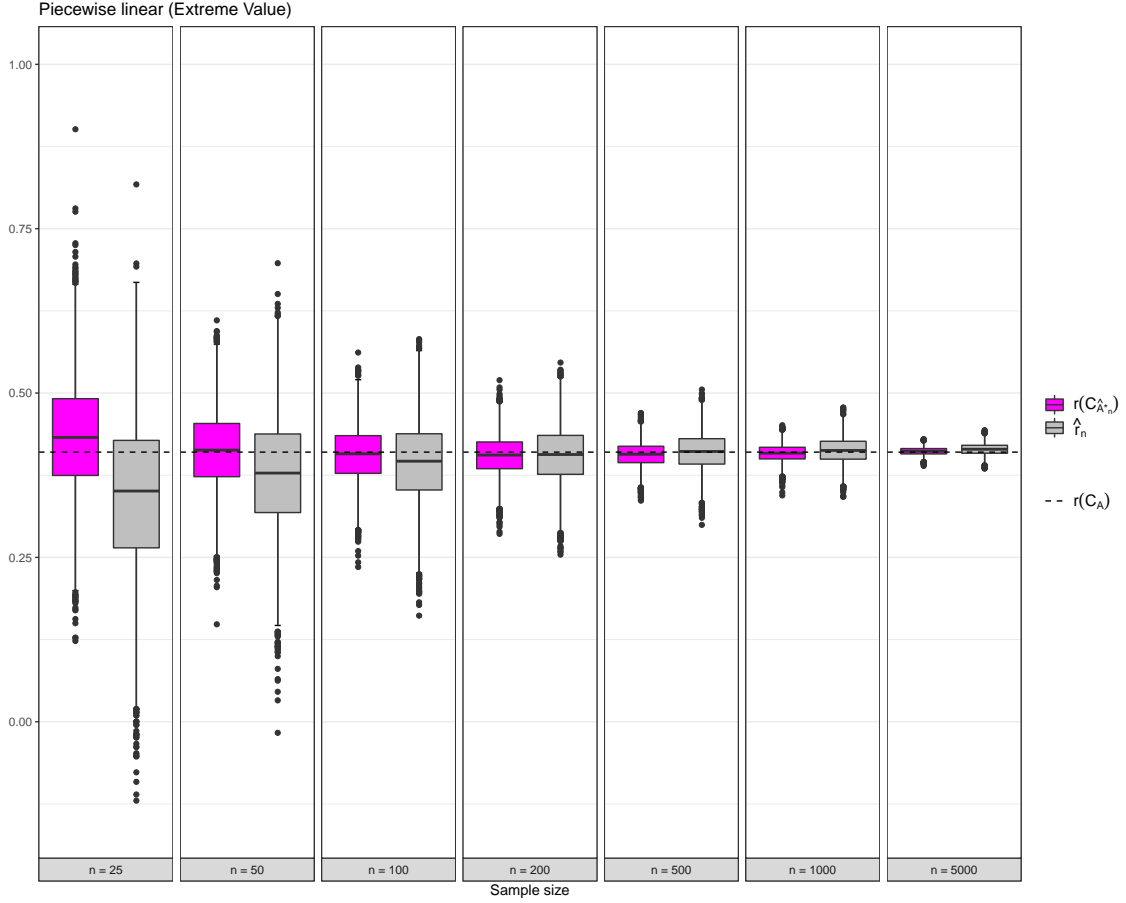


FIGURE 6. Boxplots of the obtained value of  $\hat{r}_n$  and  $r(C_{\hat{A}_n^*})$  based on samples from the copula with Pickands dependence function  $A$  according to eq. (6.1).

Not surprisingly, Figure 5 and Figure 6 show that for small to moderate sample sizes the plugin estimator  $r(C_{\hat{A}_n^*})$  (using the Extreme Value information) yields better results than  $\hat{r}_n$ , for large sample sizes both estimators perform similarly well.

**6.2. Archimedean copulas.** Turning to the Archimedean setting suppose now that  $C$  is an Archimedean copula with generator  $\varphi$  and let  $(X_1, Y_1), (X_2, Y_2), \dots$  be a random sample from  $(X, Y) \sim C$ . We will let  $\hat{F}_n^{\text{Kendall}}$  denote the estimator of the Kendall distribution function  $F^{\text{Kendall}}$  of  $C$  called  $K_{n,2}$  in [13, Lemma 1]. According to [13] (also see [1, 14]),  $\hat{F}_n^{\text{Kendall}}$  itself is a Kendall distribution function of an Archimedean copula and under mild regularity conditions the so-called empirical Kendall process  $\sqrt{n}(\hat{F}_n^{\text{Kendall}} - F^{\text{Kendall}})$  converges weakly to a centered Gaussian process. If  $\hat{F}_n^{\text{Kendall}}$  converges weakly to  $F^{\text{Kendall}}$  (to the best of authors' knowledge no sufficient conditions for this property to hold are known in the literature) then, according to Theorem 4.2, we automatically have weak conditional convergence of the sequence of corresponding Archimedean copulas  $(C_{\hat{\varphi}_n})_{n \in \mathbb{N}}$  to  $C$ , whereby  $\hat{\varphi}_n$  denotes the (normalized) generator obtained from  $\hat{F}_n^{\text{Kendall}}$ . Moreover, according to Corollary 4.9

$$\lim_{n \rightarrow \infty} |\zeta_1(C_{\hat{\varphi}_n}) - \zeta_1(C)| = 0$$

holds and the same is true for the dependence measure studied in [9], i.e., for estimating  $\zeta_1(C)$  it suffices to have a good estimator of the Kendall distribution function  $F^{\text{Kendall}}$ .

**Example 6.2.** We illustrate the afore-mentioned properties with simulations in R and consider the (normalized) generator  $\varphi(x) = (-\log(\frac{1}{2}))^{-3} \cdot (-\log(x))^3$  of the Gumbel copula  $C_\varphi$  with parameter  $\theta = 3$ . Figure 7 depicts a sample of size  $n = 10000$  from this copula as well as a two-dimensional and the corresponding marginal histograms. For both samples

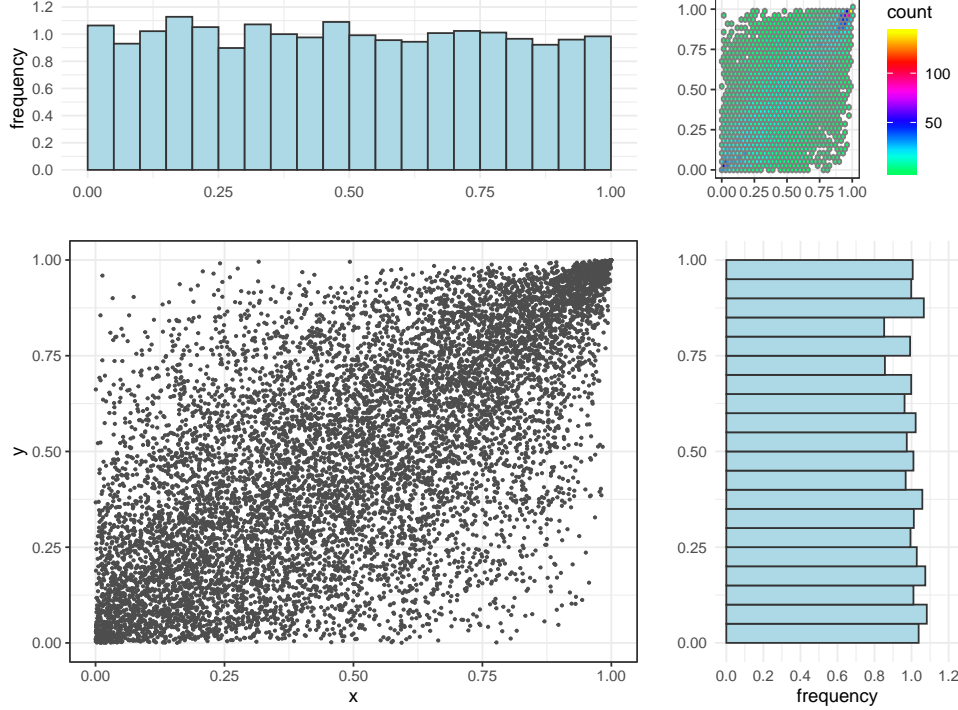


FIGURE 7. Sample of size  $n = 10000$  from the Gumbel copula with parameter  $\theta = 3$  as in Example 6.2 (lower left panel); two-dimensional histogram (upper right panel) and marginal histograms (upper left and lower right panel).

we use the R-package ‘copula’ (see [20]) to estimate the empirical Kendall distribution function  $\hat{F}_n^{\text{Kendall}}$  as described in [13]. Figure 8 (left panel) depicts the real as well as the estimated Kendall distribution function for the sample sizes  $n = 500$  and  $n = 10000$ , respectively. Based on  $\hat{F}_n^{\text{Kendall}}$  we derive the estimated (normalized) generator  $\hat{\varphi}_n$  by (again see Figure 8)

$$\hat{\varphi}_n(x) = \exp \left( \text{sign} \left( x - \frac{1}{2} \right) \int_{\min(\frac{1}{2}, x)}^{\max(\frac{1}{2}, x)} \frac{1}{t - \hat{F}_n^{\text{Kendall}}(t)} dt \right).$$

Given  $\hat{\varphi}_n$  we finally calculate the estimated Archimedean copula  $C_{\hat{\varphi}_n}$  and calculate its Markov kernel  $K_{C_{\hat{\varphi}_n}}$ . For the dependence measure  $\zeta_1$  using the R-package ‘qad’ ([17]) we obtained the following values:  $\zeta_1(C_{\hat{\varphi}_{500}}) = 0.7117$ ,  $\zeta_1(C_{\hat{\varphi}_{10000}}) = 0.7041$ ,  $\zeta_1(C_\varphi) = 0.6910$ .

As in the Extreme Value setting we perform a small simulation study comparing the performance of the plugin estimator with the estimator  $\hat{r}_n$  of the coefficient of correlation  $r$  (ignoring the Archimedean information). Generating samples of the Gumbel copula with parameter  $\theta = 3$ , calculating  $\hat{r}_n$  as well as  $r(C_{\hat{\varphi}_n}) = 6 \cdot D_2(C_{\hat{\varphi}_n}, \Pi)$  for these samples and repeating this procedure  $R = 5000$  times yielded the results depicted in Figure 9. Not surprisingly, for small to moderate sample sizes the plugin estimator  $r(C_{\hat{\varphi}_n})$  outperforms the general estimator  $\hat{r}_n$ , for large sample sizes both estimators perform comparably well.

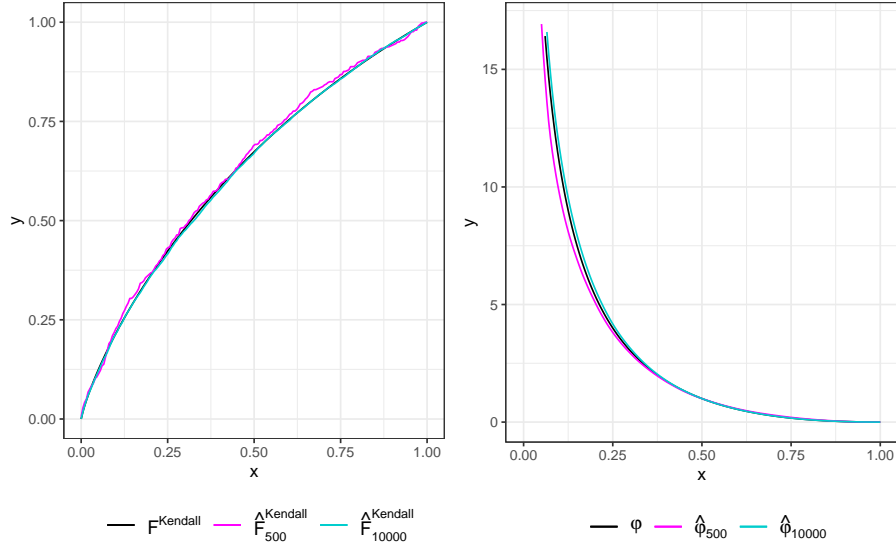


FIGURE 8. Kendall distribution function of a Gumbel copula with parameter  $\theta = 3$  (black) and its estimate  $\hat{F}_n^{\text{Kendall}}$  for  $n = 500$  and  $n = 10000$  as considered in Example 6.2 (left panel); (normalized) generator (black) and its estimates  $\hat{\varphi}_n$  for  $n = 500$  and  $n = 10000$  as considered in Example 6.2 (right panel).

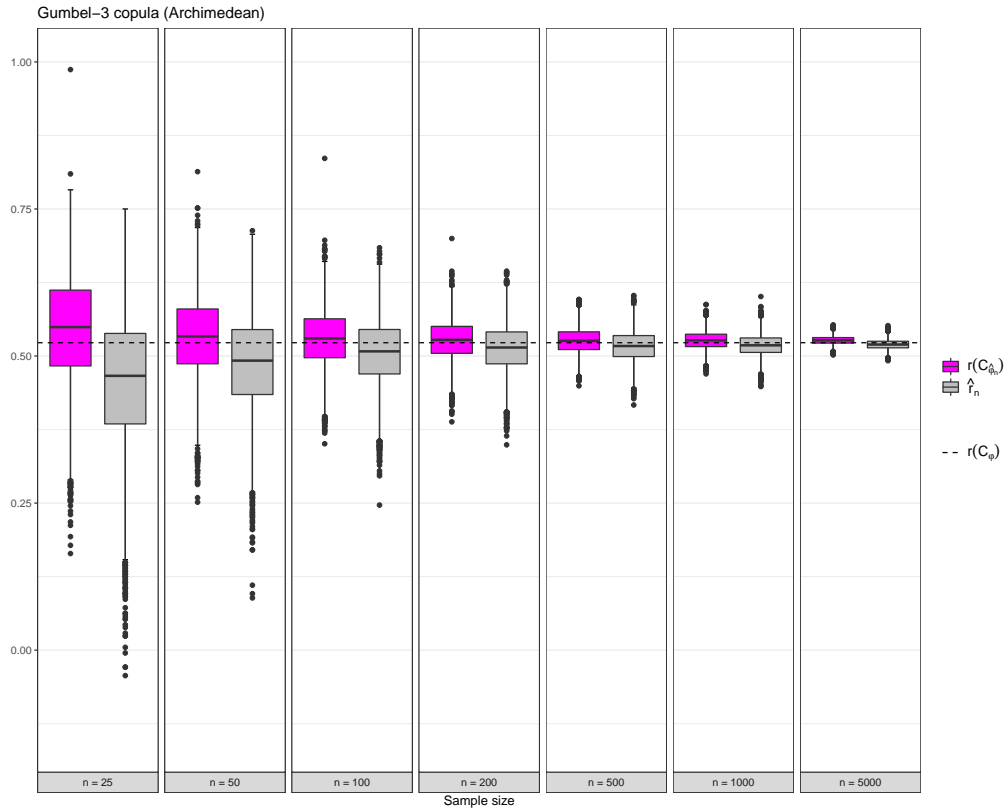


FIGURE 9. Boxplots of the obtained value of  $\hat{r}_n$  and  $r(C_{\hat{A}_n^*})$  based on samples from the Gumbel copula with parameter  $\theta = 3$ .

## ACKNOWLEDGEMENTS

The first author gratefully acknowledges the financial support from Porsche Holding Austria and Land Salzburg within the WISS 2025 project ‘KFZ’ (P1900123). Moreover, the second and the third author gratefully acknowledge the support of the WISS 2025 project ‘IDA-lab Salzburg’ (20204-WISS/225/197-2019 and 0102-F1901166-KZP).

## REFERENCES

- [1] Barbe, P., C. Genest, K. Ghouli, and B. Rémillard (1996). On Kendall’s process. *J. Multivariate Analysis* 58(2), 197 – 229.
- [2] Billingsley, P. (2013). *Convergence of Probability Measures*. Wiley Series in Probability and Statistics. Wiley.
- [3] Capéraà, P., A.-L. Fougères, and C. Genest (1997). A nonparametric estimation procedure for bivariate extreme value copulas. *Biometrika* 84, 567–577.
- [4] Charpentier, A. and J. Segers (2008). Convergence of Archimedean copulas. *Statist. Probab. Lett.* 78, 412–419.
- [5] Chatterjee, S. (2020a). A new coefficient of correlation. *Journal of the American Statistical Association* 0(0), 1–21.
- [6] Chatterjee, S. (2020b). *XICOR: Association Measurement Through Cross Rank Increments*. R package version 0.3.3.
- [7] Cherubini, U., F. Durante, and S. Mulinacci (2013). *Marshall-Olkin Distributions - Advances in Theory and Applications*. Springer International Publishing Switzerland: Springer Proceedings in Mathematics & Statistics.
- [8] de Haan, L. and S. Resnick (1977). Limit theorem for multivariate sample extremes. *Z. für Wahrscheinlichkeitstheorie* 40, 317–337.
- [9] Dette, H., K. Siburg, and P. Stoimenov (2013). A Copula-Based Non-parametric Measure of Regression Dependence. *Scand. J. Stat.* 40, 21–41.
- [10] Durante, F. and C. Sempi (2016). *Principles of copula theory*. Boca Raton FL: Taylor & Francis Group LLC.
- [11] Fernández Sánchez, J. and W. Trutschnig (2015a). Conditioning based metrics on the space of multivariate copulas, their interrelation with uniform and levelwise convergence and Iterated Function Systems. *J. Theoret. Probab.* 28, 1311–1336.
- [12] Fernández Sánchez, J. and W. Trutschnig (2015b). Singularity aspects of Archimedean copulas. *J. Math. Anal. Appl.* 432, 103–113.
- [13] Genest, C., J. Nešlehová, and J. Ziegel (2011). Inference in multivariate Archimedean copula models. *Test* 20, 223.
- [14] Genest, C. and L. Rivest (1993). Statistical inference procedures for bivariate Archimedean copulas. *J. Amer. Stat. Assoc.* 88, 1034–1043.
- [15] Genest, C. and J. Segers (2009). Rank-based inference for bivariate extreme-value copulas. *Annals of Statistics* 37, 2990–3022.
- [16] Ghouli, K., A. Khoudraji, and L.-P. Rivest (1998). Propriétés statistiques des copules de valeurs extrêmes bidimensionnelles. *Canad. J. Statist.* 26, 187–197.
- [17] Griessenberger, F., R. R. Junker, and W. Trutschnig (2020). *qad: Quantification of Asymmetric Dependence*. R package version 0.1.2.
- [18] Gudendorf, G. and J. Segers (2010). Extreme-value copulas. In P. Jaworski, F. Durante, K. Härdle, and T. Rychlik (Eds.), *Copula Theory and Its Applications*, Chapter 6, pp. 127–145. Heidelberg: Springer, Berlin, Heidelberg.
- [19] Gudendorf, G. and J. Segers (2011). Nonparametric estimation of an extreme-value copula in arbitrary dimensions. *J. Multivariate Analysis* 102(1), 37 – 47.

- [20] Hofert, M., I. Kojadinovic, M. Maechler, and J. Yan (2020). *copula: Multivariate Dependence with Copulas*. R package version 0.999-20.
- [21] Junker, R. R., F. Griessenberger, and W. Trutschnig (2020). Estimating scale-invariant directed dependence of bivariate distributions. *Computational Statistics & Data Analysis* 153(0), 0.
- [22] Kallenberg, O. (2002). *Foundations of modern probability*. New York: Springer-Verlag.
- [23] Kannan, R. and C. Krueger (2012). *Advanced Analysis: on the Real Line*. Universitext. Springer New York.
- [24] Klement, E., R. Mesiar, and E. Pap (2000). *Triangular Norms* (8 ed.). Springer Netherlands.
- [25] Klenke, A. (2008). *Wahrscheinlichkeitstheorie*. Berlin Heidelberg: Springer Lehrbuch Masterclass Series.
- [26] Lancaster, H. O. (1963, 12). Correlation and complete dependence of random variables. *Ann. Math. Statist.* 34(4), 1315–1321.
- [27] Li, X., P. Mikusinski, and M. Taylor (1998). Strong approximation of copulas. *J. Math. Anal. Appl.* 255, 608–623.
- [28] Marshall, A. W. (1970). Discussion of barlow and van zwet’s papers. In M. L. Puri (Ed.), *Nonparametric Techniques in Statistical Inference*, pp. 175–176. London: Cambridge University Press.
- [29] Mikusiński, P. and M. Taylor (2010). Some approximations of n-copulas. *Metrika* (72), 385–414.
- [30] Nelsen, R. (2006). *An Introduction to Copulas*. Berlin Heidelberg: Springer-Verlag.
- [31] Pickands, J. (1981). Multivariate extreme value distributions. *Proceedings 43rd Session International Statistical Institute* 2, 859–878.
- [32] Pollard, D. (2001). *A User’s Guide to Measure Theoretic Probability*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press.
- [33] Rockafellar, R. (1970). *Convex Analysis*. Princeton Landmarks in Mathematics and Physics. Princeton University Press.
- [34] Rudin, W. (1987). *Real and Complex Analysis, 3rd Ed.* USA: McGraw-Hill, Inc.
- [35] Sempi, C. (2004). Convergence of copulas: Critical remarks. *Radovi Matematički* 12, 241–249.
- [36] Trutschnig, W. (2011). On a strong metric on the space of copulas and its induced dependence measure. *J. Math. Anal. Appl.* 384, 690–705.
- [37] Trutschnig, W. (2012). Some results on the convergence of (quasi-) copulas. *Fuzzy Sets and Systems* 191, 113–121.
- [38] Trutschnig, W., M. Schreyer, and J. Fernández Sánchez (2016). Mass distribution of two-dimensional extreme-value copulas and related results. *Extremes* 19, 405–427.

DEPARTMENT FOR MATHEMATICS, UNIVERSITY OF SALZBURG, HELLBRUNNERSTRASSE 34, A-5020 SALZBURG, AUSTRIA

Email address: thimo.kasper@sbg.ac.at, sebastian.fuchs@sbg.ac.at, wolfgang.trutschnig@sbg.ac.at