

On differentiability and mass distributions of typical bivariate copulas

Nicolas Pascal Dietrich^a, Wolfgang Trutschnig^{a,*}

^a*University of Salzburg, Department for Artificial Intelligence and Human Interfaces, Hellbrunnerstrasse 34, Salzburg, 5020, Salzburg, Austria*

Abstract

Despite the fact that copulas are commonly considered as analytically smooth/regular objects, derivatives of copulas have to be handled with care. Triggered by a recently published result characterizing multivariate copulas via $(d - 1)$ -increasingness of their partial derivative we study the bivariate setting in detail and show that the set of non-differentiability points of a copula may be quite large. We first construct examples of copulas C whose first partial derivative $\partial_1 C(x, y)$ is pathological in the sense that for almost every $x \in (0, 1)$ it does not exist on a dense subset of $y \in (0, 1)$, and then show that the family of these copulas is dense. Since in commonly considered subfamilies more regularity might be typical, we then focus on bivariate Extreme Value copulas (EVCs) and show that a topologically typical EVC is not absolutely continuous but has degenerated discrete component, implying that in this class typically $\partial_1 C(x, y)$ exists in full $(0, 1)^2$. Considering that regularity of copulas is closely related to their mass distributions we then study mass distributions of topologically typical copulas and prove the surprising fact that topologically typical bivariate copulas are mutually completely dependent with full support. Furthermore, we use the characterization of EVCs in terms of their associated Pickands dependence measures ϑ on $[0, 1]$, show that regularity of ϑ carries over to the corresponding EVC and prove that the subfamily of all EVCs whose absolutely continuous, discrete and singular component has full support is dense in the class of all EVCs.

Keywords: Derivative, Copula, Extreme Value copula, Markov kernel, Category theory
2020 MSC: 62G32, 54E52, 28A50, 26A30

1. Introduction

Constituting the link between multivariate distribution functions and their univariate marginals (see [41]) as well as their resulting prominent role in the context of modeling stochastic dependence (see [7, 16, 17, 18, 28, 31, 32, 39]), over the past decades copulas have become an essential tool in probability theory and statistics, both from a theoretical as well as from an applied perspective. Considering that bivariate copulas are Lipschitz

*Corresponding author

Email addresses: nicolaspascal.dietrich@plus.ac.at (Nicolas Pascal Dietrich), wolfgang.trutschnig@plus.ac.at (Wolfgang Trutschnig)

continuous, Rademacher’s Theorem guarantees the existence of their partial derivatives almost everywhere in the sense of the two-dimensional Lebesgue-measure λ_2 . The exception set in Rademacher’s Theorem applied to copulas can, however, be large in the sense of even having full support, a property that sometimes seems to have been overlooked or not handled with sufficient care.

The afore-mentioned exception set is also relevant in the context of a recently published paper (see [37, Corollary 4.2]) characterizing copulas in terms of $(d - 1)$ -monotonicity of their partial derivative. It is straightforward to show that in general there exists no set $D \subseteq \mathbb{I}$ with $\lambda(D) = 1$ (λ denoting the one-dimensional Lebesgue measure) such that for every $x \in D$ the partial derivative $\partial_1 C(x, y)$ of a bivariate copula C exists for all $y \in D$. We therefore first show the existence of bivariate copulas C fulfilling that for λ -almost every $x \in \mathbb{I}$ there exists some $y_x \in \mathbb{I}$ such that the partial derivative $\partial_1 C(x, y_x)$ does not exist and then construct more pathological examples exhibiting the property that for λ -almost every $x \in \mathbb{I}$ there exists a countable dense set $\mathcal{Q}_x \subseteq \mathbb{I}$ such that $\partial_1 C(x, y)$ does not exist for any $y \in \mathcal{Q}_x$. Illustrating the fact that such pathological behavior might not be as uncommon as expected we then prove that the family of all such copulas is dense in the full class \mathcal{C} of all bivariate copulas equipped with the standard uniform metric d_∞ .

Turning towards mixed partial derivatives $\partial_1 \partial_2 C$ and $\partial_2 \partial_1 C$ of a copula C and recalling [32, Theorem 2.2.8.] it is a direct consequence of Schwarz’s theorem from calculus that in the case that $\partial_1 \partial_2 C$ and $\partial_2 \partial_1 C$ exist and are continuous everywhere on $(0, 1)^2$, they coincide. Since in the general setting the mixed partial derivatives do not even exist everywhere, the interplay between $\partial_1 \partial_2 C$ and $\partial_2 \partial_1 C$ is less obvious in general. Working with Markov kernels (regular conditional distributions) we clarify this interrelation and present a generalized version of Schwarz’s theorem for copulas extending [32, Theorem 2.2.8.].

Viewing the family of copulas from a purely topological perspective and considering that the above-mentioned pathological examples are completely dependent (or convex combinations of completely dependent) copulas the question naturally arises, whether the family of completely dependent copulas is ‘small’ or ‘large’ in the sense of being atypical or typical. Working with Baire categories (see e.g. [34]), topology offers a natural way to differentiate between ‘small’ and ‘large’ sets. We call a given subset of a topological space (\mathcal{J}, τ) nowhere dense if, and only if the interior of its closure is empty. A subset of (\mathcal{J}, τ) is referred to as meager/of first (Baire) category, if it can be covered by a countable union of nowhere dense sets. A set is of second (Baire) category if it is not meager. Moreover, a set is called co-meager if it is the complement of a meager set. Proceeding as in [2] and returning to the concept of ‘small’ and ‘large’ sets, in a complete metric space meager sets are the ‘small’ sets, sets of second category are interpreted as ‘not small’ and co-meager sets are the ‘large’ sets. Following this very interpretation, in what follows we will refer to elements of a co-meager set as typical and to elements of a meager set as atypical.

Working with the afore-mentioned topological concepts C.W. Kim (see [23]) proved the striking result that a typical bivariate copula is (mutually) completely dependent (see the subsequent section for a definition). Building upon his result, in Section 3 we prove a slightly stronger result: A typical copula C is mutually completely dependent and (its associated doubly stochastic measure μ_C) has full support. For a variety of manuscripts studying copulas in the context of Baire categories we refer to [4, 5, 6, 10] and the references therein.

Returning to the existence of $\partial_1 C(x, y)$ one might naturally conjecture that in commonly considered subclasses more regularity might be typical. Here we therefore consider the well-known family of bivariate Extreme-Value copulas \mathcal{C}_{ev} (EVC, for short) and study regularity in this class. Due to their simple algebraic form and practical aspects EVCs are particularly applied in finance and hydrology (see, e.g., [29, 40]). Tackling regularity results and mass distributions of EVCs we first study the projection of the spectral measure H , defined on the unit simplex (see [1, 8, 14, 42]), to \mathbb{I} , refer to this measure ϑ as Pickands dependence measure, and then show how singularity/regularity properties of ϑ carry over to the corresponding EVC. Doing so now allows to derive some of the results given in [27] and [45] in a simplified manner but also opens the door to proving the fact that EVCs whose discrete, singular and absolutely continuous component have full support are dense in $(\mathcal{C}_{ev}, d_\infty)$.

Finally, again working with Pickands dependence measures we prove that - contrasting the fact that typical copulas in (\mathcal{C}, d_∞) are mutually completely dependent and hence discrete in the sense introduced in the next section - typical EVCs have degenerated discrete component, are not absolutely continuous, but have full support. In particular, for typical EVCs $\partial_1 C(x, y)$ exists in full $(0, 1)^2$; EVCs can therefore be considered as typically quite regular.

The remainder of this paper is organized as follows: Section 2 contains notation and preliminaries that are used throughout the text. Section 3 studies differentiability of bivariate copulas and shows that the family of all bivariate copulas exhibiting the property that for λ -almost every $x \in (0, 1)$ there exists a countable dense set $\mathcal{Q}_x \subseteq \mathbb{I}$ such that $\partial_1 C(x, y)$ does not exist for any $y \in \mathcal{Q}_x$, is dense in (\mathcal{C}, d_∞) . The afore-mentioned extension of Schwarz's theorem for bivariate copulas is proved and the surprising and counter-intuitive fact, saying that a topologically typical copula C in (\mathcal{C}, d_∞) is mutually completely dependent and has full support, is established. Section 4 focuses on Extreme Value copulas, first recalls and then slightly extends the one-to-one-to-one interrelation between EVCs, Pickands dependence functions and Pickands measures. It is proved that various properties of the Pickands dependence measure ϑ carry over to the corresponding EVC and that the family of all EVCs whose discrete, singular and absolutely continuous component have full support are dense in $(\mathcal{C}_{ev}, d_\infty)$. Establishing the result that typical EVCs have degenerated discrete component, are not absolutely continuous, but have full support, concludes Section 4. The latter results also close the circle since they imply that for typical EVCs C the partial derivative $\partial_1 C(x, y)$ exists in full $(0, 1)^2$.

In order to simplify reading some technical lemmas and proofs have been shifted to the Appendix. Several additional examples and graphics illustrate the chosen procedures and some underlying ideas.

2. Notation and Preliminaries

In the sequel \mathcal{C} denotes the family of all bivariate copulas. Given a copula $C \in \mathcal{C}$, as usual the corresponding doubly stochastic measure will be denoted by μ_C , i.e., $\mu_C([0, x] \times [0, y]) := C(x, y)$ for all $(x, y) \in \mathbb{I}^2$ with $\mathbb{I} := [0, 1]$. A copula $C \in \mathcal{C}$ is called exchangeable if the transposed copula C^t coincides with C , i.e., if $C^t(x, y) := C(y, x) = C(x, y)$ holds for all $x, y \in \mathbb{I}$. \mathcal{C}_e denotes the family of all exchangeable copulas.

For an arbitrary topological space (S, τ) we denote the Borel σ -field on S by $\mathcal{B}(S)$, and let $\mathcal{P}(S)$ denote the family of all probability measures on $\mathcal{B}(S)$. Considering weak convergence of probability measures, we denote the topology induced by the afore-mentioned notion of convergence by τ_w . Moreover, for an arbitrary measure ν on $\mathcal{B}(S)$ the support of ν , i.e., the complement of the union of all open sets U with the property that $\nu(U) = 0$, will be denoted by $\text{supp}(\nu)$. Considering a set $E \subseteq S$, we write \overline{E} for the topological closure of the set E and $\text{int}(E)$ for the interior of E . Throughout this contribution, the support of a copula $C \in \mathcal{C}$ by definition will be the support of its corresponding doubly stochastic measure μ_C . Considering the uniform metric d_∞ on \mathcal{C} it is well-known that (\mathcal{C}, d_∞) is a compact metric space. For more background on copulas and doubly stochastic measures we refer to [7, 32].

The Lebesgue-measure on $\mathcal{B}(\mathbb{I}^2)$ is denoted by λ_2 , for the univariate Lebesgue measure we will simply write λ . For every $x \in S$ we denote the Dirac measure in $x \in S$ by δ_x . Furthermore, the space of all Lebesgue integrable functions on \mathbb{I}^d will be denoted by $L^1(\mathbb{I}^d, \mathcal{B}(\mathbb{I}^d), \lambda_d)$ for every $d \in \mathbb{N}$. Considering two metric spaces (S, d) and (S', d') , a Borel-measurable transformation $T : S \rightarrow S'$ and a probability measure $\nu \in \mathcal{P}(S)$, the push-forward (measure) ν^T of ν via T is defined by $\nu^T(F) := \nu(T^{-1}(F))$ for all $F \in \mathcal{B}(S')$.

Throughout this contribution conditional distributions and Markov kernels will play a prominent role. A *Markov kernel* from \mathbb{R} to \mathbb{R} is a map $K : \mathbb{R} \times \mathcal{B}(\mathbb{R}) \rightarrow \mathbb{I}$ fulfilling that (i) for every fixed $E \in \mathcal{B}(\mathbb{R})$ the function $x \mapsto K(x, E)$ is Borel-measurable and (ii) for every $x \in \mathbb{R}$ the map $E \mapsto K(x, E)$ is a probability measure on $\mathcal{B}(\mathbb{R})$. If in (ii) we only have that the measure $E \mapsto K(x, E)$ fulfills $K(x, \mathbb{I}) \leq 1$ (instead of $K(x, \mathbb{I}) = 1$), then K is called sub-Markov kernel.

Suppose that (X, Y) is a random vector on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Then a Markov kernel $K(\cdot, \cdot)$ will be called a regular conditional distribution of Y given X if for every set $E \in \mathcal{B}(\mathbb{R})$ the equation

$$K(X(\omega), E) = \mathbb{E}(\mathbf{1}_E \circ Y | X)(\omega)$$

holds for \mathbb{P} -almost every $\omega \in \Omega$. It is a well-established fact that for each pair (X, Y) of random variables, a regular conditional distribution $K(\cdot, \cdot)$ of Y given X exists and is unique for \mathbb{P}^X -almost every $x \in \mathbb{R}$. Assuming that (X, Y) has distribution function $C \in \mathcal{C}$ (more precisely, the natural extension of C to \mathbb{R}^2 is the distribution function of (X, Y)), we will write $(X, Y) \sim C$, let $K_C : \mathbb{I} \times \mathcal{B}(\mathbb{I}) \rightarrow \mathbb{I}$ denote (a version of) the regular conditional distribution of Y given X and call it the *Markov kernel* of C . Fixing $x \in \mathbb{I}$ and defining the x -section G_x of an arbitrary set $G \in \mathcal{B}(\mathbb{I}^2)$ by $G_x := \{y \in \mathbb{I} : (x, y) \in G\} \in \mathcal{B}(\mathbb{I})$ applying *disintegration* (see [19, Section 5] and [24, Section 8]) yields that

$$\mu_C(G) = \int_{\mathbb{I}} K_C(x, G_x) d\lambda(x). \tag{2.1}$$

In what follows, a measure ν on $(\mathbb{I}^d, \mathcal{B}(\mathbb{I}^d))$ with $d \in \mathbb{N}$ will be called singular (with respect to λ_d) if, and only if (i) ν has no point masses and (ii) there exists some set $G \in \mathcal{B}(\mathbb{I}^d)$ with $\lambda_d(G) = 0$, $\nu(G) = \nu(\mathbb{I}^d)$. Obviously the doubly stochastic measure μ_C associated with a copula $C \in \mathcal{C}$ always has degenerated discrete component (in the sense that μ_C has no point masses). Following [21], however, and using the Lebesgue-decomposition of the Markov kernel $K_C(x, \cdot)$ of C into absolutely continuous, discrete and singular sub-kernels $K_C^{abs}(\cdot, \cdot), K_C^{dis}(\cdot, \cdot), K_C^{sing}(\cdot, \cdot) : \mathbb{I} \times \mathcal{B}(\mathbb{I}) \rightarrow \mathbb{I}$, respectively, i.e.,

$$K_C(x, F) = K_C^{abs}(x, F) + K_C^{dis}(x, F) + K_C^{sing}(x, F) \quad (2.2)$$

for $x \in \mathbb{I}$ and $F \in \mathcal{B}(\mathbb{I})$, see [25], allows for a very natural definition of the absolutely continuous, the discrete and the singular component of C . In fact, working with disintegration and equation (2.2), we can define the absolutely continuous, the discrete and the singular components μ_C^{abs}, μ_C^{dis} and μ_C^{sing} of μ_C by

$$\begin{aligned} \mu_C^{abs}(E \times F) &:= \int_E K_C^{abs}(\mathbf{x}, F) d\lambda(x), \\ \mu_C^{dis}(E \times F) &:= \int_E K_C^{dis}(\mathbf{x}, F) d\lambda(x), \\ \mu_C^{sing}(E \times F) &:= \int_E K_C^{sing}(\mathbf{x}, F) d\lambda(x), \end{aligned} \quad (2.3)$$

for every $E \in \mathcal{B}(\mathbb{I})$ and every $F \in \mathcal{B}(\mathbb{I})$, and extend them to full $\mathcal{B}(\mathbb{I}^2)$ in the standard way. Throughout this paper we call a copula C absolutely continuous, discrete or singular if, and only if $\mu_C^{abs}(\mathbb{I}^2) = 1$, $\mu_C^{dis}(\mathbb{I}^2) = 1$ or $\mu_C^{sing}(\mathbb{I}^2) = 1$, respectively. Moreover, $\mu_C^{abs}, \mu_C^{dis}, \mu_C^{sing}$ will be referred to as the absolutely continuous, the discrete, and the singular components of C (or μ_C), respectively. Notice that the afore-mentioned composition of μ_C is finer than the classical Lebesgue decomposition of μ_C with respect to λ_2 - in fact, $\mu_C^{dis} + \mu_C^{sing}$ coincides with the singular component of the classical Lebesgue decomposition of μ_C . Moreover, equation (2.3) allows for the discrete component of a copula to be non-degenerated: Considering, e.g., $C = M$ we have $\mu_C^{dis}(\mathbb{I}^2) = 1$ as well as $\mu_C^{abs}(\mathbb{I}^2) = \mu_C^{sing}(\mathbb{I}^2) = 0$, so - in the sense of the decomposition (2.3) - the minimum copula M might be referred to as purely discrete (since the other two components are degenerated).

Working particularly in the context of quantifying the extent of dependence of a random variable Y on a random variable X , stronger metrics than d_∞ have to be considered in \mathcal{C} : The Markov-kernel based metrics D_p , first introduced in [43], are given by

$$D_p(A, B) := \left(\int_{\mathbb{I}^2} |K_A(x, [0, y]) - K_B(x, [0, y])|^p d\lambda_2(x, y) \right)^{\frac{1}{p}},$$

for $p \in [1, \infty)$ and

$$D_\infty(A, B) := \sup_{y \in \mathbb{I}} \int_{\mathbb{I}} |K_A(x, [0, y]) - K_B(x, [0, y])| d\lambda(x),$$

for $p = \infty$ and arbitrary $A, B \in \mathcal{C}$. As shown in [43] the topologies induced by the metrics D_p are all equivalent and the topology induced by D_p (for an arbitrary $p \in [0, \infty]$)

is strictly finer than the one induced by d_∞ . In other words: For fixed $p \in [1, \infty]$ and copulas $C, C_1, C_2, \dots \in \mathcal{C}$ we have that $D_p(C_n, C) \xrightarrow{n \rightarrow \infty} 0$ implies $d_\infty(C_n, C) \xrightarrow{n \rightarrow \infty} 0$ but not necessarily vice versa, see [43] for a counter-example.

From a statistical perspective, one might even consider a stronger but natural notion of convergence, that of weak convergence of Markov kernels (wcc, for short), first introduced in [22]: Suppose that C, C_1, C_2, \dots are copulas with associated Markov kernels $K_C, K_{C_1}, K_{C_2}, \dots$. We say that the sequence $(C_n)_{n \in \mathbb{N}}$ converges weakly conditional to C if, and only if for λ -almost every $x \in \mathbb{I}$ the sequence $(K_{C_n}(x, \cdot))_{n \in \mathbb{N}}$ of probability measures on $\mathcal{B}(\mathbb{I})$ converges weakly to the probability measure $K_C(x, \cdot)$. In the following we write $C_n \xrightarrow{wcc} C$ to denote weak conditional convergence of $(C_n)_{n \in \mathbb{N}}$ to C . It is straightforward to show that wcc implies D_p convergence but not necessarily vice versa (again see [22]).

We call a map $h: \mathbb{I} \rightarrow \mathbb{I}$ λ -preserving if, and only if $\lambda^h(F) = \lambda(F)$ holds for every $F \in \mathcal{B}(\mathbb{I})$. A copula $C \in \mathcal{C}$ is called completely dependent if there exists some λ -preserving transformation h such that C concentrates its mass on the graph of h , or, equivalently, if its corresponding Markov kernel is given by $K_C(x, F) = \mathbf{1}_F(h(x))$ for λ -almost all $x \in \mathbb{I}$ (see [43] for more equivalent formulations). If, in addition, h is bijective, we call the associated completely dependent copula mutually completely dependent. Throughout this contribution we will denote the family of all completely dependent copulas by \mathcal{C}_{cd} and the family of all mutually completely dependent copulas by \mathcal{C}_{mcd} . A λ -preserving map $h: \mathbb{I} \rightarrow \mathbb{I}$ will be called a (classical) equidistant even shuffle of M with $N \in \mathbb{N}$ stripes if, and only if h is linear with slope 1 on each interval $I_N^i := (\frac{i-1}{N}, \frac{i}{N})$, injective on $\bigcup_{i=1}^N (\frac{i-1}{N}, \frac{i}{N})$ and just permutes the intervals I_N^1, \dots, I_N^N . In the sequel \mathcal{S}_N will denote the family of all completely dependent copulas whose corresponding λ -preserving transformation is an even shuffle with N stripes. The space of all completely dependent copulas whose associated λ -preserving transformation is an equidistant shuffle will be denoted by $\mathcal{S} := \bigcup_{N \in \mathbb{N}} \mathcal{S}_N$. For more information on completely dependent copulas we refer to [32, 43].

The last concept we will need is that of a checkerboard copula: Suppose that $B \in \mathcal{C}$ and $N \in \mathbb{N}$ are arbitrary but fixed and that $T = (t_{i,j}^N)_{i,j=1}^N$ is a matrix fulfilling that $(N \cdot t_{i,j}^N)_{i,j=1}^N$ is doubly stochastic. Furthermore, for every $(i, j) \in \{1, \dots, N\}^2$ define the affine transformation $w_{i,j}^N: \mathbb{I}^2 \rightarrow R_{i,j} := \bar{I}_N^i \times \bar{I}_N^j = [\frac{i-1}{N}, \frac{i}{N}] \times [\frac{j-1}{N}, \frac{j}{N}]$ by

$$w_{i,j}^N(x, y) = \left(\frac{i-1}{N} + \frac{x}{N}, \frac{j-1}{N} + \frac{y}{N} \right).$$

Then the copula CB_T^B , defined implicitly via

$$\mu_{CB_T^B}(G) := \sum_{i,j=1}^N t_{i,j}^N \mu_B^{w_{i,j}^N}(G \cap R_{i,j}), \quad G \in \mathcal{B}(\mathbb{I}^2) \quad (2.4)$$

is called the T -checkerboard of B . To simplify notation we will let \mathcal{CB}_N^B denote the family of all CB_T^B with T being an $N \times N$ matrix as described above and write

$$\mathcal{CB}^B := \bigcup_{N \in \mathbb{N}} \mathcal{CB}_N^B.$$

For a given copula $A \in \mathcal{C}$ the N -checkerboard- B approximation of A is defined via

$$\mu_{CB_N^B(A)}(G) = \sum_{i,j=1}^N \mu_A(R_{i,j}^N) \mu_B^{w_{i,j}^N}(G \cap R_{i,j}) \quad (2.5)$$

for arbitrary $G \in \mathcal{B}(\mathbb{I}^2)$. For more background on checkerboards we refer to [22] and the references therein.

Finally we quickly recall the one-to-one correspondence between bivariate copulas and Markov operators which we will use for deriving some first Baire category results on (mutually) completely dependent copulas. A linear operator T on $L^1(\mathbb{I}, \mathcal{B}(\mathbb{I}), \lambda)$ is called *Markov operator* (see [3], [26], [43]) if it fulfills the following three properties:

1. T is positive, i.e. $T(f) \geq 0$ whenever $f \geq 0$
2. $T(\mathbf{1}_{\mathbb{I}}) = \mathbf{1}_{\mathbb{I}}$
3. $\int_{\mathbb{I}} (Tf)(x) d\lambda(x) = \int_{\mathbb{I}} f(x) d\lambda(x)$

The class of all Markov operators on $L^1(\mathbb{I}, \mathcal{B}(\mathbb{I}), \lambda)$ will be denoted by \mathcal{M} . It is straightforward to see that the operator norm of T is one, i.e. $\|T\| := \sup\{\|Tf\|_1 : \|f\|_1 \leq 1\} = 1$ holds and the operators are continuous. According to [3] and [26] there is a one-to-one correspondence between \mathcal{C} and \mathcal{M} - in fact, the mappings $\Phi : \mathcal{C} \rightarrow \mathcal{M}$ and $\Psi : \mathcal{M} \rightarrow \mathcal{C}$, defined by

$$\begin{aligned} \Phi(C)(f)(x) & : = (T_C f)(x) := \frac{d}{dx} \int_{[0,1]} C_{,2}(x,t) f(t) d\lambda(t), \\ \Psi(T)(x,y) & : = C_T(x,y) := \int_{[0,x]} (T\mathbf{1}_{[0,y]})(t) d\lambda(t) \end{aligned} \quad (2.6)$$

for every $f \in L^1(\mathbb{I})$ and $(x,y) \in \mathbb{I}^2$ ($C_{,2}$ denoting the partial derivative with respect to y), fulfill $\Psi \circ \Phi = id_{\mathcal{C}}$ and $\Phi \circ \Psi = id_{\mathcal{M}}$. Following [43], for every $C \in \mathcal{C}$ the Markov operator T_C simplifies to

$$(T_C f)(x) := \int_{\mathbb{I}} f(t) K_C(x, dt).$$

Having this it follows easily that (i) \mathcal{C}_{cd} corresponds to the family \mathcal{M}_{cd} of all Markov operators of the form $T(f)(x) = f(h(x))$ for some λ -preserving transformation $h : \mathbb{I} \rightarrow \mathbb{I}$ and that (ii) \mathcal{C}_{mcd} corresponds to all $T \in \mathcal{M}_{cd}$ for which the corresponding h is bijective; the latter class will be denoted by \mathcal{M}_{mcd} . Additional useful facts on Markov operators and their interplay with copulas are the following: If $C \in \mathcal{C}$ corresponds to $T_C \in \mathcal{M}$, then C^t corresponds to the adjoint $T_C^{adj} = T_C^t$ of T_C (also interpreted as operator on $L^1(\mathbb{I}, \mathcal{B}(\mathbb{I}), \lambda)$, see [44]). Moreover it can be shown (again see [43]) that the afore-mentioned metrics D_p with $p \in [1, \infty)$ are metrizations of the strong operator topology on \mathcal{M} , whereas the weak operator topology corresponds to d_{∞} (see [33]).

3. Differentiability of copulas and typical mass distributions

We first derive various results on differentiability and then study regularity of topologically typical copulas.

3.1. Differentiability results

Corollary 4.2 in [37] translated to the bivariate setting states that a function $C: \mathbb{I}^2 \rightarrow \mathbb{I}$ that is grounded and has uniform marginals and absolutely continuous sections is a bivariate copula if, and only if its first partial derivative $y \mapsto \partial_1 C(x, y)$ is non-decreasing for λ -almost all $x \in \mathbb{I}$. Here we focus on differentiability and start with providing examples of copulas $C \in \mathcal{C}$ fulfilling the property that for λ -almost every $x \in \mathbb{I}$ there exists some point $y = y_x \in \mathbb{I}$ such that $\partial_1 C(x, y)$ does not exist. Throughout this contribution we denote the family of all such copulas by \mathcal{C}_p . The subsequent example shows that all equidistant shuffles are elements of \mathcal{C}_p .

Example 3.1 (Non-differentiability in one point). Suppose that $C \in \mathcal{S}$. Then there exists some $N \in \mathbb{N}$ with $C \in \mathcal{S}_N$. Let $x \in \mathbb{I} \setminus \bigcup_{i=0}^N \{\frac{i}{N}\}$ be arbitrary but fixed. Then there exists some $j \in \{1, \dots, N\}$ with $x \in (\frac{j-1}{N}, \frac{j}{N})$. Letting h denote the corresponding λ -preserving transformation, setting $y = h(x)$, then using the fact that h is monotonically increasing on $(\frac{j-1}{N}, \frac{j}{N})$ and calculating the right- and left-hand partial derivative, respectively, yields

$$\partial_1^+ C(x, h(x)) = \lim_{s \downarrow 0} \frac{C(x+s, h(x)) - C(x, h(x))}{s} = 0$$

as well as

$$\partial_1^- C(x, h(x)) = \lim_{s \uparrow 0} \frac{C(x+s, h(x)) - C(x, h(x))}{s} = 1.$$

The left-hand and right-hand derivatives exist but do not coincide, implying that $\partial_1 C(x, h(x))$ does not exist by definition. Figure 1 depicts two examples of elements of shuffles and points of non-differentiability.

The family of points of non-differentiability may be much larger - much more pathological cases than the one mentioned in Example 3.1 exist: As a second example we show the existence of a copula C with the following property: There exists some set $\Lambda \in \mathcal{B}(\mathbb{I})$ with $\lambda(\Lambda) = 1$ such that for every $x \in \Lambda$ there exists a dense set $\mathcal{Q}_x \subset \mathbb{I}$ such that $\partial_1 C(x, y)$ does not exist for any $y \in \mathcal{Q}_x$. In the sequel the family of all copulas having this property will be denoted by $\mathcal{C}_\mathcal{Q}$.

Example 3.2 (Non-differentiability on a dense subset). Consider the λ -preserving rotations $R_r(x) = x + r \pmod{1}$ for $r \in [0, 1)$ and define $K: \mathbb{I} \times \mathcal{B}(\mathbb{I}) \rightarrow \mathbb{I}$ by

$$K(x, E) := \sum_{n \in \mathbb{N}} \frac{1}{2^n} \delta_{R_{r_n}(x)}(E) = \sum_{n \in \mathbb{N}} \frac{1}{2^n} \mathbf{1}_E(R_{r_n}(x))$$

for $\{r_1, r_2, \dots\} = \mathbb{Q}_{0,1} := \mathbb{Q} \cap [0, 1)$ where $r_i \neq r_j$ for all $i \neq j$. Then K is obviously measurable in the first and a probability measure in the second argument, i.e., $K(\cdot, \cdot)$ is a

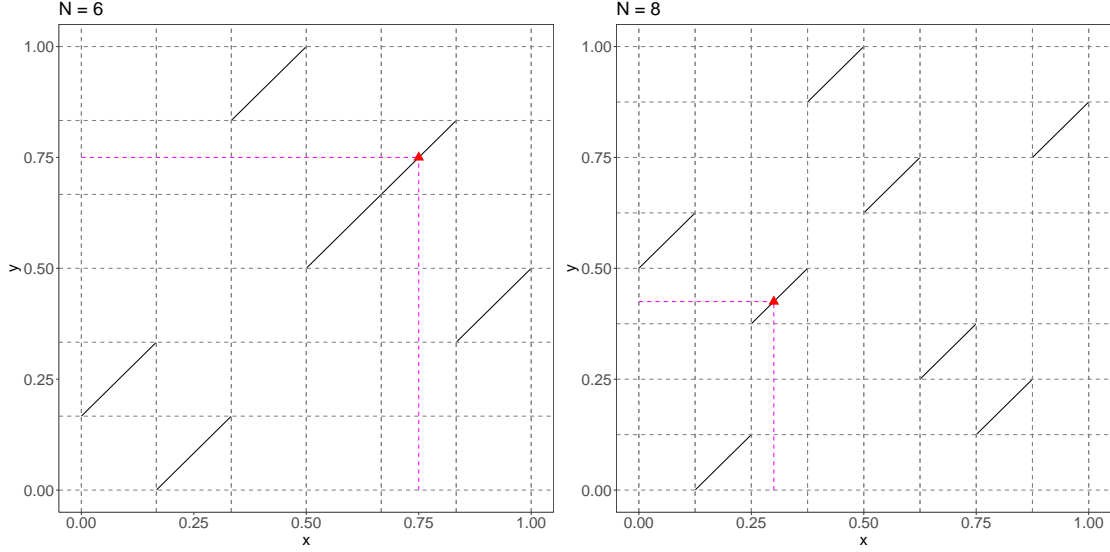


Figure 1: Two examples of equidistant shuffles as studied in Example 3.1. The red triangles depict points where the partial derivative with respect to x does not exist.

Markov kernel from \mathbb{I} to \mathbb{I} . Applying monotone convergence together with the fact that the rotations R_{r_n} are λ -preserving yields that

$$\int_{\mathbb{I}} K(x, E) d\lambda(x) = \int_{\mathbb{I}} \sum_{n \in \mathbb{N}} \frac{1}{2^n} \mathbf{1}_E(R_{r_n}(x)) d\lambda(x) = \sum_{n \in \mathbb{N}} \int_{\mathbb{I}} \frac{1}{2^n} \mathbf{1}_E(R_{r_n}(x)) d\lambda(x) = \lambda(E),$$

so $K(\cdot, \cdot)$ is the Markov kernel of a unique copula C . Fix $x_0 \in \mathbb{I} \setminus \mathbb{Q}$ and set $y_j := R_{r_j}(x_0)$. Then obviously the set $\{y_j : j \in \mathbb{N}\}$ is dense in \mathbb{I} . Writing

$$K(x, [0, y_j]) = \underbrace{\sum_{n \neq j} \frac{1}{2^n} \mathbf{1}_{[0, y_j]}(R_{r_n}(x))}_{:=L(x, y_j)} + \underbrace{\frac{1}{2^j} \mathbf{1}_{[0, y_j]}(R_{r_j}(x))}_{:=G_{r_j}(x, y_j)} \quad (3.1)$$

and using the fact that $x \mapsto \mathbf{1}_{[0, y_j]}(R_{r_n}(x))$ is continuous in x_0 for every $n \neq j$, uniform convergence of the series $L(x, y_j)$ implies that $x \mapsto L(x, y_j)$ is continuous in x_0 . Applying disintegration and using equation (3.1) yields

$$C(x, y_j) = \int_{[0, x]} K(s, [0, y_j]) d\lambda(s) = \underbrace{\int_{[0, x]} L(s, y_j) d\lambda(s)}_{:=A(x, y_j)} + \underbrace{\int_{[0, x]} G_{r_j}(s, y_j) d\lambda(s)}_{:=B(x, y_j)},$$

for every $x \in \mathbb{I}$. Calculating the right-hand partial derivative of A in x_0 , applying continuity of $x \mapsto L(x, y_j)$ in x_0 and using the fact that every point of continuity is a Lebesgue point (see [38]) yields

$$\partial_1^+ A(x_0, y_j) = \lim_{h \downarrow 0} \frac{1}{h} \int_{[x_0, x_0+h]} L(s, y_j) d\lambda(s) = L(x_0, y_j).$$

Proceeding analogously yields that the left-hand derivative of A in x_0 exists. Using the fact that $x_0 \neq 1 - r_j$ and calculating the right-hand partial derivative of B in x_0 yields

$$\partial_1^+ B(x_0, y_j) = \lim_{h \downarrow 0} \frac{1}{h} \int_{[x_0, x_0+h]} \frac{1}{2^j} \mathbf{1}_{[0, R_{r_j}(x_0)]}(R_{r_j}(s)) d\lambda(s) = 0,$$

and proceeding analogously for the left-hand partial derivative in x_0 we get

$$\partial_1^- B(x_0, y_j) = \lim_{h \downarrow 0} \frac{1}{h} \int_{[x_0-h, x_0]} \frac{1}{2^j} \mathbf{1}_{[0, R_{r_j}(x_0)]}(R_{r_j}(s)) d\lambda(s) = \frac{1}{2^j}.$$

Altogether it follows that $\partial_1 C(x_0, y_j)$ does not exist. Considering that $x_0 \in \mathbb{I} \setminus \mathbb{Q}$ and $j \in \mathbb{N}$ were arbitrary completes the proof.

Working with the previous example in combination with checkerboard copulas allows to show that $\mathcal{C}_{\mathcal{Q}}$ is dense in \mathcal{C} , implying that pathological cases of non-differentiability in the sense of Example 3.2 can be found ‘everywhere’ in \mathcal{C} even with respect to the finer topology induced by the metric D_p .

Theorem 3.3. *The set $\mathcal{C}_{\mathcal{Q}}$ is dense in (\mathcal{C}, D_p) for every $p \in [1, \infty]$.*

Proof. Fix $C \in \mathcal{C}$ and $n \in \mathbb{N}$. Choose an arbitrary $B \in \mathcal{C}_{\mathcal{Q}}$ and consider the N -checkerboard- B approximation $CB_N^B(C_n)$ of the copula $C_n := (1 - \frac{1}{n})C + \frac{1}{n}\Pi$ according to equation (2.5). Obviously C_n has full support so it follows immediately that $CB_N^B(C_n) \in \mathcal{C}_{\mathcal{Q}}$ holds for every $N \in \mathbb{N}$. Furthermore, applying [22, Theorem 3.2] yields that $CB_N^B(C_n) \xrightarrow{wcc} C_n$ as $N \rightarrow \infty$, implying $\lim_{n \rightarrow \infty} D_p(CB_N^B(C_n), C) = 0$. Finally, considering that we also have weak conditional convergence of $(C_n)_{n \in \mathbb{N}}$ to C for $n \rightarrow \infty$ the result follows. \square

Using the interrelations between D_p -convergence and convergence with respect to d_{∞} Theorem 3.3 has the following direct consequence:

Corollary 3.4. *The set $\mathcal{C}_{\mathcal{Q}}$ is dense in $(\mathcal{C}, d_{\infty})$.*

The previous example(s) built upon discontinuity of the conditional distributions - the following result shows that without discontinuities of the conditional distribution functions non-existence of the partial derivative can not happen.

Given a copula $C \in \mathcal{C}$ with corresponding Markov kernel K_C (without any additional regularity assumptions on C or K_C), according to [7, Theorem 3.4.4] for every fixed $y \in \mathbb{I}$ there exists some set $\Lambda_y \in \mathcal{B}(\mathbb{I})$ with $\lambda(\Lambda_y) = 1$ such that $\partial_1 C(x, y) = K_C(x, [0, y])$ holds for all $x \in \Lambda_y$. The next theorem shows that in the case that λ -almost all conditional distribution functions $y \mapsto K_C(x, [0, y])$ are continuous a simultaneous set Λ can be found.

Theorem 3.5. *Suppose that $C \in \mathcal{C}$ fulfills that λ -almost all conditional distribution functions $y \mapsto K_C(x, [0, y])$ are continuous. Then there exists a set $\Lambda \in \mathcal{B}(\mathbb{I})$ with $\lambda(\Lambda) = 1$ such that*

$$\partial_1 C(x, y) = K_C(x, [0, y]) \tag{3.2}$$

holds for every $x \in \Lambda$ and $y \in \mathbb{I}$.

Proof. By assumption there exists some $\Lambda_1 \in \mathcal{B}(\mathbb{I})$ with $\lambda(\Lambda_1) = 1$ such that $y \mapsto K_C(x, [0, y])$ is continuous for every $x \in \Lambda_1$. Fixing $y \in \mathbb{I}$, applying disintegration and Lebesgue's differentiation theorem (see [38]) yields the existence of some $\Lambda_y \in \mathcal{B}(\mathbb{I})$ with $\lambda(\Lambda_y) = 1$ such that

$$\partial_1 C(x, y) = K_C(x, [0, y])$$

holds for all $x \in \Lambda_y$. Defining $\Lambda_2 := \bigcap_{q \in \mathbb{Q} \cap \mathbb{I}} \Lambda_q$ implies $\lambda(\Lambda_2) = 1$ as well as

$$\partial_1 C(x, q) = K_C(x, [0, q]) \tag{3.3}$$

for every $x \in \Lambda_2$ and every $q \in \mathbb{Q} \cap \mathbb{I}$. Set $\Lambda := \Lambda_1 \cap \Lambda_2 \cap (0, 1)$ and fix $x \in \Lambda$. For every $y \in \mathbb{I}$ and $h > 0$ sufficiently small define the difference quotient $I_h(x, y)$ by

$$I_h(x, y) := \frac{C(x + h, y) - C(x, y)}{h}.$$

Then using 2-increasingness of C directly yields

$$I_h(x, q_1) \leq I_h(x, y) \leq I_h(x, q_2)$$

for arbitrary $q_1, q_2 \in \mathbb{Q} \cap \mathbb{I}$ with $q_1 \leq y \leq q_2$. According to equation (3.3) we have

$$\lim_{h \rightarrow 0} I_h(x, q_1) = K_C(x, [0, q_1])$$

as well as

$$\lim_{h \rightarrow 0} I_h(x, q_2) = K_C(x, [0, q_2]).$$

Fixing sequences $(q_n)_{n \in \mathbb{N}}, (r_n)_{n \in \mathbb{N}}$ in $\mathbb{Q} \cap \mathbb{I}$ with $q_n \downarrow y$ and $r_n \uparrow y$ and $\varepsilon > 0$ arbitrarily, using continuity of $y \mapsto K_C(x, [0, y])$ there exists some $n_0 \in \mathbb{N}$ such that

$$K_C(x, [r_n, q_n]) < \varepsilon$$

for all $n \geq n_0$. Working with the upper Dini-derivative (see [7]) yields

$$K_C(x, [0, y]) - \varepsilon < K_C(x, [0, r_n]) \leq \limsup_{h \downarrow 0} I_h(x, y) \leq K_C(x, [0, q_n]) < K_C(x, [0, y]) + \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary $\limsup_{h \downarrow 0} I_h(x, y) = K_C(x, [0, y])$ follows. Proceeding analogously for the lower Dini-derivative, we finally get that $\partial_1 C(x, y) = K_C(x, [0, y])$ for every $y \in \mathbb{I}$. Since $x \in \Lambda$ was arbitrary this completes the proof. \square

Reinterpreting Theorem 3.5 in the context of absolutely continuous, singular and discrete components of C , equation (3.2) obviously holds if C has degenerated discrete component. This implies the following result:

Corollary 3.6. *Suppose that $C \in \mathcal{C}$ fulfills $\mu_C^{dis}(\mathbb{I}^2) = 0$. Then there exists some set $\Lambda \in \mathcal{B}(\mathbb{I})$ with $\lambda(\Lambda) = 1$ such that*

$$\partial_1 C(x, y) = K_C(x, [0, y])$$

holds for all $x \in \Lambda$ and $y \in \mathbb{I}$.

Moreover, if $C \in \mathcal{C}$ is absolutely continuous with density f then there exists some set $\Lambda \in \mathcal{B}(\mathbb{I})$ with $\lambda(\Lambda) = 1$ such that for every $x \in \Lambda$ and every $y \in \mathbb{I}$ we have that

$$\partial_1 C(x, y) = \int_{[0, y]} f(x, t) d\lambda(t).$$

Proof. (i) Assume that $\mu_C^{dis}(\mathbb{I}^2) = 0$. Then according to equation (2.3), using Lebesgue's decomposition Theorem for Markov kernels, we obtain the existence of a set $\Lambda \in \mathcal{B}(\mathbb{I})$ with $\lambda(\Lambda) = 1$ such that the identity $K_C(x, [0, y]) = K_C^{abs}(x, [0, y]) + K_C^{sing}(x, [0, y])$ holds for all $y \in \mathbb{I}$ and $x \in \Lambda$. Since both $y \mapsto K_C^{abs}(x, [0, y])$ and $y \mapsto K_C^{sing}(x, [0, y])$ are continuous for every $x \in \Lambda$, the desired result follows by applying Theorem 3.5. The second assertion is a direct consequence of the fact that in the absolutely continuous case with density f a version of the Markov kernel $K_C(\cdot, \cdot)$ of C is given by $K_C(x, F) = \int_F f(x, t) d\lambda(t)$. \square

If the mixed partial derivatives $\partial_1 \partial_2 C(x, y)$ and $\partial_2 \partial_1 C(x, y)$ exist and are continuous for all $x, y \in (0, 1)$, then according to Schwarz's theorem they need to coincide, i.e.,

$$\partial_1 \partial_2 C(x, y) = \partial_2 \partial_1 C(x, y) \quad (3.4)$$

for all $x, y \in (0, 1)$ (see e.g. [32, Theorem 2.2.8.]). Turning to the general setting, the mixed partial derivatives $\partial_1 \partial_2 C$ and $\partial_2 \partial_1 C$ do not need to exist everywhere in $(0, 1)^2$ since, as pointed out in the previous examples, fixing x in a set of full λ -measure, not even the partial derivative $\partial_1 C(x, y)$ needs to exist for all $y \in \mathbb{I}$ and therefore equation (3.4) cannot hold for arbitrary $C \in \mathcal{C}$. On the other hand, for arbitrary $C \in \mathcal{C}$ with Markov kernel K_C , working again with disintegration and applying Lebesgue's differentiation theorem yields that for fixed $y \in \mathbb{I}$ there exists a set $\Lambda_y \in \mathcal{B}(\mathbb{I})$ with $\lambda(\Lambda_y) = 1$ such that

$$\partial_1 C(x, y) = K_C(x, [0, y])$$

holds for all $x \in \Lambda_y$. In other words: The partial derivative $\partial_1 C(x, y)$ coincides with the Markov kernel $K_C(\cdot, \cdot)$ on a 'good' set and thus $K_C(\cdot, \cdot)$ can be seen as a regularized version of $\partial_1 C$. Building upon this fact, working with Markov kernels we are able to generalize the result in [32, Theorem 2.2.8] to the whole family of bivariate copulas:

Theorem 3.7. *Suppose that $C \in \mathcal{C}$ is an arbitrary copula and let K_C and K_{C^t} denote Markov kernels of C and its transpose C^t , respectively. Furthermore let $f \in L^1(\mathbb{I}^2, \mathcal{B}(\mathbb{I}^2), \lambda_2)$ denote the Radon-Nikodym derivative of μ_C , i.e., the density of μ_C^{abs} . Then the identity*

$$\partial_y K_C(x, [0, y]) = f(x, y) = \partial_x K_{C^t}(y, [0, x])$$

holds for λ_2 -almost every $(x, y) \in \mathbb{I}^2$.

Proof. Working with the decomposition in equation (2.2) and aggregating the discrete and the singular component in K_C^\perp yields $K_C(x, F) = K_C^{abs}(x, F) + K_C^\perp(x, F)$ as well as $K_{C^t}(x, F) = K_{C^t}^{abs}(x, F) + K_{C^t}^\perp(x, F)$ for $x \in \mathbb{I}$ and $F \in \mathcal{B}(\mathbb{I})$. We want to show that the set

$$E := \{(x, y) \in \mathbb{I}^2 : \partial_y K_C^\perp(x, [0, y]) = 0\}$$

has full measure and proceed as follows: Being the sum of two sub-kernels of K_C , for fixed $y \in \mathbb{I}$ the function $x \mapsto K_C^\perp(x, [0, y])$ is measurable, and for fixed $x \in \mathbb{I}$ the function $y \mapsto K_C^\perp(x, [0, y])$ is non-decreasing and right-continuous. Using an argument analogous to the first part of the proof of [43, Lemma 4] shows that the map $(x, y) \mapsto$

$K_C^\perp(x, [0, y])$ is measurable. Applying the fact that $y \mapsto K_C^\perp(x, [0, y])$ is non-decreasing and right-continuous, and using [30, Theorem 1] yields that

$$\{(x, y) \in (0, 1)^2 : \partial_y K_C^\perp(x, [0, y]) \text{ exists and } \partial_y K_C^\perp(x, [0, y]) = 0\}$$

is a Borel set in $(0, 1)^2$. For fixed $x \in (0, 1)$ we have that $y \mapsto K_C^\perp(x, [0, y])$ has derivative 0 almost everywhere, so $\lambda(E_x) = 1$. Using Fubini's theorem therefore implies that

$$\lambda_2(E) = \int_{\mathbb{I}} \lambda(E_x) d\lambda(x) = 1.$$

Proceeding analogously we obtain the existence of a set $\tilde{E} \in \mathcal{B}(\mathbb{I}^2)$ with $\lambda_2(\tilde{E}) = 1$ such that $\partial_x K_{C^t}^\perp(y, [0, x]) = 0$ for all $(x, y) \in \tilde{E}$. Altogether there exists a set $Q := E \cap \tilde{E}$ with $\lambda(Q) = 1$ and $\partial_y K_C^\perp(x, [0, y]) = 0 = \partial_x K_{C^t}^\perp(y, [0, x])$ for all $(x, y) \in Q$. We will prove that

$$\partial_y K_C^{\text{abs}}(x, [0, y]) = f(x, y) = \partial_x K_{C^t}^{\text{abs}}(y, [0, x])$$

holds for λ_2 -almost all $(x, y) \in \mathbb{I}^2$ with f as in the theorem, and proceed as follows: Define a version of the absolutely continuous component K_C^{abs} of K_C by $K_C^{\text{abs}}(x, F) := \int_F f(x, y) d\lambda(y)$ for every $x \in \mathbb{I}$ and every $F \in \mathcal{B}(\mathbb{I})$. Similarly to the case for K_C^\perp , defining

$$G := \{(x, y) \in \mathbb{I}^2 : \partial_y K_C^{\text{abs}}(x, [0, y]) \text{ exists and } \partial_y K_C^{\text{abs}}(x, [0, y]) = f(x, y)\},$$

and applying the same line of argumentation as done for showing measurability of E yields Borel measurability of G . Working with the fact that $y \mapsto K_C^{\text{abs}}(x, [0, y])$ is differentiable λ -almost everywhere, applying Fubini's theorem yields $\lambda_2(G) = 1$. Considering the transposed copula C^t , using the identity $\mu_C^{\text{abs}}(F_1 \times F_2) = \mu_{C^t}^{\text{abs}}(F_2 \times F_1)$ for all $F_1, F_2 \in \mathcal{B}(\mathbb{I})$ and proceeding analogously to the previous case, there exists a set $\tilde{G} \in \mathcal{B}(\mathbb{I}^2)$ with $\lambda_2(\tilde{G}) = 1$ such that

$$\partial_x K_{C^t}^{\text{abs}}(y, [0, x]) = f(x, y)$$

holds for all $(x, y) \in \tilde{G}$. This implies that $\partial_y K_C^{\text{abs}}(x, [0, y]) = f(x, y) = \partial_x K_{C^t}^{\text{abs}}(y, [0, x])$ for all $(x, y) \in M := G \cap \tilde{G}$. Using $\lambda(M) = 1$ altogether yields $\lambda_2(Q \cap M) = 1$ as well as

$$\partial_y K_C(x, [0, y]) = \partial_y K_C^{\text{abs}}(x, [0, y]) = f(x, y) = \partial_x K_{C^t}^{\text{abs}}(y, [0, x]) = \partial_x K_C(y, [0, x]),$$

for all $(x, y) \in Q \cap M$. This completes the proof. \square

3.2. Baire category results for \mathcal{C}

The copulas proposed in Example 3.1 and Example 3.2 are completely dependent, complete dependence, however, may seem quite pathological. Nevertheless, using Baire categories it can be shown that complete dependence is much less pathological than one might assume. In fact, C.W. Kim proved in [23] that topologically typical Markov-operators are induced by a λ -preserving map. Using the fact that the family of all Markov-operators (endowed with the weak operator topology) is isomorphic to (\mathcal{C}, d_∞) (see [33]) we obtain the following translation to \mathcal{C} :

Theorem 3.8. \mathcal{C}_{mcd} and \mathcal{C}_{cd} are co-meager in (\mathcal{C}, d_∞) .

Proof. The fact that \mathcal{C}_{cd} is co-meager in (\mathcal{C}, d_∞) is an immediate consequence of [23, Theorem 1.1] and the afore-mentioned fact that the weak operator topology on \mathcal{M} corresponds to d_∞ on \mathcal{C} . According to [23] also $\mathcal{M}_{cd}^{adj} = \{T^{adj} : T \in \mathcal{M}_{cd}\}$ is co-meager in \mathcal{M} with respect to the weak operator topology. Thus, considering $\mathcal{M}_{mcd} = \mathcal{M}_{cd} \cap \mathcal{M}_{cd}^{adj}$ (again see [23]) and using the fact that finite intersections of co-meager sets are themselves co-meager, \mathcal{M}_{mcd} is co-meager in \mathcal{M} with respect to the weak operator topology, which completes the proof. \square

We conjecture that every $C \in \mathcal{C}_{mcd}$ is automatically an element of \mathcal{C}_p , i.e., that for λ -almost every $x \in \mathbb{I}$ there exists some $y \in \mathbb{I}$ such that $\partial_1 C(x, y)$ does not exist. We have, however, not been able to prove or falsify this conjecture since λ -preserving bijections may exhibit quite irregular behavior going far beyond being piecewise linear: As shown in [9] it is possible to construct a λ -preserving transformation h such that the induced completely dependent copula C_d has full support. Motivated by this result we now prove that typical bivariate copulas have full support.

Lemma 3.9. *The family $\{C \in \mathcal{C} : \text{supp}(\mu_C) = \mathbb{I}^2\}$ is co-meager in (\mathcal{C}, d_∞) .*

Proof. We define the set \mathcal{A}_N by

$$\mathcal{A}_N := \{C \in \mathcal{C} : \text{supp}(\mu_C)^c \text{ contains a cube } Q \text{ with } \lambda_2(Q) \geq \frac{1}{N}\}$$

and first show that \mathcal{A}_N is closed with respect to d_∞ . Suppose that $(C_\ell)_{\ell \in \mathbb{N}}$ is a sequence in \mathcal{A}_N with $\lim_{\ell \rightarrow \infty} d_\infty(C_\ell, C) = 0$ for some $C \in \mathcal{C}$. Knowing that for every $\ell \in \mathbb{N}$ there exists some cube $Q_\ell = (a_\ell, b_\ell) \times (c_\ell, d_\ell)$ with $\mu_{C_\ell}(Q_\ell) = 0$ and $\lambda_2(Q_\ell) \geq \frac{1}{N}$, applying Bolzano-Weierstrass yields the existence of subsequences $(a_{\ell_k})_{k \in \mathbb{N}}, (b_{\ell_k})_{k \in \mathbb{N}}, (c_{\ell_k})_{k \in \mathbb{N}}, (d_{\ell_k})_{k \in \mathbb{N}}$ and $a_0, b_0, c_0, d_0 \in \mathbb{I}$ such that $a_{\ell_k} \xrightarrow{k \rightarrow \infty} a_0, b_{\ell_k} \xrightarrow{k \rightarrow \infty} b_0, c_{\ell_k} \xrightarrow{k \rightarrow \infty} c_0$ and $d_{\ell_k} \xrightarrow{k \rightarrow \infty} d_0$. Using Lipschitz-continuity and considering the cube $Q := (a_0, b_0) \times (c_0, d_0)$ we therefore get

$$\begin{aligned} \mu_C(Q) &= C(b_0, d_0) - C(b_0, c_0) - C(a_0, d_0) + C(a_0, c_0) \\ &= \lim_{k \rightarrow \infty} [C_{\ell_k}(b_{\ell_k}, d_{\ell_k}) - C_{\ell_k}(b_{\ell_k}, c_{\ell_k}) - C_{\ell_k}(a_{\ell_k}, d_{\ell_k}) + C_{\ell_k}(a_{\ell_k}, c_{\ell_k})] \\ &= \lim_{k \rightarrow \infty} \mu_{C_{\ell_k}}(Q_{\ell_k}) = 0. \end{aligned}$$

Hence, considering $\frac{1}{N} \leq \lim_{k \rightarrow \infty} \lambda_2(Q_{\ell_k}) = \lim_{k \rightarrow \infty} (b_{\ell_k} - a_{\ell_k})(d_{\ell_k} - c_{\ell_k}) = (b_0 - a_0)(d_0 - c_0) = \lambda_2(Q)$ finally implies that \mathcal{A}_N is closed.

Taking an arbitrary copula $C \in \mathcal{C}$ and fixing $B \in \mathcal{C}$ with $\text{supp}(\mu_B) = \mathbb{I}^2$, defining C_n by $C_n := (1 - \frac{1}{n})C + \frac{1}{n}B$, obviously $\text{supp}(\mu_{C_n}) = \mathbb{I}^2$ and $d_\infty(C, C_n) \xrightarrow{n \rightarrow \infty} 0$. This shows that the family of copulas with full support is dense in \mathcal{C} , implying that \mathcal{A}_N is nowhere dense for every N . Finally, using

$$\{C \in \mathcal{C} : \text{supp}(\mu_C) \neq \mathbb{I}^2\} \subseteq \bigcup_{N \in \mathbb{N}} \mathcal{A}_N$$

yields that $\{C \in \mathcal{C} : \text{supp}(\mu_C) = \mathbb{I}^2\}$ is co-meager in (\mathcal{C}, d_∞) . \square

We close this section with the following quite counter-intuitive result, which directly follows from combining Theorem 3.8 and Lemma 3.9 and using the fact that finite and countably infinite intersections of co-meager sets are co-meager as well.

Corollary 3.10. *A typical copula $C \in (\mathcal{C}, d_\infty)$ is mutually completely dependent and has full support.*

4. Extreme Value copulas

4.1. Some preliminaries

A copula $C \in \mathcal{C}$ is called Extreme Value copula (EVC) if there exists some copula $B \in \mathcal{C}$ such that

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

for all $x, y \in \mathbb{I}$. Throughout this section the space of all bivariate EVCs will be denoted by \mathcal{C}_{ev} . It is well-known (see [45]) that $(\mathcal{C}_{ev}, d_\infty)$ is a compact metric space. Moreover, according to [7, 14, 32, 35], the following assertions are equivalent

- (i) $C \in \mathcal{C}_{ev}$
- (ii) C is max-stable, i.e., $C(x, y) = C^k(x^{\frac{1}{k}}, y^{\frac{1}{k}})$ for all $k \in \mathbb{N}$ and $x, y \in \mathbb{I}$.
- (iii) There exists a Pickands dependence function A , i.e., a convex function $A: \mathbb{I} \rightarrow \mathbb{I}$ fulfilling $\max\{1-x, x\} \leq A(x) \leq 1$ for $x \in \mathbb{I}$, such that the identity

$$C(x, y) = (xy)^{A\left(\frac{\log(x)}{\log(xy)}\right)} \quad (4.1)$$

holds for all $x, y \in (0, 1)$.

The family of all Pickands dependence functions will be denoted by \mathcal{A} and we will let C_A denote the unique EVC induced by $A \in \mathcal{A}$. A Pickands dependence function $A \in \mathcal{A}$ is called symmetric, if $A(t) = A(1-t)$ holds for all $t \in \mathbb{I}$. The family of all symmetric Pickands dependence functions will be denoted by \mathcal{A}_s . For every Pickands dependence function $A \in \mathcal{A}$ we will let $D^+A(x)$ denote the right-hand derivative of A at $x \in [0, 1)$ and $D^-A(x)$ the left-hand derivative of A at $x \in (0, 1]$. Convexity of A implies that $D^+A(x) = D^-A(x)$ holds for all but at most countably many $x \in (0, 1)$, i.e., A is differentiable outside a countable subset of $(0, 1)$, that D^+A is non-decreasing and right-continuous on $[0, 1)$ and that D^-A is non-decreasing and left-continuous on $(0, 1]$ (see [20, 36] and the references therein). Setting $D^+A(1) := D^-A(1)$ allows to view D^+A as non-decreasing and right-continuous function on the full unit interval $[0, 1]$, which, taking into account $\max\{1-x, x\} \leq A(x) \leq 1$ for all $x \in [0, 1]$, only assumes values in $[-1, 1]$. Additionally (again see [20, 36] and the references therein), we have $D^-A(x) = D^+A(x-)$ for every $x \in (0, 1)$.

In the sequel $\Delta_2 = \{(x, 1-x) : x \in \mathbb{I}\}$ denotes unit simplex in \mathbb{R}^2 . Following [1, 14, 35] a finite measure ν on $\mathcal{B}(\Delta_2)$ is called a spectral measure, if

$$\int_{\Delta_2} w_i d\nu(w_1, w_2) = 1$$

holds for $i \in \{1, 2\}$. According to [35], every Pickands dependence function A can be identified with a unique spectral measure ν . More precisely, given a Pickands dependence function $A \in \mathcal{A}$, there exists a spectral measure ν on $\mathcal{B}(\Delta_2)$ such that (again see [35])

$$A(t) = \int_{\Delta_2} \max\{tw_1, (1-t)w_2\} d\nu(w_1, w_2) \quad (4.2)$$

holds for every $t \in \mathbb{I}$. Normalizing the measure ν and projecting it in x -direction onto the unit interval \mathbb{I} (or, equivalently, writing $\Delta_2 = \iota(I)$ with $\iota(s) = (1 - s, s)$ and setting $\vartheta = \nu^{\iota^{-1}}$) yields a probability measure ϑ on $\mathcal{B}(\mathbb{I})$ fulfilling $\int_{\mathbb{I}} s d\vartheta(s) = \frac{1}{2}$ and equation (4.2) boils down to

$$A(t) = 2 \int_{\mathbb{I}} \max\{t(1 - s), (1 - t)s\} d\vartheta(s) \quad (4.3)$$

for all $t \in \mathbb{I}$. Simplifying equation (4.3) yields the following identity for ϑ (see Lemma Appendix A.2 for a proof):

$$A(t) := 1 - t + 2 \int_{[0, t]} \vartheta([0, z]) d\lambda(z). \quad (4.4)$$

Throughout this section we call such a measure ϑ a Pickands dependence measure and define the family of all Pickands dependence measures by

$$\mathcal{P}_{\mathcal{A}} := \left\{ \vartheta \in \mathcal{P}(\mathbb{I}) : \int_{\mathbb{I}} x d\vartheta(x) = \frac{1}{2} \right\}.$$

We denote the map between the spaces $(\mathcal{P}_{\mathcal{A}}, \tau_w)$ and $(\mathcal{A}, \|\cdot\|_{\infty})$ defined via equation (4.4) by Υ and prove in Lemma 4.2 that Υ is indeed a homeomorphism. Throughout this section we denote the family of all absolutely continuous, discrete and singular Pickands dependence measures by $\mathcal{P}_{\mathcal{A}}^{abs}$, $\mathcal{P}_{\mathcal{A}}^{dis}$ and $\mathcal{P}_{\mathcal{A}}^{sing}$, respectively. Moreover, we will call a function $F: \mathbb{I} \rightarrow \mathbb{I}$ a distribution function/measure generating function, if its extension via setting $F(x) = 0$ for $x < 0$ and $F(x) = 1$ for $x > 1$, is a distribution function/measure generating function. It is well known and straightforward to verify that the Pickands dependence function corresponding to the Fréchet-Hoeffding upper bound M is defined by $A_M(t) := \max\{1 - t, t\}$ and the Pickands dependence function corresponding to the independence copula Π fulfills $A_{\Pi}(t) = 1$ for all $t \in \mathbb{I}$. Moreover, the measure $\vartheta_M = \delta_{\frac{1}{2}}$ is easily seen to correspond to M and $\vartheta = \frac{1}{2}(\delta_0 + \delta_1)$ to Π .

Working with equation (4.4) it is straightforward to prove that the right-hand derivative D^+A of a Pickands dependence function $A \in \mathcal{A}$ can easily be expressed in terms of the corresponding Pickands dependence measure $\vartheta \in \mathcal{P}_{\mathcal{A}}$ - the following result holds:

Lemma 4.1. *Let $\vartheta \in \mathcal{P}_{\mathcal{A}}$ and $A \in \mathcal{A}$ denote the corresponding Pickands dependence function according to equation (4.4). Then*

$$D^+A(t) = 2\vartheta([0, t]) - 1 \text{ and} \quad (4.5)$$

$$\vartheta(\{t\}) = D^+A(t) - D^+A(t-) = D^+A(t) - D^-A(t) \quad (4.6)$$

holds for every $t \in [0, 1)$.

Proof. Defining $F_{\vartheta}(t) := \vartheta([0, t])$ for $t \in \mathbb{I}$ and fixing an arbitrary continuity point t of F_{ϑ} , applying equation (4.4) and Lebesgue's differentiation theorem (see [38]) we obtain that

$$D^+A(t) = \lim_{h \downarrow 0} 2 \int_{[t, t+h]} \vartheta([0, z]) d\lambda(z) - 1 = 2\vartheta([0, t]) - 1.$$

Using the fact that the set $\text{Cont}(F_{\vartheta})$ of all continuity points of F_{ϑ} is dense in \mathbb{I} and considering that both D^+A and F_{ϑ} are right-continuous yields the result. Equation (4.6) is an immediate consequence of equation (4.5). \square

Building upon the afore-mentioned characterization of EVC via Pickands dependence functions, the map Φ , defined via equation (4.1) maps $(\mathcal{A}, \|\cdot\|_\infty)$ to $(\mathcal{C}_{ev}, d_\infty)$. It is straightforward to see that both Υ and Φ are homeomorphisms.

Lemma 4.2. *The maps Φ , Υ and $\Phi \circ \Upsilon$ are homeomorphisms.*

Proof. Obviously, every $\vartheta \in \mathcal{P}_{\mathcal{A}}$ induces a unique $A \in \mathcal{A}$ via equation (4.4). Lemma Appendix A.2 directly yields surjectivity of Υ . Concerning injectivity of Υ , suppose that $\vartheta_1, \vartheta_2 \in \mathcal{P}_{\mathcal{A}}$ fulfill $\Upsilon(\vartheta_1) = \Upsilon(\vartheta_2) =: A$. Applying Lemma 4.1 yields

$$\vartheta_1([0, t]) = \frac{D^+ A_1(t) + 1}{2} = \frac{D^+ A_2(t) + 1}{2} = \vartheta_2([0, t])$$

for every $t \in [0, 1)$ and $\vartheta_1 = \vartheta_2$ follows. Thus, Υ is a bijection and applying Lemma Appendix A.3 immediately yields that Υ is a homeomorphism. The fact that Φ is a homeomorphism is an immediate consequence of Lemma Appendix A.3. Since compositions of homeomorphisms are homeomorphisms the proof is complete. \square

Remark 4.3. *Applying [22, Theorem 5.1], Φ is even a homeomorphism, if we equip \mathcal{C}_{ev} with the metric D_1 (or any of the metrics D_p).*

Given $A \in \mathcal{A}$ define the map $G_A: \mathbb{I} \rightarrow \mathbb{I}$ by

$$G_A(t) := A(t) + D^+ A(t)(1 - t) \quad (4.7)$$

$t \in [0, 1)$ as well as $G_A(1) := 1$. Then applying [45, Lemma 5] G_A is non-negative, right-continuous, and non-decreasing. Hence working with G_A and considering conditional distributions of an EVC $C_A \in \mathcal{C}_{ev}$ with corresponding Pickands dependence function $A \in \mathcal{A}$, according to [45] a version of the Markov-kernel of C_A is given by

$$K_A(x, [0, y]) := \begin{cases} 1, & \text{if } x \in \{0, 1\} \\ \frac{C_A(x, y)}{x} G_A\left(\frac{\log(x)}{\log(xy)}\right), & \text{if } x, y \in (0, 1) \\ y, & \text{if } (x, y) \in (0, 1) \times \{0, 1\} \end{cases} \quad (4.8)$$

for $x, y \in \mathbb{I}$. Again following [45], define $f^t(x) := x^{\frac{1}{t}-1}$ for $t \in (0, 1)$ and $x \in \mathbb{I}$, and, for a given Pickands dependence function $A \in \mathcal{A}$, set

$$L := \max\{x \in \mathbb{I}: A(x) = 1 - x\}, \quad R := \min\{x \in \mathbb{I}: A(x) = x\} \quad (4.9)$$

According to [45] the discrete component of C_A is non-degenerated if, and only if there exists some $t \in (0, 1)$ such that $\mu_{C_A}(\Gamma(f^t)) > 0$. Furthermore the discrete component is fully determined by the discontinuity points of the right-hand derivative $D^+ A$. The following Theorem shows that $\Gamma(f^t)$ with $t \in (0, 1)$ carries mass if, and only if, t is a point mass of ϑ .

Theorem 4.4. *Let $C \in \mathcal{C}_{ev}$ be an EVC with associated Pickands dependence measure $\vartheta \in \mathcal{P}_{\mathcal{A}}$. Then the following identity holds for all $t \in (0, 1)$:*

$$\mu_C(\Gamma(f^t)) = \frac{2t(1-t)}{A(t)} \vartheta(\{t\}). \quad (4.10)$$

Consequently, ϑ has a point mass in $(0, 1)$ if, and only if $\mu_C^{dis}(\mathbb{I}^2) > 0$.

Proof. Applying Lemma 4.1 it follows that ϑ has a point mass $t \in (0, 1)$ if, and only if $t \in (0, 1)$ is a point of discontinuity of D^+A . The assertion now follows by applying [45, Lemma 4]. Equation (4.10) is an immediate consequence of [45, Lemma 4] and Lemma 4.1. \square

4.2. Mass distributions of Extreme Value copulas

In what follows we characterize the existence of a non-degenerated absolutely continuous, discrete, or singular component of an EVC in terms of the corresponding Pickands dependence measures. We do not consider purely discrete or singular Pickands dependence measures since, as proved already in [45, Corollary 5], every EVC (except from M) has a non-degenerated absolutely continuous component and thus, purely discrete and singular EVC do not exist. We start with the following example illustrating why working with $(0, 1)$ instead of \mathbb{I} can't be avoided.

Example 4.5. Consider the Pickands dependence measure $\vartheta_1 = \frac{1}{2}(\delta_0 + \delta_1)$. As mentioned before, ϑ_1 induces the product copula Π , i.e., even though the EVC copula is absolutely continuous, the Pickands dependence measure is not. It is straightforward to see that Π is not the only absolutely continuous EVC whose associated Pickands dependence measure has a non-degenerated discrete component. In fact, letting ϑ_2 be an arbitrary absolutely continuous measure in $\mathcal{P}_{\mathcal{A}}$ and defining $\vartheta := \frac{\vartheta_1 + \vartheta_2}{2} \in \mathcal{P}_{\mathcal{A}}$, then ϑ is a probability measure with non-degenerated discrete component which induces an absolutely continuous EVC.

We now characterize the existence of non-degenerated discrete/singular components of C_A in terms of the Lebesgue decomposition of the Pickands dependence measure ϑ (the quite technical proof as well as some preliminary lemmas can be found in the Appendix).

Theorem 4.6. *Let $C \in \mathcal{C}_{ev}$ be an EVC and let $\vartheta \in \mathcal{P}_{\mathcal{A}}$ and $A \in \mathcal{A}$ denote the corresponding Pickands dependence measure and Pickands dependence function, respectively. Then the following equivalences hold:*

- (i) ϑ is absolutely continuous on $(0, 1)$ if, and only if C is absolutely continuous.
- (ii) ϑ has a point mass in $(0, 1)$ if, and only if C has non-degenerated discrete component.
- (iii) ϑ has a non-degenerated singular component if, and only if C has non-degenerated singular component.

The following example illustrates the previous result and considers a Pickands dependence measure with non-degenerated discrete and absolutely continuous component.

Example 4.7. Define the distribution function F_{ϑ} (see Figure 2) by

$$F_{\vartheta}(t) := \begin{cases} t, & \text{if } t \in [0, \frac{1}{2}) \\ \frac{3}{5}, & \text{if } t \in [\frac{1}{2}, \frac{3}{4}) \\ \frac{16}{35}t + \frac{1}{2}, & \text{if } t \in [\frac{3}{4}, 1) \\ 1, & \text{if } t = 1 \end{cases}$$

for $t \in \mathbb{I}$. It is straightforward to verify that F_ϑ corresponds to a unique Pickands dependence measure $\vartheta \in \mathcal{P}_{\mathcal{A}}$, which has both a non-degenerated discrete and absolutely continuous component. According to equation (4.4) the associated Pickands dependence function A is given by (again see Figure 2)

$$A(t) = \begin{cases} t^2 - t + 1, & \text{if } t \in [0, \frac{1}{2}) \\ \frac{1}{5}t + \frac{13}{20}, & \text{if } t \in [\frac{1}{2}, \frac{3}{4}) \\ \frac{16}{35}t^2 + \frac{19}{35}, & \text{if } t \in [\frac{3}{4}, 1]. \end{cases}$$

Applying Theorem 4.6, both the discrete and absolutely continuous component of ϑ propagate to the induced copula C_A . The discrete component is concentrated on the graphs of the functions $f^{\frac{1}{2}}(x) = x$ and $f^{\frac{3}{4}}(x) = \sqrt[3]{x}$, respectively. A sample of the copula C_A is depicted in Figure 3.

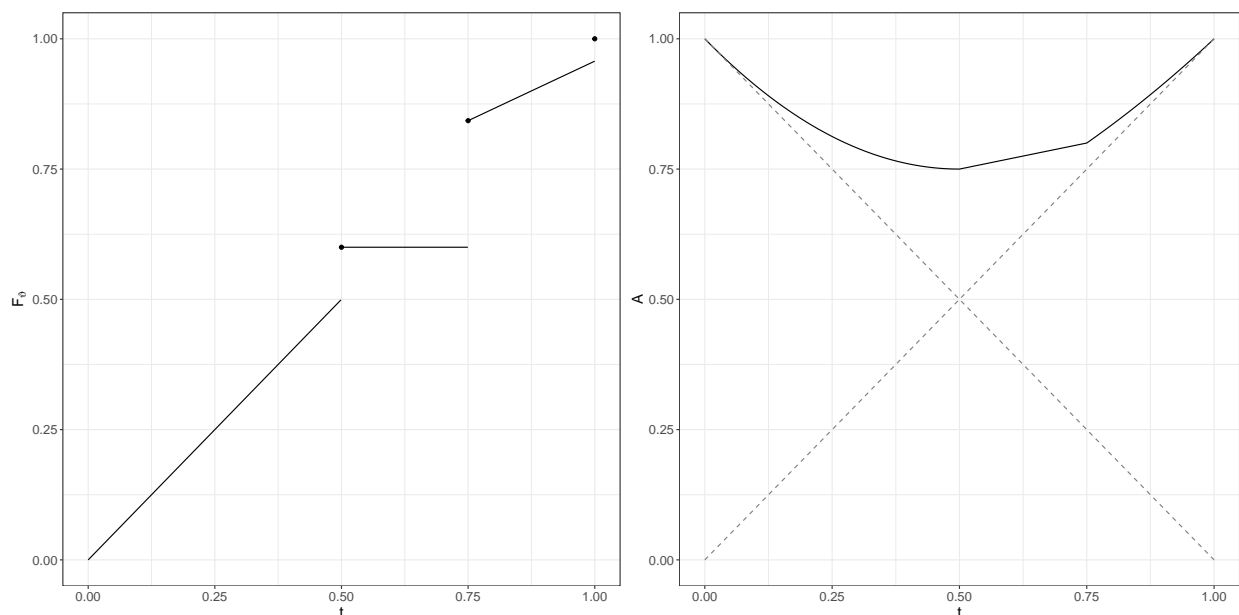


Figure 2: Graphs of the distribution function F_ϑ of the Pickands dependence measure ϑ (left) according to Example 4.7 and associated Pickands dependence function A (right) associated with it.

We have already shown that many regularity/singularity properties of the Pickands dependence measure carry over to the corresponding EVC. As next step we show that the support is no exception:

Lemma 4.8. *Let $C \in \mathcal{C}_{ev}$ and $\vartheta \in \mathcal{P}_{\mathcal{A}}$ be its associated Pickands dependence measure. Then the following assertions hold:*

- (i) *If ϑ^{dis} has full support, then $\text{supp}(\mu_C^{dis}) = \mathbb{I}^2$.*
- (ii) *If ϑ^{sing} has full support, then $\text{supp}(\mu_C^{sing}) = \mathbb{I}^2$.*
- (iii) *If ϑ fulfills $L = 0$ and $R = 1$, then $\text{supp}(\mu_C^{abs}) = \mathbb{I}^2$.*

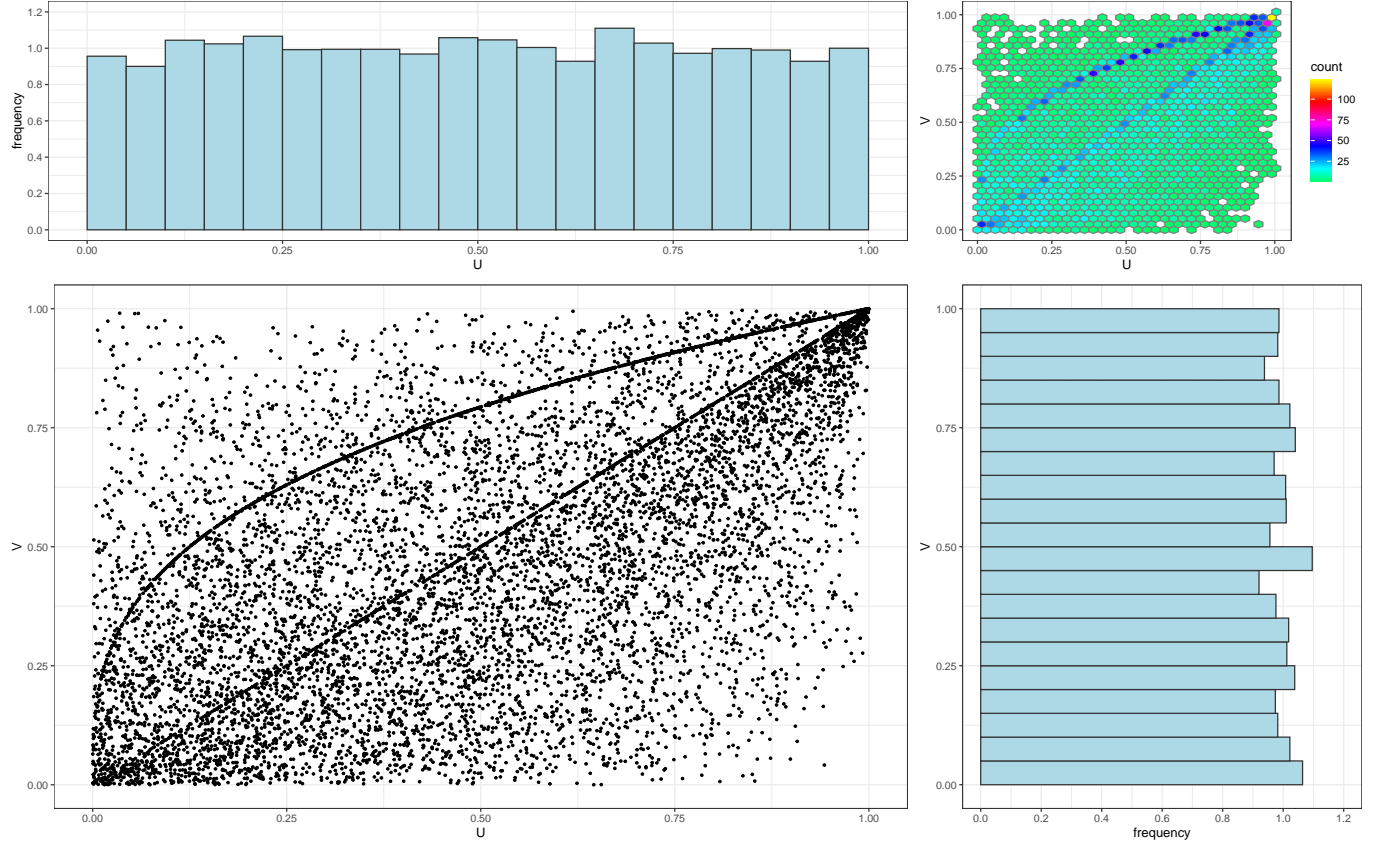


Figure 3: Sample of size 10000 of the EVC C_A , where A is the Pickands dependence function according to Example 4.7, its histogram and the two marginal histograms. The sample has been generated via conditional inverse sampling.

Proof. Suppose that $\vartheta \in \mathcal{P}_{\mathcal{A}}$ is such that ϑ^{dis} has full support. Then there exist countably infinite sets $\{q_1, q_2, \dots\} \subseteq \mathbb{I}$ and $\{\alpha_1, \alpha_2, \dots\} \subseteq (0, 1)$ with $\sum_{i \in \mathbb{N}} \alpha_i \leq 1$ such that $\vartheta^{dis} = \sum_{i \in \mathbb{N}} \alpha_i \delta_{q_i}$. Fixing $x \in (0, 1)$, applying Theorem 4.4 yields $K_A(x, \{f^t(x)\}) > 0$ if, and only if $t = q_i$ for some $i \in \mathbb{N}$. Having this it follows immediately that $\text{supp}(K_C^{dis}(x, \cdot)) = \mathbb{I}$. Since $x \in (0, 1)$ was arbitrary, applying disintegration yields $\text{supp}(\mu_C^{dis}) = \mathbb{I}^2$.

To prove the second assertion we may proceed as follows: Considering $\vartheta \in \mathcal{P}_{\mathcal{A}}$ fulfilling that its singular component ϑ^{sing} has full support, then the function $y \mapsto G_A^{sing}\left(\frac{\log(x)}{\log(xy)}\right)$ is continuous, strictly increasing, and has derivative 0 almost everywhere, whereby $G_A^{sing}(t) = D^+ A_{sing}(t)(1-t) + A_{sing}(t)$, and A_{sing} is according to equation (A.1) with the measure ϑ^{sing} being used instead of ϑ . On the other hand, the function $y \mapsto \frac{C(x,y)}{x}$ is absolutely continuous and strictly increasing so, applying the same arguments as in Lemma Appendix A.4, it follows that $K_C^{sign}(x, \cdot)$ has support \mathbb{I} and thus, disintegration yields the desired statement. The last assertion is a direct consequence of [45, Corollary 5]. \square

Finally, we prove that the family of EVCs $C \in \mathcal{C}_{ev}$ with $\text{supp}(\mu_C^{dis}) = \mathbb{I}^2$, the family fulfilling $\text{supp}(\mu_C^{sing}) = \mathbb{I}^2$ and the family with $\text{supp}(\mu_C^{abs}) = \mathbb{I}^2$ is dense in \mathcal{C}_{ev} .

Theorem 4.9. *The following assertions hold:*

(i) The family $\{C \in \mathcal{C}_{ev} : \text{supp}(\mu_C^{dis}) = \mathbb{I}^2\}$ is dense in $(\mathcal{C}_{ev}, d_\infty)$.

(ii) The family $\{C \in \mathcal{C}_{ev} : \text{supp}(\mu_C^{sing}) = \mathbb{I}^2\}$ is dense in $(\mathcal{C}_{ev}, d_\infty)$.

(iii) The family $\{C \in \mathcal{C}_{ev} : \text{supp}(\mu_C^{abs}) = \mathbb{I}^2\}$ is dense in $(\mathcal{C}_{ev}, d_\infty)$.

Proof. Combining Lemma 4.8, Lemma Appendix A.3 and Lemma Appendix A.8 yields the first two assertions. The last assertion follows by combining assertion (ii) and the fact that $\text{supp}(\mu_C) = \text{supp}(\mu_C^{abs})$ holds for every $C \in \mathcal{C}_{ev}$ (see [45, Corollary 5]). \square

Working with convex-combinations of Pickands dependence measures we finally obtain the following striking result:

Theorem 4.10. *The family $\{C \in \mathcal{C}_{ev} : \text{supp}(\mu_C^{dis}) = \text{supp}(\mu_C^{sing}) = \text{supp}(\mu_C^{abs}) = \mathbb{I}^2\}$ is dense in $(\mathcal{C}_{ev}, d_\infty)$.*

The afore-mentioned results remain valid when working with stronger notions of convergence - using Theorem 4.9 and Theorem 4.10 in combination with [22, Theorem 5.1] yields the following two corollaries:

Corollary 4.11. *The following assertions hold:*

(i) The family $\{C \in \mathcal{C}_{ev} : \text{supp}(\mu_C^{dis}) = \mathbb{I}^2\}$ is dense in (\mathcal{C}_{ev}, D_1) and in \mathcal{C}_{ev} with respect to wcc.

(ii) The family $\{C \in \mathcal{C}_{ev} : \text{supp}(\mu_C^{sing}) = \mathbb{I}^2\}$ is dense in (\mathcal{C}_{ev}, D_1) and in \mathcal{C}_{ev} with respect to wcc.

(iii) The family $\{C \in \mathcal{C}_{ev} : \text{supp}(\mu_C^{abs}) = \mathbb{I}^2\}$ is dense in (\mathcal{C}_{ev}, D_1) and in \mathcal{C}_{ev} with respect to wcc.

Corollary 4.12. *The family $\{C \in \mathcal{C}_{ev} : \text{supp}(\mu_C^{dis}) = \text{supp}(\mu_C^{sing}) = \text{supp}(\mu_C^{abs}) = \mathbb{I}^2\}$ is dense in (\mathcal{C}_{ev}, D_1) and in \mathcal{C}_{ev} with respect to wcc.*

4.3. Derivatives of Extreme Value copulas

We now revisit the results from Section 3 and show that EVC typically exhibit more regular behavior. To simplify notation we denote the family of all EVC $C \in \mathcal{C}_{ev}$ with the property that there exists some set $\Lambda \in \mathcal{B}(\mathbb{I})$ with $\lambda(\Lambda) = 1$ such that for all $x \in \Lambda$ there exists some $y = y_x \in (0, 1)$ such that $\partial_1 C(x, y)$ does not exist by $\mathcal{C}_{ev,p} := \mathcal{C}_p \cap \mathcal{C}_{ev}$. The more pathological family $\mathcal{C}_{ev,\mathcal{Q}}$ is defined in the same manner, i.e., $\mathcal{C}_{ev,\mathcal{Q}} := \mathcal{C}_{\mathcal{Q}} \cap \mathcal{C}_{ev}$. In other words: For every $C \in \mathcal{C}_{ev,\mathcal{Q}}$ there exists some set $\Lambda \in \mathcal{B}(\mathbb{I})$ with $\lambda(\Lambda) = 1$ such that for every $x \in \Lambda$ the partial derivative $\partial_1 C(x, y)$ does not exist on a dense set of $y \in \mathbb{I}$.

We start by showing that $\mathcal{C}_{ev,\mathcal{Q}}$ is non-empty.

Example 4.13 (Example of an element of $\mathcal{C}_{ev,\mathcal{Q}}$). Suppose that $Q = \{q_1, q_2, \dots\} \subseteq (0, 1)$ is dense in \mathbb{I} . Furthermore suppose that $\{\alpha_1, \alpha_2, \dots\} \subset (0, 1)$ fulfills $\sum_{i \in \mathbb{N}} \alpha_i = 1$. Then defining $\vartheta \in \mathcal{P}(\mathbb{I})$ by

$$\tilde{\vartheta} = \sum_{i=1}^{\infty} \alpha_i \delta_{q_i}$$

and, if necessary, normalizing $\tilde{\vartheta}$ in the sense of Lemma Appendix A.7 yields a discrete measure $\vartheta \in \mathcal{P}_{\mathcal{A}}$ with full support \mathbb{I} . For the sake of simplicity we will assume that $\vartheta = \tilde{\vartheta}$ holds. Let $A = A_{\vartheta} \in \mathcal{A}$ denote the Pickands dependence function induced by ϑ and C_A the corresponding EVC. We want to show that for every $x_0 \in (0, 1)$ we have that $x \mapsto C_A(x, f^{q_j}(x_0))$ is not differentiable at x_0 and proceed as follows: Let $x_0 \in (0, 1)$ be arbitrary but fixed, set $y_j := x_0^{\frac{1}{q_j}-1}$ with $q_j \in Q$ and, using equation (4.5), consider the function

$$x \mapsto D^+ A \left(\frac{\log(x)}{\log(xy_j)} \right) = 2 \left(\underbrace{\sum_{i \neq j} \alpha_i \mathbf{1}_{[0, \frac{\log(x)}{\log(xy_j)}]}(q_i)}_{=: I(x)} + \underbrace{\alpha_j \mathbf{1}_{[0, \frac{\log(x)}{\log(xy_j)}]}(q_j)}_{=: II(x)} \right) - 1.$$

Applying the fact that $\frac{\log(x_0)}{\log(x_0 y_j)} = q_j \neq q_i$ for $i \neq j$ yields that $I(x)$ is continuous in x_0 . Defining the functions

$$\xi(s, y_j) := K_A(s, [0, y_j]) - 2C_A(s, y_j) \alpha_j \mathbf{1}_{[0, \frac{\log(s)}{\log(sy_j)}]}(q_j) \frac{\log(y_j)}{s \log(sy_j)},$$

$$\varphi(x, y_j) := 2 \int_{[0, x]} C_A(s, y_j) \alpha_j \mathbf{1}_{[0, \frac{\log(s)}{\log(sy_j)}]}(q_j) \frac{\log(y_j)}{s \log(sy_j)} d\lambda(s),$$

$$\psi(x, y_j) := \int_{[0, x]} \xi(s, y_j) d\lambda(s),$$

for every $s, x \in (0, 1)$ and applying disintegration we can write the copula C_A as

$$C_A(x, y_j) = \varphi(x, y_j) + \psi(x, y_j)$$

for every $x \in (0, 1)$. Furthermore, using that x_0 is a point of continuity of the function $x \mapsto \xi(x, y_j)$ implies that the function $x \mapsto \psi(x, y_j)$ is differentiable at x_0 . Moreover, working with the right- and left-hand derivatives of φ , respectively, we obtain that

$$\partial_1^+ \varphi(x_0, y_j) = 2 \lim_{h \downarrow 0} \frac{1}{h} \int_{[x_0, x_0+h]} C_A(s, y_j) \alpha_j \mathbf{1}_{[0, \frac{\log(s)}{\log(sy_j)}]}(q_j) \frac{\log(y_j)}{s \log(sy_j)} d\lambda(s) = 0$$

as well as

$$\begin{aligned} \partial_1^- \varphi(x_0, y_j) &= \lim_{h \downarrow 0} \frac{1}{h} \int_{[x_0-h, x_0]} C_A(s, y_j) \alpha_j \mathbf{1}_{[0, \frac{\log(s)}{\log(sy_j)}]}(q_j) \frac{\log(y_j)}{s \log(sy_j)} d\lambda(s) \\ &= 2\alpha_j C_A(x_0, y_j) \frac{\log(y_j)}{x_0 \log(x_0 y_j)}, \end{aligned}$$

which altogether shows that $x \mapsto C_A(\cdot, y_j)$ is not differentiable in x_0 . Since y_j was arbitrary and the set $\{x_0^{1/q_j-1} : j \in \mathbb{N}\}$ is obviously dense in \mathbb{I} this completes the proof.

Similarly to Theorem 3.3, EVCs exhibiting pathological behavior are well spread - the following corollary holds.

Corollary 4.14. *The set $\mathcal{C}_{ev,\mathcal{Q}}$ is dense in $(\mathcal{C}_{ev}, d_\infty)$.*

Proof. Consider an arbitrary EVC C_A and let $\vartheta \in \mathcal{P}_{\mathcal{A}}$ denote its corresponding Pickands dependence measure. Then, using Lemma Appendix A.8 there exists a sequence $(\vartheta_n)_{n \in \mathbb{N}}$ in $\mathcal{P}_{\mathcal{A}}$ consisting of discrete measures with full support that converges weakly to ϑ . Applying the fact that (analogously to Example 4.13) the measures ϑ_n induce EVCs C_n which are elements of $\mathcal{C}_{ev,\mathcal{Q}}$, applying Lemma Appendix A.3 yields the desired result. \square

The following simple observation will be key for studying Baire category results for EVC in the next section.

Theorem 4.15. *Let $C \in \mathcal{C}_{ev}$. Then $\mu_C^{dis}(\mathbb{I}^2) > 0$ if, and only if $C \in \mathcal{C}_{ev,p}$.*

Proof. In case of $\mu_C^{dis}(\mathbb{I}^2) > 0$ applying Theorem 4.6 yields that $\vartheta \in \mathcal{P}_{\mathcal{A}}$ has a point mass in some point $t_0 \in (0, 1)$. For fixed $x \in (0, 1)$, setting $y_x = f^{t_0}(x)$ and proceeding analogously to Example 4.13 shows that $\partial_1^+ C(x, y_x) \neq \partial_1^- C(x, y_x)$.

For the other direction suppose that $C \in \mathcal{C}_{ev,p}$. Then there exists some set $\Lambda \in \mathcal{B}(\mathbb{I})$ with $\lambda(\Lambda) = 1$ such that for $x \in \Lambda$ we can find some $y_x \in (0, 1)$ with $\partial_1^+ C(x, y_x) \neq \partial_1^- C(x, y_x)$. It follows that $D^+ A(\frac{\log(x)}{\log(xy_x)}) \neq D^- A(\frac{\log(x)}{\log(xy_x)})$, so $\frac{\log(x)}{\log(xy_x)}$ is a point of discontinuity of $D^+ A$, and applying Theorem 4.4 yields $\mu_C^{dis}(\mathbb{I}^2) > 0$. \square

4.4. Baire category results for Extreme Value copulas

Throughout this section we again work with the afore-mentioned one-to-one-to-one correspondence between Pickands dependence functions $A \in \mathcal{A}$, Pickands dependence measures $\vartheta \in \mathcal{P}_{\mathcal{A}}$, and EVC $C \in \mathcal{C}_{ev}$. Rewriting equation (4.9) in terms of the measure $\vartheta \in \mathcal{P}_{\mathcal{A}}$ obviously yields

$$L_\vartheta = \sup\{x \in \mathbb{I} : \vartheta([0, x]) = 0\}, \quad R_\vartheta = \inf\{x \in \mathbb{I} : \vartheta([0, x]) = 1\}, \quad (4.11)$$

with the convention $\sup \emptyset := 0$ and $\inf \emptyset := 1$. According to [45] the support of an EVC $C \in \mathcal{C}_{ev}$ is fully determined by the functions f^{L_ϑ} and f^{R_ϑ} , respectively, i.e.,

$$\text{supp}(\mu_C) = \{(x, y) \in \mathbb{I}^2 : f^{L_\vartheta}(x) \leq y \leq f^{R_\vartheta}(x)\}.$$

In other words: An EVC $C \in \mathcal{C}_{ev}$ has full support if, and only if $L_\vartheta = 0$ and $R_\vartheta = 1$. We will show now that a topologically typical Pickands dependence measure has this property:

Lemma 4.16. *The set*

$$\{\vartheta \in \mathcal{P}_{\mathcal{A}} : L_\vartheta = 0 \text{ and } R_\vartheta = 1\}$$

is co-meager in $\mathcal{P}_{\mathcal{A}}$ with respect to the weak topology.

Proof. It suffices to show that the set $\{\vartheta \in \mathcal{P}_{\mathcal{A}} : L_\vartheta = 0 \text{ and } R_\vartheta = 1\}^c$ is of first category, which can be done as follows. For every $n \in \mathbb{N}$ define the sets A_n, B_n by

$$A_n := \{\vartheta \in \mathcal{P}_{\mathcal{A}} : \vartheta((0, \frac{1}{n})) = 0\} \text{ and } B_n := \{\vartheta \in \mathcal{P}_{\mathcal{A}} : \vartheta([0, 1 - \frac{1}{n}]) = 1\}.$$

Suppose that $\vartheta, \vartheta_1, \vartheta_2, \dots \in A_n$ are such that $(\vartheta_\ell)_{\ell \in \mathbb{N}}$ converges weakly to ϑ . Then applying Portmanteau's theorem yields $\vartheta \in A_n$, so the set A_n is closed with respect to the weak topology. Since an analogous argument shows that each set B_n is closed it follows that $A_n \cup B_n$ is closed. Since measures with strictly increasing distribution functions are dense in $\mathcal{P}_{\mathcal{A}}$ (see the proof of Corollary 4.14), the sets $A_n \cup B_n$ are nowhere dense in $\mathcal{P}_{\mathcal{A}}$. Considering that $\{\vartheta \in \mathcal{P}_{\mathcal{A}} : L_\vartheta > 0 \text{ or } R_\vartheta < 1\} \subseteq \bigcup_{n \in \mathbb{N}} (A_n \cup B_n)$, the desired result follows. \square

Applying Lemma 4.2 together with the previous lemma yields the following corollary, stating that typical EVCs have full support.

Corollary 4.17. *The set*

$$\{C \in \mathcal{C}_{ev} : \text{supp}(\mu_C) = \mathbb{I}^2\}$$

is co-meager in $(\mathcal{C}_{ev}, d_\infty)$. In other words: Topologically typical EVCs have full support.

In analogy with the space of all bivariate copulas \mathcal{C} (see [10]), typical EVCs are not absolutely continuous. Our proof builds upon Lemma 4.2.

Lemma 4.18. *The set of all absolutely continuous Pickands dependence measures $\mathcal{P}_{\mathcal{A}}^{abs}$ is of first Baire category in $\mathcal{P}_{\mathcal{A}}$ with respect to the weak topology.*

Proof. For every $\vartheta \in \mathcal{P}_{\mathcal{A}}^{abs}$ let k_ϑ denote its density and, for every $n \in \mathbb{N}$ define the set G_n^ϑ by

$$G_n^\vartheta := \{x \in \mathbb{I} : k_\vartheta(x) > n\}.$$

Furthermore set

$$\mathcal{W}_n := \left\{ \vartheta \in \mathcal{P}_{\mathcal{A}}^{abs} : \vartheta(G_n^\vartheta) \leq \frac{1}{4} \right\}.$$

Considering $\lambda(\bigcap_{n=1}^\infty G_n^\vartheta) = 0$, absolute continuity implies $\vartheta(\bigcap_{n=1}^\infty G_n^\vartheta) = 0$ so, using continuity from above we get that $\vartheta(G_n^\vartheta) < \frac{1}{4}$ holds for all n sufficiently large. This implies

$$\mathcal{P}_{\mathcal{A}}^{abs} \subseteq \bigcup_{n \in \mathbb{N}} \mathcal{W}_n.$$

To complete the proof it suffices to show that \mathcal{W}_n is nowhere dense in $(\mathcal{P}_{\mathcal{A}}, \tau_w)$, which can be done as follows: Consider an arbitrary discrete Pickands dependence measure $\beta \in \mathcal{P}_{\mathcal{A}}$ with only finitely many point masses, i.e., $\beta = \sum_{i=1}^N \alpha_i \delta_{x_i}$ with $x_i \in [0, 1]$, $\alpha_i \in (0, 1]$, $2 \leq N \in \mathbb{N}$ and $\sum_{i=1}^N \alpha_i = 1$. We assume that $0 \leq x_1 < \dots < x_N \leq 1$ and consider the following four cases. (i) If $x_i \in (0, 1)$ for all $1 \leq i \leq N$ set $\rho := \min\{x_1, x_2 - x_1, x_3 - x_2, \dots, x_N - x_{N-1}, 1 - x_N\}$; for (ii) $x_1 = 0$ and $x_N < 1$ or (iii) $x_1 > 0$ and $x_N = 1$ we define $\rho := \min\{x_2 - x_1, x_3 - x_2, \dots, 1 - x_N\}$ and $\rho := \min\{x_1, x_2 - x_1, \dots, x_N - x_{N-1}\}$, respectively. (iv) Finally, if $x_1 = 0$ and $x_N = 1$ we consider $\rho := \min\{x_2 - x_1, \dots, x_N - x_{N-1}\}$. In either case, setting

$$r := \frac{\rho}{8nN},$$

we obviously have $\lambda(\bigcup_{i=1}^N (x_i - r, x_i + r) \cap \mathbb{I}) \leq \frac{1}{4n}$.

Letting $f : \mathbb{I} \rightarrow [0, \infty)$ denote a continuous function fulfilling $f(x_i) = 1, f|_{(x_i - r, x_i + r) \cap \mathbb{I}} \in (0, 1]$

and $f = 0$ in $\mathbb{I} \setminus \bigcup_{i=1}^N (x_i - r, x_i + r)$, then $\int_{\mathbb{I}} f d\beta = 1$ by construction. On the other hand, for arbitrary $\vartheta \in \mathcal{W}_n$ we have that

$$\int_{\mathbb{I}} f d\vartheta = \int_{G_n^{\vartheta}} f d\vartheta + \int_{\mathbb{I} \setminus G_n^{\vartheta}} f d\vartheta \leq \frac{1}{4} + \int_{\mathbb{I} \setminus G_n^{\vartheta}} f d\vartheta \leq \frac{1}{2} < 1 = \int_{\mathbb{I}} f d\beta = 1.$$

Using the fact that (as direct consequence of Lemma Appendix A.7 in combination with Glivenko-Cantelli's theorem) discrete Pickands dependence measures with finitely many point masses are dense in $(\mathcal{P}_{\mathcal{A}}, \tau_w)$ yields that \mathcal{W}_n is nowhere dense in $\mathcal{P}_{\mathcal{A}}$, which completes the proof. \square

As direct consequence we get the following result on the family of absolutely continuous EVCs.

Theorem 4.19. *The family $\mathcal{C}_{ev,abs}$ is of first category in $(\mathcal{C}_{ev}, d_{\infty})$.*

We now return to differentiability and show that typical Pickands dependence functions are everywhere differentiable.

Lemma 4.20. *The set*

$$\{A \in \mathcal{A} : A \text{ is differentiable at every } x \in (0, 1)\}$$

is co-meager in \mathcal{A} with respect to the uniform distance $\|\cdot\|_{\infty}$ on \mathcal{A} .

Proof. We prove that the set

$$\hat{\mathcal{A}} := \{A \in \mathcal{A} : \exists x \in (0, 1) \text{ such that } D^+A(x) > D^-A(x)\}$$

is of first Baire category with respect to the topology induced by $\|\cdot\|_{\infty}$. For arbitrary $k \in \mathbb{N}$ and arbitrary $n \in \mathbb{N}$ we define the set $\mathcal{A}_{k,n}$ by

$$\mathcal{A}_{k,n} := \left\{ A \in \mathcal{A} : \exists x \in \left[\frac{1}{n}, 1 - \frac{1}{n} \right] \text{ such that } (D^+A(x) - D^-A(x)) \geq \frac{1}{k} \right\}.$$

Then, obviously $\hat{\mathcal{A}} \subseteq \bigcup_{k \in \mathbb{N}} \bigcup_{n \in \mathbb{N}_{\geq 3}} \mathcal{A}_{k,n}$. We show that the sets $\mathcal{A}_{k,n}$ are closed with respect to $\|\cdot\|_{\infty}$ and proceed as follows. Consider a sequence $(A_j)_{j \in \mathbb{N}}$ in $\mathcal{A}_{k,n}$ converging uniformly to some $A \in \mathcal{A}$. For each $j \in \mathbb{N}$ there exists some $x_j \in [\frac{1}{n}, 1 - \frac{1}{n}]$ such that $D^+A_j(x_j) - D^-A_j(x_j) \geq \frac{1}{k}$. Since $[\frac{1}{n}, 1 - \frac{1}{n}]$ is compact, there exists a subsequence $(x_{j_\ell})_{\ell \in \mathbb{N}}$ and some x^* such that $x_{j_\ell} \xrightarrow{\ell \rightarrow \infty} x^* \in [\frac{1}{n}, 1 - \frac{1}{n}]$. We can fix an arbitrary small $\Delta > 0$ fulfilling that $x^* + \Delta$ is a point of continuity of D^+A and $x^* - \Delta$ is a point of continuity of D^-A . Applying [22, Theorem 5.1] we therefore obtain that

$$\lim_{\ell \rightarrow \infty} D^+A_{j_\ell}(x^* + \Delta) = D^+A(x^* + \Delta), \quad \lim_{\ell \rightarrow \infty} D^-A_{j_\ell}(x^* - \Delta) = D^-A(x^* - \Delta).$$

Since (x_{j_ℓ}) converges to x^* there exists some $\ell_0 \in \mathbb{N}$ such that for all $\ell \geq \ell_0$ we have $|x_{j_\ell} - x^*| \leq \Delta$. Using convexity of A_{j_ℓ} yields

$$D^+A_{j_\ell}(x^* + \Delta) - D^-A_{j_\ell}(x^* - \Delta) \geq \frac{1}{k},$$

hence, considering the limit $\ell \rightarrow \infty$ yields

$$D^+A(x^* + \Delta) - D^-A(x^* - \Delta) \geq \frac{1}{k}.$$

Considering the facts that the set of continuity points of D^+A and D^-A are dense in \mathbb{I} and that D^+A and D^-A are right- and left-continuous, respectively, we may choose $\Delta > 0$ arbitrarily small and conclude that

$$D^+A(x^*) - D^-A(x^*) \geq \frac{1}{k}.$$

Since smooth Pickands dependence functions are dense in \mathcal{A} (see [13, 45]), $\mathcal{A}_{k,n}$ is nowhere dense in \mathcal{A} , and the result follows. \square

Translating to EVCs we obtain the main result of this section saying that topologically typical EVC have degenerated discrete component:

Theorem 4.21. *The set $\{C \in \mathcal{C}_{ev} : \mu_C^{dis}(\mathbb{I}^2) = 0\}$ is co-meager in \mathcal{C}_{ev} .*

Proof. Immediate consequence of Lemma 4.2, Lemma 4.20 and Theorem 4.6. \square

Remark 4.22. Viewing Theorem 4.21 in context of Theorem 4.15 implies that a typical EVC $C \in \mathcal{C}_{ev}$ does not exhibit pathological behavior, i.e., $C \notin \mathcal{C}_{ev,p}$.

Combining Corollary 4.21, Corollary 4.19 and Corollary 4.17 yields the following result on typical EVCs.

Corollary 4.23. *A topologically typical EVC C has degenerated discrete component, is not absolutely continuous and has full support. In particular, $\partial_1 C(x, y)$ exists in full $(0, 1)^2$.*

After having established the main results, we round off this section with some simple observations on exchangeable EVCs.

Lemma 4.24. *The following assertions hold.*

- (1) \mathcal{C}_{ev} is nowhere dense in (\mathcal{C}, d_∞) .
- (2) Exchangeable EVCs are nowhere dense in $(\mathcal{C}_e, d_\infty)$.

Proof. Using the fact that \mathcal{C}_{ev} is closed in (\mathcal{C}, d_∞) , for $C \in \mathcal{C}_{ev}$ setting $C_\varepsilon := (1-\varepsilon)C + \varepsilon W$ we have that $C_\varepsilon \notin \mathcal{C}_{ev}$, implying that the interior of \mathcal{C}_{ev} is empty. The second assertion follows similarly using the fact that EVCs are positive quadrant dependent and W is symmetric. \square

Lemma 4.25. *The family \mathcal{A}_s of symmetric Pickands dependence functions is nowhere dense in $(\mathcal{A}, \|\cdot\|_\infty)$.*

Proof. The family of symmetric Pickands dependence functions is obviously closed with respect to $\|\cdot\|_\infty$. Assume that \mathcal{A}_s would not be nowhere dense. Then there exists some $\varepsilon > 0$ such that $B_\varepsilon(A) \subseteq \mathcal{A}_s$, whereby $B_\varepsilon(A)$ denotes the ball with center A and radius ε with respect to $\|\cdot\|_\infty$. Fixing an arbitrary asymmetric Pickands dependence function $D \in \mathcal{A}$ and setting $E := (1 - \frac{\varepsilon}{3})A + \frac{\varepsilon}{3}D$ it follows that E is an asymmetric Pickands dependence function with $E \in B_\varepsilon(A)$ and therefore $E \in \mathcal{A}_s$, a contradiction. \square

Corollary 4.26. *The family of exchangeable EVCs is nowhere dense in $(\mathcal{C}_{ev}, d_\infty)$.*

We conclude this section with an example of an atypical EVC:

Example 4.27. Defining the probability measure

$$\vartheta := \frac{1}{5} \delta_{\frac{1}{4}} + \frac{3}{5} \delta_{\frac{1}{2}} + \frac{1}{5} \delta_{\frac{3}{4}},$$

clearly $\vartheta \in \mathcal{P}_{\mathcal{A}}$, and the corresponding Pickands dependence function is given by

$$A(t) := \begin{cases} 1 - t, & \text{if } t \in [0, \frac{1}{4}) \\ -\frac{3}{5}t + \frac{9}{10}, & \text{if } t \in [\frac{1}{4}, \frac{1}{2}) \\ \frac{3}{5}t + \frac{3}{10}, & \text{if } t \in [\frac{1}{2}, \frac{3}{4}) \\ t, & \text{if } t \in [\frac{3}{4}, 1]. \end{cases}$$

According to Corollary 4.23 the EVC C_A is atypical, since it has a non-degenerated discrete component which is concentrated on the graphs of the functions $f^{\frac{1}{4}}$, $f^{\frac{1}{2}}$ and $f^{\frac{3}{4}}$. Moreover, C_A does not have full support, since $\text{supp}(\mu_{C_A}) = \{(x, y) \in \mathbb{I}^2: f^{\frac{1}{4}}(x) \leq y \leq f^{\frac{3}{4}}(x)\}$. A sample of the copula C_A is depicted in Figure 5.

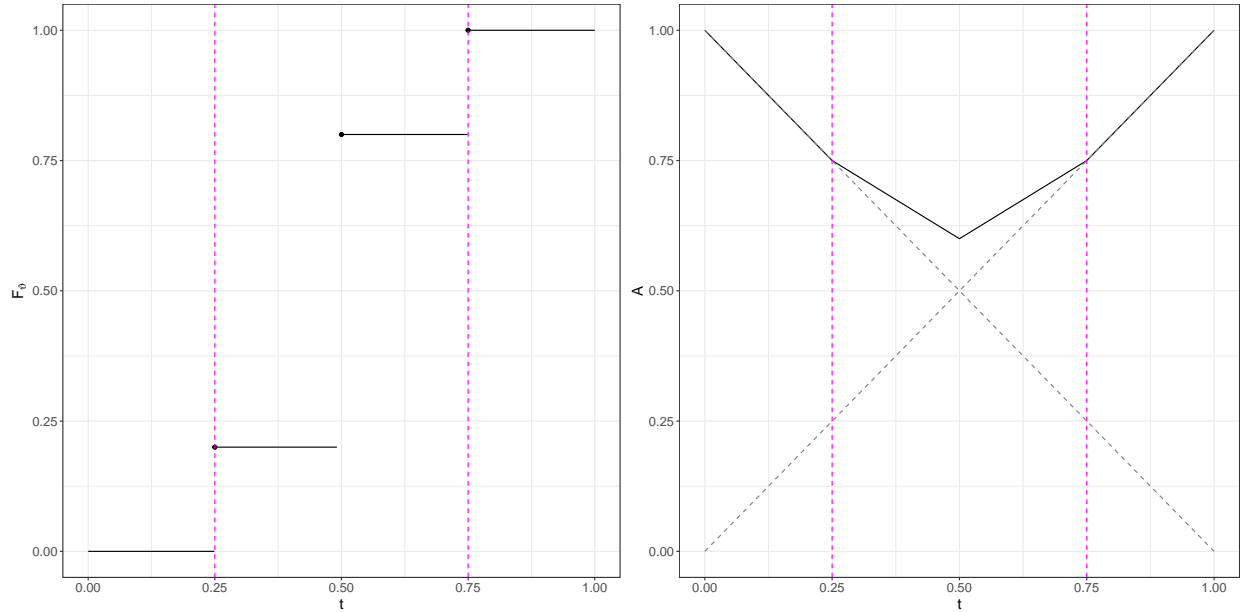


Figure 4: Graphs of the distribution function F_ϑ of the Pickands dependence measure ϑ (left) according to Example 4.27 and Pickands dependence function A (right) associated with it. The dashed black lines mark $L_\vartheta = \frac{1}{4}$ and $R_\vartheta = \frac{3}{4}$ according to equations (4.11) and (4.9), respectively.

Appendix A. Auxiliary results

The proof of the following lemma is included for the sake of completeness:

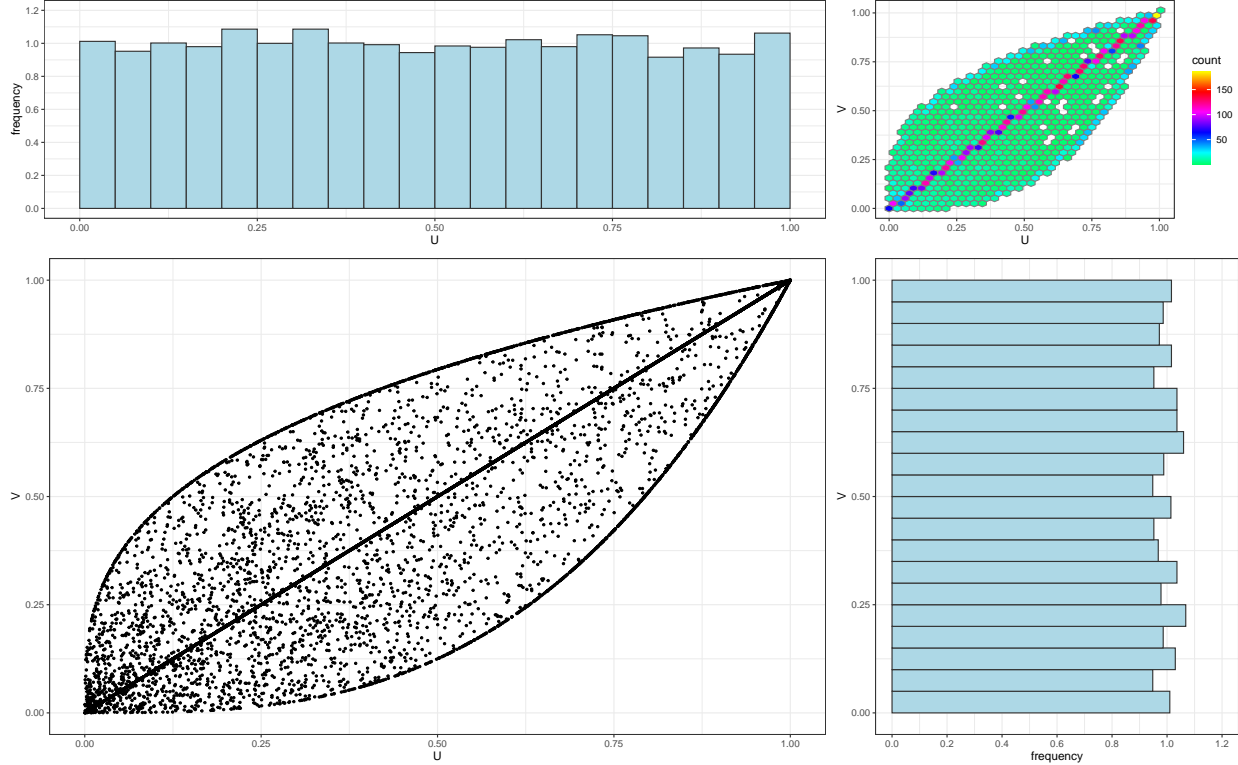


Figure 5: Sample of size 10000 of the EVC C_A , where A is the Pickands dependence function considered in Example 4.27, its histogram and the two marginal histograms. The sample has been generated via conditional inverse sampling.

Lemma Appendix A.1. *Let E_1 and E_2 be two topological spaces and $f: E_1 \rightarrow E_2$ be a homeomorphism. Then the following assertions hold:*

- (1) $G \subseteq E_1$ is nowhere dense in E_1 , if and only if $f(G)$ is nowhere dense in E_2 .
- (2) If $G \subseteq E_1$ is of first Baire category in E_1 , then $f(G)$ is of first Baire category in E_2 .
- (3) If $G \subseteq E_1$ is of second Baire category in E_1 , then $f(G)$ is of second Baire category in E_2 .
- (4) If $G \subseteq E_1$ is co-meager in E_1 , then $f(G)$ is co-meager in E_2 .

Proof. (1): Let $G \subseteq E_1$ be nowhere dense. Then $\text{int}(\overline{G}) = \emptyset$. Since homeomorphisms preserve interior and closure, we have that $\text{int}(f(\overline{G})) = \text{int}(f(\overline{G})) = f(\text{int}(\overline{G})) = f(\emptyset) = \emptyset$. Now assume that $f(G)$ is nowhere dense. Then $\emptyset = \text{int}(f(\overline{G}))$ and using the fact that f is injective we have that $\emptyset = \text{int}(\overline{G})$, implying that G is nowhere dense in E_1 .

(2): Let $G \subseteq E_1$ be of first Baire category. Then there exist countably many nowhere dense sets $(G_n)_{n \in \mathbb{N}}$ in E_1 such that $G = \bigcup_{n \in \mathbb{N}} G_n$. Since f is surjective we have that $f(G) = f(\bigcup_{n \in \mathbb{N}} G_n) = \bigcup_{n \in \mathbb{N}} f(G_n)$. Since G_n is nowhere dense in E_1 , applying (1) yields that $f(G_n)$ is nowhere dense in E_2 . Thus, $f(G)$ is the union of nowhere dense sets and therefore of first Baire category.

(3): Let $G \subseteq E_1$ be a set of second category. Suppose that $f(G)$ is not of second category, i.e., that it is of first Baire category. Thus, $f(G) = \bigcup_{n \in \mathbb{N}} A_n$ for nowhere dense sets A_1, A_2, \dots in E_2 . Since f is a homeomorphism, we can find a set G_n such that $A_n = f(G_n)$. Applying (1) yields that G_n is nowhere dense and therefore $f(G) = \bigcup_{n \in \mathbb{N}} A_n = \bigcup_{n \in \mathbb{N}} f(G_n) = f(\bigcup_{n \in \mathbb{N}} G_n)$. Again, using that f is a homeomorphism yields that $G = \bigcup_{n \in \mathbb{N}} G_n$ and therefore G would be of first Baire category in E_1 . A contradiction.

(4): Let $G \subseteq E_1$ be a co-meager set in E_1 . Then its complement G^c is of first Baire category. Since f is a homeomorphism we have that $f(G)^c \subseteq f(G^c)$. Applying (2), $f(G^c)$ is of first Baire category in E_2 and since subsets of sets of first Baire category are of first Baire category too, $f(G)^c$ is of first Baire category, implying that $f(G)$ is co-meager. \square

We prove the interrelation between the Pickands dependence function $A \in \mathcal{A}$ and the Pickands dependence measure $\vartheta \in \mathcal{P}_{\mathcal{A}}$ according to equation (4.4). Note that this interrelation goes back to [1, 14, 35], we here only include a quick proof for the sake of completeness.

Lemma Appendix A.2. *Suppose that $A \in \mathcal{A}$ and that $\vartheta \in \mathcal{P}_{\mathcal{A}}$ is the corresponding Pickands dependence measure. Then*

$$A(t) = 1 - t + 2 \int_{[0,t]} \vartheta([0, z]) d\lambda(z) \quad (\text{A.1})$$

holds for all $t \in \mathbb{I}$.

Proof. According to equation (4.3), there exists a measure $\vartheta \in \mathcal{P}_{\mathcal{A}}$ with

$$A(t) = 2 \int_{[0,1]} \max\{t(1-s), (1-t)s\} d\vartheta(s). \quad (\text{A.2})$$

Using $\vartheta \in \mathcal{P}_{\mathcal{A}}$ equation (A.2) boils down to

$$\begin{aligned} \frac{A(t)}{2} &= t \int_{[0,t]} (1-s) d\vartheta(s) + (1-t) \int_{(t,1]} s d\vartheta(s) \\ &= t \vartheta([0, t]) - t \int_{[0,t]} s d\vartheta(s) + \int_{(t,1]} s d\vartheta(s) - t \int_{(t,1]} s d\vartheta(s) \\ &= t \vartheta([0, t]) + \int_{(t,1]} s d\vartheta(s) - t \int_{[0,1]} s d\vartheta(s) \\ &= t \vartheta([0, t]) + \int_{(t,1]} s d\vartheta(s) - \frac{t}{2} \\ &= t \vartheta([0, t]) + \frac{1}{2} - \int_{[0,t]} s d\vartheta(s) - \frac{t}{2} \\ &= t \vartheta([0, t]) + \frac{1-t}{2} - \int_{[0,t]} s d\vartheta(s). \end{aligned}$$

The latter integral, however, can easily be simplified to (it is the expectation of the probability measure $\vartheta'(E) := \frac{\vartheta(E)}{\vartheta([0,t])}$ and hence can be expressed as integral over the

corresponding distribution function)

$$\begin{aligned} \int_{[0,t]} s d\vartheta(s) &= \vartheta([0,t]) \int_{[0,t]} (1 - \vartheta'([0,s])) d\lambda(s) \\ &= \vartheta([0,t])t - \int_{[0,t]} \vartheta([0,s]) d\lambda(s). \end{aligned}$$

Altogether we therefore get

$$\frac{A(t)}{2} = \frac{1-t}{2} + \int_{[0,t]} \vartheta([0,s]) d\lambda(s),$$

implying

$$A(t) = 1 - t + 2 \int_{[0,t]} \vartheta([0,s]) d\lambda(s), \quad (\text{A.3})$$

which completes the proof \square

The next lemma states that weak convergence of measures in $\mathcal{P}_{\mathcal{A}}$ is equivalent to uniform convergence of the corresponding EVC (in fact, even to weak conditional convergence, see [22]).

Lemma Appendix A.3. *Let $\vartheta, \vartheta_1, \vartheta_2, \dots \in \mathcal{P}_{\mathcal{A}}$ be Pickands dependence measures and $A, A_1, A_2, \dots \in \mathcal{A}$ as well as $C, C_1, C_2, \dots \in \mathcal{C}_{ev}$ the corresponding Pickands dependence functions and EVCs, respectively. Then the following three conditions are equivalent:*

- (i) $\vartheta_n \xrightarrow{n \rightarrow \infty} \vartheta$ weakly on \mathbb{I} ,
- (ii) $A_n \xrightarrow{n \rightarrow \infty} A$ uniformly on \mathbb{I} ,
- (iii) $C_n \xrightarrow{n \rightarrow \infty} C$ uniformly on \mathbb{I}^2 .

Proof. Considering [22, Theorem 5.1] it suffices to show the equivalence of (i) and (ii). Note that [22, Theorem 5.1] only states pointwise convergence of $(A_n)_{n \in \mathbb{N}}$ to A . Uniform convergence, however, follows via Lipschitz continuity of Pickands functions. The implication (i) \Rightarrow (ii) is a direct consequence of Lemma Appendix A.2 and dominated convergence. On the other hand, again using [22, Theorem 5.1], we have that $D^+ A_n(t) \xrightarrow{n \rightarrow \infty} D^+ A(t)$ for every continuity point t of $D^+ A$, which, applying equation (4.5) directly yields (i). \square

For establishing the regularity results for EVCs summarized in Theorem 4.6 the following technical lemma will be used:

Lemma Appendix A.4. *Let $F: \mathbb{I} \rightarrow \mathbb{I}$ be a strictly increasing, absolutely continuous distribution function, μ_G be a finite measure on $\mathcal{B}(\mathbb{I})$ with measure-generating function G and fulfilling $\mu_G((0,1)) > 0$. Furthermore let μ_H be the measure induced by the measure-generating function $H: \mathbb{I} \rightarrow \mathbb{I}$ defined by $H := F \cdot G$. Then the following three assertions hold:*

(i) If μ_G is singular, then μ_H has non-degenerated singular component.

(ii) If μ_G is discrete in $(0, 1)$, then $\mu_H^{sing}(\mathbb{I}) = 0$.

(iii) If μ_G has a point mass in $(0, 1)$, then so does μ_H .

Proof. (i) Continuity of F implies that H is continuous on \mathbb{I} . Letting f denote the density of F and considering the derivative of H yields

$$H' = F'G + FG' = fG + FG'$$

λ -almost everywhere on \mathbb{I} . Using singularity of G therefore implies $H' = fG$ λ -almost everywhere on \mathbb{I} .

If H had no singular component, it would be absolutely continuous with density $f \cdot G$. Therefore, considering an interval $(x_1, x_2) \subseteq (0, 1]$ with $G(x_1) < G(x_2)$ and using the fact that F is strictly increasing and greater than 0 on $(0, 1]$, it follows that

$$\begin{aligned} H(x_2) - H(x_1) &= \int_{[x_1, x_2]} f(s)G(s)d\lambda(s) \\ &\leq G(x_2)[F(x_1) - F(x_2)] \\ &< G(x_2)F(x_2) - G(x_1)F(x_1) \\ &= H(x_2) - H(x_1), \end{aligned}$$

a contradiction, so H has non-degenerated singular component.

(ii) Assume that μ_G is discrete. Then there exists a (finite or countably infinite) index-set I and numbers $q_i \in [0, 1)$ and $a_i \in \mathbb{I}$ for every $i \in I$, such that $\mu_G = \sum_{i \in I} a_i \delta_{q_i}$ holds. Then obviously H is given by

$$H(t) = F(t)G(t) = \sum_{i \in I} a_i F(t) \mathbf{1}_{[0, t]}(q_i)$$

for every $t \in (0, 1)$. H is continuous outside the set $\{q_i : i \in I\}$, so the discrete component H^{dis} of H is given by $H^{dis}(t) = \sum_{i \in I} a_i F(q_i) \delta_{q_i}([0, t])$. Defining $\psi(t) := H(t) - H^{dis}(t)$ and letting f again denote the density of F , working with the definition of ψ and applying Fubini's theorem for non-negative measurable functions, we obtain that

$$\begin{aligned} \psi(t) &= \sum_{i \in I} a_i [F(t) - F(q_i)] \delta_{q_i}([0, t]) \\ &= \sum_{i \in I} a_i \int_{(q_i, t]} f(s) d\lambda(s) \delta_{q_i}([0, t]) \\ &= \int_{\mathbb{I}} \sum_{i \in I} a_i f(s) \mathbf{1}_{(q_i, t]}(s) \delta_{q_i}([0, t]) d\lambda(s) \\ &= \int_{[0, t]} f(s) \sum_{q_i < s} a_i d\lambda(s). \end{aligned}$$

Observing that, defining $g(s) := f(s) \sum_{q_i < s} a_i \geq 0$ for every $s \in \mathbb{I}$, we obviously have that $g \in L^1(\mathbb{I}, \mathcal{B}(\mathbb{I}), \lambda)$ it follows that ψ is absolutely continuous on \mathbb{I} . In other words: the singular component of H is degenerated.

(iii) The third assertion directly follows from the fact that in the case of x_0 being a point mass of μ_G we have

$$\mu_H(\{x_0\}) = F(x_0) \cdot \mu_G(\{x_0\}) > 0.$$

□

Lemma Appendix A.5. *Let ν be a measure on $\mathcal{B}(\mathbb{I})$ with $\nu(\mathbb{I}) \leq 1$. Then the function $g_\nu: \mathbb{I} \rightarrow [0, \infty)$ defined by*

$$g_\nu(t) := 1 - t + 2 \int_{[0,t]} \nu([0, z]) d\lambda(z)$$

for all $t \in \mathbb{I}$ is convex and 1-Lipschitz continuous.

Proof. The function g_ν is absolutely continuous by definition and (one version of) its density $k_\nu: \mathbb{I} \rightarrow [-1, 1]$ is given by $k_\nu(t) = -1 + 2\nu([0, t]) \in [-1, 1]$, implying that

$$|g_\nu(t_1) - g_\nu(t_2)| \leq \int_{[x_1, x_2]} |k_\nu(s)| d\lambda(s) \leq |t_1 - t_2|$$

for every $t_1, t_2 \in \mathbb{I}$, whereby $x_1 = \min\{t_1, t_2\}$ and $x_2 = \max\{t_1, t_2\}$. In other words: g_ν is 1-Lipschitz continuous. Working with the fact that $z \mapsto \nu([0, z])$ is non-decreasing, the function $t \mapsto \int_{[0,t]} \nu([0, z]) d\lambda(z)$ is convex and hence, g_ν is as sum of two convex functions convex as well. □

The following lemma generalizes Lemma 5 in [45].

Lemma Appendix A.6. *Let $f: \mathbb{I} \rightarrow [0, \infty)$ be a convex function fulfilling $D^+f(0) \geq -1$ and $f(0) = 1$. Then the function $G_f: [0, 1) \rightarrow \mathbb{R}$, defined by*

$$G_f(t) := D^+f(t)(1 - t) + f(t)$$

is non-decreasing, non-negative and right-continuous.

Proof. Convexity of f yields that f is continuous on $(0, 1)$, existence of $D^+f(0)$ implies continuity of f at 0. Moreover, the right-hand derivative D^+f of f is right-continuous (see [20, 36]), hence right-continuity of G_f follows. For $0 \leq t_1 < t_2 < 1$ convexity of f implies $D^+f(t_1) \leq D^+f(t_2)$. Therefore, setting $\delta := D^+f(t_2)(1 - t_2) + f(t_2) - (D^+f(t_1)(1 - t_1) + f(t_1))$, we obtain that

$$\begin{aligned} \delta &\geq D^+f(t_2)(1 - t_2) + f(t_1) + D^+f(t_1)(t_2 - t_1) - (D^+f(t_1)(1 - t_1) + f(t_1)) \\ &\geq D^+f(t_1)(1 - t_2) + f(t_1) + D^+f(t_1)(t_2 - t_1) - (D^+f(t_1)(1 - t_1) + f(t_1)) = 0, \end{aligned}$$

implying that G_f is non-decreasing. Finally, non-negativity follows via $G_f(t) \geq G_f(0) = D^+f(0)(1 - 0) + f(0) = 1 + D^+f(0) \geq 0$. □

Proof of Theorem 4.6:

Proof. (i): If ϑ is absolutely continuous on $(0, 1)$, applying equation (4.5) yields that D^+A is absolutely continuous too. Since A is Lipschitz-continuous and finite sums of absolutely continuous functions are absolutely continuous, G_A is absolutely continuous on $[0, 1)$, i.e., there exists some function $g \in L^1(\mathbb{I}, \mathcal{B}(\mathbb{I}), \lambda)$, such that

$$G_A(t) = G_A(0) + \int_{[0,t]} g(s) d\lambda(s),$$

holds for every $t \in [0, 1)$. Using the fact that for fixed $x \in (0, 1)$ the mapping $\varphi_x : y \mapsto \frac{\log(x)}{\log(xy)}$ is a diffeomorphism of $(0, 1)$, applying change of coordinates yields that

$$G_A(\varphi_x(y)) = G_A(0) + \int_{[0,y]} g(\varphi_x(s)) \left| \frac{\log(x)}{s \log^2(xs)} \right| d\lambda(s),$$

implying that $y \mapsto G_A(\varphi_x(y))$ is absolutely continuous on $[0, 1)$. Since C_A is Lipschitz-continuous and products of absolutely continuous functions on compact intervals are absolutely continuous again (see [12, Section 7.5, Exercise 7]) the function $y \mapsto K_A(x, [0, y])$ according to equation (4.8) is absolutely continuous on $(0, 1)$. Therefore, applying disintegration shows that C_A is absolutely continuous.

Considering the reverse implication, we know that there exists some Borel set $\Lambda \subseteq (0, 1)$ with $\lambda(\Lambda) = 1$ such that for every $x \in \Lambda$ the measure $K_A(x, \cdot)$ is absolutely continuous. Fixing $x \in \Lambda$, using that $y \mapsto G_A(\varphi_x(y))$ induces a probability measure and $y \mapsto \frac{C(x,y)}{x}$ is a strictly increasing absolutely continuous distribution function, $y \mapsto G_A(\varphi_x(y))$ is absolutely continuous. In fact, if this was not the case, then $y \mapsto G_A(\varphi_x(y))$ would have a discrete or singular component and thus, using the fact that $K_A(x, \cdot)$ is absolutely continuous together with Lemma Appendix A.4 would yield a contradiction. Proceeding as in the previous case, applying change of coordinates shows absolute continuity of G_A and therefore, considering that the function A is Lipschitz-continuous and $t \mapsto 1 - t$ is smooth, absolute continuity of D^+A follows. Applying equation (4.5) now completes the proof of the first assertion.

(ii) The second assertion has already been proved in Lemma 4.4.

(iii) For proving assertion number three assume that $\vartheta \in \mathcal{P}_{\mathcal{A}}$ has a singular component, i.e., $\vartheta = \vartheta^{sing} + \mu$ with $\vartheta^{sing}(\mathbb{I}) > 0$ and μ being the non-singular component. Working with equation (4.8) for fixed $x \in (0, 1)$ yields

$$K_C(x, [0, y]) = \frac{C(x, y)}{x} \left[G_A^{sing} \left(\frac{\log(x)}{\log(xy)} \right) + G_A^\mu \left(\frac{\log(x)}{\log(xy)} \right) \right],$$

where $y \in (0, 1)$, $G_A^i(t) = D^+A_i(t)(1 - t) + A_i(t)$, $i \in \{sing, \mu\}$, and A_i according to equation (A.1) with the measures ϑ_{sing} and μ being used instead of ϑ . Applying Lemma Appendix A.5 each A_i is convex and thus, using Lemma Appendix A.6, G_A^i are non-negative, non-decreasing and right-continuous. Furthermore applying that $y \mapsto \varphi_x(y)$ is non-decreasing and right-continuous as well, G_A^i induces a measure for $i \in \{sing, \mu\}$. Moreover, using that φ_x is continuous, strictly increasing and that its derivative is bounded from below by $\ell = -\frac{e^2}{4}x \log(x) > 0$ (see [45, Example 4]), yields that $\frac{\partial}{\partial y} G_A^{sing}(\varphi_x(y)) = 0$

for λ -almost every $y \in (0, 1)$, implying that $y \mapsto G_A^{sing}(\varphi_x(y))$ is continuous and singular. Applying Lemma Appendix A.4 yields that $y \mapsto \frac{C(x,y)}{x} G_A^{sing}(\varphi_x(y))$ has a non-degenerated singular component and therefore $K_C^{sing}(x, \mathbb{I}) > 0$. Applying disintegration yields the desired result.

To prove the reverse implication assume that $\vartheta^{sing}(\mathbb{I}) = 0$, i.e., $\vartheta = \vartheta^{dis} + \vartheta^{abs}$. Let $x \in (0, 1)$ be arbitrary but fixed. Again working with equation (4.8) yields that

$$K_C(x, [0, y]) = \frac{C(x, y)}{x} \left[G_A^{abs} \left(\frac{\log(x)}{\log(xy)} \right) + G_A^{dis} \left(\frac{\log(x)}{\log(xy)} \right) \right],$$

where G_A^{abs} and G_A^{dis} are as in the previous case only for the absolutely continuous and discrete components of ϑ , respectively. Using the same arguments as in the proof of the first assertion shows that $y \mapsto \frac{C(x,y)}{x} G_A^{abs}(\varphi_x(y))$ is absolutely continuous on $(0, 1)$. Considering G_A^{dis} , notice that according to Lemma Appendix A.5 A^{dis} is Lipschitz-continuous and, since $D^+ A^{dis}$ is discrete and $t \mapsto (1-t)$ is smooth, G_A^{dis} has no singular component. Suppose $y \mapsto G_A^{dis} \left(\frac{\log(x)}{\log(xy)} \right) =: Q(y)$ has a non-degenerated singular component, i.e., $\mu_Q = \mu_Q^r + \mu_Q^s$, whereby μ_Q denotes the measure induced by Q and μ_Q^s the non-degenerated singular component of μ_Q and μ_Q^r summarizes the absolutely continuous and discrete component. Denote the measure generating functions of μ_Q^s and μ_Q^r by Q^s and Q^r , respectively. Then

$$G_A^{dis}(t) = G_A^{dis} \left(\frac{\log(x)}{\log(xx^{\frac{1}{t}-1})} \right) = Q(x^{\frac{1}{t}-1}) = Q^r(x^{\frac{1}{t}-1}) + Q^s(x^{\frac{1}{t}-1})$$

and the function $k_x(t) := x^{\frac{1}{t}-1}$ is differentiable for every $t \in (0, 1)$. Applying [38, Lemma 7.25] yields that k_x maps sets of full λ -measure to sets of full λ -measure. Applying the obvious fact that $t \mapsto Q^s(x^{\frac{1}{t}-1})$ is continuous and that

$$\frac{d}{dt} Q^s(x^{\frac{1}{t}-1}) = (Q^s)'(x^{\frac{1}{t}-1}) \frac{x^{\frac{1}{t}-1}}{t^2 \log(x)} = 0$$

for λ -almost every $t \in \mathbb{I}$, the measure induced by G_A^{dis} would have a non-degenerated singular component; a contradiction. Therefore $y \mapsto G_A^{dis}(\varphi_x(y))$ has degenerated singular component and thus, multiplying $y \mapsto G_A^{dis}(\varphi_x(y))$ with the strictly increasing, absolutely continuous distribution function $y \mapsto \frac{C(x,y)}{x}$ and working with the fact that $y \mapsto G_A^{dis}(\varphi_x(y))$ only has non-degenerated absolutely continuous and discrete component, proceeding analogously as in the proof of the first part of the Lemma and applying Lemma Appendix A.4 (ii) completes the proof of the last assertion. \square

The next lemma shows that we can normalize sequences of weakly converging probability measures with limit in $\mathcal{P}_{\mathcal{A}}$ in such a way that all elements of the sequence are in $\mathcal{P}_{\mathcal{A}}$ while preserving weak convergence to the limit.

Lemma Appendix A.7. *Let $\mu_1, \mu_2, \dots \in \mathcal{P}(\mathbb{I})$ and $\vartheta \in \mathcal{P}_{\mathcal{A}}$ fulfill that the sequence $(\mu_n)_{n \in \mathbb{N}}$ converges weakly to ϑ . Then there exist constants $\alpha_1, \alpha_2, \dots \in (0, 1)$ and $\beta_1, \beta_2, \dots \in (0, 1)$*

such that the measures

$$\vartheta_n := \begin{cases} (1 - \alpha_n)\delta_0 + \alpha_n\mu_n, & \text{if } \mathbb{E}(\mu_n) > \frac{1}{2} \\ \mu_n, & \text{if } \mathbb{E}(\mu_n) = \frac{1}{2} \\ (1 - \beta_n)\delta_1 + \beta_n\mu_n, & \text{if } \mathbb{E}(\mu_n) < \frac{1}{2} \end{cases} \quad (\text{A.4})$$

fulfill the following assertions:

(i) $\vartheta_n \in \mathcal{P}_{\mathcal{A}}$ for every $n \in \mathbb{N}$.

(ii) The sequence $(\vartheta_n)_{n \in \mathbb{N}}$ converges weakly to ϑ .

Proof. For a given μ_n set $\alpha_n := \frac{1}{2\mathbb{E}(\mu_n)}$ if $\mathbb{E}(\mu_n) > \frac{1}{2}$, $\beta_n := \frac{1}{2(1-\mathbb{E}(\mu_n))}$ if $\mathbb{E}(\mu_n) < \frac{1}{2}$ and define ϑ_n according to equation (A.4). Then for $\mathbb{E}(\mu_n) > \frac{1}{2}$ we have that

$$\mathbb{E}(\vartheta_n) = (1 - \alpha_n)\mathbb{E}(\delta_0) + \alpha_n\mathbb{E}(\mu_n) = \frac{\mathbb{E}(\mu_n)}{2\mathbb{E}(\mu_n)} = \frac{1}{2}$$

and in the case $\mathbb{E}(\mu_n) < \frac{1}{2}$, we obtain that

$$\int_{\mathbb{I}} s d\vartheta_n(s) = (1 - \beta_n) + \beta_n\mathbb{E}(\mu_n) = \left(1 - \frac{1}{2(1 - \mathbb{E}(\mu_n))}\right) + \frac{\mathbb{E}(\mu_n)}{2(1 - \mathbb{E}(\mu_n))} = \frac{1}{2}.$$

This already proves the first assertion.

To show the second one we proceed as follows: Defining

$$I^> := \{n \in \mathbb{N} : \mathbb{E}(\mu_n) > 1/2\}, I^< := \{n \in \mathbb{N} : \mathbb{E}(\mu_n) < 1/2\}, I^= := \{n \in \mathbb{N} : \mathbb{E}(\mu_n) = 1/2\}$$

we obviously have $I^> \cup I^< \cup I^= = \mathbb{N}$. If $I^>$ contains infinitely many elements, increasingly enumerating its elements by $i_1^>, i_2^>, \dots$, it follows immediately that $(\alpha_{i_j^>})_{j \in \mathbb{N}}$ converges to 1. Furthermore weak convergence implies that

$$\begin{aligned} \lim_{j \rightarrow \infty} \int_{\mathbb{I}} h(x) d\vartheta_{i_j^>}(x) &= \lim_{j \rightarrow \infty} \left[(1 - \alpha_{i_j^>}) \int_{\mathbb{I}} h(x) d\delta_0(x) + \alpha_{i_j^>} \int_{\mathbb{I}} h(x) d\mu_{i_j^>}(x) \right] \\ &= \lim_{j \rightarrow \infty} \left[(1 - \alpha_{i_j^>}) h(0) + \alpha_{i_j^>} \int_{\mathbb{I}} h(x) d\mu_{i_j^>}(x) \right] \\ &= \int_{\mathbb{I}} h(x) d\vartheta(x) \end{aligned}$$

holds for every continuous function $h: \mathbb{I} \rightarrow \mathbb{R}$. For the case that $I^<$ is infinite we proceed analogously to show that $(\beta_{i_j^<})_{j \in \mathbb{N}}$ converges to 1 and that $(\vartheta_{i_j^<})_{j \in \mathbb{N}}$ converges weakly to ϑ . Since for $n \in I^=$ we have $\vartheta_n = \mu_n$ it altogether follows that the sequence $(\vartheta_n)_{n \in \mathbb{N}}$ converges weakly to ϑ , and the proof is complete. \square

The next lemma constitutes that both, the family of Pickands dependence measures whose discrete component has full support as well as the family of Pickands dependence measures whose singular component has full support are dense in $\mathcal{P}_{\mathcal{A}}$.

Lemma Appendix A.8. *The following assertions hold:*

- (i) *The family of $\vartheta \in \mathcal{P}_{\mathcal{A}}$ with $\text{supp}(\vartheta^{dis}) = \mathbb{I}$ is dense in $(\mathcal{P}_{\mathcal{A}}, \tau_w)$.*
- (ii) *The family of $\vartheta \in \mathcal{P}_{\mathcal{A}}$ with $\text{supp}(\vartheta^{sing}) = \mathbb{I}$ is dense in $(\mathcal{P}_{\mathcal{A}}, \tau_w)$.*

Proof. To prove the first assertion we proceed as follows: Let $\vartheta \in \mathcal{P}_{\mathcal{A}}$ be arbitrary but fixed and choose some discrete measure $m_1 \in \mathcal{P}(\mathbb{I})$ with full support. Then defining $\mu_n := (1 - \frac{1}{n})\vartheta + \frac{1}{n}m_1$ yields that $\mu_n \in \mathcal{P}(\mathbb{I})$, $\text{supp}(\mu_n^{dis}) = \mathbb{I}$ and the sequence $(\mu_n)_{n \in \mathbb{N}}$ converges to ϑ weakly. Normalizing the sequence $(\mu_n)_{n \in \mathbb{N}}$ in the sense of Lemma Appendix A.7, we obtain measures $\vartheta_n \in \mathcal{P}_{\mathcal{A}}$ with $\text{supp}(\vartheta_n^{dis}) = \mathbb{I}$ such that $(\vartheta_n)_{n \in \mathbb{N}}$ converges weakly to ϑ . On the other hand, choosing a singular probability measure $m_2 \in \mathcal{P}(\mathbb{I})$ with $\text{supp}(m_2) = \mathbb{I}$ (see [11, 15] for an example for m_2) and setting $\nu_n := (1 - \frac{1}{n})\vartheta + \frac{1}{n}m_2$ yields the desired result. \square

Acknowledgement

The first author gratefully acknowledges the support of Red Bull GmbH within the ongoing Data Science collaboration with the university of Salzburg. The second author gratefully acknowledges the support of the WISS 2025 project ‘IDA-lab Salzburg’ (20204-WISS/225/197-2019 and 20102-F1901166-KZP).

References

- [1] J. Beirlant, Y. Goegebeur, J. Segers and J. Teugels, *Statistics of extremes: Theory and Applications*, Wiley Series in Probability and Statistics, John Wiley and Sons Ltd., Chichester, 2004.
- [2] A.M. Bruckner, B.S. Thomson, *Real Analysis*, Prentice-Hall, London, 1997.
- [3] W.F. Darsow, B. Nguyen, E.T. Olsen: Copulas and Markov processes, *Ill. J. Math.* 36 (1992) 600–642.
- [4] F. Durante, J. Fernández-Sánchez and C. Ignazzi, Baire category results for stochastic orders, *Rev. Real Acad. Cienc. Exactas Fis. Nat. - A: Mat.* 116 (2022) 188.
- [5] F. Durante, J. Fernández-Sánchez and W. Trutschnig, Baire category results for exchangeable copulas, *Fuzzy Sets Syst.* 284 (2016) 146-151.
- [6] F. Durante, J. Fernández-Sánchez and W. Trutschnig, Baire category results for quasi-copulas, *Depend. Model.* 4 (2016) 215-223.
- [7] F. Durante and C. Sempi, *Principles of Copula Theory*, Chapman and Hall/CRC, 2015.
- [8] M. Falk, J. Häslér, R.D. Reiss, *Laws of Small Numbers: Extremes and Rare Events*, Springer, Basel, 2011.

- [9] D. Feldman, Extreme Doubly Stochastic Measures with Full Support, *Proc. Am. Math. Soc.* 114 (1992) 919–922.
- [10] J. Fernández Sánchez and W. Trutschnig, A typical copula is singular, *J. Math. Anal. Appl.* 430 (2015) 517-527.
- [11] J. Fernández Sánchez, P. Viader, J. Paradis and M. Díaz Carrillo, A singular function with a non-zero finite derivative, *Nonlinear Anal-Theor.* 15 (2012) 5010-5014.
- [12] R.F. Gariepy and W.P. Ziemer, *Modern Real Analysis*, PWS Pub., 1995.
- [13] S. Guillotte and F. Perron, Polynomial Pickands functions, *Bernoulli* 22 (2016) 213–241.
- [14] L.d. Haan and S.I. Resnick, Limit theory for multivariate sample extremes, *Probab. Theory Relat. Fields* 40 (1977) 317-337.
- [15] E. Hewitt, K. Stromberg, *Real and Abstract Analysis, A Modern Treatment of the Theory of Functions of a Real Variable*, Springer-Verlag, New York, 1965.
- [16] P. Jaworski, F. Durante and W.K. Härdle, *Copulae in Mathematical and Quantitative Finance*, volume 213 of *Lecture Notes in Statistics - Proceedings*, Springer, Berlin Heidelberg, 2013.
- [17] P. Jaworski, F. Durante, W.K. Härdle and T. Rychlik, editors, *Copula Theory and its Applications*, volume 198 of *Lecture Notes in Statistics - Proceedings*, Springer, Berlin Heidelberg, 2010.
- [18] H. Joe, *Dependence Modeling with Copulas*, Chapman and Hall/CRC, London, 2014.
- [19] O. Kallenberg, *Foundations of modern probability*, Springer-Verlag, New York, 2002.
- [20] R. Kannan, C.K. Krueger, *Advanced analysis on the real line*. Springer Verlag, New York, 1996.
- [21] T. Kasper, N. Dietrich and W. Trutschnig, On convergence and mass distributions of multivariate Archimedean copulas and their interplay with the Williamson transform, *J. Math. Anal. Appl.* 529 (2024) 127555.
- [22] T. Kasper, S. Fuchs and W. Trutschnig On weak conditional convergence of bivariate Archimedean and Extreme Value copulas, and consequences to nonparametric estimation, *Bernoulli* 4 (2021) 2217 - 2240.
- [23] C. Kim, Uniform approximation of doubly stochastic operators, *Pac. J. Math.* 26 (1968) 515 - 527.
- [24] A. Klenke, *Wahrscheinlichkeitstheorie*, Springer Lehrbuch Masterclass Series, Berlin Heidelberg, 2008.
- [25] K. Lange, Decompositions of Substochastic Transition Functions, *Proc. Am. Math. Soc.* 37 (1973) 575 – 580.

- [26] X. Li, P. Mikusinski and M.D. Taylor, Strong approximation of copulas, *J. Math. Anal. Appl.* 255 (1998) 608-623.
- [27] J.F. Mai, M. Scherer, Bivariate extreme-value copulas with discrete Pickands dependence measure, *Extremes* 14 (2011) 311-324.
- [28] J.-F. Mai and M. Scherer. Simulating copulas: Stochastic Models, Sampling Algorithms, and Applications, volume 4 of *Series in Quantitative Finance*, Imperial College Press, London, 2012.
- [29] A. McNeil, R. Frey, P. Embrechts, *Quantitative risk management*, Princeton University Press, New Jersey, 2005.
- [30] M. Moshe, V.J. Mizel, Measurability of Partial Derivatives, *Proc. Am. Math. Soc.* 63 (1977) 236 – 38.
- [31] S. Mulinacci, F. Gobbi and U. Cherubini, *Dynamic Copula Methods in Finance*, 2011.
- [32] R.B. Nelsen, *An introduction to copulas*. Springer Series in Statistics, Springer, New York, second edition, 2006.
- [33] E.T. Olsen, W.F. Darsow, B. Nguyen, *Copulas and Markov Operators*, Lecture Notes-Monograph Series 28 (1996) 244–259.
- [34] J.C. Oxtoby, *Measure and category. A survey of the analogies between topological and measure spaces*, 2nd ed. Springer-Verlag, 1980.
- [35] J. Pickands, Multivariate extreme value distributions, in: *Proceedings 43rd Session International Statistical Institute 2*, 1981, pp. 859-878.
- [36] D. Pollard, *A User’s Guide to Measure Theoretic Probability*. Cambridge University Press, 2001.
- [37] R. Ricci, A differential characterization of the d-increasingness property, *Fuzzy Sets Syst.* 433 (2022) 79-95.
- [38] W. Rudin, *Real and Complex Analysis*, McGraw-Hill, 1974.
- [39] L. Rüschendorf, *Mathematical risk analysis*, Springer Series in Operations Research and Financial Engineering, Springer, Heidelberg, 2013.
- [40] G. Salvadori, C. De Michele, N.T. Kottegoda, and R. Rosso, *Extremes in Nature - An Approach Using Copulas*, Springer Dordrecht, 2007.
- [41] A. Sklar, Fonctions de répartition à N dimensions et leurs marges, *Publ. Inst. Statist. Univ. Paris* 8 (1959) 229–231.
- [42] J.A. Tawn, Bivariate extreme value theory: models and estimation, *Biometrika* 75 (1988) 397-415.

- [43] W. Trutschnig, On a strong metric on the space of copulas and its induced dependence measure, *J. Math. Anal. Appl.* 384 (2011) 690-705.
- [44] W. Trutschnig, On Céaro convergence of iterates of the Star Product of copulas, *Stat. Probabil. Lett.* 83 (2013) 357-365.
- [45] W. Trutschnig, M. Schreyer and J. Fernández Sánchez, Mass distributions of two-dimensional extreme-value copulas and related results, *Extremes* 19 (2016) 405-427.