On a strong metric on the space of copulas and its induced dependence measure

Wolfgang Trutschnig[a]

[a]*European Centre for Soft Computing, Edificio Científico Tecnológico*
Calle Gonzalo Gutiérrez Quirós, s/n, 3ª planta, 33600 Mieres (Asturias), Spain
Tel.: +34 985456545, Fax: +34 985456699

Abstract

Using the one-to-one correspondence between copulas and Markov operators on $L^1([0,1])$ and expressing the Markov operators in terms of regular conditional distributions (Markov kernels) allows to define a metric D_1 on the space of copulas $\mathcal{C}$ that is a metrization of the strong operator topology of the corresponding Markov operators. It is shown that the resulting metric space $(\mathcal{C}, D_1)$ is complete and separable and that the induced dependence measure ζ_1, defined as a scalar times the D_1-distance to the product copula Π, has various good properties. In particular the class of copulas that have maximum D_1-distance to the product copula is exactly the class of completely dependent copulas, i.e. copulas induced by Lebesgue-measure preserving transformations on $[0,1]$. Hence, in contrast to the uniform distance d_∞, Π can not be approximated arbitrarily well by completely dependent copulas with respect to D_1. The interrelation between D_1 and the so-called ∂-convergence by Mikusinski and Taylor as well as the interrelation between ζ_1 and the mutual dependence measure ω by Siburg and Stoimenov is analyzed. ζ_1 is calculated for some well-known parametric families of copulas and an application to singular copulas induced by certain Iterated Functions Systems is given.

Keywords: Copula, doubly stochastic measure, independence, Markov operator, Iterated Function System

2010 MSC: 62H20, 60E05, 28A80

Email address: wolfgang.trutschnig@softcomputing.es (Wolfgang Trutschnig)

1. Introduction

Considering the uniform distance d_∞ on the space $\mathcal{C}$ of two-dimensional copulas yields a compact metric space $(\mathcal{C}, d_\infty)$ in which the family of shuffles of the minimum copula M are dense (see [7], [16], [19]). If $A \in \mathcal{C}$ is a shuffle of M, μ_A denotes the corresponding doubly stochastic measure and X, Y random variables on a probability space $(\Omega, \mathcal{A}, \mathcal{P})$ with $\mathcal{P}^{X \otimes Y} = \mu_A$, then X and Y are mutually completely dependent (see [19]) and knowing X implies knowing Y and vice versa. Consequently the product copula Π (describing complete unpredictability) can be approximated arbitrary well by mutually completely predictable copulas with respect to d_∞. In other words, d_∞ does not 'distinguish between different types of statistical dependence' (see [16]) and dependence measures which are continuous w.r.t. d_∞ like Schweizer and Wolff's σ (see [19] and [23]) seem somehow unnatural.

Using the one-to-one correspondence between copulas and Markov operators on $L^1([0,1])$ allows to consider the topology $\mathcal{O}_\mathcal{M}$ on $\mathcal{C}$ which is induced by the strong operator topology on the space $\mathcal{M}$ of Markov operators (see [4], [16], [20]). Since the topology that the weak operator topology on $\mathcal{M}$ induces on $\mathcal{C}$ coincides with the topology induced by d_∞ (see [20]) it is straightforward to see that $\mathcal{O}_\mathcal{M}$ is finer than $\mathcal{O}_{d_\infty}$. Rewriting the Markov operators in terms of regular conditional distributions (Markov kernels) we will define a L^1-type metric D_1 on $\mathcal{C}$ that is based on the conditional distribution functions and show that (i) D_1 is a metrization of $\mathcal{O}_\mathcal{M}$ and that (ii) the metric space $(\mathcal{C}, D_1)$ is complete and separable. This notion of convergence induced by D_1 can be regarded both as the asymmetric version of the so-called ∂-convergence by Mikusinski and Taylor (see [17], [18]) and the asymmetric version of the Sobolev-type-metric d studied by Darsow and Olsen (see [5]) and by Siburg and Stoimenov (see [24], [25]). We will define a dependence measure $\zeta_1 : \mathcal{C} \to [0,1]$ by $\zeta_1(A) = 3\, D_1(A, \Pi)$ and show that ζ_1 exhibits various good properties, in particular that $\zeta_1(A) = 1$ if and only if A is a copula induced by a Lebesgue-measure-preserving transformation S on $[0,1]$, i.e. if $Y = S(X)$ holds almost surely (X, Y being random variables with $\mathcal{P}^{X \otimes Y} = \mu_A$). Consequently, in contrast to d_∞, all completely dependent copulas have maximum D_1-distance to Π and Π can not be approximated by such copulas w.r.t. D_1. The interrelation between ζ_1 and the mutual dependence measure ω by Siburg and Stoimenov (see [24], [25]) will be analyzed. Furthermore we will give some examples and calculate the dependence measure ζ_1 for the Farlie-Gumbel-Morgenstern family, for the Marshall-Olkin family and the Frechet family of copulas. Finally, using completeness of $(\mathcal{C}, D_1)$, we will show that the construction of copulas with fractal support given in [10] also works w.r.t. the stronger metric D_1 instead of d_∞.

2. Notation and preliminaries

Throughout the whole paper $\mathcal{C}$ will denote the family of all *two-dimensional copulas*. For every copula $A \in \mathcal{C}$ the corresponding doubly stochastic measure will be denoted by μ_A, the family of all these μ_A by $\mathcal{P}_\mathcal{C}$. For every $A \in \mathcal{C}$ A^T will denote the transposed copula, defined by $A^T(x,y) := A(y,x)$ for all $(x,y) \in [0,1]^2$, M will denote the mimimum copula, Π the product copula and W the lower Fréchet-Hoeffding bound. For properties of copulas see [8] and [19]. d_∞ will denote the uniform metric on $\mathcal{C}$, i.e.

$$d_\infty(A,B) := \max_{(x,y)\in[0,1]^2} |A(x,y) - B(x,y)|.$$

$\mathcal{B}(\mathbb{R}^d)$ denotes the Borel σ-field in $\mathbb{R}^d$, $\mathcal{B}([0,1])$ the Borel σ-field in $[0,1]$, λ^d the d-dimensional Lebesgue measure and λ the Lebesgue measure on $[0,1]$. If X, Y are real-valued random variables on a probability space $(\Omega, \mathcal{A}, \mathcal{P})$ then we will write $\mathcal{P}^{X\otimes Y}$ for their joint distribution and $\mathcal{P}^X, \mathcal{P}^Y$ for the distributions of X and Y. $\mathbf{E}(Y|X)$ will denote the *conditional expectation of Y given X*. Since by definition $\mathbf{E}(Y|X)$ is $\mathcal{A}_\sigma(X)$-measurable there exists a measurable function $g : \mathbb{R} \to \mathbb{R}$ such that $\mathbf{E}(Y|X) = g \circ X$ holds $\mathcal{P}$-almost surely; we will write $\mathbf{E}(Y|X = x) = g(x)$ and call g a *version of the conditional expectation of Y given X*. A measurable function $g : \mathbb{R} \to \mathbb{R}$ is a version of the conditional expectation of Y given X if and only if

$$\int_B g(x) d\mathcal{P}^X = \int_{X^{-1}(B)} Y d\mathcal{P} \tag{1}$$

holds for every $B \in \mathcal{B}(\mathbb{R})$. A *Markov kernel* from $\mathbb{R}$ to $\mathcal{B}(\mathbb{R})$ is a mapping $K : \mathbb{R} \times \mathcal{B}(\mathbb{R}) \to [0,1]$ such that $x \mapsto K(x,B)$ is measurable for every fixed $B \in \mathcal{B}(\mathbb{R})$ and $B \mapsto K(x,B)$ is a probability measure for every fixed $x \in \mathbb{R}$. A Markov kernel $K : \mathbb{R} \times \mathcal{B}(\mathbb{R}) \to [0,1]$ is called *regular conditional distribution of Y given X* if for every $B \in \mathcal{B}(\mathbb{R})$

$$K(X(\omega), B) = \mathbb{E}(\mathbf{1}_B \circ Y|X)(\omega) \tag{2}$$

holds $\mathcal{P}$-a.s. It is well know that for each pair (X,Y) of real-valued random variables a regular conditional distribution $K(\cdot,\cdot)$ of Y given X exists, that

$K(\cdot,\cdot)$ is unique $\mathcal{P}^X$-a.s. (i.e. unique for $\mathcal{P}^X$-almost all $x \in \mathbb{R}$) and that $K(\cdot,\cdot)$ only depends on $\mathcal{P}^{X\otimes Y}$. Hence, given $A \in \mathcal{C}$ we will denote (a version of) the regular conditional distribution of Y given X by $K_A(\cdot,\cdot)$ and refer to $K_A(\cdot,\cdot)$ simply as *regular conditional distribution of* A. Note that for every $A \in \mathcal{C}$, its conditional regular distribution $K_A(\cdot,\cdot)$, and Borel sets $E, F \in \mathcal{B}([0,1])$ we have

$$\int_F K_A(x,E)\, d\lambda(x) = \mu_A(F \times E), \tag{3}$$

so in particular

$$\int_{[0,1]} K_A(x,E)\, d\lambda(x) = \lambda(E). \tag{4}$$

For more details and properties of conditional expectation and regular conditional distributions see [14], [15], [2], [3].
A linear operator T on $L^1([0,1],\mathcal{B}([0,1]),\lambda)$ is called *Markov operator* (see [4],[16], [20]) if it fulfils the following three properties:

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

The class of all Markov operators on $L^1([0,1],\mathcal{B}([0,1]),\lambda)$ 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. According to [4] and [20] *there is a one-to-one correspondence between* $\mathcal{C}$ *and* $\mathcal{M}$ - in fact, the mappings $\Phi : \mathcal{C} \to \mathcal{M}$ and $\Psi : \mathcal{M} \to \mathcal{C}$, defined by

$$\begin{aligned} \Phi(A)(f)(x) &:= (T_A f)(x) := \frac{d}{dx}\int_{[0,1]} A_{,2}(x,t)f(t)d\lambda(t), \\ \Psi(T)(x,y) &:= A_T(x,y) := \int_{[0,x]} (T\mathbf{1}_{[0,y]})(t)d\lambda(t) \end{aligned} \tag{5}$$

for every $f \in L^1([0,1])$ and $(x,y) \in [0,1]^2$ ($A_{,2}$ denoting the partial derivative w.r.t. y), fulfil $\Psi \circ \Phi = id_{\mathcal{C}}$ and $\Phi \circ \Psi = id_{\mathcal{M}}$. Note that in case of $f := \mathbf{1}_{[0,y]}$ we have $(T_A\mathbf{1}_{[0,y]})(x) = A_{,1}(x,y)$ λ-a.s. (the a.s. existence of the partial derivative follows from the fact that for every fixed y the mapping $x \mapsto A(x,y)$ is absolutely continuous since copulas are Lipschitz continuous, see

[19], [22], [12]). According to [16] the Markov operator T_A is a version of the conditional expectation of $f \circ Y$ given X, i.e.

$$(T_A f)(x) = \mathbb{E}(f \circ Y | X = x) \tag{6}$$

holds λ-a.s. Since this result is not proved in all generality in [16] we will start with a proof in the next section. It has been shown in [20] that $\lim_{n\to\infty} d_\infty(A_n, A) = 0$ if and only if $\lim_{n\to\infty} T_n = T$ in the weak operator topology. Using (5) the strong operator topology (see [21]) on $\mathcal{M}$ induces a topology $\mathcal{O}_\mathcal{M}$ on the $\mathcal{C}$. The metric D_1 we will define in the next section is a metrization of $\mathcal{O}_\mathcal{M}$. We will show amongst other things that the resulting metric space $(\mathcal{C}, D_1)$ is complete and separable.

3. The metric space $(\mathcal{C}, D_1)$

As mentioned before we will start with the following result (already mentioned in [16] and [17]):

Lemma 1. *Suppose that $A \in \mathcal{C}$, let the Markov operator $T_A = \Phi(A)$ be defined according to (5), denote a conditional regular distribution of A by K_A and suppose that X, Y are random variables with distribution μ_A. Then for every $f \in L^1([0,1], \mathcal{B}([0,1]), \lambda)$ the function $T_A f$ is a version of the conditional expectation of $f \circ Y$ given X, i.e. the following equality holds:*

$$(T_A f)(x) = \mathbb{E}(f \circ Y | X = x) = \int_{[0,1]} f(y) K_A(x, dy) \qquad \lambda\text{-}a.s. \tag{7}$$

Proof: (I) We will use equality (1) and start with $f := \mathbf{1}_E$, $E \in \mathcal{B}([0,1])$. As first step consider $B = [\underline{b}, \overline{b}] \subseteq [0,1]$. Using the fact that the function g_f, defined by

$$g_f(x) := \int_{[0,1]} A_{,2}(x,t) f(t) d\lambda(t),$$

according to [20] is Lipschitz continuous (therefore absolutely continuous) and monotonic we get

$$\begin{aligned} L(B) := \int_B (T_A f)(x) d\lambda(x) &= \int_B \frac{\partial}{\partial x} g_f(x) d\lambda(x) = g_f(\overline{b}) - g_f(\underline{b}) \\ &= \int_E \frac{\partial}{\partial y}\Big(A(\overline{b}, t) - A(\underline{b}, t)\Big) d\lambda(t) \end{aligned}$$

$$
\begin{aligned}
&= \mu_A\big((\underline{b},\overline{b}] \times E\big) = \mu_A\big([\underline{b},\overline{b}] \times E\big) \\
&= \mathcal{P}(X \in [\underline{b},\overline{b}],\, Y \in E) \\
&= \int_{X^{-1}(B)} f \circ Y d\mathcal{P} =: R(B)
\end{aligned}
$$

Interpreting L and R as finite (positive) measure on $([0,1],\mathcal{B}[0,1])$ (the conditions are easily verified) it follows that L and R coincide on $\mathcal{B}([0,1])$ since the class of intervals generates $\mathcal{B}([0,1])$, is closed w.r.t. intersection and monotonically reaches $[0,1]$ (see [15]). Consequently $T_A f$ is a version of the conditional distribution of $f \circ Y$ given X. (II) For the general case we can proceed in the usual way: Since L and R are linear and positive in f we immediately get (7) for non-negative step functions. Using the fact that for every non-negative $f \in L^1([0,1],\mathcal{B}([0,1]),\lambda)$ we can find a sequence of non-negative step functions monotonically converging to f together with the properties of the Lebesgue integral and continuity of T_A we get the desired result for $L^1_+([0,1],\mathcal{B}([0,1]),\lambda)$. The final step to $L^1([0,1],\mathcal{B}([0,1]),\lambda)$ is clear by positivity of the Markov operator and linearity/positivity of conditional expectation. Finally, applying disintegration (see [14]) proves the second part of the equality. ■

The next step is to express convergence of the Markov operators in the strong operator topology in terms of the corresponding regular conditional distributions.

Lemma 2. *Suppose that $A, A_1, A_2, \ldots$ are copulas, let $T, T_1, T_2 \ldots$ denote the corresponding Markov operators and $K, K_1, K_2 \ldots$ the corresponding regular conditional distributions. Then the following assertions hold:*

(i) $\lim_{n\to\infty} T_n = T$ *in the strong operator topology on* $L^1([0,1],\mathcal{B}([0,1]),\lambda)$ *if and only if for every Borel set* $B \in \mathcal{B}([0,1])$ *we have*

$$\lim_{n\to\infty} \big\|K_n(\,\cdot\,,B) - K(\,\cdot\,,B)\big\|_1 = 0.$$

(ii) Suppose that Γ *is a countable dense set in* $[0,1]$. *Then* $\lim_{n\to\infty} T_n = T$ *in the strong operator topology on* $L^1([0,1],\mathcal{B}([0,1]),\lambda)$ *if and only if for every set* $B = [0,\gamma], \gamma \in \Gamma$, *we have*

$$\lim_{n\to\infty} \big\|K_n(\,\cdot\,,B) - K(\,\cdot\,,B)\big\|_1 = 0. \tag{8}$$

Proof: Suppose that $\lim_{n\to\infty} T_n = T$ in the strong operator topology on $L^1([0,1],\mathcal{B}([0,1]),\lambda)$ and that $B \in \mathcal{B}([0,1])$. Then, using Lemma 1 and setting $f := \mathbf{1}_B$ we get

$$\begin{aligned}\|K_n(\,\cdot\,,B) - K(\,\cdot\,,B)\|_1 &= \int_{[0,1]} |K_n(x,B) - K(x,B)|\, d\lambda(x) \\ &= \|T_n f - Tf\|_1 \longrightarrow 0 \quad \text{for } n \to \infty,\end{aligned}$$

which proves one implication in (i) and (ii). It suffices to prove the other implication in (ii). Suppose that Γ is as in (ii) and that (8) holds for all sets B of the form $B = [0,\gamma]$, $\gamma \in \Gamma$. According to [9] (Theorem 2.29) convergence of T_n to T with respect to the strong operator topology on $L^1([0,1],\mathcal{B}([0,1]),\lambda)$ follows if we have $\|T_n f - Tf\|_1 \longrightarrow 0$ for every $f = \mathbf{1}_{[a,b]}$ with $a,b \in \Gamma$ since the linear hull of these function is dense in $L^1([0,1],\mathcal{B}([0,1]),\lambda)$. Let $f = \mathbf{1}_{[a,b]}$ with $a,b \in \Gamma$, then

$$K_n(\,\cdot\,,[a,b]) = K_n(\,\cdot\,,[0,b]) - K_n(\,\cdot\,,[0,a]) + K_n(\,\cdot\,,\{a\})$$

for every $n \in \mathbb{N}$ and for K instead of K_n. For the last term we get

$$\int_{[0,1]} K_n(x,\{a\}) d\lambda(x) = \lambda(\{a\}) = \int_{[0,1]} K(x,\{a\})\, d\lambda(x) = 0$$

so $K_n(x,\{a\}) = K(x,\{a\}) = 0$ λ-a.s. Hence, using the triangle inequality, we get $\|T_n f - Tf\|_1 \longrightarrow 0$, which completes the proof since $a,b \in \Gamma$ were arbitrary. ■

Motivated by Lemma 1 and Lemma 2 it seems natural to consider the following metrics on $\mathcal{C}$:

$$D_\infty(A,B) := \sup_{y\in[0,1]} \int_{[0,1]} \big|K_A(x,[0,y]) - K_B(x,[0,y])\big| d\lambda(x) \tag{9}$$

$$D_1(A,B) := \int_{[0,1]} \int_{[0,1]} \big|K_A(x,[0,y]) - K_B(x,[0,y])\big| d\lambda(x)\, d\lambda(y) \tag{10}$$

Furthermore we will also use the L^2-version D_2 of D_1 to see the interrelation between D_1 and the Sobolev-type metric d considered by Darsow and Olsen (see [5]) and by Siburg and Stoimenov (see [24], [25]):

$$D_2^2(A,B) := \int_{[0,1]} \int_{[0,1]} \big(K_A(x,[0,y]) - K_B(x,[0,y])\big)^2 d\lambda(x)\, d\lambda(y) \tag{11}$$

Remark 3. Using Fubini's theorem $D_1(A, B)$ can be seen as expected L^1-distance of the conditional distribution functions.

To simplify notation we will write

$$\Phi_{A,B}(y) := \int_{[0,1]} \big|K_A(x, [0, y]) - K_B(x, [0, y])\big| d\lambda(x) \tag{12}$$

for all $A, B \in \mathcal{C}$. Before analyzing the main properties of the function $\Phi_{A,B}$ we will show that D_1, D_2 and D_∞ are metrics.

Lemma 4. *D_∞, D_1 and D_2 defined according to (9), (10) and (11), are metrics on $\mathcal{C}$.*

Proof: First of all it has to be shown that the integrand in (10) is measurable. Define H on $[0, 1]^2$ by $H(x, y) := K_A(x, [0, y])$, then H is measurable in x and non-decreasing and right-continuous in y. Fix $z \in [0, 1]$. For every $q \in \mathbb{Q} \cap [0, 1]$ define

$$A_q := \{x \in [0, 1] : H(x, q) < z\} \in \mathcal{B}([0, 1]),$$

and set

$$A := \bigcup_{q \in \mathbb{Q} \cap [0,1]} A_q \times [0, q] \in \mathcal{B}(\mathbb{R}^2).$$

Using right-continuity it is straightforward to see that $A = H^{-1}([0, z))$, from which measurability of H directly follows. Furthermore, if $D_1(A, B) = 0$ then there exists a set $\Lambda \subseteq [0, 1]^2$ with $\lambda^2(\Lambda) = 1$ such that for every $(x, y) \in \Lambda$ we have equality $K_A(x, [0, y]) = K_B(x, [0, y])$. It follows that $\lambda(\Lambda_x) = 1$ for almost every $x \in [0, 1]$. For every such x we have that the kernels coincide on a dense set, so they have to be identical. Again using disintegration (see [14]) or equation (5) shows $A = B$. The remaining properties of a metric are obviously fulfilled. The fact that D_∞ and D_2 are metrics can be shown analogously. ■

Lemma 5. *For every pair $A, B \in \mathcal{C}$ the function $\Phi_{A,B}$, defined according to (12), is Lipschitz continuous with Lipschitz constant 2 and fulfils $\Phi_{A,B}(y) \leq \min\{2y, 2(1-y)\}$ for every $y \in [0, 1]$. Moreover there exist copulas $A, B \in \mathcal{C}$ for which equality $\Phi_{A,B}(y) = \min\{2y, 2(1-y)\}$ holds for all $y \in [0, 1]$.*

Proof: Suppose that $E \in \mathcal{B}([0,1])$, then using (4) and applying Scheffé's theorem (see [6]) we get

$$\begin{aligned}\int_{[0,1]} \big|K_A(x,E) - K_B(x,E)\big| d\lambda(x) &= 2\int_G K_A(x,E) - K_B(x,E)\, d\lambda(x) \\ &\leq 2\int_{[0,1]} K_A(x,E)\, d\lambda(x) = 2\lambda(E)\end{aligned}$$

whereby $G = \{x \in [0,1] : K_A(x,E) > K_B(x,E)\}$. Since $K_A(\cdot, E^c) = 1 - K_A(\cdot, E)$ holds, considering $E = [0,y]$ implies the desired inequality. Straightforward calculations show that in case of the copulas M and W we get $\Phi_{M,W}(y) = \min\{2y, 2(1-y)\}$ for every $y \in [0,1]$.
Finally, to see Lipschitz continuity, suppose that $s > t$, then

$$\begin{aligned}|\Phi_{A,B}(s) - \Phi_{A,B}(t)| &\leq \int_{[0,1]} \big|K_A(x,(t,s]) - K_B(x,(t,s])\big|\, d\lambda(x) \\ &= 2\lambda\big((t,s]\big) = 2(s-t). \blacksquare\end{aligned}$$

Using Lemma 5 it is straightforward to show that D_1 is a metrization of $\mathcal{O}_{\mathcal{M}}$ as mentioned in the introduction:

Theorem 6. *Suppose that $A, A_1, A_2, \ldots$ are copulas and let $T, T_1, T_2, \ldots$ denote the corresponding Markov operators. Then the following four conditions are equivalent:*

(a) $\lim_{n\to\infty} D_1(A_n, A) = 0$

(b) $\lim_{n\to\infty} D_\infty(A_n, A) = 0$

(c) $\lim_{n\to\infty} \|T_n f - Tf\|_1 = 0$ *for every* $f \in L^1([0,1], \mathcal{B}([0,1]), \lambda)$

(d) $\lim_{n\to\infty} D_2(A_n, A) = 0$

Proof: For every $n \in \mathbb{N}$ define functions $f_n : [0,1] \to [0,1]$ by $f_n(y) := \Phi_{A_n,A}(y)$. Then every f_n is Lipschitz continuous with Lipschitz constant 2. Set $\|f_n\|_{C_\infty} := \max\big\{f_n(y) : y \in [0,1]\big\}$ and suppose that $f_n(y_0) = \|f_n\|_{C_\infty}$ for some $y_0 \in [0,1]$. Then the area between the graph of f_n and the x-axis (i.e. the endograph of f_n) surely has to contain the triangle Δ_L with vertices $\{(y_0 - f_n(y_0)/2\,, 0\,), (y_0, 0), (y_0, f_n(y_0))\}$ or the triangle Δ_R with vertices

$\{(y_0, 0), (y_0 + f_n(y_0)/2\,, 0), (y_0, f_n(y_0))\}$. Consequently we have

$$\|f_n\|_{C_\infty} \geq \int_{[0,1]} f_n(y)\, d\lambda(y) \geq \frac{\|f_n\|^2_{C_\infty}}{4}.$$

This shows that (a) and (b) are equivalent. Furthermore (b) implies that the sequence f_n converges uniformly to 0, from which, using Lemma 2, (c) immediately follows. Implication $(c) \Rightarrow (a)$ follows directly from Lemma 2 and Lebesgue's theorem on dominated convergence. Finally, equivalence of (a) and (d) is a direct consequence of the fact that

$$D_2^2(A, B) \leq D_1(A, B) \leq D_2(A, B)$$

holds for all $A, B \in \mathcal{C}$. ■.

Before proceeding with D_1 we will take a look at the interrelation between the above mentioned metrics, ∂-convergence analyzed by Mikusinski and Taylor (see [17], [18]), and the Sobolev-type-metric d studied by Darsow and Olsen (see [5]) as well as by Siburg and Stoimenov (see [24], [25]). It is straightforward to see that a sequence $(A_n)_{n\in\mathbb{N}}$ of copulas ∂-converges to a copula A if and only of $\lim_{n\to\infty} D_1(A_n, A) + D_1(A_n^T, A^T) = 0$. Hence the metric D_∂, defined by

$$D_\partial(A, B) := D_1(A, B) + D_1(A^T, B^T) \tag{13}$$

for all $A, B \in \mathcal{C}$, is a metrization of ∂-convergence. Furthermore it is straightforward to see that the topology $\mathcal{O}_\partial$ induced by D_∂ on $\mathcal{C}$ is finer than $\mathcal{O}_\mathcal{M}$ - in fact this is a direct consequence of Example 25 and equation (13). Moreover, Theorem 6 implies that the topology induced by the Sobolev-type metric d is exactly $\mathcal{O}_\partial$ since

$$d^2(A, B) = D_2^2(A, B) + D_2^2(A^T, B^T) \tag{14}$$

holds (using (13) this follows from [5] too).

The following lemma will be useful in Section 6:

Lemma 7. *Suppose that $A, A_1, A_2, \ldots$ are copulas with corresponding regular conditional distributions $K, K_1, K_2, \ldots$. If $K_n(x, \cdot) \to K(x, \cdot)$ weakly λ-a.s. then we have* $\lim_{n\to\infty} D_1(A_n, A) = 0$.

Proof: Let $\Lambda \subseteq [0,1]$ denote the set of all x for which the conditional distributions converge weakly and suppose that $\lambda(\Lambda) = 1$. If f is a continuous function on $[0,1]$ then we have

$$\lim_{n\to\infty} \int_{[0,1]} f(y) K_n(x, dy) = \int_{[0,1]} f(y) K(x, dy)$$

for every $x \in \Lambda$, which, using Lebesgue's theorem on dominated convergence yields

$$\lim_{n\to\infty} \|T_{A_n} f - T_A f\|_1 = 0.$$

Since the space $C_\infty([0,1])$ of all continuous functions on $[0,1]$ is dense in $L^1([0,1], \mathcal{B}([0,1]), \lambda))$ this completes the proof. ■

It is well known that $(\mathcal{C}, d_\infty)$ is a compact metric space. Since the topology induced by D_1 is strictly finer than that induced by d_∞ (see [16] or Proposition 14) we can not expect the metric space $(\mathcal{C}, D_1)$ to be compact. The next theorem, however, shows that $(\mathcal{C}, D_1)$ is still topologically rich:

Theorem 8. *The metric space $(\mathcal{C}, D_1)$ is complete and separable.*

Proof: Suppose that $(A_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $(\mathcal{C}, D_1)$. For every $n \in \mathbb{N}$ let $K_n(\cdot,\cdot)$ denote the corresponding regular conditional distribution and H_n the function on $[0,1]^2$, defined by $H_n(x,y) := K_n(x, [0,y])$. Since we have

$$\begin{aligned} D_1(A_n, A_m) &= \int_{[0,1]} \int_{[0,1]} \big|H_n(x,y) - H_m(x,y)\big| d\lambda(x)\; d\lambda(y) \\ &= \|H_n - H_m\|_{L^1([0,1]^2,\, \mathcal{B}([0,1]^2),\, \lambda^2)} \end{aligned}$$

$(H_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $L^1([0,1]^2, \mathcal{B}([0,1]^2), \lambda^2)$, so there exists a L^1-limit $H \in L^1([0,1]^2, \mathcal{B}([0,1]^2), \lambda^2)$. According to the theorem of Riesz-Fischer (see [9], [22]) we can find a subsequence $(H_{n_j})_{j\in\mathbb{N}}$ and a Borel set $\Delta \subseteq [0,1]^2$ with $\lambda^2(\Delta) = 1$ and $\lim_{j\to\infty} H_{n_j}(x,y) = H(x,y)$ for all $(x,y) \in \Delta$. W.l.o.g. we may assume that $H(x,1) = 1$ for every $x \in [0,1]$. We will show that we can find a measurable function $G : [0,1]^2 \to [0,1]$ with the following two properties: (i) $G = H$ λ^2-a.s. and (ii) $K(x,[0,y]) := G(x,y)$ is again a regular conditional distribution of a copula $A \in \mathcal{C}$.
Using Fubini's theorem (see [9], [22]) it follows that $\lambda(\Delta_y) = \lambda(\{x \in [0,1] : (x,y) \in \Delta\}) = 1$ for λ-almost all $y \in [0,1]$. Consequently we can find a

countable set $Q = \{y_1, y_2, \ldots\} \subseteq [0,1]$ with $1 \in Q$ and a set $\Lambda_0 \subseteq [0,1]$ with $\lambda(\Lambda_0) = 1$ such that $\lim_{j\to\infty} H_{n_j}(x, y_i) = H(x, y_i)$ holds for every $y_i \in Q$ and every $x \in \Lambda_0$. Again using Fubini we can find a subset $\Lambda \subseteq \Lambda_0$ such that $\lambda(\Delta_x) = \lambda(\{y \in [0,1] : (x,y) \in \Delta\}) = 1$ for every $x \in \Lambda$. Define a new function $G : [0,1]^2 \to [0,1]$ by $G(x,y) = 1$ if $y = 1$ and

$$G(x,y) := \inf_{y_i \in Q, y_i > y} H(x, y_i)\, \mathbf{1}_\Lambda(x) \,+\, \mathbf{1}_{[0,1]}(y)\, \mathbf{1}_{\Lambda^c}(x).$$

It is straightforward to see that $G(\cdot,\cdot)$ is measurable in x for fixed y and a distribution function on $[0,1]$ in y for fixed x. In particular G is measurable (same argument as in Lemma 4) and G induces a Markov kernel $K(\cdot,\cdot)$ by setting $K(x,[0,y]) := G(x,y)$ and, for every x, uniquely extending the probability measure $K(x,\cdot)$ from the class of all intervals $[0,y]$ to $\mathcal{B}([0,1])$ in the standard way (see [9], [14]).
For every fixed $x \in \Lambda$ define (measurable) functions $g_x, h_x : [0,1] \to [0,1]$ by $g_x(y) := G(x,y)$, $h_x(y) := H(x,y)$ and set $\Pi_x := \{y \in [0,1] : g_x(y) \neq h_x(y)\}$. Using monotonicity it follows that $\Pi_x \subseteq \Delta_x^c \cup \mathcal{DC}(g_x)$, whereby $\mathcal{DC}(g_x)$ denotes the (at most) countably infinite set of discontinuities of g_x. Consequently, setting $\Pi := \{(x,y) \in [0,1]^2 : G(x,y) \neq H(x,y)\}$ and again using Fubini we get

$$\lambda^2(\Pi) = \int_{[0,1]} \lambda(\Pi_x)\, d\lambda(x) = \int_\Lambda \lambda(\Pi_x)\, d\lambda(x) = 0,$$

which implies $\lim_{n\to\infty} \|H_n - G\|_{L^1([0,1]^2, \mathcal{B}([0,1]^2), \lambda^2)} = 0$. It remains to show that $K(x,[0,y])$ is a regular conditional distribution of a copula $A \in \mathcal{C}$. Fix $y \in [0,1]$, then there exists a monotonically decreasing sequence $(z_l)_{l\in\mathbb{N}}$ in Q with limit y. Applying Lebesgue's theorem on dominated convergence shows

$$\begin{aligned}\int_{[0,1]} K(x,[0,y]) d\lambda(x) &= \int_{[0,1]} G(x,y) d\lambda(x) = \lim_{i\to\infty} \int_{[0,1]} H(x,z_i) d\lambda(x) \\ &= \lim_{i\to\infty} \lim_{j\to\infty} \int_{[0,1]} H_{n_j}(x,z_i) d\lambda(x) = \lim_{i\to\infty} z_i = y\end{aligned}$$

Hence there exists a copula $A \in \mathcal{C}$ such that $K(\cdot,\cdot) = K_A(\cdot,\cdot)$. This completes the proof of the first part of the theorem.
In order to show separability we can proceed as follows: For every $n \geq 2$ define subsets $\mathcal{S}_n$ and $\mathcal{SQ}_n$ of $\mathcal{C}$ as follows: $\mathcal{S}_n$ is the class of all $B \in \mathcal{C}$ whose

mass μ_B is uniformly distributed on each rectangle R_{ij} of the form $R_{ij} = [(i-1)/n, i/n] \times [(j-1)/n, j/n]$. Denote by $\mathcal{SQ}_n$ the subset of all $B \in \mathcal{S}_n$ that also fulfil $\mu_B(R_{ij}) \in \mathbb{Q}$ for all $i, j \in \{1, \ldots, n\}$. Since $\mathcal{SQ}_n$ is countably infinity $\mathcal{SQ} := \cup_{n=2}^{\infty} \mathcal{SQ}_n \subseteq \mathcal{C}$ is countably infinite too. Using the results in [16] $\mathcal{S}_n$ is dense in $\mathcal{C}$ with respect to the strong operator topology, so, by Theorem 6, $\mathcal{S}_n$ is dense in the metric space $(\mathcal{C}, D_1)$. Fix an arbitrary $B \in \mathcal{S}_n$ and let $\varepsilon > 0$. Obviously the family $\mathcal{S}_n$ is isomorphic to the class Ω_n of all doubly stochastic matrices. According to Birkhoff's theorem on doubly stochastic matrices (see [11]) every element $M \in \Omega_n$ is the convex combination of m $(\leq n^2+1)$ permutation matrices $(P_i)_{i=1}^m$, i.e. $M = \sum_{i=1}^m \alpha_i P_i$ with $\alpha_i \geq 0$ and $\sum_{i=1}^m \alpha_i = 1$. Since $\mathbb{Q}$ is dense in $[0, 1]$ we can find a vector $(\beta_1, \ldots, \beta_m) \in \mathbb{Q}^m$ such that both $\max_{i=1\ldots m} |\alpha_i - \beta_i| < \varepsilon/(n^2+1)$ and $\sum_{i=1}^m \beta_i = 1$ holds. Returning to B this implies the existence of an element $\hat{B} \in \mathcal{SQ}_n$ such that $\max_{i,j=1\ldots m} |\mu_B(R_{ij}) - \mu_{\hat{B}}(R_{ij})| < \varepsilon/(n^2+1)$. It follows immediately that $D_1(B, \hat{B}) < \varepsilon$ and we have shown that $\mathcal{SQ}_n$ is dense in $\mathcal{S}_n$, which completes the proof. ■

4. The dependence measure ζ_1 induced by D_1

As mentioned in the introduction we want to analyze the dependence measure ζ_1 defined as a scalar times the D_1-distance to the product copula Π. Intuitively it seems natural that completely dependent copulas (in the sense mentioned in the introduction, for a precise definition see below) should be assigned maximum dependence degree since they describe a (unidirectional) deterministic interrelation between X and Y (i.e. knowing X implies knowing Y, but in general not vice versa), whereas Π describes the other extreme in which knowing X does not at all improve our a-priori-knowledge on Y. Theorem 14 states that our dependence measure ζ_1 fulfils this property.
We will start with the following definition of completely dependent copulas and afterwards give equivalent conditions justifying the name *completely dependent*:

Definition 9. A copula $A \in \mathcal{C}$ is called *completely dependent* if there exists a λ-preserving transformation $S : [0, 1] \to [0, 1]$ such that the corresponding Markov operator T_A has the form $T_A f = f \circ S$ λ-a.s. for every $f \in L^1([0, 1], \mathcal{B}([0, 1]), \lambda)$. The class of all completely dependent copulas will be denoted by $\mathcal{C}_d$. A copula is called *mutually completely dependent* if and only if $A, A^T \in \mathcal{C}_d$ holds.

Lemma 10. *Given $A \in \mathcal{C}$ the following conditions are equivalent:*

(d1) $A \in \mathcal{C}_d$

(d2) There exists a λ-preserving transformation $S : [0,1] \to [0,1]$ such that $A(x,y) = \lambda\big([0,x] \cap S^{-1}([0,y])\big)$ for all $(x,y) \in [0,1]^2$.

(d3) There exists a λ-preserving transformation $S : [0,1] \to [0,1]$ such that $K(x,E) := \mathbf{1}_E(Sx) = \delta_{Sx}(E)$ is a regular conditional distribution of A.

(d4) There exists a λ-preserving transformation $S : [0,1] \to [0,1]$ such that $\mu_A(\Gamma(S)) = 1$, whereby $\Gamma(S) := \{(x,Sx) : x \in [0,1]\} \in \mathcal{B}([0,1]^2)$ denotes the graph of S.

Proof: $(d1) \Rightarrow (d2)$: Using the interrelation between Markov operators and copulas formulated in (5) we immediately get

$$\begin{aligned} A(x,y) &= \int_{[0,x]} (T_A \mathbf{1}_{[0,y]})(z)\, d\lambda(z) = \int_{[0,x]} \mathbf{1}_{[0,y]} \circ S(z)\, d\lambda(z) \\ &= \lambda\big([0,x] \cap S^{-1}([0,y])\big) \end{aligned}$$

for all $(x,y) \in [0,1]^2$.
$(d2) \Rightarrow (d3)$: It is clear that if $S : [0,1] \to [0,1]$ is a λ-preserving transformation, then $K(x,E)$ defined as in $(d3)$ is a Markov kernel. Suppose that $X, Y : \Omega \to [0,1]$ are random variables on a probability space $(\Omega, \mathcal{A}, \mathcal{P})$ such that $\mathcal{P}^{X \otimes Y} = \mu_A$ holds. If $E, F \in \mathcal{B}([0,1])$, then, using the extension theorem for measures, we have

$$\int_{X^{-1}(F)} \mathbf{1}_E \circ Y \, d\mathcal{P} = \mathcal{P}\big(X \in F,\, Y \in E\big) = \lambda\big(F \cap S^{-1}(E)\big) = \int_F \mathbf{1}_E(Sx)\, d\lambda(x),$$

so $K(x,E) := \mathbf{1}_E(Sx) = \delta_{Sx}(E)$ is a regular conditional distribution of A.
$(d3) \Rightarrow (d1)$: Using Lemma 1 we get $(T_A f)(x) = \int_{[0,1]} f(y) K_A(x,dy) = f(Sx)$ for λ-a.s..
$(d3) \Rightarrow (d4)$: Using disintegration (see [14]) we directly get

$$\mu_A(\Gamma(S)) = \int_{[0,1]} K_A(x, (\Gamma(S))_x)\; d\lambda(x) = \int_{[0,1]} \mathbf{1}_{\{Sx\}}(Sx)\, d\lambda(x) = 1$$

$(d4) \Rightarrow (d2)$: In case the graph of S has full mass we have $K_A(x, \{Sx\}) = 1$ for λ-almost all $x \in [0,1]$. Consequently, using disintegration and again

Lemma 1 we finally get

$$\begin{aligned} A(x_0,y_0) &= \int_{[0,x_0]} \big(T_A \mathbf{1}_{[0,y_0]}\big)(z)\, d\lambda(z) = \int_{[0,x_0]} K_A(z,[0,y_0])\, d\lambda(z) \\ &= \int_{[0,x_0]} \mathbf{1}_{[0,y_0]} \circ S(z)\, d\lambda(z) = \lambda\big([0,x_0] \cap S^{-1}([0,y_0])\big). \end{aligned}$$

This completes the proof. ■

Remark 11. Lemma 10 in particular shows that $\mathcal{C}_d$ contains all *shuffles of Min*, i.e. copulas induced by interval exchange transformations on $[0,1]$ (see [7]). Point (d4) implies that Definition 9 of complete dependence is equivalent to the original one given by Lancaster (see [13] and [25]), and point (d3) that a copula A is completely dependent if and only if it is left-invertible w.r.t. the $*$-product (see [5] and [25]).

The following lemma essentially answers the question about which copulas have maximum D_1-distance to Π:

Lemma 12. *For every $A \in \mathcal{C}$ the function $\Phi_{A,\Pi}$ fulfils $\Phi_{A,\Pi}(y) \leq 2y(1-y)$ for all $y \in [0,1]$. Furthermore equality $\Phi_{A,\Pi}(y) = 2y(1-y)$ holds for every $y \in [0,1]$ if and only if A is a completely dependent copula.*

Proof: Because of $\Phi_{A,\Pi}(0) = \Phi_{A,\Pi}(1) = 0$ if suffices to consider $y \in (0,1)$. Define

$$\mathfrak{D}_y := \Big\{ f : [0,1] \to [0,1], f \text{ measurable and } \int_{[0,1]} f(x) d\lambda(x) = y \},$$

then obviously $K_A(\,\cdot\,,[0,y]) \in \mathfrak{D}_y$ for every copula $A \in \mathcal{C}$. Using Scheffé's theorem (see [6]) we have

$$\int_{[0,1]} |f(x) - y|\, d\lambda(x) = 2 \int_{E_f} (f(x) - y)\, d\lambda(x) = 2 \int_{E_f^c} (y - f(x))\, d\lambda(x) \quad (15)$$

for every $f \in \mathfrak{D}_y$, whereby $E_f := \{x \in [0,1] : f(x) > y\}$. We will show that the left hand side of (15) becomes maximal if and only if there exists a set E such that $f = \mathbf{1}_E$ λ-a.s.:
(i) If $\int_{E_f^c} f(x) d\lambda(x) > 0$ then the function H, defined by

$$H(x) := \int_{[0,x] \cap E_f^c} f(z) d\lambda(z) - \int_{[x,1] \cap E_f^c} (1 - f(z)) d\lambda(z), \qquad x \in [0,1]$$

is absolutely continuous and fulfils $H(0) \leq -(1-y)\lambda(E_f^c) < 0$ and $H(1) = \int_{E_f^c} f(x)d\lambda(x) > 0$. Consequently we can find $x_0 \in (0,1)$ such that $H(x_0) = 0$ holds. Define a new function $f^\star$ by $f^\star := f\,\mathbf{1}_{E_f} + \mathbf{1}_{E_f^c \cap [x_0,1]}$. It is straightforward to see that $f^\star \in \mathfrak{D}_y$ and, using the first equality in (15), that $\int_{[0,1]} |f(x)-y|\,d\lambda(x) < \int_{[0,1]} |f^\star(x)-y|\,d\lambda(x)$.
(ii) If $\int_{E_f^c} f(x)d\lambda(x) = 0$ but $\int_{E_f} 1 - f(x)dx > 0$ then we can proceed analogously and define a function H by

$$H(x) := \int_{[0,x]\cap E_f} f(z)d\lambda(z) - \int_{[x,1]\cap E_f} (1-f(z))d\lambda(z), \qquad x \in [0,1].$$

H is absolutely continuous and fulfils both $H(0) = -\int_E (1-f(x))d\lambda(x) < 0$ as well as $H(1) = \int_{E_f} f(x)d\lambda(x) = y > 0$, so we can find $x_0 \in (0,1)$ such that $H(x_0) = 0$ holds. Define a new function $f^\star$ by $f^\star := f\,\mathbf{1}_{E_f^c} + \mathbf{1}_{E_f \cap [x_0,1]}$. Again it is straightforward to see that $f^\star \in \mathfrak{D}_y$ and, using the second equality in (15), that $\int_{[0,1]} |f(x)-y|\,d\lambda(x) < \int_{[0,1]} |f^\star(x)-y|\,d\lambda(x)$.
In case neither (i) nor (ii) holds we immediately get $f = \mathbf{1}_{E_f}$ λ-a.s. as well as $\lambda(E_f) = y$, which in turn implies $\int_{[0,1]} |f(x)-y|\,d\lambda(x) = 2y(1-y)$. This completes the proof of the first part of Lemma 12.
If $A \in \mathcal{C}$ then according to (d3) in Lemma 10 there exists a λ-preserving transformation $S : [0,1] \to [0,1]$ such that $K(x,E) := \mathbf{1}_E(Sx) = \delta_{Sx}(E)$ is a regular conditional distribution of A. Hence

$$\begin{aligned}\Phi_{A,\Pi}(y) &= \int_{[0,1]} \left|K_A(x,[0,y]) - y\right| d\lambda(x) = \int_{[0,1]} \left|\mathbf{1}_{[0,y]}(Sx) - y\right| d\lambda(x) \\ &= \int_{[0,1]} \left|\mathbf{1}_{[0,y]}(x) - y\right| d\lambda(x) = 2y(1-y)\end{aligned}$$

holds for every $y \in [0,1]$.
To prove the other implication suppose that $A \in \mathcal{C}$, that $K_A(\,\cdot\,,\,\cdot\,)$ is a regular conditional distribution of A and that $\Phi_{A,\Pi}(y) = 2y(1-y)$ holds for every $y \in [0,1]$. It follows from the first part of the proof that for every $y \in [0,1]$ there exists a set E_y with $\lambda(E_y) = y$ and $K_A(x,[0,y]) = \mathbf{1}_{E_y}(x)$ for λ-almost every $x \in [0,1]$. Consequently we can find a measurable set $M \subseteq [0,1]$ fulfilling $\lambda(M) = 1$ such that for every $x \in M$ we have $K_A(x,[0,y]) = \mathbf{1}_{E_y}(x)$ for every $y \in [0,1] \cap \mathbb{Q}$. Define a transformation $S : [0,1] \to [0,1]$ by

$$Sx := \mathbf{1}_M(x) \inf\big\{y \in \mathbb{Q} \cap [0,1] : K_A(x,[0,y]) = 1\big\}.$$

Using right-continuity of distribution functions it follows that on M we have $K_A(x, [0, y_0]) = 1$ if and only if $Sx \leq y_0$, i.e. if $\mathbf{1}_{[0,y_0]}(Sx) = 1$. This implies that S is measurable since

$$\{x \in [0, 1] : Sx \leq y_0\} = M^c \cup \{x \in M : K_A(x, [0, y_0]) = 1\} \in \mathcal{B}([0, 1])$$

holds for every $y_0 \in [0, 1]$. Furthermore

$$\lambda^S\big([0, y_0]\big) = \lambda\big(\{x \in [0, 1] : K_A(x, [0, y_0]) = 1\big)\} = \lambda(E_{y_0}) = y_0,$$

so S is also λ-preserving. Since on M we have $K_A(x, [0, y_0]) = \mathbf{1}_{[0,y_0]}(Sx) = \delta_{Sx}([0, y_0])$ we have $K_A(x, E) = \delta_{Sx}(E)$ for every Borel set E which shows that $(x, E) \mapsto \delta_{Sx}(E)$ is a regular conditional distribution of A. Applying Lemma 10 completes the proof. ■

Using Lemma 12 and the fact that $\int_{[0,1]} 2y(1 - y)dy = 1/3$ we finally define the dependence measure $\zeta_1 : \mathcal{C} \to [0, 1]$ by

$$\zeta_1(A) := 3D_1(A, \Pi), \qquad A \in \mathcal{C}. \tag{16}$$

Remark 13. Looking back at Remark 3 the dependence measure $\zeta_1(A)$ can, up to a scalar, be interpreted as expected L^1-distance between the conditional distribution function of A and the distribution function of the uniform distribution $\mathcal{U}_{[0,1]}$.

Lemma 12 implies the following result.

Theorem 14. *Suppose that $A \in \mathcal{C}$ and let ζ_1 be defined according to (16). Then $\zeta_1(A) \in [0, 1]$. Furthermore $\zeta_1(A) = 1$ if and only if $A \in \mathcal{C}_d$, i.e. all completely dependent copulas have maximum dependence measure.*

Proposition 15. *The following assertions hold:*

(i) The family $\mathcal{C}_d$ is closed with respect to D_1.

(ii) Suppose that S_1, S_2 are λ-preserving transformations on $[0, 1]$ and let A_1, A_2 denote the corresponding completely dependent copulas. Then we have $D_2(A_1, A_2) = D_1(A_1, A_2) = \|S_1 - S_2\|_1$.

Proof: Since only completely dependent copulas have maximum D_1-distance 1/3 from Π (i) immediately follows from the fact that metrics are continuous in each argument. Point (ii) can be proved as follows:

$$\begin{aligned}
D_2^2(A_1, A_2) &= \int_{[0,1]} \int_{[0,1]} \big(\mathbf{1}_{[0,y]}(S_1 x) - \mathbf{1}_{[0,y]}(S_2 x)\big)^2 d\lambda(x) d\lambda(y) \\
&= \int_{[0,1]} \int_{[0,1]} \big|\mathbf{1}_{[0,y]}(S_1 x) - \mathbf{1}_{[0,y]}(S_2 x)\big| d\lambda(x) d\lambda(y) = D_1(A_1, A_2) \\
&= \int_{[0,1]} \int_{[0,1]} \big|\mathbf{1}_{[0,y]}(S_1 x) - \mathbf{1}_{[0,y]}(S_2 x)\big| d\lambda(y) d\lambda(x) \\
&= \|S_1 - S_2\|_1. \blacksquare
\end{aligned}$$

Remark 16. Independence of two random variables is a symmetric concept (knowing X does not change our knowledge about Y and vice versa) - nevertheless, from the authors point of view, notions 'measuring' dependence are not necessarily symmetric since in many situations the dependence structure might be strongly asymmetric as it is, for instance, the case in Example 25. Furthermore, having a unidirectional (i.e. non-mutual) dependence measure one can easily construct a mutual one (see, for instance, equation (17) below).

Remark 17. The mutual dependence measure ω studied by Siburg and Stoimenov (see [25]) is defined by

$$\omega^2(A) := 3\, d^2(A, \Pi) = 3\,(D_2^2(A, \Pi) + D_2^2(A^T, \Pi)). \tag{17}$$

Arguments analogous to the ones used in the proof of Lemma 12 show that $D_2^2(A, \Pi) \leq 1/6$ with equality if and only if $A \in \mathcal{C}_d$. Therefore, using (17) and Proposition 15 it follows immediately that $\omega(A) = 1$ if and only if A is invertible and that the class of all invertible copulas is closed in $(\mathcal{C}, d)$ (already proved in a different manner in [25]).

We will conclude this section with an example that shows the existence of λ-preserving transformations S, S_1,$S_2, \ldots$ on $[0, 1]$ such that $(S_n(x))_{n\in\mathbb{N}}$ does not converge to $S(x)$ in any point $x \in [0, 1]$ although at the same time $\lim_{n\to\infty} D_1(A_n, A) = 0$ holds.

Example 18. For every $m \in \mathbb{N}$ and $j \in \{1, \ldots, 2^{m-1}\}$ define an interval-exchange transformation (see [7]) $S_{2^{m-1}+j} : [0, 1] \to [0, 1]$ as follows (see

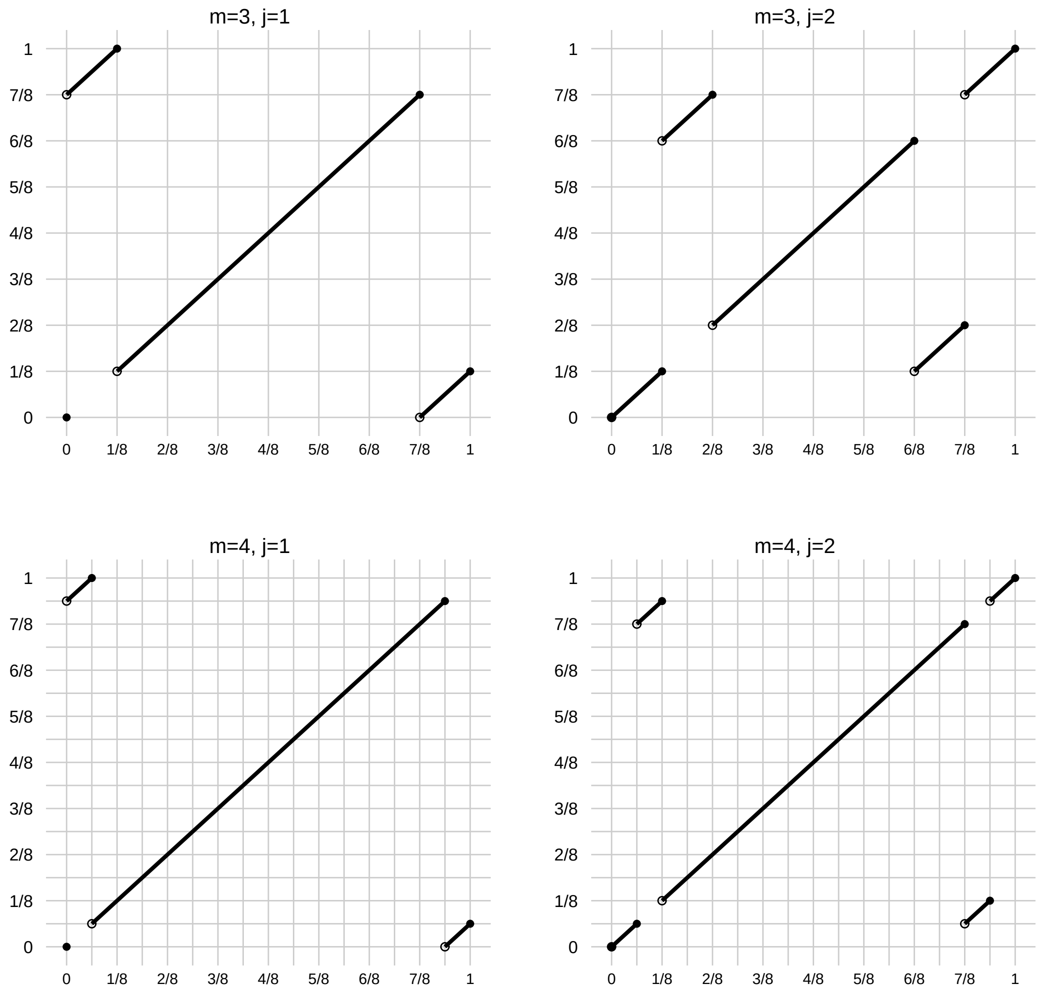

Figure 1: Interval exchange transformations used in Example 18

Figure 1):

$$S_{2^{m-1}+j}(x) = \begin{cases} x + \left(1 - \frac{2j-1}{2^m}\right) & \text{if } x \in \left(\frac{j-1}{2^m}, \frac{j}{2^m}\right] \\ x - \left(1 - \frac{2j-1}{2^m}\right) & \text{if } x \in \left(1 - \frac{j}{2^m}, 1 - \frac{j-1}{2^m}\right] \\ x & \text{otherwise} \end{cases}$$

Since every $n \in \mathbb{N}$ can uniquely be expressed in the form $n = 2^m + j$ with $m \in \mathbb{N}$ and $j \in \{1, \ldots, 2^{m-1}\}$ this defines a sequence $(S_n)_{n\in\mathbb{N}}$ of λ-preserving transformations on $[0,1]$. Let S denote the identity on $[0,1]$ and $M, A_1, A_2 \dots$ the corresponding completely dependent copulas in $\mathcal{C}_d$. Since

$$\|S_{2^{m-1}+j} - S\|_1 \leq \frac{1}{2^{m-1}}$$

holds we have $\lim_{n\to\infty} \|S_n - S\|_1 = 0$. Consequently, using Proposition 15, $\lim_{n\to\infty} D_1(A_n, A) = 0$ follows. Suppose now that $x \in (0, 1/2)$. Then for every $m \in \mathbb{N}$ there exists a unique $j_m^x \in \{1, \dots, 2^{m-1}\}$ such that $x \in \left(\frac{j_m^x - 1}{2^m}, \frac{j_m^x}{2^m}\right]$ holds. Set $\varepsilon = 1/2 - x > 0$, then it follows that

$$\lim_{m\to\infty} S_{2^{m-1}+j_m^x}(x) = \lim_{m\to\infty} \left(x + 1 - \frac{2j_m^x - 1}{2^m}\right) = x + 1 - 2x = 1 - x > x + \varepsilon,$$

which shows that $(S_n(x))_{n\in\mathbb{N}}$ can not converge to $S(x) = x$. Analogous arguments show that $(S_n(x))_{n\in\mathbb{N}}$ does not converge to $S(x) = x$ for every $x \in (0,1]$. The only two points where $(S_n)_{n\in\mathbb{N}}$ converges to S are 0 and 1. If we modify S on these two points this changes neither the induced copula M nor L^1 convergence of $(S_n)_{n\in\mathbb{N}}$ to S. Hence we have constructed a sequence $(S_n)_{n\in\mathbb{N}}$ of measure preserving transformation that converges nowhere to S.

5. Examples: ζ_1 for some parametric classes of copulas

The aim of this section is to calculate ζ_1 for some well known parametric classes of copulas.

Example 19 (Farlie-Gumbel-Morgenstern family). The FGM family $(G_\theta)_{\theta\in[-1,1]}$ of copulas is defined by (see [19])

$$G_\theta(x, y) = xy + \theta xy(1 - x)(1 - y). \tag{18}$$

G_θ is absolutely continuous so $K_\theta(\cdot, \cdot)$, defined by

$$K_{G_\theta}(x, [0, y]) := y + \theta y(1 - 2x)(1 - y) \quad \forall (x, y) \in [0, 1]^2, \tag{19}$$

is a regular conditional distribution of G_θ. Using Lemma 7 it follows immediately that the family $(G_\theta)_{\theta\in[-1,1]}$ is continuous in θ with respect to D_1. Furthermore it is straightforward to verify that $D_1(G_\theta, \Pi) = \frac{|\theta|}{12}$, so $\zeta_1(G_\theta) = \frac{|\theta|}{4}$ holds for every $\theta \in [-1, 1]$.

Example 20 (Marshall-Olkin family). The MO family $(M_{\alpha,\beta})_{(\alpha,\beta)\in[0,1]^2}$ of copulas (see [19]) is defined by

$$M_{\alpha,\beta}(x,y) = \begin{cases} x^{1-\alpha}\, y & \text{if } x^\alpha \geq y^\beta \\ x\, y^{1-\beta} & \text{if } x^\alpha \leq y^\beta. \end{cases} \tag{20}$$

It contains Π ($\alpha = 0$ or $\beta = 0$) as well as M ($\alpha = \beta = 1$). Suppose that $\alpha, \beta > 0$ then a regular conditional distribution $K_{A_{\alpha,\beta}}(\cdot,\cdot)$ of $A_{\alpha,\beta}$ is given by ($x \in (0,1], y \in [0,1]$)

$$K_{A_{\alpha,\beta}}(x,[0,y]) = \begin{cases} (1-\alpha)x^{-\alpha}\, y & \text{if } y < x^{\frac{\alpha}{\beta}} \\ y^{1-\beta} & \text{if } y \geq x^{\frac{\alpha}{\beta}}. \end{cases} \tag{21}$$

Again using Lemma 7 and the before-mentioned boundary cases it follows immediately that the family is continuous in (α,β) with respect to D_1. Straightforward but laborious calculations show that in case of $\alpha, \beta > 0$

$$\zeta_1(M_{\alpha,\beta}) = 3\alpha\,(1-\alpha)^z + \frac{6}{\beta}\,\frac{1-(1-\alpha)^z}{z} - \frac{6}{\beta}\,\frac{1-(1-\alpha)^{z+1}}{z+1} \tag{22}$$

holds, whereby $z = \frac{1}{\alpha} + \frac{2}{\beta} - 1$. Figure 2 is an image plot of the function $(\alpha,\beta) \mapsto \zeta_1(M_{\alpha,\beta})$.

Example 21 (Frechet family). The Frechet family $(F_{\alpha,\beta})$ with $(\alpha,\beta) \in [0,1]^2$ and $\alpha + \beta \leq 1$ (see [19]) is defined by

$$F_{\alpha,\beta}(x,y) := \alpha M(x,y) + \beta W(x,y) + (1-\alpha-\beta)\Pi(x,y). \tag{23}$$

Being a convex combination of the M, W and Π obviously $K_{F_{\alpha,\beta}}(\cdot,\cdot)$, defined by

$$K_{F_{\alpha,\beta}}(x,[0,y]) = \alpha \mathbf{1}_{[0,y]}(x) + \beta \mathbf{1}_{[0,y]}(1-x) + (1-\alpha-\beta)y \tag{24}$$

for all $(x,y) \in [0,1]^2$ is a regular conditional distribution of $F_{\alpha,\beta}$. As in the previous examples the family is continuous in (α,β) with respect to D_1. Furthermore it follows that

$$\zeta_1(F_{\alpha,\beta}) = \frac{1}{2}\,\frac{3\alpha^3 + 3\alpha\beta^2 + 2\beta^3}{(\alpha+\beta)^2} = \zeta_1(F_{\beta,\alpha}) \tag{25}$$

whenever $\alpha \leq \beta$ and $\alpha + \beta > 0$. In case $\alpha + \beta = 0$ we have $\zeta_1(F_{0,0}) = 0$ since $F_{0,0} = \Pi$ - which is also the limit of (25) for $\alpha, \beta \longrightarrow 0+$. Also note that for fixed $\gamma \in (0,1]$, $\alpha \in [0,\gamma]$ and $\beta = \gamma - \alpha$ the dependence measure ζ_1 becomes mimimal in case of $\alpha = \beta = \gamma/2$. Figure 3 is an image plot of the function $(\alpha,\beta) \mapsto \zeta_1(F_{\alpha,\beta})$.

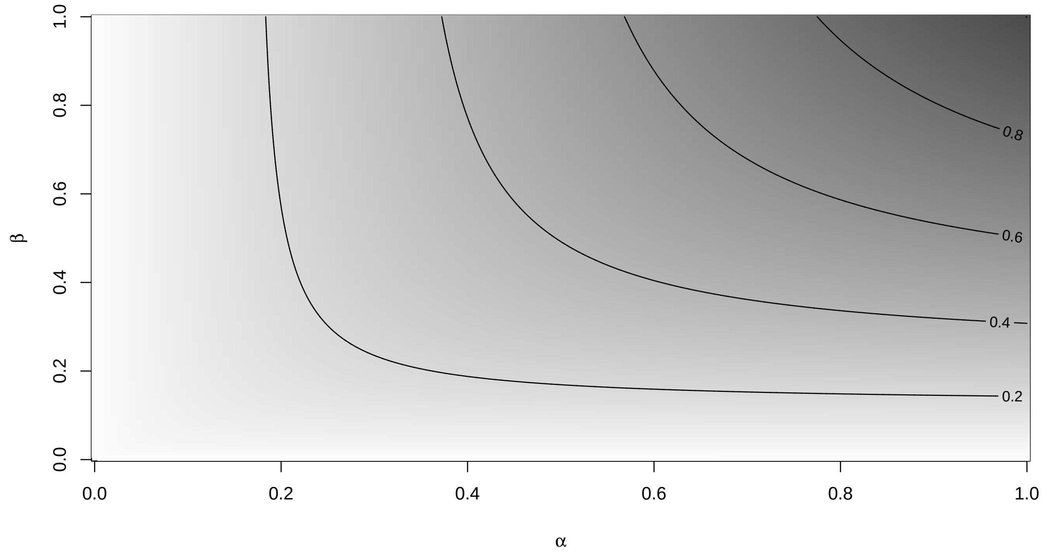

Figure 2: Image plot of the function $(\alpha, \beta) \mapsto \zeta_1(M_{\alpha,\beta})$

6. An application to copulas induced by special Iterated Function Systems

We will now take a look to the construction of copulas with fractal support via Iterated Function System given in [10] and show that the mentioned convergence results w.r.t. d_∞ also hold w.r.t. the much stronger metric D_1. Before analyzing the general case we recall the definition of an Iterated Function System (see [1]) and start with a simple example.

Definition 22. Suppose that (Ω, d) is a metric space and that $n \in \mathbb{N}$. A mapping $w : \Omega \to \Omega$ is called *contraction* if there exists a constant $L < 1$ such that $d(w(x), w(y)) \leq Ld(x, y)$ holds for all $x, y \in \Omega$. A family $(w_l)_{l=1}^n$ of contractions on Ω together with a vector $(p_l)_{l=1}^n \in [0, 1]^n$ fulfilling $\sum_{l=1}^n p_l = 1$ is called an *Iterated Function System with probabilities* (IFS for short). We will denote IFSs by $\{(w_l)_{l=1}^n, (p_l)_{l=1}^n\}$.

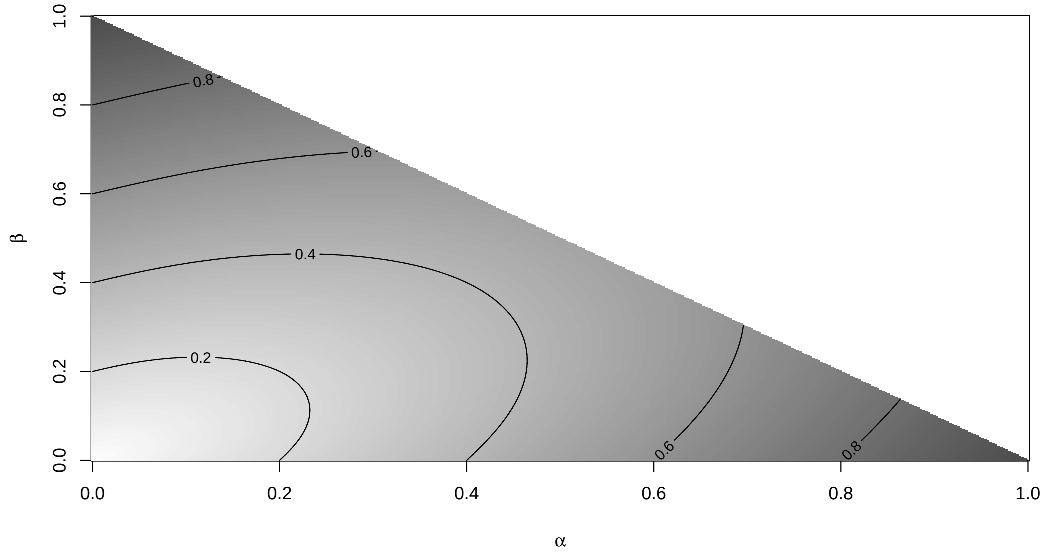

Figure 3: Image plot of the function $(\alpha, \beta) \mapsto \zeta_1(F_{\alpha,\beta})$

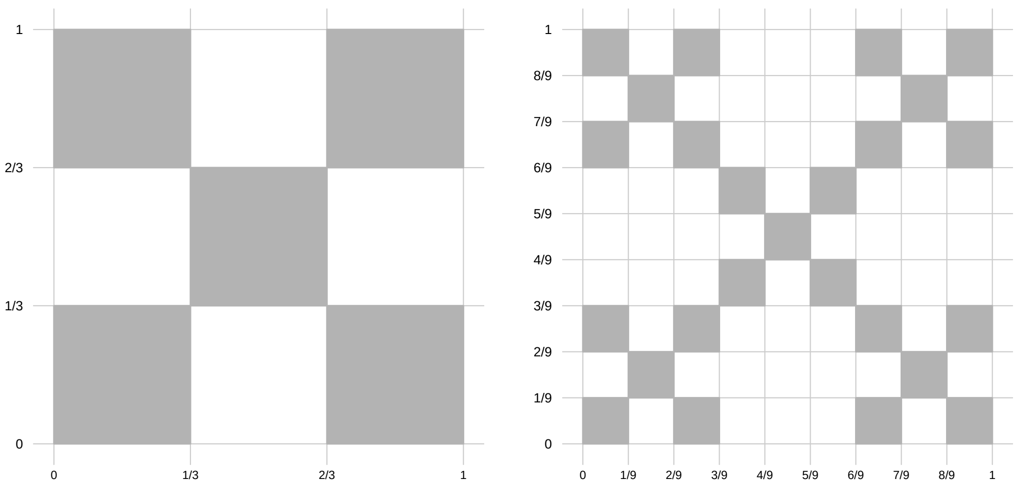

Figure 4: Support of $V\Pi$ and $V^2\Pi$ in Example 23

Example 23. Consider the matrix $M = (t_{ij})_{i,j=1}^3$ defined by

$$M = \begin{pmatrix} \frac{1}{6} & 0 & \frac{1}{6} \\ 0 & \frac{1}{3} & 0 \\ \frac{1}{6} & 0 & \frac{1}{6} \end{pmatrix},$$

set $a = b = (0, 1/3, 2/3, 1)$ and $R_{ji} := [a_{j-1}, a_j] \times [b_{i-1}, b_i]$, $1 \le i, j \le 3$. M induces an IFS $\{(w_{ji})_{i,j=1}^{3}, (t_{ji})_{i,j=1}^{3}\}$, whereby the affine contractions $w_{ji} : [0,1]^2 \to R_{ji}$, $1 \le i, j \le 3$ are defined by

$$w_{ji}(x, y) = (a_{j-1} + x(a_j - a_{j-1}) \, , \, b_{i-1} + y(b_i - b_{i-1})).$$

Let $\mathcal{P}([0,1]^2)$ denoting the set of all probability measures on $([0,1]^2, \mathcal{B}([0,1]^2))$. It straightforward to verify (see [10]) that the operator $V : \mathcal{P}([0,1]^2) \mapsto \mathcal{P}([0,1]^2)$, defined by

$$V(\mu) := \sum_{i,j=1}^{3} t_{ij} \, \mu^{w_{ji}}, \tag{26}$$

maps $\mathcal{P}_{\mathcal{C}}$ to $\mathcal{P}_{\mathcal{C}}$, so we can also see it as operator on $\mathcal{C}$ (see Figure 4). Suppose now that $A \in \mathcal{C}$, that $\mu_A \in \mathcal{P}_{\mathcal{C}}$ is the corresponding doubly stochastic measure and that $K_A(\cdot, \cdot)$ denotes a regular conditional distribution of A. It is straightforward to see that the Markov kernel $K_{VA}(\cdot, \cdot)$, defined by (27), is a regular conditional distribution of VA (again see Figure 4):

$$\begin{aligned}
y \in \Big[0, \frac{1}{3}\Big] &: \; K_{VA}\big(x, [0, y]\big) = \frac{1}{2} K_A\big(3x, [0, 3y]\big) \mathbf{1}_{[0, \frac{1}{3}]}(x) \; + \\
&\qquad\qquad + \frac{1}{2} K_A\big(3x - 2, [0, 3y]\big) \mathbf{1}_{[\frac{2}{3}, 1]}(x) \\
y \in \Big(\frac{1}{3}, \frac{2}{3}\Big] &: \; K_{VA}\big(x, [0, y]\big) = \frac{1}{2} \mathbf{1}_{[0, \frac{1}{3}] \cup (\frac{2}{3}, 1]}(x) \; + \\
&\qquad\qquad + K_A\big(3x - 1, [0, 3y - 1]\big) \mathbf{1}_{(\frac{1}{3}, \frac{2}{3}]}(x) \\
y \in \Big(\frac{2}{3}, 1\Big] &: \; K_{VA}\big(x, [0, y]\big) = \Big(\frac{1}{2} + \frac{1}{2} K_A\big(3x, [0, 3y - 2]\big)\Big) \mathbf{1}_{[0, \frac{1}{3}]}(x) + \mathbf{1}_{(\frac{1}{3}, \frac{2}{3}]}(x) \\
&\qquad\qquad + \Big(\frac{1}{2} + \frac{1}{2} K_A\big(3x - 2, [0, 3y - 2]\big)\Big) \mathbf{1}_{[\frac{2}{3}, 1]}(x)
\end{aligned} \tag{27}$$

Using (27) straightforward calculations show that for every $A, B \in \mathcal{C}$ the following relation between $\Phi_{VA,VB}$ and $\Phi_{A,B}$ holds:

$$\begin{aligned}
\Phi_{VA,VB}(3y) &= \frac{1}{3} \Phi_{A,B}(3y) \mathbf{1}_{[0, \frac{1}{3}]}(y) + \frac{1}{3} \Phi_{A,B}(3y - 1) \mathbf{1}_{(\frac{1}{3}, \frac{2}{3}]}(y) + \\
&\quad + \frac{1}{3} \Phi_{A,B}(3y - 2) \mathbf{1}_{(\frac{2}{3}, 1]}(y)
\end{aligned}$$

Hence we get

$$D_1(VA, VB) = \int_{[0,1]} \Phi_{VA,VB}(y) \, dy = 3 \frac{1}{3} \frac{1}{3} \int_{[0,1]} \Phi_{A,B}(y) \, dy = \frac{1}{3} D_1(A, B),$$

showing that V is a contraction on $(\mathcal{C}, D_1)$ with $L = 1/3$. Applying Banach's fixed point theorem and Theorem 8 it therefore follows that there is a (unique) globally attractive fixed point $A^\star \in \mathcal{C}$ of V, i.e. for every copula $B \in \mathcal{C}$ we have $D_1(V^n B, A^\star) \to 0$ for $n \to \infty$. Since convergence w.r.t. D_1 implies convergence w.r.t. d_∞ the copula $A^\star$ coincides with the fixed point w.r.t. d_∞, so $\mu_{A^\star}$ is a singular measure whose support has Hausdorff dimension $dim_H(supp(\mu_{A^\star})) = \ln(5)/\ln(3)$ (see [10]).

We will analyze the mapping $V : \mathcal{C} \to \mathcal{C}$ and its properties now in the general case. Suppose that $M = (t_{ij})_{i=1\ldots n,\, j=1\ldots m}$ is a matrix with $n \geq 2$ rows and m columns fulfilling the following three conditions: (i) All entries are non-negative, (ii) $\sum t_{ij} = 1$, and (iii) no row or column has all entries 0. According to [10] we will call such a matrix M *transformation matrix*. Given M we define the vectors $(a_j)_{j=0}^m, (b_i)_{i=0}^n$ of cumulative column and row sums by

$$\begin{aligned} a_0 &= b_0 = 0 \\ a_j &= \sum_{j_0 \leq j} \sum_{i=1}^{n} t_{ij} \quad j \in \{1, \ldots, m\} \\ b_i &= \sum_{i_0 \leq i} \sum_{j=1}^{m} t_{ij} \quad i \in \{1, \ldots, n\}. \end{aligned} \tag{28}$$

Since M is a transformation matrix both $(a_j)_{j=0}^m$ and $(b_i)_{i=0}^n$ are strictly increasing. Consequently $R_{ji} := [a_{j-1}, a_j] \times [b_{i-1}, b_i]$ are compact non-empty rectangles for every $j \in \{1, \ldots, m\}$ and $i \in \{1, \ldots, n\}$. Consider the IFS $\{(w_{ji})_{j=1\ldots m, i=1\ldots n}, (t_{ij})_{j=1\ldots m, i=1\ldots n}\}$, whereby the contraction $w_{ji} : [0,1]^2 \to R_{ji}$ is defined by

$$w_{ji}(x, y) = \big(a_{j-1} + x(a_j - a_{j-1})\,,\, b_{i-1} + x(b_i - b_{i-1})\big).$$

The induced operator V on $\mathcal{P}([0,1]^2)$ is defined by

$$V(\mu) := \sum_{j=1}^{m} \sum_{i=1}^{n} t_{ij}\, \mu^{w_{ji}}. \tag{29}$$

Again it is straightforward too see that V maps $\mathcal{P}_\mathcal{C}$ into itself (see [10]). Fix an arbitrary $A \in \mathcal{C}$ and let K_A denote a regular conditional distribution of

A. Then $K_{VA}(\cdot,\cdot)$ is given by (empty sums are zero by definition)

$$K_{VA}(x,[0,y]) := \frac{\sum_{i_0<i} t_{i_0 j}}{\sum_{i_0=1}^{n} t_{i_0 j}} + \frac{t_{ij}}{\sum_{i_0=1}^{n} t_{i_0 j}} K_A\Big(\frac{x-a_{j-1}}{a_j-a_{j-1}}\,,\,\Big[0,\frac{y-b_{i-1}}{b_i-b_{i-1}}\Big]\Big) \quad (30)$$

for every $x,y \in R_{ji} = [a_{j-1},a_j] \times [b_{i-1},b_i]$ - we will use the smallest index j and the greatest index i such that $(x,y) \in R_{ji}$ to assures that K_{VA} is well-defined also on the intersections of the rectangles and to make sure that $y \mapsto K_{VA}(x,[0,y])$ is a distribution function for every $x \in [0,1]$. Suppose now that $A,B \in \mathcal{C}$ and that $y \in (b_{i-1},b_i)$, then the following interrelation between $\Phi_{VA,VB}(y)$ and $\Phi_{A,B}(y)$ holds:

$$\begin{aligned}
\Phi_{VA,VB}(y) &= \int_{[0,1]} \big|K_{VA}(x,[0,y]) - K_{VB}(x,[0,y])\big| d\lambda(x) \\
&= \sum_{j=1}^{m} \int_{[a_{j-1},a_j]} \frac{t_{ij}}{a_j-a_{j-1}} \Big| K_A\Big(\frac{x-a_{j-1}}{a_j-a_{j-1}}\,,\,\Big[0,\frac{y-b_{i-1}}{b_i-b_{i-1}}\Big]\Big) - \\
&\qquad\qquad K_B\Big(\frac{x-a_{j-1}}{a_j-a_{j-1}}\,,\,\Big[0,\frac{y-b_{i-1}}{b_i-b_{i-1}}\Big]\Big)\Big| \, d\lambda(x) \\
&= \sum_{j=1}^{m} t_{ij} \int_{[0,1]} \Big| K_A\Big(x,\Big[0,\frac{y-b_{i-1}}{b_i-b_{i-1}}\Big]\Big) - K_B\Big(x,\Big[0,\frac{y-b_{i-1}}{b_i-b_{i-1}}\Big]\Big)\Big| \, d\lambda(x) \\
&= \sum_{j=1}^{m} t_{ij} \Phi_{A,B}\Big(\frac{y-b_{i-1}}{b_i-b_{i-1}}\Big) = (b_i-b_{i-1})\Phi_{A,B}\Big(\frac{y-b_{i-1}}{b_i-b_{i-1}}\Big)
\end{aligned}$$

Since, according to Lemma 5, $\Phi_{A,B}$ is Lipschitz continuous on $[0,1]$ and zero on $\{0,1\}$ it follows that

$$\Phi_{VA,VB}(y) = \sum_{i=1}^{n} (b_i-b_{i-1})\Phi_{A,B}\Big(\frac{y-b_{i-1}}{b_i-b_{i-1}}\Big)\mathbf{1}_{(b_{i-1},b_i]}(y)$$

for all $y \in [0,1]$. Hence

$$\begin{aligned}
D_1(VA,VB) &= \sum_{i=1}^{n} \int_{(b_{i-1},b_i]} (b_i-b_{i-1})\Phi_{A,B}\Big(\frac{y-b_{i-1}}{b_i-b_{i-1}}\Big)\, d\lambda(y) \\
&= \sum_{i=1}^{n} (b_i-b_{i-1})^2 \int_{(0,1]} \Phi_{AB}(y)\, d\lambda(y) \\
&= \sum_{i=1}^{n} (b_i-b_{i-1})^2\, D_1(A,B),
\end{aligned}$$

which shows that V is a contraction on $(\mathcal{C}, D_1)$ since $\sum_{i=1}^{n}(b_i - b_{i-1})^2 < \sum_{i=1}^{n}(b_i - b_{i-1}) = 1$. Since M was an arbitrary transformation matrix we have proved the following result (see [10] for the analogous result with respect to the uniform distance d_∞):

Theorem 24. *Suppose that M is a transformation matrix and let the operator V be defined according to (29). Then V is a contraction on the metric space $(\mathcal{C}, D_1)$ and there exists a unique copula $A^\star$ such that $VA^\star = A^\star$ and for every $B \in \mathcal{C}$ we have $\lim_{n\to\infty} D_1(V^n B, A^\star) = 0$.*

Example 25. For every $n \in \mathbb{N}_0$ define λ-preserving transformations $S_n : [0,1] \to [0,1]$ by

$$S_n(x) = 2^n x \, (mod 1)$$

and denote the corresponding completely dependent copulas by A_n. Since $A_n \in \mathcal{C}_d$ we have $D_1(A_n, \Pi) = 1/3$. Consider the transformation matrix M defined by

$$M = \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}$$

and let V denote the corresponding operator defined according to (29). Then it follows that

$$D_1(A_n^T, \Pi) = D_1(V^n A_0^T, V^n \Pi) = \frac{1}{2^n} D_1(M, \Pi) = \frac{1}{2^n} \frac{1}{3}$$

which shows that $\lim_{n\to\infty} D_1(A_n^T, \Pi) = 0$.

7. Conclusion and future work

We have introduced a metric D_1 on the space $\mathcal{C}$ that is a metrization of the topology $\mathcal{O}_\mathcal{M}$ induced by the strong operator topology on the space $\mathcal{M}$ of corresponding Markov operators. It has been shown that the metric space $(\mathcal{C}, D_1)$ is complete and separable and that the family $\mathcal{C}_d$ of completely dependent copulas is a closed subset of $\mathcal{C}$ having maximum D_1-distance to the product copula Π. As a consequence ζ_1 assigns all elements in $\mathcal{C}_d$ maximum dependence measure one. ζ_1 has been calculated for three parametric families of copulas and an application to copulas induced by special Iterated Functions Systems has been given.
As future work it seems reasonable to explore further properties of the dependence measure ζ_1 and the metric spaces $(\mathcal{C}, D_1)$ and $(\mathcal{C}, D_2)$ in general. In

particular it should be analyzed how well Π can be approximated by copulas induced by n λ-preserving transformations on $[0,1]$.

[1] M.F. Barnsley: *Fractals everywhere*, Academic Press, Cambridge, 1993

[2] R. Bhattacharya, E.C. Waymire: *A Basic Course in Probability Theory*, Springer Verlag New York 2007

[3] H. Bauer: *Wahrscheinlichkeitstheorie*, de Gruyter, Berlin New York 2002 (Fifth Edition)

[4] W.F. Darsow, B. Nguyen, E.T. Olsen: Copulas and Markov processes, *Illinois Journal of Mathematics* **36**, Number 4, 600-642 (1992)

[5] W.F. Darsow, E.T. Olsen: Norms for copulas, *International Journal of Mathematics and Mathematical Sciences* **18**, Number 3, 417-436 (1995)

[6] L. Devroye: *A Course in Density Estimation*, Birkhäuser, Boston Basel Stuttgart, 1987

[7] F. Durante, P. Sarkoci, C. Sempi: Shuffles of copulas, *Journal of Mathematical Analysis and Applications* **352**, 914-921 (2009)

[8] F. Durante, C. Sempi: Copula theory: an introduction, in P. Jaworski, F. Durante, W. Härdle, T. Rychlik, Eds., *Copula theory and its applications, Lecture Notes in Statistics*–Proceedings, vol 198, pp. 1–31, Springer, Berlin (2010)

[9] J. Elstrodt: *Maß- und Integrationstheorie*, Springer Verlag, Berlin Heidelberg New York, 1999

[10] G.A. Fredricks, R.B. Nelsen, J.A. Rodríguez-Lallena: Copulas with fractal supports, *Insur. Math. Econ.* **37** (2005) 42-48

[11] P.M. Gruber: *Convex and Discrete Geometry*, Springer-Verlag Berlin Heidelberg, 2007

[12] E. Hewitt, K. Stromberg: *Real and Abstract Analysis*, Springer Verlag, Berlin Heidelberg, 1965

[13] H.O. Lancaster, Correlation and complete dependence of random variables, *Annals of Mathematical Statistics* **34**, 1315-1321 (1963)

[14] O. Kallenberg: *Foundations of modern probability*, Springer Verlag, New York Berlin Heidelberg, 1997

[15] A. Klenke: *Probability Theory - A Comprehensive Course*, Springer Verlag Berlin Heidelberg 2007

[16] X. Li, P. Mikusinski, M.D. Taylor: Strong approximation of copulas, *Journal of Mathematical Analysis and Applications* **255**, 608-623 (1998)

[17] P. Mikusinski, M.D. Taylor: Markov operators and n-copulas, *Annales Polonici Mathematici* **96**, Nr. 1, 75-95 (2009)

[18] P. Mikusinski, M.D. Taylor: Some approximations of n-copulas, *Metrika* **72**, 385-414 (2010)

[19] R.B. Nelsen: *An Introduction to Copulas*, Springer, New York, 2006

[20] E.T. Olsen, W.F. Darsow, B. Nguyen: Copulas and Markov operators, in *Proceedings of the Conference on Distributions with Fixed Marginals and Related Topics, IMS Lecture Notes*, Monograph Series Vol. 28, pp. 244-259 (1996)

[21] W. Rudin: *Functional Analysis*, McGraw-Hill International Editions, Singapore, 1991

[22] W. Rudin: *Real and Complex Analysis*, McGraw-Hill International Editions, Singapore, 1987

[23] B. Schweizer, E. F. Sklar: On nonparametric measures of dependence for random variables, *Annals of Statistics* **9**, No. 4, 879-885 (1981)

[24] K.F. Siburg, P.A. Stoimenov: A scalar product for copulas, *Journal of Mathematical Analysis and Applications* **344**, 429-439 (2008)

[25] K.F. Siburg, P.A. Stoimenov: A measure of mutual complete dependence, *Metrika* **71**, 239-251 (2010)