Journal of Theoretical Probability manuscript No.
(will be inserted by the editor)

Conditioning based metrics on the space of multivariate copulas, their interrelation with uniform and levelwise convergence and Iterated Function Systems

Juan Fernández Sánchez · Wolfgang Trutschnig*

Received: date / Accepted: date

Abstract Using the one-to-one correspondence between copulas and special Markov kernels three strong metrics on the class $\mathcal{C}_\rho$ of ρ-dimensional copulas with $\rho \geq 3$ are studied. Being natural extensions of the two-dimensional versions introduced by Trutschnig (2011) these metrics exhibit various good properties. In particular it can be shown that the resulting metric spaces are separable and complete, which, as byproduct, offers a simple separable and complete metrization of the so-called ∂-convergence studied by Mikusinski and Taylor (2009, 2010). As additional consequence of completeness it is proved that the construction of singular copulas with fractal support via special Iterated Function Systems also converges with respect to any of the three introduced metrics. Moreover, the interrelation with the uniform metric d_∞ is studied and convergence with respect to d_∞ is characterized in terms of level-set and endograph convergence with respect to the Hausdorff metric.

Keywords Copula · Stochastic measure · Markov kernel · Iterated Function System · Level set · Endograph

Mathematics Subject Classification (2000) 62H20 · 60E05 · 28A80

Juan Fernández Sánchez
Grupo de Investigación de Análisis Matemático
Universidad de Almería
La Cañada de San Urbano, Almería, Spain
E-mail: juanfernandez@ual.es

Wolfgang Trutschnig
Department for Mathematics
University Salzburg
Hellbrunnerstrasse 34
5020 Salzburg, Austria
E-mail: wolfgang@trutschnig.net

*corresponding author

1 Introduction

The one-to-one correspondence between copulas and Markov operators going back to Darsow et al. (see [7], [28]) can be used to define various metrics on the space $\mathcal{C}_2$ of two-dimensional copulas whose induced topology is strictly finer than the one induced by the standard uniform metric d_∞. In [32] three metrics called D_1, D_2 and D_∞ of this kind have been introduced and studied. Among other things it has been proved (i) that all three metrics induced the same topology, (ii) that D_1 is a metrization of the strong operator topology of the corresponding Markov operators and (iii) that the resulting metric space $(\mathcal{C}_2, D_1)$ is separable and complete. D_1 has been shown to be strongly related to the so-called ∂-convergence studied by Mikusinski and Taylor (see [24], [25] as well as [23]). Additionally, using the fact that the class of completely dependent copulas has maximum D_1-distance to the product copula Π_2, D_1 induces a natural dependence measure τ_1 with various good properties.
Main content of the present paper is the study of the generalizations of D_1, D_2 and D_∞ to the space $\mathcal{C}_\rho$ of ρ-dimensional copulas for arbitrary $\rho \geq 3$. Although the multivariate case requires technically more tedious proofs we will show that many (but not all) properties valid in dimension two also hold in the ρ-dimensional setting. In particular we will prove that the metric space $(\mathcal{C}_\rho, D_1)$ is separable and complete, a result which, as direct by-product, offers a simple separable and complete metrization of ∂-convergence on $\mathcal{C}_\rho$. As additional consequence of completeness we will state a (not very restrictive and easily verifiable) condition under which the construction of copulas with fractal support via Iterated Function Systems (IFSs) also converges w.r.t. the metric D_1 for every dimension $\rho \geq 3$ and give an example illustrating that the condition is sufficient but not necessary. Although in the general setting the class of so-called completely dependent copulas (see [21], [32]) is strictly separated from the product copula Π_ρ too, we will point out why a direct and seemingly natural extension of the dependence measure τ_1 introduced in [32] to the multivariate setting yields no reasonable notion quantifying dependence. Apart from the just mentioned results we include a simple proof of the fact that every completely dependent copula can be approximated by shuffles of the Minimum copula w.r.t. the metric D_1 (see [13]), show that convergence in $(\mathcal{C}_\rho, D_1)$ implies uniform convergence but not vice versa and characterize uniform convergence in $\mathcal{C}_\rho$ in terms of level-set-convergence and endograph convergence w.r.t.t. Hausdorff metric δ_H.
We would like to emphasize that our main motivation for the study of conditioning-based metrics on $\mathcal{C}_\rho$ was the need to have at hand strong metrics that clearly distinguish between different types of statistical dependence, which the standard uniform metric does not (see [23]). Loosely speaking, given $\mathcal{U}(0,1)$-distributed random variables $X, Y_1, Y_2, \ldots, Y_{\rho'}, Z_1, Z_2, \ldots, Z_{\rho'}$, the introduced metrics measure how different the conditional distributions of $(Y_1, Y_2, \ldots, Y_{\rho'})$ and $(Z_1, Z_2, \ldots, Z_{\rho'})$ given X are (relevant, for instance, when analyzing how different the dependence of two portfolios on another quantity X is). In many situations it seems natural to approximate a copula $A \in \mathcal{C}_\rho$ by elements $\hat{A}_n$

from sufficiently general but easily manageable families (e.g. checkerboards) in a such a way that dependence properties of A (in our case dependence on the first coordinate) also have to be valid for the approximations $\hat{A}_n$. The facts that the metrics offer, firstly, a separable and complete metrization of ∂-convergence and, secondly, a simple way to prove that every completely dependent copula can be approximated by shuffles of the minimum copula w.r.t. ∂-convergence are nice by-products.

The rest of the paper is organized as follows: Section 2 gathers some preliminaries and notations that will be used throughout the paper. Section 3 is devoted to the weakest types of convergence studied in the paper - we prove the afore-mentioned characterization of uniform convergence in $\mathcal{C}_\rho$ in terms of level-set-convergence and endograph-convergence w.r.t.t. Hausdorff metric δ_H. In Section 4 the metrics D_1, D_2 and D_∞ on $\mathcal{C}_\rho$ are introduced, their interrelation and main properties are analyzed. The main content of Section 5 is the convergence of the IFS construction of copulas with fractal support w.r.t. any of the three new metrics; two concrete examples/graphics illustrate the approach. Finally, Section 6 contains open problems and future work.

2 Notation and preliminaries

Throughout the paper ρ will denote the dimension and ρ' is defined by $\rho' := \rho - 1$. Bold symbols will denote vectors, i.e. $\mathbf{y} = (y_1, y_2, \ldots, y_{\rho'}) \in \mathbb{R}^{\rho'}$, and, slightly misusing notation, we will also write $(x, \mathbf{y})$ for $(x, y_1, \ldots, y_{\rho'}) \in \mathbf{R}^\rho$. Analogously the notation $[\mathbf{0}, \mathbf{y}] := \times_{i=1}^{\rho'} [0, y_i]$ will be used for all $\mathbf{y} \in [0,1]^{\rho'}$. For every pair $\mathbf{y}, \mathbf{z} \in [0,1]^\rho$ we will write $y \succ z$ if and only if $y_i > z_i$ for all $i \in \{1, \ldots, \rho\}$. Furthermore we will write $\mathbf{y} \gg \mathbf{z}$ if and only if for every $i \in \{1, \ldots, \rho\}$ we have both $y_i \geq z_i$ as well as $y_i > z_i$ whenever $z_i < 1$.
For every metric space (Ω, d), every subset $E \subseteq \Omega$, and every $r \geq 0$, $B_d(E, r)$ and $\overline{B}_d(E, r)$ denote the open and closed r-neighbourhood of E respectively, i.e.

$$B_d(E, r) := \{y \in \Omega : \exists x \in E \text{ such that } d(x, y) < r\} \quad \text{and}$$

$$\overline{B}_d(E, r) := \{y \in \Omega : \exists x \in E \text{ such that } d(x, y) \leq r\}.$$

$\mathcal{K}(\Omega)$ denotes the family of all non-empty compact subsets of Ω, δ_H the Hausdorff metric on $\mathcal{K}(\Omega)$ (see, for instance, [26]), and $\mathcal{B}(\Omega)$ the Borel σ-field in Ω. $\mathcal{P}(\Omega)$ denotes the family of all probability measures on $(\Omega, \mathcal{B}(\Omega))$ and, in case of $\Omega = [0,1]^\rho$, $\rho \geq 2$, $\mathcal{P}_\mathcal{C}(\Omega)$ the class of all probability measure for which the corresponding distribution function is a copula (i.e. probability measures for which all one-dimensional marginals coincide with the Lebesgue measure λ on $[0,1]$). $\mathcal{C}_\rho$ will denote the family of all ρ-dimensional copulas, d_∞ the uniform metric on $\mathcal{C}_\rho$. It is well known that $(\mathcal{C}_\rho, d_\infty)$ is a compact metric space. For very $A \in \mathcal{C}_\rho$, μ_A will denote the corresponding element in $\mathcal{P}_\mathcal{C}([0,1]^\rho)$. λ^ρ will denote the ρ-dimensional Lebesgue measure on $[0,1]^\rho$ (in case of $\rho = 1$ we will also write λ instead of λ^1), Π_ρ and M_ρ the ρ-dimensional product and the ρ-dimensional minimum copula respectively, i.e. $\Pi_\rho(x_1, \ldots, x_\rho) = \prod_{i=1}^\rho x_i$

and $M_\rho(x_1,\ldots,x_\rho) = \min\{x_1,\ldots,x_\rho\}$. For properties of copulas we refer to [11], [27].
A *Markov kernel* from $\mathbb{R}$ to $\mathcal{B}(\mathbb{R}^{\rho'})$ is a mapping $K : \mathbb{R} \times \mathcal{B}(\mathbb{R}^{\rho'}) \to [0,1]$ such that $x \mapsto K(x,B)$ is measurable for every fixed $B \in \mathcal{B}(\mathbb{R}^{\rho'})$ and $B \mapsto K(x,B)$ is a probability measure for every fixed $x \in \mathbb{R}$. A Markov kernel $K : \mathbb{R} \times \mathcal{B}(\mathbb{R}^{\rho'}) \to [0,1]$ is called *regular conditional distribution of* $\mathbf{Y}$ *given* X ($\mathbf{Y} : \Omega \to \mathbb{R}^{\rho'}$ and $X : \Omega \to \mathbb{R}$ random variables on a probability space $(\Omega, \mathcal{A}, \mathcal{P})$) if for every $B \in \mathcal{B}(\mathbb{R}^{\rho'})$

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

holds $\mathcal{P}$-a.s. It is well know that for each random vector $(X, \mathbf{Y})$ a regular conditional distribution $K(\cdot,\cdot)$ of $\mathbf{Y}$ 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 the distribution $\mathcal{P}^{X\otimes\mathbf{Y}}$. Hence, given $A \in \mathcal{C}_\rho$ we will denote (a version of) the regular conditional distribution of $\mathbf{Y}$ given X by $K_A(\cdot,\cdot)$ and refer to $K_A(\cdot,\cdot)$ simply as *regular conditional distribution* or *Markov kernel of* A. Note that for every $A \in \mathcal{C}_\rho$, its conditional regular distribution $K_A(\cdot,\cdot)$, and a Borel set $F \in \mathcal{B}([0,1]^\rho)$ we have (with $F_x = \{\mathbf{y} \in [0,1]^{\rho'} : (x,\mathbf{y}) \in F\}$)

$$\int_{[0,1]} K_A(x, F_x)\, d\lambda(x) = \mu_A(F), \tag{2}$$

so in particular

$$\int_{[0,1]} K_A(x, F)\, d\lambda(x) = \lambda(F_{i_0}) \tag{3}$$

in case $F = \times_{i=1}^{\rho} F_i$ and $F_i = [0,1]$ for all $i \neq i_0$. On the other hand, every Markov kernel $K : [0,1] \times \mathcal{B}(\mathbb{R}^{\rho'}) \to [0,1]$ fulfilling (3) is the regular conditional distribution of a copula $A \in \mathcal{C}_\rho$. For more details and properties of conditional expectation and regular conditional distributions see [19], [20] and [2], [3].
As next step we recall the definition of an Iterated Function System (IFS) and some main results about IFSs (for more details see [1], [4], [14]). Suppose for the following that (Ω, d) is a compact metric space. A mapping $w : \Omega \to \Omega$ is called *contraction* if there exists a constant $L < 1$ such that $d(w(x), w(y)) \leq L d(x,y)$ holds for all $x, y \in \Omega$. A family $(w_l)_{l=1}^n$ of $n \geq 2$ contractions on Ω is called *Iterated Function System* (IFS for short) and will be denoted by $\{\Omega, (w_l)_{l=1}^n\}$. An IFS together with a vector $(p_l)_{l=1}^n \in [0,1]^n$ fulfilling $\sum_{l=1}^n p_l = 1$ is called *Iterated Function System with probabilities* (IFSP for short). We will denote IFSPs by $\{\Omega, (w_l)_{l=1}^n, (p_l)_{l=1}^n\}$. Every IFS(P) induces the so-called *Hutchinson operator* $\mathcal{H} : \mathcal{K}(\Omega) \to \mathcal{K}(\Omega)$, defined by

$$\mathcal{H}(Z) := \bigcup_{i \leq n:\, p_i > 0} w_i(Z). \tag{4}$$

It can be shown (see [1]) that $\mathcal{H}$ is a contraction on the compact metric space $(\mathcal{K}(\Omega), \delta_H)$, so Banach's Fixed Point theorem implies the existence of a unique, globally attractive fixed point $Z^\star$ of $\mathcal{H}$, i.e. for every $R \in \mathcal{K}(\Omega)$ we have

$$\lim_{n\to\infty} \delta_H\big(\mathcal{H}^n(R), Z^\star\big) = 0.$$

On the other hand every IFSP also induces an operator $\mathcal{V} : \mathcal{P}(\Omega) \to \mathcal{P}(\Omega)$, defined by

$$\mathcal{V}(\mu) := \sum_{i=1}^{n} p_i \, \mu^{w_i}, \tag{5}$$

whereby μ^{w_i} denotes the push-forward of μ via the contraction w_i for every $i \in \{1, \ldots, n\}$. The so-called *Hutchinson metric* h (also called Kantorovich or Wasserstein metric) on $\mathcal{P}(\Omega)$ is defined by

$$h(\mu, \nu) := \sup \left\{ \int_\Omega f \, d\mu - \int_\Omega f \, d\nu : \, f \in Lip_1(\Omega, \mathbb{R}) \right\}, \tag{6}$$

whereby $Lip_1(\Omega, \mathbb{R})$ is the class of all non-expanding functions $f : \Omega \to \mathbb{R}$, i.e. functions fulfilling $|f(x) - f(y)| \leq d(x, y)$ for all $x, y \in \Omega$. It is not difficult to show that $\mathcal{V}$ is a contraction on $(\mathcal{P}(\Omega), h)$, that h is a metrization of the topology of weak convergence on $\mathcal{P}(\Omega)$ and that $(\mathcal{P}(\Omega), h)$ is a compact metric space (see [1], [9]). Consequently, again by Banach's Fixed Point theorem, it follows that there is a unique, globally attractive fixed point $\mu^\star \in \mathcal{P}(\Omega)$ of $\mathcal{V}$, i.e. for every $\nu \in \mathcal{P}(\Omega)$ we have

$$\lim_{n \to \infty} h\big(\mathcal{V}^n(\nu), \mu^\star\big) = 0.$$

Furthermore $Z^\star$ is the support of $\mu^\star$ (again see [1], [4]).

Finally we will recall some properties of the Hausdorff metric that will be used in Section 3. $\mathcal{K}_{pc}([0,1]^\rho)$ will denote the class of all pathwise connected elements in $\mathcal{K}([0,1]^\rho)$. It is not difficult to see that in $\mathcal{K}_{pc}([0,1]^\rho)$ convergence w.r.t. δ_H coincides with *Painlevé-Kuratowski-(PK for short) convergence* of closed sets (see Proposition 12 in [31]): A sequence $(E_n)_{n \in \mathbb{N}}$ of subsets of an arbitrary metric space (Ω, d) is said to be convergent in the PK-sense if the topological limit inferior ${}^t\!\liminf_{n \to \infty} E_n$ and the topological limit superior ${}^t\!\limsup_{n \to \infty} E_n$ coincide, whereby

$$\begin{aligned} {}^t\!\liminf_{n \to \infty} E_n &:= \{x \in \Omega : \exists (x_n)_{n \in \mathbb{N}} \text{ with } \forall n \, x_n \in E_n \text{ and } \lim_{n \to \infty} x_n = x\} \\ {}^t\!\limsup_{n \to \infty} E_n &:= \{x \in \Omega : \exists (x_{n_k})_{k \in \mathbb{N}} \text{ with } \forall k \, x_{n_k} \in E_{n_k} \text{ and } \lim_{k \to \infty} x_{n_k} = x\} \end{aligned} \tag{7}$$

and $(n_k)_{k \in \mathbb{N}}$ denotes a strictly increasing sequence in $\mathbb{N}$. One additional property of the Hausdorff metric δ_H on a compact metric space (Ω, d) that we will use later is that

$$\lim_{n \to \infty} \delta_H \left(E_n, \overline{\bigcup_{n=1}^{\infty} E_n} \right) = 0 \tag{8}$$

holds for very increasing sequence $(E_n)_{n \in \mathbb{N}}$ of non-empty compact subsets of Ω ($\overline{E}$ denoting the topological closure of E). For more details on δ_H and PK-convergence see [26].

3 Uniform, levelwise and endograph convergence

In this short section we will take a closer look to convergence w.r.t. d_∞ and generalize the two-dimensional results in [33] to the ρ-dimensional setting ($\rho \geq 2$). In particular it will be proved that uniform convergence of a sequence $(A_n)_{n\in\mathbb{N}}$ of copulas to a copula $A \in \mathcal{C}_\rho$ is equivalent (i) to the convergence of the corresponding endographs, and (ii) to the convergence of the corresponding upper (or lower) α-level-sets for all but at most countably many α in $[0,1]$ (all with respect to the Hausdorff metric in the corresponding dimensions).
We start with the definition of the endograph and the upper and lower level sets of a copula $A \in \mathcal{C}_\rho$ and only consider the case $\rho \geq 3$. For every $A \in \mathcal{C}_\rho$ the endograph $end(A)$ is defined by

$$end(A) := \big\{(\mathbf{x}, t) \in [0,1]^{\rho+1} : A(\mathbf{x}) \leq t\big\}. \tag{9}$$

The mapping $A \mapsto end(A)$ obviously is an embedding of $\mathcal{C}_\rho$ in the metric space $(\mathcal{K}_{pc}([0,1]^{\rho+1}), \delta_H)$ of all non-empty compact pathwise connected subsets of $[0,1]^{\rho+1}$ endowed with the Hausdorff metric δ_H. Since the topology generated by the Hausdorff metric is independent of the concrete chosen metrization of the underlying space (see [26])) we can, w.l.o.g., work with the metric $\beta_{\rho+1}$, defined by

$$\beta_{\rho+1}\big((\mathbf{x}, t), (\mathbf{y}, s)\big) := \max\big\{\beta(\mathbf{x}, \mathbf{y}), |t - s|\big\}$$

on $[0,1]^{\rho+1}$, whereby β denotes the Euclidean metric on $[0,1]^\rho$. The before-mentioned embedding allows to consider the so-called *endograph metric* d_{end} on $\mathcal{C}_\rho$, defined by

$$d_{end}(A, B) := \delta_H(end(A), end(B)) \tag{10}$$

for all $A, B \in \mathcal{C}_\rho$. We will see that d_{end} and d_∞ are equivalent metrics.
Closely related to the endograph of a copula A is the family of its upper and lower level sets, $([A]_\alpha)_{\alpha\in[0,1]}$ and $([A]^\alpha)_{\alpha\in[0,1]}$ respectively, which are defined as

$$[A]_\alpha = \big\{\mathbf{x} \in [0,1]^\rho : A(\mathbf{x}) \geq \alpha\big\} \tag{11}$$
$$[A]^\alpha = \big\{\mathbf{x} \in [0,1]^\rho : A(\mathbf{x}) \leq \alpha\big\} \tag{12}$$

for every $\alpha \in [0,1]$. Note that due to continuity and monotonicity all these level sets are elements of the metric space $(\mathcal{K}_{pc}([0,1]^\rho), \delta_H)$. Hence the *upper and lower level functions*

$$\Phi_A(\alpha) := [A]_\alpha \quad \text{and} \quad \Psi_A(\alpha) := [A]^\alpha \tag{13}$$

map $[0,1]$ to $\mathcal{K}_{pc}([0,1]^\rho)$.

As in the two-dimensional setting (see Lemma 1 in [33]) it is straightforward to prove that a point $\alpha_0 \in (0,1)$ is a discontinuity point of Φ_A if and only if $A^{-1}(\{\alpha_0\})$ has non-empty interior and that the same holds for Ψ_A. Since, firstly, $\alpha_1 \neq \alpha_2$ implies $A^{-1}(\{\alpha_1\}) \cap A^{-1}(\{\alpha_2\}) = \emptyset$ and, secondly, $[0,1]^\rho$ is separable there can only be countably many α_0 for which $A^{-1}(\{\alpha_0\})$ has non-empty interior. Hence one directly gets the following result.

Theorem 1 *For every $A \in \mathcal{C}_\rho$ both the upper and the lower level function Φ_A and Ψ have at most countably many discontinuities and the discontinuities in $(0,1)$ are the same for both functions.*

The proofs of the next two results are essentially the same as the corresponding ones for the two-dimensional setting in [33] - we only include them for the sake of completeness.

Proposition 1 *Suppose that $(A_n)_{n\in\mathbb{N}}$ is a sequence of copulas that converges pointwise to $A \in \mathcal{C}_\rho$. If $\alpha \in (0,1)$ is a continuity point of Φ_A, then*

$$\lim_{n\to\infty} \delta_H\big([A_n]_\alpha, [A]_\alpha\big) = \lim_{n\to\infty} \delta_H\big([A_n]^\alpha, [A]^\alpha\big) = 0$$

holds.

Proof: Suppose that $\alpha \in (0,1)$ is a continuity point of Φ_A and let $\varepsilon > 0$. Then there exists $\delta \in (0,\varepsilon)$ such that $\delta_H([A]_\alpha, [A]_\beta) \le \varepsilon$ whenever $|\alpha-\beta| \le \delta$. Since pointwise convergence of $(A_n)_{n\in\mathbb{N}}$ to A implies uniform convergence (see [27]) there exists an index $n_0 \in \mathbb{N}$ with $d_\infty(A_n, A) \le \delta$ for every $n \ge n_0$. If $\mathbf{x} \in [A]_\alpha$ then we can find $\mathbf{z} \in [A]_{\alpha+\delta}$ fulfilling $\beta(\mathbf{x},\mathbf{z}) \le \varepsilon$. Consequently $A_n(\mathbf{z}) \ge \alpha$ and $\mathbf{z} \in [A_n]_\alpha$ holds for all $n \ge n_0$. Since $\mathbf{x} \in [A]_\alpha$ was arbitrary $[A]_\alpha \subseteq \overline{B}([A_n]_\alpha, \varepsilon)$ is fulfilled for every $n \ge n_0$. On the other hand, if $\mathbf{x} \in [A_n]_\alpha$ and $n \ge n_0$ then $A(\mathbf{x}) \ge \alpha - \delta$ and $\mathbf{x} \in [A]_{\alpha-\delta}$ follows. We can find $\mathbf{z} \in [A]_\alpha$ such that $\beta(\mathbf{x},\mathbf{z}) \le \varepsilon$ holds. Since $\mathbf{x} \in [A_n]_\alpha$ was arbitrary we get $[A_n]_\alpha \subset \overline{B}([A]_\alpha, \varepsilon)$. Altogether this shows that $\delta_H([A_n]_\alpha, [A]_\alpha) \le \varepsilon$ for every $n \ge n_0$. Convergence of the lower α-level sets can be proved analogously. ■

Remark 1 Examples of copulas $A, A_1, A_2, \ldots \in \mathcal{C}_\rho$ for which we have both $\delta_H([A_n]_\alpha, [A]_\alpha) \not\to 0$ for infinitely many $\alpha \in (0,1)$ and, at the same time, $\lim_{n\to\infty} d_\infty(A_n, A) = 0$ can be constructed analogously to the two-dimensional setting by using, for instance, the IFS construction.

Proposition 2 *Suppose that $A, A_1, A_2, \ldots \in \mathcal{C}_\rho$ and that one of the following two conditions is fulfilled:*

(a) There exists a set $\Lambda \subseteq [0,1]$ of Lebesgue measure 1 such that for all $\alpha \in \Lambda$ the equality $\lim_{n\to\infty} \delta_H([A_n]_\alpha, [A]_\alpha) = 0$ holds.
(b) There exists a set $\Gamma \subseteq [0,1]$ of Lebesgue measure 1 such that for all $\alpha \in \Gamma$ the equality $\lim_{n\to\infty} \delta_H([A_n]^\alpha, [A]^\alpha) = 0$ holds.

Then $\lim_{n\to\infty} d_\infty(A_n, A) = 0$ follows.

Proof: Suppose that (a) holds. Fix $\mathbf{x} \in [0,1]^\rho$ and set $\alpha = A(\mathbf{x})$. If $\alpha > 0$ and $k \in \mathbb{N}$ then there exists $\gamma \in (\alpha - 1/k, \alpha] \cap \Lambda$ since Λ is dense in $[0,1]$. Using $\mathbf{x} \in [A]_\alpha \subseteq [A]_\gamma$, $\lim_{n\to\infty} \delta_H([A_n]_\gamma, [A]_\gamma) = 0$, and (7) therefore shows the existence of a sequence $(\mathbf{x}^n)_{n\in\mathbb{N}}$ converging to $\mathbf{x}$ and fulfilling $\mathbf{x}^n \in [A_n]_\gamma$ for every $n \in \mathbb{N}$. Lipschitz continuity (see [27]) implies

$$A_n(\mathbf{x}) \ge A_n(\mathbf{x}^n) - \sqrt{\rho}\,\beta\big(\mathbf{x}^n, \mathbf{x}\big) \ge \gamma - \sqrt{\rho}\beta\big(\mathbf{x}^n, \mathbf{x}\big)$$

from which $\liminf_{n\to\infty} A_n(\mathbf{x}) \geq \gamma$ follows. Since k was arbitrary we get $\liminf_{n\to\infty} A_n(\mathbf{x}) \geq \alpha$. In case of $\alpha = 0$ this inequality is clearly valid. Assume that there exists $r > 0$ such that $A_n(\mathbf{x}) \geq \alpha + r$ holds for infinitely many $n \in \mathbb{N}$. For every $\gamma \in (\alpha, \alpha + r) \cap \Lambda$ we get $\mathbf{x} \in [A_n]_\beta$ infinitely often, hence $\mathbf{x} \in {}^t\limsup_{n\to\infty}[A_n]_\gamma = [A]_\gamma$ follows, which contradicts $A(\mathbf{x}) = \alpha$. Consequently $\limsup_{n\to\infty} A_n(\mathbf{x}) \leq \alpha$, and $\lim_{n\to\infty} A_n(\mathbf{x}) = \alpha$ holds. If (b) holds pointwise convergence can be proved analogously. ■

We close this section with the following remark.

Remark 2 It is well known that the empirical copula $\hat{E}_n$ (for the two-dimensional case, see, for instance [27,18]) for i.i.d. data from A is a strongly consistent estimator of A w.r.t. the uniform metric d_∞. As direct consequence of Proposition 1 we get almost sure convergence of $[\hat{E}_n]_\alpha$ to $[A]_\alpha$ w.r.t. δ_H for every α being a continuity point of Φ_A without any further assumptions. For a general discussion on the estimation of level sets of general (two-dimensional) distribution functions we refer to [8].

4 Three conditioning based metrics on $\mathcal{C}_\rho$ and their main properties

In the following we will only consider the case $\rho \geq 3$, the case $\rho = 2$ has already been studied in [32]. Using the one-to-one correspondence between ρ-dimensional copulas and Markov kernels fulfilling equation (3) we will consider the following three metrics:

$$D_\infty(A,B) := \sup_{\mathbf{y}\in[0,1]^{\rho'}} \int_{[0,1]} \big|K_A(x,[\mathbf{0},\mathbf{y}]) - K_B(x,[\mathbf{0},\mathbf{y}])\big| d\lambda(x) \tag{14}$$

$$D_1(A,B) := \int_{[0,1]^{\rho'}} \int_{[0,1]} \big|K_A(x,[\mathbf{0},\mathbf{y}]) - K_B(x,[\mathbf{0},\mathbf{y}])\big| d\lambda(x)\, d\lambda^{\rho'}(\mathbf{y}) \tag{15}$$

$$D_2(A,B)^2 := \int_{[0,1]^{\rho'}} \int_{[0,1]} \big|K_A(x,[\mathbf{0},\mathbf{y}]) - K_B(x,[\mathbf{0},\mathbf{y}])\big|^2 d\lambda(x)\, d\lambda^{\rho'}(\mathbf{y}) \tag{16}$$

To simplify notation we will write

$$\Phi_{A,B}(\mathbf{y}) := \int_{[0,1]} \big|K_A(x,[\mathbf{0},\mathbf{y}]) - K_B(x,[\mathbf{0},\mathbf{y}])\big| d\lambda(x) \tag{17}$$

for all $A, B \in \mathcal{C}_\rho$ and all $\mathbf{y} \in [0,1]^{\rho'}$. Before analyzing the main properties of the function $\Phi_{A,B}$ it has to be shown that D_1, D_2, D_∞ are metrics.

Lemma 1 *D_1, D_2 and D_∞ defined according to (14), (15), (16) are metrics on the space $\mathcal{C}_\rho$ of ρ-dimensional copulas.*

Proof: First of all we shown that the integrand in (15) and (16) is measurable. Define H on $[0,1]^\rho$ by $H(x,\mathbf{y}) := K_A(x,[\mathbf{0},\mathbf{y}])$, then H is measurable in x and non-decreasing and right-continuous in $\mathbf{y}$. Fix $z \in [0,1]$. For every $\mathbf{q} \in \mathcal{Q} := \mathbb{Q}^{\rho'} \cap [0,1]^{\rho'}$ define

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

and set

$$E := \bigcup_{q \in \mathcal{Q}} E_{\mathbf{q}} \times [\mathbf{0}, \mathbf{q}] \in \mathcal{B}([0,1]^{\rho}).$$

Using right-continuity it is straightforward to see that $E = H^{-1}([0,z))$, from which measurability of H directly follows. Furthermore, if $D_1(A,B) = 0$ then there exists a set $\Lambda \subseteq [0,1]^{\rho}$ with $\lambda^{\rho}(\Lambda) = 1$ such that for every $(x,\mathbf{y}) \in \Lambda$ we have equality $K_A(x,[\mathbf{0},\mathbf{y}]) = K_B(x,[\mathbf{0},\mathbf{y}])$. It follows that $\lambda^{\rho'}(\Lambda_x) = 1$ 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. Using disintegration (see [19]) or equation (2) directly yields $A = B$. The remaining properties of a metric are obviously fulfilled. The fact that D_∞ and D_2 are metrics can be shown analogously. ■

As in the two-dimensional setting $\Phi_{A,B}$ is Lipschitz continuous:

Lemma 2 *For every pair $A, B \in \mathcal{C}_\rho$ the function $\Phi_{A,B}$, defined according to (17), is Lipschitz continuous (with Lipschitz constant 2) w.r.t. $\|\cdot\|_1$ on $[0,1]^{\rho'}$. Furthermore $\Phi_{A,B}$ fulfills*

$$\Phi_{A,B}(\mathbf{y}) \leq 2 \min\left\{ \min_{i=1\ldots,\rho'} y_i, \sum_{i=1}^{\rho'} (1 - y_i) \right\} \tag{18}$$

for every $\mathbf{y} \in [0,1]^{\rho'}$.

Proof: Consider the measurable rectangle $E := \times_{i=1}^{\rho'} E_i$, then we get

$$\begin{aligned} \int_{[0,1]} |K_A(x,E) - K_B(x,E)| \, d\lambda(x) &\leq \int_{[0,1]} K_A(x,E) + K_B(x,E) d\lambda(x) \\ &\leq 2 \min_{i=1\ldots,\rho'} \lambda(E_i) \end{aligned}$$

as well as

$$\begin{aligned} \int_{[0,1]} |K_A(x,E) - K_B(x,E)| \, d\lambda(x) &= \int_{[0,1]} |K_A(x,E^c) - K_B(x,E^c)| \, d\lambda(x) \\ &\leq \int_{[0,1]} K_A(x,E^c) + K_B(x,E^c) d\lambda(x). \end{aligned}$$

Set $R := \int_{[0,1]} K_A(x,E^c) d\lambda(x)$. Using the fact that

$$E^c = \bigcup_{i=1}^{\rho'} \left(E_1 \times E_2 \times \ldots \times E_{i-1} \times E_i^c \times [0,1] \times \cdots \times [0,1] \right)$$

it follows that

$$\begin{aligned}
R &\leq \lambda(E_1^c) + \min\left\{\lambda(E_1), \lambda(E_2^c)\right\} + \min\left\{\lambda(E_1), \lambda(E_2), \lambda(E_3^c)\right\} + \ldots + \\
&\quad + \min\left\{\lambda(E_1), \lambda(E_2), \lambda(E_3), \ldots, \lambda(E_{\rho'-1}), \lambda(E_{\rho'}^c)\right\} \\
&\leq \sum_{i=1}^{\rho'} (1 - \lambda(E_i))
\end{aligned}$$

from which (18) is a direct consequence. Suppose now that $\mathbf{y}, \mathbf{z} \in [0,1]^{\rho'}$ and set $U := \times_{i=1}^{\rho'} \big[0, \min\{y_i, z_i\}\big]$. Then, expressing U^c as union of ρ' rectangles (in the same way as E^c before), using the fact that A, B are copulas, and setting $L := |\Phi_{A,B}(\mathbf{y}) - \Phi_{A,B}(\mathbf{z})|$ yields

$$\begin{aligned}
L &\leq \int_{[0,1]} \big|K_A(x, [\mathbf{0}, \mathbf{y}]) - K_B(x, [\mathbf{0}, \mathbf{y}]) - K_A(x, [\mathbf{0}, \mathbf{z}]) + K_B(x, [\mathbf{0}, \mathbf{z}])\big| d\lambda(x) \\
&\leq \int_{[0,1]} \big|K_A(x, [\mathbf{0}, \mathbf{y}] \cap U^c) - K_B(x, [\mathbf{0}, \mathbf{y}] \cap U^c)\big| d\lambda(x) \; + \\
&\quad + \int_{[0,1]} \big|K_A(x, [\mathbf{0}, \mathbf{z}] \cap U^c) - K_B(x, [\mathbf{0}, \mathbf{z}] \cap U^c)\big| d\lambda(x) \\
&\leq 2\Big(y_1 - \min\{y_1, z_1\} \, + \, y_2 - \min\{y_2, z_2\} \, + \ldots + \, y_{\rho'} - \min\{y_{\rho'}, z_{\rho'}\}\Big) + \\
&\quad + 2\Big(z_1 - \min\{y_1, z_1\} \, + \, z_2 - \min\{y_2, z_2\} \, + \ldots + \, z_{\rho'} - \min\{y_{\rho'}, z_{\rho'}\}\Big) \\
&= 2 \sum_{i=1}^{\rho'} |y_1 - z_i| = 2\|\mathbf{y} - \mathbf{z}\|_1. \quad \blacksquare
\end{aligned}$$

Lemma 3 *For all $A, B \in \mathcal{C}_\rho$ the following two inequalities hold:*

$$\begin{aligned}
D_\infty(A, B) \geq D_1(A, B) \; &\geq \; \left(\frac{1}{2\rho'}\right)^{\rho'} \frac{D_\infty(A, B)^\rho}{\rho} \\
D_2(A, B) \geq D_1(A, B) \; &\geq \; D_2^2(A, B)
\end{aligned} \tag{19}$$

Additionally we have $D_\infty(A, B) \geq d_\infty(A, B)$.

Proof: Since $D_2(A, B) \geq D_1(A, B)$ is a direct consequence of Cauchy-Schwarz inequality and $D_1(A, B) \geq D_2^2(A, B)$ obviously holds it suffices to prove the first inequality in (19), which can be done as follows. Fix $A, B \in \mathcal{C}_\rho$ and choose $\mathbf{z} \in [0,1]^{\rho'}$ such that $\Phi_{A,B}(\mathbf{z}) = D_\infty(A, B)$. Then there exists a ρ'-dimensional hypercube $W \subset [0,1]^{\rho'}$ with side length $\frac{\Phi_{A,B}(\mathbf{z})}{2\rho'}$ having $\mathbf{z}$ as one of its $2^{\rho'}$ vertices $\{\mathbf{v}^1, \mathbf{v}^2, \ldots, \mathbf{v}^{2^{\rho'}}\}$. For every $j \in \{1, \ldots, 2^{\rho'}\}$ define a point in $[0,1]^\rho$ via $\mathbf{w}^j := (v_1^j, v_2^j, \ldots, v_{\rho'}^j, 0)$, set $\mathbf{w}^{2^{\rho'}+1} := (z_1, z_2, \ldots, z_{\rho'}, f(\mathbf{z}))$ and let $P \subseteq [0,1]^\rho$ denote the convex hull of the points $\{\mathbf{w}^1, \ldots, \mathbf{w}^{2^{\rho'}+1}\}$. Then P

is contained in the endograph of $\Phi_{A,B}$ (the volume between the graph of $\Phi_{A,B}$ and the hyperplane $x_\rho = 0$) and, using Fubini's theorem, we finally get

$$\begin{aligned} D_\infty(A,B) \geq D_1(A,B) &= \int_{[0,1]^{\rho'}} \Phi_{A,B}(\mathbf{y}) d\lambda^{\rho'}(\mathbf{y}) \geq \lambda^\rho(P) \\ &= \int_{[0,\Phi_{A,B}(\mathbf{z})]} \left(\frac{x}{2\rho'}\right)^{\rho'} d\lambda(x) = \left(\frac{1}{2\rho'}\right)^{\rho'} \frac{\Phi_{A,B}(\mathbf{z})^\rho}{\rho}, \end{aligned}$$

which completes the proof of (19). The last assertion of the lemma follows directly from the fact that for all $(x, \mathbf{y}) \in [0,1]^\rho$ we have

$$\begin{aligned} |A(x,\mathbf{y}) - B(x,\mathbf{y})| &= \left| \int_{[0,x]} K_A(t,[\mathbf{0},\mathbf{y}]) - K_A(t,[\mathbf{0},\mathbf{y}])\, d\lambda(t) \right| \\ &\leq \int_{[0,x]} \big| K_A(t,[\mathbf{0},\mathbf{y}]) - K_A(t,[\mathbf{0},\mathbf{y}]) \big|\, d\lambda(t) \leq \Phi_{A,B}(\mathbf{y}) \\ &\leq D_\infty(A,B) \quad \blacksquare \end{aligned}$$

As direct consequence of Lemma 3 we get the following result.

Theorem 2 *Suppose that $A, A_1, A_2, \ldots$ are elements of $\mathcal{C}_\rho$. Then the following three conditions are equivalent:*

(a) $\lim_{n\to\infty} D_\infty(A_n, A) = 0$
(b) $\lim_{n\to\infty} D_1(A_n, A) = 0$
(c) $\lim_{n\to\infty} D_2(A_n, A) = 0$

Before proceeding with the three metrics D_1, D_2 and D_∞ we take a look to the interrelation between D_1 and the so-called ∂-convergence in the sense of Mikusinski and Taylor (see [24], [25]): Suppose that $A, A_1, A_2, \ldots \in \mathcal{C}_\rho$. Then, by definition, $(A_n)_{n\in\mathbb{N}}$ ∂-converges to A if and only if for every $i \in \{1,\ldots,\rho\}$ we have (t in the i-th coordinate)

$$\lim_{n\to\infty} \int_{[0,1]} \left| \frac{\partial A_n}{\partial x_i}(x_1,\ldots,t,\ldots,x_\rho) - \frac{\partial A}{\partial x_i}(x_1,\ldots,t,\ldots,x_\rho) \right| d\lambda(t) = 0. \quad (20)$$

For every $i \in \{1,\ldots,\rho\}$ and every $A \in \mathcal{C}_\rho$ let $\pi_i(A)$ denote the copula defined by

$$\pi_i(A)(x_1,\ldots,x_\rho) = A(x_i, x_2, \ldots, x_{i-1}, x_1, x_i, \ldots, x_\rho). \quad (21)$$

Using this notation the following result holds:

Lemma 4 *The metric D_∂, defined by*

$$D_\partial(A,B) := \sum_{i=1}^{\rho} D_1(\pi_i(A), \pi_i(B))$$

is a metrization of ∂-convergence on $\mathcal{C}_\rho$.

Proof: Fix $A \in \mathcal{C}_\rho$, (a version of) the corresponding regular conditional distribution K_A and let $\mathbf{y} \in [0,1]^{\rho'}$ be arbitrary. Then the set Λ of all $x \in (0,1)$ at which the function $t \mapsto A(t,\mathbf{y})$ is differentiable and which are at the same time Lebesgue points (see [29]) of the L^1-function $t \mapsto K_A(t,[\mathbf{0},\mathbf{y}])$. fulfills $\lambda(\Lambda) = 1$. For every $t_0 \in \Lambda$ we have

$$\begin{aligned}\frac{\partial A}{\partial x_1}(t_0,\mathbf{y}) &= \lim_{n\to\infty} \frac{A(t_0+1/n,\mathbf{y}) - A(t_0-1/n,\mathbf{y})}{2/n} \\ &= \lim_{n\to\infty} \frac{\mu_A\Big((t_0-1/n,t_0+1/n] \times [\mathbf{0},\mathbf{y}]\Big)}{2/n} \\ &= \lim_{n\to\infty} \frac{\int_{(t_0-1/n,t_0+1/n)} K_A(t,[\mathbf{0},\mathbf{y}])d\lambda(t)}{2/n} = K_A(t_0,[\mathbf{0},\mathbf{y}]),\end{aligned}$$

from which the result easily follows. ■.

The following lemma will be useful in Section 5.

Lemma 5 *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 for λ-almost every $x \in [0,1]$ then we have $\lim_{n\to\infty} D_1(A_n,A) = 0$. The same result holds for the metrics D_2 and D_∞.*

Proof: Fix $\mathbf{y} \in [0,1]^{\rho'}$, then the boundary of $[\mathbf{0},\mathbf{y}]$, denoted by $\mathrm{bd}([\mathbf{0},\mathbf{y}])$, consists of finitely many hyperplanes. Therefore, using the fact that $A \in \mathcal{C}_\rho$ and disintegration, we have

$$0 = \mu_A\big([0,1] \times \mathrm{bd}([\mathbf{0},\mathbf{y}])\big) = \int_{[0,1]} K\big(x, \mathrm{bd}([\mathbf{0},\mathbf{y}])\big) d\lambda(x).$$

Hence, denoting the set of all $x \in [0,1]$ such that both $K\big(x,\mathrm{bd}([\mathbf{0},\mathbf{y}])\big) = 0$ and $K_n(x,\cdot) \to K(x,\cdot)$ weakly by Λ, it follows that $\lambda(\Lambda) = 1$ and that for every $x \in \Lambda$ we have $\lim_{n\to\infty} K_n(x,[\mathbf{0},\mathbf{y}]) = K(x,[\mathbf{0},\mathbf{y}])$. Lebesgue's theorem on dominated converges implies $\lim_{n\to\infty} \Phi_{A_n,A}(\mathbf{y}) = 0$ which completes the proof for D_1 since $\mathbf{y}$ was arbitrary. The analogous results for D_2 and D_∞ are a direct consequence of Theorem 2. ■

The following theorem is one of the main results of this section.

Theorem 3 *The metric space $(\mathcal{C}_\rho, D_1)$ is complete and separable.*

Proof: We extend the proof of the two-dimensional result given in [32] to the multivariate setting: 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 $[0,1]^\rho$, defined by $H_n(x,\mathbf{y}) := K_n(x,[0,\mathbf{y}])$. Since we have

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

$(H_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $L^1([0,1]^\rho, \mathcal{B}([0,1]^\rho), \lambda^\rho)$, so there exists a L^1-limit $H \in L^1([0,1]^\rho, \mathcal{B}([0,1]^\rho), \lambda^\rho)$. According to the theorem of Weyl (see [15], [29]) we can find a subsequence $(H_{n_j})_{j\in\mathbb{N}}$ and a Borel set $\Delta \subseteq [0,1]^\rho$ with $\lambda^\rho(\Delta) = 1$ and $\lim_{j\to\infty} H_{n_j}(x,\mathbf{y}) = H(x,\mathbf{y})$ for all $(x,\mathbf{y}) \in \Delta$. W.l.o.g. we may assume that $H(x,\mathbf{1}) = 1$ for every $x \in [0,1]$. We will show that we can find a measurable function $G : [0,1]^\rho \to [0,1]$ with the following two properties: (i) $G = H$ λ^ρ-a.s. and (ii) $K(x,[\mathbf{0},\mathbf{y}]) := G(x,\mathbf{y})$ is again a regular conditional distribution of a copula $A \in \mathcal{C}_\rho$.
Using Fubini's theorem (see [15], [29]) it follows that $\lambda(\Delta_{\mathbf{y}}) = \lambda(\{x \in [0,1] : (x,\mathbf{y}) \in \Delta\}) = 1$ for $\lambda_{\rho'}$-almost all $\mathbf{y} \in [0,1]^{\rho'}$. Consequently we can find a countable dense set $Q = \{\mathbf{q}_1, \mathbf{q}_2, \ldots\} \subseteq [0,1]^{\rho'}$ with $\mathbf{1} \in Q$ and a set $\Lambda_0 \subseteq [0,1]$ with $\lambda(\Lambda_0) = 1$ such that $\lim_{j\to\infty} H_{n_j}(x,\mathbf{q}) = H(x,\mathbf{q})$ holds for every $\mathbf{q} \in Q$ and every $x \in \Lambda_0$. Again using Fubini we can find a subset $\Lambda \subseteq \Lambda_0$ such that $\lambda_{\rho'}(\Delta_x) = \lambda_{\rho'}(\{\mathbf{y} \in [0,1] : (x,\mathbf{y}) \in \Delta\}) = 1$ for every $x \in \Lambda$. Fix $x \in \Lambda$ and consecutively define two new functions $g'_x, g_x : [0,1]^{\rho'} \to [0,1]$ by

$$g'_x(\mathbf{y}) = \begin{cases} 0 \text{ if } \min\{y_1,\ldots,y_{\rho'}\} = 0, \\ \sup_{\mathbf{q}\in Q:\mathbf{y}\succ\mathbf{q}} H(x,\mathbf{q}) \text{ otherwise} \end{cases}$$

and

$$g_x(\mathbf{y}) = \begin{cases} 1 \text{ if } \mathbf{y} = \mathbf{1}, \\ \inf_{\mathbf{q}\in\mathbb{Q}_{\rho'}:\mathbf{q}\gg\mathbf{y}} g'_x(\mathbf{q}) \text{ otherwise} \end{cases}$$

whereby $\mathbb{Q}_{\rho'} := \mathbb{Q}^{\rho'} \cap [0,1]^{\rho'}$ denotes the vectors with rational coordinates in $[0,1]^{\rho'}$. It follows directly from the definition that g'_x is non-decreasing (in each coordinate), bounded by 1 and left-continuous. Furthermore it is straightforward to show (see, for instance, [35] Lemma 2.5) that for each $\mathbf{y}$ being a continuity point of g'_x we have $\lim_{j\to\infty} H_{n_j}(x,\mathbf{y}) = g'_x(\mathbf{y})$. Using monotonicity of g'_x according to [22] $\lambda_{\rho'}$-almost all $\mathbf{y} \in [0,1]^{\rho'}$ are continuity points of g'_x, so $H(x,\mathbf{y}) = g'_x(\mathbf{y})$ for $\lambda_{\rho'}$-almost all $\mathbf{y} \in [0,1)^{\rho'}$ holds. Using the same argumentation it follows that g_x is non-decreasing (in each coordinate), bounded by 1 and right-continuous, and that $H(x,\mathbf{y}) = g_x(\mathbf{y})$ for $\lambda_{\rho'}$-almost all $\mathbf{y} \in [0,1)^{\rho'}$. Furthermore, Helly's theorem (see [5]) implies the existence of a subsequence $(H_{n_{j_l}})_{l\in\mathbb{N}}$ and a distribution function h_x on $[0,1]^{\rho'}$ such that $(H_{n_{j_l}})$ converges weakly to h_x. Hence, using monotonicity and right-continuity, altogether we conclude that g_x is a distribution function on $[0,1]^{\rho'}$ and that $H(x,\mathbf{y}) = g_x(\mathbf{y})$ for $\lambda_{\rho'}$-almost all $\mathbf{y} \in [0,1]^{\rho'}$. Having this define a new function by $G : [0,1]^\rho \to [0,1]$ by

$$G(x,\mathbf{y}) := \mathbf{1}_\Lambda(x)\, g_x(\mathbf{y}) \,+\, \mathbf{1}_{\Lambda^c}(x)\, \mathbf{1}_{[0,1]^{\rho'}}(\mathbf{y}).$$

It is straightforward to see that $G(\cdot,\cdot)$ is measurable in x for fixed $\mathbf{y}$ and, by construction, a distribution function in $\mathbf{y}$ on $[0,1]^{\rho'}$ for fixed x. Hence G induces a Markov kernel $K(\cdot,\cdot)$ by setting $K(x,[\mathbf{0},\mathbf{y}]) := G(x,\mathbf{y})$, and, for every x, uniquely extending the probability measure $K(x,\cdot)$ from the class of all rectangles $[\mathbf{0},\mathbf{y}]$ to $\mathcal{B}([0,1]^{\rho'})$ in the standard way (see [15], [19]).
Furthermore it follows from the construction and Fubini's theorem that G

coincides with H λ_ρ-almost everywhere on $[0,1]^\rho$. It remains to show that $K(x,[\mathbf{0},\mathbf{y}])$ is a regular conditional distribution of a copula $A \in \mathcal{C}_\rho$, which can be done as follows: Define a function $A : [0,1]^\rho \to [0,1]$ by

$$A(x,\mathbf{y}) := \int_{[0,x]} G(t,\mathbf{y}) d\lambda(t),$$

then A is absolutely continuous as function of x for fixed $\mathbf{y}$ and right-continuous in $\mathbf{y}$ for fixed x. Furthermore on a dense set $\mathcal{D} \subseteq [0,1]^{\rho'}$ we have

$$\lim_{n\to\infty} A_{n_j}(x,\mathbf{y}) = A(x,\mathbf{y}).$$

. Since all A_n are copulas, using equicontinuity, it follows that $(A_{n_j})_{j\in\mathbb{N}}$ converges pointwise to a copula $\tilde{A}$, which, according to the above-mentioned continuity properties of A implies $\tilde{A} = A$. 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 $\mathcal{C}_\rho$ (so called-checkerboard copulas) as follows: $\mathcal{S}_n$ is the class of all $B \in \mathcal{C}_\rho$ whose mass μ_B is uniformly distributed on each rectangle $R_\mathbf{i}$ of the form $R_\mathbf{i} = \times_{j=1}^{\rho}[(i_j-1)/n, i_j/n]$, whereby $\mathbf{i} \in \{1,\dots,n\}^\rho$. Denote by $\mathcal{SQ}_n$ the subset of all $B \in \mathcal{S}_n$ that also fulfill $\mu_B(R_\mathbf{i}) \in \mathbb{Q}$ for every $\mathbf{i}$. Since $\mathcal{SQ}_n$ is countably infinity $\mathcal{SQ} := \cup_{n=2}^{\infty}\mathcal{SQ}_n \subseteq \mathcal{C}_\rho$ is countably infinite too. According to the results in [25] $\mathcal{S}_n$ is dense in $\mathcal{C}_\rho$ with respect to the topology induced by ∂-convergence on $\mathcal{C}_\rho$. Hence, by Lemma 4, $\mathcal{S}_n$ is dense in the metric space $(\mathcal{C}_\rho, D_1)$. Fix an arbitrary $B \in \mathcal{S}_n$ and let $\varepsilon > 0$. It is straightforward to see that $\mathcal{S}_n$ is isomorphic to a polytope $\Omega_n \subseteq [0,1]^{n^\rho}$ (e.g. sort all rectangles $R_\mathbf{i}$ lexicographically and then consider the vector of masses of these rectangles as embedding $\varepsilon : S_n \to [0,1]^{n^\rho}$). Being a polytope Ω_n the set of extreme points E_n of Ω_n is finite and every point $\mathbf{a}$ in Ω_n is a convex combination of m $(\leq n^\rho + 1)$ elements of E_n (see [17]), i.e. $\mathbf{a} = \sum_{i=1}^m \alpha_i \mathbf{e}^i$ with $\alpha_i \geq 0$, $\sum_{i=1}^m \alpha_i = 1$, and $\mathbf{e}^1,\dots,\mathbf{e}^m \in E_n$. Since $\mathbb{Q}$ is dense in $[0,1]$ we can find a vector $(\beta_1,\dots,\beta_m) \in \mathbb{Q}^m$ such that both $\max_{i=1\dots m} |\alpha_i - \beta_i| < \varepsilon/(n^\rho+1)$ and $\sum_{i=1}^m \beta_i = 1$ holds. Returning to $B \in \mathcal{S}_n$ this implies the existence of an element $\hat{B} \in \mathcal{SQ}_n$ such that

$$\max_{\mathbf{i}\in\{1,\dots,n\}^\rho} |\mu_B(R_\mathbf{i}) - \mu_{\hat{B}}(R_\mathbf{i})| < \varepsilon/(n^\rho+1).$$

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. ∎

Since the metrics D_1, D_2 and D_∞ induce the same topology Theorem 3 also holds for $(\mathcal{C}_\rho, D_2)$ and $(\mathcal{C}_\rho, D_\infty)$:

Proposition 3 *The metric spaces $(\mathcal{C}_\rho, D_2)$, $(\mathcal{C}_\rho, D_\infty)$ are separable and complete.*

Theorem 4 *The metric space $(\mathcal{C}_\rho, D_\partial)$ is separable and complete, i.e. D_∂ is a separable and complete metrization of ∂-convergence.*

Proof: If $(A_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $(\mathcal{C}_\rho, D_\partial)$ then $\pi_i(A_n)$ (see equation (21)) is a Cauchy sequence in $(\mathcal{C}_\rho, D_1)$ for each $i \in \{1,\dots,\rho\}$, so, according to Theorem 3, we have $\lim_{n\to\infty} D_1(\pi_i(A_n), B^i)$ for some copula $B^i \in \mathcal{C}_\rho$. Using Lemma 3 $\lim_{n\to\infty} d_\infty(\pi_i(A_n), B^i) = 0$ follows, which, in turn, implies $\lim_{n\to\infty} d_\infty(A_n, \pi_i(B^i)) = 0$. The d_∞-limit is unique so there exists $B \in \mathcal{C}_\rho$ with $B = \pi_i(B^i)$ for all $i \in \{1,\dots,\rho\}$, from which

$$\lim_{n\to\infty} D_\partial(A_n, B) = \lim_{n\to\infty} \sum_{i=1}^{\rho} D_1(\pi_i(A_n), \pi_i(B)) = 0$$

immediately follows. The fact that $(\mathcal{C}_\rho, D_\partial)$ is separable is already contained in the last part of the proof of Theorem 3. ■

The following is a slight generalization of (non-mutual) complete dependence (on the first coordinate) to the multivariate setting (see [21], [32]):

Definition 1 A copula $A \in \mathcal{C}_\rho$ is called *completely dependent* (w.r.t. the first coordinate) if there exist λ-preserving transformations $S_1, S_2, \dots, S_{\rho'} : [0,1] \to [0,1]$ such that

$$K(x, E) := \mathbf{1}_E(S_1x, S_2x, \dots, S_{\rho'}x) = \delta_{(S_1x, S_2x, \dots, S_{\rho'}x)}(E)$$

is a regular conditional distribution of A. The class of all completely dependent copulas will be denoted by $\mathcal{C}_\rho^d$. In case all S_i are bijections with measurable inverse then A will be called *generalized shuffle of the minimum copula* M_ρ, if, additionally, all S_i are piecewise linear then A will be called *shuffle of the minimum copula* M_ρ (see [10] as well as [12], [13]).

The following lemma can be proved in completely the same manner as the two-dimensional result in [32].

Lemma 6 *Given $A \in \mathcal{C}_\rho$ the following conditions are equivalent:*

(d1) $A \in \mathcal{C}_\rho^d$.

(d2) There exist λ-preserving transformations $S_1, S_2, \dots, S_{\rho'} : [0,1] \to [0,1]$ such that $A(x, y_1, y_2, \dots, y_{\rho'}) = \lambda([0,x] \cap S_1^{-1}([0,y_1]) \cap \dots \cap S_{\rho'}^{-1}([0,y_{\rho'}])$ holds for all $(x, y_1, y_2, \dots, y_{\rho'})$.

(d3) There exists λ-preserving transformations $S_1, S_2, \dots, S_{\rho'} : [0,1] \to [0,1]$ such that $\mu_A(\Gamma(S_1, \dots, S_{\rho'})) = 1$, whereby

$$\Gamma(S_1, \dots, S_{\rho'}) := \{(x, S_1x, S_2x, \dots, S_{\rho'}x) : x \in [0,1]\} \in \mathcal{B}([0,1]^\rho)$$

denotes the graph of $(S_1, S_2, \dots, S_{\rho'})$.

The main content of the subsequent Theorem is that, given copulas $A, B \in \mathcal{C}_\rho^d$ with corresponding measure preserving transformations $(S_1, \dots, S_{\rho'})$ and $(T_1, \dots, T_{\rho'})$, the distance $D_1(A, B)$ is small if $\|S_i - T_i\|$ is small for every $i \in \{1, \dots, \rho'\}$. Furthermore it shows that $\mathcal{C}_\rho^d$ and Π_ρ are strongly separated w.r.t. all three metrics D_∞, D_1, D_2.

Theorem 5 *Suppose that $A, B \in \mathcal{C}_\rho^d$ and that $(S_1, \ldots, S_{\rho'})$ and $(T_1, \ldots, T_{\rho'})$ are (versions of) the corresponding λ-preserving transformations, then the following two inequalities hold:*

$$D_1(A,B) \leq \sum_{i=1}^{\rho'} \|S_i - T_i\|_1$$
$$D_\infty(A, \Pi_\rho) \geq 2^{-\rho}$$

Proof: Using Fubini's theorem we may express $D_1(A,B)$ in the form

$$\begin{aligned} D_1(A,B) &= \int_{[0,1]} \int_{[0,1]^{\rho'}} \big|K_A([\mathbf{0},\mathbf{y}]) - K_B(x,[\mathbf{0},\mathbf{y}])\big| \, d\lambda_{\rho'}(\mathbf{y}) d\lambda(x) \\ &= \int_{[0,1]} \int_{[0,1]^{\rho'}} \big|\mathbf{1}_{\times_{i=1}^{\rho'}(S_i x,1]}(\mathbf{y}) - \mathbf{1}_{\times_{i=1}^{\rho'}(T_i x,1]}(\mathbf{y})\big| \, d\lambda_{\rho'}(\mathbf{y}) d\lambda(x). \end{aligned}$$

For every $i \in \{1, \ldots, \rho'\}$ define a rectangle R_i by $R_i := \times_{j=1}^{\rho'} E_j$ whereby $E_j = (\min\{S_j x, T_j x\}, \max\{S_j x, T_j x\}]$ for $j = i$ and $E_j = [0,1]$ for all $j \neq i$. Obviously we have

$$\Big\{\mathbf{y} \in [0,1]^{\rho'} : \big|\mathbf{1}_{\times_{i=1}^{\rho'}(S_i x,1]}(\mathbf{y}) - \mathbf{1}_{\times_{i=1}^{\rho'}(T_i x,1]}(\mathbf{y})\big| = 1\Big\} \subseteq \bigcup_{i=1}^{\rho'} R_i,$$

which directly implies that

$$D_1(A,B) \leq \int_{[0,1]} \sum_{i=1}^{\rho'} |S_i x - T_i x| d\lambda(x) = \sum_{i=1}^{\rho'} \|S_i - T_i\|_1.$$

To prove the second assertion suppose that $A \in \mathcal{C}_\rho$ and that $S_1, \ldots, S_{\rho'}$ are the corresponding λ-preserving transformations. Then for every $\mathbf{y} \in [0,1]^{\rho'}$ we have

$$\begin{aligned} \Phi_{A,\Pi_\rho}(\mathbf{y}) &= \int_{[0,1]} \Big|\mathbf{1}_{[\mathbf{0},\mathbf{y}]}(S_1 x, S_2 x, \ldots, S_{\rho'} x) - \lambda^{\rho'}([\mathbf{0},\mathbf{y}])\Big| d\lambda(x) \\ &\geq \lambda^{\rho'}([\mathbf{0},\mathbf{y}])\, \lambda\Big(\bigcup_{i=1}^{\rho'} S_i^{-1}\big((y_i,1]\big)\Big) \geq \prod_{i=1}^{\rho'} y_i \max_{i=1,\ldots,\rho'} (1-y_i), \end{aligned}$$

which implies $D_\infty(A, \Pi_\rho) \geq 2^{-\rho}$ and completes the proof. ■

It has already been proved in [13] that the family of all shuffles of M_ρ are dense in the family of all generalized shuffles w.r.t. ∂-convergence. The second part of the following theorem points in the same direction and allows for a very simple proof:

Theorem 6 *$\mathcal{C}_\rho^d$ is closed in the metric space $(\mathcal{C}_\rho, D_1)$. Furthermore for every copula $A \in \mathcal{C}_\rho^d$ there exists a sequence $(B_n)_{n\in\mathbb{N}}$ of shuffles of the minimum copula M_ρ with $\lim_{n\to\infty} D_1(B_n, A) = 0$.*

Proof: Suppose that $(A_n)_{n\in\mathbb{N}}$ is a sequence in $\mathcal{C}_\rho^d$ converging to $A \in \mathcal{C}_\rho$ w.r.t. D_1 and let $(S_1^n, \ldots, S_{\rho'}^n)$ denote the corresponding family of λ-preserving transformations. Since convergence w.r.t. D_1 implies convergence w.r.t. D_∞ there is a subsequence $(A_{n_j})_{j\in\mathbb{N}}$ and a Borel set $\Lambda \subseteq [0,1]$ with $\lambda(\Lambda) = 1$ such that for every $\mathbf{q} \in \mathcal{Q} := \mathbb{Q}^{\rho'} \cap [0,1]^{\rho'}$ and every $x \in \Lambda$ we have

$$\lim_{j\to\infty} K_{A_{n_j}}(x, [\mathbf{0}, \mathbf{q}]) = K_A(x, [\mathbf{0}, \mathbf{q}])$$

from which $K_A(x, [\mathbf{0}, \mathbf{q}]) \in \{0, 1\}$ follows immediately. For every $x \in \Lambda$ define a vector $(S_1x, \ldots, S_{\rho'}x)$ by

$$(S_1x, \ldots, S_{\rho'}x) = \mathbf{1}_\Lambda(x) \inf \big\{\mathbf{q} \in \mathcal{Q} : K_A(x, [\mathbf{0}, \mathbf{q}]) = 1\big\}.$$

Having this is follows easily that $A \in \mathcal{C}_\rho^d$, which completes the proof of the first part of the theorem. The second part is direct consequence of Lemma 11 and the fact that every λ-preserving transformation on $[0,1]$ is the (almost everywhere) limit of a sequence of piecewise linear λ-preserving bijections on $[0,1]$ (see [6]). ■

Remark 3 In [25] examples for copulas $A, A_1, A_2, \ldots \in \mathcal{C}_\rho$ are given such that $\lim_{n\to\infty} d_\infty(A_n, A) = 0$ but $\limsup_{n\to\infty} D_1(A_n, A) > 0$. Alternatively one may also use the second assertion of Theorem 5 and the fact that, according to [25], the class of multivariate shuffles of M_ρ are dense in $(\mathcal{C}_\rho, d_\infty)$.

Remark 4 The dependence measure $\tau_1 : \mathcal{C}_2 \to [0,1]$ in [32] was defined by $\tau_1(A) = 3D_1(A, \Pi_2)$. τ_1 was shown to fulfill the property that completely dependent copulas (describing the situation of two $\mathcal{U}_{0,1}$-distributed random variables X, Y such that knowing X implies knowing Y) are assigned maximum dependence - a seemingly natural property since independence describes exactly the other extreme in which knowing X does not improve our a-priori knowledge about Y. Since, according to the results in this section, most properties of D_1 in dimension two also hold in the general ρ-dimensional setting it might seem natural to simply consider

$$\tau_1(A) := a\, D_1(A, \Pi_\rho),$$

for all $A \in \mathcal{C}_\rho$ (a being a normalizing constant). It is, however, straightforward to see that this yields no reasonable notion of a dependence measure. In fact, we would also have $\tau_1(A) > 0$ for copulas $A \in \mathcal{C}_\rho$ describing independence of X and $\mathbf{Y} = (Y_1, \ldots, Y_{\rho'})$, i.e. for copulas whose associated ρ-stochastic measure μ_A is of the form $\mu_A = \lambda \otimes \mu_{A'}$ with $A' \in \mathcal{C}_{\rho'} \setminus \{\Pi_{\rho'}\}$. For $\rho = 3$, for instance, considering the copula $A \in \mathcal{C}_3$ with $\mu_A = \lambda \otimes \mu_{A'}$ and $A' \neq \Pi_2$ yields

$$\begin{aligned}\Phi_{A,\Pi_3}(y_1, y_2) &= \int_{[0,1]} \big|K_A(x, [0, y_1] \times [0, y_2]) - y_1y_2\big|\, dx \\ &= \big|A'(y_1, y_2) - y_1y_2\big|\end{aligned}$$

implying $D_\infty(A, \Pi_3) = d_\infty(A', \Pi_2) > 0$ as well as $D_1(A, \Pi_3) > 0$. For $A' = M_2$ or $A' = W_2$ we even get $D_\infty(A, \Pi_3) = 1/4$.

One idea to overcome this problem and extend τ_1 to $\mathcal{C}_\rho$ could be to consider the two-dimensional projections $\hat{A}_i$, $i \in \{2,\ldots,\rho\} \in \mathcal{C}_2$, given by

$$\hat{A}_i(x,y) = A(x,1,\ldots,1\underbrace{,\,y\,,}_{i\text{-th coordinate}}1,\ldots,1)$$

for all $x,y \in [0,1]^2$, and set

$$\tau_1(A) := \frac{3}{\rho'} \sum_{i=2}^{\rho} D_1(\hat{A}_i, \Pi_2). \tag{22}$$

Using the results in [32] it would follow immediately that $\tau_1(A) = 1$ is maximal if and only if $\hat{A}_i \in \mathcal{C}_2^d$ for every $i \in \{2,\ldots,\rho\}$. Also note that for all afore-mentioned copulas $A \in \mathcal{C}_\rho$ with $\mu_A = \lambda \otimes \mu_{A'}$ using τ_1 according to (22) we would get $\tau_1(A) = 0$.
We finally remark that the idea behind equation (22) is closely connected with the method of constructing multivariate concordance measures by averaging the concordance values of the two-dimensional 'projections' and refer to [30] and the reference therein for an interesting and detailed discussion.

We conclude this section with a theorem gathering the interrelations/equivalences between all mentioned concepts of convergence on $\mathcal{C}_\rho$:

Theorem 7 *Suppose that $A, A_1, A_2, \ldots \in \mathcal{C}_\rho$ and consider the following conditions:*

(a) $\lim_{n\to\infty} D_\infty(A_n, A) = 0$.
(b) $\lim_{n\to\infty} D_1(A_n, A) = 0$.
(c) $\lim_{n\to\infty} D_2(A_n, A) = 0$.
(d) $\lim_{n\to\infty} \big|A_n(\mathbf{x}) - A(\mathbf{x})\big| = 0$ *for every* $\mathbf{x} \in [0,1]^\rho$.
(e) $\lim_{n\to\infty} d_\infty(A_n, A) = 0$.
(f) The sequence $(\mu_{A_n})_{n\in\mathbb{N}}$ *converges weakly to* μ_A.
(g) $\lim_{n\to\infty} d_{end}(A_n, A) = 0$.
(h) There exists a set $\Lambda \subseteq [0,1]$ of Lebesgue measure 1 such that for all $\alpha \in \Lambda$ the equality $\lim_{n\to\infty} \delta_H([A_n]_\alpha, [A]_\alpha) = 0$ *holds.*
(i) There exists a set $\Gamma \subseteq [0,1]$ of Lebesgue measure 1 such that for all $\alpha \in \Gamma$ the equality $\lim_{n\to\infty} \delta_H([A_n]^\alpha, [A]^\alpha) = 0$ *holds.*
(j) For every $\alpha \in (0,1)$ such that $A^{-1}(\{\alpha\})$ has empty interior the equality $\lim_{n\to\infty} \delta_H([A_n]_\alpha, [A]_\alpha) = 0$ *holds.*

Then the conditions (d)-(j) are equivalent, conditions (a)-(c) are equivalent, and each of the conditions (a)-(c) implies all other conditions.

Proof: It only remains to prove the equivalence of (e) and (g). To do so we will show that the following inequality holds for all $A, B \in \mathcal{C}_\rho$.

$$d_{end}(A,B) \le d_\infty(A,B) \le (1+\sqrt{\rho})\, d_{end}(A,B) \tag{23}$$

Since for $(\mathbf{x},t) \in end(A)$ there exists $(\mathbf{x},s) \in end(B)$ such that $|t-s| \le |A(\mathbf{x}) - B(\mathbf{x})| \le d_\infty(A,B)$, the first part of (23) is obvious.

To prove the second part fix $\mathbf{x} \in [0,1]^\rho$ and set $\Delta := d_{end}(A,B)$. Assume that $A(\mathbf{x}) \geq B(\mathbf{x})$. Then, since $(\mathbf{x}, A(\mathbf{x})) \in end(A)$, there exists $(\mathbf{y}, s) \in end(B)$ such that $\beta(\mathbf{x},\mathbf{y}), |A(x,y)-s| \leq \Delta$. Using Lipschitz continuity $|B(\mathbf{x})-B(\mathbf{y})| \leq \sqrt{\rho}\,\beta(\mathbf{x},\mathbf{y}) \leq \sqrt{\rho}\,\Delta$, and therefore

$$B(\mathbf{x}) \geq B(\mathbf{y}) - \sqrt{\rho}\,\Delta \geq A(\mathbf{x}) - (1+\sqrt{\rho})\,\Delta$$

follows. This shows that in this case $0 \leq A(\mathbf{x}) - B(\mathbf{y}) \leq (1+\sqrt{\rho})\,\Delta$ holds, from which the second part of (23) follows. ■

5 An application to singular copulas with fractal support induced by special Iterated Function Systems (IFS)

In the following we recall the definition of a generalized transformation matrix given in [34] (also see [16]) and show that under quite general conditions the IFS construction of copulas with fractal support also converges w.r.t. the three metrics mentioned in the previous section. Fix $m_1, \ldots, m_\rho \in \mathbb{N}$ and set

$$\mathcal{I}_\rho := \times_{i=1}^{\rho} I_i, \quad \text{whereby } I_i = \{1,\ldots,m_i\} \text{ for every } i \in \{1,\ldots,d\}. \tag{24}$$

We will denote elements in $\mathcal{I}_\rho$ in the form $\mathbf{i} = (i_1,\ldots,i_\rho)$, and, for every probability distribution τ on $(\mathcal{I}_\rho, 2^{\mathcal{I}_\rho})$ write $\tau(\mathbf{i}) := \tau(\{\mathbf{i}\})$ for the point mass in $\mathbf{i}$.

Definition 2 Suppose that $\rho \geq 2$, that $m_1,\ldots,m_\rho \in \mathbb{N}$, $\max_j m_j \geq 2$, and let $\mathcal{I}_\rho$ be defined according to (24). A probability distribution τ on $(\mathcal{I}_\rho, 2^{\mathcal{I}_\rho})$ is called *generalized transformation matrix* if for every $j \in \{1,\ldots,\rho\}$

$$\Delta_j^k := \sum_{\mathbf{i} \in \mathcal{I}_\rho:\ i_j = k} \tau(\mathbf{i}) > 0$$

holds for every $k \in I_j$. The class of all generalized transformation matrices for fixed $\rho \geq 2$ will be denoted by $\mathcal{T}_\rho$.

Every $\tau \in \mathcal{T}_\rho$ induces a partition of $[0,1]^\rho$ in the following way: For each $j \in \{1,\ldots,\rho\}$ define $a_0^j := 0$,

$$a_k^j := \sum_{\mathbf{i} \in \mathcal{I}_\rho:\ i_j \leq k} \tau(\mathbf{i}),$$

and $E_k^j := [a_{k-1}^j, a_k^j]$ for every for every $k \in I_j$. Then $\bigcup_{k \in I_j} E_k^j = [0,1]$ and $E_{k_1}^j \cap E_{k_2}^j$ is empty or consists of exactly one point whenever $k_1 \neq k_2$. Setting $R_{\mathbf{i}} := \times_{j=1}^{d} E_{i_j}^j$ for every $\mathbf{i} \in \mathcal{I}_d$ therefore yields a family of compact rectangles $(R_{\mathbf{i}})_{\mathbf{i} \in \mathcal{I}_\rho}$ whose union is $[0,1]^\rho$ and which additionally fulfills that $R_{\mathbf{i}_1} \cap R_{\mathbf{i}_2}$ is empty or a set of λ_ρ-measure zero whenever $\mathbf{i}_1 \neq \mathbf{i}_2$.

To complete the construction of the IFSP induced by $\tau \in \mathcal{T}_\rho$ define affine contractions $w_{\mathbf{i}} : [0,1]^\rho \to R_{\mathbf{i}}$ by

$$w_{\mathbf{i}}(x_1,\dots,x_d) = \begin{pmatrix} a^1_{i_1-1} \\ a^2_{i_2-1} \\ \vdots \\ a^d_{i_d-1} \end{pmatrix} + \begin{pmatrix} (a^1_{i_1} - a^1_{i_1-1})\, x_1 \\ (a^2_{i_2} - a^2_{i_2-1})\, x_2 \\ \vdots \\ (a^d_{i_d} - a^d_{i_d-1})\, x_d \end{pmatrix}.$$

Since the j-th coordinate of $w_{\mathbf{i}}(x_1,\dots,x_d)$ only depends on i_j and x_j we will also denote it by $w^j_{i_j}$, i.e. $w^j_{i_j} : [0,1] \to E^j_{i_j}$, $w^j_{i_j}(x_j) := a^j_{i_j-1} + (a^j_{i_j} - a^j_{i_j-1})\, x_j$. It follows directly from the construction that

$$\Big([0,1]^d, (w_{\mathbf{i}})_{\mathbf{i}\in\mathcal{I}_d}, \tau(\mathbf{i})_{\mathbf{i}\in\mathcal{I}_d}\Big) \tag{25}$$

is an IFSP. The IFSP induces an operator $V_\tau : \mathcal{P}([0,1]^\rho) \to \mathcal{P}([0,1]^\rho)$, defined by

$$V_\tau(\mu) := \sum_{\mathbf{i}\in\mathcal{I}_\rho} \tau(\mathbf{i}) \mu^{w_{\mathbf{i}}} \tag{26}$$

and it is straightforward to verify that V_τ maps $\mathcal{P}_\mathcal{C}([0,1]^\rho)$ into $\mathcal{P}_\mathcal{C}([0,1]^\rho)$. Furthermore it follows from the construction that the support of the limit copula $A^\star$ has λ^ρ measure zero if there exists (at least one) $\mathbf{i} \in \mathcal{I}_\rho$ with $\tau(\mathbf{i}) = 0$. Analogous to [32] we will show now how V_τ acts on regular conditional distributions. Suppose that $\mathbf{i} \in \mathcal{I}_\rho$ and that K_A is a regular conditional distribution of the copula $A \in \mathcal{C}_\rho$. If $(x, \mathbf{y}) \in R_{\mathbf{i}}$ then we have

$$\begin{aligned} K_{V_\tau A}(x, [\mathbf{0}, \mathbf{y}]) = {} & \frac{1}{\sum_{\mathbf{j}\in\mathcal{I}_\rho : j_1 = i_1} \tau(\mathbf{j})} \sum_{\mathbf{j}\in\mathcal{I}_\rho : j_1 = i_1, j_2 < i_2, \dots, j_\rho < i_\rho} \tau(\mathbf{j}) + \frac{1}{\sum_{\mathbf{j}\in\mathcal{I}_\rho : j_i = i_1} \tau(\mathbf{j})} \times \\ & \times \sum_{\mathbf{j}\in\mathcal{I}_\rho : \mathbf{j}\preceq\mathbf{i}, j_1 = i_1, \exists l \geq 2 : j_l = i_l} \tau(\mathbf{j}) K_A\left(\frac{x - a^1_{i_1-1}}{a^1_{i_1} - a^1_{i_1-1}}, \left[\mathbf{0}, \hat{\mathbf{y}}^{\mathbf{j}}\right]\right) \end{aligned}$$

whereby

$$\left[\mathbf{0}, \hat{\mathbf{y}}^{\mathbf{j}}\right] = \left[0, \min\left\{\frac{y_1 - a^2_{j_2-1}}{a^2_{j_2} - a^2_{j_2-1}}, 1\right\}\right] \times \cdots \times \left[0, \min\left\{\frac{y_{\rho'} - a^\rho_{j_\rho-1}}{a^\rho_{j_\rho} - a^\rho_{j_\rho-1}}, 1\right\}\right].$$

It follows that

$$\begin{aligned} & \int_{[a^1_{i_1-1}, a^1_{i_1}]} \big| K_{V_\tau A}(x, [\mathbf{0}, \mathbf{y}]) - K_{V_\tau B}(x, [\mathbf{0}, \mathbf{y}]) \big| d\lambda(x) \\ & \quad \leq \sum_{\mathbf{j}\in\mathcal{I}_\rho : \mathbf{j}\preceq\mathbf{i}, j_1 = i_1, \exists l \geq 2 : j_l = i_l} \tau(\mathbf{j}) \Phi_{A,B}(\hat{\mathbf{y}}^{\mathbf{j}}) \\ & \quad \leq D_\infty(A, B) \sum_{\mathbf{j}\in\mathcal{I}_\rho : \mathbf{j}\preceq\mathbf{i}, j_1 = i_1, \exists l \geq 2 : j_l = i_l} \tau(\mathbf{j}) \end{aligned}$$

which, in turn, implies that

$$\begin{aligned}\Phi_{V_\tau A, V_\tau B}(\mathbf{y}) &= \int_{[0,1]} \big| K_{V_\tau A}(x,[\mathbf{0},\mathbf{y}]) - K_{V_\tau B}(x,[\mathbf{0},\mathbf{y}]) \big| d\lambda(x) \\ &= D_\infty(A,B) \sum_{i_1=1}^{m_1} \sum_{\mathbf{j}\in\mathcal{I}_\rho : \mathbf{j}\preceq\mathbf{i}, j_1=i_1, \exists l\geq 2: j_l=i_l} \tau(\mathbf{j}) \\ &= D_\infty(A,B) \underbrace{\sum_{\mathbf{j}\in\mathcal{I}_\rho : \forall l\geq 2: j_l\leq i_l, \exists l\geq 2: j_l=i_l} \tau(\mathbf{j})}_{:=b(i_2,i_3,\ldots,i_\rho)} .\end{aligned} \tag{27}$$

(i) If there exists $l \geq 2$ such that $i_l < m_l$ then we have (with the notation from Definition 2)

$$b(i_2,i_3,\ldots,i_\rho) \leq 1 - \sum_{\mathbf{j}\in\mathcal{I}_\rho : j_l=m_l} \tau(\mathbf{j}) \leq 1 - \min_{l=2\ldots\rho} \Delta_l^{m_l} < 1.$$

(ii) If $i_l = m_l$ for all $l \in \{2,\ldots,\rho\}$ then

$$b(i_2,i_3,\ldots,i_\rho) = \sum_{\mathbf{j}\in\mathcal{I}_\rho : \exists l\geq 2: j_l=m_l} \tau(\mathbf{j}) := \Delta'_\tau. \tag{28}$$

Hence, if we assure that the latter is smaller than one then, using Banach's fixed point theorem, we have the following result:

Theorem 8 *Suppose that $\rho \geq 3$, that $\tau \in \mathcal{T}_\rho$ is a generalized transformation matrix such that Δ'_τ in equation (28) fulfills $\Delta'_\tau < 1$, and let the operator V_τ be defined according to (26). Then V_τ is a contraction on the metric space $(\mathcal{C}_\rho, D_\infty)$. Furthermore there exists a unique copula $A^\star$ fulfilling $V_\tau A^\star = A^\star$ and for every $B \in \mathcal{C}_\rho$ we have $\lim_{n\to\infty} D_\infty(V_\tau^n B, A^\star) = 0$.*

Proof: Using equation (27) it follows immediately that

$$\Phi_{V_\tau A, V_\tau B}(\mathbf{y}) \leq \Phi_{A,B}(\mathbf{y}) L$$

whereby $L := \max\left\{\Delta', 1 - \min_{l=2\ldots\rho} \Delta_l^{m_l}\right\} < 1$. This completes the proof of the first part of the theorem. The second part is a direct consequence of Banach's fixed point theorem in combination with Theorem 3 and Proposition 3. ■.

Proposition 4 *Under the assumptions of Theorem 8 there exists a unique copula $A^\star \in \mathcal{C}_\rho$ fulfilling $V_\tau A^\star = A^\star$ and for every $B \in \mathcal{C}_\rho$ we have*

$$\lim_{n\to\infty} D_1(V_\tau^n B, A^\star) = \lim_{n\to\infty} D_2(V_\tau^n B, A^\star) = 0.$$

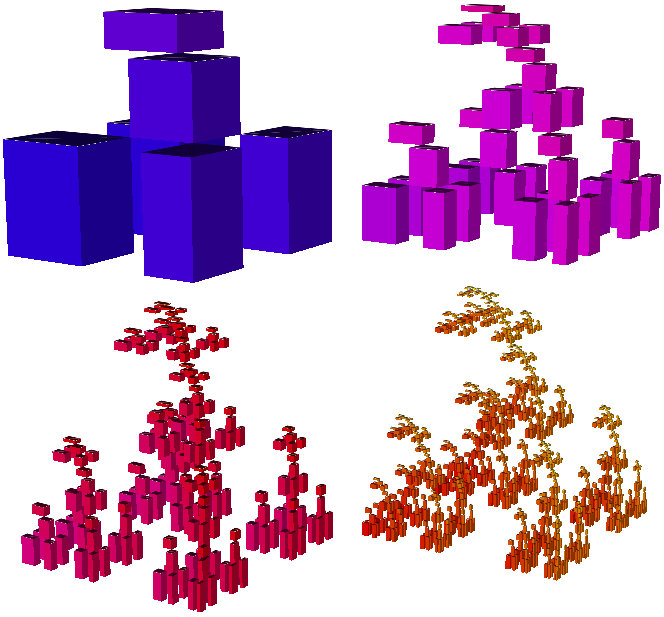

Fig. 1 Densities of $V_\tau^1(\Pi_3)$, $V_\tau^2(\Pi_3)$, $V_\tau^3(\Pi_3)$ and $V_\tau^4(\Pi_3)$ - colors ranging from blue (low density) via red to yellow (high density), τ according to Example 1.

Example 1 Using the IFS approach we can easily construct a three-dimensional copula whose support is a *dendrit-like* structure: For every $j \in \{1,2,3\}$ set $m_j := 3$ and define $\tau \in \mathcal{T}_3$ by

$$\tau(\mathbf{i}) = \begin{cases} \frac{2}{12} & \text{if } \mathbf{i} \in \{(1,1,1),(3,1,1),(1,3,3)\} \\ \frac{1}{12} & \text{if } \mathbf{i} \in \{(1,3,1),(3,3,1)\} \\ \frac{1}{3} & \text{if } \mathbf{i} = (2,2,2) \\ 0 & \text{otherwise.} \end{cases}$$

Then we get $a^1 = (0,\, 5/12,\, 3/4, 1)$, $a^2 = (0, 1/3,\, 2/3,\, 1)$ as well as $a^3 = (0, 1/2,\, 5/6,\, 1)$, so the corresponding IFS consists of 27 contractions, 6 of which have positive probability $\tau(\mathbf{i}) > 0$. Figure 1 depicts the density of $\mathcal{V}_\tau^n(\Pi)$ for $n \in \{1,2,3,4\}$. Since τ obviously fulfills $\Delta'_\tau < 1$ according to Proposition 4 we also have $\lim_{n\to\infty} D_1(V_\tau^n B, A^\star) = 0$ for every $B \in \mathcal{C}_3$.

Example 2 As second example we consider the case $m_1 = m_2 = 3$, $m_3 = 2$ and define and define $\tau \in \mathcal{T}_3$ by

$$\tau(\mathbf{i}) = \begin{cases} \frac{1}{4} & \text{if } \mathbf{i} \in \{(1,1,1),(2,2,1)\} \\ \frac{1}{12} & \text{if } \mathbf{i} \in \{(1,3,1),(3,3,1)\} \\ \frac{1}{3} & \text{if } \mathbf{i} = (1,3,2) \\ 0 & \text{otherwise.} \end{cases}$$

We get $a^1 = (0, 2/3, 11/12, 1)$, $a^2 = (0, 1/4, 1/2, 1)$, $a^3 = (0, 2/3, 1)$, so the corresponding IFS consists of 18 contractions, 5 of which have positive probability $\tau(\mathbf{i}) > 0$. Figure 2 depicts the density of $\mathcal{V}_\tau^n(\Pi)$ for $n \in \{1, 2, 3, 4\}$. Since τ obviously fulfills $\Delta'_\tau < 1$ according to Proposition 4 we again have $\lim_{n\to\infty} D_1(V_\tau^n B, A^\star) = 0$ for every $B \in \mathcal{C}_3$.

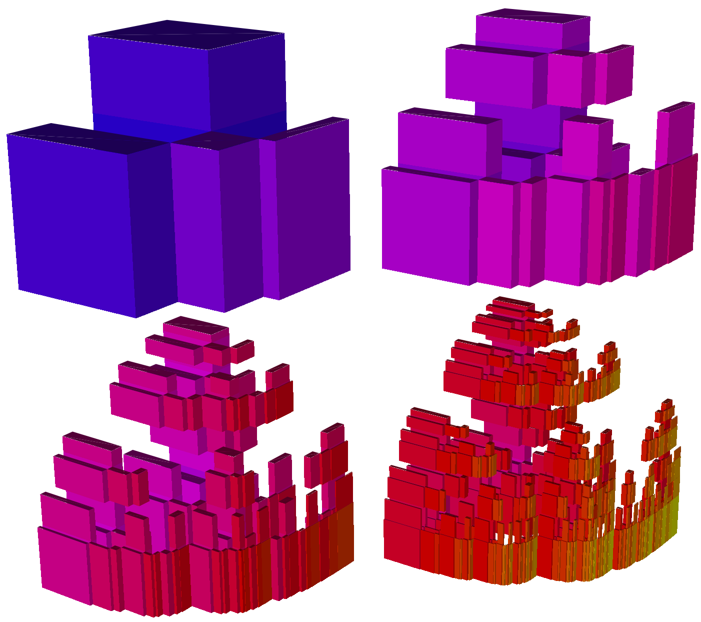

Fig. 2 Densities of $V_\tau^1(\Pi_3)$, $V_\tau^2(\Pi_3)$, $V_\tau^3(\Pi_3)$ and $V_\tau^4(\Pi_3)$ - colors ranging from blue (low density) via red to yellow (high density), τ according to Example 2.

We conclude this section with a simple example showing that we can also have convergence of the IFS construction w.r.t. D_1 although the condition $\Delta'_\tau < 1$ in Theorem 8 is not fulfilled:

Example 3 We consider the case $\rho = 3$, $m_1 = 1, m_2 = m_3 = 2$ and the generalized transformation matrix $\tau \in \mathcal{T}_3$, defined by $\tau(1,1,1) = 0$ and $\tau(1,1,2) = \tau(1,2,1) = \tau(1,2,2) = 1/3$. Obviously τ fulfills $\Delta'_\tau = 1$ so Theorem 8 is not applicable. Nevertheless (see [34]) there exists a unique copula $A^\star \in \mathcal{C}_3$ such that $V_\tau A^\star = A^\star$ and $\lim_{n\to\infty} d_\infty(V_\tau^n B, A^\star) = 0$ holds for every $B \in \mathcal{C}_3$. To simplify notation we will write $H_B(x, y, z) := K_B(x, [0, y] \times [0, z])$ for every $B \in \mathcal{C}_3$ and all $(x, y, z) \in [0, 1]^3$. Fix an arbitrary $A \in \mathcal{C}_3$. For every $x \in [0, 1]$ obviously we have

$$H_{V_\tau A}(x, 1, 1) = H_{V_\tau A^\star}(x, 1, 1) = H_{A^\star}(x, 1, 1) = 1$$

$$\begin{aligned} H_{V_\tau A}\Big(x,\frac{1}{3},\frac{1}{3}\Big) &= H_{V_\tau A^\star}\Big(x,\frac{1}{3},\frac{1}{3}\Big) = H_{A^\star}\Big(x,\frac{1}{3},\frac{1}{3}\Big) \\ H_{V_\tau A}\Big(x,\frac{1}{3},1\Big) &= H_{V_\tau A^\star}\Big(x,\frac{1}{3},1\Big) = H_{A^\star}\Big(x,\frac{1}{3},1\Big) \\ H_{V_\tau A}\Big(x,1,\frac{1}{3}\Big) &= H_{V_\tau A^\star}\Big(x,1,\frac{1}{3}\Big) = H_{A^\star}\Big(x,1,\frac{1}{3}\Big). \end{aligned}$$

Define a new IFS $\big([0,1]^2,(f_i)_{i=1}^4\big)$ by

$$\begin{aligned} f_1(y,z) &= \Big(\frac{y}{3},\frac{z}{3}\Big),\ f_2(y,z) = \Big(\frac{y+2}{3},\frac{z}{3}\Big) \\ f_3(y,z) &= \Big(\frac{y}{3},\frac{z+2}{3}\Big),\ f_4(y,z) = \Big(\frac{y+2}{3},\frac{z+2}{3}\Big) \end{aligned}$$

and let $\mathcal{H}:\mathcal{K}([0,1]^2)\to\mathcal{K}([0,1]^2)$ denote the corresponding Hutchinson operator. Furthermore set $E_0 := \{(1,1)\}$ and $E_n := \mathcal{H}^n(E_0)$ for every $n\in\mathbb{N}$. Then, according to the above mentioned four equalities, we have $H_{V_\tau A}(x,\mathbf{e}) = H_{A^\star}(x,\mathbf{e})$ for all $\mathbf{e}\in E_1$. In fact it is straightforward to verify that for every $n\in\mathbf{N}$ equality $H_{V_\tau^n A}(x,\mathbf{e}) = H_{A^\star}(x,\mathbf{e})$ holds for all $\mathbf{e}\in E_n$, that the sequence $(E_n)_{n\in\mathbb{N}}$ is monotonically increasing, $E := \bigcup_{n=1}^\infty E_n$ is dense in $[0,1]^2$, and that for every $\mathbf{e}\in E$ and $x\in[0,1]$ we have

$$\lim_{n\to\infty} H_{V_\tau^n A}(x,\mathbf{e}) = H_{A^\star}(x,\mathbf{e}).$$

From this it follows immediately that $K_{V_\tau^n A}(x,\cdot)\to K_{A^\star}(x,\cdot)$ weakly for $n\to\infty$, which, according to Lemma 5, implies $\lim_{n\to\infty} D_1(V_\tau^n A, A^\star) = 0$.

Remark 5 The 'projection' $\hat{\tau}\in\mathcal{T}_2$ of τ in the previous example, defined by $\hat{\tau}(i,j) := \tau(1,i,j)$ for $1\le i,j\le 2$, is a two-dimensional transformation matrix. So, according to [16], [32], there exists a unique fixed point $\hat{A}^\star\in\mathcal{C}_2$ of $V_{\hat{\tau}}$. Note that the fixed point $A^\star$ in Example 3 and $\hat{A}^\star$ fulfill $\mu_{A^\star} = \lambda\otimes\mu_{\hat{A}^\star}$.

6 Conclusion and future work

We have studied the main properties of three conditioning based metrics on the space $\mathcal{C}_\rho$ of ρ-dimensional copulas, the most important result being that the corresponding metric spaces are complete and separable and that all three metrics induce the same topology. As application of completeness we have (i) found a simple complete and separable metrization of ∂-convergence studied by Mikusinski and Taylor, and (ii) have shown that the IFS construction of copulas with fractal support also converges w.r.t. any of these metrics under quite mild conditions. Furthermore we have characterized uniform convergence in $\mathcal{C}_\rho$ in terms of convergence of λ-almost all α-level-sets w.r.t.t. Hausdorff metric. As direct consequence of this equivalence and the fact that D_1 is strictly stronger than d_∞ the question arises whether it is possible to give a characterization of D_1-convergence in terms of convergence of level sets w.r.t.t. Hausdorff metric too, a question which, to the best of the authors' knowledge, has not been answered yet. More importantly the authors are also interested in extending the dependence measure introduced in [32] to the case where the conditioning is done w.r.t. several 'input' variables.

Acknowledgements The first author gratefully acknowledges the support of the grant MTM2011-22394 from the Spanish Ministry of Science and Innovation.

References

1. Barnsley, M.F.: *Fractals everywhere.* Academic Press, Cambridge (1993)
2. Bauer, H.: *Wahrscheinlichkeitstheorie.* de Gruyter, Berlin, New York (2002)
3. Bhattacharya, R., Waymire, E.C.: *A Basic Course in Probability Theory.* Springer Verlag, New York (2007)
4. Berger, A.: *Chaos and Chance - An Introduction to Stochastic Aspects of Dynamics.* de Gruyter, Berlin, New York (2001)
5. Billingsley, P.: *Convergence of Probability Measures.* John Wiley & Sons (1968)
6. Chou, S.H., Nguyen, T.T.: On Fréchet theorem in the set of measure preserving functions over the unit interval. *Internat. J. Math. Math. Sci.* **13** (2), 373-378 (1990)
7. Darsow, W.F., Nguyen, B., Olsen, E.T.: Copulas and Markov processes. *Illinois J. Math.* **36** (4), 600-642 (1992)
8. Di Bernardino, E., Laloë, T., Maume-Deschamps, V., Prieur, C.: Plug-in estimation of level sets in a non-compact setting with applications in multivariate risk theory. *ESAIM Probab. Stat.*, **17**, 236-256 (2013)
9. Dudley, R.M.: *Real Analysis and Probability.* Cambridge University Press (2002)
10. Durante, F., Sarkoci, P., Sempi, C.: Shuffles of copulas. *J. Math. Anal. Appl.* **352**, 914-921 (2009)
11. Durante, F., Sempi, C.: 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)
12. Durante, F., Fernández Sánchez, J.: Multivariate shuffles and approximation of copulas. *Stat. Probabil. Lett.* **80** (23-24), 1827-1834 (2010)
13. Durante, F., Fernández Sánchez, J.: On the approximation of copulas via shuffles of Min. *Stat. Probab. Lett.* **82** (10), 1761-1767 (2012)
14. Edgar, G.: *Measure, Topology, and Fractal Geometry.* Springer Verlag, New York (2008)
15. Elstrodt, J.: *Maß- und Integrationstheorie.* Springer Verlag, Berlin, Heidelberg, New York (1999)
16. Fredricks, G.A., Nelsen, R.B., Rodríguez-Lallena, J.A.: Copulas with fractal supports. *Insur. Math. Econ.* **37**, 42-48 (2005)
17. Gruber, P.M.: *Convex and Discrete Geometry.* Springer-Verlag, Berlin, Heidelberg (2007)
18. Janssen, P., Swanepoel, J., Veraverbeke, N.: Large sample behavior of the Bernstein copula estimator. *J. Stat. Plan. Infer.* **142**, 1189-1197 (2012)
19. Kallenberg, O.: *Foundations of modern probability.* Springer Verlag, New York, Berlin, Heidelberg (1997)
20. Klenke, A.: *Probability Theory - A Comprehensive Course.* Springer Verlag, Berlin, Heidelberg (2007)
21. Lancaster, H.O.: Correlation and complete dependence of random variables. *Ann. Math. Stat.* **34**, 1315-1321 (1963)
22. Lavric, B.: Continuity of monotone functions. *Archivum Mathematicum* **29**, 1-4 (1993)
23. Li, X., Mikusinski, P., Taylor, M.D.: Strong approximation of copulas. *J. Math. Anal. Appl.* **255**, 608-623 (1998)
24. Mikusinski, P., Taylor, M.D.: Markov operators and n-copulas. *Annales Polonici Mathematici* **96**, 75-95 (2009)
25. Mikusinski, P., Taylor, M.D.: Some approximations of n-copulas. *Metrika* **72**, 385-414 (2010)

26. Molchanov, I.: *Theory of Random Sets.* Springer, London (2005)
27. Nelsen, R.B.: *An Introduction to Copulas.* Springer, New York (2006)
28. Olsen, E.T., Darsow, W.F., Nguyen, B.: 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)
29. Rudin, W.: *Real and Complex Analysis.* McGraw-Hill International Editions, Singapore (1987)
30. Taylor, M.D.: Multivariate measures of concordance. *The Institute of Statistical Mathematics* **59**, 789-806 (2006)
31. Trutschnig, W.: Characterization of the sendograph-convergence of fuzzy sets by means of their L_p- and levelwise convergence. *Fuzzy Set Syst.* **161** (8), 1064-1077 (2010)
32. Trutschnig, W.: On a strong metric on the space of copulas and its induced dependence measure. *J. Math. Anal. Appl.* **384**, 690-705 (2011)
33. Trutschnig, W.: Some results on the convergence of (quasi-)copulas. *Fuzzy Set Syst.* **191**, 113-121 (2012)
34. Trutschnig, W., Fernández-Sánchez, J.: Idempotent and multivariate copulas with fractal support. *J. Stat. Plan. Infer.* **142**, 3086-3096 (2012)
35. van der Vaart, A. W.: *Asymptotic Statistics.* Cambridge University Press (2000)