Chapter 14
Complete dependence everywhere?

Wolfgang Trutschnig

Abstract Describing the situation of full predictability of a random variable Y given the value of another random variable X, the notion of complete dependence might seem far too restrictive to be of any practical importance at first glance. Recent results have shown, however, that complete dependence naturally appears in various settings. This chapter will therefore sketch some problems related to complete dependence. Doing so, the focus will be on dependence measures strictly separating extreme dependence concepts, on a problem related to joint-default maximization, on a question from uniform distribution theory, and on the relationship between the two most well-known measures of concordance, Kendall's τ and Spearman's ρ. A short excursion to topology showing that complete dependence is not atypical at all complements the chapter.

14.1 Introduction

Given two random variables X,Y we call Y completely dependent on X if there exists a measurable function f such that $Y = f(X)$ holds with probability one (see [21] for the original definition). In other words: Knowing X means knowing Y but not necessarily vice versa. Although a dependence structure describing full predictability seems very pathological at first, research in the field of dependence modeling conducted in the last years clearly points in the direction that complete dependence is much more important than reflected by textbooks so far. Main objective of this chapter is to illustrate this observation by means of some fairly recent results.

The rest of this chapter is organized as follows: Section 2 gathers notation and preliminaries that will be used in the sequel and states various properties equivalent to complete dependence. Section 3 recalls one possible way to construct metrics that

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

M. Úbeda Flores et al. (eds.), *Copulas and Dependence Models with Applications*,
DOI 10.1007/978-3-319-64221-5_14

clearly distinguish the two extremal dependence concepts, that of complete dependence and that of independence. Based on these metrics a new dependence measure ζ_2 is introduced and analyzed, and alternative approaches from the literature are mentioned. Section 4 first sketches why extreme points naturally comes into play in the context of optimization problems and then shows two concrete examples in both of which completely dependent copulas constitute solutions of the maximization problem. Section 5 contains a short excursion to topology showing that (in the language of Baire categories) with respect to weak convergence complete dependence is all but an atypical property of a copula. Finally, Section 6 focuses on recently established sharp inequalities between Kendall's τ and Spearman's ρ and points out that mutually completely dependent random variable cover all possible constellations of τ and ρ.

14.2 Notation and preliminaries

In the sequel $\mathscr{C}$ will denote the family of all two-dimensional *copulas*, $\mathscr{P}_{\mathscr{C}}$ the family of all *doubly stochastic measures*, i.e. the family of all probability measures on $[0,1]^2$ whose marginals are uniformly distributed on $[0,1]$. M will denote the lower Fréchet-Hoeffding bound, W the lower Fréchet-Hoeffding bound and Π the product copula; for background on copulas we refer to [7] and [25]. For every $C \in \mathscr{C}$ the corresponding doubly stochastic measure will be denoted by μ_C. Letting d_∞ denote the uniform metric on $\mathscr{C}$ it is well known that $(\mathscr{C}, d_\infty)$ is a compact metric space.

For every metric space (Ω, d) the Borel σ-field on Ω will be denoted by $\mathscr{B}(\Omega)$, δ_x will denote the Dirac measure (concentrated) at $x \in \Omega$. λ and λ_2 will denote the Lebesgue measure on $\mathscr{B}(\mathbb{R})$ and $\mathscr{B}(\mathbb{R}^2)$ respectively. For every probability measure ν on $\mathscr{B}(\Omega)$ the support of ν, i.e. the complement of the union of all open sets U fulfilling $\nu(U) = 0$, will be denoted by $Supp(\nu)$.

Suppose that (Ω_1, d_1) and (Ω_2, d_2) are metric spaces. A *Markov kernel* from Ω_1 to $\mathscr{B}(\Omega_2)$ is a mapping $K : \Omega_1 \times \mathscr{B}(\Omega_2) \to [0,1]$ such that $x \mapsto K(x,B)$ is measurable for every fixed $B \in \mathscr{B}(\Omega_2)$ and $B \mapsto K(x,B)$ is a probability measure for every fixed $x \in \Omega_1$. Given real-valued random variables X, Y on a probability space $(\Omega, \mathscr{A}, \mathscr{P})$, a Markov kernel $K : \mathbb{R} \times \mathscr{B}(\mathbb{R}) \to [0,1]$ is called a *regular conditional distribution of Y given X* if for every $B \in \mathscr{B}(\mathbb{R})$

$$K(X(\omega), B) = \mathbb{E}(\mathbf{1}_B \circ Y | X)(\omega) \tag{14.1}$$

holds $\mathscr{P}$-a.e. It is well known 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 $\mathscr{P}^X$-a.s. (i.e. unique for $\mathscr{P}^X$-almost all $x \in \mathbb{R}$) and that $K(\cdot,\cdot)$ only depends on $\mathscr{P}^{X \otimes Y}$. Hence, given $A \in \mathscr{C}$ we will denote (a version of) the regular conditional distribution of Y given X by $K_A(\cdot,\cdot)$, directly view it as Markov kernel from $[0,1]$ to $\mathscr{B}([0,1])$, and refer to $K_A(\cdot,\cdot)$ simply as *regular conditional distribution of A* or as

Markov kernel of A. Note that for every $A \in \mathscr{C}$, its regular conditional distribution $K_A(\cdot,\cdot)$, and every Borel set $G \in \mathscr{B}([0,1]^2)$ we have the following *disintegration* (here $G_x := \{y \in [0,1] : (x,y) \in G\}$ denotes the x-section of G for every $x \in [0,1]$)

$$\int_{[0,1]} K_A(x, G_x)\, d\lambda(x) = \mu_A(G), \tag{14.2}$$

so in particular

$$\int_{[0,1]} K_A(x, F)\, d\lambda(x) = \lambda(F) \tag{14.3}$$

for every $F \in \mathscr{B}([0,1])$. On the other hand, every Markov kernel $K : [0,1] \times \mathscr{B}([0,1]) \to [0,1]$ fulfilling (14.3) induces a unique element $\mu \in \mathscr{P}_{\mathscr{C}}([0,1]^2)$ via (14.2). For every $A \in \mathscr{C}$ and $x \in [0,1]$ the function $y \mapsto F_x^A(y) := K_A(x,[0,y])$ will be called *conditional distribution function of A at x*. For more details and properties of conditional expectation, regular conditional distributions, and general disintegration see [15] and [18].

Viewing copulas as special Markov kernels (fulfilling eq. (14.3)) has proved surprisingly useful in the past. As an example, translating the so-called $*$-product of copulas introduced in [3] to the Markov kernel setting directly yields that the Markov kernel K_{A*B} of $A*B$ is nothing else but the standard composition of the Markov kernels K_A and K_B as well known in the context Markov processes, i.e.

$$K(x,F) := \int_{[0,1]} K_B(y,F) K_A(x,dy) \tag{14.4}$$

is a Markov kernel of $A*B$ (see [36]). For additional examples underlining the usefulness of Markov kernels we refer, for instance, to [7, 10, 36, 37] and the references therein.

A copula A is called *completely dependent* (see [21, 35]) if there exists a λ-preserving transformation $h : [0,1] \to [0,1]$ such that $K(x,E) = \mathbf{1}_E(h(x))$ is a Markov kernel of A. Slightly extending [35, Lemma 10] the subsequent characterization of complete dependence can be proved:

Lemma 14.1. *The following assertions are equivalent:*

(d1) A is completely dependent .
(d2) For $\mathscr{P}^{X \otimes Y} = \mu_A$ *the conditional variance* $\mathbb{V}(Y|X=x)$ *of* Y *given* X *fulfills* $\mathbb{V}(Y|X=x) = 0$ *for* λ*-a.e.* $x \in [0,1]$.
(d3) For λ*-a.e.* $x \in [0,1]$ *the conditional distribution function* F_x^A *is* $\{0,1\}$*-valued.*
(d4) There exists a λ*-preserving transformation* $h : [0,1] \to [0,1]$ *such that* $A(x,y) = \lambda\big([0,x] \cap h^{-1}([0,y])\big)$ *for all* $(x,y) \in [0,1]^2$.
(d5) There exists a λ*-preserving transformation* $h : [0,1] \to [0,1]$ *with* $\mu_A(\Gamma(h)) = 1$, *whereby* $\Gamma(h) = \{(x,h(x)) : x \in [0,1]\} \in \mathscr{B}([0,1]^2)$ *denotes the graph of h.*
(d6) A is left invertible w.r.t. the star product, i.e. there exists a copula $B \in \mathscr{C}$ *with* $B*A = M$.

In the sequel $\mathscr{T}$ will denote the family of all λ-preserving transformations h on $[0,1]$, $\mathscr{T}_b$ the subclass of all λ-preserving bijections, and $\mathscr{T}_s$ the subclass of all

λ-preserving, piecewise linear bijections. For every $h \in \mathscr{T}$ we will let A_h denote the corresponding completely dependent copula, $\mathscr{C}_d$ will denote the family of all completely dependent copulas. In case of $h \in \mathscr{T}_s$ we will refer to A_h as a (classical) *shuffle of M* and in case of $h \in \mathscr{T}_b$ the copula A_h will be called *mutually completely dependent*.

14.3 Quantifying dependence

According to [25] the class of all shuffles of M is dense in the compact metric space $(\mathscr{C}, d_\infty)$. Hence, as pointed out explicitly already in [22], the uniform distance d_∞ is not clearly 'distinguishing different types of statistical dependence' and the same holds for every dependence measure that is continuous w.r.t. d_∞, including Schweizer and Wolffs σ (see [25, 30]). Viewing copulas as Markov kernels allows for a simple way to construct stronger metrics on $\mathscr{C}$ that strictly separate extremal dependence concepts, i.e. that of independence and that of complete dependence. Following [35] and setting

$$D_p^p(A,B) := \int_{[0,1]^2} \big|K_A(x,[0,y]) - K_B(x,[0,y])\big|^p d\lambda_2(x,y) \tag{14.5}$$

defines a metric D_p on $\mathscr{C}$ for every $p \in [1,\infty)$. For a generalization to the multivariate setting we refer to [7, 9]. According to [35] the metrics D_2 and D_1 induce the same topology on $\mathscr{C}$ and the resulting metric spaces $(\mathscr{C}, D_1)$ and $(\mathscr{C}, D_2)$ are complete and separable. It is straightforward to verify that the same assertions hold for $(\mathscr{C}, D_p)$ and arbitrary $p \in [1,\infty)$. In fact, using Hölder inequality and considering $\big|K_A(x,[0,y]) - K_B(x,[0,y])\big| \leq 1$ directly yields

$$D_p^p(A,B) \leq D_1(A,B) \leq D_p(A,B), \tag{14.6}$$

from which separability and completeness of $(\mathscr{C}, D_p)$ directly follows from separability and completeness of $(\mathscr{C}, D_1)$. Although all metrics D_p induce the same topology on $\mathscr{C}$ they are not equivalent - the following result holds:

Lemma 14.2. *For any pair $p,q \in [1,\infty)$ with $p \neq q$ the metrics D_p and D_q are not equivalent.*

Proof: We start by showing that D_p^p and D_q^q coincide on $\mathscr{C}_d$ and consider $h_1, h_2 \in \mathscr{T}$:

$$\begin{aligned} D_p^p(A_{h_1}, A_{h_2}) &= \int_{[0,1]^2} \big|\mathbf{1}_{[0,y]}(h_1(x)) - \mathbf{1}_{[0,y]}(h_2(x))\big|^p d\lambda_2(x,y) = \|h_1 - h_2\|_1 \\ &= \int_{[0,1]^2} \big|\mathbf{1}_{[0,y]}(h_1(x)) - \mathbf{1}_{[0,y]}(h_2(x))\big|^q d\lambda_2(x,y) = D_q^q(A_{h_1}, A_{h_2}) \end{aligned}$$

From this, considering $q = 1$ it also follows that the first inequality in (14.6) can not be improved. For every $n \in \mathbb{N}$ define $h_n \in \mathscr{T}_s$ (see Figure 14.1) by

$$S_n(x) = \begin{cases} x + \left(1 - \frac{1}{2^n}\right) & \text{if } x \in \left(0, \frac{1}{2^n}\right] \\ x - \left(1 - \frac{1}{2^n}\right) & \text{if } x \in \left(1 - \frac{1}{2^n}, 1\right] \\ x & \text{otherwise,} \end{cases}$$

and set $h = id_{[0,1]} \in \mathscr{T}_s$. Then we get

$$D_p^p(A_{h_n}, M) = D_q^q(A_{h_n}, M) = \|h_n - h\|_1 = 2\int_{[0,\frac{1}{2^n}]} \left(1 - \frac{1}{2^n}\right) d\lambda = \frac{1}{2^{n-1}}\left(1 - \frac{1}{2^n}\right). \tag{14.7}$$

Suppose now that $p > q$. Then eq. (14.7) implies that the quotient $\frac{D_p(A_n,M)}{D_q(A_n,M)}$ is unbounded in n, so D_q and D_p can not be equivalent metrics. ■

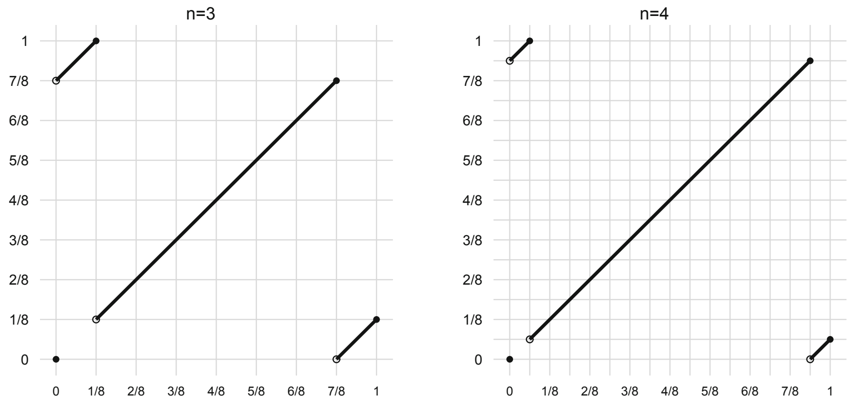

Fig. 14.1: The transformations h_n used in the proof of Lemma 14.2

In [35] the metric D_1 mainly served as a vehicle to construct the dependence measure $\zeta_1(A) = 3D_1(A, \Pi)$ for every $A \in \mathscr{C}$. The most important properties of ζ_1 are summarized in the following theorem.

Theorem 14.1 ([35]). *For every $A \in \mathscr{C}$ we have $\zeta_1(A) \in [0,1]$. Additionally, $\zeta_1(A) = 1$ holds if and only if $A \in \mathscr{C}_d$, and $\zeta_1(A) = 0$ implies $A = \Pi$. In other words: Exclusively all completely dependent copulas are assigned maximum dependence measure and the product copula is the only copula with zero dependence.*

Taking into account eq. (14.6) a similar result can be expected for the dependence measure ζ_p, defined by $\zeta_p(A) = c_p D_p(A, \Pi)$ for every $p \in [1, \infty)$. Thereby c_p is a normalizing constant assuring $\max_{A \in \mathscr{C}} \zeta_p(A) = 1$. In the sequel we will state and prove the result for the case $p = 2$ which allows for a more elegant proof than the original one for D_1.

Theorem 14.2. *For every* $A \in \mathscr{C}$ *we have* $D_2^2(A,\Pi) \leq 1/6$ *with equality if and only if* $A \in \mathscr{C}_d$.

Proof: Fix $A \in \mathscr{C}$ and $y \in [0,1]$, and define a random variable Z_y on the probability space $([0,1], \mathscr{B}([0,1], \lambda) \to [0,1]$ by

$$Z_y(x) := K_A(x,[0,y]).$$

Then

$$\mathbb{E}(Z_y) = \int_{[0,1]} K_A(x,[0,y])d\lambda(x) = \mu_A\big([0,1]\times[0,y]\big) = y$$

holds. Considering

$$\begin{aligned}\int_{[0,1]} \big(K_A(x,[0,y]) - y\big)^2 d\lambda(x) &= \mathbb{V}(Z_y) = \mathbb{E}(Z_y^2) - (\mathbb{E}(Z_y))^2 = \mathbb{E}(Z_y^2) - y^2 \\ &\leq \mathbb{E}(Z_y) - y^2 = y - y^2 \end{aligned} \quad (14.8)$$

we directly get $D_2^2(A,\Pi) \leq \int_{[0,1]}(y-y^2)d\lambda(y) = \frac{1}{6}$, which completes the proof of the first assertion.

Since ineq. (14.8) becomes an equality if and only if $\mathbb{E}(Z_y^2) = \mathbb{E}(Z_y)$ holds, it follows that $D_2^2(A,\Pi) = \frac{1}{6}$ is equivalent to the condition that $Z_y^2 = Z_y$ holds λ-a.e. The latter, however, is obviously equivalent to the existence of a set $\Lambda_y \in \mathscr{B}([0,1])$ with $\lambda(\Lambda_y) = 1$ such that $Z_y(x) = K_A(x,[0,y]) \in \{0,1\}$ for every $x \in \Lambda_y$.

Suppose now that $D_2^2(A,\Pi) = 1/6$. Repeating the last argument we can find a set $\Lambda \in \mathscr{B}([0,1])$ fulfilling $\lambda(\Lambda) = 1$ such that $F_x^A(y) = K_A(x,[0,y]) \in \{0,1\}$ holds for every $x \in \Lambda$ and every $y \in \mathbb{Q} \cap [0,1]$. Using right-continuity of distribution functions we immediately get that λ-a.e. conditional distribution functions F_x^A are $\{0,1\}$-valued, so Lemma 14.1 implies that A is completely dependent. Since, on the other hand, it is straightforward to verify $D_2^2(A_h,\Pi) = \frac{1}{6}$ for every $h \in \mathscr{T}$, the proof is complete. ■

As direct consequence of Theorem 14.1, setting $\zeta_2(A) = \sqrt{6}D_2(A,\Pi)$ we get the following result:

Proposition 14.1. *For every* $A \in \mathscr{C}$ *we have* $\zeta_2(A) \in [0,1]$. *Additionally,* $\zeta_2(A) = 1$ *holds if and only if* $A \in \mathscr{C}_d$ *and* $\zeta_2(A) = 0$ *implies* $A = \Pi$. *In other words: Exclusively all completely dependent copulas are assigned maximum* ζ_2*-value and the product copula is the only copula with* $\zeta_2(A) = 0$.

Independence of two random variables X,Y is a symmetric concept (knowing X does not change our knowledge about Y and vice versa) - nevertheless, from the author's point of view, notions quantifying dependence should not automatically be symmetric since in many situations one might also be interested in understanding causal effects between X and Y and the dependence structure might be strongly asymmetric. The latter is the case, for instance, for the copula $A_h \in \mathscr{C}_d$ with $h \in \mathscr{T}$ being the transformation $h(x) = 2^n(mod 1)$ for large $n \in \mathbb{N}$. Furthermore, having a unidirectional (i.e. non-mutual) dependence measure one can easily construct a mutual one: The mutual dependence measure ω studied by Siburg and Stoimenov (see [33]), for instance, can easily be expressed in terms of ζ_2 as

$$\omega^2(A) = 3\left(\zeta_2^2(A) + \zeta_2^2(A^t)\right), \tag{14.9}$$

whereby A^t denotes the transpose of A, defined by $A^t(x,y) = A(x,y)$. Proposition 14.1 directly yields that $\omega(A) = 1$ if and only if both A and A^t are completely dependent, i.e. if and only if A is mutually completely dependent.

Recently, dependence measures for (absolutely continuous) random vectors have been introduced using similar ideas as the afore-mentioned ones, see [2] and the references therein. For a dependence measure based on conditional variance we refer to [16].

14.4 Complete dependence in the context of optimization

Remember that a point x in a convex set Ω is called an extreme point of Ω if it is not an interior point of any line segment lying entirely in Ω, i.e. if $x = \alpha y + (1-\alpha)z$ for $y,z \in \Omega$ and $\alpha \in [0,1]$ implies $x = y$ or $x = z$.

It is straightforward to show that every completely dependent copula is an extreme point of $\mathscr{C}$. As a consequence, in the metric space $(\mathscr{C}, d_\infty)$ the set $Ex(\mathscr{C})$ of all extreme points of $\mathscr{C}$ is dense. Although a full and handy characterization of the set $Ex(\mathscr{C})$ seems out of reach it is known, that there are extreme points which are not completely dependent. In fact, in [31] (also see [26, 32]) so-called hairpin copulas, which concentrate their mass on the graphs of two functions were studied and shown to be elements of $Ex(\mathscr{C})$. For the generalization of hairpin copulas to the multivariate setting we refer to [5]. To the best of the author's knowledge the most striking example of an extreme point of $\mathscr{C}$ was given in [23], where the author proved the existence of a copula $A \in \mathscr{C}$ such that $Supp(\mu_A) = [0,1]^2$.

Extreme points of $\mathscr{C}$ naturally come into play in the context of optimization problems of the form

$$\overline{M}_H := \sup_{A \in \mathscr{C}} \int_{[0,1]^2} H(x,y) d\mu_A(x,y) \tag{14.10}$$

whereby H is a non-negative measurable function on $[0,1]^2$. In fact, if there is a unique $A \in \mathscr{C}$ attaining $\overline{M}_H$ then A has to be an extreme point of $\mathscr{C}$. Additionally, if H is continuous (hence bounded) then $\overline{M}_H$ is attained and, according to the Bauer Maximum Principle (see [1]), the maximum is also attained by an extreme point.

14.4.1 Distributions with fixed marginals maximizing the mass of the endograph of a function

Suppose that F and G are (continuous) distribution functions of two default times and let $\mathscr{F}_{F,G}$ denote the Fréchet class of F,G (i.e. the family of all two-dimensional distribution functions having F and G as marginals). It is well known from coupling theory (see [34]) that there exists a maximal coupling, i.e. a two-dimensional distribution function $H \in \mathscr{F}_{F,G}$ such that for the case of $(X,Y) \sim H$ the probability of a joint default $\mathbb{P}(X = Y)$ is maximal. Translating to the class of copulas maximizing the probability of a joint default means calculating $\sup_{A\in\mathscr{C}} \mu_A(\Gamma(T))$ for $T : [0,1] \to [0,1]$ being defined by $T = G \circ F^-$, F^- denoting the quasi-inverse of F and $\Gamma(T)$ the graph of T. Using coupling theory we can find a (not necessarily unique) copula A_0 with

$$\overline{M}_{\mathbf{1}_{\Gamma(T)}} = \sup_{A\in\mathscr{C}} \int_{[0,1]^2} \mathbf{1}_{\Gamma(T)} d\mu_A(x,y) = \sup_{A\in\mathscr{C}} \mu_A(\Gamma(T)) = \mu_{A_0}(\Gamma(T)) \tag{14.11}$$

that can even be computed in closed form. Additionally, applying the results from [28] or via manual calculations a very simple formula for $\overline{M}_{\mathbf{1}_{\Gamma(T)}}$ can be derived. Returning to the original problem of maximizing the probability of a joint default, considering $(U,V) \sim A_0$ and setting $(X,Y) = (F^- \circ U, G^- \circ V)$, it follows that the pair (X,Y) has marginal distribution functions F and G and maximizes the joint default probability. In general, A_0 is not completely dependent unless F and G coincide.

Slightly modifying the optimization problem and maximizing $\mathbb{P}(Y \leq X)$ instead of $\mathbb{P}(Y = X)$ brings us back to complete dependence. In fact, proceeding analogously as before and setting

$$\Gamma^{\leq}(T) = \big\{(x,y) \in [0,1]^2 : y \leq T(x)\big\} \in \mathscr{B}([0,1]^2) \tag{14.12}$$

the following results can be derived manually or using the results in [28]:

Theorem 14.3 ([24]). *For every non-decreasing* $T : [0,1] \to [0,1]$ *we have*

$$\overline{M}_{\mathbf{1}_{\Gamma^{\leq}(T)}} = 1 + \inf_{x\in[0,1]} (T(x) - x). \tag{14.13}$$

Additionally, there exists a shuffle $A_R \in \mathscr{C}_d$ *fulfilling* $\mu_{A_R}(\Gamma^{\leq}(T)) = \overline{M}_{\mathbf{1}_{\Gamma^{\leq}(T)}}$.

In other words, given continuous distribution functions F and G, considering $(U,V) \sim A_R$ and setting $(X,Y) = (F^- \circ U, G^- \circ V)$, it follows that for the pair (X,Y) the quantity $\mathbb{P}(Y \leq X)$ is maximal.

Example 14.1. We consider a very simple situation illustrating Theorem 14.3: Choosing F as the distribution function of $\mathscr{U}(0,1)$ and $G = \Phi$ as the distribution function of $\mathscr{N}(0,1)$ we immediately get $T = \Phi$ as well as $\overline{M}_{\mathbf{1}_{\Gamma^{\leq}(T)}} = \Phi(1) \approx 0.841$. Figure 14.2 denotes a sample of the random vector (X,Y) with marginal distribution functions F and G for which $\mathbb{P}(Y \leq X)$ is maximal.

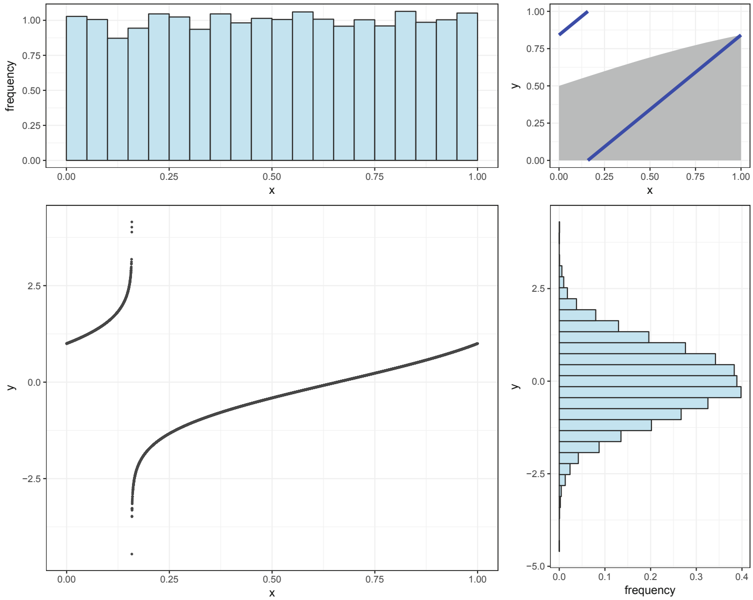

Fig. 14.2: Sample of size $n = 10.000$ of (X,Y) as in Example 14.2 and the corresponding marginal histrograms; the upper right panel depicts the endograph of T (gray) and the shuffle A_R according to Theorem 14.3 (blue).

14.4.2 A maximization problem from uniform distribution theory

Recall that a sequence $(x_n)_{n\in\mathbb{N}}$ in $[0,1]$ is called uniformly distributed if the induced empirical measure $\vartheta_n := \frac{1}{n}\sum_{i=1}^n \delta_{x_i}$ converges weakly to λ on $[0,1]$ for $n \to \infty$. For background on uniform distribution theory we refer to [4, 20].

Following [14] and the references therein one particularly interesting problem in the context of uniform distribution theory is the following one: Given uniformly distributed sequences $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ in $[0,1]$ and a real-valued continuous function H on $[0,1]^2$, determine

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{i=1}^n H(x_i,y_i) \quad \text{and} \quad \liminf_{n\to\infty} \frac{1}{n}\sum_{i=1}^n H(x_i,y_i). \tag{14.14}$$

It is a straightforward exercise to show that for every accumulation point a of the sequence $(\frac{1}{n}\sum_{i=1}^n H(x_i,y_i))_{n\in\mathbb{N}}$ there exists a copula $A \in \mathscr{C}$ such that

$$a = \int_{[0,1]^2} H\, d\mu_A$$

holds. In other words, for calculating the quantities in eq (14.14) it suffices to calculate $\overline{M}_H$ and $-\overline{M}_{-H}$.

We will proceed as in [14] and consider the following special case for H: Fix $y_0 \in (0,1)$ and, suppose that H is continuous on $[0,1]^2$ and that $\frac{\partial^2 H(x,y)}{\partial y \partial x} > 0$ on $(0,1)\times(0,y_0)$ as well that $\frac{\partial^2 H(x,y)}{\partial y \partial x} < 0$ on $(0,1)\times(y_0,1)$ holds. Moreover we will let $\mathscr{S}_{y_0}$ denote the class of all y_0-sections of copulas, i.e. the family of all maps of the form $x \mapsto C(x,y_0)$, $x \in [0,1]$, with $C \in \mathscr{C}$. It is straightforward to verify that $s \in \mathscr{S}_{y_0}$ if and only if s fulfills the following three properties:

- $s(0) = 0, s(1) = y_0$
- s is non-decreasing and Lipschitz continuous with Lipschitz constant $L = 1$
- s fulfills $s(x) \in [W(x,y_0), M(x,y_0)]$ for all $x \in [0,1]$

For every $s \in \mathscr{S}_{y_0}$ in the rest of this section the copula $C^s \in \mathscr{C}$ be defined by

$$C^s(x,y) = \begin{cases} M(s(x),y) & \text{if } (x,y) \in [0,1]\times[0,y_0] \\ s(x) + (1-y_0)W\left(\frac{x-s(x)}{1-y_0}, \frac{y-y_0}{1-y_0}\right) & \text{if } (x,y) \in [0,1]\times(y_0,1]. \end{cases} \tag{14.15}$$

Obviously the y_0-section of C^s coincides with s and, setting $\overline{s}(x) = 1-(x-s(x))$ the copula C^s concentrates its mass on $\Gamma(s)\cup\Gamma(\overline{s})$ in the sense that

$$\mu_{C_s}(\Gamma(s)\cup\Gamma(\overline{s})) = 1.$$

The following reduction result can be shown:

Theorem 14.4 ([14]). *Under the afore-mentioned assumptions on H the following equality holds:*

$$\overline{M}_H = \max_{s\in\mathscr{S}_{y_0}} \int_{[0,1]^2} H \, d\mu_{C^s} \tag{14.16}$$

For general $s \in \mathscr{S}_{y_0}$ obviously the copula C^s need not be completely dependent. If, however, s is strictly increasing with $s' < 1$ then we get $(C^s)^t \in \mathscr{C}_d$, i.e. the transpose of C^s is completely dependent (also see Figure 14.3). We conclude this section with an example illustrating that a copula of the latter type may even be the unique maximizer.

Example 14.2. Consider $y_0 \in [\frac{1}{2},1)$ and suppose that H is given by

$$H = \begin{cases} xy & \text{if } (x,y) \in [0,1]\times[0,y_0] \\ xy_0 - x(y-y_0) & \text{if } (x,y) \in [0,1]\times(y_0,1]. \end{cases} \tag{14.17}$$

For arbitrary $s \in \mathscr{S}_{y_0}$ applying Theorem 14.4 and using integration by parts we finally get

$$\int_{[0,1]^2} H \, d\mu_{C^s} = y_0^2 - \frac{1}{2}\left\{ \int_{[0,1]} \left(s^2(x) + (2y_0 - 1 + x - s(x))^2\right) d\lambda(x) \right\}.$$

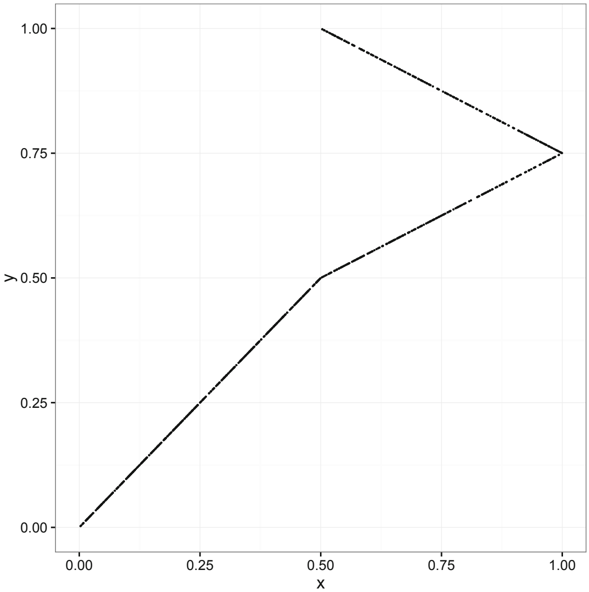

Fig. 14.3: Sample of size $n = 1.000$ of the unique maximizer C^s in Example 14.2.

For fixed x the integrand becomes minimal if $s(x) = y_0 - \frac{1}{2} + \frac{x}{2}$. The function $s_1 : x \mapsto y_0 - \frac{1}{2} + \frac{x}{2}$ is a global minimizer of the integral which, however, only lies in $\mathscr{S}_{y_0}$ for $y_0 = \frac{1}{2}$. It is straightforward to verify that for $y_0 \geq \frac{1}{2}$ the (piecewise linear) function h, defined by

$$s(x) = \begin{cases} x \text{ if } x \in [0, 2y_0 - 1] \\ y_0 - \frac{1}{2} + \frac{x}{2} \text{ if } x \in (2y_0 - 1, 1] \end{cases}$$

is the unique minimizer of the integral in eq. (14.17). As a consequence, the corresponding copula C^s, which fulfills $(C^s)^t \in \mathscr{C}_d$ is the unique copula attaining $\overline{M}_H$. Figure 14.3 depicts a sample of the corresponding copula C^s for the case $y_0 = \frac{3}{4}$.

14.5 A typical copula is (mutually) completely dependent

As already mentioned in Section 3 the set of all shuffles of M is dense in the metric space $(\mathscr{C}, d_\infty)$ (but nowhere dense in the metric space $(\mathscr{C}, D_1)$, see [35]). On the one hand, non-absolutely continuous copulas naturally appear in various problems, on the other hand, possibly due to their handy structure, absolutely continuous copulas are certainly not underrepresented in the literature.

Topology offers a way to quantify the size of sets in a binary manner through Baire categories (see [27]): A subset E of a general metric space (Ω, d) is considered small if it is of first category or meager, i.e. if it is the countable union of sets E_i whose topological closer has empty interior. If E is not of first category then, by definition, E is said to be of second category. Finally, if E is meager then E^c is considered big and referred to as co-meager. Following [6] we will call elements of a meager set E *atypical* and elements of a co-meager set *typical*. With this topological notions various interesting results can be shown, e.g. that the family $\mathscr{C}_{abs}$ of all absolutely continuous copulas is of first category both in $(\mathscr{C}, d_\infty)$ and in $(\mathscr{C}, D_1)$. Considering completeness of $(\mathscr{C}, d_\infty)$ and $(\mathscr{C}, D_1)$ it directly follows that the family $\mathscr{C}_{abs}^c = \mathscr{C} \setminus \mathscr{C}_{abs}$ of all copulas with non-degenerated singular component is co-meager and of second category in $(\mathscr{C}, d_\infty)$ and in $(\mathscr{C}, D_1)$, i.e. a typical copula has a non-degenerated singular component. As a matter of fact, the following much stronger result holds (and can be extended to the general multivariate setting):

Theorem 14.5 ([6]). *The family $\mathscr{C}_{sing}$ of all purely singular copulas is co-meager (hence of second category) in $(\mathscr{C}, d_\infty)$.*

In other words: A typical copula in $(\mathscr{C}, d_\infty)$ has no absolutely continuous component. It remains an open question if $\mathscr{C}_{sing}$ is also co-meager in $(\mathscr{C}, D_1)$. As the authors of [6] discovered recently, Theorem 14.5 is not even close to the end of the story - the following striking result was proved already in 1968:

Theorem 14.6 ([17]). *$\mathscr{C}_d$ is co-meager (hence of second category) in $(\mathscr{C}, d_\infty)$.*

Based on the elegant proof of Theorem 14.6 as given in [17] one gets the following even more striking corollary, saying that *in $(\mathscr{C}, d_\infty)$ a typical copula is mutually completely dependent*, without any difficulty:

Corollary 14.1. *The family of all mutually completely dependent copulas is co-meager (hence of second category) in $(\mathscr{C}, d_\infty)$.*

14.6 Sharp inequalities between Kendall's τ and Spearman's ρ

This section first recalls the main results from [29] and then sketches why 'complete dependence everywhere' particularly holds true in the situation of Kendall's τ and Spearman's ρ.

Suppose that X, Y are random variables with continuous distribution functions F and G respectively. Then Spearman's ρ is defined as the Pearson correlation coefficient of the $\mathscr{U}(0,1)$-distributed random variables $U := F \circ X$ and $V := G \circ Y$ and Kendall's τ is given by the probability of concordance minus the probability of discordance, i.e.

$$\begin{aligned} \rho(X,Y) &= 12\big(\mathbb{E}(UV) - \tfrac{1}{4}\big) \\ \tau(X,Y) &= \mathbb{P}\big((X_1 - X_2)(Y_1 - Y_2) > 0\big) - \mathbb{P}\big((X_1 - X_2)(Y_1 - Y_2) < 0\big), \end{aligned}$$

where (X_1,Y_1) and (X_2,Y_2) are independent and have the same distribution as (X,Y). Clearly τ and ρ are the two most famous nonparametric measures of concordance. Both measures are scale invariant and only depend on the underlying (uniquely determined) copula A of (X,Y). It is well known and straightforward to verify (see [25]) that $\tau(X,Y)$ and $\rho(X,Y)$ can be expressed in terms of the underlying copula A as

$$\tau(X,Y) = 4\int_{[0,1]^2} A(x,y)\, d\mu_A(x,y) - 1 =: \tau(A) \tag{14.18}$$

$$\rho(X,Y) = 12\int_{[0,1]^2} xy\, d\mu_A(x,y) - 3 =: \rho(A) \tag{14.19}$$

Considering that τ and ρ quantify different aspects of the underlying dependence structure, it is natural to ask how much they can differ. Since the 1950s two universal inequalities between τ and ρ are known - the first one goes back to Daniels ([3]), the second one to Durbin and Stuart ([8]); for proofs alternative to the original ones see [19, 12, 25].

$$|3\tau - 2\rho| \leq 1 \tag{14.20}$$

$$\frac{(1+\tau)^2}{2} - 1 \leq \rho \leq 1 - \frac{(1-\tau)^2}{2} \tag{14.21}$$

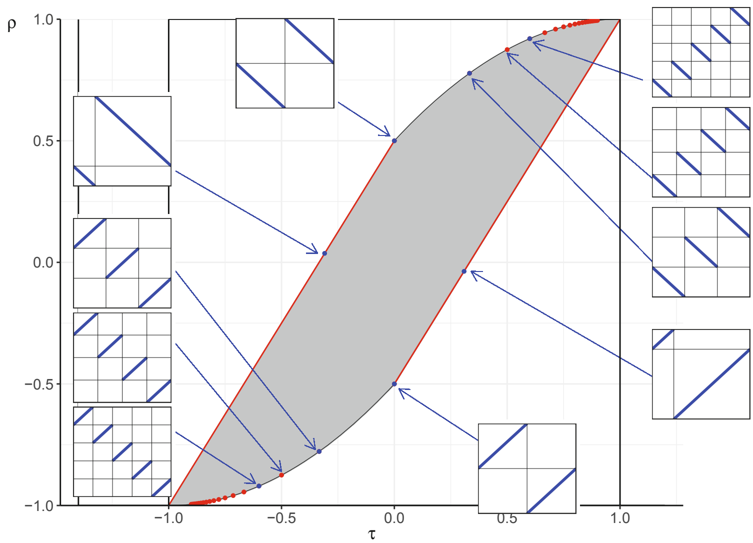

Fig. 14.4: The classical τ-ρ-region Ω_0 and some copulas (distributing mass uniformly on the blue segments) for which the inequality by Durbin and Stuart is sharp.

The inequalities together yield the set Ω_0 (see Figure 14.4) which, following [29] we will refer to as *classical τ-ρ region* in the sequel. Daniels' inequality was known to be sharp [25] whereas the first part of the inequality by Durbin and Stuart was only known to be sharp at the points $p_n = (-1+\frac{2}{n}, -1+\frac{2}{n^2})$ with $n \geq 2$ (which, using symmetry, is to say that the second part is sharp at the points $-p_n$). Although both inequalities were known since the 1950s and the interrelation between τ and ρ keeps receiving much attention in recent years (particularly in the context of the Hutchinson-Lai conjecture [11, 13]), only very recently a full characterization of the *exact τ-ρ region* Ω, defined by

$$\begin{aligned} \Omega &= \big\{(\tau(X,Y),\rho(X,Y)) : X,Y \text{ continuous random variables}\big\} \qquad (14.22) \\ &= \big\{(\tau(A),\rho(A)) : A \in \mathscr{C}\big\}, \end{aligned}$$

was given in [29]. One direct consequence of this characterization is the fact that inequality by Durbin and Stuart is not sharp outside the points $\pm p_n$.

Throughout the entire proof shuffles (hence complete dependence) played a crucial role: The authors first calculated τ and ρ for so-called prototypes, which, loosely speaking, are shuffles consisting of $n-1$ segments of equal length and a shorter one, arranged in decreasing order similar to the shuffles depicted in Figure 14.4. Based on these prototypes they defined $\Phi_n : [-1+\frac{2}{n}, 1] \to [-1,1]$ by

$$\Phi_n(x) = -1 - \frac{4}{n^2} + \frac{3}{n} + \frac{3x}{n} - \frac{n-2}{\sqrt{2}n^2\sqrt{n-1}}(n-2+nx)^{3/2} \qquad (14.23)$$

and then set

$$\Phi(x) = \begin{cases} -1 & \text{if } x = -1, \\ \Phi_n(x) & \text{if } x \in \left[\frac{2-n}{n}, \frac{2-(n-1)}{n-1}\right] \text{ for some } n \geq 2. \end{cases} \qquad (14.24)$$

Based on Φ the set Ω can be characterized as follows (this characterization was already conjectured by Manuel Úbeda-Flores in an unpublished working paper in 2011):

Theorem 14.7 ([29]). *The following equality holds:*

$$\Omega = \big\{(x,y) \in [-1,1]^2 : \Phi(x) \leq y \leq -\Phi(-x)\big\} \qquad (14.25)$$

In particular, Ω is compact but not convex. For an animation showing for which copulas A we have $(\tau(A),\rho(A)) \in \partial\Omega$, where $\partial\Omega$ denotes the topological boundary of Ω, we refer to http://www.trutschnig.net/tau-rho-boundary.pdf

Returning to complete dependence, notice that continuity of τ and ρ with respect to d_∞ directly yields that $\{(\tau(A_h),\rho(A_h)) : h \in \mathscr{T}_s\}$ is dense in Ω. The proof of Theorem 14.7, however, produced the by-product that only for prototypes $A \in \mathscr{C}_d$ we can have $(\tau(A),\rho(A)) \in \partial\Omega$. In fact, using a homotopy argument it was possible to show the following corollary which underlines the importance of (mutual) complete dependence yet again.

Corollary 14.2. *For every point* $(x,y) \in \Omega$ *there exists a transformation* $h \in \mathcal{T}_s$ *such that we have* $(\tau(A_h), \rho(A_h)) = (x,y)$.

As pointed out in [29, Section 6] characterizing the exact τ-ρ-region for standard subclasses of copulas may in some cases be even more difficult than determining Ω was. The main reason for this fact is that not in all subclasses of $\mathscr{C}$ we may find dense subsets consisting of elements B for which $\tau(B)$ and $\rho(B)$ reduce to handy formulas (as it is the case for shuffles of M). The author conjectures, however, that the classical Hutchinson-Lai inequalities are not sharp for the class of all extreme-value copulas and that it might be possible to derive sharper inequalities in the near future.

Acknowledgements The authory gratefully acknowledge the support of the grant MTM2014-60594-P (partially supported by FEDER) from the Spanish Ministry of Economy and Competitiveness.

References

1. Aliprantis, C.D., Border, K.C.: Infinite dimensional analysis, third edn. Springer, Berlin (2006). A hitchhiker's guide
2. Boonmee, T., Tasena, S.: Measure of complete dependence of random vectors. J. Math. Anal. Appl. **443**(1), 585–595 (2016)
3. Daniels, H.E.: Rank correlation and population models. J. Roy. Statist. Soc. Ser. B. **12**, 171–181 (1950)
4. Drmota, M., Tichy, R.F.: Sequences, discrepancies and applications, *Lecture Notes in Mathematics*, vol. 1651. Springer-Verlag, Berlin (1997)
5. Durante, F., Fernández Sánchez, J., Trutschnig, W.: Multivariate copulas with hairpin support. J. Multivariate Anal. **130**, 323–334 (2014)
6. Durante, F., Fernández-Sánchez, J., Trutschnig, W.: A typical copula is singular. J. Math. Anal. Appl. **430**(1), 517–527 (2015)
7. Durante, F., Sempi, C.: Principles of copula theory. CRC Press, Boca Raton, FL (2016)
8. Durbin, J., Stuart, A.: Inversions and rank correlation coefficients. J. Roy. Statist. Soc. Ser. B. **13**, 303–309 (1951)
9. Fernández Sánchez, J., Trutschnig, W.: Conditioning-based metrics on the space of multivariate copulas and their interrelation with uniform and levelwise convergence and iterated function systems. J. Theoret. Probab. **28**(4), 1311–1336 (2015)
10. Fernández Sánchez, J., Trutschnig, W.: Singularity aspects of Archimedean copulas. J. Math. Anal. Appl. **432**(1), 103–113 (2015)
11. Fredricks, G.A., Nelsen, R.B.: On the relationship between Spearman's rho and Kendall's tau for pairs of continuous random variables. J. Statist. Plann. Inference **137**(7), 2143–2150 (2007)
12. Genest, C., Nešlehová, J.: Analytical proofs of classical inequalities between Spearman's ρ and Kendall's τ. J. Statist. Plann. Inference **139**(11), 3795–3798 (2009)
13. Hürlimann, W.: Hutchinson-Lai's conjecture for bivariate extreme value copulas. Statist. Probab. Lett. **61**(2), 191–198 (2003)
14. Iacò, M.R., Thonhauser, S., Tichy, R.F.: Distribution functions, extremal limits and optimal transport. Indag. Math. (N.S.) **26**(5), 823–841 (2015)
15. Kallenberg, O.: Foundations of modern probability, second edn. Probability and its Applications (New York). Springer-Verlag, New York (2002)

16. Kamnitui, N., Santiwipanont, T., Sumetkijakan, S.: Dependence measuring from conditional variances. Depend. Model. **3**(1), 98–112 (2015)
17. Kim, C.W.: Uniform approximation of doubly stochastic operators. Pacific J. Math. **26**, 515–527 (1968)
18. Klenke, A.: Probability theory, second edn. Universitext. Springer, London (2014)
19. Kruskal, W.H.: Ordinal measures of association. J. Amer. Statist. Assoc. **53**, 814–861 (1958)
20. Kuipers, L., Niederreiter, H.: Uniform distribution of sequences. Wiley-Interscience [John Wiley & Sons], New York-London-Sydney (1974). Pure and Applied Mathematics
21. Lancaster, H.O.: Correlation and complete dependence of random variables. Ann. Math. Statist. **34**, 1315–1321 (1963)
22. Li, X., Mikusiński, P., Taylor, M.D.: Strong approximation of copulas. J. Math. Anal. Appl. **225**(2), 608–623 (1998)
23. Losert, V.: Counterexamples to some conjectures about doubly stochastic measures. Pacific J. Math. **99**(2), 387–397 (1982)
24. Mroz, T., Trutschnig, W., Fernández Sánchez, J.: Distributions with fixed marginals maximizing the mass of the endograph of a function. preprint on aArXiv: 1602.05807
25. Nelsen, R.B.: An Introduction to Copulas, second edn. Springer Series in Statistics. Springer, New York (2006)
26. Nelsen, R.B., Fredricks, G.A.: Diagonal copulas. In: Distributions with given marginals and moment problems (Prague, 1996), pp. 121–128. Kluwer Acad. Publ., Dordrecht (1997)
27. Oxtoby, J.C.: Measure and category. A survey of the analogies between topological and measure spaces. Springer-Verlag, New York-Berlin (1971). Graduate Texts in Mathematics, Vol.2
28. Rüschendorf, L.: Random variables with maximum sums. Adv. in Appl. Probab. **14**(3), 623–632 (1982)
29. Schreyer, M., Paulin, R., Trutschnig, W.: On the exact region determined by Kendall's tau and Spearman's rho. J. Roy. Statist. Soc. Ser. B. **79**(2), 613–633 (2017)
30. Schweizer, B., Wolff, E.F.: On nonparametric measures of dependence for random variables. Ann. Statist. **9**(4), 879–885 (1981)
31. Seethoff, T.L., Shiflett, R.C.: Doubly stochastic measures with prescribed support. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **41**(4), 283–288 (1977/78)
32. Sherwood, H., Taylor, M.D.: Doubly stochastic measures with hairpin support. Probab. Theory Related Fields **78**(4), 617–626 (1988)
33. Siburg, K.F., Stoimenov, P.A.: A measure of mutual complete dependence. Metrika **71**(2), 239–251 (2010)
34. Thorisson, H.: Coupling, stationarity, and regeneration. Probability and its Applications (New York). Springer-Verlag, New York (2000)
35. Trutschnig, W.: On a strong metric on the space of copulas and its induced dependence measure. J. Math. Anal. Appl. **384**(2), 690–705 (2011)
36. Trutschnig, W., Fernández Sánchez, J.: Idempotent and multivariate copulas with fractal support. J. Statist. Plann. Inference **142**(12), 3086–3096 (2012)
37. Trutschnig, W., Fernández Sánchez, J.: Copulas with continuous, strictly increasing singular conditional distribution functions. J. Math. Anal. Appl. **410**(2), 1014–1027 (2014)