Copulas with continuous, strictly increasing singular conditional distribution functions

Wolfgang Trutschnig[a,*], Juan Fernández Sánchez[b]

[a] *Department of Mathematics, University of Salzburg, Hellbrunner Strasse 34, 5020 Salzburg, Austria, Tel.: +43 662 8044-5312, Fax: +43 662 8044-137*
[b] *Grupo de Investigación de Análisis Matemático, Universidad de Almería, La Cañada de San Urbano, Almería, Spain*

Abstract

Using Iterated Function Systems induced by so-called modifiable transformation matrices T and tools from Symbolic Dynamical Systems we first construct mutually singular copulas $A_T^\star$ with identical (possibly fractal or full) support that are at the same time singular with respect to the Lebesgue measure λ_2 on $[0,1]^2$. Afterwards the established results are utilized for a simple proof of the existence of singular copulas $A_T^\star$ with full support for which all conditional distribution functions $y \mapsto F_x^{A_T^\star}(y)$ are continuous, strictly increasing and have derivative zero λ-almost everywhere. This result underlines the fact that conditional distribution functions of copulas may exhibit surprisingly irregular analytic behavior. Finally, we extend the notion of empirical copula to the case of non i.i.d. data and prove uniform convergence of the empirical copula E_n' corresponding to almost all orbits of a Markov process usually referred to as chaos game to the singular copula $A_T^\star$. Several examples and graphics illustrate both the chosen approach and the main results. XXX

Keywords: Copula, Doubly stochastic measure, Singular function, Markov kernel, Symbolic Dynamical System

2010 MSC: 62H20, 60E05, 28A80, 26A30

---

*Corresponding author
Email addresses: wolfgang@trutschnig.net (Wolfgang Trutschnig), juanfernandez@ual.es (Juan Fernández Sánchez)

1. Introduction

The construction of copulas with fractal support via Iterated Function Systems (IFSs) induced by so-called transformation matrices goes back to Fredricks et al. in [17]. Among other things the authors proved the existence of families $(A_r)_{r\in(0,1/2)}$ of two-dimensional copulas fulfilling that for every $s \in (1,2)$ there exists $r_s \in (0,1/2)$ such that the Hausdorff dimension of the support Z_{r_s} of A_{r_s} is s. Using the fact that the same IFS-construction also converges with respect to the strong metric D_1 (a metrization of the strong operator topology of the corresponding Markov operators, see [30]) on the space $\mathcal{C}$ of two-dimensional copulas Trutschnig and Fernández-Sánchez [31] showed that the same result holds for the subclass of idempotent copulas. Thereby idempotent means idempotent with respect to the star-product introduced by Darsow et al. in [9], i.e. $A \in \mathcal{C}$ is idempotent if $A * A = A$. Families $(A_r)_{r\in(0,1/2)}$ of copulas with fractal support were also studied by de Amo et al. in [1] and in [2]. In the latter paper, using techniques from Probability and Ergodic Theory, the authors discussed properties of subsets of the corresponding fractal supports and constructed mutually singular copulas having the same fractal set as support. Moments of these copulas were calculated in [4]; some surprising properties of homeomorphisms between fractal supports of copulas were studied in [5].

In the current paper we first generalize some results concerning the construction of mutually singular copulas with identical (fractal or full) support by a different method of proof than the one chosen in [2]. In particular we show that for each so-called modifiable transformation matrix T with corresponding invariant copula $A^\star_T$ and attractor $Z^\star_T$ we can find (uncountable many) copulas B having the same support $Z^\star_T$ but being singular w.r.t. $A^\star_T$. Afterwards in Section 4 we focus on transformation matrices T having non-zero entries (hence being modifiable) and the corresponding singular copulas $A^\star_T$ with full support $[0,1]^2$ and study singularity properties of their conditional distribution functions $y \mapsto F_x^{A^\star_T}(y) = K_{A^\star_T}(x,[0,y])$ of $A^\star_T$. Using the one-to-one correspondence between copulas and Markov kernels having the Lebesgue measure λ on $[0,1]$ as fixed point and the fact that the IFS construction can easily be expressed as operation on the corresponding Markov kernels we prove that (λ-almost) all conditional distribution functions $y \mapsto F_x^{A^\star_T}(y) = K_{A^\star_T}(x,[0,y])$ of $A^\star_T$ are continuous, strictly increasing, and have derivative zero λ-almost everywhere. In other words, we prove the existence of copulas $A^\star_T$ for which all conditional distribution functions are

continuous, strictly increasing and singular in the sense of [16, 26] as well as [19] (pp. 278-282). Note that this complements some results in [12] and [13] since the singular copulas with full support considered therein have discrete conditional distributions. For a general study of the interrelation between 2-increasingness and differential properties of copulas we refer to [18].
Finally, in Section 5 we first extend the notion of empirical copulas to non i.i.d. data and then consider sequences of empirical copulas $(\hat{E}_n(\mathbf{k}))_{n\in\mathbb{N}}$ induced by orbits $(Y_n(\mathbf{k}))_{n\in\mathbb{N}}$ of the so-called chaos game (a Markov process induced by transformation matrices T, see [15, 24]). We prove that, with probability one, $(\hat{E}_n(\mathbf{k}))_{n\in\mathbb{N}}$ converges uniformly to the copula $A_T^\star$. Several examples and graphics illustrate the main results.

2. Notation and preliminaries

For every metric space (Ω, ρ) the family of all non-empty compact sets is denoted by $\mathcal{K}(\Omega)$, the Borel σ-field by $\mathcal{B}(\Omega)$ and the family of all probability measures on $\mathcal{B}(\Omega)$ by $\mathcal{P}(\Omega)$. We will call two probability measures μ_1, μ_2 on $\mathcal{B}(\Omega)$ singular with respect to each other (and will write $\mu_1 \perp \mu_2$) if there exist disjoint Borel sets $E, F \in \mathcal{B}(\Omega)$ with $\mu_1(E) = 1 = \mu_2(F)$. λ and λ_2 will denote the Lebesgue measure on $\mathcal{B}([0,1])$ and $\mathcal{B}([0,1]^2)$ respectively. For every set E the cardinality of E will be denoted by $\#E$. $\mathcal{C}$ will denote the family of all two-dimensional *copulas*, see [11, 25, 29], Π will denote the product copula. d_∞ will denote the uniform distance on $\mathcal{C}$; it is well known that $(\mathcal{C}, d_\infty)$ is a compact metric space. For every $A \in \mathcal{C}$ μ_A will denote the corresponding *doubly stochastic measure* defined by $\mu_A([0,x] \times [0,y]) := A(x,y)$ for all $x, y \in [0,1]$, $\mathcal{P}_\mathcal{C}$ the class of all these doubly stochastic measures. A *Markov kernel* from $\mathbb{R}$ to $\mathcal{B}(\mathbb{R})$ is a mapping $K : \mathbb{R} \times \mathcal{B}(\mathbb{R}) \to [0,1]$ such that $x \mapsto K(x,B)$ is measurable for every fixed $B \in \mathcal{B}(\mathbb{R})$ and $B \mapsto K(x,B)$ is a probability measure for every fixed $x \in \mathbb{R}$. Suppose that X, Y are real-valued random variables on a probability space $(\Omega, \mathcal{A}, \mathcal{P})$, then a Markov kernel $K : \mathbb{R} \times \mathcal{B}(\mathbb{R}) \to [0,1]$ is called a *regular conditional distribution of Y given X* if for every $B \in \mathcal{B}(\mathbb{R})$

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

holds $\mathcal{P}$-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 $\mathcal{P}^X$-a.s. (i.e. unique for $\mathcal{P}^X$-almost all $x \in \mathbb{R}$) and that

$K(\cdot,\cdot)$ only depends on $\mathcal{P}^{X\otimes Y}$. Hence, given $A \in \mathcal{C}$ we will denote (a version of) the regular conditional distribution of Y given X by $K_A(\cdot,\cdot)$ and refer to $K_A(\cdot,\cdot)$ simply as *regular conditional distribution of* A or as *Markov kernel of* A. Note that for every $A \in \mathcal{C}$, its conditional regular distribution $K_A(\cdot,\cdot)$, and every Borel set $G \in \mathcal{B}([0,1]^2)$ we have ($G_x := \{y \in [0,1] : (x,y) \in G\}$ denoting the x-section of G for every $x \in [0,1]$)

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

so in particular

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

for every $F \in \mathcal{B}([0,1])$. On the other hand, every Markov kernel $K : [0,1] \times \mathcal{B}([0,1]) \to [0,1]$ fulfilling (3) induces a unique element $\mu \in \mathcal{P}_{\mathcal{C}}([0,1]^2)$ via (2). For every $A \in \mathcal{C}$ 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 disintegration see [21, 22].

Expressing copulas in terms of their corresponding regular conditional distributions a metric D_1 on $\mathcal{C}$ can be defined as follows:

$$D_1(A,B) := \int_{[0,1]} \int_{[0,1]} \big|K_A(x,[0,y]) - K_B(x,[0,y])\big| d\lambda(x)\, d\lambda(y) \tag{4}$$

It can be shown that $(\mathcal{C}, D_1)$ is a complete and separable metric space and that the topology induced by D_1 is strictly finer than the one induced by d_∞ (for an interpretation and various properties of D_1 see [30]).

Before sketching the construction of copulas with fractal support via so-called transformation matrices we recall the definition of an Iterated Function System (IFS for short) and some main results about IFSs (for more details see [6, 14, 24]). Suppose for the following that (Ω, ρ) is a compact metric space and let δ_H denote the Hausdorff metric on $\mathcal{K}(\Omega)$. A mapping $w : \Omega \to \Omega$ is called a *contraction* if there exists a constant $L < 1$ such that $\rho(w(x), w(y)) \leq L\rho(x,y)$ 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* 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 IFSP induces the so-called *Hutchinson operator* $\mathcal{H} : \mathcal{K}(\Omega) \to \mathcal{K}(\Omega)$, defined by

$$\mathcal{H}(Z) := \bigcup_{l=1}^{N} w_l(Z). \tag{5}$$

It can be shown (see [6, 24]) 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 $Z^\star \in \mathcal{K}(\Omega)$ of $\mathcal{H}$. Hence, for every $R \in \mathcal{K}(\Omega)$, we have

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

The *attractor* $Z^\star$ will be called *self-similar* if all contractions in the IFS are similarities. An IFS $\{\Omega, (w_l)_{l=1}^N\}$ is called *totally disconnected* (or disjoint) if the sets $w_1(Z^\star), w_2(Z^\star), \ldots, w_N(Z^\star)$ are pairwise disjoint. $\{\Omega, (w_l)_{l=1}^N\}$ will be called *just touching* if it is not totally disconnected but there exists a non-empty open set $U \subseteq \Omega$ such that $w_1(U), w_2(U), \ldots, w_N(U)$ are pairwise disjoint. Additionally to the operator $\mathcal{H}$ every IFSP also induces a (Markov) operator $\mathcal{V} : \mathcal{P}(\Omega) \to \mathcal{P}(\Omega)$, defined by

$$\mathcal{V}(\mu) := \sum_{i=1}^{N} p_i\, \mu^{w_i}. \tag{6}$$

The so-called *Hutchinson metric* h (sometimes also called Kantorovich or Wasserstein metric) on $\mathcal{P}(\Omega)$ is defined by

$$h(\mu, \nu) := \sup\left\{ \int_\Omega f\, d\mu - \int_\Omega f\, d\nu : \; f \in Lip_1(\Omega, \mathbb{R}) \right\}. \tag{7}$$

Hereby $Lip_1(\Omega, \mathbb{R})$ is the class of all non-expanding functions $f : \Omega \to \mathbb{R}$, i.e. functions fulfilling $|f(x) - f(y)| \le \rho(x, y)$ 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 [6, 10]). 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

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

$\mu^\star$ will be called *invariant measure* - it is well known that the support of $\mu^\star$ is exactly the attractor $Z^\star$. The measure $\mu^\star$ will be called *self-similar* if $Z^\star$ is self-similar, i.e. if all contractions in the IFSP are similarities.

Attractors of IFSs are strongly interrelated with symbolic dynamics via the so-called *address map* (see [6, 24]): For every $N \in \mathbb{N}$ the *code space of N symbols* will be denoted by Σ_N, i.e.

$$\Sigma_N := \{1, 2, \dots, N\}^{\mathbb{N}} = \big\{(k_i)_{i\in\mathbb{N}} : 1 \le k_i \le N \ \forall i \in \mathbb{N}\big\}.$$

Bold symbols will denote elements of Σ_N. σ will denote the (left-) shift operator on Σ_N, i.e. $\sigma((k_1, k_2, \dots)) = (k_2, k_3, \dots)$. Define a metric ρ on Σ_N by setting

$$\rho(\mathbf{k}, \mathbf{l}) := \begin{cases} 0 & \text{if } \mathbf{k} = \mathbf{l} \\ 2^{1-\min\{i : k_i \neq l_i\}} & \text{if } \mathbf{k} \neq \mathbf{l}, \end{cases}$$

then it is straightforward to verify that (Σ_N, ρ) is a compact ultrametric space and that ρ is a metrization of the product topology. Suppose now that $\{\Omega, (w_l)_{l=1}^N\}$ is an IFS with attractor $Z^\star$, fix an arbitrary $x \in \Omega$ and define the *address map* $G : \Sigma_N \to \Omega$ by

$$G(\mathbf{k}) := \lim_{m\to\infty} w_{k_1} \circ w_{k_2} \circ \dots \circ w_{k_m}(x), \tag{8}$$

then (see [24]) $G(\mathbf{k})$ is independent of x, $G : \Sigma_N \to \Omega$ is Lipschitz continuous and $G(\Sigma_N) = Z^\star$. Furthermore G is injective (and hence a homeomorphism) if and only if the IFS is totally disconnected. Given $z \in Z^\star$ every element of the preimage $G^{-1}(\{z\})$ will be called *address* of z. Considering an IFSP $\{\Omega, (w_l)_{l=1}^N, (p_l)_{l=1}^N\}$ with attractor $Z^\star$ and invariant measure $\mu^\star$ we can define a probability measure P on $\mathcal{B}(\Sigma_N)$ by setting

$$P\Big(\big\{\mathbf{k} \in \Sigma_N : k_1 = i_1, k_2 = i_2, \dots, k_m = i_m\big\}\Big) = \prod_{j=1}^{m} p_{i_j} \tag{9}$$

and extending in the standard way to full $\mathcal{B}(\Sigma_N)$. According to [24] $\mu^\star$ is the push-forward of P via the address map, i.e. $P^G(B) := P(G^{-1}(B)) = \mu^\star(B)$ holds for each $B \in \mathcal{B}(Z^\star)$.

Throughout the rest of the paper we will consider IFSP induced by so-called *transformation matrices* - for the original definition see [17], for the generalization to the multivariate setting we refer to [31].

Definition 1 ([17]). A $n\times m$-matrix $T=(t_{ij})_{i=1\ldots n,\, j=1\ldots m}$ is called *transformation matrix* if it fulfills the following four conditions: (i) $\max(n,m)\geq 2$, (ii) all entries are non-negative, (iii) $\sum_{i,j} t_{ij}=1$, and (iv) no row or column has all entries 0. $\mathcal{T}$ will denote the family of all transformations matrices.

Given $T\in\mathcal{T}$ we define the vectors $(a_j)_{j=0}^m, (b_i)_{i=0}^n$ of cumulative column and row sums by $a_0=b_0=0$ and

$$\begin{aligned} a_j &= \sum_{j_0\leq j}\sum_{i=1}^{n} t_{ij_0} \quad j\in\{1,\ldots,m\} \\ b_i &= \sum_{i_0\leq i}\sum_{j=1}^{m} t_{i_0 j} \quad i\in\{1,\ldots,n\}. \end{aligned}$$

Since T is a transformation matrix both $(a_j)_{j=0}^m$ and $(b_i)_{i=0}^n$ are strictly increasing and $R_{ji}:=[a_{j-1},a_j]\times[b_{i-1},b_i]$ is a non-empty compact rectangle for all $j\in\{1,\ldots,m\}$ and $i\in\{1,\ldots,n\}$. Set $\widetilde{I}:=\{(i,j): t_{ij}>0\}$ and consider the IFSP $\{[0,1]^2,(f_{ji})_{(i,j)\in\widetilde{I}},(t_{ij})_{(i,j)\in\widetilde{I}}\}$, whereby the affine contraction $f_{ji}:[0,1]^2\to R_{ji}$ is given by

$$f_{ji}(x,y)=\big(a_{j-1}+x(a_j-a_{j-1})\,,\,b_{i-1}+y(b_i-b_{i-1})\big). \tag{10}$$

$Z_T^\star\in\mathcal{K}([0,1]^2)$ will denote the attractor of the IFSP. The induced operator $\mathcal{V}_T$ on $\mathcal{P}([0,1]^2)$ is defined by

$$\mathcal{V}_T(\mu):=\sum_{j=1}^{m}\sum_{i=1}^{n} t_{ij}\,\mu^{f_{ji}}=\sum_{(i,j)\in\widetilde{I}} t_{ij}\,\mu^{f_{ji}}. \tag{11}$$

It is straightforward to see that $\mathcal{V}_T$ maps $\mathcal{P}_\mathcal{C}$ into itself so we may view $\mathcal{V}_T$ also as operator on $\mathcal{C}$ (see [17]). According to the before-mentioned facts there is exactly one copula $A_T^\star\in\mathcal{C}$, to which we will refer to as *invariant copula*, such that $\mathcal{V}_T(\mu_{A_T^\star})=\mu_{A_T^\star}$ holds. In the sequel we will also write $\mu_T^\star$ instead of $\mu_{A_T^\star}$. Considering that $\mathcal{V}_T$ is a contraction on the complete metric space $(\mathcal{C},D_1)$ (see [30]) it follows that

$$\lim_{n\to\infty} D_1(\mathcal{V}_T^n B, A_T^\star)=0$$

for every copula $B\in\mathcal{C}$, i.e. the IFSP construction also converges to $A_T^\star$ w.r.t. D_1 for every starting copula B. If T contains at least one zero then $\lambda_2(Z_T^\star)=0$ so $\mu_T^\star\perp\lambda_2$. It has already been mentioned in [2, 8] that T containing at least one zero is a sufficient but not necessary condition for $\mu_T^\star\perp\lambda_2$.

3. Mutually singular copulas with identical (fractal) support

In this section we extend some ideas from [2] to the setting of arbitrary transformation matrices $T \in \mathcal{T}$ not necessarily inducing IFSPs that only contain similarities. Doing so we do not follow the approach in [2] but directly prove and utilize the fact that the dynamical systems $(\Sigma_N, \mathcal{B}(\Sigma_N), P_T, \sigma)$ and $(Z_T^\star, \mathcal{B}(Z_T^\star), \mu_T^\star, \Phi_T)$ are isomorphic (Φ_T to be defined subsequently in equation (13)) for every $T \in \mathcal{T}$.

Fix $T \in \mathcal{T}$, consider the IFSP $\{[0,1]^2, (f_{ji})_{(i,j)\in\widetilde{I}}, (t_{ij})_{(i,j)\in\widetilde{I}}\}$ induced by T, and let $Z_T^\star \in \mathcal{K}([0,1]^2)$ denote the corresponding attractor. To simplify notation sort the functions $(f_{ji})_{(i,j)\in\widetilde{I}}$ lexicographically and rename them as $w_1, \ldots, w_N$ with $N := \#\widetilde{I}$; analogously rename the rectangles R_{ji} by $Q_1, \ldots, Q_N$ and the probabilities t_{ij} by $p_1, \ldots p_N$. Using the fact that the IFSP $\{[0,1]^2, (w_i)_{i=1}^N, (p_i)_{i=1}^N\}$ is either totally disconnected or just touching we can show that $\#G^{-1}(\{(x,y)\}) = 1$ for μ_A-almost every $(x,y) \in Z_T^\star$ and arbitrary $A \in \mathcal{C}$. In fact, setting

$$D_m = \bigcup_{\mathbf{i},\mathbf{j}\in\Sigma_N:\,(i_1,\ldots,i_m)\neq(j_1,\ldots,j_m)} \big(w_{i_1}\circ\ldots\circ w_{i_m}(Z_T^\star)\big)\cap\big(w_{j_1}\circ\ldots\circ w_{j_m}(Z_T^\star)\big) \quad (12)$$

for every $m \in \mathbb{N}$ as well as $D := \bigcup_{m=1}^\infty D_m$ the following result holds:

Lemma 2. *Suppose that $T \in \mathcal{T}$. Then $(x,y) \in Z_T^\star$ has only one address if and only if $(x,y) \in D^c$. Furthermore for every $A \in \mathcal{C}$ we have $\mu_A(D^c) = 1$.*

Proof: If $(x,y) \in w_{i_1} \circ \ldots \circ w_{i_m}(Z_T^\star) \cap w_{j_1} \circ \ldots \circ w_{j_m}(Z_T^\star)$ for some $m \in \mathbb{N}$ and $(i_1,\ldots,i_m) \neq (j_1,\ldots,j_m)$ then there are $\mathbf{k}, \mathbf{l} \in \Sigma_N$ with $(x,y) = w_{i_1} \circ \ldots \circ w_{i_m}(G(\mathbf{k})) = w_{j_1} \circ \ldots \circ w_{j_m}(G(\mathbf{l}))$. Hence (x,y) has at least the two addresses $(i_1,\ldots,i_m,k_1,k_2,\ldots), (j_1,\ldots,j_m,l_1,l_2,\ldots) \in \Sigma_N$. Suppose now that $G(\mathbf{k}) = G(\mathbf{l}) = (x,y)$ for $\mathbf{k} \neq \mathbf{l}$ and let m denote the smallest $i \in \mathbb{N}$ with $k_i \neq l_i$. Then it follows immediately that $(x,y) \in w_{k_1} \circ \ldots \circ w_{k_m}(Z_T^\star) \cap w_{l_1} \circ \ldots \circ w_{l_m}(Z_T^\star) \subseteq D_m$ completing the proof that D is exactly the set of all points with at least two addresses.

To prove the second assertion note that for every copula $A \in \mathcal{C}$ we even have $\mu_A\big(w_{i_1}\circ\ldots\circ w_{i_m}([0,1]^2)\cap w_{j_1}\circ\ldots\circ w_{j_m}([0,1]^2)\big) = 0$ whenever $(i_1,\ldots,i_m) \neq (j_1,\ldots,j_m)$ since $\mu_A(\{x\}\times[0,1]) = \mu_A([0,1]\times\{y\}) = 0$ for all $x,y \in [0,1]$. This implies $\mu_A(D_m) = 0$ from which $\mu_A(D) = 0$ follows immediately. ■

Define a measurable function $\Phi_T : Z_T^\star \to Z_T^\star$ by

$$\Phi_T(x,y) = \sum_{i=1}^{N} w_i^{-1}(x,y)\mathbf{1}_{Z_T^\star \cap (Q_i \setminus \bigcup_{j=1}^{i-1} Q_j)}(x,y). \tag{13}$$

and, as before, let $\mu_T^\star \in \mathcal{P}_\mathcal{C}$ denote the invariant measure of the IFSP. The following result holds (P_T as in equation (9)):

Theorem 3. *For every $T \in \mathcal{T}$ the dynamical systems $(\Sigma_N, \mathcal{B}(\Sigma_N), P_T, \sigma)$ and $(Z_T^\star, \mathcal{B}(Z_T^\star), \mu_T^\star, \Phi_T)$ are isomorphic.*

Proof: We prove the result in three steps.
(S1) Suppose that $\mathbf{k} \in \Sigma_N$ and that there exists $\mathbf{l} \neq \sigma(\mathbf{k})$ with $G(\mathbf{l}) = G(\sigma(\mathbf{k}))$. Applying w_{k_1} to both sides yields $G((k_1, l_1, l_2, \ldots)) = G(\mathbf{k}) = (x,y)$, so (x,y) has at least two different addresses. Having this, $\sigma(G^{-1}(D^c)) \subseteq G^{-1}(D^c)$ follows immediately. Furthermore for every $(x,y) \in D^c$ with address $\mathbf{k} \in G^{-1}(D^c)$ we obviously have

$$\Phi_T \circ G(\mathbf{k}) = G \circ \sigma(\mathbf{k}). \tag{14}$$

Hence, using $\sigma(G^{-1}(D^c)) \subseteq G^{-1}(D^c)$, $\Phi_T(x,y) = \Phi_T \circ G(\mathbf{k}) \in D^c$ follows, which shows $\Phi_T(D^c) \subseteq D^c$.
(S2) Obviously the shift-operator σ is P_T-preserving, i.e. we have $P_T^\sigma = P_T$. Showing that Φ_T preserves $\mu_T^\star$ can be done as follows: For every Borel set $B \in \mathcal{B}([0,1]^2)$ and every $j \in \{1, \ldots, N\}$ considering that $\mu_T^\star$ is $\mathcal{V}_T$-invariant and that $\mu_T^\star(D) = 0$ implies (measurability of $w_j(B)$ follows from the fact that w_j is an affine bijection by construction)

$$\mu_T^\star(w_j(B)) = \sum_{i=1}^{N} p_i \, \mu_T^\star\big(w_i^{-1}(w_j(B))\big) = p_j \, \mu_T^\star(B).$$

Hence, for every $B \in \mathcal{B}(Z_T^\star)$ we get

$$\begin{aligned}
\mu_T^\star(\Phi_T^{-1}(B)) &= \mu_T^\star(D^c \cap \Phi_T^{-1}(B)) = \sum_{i=1}^{N} \mu_T^\star\big(D^c \cap \Phi_T^{-1}(B) \cap w_i(Z_T^\star)\big) \\
&= \sum_{i=1}^{N} \mu_T^\star\big(D^c \cap w_i(B)\big) = \sum_{i=1}^{N} \mu_T^\star\big(w_i(B)\big) = \sum_{i=1}^{N} p_i \, \mu_T^\star(B) \\
&= \mu_T^\star(B),
\end{aligned}$$

implying that Φ_T is $\mu_T^\star$-preserving.

(S3) As last step we show that the inverse $G^{-1} : D^c \to G^{-1}(D^c)$ of the bijection $G : G^{-1}(D^c) \to D^c$ is measurable (w.r.t. the σ-fields $\mathcal{B}(Z_T^\star) \cap D^c$ and $\mathcal{B}(\Sigma_N) \cap G^{-1}(D^c)$) and measure-preserving. Suppose that $m \in \mathbb{N}$, that $l_1, \ldots, l_m \in \{1, \ldots N\}$ and set $Y := \{\mathbf{k} \in \Sigma_N : k_1 = l_1, \ldots, k_m = l_m\} \cap G^{-1}(D^c)$. Then we have

$$(G^{-1})^{-1}(Y) = \big\{(x,y) \in D^c : G^{-1}(x,y) \in Y\big\} = D^c \cap w_{l_1} \circ w_{l_2} \circ \cdots \circ w_{l_m}(Z_T^\star)$$

from which measurability follows since m and $l_1, \ldots, l_m \in \{1, \ldots N\}$ were arbitrary. The fact that $G^{-1} : D^c \to G^{-1}(D^c)$ is measure-preserving now follows easily from $P_T^G = \mu_T^\star$. ■

The subsequent diagram depicts the measure-preserving maps studied in the previous Theorem (vertical two-headed arrows symbolize invertible measure-preserving transformations outside sets of measure zero):

$$\begin{array}{ccc} (\Sigma_N, P_T) & \xrightarrow{\ \sigma\ } & (\Sigma_N, P_T) \\ {\scriptstyle G}\big\downarrow & & \big\downarrow{\scriptstyle G} \\ (Z_T^\star, \mu_T^\star) & \xrightarrow[\ \Phi_T\ ]{} & (Z_T^\star, \mu_T^\star) \end{array} \tag{15}$$

As direct consequence of Theorem 3 we get the following corollary (see [32]).

Corollary 4. *The dynamical system $(Z_T^\star, \mathcal{B}(Z_T^\star), \mu_T^\star, \Phi_T)$ is strongly mixing and its entropy $h(\Phi_T)$ is given by $h(\Phi_T) = -\sum_{(i,j)\in\widetilde{I}} t_{ij} \log(t_{ij})$.*

Theorem 3 offers a simple way for the construction of mutually singular copulas $A, B \in \mathcal{C}$ having the same (fractal) support whenever, loosely speaking, it is possible to modify $T \in \mathcal{T}$ without changing the induced IFS. This motivates the following definition:

Definition 5. A transformation matrix $T \in \mathcal{T}$ will be called *modifiable* if and only if we can find $T' \in \mathcal{T}$ with $T' \neq T$ such that T and T' induce the same IFS (but not the same IFSP). For every $T \in \mathcal{T}$ the family of all such T' will be denoted by $\mathcal{M}_T$.

Obviously $\mathcal{M}_T$ is either empty or contains uncountably many elements. A simple sufficient (but not necessary) condition for $\mathcal{M}_T \neq \emptyset$ is the existence of $i_1 < i_2$ and $j_1 < j_2$ such that $t_{ij} > 0$ for all $(i,j) \in \{i_1, i_2\} \times \{j_1, j_2\}$.

Example 6. Consider the following three modifiable transformation matrices T_1, T_2, T_3, defined by

$$T_1 = \begin{pmatrix} \frac{1}{3} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{3} \end{pmatrix}, \quad T_2 = \begin{pmatrix} \frac{3}{16} & \frac{5}{16} \\ \frac{3}{16} & \frac{5}{16} \end{pmatrix}, \quad T_3 = \begin{pmatrix} \frac{2}{16} & \frac{6}{16} \\ \frac{4}{16} & \frac{4}{16} \end{pmatrix}.$$

Then obviously $T_2 \in \mathcal{M}_{T_3}$ and vice versa. Figure 1 and Figure 2 depict $\mathcal{V}^n_{T_1}(\Pi)$ and $\mathcal{V}^n_{T_2}(\Pi)$ for $n \in \{1, 2, 4, 6\}$.

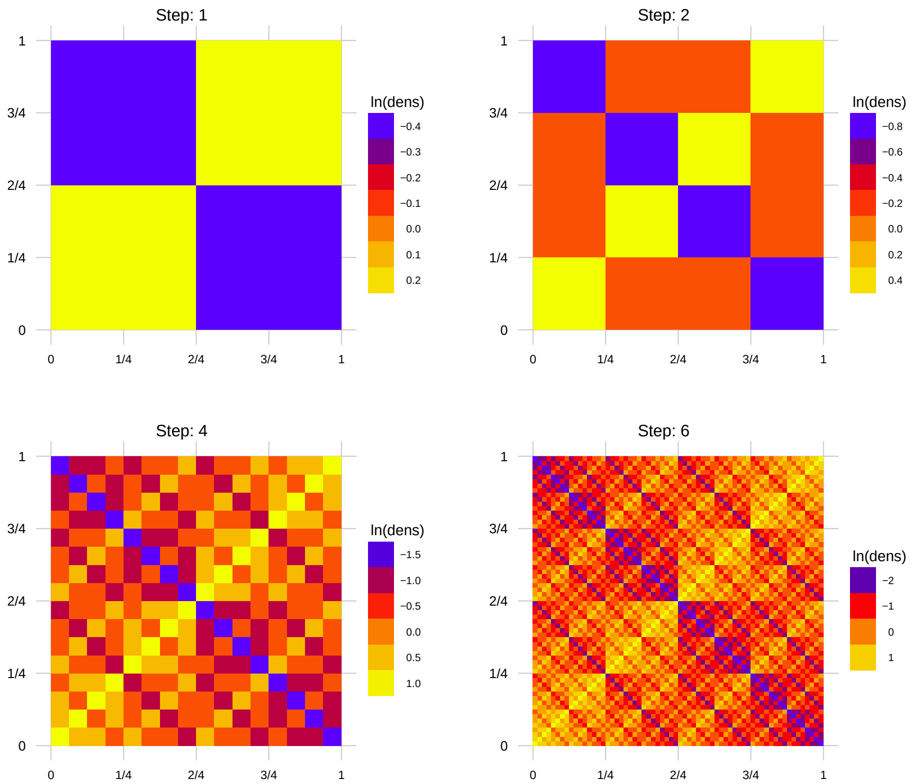

Figure 1: Image plot of the (natural) logarithm of the density of $\mathcal{V}^n_{T_1}(\Pi)$ for $n \in \{1, 2, 4, 6\}$, T_1 according to Example 6

Theorem 7. *Suppose that T is a modifiable transformation matrix. Then there exist uncountably many copulas with support $Z^\star_T$ that are pairwise mutually singular with respect to each other.*

Proof: Choose $T' \in \mathcal{M}_T$, let $p'_1, \dots, p'_N > 0$ denote the corresponding probabilities and $P_{T'}$ the corresponding probability measure on Σ_N induced by $p'_1, \dots, p'_N$ according to equation (9). Then, using ergodicity of σ both w.r.t. $P_{T'}$ and with respect to P_T (see [32]), we have $P_T \perp P_{T'}$, so we can find $\Lambda, \Lambda' \in \mathcal{B}(\Sigma_N)$ fulfilling $\Lambda \cap \Lambda' = \emptyset$ as well as $P_T(\Lambda) = P_{T'}(\Lambda') = 1$. Since, according to Lemma 2, we have $P_T(G^{-1}(D^c)) = \mu_T^\star(D^c) = 1 = \mu_{T'}^\star(D^c) = P_{T'}(G^{-1}(D^c))$ this implies $P_T(\Lambda \cap G^{-1}(D^c)) = P_{T'}(\Lambda' \cap G^{-1}(D^c)) = 1$. Set $\Gamma = G(\Lambda \cap G^{-1}(D^c))$ and $\Gamma' = G(\Lambda' \cap G^{-1}(D^c))$, then using Theorem 3, $\Gamma, \Gamma' \in \mathcal{B}(Z_T^\star)$ and $\Gamma \cap \Gamma' = \emptyset$ follows, and we have

$$\mu_T^\star(\Gamma) = P_T(G^{-1}(\Gamma)) = 1 = P_{T'}(G^{-1}(\Gamma')) = \mu_{T'}^\star(\Gamma')$$

which completes the proof. ■

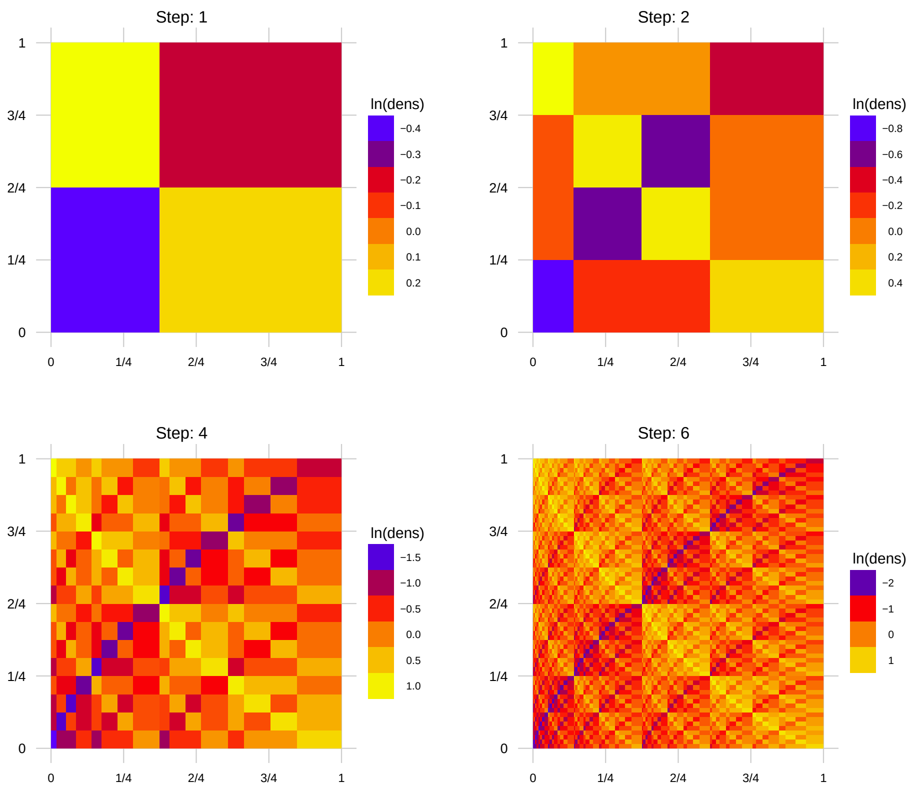

Figure 2: Image plot of the (natural) logarithm of the density of $\mathcal{V}_{T_2}^n(\Pi)$ for $n \in \{1, 2, 4, 6\}$, T_2 according to Example 6

216

Remark 8. Theorem 7 also holds if, for some $m \in \mathbb{N}$, the Kronecker product $T^{*m} := T * T * \cdots * T$ of T with itself fulfills $\mathcal{M}_{T^{*m}} \neq \emptyset$.

Remark 9. If T is a transformation matrix containing no zeros then obviously T is modifiable. If T additionally fulfills $t_{ij} \neq (a_{j+1} - a_j)(b_{i+1} - b_i)$ for at least one (i, j) then $\mathcal{M}_T$ contains the transformation matrix E with $e_{ij} = (a_{j+1} - a_j)(b_{i+1} - b_i)$ for which obviously $\mu_E^\star = \mu_\Pi$ holds. Hence for all these T we have that $A_T^\star$ has full support although at the same time $\mu_T^\star \perp \lambda_2$, generalizing some results contained in [8] and [2].

We close this section with the following result:

Corollary 10. *Suppose that $T \in \mathcal{T}$ fulfills $\mu_T^\star \neq \lambda_2$. Then for λ-almost every $x \in [0, 1]$ the probability measure $K_{A_T^\star}(x, \cdot)$ is singular w.r.t. λ and the conditional distribution function $y \mapsto F_x^{A_T^\star}(y) := K_{A_T^\star}(x, [0, y])$ has derivative zero λ-almost everywhere.*

Proof: Using the afore-mentioned results $\mu_T^\star \neq \lambda_2$ implies $\mu_T^\star \perp \lambda_2$, so there exists $\Gamma \in \mathcal{B}([0, 1]^2)$ with $\mu_T^\star(\Gamma) = 1$ and $\lambda_2(\Gamma) = 0$. Disintegration immediately yields $\lambda(\Gamma_x) = 0$ as well as $K_{A_T^\star}(x, \Gamma_x) = 1$ for λ-almost all $x \in [0, 1]$. Having this the remaining assertion follows directly from the fact that the derivative of a singular measure w.r.t. λ is zero almost everywhere (see [27]). ■

4. Singular copulas with full support whose conditional distribution functions are continuous, strictly increasing and singular

Singular copulas A with full support have already been studied in the literature, see [12, 13]. In both constructions the conditional distributions $K_A(x, \cdot)$ are concentrated on countable sets, i.e. they are discrete probability measures. We will use the results of the previous section now in order to prove the existence of copulas A for which all conditional distribution functions $F_x^A : y \mapsto K_A(x, [0, y])$ are continuous, strictly increasing, and fulfill $\frac{dF_x^A(y)}{dy} = 0$ for λ-almost every $y \in [0, 1]$. To simplify notation we will only consider elements of the class $\hat{\mathcal{T}}$ of all transformation matrices T (i) containing no zeros, (ii) fulfilling that the row sums and column sums through every t_{ij} are identical and (iii) $\mu_T^\star \neq \lambda_2$. It is straightforward to verify that $\hat{\mathcal{T}}$ coincides with the class of all transformation matrices of the form $T = \frac{1}{m} S$, whereby

$m \in \{2,3,\dots\}$ and S is a m-dimensional doubly stochastic matrix containing no zeros and having at least two different entries. The reason for considering $\hat{\mathcal{T}}$ is that, firstly, in this case $\mu_T^\star$ has full support and fulfills $\mu_T^\star \perp \lambda_2$ and that, secondly, $a_i = b_i = \frac{i}{m}$ for every $i \in \{0,\dots,m\}$. Hence writing

$$g_i(x) = a_{i-1} + (a_i - a_{i-1})x = \frac{i-1}{m} + \frac{1}{m}\,x \tag{16}$$

and setting $h_i := g_i^{-1}$ for every $i \in \{1,\dots,m\}$ the contractions (f_{ij}) in equation (10) can be expressed as

$$f_{ij}(x,y) = \big(g_i(x), g_j(y)\big).$$

Obviously the corresponding IFSP $\{[0,1]^2, (w_i)_{i=1}^N, (p_i)_{i=1}^N\}$ with $N = m^2$ only contains similarities, so $\mu_T^\star$ is self-similar.
Before studying the general case we have a look at the transformation matrix $T_1 \in \hat{\mathcal{T}}$ from Example 6. For every $A \in \mathcal{C}$ the kernel $K_{\mathcal{V}_{T_1}A}$ can directly be calculated from K_A - in fact, extending the definition of K_A to $[0,1] \times \mathcal{B}(\mathbb{R})$ by setting $K_A(\cdot, E) := 0$ if $E \cap [0,1] = \emptyset$, and fixing $z \in [0,1], E \in \mathcal{B}([0,1])$ we have

$$\begin{pmatrix} K_{\mathcal{V}_{T_1}A}\big(g_1(z),E\big) \\ K_{\mathcal{V}_{T_1}A}\big(g_2(z),E\big) \end{pmatrix} = \underbrace{\begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}}_{=2\,T_1^t} \begin{pmatrix} K_A\big(z,h_1(E)\big) \\ K_A\big(z,h_2(E)\big) \end{pmatrix}. \tag{17}$$

If the conditional distribution function $y \mapsto F_z^A(y) := K_A(x,[0,y])$ is continuous then equation (17) implies continuity of $F_{g_i(z)}^{\mathcal{V}_{T_1}A}$. Let $\mathcal{F}$ denote the family of all continuous non-decreasing functions F on $[0,1]$ fulfilling $F(0) = 0$ and $F(1) = 1$ and ρ_∞ the uniform distance. Then $(\mathcal{F}, \rho_\infty)$ is a complete metric space (see [19, 27]). Define two functions $\varphi_1, \varphi_2 : \mathcal{F} \to \mathcal{F}$ by

$$(\varphi_1 \circ F)(y) = \begin{cases} \frac{2}{3}\,F(2y) & \text{if } y \in \big[0, \frac{1}{2}\big] \\ \frac{2}{3} + \frac{1}{3}\,F(2y-1) & \text{if } y \in \big(\frac{1}{2}, 1\big] \end{cases} \tag{18}$$

and

$$(\varphi_2 \circ F)(y) = \begin{cases} \frac{1}{3}\,F(2y) & \text{if } y \in \big[0, \frac{1}{2}\big] \\ \frac{1}{3} + \frac{2}{3}\,F(2y-1) & \text{if } y \in \big(\frac{1}{2}, 1\big], \end{cases} \tag{19}$$

then we have

$$F_{g_i(z)}^{\mathcal{V}_T A}(y) = \big(\varphi_i \circ F_z^A\big)(y) \tag{20}$$

for every $y \in [0,1]$ and $z \in [0,1]$. The next lemma gathers some properties of $\varphi_1, \varphi_2 : \mathcal{F} \to \mathcal{F}$:

Lemma 11. *Let φ_1, φ_2 be defined according to equation (18) and (19). Then setting*

$$\Psi(\mathbf{k}) = \lim_{n\to\infty} \varphi_{k_1} \circ \varphi_{k_2} \circ \cdots \circ \varphi_{k_n}(F)$$

for $\mathbf{k} \in \Sigma_2$ and arbitrary $F \in \mathcal{F}$ defines a continuous mapping $\Psi : \Sigma_2 \to \mathcal{F}$. Moreover the function $\Psi(\mathbf{k}) \in \mathcal{F}$ is strictly increasing for every $\mathbf{k} \in \Sigma_2$.

Proof: The functions φ_1, φ_2 are contractions on $(\mathcal{F}, \rho_\infty)$ with Lipschitz constant $L = \frac{2}{3}$. Fix $\mathbf{k} \in \Sigma_2$. As Cauchy sequence $\big(\varphi_{k_1} \circ \cdots \circ \varphi_{k_n}(F)\big)_{n\in\mathbb{N}}$ has a limit $\Psi(\mathbf{k}, F) \in \mathcal{F}$. Using the fact that φ_1, φ_2 are contractions it follows that $\Psi(\mathbf{k}, F) \in \mathcal{F}$ is independent of the function F, i.e. $\Psi : \Sigma_2 \to \mathcal{F}$ is well-defined and, without loss of generality, we may choose $F = id_{[0,1]}$. Since continuity of Ψ directly follows from the construction it remains to show that $\Psi(\mathbf{k})$ is strictly increasing for every $\mathbf{k} \in \Sigma_2$. For every $l \in \mathbb{N}$ set $S_l := \big\{0, \frac{1}{2^l}, \frac{2}{2^l}, \ldots, \frac{2^l-1}{2^l}, 1\big\}$. Considering that the construction of Ψ implies $\varphi_{k_1} \circ \varphi_{k_2} \circ \cdots \circ \varphi_{k_n}(id)(z) = \varphi_{k_1} \circ \varphi_{k_2} \circ \cdots \circ \varphi_{k_l}(id)(z)$ for every $z \in S_l$ and $n \geq l$ we have

$$\Psi(\mathbf{k})(z) = \varphi_{k_1} \circ \varphi_{k_2} \circ \cdots \circ \varphi_{k_l}(id)(z)$$

for every $z \in S_l$. Since $\varphi_{k_1} \circ \varphi_{k_2} \circ \cdots \circ \varphi_{k_l}(id)$ is strictly increasing on S_l for every $l \in \mathbb{N}$ it follows that $\Psi(\mathbf{k})$ is strictly increasing on $S := \bigcup_{l=1}^\infty S_l$ which completes the proof. ■

Let Λ denote the set of all $x \in [0,1]$ for which there is a unique $\mathbf{k} \in \Sigma_2$ (which we will also refer to as 'address' of x) such that

$$x = \lim_{n\to\infty} g_{k_1} \circ g_{k_2} \circ \cdots \circ g_{k_n}(1), \tag{21}$$

then Λ^c is countable. Considering $F_x^\Pi(y) = y$ for every $x \in [0,1]$ Lemma 11 and equation (20) imply that for every $x \in \Lambda$ we have

$$K(x, [0,y]) := \lim_{n\to\infty} K_{\mathcal{V}_{T_1}^n \Pi}(x, [0,y]) = \lim_{n\to\infty} F_{g_{k_1}\circ\cdots\circ g_{k_n}(1)}^{\mathcal{V}_{T_1}^n \Pi}(y) = \Psi(\mathbf{k})(y). \tag{22}$$

Applying Lebesgue's theorem on dominated convergence shows

$$A_{T_1}^\star(a, y) = \lim_{n\to\infty} \int_{[0,a]} K_{\mathcal{V}_{T_1}^n \Pi}(x, [0,y])\, d\lambda(x) = \int_{[0,a]} K(x, [0,y])\, d\lambda(x)$$

for every $a \in [0,1]$ and $y \in [0,1]$. Hence $K(\cdot,\cdot)$ is (a version of) the Markov kernel $K_{A_{T_1}^\star}(\cdot,\cdot)$ of $A_{T_1}^\star$. Using Corollary 10, we have altogether shown that for λ-almost every $x \in [0,1]$ the conditional distribution function $y \mapsto F_x^{A^\star_{T_1}}(y)$ is continuous and strictly increasing with derivative zero λ-almost everywhere. Since kernels are only unique λ-almost everywhere we may, without loss of generality, assume that $y \mapsto F_x^{A^\star_{T_1}}(y)$ is continuous and strictly increasing with derivative zero λ-almost everywhere for every $x \in [0,1]$. In other words $A^\star_T$ fulfills all singularity properties stated at the beginning of this section. Figure 3 depicts $F_x^{\mathcal{V}^i_{T_1}\Pi}$ for some values of i and three different x, Figure 4 is an image- and $3d$-plot of $(x,y) \mapsto F_x^{\mathcal{V}^8_{T_1}\Pi}(y)$.

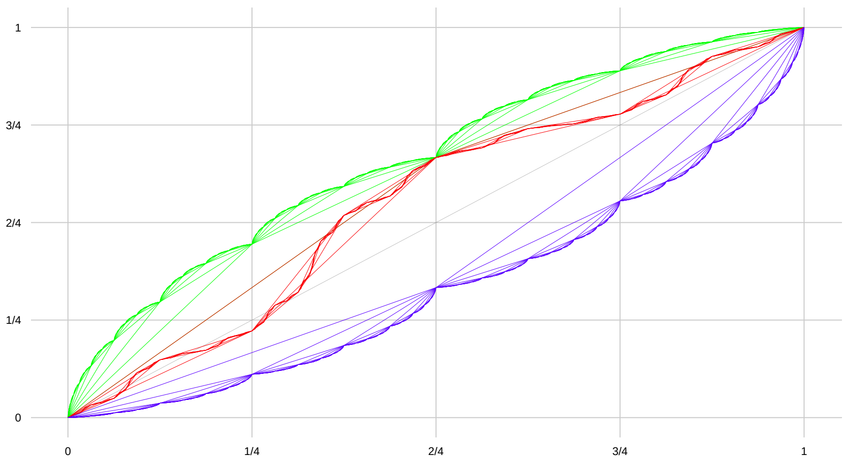

Figure 3: $y \mapsto F_x^{\mathcal{V}^i_{T_1}(\Pi)}$ for $i \in \{1,2,\ldots,12\}$ and x having address $\mathbf{k}$ of the form $(1,1,\ldots,1,*,*)$ (green), of the form $(2,2,\ldots,2,*,*)$ (blue), and of the form $(1,2,1,2,\ldots,1,2,*,*)$ (red); the $*$-entries start at coordinate 13.

Remark 12. Using the fact that $\varphi_1(\mathcal{F}) \cap \varphi_2(\mathcal{F}) = \emptyset$ together with injectivity of $\varphi_1, \varphi_2 : \mathcal{F} \to \mathcal{F}$ it can be shown that the mapping Ψ in Lemma 11 is injective, implying that the conditional distribution functions $(F_x^{A^\star_{T_1}})_{x \in [0,1]}$ are pairwise different for all x outside a set of λ-measure zero.

We now state and prove the general result for arbitrary elements in $\hat{\mathcal{T}}$:

Theorem 13. *Suppose that $T \in \hat{\mathcal{T}}$ and let $A^\star_T$ denote the corresponding singular copula with full support. Then for λ-almost every $x \in [0,1]$ the conditional distribution function $y \mapsto F_x^{A^\star_T}(y) = K_{A^\star_T}(x,[0,y])$ is continuous, strictly increasing and has derivative zero λ-almost everywhere.*

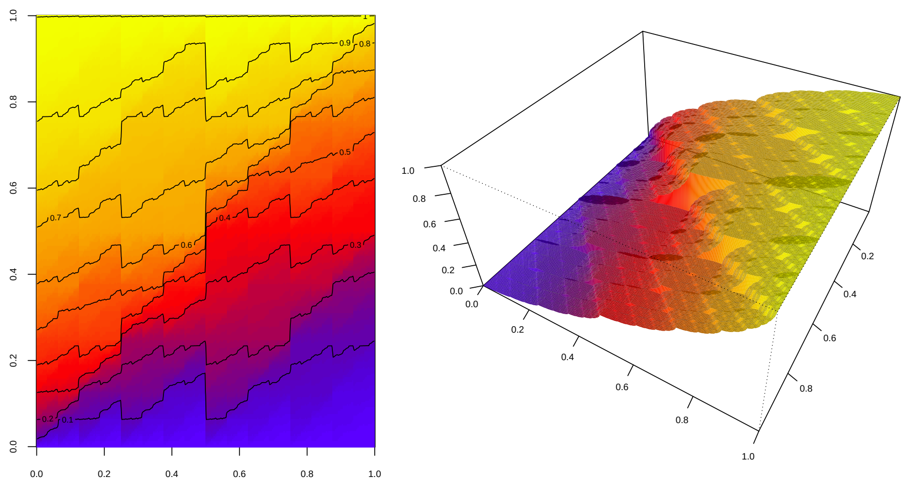

Figure 4: Image- and 3d-plot of the function $(x,y) \mapsto F_x^{\mathcal{V}_{T_1}^8 \Pi}(y)$

Proof: We proceed in four main steps.
(S1) For every $A \in \mathcal{C}$ extend the definition of the kernel K_A to $[0,1] \times \mathcal{B}(\mathbb{R})$ by setting $K_A(\cdot, E) := 0$ if $E \cap [0,1] = \emptyset$. Then for $z \in [0,1], E \in \mathcal{B}([0,1])$ we have

$$\begin{pmatrix} K_{\mathcal{V}_T A}\big(g_1(z), E\big) \\ K_{\mathcal{V}_T A}\big(g_2(z), E\big) \\ \vdots \\ K_{\mathcal{V}_T A}\big(g_m(z), E\big) \end{pmatrix} = m \underbrace{\begin{pmatrix} t_{11} & t_{21} & \dots & t_{m1} \\ t_{12} & t_{22} & \dots & t_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ t_{1m} & t_{2m} & \dots & t_{mm} \end{pmatrix}}_{=m T^t} \begin{pmatrix} K_A\big(z, h_1(E)\big) \\ K_A\big(z, h_2(E)\big) \\ \vdots \\ K_A\big(z, h_m(E)\big) \end{pmatrix}. \tag{23}$$

For $y \in (b_{i_0-1}, b_{i_0}]$ and $E = [0,y]$, we get (empty sums are zero by definition)

$$K_{\mathcal{V}_T A}(g_j(z), [0,y]) = m \sum_{i<i_0} t_{ij} \, + \, m \, t_{i_0 j} K_A\Big(z, \Big[\frac{y - b_{i_0-1}}{b_{i_0} - b_{i_0-1}}\Big]\Big). \tag{24}$$

(S2) For every $j \in \{1, \dots, m\}$ define a function $\varphi_j : \mathcal{F} \to \mathcal{F}$ by

$$(\varphi_j \circ F)(y) = m \sum_{i<i_0} t_{ij} \, + \, m \, t_{i_0 j} F\Big(\frac{y - b_{i_0-1}}{b_{i_0} - b_{i_0-1}}\Big). \tag{25}$$

whenever $y \in (b_{i_0-1}, b_{i_0}]$. Then each φ_j fulfills

$$\rho_\infty(\varphi_j \circ F, \varphi_j \circ G) \leq \underbrace{\big(m \max_{i,j} t_{ij}\big)}_{<1} \rho_\infty(F, G)$$

and the function $\Psi : \Sigma_m \to \mathcal{F}$ given by

$$\Psi(\mathbf{k}) = \lim_{n\to\infty} \varphi_{k_1} \circ \varphi_{k_2} \circ \cdots \circ \varphi_{k_n}(F)$$

is well-defined, independent of the concrete choice of F, and continuous.
(S3) For every $l \in \mathbb{N}$ set $S_l := \big\{0, \frac{1}{m^l}, \frac{2}{m^l}, \ldots, \frac{m^l-1}{m^l}, 1\big\}$. The construction of Ψ implies $\varphi_{k_1} \circ \varphi_{k_2} \circ \cdots \circ \varphi_{k_n}(id)(z) = \varphi_{k_1} \circ \varphi_{k_2} \circ \cdots \circ \varphi_{k_l}(id)(z)$ for every $z \in S_l$ and $n \geq l$ we have

$$\Psi(\mathbf{k})(z) = \varphi_{k_1} \circ \varphi_{k_2} \circ \cdots \circ \varphi_{k_l}(id)(z)$$

for every $z \in S_l$. Since $\varphi_{k_1} \circ \varphi_{k_2} \circ \cdots \circ \varphi_{k_l}(id)$ is strictly increasing on S_l for every $l \in \mathbb{N}$ it follows that $\Psi(\mathbf{k})$ is strictly increasing on $S := \bigcup_{l=1}^\infty S_l$ implying that $\Psi(\mathbf{k})$ is strictly increasing on $[0, 1]$ since S is dense in $[0, 1]$.
(S4) Let Λ_m denote the set of all $x \in [0, 1]$ for which there is a unique $\mathbf{k} \in \Sigma_m$ such that

$$x = \lim_{n\to\infty} g_{k_1} \circ g_{k_2} \circ \cdots \circ g_{k_n}(1). \tag{26}$$

Then Λ_m^c is countable. Considering that for every $x \in \Lambda_m$ with address $\mathbf{k} \in \Lambda_m$ we have

$$K(x, [0, y]) := \lim_{n\to\infty} K_{\mathcal{V}_T^n \Pi}(x, [0, y]) = \lim_{n\to\infty} F^{\mathcal{V}_T^n \Pi}_{g_{k_1} \circ \cdots \circ g_{k_n}(1)}(y) = \Psi(\mathbf{k})(y), \tag{27}$$

it follows in the same way as before that $K(\cdot, \cdot)$ is (a version of) the Markov kernel $K_{A_T^\star}(\cdot, \cdot)$. Using Proposition 10 therefore completes the proof. ■

Corollary 14. *Suppose that $T \in \hat{\mathcal{T}}$ and that $T' \in \mathcal{M}_T$ has at least two different entries. Then $\mu_T^\star, \mu_{T'}^\star \perp \lambda_2$ and $\mu_T^\star \perp \mu_{T'}^\star$. Furthermore we can find a set $\Lambda \in \mathcal{B}([0, 1])$ with $\lambda(\Lambda) = 1$ such that for every $x \in \Lambda$ we have $K_{A_T^\star}(x, \cdot) \perp K_{A_{T'}^\star}(x, \cdot)$ and the conditional distribution functions $F_x^{A_T^\star}, F_x^{A_{T'}^\star}$ are continuous, strictly increasing with derivative zero λ-almost everywhere.*

5. Uniform convergence of empirical copulas induced by orbits of a special Markov process to singular copulas

It is well known that the empirical copula E'_n for i.i.d. data $((x_i, y_i))_{i=1}^n$ from $A \in \mathcal{C}$ is a strongly consistent estimator of A w.r.t. d_∞. In fact, according to [20, 28], we even have

$$d_\infty(E'_n, A) = O\left(\sqrt{\frac{\log \log n}{n}}\right)$$

almost surely for $n \to \infty$. The purpose of this section is to show that the empirical copula based on (non i.i.d.) samples of the so-called chaos game (a Markov process defined subsequently) induced by IFSP coming from transformation matrices $T \in \mathcal{T}$ is a strongly consistent estimator of $A_T^\star$ w.r.t. d_∞ too. We start with an extension of the notion of empirical copula to the case of arbitrary samples $((x_i, y_i))_{i=1}^n$ for which not necessarily all x_i or all y_i are different. Consider a finite sample $((x_i, y_i))_{i=1}^n \in \mathbb{R}^2$ and define the empirical distribution function (ecdf, for short) H_n and the marginal empirical distribution functions F_n, G_n in the usual way, i.e.

$$H_n(x, y) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{(-\infty,x]\times(-\infty,y]}(x_i, y_i)$$

as well as

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{(-\infty,x]}(x_i)\,, \quad G_n(y) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{(-\infty,y]}(y_i).$$

Then the function $E_n : \mathrm{Rg}(F_n) \times \mathrm{Rg}(G_n) \to [0, 1]$, defined on the range of (F_n, G_n) by

$$E_n(F_n(x), G_n(y)) = H_n(x, y) \tag{28}$$

is easily seen to be a subcopula that can be extended to a copula E'_n via bilinear interpolation (see [25] as well as [3, 7] for other possible extensions). Note that bilinear interpolation yields a checkerboard copula (see [23]), i.e. the corresponding doubly stochastic measure μ'_n is absolutely continuous and its density is constant on each of the rectangles induced by the grid $\mathrm{Rg}(F_n) \times \mathrm{Rg}(G_n)$. In the sequel we will refer to E'_n as *empirical copula* (ecop for short) of the sample $((x_i, y_i))_{i=1}^n$. Figure 5 shows the empirical copula E'_8 and the mass distribution of μ'_8 for a (non i.i.d.) sample of size eight.

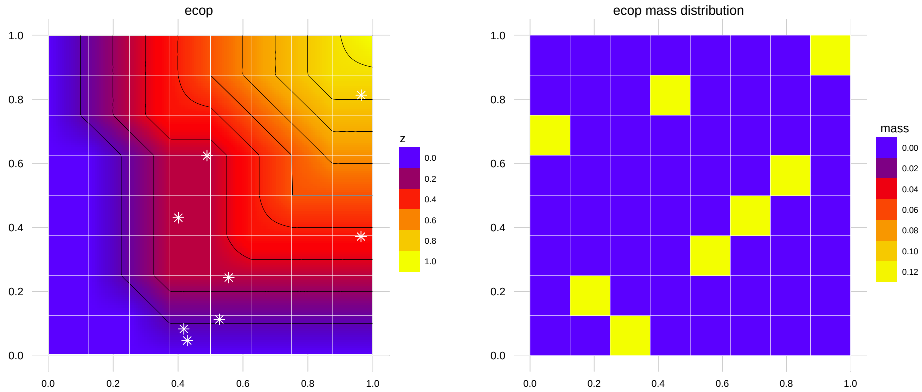

Figure 5: ecop and mass distribution for a sample of size eight (the black lines in the left plot depict the $0.1, 0.2, \dots, 0.9$-contour lines, the white asterisks depict the sample).

We now recall the definition of the chaos game for the specific case of IFSPs coming from transformation matrices and then prove the afore-mentioned consistency result. Suppose that $T \in \mathcal{T}$ and let $\{[0,1]^2, (w_i)_{i=1}^N, (p_i)_{i=1}^N\}$ denote the induced IFSP and P_T the probability measure on $\mathcal{B}(\Sigma_N)$ defined according to equation (9). Fix a point $(x_0, y_0) \in [0,1]^2$ (not necessarily an element of $Z_T^\star$) and define a $[0,1]^2$-valued Markov process $(Y_n^{(x_0,y_0)})_{n\in\mathbb{N}}$ on $(\Sigma_N, \mathcal{B}(\Sigma_N), P_T)$ by setting

$$Y_n^{(x_0,y_0)}(\mathbf{k}) := w_{k_n} \circ w_{k_{n-1}} \circ \cdots \circ w_{k_1}(x_0, y_0) \tag{29}$$

for every $n \in \mathbb{N}$. In other words, the (one-step) transition probabilities $H(\cdot,\cdot)$ of the process are given by (see [15])

$$H\big((x,y), B\big) = \sum_{j=1}^{N} p_j \, \mathbf{1}_B(w_j(x,y)).$$

for all $(x,y) \in [0,1]^2$ and $B \in \mathcal{B}([0,1]^2)$. The Markov process $(Y_n^{(x_0,y_0)})_{n\in\mathbb{N}}$ will be called *chaos game starting at* (x_0, y_0) (see [6, 15, 24]). According to Elton's ergodic theorem (see [15, 24]) it follows that for P_T-almost all $\mathbf{k} \in \Sigma_N$ the empirical measure ϑ_n converges weakly to the invariant measure $\mu_T^\star$, i.e.

$$\vartheta_n := \frac{1}{n} \sum_{i=1}^{n} \delta_{Y_i^{(x_0,y_0)}(\mathbf{k})} \longrightarrow \mu_T^\star \quad \text{weakly for } n \to \infty. \tag{30}$$

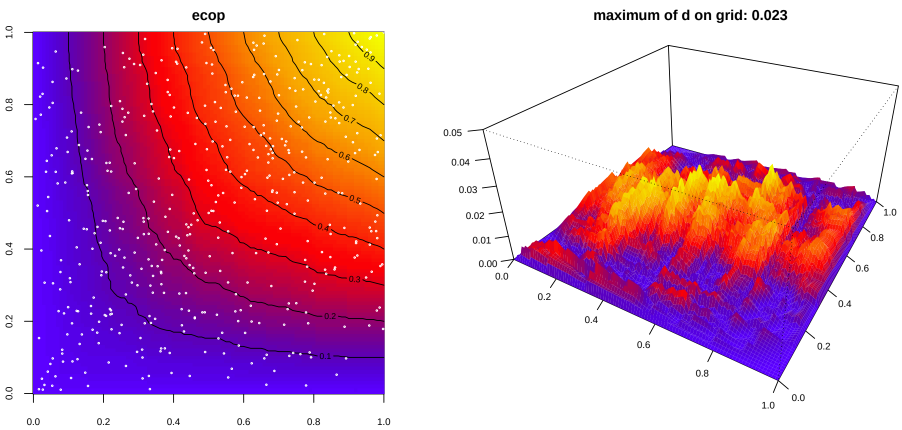

Figure 6: E'_{500} and $d(x,y)$ according to Example 18 (a)

Remark 15. As pointed out in [24] the chaos game provides a very and simple and efficient way (random iteration of functions) to approximate the invariant measure of an IFSP also in the general setting.

As direct consequence of (30) we get that the empirical distribution function H_n of the sample $Y_1^{(x_0,y_0)}(\mathbf{k})$, $Y_2^{(x_0,y_0)}(\mathbf{k}), \ldots, Y_n^{(x_0,y_0)}(\mathbf{k})$ converges pointwise to the copula $A_T^\star$ for P_T-almost all $\mathbf{k} \in \Sigma_N$. Having this we can prove the following result:

Theorem 16. *Suppose that $T \in \mathcal{T}$, let $\{[0,1]^2, (w_i)_{i=1}^N, (p_i)_{i=1}^N\}$ denote the corresponding IFSP and fix $(x_0, y_0) \in [0,1]^2$. Then there exists a set $\Lambda \subseteq \Sigma_N$ with $P_T(\Lambda) = 1$ such that for every $\mathbf{k} \in \Lambda$ the sequence $(E'_n)_{n\in\mathbb{N}}$ of empirical copulas of the sample $Y_1^{(x_0,y_0)}(\mathbf{k}), Y_2^{(x_0,y_0)}(\mathbf{k}), \ldots, Y_n^{(x_0,y_0)}(\mathbf{k})$ fulfills*

$$\lim_{n\to\infty} d_\infty(E'_n, A_T^\star) = 0.$$

In other words: with probability one the empirical copula converges uniformly to the copula $A_T^\star$.

Proof: Define $\Lambda \subseteq \Sigma_N$ as the set of all $\mathbf{k}$ fulfilling Equation (30), then Elton's ergodic theorem (see [15]) implies $P_T(\Lambda) = 1$. As at the beginning of this section let H_n, F_n, G_n denote the empirical distribution functions of the sample $Y_1^{(x_0,y_0)}(\mathbf{k}), Y_2^{(x_0,y_0)}(\mathbf{k}), \ldots, Y_n^{(x_0,y_0)}(\mathbf{k})$ for every $n \in \mathbf{N}$.

Equation (30) implies $\lim_{n\to\infty} H_n(x,y) = A_T^\star(x,y)$ for all $x, y \in [0,1]$. Fix $(x,y) \in [0,1]^2$. Then, using $H_n(x,y) = E_n'(F_n(x), G_n(y))$ and setting $R_n := \big|E_n'(F_n(x), G_n(y)) - A_T^\star(x,y)\big|$ we have $\lim_{n\to\infty} R_n = 0$ and it follows that

$$\Big|\, \big|E_n'(F_n(x), G_n(y)) - E_n'(x,y)\big| - \big|A_T^\star(x,y) - E_n'(x,y)\big| \,\Big| \leq R_n \longrightarrow 0$$

for $n \to \infty$. Taking into account that (30) also implies

$$\begin{aligned} \limsup_{n\to\infty} \big|E_n'(F_n(x), G_n(y)) - E_n'(x,y)\big| &\leq \limsup_{n\to\infty} \big(|F_n(x) - x| + |G_n(y) - y|\big) \\ &= 0 \end{aligned}$$

altogether we get $\lim_{n\to\infty} E_n'(x,y) = A_T^\star(x,y)$. Since $(x,y) \in [0,1]^2$ was arbitrary and in $\mathcal{C}$ pointwise and uniform convergence are equivalent the proof is complete. ∎

Remark 17. It is straightforward to verify that we do not have convergence of E_n' to $A_T^\star$ w.r.t. D_1 if T has at least one column with two non-zero entries.

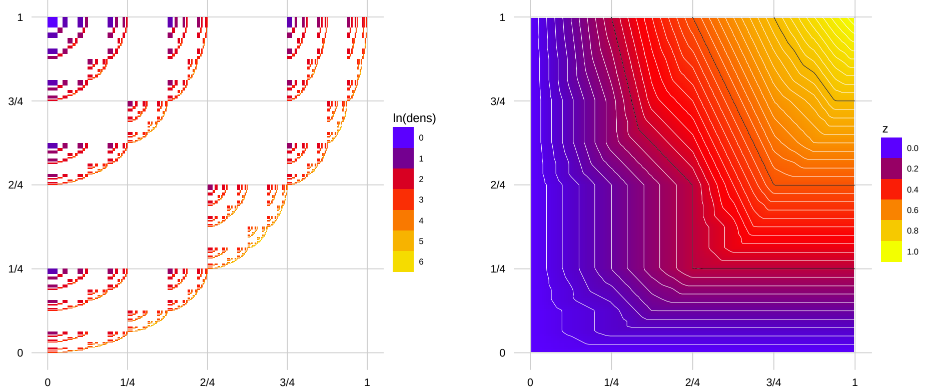

Figure 7: Image plot of the (natural) logarithm of the density of $\mathcal{V}_{T_4}^5(\Pi)$ (left) as well as image plot of the copula $\mathcal{V}_{T_4}^5(\Pi)$ (right), T_4 from Example 18 (b).

We conclude the paper with the following example:

Example 18. **(a)** Consider again the transformation matrix T_1 from Example 6. In this case the invariant copula $A_T^\star$ is self-similar, singular and has full support. Figure 6 depicts E_{500}' for an Orbit $(Y_i^{(1,1)}(\mathbf{k}))_{i\in\mathbb{N}}$ as well as the

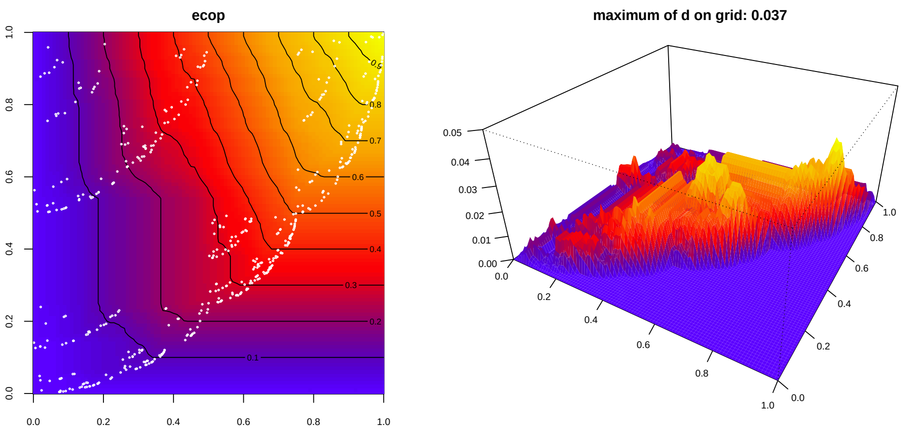

Figure 8: E'_{500} and $d(x,y)$ according to Example 18 (b)

function $d(x,y) := |E'_{500}(x,y) - A^{\star}_{T_1}(x,y)|$ on an equidistant grid of 101×101 points.

(b) Consider the transformation matrix T_4, defined by

$$T_4 = \begin{pmatrix} 1/4 & 0 & 0 \\ 0 & 1/4 & 0 \\ 1/4 & 0 & 1/4 \end{pmatrix}.$$

In this case the invariant copula $A^{\star}_{T_4}$ is not self-similar. Figure 7 depicts the (natural) logarithm of the density of $\mathcal{V}^5_{T_4}(\Pi)$ as well as $\mathcal{V}^5_{T_4}(\Pi)$. Figure 8 shows E'_{500} for an Orbit $(Y_i^{(1,1)}(\mathbf{k}))_{i\in\mathbb{N}}$ as well as the function $d(x,y) := |E'_{500}(x,y) - A^{\star}_{T_4}(x,y)|$ on an equidistant grid of 101×101 points.

Acknowledgments

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

References

[1] E. de Amo, M. Díaz Carrillo, J. Fernández-Sánchez, Measure-Preserving Functions and the Independence Copula, *Mediterr. J. Math.* **8** (2011) 431–450.

[2] E. de Amo, M. Díaz Carrillo, J. Fernández-Sánchez, Copulas and associated fractal sets, *J. Math. Anal. Appl.* **386** (2012) 528–541.

[3] E. de Amo, M. Díaz Carrillo, J. Fernández-Sánchez, Characterization of all copulas associated with non-continuous random variables, *Fuzzy Set Syst.* **191** (2012) 103–112.

[4] E. de Amo, M. Díaz Carrillo, J. Fernández-Sánchez, A. Salmerón, Moments and associated measures of copulas with fractal support, *Appl. Math. Comput.* **218** (2012) 8634-8644.

[5] E. de Amo, M. Díaz Carrillo, J. Fernández-Sánchez, W. Trutschnig, Some Results on Homeomorphisms Between Fractal Supports of Copulas, *Nonlinear Anal-Theor.* **85** (2013) 132–144.

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

[7] H. Carley, Maximum and minimum extensions of finite subcopulas, *Commun. Stat. A-Theor.* **31** (2002) 2151-2166.

[8] I. Cuculescu, R. Theodorescu, Copulas: diagonals, tracks, *Rev. Roum. Math Pure A.* **46** (2001) 731-742.

[9] W.F. Darsow, B. Nguyen, E.T. Olsen, Copulas and Markov processes, *Illinois J. Math.* **36** (1992) 600-642.

[10] R.M. Dudley, *Real Analysis and Probability*, Cambridge University Press, 2002.

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

[12] F. Durante, J. Fernández-Sánchez, C. Sempi, A note on the notion of singular copula, *Fuzzy Set Syst.* **211** (2013) 120-122.

[13] F. Durante, P. Jaworski, R. Mesiar, Invariant dependence structures and Archimedean copulas, *Stat. Probabil. Lett.* **81** (2011) 1995-2003.

[14] G.A. Edgar, *Measure, Topology and Fractal Geometry*, Undergrad. Texts in Math., Springer, Berlin Heidelberg New York, 1990.

[15] J.F. Elton, An ergodic theorem for iterated maps, *Ergod. Theor. Dyn. Syst.* **7** (1987) 481-488.

[16] J. Fernández Sánchez, P. Viader, J. Paradis, M. Díaz Carrillo, A singular function with a non-zero finite derivative, *Nonlinear Anal-Theor.* **15** (2012) 5010-5014.

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

[18] R. Ghiselli Ricci, On differential properties of copulas, *Fuzzy Set Syst.* **220** (2013) 78-87.

[19] E. Hewitt, K. Stromberg, *Real and Abstract Analysis, A Modern Treatment of the Theory of Functions of a Real Variable*, Springer-Verlag, New York, 1965.

[20] P. Janssen, J. Swanepoel, N. Veraverbeke, Large sample behavior of the Bernstein copula estimator, *J. Stat. Plan. Infer.* **142** (2012) 1189–1197.

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

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

[23] T. Kulpa, On Approximation of Copulas, *Internat. J. Math. & Math. Sci.* **22** (1999) 259-269.

[24] H. Kunze, D. La Torre, F. Mendivil, E.R Vrscay, *Fractal Based Methods in Analysis*, Springer, New York Dordrecht Heidelberg London, 2012.

[25] R.B. Nelsen, *An Introduction to Copulas* (2nd edn.), Springer Series in Statistics, New York, 2006.

[26] J. Paradís, P. Viader, L. Bibiloni, Riesz-Nágy singular functions revisited, *J. Math. Anal. Appl.* **329** (2007) 592–602.

[27] W. Rudin, *Real and Complex Analysis*, McGraw-Hill, New York, 1966.

[28] J. Swanepoel, J. Allison, Some new results on the empirical copula estimator with applications, *Stat. Probabil. Lett.* **83** (2013) 1731-1739.

[29] C. Sempi, Copulæ: Some Mathematical Aspects, *Appl. Stoch. Model Bus.* **27** (2010) 37–50.

[30] W. Trutschnig, On a strong metric on the space of copulas and its induced dependence measure, *J. Math. Anal. Appl.* **384** (2011) 690-705.

[31] W. Trutschnig, J. Fernández-Sánchez, Idempotent and multivariate copulas with fractal support, *J. Stat. Plan. Infer.* **142** (2012) 3086-3096.

[32] P. Walters, *An Introduction to Ergodic Theory*, Springer, New York, 1982.