Some results on shuffles of two-dimensional copulas

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

[a] *Research Unit for Intelligent Data Analysis and Graphical Models, European Centre for Soft Computing, Edificio Científico Tecnológico, Calle Gonzalo Gutiérrez Quirós, 33600 Mieres (Asturias), Spain, Tel.: +34 985456545*

[b] *Grupo de Investigación de Análisis Matemático, Universidad de Almería, La Cañada de San Urbano, Almería, Spain*

Abstract

Using the one-to-one correspondence between two-dimensional copulas and special Markov kernels allows to study properties of T-shuffles of copulas, T being a general Lebesgue-measure-preserving transformation on $[0, 1]$, in terms of the corresponding operation on Markov kernels. As one direct consequence of this fact the asymptotic behaviour of iterated T-shuffles $\mathcal{S}_{T^n}(A)$ of a copula $A \in \mathcal{C}$ can be characterized through mixing properties of T. In particular it is shown that $\mathcal{S}_{T^n}(A)$ $(\frac{1}{n}\sum_{i=1}^{n}\mathcal{S}_{T^i}(A))$ converges uniformly to the product copula Π for every copula A if and only if T is strongly mixing (ergodic). Moreover working with Markov kernels also allows, firstly, to give a short proof of the fact that the mass of the singular component of $\mathcal{S}_T(A)$ cannot be bigger than the mass of the singular component of A, secondly, to introduce and study another operator $\mathcal{U}_T : \mathcal{C} \to \mathcal{C}$ fulfilling $\mathcal{S}_T \circ \mathcal{U}_T(A) = A$ for all $A \in \mathcal{C}$, and thirdly to express $\mathcal{S}_T(A)$ and $\mathcal{U}_T(A)$ as $*$-product of A with the completely dependent copula C_T induced by T.

Keywords: Copula, Doubly stochastic measure, Markov kernel, shuffle, star product, Frobenius Perron operator

2010 MSC: 62H20, 60E05, 28D05

---

*Corresponding author

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

1. Introduction

The study of shuffles of copulas probably started with the paper [13] by Mikusínski et al. in 1992 on shuffles of the minimum copula M. Informally (see [13]) a (straight) shuffle of M can be obtained by (i) considering the doubly stochastic measure μ_M corresponding to M on $[0,1]^2$, (ii) cutting $[0,1]^2$ into a finite number of (vertical) strips and (iii) permutating (shuffling) the strips. In other words (see [6]) one chooses an interval-exchange transformation $T:[0,1]\to[0,1]$ and the shuffle $\mathcal{S}_T(M)$ of M is then defined in terms of the corresponding doubly stochastic measure $\mu_{\mathcal{S}_T(M)}$ by

$$\mu_{\mathcal{S}_T(M)}(E\times F)=\mu_M(T^{-1}(E)\times F)$$

for all Borel sets $E,F\in\mathcal{B}([0,1])$. Durante et al. (see [6]) generalized this concept in two directions - they replaced, firstly, M by arbitrary copulas $A\in\mathcal{C}$ and, secondly, the interval-exchange transformation T by arbitrary λ-preserving transformations $T:[0,1]\to[0,1]$ (λ denoting the Lebesgue measure on $[0,1]$). Their article [6] contains various interesting results - in particular the authors showed (Theorem 14) that there exists a bijective $T:[0,1]\to[0,1]$ with T,T^{-1} λ-preserving and a strictly increasing sequence $(n_j)_{j\in\mathbf{N}}$ in $\mathbf{N}$ such that for every $A\in\mathcal{C}$ the sequence $(\mathcal{S}_{T^{n_j}}(A))_{j\in\mathbf{N}}$ converges uniformly to the product copula Π. According to [19] the same result can not hold for the stronger metric D_1 since D_1 strictly separates Π from the class of completely dependent copulas.
In the current paper we will consider general (i.e. not necessarily bijective) λ-preserving transformations, use the one-to-one correspondence between the class $\mathcal{C}$ of two-dimensional copulas and the class $\mathcal{K}$ of special Markov kernels (see [19], [20]), and show that on $\mathcal{K}$ the operator $A\mapsto\mathcal{S}_T(A)$ coincides with the well known Frobenius Perron operator induced by T (see Section 2 and [12]). This interrelation does not only offer a short proof of the above-mentioned theorem, it also allows to characterize the asymptotic behaviour of iterated T-shuffles $\mathcal{S}_{T^n}(A)$ in terms of mixing properties of the transformation T. In particular we will show that $\mathcal{S}_{T^n}(A)$ ($\frac{1}{n}\sum_{i=1}^n\mathcal{S}_{T^i}(A)$) converges uniformly to the product copula Π for every copula A if and only if T is strongly mixing (ergodic) and that we even have $\lim_{n\to\infty}D_1(\mathcal{S}_{T^n}(A),\Pi)=0$ if T is exact (D_1 is the metric studied in [19], see Section 2). Additionally we will show that the operation $A\mapsto\mathcal{S}_T(A)$ is smoothing in the sense that the mass of the singular component of $\mathcal{S}_T(A)$ cannot be bigger than the mass of the singular component of A and state some properties of another operator

$\mathcal{U}_T : \mathcal{C} \to \mathcal{C}$ fulfilling $\mathcal{S}_T \circ \mathcal{U}_T(A) = A$ for all A and for all T.
The rest of the paper is organized as follows: Section 2 gathers some preliminaries and notations. Section 3 contains the description of T-shuffles on the class $\mathcal{K}$ of associated Markov kernels and three examples illustrating the different asymptotic behaviour of $\mathcal{S}_{T^n}(A)$ given different mixing properties of the transformation T. All main results are presented in Section 4.

2. Notation and preliminaries

As already mentioned before $\mathcal{C}$ will denote the family of all two-dimensional *copulas*, see [7], [14], [17]. M and Π will denote the minimum and the product copula respectively. Given $A \in \mathcal{C}$ the *transpose* $A^t \in \mathcal{C}$ of A is defined by $A^t(x, y) := A(y, x)$ for all $x, y \in [0, 1]$. 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. λ and λ_2 will denote the Lebesgue measure on $[0, 1]$ and $[0, 1]^2$ respectively, $\mathcal{B}([0, 1])$ and $\mathcal{B}([0, 1]^2)$ the Borel σ-fields in $[0, 1]$ and $[0, 1]^2$. $\mathcal{T}$ will denote the class of all λ-preserving transformations $T : [0, 1] \to [0, 1]$, $\mathcal{T}_p$ the subset of all bijective $T \in \mathcal{T}$ fulfilling $T^{-1} \in \mathcal{T}$.
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 more details and properties of conditional expectation, regular conditional distributions, and disintegration see [8] and [9].

A copula $A \in \mathcal{C}$ will be called *completely dependent* if and only if there exists $T \in \mathcal{T}$ such that $K(x, E) := \mathbf{1}_E(Tx)$ is a regular conditional distribution of A (see [11] and [19] for equivalent definitions and main properties). For every $T \in \mathcal{T}$ the induced completely dependent copula will be denoted by C_T throughout the rest of the paper.

A linear operator V on $L^1([0,1]) := L^1([0,1], \mathcal{B}([0,1]), \lambda)$ is called *Markov operator* (see [2] and [15]) if it fulfills the following three properties:

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

As mentioned in the introduction $\mathcal{M}$ will denote the class of all Markov operators on $L^1([0,1])$. It is straightforward to see that the operator norm of V is one, i.e. $\|V\| := \sup\{\|Vf\|_1 : \|f\|_1 \leq 1\} = 1$ holds. According to [2] and [15] *there is a one-to-one correspondence between* $\mathcal{C}$ *and* $\mathcal{M}$ - in fact, the mappings $\Phi : \mathcal{C} \to \mathcal{M}$ and $\Psi : \mathcal{M} \to \mathcal{C}$, defined by

$$\begin{aligned} \Phi(A)(f)(x) &:= (V_A f)(x) := \frac{d}{dx} \int_{[0,1]} A_{,2}(x,t) f(t) d\lambda(t), \\ \Psi(T)(x,y) &:= A_V(x,y) := \int_{[0,x]} (T\mathbf{1}_{[0,y]})(t) d\lambda(t) \end{aligned} \tag{4}$$

for every $f \in L^1([0,1])$ and $(x,y) \in [0,1]^2$ ($A_{,2}$ denoting the partial derivative w.r.t. y), fulfill $\Psi \circ \Phi = id_{\mathcal{C}}$ and $\Phi \circ \Psi = id_{\mathcal{M}}$. Note that in case of $f := \mathbf{1}_{[0,y]}$

we have $(V_A \mathbf{1}_{[0,y]})(x) = A_{,1}(x,y)$ λ-a.s. According to [19] the first equality in (4) can be simplified to

$$(V_A f)(x) = \mathbb{E}(f \circ Y | X = x) = \int_{[0,1]} f(y) K_A(x, dy) \qquad \lambda\text{-a.s.} \tag{5}$$

Expressing copulas in terms of their corresponding regular conditional distributions two metrics D_1 and D_∞ on $\mathcal{C}$ can be defined as follows:

$$D_1(A,B) := \int_{[0,1]} \int_{[0,1]} \big| K_A(x,[0,y]) - K_B(x,[0,y]) \big| d\lambda(x)\, d\lambda(y) \tag{6}$$

$$D_\infty(A,B) := \sup_{y\in[0,1]} \int_{[0,1]} \big| K_A(x,[0,y]) - K_B(x,[0,y]) \big| d\lambda(x) \tag{7}$$

It can be shown that $(\mathcal{C}, D_1)$ as well as $(\mathcal{C}, D_\infty)$ are complete and separable metric spaces and that, given copulas $A, A_1, A_2 \ldots$ and their corresponding Markov operators $V_A, V_{A_1}, V_{A_2} \ldots$, the following three conditions are equivalent (see [19]):

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

(b) $\lim_{n\to\infty} \|V_{A_n} f - V_A f\|_1 = 0$ for every $f \in L^1([0,1])$

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

In other words D_1 and D_∞ are (complete and separable) metrizations of the strong operator topology on $\mathcal{M}$. Furthermore it is straightforward to show that the topology induced by D_∞ is strictly finer than the one induced by d_∞ (again see [19]).

Given $A, B \in \mathcal{C}$ the *star product* $A * B \in \mathcal{C}$ is defined by (see [2], [15])

$$(A * B)(x,y) := \int_{[0,1]} A_{,2}(x,t) B_{,1}(t,y) d\lambda(t) \tag{8}$$

and fulfills

$$V_{A*B} = \Phi_{A*B} = \Phi(A) \circ \Phi(B) = V_A \circ V_B, \tag{9}$$

so the mapping Φ in (4) actually is an isomorphism (see [15]). Additionally Φ maps transposes to adjoints, i.e. $\Phi(A^t) = V_A^{adj}$ holds for every $A \in \mathcal{C}$ ($V_A^{adj} : L^\infty([0,1]) \to L^\infty([0,1])$ denoting the adjoint operator of V_A).

The following result, stating that the Markov kernel of $A * B$ is just the standard composition of the Markov kernels of A and B will prove useful in the sequel:

Lemma 1 ([20]). *Suppose that $A, B \in \mathcal{C}$ and let K_A, K_B denote regular conditional distributions of A and B. Then the Markov kernel $K_A \circ K_B$, defined by*

$$(K_A \circ K_B)(x, F) := \int_{[0,1]} K_B(y, F) K_A(x, dy), \tag{10}$$

is a regular conditional distribution of $A * B$.

Finally we extend the definition of general T-shuffles of copulas (see [6]) to the case of arbitrary $T \in \mathcal{T}$ and only stick to the name 'shuffle' for the sake of simplicity. For every $T \in \mathcal{T}$ and every copula $A \in \mathcal{C}$ let $\mathcal{S}_T(A) \in \mathcal{C}$ denote the (generalized) T*-shuffle of* A, defined via the corresponding doubly stochastic measures by

$$\mu_{\mathcal{S}_T(A)}(E \times F) := \mu_A(T^{-1}(E) \times F) \tag{11}$$

for all $E, F \in \mathcal{B}([0,1])$. As mentioned before we will study T-shuffles in terms of the corresponding Markov kernels - doing so our main tool will be the so-called *Frobenius-Perron operator* (FPO for short). Suppose that $T \in \mathcal{T}$, then the FPO P_T is the unique linear operator on $L^1([0,1])$ fulfilling

$$\int_E P_T f(x) d\lambda(x) = \int_{T^{-1}(E)} f(x) d\lambda(x) \tag{12}$$

for every $E \in \mathcal{B}([0,1])$ and every $f \in L^1([0,1])$. For the definition of the FPO in the general setting and its main properties we refer to [12].

3. Expressing shuffles in terms of Markov kernels and the star product - basic interrelations and three motivating examples

We start with the following simple but useful lemma that expresses the T-shuffle of a copula A both in terms of a transformation of the corresponding Markov kernel K_A and in terms of the star product of A with the completely dependent copula C_T.

Lemma 2. *Suppose that $T \in \mathcal{T}$ and that $A \in \mathcal{C}$. Then $\mathcal{S}_T(A) = (C_T)^t * A$ and for every $F \in \mathcal{B}([0,1])$ the following interrelation holds for λ-almost all $x \in [0,1]$:*

$$K_{\mathcal{S}_T(A)}(x, F) = \frac{d}{dx} \int_{T^{-1}([0,x])} K_A(t, F)\, d\lambda(t) \tag{13}$$

Proof: Equation (11) implies

$$\int_E K_{\mathcal{S}_T(A)}(t,F)d\lambda(t) = \int_{T^{-1}(E)} K_A(t,F)d\lambda(t) \tag{14}$$

for all $E, F \in \mathcal{B}([0,1])$. Setting $E := [0,x]$ and viewing both sides in (14) as bounded monotonic (hence λ-almost everywhere differentiable, see [16]) functions it follows that (13) holds for λ-almost every $x \in [0,1]$. To prove the fact that $\mathcal{S}_T(A) = (C_T)^t * A$ note that Lemma 1 implies

$$\begin{aligned} K_{B*C_T}(x,E) &= \int_{[0,1]} K_{C_T}(y,E)K_B(x,dy) = \int_{[0,1]} \mathbf{1}_E(Ty)K_B(x,dy) \\ &= K_B(x,T^{-1}(E)) \end{aligned}$$

for every copula $B \in \mathcal{C}$, every $E \in \mathcal{B}([0,1])$ and λ-almost all $x \in [0,1]$. Having this it follows immediately that for all $E, F \in \mathcal{B}([0,1])$ we have

$$\begin{aligned} \mu_{\mathcal{S}_T(A)}(E \times F) &= \mu_A(T^{-1}(E) \times F) = \mu_{A^t}(F \times T^{-1}(E)) \\ &= \int_F K_{A^t}\big(x, T^{-1}(E)\big)d\lambda(x) = \int_F K_{A^t * C_T}(x,E)d\lambda(x) \\ &= \mu_{A^t * C_T}(F \times E) = \mu_{(C_T)^t * A}(E \times F), \end{aligned}$$

which completes the proof. ■

Since for $T \in \mathcal{T}_p$ equation (13) can be simplified to

$$\begin{aligned} K_{\mathcal{S}_T(A)}(x,F) &= \frac{d}{dx}\int_{T^{-1}([0,x])} K_A(t,F)\,d\lambda(t) \\ &= \frac{d}{dx}\int_{[0,1]} \mathbf{1}_{T^{-1}([0,x])}(T^{-1}t)K_A(T^{-1}t,F)\,d\lambda(t) \\ &= \frac{d}{dx}\int_{[0,1]} \mathbf{1}_{[0,x]}(t)K_A(T^{-1}t,F)\,d\lambda(t) = K_A(T^{-1}x,F) \end{aligned}$$

the follow proposition holds.

Proposition 3. *Suppose that $T \in \mathcal{T}_p$ then for every $A \in \mathcal{C}$ and every Borel set $F \in \mathcal{B}([0,1])$ the following interrelation holds for λ-almost all $x \in [0,1]$:*

$$K_{\mathcal{S}_T(A)}(x,F) = K_A(T^{-1}x,F) \tag{15}$$

Note that equation (14) implies that, for given $F \in \mathcal{B}([0,1])$, $K_{\mathcal{S}_{T(A)}}(\cdot, F)$ is the Frobenius Perron operator P_T applied to the function $K_A(\cdot, F)$ (see [12] and [5]). Asymptotics of the FPO have been studied extensively (see [12]) - we will see in the next section how these properties may be used to prove a general result about the asymptotic behaviour of iterated T-shuffles $\mathcal{S}_{T^n}(A)$. Before doing so we take a look to three examples:

Example 1. Consider the transformation $T_2 x := 2x\,(mod 1)$. Fix $A \in \mathcal{C}$ and $E \in \mathcal{B}([0,1])$. Using (13) obviously we have

$$K_{\mathcal{S}_{T_2}(A)}(x, E) = \frac{1}{2} K_A\Big(\frac{x}{2}, E\Big) + \frac{1}{2} K_A\Big(\frac{x+1}{2}, E\Big),$$

and it is straightforward to see that for each $n \in \mathbf{N}$ the following equality holds for λ-almost all $x \in (0,1)$:

$$K_{\mathcal{S}_{T_2^n}(A)}(x, E) = \frac{1}{2^n} \sum_{i=0}^{2^n-1} K_A\left(\frac{x}{2^n} + \frac{i}{2^n}, E\right)$$

If we assume that $A \in \mathcal{C}$ fulfills the property that for each $y \in [0,1]$ the functions $x \mapsto K_A(x, [0,y])$ are Riemann integrable, then it follows that for λ-almost all $x \in (0,1)$ we have

$$\lim_{n\to\infty} K_{\mathcal{S}_{T_2^n}(A)}(x, [0,y]) = \int_{[0,1]} K_A(t, [0,y])\, d\lambda(t) = y.$$

Hence, using the results in [19], we have $\lim_{n\to\infty} D_1(\mathcal{S}_{T_2^n}(A), \Pi) = 0$, and, as direct consequence $\lim_{n\to\infty} d_\infty(\mathcal{S}_{T_2^n}(A), \Pi) = 0$. Since the minimum copula M obviously fulfills the just mentioned regularity condition we also have $\lim_{n\to\infty} D_1(\mathcal{S}_{T_2^n}(M), \Pi) = 0$. Figure 1 depicts image plots of the kernels of $A, \mathcal{S}_{T_2}(A)$ and $\mathcal{S}_{T_2^2}(A)$ for A being the Marshall Olkin copula with parameters $\alpha = 0.2, \beta = 0.5$ (see [14]), which obviously fulfills the regularity condition. Considering the copula A corresponding to the uniform distribution on the set $[0,1/2]^2 \cup [1/2,1]^2$ we directly get $\mathcal{S}_{T_2}(A) = \Pi$, which shows that in general $\mathcal{S}_T$ is not injective.

Example 2. For every $\theta \in (0,1)$ define a an element $G_\theta \in \mathcal{T}$ by

$$G_\theta(x) = \frac{x}{\theta} \mathbf{1}_{[0,\theta]}(x) \; + \; \frac{x-\theta}{1-\theta} \mathbf{1}_{(\theta,1]}(x).$$

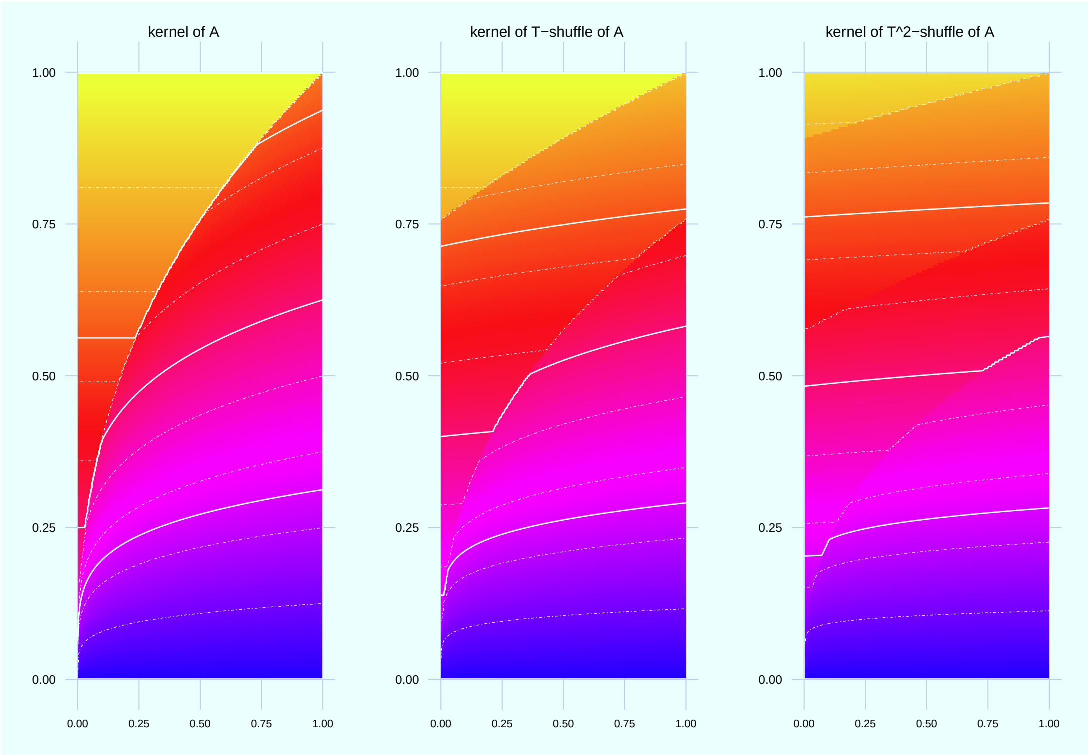

Figure 1: The functions $(x, y) \mapsto K_A(x, [0, y]), K_{\mathcal{S}_{T_2}(A)}(x, [0, y]), K_{\mathcal{S}_{T_2^2}(A)(x,[0,y])}$ for A being the Marhsall Olkin copula $M_{0.2,0.5}$; solid white lines depict the $0.25, 0.5, 0.75$-contours, dashed lines the $0.1, 0.2, \ldots, 0.9$-contours

Using arguments analogous to Example 1 it follows that for each copula $A \in \mathcal{C}$ such that the functions $x \mapsto K_A(x, [0, y])$ are Riemann integrable for each $y \in [0, 1]$ we have

$$\lim_{n\to\infty} D_1\big(\mathcal{S}_{G_\theta^n}(A), \Pi\big) = 0.$$

161 **Example 3.** For every $a \in (0, 1)$ define a transformation $\tau_a : [0, 1] \to [0, 1]$
162 by

$$\tau_a(x) = \begin{cases} 0 & \text{if } x = 0 \\ x + a & \text{if } x \in (0, 1 - a] \\ x + a - 1 & \text{if } x \in (1 - a, 1]. \end{cases}$$

It is straightforward to see that $\tau_a \in \mathcal{T}_p$ and that $\tau_a^{-1} = \tau_{1-a}$. Furthermore (see [12],[21]) it is well-known that τ_a is ergodic if and only if a is irrational.

We will consider two different situations. **(i)** $a = 0.5$**:** Setting $T := \tau_{0.5}$ we have $T^2 = T$ as well as $\mathcal{S}_{T^{2n}}(A) = A$ for every $n \in \mathbb{N}$. Hence, in case of $A \neq \Pi$, $\limsup_{n\to\infty} d_\infty(\mathcal{S}_{T^n}(A), \Pi) > 0$ follows, i.e. contrary to the previous examples the iterated T-shuffles of A do not converge to the product copula Π w.r.t. d_∞. Considering, for instance, the copula A corresponding to the uniform distribution on the set $[0, 3/4]^2 \cup [3/4, 1]^2$ shows that we may even have

$$\limsup_{n\to\infty} d_\infty\left(\frac{1}{n}\sum_{i=1}^{n} \mathcal{S}_{T^i}(A), \Pi\right) > 0.$$

(ii) a **irrational:** Set $T := \tau_a$ and $T' := \tau_{1-a}$. Fix $A \in \mathcal{C}$ and $F \in \mathcal{B}([0, 1])$, then, using Birkhoff's ergodic theorem (see [21]), it follows immediately that

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} K_{\mathcal{S}_{T^i}(A)}(x, F) = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} K_A(T'^i x, F) = \lambda(F)$$

holds for λ-almost all $x \in [0, 1]$. Using Lebesgue's theorem on dominated convergence this yields

$$\lim_{n\to\infty} \int_{[0,1]} \left| \frac{1}{n}\sum_{i=1}^{n} K_{\mathcal{S}_{T^i}(A)}(x, [0, y]) - y \right| d\lambda(x) = 0$$

for every $y \in [0, 1]$ from which (see [19])

$$\lim_{n\to\infty} D_1\left(\frac{1}{n}\sum_{i=1}^{n} \mathcal{S}_{T^i}(A), \Pi\right) = 0 \quad \text{and} \quad \lim_{n\to\infty} d_\infty\left(\frac{1}{n}\sum_{i=1}^{n} \mathcal{S}_{T^i}(A), \Pi\right) = 0$$

follows. Figure 2 depicts image plots of the kernels of A, $B_2 := \frac{1}{2}\sum_{i=1}^{2} \mathcal{S}_{T^i}(A)$ $B_{10} := \frac{1}{10}\sum_{i=1}^{10} \mathcal{S}_{T^i}(A)$ whereby A is again the Marshall Olkin copula with parameters $\alpha = 0.2, \beta = 0.5$ and $T = \tau_{\sqrt{2}/2}$.

4. Main results

The long-time behaviour of the transformations mentioned in Examples 1-3 is very different - the transformations in the first two examples are exact (hence strongly mixing and ergodic), the irrational translation τ_a considered in part (ii) of Example 3 is ergodic (but not strongly mixing) and the rational translation $\tau_{0.5}$ considered in part (i) of Example 3 is only measure-preserving

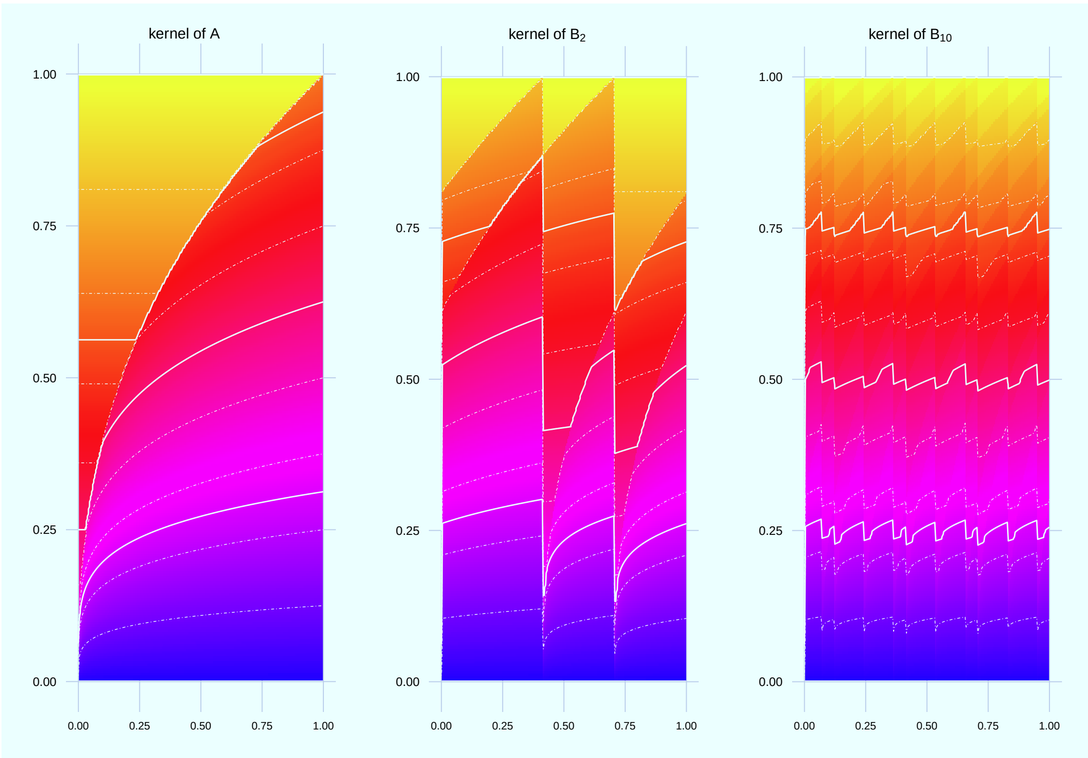

Figure 2: The functions $(x,y) \mapsto K_A(x,[0,y]), K_{B_2}(x,[0,y]), K_{B_{10}}(x,[0,y])$ in Example 3; solid white lines depict the $0.25, 0.5, 0.75$-contours, dashed lines the $0.1, 0.2, \ldots, 0.9$-contours

(but not ergodic). For an overview of mixing/ergodic properties of measure preserving transformations and precise definitions see, for instance, [21] and [12]. According to [12] the just mentioned mixing properties can fully be characterized in terms of the corresponding FPO - the following results hold (for the definition of weak and strong convergence see [12]):

1. T is ergodic if and only if $\frac{1}{n}\sum_{i=1}^{n} P_{T^i} f \to 1$ weakly in $L^1([0,1])$ for all $f \in \mathcal{D}([0,1])$.
2. T is strongly mixing if and only if $P_{T^n} f \to 1$ weakly in $L^1([0,1])$ for all $f \in \mathcal{D}([0,1])$.
3. T is exact if and only if $P_{T^n} f \to 1$ strongly in $L^1([0,1])$ for all $f \in \mathcal{D}([0,1])$.

Here $\mathcal{D}([0,1])$ denotes the family of all probability densities in $L^1([0,1])$). Since for given $F \in \mathcal{B}([0,1])$

$$K_{\mathcal{S}_T(A)}(\cdot, F) = P_T\big(K_A(\cdot, F)\big)$$

holds we can use these results to prove the following theorem.

Theorem 4. *Suppose that $T \in \mathcal{T}$, then the following assertions hold:*

1. *T is ergodic if and only if we have* $\lim_{n\to\infty} d_\infty\Big(\frac{1}{n}\sum_{i=1}^n \mathcal{S}_{T^i}(A), \Pi\Big) = 0$ *for every $A \in \mathcal{C}$.*
2. *T is strongly mixing if and only if we have* $\lim_{n\to\infty} d_\infty\big(\mathcal{S}_{T^n}(A), \Pi\big) = 0$ *for every $A \in \mathcal{C}$.*
3. *If T is exact then* $\lim_{n\to\infty} D_1(\mathcal{S}_{T^n}(A), \Pi) = 0$ *for every $A \in \mathcal{C}$.*

Proof: Suppose that $A \in \mathcal{C}$ and that T is strongly mixing, then for fixed $y \in [0,1]$ we have

$$K_{\mathcal{S}_{T^n}(A)}(\cdot, [0,y]) = P_{T^n}(K_A(\cdot, [0,y])) \longrightarrow \int_{[0,1]} K_A(\cdot, [0,y])d\lambda = y$$

weakly in $L^1([0,1])$, so, in particular for every fixed $x \in [0,1]$

$$\mathcal{S}_{T^n}(A)(x,y) = \int_{[0,x]} K_{\mathcal{S}_{T^n}(A)}(t, [0,y])\, d\lambda(t) \longrightarrow xy = \Pi(x,y)$$

follows. Since x and y were arbitrary this implies $\lim_{n\to\infty} d_\infty(\mathcal{S}_{T^n}(A), \Pi) = 0$. On the other hand if $\lim_{n\to\infty} d_\infty(\mathcal{S}_{T^n}(A), \Pi) = 0$ for every $A \in \mathcal{C}$, then considering $A := M$ and intervals $E := [\underline{e}, \overline{e}], F := [\underline{f}, \overline{f}]$ we get

$$\lambda(T^{-n}(E) \cap F) = \mu_{\mathcal{S}_{T^n}(M)}(E \times F) \longrightarrow \mu_\Pi(E \times F) = \lambda(E)\lambda(F)$$

from which it follows that T is strongly mixing since the class of intervals is a semi-algebra generating $\mathcal{B}([0,1])$ (see [21]). The first assertion of the theorem can be proved analogously.
To prove the last assertion assume that T is exact, then according to the before-mentioned properties of the FPO we have

$$\lim_{n\to\infty} \int_{[0,1]} \big|K_{\mathcal{S}_{T^n}(A)}(x, [0,y]) - y\big| d\lambda(x) = 0.$$

Since y was arbitrary $\lim_{n\to\infty} D_1(\mathcal{S}_{T^n}(A), \Pi) = 0$ follows (see [19]). ■

The next result is a weak converse of the third assertion of Theorem 4:

Lemma 5. *Suppose that $T \in \mathcal{T}$ fulfills $T(B) \in \mathcal{B}([0,1])$ for every $B \in \mathcal{B}([0,1])$. If $\lim_{n\to\infty} D_1(\mathcal{S}_{T^n}(A), \Pi) = 0$ holds for every $A \in \mathcal{C}$ then for every interval I of the form $I = (a, b] \subseteq [0,1]$ with $b > a$ we have*

$$\lim_{n\to\infty} \lambda(T^n(I)) = 1.$$

Proof: Suppose that $I := (a, b]$ with $b > a$. Using, for instance, the ordinal sum of Π w.r.t. the partition $\{[0,a], (a,b], (b,1]\}$ (see [14]) we can construct a copula A whose Markov kernel K_A fulfills $K_A(x, I) = \mathbf{1}_I(x)$. For every $n \in \mathbf{N}$ and every $y \in [0,1]$ set

$$\Phi_{\mathcal{S}_{T^n}(A),\Pi}(y) = \int_{[0,1]} \left| K_{\mathcal{S}_{T^n}(A)}(x, [0,y]) - y \right| d\lambda(x).$$

Then using (see [12])

$$\lambda(I)\lambda(T^n I) = \int_{T^n I} K_{\mathcal{S}_{T^n}(A)}(x, I)\, d\lambda(x) - \underbrace{\int_{T^n I} \Big(K_{\mathcal{S}_{T^n}(A)}(x, I) - \lambda(I) \Big) d\lambda(x)}_{\delta_n}$$

together with the fact that $|\delta_n| \leq \Phi_{\mathcal{S}_{T^n}(A),\Pi}(a) + \Phi_{\mathcal{S}_{T^n}(A),\Pi}(b) \longrightarrow 0$ for $n \to \infty$ we get

$$\begin{aligned} \liminf_{n\to\infty} \lambda(I)\lambda(T^n I) &\geq \liminf_{n\to\infty} \int_{T^n I} K_{\mathcal{S}_{T^n}(A)}(x, I)\, d\lambda(x) \\ &= \liminf_{n\to\infty} \int_{T^{-n}(T^n I)} K_A(x, I)\, d\lambda(x) \geq \lambda(I). \end{aligned}$$

This completes the proof. ■

Remark 1. An alternative way to prove the first two assertions of Theorem 4 would be to use

$$\mathcal{S}_{T^n}(A) = (\underbrace{C_T * C_T * \cdots * C_T}_{n \text{ times}})^t * A = (C_{T^n})^t * A,$$

together with the fact that, firstly, $\lim_{n\to\infty} d_\infty(C_{T^n}, \Pi) = 0$ if and only if T is strongly mixing and secondly, that $\lim_{n\to\infty} d_\infty(\frac{1}{n}\sum_{i=1}^n C_{T^i}, \Pi) = 0$ if and only if T is ergodic.

We can use the same technique to give an alternative short proof of the following result in [6]:

Theorem 6 ([6]). *For every copula C the independence copula Π can be approximated uniformly by elements of $\mathcal{T}_p(C)$.*

Proof: Choose a weakly mixing $T \in \mathcal{T}_p$ (see [1]), then T^{-1} is weakly mixing too. Consequently, by a direct application of Lemma 15 in [6] (also see [21]), there exists a strictly increasing sequence $(n_j)_{j\in\mathbb{N}}$ in $\mathbf{N}$ such that for all $x, y \in [0,1]$

$$\begin{aligned}\lim_{j\to\infty} \mathcal{S}_{T^{n_j}}(A)(x,y) &= \lim_{j\to\infty} \int_{[0,x]} K_{\mathcal{S}_{T^{n_j}}(A)}(t,[0,y])d\lambda(t) \\ &= \lim_{j\to\infty} \int_{[0,x]} K_A(T^{-n_j}t,[0,y])d\lambda(t) = xy. \blacksquare\end{aligned}$$

Before taking a look to another mapping $\mathcal{U}_T : \mathcal{C} \to \mathcal{C}$ which is strongly connected to $\mathcal{S}_T$ we will show that for every copula $A \in \mathcal{C}$ and every $T \in \mathcal{T}$ the mass of the singular component of $\mathcal{S}_T(A)$ cannot be greater than the mass of singular component of A (already proved in [6] for the special case of absolutely continuous copulas and $T \in \mathcal{T}_p$). For every $A \in \mathcal{C}$ we will write $\mu_A = \mu_A^{abs} + \mu_A^{sing}$ for the Lebesgue decomposition of μ_A w.r.t. λ_2 and denote the mass of the singular component μ_A^{sing} by $sing(A)$, i.e.

$$sing(A) := \mu_A^{sing}([0,1]^2).$$

In general neither $\frac{\mu_A^{abs}}{1-sing(A)}$ nor $\frac{\mu_A^{abs}}{sing(A)}$ is a doubly stochastic measure, so A cannot be expressed as a convex combination of an absolutely continuous and a singular copula.

Theorem 7. *For every $T \in \mathcal{T}$ and $A \in \mathcal{C}$ $sing(\mathcal{S}_T(A)) \leq sing(A)$ holds.*

Proof: Fix $T \in \mathcal{T}$ and define a measurable mapping $\Phi_T : [0,1]^2 \to [0,1]^2$ by

$$\Phi_T(x,y) = (x, Ty).$$

Since T is λ-preserving Φ_T is λ_2-preserving. It is straightforward to verify that for every $N \in \mathcal{B}([0,1]^2)$ and every $x \in [0,1]$ we have $\left(\Phi_T^{-1}(N)\right)_x =$

$T^{-1}(N_x)$. Hence, for every $N \in \mathcal{B}([0,1]^2)$ with $\lambda_2(N) = 0$, using Lemma 2, disintegration and $\lambda_2(\Phi_T^{-1}(N)) = 0$ it follows that

$$\begin{aligned}
\mu_{A^t * C_T}(N) &= \int_{[0,1]} K_{A^t * C_T}(x, N_x) d\lambda(x) = \int_{[0,1]} K_{A^t}(x, T^{-1}(N_x)) d\lambda(x) \\
&= \int_{[0,1]} K_{A^t}\big(x, (\Phi_T^{-1}(N))_x\big) d\lambda(x) = \mu_{A^t}\big(\Phi_T^{-1}(N)\big) \\
&\leq sing(A^t) = sing(A),
\end{aligned}$$

which completes the proof since N was an arbitrary Borel set with $\lambda_2(N) = 0$ and $sing(B) = sing(B^t)$ for all $B \in \mathcal{C}$. ■

Remark 2. It is not difficult to construct examples showing that in general we do not have equality in $sing(\mathcal{S}_T(A)) \leq sing(A)$: According to [3] (also see [4] and [10]) we can find λ-preserving functions $f, g \in \mathcal{T}$ such that (same notation as in [3] and [4])

$$\Pi(x, y) = \lambda\big(f^{-1}([0, x]) \cap g^{-1}([0, y])\big) := A_{f,g}(x, y)$$

holds for all $x, y \in [0, 1]$. Set $A := A_{id,g}$, then we have (see [3])

$$S_f(A) = (C_f)^t * A = A_{f;id} * A_{id;g} = A_{f;g} = \Pi$$

so $sing(S_f(A)) = 0$ although $sing(A) = 1$.

As final step we introduce another operator $\mathcal{U}_T : \mathcal{C} \to \mathcal{C}$ that has already appeared in Proposition 3, Example 3 and the proof of Theorem 6. Again consider $T \in \mathcal{T}$ and define $\mathcal{U}_T$ via the following transformation of the corresponding kernels:

$$K_{\mathcal{U}_T(A)}(x, E) := K_A(Tx, E)$$

It straightforward to verify that $\mathcal{U}_T$ is well-defined (i.e. that $K_{\mathcal{U}_T(A)}$ really is the Markov kernel of a copula) and that $\mathcal{U}_T$ is an isometry on the space $(\mathcal{C}, D_1)$. According to [12] $K_{\mathcal{U}_T(A)}(\cdot, E)$ is just the so-called *Koopman operator* applied to the function $K_A(\cdot, E)$. It is well known that the Koopman operator is the adjoint of the FPO (again see [12]). The following theorem describes the interrelation between $\mathcal{S}_T$ and $\mathcal{U}_T$ in our setting:

Theorem 8. *For every $T \in \mathcal{T}$ the following assertions hold:*

1. $\mathcal{S}_T \circ \mathcal{U}_T(A) = A$ *for every* $A \in \mathcal{C}$.

2. *For all Borel sets E, F we have*

$$\mu_{\mathcal{U}_T \circ \mathcal{S}_T(A)}(T^{-1}(E) \times F) = \mu_A(T^{-1}(E) \times F)$$

Consequently in case the σ-field $T^{-1}(\mathcal{B}([0,1]))$ generated by T coincides with $\mathcal{B}([0,1])$ also $\mathcal{U}_T \circ \mathcal{S}_T(A) = A$ holds for every $A \in \mathcal{C}$.

3. *$sing(\mathcal{U}_T(A)) \geq sing(A)$ for every $A \in \mathcal{C}$.*
4. *$\mathcal{U}_T(A) = C_T * A$ for every $A \in \mathcal{C}$.*

Proof: The first assertion is a direct consequence of the following equation, which holds for all $E, F \in \mathcal{B}([0,1])$:

$$\begin{aligned}
\int_E K_{\mathcal{S}_T \circ \mathcal{U}_T(A)}(x, F) d\lambda(x) &= \int_{T^{-1}(E)} K_{\mathcal{U}_T(A)}(x, F) d\lambda(x) \\
&= \int_{T^{-1}(E)} K_A(Tx, F) d\lambda(x) \\
&= \int_{[0,1]} \mathbf{1}_E(Tx) K_A(Tx, F) d\lambda(x) \\
&= \int_E K_A(x, F) d\lambda(x).
\end{aligned}$$

To prove the second assertion note that for all $E, F \in \mathcal{B}([0,1])$ we have

$$\begin{aligned}
\mu_{\mathcal{U}_T \circ \mathcal{S}_T(A)}(T^{-1}(E) \times F) &= \int_{T^{-1}(E)} K_{\mathcal{U}_T \circ \mathcal{S}_T(A)}(x, F) d\lambda(x) \\
&= \int_{[0,1]} \mathbf{1}_E(Tx) K_{\mathcal{S}_T(A)}(Tx, F) d\lambda(x) \\
&= \int_{[0,1]} \mathbf{1}_E(x) K_{\mathcal{S}_T(A)}(x, F) d\lambda(x) \\
&= \mu_A(T^{-1}(E) \times F).
\end{aligned}$$

The third assertion is a direct consequence of the first one and Theorem 7 and assertion four follows from (δ_{Tx} denoting the Dirac measure in Tx)

$$\begin{aligned}
K_{C_T * A}(x, E) &= \int_{[0,1]} K_A(y, E) K_{C_T}(x, dy) = \int_{[0,1]} K_A(y, E) \, d\delta_{Tx}(y) \\
&= K_A(Tx, E) \quad \blacksquare
\end{aligned}$$

Given the strong interrelation between $\mathcal{S}_T$ and $\mathcal{U}_T$ the following result is not surprising:

Theorem 9. *Suppose that $T \in \mathcal{T}$, then the following two assertions hold:*

1. *T is ergodic if and only if we have $\lim_{n\to\infty} d_\infty\Big(\frac{1}{n}\sum_{i=1}^n \mathcal{U}_{T^i}(A), \Pi\Big) = 0$ for every $A \in \mathcal{C}$.*
2. *T is strongly mixing if and only if we have $\lim_{n\to\infty} d_\infty\big(\mathcal{U}_{T^n}(A), \Pi\big) = 0$ for every $A \in \mathcal{C}$.*

Proof: We will only prove the first assertion, the second one can be proved analogously. If T is ergodic, then for every fixed $y \in [0,1]$ Birkhoff's ergodic theorem (see [21]) implies

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n K_{\mathcal{U}_{T^i}(A)}(x,[0,y]) = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n K_A(T^i x,[0,y]) = y$$

for λ-almost all $x \in [0,1]$ for which, using Lebesgue's theorem on dominated converges, one implication immediately follows. To prove the other one consider intervals $E := [\underline{e}, \overline{e}], F := [\underline{f}, \overline{f}]$. Then the implication directly follows from

$$\begin{aligned} \frac{1}{n}\sum_{i=1}^n \lambda\big(T^{-i}(E)\cap F\big) &= \frac{1}{n}\sum_{i=1}^n \int_{[0,1]} \mathbf{1}_E(T^i x)\mathbf{1}_F(x)d\lambda(x) \\ &= \frac{1}{n}\sum_{i=1}^n \int_{[0,1]} K_M(T^i x, E)\mathbf{1}_F(x)d\lambda(x) \\ &= \frac{1}{n}\sum_{i=1}^n \mu_{\mathcal{U}_{T^i}(M)}(F\times E) \longrightarrow \lambda(F)\lambda(E). \ \blacksquare \end{aligned}$$

The following final example shows that a result analogous to point three in Theorem 4 does not hold for $\mathcal{U}_T$.

Example 4. Consider the exact transformations G_θ from Example 2. Then, according to Theorem 9 we have $\lim_{n\to\infty} d_\infty\big(\mathcal{U}_{D_\theta^n}(A), \Pi\big) = 0$ for every $A \in \mathcal{C}$. Nevertheless we cannot expect convergence w.r.t. D_1 for all $A \in \mathcal{C}$: If we choose $A := C_T$ for some $T \in \mathcal{T}$ then $\mathcal{U}_{G_\theta}(A)$ is the gluing copula of A with A along the line $x = \theta$ (see [18]) and we have $\mathcal{U}_{G_\theta^n}(A) = C_{T\circ G_\theta^n}$ for every $n \in \mathbb{N}$. Hence $\mathcal{U}_{G_\theta^n}(A)$ is completely dependent and, according to [19], $D_1(\mathcal{U}_{G_\theta^n}(A), \Pi) = 1/3$ holds for every n.

Acknowledgments

The authors gratefully acknowledge the support of the grants MTM2009-09440-C02-02 and MTM2011-22394 from the Spanish Ministry of Science and Innovation.

References

[1] A. Avila, G. Forni: Weak mixing for interval exchange transformations and translation flows, *Ann. Math.* **165**, 637-664 (2007)

[2] W.F. Darsow, B. Nguyen, E.T. Olsen: Copulas and Markov processes, *Illinois J. Math.* **36**, no. 4, 600-642 (1992)

[3] W.F. Darsow, E.T Olsen: Characterization of idempotent 2-copulas, *Note Mat.* **30**, no. 1, 147-177 (2010)

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

[5] J. Ding, Q. Du, T.Y. Li: The Spectral Properties of the Frobenius Perron Operator, *J. Math. Anal. Appl.* **184**, 285-301 (1994)

[6] F. Durante, P. Sarkoci, C. Sempi: Shuffles of copulas, *J. Math. Anal. Appl.* **352**, 914-921 (2009)

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

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

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

[10] A. Kolesárová, R. Mesiar, C. Sempi: Measure-preserving transformations, copulae and compatibility, *Mediterr. J. Math.* **5**, 325-339 (2008)

[11] H.O. Lancaster, Correlation and complete dependence of random variables, *Ann. Math. Stat.* **34**, 1315-1321 (1963)

[12] A. Lasota, M.C. Mackey: *Chaos, Fractals and Noise - Stochastic Aspects of Dynamics*, Springer New York, 1994

[13] P. Mikusínski, H. Sherwood, M.D. Taylor: Shuffles of Min, *Stochastica* **13**, 61-74 (1992).

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

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

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

[17] C. Sempi: Copulae: Some mathematical aspects, *Appl. Stoch. Model. Bus.* **27**, 37-50 (2011)

[18] K.F. Siburg, P.A. Stoimenov: Gluing copulas, *Comm. Statist. A - Theory Methods* **37**, 3124-3134 (2008)

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

[20] W. Trutschnig, J. Fernández Sánchez: Idempotent and multivariate copulas with fractal support, *J. Stat. Plan. Infer.*, to appear

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