On Bertino copulas and the Markov product

Henrik Kaiser[a], Wolfgang Trutschnig[a,*]

[a] *Department for Artificial Intelligence and Human Interfaces, University Salzburg,*
Hellbrunnerstrasse 34, A-5020 Salzburg, Austria
Tel: +43 662 8044 5326 Fax: +43 662 8044 137

Abstract

We revisit the family of bivariate Bertino copulas, first studied by Fredricks and Nelsen in 2002 and characterized as the minimal elements in the family of all copulas with given diagonal, and derive several (partially surprising) results on their behavior under the so-called Markov product (also known as star product) of copulas. Using tools from measure theory and Markov chains in discrete time we first show that for every Bertino copula B_δ with diagonal δ all Markov product iterates $B_\delta^{*n} = B_\delta * \cdots * B_\delta$ are Bertino copulas too. Having this, we then prove that for each Bertino copula B_δ whose diagonal fulfills a simple but essential regularity/slope condition, the sequence $(B_\delta^{*n})_{n\in\mathbb{B}}$ converges uniformly (in fact even weakly conditional) to some copula C, and show that C again is a Bertino copula which, additionally, is idempotent. Rounding off these results we provide a handy characterization of idempotence in the Bertino family.

Keywords: Copula, Doubly stochastic measure, Markov product, Markov kernel, Idempotence, Markov chain

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

---

1. Introduction

Based on the observation that the distribution function H of each pair (X, Y) of continuous random variables X, Y can be factorized in the marginal distribution functions F_X and F_Y, respectively, and a unique bivariate copula C in terms of $H(x, y) = C(F_X(x), F_Y(y))$, over the past decades copu-

---

[*] Corresponding author

Email addresses: `henrik.kaiser@plus.ac.at` (Henrik Kaiser), `wolfgang@trutschnig.net` (Wolfgang Trutschnig)

las have become one of the key concepts for modeling dependence of random quantities. As pointed out by Jaworski in [14] there are numerous reasons for the study of bivariate copulas with given diagonal $\delta : [0,1] \to [0,1]$ - the facts that, firstly, tail dependence of a copula C only considers the copula's diagonal and that, secondly, for a pair (X,Y) of uniformly $[0,1]$-distributed random variables X,Y with copula C the maximum $\max(X,Y)$ has distribution function $\delta_C(x) = C(x,x)$ potentially being the two most important ones. One natural question arising in this context is, how much flexibility we have in the class $\mathcal{C}_\delta$ of all copulas with given diagonal δ, or, more generally, how 'large' $\mathcal{C}_\delta$ is. In [10] Fredricks and Nelsen showed that the Bertino copula B_δ is the minimal element in $\mathcal{C}_\delta$, i.e., every copula $C \in \mathcal{C}_\delta$ fulfills $B_\delta(x,y) \leq C(x,y)$ for all $x,y \in [0,1]^2$. Furthermore (see [19, 20]) it is known that the diagonal copula is the maximal element in the subclass of all symmetric copulas in $\mathcal{C}_\delta$. Durante and Jaworski ([5]) formulated necessary and sufficient conditions for a diagonal δ to be the diagonal of an absolutely continuous copula, Durante et al. constructed asymmetric elements of $\mathcal{C}_\delta$ via so-called patchworks (see [7]).

We here revisit the family $\mathcal{C}_{Ber}$ of all bivariate Bertino copulas (for a multivariate extension of Bertino copulas see [1]) and study the behavior of Bertino copulas under the so-called Markov product $*$ of copulas. The Markov product was introduced by Darsow et el. in 1992 under the name 'star product' (see [4]). Four years later Olsen et al. (see [12]) closely linked the family $(\mathcal{C}, *)$ of all bivariate copulas with the Markov product as binary operation and the space $(\mathcal{M}, \circ)$ of Markov operators on $L^1([0,1])$ with the standard composition as binary operation by showing that these space are isomorphic. A second reason for the name 'Markov product' being adequate was provided in [24], where the authors showed that calculating the Markov product $A * B$ of copulas $A, B \in \mathcal{C}$ corresponds to considering the standard composition of the Markov kernels $K_A(\cdot,\cdot)$ and $K_B(\cdot,\cdot)$ (transition probabilities) well known in the context of Markov chains in discrete time. In particular, we have the following: if $X_1, X_2, X_3, \ldots$ is a stationary Markov chain of uniformly $[0,1]$-distributed random variables with state space $[0,1]$ and transition probability $K_C(\cdot,\cdot)$, then the 2-step transition probability is given by $K_{C*C}(\cdot,\cdot)$.

Looking at frequently used subfamilies of $\mathcal{C}$ like the Extreme-Value or the Archimedean class $\mathcal{C}_{EV}$ and $\mathcal{C}_{Ar}$, respectively, it can be shown that neither of them is closed with respect to the Markov product, i.e., the Markov product of two Extreme-Value copulas is not necessarily an Extreme-Value copula, and the same is true for $\mathcal{C}_{Ar}$. In fact, as shown recently in [18], apart from

(i) parametric classes like the Fréchet and the Gauss family, (ii) the family of all completely dependent copulas, and (iii) the family of all checkerboards with fixed resolution, to the best of the authors' knowledge so-called lower semilinear copulas (which, analogous to Bertino copulas are characterized via diagonals) constitute the only family being closed with respect to the Markov product.

In the current paper we will show that the Markov product $B_{\delta_1} * B_{\delta_2}$ of two Bertino copulas is not necessarily a Bertino copula, so (analogously to $\mathcal{C}_{Ev}$ and $\mathcal{C}_{Ar}$) the Bertino family $\mathcal{C}_{Ber}$ is not closed. The situation, however, changes when we consider the Markov product $B_\delta^{*2} := B_\delta * B_\delta$ of a Bertino copula B_δ with itself - in this case B_δ^{*2} is indeed a Bertino copula too. In fact, much more is true: we will show that for every $n \in \mathbb{N}$ and every Bertino copula B_δ the iterated Markov product $B_\delta^{*n} = B_\delta * \ldots * B_\delta$ fulfills $B_\delta^{*n} \in \mathcal{C}_{Ber}$. Moreover, for every Bertino copula B_δ whose diagonal fulfills a certain regularity condition the sequence $(B_\delta^{*n})_{n\in\mathbb{N}}$ will be shown to converge uniformly to an idempotent Bertino copula, i.e., to a Bertino copula C fulfilling $C * C = C$. Since the limit C is idempotent we complement the afore-mentioned results by providing a simple characterization of idempotence in the Bertino family.

We remark that these convergence results are quite surprising in so far as for general copulas $A \in \mathcal{C}$ iterates of the Markov product A^{*n} do not need to converge (see [25]) and - even if they do - determining the limit is often out of reach.

The remainder of this paper is organized as follows: Section 2 gathers some notations and preliminaries, and recalls some facts about Bertino copulas already established in the literature. After motivating how the limit behavior of $(B_\delta^{*n})_{n\in\mathbb{N}}$ might look like, all afore-mentioned results are stated and proved in Section 3. Several examples and graphics illustrate the studied procedures and underlying ideas.

2. Notation and preliminaries

For every metric space (Ω, ρ) the Borel σ-field on Ω will be denoted by $\mathcal{B}(\Omega)$, the family of all probability measures on $\mathcal{B}(\Omega)$ by $\mathcal{P}(\Omega)$. For every $\vartheta \in \mathcal{P}(\Omega)$ the support of ϑ, defined as the set of all $x \in \Omega$ such that every non-empty open ball $B(x, r)$ of radius $r > 0$ around x has positive mass, will be denoted by $\mathrm{supp}(\vartheta)$. For every $z \in \Omega$ the Dirac measure at z will be denoted by ϵ_z, furthermore the Lebesgue measure on $\mathcal{B}([0,1]^2)$ and $\mathcal{B}([0,1])$ will be denoted by λ_2 and λ, respectively.

As already mentioned before, $\mathcal{C}$ will denote the family of all two-dimensional *copulas*, i.e., the family of distribution functions (restricted to $[0,1]^2$) of random vectors (X,Y) on a probability space $(\Omega,\mathcal{A},\mathbb{P})$, fulfilling that the marginal distributions $\mathbb{P}^X,\mathbb{P}^Y \in \mathcal{P}(\mathbb{R})$ coincide with the Lebesgue measure λ on $[0,1]$. For any fixed $C \in \mathcal{C}$ we will write $(U,V) \sim C$ if C is the distribution function of (U,V) restricted to $[0,1]^2$; in this case we obviously have that U,V are uniformly distributed on $[0,1]$. Letting d_∞ denote the uniform metric on $\mathcal{C}$ it is well known that $(\mathcal{C},d_\infty)$ is a compact metric space and that (due to Lipschitz continuity) in $\mathcal{C}$ pointwise and uniform convergence are equivalent. M will denote the minimum copula, Π the product copula and W the lower Fréchet-Hoeffding bound. Given $C \in \mathcal{C}$ the transpose $C^t \in \mathcal{C}$ of C is defined by $C^t(x,y) := C(y,x)$. For every $C \in \mathcal{C}$ the corresponding doubly stochastic measure will be denoted by μ_C and $\mathcal{P}_{\mathcal{C}} \subset \mathcal{P}([0,1]^2)$ will denote the family of all *doubly stochastic measures.*

Definition 1 (Diagonal). A function $\delta : [0,1] \to [0,1]$ is called a *diagonal* if, and only if the following four conditions are fulfilled:

1. $\delta(0) = 0$ and $\delta(1) = 1$.
2. δ is non-decreasing.
3. δ is Lipschitz continuous with Lipschitz constant $L = 2$.
4. $\delta(t) \leq t$ for all $t \in [0,1]$.

In the sequel $\mathcal{D}$ will denote the family of all diagonals.

It is well-known (see [8, 19] and the references therein) that for every $\delta \in \mathcal{D}$ there exists some copula $C \in \mathcal{C}$ such that $\delta(x) = C(x,x)$ for all $x \in [0,1]$, and that, vice versa, for every $C \in \mathcal{C}$ the function $x \mapsto C(x,x)$ is an element of $\mathcal{D}$. As a direct consequence, for every $\delta \in \mathcal{D}$ the class $\mathcal{C}_\delta \subseteq \mathcal{C}$ of all copulas $C \in \mathcal{C}$ having diagonal δ is non-empty. For more background on copulas and doubly stochastic measure we refer to the textbooks [8] and [19].

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 $x \in \mathbb{R}$. Given real-valued random variables X,Y on a joint probability space $(\Omega,\mathcal{A},\mathbb{P})$, 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 for $\mathbb{P}$-almost every $\omega \in \Omega$. 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 $\mathbb{P}^X$-almost everywhere (i.e., unique for $\mathbb{P}^X$-almost all $x \in \mathbb{R}$) and that $K(\cdot,\cdot)$ only depends on the joint distribution $\mathbb{P}^{(X,Y)} \in \mathcal{P}(\mathbb{R}^2)$ of (X,Y). Hence, given $C \in \mathcal{C}$ we will denote (a version of) the regular conditional distribution of Y given X by $K_C(\cdot,\cdot)$, view it as a function mapping $[0,1] \times \mathcal{B}([0,1])$ to $[0,1]$ and refer to $K_C(\cdot,\cdot)$ simply as *regular conditional distribution of* C or as *Markov kernel of* C. Note that for every $C \in \mathcal{C}$, its Markov kernel $K_C(\cdot,\cdot)$, 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 $x \in [0,1]$)

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

So in particular

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

holds 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 equation (3) induces a unique element $\mu \in \mathcal{P}_{\mathcal{C}}$ via equation (2). Notice that for fixed $y \in [0,1]$, λ-almost every $x \in [0,1]$ is a Lebesgue point (see [22]) of the mapping $x \mapsto K(x,[0,y])$, implying that the identity

$$\frac{\partial C(x,y)}{\partial x} = K(x,[0,y]) \tag{4}$$

holds (for such x). For more details and properties of conditional expectation, regular conditional distributions, and disintegration we refer to the textbooks [15] and [17].

Viewing bivariate copulas in terms of their Markov kernels/conditional distributions allows to define notions of convergence on $\mathcal{C}$ that are (much) stronger than uniform convergence. Following [16] a sequence $(C_n)_{n\in\mathbb{N}}$ of copulas converges *weakly conditional* to $C \in \mathcal{C}$ if, and only if for λ-almost every $x \in [0,1]$ we have that the sequence $(K_{C_n}(x,\cdot))_{n\in\mathbb{N}}$ of probability measures on $\mathcal{B}([0,1])$ converges weakly to the probability measure $K_C(x,\cdot)$. Moreover, following [23] and setting

$$D_1(A,B) := \int_{[0,1]^2} |K_A(x,[0,y]) - K_B(x,[0,y])|\, d\lambda_2(x,y)$$

for all $A, B \in \mathcal{C}$ yields a complete metric space $(\mathcal{C}, D_1)$. It is straightforward to verify (see [23, Lemma 7]) that (i) weak conditional convergence implies convergence w.r.t. D_1, that (ii) convergence w.r.t. D_1 implies convergence w.r.t. d_∞ (again see [23]), and that (iii) the reverse implications do not hold in general.

In the sequel, the so-called *Markov product* (a.k.a. *star product*) $A * B$ of copulas $A, B \in \mathcal{C}$ will play a prominent role. Letting ∂_i denote the partial derivative with respect to the i-th coordinate, the Markov product $A * B$ is defined by

$$(A * B)(x, y) := \int\limits_{[0,1]} \partial_2 A(x, s) \cdot \partial_1 B(s, y) \, d\lambda(s) \tag{5}$$

for all $x, y \in [0, 1]$. It is well known (see [19]) that $A*B$ is again a copula, that the star product is in general not commutative, i.e., $A * B \neq B * A$ can hold, that Π is the null- and M the unit element in $(\mathcal{C}, *)$, i.e., $A * \Pi = \Pi * A = \Pi$ and $A * M = M * A = A$ for every $A \in \mathcal{C}$. Moreover (see [4, 19]), for each pair $A, B \in \mathcal{C}$ we have $(A * B)^t = B^t * A^t$. A copula A is called *idempotent*, if $A * A = A$ holds. Using Markov kernels and eq. (4) it is straightforward to verify that the following identity holds:

$$(A * B)(x, y) = \int\limits_{[0,1]} K_{A^t}(s, [0, x]) K_B(s, [0, y]) \, d\lambda(s). \tag{6}$$

Moreover, as shown in [26], using disintegration it follows that a (version of the) Markov kernel of $A * B$ is given by the standard composition of the Markov kernels of A and B, a concept well-known in the context of Markov chains in discrete time. In other words, $K_A \circ K_B$, defined by

$$(K_A \circ K_B)(x, F) = \int\limits_{[0,1]} K_B(s, F) K_A(x, dy) \tag{7}$$

for every $x \in [0, 1]$ and $F \in \mathcal{B}[0, 1]$ is a (version of the) Markov kernel of $A*B$. Translating to the Markov chain setting, if $(X_n)_{n\in\mathbb{N}}$ is a Markov chain of random variables uniformly distributed on $[0, 1]$ with one-step transition probability $K_C(\cdot, \cdot)$ for some fixed $C \in \mathcal{C}$, then the two-step transition probabilities are given by $K_{C*C}(\cdot, \cdot) = (K_C \circ K_C)(\cdot, \cdot)$. As mentioned in the previous section, this very fact is the main reason for referring to $A * B$ as the Markov (instead of the star product) of $A, B \in \mathcal{C}$. It can be shown that the Markov product is jointly continuous w.r.t. D_1, i.e., for sequences $(A_n)_{n\in\mathbb{N}}, (B_n)_{n\in\mathbb{N}}$ converging to $A, B \in \mathcal{C}$ w.r.t. D_1 we also have

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

but not jointly continuous w.r.t. d_∞ (see [4]). For more background on the Markov product we refer to [4, 19, 25, 26] and the references therein.

We close this section by recalling some basic facts about *Bertino copulas* already established in the literature and start with additional properties of diagonals. For every diagonal $\delta \in \mathcal{D}$ let $\hat{\delta} : [0,1] \to [0, \frac{1}{2}]$ be defined by

$$\hat{\delta}(x) = x - \delta(x).$$

It is straightforward to verify that for arbitrary $\delta \in \mathcal{D}$ the function $\hat{\delta} : [0,1] \to [0, \frac{1}{2}]$ satisfies $\hat{\delta}(0) = \hat{\delta}(1) = 0$, that $\hat{\delta}$ is Lipschitz continuous, with Lipschitz constant 1, and that $0 \leq \hat{\delta}(t) \leq \min\{t, 1-t\}$ holds for all $t \in [0,1]$. Due to Lipschitz continuity both δ and $\hat{\delta}$ are differentiable λ-almost everywhere and we can find Borel measurable functions $w_\delta : [0,1] \to [0,2]$ and $\hat{w}_\delta : [0,1] \to [-1,1]$ fulfilling $\delta'(x) = w_\delta(x)$ and $\hat{\delta}'(x) = \hat{w}_\delta(x)$ for λ-almost every $x \in [0,1]$ (see, e.g., [22]). In the sequel we will refer to w_δ and $\hat{w}_\delta$ as (measurable versions of) the derivative of δ and $\hat{\delta}$, respectively. Despite being analytically regular in the sense of Lipschitz continuity, according to [11, Theorem 3.2] it is possible to construct a diagonal δ_0 exhibiting the property that $\hat{\delta}_0$ is not monotonic on any non-empty open interval since

$$\lambda\Big(\{x \in (a,b) : \hat{\delta}_0'(x) = 1\}\Big) > 0, \quad \lambda\Big(\{x \in (a,b) : \hat{\delta}_0'(x) = -1\}\Big) > 0$$

for every non-empty open interval $(a,b) \subseteq [0,1]$. We will refer to this diagonal several times in what follows since it illustrates, why other seemingly natural (interval-based) methods of proof may fail.
Following [10] in the sequel we will work with the subsequent definition of a Bertino copula (also see [2] for the original, more general construction by S. Bertino).

Definition 2 (Bertino copula). For every diagonal $\delta \in \mathcal{D}$ the Bertino copula B_δ is defined by

$$B_\delta(x,y) := M(x,y) - \min\Big\{\hat{\delta}(t) : t \in \big[\min\{x,y\}, \max\{x,y\}\big]\Big\}. \tag{8}$$

Obviously $B_\delta \in \mathcal{C}_\delta$ holds. As mentioned in the previous section, $\mathcal{C}_{Ber}$ will denote the family of all Bertino copulas. According to [10], B_δ is the minimal element in $\mathcal{C}_\delta$, i.e., for every $C \in \mathcal{C}_\delta$ we have $B_\delta(x,y) \leq C(x,y)$ for all $x, y \in [0,1]$. Defining the functions $l, u : [0,1] \to [0,1]$ by

$$\begin{aligned} u(x) &:= \max\big\{y \geq x : \hat{\delta}(t) \geq \hat{\delta}(x) \text{ for all } t \in [x,y]\big\} \\ l(x) &:= \min\big\{y \leq x : \hat{\delta}(t) \geq \hat{\delta}(x) \text{ for all } t \in [y,x]\big\} \end{aligned} \tag{9}$$

we have that u is upper semicontinuous and l lower semicontinuous (see [11]). Moreover (again see [11]) a version of the Markov kernel K_{B_δ} of B_δ is given by

$$K_{B_\delta}(x, E) = \begin{cases} (1 - \hat{w}_\delta(x))\epsilon_x(E) \,+\, \hat{w}_\delta(x)\epsilon_{u(x)}(E) & \text{if } \hat{w}_\delta(x) > 0 \\ (1 + \hat{w}_\delta(x))\epsilon_x(E) \,-\, \hat{w}_\delta(x)\epsilon_{l(x)}(E) & \text{if } \hat{w}_\delta(x) \leq 0, \end{cases} \tag{10}$$

whereby $\hat{w}_\delta : [0, 1] \to [-1, 1]$ denotes a measurable version of the derivative of $\hat{\delta}$. Notice that we only need the function $\hat{w}_\delta$ in order to assure that the kernel in eq. (10) has the property that the mapping $x \mapsto K_{B_\delta}(x, E)$ is measurable for every fixed $E \in \mathcal{B}([0, 1])$. Considering that for $\hat{w}_\delta(x) = 0$ the kernel fulfills $K_{B_\delta}(x, \{x\}) = 1$ it follows that (see [11]) the support of μ_{B_δ} is contained in the union of the diagonal $\Delta := \{(x, x) : x \in [0, 1]\}$ and the closure of the graph $\Gamma(S) = \{(x, S(x)) : x \in [0, 1]\}$ of the function $S : [0, 1] \to [0, 1]$, defined by

$$S := \begin{cases} u(x) & \text{if } \hat{w}_\delta(x) > 0 \\ l(x) & \text{if } \hat{w}_\delta(x) \leq 0. \end{cases} \tag{11}$$

Notice that measurability of S is a direct consequence of measurability of u, l and $\hat{w}_\delta$.

In [10, Section 3] Fredricks and Nelsen provided the following stochastic characterization of Bertino copulas, which will be useful in the next section:

Theorem 3 ([10]). *For $(U, V) \sim C$ the following two conditions are equivalent:*

1. *C is a Bertino copula.*
2. *For all $u, v \in [0, 1]$ there exists some $t \in [\min(u, v), \max(u, v)]$ such that*

$$\begin{aligned} &\mathbb{P}\left(\min(U, V) \leq \min(u, v),\, \max(U, V) > \max(u, v)\right) \\ &\quad = \mathbb{P}\left(\min(U, V) \leq t < \max(U, V)\right). \end{aligned} \tag{12}$$

In the sequel the following symmetric version of Theorem 3, formulated directly for the corresponding doubly stochastic measure, will be key.

Corollary 4. *Suppose that $C \in \mathcal{C}$ is symmetric. Then the identity (12) holds if, and only if for $0 \leq u \leq v \leq 1$ there exists some $t \in [u, v]$ with*

$$\mu_C\left([0, t] \times (t, v]\right) = 0 = \mu_C\left([u, t] \times (t, 1]\right). \tag{13}$$

Proof. Without loss of generality, suppose that $u \leq v$ and that $t \in [u, v]$. Using symmetry of C, the left hand side of eq. (12) can be written in the form

$$\begin{aligned}\mathbb{P}\left(\min(U,V) \leq u,\, \max(U,V) > v\right) &= 2\,\mathbb{P}\left(U < V, U \leq u,\, V > v\right)\\ &= 2\,\mathbb{P}\left(U \leq u,\, V > v\right)\\ &= 2\mu_C\left([0,u] \times (v,1]\right).\end{aligned}$$

Proceeding analogously, the right-hand side of eq. (12) simplifies to

$$\mathbb{P}\left(\min(U,V) \leq t < \max(U,V)\right) = 2\mu_C([0,t] \times (t,1]).$$

Altogether it follows that eq. (12) holds if, and only if we have

$$\mu_C\left([0,u] \times (v,1]\right) = \mu_C([0,t] \times (t,1]).$$

Considering $[0,u] \times (v,1] \subseteq [0,t] \times (t,1]$ the latter, however is equivalent to

$$\mu_C([0,t] \times (t,v]) + \mu_C((u,t] \times (t,1]) = 0,$$

which completes the proof. $\square$

3. Novel results on the Markov product of Bertino copulas

In order to facilitate the proofs of the main results we start with some notation and regularity results. Suppose that $\delta \in \mathcal{D}$ is an arbitrary diagonal. Then by construction there exists some set $\Lambda_\delta = \Lambda \in \mathcal{B}([0,1])$ fulfilling $\lambda(\Lambda) = 1$ such that the derivative $\hat{\delta}'(x)$ exists and $\hat{w}_\delta(x) = \hat{\delta}'(x)$ holds for every $x \in \Lambda$. In what follows we will consider the pairwise disjoint Borel sets

$$\begin{aligned}\hat{\Lambda}^+ &:= \{x \in \Lambda :\, \hat{w}_\delta(x) > 0\}\\ \hat{\Lambda}^- &:= \{x \in \Lambda :\, \hat{w}_\delta(x) < 0\}\\ \hat{\Lambda}^0 &:= \{x \in \Lambda :\, \hat{w}_\delta(x) = 0\},\end{aligned}$$

obviously fulfilling $\hat{\Lambda}^0 \cup \hat{\Lambda}^+ \cup \hat{\Lambda}^- = \Lambda$. Moreover, instead of working with the transformation S according to eq. (11) we will consider the transformation $T : [0,1] \to [0,1]$, defined by

$$T_\delta(x) = T(x) = \begin{cases} u(x) & \text{if } x \in \hat{\Lambda}^+ \\ l(x) & \text{if } x \in \hat{\Lambda}^- \\ x & \text{if } x \in \hat{\Lambda}^0 \cup \Lambda^c. \end{cases} \tag{14}$$

Notice that T is measurable since the functions u, l are measurable and since $\hat{\Lambda}^+, \hat{\Lambda}^-, \hat{\Lambda}^0$ and $\hat{\Lambda}$ are Borel sets. In terms of T, for λ-almost every $x \in [0,1]$ the Markov kernel (10) becomes

$$K_{B_\delta}(x, E) = |\hat{\delta}'(x)| \epsilon_{T(x)}(E) + \Big(1 - |\hat{\delta}'(x)|\Big)\, \epsilon_x(E). \tag{15}$$

It is straightforward to verify (again see [11]) that for $x \in \hat{\Lambda}^+$ we have $T(x) > x$, whereas $T(x) < x$ holds for every $x \in \hat{\Lambda}^-$. Moreover, on the set

$$(\hat{\Lambda}^+ \cap T^{-1}(\hat{\Lambda}^-)) \cup (\hat{\Lambda}^- \cap T^{-1}(\hat{\Lambda}^+)) \cup \hat{\Lambda}^0 \cup \Lambda^c$$

the transformation T is an involution, i.e., $T(T(x)) = x$ holds. More generally, for all $x \in [0,1]$ we have $T(T(x)) \in \{x, T(x)\}$. Considering that in the sequel we will work with expressions of the form $\hat{\delta}'(T(x))$, the following technical lemma about T will be useful since it assures that for λ-almost every $x \in \Lambda$ we also have $T(x) \in \Lambda$, implying that $\hat{\delta}'(T(x))$ is well defined for such x.

Lemma 5. *The transformation T is non-singular, i.e., for every set $N \in \mathcal{B}([0,1])$ with $\lambda(N) = 0$ we have $\lambda(T^{-1}(N)) = 0$. As a direct consequence, $\lambda(\Lambda \cap T^{-1}(\Lambda)) = 1$ holds.*

Proof. Considering $\lambda(\Lambda) = 1$ as well as the decomposition

$$\begin{aligned} T^{-1}(N) \;&=\; \Big(T^{-1}(N) \cap \hat{\Lambda}^+\Big) \cup \Big(T^{-1}(N) \cap \hat{\Lambda}^-\Big) \cup \Big(T^{-1}(N) \cap \hat{\Lambda}^0\Big) \\ &\quad \cup \big(T^{-1}(N) \cap \Lambda^c\big) \end{aligned}$$

of $T^{-1}(N)$ in four pairwise disjoint Borel sets it suffices to show that each of the four sets has λ-measure zero, which can be done as follows.
(i) For every $x \in \hat{\Lambda}^+$ we have $T(x) > x$ as well as $K_{B_\delta}(x, \{T(x)\}) > 0$. Using disintegration and the fact that μ_{B_δ} is doubly stochastic yields

$$\begin{aligned} \int_{T^{-1}(N) \cap \hat{\Lambda}^+} \underbrace{K_{B_\delta}(x, \{T(x)\})}_{>0} \, d\lambda(x) \;&\leq\; \int_{T^{-1}(N) \cap \hat{\Lambda}^+} K_{B_\delta}(x, N) d\lambda(x) \\ &=\; \mu_{B_\delta}\Big((T^{-1}(N) \cap \hat{\Lambda}^+) \times N\Big) \\ &\leq\; \mu_{B_\delta}\left([0,1] \times N\right) = \lambda(N) = 0, \end{aligned}$$

implying $\lambda(T^{-1}(N) \cap \hat{\Lambda}^+) = 0$.
(ii) Proceeding in the same manner yields $\lambda(T^{-1}(N \cap \hat{\Lambda}^-)) = 0$.
(iii) For every $x \in \hat{\Lambda}^0$ by construction we have $T(x) = x$, which implies

$$T^{-1}(N) \cap \hat{\Lambda}^0 = N \cap \hat{\Lambda}^0 \subseteq N,$$

so $\lambda(T^{-1}(N) \cap \hat{\Lambda}^0) = 0$ is trivial.
(iv) Finally, considering

$$\lambda(T^{-1}(N) \cap \Lambda^c) \leq \lambda(\Lambda^c) = 0$$

completes the proof for $\lambda(T^{-1}(N)) = 0$.
The remaining assertion is now a direct consequence of $\lambda(T^{-1}(\Lambda^c)) = 0$. $\square$

We now focus on the main topic of this section, the behavior of the Bertino class under the Markov product, and start by calculating the Markov kernel $K_{B_\delta * B_\gamma}(\cdot,\cdot)$ for two Bertino copulas B_δ and B_γ. Letting $\Lambda_\delta, \Lambda_\gamma$ denote the corresponding sets and T_δ, T_γ the corresponding transformations it follows that the set $G_{\delta,\gamma} := \Lambda_\delta \cap \Lambda_\gamma \cap T_\delta^{-1}(\Lambda_\delta) \cap T_\gamma^{-1}(\Lambda_\gamma))$ fulfills $\lambda(G_{\delta,\gamma}) = 1$. Moreover, for every $x \in G_{\delta,\gamma}$ we get

$$\begin{aligned} K_{B_\delta * B_\gamma}(x,E) &= (K_{B_\delta} \circ K_{B_\gamma})(x,E) = \int_{[0,1]} K_{B_\gamma}(z,E) K_{B_\delta}(x,dz) \\ &= |\hat{\delta}'(x)| K_{B_\gamma}(T_\delta(x),E) \;+\; (1-|\hat{\delta}'(x)|) K_{B_\gamma}(x,E) \\ &= |\hat{\delta}'(x)| \left(|\hat{\gamma}'(T_\delta(x))| \epsilon_{T_\gamma \circ T_\delta(x)}(E) \;+\; (1-|\hat{\gamma}'(T_\delta(x))|) \epsilon_{T_\delta(x)}(E) \right) \\ &\quad + (1-|\hat{\delta}'(x)|) \left(|\hat{\gamma}'(x)| \epsilon_{T_\gamma(x)}(E) + (1-|\hat{\gamma}'(x)|) \epsilon_x(E) \right). \end{aligned} \tag{16}$$

The following example shows that the Bertino class is not closed w.r.t. the Markov product, i.e., there are Bertino copulas B_δ, B_γ fulfilling that $B_\delta * B_\gamma$ is not a Bertino copula.

Example 6. We consider $\delta, \gamma \in \mathcal{D}$, defined by

$$\begin{aligned} \hat{\delta}(x) &:= \tfrac{x}{2} \mathbf{1}_{[0,\frac{1}{2}]}(x) + \tfrac{1-x}{2} \mathbf{1}_{(\frac{1}{2},1]}(x), \\ \hat{\gamma}(x) &:= \tfrac{3x}{7} \mathbf{1}_{[0,\frac{1}{3}]}(x) + \tfrac{3}{14}(1-x) \mathbf{1}_{(\frac{1}{3},1]}(x). \end{aligned}$$

Then $\hat{\delta}'(x) \neq 0$ as well as $\hat{\gamma}'(x) \neq 0$ for all but finitely many points $x \in [0,1]$. The corresponding transformations are given by $T_\delta(x) = 1-x$ as well as

$$T_\gamma(x) = (1-2x) \mathbf{1}_{[0,\frac{1}{3}]}(x) + \tfrac{1-x}{2} \mathbf{1}_{(\frac{1}{3},1]}(x).$$

Notice that B_δ coincides with $\frac{1}{2}(M+W)$, so it is a convex combination of M and W. It is therefore easy to verify that B_δ is idempotent (also see Theorem 15). For all but at most finitely many $x \in [0,1]$ we have that $x, T_\delta(x), T_\gamma(x), T_\gamma \circ T_\delta(x)$ are pairwise disjoint, which, using eq. (16) shows that for all but finitely many x the Markov kernel $K_{B_\delta * B_\gamma}(x,\cdot)$ has four

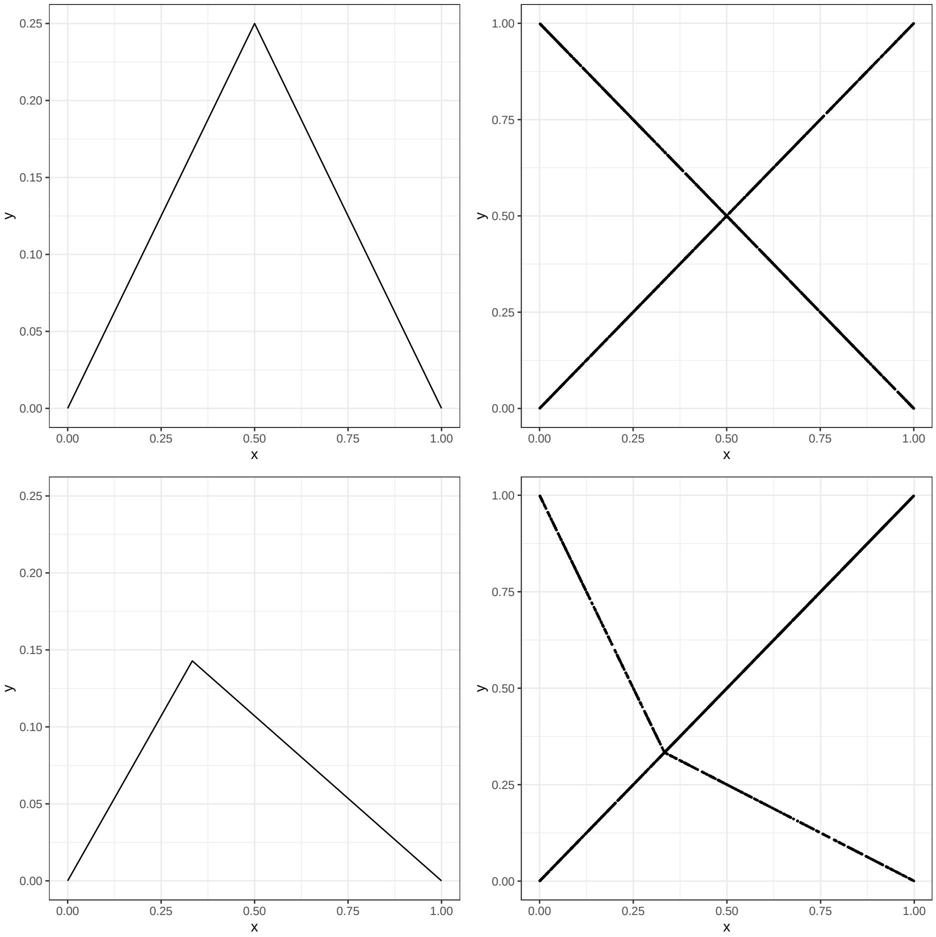

Figure 1: The functions $\hat{\delta}$ and $\hat{\gamma}$ (left panels) and samples of the corresponding Bertino copulas B_δ and B_γ, respectively (right panels) as considered in Example 6.

distinct point masses, implying that $B_\delta * B_\gamma$ can not be a Bertino copula. Again using eq. (16) it follows that the support of $B_\delta * B_\gamma$ is given by

$$\mathrm{supp}(B_\delta * B_\gamma) = \Delta \cup \Gamma(T_\delta) \cup \Gamma(T_\gamma) \cup \Gamma(T_\gamma \circ T_\delta).$$

Figure 1 depicts the functions $\hat{\delta}$ and $\hat{\gamma}$ and samples of the corresponding Bertino copulas, Figure 2 contains a sample of size $n = 2.000$ of the copula $B_\delta * B_\gamma$.

The situation changes, if we consider $B_\delta^{*2} := B_\delta * B_\delta$, i.e., if we consider

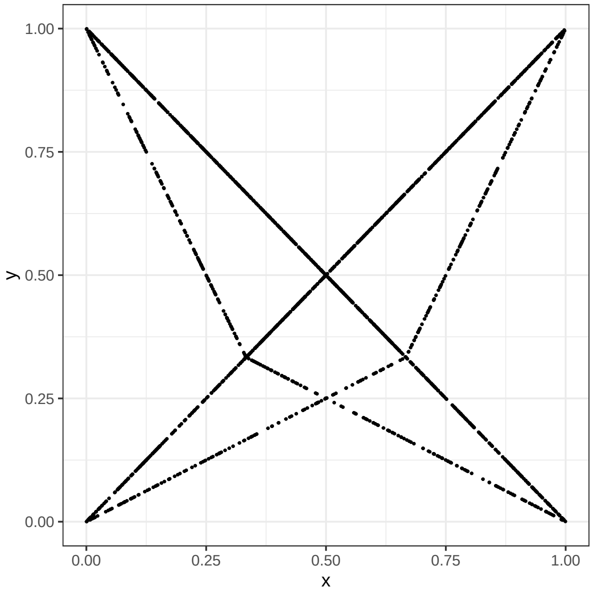

Figure 2: Sample of size $n = 2.000$ of the copula $B_\delta * B_\gamma$ considered in Example 6. The four graphs are clearly visible.

the Markov product of a Bertino copula with itself. In this case the Markov kernel (16) for $x \in \Lambda \cap T^{-1}(\Lambda)$ boils down to

$$\begin{aligned} K_{B_\delta^{*2}}(x,E) &= |\hat{\delta}'(x)| \left(2 - |\hat{\delta}'(T(x))| - |\hat{\delta}'(x)|\right) \epsilon_{T(x)}(E) \\ &\quad + \left(|\hat{\delta}'(x)||\hat{\delta}'(T(x))| + (1 - |\hat{\delta}'(x)|)^2\right) \epsilon_x(E). \end{aligned} \tag{17}$$

Notice that the reason for $\epsilon_{T(T(x))}$ not appearing in eq. (17) is that for $x \in (\hat{\Lambda}^+ \cup \hat{\Lambda}^-) \cap T^{-1}(\hat{\Lambda}^+ \cup \hat{\Lambda}^-)$ we have $T(T(x)) = x$ and for the case $x \in (\hat{\Lambda}^+ \cup \hat{\Lambda}^-) \cap T^{-1}(\hat{\Lambda}^0)$ we have $T(T(x)) = T(x)$. Obviously the copula B_δ^{*2} concentrates its mass on the set $\Delta \cup \Gamma(T)$, analogously to B_δ.

The following two examples insinuate, which behavior the sequence of Markov product iterates $B_\delta, B_\delta^{*2} = B_\delta * B_\delta, B_\delta^{*3} = B_\delta * B_\delta * B_\delta, \ldots$ might exhibit.

Example 7. Consider the diagonal $\delta \in \mathcal{D}$ whose corresponding function $\hat{\delta}$ is given by

$$\hat{\delta}(x) = x\, \mathbf{1}_{[0,\frac{1}{2}]}(x) + (1-x)\, \mathbf{1}_{(\frac{1}{2},1]}(x).$$

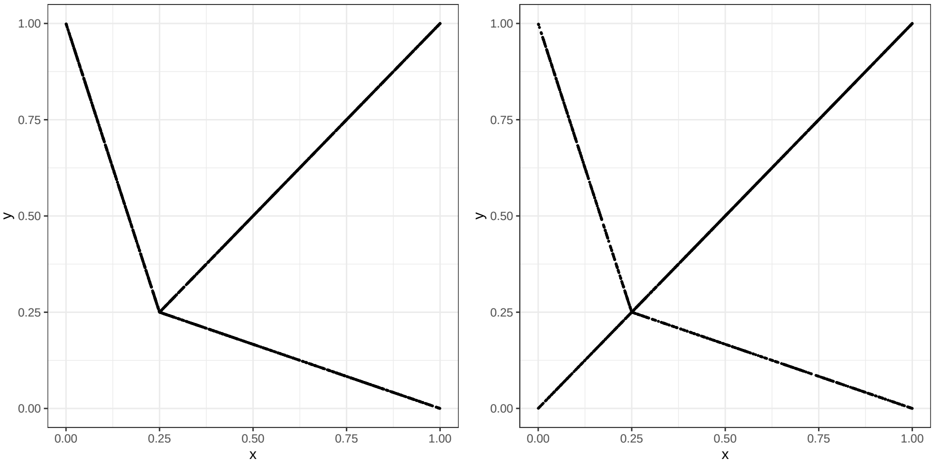

Figure 3: Sample of size $n = 2.000$ of the copula B_δ (left panel) and the copula $B_\delta * B_\delta$ (right panel) as considered in Example 8.

Then a straightforward calculation yields that the corresponding Bertino copula B_δ coincides with the lower Fréchet-Hoeffding bound W. As a direct consequence we have that $B_\delta^{*n} = M$ for $n \in \mathbb{N}$ even and $B_\delta^{*n} = W$ for $n \in \mathbb{N}$ odd. In other words, each iterate B_δ^{*n} is again a Bertino copula, but the sequence $(B_\delta^{*n})_{n\in\mathbb{N}}$ has period 2 and does not converge.

The next example shows that $(B_\delta^{*n})_{n\in\mathbb{N}}$ may indeed converge w.r.t. d_∞ to a limit which is itself an idempotent Bertino copula. The limit result is established in two ways, both of which we will also use in the general setting for proving our main results later.

Example 8. Consider the diagonal $\delta \in \mathcal{D}$ whose corresponding function $\hat{\delta}$ is given by

$$\hat{\delta}(x) = x\,\mathbf{1}_{[0,\frac{1}{4}]}(x) + \tfrac{1-x}{3}\,\mathbf{1}_{(\frac{1}{4},1]}(x).$$

Then we obviously have $\hat{\Lambda}^+ = (0, \frac{1}{4}), \hat{\Lambda}^- = (\frac{1}{4}, 1)$ and it is straightforward to verify that the corresponding transformation $T = T_\delta$ fulfills

$$T(x) = (1-3x)\,\mathbf{1}_{[0,\frac{1}{4}]}(x) + \tfrac{1-x}{3}\,\mathbf{1}_{(\frac{1}{4},1]}(x)$$

as well as $T(\hat{\Lambda}^+) = \hat{\Lambda}^-$ and vice versa. A version of the Markov kernel for $x \in \Lambda$ is given by

$$K_{B_\delta}(x, E) = \begin{cases} \epsilon_{T(x)}(E) & \text{if } x \in [0, \frac{1}{4}] \\ \frac{1}{3}\epsilon_{T(x)}(E) + \frac{2}{3}\epsilon_x(E) & \text{if } x \in (\frac{1}{4}, 1]. \end{cases}$$

Applying eq. (17) yields that the Markov Kernel $K_{B_\delta * B_\delta}(x,\cdot)$ is of the form

$$K_{B_\delta * B_\delta}(x,E) = \begin{cases} \frac{2}{3}\epsilon_{T(x)}(E) + \frac{1}{3}\epsilon_x(E) & \text{if } x \in [0,\frac{1}{4}] \\ \frac{2}{9}\epsilon_{T(x)}(E) + \frac{7}{9}\epsilon_x(E) & \text{if } x \in (\frac{1}{4},1]. \end{cases} \tag{18}$$

Obviously the diagonal γ of $B_\delta * B_\delta$ fulfills

$$\gamma(x) = (B_\delta * B_\delta)(x,x) = \begin{cases} \frac{x}{3} & \text{if } x \in [0,\frac{1}{4}] \\ \frac{11x-2}{9} & \text{if } x \in (\frac{1}{4},1], \end{cases}$$

so

$$\hat{\gamma}(x) = x - \gamma(x) = \begin{cases} \frac{2x}{3} & \text{if } x \in [0,\frac{1}{4}] \\ \frac{2-2x}{9} & \text{if } x \in (\frac{1}{4},1]. \end{cases}$$

Considering that this implies $\hat{\gamma} = \frac{2}{3}\hat{\delta}$ it follows immediately that for the Bertino copula B_γ the transformation T_γ coincides with T_δ. Moreover, for λ-almost every $x \in [0,1]$ we have $K_{B_\gamma}(x,\cdot) = K_{B_\delta * B_\delta}(x,\cdot)$, which implies that $B_\delta^{*2} = B_\gamma$ is again an element of the Bertino class. Figure 3 depicts samples of B_δ and B_δ^{*2}.

In order to tackle higher order iterates $B_\delta^{*n} = B_\delta * B_\delta * \cdots * B_\delta$ (n times) we return to B_δ^{*2} and view the composition of the Markov kernels in a more structured way: For $x \in (0,\frac{1}{4})$ we have $T(x) > x$ as well as $T(T(x)) = x$. As a consequence, the transitions corresponding to $K_{B_\delta}(x,\cdot)$ for such x can equivalently be viewed as a discrete Markov chain with state space $\mathcal{S} = \{x, T(x)\}$ and transition matrix π, given by (see, e.g., [13, 21] for background on Markov chains)

$$\pi = \begin{pmatrix} 0 & 1 \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix} = \begin{pmatrix} K_{B_\delta}(x,\{x\}) & K_{B_\delta}(x,\{T(x)\}) \\ K_{B_\delta}(T(x),\{x\}) & K_{B_\delta}(T(x),\{T(x)\}) \end{pmatrix}.$$

Notice that the matrix π also contains all relevant transition probabilities for points $x \in (\frac{1}{4},1)$ since we have $T(\hat{\Lambda}^+) = \hat{\Lambda}^-$ as well as $T(\hat{\Lambda}^-) = \hat{\Lambda}^+$. As a direct and handy consequence, higher order iterates

$$K_{B_\delta^{*n}}(x,\cdot) = \left(K_{B_\delta} \circ K_{B_\delta} \circ \cdots \circ K_{B_\delta}\right)(x,\cdot)$$

can simply be calculated via the n-step transition matrix π^n. For $n = 2$ we get (compare with eq. (18))

$$\pi^2 = \begin{pmatrix} 0 & 1 \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{2}{9} & \frac{7}{9} \end{pmatrix}.$$

For determining π^n we can use the eigendecomposition (diagonalization) $\pi = UDU^{-1}$ of the matrix π, given by

$$\pi = \begin{pmatrix} -3 & 1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} -\frac{1}{3} & 0 \\ 0 & 1 \end{pmatrix} \cdot \frac{1}{4} \begin{pmatrix} -1 & 1 \\ 1 & 3 \end{pmatrix}$$

and conclude that

$$\pi^n = \begin{pmatrix} -3 & 1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} \left(-\frac{1}{3}\right)^n & 0 \\ 0 & 1 \end{pmatrix} \cdot \frac{1}{4} \begin{pmatrix} -1 & 1 \\ 1 & 3 \end{pmatrix}.$$

Considering $n \to \infty$ this directly yields

$$\lim_{n\to\infty} \pi^n = \begin{pmatrix} \frac{1}{4} & \frac{3}{4} \\ \frac{1}{4} & \frac{3}{4} \end{pmatrix} =: \overline{\pi}. \tag{19}$$

In other words, for every $x \in (0, \frac{1}{4})$ we have that the sequence $(K_{B_\delta^{*n}}(x, \cdot))_{n\in\mathbb{N}}$ of probability measures converges in total variation on $\mathcal{B}([0,1])$ to the probability measure $K(x,E) := \frac{3}{4}\epsilon_{T(x)}(E) + \frac{1}{4}\epsilon_x(E)$. Analogously, for $x \in (\frac{1}{4}, 1)$ the sequence $(K_{B_\delta^{*n}}(x, \cdot))_{n\in\mathbb{N}}$ converges in total variation to $K(x,E) := \frac{1}{4}\epsilon_{T(x)}(E) + \frac{3}{4}\epsilon_x(E)$. The mapping $K : [0,1] \times \mathcal{B}([0,1]) \to [0,1]$, given by

$$K(x,E) = \begin{cases} \frac{3}{4}\epsilon_{T(x)}(E) + \frac{1}{4}\epsilon_x(E) & \text{if } x \in [0, \frac{1}{4}] \\ \frac{1}{4}\epsilon_{T(x)}(E) + \frac{3}{4}\epsilon_x(E) & \text{if } x \in (\frac{1}{4}, 1]. \end{cases} \tag{20}$$

obviously is a Markov kernel fulfilling eq. (3), i.e., $K(\cdot,\cdot)$ is the Markov kernel of a unique copula C and we may write $K_C(\cdot,\cdot)$ instead of $K(\cdot,\cdot)$. Letting η denote the diagonal of C we have

$$\hat{\eta}(x) = x - \eta(x) = \begin{cases} \frac{3x}{4} & \text{if } x \in [0, \frac{1}{4}] \\ \frac{1-x}{4} & \text{if } x \in (\frac{1}{4}, 1], \end{cases}$$

implying $\hat{\eta} = \frac{3}{4}\hat{\delta}$ and it follows that C is again an element of the Bertino family. Moreover, considering $\overline{\pi}^2 = \overline{\pi}$ (or using eq. (17)) it follows that C is idempotent. Finally, using the fact that convergence in total variation implies that for λ-almost every $x \in [0,1]$ the sequence of conditional distribution functions $(F^x_{B_\delta^{*n}})_{n\in\mathbb{N}}$ converges pointwise to F^C_x, i.e., for every $y \in [0,1]$ we have

$$\lim_{n\to\infty} F^x_{B_\delta^{*n}}(y) = \lim_{n\to\infty} K_{B_\delta^{*n}}(x, [0,y]) = K_C(x,[0,y]) = F^C_x(y).$$

Having this, applying disintegration and dominated convergence directly yields

$$\begin{aligned}\lim_{n\to\infty} B_\delta^{*n}(x,y) &= \lim_{n\to\infty}\int_{[0,x]} K_{B_\delta^{*n}}(s,[0,y])\,d\lambda(s) = \int_{[0,x]} K_C(s,[0,y])\,d\lambda(s)\\ &= C(x,y).\end{aligned}$$

Summing up, we have shown that the sequence $(B_\delta^{*n})_{n\in\mathbb{N}}$ converges uniformly to an idempotent Bertino copula.
Instead of working with the eigendecomposition of π we can also directly tackle the limit eq. (19) using asymptotic results from the theory of Markov chains. In fact, our transition matrix π is easily seen to be irreducible and aperiodic (all entries of π^2 are positive) so (see [13, 21]) the limit $\overline{\pi}$ of $(\pi^n)_{n\in\mathbb{N}}$ is fully determined by the (unique) invariant distribution κ of π in the sense that

$$\begin{aligned}\overline{\pi} &= \lim_{n\to\infty}\pi^n = \begin{pmatrix} \kappa(\{x\}) & \kappa(\{T(x)\}) \\ \kappa(\{x\}) & \kappa(\{T(x)\}) \end{pmatrix}\\ &= \begin{pmatrix} K_C(x,\{x\}) & K_C(x,\{T(x)\}) \\ K_C(T(x),\{x\}) & K_C(T(x),\{T(x)\}) \end{pmatrix}\end{aligned}$$

holds. For determining κ we need to solve the fixed point equation

$$\big(\kappa(\{x\}),\kappa(\{T(x)\}\big)\cdot\begin{pmatrix} 0 & 1 \\ \frac{1}{3} & \frac{2}{3}\end{pmatrix} = \big(\kappa(\{x\}),\kappa(\{T(x)\}\big),$$

which directly yields $\kappa(\{x\}) = \frac{1}{4}$ and $\kappa(\{T(x)\}) = \frac{3}{4}$.

Motivated by the above examples we now tackle the first main result of this paper.

Theorem 9. *For every Bertino copula B_δ the Markov product $B_\delta * B_\delta$ is a Bertino copula too.*

Proof. We will use Theorem 3 and proceed as follows. Considering $B_\delta^t = B_\delta$ we have $(B_\delta^{*2})^t = (B_\delta)^t * (B_\delta)^t = B_\delta * B_\delta = B_\delta^{*2}$, so B_δ^{*2} is symmetric too. Moreover we already know that

$$\mu_{B_\delta^{*2}}\left(\Delta \cup \Gamma(T)\right) = 1$$

holds. Applying Theorem 3 it therefore suffices to show that for arbitrary $0 \le u \le v \le 1$ there exists some $t \in [u,v]$ with

$$\mu_{B_\delta^{*2}}\left([0,t]\times(t,v]\right) = 0 = \mu_{B_\delta^{*2}}\left([u,t]\times(t,1]\right).$$

Using disintegration the latter identity further boils down to

$$\begin{aligned}\int_{[0,t]\cap T^{-1}((t,v])} K_{B_\delta^{*2}}(s,\{T(s)\})d\lambda(s) &= \mu_{B_\delta^{*2}}([0,t]\times(t,v]) = 0 \\ \int_{[u,t]\cap T^{-1}((t,1])} K_{B_\delta^{*2}}(s,\{T(s)\})d\lambda(s) &= \mu_{B_\delta^{*2}}([u,t]\times(t,1]) = 0. \end{aligned} \tag{21}$$

Let $0 \le u \le v \le 1$ by arbitrary but fixed and choose $t \in [u,v]$ so that eq. (13) holds for C replaced by B_δ. We want to show that for this very t also eq. (21) is fulfilled. Notice that according to eq. (17) for $x \in \hat{\Lambda}^+ \cup \hat{\Lambda}^-$ we can only have $K_{B_\delta^{*2}}(x,\{T(x)\}) > 0$ if $|\hat{\delta}'(x)| > 0$, i.e., if $K_{B_\delta}(x,\{T(x)\}) > 0$ holds. Therefore,

$$\int_{[0,t]\cap T^{-1}((t,v])} K_{B_\delta^{*2}}(s,\{T(s)\})d\lambda(s) > 0$$

would directly yield

$$\int_{[0,t]\cap T^{-1}((t,v])} K_{B_\delta}(s,\{T(s)\})d\lambda(s) > 0,$$

a contradiction. This shows the first line in eq. (21), the second one follows analogously. Since $0 \le u \le v \le 1$ was arbitrary this completes the proof. □

Remark 10. It might seem natural to tackle the proof of Theorem 9 without Theorem 3 by simply calculating the diagonal δ_2 of B_δ^{*2} and then showing that $B_{\delta_2} = B_\delta^{*2}$ holds. Using disintegration it can be shown that the diagonal δ_2 is given by

$$\hat{\delta}_2(x) = \int_{[0,x]\cap T^{-1}((x,1])} \hat{\delta}'(s)\left(2 + \hat{\delta}'(T(s)) - \hat{\delta}'(s)\right) d\lambda(s).$$

Calculating the transformation T_2 corresponding to the Bertino copula B_{δ_2} and showing that it coincides with T, however, seems far from straightforward, which is why we tackled the proof via Theorem 3.

Extending Theorem 9 to higher order iterates B_δ^{*n}, for $x \in \Lambda \cap T^{-1}(\Lambda)$ we will write

$$\begin{aligned} \sigma_\delta(x) &:= |\hat{\delta}'(x)| + |\hat{\delta}'(T(x))| \\ r_n(x) &:= 1 - (1-\sigma_\delta(x))^n. \end{aligned} \tag{22}$$

Theorem 11. *For every Bertino copula B_δ and every $n \in \mathbb{N}$ the copula B_δ^{*n} is a Bertino copula. Moreover, the Markov kernel of B_δ^{*n} fulfills $K_{B_\delta^{*n}}(x, \{x\}) = 1$ for $x \in \hat{\Lambda}^0$, as well as*

$$K_{B_\delta^{*n}}(x, E) = \frac{|\hat{\delta}'(x)|}{\sigma_\delta(x)} r_n(x) \epsilon_{T(x)}(E) + \left\{ 1 - \frac{|\hat{\delta}'(x)|}{\sigma_\delta(x)} r_n(x) \right\} \epsilon_x(E) \tag{23}$$

for $x \in (\hat{\Lambda}^+ \cup \hat{\Lambda}^-) \cap T^{-1}(\Lambda)$.

Proof. We first calculate the Markov kernel of B_δ^{*n} and then proceed analogous to the proof of Theorem 9. Since the case $x \in \hat{\Lambda}^0$ is trivial it suffices to consider $x \in (\hat{\Lambda}^+ \cup \hat{\Lambda}^-) \cap T^{-1}(\Lambda)$ for which we obviously have $\sigma_\delta(x) > 0$. Calculating the first few iterates (see eq. (17) for $n = 2$) and using induction on n it follows that the sequence of Markov kernels has the structure

$$K_{B_\delta^{*n}}(x, E) = f_n(x) \epsilon_{T(x)}(E) + (1 - f_n(x)) \epsilon_x(E), \tag{24}$$

where $f_n(x) := f_{n-1}(x) + f_1(x)\{1 - f_{n-1}(x) - f_{n-1}(T(x))\}$, with $f_0(x) := 0$ and $f_1(x) := |\hat{\delta}'(x)|$. For $(x, T(x)) \in (\hat{\Lambda}^+ \cup \hat{\Lambda}^-) \times \hat{\Lambda}^0$ we have $T(T(x)) = T(x)$ so $f_n(T(x)) = 0$ for all $n \in \mathbb{N}$. Moreover, recall that for $(x, T(x)) \in (\hat{\Lambda}^+ \times \hat{\Lambda}^-) \cup (\hat{\Lambda}^- \times \hat{\Lambda}^+)$ we have $T(T(x)) = x$. Hence, through summation, setting $r_n(x) := f_n(x) + f_n(T(x))$, for each $x \in (\hat{\Lambda}^+ \cup \hat{\Lambda}^-) \cap T^{-1}(\Lambda)$ the recursion

$$1 - r_n(x) = (1 - r_1(x))(1 - r_{n-1}(x))$$

follows (which equals eq. (22) in non-recursive form). Accordingly, $f_n(x) = f_{n-1}(x) + f_1(x)(1 - \sigma_\delta(x))^{n-1}$ and this last recursion, in turn, is easily seen to admit the representation $f_n(x) = f_1(x) \sum_{j=0}^{n-1} (1 - \sigma_\delta(x))^j$, so that eq. (24) is equivalent to eq. (23).
According to eq. (23), the support of the Markov kernel $K_{B_\delta^{*n}}(x, \cdot)$ is contained in $\{x, T(x)\}$ and it remains to verify that B_δ^{*n} is indeed a Bertino copula. For this, however, we may use symmetry of B_δ^{*n} and proceed analogous to the proof of Theorem 9 (just replacing B_δ^{*2} by B_δ^{*n}). □

The following theorem focusing on convergence of the iterates $(B_\delta^{*n})_{n \in \mathbb{N}}$ is the second main result of this contribution.

Theorem 12. *Suppose that B_δ is an arbitrary Bertino copula fulfilling the condition that $(\hat{\delta}'(x), \hat{\delta}'(T(x))) = (1, -1)$ only holds for a set of λ-measure 0. Then there exists some idempotent copula $C \in \mathcal{C}$ such that the sequence $(B_\delta^{*n})_{n \in \mathbb{N}}$ of Markov product iterates of B_δ converges weakly conditional, hence uniformly, to C.*

Proof. Suppose that B_δ is an arbitrary Bertino copula with corresponding sets $\hat{\Lambda}^+, \hat{\Lambda}^-, \hat{\Lambda}^0$ and transformation $T = T_\delta$. Without loss of generality we may assume that both $T(\hat{\Lambda}^+) \subseteq \hat{\Lambda}^- \cup \hat{\Lambda}^0$ as well as $T(\hat{\Lambda}^-) \subseteq \hat{\Lambda}^+ \cup \hat{\Lambda}^0$ holds (otherwise we may use Lemma 5 and redefine the sets $\hat{\Lambda}^+, \hat{\Lambda}^-, \hat{\Lambda}^0$ to assure this property, or, alternatively argue for λ-almost every $x \in \hat{\Lambda}^+, x \in \hat{\Lambda}^-$, etc.). We need to show the existence of some copula C such that for λ-almost every $x \in [0,1]$ the sequence $(K_{B_\delta^{*n}}(x,\cdot))_{n\in\mathbb{N}}$ converges weakly to $K_C(x,\cdot)$ and do so again using results from Markov chains. We distinguish the following five situations.

(i) For λ-almost every $x \in \hat{\Lambda}^0 \cup \Lambda^c$ we have $K_{B_\delta}(x,\{x\}) = 1$, implying $K_{B_\delta^{*n}}(x,\{x\}) = 1$. In other words, the sequence $K_{B_\delta^{*n}}(x,\cdot)$ is constant (hence convergent in total variation and weakly conditionally convergent to ϵ_x).

(ii) If $x \in \hat{\Lambda}^+$ with $T(x) \in \hat{\Lambda}^-$ we have $T(x) > x$ as well $\hat{\delta}'(T(x)) < 0$ and may proceed as follows: Consider the discrete Markov chain with state space $\mathcal{S} = \{x, T(x)\}$ and transition matrix π_x, given by

$$\pi_x = \begin{pmatrix} 1-\hat{\delta}'(x) & \hat{\delta}'(x) \\ -\hat{\delta}'(T(x)) & 1+\hat{\delta}'(T(x)) \end{pmatrix}. \tag{25}$$

(a) If $\hat{\delta}'(x) < 1$ then π_x^2 (for $\hat{\delta}'(T(x)) > -1$ even π_x) only contains positive entries, implying that π_x is irreducible and aperiodic and that the limit $\overline{\pi}_x$ of π_x^n is fully determined by the (unique) invariant distribution κ of π_x, i.e.,

$$\overline{\pi}_x = \lim_{n\to\infty} \pi_x^n = \begin{pmatrix} \kappa(\{x\}) & \kappa(\{T(x)\}) \\ \kappa(\{x\}) & \kappa(\{T(x)\}) \end{pmatrix}$$

holds. For determining κ we need to solve the equation

$$\big(\kappa(\{x\}), \kappa(\{T(x)\}\big) \cdot \begin{pmatrix} 1-\hat{\delta}'(x) & \hat{\delta}'(x) \\ -\hat{\delta}'(T(x)) & 1+\hat{\delta}'(T(x)) \end{pmatrix} = \big(\kappa(\{x\}), \kappa(\{T(x)\}\big),$$

within the family of all probability measures on $\mathcal{S}$, which directly yields

$$\big(\kappa(\{x\}), \kappa(\{T(x)\}\big) = \frac{1}{\hat{\delta}'(x) - \hat{\delta}'(T(x))}\big(-\hat{\delta}'(T(x)), \hat{\delta}'(x)\big). \tag{26}$$

This shows that the sequence $(K_{B_\delta^{*n}}(x,\cdot))_{n\in\mathbb{N}}$ converges weakly (in fact even in total variation) to the probability measure $\kappa(\{x\})\epsilon_x + \kappa(\{T(x)\})\epsilon_{T(x)}$ on $\mathcal{B}([0,1])$.

(b) If $\hat{\delta}'(x) = 1$ and $\hat{\delta}'(T(x)) \in (-1,0)$ then the transition matrix π_x according to eq. (25) has the property, that for all $l \geq 2$ all entries of π_x^l are positive, so we may proceed as in (a) to conclude that also in this case the

sequence $(K_{B_\delta^{*n}}(x,\cdot))_{n\in\mathbb{N}}$ converges weakly to $\kappa(\{x\})\epsilon_x + \kappa(\{T(x)\})\epsilon_{T(x)}$.
(iii) For $x \in \hat\Lambda^+$ with $T(x) \in \hat\Lambda^0$ we have $T(x) > x$ as well $\hat\delta'(T(x)) = 0$ and the corresponding Markov chain with state space $\mathcal{S} = \{x, T(x)\}$ has transition matrix

$$\pi_x = \begin{pmatrix} 1-\hat\delta'(x) & \hat\delta'(x) \\ 0 & 1 \end{pmatrix}. \tag{27}$$

Contrary to (ii) π_x is not irreducible, convergence of $(\pi_x^n)_{n\in\mathbb{N}}$ can nevertheless be shown as follows: (a) if $\hat\delta'(x) < 1$ then a straightforward calculation yields

$$\overline{\pi}_x = \lim_{n\to\infty} \pi_x^n = \begin{pmatrix} (1-\hat\delta'(x))^n & 1-(1-\hat\delta'(x))^n \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix},$$

implying that $(K_{B_\delta^{*n}}(x,\cdot))_{n\in\mathbb{N}}$ converges in total variation to $\epsilon_{T(x)}$.
(b) If $\hat\delta'(x) = 1$ then π_x even fulfills $\pi_x^n = \pi_x$ and we may infer the same conclusion as in (a).
(iv) Suppose that $x \in \hat\Lambda^-$ and that $T(x) \in \hat\Lambda^+$ holds. Then we may either use the property $T(T(x)) = x$ or argue analogously to case (ii) as follows: the transition matrix π_x now takes the form

$$\pi_x = \begin{pmatrix} 1+\hat\delta'(x) & -\hat\delta'(x) \\ \hat\delta'(T(x)) & 1-\hat\delta'(T(x)) \end{pmatrix}.$$

Distinguishing the situations $\hat\delta'(x) > -1$ and $\hat\delta'(x) = -1$ yields that the invariant distribution κ on $\mathcal{S}$ is now given by

$$\big(\kappa(\{x\}), \kappa(\{T(x)\})\big) = \frac{1}{\hat\delta'(T(x)) - \hat\delta'(x)} \big(\hat\delta'(T(x)), -\hat\delta'(x)\big), \tag{28}$$

which implies that the sequence $(K_{B_\delta^{*n}}(x,\cdot))_{n\in\mathbb{N}}$ converges in total variation and hence weakly to the probability measure $\kappa(\{x\})\epsilon_x + \kappa(\{T(x)\})\epsilon_{T(x)}$.
(v) The case $x \in \hat\Lambda^-$ with $T(x) \in \hat\Lambda^0$ can be handled analogously to (iii).
It remains to show the existence of some idempotent $C \in \mathcal{C}$ fulfilling that its Markov kernel $K_C(\cdot,\cdot)$ coincides with limit of the sequences $(K_{B_\delta^{*n}}(x,\cdot))_{n\in\mathbb{N}}$ form before. Defining, however, the probability measure $K(x,\cdot)$ as the weak limit of the sequence $(K_{B_\delta^{*n}}(x,\cdot))_{n\in\mathbb{N}}$ for every $x \in [0,1]$ it follows that $K(x,\cdot)$ is a probability measure and that $x \mapsto K(x,E)$ is measurable for every $E \in \mathcal{B}([0,1])$, i.e., $K : [0,1] \times \mathcal{B}([0,1])$ is a Markov kernel. For every $(x,y) \in [0,1]^2$ define $C(x,y)$ by

$$C(x,y) := \int_{[0,x]} K(t,[0,y])\, d\lambda(t).$$

Then using dominated convergence it is straightforward to verify that C is indeed a copula with $K = K_C$ being it's Markov kernel. Finally, idempotence of C is a direct consequence of idempotence of the identity matrix and the limit transition matrices $\overline{\pi}_x$ (in all five cases) studied above. □

Combining Theorem 11 and Theorem 12 directly yields the following result:

Theorem 13. *Suppose that B_δ is an arbitrary Bertino copula fulfilling the condition that $(\hat{\delta}'(x), \hat{\delta}'(T(x))) = (1, -1)$ only holds for a set of λ-measure 0. Then B_δ^{*n} is a Bertino copula for every $n \in \mathbb{N}$ and the sequence $(B_\delta^{*n})_{n\in\mathbb{N}}$ of Markov product iterates of B_δ converges weakly conditional, hence uniformly, to an idempotent Bertino copula C.*

Proof. We already know that under the given assumptions each B_δ^{*n} is a Bertino copula and that the limit copula C is idempotent. It hence suffices to show that family $\mathcal{C}_{Ber}$ of all Bertino copulas is closed in $(\mathcal{C}, d_\infty)$, which can be done as follows: Using Lipschitz continuity, the Arzela-Ascoli theorem (see [22]) implies that the family $\mathcal{D}$ is compact with respect to the uniform norm $\| \cdot \|_\infty$ on $[0, 1]$. Let $\iota : (\mathcal{D}, \|) \cdot \|_\infty \to (\mathcal{C}_{Ber}, d_\infty)$ denote the mapping assigning each diagonal δ the corresponding Bertino copula B_δ. We want to show that ι is a homeomorphism. Considering $B_\delta \in \mathcal{C}_\delta$ obviously ι is bijective. For arbitrary diagonals $\delta_1, \delta_2 \in \mathcal{D}$ and $x \le y$ we have

$$\Big| \min\left\{\hat{\delta}_1(t) : t \in [x, y]\right\} - \min\left\{\hat{\delta}_2(t) : t \in [x, y]\right\} \Big| \le \|\delta_1 - \delta_2\|_\infty,$$

which yields Lipschitz continuity of ι, i.e., $d_\infty(B_{\delta_1}, B_{\delta_2}) \le \|\delta_1 - \delta_2\|_\infty$ holds. Compactness of $(\mathcal{D}, \| \cdot \|_\infty)$ yields continuity of ι^{-1}, so ι is a homeomorphism. Compactness of the Bertino family now is a direct consequence of the fact that continuity preserves compactness. □

The following theorem shows that the condition that $(\hat{\delta}'(x), \hat{\delta}'(T(x))) = (1, -1)$ only holds for a set of λ-measure 0 can not be avoided in order to get convergence, i.e., Theorem 13 is best possible.

Theorem 14. *Suppose that B_δ is a Bertino for which the set*

$$E := \{x \in [0, 1] : (\hat{\delta}'(x), \hat{\delta}'(T(x))) = (1, -1)\} \in \mathcal{B}([0, 1])$$

has positive λ-measure. Then the sequence $(B_\delta^{*n})_{n\in\mathbb{N}}$ of Markov product iterates of B_δ does not even converge w.r.t. d_∞.

Proof. First of all notice that for every $x \in E$ we have that the corresponding transition matrix π_x on $\mathcal{S} = \{x, T(x)\}$ is given by

$$\pi_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Obviously the sequence $(\pi_x^n)_{n\in\mathbb{N}}$ has period 2, implying that the sequence $(K_{B_\delta^{*n}}(x,\cdot))_{n\in\mathbb{N}}$ does not converge, neither in total variation nor weakly. This already yields that $(B_\delta^{*n})_{n\in\mathbb{N}}$ does not converge in the weak conditional sense. We want to show, however, that we do not even have convergence w.r.t. d_∞. Notice that (see [23]) there are examples of copulas $A, A_1, A_2, \ldots$ such that $\lim_{n\to\infty} d_\infty(A_n, A) = 0$ although for λ-almost every $x \in \mathbb{N}$ the probability measures $K_{A_n}(x,\cdot)$ do not converge to $K_A(x,\cdot)$ weakly. For Bertino copulas, however, the situation is different.
Suppose now that $\lambda(E) > 0$ holds. Considering that $\mathbf{1}_E$ as well as T are integrable we have that λ-almost every $x \in [0,1]$ is a Lebesgue point of both functions (see [22]). Let $x_0 \in E$ be a Lebesgue point of both functions and set

$$r := \left(\frac{T(x_0) - x_0}{8}\right)^2 \in \left(0, \frac{1}{64}\right).$$

By construction there exists some $h_0 \in (0, r)$ such that for every $h \in (0, h_0]$ we have

$$\begin{aligned} \frac{1}{2h}\int_{[x_0-h,x_0+h]} |T(s) - T(x_0)|\, d\lambda(s) &< r \\ \underbrace{\frac{1}{2h}\int_{[x_0-h,x_0+h]} |\mathbf{1}_E(s) - 1|\, d\lambda(s)}_{=\frac{1}{2h}\lambda(\{s\in[x_0-h,x_0+h]:\, s\notin E\})} &< r. \end{aligned}$$

Our objective is to show that for the rectangle

$$R = [x_0 - h_0, x_0 + h_0] \times [T(x_0) - \sqrt{r}, T(x_0) + \sqrt{r}]$$

the values $\mu_{B_\delta^{*n}}(R)$ do not converge, implying that the sequence $(B_\delta^{*n})_{n\in\mathbb{N}}$ can not even converge pointwisely.
A straightforward application of the Markov inequality (see, e.g., [17]) yields

$$\frac{1}{2h_0}\lambda\left(\{s \in [x_0 - h_0, x_0 + h_0] : \, |T(s) - T(x_0)| \geq \sqrt{r}\}\right) \quad < \quad \sqrt{r},$$

so altogether we get

$$\begin{aligned}\frac{1}{2h_0}\lambda\left(\{s\in[x_0-h_0,x_0+h_0]:\,|T(s)-T(x_0)|<\sqrt{r}\}\right) &> 1-\sqrt{r},\\ \frac{1}{2h_0}\lambda\left(\{s\in[x_0-h_0,x_0+h_0]:\,s\in E\}\right)>1-r &> 1-\sqrt{r}.\end{aligned}$$

Having this, applying Bonferroni inequality as well as $1-2\sqrt{r}>\frac{3}{4}$

$$\frac{1}{2h_0}\lambda\left(\{s\in[x_0-h_0,x_0+h_0]:\,s\in E \text{ and } |T(s)-T(x_0)|<\sqrt{r}\}\right)>\frac{3}{4} \tag{29}$$

follows. For every $n\in\mathbb{N}$ using disintegration we have

$$\mu_{B_\delta^{*n}}(R)=\int_{[x_0-h_0,x_0+h_0]}K_{B_\delta^{*n}}(s,[T(x_0)-\sqrt{r},T(x_0)+\sqrt{r}])\,d\lambda(s). \tag{30}$$

Considering periodicity of $K_{B_\delta^{*n}}(s,\cdot)$ for $s\in E$ for odd $n\in\mathbb{N}$ we get

$$\begin{aligned}\mu_{B_\delta^{*n}}(R) &= \int_{[x_0-h_0,x_0+h_0]}K_{B_\delta^{*n}}(s,[T(x_0)-\sqrt{r},T(x_0)+\sqrt{r}])\,d\lambda(s)\\ &\geq \int_{[x_0-h_0,x_0+h_0]\cap E}K_{B_\delta}(s,[T(x_0)-\sqrt{r},T(x_0)+\sqrt{r}])\,d\lambda(s)\\ &\geq 2h_0\,\frac{3}{4},\end{aligned}$$

i.e., $\mu_{B_\delta^{*n}}$ assign at least $\frac{3}{4}$ of the total mass of the stripe $[x_0-h_0,x_0+h_0]\times[0,1]$ to R. On the other hand, for even $n\in\mathbb{N}$ using the fact that $K_{B_\delta^{*2}}(s,[T(x_0)-\sqrt{r},T(x_0)+\sqrt{r}])=0$ holds for every $s\in E\cap[x_0-h_0,x_0+h_0]$, it follows that

$$\begin{aligned}\mu_{B_\delta^{*n}}(R) &= \int_{[x_0-h_0,x_0+h_0]}K_{B_\delta^{*n}}(s,[T(x_0)-\sqrt{r},T(x_0)+\sqrt{r}])\,d\lambda(s)\\ &= \int_{[x_0-h_0,x_0+h_0]\setminus E}K_{B_\delta^{*n}}(s,[T(x_0)-\sqrt{r},T(x_0)+\sqrt{r}])\,d\lambda(s)\\ &\leq \lambda\left([x_0-h_0,x_0+h_0]\setminus E\right)\leq 2h_0 r\leq 2h_0\frac{1}{64}.\end{aligned}$$

□

The previous results show that the limit copula C is an idempotent Bertino copula. We therefore close this section by providing a characterization of idempotence in the Bertino class $\mathcal{C}_{Ber}$.

Theorem 15. *For a Bertino copula B_δ the following two conditions are equivalent:*

1. *B_δ is idempotent.*
2. *For every $x \in [0,1]$ with $(x, T(x)) \in (\hat{\Lambda}^+ \cup \Lambda^-) \times \Lambda$ we have*

$$|\hat{\delta}'(T(x))| + |\hat{\delta}'(x)| = 1.$$

Proof. By λ-almost everywhere uniqueness of the Markov kernel we have that B_δ is idempotent, if and only if

$$(K_{B_\delta} \circ K_{B_\delta})(x, \cdot) = K_{B_\delta}(x, \cdot) \tag{31}$$

for λ-almost every $x \in [0,1]$. Since for $x \in \hat{\Lambda}^0$ this identity obviously holds using eq. (16) idempotence of B_δ is therefore equivalent to the equality

$$\begin{aligned} K_{B_\delta * B_\delta}(x, E) &= |\hat{\delta}'(x)| \Big(|\hat{\delta}'(T(x))| \epsilon_{T \circ T(x)}(E) \; + \; (1 - |\hat{\delta}'(T(x))|) \epsilon_{T(x)}(E) \Big) \\ &\quad + (1 - |\hat{\delta}'(x)|) \Big(|\hat{\delta}'(x)| \epsilon_{T(x)}(E) + (1 - |\hat{\delta}'(x)|) \epsilon_x(E) \Big) \\ &= |\hat{\delta}'(x)| \epsilon_{T(x)}(E) + \Big(1 - |\hat{\delta}'(x)| \Big) \epsilon_x(E) = K_{B_\delta}(x, E) \end{aligned} \tag{32}$$

to hold for every $x \in [0,1]$ with $(x, T(x)) \in (\hat{\Lambda}^+ \cup \Lambda^-) \times \Lambda$ and every $E \in \mathcal{B}([0,1])$.
(i) For $(x, T(x)) \in (\hat{\Lambda}^+ \cup \Lambda^-)^2$ we have $T(T(x)) = x \neq T(x)$ as well as $|\hat{\delta}'(x)| > 0$, so (comparing the coefficients of ϵ_x) eq. (32) is equivalent to $|\hat{\delta}'(T(x))| + |\hat{\delta}'(x)| = 1$.
(ii) For $(x, T(x)) \in (\hat{\Lambda}^+ \cup \Lambda^-) \times \hat{\Lambda}^0$ we have $T(T(x)) = T(x) \neq x$ as well as $|\hat{\delta}'(x)| > 0 = \hat{\delta}'(T(x))$, so (again comparing the coefficients of ϵ_x) eq. (32) is equivalent to $|\hat{\delta}'(x)| = 1$. This completes the proof. □

Remark 16. Looking at the proof of Theorem 12 in the light of Theorem 15 shows that eq. (26) and eq. (28) precisely correspond to the condition $|\hat{\delta}'(T(x))| + |\hat{\delta}'(x)| = 1$.

Acknowledgements

Both authors gratefully acknowledge the support of the WISS 2025 project ‘IDA-lab Salzburg’ (20204-WISS/225/197-2019 and 20102-F1901166-KZP).

References

[1] J.J. Arias García, H. De Meyer, B. De Baets: Multivariate Bertino copulas, *J. Math. Anal. Appl.* 434(2), 1346-1364, 2016.

[2] S. Bertino: Sulla dissomiglianza tra mutabili cicliche, *Metron* 35, 53-88, 1977.

[3] I. Cuculescu, R. Theodorescu, Copulas: diagonals, tracks, *Rev. Roumaine Math. Pures Appl.* 46, 731-742, 2001.

[4] W.F. Darsow, B. Nguyen, E.T. Olsen: Copulas and Markov Processes, *Illinois J. Math.* 36(4), 600-642, 1992.

[5] F. Durante, P. Jaworski, Absolutely continuous copulas with given diagonal sections, *Comm. Statist. Theory Methods* 37, 2924-2942, 2008.

[6] F. Durante, R. Mesiar, C. Sempi: On a family of copulas constructed from the diagonal section *Soft Comput.* 10, 490-494, 2006.

[7] F. Durante, A. Kolesárová, R. Mesiar, C. Sempi: Copulas with given diagonal sections: novel constructions and applications, *Int. J. Uncertain. Fuzziness Knowlege-Based Syst.* 15, 397-410, 2007.

[8] F. Durante, C. Sempi: *Principles of Copula Theory*, Chapman and Hall/CRC, 2015.

[9] H.K. Farahat: The semigroup of doubly-stochastic matrices, *Proc. Glasgow Math. Ass.* 7, 178-183, 1966.

[10] G.A. Fredricks, R.B. Nelsen: The Bertino family of copulas, in C.M. Cuadras, J. Fortiana, J.A. Rodriguez-Lallena (eds): *Distributions with given Marginals and Statistical Modeling*, pp. 81-91, Springer Science+Business Media Dordrecht, 2002.

[11] J. Fernández Sánchez, W. Trutschnig: Some members of the class of (quasi-) copulas with given diagonal from the Markov kernel perspective, *Comm. Statist. Theory Methods*, 45, 1508-1526, 2016.

[12] 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.

[13] O. Häggström: *Finite Markov Chains and Algorithmic Applications*, Cambridge University Press, 2002.

[14] P. Jaworski, On copulas and their diagonals, *Inform. Sci.* 179, 2863-2871, 2009.

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

[16] T. Kasper, S. Fuchs, W. Trutschnig: On weak conditional convergence of bivariate Archimedean and Extreme Value copulas, and consequences to nonparametric estimation, *Bernoulli* 27(4), 2217-2240, 2021.

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

[18] L. Maislinger, W. Trutschnig: On bivariate lower semilinear copulas and the star product, *Int. J. Approx. Reason.* 179, 109366, 2025.

[19] R.B. Nelsen, *An Introduction to Copulas*, 2nd edition, Springer Series in Statistics, New York, 2006.

[20] R.B. Nelsen, J.J. Quesada-Molina, J.A. Rodríguez-Lallena, M. Ubeda-Flores: On the construction of copulas and quasi-copulas with given diagonal sections, *Insur.: Math. Econ.* 42, 473-483, 2008.

[21] J. Norris: *Markov Chains*, Cambridge University Press, 1997.

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

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

[24] W. Trutschnig, J. Fernández Sánchez: Idempotent and multivariate copulas with fractal support, *J. Statist. Plann. Inference* 142, 3086-3096, 2012.

[25] W. Trutschnig: On Cesáro convergence of iterates of the Star Product of copulas, *Stat. Probab. Lett.* 83, 357-365, 2013.

[26] W. Trutschnig, J. Fernández Sánchez: Idempotent and multivariate copulas with fractal support, *J. Stat. Plan. Inference* 142, 3086-3096, 2012.