Table of Contents

Editorial
Special Issue in memory of Abe Sklar Open Access

Giovanni Puccetti*

Special Issue on copulas in memory of Abe Sklar (1925-2020)

https://doi.org/10.1515/demo-2021-0109
Received June 30, 2021; accepted July 1, 2021

Abe (born Abraham) Sklar, one of the milestone personalities in the field of copulas and dependence modeling, passed away on October 30, 2020 at the age of 94. Soon after, I received a request from some of his friends and collaborators to publish a Special Issue devoted to his memory. I am very glad that many eminent colleagues made this Special Issue not only possible but also a gratifying reading. For the occasion, I left the handling of the review process in the safe hands of Roger Nelsen and Carlo Sempi, whom I thank for their kind cooperation.
In 2015, when *Dependence Modeling* started its Interview Article series, the name of Abe was the first on the list. Some members of the editorial board tried to contact him, and even managed to send by post a list of questions to be answered. Unfortunately, we never heard back. Personally, I always felt this was a missing tile in our journal, almost a shame. Finally, this gap has been bridged by Christian Genest, who wrote a magnificent tribute that will open the Special Issue.

Presumably, if you are reading these words, you have already encountered Sklar's theorem in your career. Can you remember the first occurrence? On my side, it happened as an undergraduate student in Florence. I remember that I was a bit sick one day and could not wait to leave the class and go home. One professor urged me to stay since the department was having a distinguished speaker that day: Paul Embrechts. In the early days of Risk Aggregation via Copulas, that was the first time I noticed the title *Sklar's Theorem* on one slide. Incidentally, Paul delivered the first interview in the series, years later.

Below is what Roger and Carlo had to say. And you, can you remember your first encounter with Abe?

Roger Nelsen: During the 1983-84 academic year I had a "sabbatical lectureship" at the University of Massachusetts in Amherst. It was during that year that I met Berthold Schweizer, who not only introduced me to copulas, but also to several visitors to Amherst that year, including Abe Sklar and Jerry Frank (and others). That year changed my life.

Carlo Sempi: I well remember my first meeting with Abe; it was on the spring of 1972 in Waterloo, where Abe usually came to visit alone or accompanied with Berthold Schweizer or Jerry Frank. He used to come there to talk with János Aczél, Bruno Forte (my PhD supervisor) and Michael McKiernan.

***Corresponding Author: Giovanni Puccetti:** Dipartimento di Economia, Management e Metodi Quantitativi, Università di Milano, Italy. E-mail: giovanni.puccetti@unimi.it
Photo courtesy of J.A. Rodríguez-Lallena.

Depend. Model. 2021; 9:200–224 DE GRUYTER

Interview Article
Special Issue in memory of Abe Sklar
Open Access

Christian Genest*

A tribute to Abe Sklar

https://doi.org/10.1515/demo-2021-0110
Received June 27, 2021; accepted July 2, 2021

This paper gives an account of the life and works of the American mathematician Abe Sklar. Born in Chicago on November 17, 1925, Sklar completed his PhD at the California Institute of Technology in 1956. He then joined the Illinois Institute of Technology, where he taught mathematics until his retirement in 1995. With his close friend and lifelong collaborator Berthold Schweizer (1929–2010), he was a pioneer of the theory of probabilistic metric spaces, which were introduced in 1942 by the Austro-American mathematician Karl Menger (1902–85). Together, Schweizer and Sklar made important contributions to the algebra of functions, the study of t-norms, and distributional chaos. Sklar is also credited for the notion of copula and for showing that any multivariate distribution function can be expressed in terms of its univariate margins and a copula. This result, known as Sklar's representation theorem, is the bedrock of a widespread data analytical technique called copula modeling. Sklar passed away in Chicago on October 30, 2020.

By now, just about anyone who is conscious of the role of dependence in data analysis has heard of copulas as a powerful and flexible tool for modeling association and assessing its impact on inference, decision making, and risk management. This approach is rooted in a 3-page note, written in French, which appeared in 1959 in the *Publications de l'Institut de statistique de l'Université de Paris*. This paper [S6], attributed only to "M. Sklar" (M. for Mr.), without address or affiliation, claimed without proof that given any d-variate cumulative distribution function H with one-dimensional margins $F_1, \ldots, F_d$, a function $C : [0, 1]^d \to [0, 1]$ having specific analytical properties can always be found such that

$$H = C(F_1, \ldots, F_d). \tag{1}$$

The author, Abe Sklar, called C a 'copula' by analogy with the grammatical term for a word or expression that links a subject and a predicate [S57]. He also pointed out that any choice of copula C and marginal distributions $F_1, \ldots, F_d$ in formula (1) constitutes a *bona fide* multivariate distribution function.

Eq. (1), now widely referred to as Sklar's representation theorem, is the corner stone of the popular copula approach to modeling. According to *MathSciNet*, Sklar's seminal paper [S6] has been referenced over 450 times since its publication; *Google Scholar* lists 10 times as many citations. What a success story for this note which stemmed from correspondence between Sklar, who was then a modest Assistant Professor of Mathematics in Chicago, and the famous French mathematician Maurice Fréchet!

This paper surveys the life and works of Abe Sklar (1925–2020), a talented and discreet man with a singular passion for mathematics. His biography is sketched in Section 1 and an overview of his scientific contributions is given in Section 2. His spirit continues to live through his students, whose own careers are outlined in Section 3. A few testimonies are collected in Section 4 and Appendix 1 contains a (hopefully complete) list of Sklar's publications. Finally, an annotated transcript of an interview that Sklar granted to Bonnie Schulman in 2006 is included in Appendix 2.

***Corresponding Author: Christian Genest:** Department of Mathematics and Statistics, McGill University, Montréal (Québec) Canada. E-mail: christian.genest@mcgill.ca

1 Abe's biography

Abe Sklar was born in Chicago, Illinois, on November 17, 1925. His parents were Jewish immigrants of humble extraction from (modern-day) Ukraine. Abe, who was their only child, graduated from Von Steuben High School in Chicago on June 25, 1942 and entered the University of Chicago at the age of 16. He was mainly interested in physics but after two years, he decided to interrupt his studies and join the war effort. He enlisted in the US Navy on June 20, 1944, and proudly served as a radio technician in the Philippines until his honorable discharge on June 2, 1946; his final rank was RT2/c.

Upon returning home, now aged 21, Abe resumed his studies at the University of Chicago but switched to a major in mathematics. Having completed a master's degree in Fall 1948, he worked as a physics lab technician at the university while taking, on and off, additional courses in mathematics, science, and the arts until Fall 1951. Allen Strehler (1921–78), an educator [43] and later dean at the Carnegie Institute of Technology, kindled and sustained Abe's interest in number theory, leading to his first paper published in 1952 in the *Notices of the American Mathematical Society* [S1]. The same year, Abe registered for a PhD in mathematics at the California Institute of Technology.

In moving to California, Abe's initial intention was to contribute to the Bateman Project, a major effort at collation and encyclopedic compilation of the theory of special functions. However, the project was nearing completion and he wound up working under the guidance of analytic number theorist Tom Apostol (1923–2016), best known for his widely used textbooks. Abe was just two years younger than his supervisor and was one of his first PhD students. He earned his PhD is mathematics in 1956 with a thesis entitled "Summation Formulas Associated with a Class of Dirichlet Series" [S2].

In February 1956, Haim Reingold (1910–2003), who served for many years as Chair of the Department of Mathematics at the Illinois Institute of Technology (IIT), persuaded Abe to return to Chicago and join his department. Except for sabbatical leaves, Abe's entire academic career was spent at IIT. He was promoted to Assistant Professor in 1957, to Associate Professor in 1960, and to Professor of Mathematics in 1967. He retired in 1995 and was appointed Professor Emeritus in 1996. He passed away in Chicago on October 30, 2020, after a short illness.


Illinois Tech Appoints Five New Instructors

The appointment of five new instructors has been announced by the Illinois Institute of Technology. They are Michael A. McKiernan, 1056 Ardmore av.; Barth Pollak, 7437 South Shore dr.; Abe Sklar, 4834 N. Troy st.; Claus P. Jacobson, 5400 S. Greenwood av., and Dennis W. Prosser, 2108 Ainslie av.

Chicago Tribune, May 20, 1956

At IIT, Abe was greatly influenced by Karl Menger (1902–85), a prominent Austrian mathematician who had emigrated to the United States in 1937 to flee nazism and whose longest and last academic appointment was at IIT from 1946 to 1971. Menger is considered as one of the founders of distance geometry, which is the study of sets of points based only on given values of the distances between pairs. He is also well-known for Menger's theorem in graph theory and the so-called Menger sponge, which is a three-dimensional generalization of the classical Cantor set (have a look on *Wikipedia*!).

In 1942, Menger initiated a study of probabilistic metric spaces [24], in which distances are represented by distributions rather than by numbers. Abraham Wald (1902–50), a well-known statistician and decision theorist who had been Menger's student in Vienna, showed interest [46] and upon completing his own thesis with Menger at IIT, Bert Schweizer (1929–2010) also started investigating the topic.

Abe joined these efforts shortly after coming to IIT (see Figure 1). He nurtured a lifelong friendship and collaboration with Bert, with whom he developed the theory of probabilistic metric spaces, and contributed to ergodic theory as well as distributional chaos to much acclaim. Together they wrote over 40 papers and an influential monograph on probabilistic metric spaces [S44] published in 1983. For over 60 years, Abe used his vast knowledge of analysis and functional equation theory to explore these and other themes. In addition, he invested much time and effort in memorializing Menger and his work.

During his career, Abe enjoyed (at least) two sabbatical leaves. One was in Spring 1962, when he taught courses at the University of Arizona [S41] while collaborating with Bert, who was then on faculty there. He had a second leave during the 1982–83 academic year. He spent part of that year working with Luigi Paganoni at the Università di Milano, in Italy. Abe's paper [S62] with a coauthor of Paganoni, Jaroslav Smítal (1942–), also reflects research that he conducted on a subsequent visit there.

Figure 1: Karl Menger with some of his students* and colleagues† at IIT circa 1956. From left to right: Abe Sklar†, Bert Schweizer*†, Karl Menger, Tom Erber†, and Michael McKiernan.*†

2 An overview of Abe's work

In a paper published in 1980 in *The American Mathematical Monthly* [S41], Abe cited analytic number theory, functional equations, function systems, and probabilistic geometry as his research areas. At the 1993 conference on distributions with fixed marginals held in Seattle [S57], he explained how he actually became interested in probability through number theory, and how Bert convinced him to join his project to develop probabilistic metric spaces soon after the two met at IIT.

Motivated by his youthful encounters with Albert Einstein (1879–1955), who was then a neighbor, Bert had come to Chicago to explore probabilistic metric spaces after reading Menger's contribution to a book written in appreciation of Einstein's work [25]. In Menger's opinion, however, this topic was unsuitable for a thesis, and hence Bert could only delve into the matter after completing his degree [37].

Thus began a lifelong friendship and collaboration between Bert and Abe, who also remained close to Karl Menger. When the latter died in 1985, Abe was entrusted with his archives and, with the help of Bert and others, he published a series of volumes devoted to Menger's work [29, 31, 32] and arranged for Menger's calculus textbook [27] to be reissued by Dover in 2007.

Abe's extensive knowledge of Karl Menger's life and thinking is apparent from his interview with Bonnie Shulman [42] which is reproduced in Appendix 2. This discussion is included here in full, as it provides a faithful reflection of Abe's humility, wit, and good humor. It also illustrates his close relationship with both Bert and Karl Menger – so close, in fact, that it is difficult to talk about his research without describing also some of Bert's and Menger's contributions.

As a full review of Abe's work would be too ambitious, Sections 2.1–2.3 are limited to three major themes that occupied his mind over his long and fruitful career: probabilistic metric spaces, derivability issues and bounds, and distributional chaos. One distinct omission is the algebra of functions, another theme pioneered by Menger which was among Abe's finest achievements and about which a book project with Bert remained unfinished. Section 2.4 gives a glimpse of Abe's *modus operandi* as a researcher.

Inspiration for this section was drawn in part from the "personal look backward and forward" that Abe presented at the 1993 meeting in Seattle [S57]. Also useful were two limpid survey papers written by Bert: the first for the 1990 conference on distributions with fixed marginals held in Rome [38], and the second, for

Bert's talk in the series *Seminario Matematico e Fisico di Milano* [39]. For Section 2.3, I consulted the survey of distributional chaos and its measurement that Bert and Abe wrote with their Czech collaborator Jaroslav Smítal for *Real Analysis Exchange* in 2000 [S63].

For precise statements, additional details and further results, readers should consult these four papers and references therein. In particular, note that following an old European tradition, distribution functions were taken to be *left-continuous* in many of the papers cited below.

2.1 Probabilistic metric spaces

Consider a set S and a map $\mathcal{F} : S \times S \to \Delta_0$ such that, for every pair $(x, y) \in S \times S$, $\mathcal{F}(x, y) = F_{xy}$ is an element in the space Δ_0 of cumulative distribution functions that vanish on $(-\infty, 0)$. For any $F, G \in \Delta_0$, write $F \leq G$ whenever $F(t) \leq G(t)$ for all $t \in \mathbb{R} = (-\infty, \infty)$.

Following the authoritative book that Bert and Abe on probabilistic metric spaces [S44], a map $\tau : \Delta_0 \times \Delta_0 \to \Delta_0$ is called a *triangle function* if, for all $F, G, H, K \in \Delta_0$, (i) $\tau(F, G) = \tau(G, F)$; (ii) $\tau\{\tau(F, G), H\} = \tau\{F, \tau(G, H)\}$; (iii) $\tau(F, G) \leq \tau(H, K)$ whenever $F \leq H$ and $G \leq K$; (iv) $\tau(F, \delta_0) = F$, where δ_0 is the distribution function of a random variable which is identically equal to 0.

Because of its properties, the map τ can be viewed as a commutative and associative operation on Δ_0. The pair $(S, \mathcal{F})$ is then said to form a *probabilistic metric space* with respect to τ if, for all $x, y, z \in S$, (i) $F_{xy} = F_{yx}$; (ii) $x = y \Leftrightarrow F_{xy} = \delta_0$; (iii) $F_{xz} \geq \tau(F_{xy}, F_{yz})$ holds point-wise, so that F_{xz} is stochastically smaller than $\tau(F_{xy}, F_{yz})$. Condition (iii) is called Menger's triangle inequality.

A natural choice of triangle function τ is convolution, as noted by Wald [46]. However, Menger [24] viewed this choice as too stringent and proposed to generate τ from a *triangular norm* or *t-norm* T, i.e., a binary algebraic operation such that for all $a, b, c, d \in [0, 1]$, (i) $T(a, 1) = a$; (ii) $T(a, b) = T(b, a)$; (iii) $T(a, b) \leq T(c, d)$ whenever $a \leq c$ and $b \leq d$; (iv) $T\{T(a, b), c\} = T\{a, T(b, c)\}$. A triangle function τ is then induced by T upon setting, for all $F, G \in \Delta_0$ and $t \in [0, \infty)$,

$$\tau_T(F, G)(t) = \sup\left\{T\{F(u), G(v)\} : u + v = t\right\}. \tag{2}$$

Simple examples of t-norms include T_W, T_M, and T_Π, respectively defined for all $a, b \in [0, 1]$, by $T_W(a, b) = \max(0, a + b - 1)$, $T_M(a, b) = \min(a, b)$, and $T_\Pi(a, b) = ab$. The first two are known in the copula literature as the Fréchet–Hoeffding bounds, and T_Π is the independence (or product) copula.

For nearly 20 years, Bert and Abe studied these and many other aspects of the theory of probabilistic metric spaces with their students, colleagues, and a few collaborators. They worked in relative isolation and at their own pace. In his 1991 review paper [38, p. 29], Bert stated that in hindsight, this...

> "... let the ideas and techniques mature in an environment uninfected by the contentious competition that seems to surround so many 'hot' topics."

In particular, the statistical community would only start noticing in the 1980s, after Bert and his PhD student Ed Wolff made a connection with measures of dependence in an *Annals of Statistics* paper [41].

One very fruitful line of investigation is exemplified by a 1961 paper in *Publicationes Mathematicae Debrecen* [S11], where Bert and Abe used results from the theory of functional equations to show that if a t-norm T is continuous and strictly increasing, it can be expressed, for all $a, b \in (0, 1]$, in the form

$$T(a, b) = \phi^{-1}\{\phi(u) + \phi(v)\}. \tag{3}$$

for some continuous and strictly decreasing function $\phi : [0, 1] \to [0, \infty]$ with $\phi(1) = 0$ and quasi-inverse ϕ^{-1}. In the same paper, they also proved the familiar fact that T is then a copula if and only if ϕ is convex; see [10] or Theorem 4.1.4 in [34]. Abe's first PhD student, Cho-hsin (Joyce) Ling [23], later published in the same journal a proof that if a continuous t-norm T is Archimedean, i.e., $T(a, a) < a$ for all $a \in (0, 1)$, then T is necessarily of the form (3). And so were born *Archimedean copulas*.

Among many other things, this line of research led Bert and Abe to propose various one-parameter classes of strict t-norms, the most prominent of which is the *Schweizer–Sklar family* which includes what are now known as Clayton copulas. In this context, the Gumbel copulas are Aczél–Alsina t-norms and the Hoeffding–Fréchet upper and lower bounds are called the Gödel and Łukasiewicz t-norm, respectively.

Also, because t-norms can be viewed as an extension of the usual two-valued logical conjunction, the work of Bert and Abe in this area bore fruit in fuzzy logic. When Lotfi Zadeh (1921–2017) first introduced fuzzy mathematics [47], back in 1965, max and min were the only (associative) operations used to perform generalized unions and intersections. Bert and Abe's contributions eventually made it possible to replace these operations by t-norms and their duals called t-conorms. These concepts have since become essential tools to represent logical disjunction in fuzzy logic and union in fuzzy set theory.

Years later, new and fruitful definitions for a probabilistic normed space [S56] and a probabilistic inner product space [S58] were found by Bert and Abe in collaboration with Claudi Alsina (1952–), who was Bert's postdoc at the University of Massachusetts in Amherst before joining Universitat Politècnica de Catalunya; Carlo Sempi (1948–), from the Università del Salento [13] also coauthored [S58]. For more about t-norms, refer to the monographs by Alsina et al. [1] and Klement et al. [20].

2.2 Derivability issues and bounds

Abe's work with Bert Schweizer led them to look into operations on distribution functions that could be derived from operations on random variables. Formally, a binary operation $\star$ on the class Δ of distributions is said to be derivable from a (Borel-measurable) function Ψ on random variables if for any $F, G \in \Delta$, random variables X and Y with these distributions can be constructed on a common probability space and $\Psi(X, Y)$ has distribution $F \star G$. Convolution is an example of an operation which is derivable, as it corresponds to the addition of random variables.

In their remarkable 1974 paper [S34], Bert and Abe showed that if T is any continuous t-norm other than T_M, then the algebraic operation τ_T defined in Eq. (2) is not derivable from any operation on random variables. This led them to the insightful conclusion, stated in Section 7.6 of [S44], that

> "the distinction between working directly with distribution functions... and working with them indirectly, via random variables, is intrinsic and not just a matter of taste."

This observation has an analog in dependence modeling: notwithstanding Sklar's representation theorem, models that are easily expressible in terms of copulas may not be interpretable in terms of operations on random variables, and vice versa. The two approaches are useful, and complementary.

In their search for classes of triangle functions, Bert and Abe further noticed very early that an extension of Eq. (2) would arise if addition, viz. $L(u, v) = u + v$, were replaced by any other two-place commutative and associative function L satisfying $L(0, 0) = 0$. In particular, for any (bivariate) copula C and marginal distributions $F, G \in \Delta_0$, one can set

$$\tau_{C,L}(F, G)(t) = \sup\left\{C\{F(u), G(v)\} : L(u, v) = t\right\}$$

and

$$\rho_{C,L}(F, G)(t) = \inf\left\{\bar{C}\{F(u), G(v)\} : L(u, v) = t\right\},$$

where $\bar{C}$ stands for the dual copula defined, for all $u, v \in [0, 1]$, by $\bar{C}(u, v) = u + v - C(u, v)$.

In the 1970s, one of Abe's PhD students, Jerry Frank, investigated yet another class of binary operations defined, for every copula C and marginal distributions $F, G \in \Delta_0$, by

$$\sigma_{C,L}(F, G)(t) = \iint \mathbf{1}\{L(u, v) \le t\} dC\{F(u), G(v)\},$$

where for any set A, $\mathbf{1}(A) = 1$ whenever A holds true, and 0 otherwise. The operation $\sigma_{C,L}$ reduces to standard convolution when L is the sum and C is the product copula. It is Frank who came to the surprising conclusion that in order for the latter operation to be associative, C must be what later came to be known as a Frank copula [7, 9]. This is not sufficient, however, as he later showed [8].

Figure 2: Abe Sklar (left) and Bert Schweizer (right) at the 35th International Symposium on Functional Equations held September 7–14, 1997 in Graz, Austria (participation in this meeting was by invitation only).

In joint work with a PhD student, Richard Moynihan [33], Bert and Abe were able to relate these different operations and showed in [S37] that for a large choice of functions L, the string of inequalities

$$\tau_{W,L} \le \tau_{C,L} \le \sigma_{C,L} \le \rho_{C,L} \le \rho_{W,L}$$

holds. The special case of the sum had first been treated by Abe alone [S32].

These inequalities, which are analogous to the Fréchet–Hoeffding bounds on any (bivariate) copula C, viz. $T_W \le C \le T_M$, are useful in exploring the geometric notion of betweenness in the context of a probabilistic metric space S with triangle function τ, where for any distinct elements x, y, z of S, y is said to lie between x and z if and only if $F_{xz} = \tau(F_{xy}, F_{yz})$.

2.3 Distributional chaos

The associativity of a t-norm $T : [0,1] \to [0,1]$ leads naturally to the consideration of serial iterates. The latter are defined recursively, for every integer $n \ge 1$ and scalars $a_1, a_2, \ldots \in [0,1]$, by

$$T^{n+1}(a_1, \ldots, a_{n+2}) = T\{T^n(a_1, \ldots, a_{n+1}), a_{n+2}\}, \tag{4}$$

where $T^1 = T$. It was Abe who first noted in [S32] that if T is a strict t-norm such that Eq. (4) yields an n-copula for every integer $n \ge 1$, then T must be of the form (3), i.e., an Archimedean copula with a completely monotonic generator ϕ^{-1}; furthermore, $T \ge T_\Pi$ point-wise, i.e., T is positive quadrant dependent [34]. Another of Abe's PhD students, Clark Kimberling, used this result to study sets of exchangeable random variables and give a probabilistic interpretation of complete monotonicity [19].

As recounted by Bert in the survey paper [39], the study of function iterates ultimately led Abe to study with him distributional chaos. To avoid technicalities, consider a compact, separable metric space (S, d) and a continuous function $f : S \to S$. For any $x, y \in S$ and every integer $n \ge 0$, set $\delta_{xy}(n) = d\{f^n(x), f^n(y)\}$, where

$f^0(x) = x$. The sequence $f^n(x)$ can be viewed as the trajectory of x in S, and for any real $t \in \mathbb{R}$, the proportion $F^n_{xy}(t)$ of points in the set $\{\delta_{xy}(0), \ldots, \delta_{xy}(n-1)\}$ which are less than or equal to t defines a distribution function F^n_{xy} in Δ_0. Next set, for all $t \in \mathbb{R}$,

$$F_{xy}(t) = \liminf_{n\to\infty} F^n_{xy}(t), \quad F^\star_{xy}(t) = \limsup_{n\to\infty} F^n_{xy}(t).$$

It can be seen that both F_{xy} and $F^\star_{xy}$ are distribution functions, and that the former satisfies Menger's triangle inequality. Together with the map $\mathcal{F} : S \times S \to \Delta_0$ defined by $\mathcal{F}(x, y) = F_{xy}$, the set S then forms what is called a "transformation generated space."

Bert and Abe initiated a study of this construction in the early 1970s. To describe their main results, assume that f is measure preserving with respect to some (non-singular) probability measure P on S, i.e., $P\{f^{-1}(A)\} = P(A)$ for any (measurable) set $A \subseteq S$. In the paper [S31] published in 1973, they used the Birkhoff–Khinchin ergodic theorem to show that, for any such transformation f,

$$F_{xy} = F^\star_{xy} \tag{5}$$

for almost all pairs $(x, y) \in S \times S$ endowed with the measure $P \times P$. The same year [S30], they established with Tom Erber, a mathematical physicist at IIT (see again Figure 1), that the limiting distribution in Eq. (5) is independent of x and y when in addition to being measure preserving, the transformation f is mixing, so that for any (measurable) subsets A and B of S, $P\{f^{-n}(A) \cup B\} \to P(A)P(B)$ as $n \to \infty$. Abe carried out calculations in a special case with Tom Erber [S33] and he studied the recurrence and dispersive behavior of Čebyšev polynomials with another IIT physicist, Porter Johnson [S35].

These results allowed Abe, Bert and their collaborators to shed new light on a famous controversy between Ludwig Boltzmann (1844–1906) and Ernst Zermelo (1871–1953) about the relevance to physics of Poincaré's recurrence theorem. The latter result states that a broad class of systems will, after a finite time, return to a state arbitrarily close to their initial state. While this conclusion appears to contradict thermodynamic principles, it relies critically on the assumption that the state of the system is known exactly. The results of Erber et al. [S30] vindicate Boltzmann's intuition that due to the uncertainty induced by experimentally indistinguishable states, recurrence may actually fail to occur.

These considerations eventually led Bert and Abe to elaborate a theory of distributional chaos to which Jaroslav Smítal, Professor of Mathematics at Slezská Univerzita in Opava, Czechia, was associated from its inception in 1985; see [39] and Figure 3. Specifically, a pair $(x, y) \in S \times S$ is then said to exhibit distributional chaos if and only if there exists an interval I of positive length such that, for all $t \in I$, $F_{xy}(t) < F^\star_{xy}(t)$. This leads in turn to a measure of the chaos induced by f, viz.

$$\mu_P(f) = \sup_{x,y\in S} \frac{1}{d_S} \int_0^\infty \{F^\star_{xy}(t) - F_{xy}(t)\}dt,$$

where the normalizing factor d_S, which ensures that $\mu_P(f) \le 1$, is assumed to be finite.

As shown by Schweizer and Smítal [40], this notion of distributional chaos implies chaos in the classical sense of Li and Yorke [22]. With their collaborators, Bert and Abe then continued to contribute to chaos theory, which studies dynamical systems whose apparently random states of disorder and irregularities are actually governed by underlying patterns and deterministic laws that are highly sensitive to initial conditions. Three of Abe's last five papers, published in the early 2000s, are concerned with the elucidation of this concept and the conditions on f required for distributional chaos to exist; see [S62], [S63], and [S65]. All of them were coauthored by his friend Jaroslav Smítal (see again Figure 3). Ultimately, the notion of distributional chaos introduced in this body of work makes it possible to clearly distinguish ergodic behavior from genuinely chaotic behavior.

2.4 Modus operandi

Judging from the meetings he attended regularly, Abe's scientific home was the community of functional equation specialists that was first brought together by János Aczél (1924–2020), Otto Haupt (1887–1988) and

Figure 3: Abe Sklar and Jaroslav Smítal featured in the Czech newspaper *Opavský a Hlučínský Deník*, June 22, 2008. The 46th International Symposium on Functional Equations was held in Opava June 22–29, 2008.

Alexander Ostrowski (1893–1986) at the *Mathematisches Forschungsinstitut Oberwolfach* in 1962. Annual meetings known as the *International Symposium on Functional Equations* (ISFE) have been held since then (with the exception of 1964). Participation is by invitation only. Abe was present in 1962 and remained a regular attendee and member of the scientific committee thereafter. He often used these meetings as an opportunity to have intensive working sessions with Bert. The picture shown in Figure 2 was taken on one such occasion. Reports of these meetings have appeared in *Aequationes Mathematicae* since this journal was founded by Aczél in 1968. Abstracts of Abe's talks published therein provide a remarkable record of what was on his mind at any given time; see, e.g., Sklar [S39], cited in [S55].

3 Sklar's students

Part of any researcher's intellectual legacy is in the graduate students he/she supervised, and what they became. In this regard also, Abe's career was remarkable. A list of his PhD students is given in Table 1. It is based in part on, but more complete than, the data provided by the *Mathematical Genealogy Project*.

Looking at Table 1, copula modelers will immediately recognize Jerry Frank (1942–), of Frank copula fame, who was already mentioned in Section 2. After graduation, he had a long career as a professor in, and for some time chair of, the Department of Mathematics at IIT. In addition to publishing a dozen papers on functional equations, distribution theory, and non-linear dynamics, he coauthored with Claudi Alsina and Bert Schweizer a monograph on t-norms and copulas [1]; their work, published in 2006, nicely complements the 2000 book on the same topic by Klement et al. [20].

Table 1: A list of Abe Sklar's 14 PhD students at the Illinois Institute of Technology.

Name	Year	Name	Year
Ling, Cho-hsin (Joyce)	1964	Frank, Maurice J.	1972
Penner, Sidney	1964	Jurschak, Jerome	1974
Senechal, Marjorie	1965	Ranade, Mary S.	1974
Calabrese, Philip G.	1968	Thoresen, Lee L.	1978
Harkin, Joseph B.	1968	Terrana, Victor E.	1979
Marcus, Philip	1968	Krause, Gerianne M.	1981
Kimberling, Clark	1970	Ghebremeskel, Kuflu	1989

From the angle of dependence modeling, another name that stands out is Clark Kimberling (1942–), mentioned earlier for his probabilistic interpretation of complete monotonicity in Archimedean copulas [19]. He too had a successful career as a professor of mathematics, publishing extensively in number theory and geometry. He has been affiliated with the University of Evansville, Indiana, since earning his PhD. A man of multiple talents, he maintains an extensive website (link) that features, among others, the world's largest internet collection of triangle centers, facts and problems about integer sequences and arrays, but also many biographical studies (starting with [18]) and musical compositions of his own.

The name of Marjorie Senechal (1939–) is also intimately linked with geometry. A Professor Emerita in Mathematics and History of Science and Technology at Smith College, Northampton, Massachusetts, Senechal is well known for her work on tessellations and quasicrystals. She also taught a course on ancient inventions and published several books about silk. Moreover, she served as co-editor-in-chief (2005–13) and later sole editor-in-chief (2013–20) of *The Mathematical Intelligencer* (TMI) in addition to editing the column "Mathematical Communities" since 1997. A fascinating account of her life and career is given in her *Adventures of an Amateur Crystallographer*, which may be found on the website of the American Crystallographic Association (link). See also her interview in TMI (link).

Many other students of Abe had careers teaching mathematics in colleges and universities. Sidney Penner, raised in the Bronx, returned there after completing his PhD and teaching mathematics for a few years at SUNY Buffalo. Affiliated to the Bronx Community College at the time of his premature death in May 1980, he had extensive journal contributions to problems and problem-solving, and was Problem Editor of the *New York State Mathematics Teachers Journal* [2, 14].

Pursing an idea sketched by Abe [S19], Joseph Harkin (1926–2006) wrote a thesis on "Sklar–Stieltjes integrals" [16] and then taught mathematics for 32 years (1971–2003) at SUNY Brockport. After graduation, Philip Marcus (1936–2020) was successively affiliated with Shimer College, Indiana University, the University of Kentucky, Christian Brothers College, and Eureka College. Similarly, Victor Terrana (1945–) taught mathematics at Indiana University Northwest, Queens University of Charlotte, North Carolina, and then at Newberry College, South Carolina, from 1995 to 2011. As for Gerianne Krause (1951–), she taught discrete mathematics at San Francisco State University for most of her career.

Other students of Abe's had careers in industry. For example, Jerry Jurschak (1948–) was a college teacher for three years in Chicago before holding various executive positions within the insurance brokerage, underwriting, and reinsurance business (e.g., Marsh McLennan, CNA Insurance) in Chicago and in New York City. Now in retirement, he is running a charitable foundation with his wife out of Katonah, New York. Another example is Mary Ranade (1941–), who taught college level math and computer/information science part and full time at several colleges and universities in Illinois, North Carolina, and Maryland. Most of her career was devoted to the design, administration, and evaluation of large database systems; she also developed software tools for relational database administrators.

Two of Abe's students had a major spiritual part to their lives. Philip Calabrese (1941–) initially taught mathematics in California (Naval Postgraduate School, Monterey; California State University, Bakersfield; Humboldt State University, Arcata). He then had an aerospace career, founded a consulting firm called Data Synthesis, and retired in 2004 as senior scientist at the Space and Naval Warfare Systems Center, San Diego. His work synthesizing logic and conditional probability culminated with his 2017 book [4]. Follow this link for a 2-hour interview with him about his "axiom of free will vs. mechanism."

Kuflu Ghebremeskel returned to Eritrea after completing his PhD. He was a mathematics lecturer at the University of Asmara for 10 years [S61]. He then became a full-time pastor and a leading figure of the Full Gospel Church of Eritrea before he was arrested in May 2004 because of his religious beliefs. He has been detained incommunicado since then and is presumed dead.

Finally, Huse Fatkić, a professor of mathematics at Univerzitet u Sarajevu, in Bosnia and Herzegovina, is also listed on the *Mathematical Genealogy Project* as having been supervised by Bert and Abe. As explained in his thesis [6], they mentored him and inspired his research topic while he visited the University of Massachusetts at Amherst on a fellowship in 1985–86. Although neither Bert nor Abe appears to have been his supervisor in a formal sense, Abe's last paper was with Fatkić [S68].

Alas, I was unable to retrace two of Abe's PhD students: Cho-hsin (Joyce) Ling and Lee L. Thoresen.

4 Personal testimonies

Abe, who was single, had a modest life at home but a passion for traveling which he lived out every summer. We crossed paths a number of times over the years. We were introduced in April 1990 at the inaugural *Symposium on Distributions With Given Marginals* held in Rome, and we met again at two subsequent conferences in this series: Barcelona (2000), and Québec (2004). I last saw him in 2009 at the meeting organized in Lecce, Italy, to mark the 50th anniversary of his seminal 1959 paper on copulas.

In our encounters, Abe never revealed much about his private life but besides his love of mathematics and his broad interest for science and philosophy, it was obvious that he enjoyed Italian cuisine and the opera. He also had a predilection for words and languages, and he was a strong advocate of Esperanto as a means of international communication and as an instrument of peace. However, Abe's command of English, coupled with his wry sense of humor, did not always go down well with people. Some initially frowned at his choice of the word 'copula' for distribution functions with uniform margins. He often proudly referred to it as a *sexy name*, a fact on which Jock MacKay and I later capitalized [11].

In preparation for this memorial article, I interviewed several of Abe's colleagues, students, and friends, all of whom are gratefully thanked in the Acknowledgments. One of them is Ty Olsen (1942–), a 1983 PhD graduate of IIT who taught mathematics there before retiring in 1995. Ty, who had a first career as a lawyer, was named executor of Abe's estate in his will. He dug out many facts about Abe's youth for me by sifting through boxes of papers, but he could find no curriculum vitae. He wrote to me:

> "Besides being a first-rate mathematician, Abe was a wonderful colleague and mentor, the backbone of our department at IIT for many years. He had a tremendous breadth of knowledge that he was happy to share."

Arthur Lubin, a current Professor of Applied Mathematics at IIT, also commented as follows:

> "Abe gladly gave advice and guidance to people working on outgrowths of his work. He could immediately cite relevant references, often providing a copy of the papers the next day. He was happy to spend time helping students, graduate or undergraduate, embarking on research projects. He was highly principled, and exceedingly modest. Many of his strongest papers appeared in somewhat obscure journals, due to his modesty and abnegation of fame."

Another longtime colleague and friend of Abe's, Tom Erber (1930–), Distinguished Professor Emeritus of Physics and Mathematics at IIT (see again Figure 1), had this to say:

> "Abe's strong abnegation to prominence and 'fame' meant that he scattered many basic results in somewhat obscure mathematics journals. His one contact with 'fame' occurred as a consequence of having invented 'copulas'... Beneath Abe's placid exterior there was a hard unyielding commitment to certain principles. For instance, when IIT resurrected its mathematical activities within the framework of an 'Applied Mathematics Department', Abe promptly resigned because he refused to carry the title of Professor of *Applied* Mathematics."

A surprising aspect of Abe's career is that except for Bert and his PhD students, he had next to no collaborator outside IIT. Claudi Alsina, a prolific functional equation researcher and mathematical popularizer, who is one of Abe's rare foreign coauthors (with Smítal), offered the following explanation:

> "While Abe liked to think about math and scribble notes in pencil on small pieces of paper, he did not enjoy writing formal papers or books. This is why, in particular, it took him so many years to publish his proof of what is known today as Sklar's representation theorem about copulas. In that regard, Abe was lucky to have an accomplished expositor and a perfectionist like Bert as his coauthor; they complemented one another very well. A problem with Abe was that it was easier to meet him anywhere in the world than to receive a letter from him. Bert really laughed when I once said that IIT is like a black hole: you may send anything there, but nothing ever comes out of it."

Claudi also pointed out to me that Abe was proud to have an Erdős number of 2. This means that Abe never wrote a paper with the renowned Hungarian mathematician Paul Erdős (1913–96), who was one of the most prolific mathematicians of the 20th century, but that he wrote a paper with someone who did, namely his PhD advisor Tom Apostol (see [S3] but I could not locate the Apostol–Erdős paper).

On the subject of fame, which does have its drawbacks, it is perhaps fitting to conclude this tribute with an anecdote recounted by Tom Erber. To give a bit of context first, the monthly American magazine *Wired* ran a cover story entitled “Recipe for disaster: The formula that killed Wall Street” on February 23, 2009. Later reissued in *Significance* [36], this award-winning paper, written by the financial journalist Felix Salmon (1972–), described how the unwise use of the Gaussian copula by some quants played a role in the onset of the 2007–08 financial crisis. As a result of this negative publicity, undiscriminating eyes suddenly came to blame copulas and copula modeling for this global crisis that brought about a great deal of suffering. Erber reported as follows on the collateral damage this inflicted upon Abe:

> “Unbeknownst to Abe, some ‘quants’ adapted copulas to predicting the behavior of financial markets. Since big money was involved some subset of quants became very angry with Abe, and one day a group of vociferous demonstrators showed up at Abe’s office at IIT. It took some time to quiet the situation down, prevent physical harm, and assure the crowd that although Abe had invented copulas, he had no idea about the financial ramifications.”

More discussion on copulas and the financial crisis can be found in the interviews of Paul Embrechts and David Li in this journal [5, 35], and in Paul’s interview in *International Statistical Review* [12].

Appendix 1: A partial list of Abe Sklar’s scientific writings

Below is a possibly incomplete list of publications authored or coauthored by Abe Sklar. Lacking a curriculum vitae, I constructed this list to the best of my ability using mainly, but not exclusively, *Google Scholar*, *MathSciNet*, *Research Gate*, and *Web of Science*. I regret any involuntary omission.

The list comprises 71 items, grouped in yearly blocks to give a visual clue of Abe’s research productivity over the course of his career. It excludes the works of Karl Menger [29, 31, 32] which were coedited by Abe and published many years after Menger’s death.

In [29], Menger reminisced about the celebrated *Wiener Kreis* (Vienna Circle), a group of philosophers and scientists chaired by Moritz Schlick (1882–1936) that met regularly from 1924 to 1936 at Universität Wien. Abe edited this book with Louise Ahrndt Golland (1942–), an independent scholar in the history of mathematics, and the famous Wittgenstein scholar Brian McGuinness (1927–2019).

The other two posthumous books by Menger [31, 32] represent his collected works, coedited by seven scholars, including Bert and Abe who were also the driving force behind a new release by Dover of Menger’s calculus [27], for which they wrote a foreword.

Abe Sklar’s scientific writings in chronological order

[S1] Sklar, A. (1952). On the factorization of squarefree integers. *Proc. Amer. Math. Soc. 3*(5), 701–705.

[S2] Sklar, A. (1956). *Summation Formulas Associated with a Class of Dirichlet Series*. PhD thesis, California Institute of Technology, Pasadena CA.

[S3] Apostol, T.M. and A. Sklar (1957). The approximate functional equation of Hecke’s Dirichlet series. *Trans. Amer. Math. Soc. 86*(2), 446–462.

[S4] Schweizer, B. and A. Sklar (1958). Espaces métriques aléatoires. *C. R. Acad. Sci. Paris 247*(2), 2092–2094.

[S5] Menger, K., B. Schweizer, and A. Sklar (1959). On probabilistic metrics and numerical metrics with probability 1. *Czechoslovak Math. J. 9*(3), 459–466.

[S6] Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris 8*, 229–231.

[S7] Schweizer, B. and A. Sklar (1960). Statistical metric spaces. *Pacific J. Math. 10*(1), 313–334.

[S8] Schweizer, B. and A. Sklar (1960). The algebra of functions. *Math. Ann. 139*(5), 366–382.

[S9] Schweizer, B., A. Sklar, and E. Thorp (1960). The metrization of statistical metric spaces. *Pacific J. Math. 10*(2), 673–675.

[S10] Sklar, A. (1960). On the definition of the Riemann integral. *Amer. Math. Monthly 67*(9), 897–900.

[S11] Schweizer, B. and A. Sklar (1961). Associative functions and statistical triangle inequalities. *Publ. Math. Debrecen 8*(1), 169–186.

[S12] Schweizer, B. and A. Sklar (1961). The algebra of functions. II. *Math. Ann. 143*(5), 440–447.

[S13] Schweizer, B. and A. Sklar (1961). Topology and Tchebycheff. *Amer. Math. Monthly 68*(8), 760–762.

[S14] Schweizer, B. and A. Sklar (1962). A mapping-algebra with infinitely many operations. *Colloq. Math. 9*(1), 33–38.

[S15] Schweizer, B. and A. Sklar (1962). Statistical metric spaces arising from sets of random variables in Euclidean *n*-space. *Teor. Verojatnost. i Primenen. 7*(4), 456–465.

[S16] Schweizer, B. and A. Sklar (1963). Associative functions and abstract semigroups. *Publ. Math. Debrecen 10*(1), 69–81.

[S17] Schweizer, B. and A. Sklar (1963). Triangle inequalities in a class of statistical metric spaces. *J. London Math. Soc. 38*(1), 401–406.

[S18] Schweizer, B. and A. Sklar (1963). Inequalities for the confluent hypergeometric function. *J. Math. and Phys. 42*, 329–330.

[S19] Sklar, A. (1964). Uniform Stieltjes integrals. *Notices Amer. Math. Soc. 11*(3), 342–343.

[S20] Sklar, A. (1964). On some exact formulæ in analytic number theory. In Marsch, D.C.B. (Ed.), *Report of the Institute in the Theory of Numbers*, pp. 104–110, University of Colorado, Boulder CO.

[S21] Schweizer, B. and A. Sklar (1965). The algebra of functions. III. *Math. Ann. 161*(3), 171–196.

[S22] Schweizer, B. and A. Sklar (1967). Function systems. *Math. Ann. 172*(1), 1–16.

[S23] Schweizer, B. and A. Sklar (1967). The algebra of multiplace vector-valued functions. *Bull. Amer. Math. Soc. 73*(4), 510–515.

[S24] Schweizer, B. and A. Sklar (1968). A grammar of functions. I. *Aequationes Math. 2*(1), 62–85.

[S25] Schweizer, B. and A. Sklar (1969). A grammar of functions. II. *Aequationes Math. 3*(1–2), 15–43.

[S26] Schweizer, B. and A. Sklar (1969). Mesures aléatoires de l'information. *C. R. Acad. Sci. Paris Sér. A-B 269*, A721–A723.

[S27] Sklar, A. (1969). Canonical decompositions, stable functions, and fractional iterates. *Aequationes Math. 3*(1–2), 118–129.

[S28] Erber, T., Schweizer, B., and A. Sklar (1971). Probabilistic metric spaces and hysteresis systems. *Comm. Math. Phys. 20*(3), 205–219.

[S29] Schweizer, B. and A. Sklar (1971). Mesure aléatoire de l'information et mesure de l'information par un ensemble d'observateurs. *C. R. Acad. Sci. Paris Sér. A-B 272*, A149–A152.

[S30] Erber, T., B. Schweizer, and A. Sklar (1973). Mixing transformations on metric spaces. *Comm. Math. Phys. 29*(4), 311–317.

[S31] Schweizer, B. and A. Sklar (1973). Probabilistic metric spaces determined by measure preserving transformations. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 26*(3), 235–239.

[S32] Sklar, A. (1973). Random variables, joint distribution functions, and copulas. *Kybernetika 9*(6), 449–460.

[S33] Erber, T. and A. Sklar (1974). Macroscopic irreversibility as a manifestation of micro-instabilities. In Gal-Or, B. (Ed.), *Modern Developments in Thermodynamics*, pp. 281–301. Israel Universities Press, Jerusalem.

[S34] Schweizer, B. and A. Sklar (1974). Operations on distribution functions not derivable from operations on random variables. *Studia Math. 52*(1), 43–52.

[S35] Johnson, P. and A. Sklar (1976). Recurrence and dispersion under iteration of Čebyšev polynomials. *J. Math. Anal. Appl. 54*(3), 752–771.

[S36] Schweizer, B. and A. Sklar (1977). The axiomatic characterization of functions. *Z. Math. Logik Grundlagen Math. 23*(5), 373–382.

[S37] Moynihan, R., B. Schweizer, and A. Sklar (1978). Inequalities among operations on probability distribution functions. In Beckenbach, E.F. (Ed.), *General Inequalities 1 / Allgemeine Ungleichungen 1* (Proc. First Internat. Conf., Math. Res. Inst., Oberwolfach, 1976), I, pp. 133–149. Birkhäuser, Basel.

[S38] Schweizer, B. and A. Sklar (1978). Probabilistic information spaces. *Information Theory* (Proc. Internat. CNRS Colloq., Cachan, 1977), pp. 485–494, Colloq. Internat. CNRS, 276, CNRS, Paris, 1978.

[S39] Sklar, A. (1978). The embedding of functions in flows. *Aequationes Math. 17*(1), 367.

[S40] Sklar, A. and B. Schweizer (1978). Unsolvability of the general "problème réciproque" for probabilistic and random information spaces. *Information Theory* (Proc. Internat. CNRS Colloq., Cachan, 1977), pp. 495–501, Colloq. Internat. CNRS, 276, CNRS, Paris.

[S41] Rice, R.E., B. Schweizer, and A. Sklar (1980). When is $f(f(z)) = az^2 + bz + c$? *Amer. Math. Monthly 87*(4), 252–263.

[S42] Schweizer, B. and A. Sklar (1980). How to derive all L_p-metrics from a single probabilistic metric. In Beckenbach, E.F. (Ed.), *General Inequalities*, 2 (Proc. Second Internat. Conf., Oberwolfach, 1978), pp. 429–434, Birkhäuser, Basel.

[S43] Erber, T., T.M. Rynne, and A. Sklar (1981). Quantum mechanics and mixing transformations. *Acta Phys. Austriaca 53*(3), 145–155.

[S44] Schweizer, B. and A. Sklar (1983). *Probabilistic Metric Spaces*. North-Holland Publishing Co., New York.

[S45] Schweizer, B. and A. Sklar (1983). All trapezoid functions are conjugate. *C. R. Math. Rep. Acad. Sci. Canada* 5(6), 275–280.

[S46] Sklar, A. (1984). Confluence and stability of orbits of quadratic polynomials. *C. R. Math. Rep. Acad. Sci. Canada 6*(5), 291–296.

[S47] Sklar, A. (1984). On some unexpected theoretical and practical applications of functional equations. In Aczél, J. (Ed.), *Functional Equations: History, Applications and Theory*, pp. 55–56. Reidel, Dordrecht.

[S48] Sklar, A. (1984). The cycle theorem for flows. In Aczél, J. (Ed.), *Functional Equations: History, Applications and Theory*, pp. 227–229. Reidel, Dordrecht.

[S49] Sklar, A. (1984). A problem concerning special functions. In Walter, W. (Ed.), *General Inequalities 4: In Memoriam Edwin F. Beckenbach 4th International Conference on General Inequalities, Oberwolfach, May 8–14, 1983*, pp. 425–426. Springer, Basel.

[S50] Schweizer, B. and A. Sklar (1985). Continuous functions that conjugate trapezoid functions. *Aequationes Math. 28*(3), 300–304.

[S51] Schweizer, B. and A. Sklar (1986). Probability distributions with given margins: Note on a paper by P.D. Finch and R. Groblicki ["Bivariate probability densities with given margins," *Found. Phys. 14*(6), 549–552]. *Found. Phys. 16*(10), 1061–1064.

[S52] Alsina, C. and A. Sklar (1987). A characterization of continuous associative operations whose graphs are ruled surfaces. *Aequationes Math. 33*(1), 114–119.
[S53] Sklar, A. (1987). The structure of one-dimensional flows with continuous trajectories. *Rad. Mat. 3*(1), 111–142.

[S54] Schweizer, B. and A. Sklar (1988). Invariants and equivalence classes of polynomials under linear conjugacy. In Dorninger, D. (Ed.), *Contributions to General Algebra 6: Dedicated to the Memory of Wilfried Nöbauer*, pp. 253–257, Hölder–Pichler–Tempsky, Vienna.

[S55] Schweizer, B. and A. Sklar (1990). The baker's transformation is not embeddable. *Found. Phys. 20*(7), 873–879.

[S56] Alsina, C., B. Schweizer, and A. Sklar (1993). On the definition of a probabilistic normed space. *Aequationes Math. 46*(1–2), 91–98.

[S57] Sklar, A. (1996). Random variables, distribution functions, and copulas – a personal look backward and forward. In Rüschendorf, L., B. Schweizer, and M.D. Taylor (Eds.), *Distributions With Fixed Marginals and Related Topics*, pp. 1–14. Inst. Math. Statist., Hayward CA.

[S58] Alsina, C., B. Schweizer, C. Sempi, and A. Sklar (1997). On the definition of a probabilistic inner product space. *Rend. Mat. Appl. (7)17*, 115–127.
[S59] Alsina, C., B. Schweizer, and A. Sklar (1997). Continuity properties of probabilistic norms. *J. Math. Anal. Appl. 208*(2), 446–452.
[S60] Schweizer, B. and A. Sklar (1997). Copula. In Hazewinkel, M. (Ed.), *Encyclopedia of Mathematics*, *Supplement 1*, pp. 199–201. Springer, Berlin.

[S61] Ghebremeskel, K. and A. Sklar (1999). Ergodic automorphisms of the n-torus are not embeddable. *Proceedings of the European Conference on Iteration Theory (ECIT 1996)*, pp. 189–194. Karl-Franzens-Universität, Graz, Austria.

[S62] Sklar, A. and J. Smítal (2000). Distributional chaos on compact metric spaces via specification properties. *J. Math. Anal. Appl. 241*(2), 181–188.

[S63] Schweizer, B., A. Sklar, and J. Smítal (2000/01). Distributional (and other) chaos and its measurement. *Real Anal. Exchange 26*(2), 495–524.

[S64] Sklar, A. (2002). On certain dense sets of rationals, simultaneous Schröder and Böttcher equations, and characterizations of functions. *Aequationes Math. 64*(3), 232–247.

[S65] Balibrea, F., B. Schweizer, A. Sklar, and J. Smítal (2003). Generalized specification property and distributional chaos. *Internat. J. Bifur. Chaos 13*(7), 1683–1694.

[S66] Riedel, T., M. Sablik, and A. Sklar (2005). Polynomials and divided differences. *Publ. Math. Debrecen 66*(3–4), 313–326.

[S67] Schweizer, B. and A. Sklar (2008). More on j (Letter to the Editor). *Math. Intell. 30*, 4.

[S68] Fatkić, H., A. Sklar, and H. Fatkić (2015). Dispersion under iteration of strongly mixing transformations on metric spaces. *Intern. J. Appl. Math. Electron. Comput. 3*(3), 150–154.

[Also published in Bay, O.F. and I. Saritas (Eds.), *Proceedings of the International Conference on Advanced Technology & Sciences (ICAT'14)*, pp. 772–777.]

Other professional writings by Abe Sklar

[S69] Sklar, A. (1964). On category overlapping in taxonomy. In Gregg, J.R. and F.T.C. Harris (Eds.), *Form and Strategy in Science*, pp. 395–401. Reidel, Dordrecht.

[S70] Sklar, A. (1983). Book review of *Topics in Iteration Theory*, by György Targonski, Studia Mathematica, Skript 6, Vandenhoeck & Ruprecht, Gottingen and Zürich, 1981, 292 pp., (kart. DM 45, – –), ISBN 3-5244-0126-9. *Bull. Amer. Math. Soc. 9*(3), 345–348.

[S71] Kass, S., B. Schweizer, A. Sklar, and A. Alt (1997). Interview of Franz Alt, New York, May 17, 1997.

Appendix 2: Sklar's interview with Bonnie Shulman

Below is a transcription of an interview of Abe Sklar conducted and taped in August 2006 by Bonnie Shulman, now a Professor Emerita of Mathematics at Bates College in Lewiston, Maine. This document, stored in the Duke University archives [42], is published here with her permission. The footnotes, which contain additional information and clarifications, are mine.

This is Bonnie Shulman (BS). It's August 25, 2006. I'm talking with Abe Sklar (AS) in Chicago, IL.

BS: So you worked after you got out of the army, you went to school at the University of Chicago, and from your major in physics you moved on to a major in mathematics.

AS: Mmm, hmm, and got a master's degree and then it turned out there was nobody on the faculty who was doing things that really interested me and I was not by any appearances going to be a mathematical genius, so nobody on the faculty particularly wanted me as a graduate student. So I hung around for a while and then we had a visiting professor who was interested in number theory. He and I got along very well and one result of that was that I had my first published paper. It was in number theory and appeared the *Notices of the AMS*.[1] And then although he left, another visiting professor came in and this was a rather famous person, Professor V.J. Rajagopal from India who at that time was the head of the Ramanujan Institute.[2] He and I got along very well and he encouraged me, and then he had again a very well-known colleague drop in for a short time – Professor Chowla[3] – and Chowla asked me about my interests and my work, and I told him of another result that I had – different from the published one, and he said "oh, that's very beautiful."

Anyway, so as it happened one of the more senior graduate students had gotten his PhD and had gone to Cal Tech and was working on the so-called Bateman Project. He wrote me and said, would you be interested in coming down to work on the Bateman Project? I said yes and so, anyway, I went to Mac Lane[4] and said there's this possibility and he essentially said "go for it" and I'm pretty sure he wrote me a good recommendation – if only to get rid of me!

So anyway, I went to Cal Tech and as it turned out they were in the process really of finishing off the Bateman Project so they didn't really need anyone at the time. So instead I got a – actually I guess it is called a fellowship, a research fellowship. So I didn't have teaching duties and I was working with Tom Apostol and I was one of his first two PhD students.[5] Of course he was not that much older than I was.[6] I got my PhD with

1 This is paper [S1] in Sklar's bibliography; see Appendix 1.

2 The Ramanujan Institute of Mathematics was founded in 1950. Here, Sklar presumably refers to C.T. – not V.J. – Rajagopal (1903–78), who joined the institute in 1951 but only became its director in 1955.

3 Sarvadaman D.S. Chowla (1907–95) was a British-born Indian American number theorist.

4 Saunders Mac Lane (1909–2005) was an American mathematician who co-founded category theory with Samuel Eilenberg. He was Chair of the Department of Mathematics at the University of Chicago from 1952 to 1957.

5 The other was Basil Gordon (1931–2012), a professor at UCLA, specializing in number theory and combinatorics.

6 Apostol was born August 20, 1923, which made him 820 days older than Abe Sklar.

him and of course had the chance to meet and interact with some really top notch mathematicians. Well, Cal Tech is a small place and the math faculty was small.

BS: And your first job was at IIT?

AS: My family, both my parents were in Chicago and at that time it was very easy to get jobs.

BS: Those were the days!

AS: So I went to IIT and actually I had contact with IIT before, because a lot of graduate students at Chicago were hired as teaching assistants at IIT. Not at Chicago, of course.

BS: So you began there in 1956 and that was the year Menger had his first thesis student, Berthold Schweizer?

AS: Well of course he'd had some in Vienna and he had students at Notre Dame.[7] I believe Joe Landin[8] was one of his students at Notre Dame. Rufus Isaacs[9] may have been another. He was at Notre Dame and certainly had contact with Menger. He's listed in what is now a quite famous paper.[10] He answered a question that Menger had raised, presumably either in class or in a general lecture, namely to find – if such a thing exists – a function, real function, whose second iterate was the negative of the identity.

BS: Kind of like a *i*, like a square root of −1 or something.

AS: Well if you do it in the complex plane, there's no problem, but if you want a real function...

BS: Right, that's what's so hard.

AS: Well, it turns out that there are in fact lots and lots of real functions that do it. None of them, of course, are continuous or monotonic. You can get piecewise continuous, but not – you have to have discontinuities.

BS: So he answered that.

AS: Yes.

BS: So, I am most interested in, or I'm going to try and focus on Menger in America. And I have this quote from a biography that van Dalen wrote of Brouwer.[11] There are two volumes.[12] The second volume just came out and I reviewed it. I reviewed both volumes, and I've had some correspondence with van Dalen about this. And there's a short paragraph I wanted to read you on Menger from that biography:

> "Menger's fate was typical of the middle European scientist of the pre-War era, highly successful, versatile, cultured, with many contacts, who through a forced emigration suddenly became separated from his scientific cultural background and lived on in relative obscurity. He would never reach the heights of his European period again."

That's van Dalen speaking.

AS: Mmm, hmm, mmm, hmm.

BS: So a question that I have for you is, I think it's not to be disputed that although he may have done – and I think he did do much fine mathematical and other kinds of work while here – he never did recapture the – I don't know what to call them – the glory days in the Vienna Circle. He didn't achieve the same prominence here that say, von Neumann...[13]

AS: No?

BS: ... or Teller[14] or you know, some of the other émigrés did.

AS: Mmm, hmm.

7 The University of Notre Dame is a famous private Catholic research university located near South Bend, Indiana.

8 Joseph Landin (1913–2000) was a professor of mathematics at the University of Illinois at Chicago who is best known for his textbooks on algebraic structures and set theory. He completed his PhD under Karl Menger's tutelage in 1946.

9 Rufus Isaacs (1914–1981) was a game theorist, prominent in the 1950s and 1960s with his work on differential games.

10 The paper in question appears to be [17].

11 L.E.J. Brouwer (1881–1966) was a Dutch mathematician and philosopher, who is known as the founder of modern topology; a famous fixed-point theorem bears his name.

12 See entries [44, 45] in the bibliography.

13 John von Neumann (1903–57) was a Hungarian-American mathematician, physicist, computer scientist, engineer and polymath. He is generally regarded as the foremost mathematician of his time.

14 Edward Teller (Hungarian: Teller Ede), born in 1908, deceased in 2003, was a Hungarian-American theoretical physicist who is known colloquially as "the father of the hydrogen bomb."

BS: And that was true for some other émigrés as well. But I just wonder what you had to say about that, and why you think that might be the case.

AS: Well I think it's basically as van Dalen says, that he had lost a very favorable milieu and of course it was a time and a place where many exciting things were going on and in fact of course before the Nazi induced immigration, the United States just didn't compare with Europe in the sciences in general, and I'd say in mathematics in particular, although there were some very brilliant people in the United States. In fact in a sense Menger adjusted to the United States much better than probably most of the European émigrés. He liked the United States, I think he liked South Bend,[15] although he certainly liked Chicago more.

I know one very well-known mathematician came to the States from Frankfurt, of course a full professor, and went to Urbana,[16] and OK so he was professor at Urbana, but being professor at Urbana was not the same thing as being full professor in Frankfurt. And in particular he really resented the fact that he was asked to teach trigonometry. Whereas Menger started writing about his experiences at IIT and he says then that when he came to IIT, I think after the first semester he was asked – very reluctantly –

BS: To teach at night school?

AS: Yes!

BS: I've read some of those reminiscences from his papers.

AS: Aha!

BS: And nobody else wanted to do it, and he in fact embraced it; he enjoyed that population.

AS: Yes he did. And he kept on enjoying it. And teaching so-called elementary courses.

BS: So do you think that some of this was that he put his heart and soul and energy into the teaching and put a lot of thought – somewhat maybe like Lagrange[17] when he ended up writing his book based on his experience trying to teach young cadets? Do you think there's maybe some analogy there, with how Menger put his creative energies, his thinking about mathematics, influenced by his heavy teaching load? Didn't he train a lot of Navy cadets?

AS: During the war, yes. B-12 cadets at Notre Dame. Very large classes, presumably more than one of them. And the idea was to teach them quickly, cause they really weren't supposed to spend the entire duration of the war in college. Anyway, but one thing was that it would have made a difference, say if he had gone from Europe to Harvard, Yale, or Princeton.

BS: As von Neumann, for instance, did.

AS: Yes. But as it was, he was at Notre Dame, and then at IIT, where – at Notre Dame of course he tried to revive the Colloquium, not the Circle, but the –

BS: The Mathematics Colloquium.

AS: Yes.

BS: He published –

AS: Yes. But of course it just wasn't the same. And of course as Bert Schweizer says, "how could it have been?"

BS: How did he make that transition? Was he disappointed? Did he regret that for the rest of his life? Did he move on with that? How was that for him?

AS: I don't get the feeling that he regretted that.[18] He was still of course in contact with a great many people, particularly at Notre Dame. After all, one of the Notre Dame lecture series was the original place for Artin's Galois Theory.[19] He was never one for doing what was fashionable. So he just did things that he wanted to do.

15 South Bend, Indiana, where Menger worked at the University of Notre Dame from 1937 to 1946.

16 Urbana is a town in the state of Illinois which shares the campus of the University of Illinois at Urbana–Champaign with its sister city of Champaign.

17 Joseph-Louis Lagrange (1736–1813), famous Italo-French mathematician and astronomer.

18 Nevertheless, Bert Schweizer (private comm.) reported that after the war, Menger asked Universität Wien to be reinstated. As he had resigned of his own, however, his request was turned down, which deeply disappointed him.

19 Reference to the book by Emil Artin [3] published in the Notre Dame Lecture Series.

BS: And he dug into the work that was in front of him.

AS: Yes. If he'd had enough students of the caliber that he would have had, had he gone to Harvard or Princeton, say, then it might have been a different story, because he would really have nurtured the sort of students that he had had in Vienna.

BS: But it doesn't sound like he begrudged –

AS: No.

BS: He dealt with the students he had.

AS: He never complained. He never begrudged anything. Unlike the professor I mentioned –

BS: Who shall remain nameless.

AS: Oh I think you could find out easily. In fact you probably already know.

BS: I suspect.

BS: So, what was he like as a colleague? So you came there, not young chronologically, but a young – newly minted – PhD. How close were you in age to him, how much senior to you was he?

AS: Well, basically, he was 24 years older.[20]

BS: I haven't done the math, is why I asked. So yes, he was quite a bit older – was he a mentor to you?

AS: Not when I just came to IIT because for a while, I did work in number theory. And in fact I did write a paper.[21] But actually what happened was that I'd come to IIT, I was getting settled in, when one of the other instructors came, and that was Bert Schweizer,[22] and he was looking for people to work with him on the subject of what we later called probabilistic metric spaces – at the time it was statistical metric spaces – which was something that Menger had started actually. And anyway, the subject intrigued me, so we started working together. And it was through Bert that I really got in contact with Menger. And then since...

(END OF TAPE 1, SIDE 1)

[Bert] left Chicago basically for healthier climates that turned out not to really be not really that much healthier. Let's see; he went to Los Angeles.[23] Anyway, so I remained the one person physically closest to Menger. So of course we kept up contact and...

BS: So, what was he like? I heard him described when he was younger as various things, from brilliant, ambitious, introverted, some mixture... some of those seem contradictory to each other.

AS: Yes, in fact I think those characterizations are from van Dalen. Van Dalen corresponded with Bert while writing the second volume, essentially because of the connection of Menger through Brouwer. So I got actually to read some of the biography before it was published. And I remember that characterization, which Bert particularly said wouldn't have called – nor would I have called Menger introverted. Actually he was, if anything, quite the opposite.

BS: Could this have been different when he was younger?

AS: Uh, I don't think so. I got the impression that he was a rather brash young man.

BS: Yes, that came through in van Dalen's treatment of Brouwer as well. They clashed.

AS: Yes. Because after all he was the son of a quite famous person[24] and –

BS: And he was smart.

AS: Yes. And he would have had enough self-confidence and self-assurance to be able as a beginning college student to go and talk to a professor after the first lecture. So no, I wouldn't call him introverted at all.

BS: It sounds like he was very kind.

AS: He could be very kind. He could also get angry and be very angry, but it didn't last.

---

20 Karl Menger was born on January 13, 1902, and Abe on November 17, 1925; so their age difference is exactly 23 years, 10 months, 4 days (or 8709 days in total) excluding the end date.

21 This is entry [S3] in Sklar's bibliography.

22 After completing his PhD under Menger's tutelage, Bert Schweizer was an instructor at IIT in 1956–57.

23 Bert Schweizer was an Assistant Professor at San Diego State College in 1957–58 and at UCLA from 1958 to 1960.

24 Karl Menger's father was the famous Austrian economist Carl Menger (1840–1921), who is regarded as the founder of the Austrian school of economics.

BS: So, as well as not begrudging, he also didn't hold a grudge?

AS: Right. Olga Taussky,[25] who of course had known Menger very well, once remarked that with him "everything is either plus infinity or minus infinity."

BS: So he was effusive. And he was a man maybe of extreme passions or emotions, felt strongly about issues.

AS: Yes he did.

BS: And did he change his mind, was he willing to change his mind about things? Was he rigid?

AS: He could be, I'd say if it was in connection with something that he himself had, say, originated. But if he wasn't really –

BS: – invested, his self wasn't invested in it –

AS: Yes.

BS: So you could imagine his fight with Brouwer,[26] how that –

AS: Yes, yes. Well they were, let me find another adjective – I'd say both Brouwer and Menger were very touchy. That was an adjective I first heard applied to Menger by John Dawson,[27] who edited the Gödel papers, and who had interviewed Menger in connection with Gödel.[28]

BS: So I imagine the reaction to his calculus textbook[29] must have disappointed him – did it anger him? I mean this was something he was very much invested in –

AS: Very much so.

BS: And it never really took off.

AS: Right. Not only of course did he himself put a lot of work into it, but he got his children involved.[30]

BS: In helping proofread it?

AS: Oh yes.

BS: Right, no word processors in those days.

AS: And in actual physical production because the first editions were mimeographed.

BS: I've seen one of those copies, yes.

AS: And you can visualize the kids doing that. Also the – I don't know who did secretarial work for him. Because although he would type himself, he was not a particularly good typist and often he would type something and then change it, sometimes make handwritten changes, or sometimes take typed pages and cut and paste. This meant that actually converting something like that into a decent manuscript took a great deal of work. So although probably the department secretary did part of it, I wouldn't be surprised if the kids also got in on that.

BS: I'll have to ask them. So you did say you thought it was going to be reissued by Dover?

AS: Yes.

BS: That's interesting. Who's responsible for that happening?

AS: Well, it's basically again Bert who – Bert and I of course had published a book on probabilistic metric spaces in '83 and it went out of print.[31] Then a few years ago Bert thought what the hell, I'll see if I can get Dover to reprint it. Which he did, it's been reprinted.[32] Anyway, apparently Dover likes it, one reason being that of course we added material to it and that was done by a – the actual production, of course everything was photocopied by Dover – the actual production of the notes and the supplementary bibliography and the

25 Olga Taussky-Todd (1906–95) was a Czech-American mathematician, famous for her abundant work in algebraic number theory, integral matrices, and matrices in algebra and analysis. She completed her PhD at Universität Wien in 1930 under the supervision of Philipp Furtwängler (1869–1940). Menger was a professor there as of 1927.

26 For additional information about the relationship between Brouwer and Menger from the latter's perspective, see [21], pp. 5–6.

27 John W. Dawson Jr. (1944–), Professor of Mathematics, Emeritus at Pennsylvania State University at York.

28 Reference to the monumental work of the logician, mathematician, and philosopher Kurt Gödel (1906–78).

29 See entries [26, 27] in the bibliography.

30 In 1935 Karl Menger had married Hilda Axamit, an actuarial student. They had four children. Karl Jr., born in 1936, Rosemary and Fred, twins born in 1937, and Eve, born in 1942.

31 This is reference [S44] in the bibliography.

32 It appeared in 2005.

preface and so on were done by a lady in California – very very good – in fact so good that Dover has used her in other projects. So anyway, Bert finally had the idea of approaching the editor at Dover and – or maybe it was really the other way around, that the editor said, you know "I looked at Menger's thing and thought that it might be a worthwhile thing for Dover to reprint," and whichever way it went. So there is an agreement although contracts have yet to be – probably will be in a month or so.[33]

BS: How do you think it will fare in this current market? Do you think it will kind of be a specialty item? Do you anticipate anyone actually using it as a calculus text?

AS: Probably not, although it's been done.

BS: I noticed in his papers, for instance, there were classes being taught at IIT, and there was a notice of where to obtain it at the bookstore. So it was clear that at least he – and I don't know if any other instructors – had actually used it, there.

AS: Yes, yes. He used it there, and there are people, in particular one of my PhD students, who used it as a text in other places. Unfortunately, he died recently, so he could have told you more about [how] it worked as a text.[34] He thought it worked very well. If you were –

BS: – on board with the program.

AS: Yes. The thing with that as a text is that of course it is different enough so that using it as a text, it would be very hard to deviate from it. And of course it would be just as hard, maybe harder, to use it as a supplement to a standard text.

BS: So it's not clear who the audience would be.

AS: I would say the audience resides in, besides people interested in it as a historical document, would be some future writers of calculus textbooks.

BS: It would be – I think he had some good ideas.

AS: Very much so.

BS: It would be nice to see them get back into the literature.

AS: Yes.

BS: What role do you think his earlier – the impact of the intellectual issues in the Vienna Circle, you know the emphasis on clarity of language, clarity of thought, and so on, had on him. I know he also critiqued many of the discussions in the Vienna Circle, but how the emphasis on clear thought and expression, how did that play out in his later work, perhaps in the calculus text and in other ways?

AS: I think there was a very definite influence. As you say, striving to be clear, which means in mathematics in particular, that you say essentially clearly, what you are assuming and taking for granted and if you – if you're doing something like writing a text – you're saying well, as essentially he does with the mean value theorem, this is an important notion, but at this point you don't really have enough background to do it properly.

BS: So it's kind of an honesty.

AS: Yes. Probably honesty is a better term than clarity. Yes. I mean, of course as far as clarity goes, don't confuse notions which are different.

BS: So, making distinctions – I mean that's what language is really about, categorizing, naming things. Thinking of Wittgenstein's influence perhaps there.[35] But that it's very important that you say what you mean and that you aren't fuzzy about things that are really different.

AS: But of course besides this there are things in the calculus book, aside from the general mode of presentation, which are very beautiful mathematical things that you don't find in other places. The graphical composition.

BS: So there might be parts of this [that] could be used out of the context of the entire book.

33 The Dover edition, published in 2007, features a new preface and guide to further reading by Bert and Abe.

34 This is Joe Harkin (1926–2006), who taught mathematics for 32 years at SUNY Brockport; see Section 3.

35 Ludwig Wittgenstein (1889–1951) was an Austrian-British philosopher who worked primarily in logic, the philosophy of mathematics, language and mind. He is considered to be one of the greatest philosophers of the 20th century.

AS: Oh yes, very much so. One thing that he did get from the Circle, but of course also to some extent from the general philosophic milieu of the time, applies more to his attitude towards science than mathematics as such – was a positivistic attitude, where you try not to deal with things that are not concrete. And this of course is out of favor in science.

BS: Today, you think?

AS: Yes and of course part of this comes from Mach.[36] Menger had great – well, the whole Circle did – great respect for Mach. But you could say this got Mach into trouble, scientifically. Mach, until his dying day, refused to believe in such things as atoms.

BS: So I know metaphysics was a bad word.

AS: Yes, in the Circle.

BS: And Wittgenstein saying, you know, if we can talk about it, it's not important. You know, the really important stuff is what we can't even talk about.

AS: Of course Menger wouldn't agree with that.

BS: Yes, that's right he definitely did not agree with that. Given that, though, I don't know if this is a relevant question or not, but would you say that Menger had a spiritual life? Did he have any part of his life that was intangible, not concrete?

AS: Well, inasmuch as (laughs) inasmuch as mathematics is intangible? But otherwise, neither he nor the family were say religious in any sense. He was interested in theology.

BS: From a philosophical point of view?

AS: Yes, I would say so. And in philosophy he had read Kant,[37] and although I don't think he agreed with most of what Kant said, but he had read him.

BS: How about Spinoza, did he read Spinoza?[38]

AS: I don't remember him ever mentioning Spinoza. He may very well have.

BS: The only reason I'm thinking of that is, well first of all I recently read Rebecca Goldstein's book *Betraying Spinoza*[39] but also his – Spinoza's Ethics where he talks about –

AS: Oh, I'm sorry, yes, in fact he – a very long – he had read Spinoza and in fact he criticizes Spinoza, and in particular his attempt to treat ethics –

BS: – as Euclid's postulates. Is that in his Morality – I might have read that in –

AS: I don't think it's there. Let me see, I (rummages in bag). Do you know this?

BS: Yes, I have a copy of that, the *Reminiscences*.[40] So maybe that's where he talks about it?

AS: He talks about Spinoza at some length, I don't know if it's here.

BS: Where else might it be? Published, something published?

AS: Yes, p. 117.

BS: OK, so it's in the *Reminiscences*.

(END OF SIDE 2, TAPE 1)

BS: So how did you come – were you the last person to have his papers. And you worked on editing some of his work. How did that come about?

AS: The person who essentially had a complete collection was Bert Schweizer. And Menger had talked to Bert, oh, probably something like 10 years before his death, picking out the papers that he would like to see in a *Selecta*. And as I say this was some time before his death. And then he did go ahead and publish the *Selected*

36 Ernst Mach (1838–1916) was an Austrian physicist and philosopher who studied shock waves, inter alia, and who influenced logical positivism. His criticism of Newton's theories foreshadowed Einstein's theory of relativity.

37 Immanuel Kant (1724–1804) was a German philosopher and one of the central Enlightenment thinkers. He is considered one of the most influential figures in modern Western philosophy.

38 Baruch (de) Spinoza (1632–77) was a Dutch philosopher of Portuguese Sephardi origin. He is regarded as one of the great rationalists of 17th-century philosophy.

39 The book's full title is *Betraying Spinoza: The Renegade Jew Who Gave Us Modernity* [15].

40 See entry [29] in the bibliography.

Papers, what do you call it, in logic, didactics, and so on...[41] So this was a selection from the *Selection*. So he did that. And then in his will he mentioned some things that he wanted to have published, and one was a revised and expanded version of *Calculus*, one was a *Selecta*, and one was a biography of his father which he started but never got very far into and the last one was, I could look it up, but I don't remember at the moment. Of course he would have liked to have had those things published within 10 years of his death. But that was quite impossible. But of course now the *Selecta* has appeared and I think very well done. Because we got some really top notch people to do commentaries, including van Dalen.

BS: So I don't know if I'm aware of this volume...

AS: Two volumes, published by Springer.

BS: Called?

AS: *Selecta Mathematica*. Karl Menger: *Selecta Mathematica*. And for Springer books, well, I'll show you, for Springer books, they're not astronomically expensive.

BS: Great!

AS: Which is maybe why Springer hasn't pushed them as much as we think they should.

BS: So you were instrumental in seeing that?

AS: Yeah, actually the most instrumental person was Professor Karl Sigmund[42] in Vienna. But he and Bert and I basically were the editors. We got the other people to write the commentaries.

BS: What about the biography of his father, what's up with that?

AS: He started it, as I say, he couldn't have gotten very far, if really anything still exists, it would probably be with Rosemary or Eve who would have it.[43] Of course he also started an autobiography, and again he didn't get too far into that, but I have a copy of that as far as it goes. He doesn't really get out of his childhood.

BS: So, you were close to him in later years. You guys developed a relationship over time?

AS: Yeah. He of course had developed a heart problem in his life and in '83 I got a phone call from him, I think it was close to midnight.[44] So I went up to the place and we talked and finally convinced him that he should go to a hospital. And I took him to Swedish Covenant.[45] So he went into the Emergency Room, I think it was about 3 in the morning. And he was there for an hour or two. Finally they officially admitted him. I left the hospital and in the morning I called Rosemary.[46] Anyway he was in the hospital for something like a week, then was released and he went to stay with Rosemary and Dick in Highland Park; he was there for two weeks or so and apparently was doing all right, and then died in his sleep.

BS: So you were the first person that he called...

AS: Yes. Well, I was close.

BS: Physically, geographically close. But also... yeah. I guess, I mean I could ask you a hundred more questions, but those are pretty much the ones I had written down. What kinds of materials did you bring along?

AS: Basically just that book (*Reminiscences*).

BS: And we did consult it!

AS: And a reminder of the other books that I have related to Menger. Let's see, well beside the *Selecta* and the *Selected Papers*, there is the *Ergebnisse Kolloquiums*[47] which has been republished by Springer and that was essentially Sigmund's doing. That is very worthwhile. In connection with that I know mathematicians in Graz[48] (laughs) I have another thing to say about Graz – but anyway, one of the leading mathematicians there found out that I had copies of the *Ergebnisse eines Mathematischen Kolloquiums* and he told me that they didn't have anything in Graz, because apparently the Nazis had taken care of that. So I sent him a set and when

41 The exact title is *Selected Papers in Logic and Foundations, Didactics, Economics*; see [28].

42 Karl Sigmund (1945–) is a Professor of Mathematics at Universität Wien and a pioneer of evolutionary game theory.

43 These are Menger's two daughters already mentioned in footnote no 30.

44 Here, Sklar actually means 1985 as Karl Menger died on October 5, 1985.

45 Swedish Hospital (formerly Swedish Covenant Hospital) is a teaching hospital located on the north side of Chicago.

46 This is Karl Menger's first daughter; her husband Dick Gilmore is mentioned in the subsequent sentence.

47 See entry [30] in the bibliography.

48 Graz is the second largest city in Austria after Vienna.

he got it he wrote back and he said "we were very pleased to get this." He said, "Vienna doesn't have one." So obviously that's one reason why Sigmund (who's at Vienna) wanted to get the *Ergebnisse* republished, which they have done, very nicely.

I mentioned Graz – I remember one thing he said in connection with Graz, and that was about Schrödinger.[49] Schrödinger had been at Cambridge[50] and ultimately had to leave Cambridge basically because of the two families.[51] So he went back to Austria, but not to Vienna, he went to Graz. Again, apparently because Vienna may have – this was in the middle 30s – basically again because Vienna was a bit puritanical. So he went to Graz and Menger said that – who apparently wasn't aware of Schrödinger's –

BS: – unconventional arrangements?

AS: Unconventional arrangements. He said it was strange that he went to Graz because although Vienna at the time was 80% Nazi, Graz was 100% Nazi. And of course after the Anschluß, Schrödinger was in very deep trouble in Austria and he was essentially rescued out of Austria by de Valera in Ireland,[52] who set up the Institute for Advanced Study in Dublin, especially for Schrödinger.[53]

BS: So, was Menger charismatic? I mean he seems to have inspired devotion from a number of people, seeing that his legacy continues, and in his lifetime as well. What was it about him that inspired that?

AS: I think one reason for that is that he was always interested in what you were doing. And it didn't matter, it could be – of course this is in connection primarily with mathematics – but in other things as well. And it didn't matter whether you were a student or professor or whatever. When he'd meet someone, even after only a day, he'd say, "Well, what's the news?" As if to say, well what fabulous discoveries have you made? And of course he could be very charming. When he got angry, of course, he could be very un-charming.

BS: And did you ever have an altercation with him? Were you ever the recipient of his anger?

AS: Yes. I think everybody at some point. But they were rare.

BS: And it would flare up and then subside.

AS: Yeah, yeah.

BS: Well, thank you.

AS: Oh, you're very welcome.

Acknowledgments and credits. People who graciously shared personal memories about Abe Sklar with me include Claudi Alsina, Philip Calabrese, Thomas Erber, Jerry Frank, Jerry Jurschak, Clark Kimberling, Arthur Lubin, Roger Nelsen, Ty Olsen, Mary Ranade, David Schweizer, Marjorie Senechal, and Victor Terrana. May they find here the expression of my gratitude for their availability and candid comments. I am also indebted to Bonnie Shulman for permission to include, and annotate, her interview of Abe Sklar. Additional thanks go to Carlo Sempi and David Schweizer, who sent me photographs, and to the Czech newspaper *Opavský a Hlučínský Deník* for authorizing the use of Figure 3.

References

[1] Alsina, C., M.J. Frank, and B. Schweizer (2006). *Associative Functions: Triangular Norms and Copulas*. World Scientific Publ., Hackensack NJ.

[2] Anonymous (1981). Sidney Penner – In memoriam. *The Two-Year College Math. J. 12*(1), 69.

49 Erwin Schrödinger (1887–1961) was a Nobel Prize-winning Austrian physicist who developed fundamental results in quantum theory: the Schrödinger equation gives a way of computing a system's wave function and its dynamic change.

50 Here, Sklar seems to have mixed up Cambridge and Oxford. Schrödinger had a position at Oxford in 1933–34.

51 Schrödinger shared living quarters with his wife and his mistress, a fact that did not meet with acceptance.

52 Éamon de Valera (1882–1975), who was trained as a mathematician, is an Irish statesman who was, inter alia, President of the Republic of Ireland.

53 The Dublin Institute for Advanced Studies is a statutory, independent research institute which was established in 1940 by Éamon de Valera, who was then the "Taoiseach," i.e., the Prime Minister of Ireland.

[3] Artin, E. (1971). Galois Theory: Lectures Delivered at the University of Notre Dame. In Milgram, A.N. (Ed.), *Notre Dame Math. Lectures 2*, North Bend IN.

[4] Calabrese, P.G. (2017). *Logic and Conditional Probability: A Synthesis*. College Publ., London.

[5] Durante, F., G. Puccetti, and M. Scherer (2015). Building bridges between mathematics, insurance and finance: An interview with Paul Embrechts. *Depend. Model. 3*, 17–28.

[6] Fatkić, H. (2000). *O Vjerovatnosnim Metričkim Prostorima i Ergodičnim Transformacijama*. PhD Thesis, Univerzitet u Sarajevu, Bosnia and Herzgegovina.

[7] Frank, M.J. (1979). On the simultaneous associativity of $F(x, y)$ and $x + y - F(x, y)$. *Aequationes Math. 19*(1), 194–226.

[8] Frank, M.J. (1991). Convolutions for dependent random variables. In Dall'Aglio, G., S. Kotz, and G. Salinetti (Eds.), *Advances in Probability Distributions with Given Marginals. Mathematics and Its Applications*, pp. 75–93. Springer, Dordrecht.

[9] Genest, C. (1987). Frank's family of bivariate distributions. *Biometrika 74*(3), 549–555.

[10] Genest, C. and R.J. MacKay (1986). Copules archimédiennes et familles de lois bidimensionnelles dont les marges sont données. *Canad. J. Statist. 14*(2), 145–159.

[11] Genest, C. and R.J. MacKay (1986). The joy of copulas: Bivariate distributions with uniform marginals. *Amer. Statist. 40*(4), 280–283.

[12] Genest, C. and J.G. Nešlehová (2020). A conversation with Paul Embrechts. *Internat. Stat. Rev. 88*(4), 521–547.

[13] Genest, C. and M. Scherer (2020). The gentleman copulist: An interview with Carlo Sempi. *Depend. Model. 8*, 34–44.

[14] Goldberg, K. (1980). Sidney Penner. *New York State Math. Teachers' J. 30*, 168.

[15] Goldstein, R.N. (2006). *Betraying Spinoza: The renegade Jew who gave us modernity*. Schocken Books/Nextbook Press, New York.

[16] Harkin, J.B. (1972). Sklar–Stieltjes integrals. *Illinois J. Math. 16*(2), 196–202.

[17] Isaacs, R. (1950). Iterates of fractional order. *Canad. J. Math. 2*, 409–416.

[18] Kimberling, C.H. (1972). Emmy Noether. *Amer. Math. Monthly 79*(2), 136–149. [Addendum: *Amer. Math. Monthly 79*(7), 755.]

[19] Kimberling, C.H. (1974). A probabilistic interpretation of complete monotonicity. *Aequationes Math. 10*(2–3), 152–164.

[20] Klement, E.P., R. Mesiar, and E. Pap (2000). *Triangular Norms*. Kluwer, Dordrecht.

[21] Leonard, R.J. (1998). Ethics and the excluded middle: Karl Menger and social science in interwar Vienna. *Isis 89*(1), 1–26.

[22] Li, T.-Y. and J.A. Yorke (1975). Period three implies chaos. *Amer. Math. Monthly 82*(10), 985–992.

[23] Ling, C.H. (1965). Representation of associative functions. *Publ. Math. Debrecen 12*, 189–212.

[24] Menger, K. (1942). Statistical metrics. *Proc. Nat. Acad. Sci. 28*(12), 535–537.

[25] Menger, K. (1949). The theory of relativity and geometry. In Schilpp, P.A. (Ed.), *Albert Einstein, Philosopher-Scientist*, pp. 459–474. Library of Living Philosophers, Evanston IL.

[26] Menger, K. (1953). *Calculus: A Modern Approach*. Second edition. The Bookstore, Illinois Institute of Technology, Chicago IL.

[27] Menger, K. (1955). *Calculus: A Modern Approach*. Ginn and Company, New York.

[28] Menger, K. (1979). *Selected Papers in Logic and Foundations, Didactics, Economics* (Mulder, H.L., Ed.). Springer, Dordrecht.

[29] Menger, K. (1994). *Reminiscences of the Vienna Circle and the Mathematical Colloquium* (Golland, L., B.F. McGuinness, and A. Sklar, Eds.). Springer, Vienna.

[30] Menger, K. (1998). *Ergebnisse eines Mathematischen Kolloquiums* (Dierker, E. and K. Sigmund, Eds.). Springer, Vienna.

[31] Menger, K. (2002). *Selecta Mathematica*, vol. 1 (Schweizer, B., A. Sklar, K. Sigmund, L. Schmetterer, P. Gruber, E. Hlawka, and L. Reich, Eds.). Springer, Vienna.

[32] Menger, K. (2003). *Selecta Mathematica*, vol. 2 (Schweizer, B., A. Sklar, K. Sigmund, P. Gruber, E. Hlawka, L. Reich, and L. Schmetterer, Eds.). Springer, Vienna.

[33] Moynihan, R.A. (1975). *Conjugate Transforms for τ_t-Semigroups*. PhD Thesis, University of Massachusetts, Amherst MA.

[34] Nelsen, R.B. (1999). *An Introduction to Copulas*, Springer, New York.

[35] Puccetti, G. and M. Scherer (2018). Copulas, credit portfolios, and the broken heart syndrome: An interview with David X. Li. *Depend. Model. 6*, 114–130.

[36] Salmon, F. (2012). The formula that killed Wall Street. *Significance 9*(1), 16–20.

[37] Schweizer, B. (1956). *Jacobi Series and the Numerical Solution of Eigenvalue Problems*. PhD Thesis, Illinois Institute of Technology, Chicago IL.

[38] Schweizer, B. (1991). Thirty years of copulas. In Dall'Aglio, G., S. Kotz, and G. Salinetti (Eds.), *Advances in Probability Distributions with Given Marginals. Mathematics and Its Applications*, pp. 14–50. Springer, Dordrecht.

[39] Schweizer, B. (1996). The genesis of the notion of distributional chaos. *Seminario Mat. e. Fis. di Milano, 66*(1), 159–167.

[40] Schweizer, B. and J. Smítal (1994). Measures of chaos and a spectral decomposition of dynamical systems on the interval. *Trans. Amer. Math. Soc. 344*(2), 737–754.

[41] Schweizer, B. and E.F. Wolff (1981). On nonparametric measures of dependence for random variables. *Ann. Statist. 9*(4), 879–885.

[42] Shulman, B.J. (2007). *Interview with Abe Sklar. August, 2006.* Tape transcribed and deposited in Duke University archives, Karl Menger folder, June 2007, Durham NC.

[43] Strehler, A.F. (1963). New mathematics – a controversy. *The Quarterly 9*(4), 14–17.

[44] van Dalen, D. (1999). *Mystic, Geometer, and Intuitionist: The Life of L.E.J. Brouwer, Volume 1: The Dawning Revolution*. Oxford University Press.
[45] van Dalen, D. (2004). *Mystic, Geometer, and Intuitionist: The Life of L.E.J. Brouwer, Volume 2: Hope and Disillusion*. Oxford University Press.
[46] Wald, A. (1943). On a statistical generalization of metric spaces. *Proc. Nat. Acad. Sci. 29*(6), 196–197.
[47] Zadeh, L.A. (1965). Fuzzy sets. *Inf. Control 8*(3), 338–353.

DE GRUYTER Depend. Model. 2021; 9:225–242

Research Article Open Access
Special Issue in memory of Abe Sklar

Claude Lefèvre*

On partially Schur-constant models and their associated copulas

https://doi.org/10.1515/demo-2021-0111
Received March 25, 2021; accepted May 10, 2021

Abstract: Schur-constant vectors are used to model duration phenomena in various areas of economics and statistics. They form a particular class of exchangeable vectors and, as such, rely on a strong property of symmetry. To broaden the field of applications, partially Schur-constant vectors are introduced which correspond to partially exchangeable vectors. First, their copulas of survival, said to be partially Archimedean, are explicitly obtained and analyzed. Next, much attention is devoted to the construction of different partially Schur-constant models with two groups of exchangeable variables. Finally, partial Schur-constancy is briefly extended to the modeling of nested and multi-level dependencies.

Keywords: Schur-constant model, Archimedean copula, partial exchangeability, multivariate monotonicity, bivariate survival functions

MSC: 60G09, 62H05, 62H10

1 Introduction

Schur-constant vectors play a central role in modeling lifetime data in actuarial science, reliability and survival analysis. The case traditionally considered is that where lifetimes are absolutely continuous random variables with values in $\mathbb{R}_+$. Continuous Schur-constant models have been discussed by many authors including [4, 6, 13, 35, 38, 46, 47]. Recently, Schur-constancy for discrete lifetimes with values in $\mathbb{N}_0$ has been studied by [8] and then in [7, 21, 28, 36].

By definition, a positive random vector $\mathbf{X} = (X_1, \dots, X_n)$ is Schur-constant if its joint survival function $\Pr(X_1 \geq x_1, \dots, X_n \geq x_n)$ depends on the vector $(x_1, \dots, x_n)$ only through its sum $x_1 + \dots + x_n$, via a univariate function (generator) S. Thus, the distribution of $\mathbf{X}$ is of the exchangeable type (in the sense of [15]) with a particular expression. By exchangeability, the n variables X_i have the same distribution and the correlations between any pair of variables (X_i, X_j) are all equal. Such a property of symmetry is very restrictive, and often unrealistic, for a large number of real situations. For example, in insurance or finance, a Schur-constant portfolio would necessarily consist of n contracts or assets that are similar in all respects.

To break this symmetry, it is natural to widen the Schur-constant model by making it only partially exchangeable (as defined by de Finetti [16]). This leads us to introduce a generalized more flexible dependency model called partially Schur-constant. More precisely, the vector $\mathbf{X}$ is now partitioned into $m \geq 2$ groups $\mathbf{X}_1$ of size n_1, …, $\mathbf{X}_m$ of size n_m. Its survival function $P(\mathbf{X}_1 \geq \mathbf{x}_1, \dots, \mathbf{X}_m \geq \mathbf{x}_m)$ is then assumed to depend on the vector $(\mathbf{x}_1, \dots, \mathbf{x}_m)$ only through the vector of corresponding sums $(|\mathbf{x}_1|, \dots, |\mathbf{x}_m|)$, via an m-variates generator S. In this way, we create a new model composed of m groups of variables which are homogeneous with different Schur-constant structures and which can be interdependent with different intensity levels. A discrete version of such a vector has been proposed and applied to risk management by Castañer et al. [9]. In

***Corresponding Author: Claude Lefèvre:** Université Libre de Bruxelles, Département de Mathématique, Campus de la Plaine C.P. 210, B-1050 Bruxelles, Belgium, E-mail: Claude.Lefevre@ulb.be

what follows, we focus on the continuous version, that is when $\mathbf{X}$ is an absolutely continuous random vector over $\mathbb{R}_+^n$, as well as the associated copula.

It is well-known that continuous Schur-constant models are closely linked to Archimedean copulas (see e.g. [13, 29, 38]). In the current context, we will establish a connection between continuous partially Schur-constant vectors and an associated family of copulas, said to be partially Archimedean. These copulas have the practical property of having a partially exchangeable distribution instead of being simply exchangeable. We point out that this family of copulas was recently derived by Ressel [42] following his comprehensive papers on the multiple monotonicity of multivariate functions. At the beginning, we will refer moreover to several important results that he obtained. Our approach, which is different, comes from our desire to generalize the property of symmetry in a Schur-constant model. It is therefore intrinsically probabilistic by nature. We also mention a related article by [14] who discussed an Archimedean copula of similar type based on a particular model of multivariate compound distributions.

The paper is organized as follows. In Section 2, we introduce the class of partially Schur-constant models and we present their main properties. In Section 3, we define a family of survival copulas, called partially Archimedean, and we show their close links with the partially Schur-constant vectors. The following four sections concern the construction of different partially Schur-constant models with two groups of exchangeable variables. This step will lead us to consider various classical bivariate survival functions. In Section 4, we choose as generator a survival function which corresponds to a univariate Laplace transform. Several notable bivariate distributions are discussed in detail. In Section 5, this time we take as generator the Laplace transform of a bivariate random vector. Special attention is paid here to a bivariate distribution of gamma type. In Section 6, we use as generator the Gumbel bivariate exponential survival function. It is a more complex case because the survival does not correspond to a Laplace transform. In Section 7, we define a generator from the bivariate Williamsom transform. This construction is however often little convenient as illustrated by two particular models. Finally, we return in Section 8 to the general partial Schur-constancy to briefly show that this modeling covers the usual nested Archimedean copulas and how it can be extended to multi-level dependencies.

2 Partially Schur-constant vectors

In this Section, we introduce the notion of partially Schur-constant (continuous) vector and we present its main properties. Let us start by recalling the simple Schur-constancy (see e.g. Nelsen [38]).

Schur-constant model. A vector $\mathbf{X} = (X_1, \dots, X_n)$ on $\mathbb{R}_+^n$ is Schur-constant if its joint survival function can be written in the form

$$P(X_1 > x_1, \dots, X_n > x_n) = S(x_1 + \dots + x_n), \quad x_1, \dots, x_n \geq 0, \tag{2.1}$$

for some univariate function $S(x) : \mathbb{R}_+ \to [0, 1]$, called generator.

From (2.1), we directly notice that the vector $(X_1, \dots, X_n)$ has an exchangeable distribution of particular structure. As a result, the n variables X_i have the same marginal distributions and their interdependencies are identical. Such properties can considerably limit the practical field of application of Schur-constant vectors. To remedy this, we introduce a more general model, called partially Schur-constant, which is based on the property of partial exchangeability of a random vector ([16]).

Partially Schur-constant model. Suppose that the vector $\mathbf{X} = (X_1, \dots, X_n)$ can be partitoned into $m \geq 2$ sub-vectors $\mathbf{X}_j = (X_{j,1}, \dots, X_{j,n_j})$, $1 \leq j \leq m$, with $n_j \geq 1$ and $n_1 + \dots + n_m = n$.

Definition 2.1. *A vector $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_m)$ on $\mathbb{R}_+^n$ is partially Schur-constant if its joint survival function can be written in the form*

$$P(\mathbf{X}_1 \geq \mathbf{x}_1, \dots, \mathbf{X}_m \geq \mathbf{x}_m) = S\left(|\mathbf{x}_1|, \dots, |\mathbf{x}_m|\right), \quad \text{for all } \mathbf{x}_j \in \mathbb{R}_+^{n_j}, \tag{2.2}$$

for some m-variates generator $S(x_1,\dots,x_m):\mathbb{R}_+^m\to[0,1]$, *and using the notation*

$$|\mathbf{x}_j| = x_{j,1}+\ldots+x_{j,n_j},\quad 1\le j\le m.$$

Clearly, $\mathbf{X}$ is a partially exchangeable vector such that the sub-vectors $\mathbf{X}_j$, $1\le j\le m$, are Schur-constant of univariate generator

$$S_j(x_j) = S(0,\dots,0,x_j,0,\dots,0),\quad x_j\in\mathbb{R}_+. \tag{2.3}$$

In particular, inside each $\mathbf{X}_j$, the variables $X_{j,1},\dots,X_{j,n_j}$ have the same mean μ_j, variance σ_j^2 and correlation coefficient ρ_j (if it exists). They can be computed from the survival function S_j since

$$\mu_j = E(X_{j,1}) = \int_0^\infty S_j(x_j)dx,\quad E(X_{j,1}^2) = 2\int_0^\infty x_jS_j(x_j)dx_j,$$

$$E(X_{j,1}X_{j,2}) = \int_0^\infty\int_0^\infty S_j(x_j+y_j)dx_jdy_j = \int_0^\infty t_jS_j(t_j)dt_j = E(X_{j,1}^2)/2, \tag{2.4}$$

so that, after some calculations,

$$\rho_j = (v_j^2-1)/2v_j^2,$$

where $v_j=\sigma_j/\mu_j$ is the usual coefficient of variation. Note that $\rho_j>0$ is equivalent to $v_j<1$ (the standard deviation of $X_{j,1}$ does not exceed its mean). As for the dependence between two different groups $j<k$, all the pairs of variables have the same correlation coefficient $\rho_{j,k}$ (if it exists), positive or not. Considering the bivariate survival function

$$S_{j,k}(x_j,x_k) = S(0,\dots,0,x_j,0,\dots,0,x_k,0,\dots,0),\quad x_j,x_k\in\mathbb{R}_+,$$

we can obtain $\rho_{j,k}$ by calculating the expectation

$$E(X_{j,1}X_{k,1}) = \int_0^\infty\int_0^\infty S_{j,k}(x_j,x_k)dx_jdx_k. \tag{2.5}$$

A central problem is how to construct a partially Schur-constant vector. For this, we will use a standard vectorial notation. Let $\mathbf{k}=(k_1,\dots,k_m)$, $\mathbf{l}=(l_1,\dots,l_m)\in\mathbb{N}_0^m$, and $\mathbf{x}=(x_1,\dots,x_m)$, $\mathbf{t}=(t_1,\dots,t_m)\in\mathbb{R}_+^m$. We set $\mathbf{k}!=k_1!\dots k_m!$, $\mathbf{t}^{\mathbf{k}}=t_1^{k_1}\dots t_m^{k_m}$, $\sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{l}}=\sum_{k_1=0}^{l_1}\dots\sum_{k_m=0}^{l_m}$ and $\mathbf{k}\le(\lneqq)\,\mathbf{l}$ if $k_j\le l_j$ for all j (with $k_j<l_j$ for at least one j). We also write $S^{(\mathbf{k})}(\mathbf{x})=\partial^{k_1+\ldots+k_m}S(x_1,\dots,x_m)/\partial x_1^{k_1}\dots\partial x_m^{k_m}$.

First, the generator S is characterized through the property of multiple monotonicity for a multivariate function. A thorough study of this property is provided by [40, 41]. Let f be any function f on $\mathbb{R}_+^m$ and, for simplicity, suppose that it is $\mathbf{n}$ times differentiable. Then, f is said to be $\mathbf{n}$-monotone if for all $\mathbf{x}\in\mathbb{R}_+^m$,

$$(-1)^{|\mathbf{k}|}S^{(\mathbf{k})}(\mathbf{x})\ge 0,\quad \mathbf{0}\lneqq\mathbf{k}\le\mathbf{n}. \tag{2.6}$$

Note that monotonicity is a standard property in the univariate case; see e.g. [27, 34, 49].

From the definition (2.2), the generator S is an m-variates function which corresponds to a survival function on $\mathbb{R}_+^m$. Theorem 10 of [40] states the following characterization of such a function S.

Proposition 2.2. *(Monotonicity of S)*
A survival function $S:\mathbb{R}_+^m\to[0,1]$ *may generate a partially Schur-constant vector (2.2) if and only if the function* S *is* $\mathbf{n}=(n_1,\dots,n_m)$*-monotone on* $\mathbb{R}_+^m$.

Now, a vector that is partially Schur-constant can be represented in two different ways. Both representations are determined from the m sums of all the variables inside the m sub-vectors $\mathbf{X}_1,\dots,\mathbf{X}_m$. So, denote this vector by $\mathbf{T}=(T_1,\dots,T_m)$ with $T_j=X_{j,1}+\ldots+X_{j,n_j}$, $1\le j\le m$. Given its importance, we derive below the distribution of $\mathbf{T}$.

Proposition 2.3. *(Distribution of* $\mathbf{T}$*)*
The density function of $\mathbf{T}$ *is given by*

$$f_{\mathbf{T}}(\mathbf{t}) = (-1)^{|\mathbf{n}|}\, S^{(\mathbf{n})}(\mathbf{t}) \frac{\mathbf{t}^{\mathbf{n}-\mathbf{1}}}{(\mathbf{n}-\mathbf{1})!}, \tag{2.7}$$

and the survival function by

$$P(\mathbf{T} > \mathbf{t}) = \sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{n}-\mathbf{1}} (-1)^{|\mathbf{k}|+m}\, S^{(\mathbf{k})}(\mathbf{t}) \frac{\mathbf{t}^{\mathbf{k}}}{\mathbf{k}!}, \quad \mathbf{t} \in \mathbb{R}_+^m.$$

Proof. Consider all the possible partial sums $T_{j,k_j} = X_{j,1} + \ldots + X_{j,k_j}$, for $k_j = 1, \ldots, n_j$ and $j = 1, \ldots, m$. From (2.1), we see that their joint density function is given by

$$\begin{aligned} & f_{(T_{1,1},\ldots,T_{1,n_1};\ldots;T_{m,1},\ldots,T_{m,n_m})}(t_{1,1}, \ldots, t_{1,n_1}; \ldots; t_{m,1}, \ldots, t_{m,n_m}) \\ &= (-1)^{n_1+\ldots+n_m} S^{(n_1,\ldots,n_m)}(t_{1,n_1}, \ldots, t_{m,n_m}), \quad 0 \le t_{1,1} \le \ldots \le t_{1,n_1}; \ldots; 0 \le t_{m,1} \le \ldots \le t_{m,n_m}. \end{aligned}$$

Integrating over the range of $(t_{1,1}, \ldots, t_{1,n_1-1}; \ldots; t_{m,1}, \ldots, t_{m,n_m-1})$, we then get

$$\begin{aligned} & f_{(T_{1,n_1},\ldots,T_{m,n_m})}(t_{1,n_1}, \ldots, t_{m,n_m}) \\ &= (-1)^{n_1+\ldots+n_m} S^{(n_1,\ldots,n_m)}(t_{1,n_1}, \ldots, t_{m,n_m}) \frac{t_{1,n_1}^{n_1-1} \ldots t_{m,n_m}^{n_m-1}}{(n_1-1)! \ldots (n_m-1)!}, \quad t_{1,n_1}, \ldots, t_{m,n_m} \ge 0. \end{aligned}$$

As each T_{j,n_j} corresponds to T_j, this gives precisely the formula (2.7) for the density function of $\mathbf{T}$. The associated survival function follows easily by integration. ◇

An important theorem of Williamson [49] states that any n-monotone function on $\mathbb{R}_+$ can be expressed as a mixture of functions of the form $(1 - rx)_+^{n-1}$ with $r > 0$. A generalization of this theorem with positive multivariate functions is established by [42], Theorem 3. Applied to the generator S, it reads as follows.

Proposition 2.4. *(Representation of* S*)*
A function $S : \mathbb{R}_+^m \to [0, 1]$ *is the generator of a partially Schur-constant vector (2.2) if and only if it admits the integral representation*

$$S(\mathbf{x}) = E\left[\prod_{j=1}^{m} \left(1 - \frac{x_j}{Z_j}\right)_+^{n_j-1}\right], \quad \mathbf{x} \in \mathbb{R}_+^m, \tag{2.8}$$

where the random vector $(Z_1, \ldots, Z_m)$ *is distributed as* $\mathbf{T} = (T_1, \ldots, T_m)$ *(of density (2.7)).*

Using (2.8), it can be seen that a Schur-constant vector $\mathbf{X}$ has the remarkable distributional representation (2.9), (2.10) below.

Proposition 2.5. *(Representation of* $\mathbf{X}$*)*
A Schur-constant vector $\mathbf{X}$ *can be represented in distribution as*

$$(\mathbf{X}_1, \ldots, \mathbf{X}_m) =_d [Z_1(U_{1,1}, \ldots, U_{1,n_1}), \ldots, Z_m(U_{m,1}, \ldots, U_{m,n_m})], \tag{2.9}$$

where the random vector $(Z_1, \ldots, Z_m)$ *is distributed as* $\mathbf{T} = (T_1, \ldots, T_m)$ *(of density (2.7)), and the random vectors* $(U_{j,1}, \ldots, U_{j,n_j})$, $1 \le j \le m$, *are independent of each other and of* $(Z_1, \ldots, Z_m)$ *and each forms a Schur-constant vector of joint survival function*

$$P(U_{j,1} > u_{j,1}, \ldots, U_{j,n_j} > u_{j,n_j}) = [1 - (u_{j,1} + \ldots + u_{j,n_j})]_+^{n_j-1}, \quad u_{j,1}, \ldots, u_{j,n_j} \in (0, 1). \tag{2.10}$$

Note that when if $n_j = 1$, i.e. when the group j contains a single variable $X_{j,1}$, then (2.8) or (2.9) gives $Z_j =_d X_{j,1}$. From (2.9) and (2.10), a partially Schur-constant model can be constructed in two steps: first a random vector $(Z_1, \ldots, Z_m)$ gives the total sums of variables in the m groups, and then each sum Z_j is independently distributed on the standard simplex using the vector $(U_{j,1}, \ldots, U_{j,n_j})$.

From (2.9), the moments and correlations of $\mathbf{X}$ can be directly expressed as a function of those of the vector $\mathbf{Z}$ ($=_d \mathbf{T}$). Inside each group j, we get, in obvious notation,

$$\mu_j = \mu_{Z_j}/n_j, \quad \sigma_j^2 = 2\sigma_{Z_j}^2/n_j(n_j+1) + (n_j-1)\mu_{Z_j}^2/n_j^2(n_j+1),$$
$$\rho_j = (n_j v_{Z_j}^2 - 1)/(2n_j v_{Z_j}^2 + n_j - 1). \tag{2.11}$$

Between two different groups $j \neq k$, we have

$$\rho_{j,k} = \rho_{Z_j,Z_k}(v_{Z_j} v_{Z_j}/v_j v_k), \tag{2.12}$$

and as $v_j \geq v_{Z_j}$, this implies that $\rho_{j,k} \leq \rho_{Z_j,Z_k}$, which is in agreement with intuition.

To conclude, we recall that a product ZU with Z positive and U uniform on $(0,1)$ is the definition due to Khintchine [22] for the unimodality on $\mathbb{R}_+$ with mode at 0. So, each Schur-constant vector $Z_j(U_{j,1},\dots,U_{j,n_j})$ could be considered as defining a unimodal distribution on $\mathbb{R}_+^{n_j}$. By (2.9), a partially Schur-constant vector would then represent a kind of multimodal distribution on $\mathbb{R}_+^{\mathbf{n}}$.

3 Associated survival copulas

As a corollary, we will now build an associated family of survival copulas. These are called partially Archimedean because they generalize classical Archimedean copulas by assuming the partial exchangeability of the uniform vector involved. Next, we will show that a partially Schur-constant vector has a survival copula which is precisely partially Archimedean with the same generator, and conversely. Let us start by recalling the definition of an Archimedean copula and, before that, of a copula in general (see e.g. [20, 39]).

Let $\mathbf{U} = (U_1,\dots,U_n)$, $n \geq 2$, be a random vector over the unit cube $[0,1]^n$ with uniform marginals. Its survival copula $C : [0,1]^n \to [0,1]$ is defined by

$$C(u_1,\dots,u_n) = P(U_1 > 1-u_1,\dots,U_n > 1-u_n), \quad u_1,\dots,u_n \in [0,1]. \tag{3.1}$$

Thus, the survival function of $(U_1,\dots,U_n)$ is given by $C(1-u_1,\dots,1-u_n)$. Now, consider a continuous random vector $(X_1,\dots,X_n)$. Its distribution can be represented in terms of its marginals and a copula: this is the famous Sklar theorem [45]. Specifically, let $\bar{F}_i(x_i) = P(X_i > x_i)$ be the marginal survival functions, $1 \leq i \leq n$. Then, there exists a unique survival copula C such that the survival function of $(X_1,\dots,X_n)$ is expressed as

$$P(X_1 > x_1,\dots,X_n > x_n) = C[\bar{F}_1(x_1),\dots,\bar{F}_n(x_n)], \quad x_1,\dots,x_n \in \mathbb{R}. \tag{3.2}$$

Archimedean copula. A survival copula C is Archimedean if (3.1) can be simplified in the form

$$C(u_1,\dots,u_n) = \psi\left[\psi^{-1}(u_1) + \dots + \psi^{-1}(u_n)\right], \quad u_1,\dots,u_n \in [0,1], \tag{3.3}$$

for some univariate function $\psi(x) : \mathbb{R}_+ \to [0,1]$, called generator, which satisfies $\lim_{x\to\infty}\psi(x) = 0$ and $\psi(0) = 1$, and where ψ^{-1} denotes the generalized inverse of ψ (i.e. $\psi^{-1}(u) = \inf\{x : \psi(x) \leq u\}$).

From (3.3), the vector $(U_1,\dots,U_n)$ has again a particular exchangeable distribution. Arguing as in Section 2, we break the symmetry of this distribution by making this vector only partially exchangeable.

Partially Archimedean copula. Let us partition the vector $\mathbf{U} = (U_1,\dots,U_n)$ into $m \geq 2$ sub-vectors $\mathbf{U}_j = (U_{j,1},\dots,U_{j,n_j})$, $1 \leq j \leq m$, with $n_j \geq 1$ and $n_1 + \dots + n_m = n$. We then introduce an m-variates function $\psi(x_1,\dots,x_m) : \mathbb{R}_+^m \to [0,1]$ which satisfies the border conditions $\psi(0,\dots,0) = 1$ and $\psi(x_1,\dots,x_m) \to 0$ when any $x_j \to \infty$.

Definition 3.1. *A survival copula $C(\mathbf{u}) = C(\mathbf{u}_1,\dots,\mathbf{u}_m)$ is partially Archimedean of generator ψ if (3.1) can be written in the form*

$$C(\mathbf{u}_1,\dots,\mathbf{u}_m) = \psi\left[|\psi_1^{-1}(\mathbf{u}_1)|,\dots,|\psi_m^{-1}(\mathbf{u}_m)|\right], \quad \textit{for all } \mathbf{u}_j \in [0,1]^{n_j}, \tag{3.4}$$

where the univariate functions $\psi_j : \mathbb{R}_+ \to [0,1]$, $1 \le j \le m$, are the m marginal projections of ψ given by

$$\psi_j(x_j) = \psi(0,\dots,0,x_j,0,\dots,0), \quad x_j \in \mathbb{R}_+, \tag{3.5}$$

and using the notation

$$|\psi_j^{-1}(\mathbf{u}_j)| = \psi_j^{-1}(u_{j,1}) + \dots + \psi_j^{-1}(u_{j,n_j}).$$

The idea of defining a generalized Archimedean copula by the formula (3.4) is not entirely new and was recently proposed by [42] in a different context. In Section 4 of that paper, it is also proved that the generator ψ can be characterized as in the partially Schur-constant model.

Proposition 3.2. *(Characterization of ψ)*
A survival function $\psi : \mathbb{R}_+^{\mathrm{m}} \to [0,1]$ may generate a partially Archimedean copula (3.4) if and only if the function ψ is $\mathbf{n} = (n_1,\dots,n_m)$-monotone on $\mathbb{R}^{\mathrm{m}}$. In that case, ψ can be expressed as the Williamson $\mathbf{n}$-transform given by (2.8).

Schur-constant models and Archimedean copulas are known to be closely related (see e.g. [13, 29, 38]). We show below that a similar link is also valid in the partially exchangeable case.

Proposition 3.3. *(Correspondence $S \leftrightarrow \psi$)*
(i) Let $\mathbf{X} = (\mathbf{X}_1,\dots,\mathbf{X}_m)$ be a partially Schur-constant vector of generator S. Then, the associated survival copula is a partially Archimedean copula of generator ψ with $\psi = S$.
(ii) Let $\mathbf{U} = (\mathbf{U}_1,\dots,\mathbf{U}_m)$ be a vector over $[0,1]^n$ with uniform marginals whose survival copula is partially Archimedean of generator ψ with marginal projections ψ_j, $1 \le j \le m$. Then, the vector $[\psi_1^{-1}(1-\mathbf{U}_1),\dots,\psi_m^{-1}(1-\mathbf{U}_m)]$ formed by the m sub-vectors

$$\psi_j^{-1}(1-\mathbf{U}_j) = [\psi_j^{-1}(1-U_{j,1}),\dots,\psi_j^{-1}(1-U_{j,n_j}],$$

is a partially Schur-constant model of generator S with $S = \psi$.

Proof. The reasoning is roughly similar to that followed in the simple Schur-constant model. For this case, a complete proof is provided via Proposition 4.5 of [29]. For simplicity, we assume here that the functions S_j andψ_j have an ordinary inverse function S_j^{-1} and ψ_j^{-1}. Otherwise, the proof method requires working with the generalized inverses.
Let $\mathbf{X}$ be partially Schur-constant of generator S. Its distribution function is given by (2.2). From (3.2), we also know that

$$P(\mathbf{X}_1 \ge \mathbf{x}_1,\dots,\mathbf{X}_m \ge \mathbf{x}_m) = C\left\{[S_j(x_{j,1}),\dots,S_j(x_{j,n_j})],\ 1 \le j \le m\right\}.$$

We thus get the identity

$$S(|\mathbf{x}_1|,\dots,|\mathbf{x}_m|) = C\left\{[S_j(x_{j,1}),\dots,S_j(x_{j,n_j})],\ 1 \le j \le m\right\}. \tag{3.6}$$

In each group j, we write $x_{j,i} = S_j^{-1}(u_{j,i})$ with $u_{j,i} \in (0,1)$, $1 \le i \le n_j$. Then, (3.6) becomes

$$S\left[S_j^{-1}(u_{j,1}) + \dots + S_j^{-1}(u_{j,n_j}),\ 1 \le j \le m\right] = C\left[(u_{j,1},\dots,u_{j,n_j}),\ 1 \le j \le m\right].$$

By virtue of (3.4), this means that the copula C is partially Archimedean of generator S.
Consider now the vector $[\psi_1^{-1}(1-\mathbf{U}_1),\dots,\psi_m^{-1}(1-\mathbf{U}_m)]$. Its distribution function is

$$\begin{aligned} &P\left\{[\psi_j^{-1}(1-U_{j,1}) > x_{j,1},\dots,\psi_j^{-1}(1-U_{j,n_j}) > x_{j,n_j}],\ 1 \le j \le m\right\} \\ &\quad = P\left\{[U_{j,1} > 1-\psi_j(x_{j,1}),\dots,U_{j,n_j} > 1-\psi_j(x_{j,n_j})],\ 1 \le j \le m\right\}. \end{aligned} \tag{3.7}$$

By assumption, the copula of $\mathbf{U}$ is partially Archimedean of generator ψ. From (3.1) and (3.4), we can thus express the r.h.s. of (3.7) as

$$\begin{aligned} P\left\{[U_{j,1} > 1-\psi_j(x_{j,1}),\dots,U_{j,n_j} > 1-\psi_j(x_{j,n_j})],\ 1\le j\le m\right\} \\ = \psi\left[\psi_j^{-1}(\psi_j(x_{j,1}))+\dots+\psi_j^{-1}(\psi_j(x_{j,n_j})),\ 1\le j\le m\right] \\ = \psi\left(x_{j,1}+\dots+x_{j,n_j},\ 1\le j\le m\right). \end{aligned} \tag{3.8}$$

From (3.7), (3.8) and (2.2), we deduce that $[\psi_1^{-1}(1-\mathbf{U}_1),\dots,\psi_m^{-1}(1-\mathbf{U}_m)]$ is partially Schur-contant of generator ψ. $\diamond$

In the next four sections, we focus on partially Schur-constant models built for two groups of variables. So, $\mathbf{X} = [\mathbf{X}_1 = (X_{1,1},\dots,X_{1,n_1}), \mathbf{X}_2 = (X_{2,1},\dots,X_{2,n_2})]$, and from (2.2), there is a bivariate generator $S(x_1,x_2) : \mathbb{R}_+^2 \to [0,1]$ such that

$$\begin{aligned} P[(X_{1,1} > x_{1,1},\dots,X_{1,n_1} > x_{1,n_1}),(X_{2,1} > x_{2,1},\dots,X_{2,n_2} > x_{2,n_2})] \\ = S(x_{1,1}+\dots+x_{1,n_1}, x_{2,1}+\dots+x_{2,n_2}), \quad \text{for all } x_{j,i}\ge 0. \end{aligned} \tag{3.9}$$

This generator S is a bivariate survival function and by Proposition 2.2, it must be (n_1,n_2)-monotone. Our goal is to present different possible functions for S and to specify, if possible, which are the corresponding probability distributions (for a vector (X_1,X_2) say).

4 From univariate Laplace transforms

We begin with the simplest case where $S(x_1,x_2)$ is the Laplace transform of a positive variable Λ of argument $\zeta_1x_1+\zeta_2x_2$ with $\zeta_1,\zeta_2>0$. So, $S(x_1,x_2)$ is of the form

$$S(x_1,x_2) = E(e^{-\Lambda(\zeta_1x_1+\zeta_2x_2)}), \quad x_1,x_2\ge 0. \tag{4.1}$$

Of course, $(\zeta_1X_1,\zeta_2X_2) =_d (Y_1/\Lambda, Y_2/\Lambda)$ where Y_1,Y_2 are two independent exponentials of parameter 1. In fact, (ζ_1X_1,ζ_2X_2) is Schur-constant and is distributed as the vector $(Z/\Lambda)(U,1-U)$ where U is uniform on $(0,1)$, and Z is distributed as Y_1+Y_2 independently of U, i.e. Z has an Erlang distribution of density $x\exp(-x)$, $x\ge 0$.

Obviously, S defined by (4.1) is infinitely monotone in x_1 and x_2. From (3.9), the associated partially Schur-constant vector $\mathbf{X}$ is of survival function

$$P(\mathbf{X}_1\ge\mathbf{x}_1,\mathbf{X}_2\ge\mathbf{x}_2) = E(e^{-\Lambda(\zeta_1|\mathbf{x}_1|+\zeta_2|\mathbf{x}_2|)}). \tag{4.2}$$

For the main moments of the model, we get using (2.4), (2.5) that inside each group j, $\mu_j = E(X_{j,1}) = (1/\zeta_j)E(1/\Lambda)$, $E(X_{j,1}^2) = (2/\zeta_j^2)E(1/\Lambda^2)$, $E(X_{j,1}X_{j,2}) = E(X_{j,1}^2)/2$, and between the two groups, $E(X_{1,1}X_{2,1}) = (1/\zeta_1\zeta_2)E(1/\Lambda^2)$. Thus, $\rho_j = \rho_{1,2} \equiv \rho$ with

$$\rho = \frac{E(1/\Lambda^2)-[E(1/\Lambda)]^2}{2E(1/\Lambda^2)-[E(1/\Lambda)]^2},$$

i.e., all the correlations within and between the groups are identical and valued in $(0,1/2)$, independently of ζ_1,ζ_2. In other words, the generator (4.1) produces two Schur-constant groups with different distributions but keeps the correlations all equal.

The corresponding partially Archimedean copula (3.4) shows that the dependency structure is not modified by (4.1). Indeed, $C(\mathbf{U})$ in this case is reduced to

$$C(\mathbf{u}_1,\mathbf{u}_2) = \mathcal{L}_\Lambda[|\mathcal{L}_\Lambda^{-1}(\mathbf{u}_1)|+|\mathcal{L}_\Lambda^{-1}(\mathbf{u}_2)|], \quad u_{j,i}\in[0,1], \tag{4.3}$$

where $|\mathcal{L}_\Lambda^{-1}(\mathbf{u}_j)| = \mathcal{L}_\Lambda^{-1}(u_{j,1})+\dots+\mathcal{L}_\Lambda^{-1}(u_{j,n_j})$ and $\mathcal{L}_\Lambda(x) = E(e^{-\Lambda x})$ is the Laplace transform of Λ of parameter $x\ge 0$. The formula (4.3) is simply that of a Schur-constant vector of dimension $|\mathbf{n}| = n_1+n_2$ and of generator $\mathcal{L}_\Lambda$. For illustration, we consider several particular distributions for Λ.

(1) Λ *has a gamma distribution with positive parameters* $(1/\theta, 1/\beta)$. Its Laplace transform is $\mathcal{L}_\Lambda(x) = 1/(1+\beta x)^{1/\theta}$, so that from (4.1),

$$S(x_1, x_2) = \frac{1}{(1+\beta_1 x_1 + \beta_2 x_2)^{1/\theta}}, \quad x_1, x_2 \geq 0, \tag{4.4}$$

where $\beta_j = \beta\zeta_j$, $j = 1, 2$. Thus, (X_1, X_2) has a bivariate Lomax (Pareto Type II) distribution discussed by [31] and [37] (see also e.g. [50]).

Consider a partially Schur-constant vector $\mathbf{X}$ built from (4.4). Its survival function (4.2) then becomes

$$P(\mathbf{X}_1 \geq \mathbf{x}_1, \mathbf{X}_2 \geq \mathbf{x}_2) = \frac{1}{(1+\beta_1|\mathbf{x}_1| + \beta_2|\mathbf{x}_2|)^{1/\theta}}, \quad x_{j,i} \geq 0.$$

If $\theta < 1/2$, we have $\mu_j = (1/\beta_j)\theta/(1-\theta)$, $\sigma_j^2 = (1/\beta_j^2)\theta^2/(1-\theta)^2(1-2\theta)$, and $\rho = \theta$ which increases with θ from 0 to 1/2. The survival copula (4.3) is the Clayton copula

$$C(\mathbf{u}_1, \mathbf{u}_2) = \frac{1}{(|\mathbf{u}_1^{-\theta}| + |\mathbf{u}_2^{-\theta}| - n_1 - n_2 + 1)_+^{1/\theta}}, \quad u_{j,i} \in [0, 1],$$

after denoting $|\mathbf{u}_j^{-\theta}| = u_{j,1}^{-\theta} + \ldots + u_{j,n_j}^{-\theta}$.

(2) Λ *has a one-sided stable distribution with parameter* $\theta \geq 1$. Its Laplace transform is $\mathcal{L}_\Lambda(x) = \exp(-x^{1/\theta})$, so that from (4.1),

$$S(x_1, x_2) = e^{-(\zeta_1 x_1 + \zeta_2 x_2)^{1/\theta}}, \quad x_1, x_2 \geq 0. \tag{4.5}$$

Thus, (X_1, X_2) has a bivariate Weibull distribution discussed in Section 5 of [25] (see also e.g. [19] and [32]).

A partially Schur-constant vector $\mathbf{X}$ built from (4.5) is of survival function

$$P(\mathbf{X}_1 \geq \mathbf{x}_1, \mathbf{X}_2 \geq \mathbf{x}_2) = e^{-(\zeta_1|\mathbf{x}_1| + \zeta_2|\mathbf{x}_2|)^{1/\theta}}, \quad x_{j,i} \geq 0.$$

We get $\mu_j = (1/\zeta_j)\,\theta\Gamma(\theta)$, $\sigma_j^2 = (1/\zeta_j^2)[2\theta\Gamma(2\theta) - \theta^2\Gamma^2(\theta)]$, and

$$\rho = \frac{\theta\Gamma(2\theta) - \theta^2\Gamma^2(\theta)}{2\theta\Gamma(2\theta) - \theta^2\Gamma^2(\theta)} = 1 - \frac{1}{2 - \theta\Gamma^2(\theta)/\Gamma(2\theta)}.$$

which increases with θ from 0 for $\theta = 1$ to 1/2 as $\theta \to \infty$. The associated survival copula is the Gumbel copula

$$C(\mathbf{u}_1, \mathbf{u}_2) = e^{-\{[-\ln(|\mathbf{u}_1|)]^\theta + [-\ln(|\mathbf{u}_2|)]^\theta\}^{1/\theta}}, \quad u_{j,i} \in [0, 1],$$

where $[-\ln(|\mathbf{u}_j|)]^\theta$ denotes $[-\ln(u_{j,1})]^\theta + \ldots + [-\ln(u_{j,n_j})]^\theta$.

(3) Λ *has a shifted geometric distribution of parameter* $\theta \in [0, 1)$. Its Laplace transform is $\mathcal{L}_\Lambda(x) = (1-\theta)/(\exp(x) - \theta)$, so that from (4.1),

$$S(x_1, x_2) = \frac{1-\theta}{\exp(\zeta_1 x_1 + \zeta_2 x_2) - \theta}, \quad x_1, x_2 \geq 0, \tag{4.6}$$

This survival function does not seem to be explicitly referenced in the literature.

A partially Schur-constant vector $\mathbf{X}$ built from (4.6) is such that

$$P(\mathbf{X}_1 \geq \mathbf{x}_1, \mathbf{X}_2 \geq \mathbf{x}_2) = \frac{1-\theta}{\exp(\zeta_1|\mathbf{x}_1| + \zeta_2|\mathbf{x}_2|) - \theta}, \quad x_{j,i} \geq 0.$$

We find that $\mu_j = [(1-\theta)/\theta\zeta_j]Li_1(\theta)$, $\sigma_j^2 = [2(1-\theta)/\theta\zeta_j^2]Li_2(\theta) - \mu_j^2$, and

$$\rho = \frac{\theta Li_2(\theta) - (1-\theta)[Li_1(\theta)]^2}{2\theta Li_2(\theta) - (1-\theta)[Li_1(\theta)]^2} = 1 - \frac{1}{2 - (1-\theta)[Li_1(\theta)]^2/\theta Li_2(\theta)},$$

in which the function $Li_s(\theta) = \sum_{i=1}^\infty \theta^i/i^s$ is the polylogarithm of order $s = 1, 2$. Recall that for $s = 1$, $Li_1(\theta) = -\ln(1-\theta)$, and for $s \geq 2$, $\theta d[Li_s(\theta)]/d\theta = Li_{s-1}(\theta)$. It can be proved that the correlation increases with θ from 0 to 1/2. The associated survival copula is the Ali-Mikhail-Haq copula

$$C(\mathbf{u}_1, \mathbf{u}_2) = \frac{\prod_{i=1}^{n_1} u_{1,i} \prod_{i=1}^{n_2} u_{2,i}}{1 - \theta\prod_{i=1}^{n_1}(1-u_{1,i})\prod_{i=1}^{n_2}(1-u_{2,i})}, \quad u_{j,i} \in [0, 1].$$

(4) Λ *has a logarithmic distribution of parameter* $p \in (0,1)$. Its Laplace transform is $\mathcal{L}_\Lambda(x) = \ln(1 - pe^{-x})/\ln(1-p)$, so that from (4.1),

$$S(x_1, x_2) = \ln(1 - pe^{-\zeta_1 x_1 - \zeta_2 x_2})/\ln(1-p), \quad x_1, x_2 \geq 0, \tag{4.7}$$

which does not seem to be standard either, to our knowledge.

A partially Schur-constant vector $\mathbf{X}$ built from (4.7) is of survival function

$$P(\mathbf{X}_1 \geq \mathbf{x}_1, \mathbf{X}_2 \geq \mathbf{x}_2) = \ln(1 - pe^{-\zeta_1|\mathbf{x}_1| - \zeta_2|\mathbf{x}_2|})/\ln(1-p), \quad x_{j,i} \geq 0.$$

We get $\mu_j = (1/\zeta_j)Li_2(p)/Li_1(p)$, $\sigma_j^2 = (2/\zeta_j^2)Li_3(p)/Li_1(p) - \mu_j^2$, and

$$\rho = \frac{Li_1(p)Li_3(p) - [Li_2(p)]^2}{2Li_1(p)Li_3(p) - [Li_2(p)]^2} = 1 - \frac{1}{2 - [Li_2(p)]^2/Li_1(p)Li_3(p)},$$

which increases with p from 0 to 1/2 (numerical check). Setting $p = 1 - e^{-\theta}$ ($\theta > 0$), the associated survival copula is the Frank copula

$$C(\mathbf{u}_1, \mathbf{u}_2) = -(1/\theta)\ln[1 - (1 - e^{-\theta|\mathbf{u}_1|})(1 - e^{-\theta|\mathbf{u}_2|})/(1 - e^{-\theta})], \quad u_{j,i} \in [0,1].$$

5 From bivariate Laplace transforms

We pass here to the interesting case, more complex, where $S(x_1, x_2)$ is the Laplace transform of a positive random vector (Λ_1, Λ_2) of arguments (x_1, x_2). So, $S(x_1, x_2)$ is of the form

$$S(x_1, x_2) = E(e^{-\Lambda_1 x_1 - \Lambda_2 x_2}), \quad x_1, x_2 \geq 0, \tag{5.1}$$

The function S defined by (5.1) is again infinitely monotone in x_1 and x_2. From (3.9), the partially Schur-constant vector $\mathbf{X}$ is of survival function

$$P(\mathbf{X}_1 \geq \mathbf{x}_1, \mathbf{X}_2 \geq \mathbf{x}_2) = E(e^{-\Lambda_1|\mathbf{x}_1| - \Lambda_2|\mathbf{x}_2|}). \tag{5.2}$$

For the main moments, we get from (2.4), (2.5) that inside each group j, $\mu_j = E(X_{j,1}) = E(1/\Lambda_j)$, $E(X_{j,1}^2) = 2E(1/\Lambda_j^2)$ and $E(X_{j,1}X_{j,2}) = E(X_{j,1}^2)/2$, and between the two groups, $E(X_{1,1}X_{2,1}) = E(1/\Lambda_1\Lambda_2)$. Thus, ρ_j has values in $(0, 1/2)$ and is equal to

$$\rho_j = \frac{E(1/\Lambda_j^2) - [E(1/\Lambda_j)]^2}{2E(1/\Lambda_j^2) - [E(1/\Lambda_j)]^2},$$

while $\rho_{1,2} \equiv \rho(X_{1,1}, X_{2,1})$ can be positive or not and is given by

$$\rho_{1,2} = \frac{E(1/\Lambda_1\Lambda_2) - E(1/\Lambda_1)E(1/\Lambda_2)}{\{2E(1/\Lambda_1^2) - [E(1/\Lambda_1)]^2\}^{1/2}\{2E(1/\Lambda_2^2) - [E(1/\Lambda_2)]^2\}^{1/2}}.$$

The associated partially Archimedean copula (3.4) is reduced to

$$C(\mathbf{u}_1, \mathbf{u}_2) = \mathcal{L}_{(\Lambda_1,\Lambda_2)}[|\mathcal{L}_{\Lambda_1}^{-1}(\mathbf{u}_1)| + |\mathcal{L}_{\Lambda_2}^{-1}(\mathbf{u}_2)|], \quad u_{j,i} \in [0,1], \tag{5.3}$$

where $\mathcal{L}_{\Lambda_1}(x_1)$, $\mathcal{L}_{\Lambda_2}(x_2)$, $\mathcal{L}_{(\Lambda_1,\Lambda_2)}(x_1, x_2)$ are the Laplace transforms of $\Lambda_1, \Lambda_2, (\Lambda_1, \Lambda_2)$ of parameters $x_1, x_2 \geq 0$. For illustration, we consider several particular distributions for (Λ_1, Λ_2).

(1) (Λ_1, Λ_2) *has a bivariate gamma distribution of positive parameters* $(\alpha, \zeta_1, \zeta_2, \zeta_3)$ *with* $\zeta_3 \leq \zeta_1\zeta_2$. Following [10, 12, 30], this means that the Laplace transform of (Λ_1, Λ_2) is of the form

$$E(e^{-\Lambda_1 x_1 - \Lambda_2 x_2}) = \frac{1}{(1 + \zeta_1 x_1 + \zeta_2 x_2 + \zeta_3 x_1 x_2)^\alpha}, \quad x_1, x_2 \geq 0. \tag{5.4}$$

As proved in [10], getting an effective Laplace transform requires that $\zeta_3 \leq \zeta_1\zeta_2$. Note that each Λ_j has a univariate gamma distribution of parameters (α, ζ_j), and they are independent if $\zeta_3 = \zeta_1\zeta_2$. In the literature, the distribution of (Λ_1, Λ_2) is often called the Kibble and Moran distribution (see e.g. the books [3] and [23]). We see that $E(\Lambda_j) = \alpha\zeta_j$, $\sigma^2(\Lambda_j) = \alpha\zeta_j^2$ (the two coefficients of variation are identical and equal to $1/\sqrt{\alpha}$), and $\rho(\Lambda_1, \Lambda_2) = 1 - \zeta_3/\zeta_1\zeta_2 \geq 0$ and independent of α. As with the bivariate Normal distribution, the two variables are independent when their correlation is 0.

Let us insert (5.4) in the survival function $S(x_1, x_2)$ defined by (5.1). The corresponding vector (X_1, X_2) has a bivariate Lomax (Pareto Type II) distribution of survival function

$$S(x_1, x_2) = \frac{1}{(1 + \zeta_1 x_1 + \zeta_2 x_2 + \zeta_3 x_1 x_2)^\alpha}, \quad x_1, x_2 \geq 0. \tag{5.5}$$

The condition on the parameters for the existence of a Lomax distribution is $\zeta_3 \leq (1 + \alpha)\zeta_1\zeta_2$. Recall that this distribution comes from a Laplace transform under the more restrictive constraint $\zeta_3 \leq \zeta_1\zeta_2$. This bivariate Lomax distribution was proposed by [17] in the case where $\zeta_1 = \zeta_2 = 1$. The general form was examined by [44] and revisited by [43] as a bivariate Pareto model; see also e.g. [2], formula (16). Obviously, if (X_1, X_2) has this Lomax distribution of parameters $(\zeta_1, \zeta_2, \zeta_3, \alpha)$, then $(\zeta_1 X_1, \zeta_2 X_2)$ has the Lomax distribution of parameters $(1, 1, \zeta_3/\zeta_1\zeta_2, \alpha)$ (considered by [17]). Note that each X_j has a univariate Lomax distribution of parameters (α, μ_j), and they are independent if $\zeta_3 = \zeta_1\zeta_2$.

A partially Schur-constant vector $\mathbf{X}$ built from (5.5) is such that

$$P(\mathbf{X}_1 \geq \mathbf{x}_1, \mathbf{X}_2 \geq \mathbf{x}_2) = \frac{1}{(1 + \zeta_1|\mathbf{x}_1| + \zeta_2|\mathbf{x}_2| + \zeta_3|\mathbf{x}_1||\mathbf{x}_2|)^\alpha}, \quad x_{j,i} \geq 0.$$

For each group, if $\alpha > 2$, we get $\mu_j = 1/\zeta_j(\alpha - 1)$, $\sigma_j^2 = \alpha/\zeta_j^2(\alpha - 1)^2(\alpha - 2)$, and $\rho_j = 1/\alpha$, the same for the two groups, which decreases with α from 1/2 to 0. Moreover, we see that if $\zeta_3 \leq (\geq)\, \zeta_1\zeta_2$, then $S(x_1, x_2) \geq (\leq) S(x_1, 0)S(0, x_2)$ and thus $\rho_{1,2} \geq (\leq)\, 0$, which is in agreement with Figure 1 in [43]. In the present case, however, we assumed $\zeta_3 \leq \zeta_1\zeta_2$, so that $\rho_{1,2} \geq 0$ (as $\rho(\Lambda_1, \Lambda_2)$). In [24], it is shown that when $\alpha > 2$,

$$\rho_{1,2} = [(1 - \zeta)(\alpha - 2)/\alpha^2]\, F(1, 2; \alpha + 1; 1 - \zeta),$$

where $\zeta = \zeta_3/\zeta_1\zeta_2$ (≤ 1 in our case) and $F(a, b; c; z)$ is the Gauss hypergeometric function (see e.g. Chapter 15 of [1]). Thus, $\rho_{1,2}$ decreases with ζ, from $1/\alpha$ when $\zeta = 0$, i.e. $\zeta_3 = 0$ (since $F(1, 2; \alpha + 1; 1) = \alpha/(\alpha - 2)$) to 0 when $\zeta = 1$, i.e. $\zeta_3 = \zeta_1\zeta_2$. The associated survival copula (5.3) is

$$C(\mathbf{u}_1, \mathbf{u}_2) = \frac{1}{[|\mathbf{u}_1^{-1/\alpha}| + |\mathbf{u}_2^{-1/\alpha}| - n_1 - n_2 + 1 + (\zeta_3/\zeta_1\zeta_2)(|\mathbf{u}_1^{-1/\alpha}| - n_1)(|\mathbf{u}_2^{-1/\alpha}| - n_2)]_+^\alpha}, \quad u_{j,i} \in [0, 1].$$

(2) (Λ_1, Λ_2) *has a bifactor gamma distribution of positive parameters* $(\alpha_1, \alpha_2, \zeta_1, \zeta_2, \zeta_3)$ *with* $\zeta_3 \leq \zeta_1\zeta_2$ *and* $\alpha_1 \leq \alpha_2$. Following [5, 11], this means that $(\Lambda_1, \Lambda_2) =_d (\Lambda_{[1]}, \Lambda_{[2]} + \Lambda)$ with $(\Lambda_{[1]}, \Lambda_{[2]})$ has a bivariate gamma distribution of positive parameters $\alpha_1, \zeta_1, \zeta_2, \zeta_3 \leq \zeta_1\zeta_2$ and Λ has an independent gamma distribution of positive parameters $(\alpha_2 - \alpha_1, \zeta_2)$. Equivalently, the Laplace transform of (Λ_1, Λ_2) is given by

$$E(e^{-\Lambda_1 x_1 - \Lambda_2 x_2}) = \frac{1}{(1 + \zeta_1 x_1 + \zeta_2 x_2 + \zeta_3 x_1 x_2)^{\alpha_1}} \frac{1}{(1 + \zeta_2 x_2)^{\alpha_2 - \alpha_1}}, \quad x_1, x_2 \geq 0. \tag{5.6}$$

Note that each Λ_j has a univariate gamma distribution of parameters (α_j, ζ_j), i.e. the parameters α_j can also be different here. We see that $\rho(\Lambda_1, \Lambda_2) = \rho(\Lambda_{[1]}, \Lambda_{[2]})\, \sigma(\Lambda_{[2]})/\sigma(\Lambda_2)$ with $\sigma^2(\Lambda_2) = \sigma^2(\Lambda_{[2]}) + \sigma^2(\Lambda)$, hence $\rho(\Lambda_1, \Lambda_2) = (1 - \zeta_3/\zeta_1\zeta_2)\sqrt{\alpha_1/\alpha_2} \in [0, \sqrt{\alpha_1/\alpha_2}]$.

Combining (5.1), (5.6) leads us to define a vector (X_1, X_2) of survival function

$$S(x_1, x_2) = \frac{1}{(1 + \zeta_1 x_1 + \zeta_2 x_2 + \zeta_3 x_1 x_2)^{\alpha_1}} \frac{1}{(1 + \zeta_2 x_2)^{\alpha_2 - \alpha_1}}, \quad x_1, x_2 \geq 0. \tag{5.7}$$

This implies that $(X_1, X_2) =_d [X_{[1]}, \min(X_{[2]}, X)]$ where $(X_{[1]}, X_{[2]})$ is a bivariate Lomax vector of survival function (5.5) with α_1 substitued for α, and X is an independent Lomax variable of parameters $(\alpha_2 - \alpha_1, \zeta_2)$. Indeed, thanks to the assumptions made, we have

$$P[X_{[1]} > x_1, \min(X_{[2]}, X) > x_2] = P(X_{[1]} > x_1, X_{[2]} > x_2)P(X > x_2) = S(x_1, x_2).$$

Note that X_2 is a Lomax variable of parameters (α_2, ζ_2). The survival function (5.7) does not seem to be referenced in the literature.

We associate with (5.7) a partially Schur-constant vector $\mathbf{X}$ defined by

$$P(\mathbf{X}_1 \geq \mathbf{x}_1, \mathbf{X}_2 \geq \mathbf{x}_2) = \frac{1}{(1+\zeta_1|\mathbf{x}_1| + \zeta_2|\mathbf{x}_2| + \zeta_3|\mathbf{x}_1||\mathbf{x}_2|)^{\alpha_1}} \frac{1}{(1+\zeta_2|\mathbf{x}_2|)^{\alpha_2-\alpha_1}}, \quad x_{j,i} \geq 0.$$

If $\alpha_1, \alpha_2 > 2$, we have $\mu_j = 1/\zeta_j(\alpha_j - 1)$, $\sigma_j^2 = \alpha_j/\zeta_j^2(\alpha_j-1)^2(\alpha_j-2)$, and $\rho_j = 1/\alpha_j$, different for the two groups. Moreover, $\rho_{1,2} = \rho(X_{[1]}, X_{[2]})\, \sigma(X_{[2]})/\sigma(X_2)$ with $\sigma^2(X_2) = \sigma_2^2$ and $\sigma^2(X_{[2]}) = \alpha_1/\zeta_2^2(\alpha_1-1)^2(\alpha_1-2)$, hence

$$\rho_{1,2} = \rho(X_{[1]}, X_{[2]}) \sqrt{\alpha_1(\alpha_2-1)^2(\alpha_2-2)/\alpha_2(\alpha_1-1)^2(\alpha_1-2)}.$$

This time, the survival copula (5.3) is

$$C(\mathbf{u}_1, \mathbf{u}_2) = \frac{1}{[|\mathbf{u}_1^{-1/\alpha_1}| + |\mathbf{u}_2^{-1/\alpha_2}| - n_1 - n_2 + 1 + (\zeta_3/\zeta_1\zeta_2)(|\mathbf{u}_1^{-1/\alpha_1}| - n_1)(|\mathbf{u}_2^{-1/\alpha_2}| - n_2)]_+^{\alpha_1}} \frac{1}{[|\mathbf{u}_2^{-1/\alpha_2}| - n_2 + 1]_+^{\alpha_2-\alpha_1}}, \quad u_{j,i} \in [0,1].$$

(3) *Using an exponential transform of positive powers* $\alpha_1, \alpha_2 \leq 1$. Suppose that the vector $(\zeta_1 X_1, \zeta_2 X_2)$, $\zeta_1, \zeta_2 > 0$, is distributed as the vector $[(Y_1/\Lambda)^{1/\alpha_1}, (Y_2/\Lambda)^{1/\alpha_2}]$, $\alpha_1, \alpha_2 > 0$, where Λ is some positive variable and Y_1, Y_2 are two independent exponentials of parameter 1. Note that the model (4.1) corresponds to the particular case where $\alpha_1 = \alpha_2 = 1$. Then, $S(x_1, x_2) = P(X_1 > x_1, X_2 > x_2)$ is given by

$$S(x_1, x_2) = E(e^{-\Lambda[(\zeta_1 x_1)^{\alpha_1} + (\zeta_2 x_2)^{\alpha_2}]}), \quad x_1, x_2 \geq 0. \tag{5.8}$$

This function is a bivariate Laplace transform if it is infinitely monotone in x_1 and x_2. We observe that this is verified with the additional conditions on the parameters $\alpha_1, \alpha_2 \leq 1$.

For example, if Λ as a one-sided stable distribution with parameter $\theta \geq 1$ (as for (4.5)), we get

$$S(x_1, x_2) = e^{-[(\zeta_1 x_1)^{\alpha_1} + (\zeta_2 x_2)^{\alpha_2}]^{1/\theta}}, \quad x_1, x_2 \geq 0. \tag{5.9}$$

Thus, (X_1, X_2) has a bivariate Weibull distribution discussed by [26] (see also e.g. [32]). For $\zeta_1 = \zeta_2 = 1$, this distribution is chosen by [42] as an example of a bivariate Laplace transform under the conditions $\alpha_1, \alpha_2 \leq 1$.

If Λ has a gamma distribution with positive parameters $(1/\theta, 1)$ (as for (4.4)),

$$S(x_1, x_2) = \frac{1}{[1 + (\zeta_1 x_1)^{\alpha_1} + (\zeta_2 x_2)^{\alpha_2}]^{1/\theta}}, \quad x_1, x_2 \geq 0. \tag{5.10}$$

Thus, (X_1, X_2) has a bivariate Burr (Pareto Type IV) distribution defined by [48] (see also e.g. [23]).

The partially Schur-constant vector built from (5.9) is of survival function

$$P(\mathbf{X}_1 \geq \mathbf{x}_1, \mathbf{X}_2 \geq \mathbf{x}_2) = E(e^{-\Lambda[(\zeta_1|\mathbf{x}_1|)^{\alpha_1} + (\zeta_2|\mathbf{x}_2|)^{\alpha_2}]}), \tag{5.11}$$

For the moments, since

$$\int_0^\infty x^m e^{-\beta x^\alpha} dx = \frac{1}{\alpha \beta^{(m+1)/\alpha}} \Gamma[(m+1)/\alpha], \quad m \geq 0, \alpha, \beta > 0,$$

we find that $\mu_j = E(X_{j,1}) = (1/\zeta_j\alpha_j)\Gamma(1/\alpha_j)E(1/\Lambda^{1/\alpha_j})$, $E(X_{j,1}^2) = (2/\zeta_j^2\alpha_j)\Gamma(2/\alpha_j)E(1/\Lambda^{2/\alpha_j})$, $E(X_{j,1}X_{j,2}) = E(X_{j,1}^2)/2$, and $E(X_{1,1}X_{2,1}) = (1/\zeta_1\alpha_1\zeta_2\alpha_2)\Gamma(1/\alpha_1)\Gamma(1/\alpha_2)E(1/\Lambda^{1/\alpha_1+1/\alpha_2})$. A formula for ρ_j and $\rho_{1,2}$ follows directly. For the associated survival copula, since $[\mathcal{L}_\Lambda(x_j^{\alpha_j})]^{-1} = [\mathcal{L}_\Lambda^{-1}(x_j)]^{1/\alpha_j}$, we get

$$C(\mathbf{u}_1, \mathbf{u}_2) = \mathcal{L}_\Lambda\{[(\mathcal{L}_\Lambda^{-1}(u_{1,1}))^{1/\alpha_1} + \ldots + (\mathcal{L}_\Lambda^{-1}(u_{1,n_1}))^{1/\alpha_1}]^{\alpha_1} + [(\mathcal{L}_\Lambda^{-1}(u_{2,1}))^{1/\alpha_2} + \ldots + (\mathcal{L}_\Lambda^{-1}(u_{2,n_2}))^{1/\alpha_2}]^{\alpha_2}\}, \quad u_{j,i} \in [0,1].$$

6 A finitely monotone survival function

So far, we have presented examples of survival functions $S(x_1, x_2)$ which correspond to Laplace transforms and are therefore infinitely monotone in x_1 and x_2. In this Section, we want to put forward a particular classical distribution for which $S(x_1, x_2)$ is not a Laplace transform.

Among various possible cases, we choose the bivariate exponential distribution of Gumbel [17] with positive parameters $(\zeta_1, \zeta_2, \zeta_3)$. Its survival function is given by

$$S(x_1, x_2) = e^{-\zeta_1 x_1 - \zeta_2 x_2 - \zeta_3 x_1 x_2}, \quad x_1, x_2 \geq 0, \tag{6.1}$$

and as the density shows, it is actually a distribution if $0 \leq \zeta_3 \leq \zeta_1 \zeta_2$. Obviously, $(\zeta_1 X_1, \zeta_2 X_2)$ has a bivariate exponential distribution of parameters $(1, 1, \zeta)$ where $0 \leq \zeta \equiv \zeta_3/\zeta_1\zeta_2 \leq 1$.

Let us consider (6.1) as a possible generator for a partially Schur-constant model (3.9). According to Proposition 2.2, we must verify that $S(x_1, x_2)$ is indeed a (n_1, n_2)-monotone function. The proposition below illustrates that this property depends on the value of the parameters and the size of the groups.

Proposition 6.1. *$S(x_1, x_2)$ defined by (6.1) is $(1, n)$-monotone if and only if*

$$\zeta_3 \leq (1/n)\zeta_1\zeta_2, \quad n \geq 1. \tag{6.2}$$

To be $(2, n)$-monotone ($n \geq 2$), a sufficient condition is $\zeta_3 \leq [1 - (1 - 1/n)^{1/2}]\zeta_1\zeta_2$. To be $(3, n)$-monotone ($n \geq 3$), it suffices that $\zeta_3 \leq [1 - (1 - 1/n(n-1)(n-2))^{1/3}]\zeta_1\zeta_2$.

Proof. From (2.6), the function S is $(1, n)$-monotone when

$$(-1)^{i+j} S^{(i,j)}(x_1, x_2) \geq 0, \quad \text{for } (i, j) = (0, 1), \ldots (0, n); (1, 0), \ldots, (1, n), \tag{6.3}$$

and all $x_1, x_2 \geq 0$. For $i = 0$, we have

$$S^{(0,j)}(x_1, x_2) = (-1)^j (\zeta_2 + \zeta_3 x_1)^j S(x_1, x_2),$$

hence (6.3) is satisfied for all j. For $i = 1$, we get

$$S^{(1,j)}(x_1, x_2) = (-1)^j (\zeta_1 + \zeta_3 x_2)^{j-1} [j\zeta_3 - (\zeta_1 + \zeta_3 x_2)(\zeta_2 + \zeta_3 x_1)] S(x_1, x_2),$$

so that (6.3) is satisfied for any given j if and only if

$$j\zeta_3 \leq (\zeta_1 + \zeta_3 x_2)(\zeta_2 + \zeta_3 x_1), \quad x_1, x_2 \geq 0,$$

which is equivalent to $\zeta_3 \leq (1/j)\zeta_1\zeta_2$. Since it must be true for $1 \leq j \leq n$, the global condition becomes (6.2). The $(2, n)$-monotonicity of S, $n \geq 2$, requires that the condition (6.3) is also satisfied when $(i, j) = (2, 0), \ldots, (2, n)$. First, consider the case $j \geq 2$. Then, we have

$$S^{(2,j)}(x_1, x_2) = (-1)^j (\zeta_2 + \zeta_3 x_1)^{j-2} \\ [j(j-1)\zeta_3^2 - 2j\zeta_3(\zeta_1 + \zeta_3 x_2)(\zeta_2 + \zeta_3 x_1) + (\zeta_1 + \zeta_3 x_2)^2 (\zeta_2 + \zeta_3 x_1)^2] S(x_1, x_2),$$

so that (6.3) for any given j is fulfilled if and only if

$$j(j-1)\zeta_3^2 - 2j\zeta_3(\zeta_1 + \zeta_3 x_2)(\zeta_2 + \zeta_3 x_1) + (\zeta_1 + \zeta_3 x_2)^2 (\zeta_2 + \zeta_3 x_1)^2 \geq 0,$$

which can be rewritten as

$$j[(\zeta_1 + \zeta_3 x_2)(\zeta_2 + \zeta_3 x_1) - \zeta_3]^2 + (j^2 - 2j)\zeta_3^2 - (j-1)(\zeta_1 + \zeta_3 x_2)^2 (\zeta_2 + \zeta_3 x_1)^2 \geq 0.$$

Since $j^2 - 2j \geq 0$ (because $j \geq 2$), a sufficient condition is

$$j[(\zeta_1 + \zeta_3 x_2)(\zeta_2 + \zeta_3 x_1) - \zeta_3]^2 \geq (j-1)(\zeta_1 + \zeta_3 x_2)^2 (\zeta_2 + \zeta_3 x_1)^2, \quad x_1, x_2 \geq 0,$$

and taking the square root then leads to the condition $\zeta_3 \le [1-(1-1/j)^{1/2}]\zeta_1\zeta_2$. Since the upper bound is decreasing with j, the overall condition for $2 \le j \le n$ becomes $\zeta_3 \le [1-(1-1/n)^{1/2}]\zeta_1\zeta_2$. Now, consider the cases $j = 0$ and $j = 1$. For $(2, 0)$, there is no condition on the parameters. For $(2, 1)$, the condition is, as for $(1, 2)$, $\zeta_3 \le (1/2)\zeta_1\zeta_2$, and since $1/2 > 1-(1-1/n)^{1/2}$ (because $n \ge 2$), the condition $\zeta_3 \le [1-(1-1/n)^{1/2}]\zeta_1\zeta_2$ still applies. Finally, it remains to combine this result with the two other cases $i = 0$ and $i = 1$. As $1-(1-1/n)^{1/2} \le 1/n$, we keep the same sufficient condition.

For the $(3, n)$-monotonicity of S, $n \ge 3$, we obtain the sufficient condition announced using a similar argument. Details are omitted. $\diamond$

If $S(x_1, x_2)$ is (n_1, n_2)-monotone, we can build a partially Schur-constant vector of survival function

$$P(\mathbf{X}_1 \ge \mathbf{x}_1, \mathbf{X}_2 \ge \mathbf{x}_2) = e^{-\zeta_1|\mathbf{x}_1|-\zeta_2|\mathbf{x}_2|-\zeta_3|\mathbf{x}_1||\mathbf{x}_2|}, \quad x_{j,i} \ge 0. \tag{6.4}$$

For example, when $(n_1, n_2) = (1, n)$, we know from Proposition 6.1 that $S(x_1, x_2)$ is $(1, n)$-monotone iff $\zeta_3 \le (1/n)\zeta_1\zeta_2$. Under this condition, the vector $\mathbf{X} = [X_1, \mathbf{X}_2 = (X_{2,1}, \dots, X_{2,n})]$ is $(1, n)$-partially Schur-constant with

$$P(X_1 \ge x_1, \mathbf{X}_2 \ge \mathbf{x}_2) = e^{-\zeta_1 x_1-(\zeta_2+\zeta_3 x_1)(x_{2,1}+\ldots+x_{2,n})}, \quad x_1, x_{2,i} \ge 0.$$

Each vector $\mathbf{X}_j$ is formed wih n_j independent exponentials of parameter ζ_j, so that $\mu_j = 1/\zeta_j$, $\sigma_j^2 = 1/\zeta_j^2$ and $\rho_j = 0$. The correlation $\rho_{1,2} \equiv \rho_{1,2}(\zeta)$ is given by the formula (2.12) of [17] in tems of the so-called logarithmic integral. Using the related exponential integral $E_1(x) = \int_x^\infty e^{-t}/t dt$ (see e.g. Chapter 5.1 in [1]), it becomes

$$\rho_{1,2}(\zeta) = -1 + (1/\zeta)e^{1/\zeta}E_1(1/\zeta),$$

which shows that $\rho_{1,2}(\zeta) \le 0$ and decreases with ζ from $\rho_{1,2}(0) = 0$ to $\rho_{1,2}(1) = -0.4036527$. Furthermore, we can use a known approximation for the function $xe^x E_1(x)$ (formula 5.1.54 in Chapter 5) to obtain

$$\rho_{1,2}(\zeta) \approx -1 + \frac{1/\zeta^2 + a_1/\zeta + a_2}{1/\zeta^2 + b_1/\zeta + b_2}, \quad \text{with an error of } 5/10^5,$$

where $a_1 = 2.334733$, $a_2 = 0.250621$, $b_1 = 3.330657$, $b_2 = 1.681534$. If the model (6.4) is well defined, the associated survival copula is given by

$$C(\mathbf{u}_1, \mathbf{u}_2) = \prod_{i=1}^{n_1} u_{1,i} \, (\prod_{i=1}^{n_2} u_{2,i})^{1-\zeta\sum_{i=1}^{n_1}\ln(u_{1,i})}, \quad u_{j,i} \in [0, 1].$$

Other cases of finitely monotone survival functions are provided by some distributions presented before, but this time asking that $S(x_1, x_2)$ be only (n_1, n_2)-monotone (instead of infinitely monotone). So, we have seen that the bivariate Lomax survival function exists if $\zeta \le 1 + \alpha$ and coresponds to a bivariate Laplace transform if $\zeta \le 1$. Now, it can be shown, for example, that $S(x_1, x_2)$ defined by (5.5) is $(1, n)$-monotone if and only if

$$\zeta \le 1 + \alpha/n, \quad n \ge 1,$$

while it is $(2, n)$-monotone ($n \ge 2$) under the stronger condition $\zeta \le \sqrt{1 + \alpha/n}$.

7 From bivariate Williamson transforms

By (2.8), a survival function $S(x_1, x_2)$ which is (n_1, n_2)-monotone can be represented as

$$S(x_1, x_2) = E\left[\left(1 - x_1/Z_1\right)_+^{n_1-1}\left(1 - x_2/Z_2\right)_+^{n_2-1}\right], \quad x_1, x_2 \ge 0, \tag{7.1}$$

where Z_j is distributed as the sum of the variables in group j. Thus, it suffices to choose any distribution for (Z_1, Z_2) to obtain a partially Schur-constant model.

In practice, however, the formula (7.1) turns out to be less convenient than it looks. Let us give two simple illustrations.

(1) Suppose that (Z_1, Z_2) is a continuous Schur-constant vector of univariate generator $\hat{S}(x)$. This means that (Z_1, Z_2) is distributed as the vector $Z(U, 1-U)$ where U is a uniform variable on $(0, 1)$ and Z is an independent variable distributed as $Z_1 + Z_2$ (with density $f_Z(z) = z\hat{S}^{(2)}(z)$) (see e.g. [13]). By insertion in (7.1), we get

$$S(x_1, x_2) = \int_{u=0}^{1} du \int_{z=g_u(x_1,x_2)}^{\infty} (1 - x_1/zu)^{n_1-1}\,[1 - x_2/z(1-u)]^{n_2-1} f_Z(z)dz,$$

where $g_u(x_1, x_2) = \max\{x_1/u, x_2/(1-u)\}$. Applying the binomial rule then leads to

$$\begin{aligned} S(x_1, x_2) &= \int_{u=0}^{1} du \int_{z=g_u(x_1,x_2)}^{\infty} \sum_{i=0}^{n_1-1} \binom{n_1-1}{i} (\frac{-x_1}{zu})^i \sum_{j=0}^{n_2-1} \binom{n_2-1}{j} (\frac{-x_2}{z(1-u)})^j f_Z(z)dz \\ &= \sum_{i=0}^{n_1-1}\sum_{j=0}^{n_2-1} \binom{n_1-1}{i}\binom{n_2-1}{j} (-x_1)^i(-x_2)^j \int_{u=0}^{1} \frac{1}{u^i(1-u)^j} du \int_{z=g_u(x_1,x_2)}^{\infty} z^{-i-j} f_Z(z)dz. \end{aligned} \tag{7.2}$$

For example, choose $\hat{S}(x) = \exp(-\lambda x)$, $\lambda > 0$. Thus, Z has an Erlang distribution of parameters $(2, \lambda)$ with density $f_Z(z) = \lambda^2 z \exp(-\lambda z)$. In that case, the last integral in (7.2) can be rewritten as

$$\lambda^2 \int_{z=g_u(x_1,x_2)}^{\infty} z^{-i-j+1} e^{-\lambda z} dz = (\frac{1}{\lambda})^{i+j}\,\Gamma[-i-j+2, \lambda g_u(x_1, x_2)],$$

where $\Gamma(a, x)$ is the upper incomplete gamma function ($= \int_x^\infty t^{a-1}e^{-t}dt$, for $x > 0$, $a \in \mathbb{R}$). Substituting this in formula (7.2) then gives us the expression of $S(x_1, x_2)$.

As another example, choose $1/Z$ of gamma distribution with parameter $1/\theta$, $\theta > 0$. Thus, Z is inverse gamma with density $f_Z(z) = [1/\Gamma(\theta)]z^{-\theta-1}\exp(-1/z)$. The last integral in (7.2) then becomes

$$\begin{aligned} \frac{1}{\Gamma(\theta)} \int_{z=g_u(x_1,x_2)}^{\infty} z^{-i-j-\theta-1} e^{-1/z} dz &= \frac{1}{\Gamma(\theta)} \int_{t=0}^{1/g_u(x_1,x_2)} t^{i+j+\theta-1} e^{-t} dt \\ &= \frac{1}{\Gamma(\theta)}\,\gamma[i + j + \theta, 1/g_u(x_1, x_2)], \end{aligned}$$

where $\gamma(a, x)$ is the lower incomplete gamma function ($= \int_0^x t^{a-1}e^{-t}dt$, for $x > 0$, $a > 0$).

Due to the complexity of the function S, the partially Schur-constant vector itself is difficult to use. Its moments and correlations, however, can be easily calculated. First, concerning the two variables Z_j, we apply (2.11) with $n_j = 2$ to get μ_{Z_j}, $\sigma^2_{Z_j}$, ρ_{Z_1,Z_2}; for example, $\mu_{Z_j} = \mu_Z/2$. Then, passing to the variables $X_{j,i}$ of the model, we use (2.11), (2.12), this time with (n_1, n_2), to determine μ_j, σ_j^2, ρ_j and $\rho_{1,2}$; for example, $\mu_j = \mu_{Z_j}/n_j = \mu_Z/2n_j$.

(2) Suppose that (Z_1, Z_2) is distributed as the bivariate gamma vector (Λ_1, Λ_2) of Laplace transform (5.4). From (7.1), we have

$$\begin{aligned} S(x_1, x_2) &= \int_{z_1=0}^{\infty}\int_{z_2=0}^{\infty} (1 - x_1/z_1)_+^{n_1-1}\,(1 - x_2/z_2)_+^{n_2-1} f_{(Z_1,Z_2)}(z_1, z_2)dz_1dz_2 \\ &= \sum_{i=0}^{n_1-1}\sum_{j=0}^{n_2-1} \binom{n_1-1}{i}\binom{n_2-1}{j} x_1^i x_2^j \int_{z_1=x_1}^{\infty}\int_{z_2=x_2}^{\infty} z_1^{-i} z_2^{-j} f_{(Z_1,Z_2)}(z_1, z_2)dz_1dz_2, \end{aligned} \tag{7.3}$$

in which $f_{(Z_1,Z_2)}$ is the corresponding density and is given by

$$f_{(Z_1,Z_2)}(z_1, z_2) = [1/\zeta_3^\alpha \Gamma(\alpha)]\,(z_1z_2)^{\alpha-1} e^{-(\zeta_2/\zeta_3)z_1-(\zeta_1/\zeta_3)z_2} f_\alpha(cz_1z_2),$$

where $c = (\zeta_1\zeta_2 - \zeta_3)/\zeta_3^2$ and $f_\alpha(z) = \sum_{k=0}^\infty z^k/\Gamma(\alpha+k)k!$ (see [10]). The double integral $\int_{z_1=x_1}^\infty \int_{z_2=x_2}^\infty$ in (7.3) is explicitly written as

$$\frac{1}{\zeta_3^\alpha \Gamma(\alpha)} \sum_{k=0}^{\infty} \frac{c^k}{\Gamma(\alpha+k)k!} \int_{z_1=x_1}^{\infty} \int_{z_2=x_2}^{\infty} z_1^{\alpha+k-i-1} z_2^{\alpha+k-j-1} e^{-(\zeta_2/\zeta_3)z_1-(\zeta_1/\zeta_3)z_2} dz_1 dz_2,$$

and the two integrals above can be factorized and then expressed as

$$(\frac{\zeta_3}{\zeta_2})^{\alpha+k-i}(\frac{\zeta_3}{\zeta_1})^{\alpha+k-j}\Gamma[\alpha+k-i,(\zeta_2/\zeta_3)x_1]\Gamma[\alpha+k-j,(\zeta_1/\zeta_3)x_2].$$

By inserting this in (7.3), we obtain the survival function $S(x_1, x_2)$.

The partially Schur-constant model is again little convenient. Still, as previously indicated, the formulas for μ_{Z_j}, $\sigma^2_{Z_j}$, ρ_{Z_1,Z_2} are particularly simple in the case of a bivariate gamma vector (Z_1, Z_2). Thus, for the variables $X_{j,i}$ of the model, we easily obtain from (2.11), (2.12) the parameters μ_j, σ_j^2, ρ_j and $\rho_{1,2}$; for example, $\mu_j = \mu_{Z_j}/n_j = \alpha\zeta_j/n_j$.

8 Building nested and multi-level models

Partially exchangeable models provide a flexible framework for representing the dependence between groups of random variables. We briefly show below that this approach covers the well-known nested Archimedean copulas, and that it can be easily generalized to model dependence on several levels.

Nested Schur-constant models. Hierarchical Archimedean copulas are an alternative to overcome the symmetry present in simple Archimedean copulas. A popular hierarchical method consists in nesting several Archimedean copulas one inside the other, with a certain nesting constraint in order to obtain a proper copula. It was originally developed by Joe [20] and studied in detail by e.g. [18] and [33]. We show below that this approach, although different, has links with the partially Archimedean copulas discussed here.

Let us partition the vector $\mathbf{U} = (U_1, \ldots, U_n)$ as in Section 3, i.e.

$$\mathbf{U} = (\mathbf{U}_1, \ldots, \mathbf{U}_m), \quad \text{where} \quad \mathbf{U}_j = (U_{j,1}, \ldots, U_{j,n_j}), \quad 1 \le j \le m.$$

A nested Archimedean copula with 1 level and m children is defined by

$$\begin{aligned} C(\mathbf{u}_1, \ldots, \mathbf{u}_m) &= C_0[C_1(\mathbf{u}_1), \ldots, C_m(\mathbf{u}_m)] \\ &= \psi_0\left\{\psi_0^{-1}\left[\psi_1\left(|\psi_1^{-1}(\mathbf{u}_1)|\right)\right] + \ldots + \psi_0^{-1}\left[\psi_m\left(|\psi_m^{-1}(\mathbf{u}_m)|\right)\right]\right\}, \end{aligned} \tag{8.1}$$

where ψ_0 is the generator of the parent copula C_0, and $\psi_1, \ldots, \psi_m$ are the generators of the m child copulas $C_1, \ldots, C_m$.

Now, consider a partially Archimedean copula (3.4) of generator $\psi(x_1, \ldots, x_m)$, and let $\psi_j(x_j)$, $1 \le j \le m$, be the m marginals of ψ.

Proposition 8.1. *If ψ can be expressed in the particular form*

$$\psi(x_1, \ldots, x_m) = \psi_0[\psi_0^{-1}(\psi_1(x_1)) + \ldots + \psi_0^{-1}(\psi_m(x_m))], \quad x_1, \ldots, x_m \ge 0, \tag{8.2}$$

where the function $\psi_0(x) : \mathbb{R}_+ \to [0,1]$ satisfies $\psi_0^{-1}(\psi(0, \ldots, 0)) = 0$, then the copula (3.4) corresponds to the nested Archimedean copula (8.1).

Proof. By (3.4), the partially Archimedean copula is defined as $\psi[|\psi_1^{-1}(\mathbf{u}_1)|, \ldots, |\psi_m^{-1}(\mathbf{u}_m)|]$. Inserting the expression (8.2) for ψ then gives the copula (8.1). Note that $\psi_0^{-1}(\psi(0, \ldots, 0)) = 0$ guarantees that ψ_j, $1 \le j \le m$, are the marginals of ψ as it was supposed. ◇

We emphasize that (8.2) is a special form for ψ, not the general form. In addition, since a partially Archimedean generator is necessarily $(n_1, \ldots, n_m)$-monotone, the function ψ_0 in (8.2) must be such that this condition is effectively fulfilled, which is not an easy task to verify. Partially Archimedean copulas seem to be more convenient to handle.

By Proposition 3.3 (ii), the partially Schur-constant model $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_m)$ whose generator $S = \psi$ is given by (8.2) is provided by the vector

$$[\mathbf{X}_1 = \psi_1^{-1}(1 - \mathbf{U}_1), \ldots, \mathbf{X}_m = \psi_m^{-1}(1 - \mathbf{U}_m)]. \tag{8.3}$$

In other words, (8.3) represents the Schur-constant model associated with the nested Archimedean copula (8.1).

Multi-level Schur-constant models. Partially Schur-constant models can be generalized to describe a form of dependence on several levels. This leads us to the introduction of a notion of partial exchangeability at several levels which, to our knowledge, is not standard in the literature.

Specifically, consider 2 levels. The vector $\mathbf{X}$ is partitioned into l groups which are partially exchangeable (instead of simply exchangeable). Each group k, $1 \le k \le l$, is formed of m_k subgroups (k, j) of $n_{k,j}$ random variables, hence the notation

$$(\mathbf{X}_{k,1}, \ldots, \mathbf{X}_{k,m_k}) \text{ where } \mathbf{X}_{k,j} = (X_{k,j,1}, \ldots, X_{k,j,n_{k,j}}), \quad 1 \le j \le m_k.$$

The whole vector $\mathbf{X} = (X_1, \ldots, X_n)$ is thus represented as

$$\mathbf{X} = [(\mathbf{X}_{1,1}, \ldots, \mathbf{X}_{1,m_1}), \ldots, (\mathbf{X}_{l,1}, \ldots, \mathbf{X}_{l,m_l})],$$

where, of course, $n = n_{1,1} + \ldots + n_{1,m_1} + \ldots + n_{l,1} + \ldots + n_{l,m_l}$.

Let us introduce a multivariate survival function $S[(x_{1,1}, \ldots, x_{1,m_1}), \ldots, (x_{l,1}, \ldots, x_{l,m_l})] : \mathbb{R}_+^{m_1+\ldots+m_l} \to [0, 1]$. We define the joint survival function of $\mathbf{X}$ by

$$\begin{aligned} P[(\mathbf{X}_{1,1} > \mathbf{x}_{1,1}, \ldots, \mathbf{X}_{1,m_1} > \mathbf{x}_{1,m_1}), \ldots, (\mathbf{X}_{l,1} > \mathbf{x}_{l,1}, \ldots, \mathbf{X}_{l,m_l} > \mathbf{x}_{l,m_l})] \\ = S\left[(|\mathbf{x}_{1,1}|, \ldots, |\mathbf{x}_{1,m_1}|), \ldots, (|\mathbf{x}_{l,1}|, \ldots, |\mathbf{x}_{l,m_l}|)\right], \end{aligned} \tag{8.4}$$

using the previous notation

$$|\mathbf{x}_{k,j}| = x_{k,j,1} + \ldots + x_{k,j,n_{k,j}}, \quad 1 \le k \le l, 1 \le j \le m_k.$$

We also consider the marginals of the l groups in the function S, with the notation

$$S_k(x_{k,1}, \ldots, x_{k,m_k}) = S[(0, \ldots, 0), \ldots, (0, \ldots, 0), (x_{k,1}, \ldots, x_{k,m_k}), (0, \ldots, 0), \ldots, (0, \ldots, 0)].$$

Then, we have

$$P(\mathbf{X}_{k,1} > \mathbf{x}_{k,1}, \ldots, \mathbf{X}_{k,m_k} > \mathbf{x}_{k,m_k}) = S_k(|\mathbf{x}_{k,1}|, \ldots, |\mathbf{x}_{k,m_k}|), \quad 1 \le k \le l,$$

which is asked to be a partially Schur-constant vector. Thus, each function $S_k : \mathbb{R}_+^{m_k} \to [0, 1]$ is $(n_{k,1}, \ldots, n_{k,m_k})$-monotone and admits the representation (2.8).

Now, we construct a partially Archimedean survival copula associated with the model (8.4). For that, we introduce again a function $\psi[(x_{1,1}, \ldots, x_{1,m_1}), \ldots, (x_{l,1}, \ldots, x_{l,m_l})] : \mathbb{R}_+^{m_1+\ldots+m_l} \to [0, 1]$ whose marginals are denoted by

$$\psi_{k,j}(x_{k,j}) = \psi(0, \ldots, 0, x_{k,j}, 0, \ldots, 0),$$

and the marginals of the l groups by

$$\psi_k(x_{k,1}, \ldots, x_{k,m_k}) = \psi[(0, \ldots, 0), \ldots, (0, \ldots, 0), (x_{k,1}, \ldots, x_{k,m_k}), (0, \ldots, 0), \ldots, (0, \ldots, 0)].$$

We partition the vector $\mathbf{U} = (U_1, \ldots, U_n)$ as $\mathbf{X}$ into l groups formed of m_k subgroups (k, j) of sizes $n_{k,j}$. For 2 levels, a partially Archimedean copula $C(\mathbf{u})$ is then defined by

$$\begin{aligned} C[(\mathbf{u}_{1,1}, \ldots, \mathbf{u}_{1,m_1}), \ldots, (\mathbf{u}_{l,1}, \ldots, \mathbf{u}_{l,m_l})] \\ = \psi\left[\left(|\psi_{1,1}^{-1}(\mathbf{u}_{1,1})|, \ldots, |\psi_{1,m_1}^{-1}(\mathbf{u}_{1,m_1})|\right), \ldots, \left(|\psi_{l,1}^{-1}(\mathbf{u}_{l,1})|, \ldots, |\psi_{l,m_l}^{-1}(\mathbf{u}_{l,m_l})|\right)\right], \end{aligned} \tag{8.5}$$

where

$$|\psi_{k,j}^{-1}(\mathbf{u}_{k,j})| = \psi_{k,j}^{-1}(u_{k,j,1}) + \ldots + \psi_{k,j}^{-1}(u_{k,j,n_{k,j}}), \quad 1 \le k \le l,\ 1 \le j \le m_k.$$

When all the $\mathbf{u}_{i,j} = 0$ except in group k, $C(\mathbf{u})$ is reduced to

$$C_k(\mathbf{u}_{k,1}, \ldots, \mathbf{u}_{k,m_k}) = \psi_k \left[|\psi_{k,1}^{-1}(\mathbf{u}_{k,1})|, \ldots, |\psi_{k,m_k}^{-1}(\mathbf{u}_{k,m_k})| \right], \quad 1 \le k \le l,$$

in agreement with the definition (3.4).

Finally, compare with a nested Archimedean copula for 2 levels defined by

$$\begin{aligned} &C[(\mathbf{u}_{1,1}, \ldots, \mathbf{u}_{1,m_1}), \ldots, (\mathbf{u}_{l,1}, \ldots, \mathbf{u}_{l,m_l})] \\ &\qquad = \psi_0\{\psi_0^{-1}[C_1(\mathbf{u}_{1,1}, \ldots, \mathbf{u}_{1,m_1})] + \ldots + \psi_0^{-1}[C_l(\mathbf{u}_{l,1}), \ldots, C_{l,m_l}(\mathbf{u}_{l,m_l})]\}, \end{aligned} \tag{8.6}$$

in which $\psi_0 : \mathbb{R}_+ \to [0, 1]$ is the generator of the copula for the l groups, and each of them is of copula defined as in (8.1) by

$$\begin{aligned} C_k(\mathbf{u}_{k,1}, \ldots, \mathbf{u}_{k,m_k}) &= C_{k,0}\left[C_{k,1}(\mathbf{u}_{k,1}), \ldots, C_{k,m_k}(\mathbf{u}_{k,m_k})\right] \\ &= \psi_{k,0}\left\{\psi_{k,0}^{-1}\left[\psi_{k,1}(|\psi_{k,1}^{-1}(\mathbf{u}_{k,1})|)\right] + \ldots + \psi_{k,0}^{-1}\left[\psi_{k,m_k}(|\psi_{k,m_k}^{-1}(\mathbf{u}_{k,m_k})|)\right]\right\}, \quad 1 \le k \le l, \end{aligned}$$

where $\psi_{k,0} : \mathbb{R}_+^{\mathrm{m}_k} \to [0, 1]$ and $\psi_{k,j} : \mathbb{R}_+ \to [0, 1]$, $j = 1, \ldots, m_k$, are the generators of the parent and child copulas, respectively. It is easily seen that Proposition 8.1 can be extended to 2 levels. In other words, the nested Archimedean copula (8.6) is a partially Archimedean copula (8.5) for which the generator ψ has a particular expression generalizing (8.2).

Acknowledgements: I thank the editor and the referees for their valuable comments. I am grateful to M.M. Claramunt (Universitat de Barcelona) and S. Loisel (ISFA, Université de Lyon) for interesting and stimulating discussions. This research received partial support via the *DIALog* Chair sponsored by CNP Assurances.

Conflict of interest statement: Author states no conflict of interest.

References

[1] Abramowitz, M. and I. A. Stegun (1964). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover, New York.

[2] Arnold, B. C. (1990). A flexible family of multivariate Pareto distributions. *J. Statist. Plann. Inference 24*(2), 249–258.

[3] Balakrishnan, N. and C.-D. Lai (2009). *Continuous Bivariate Distributions*. Second edition. Springer, New York.

[4] Barlow, R. E. and M. B. Mendel (1993). Similarity as a probabilistic characteristic of aging. In R. E. Barlow, C. A. Clarotti, and F. Spizzichino (Eds.), *Reliability and Decision Making*, pp. 233–244. Chapman & Hall, London.

[5] Bernardoff, P. (2018). Marshall-Olkin Laplace transform copulas of multivariate gamma distributions. *Comm. Statist. Theory Methods 47*(3), 655–670.

[6] Caramellino, L. and F. Spizzichino (1994). Dependence and aging properties of lifetimes wih Schur-constant survival functions. *Probab. Engrg. Inform. Sci. 8*(1), 103–111.

[7] Castañer, A. and M. M. Claramunt (2019). Equilibrium distributions and discrete Schur-constant models. *Methodol. Comput. Appl. Probab. 21*, 449–459.

[8] Castañer, A., M. M. Claramunt, C. Lefèvre, and S. Loisel (2015). Discrete Schur-constant models. *J. Multivariate Anal. 140*, 343–362.

[9] Castañer, A., M. M. Claramunt, C. Lefèvre, and S. Loisel (2019). Partially Schur-constant models. *J. Multivariate Anal. 172*, 47–58.

[10] Chatelain, F. (2007). *Lois Gamma Multivariées pour le Traitement d'Images Radar*. PhD thesis, Institut National Polytechnique de Toulouse, France.

[11] Chatelain, F., J.-Y. Tourneret, and J. Inglada (2008). Change of detection in multisensor SAR images using bivariate gamma distributions. *IEEE Trans. Image Process. 17*(3), 249–258.

[12] Chatelain, F., J.-Y. Tourneret, J. Inglada, and A. Ferrari (2007). Bivariate gamma distributions for image registration and change detection. *IEEE Trans. Image Process. 16*(7), 1796–1806.

[13] Chi, Y., J. Yang, and Y. Qi (2009). Decompositon of a Schur-constant model and its applications. *Insurance Math. Econom. 44*(3), 398–408.

[14] Cossette, H., S.-P. Gadoury, E. Marceau, and I. Mtalai (2017). Hierarchical Archimedean copulas through multivariate compound distributions. *Insurance Math. Econom. 76*, 1–13.

[15] de Finetti, B. (1930). Funzione caratteristica di un fenomeno aleatorio. *Mem. R. Accad. Naz. Lincei 4*, 86–133.

[16] de Finetti, B. (1938). Sur la condition d'équivalence partielle. In *Actualités Scientifiques et Industrielles 739*, 5–18. Hermann & Cie, Paris.

[17] Gumbel, E. J. (1960). Bivariate exponential distributions. *J. Amer. Statist. Assoc. 55*(292), 698–707.

[18] Hofert, M. (2011). Efficiently sampling nested Archimedean copulas. *Comput. Statist. Data Anal. 55*(1), 57–70.

[19] Hougaard, P. (1986). A class of multivariate failure time distributions. *Biometrika 73*(3), 671–678.

[20] Joe, H. (1997). *Multivariate Models and Dependence Concepts*. Chapman & Hall, London.

[21] Jones, M. C. and É. Marchand (2019). Multivariate discrete distributions via sums and shares. *J. Multivariate Anal. 171*, 83–93.

[22] Khintchine, A. Y. (1938). On unimodal distributions. *Izv. Nauch. Inst. Mat. Mekh. 2*, 1–7 (in russian).

[23] Kotz, S., N. Balakrishnan, and N. L. Johnson (2000). *Continuous Multivariate Distributions*. Second edition. John Wiley & Sons, New York.

[24] Lai, C. D., M. Xie, and I. G. Bairamov (2001). Dependence and ageing properties of bivariate Lomax distribution. In Y. Hayakawa, T. Irony, and M. Xie (Eds.), *System and Bayesian Reliability*, pp. 243–256. World Scientific, Singapore.

[25] Lee, L. (1979). Multivariate distributions having Weibull properties. *J. Multivariate Anal. 9*(2), 267–277.

[26] Lee, C. and M.-J. Wen (2009). A multivariate Weibull distribution. *Pakistan J. Statist. Oper. Res. 5*(2), 55–66.

[27] Lefèvre, C. and S. Loisel (2013). On multiply monotone distributions, continuous or discrete, with applications. *J. Appl. Probab. 50*(3), 827–847.

[28] Lefèvre, C., S. Loisel, and S. Utev (2018). Markov property in discrete Schur-constant models. *Methodol. Comput. Appl. Probab. 20*, 1003–1012.

[29] Lefèvre, C. and M. Simon (2021). Schur-constant and related dependence models, with application to ruin probabilities. *Methodol. Comput. Appl. Probab. 23*, 317–339.

[30] Letac, G., J.-Y. Tourneret, and F. Chatelain (2007). The bivariate gamma distribution with Laplace transform $(1+as+bt+cst)^{-q}$: history, characterizations, estimation. Available at http://www.math.univ-toulouse.fr/~letac/bivgamma.pdf.

[31] Lindley, D. V. and N. D. Singpurwalla (1986). Multivariate distributions for the life lengths of components of a system sharing a common environment. *J. Appl. Probab. 23*(2), 418–431.

[32] Lu, J.-C. and G. K. Bhattacharyya (1990). Some new constructions of bivariate Weibull models. *Ann. Inst. Statist. Math. 42*, 543–559.

[33] McNeil, A. J. (2008). Sampling nested Archimedean copulas. *J. Stat. Comput. Simul. 78*(6), 567–581.

[34] McNeil, A. J. and J. Nešlehová (2009). Multivariate Archimedean copulas, d-monotone functions and l_1-norm symmetric distributions. *Ann. Statist. 37*(5B), 3059–3097.

[35] Nair, N. U. and P. G. Sankaran (2014). Modelling lifetimes with bivariate Schur-constant equilibrium distributions from renewal theory. *Metron 72*, 331–349.

[36] Nair, N. U, P. G. Sankaran, and N. Balakrishnan (2018). *Reliability Modelling and Analysis in Discrete Time*. Academic, San Diego.

[37] Nayak, T. K. (1987). Multivariate Lomax distribution: properties and usefulness in reliability theory. *J. Appl. Probab. 24*(1), 170–177.

[38] Nelsen, R. B. (2005). Some properties of Schur-constant survival models and their copulas. *Braz. J. Probab. Stat. 19*(2), 179–190.

[39] Nelsen, R. B. (2006). *An Introduction to Copulas*. Second edition. Springer, New York.

[40] Ressel, P. (2011). Monotonicity properties of multivariate distribution and survival functions - with an application to Lévy-frailty copulas. *J. Multivariate Anal. 102*(3), 393–404.

[41] Ressel, P. (2014). Higher order monotonic functions of several variables. *Positivity 18*, 257–285.

[42] Ressel, P. (2018). A multivariate version of Williamson's theorem, l^1-symmetric survival functions, and generalized Archimedean copulas. *Depend. Model. 6*, 356–368.

[43] Sankaran, P. G. and D. Kundu (2014). On a bivariate Pareto model. *Statistics 48*(2), 241–255.

[44] Sankaran, P. G. and N. U. Nair (1993). A bivariate Pareto model and its applications to reliability. *Naval Res. Logist. 40*(7), 1013–1020.

[45] Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris 8*, 229–231.

[46] Ta, B. Q., D. S. Le, M. B. Ha, and X. D. Tran (2018). On characterizations of bivariate Schur-constant models and applications. In L. H. Anh, L. S. Dong, V. Kreinovich, and N. N. Thach (Eds.), *Econometrics for Financial Applications*, pp. 890–901. Springer, Cham.

[47] Ta, B. Q. and C. P. Van (2017). Some properties of bivariate Schur-constant distributions. *Statist. Probab. Lett. 124*, 69–76.

[48] Takahasi, K. (1965). Note on multivariate Burr's distribution. *Ann. Inst. Statist. Math. 17*, 257–260.

[49] Williamson, R. E. (1956). Multiply monotone functions and their Laplace transforms. *Duke Math. J. 23*(2), 189–207.

[50] Yeh, H.-C. (2004). Some properties and characterizations for generalized multivariate Pareto distributions. *J. Multivariate Anal. 88*(1), 47–60.

DE GRUYTER Depend. Model. 2021; 9:243–266

Research Article **Open Access**
Special Issue in memory of Abe Sklar

Piotr Jaworski* and Marcin Krzywda

On copulas of self-similar Ito processes

https://doi.org/10.1515/demo-2021-0112
Received February 13, 2021; accepted May 11, 2021

Abstract: We characterize the cumulative distribution functions and copulas of two-dimensional self-similar Ito processes, with randomly correlated Wiener margins, as solutions of certain elliptic partial differential equations.

Keywords: copula, Ito diffusion, stochastic differential equations, self-similar processes, elliptic partial differential equations

MSC: 62H05, 60G18, 60H10, 60E05, 60J60

1 Introduction

The paper deals with the borderline of the copula theory and stochastic processes. It concerns the vector valued stochastic processes $X_t = (X_t^1, \dots, X_t^n)$, $t \in T$. The goal is to describe the evolution of interdependences between $X_t^1, \dots, X_t^n$ in terms of copulas and copula processes. In a way it is a continuation of papers by Sempi [31], Choe etal. [7], Jaworski and Krzywda [16] and Jaworski [15]. Sempi in ([31]) was studying the possibility of coupling two Wiener processes by using a given copula. In [16] we started to investigate copulas of self-similar processes. In [7] the partial differential equation for copulas of the Ito processes is derived but under very tight technical assumptions. In [15] it is achieved but in a general case. It is shown that the copula process is a weak solution of a parabolic partial differential equation.

In this paper we study the evolution of the dependence between two Wiener processes with random quadratic covariation. Specifically we extend the research on two-dimensional 1/2-self-similar Ito diffusions whose margins are Wiener processes and their interdependencies at every time moment are described by a given copula, that we have initiated in [16]. We drop the assumption of differentiability of cumulative distribution functions and corresponding copulas. We replace the classical derivatives by weak derivatives in Sobolev sense, so called distributional derivatives, as in [15]. This allows us to present new and more general results. We consider 1/2-self-similar processes as solutions of certain stochastic differential equations (SDE) and show how to construct 1/2-self-similar solution basing on a non-self-similar solution.

In more details, we consider a pair of stochastic processes $(X_t, Y_t)_{t\geq 0}$, where $X = (X_t)_{t\geq 0}$ and $Y = (Y_t)_{t\geq 0}$ are one dimensional Wiener processes and their quadratic covariation is homogeneous for $t \geq 1$

$$d\langle X, Y\rangle_t = A\left(\frac{X_t}{\sqrt{t}}, \frac{Y_t}{\sqrt{t}}\right) dt, \tag{1.1}$$

where the real valued function A is defined on the real plane and $|A|$ is a bounded by 1. Such processes arise as solutions of the stochastic differential equations (SDE) with homogeneous coefficients

$$\begin{aligned} dX_t &= dW_t^1, \\ dY_t &= A\left(\frac{X_t}{\sqrt{t}}, \frac{Y_t}{\sqrt{t}}\right) dW_t^1 + B\left(\frac{X_t}{\sqrt{t}}, \frac{Y_t}{\sqrt{t}}\right) dW_t^2, \end{aligned} \tag{1.2}$$

***Corresponding Author: Piotr Jaworski:** Institute of Mathematics, University of Warsaw, Warszawa, Poland, ORCID: 0000-0001-8540-2698, E-mail: P.Jaworski@mimuw.edu.pl
Marcin Krzywda: Institute of Mathematics, Jagiellonian University, Kraków, Poland, E-mail: marcin@krzywda.net

where $B(x, y) = \sqrt{1 - A(x, y)^2}$.

We begin with showing that for A and B Lipschitz, any pair of Wiener processes fulfilling 1.2 gives rise to 1/2-self similar solutions. They appear as limits of time-rescaled (time-shifted) processes $(X_{t\tau}/\sqrt{\tau}, Y_{t\tau}/\sqrt{\tau})_{t\geq 1}$ when τ tends to infinity or as weighted generalized mixtures of time-rescaled processes. Then we derive a partial differential equation (PDE) (satisfied in a weak sense) that describes the copula of the initial value, i.e. the copula C of the random pair (X_1, Y_1), giving rise to 1/2-self similar solution.

$$\begin{aligned}&\varphi(\Phi^{-1}(u))^2 D^{2,0}C(u, v) + \varphi(\Phi^{-1}(v))^2 D^{0,2}C(u, v)\\&+2A(\Phi^{-1}(u), \Phi^{-1}(v))\varphi(\Phi^{-1}(u))\varphi(\Phi^{-1}(v))D^{1,1}C(u, v) = 0,\end{aligned} \tag{1.3}$$

where Φ and φ denote the cumulative distribution function and the density of the standard normal distribution $N(0, 1)$ and $D^{i,j}$ denote the partial derivatives. We also present the opposite result by which if any copula C fulfills the previously obtained PDE in a weak sense with coefficients satisfying certain regularity conditions then there exists a 1/2-self-similar process $(X_t, Y_t)_{t\geq 1}$ such that the copula of $(X_t, Y_t)_{t\geq 1}$ is equal to C. Furthermore basing on the maximum principle for the elliptic PDE we show that the point-wise dominance of the quadratic covariations A implies the concordance dominance of the corresponding copulas.

The structure of the paper is as follows: On the start we shortly recall basics on copulas, self-similarity and weak derivatives (section 2). In section 3 we introduce the underlying stochastic processes X_t and Y_t and state the results concerning their self-similarity. First concerning the existence of 1/2-self-similar solutions among solutions of certain stochastic differential equations, next concerning their differential characterization. The proofs are provided in section 4. In the last section we discuss some examples based on Gaussian, FGM and Archimedean copulas.

2 Notation

We recall the basic concepts.

A (bivariate) *copula* is a restriction to $[0, 1]^2$ of a distribution function whose univariate margins are uniformly distributed on $[0, 1]$. Specifically, $C\colon [0, 1]^2 \to [0, 1]$ is a copula if it satisfies the following properties:
(C1) $C(x, 0) = C(0, x) = 0$ for every $x \in [0, 1]$,
(C2) $C(x, 1) = C(1, x) = x$ for every $x \in [0, 1]$,
(C3) C is 2-*increasing*, that is, the C-volume V_C of any rectangle $[x_1, x_2] \times [y_1, y_2]$ of $[0, 1]^2$ is positive, i.e.

$$V_C([x_1, x_2] \times [y_1, y_2]) = C(x_1, y_1) + C(x_2, y_2) - C(x_1, y_2) - C(x_2, y_1) \geq 0.$$

Due to the celebrated *Sklar's Theorem*, the joint distribution function F of any pair (X, Y) of random variables defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ can be written as a composition of a copula C and the univariate marginals F_1 and F_2, i.e. for all $(x, y) \in \mathbb{R}^2$, $F(x, y) = C(F_1(x), F_2(y))$. Moreover, if X, Y are random variables with continuous cumulative distribution functions, then the copula C is uniquely determined.
We will denote the set of all two-dimensional copulas by $\mathcal{C}^2$. Note, that $\mathcal{C}^2$ is a bounded, compact, convex subset of the Banach space of the continuous functions on the unit square endowed with the supremum metric. For copulas C_1 and C_2 we put

$$||C_2 - C_1||_\infty = \sup\{|C_2(u, v) - C_1(u, v)| : (u, v) \in [0, 1]^2\}. \tag{2.1}$$

Specifically any sequence of copulas contains a convergent subsequence, i.e.

$$\forall (C_n)_{n=1}^\infty \quad \exists C_\star \quad \exists (n_k)_{k=1}^\infty \quad n_k \to \infty, \quad ||C_{n_k} - C_\star||_\infty \to 0. \tag{2.2}$$

The copulas $C_\star$, being the limits of subsequences, are referred to as *cluster points* of the sequence C_n. For more details about copula theory and some of its applications, we refer to [6, 9, 17–20, 26].

A random process is self-similar if its distributions scale. Specifically (X_t, Y_t), $t \geq t_0 \geq 0$, is called H-self-similar (with $H \geq 0$) when

$$(X_{at}, Y_{at}) \sim a^H (X_t, Y_t) \qquad \text{for all } a \geq 1,$$

where $\sim$ denotes equality of joint distributions. We call H the exponent of self-similarity of the process. For example standard Brownian Motion is $1/2$-self-similar. For the details on self-similarity we refer to [10] and [33]. Please note that our definition is a bit weaker than the common literature definition, ie. we do not require the equality of distributions of stochastic processes but only those of random vectors at each time $t \geq t_0$.

We can re-write the definition of self-similarity as follows

$$F_t(x, y) = F(t^{-H}x, t^{-H}y),$$

where F_t is a cumulative distribution function of the pair (X_t, Y_t) and $F = F_{t_0}$. By Sklar's theorem applied to F

$$F_t(x, y) = C(\Phi(t^{-H}x), \Psi(t^{-H}y)),$$

where we use $\Phi = \Phi_{t_0}$ and $\Psi = \Psi_{t_0}$ to denote univariate cumulative distribution functions of X_{t_0} and Y_{t_0} which we assume to be continuous. Since the processes X_t and Y_t also are self-similar analogously we may write

$$\Phi_t(x) = \Phi(t^{-H}x), \qquad \Psi_t(y) = \Psi(t^{-H}y).$$

Once again by Sklar's theorem for each $t \geq t_0$ we obtain a copula C_t such that

$$F_t(x, y) = C_t(\Phi_t(x), \Psi_t(y)).$$

Therefore, by the uniqueness of copula for random variables with continuous distribution functions, we have $C_t(x, y) = C_{t_0}(x, y)$ for each $t \geq t_0$.

We conclude that, when X_t and Y_t have continuous distribution functions for each $t \geq t_0$, then the process $(X_t, Y_t)_{t \geq t_0}$ is self-similar if and only if its copula is constant and both $(X_t)_{t \geq t_0}$ and $(Y_t)_{t \geq t_0}$ are self-similar.

In the following we shall deal with $H = \frac{1}{2}$.

To characterize the copulas of 1/2-self-similar processes we will need to substantially weaken the notion of partial derivatives (see [4, 11, 13]).

Definition 2.1. *Suppose that U is an open subset of $\mathbb{R}^n$, functions u and v are locally integrable on U and $\alpha = (\alpha_1, \dots, \alpha_n)$ is a multi-index.*

We say that v is α^{th} weak partial derivative of u, written

$$D^\alpha u = v,$$

provided that

$$\int_U u(x) \frac{\partial^{|\alpha|} h(x)}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}} dx_1 \dots dx_n = (-1)^{|\alpha|} \int_U v(x) h(x) dx_1 \dots dx_n$$

for all test functions $h \in C^\infty(U)$ with compact support and $|\alpha| = \alpha_1 + \dots + \alpha_n$.

Weak partial derivatives, if they exist, are unique up to a set of measure zero. Note that, since copulas are Lipschitz functions, they are weakly differentiable. As an example of weak partial derivatives of a copula $C(u, v)$ one may consider one of the Dini derivatives ([8, 24]), say right-side upper one, applied respectively to both variables

$$D^{1,0}\, C(u,v) = \limsup_{h\to 0^+} \frac{C(u+h,v)-C(u,v)}{h}, \quad \text{for} \quad u\in[0,1), v\in[0,1], \tag{2.3}$$

$$D^{0,1}\, C(u,v) = \limsup_{h\to 0^+} \frac{C(u,v+h)-C(u,v)}{h}, \quad \text{for} \quad u\in[0,1], v\in[0,1). \tag{2.4}$$

The choice of the version of the weak derivative is related to the choice of the version of the conditional probability. In more details, if C is a copula of the random variables X and Y, then the formula

$$F_{X|Y=y}(x) = \lim_{\xi\to 0^+} D^{0,1}C(F_X(\xi), F_Y(y)), \tag{2.5}$$

expressing the conditional distribution function of X in terms of the right continuous modification of the weak derivative (2.4), determines the version of the conditional probability $P(\cdot|\sigma(Y))$ on the σ-field $\sigma(X)$ (compare [2, 14]). Since the modifications occur only at "jumps" of $D^{0,1}C(F_X(\cdot), F_Y(y))$, we have for fixed y, such that $F_Y(y) < 1$,

$$F_{X|Y=y}(x) \overset{ae}{=} D^{0,1}C(F_X(x), F_Y(y)). \tag{2.6}$$

3 Main results

We consider the solutions $(X_t, Y_t)_{t\geq 1}$ of the following system of stochastic differential equations:

$$\begin{aligned} dX_t &= dW_t^1, \\ dY_t &= A\left(\frac{X_t}{\sqrt{t}}, \frac{Y_t}{\sqrt{t}}\right) dW_t^1 + B\left(\frac{X_t}{\sqrt{t}}, \frac{Y_t}{\sqrt{t}}\right) dW_t^2, \end{aligned} \tag{3.1}$$

where
A1. W^1 and W^2 are two independent Wiener processes defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$,
A2. the functions $A, B : \mathbb{R}^2 \to \mathbb{R}$ are Lipschitz,
A3. for all $(x,y) \in \mathbb{R}^2$ we have $B(x,y) = \sqrt{1-A(x,y)^2}$ and $|A(x,y)| < 1$,
A4. A is differentiable with respect to the second variable.

By M_A and L_A we shall denote respectively the supremum of the modulus and the Lipschitz coefficient of A with respect to the euclidean distance

$$M_A = \sup\{|A(x,y)| : (x,y)\in\mathbb{R}^2\} \tag{3.2}$$

$$L_A = \sup\left\{ \frac{|A(x_2,y_2)-A(x_1,y_1)|}{\sqrt{(x_2-x_1)^2+(y_2-y_1)^2}} : (x_1,y_1)\neq(x_2,y_2)\right\} \tag{3.3}$$

Furthermore we denote by L the square of the Lipschitz coefficient of the vector (A, B)

$$L = \sup\left\{ \frac{|A(x_1,y_1)-A(x_2,y_2)|^2 + |B(x_1,y_1)-B(x_2,y_2)|^2}{(x_2-x_1)^2+(y_2-y_1)^2} : (x_1,y_1)\neq(x_2,y_2)\right\} \tag{3.4}$$

Note that when $M_A < 1$, we have a bound

$$L \leq \frac{L_A^2}{1-M_A^2}. \tag{3.5}$$

In fact we understand (3.1) as an integral equation:

$$\begin{aligned} X_t &= X_1 + W_t^1 - W_1^1, \\ Y_t &= Y_1 + \int_1^t A\left(\frac{X_s}{\sqrt{s}}, \frac{Y_s}{\sqrt{s}}\right) dW_s^1 + \int_1^t B\left(\frac{X_s}{\sqrt{s}}, \frac{Y_s}{\sqrt{s}}\right) dW_s^2, \end{aligned} \tag{3.6}$$

where the initial values, i.e. the random pair (X_1, Y_1), and the two-dimensional Wiener process $(W_t^1 - W_1^1, W_t^2 - W_1^2)_{t\geq 1}$ are independent. The process $(X_t, Y_t)_{t\geq 1}$ is adapted to the filtration defined by the Wiener processes W^1 and W^2 and the initial values

$$\mathcal{F}_t = \sigma(X_1, Y_1, W_s^1 - W_1^1, W_s^2 - W_1^2, s \in [1, t]). \tag{3.7}$$

Assumption **A1.** ensures that the Ito integrals in (3.6) are well defined and the quadratic covariation of any solution of SDE (3.1) $(X, Y) = (X_t, Y_t)_{t\geq 1}$ fulfills

$$d\langle X, Y\rangle_t = A\left(\frac{X_t}{\sqrt{t}}, \frac{Y_t}{\sqrt{t}}\right) dt. \tag{3.8}$$

The second one, **A2.**, implies the existence and uniqueness of the solution of SDE (3.1). Due to **A3.**, the quadratic variation of $Y = (Y_t)_{t\geq 1}$ is given by

$$d\langle Y, Y\rangle_t = (A^2 + B^2)dt = 1 \cdot dt, \tag{3.9}$$

which implies that $(Y_{t+1} - Y_1)_{t\geq 0}$ is a Wiener process.

Our goal is to study the existence and properties of $1/2$-self-similar solutions of SDE (3.1). Note that the self-similarity of the process $(X_t, Y_t)_{t\geq 1}$ does not depend on the choice of the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ neither on the choice of the Wiener processes. It depends only on the joint distribution of the of the pair (X_1, Y_1). It is a corollary from the uniqueness of the weak solutions of the SDE, see [5]. In more details, when the seven tuple

$$\left(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}}, (\widetilde{X_1}, \widetilde{Y_1}), (\widetilde{W_t^1}, \widetilde{W_t^2})_{t\geq 1}, (\widetilde{X}_t, \widetilde{Y}_t)_{t\geq 1}\right)$$

is a weak solution of equations (3.1) and

$$(\widetilde{X_1}, \widetilde{Y_1}) \sim (X_1, Y_1),$$

then

$$(\widetilde{X}_t, \widetilde{Y}_t) \sim (X_t, Y_t).$$

Please also take note of the fact that the margins $(X_t)_{t\geq 1}$ and $(Y_t)_{t\geq 1}$ are $1/2$-self-similar if and only if they coincide with one-dimensional Wiener processes.

Proposition 3.1. *The processes $(X_t)_{t\geq 1}$ and $(Y_t)_{t\geq 1}$ are $1/2$-self-similar if and only if the beginning values X_1 and Y_1 have the standard normal distribution*

$$X_1 \sim Y_1 \sim N(0, 1).$$

The proof is elementary, we refer to section 4.1. Throughout the rest of this paper we denote by Φ and φ the cumulative distribution function and density of the standard normal distribution $N(0, 1)$.

On the other hand, as it is shown in section 4.3, the set of $1/2$-self-similar solutions is not empty. Indeed for any functions A and B fulfilling **A2** and **A3** there exits a selfsimilar solution.

Theorem 3.2. *Assume **A1**, **A2** and **A3**, then for properly chosen initial values the SDE 3.1 have a $1/2$-self-similar solution.*

Such $1/2$-self-similar solutions arise as a weighted generalized mixtures of τ-rescaled processes $(X_{\tau t}/\sqrt{\tau}, Y_{\tau t}/\sqrt{\tau})_{t\geq 1},\ \tau > 1$.

When the functions A and $B = \sqrt{1 - A^2}$, i.e. the volatility of the solutions of (3.1), are varying in a moderate way then the probability law of "properly chosen" initial values can be described just as a limit with respect to convergence in distribution. Let C_t^C denote a copula of the pair (X_t, Y_t) for $t \geq 1$, where the process $(X_t, Y_t)_{t\geq 1}$ is a solution of eq. (3.1), with initial values X_1 and Y_1 having the standard normal distribution $X_1 \sim Y_1 \sim N(0, 1)$ and linked by a copula C.

Proposition 3.3. *If L, the square of the Lipschitz coefficient of (A, B), is smaller than 1, then for any copula C a solution of eq. (3.1), with initial values having a cumulative bivariate distribution $C^*(\Phi(x), \Phi(y))$, where*

$$C^*(u, v) = \lim_{t \to \infty} C_t^C(u, v), \tag{3.10}$$

is $1/2$-self-similar.

The proof is provided in section 4.3. Note that $L < 1$ when for example $L_A^2 + A_M^2 < 1$. Under assumption $L < 1$, Proposition 3.3 implies, that the rescalings (shiftings) of any solution of eq. (3.1) are converging in distribution to a self similar solution. This has a certain practical impact. When we observe some phenomena driven by a solution of SDE (3.1), with $L < 1$, for a sufficiently long time we may approximate the "true" solution by a 1/2-self-similar one. In the last part of the section 5.1 we provide an illustration of this effect.

After this short discussion on existence we state our results on differential characterization of copulas of $1/2$-selfsimilar solutions.

Theorem 3.4. *Assume A1,A2, A3 and A4, then:*
If the process $(X_t, Y_t)_{t \geq 1}$ is a $1/2$-self-similar solution of eq. (3.1), then the copula of (X_1, Y_1) which we denote by $C(u, v)$ is twice differentiable in a weak sense in $(0, 1)^2$ and almost everywhere in $(0, 1)^2$ fulfills the equation

$$\begin{aligned} &D^{2,0}C(u, v)\varphi(\Phi^{-1}(u))^2 + D^{0,2}C(u, v)\varphi(\Phi^{-1}(v))^2 \\ &+2A(\Phi^{-1}(u), \Phi^{-1}(v))D^{1,1}C(u, v)\varphi(\Phi^{-1}(u))\varphi(\Phi^{-1}(v)) = 0. \end{aligned} \tag{3.11}$$

the cumulative distribution function of (X_1, Y_1) which we denote by $F(x, y)$ is twice differentiable in a weak sense and fulfills the equation

$$\begin{aligned} &D^{2,0}F(x, y) + D^{0,2}F(x, y) + 2A(x, y)D^{1,1}F(x, y) \\ &\quad +xD^{1,0}F(x, y) + yD^{0,1}F(x, y) = 0 \end{aligned} \tag{3.12}$$

almost everywhere in $\mathbb{R}^2$.

As a consequence of the maximum principle for elliptic PDE, we get:

Theorem 3.5. *Let A_i, $i = 1, 2$, be bounded Lipschitz functions, $|A_i| < 1$. If copulas C_1 and C_2 fulfill PDE (3.11) with A equals respectively to A_1 and A_2 and*

$$\forall (x, y) \in \mathbb{R}^2 \quad A_1(x, y) \geq A_2(x, y),$$

then copula C_1 dominates in concordance ordering

$$\forall (u, v) \in [0, 1]^2 \quad C_1(u, v) \geq C_2(u, v),$$

Theorem 3.5 implies two important corollaries.

Corollary 3.6. *Any two $1/2$-self-similar solutions of eq. (3.1), with coefficients fulfilling **A2**, **A3** and **A4**, are law equivalent. Specifically, the copulas of the initial values of any two such $1/2$-self-similar solutions of eq. (3.1) are equal to each other.*

Corollary 3.7. *Let a twice weakly differentiable copula C be a solution of PDE (3.11) with coefficient A and Ga_ρ, $\rho \in (-1, 1)$, be a Gaussian copula with correlation coefficient ρ.*
• *If $A(x, y) \leq \rho$ for all $(x, y) \in \mathbb{R}^2$ then*

$$\forall (u, v) \in [0, 1]^2 \quad C(u, v) \leq Ga_\rho(u, v);$$

- *If $A(x,y) \geq \rho$ for all $(x,y) \in \mathbb{R}^2$ then*

$$\forall (u,v) \in [0,1]^2 \quad C(u,v) \geq Ga_\rho(u,v);$$

- *If $A(x,y)$ is nonnegative then the copula C is PQD*

$$\forall (u,v) \in [0,1]^2 \quad C(u,v) \geq \Pi(u,v) = uv;$$

- *If $A(x,y)$ is nonpositive then the copula C is NPQD*

$$\forall (u,v) \in [0,1]^2 \quad C(u,v) \leq \Pi(u,v) = uv;$$

The "existence" Theorem 3.2 and the "uniqueness" Corollary 3.6 imply the following theorem.

Theorem 3.8. *If a copula $C(x,y)$ is twice differentiable in a weak sense and fulfills the equation (3.11) with $A(x,y)$ bounded, $|A| < 1$, A and $\sqrt{1-A^2}$ Lipschitz and A differentiable with respect to the second variable, then there exists a $1/2$-self-similar solution $(X_t, Y_t)_{t\geq 1}$ of eq. (3.1), such that C is the copula of (X_t, Y_t) for all $t \geq 1$.*

The "uniqueness" Corollary 3.6 allows us to restate Theorem 3.2 in a more effective way.

Theorem 3.9. *Assume A1,A2, A3 and A4, then for any copula C, a solution of eq. (3.1), with initial values having a cumulative bivariate distribution $C^*(\Phi(x), \Phi(y))$, where*

$$C^*(u,v) = \lim_{n\to\infty} \frac{1}{n} \int_1^{e^n} \frac{1}{s} C_s^C(u,v)ds, \tag{3.13}$$

is $1/2$-self-similar.

The proofs of the subsequent theorems and corollaries stated above are provided in sections 4.5, and 4.6. With the exception of Corollary 3.7 which is based on Section 5.1

The weak derivatives proved also to be very useful in a study concerning dynamics of copulas of more general Ito processes, see [15].

In section 5, we will show that the set of solutions of equations (3.11) contains Gaussian copulas with $\rho^2 < 1$, FGM copulas with $\alpha^2 < 1$ and some but not all Frank copulas and does not contain Clayton copulas. In the first case A is constant, it equals to correlation coefficient. In the second and third A may vary.

4 Proofs and auxiliary results

4.1 Margins - proof of Proposition 3.1

Let $X_t = X_1 + W_t^1 - W_1^1$, where X_1 and $W_t^1 - W_1^1$ are independent. If $X_1 \sim N(0,1)$, then X_t, for $t > 1$, has normal distribution, with zero mean, as well. Furthermore

$$\sigma^2(X_t) = \sigma^2(X_1) + \sigma^2(W_t - W_1) = 1 + (t-1) = t.$$

Hence $(X_t)_{t\geq 1}$ is $1/2$-self-similar, ie,: $X_t \sim \sqrt{t}X_1$.

Now let $(X_t)_{t\geq 1}$ be $1/2$-self-similar. We need to show that $X_1 \sim N(0,1)$. We will base on the properties of characteristic functions (see [3] §26). Since

$$X_1 \sim \frac{1}{\sqrt{t}} X_t = \frac{1}{\sqrt{t}} X_1 + \frac{1}{\sqrt{t}} (W_t - W_1),$$

we get for fixed s

$$\mathbb{E}(\exp(isX_1)) = \mathbb{E}\left(\exp\left(is\frac{X_1}{\sqrt{t}}\right)\right)\exp\left(-\frac{t-1}{t}\frac{s^2}{2}\right).$$

Since $\frac{X_1}{\sqrt{t}}$ converges to 0 when $t \to +\infty$, we obtain by Lebesgue dominated convergence Theorem

$$\mathbb{E}(\exp(isX_1)) = \lim_{t\to\infty}\mathbb{E}\left(\exp\left(is\frac{X_1}{\sqrt{t}}\right)\right)\exp\left(-\frac{t-1}{t}\frac{s^2}{2}\right) = \exp\left(-\frac{s^2}{2}\right).$$

Hence X_1 has standard normal distribution.

When it comes to the second variable, the stochastic process

$$W_t' = \int_1^t A\left(\frac{X_s}{\sqrt{s}}, \frac{Y_s}{\sqrt{s}}\right) dW_s^1 + \int_1^t B\left(\frac{X_s}{\sqrt{s}}, \frac{Y_s}{\sqrt{s}}\right) dW_s^2,$$

where $B(x, y) = \sqrt{1 - A(x, y)^2}$, is a (time-rescaled) Brownian Motion, therefore re-writing $Y_t = Y_1 + W_t'$, we can see that by repeating the above reasoning we may conclude that Y_t is $1/2$-self-similar if and only $Y_1 \sim N(0, 1)$.

4.2 Semigroup property

Let C_t^C denote a copula of the pair (X_t, Y_t) for $t \geq 1$, where the process $(X_t, Y_t)_{t\geq1}$ is a solution of eq. (3.1), with initial values X_1 and Y_1 having the standard normal distribution $X_1 \sim Y_1 \sim N(0, 1)$ and linked by a copula C. Basing on Proposition 4.2 we show that the mapping

$$H : [1, \infty) \times \mathcal{C}^2 \longrightarrow \mathcal{C}^2, \quad H(t, C) = C_t^C, \tag{4.1}$$

is a representation of a multiplicative semigroup $\big([1, \infty), \cdot, 1\big)$ (see section 4.3 for more details). Observe that copula of a 1/2-self-similar process is a fixed point of H. It shows that such copulas can be obtained as generalized weighted averages along the orbit of H.

We start with the following auxiliary proposition which will be essential in the course of proving Theorem 3.2. We show that the rescaling (shifting) of a strong solution of SDE (3.1) give rise to a weak solution of the same equations.

Proposition 4.1. *If a stochastic process $(X_t, Y_t)_{t\geq1}$ is a solution of the set of equations (3.1) with initial values (X_1, Y_1) then for any $\tau > 1$ the 7-tuple*

$$\left(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_{t\tau}\}_{t\geq1}, \left(\frac{X_\tau}{\tau^{1/2}}, \frac{Y_\tau}{\tau^{1/2}}\right), \left(\frac{W_{t\tau}^1}{\tau^{1/2}}, \frac{W_{t\tau}^2}{\tau^{1/2}}\right)_{t\geq1}, \left(\frac{X_{t\tau}}{\tau^{1/2}}, \frac{Y_{t\tau}}{\tau^{1/2}}\right)_{t\geq1}\right)$$

is a (weak) solution of (3.1).

Proof.

First we observe that for $\tau > 1$

$$\begin{aligned} X_{t\tau} &= X_\tau + W_{t\tau}^1 - W_\tau^1, \\ Y_{t\tau} &= Y_\tau + \int_\tau^{t\tau} A\left(\frac{X_s}{\sqrt{s}}, \frac{Y_s}{\sqrt{s}}\right) dW_s^1 + \int_\tau^{t\tau} B\left(\frac{X_s}{\sqrt{s}}, \frac{Y_s}{\sqrt{s}}\right) dW_s^2, \end{aligned}$$

where $B(x, y) = \sqrt{1 - A(x, y)^2}$.

Next by the change of variables ($s := u\tau$) we get

$$\begin{aligned}\frac{X_{t\tau}}{\sqrt{\tau}} &= \frac{X_\tau}{\sqrt{\tau}} + \frac{1}{\sqrt{\tau}} W^1_{t\tau} - \frac{1}{\sqrt{\tau}} W^1_\tau, \\ \frac{Y_{t\tau}}{\sqrt{\tau}} &= \frac{Y_\tau}{\sqrt{\tau}} + \int_1^t A\left(\frac{X_{u\tau}}{\sqrt{u\tau}}, \frac{Y_{u\tau}}{\sqrt{u\tau}}\right) \frac{1}{\sqrt{\tau}} dW^1_{u\tau} + \int_1^t B\left(\frac{X_{u\tau}}{\sqrt{u\tau}}, \frac{Y_{u\tau}}{\sqrt{u\tau}}\right) \frac{1}{\sqrt{\tau}} dW^2_{u\tau}.\end{aligned}$$

Since for fixed τ

$$\tilde{W}^1_t = \frac{1}{\sqrt{\tau}} W^1_{t\tau} \quad \text{and} \quad \tilde{W}^2_t = \frac{1}{\sqrt{\tau}} W^2_{t\tau}$$

are independent Wiener processes (defined on the same probability space as W^1 and W^2), we conclude that indeed we obtained a solution of the stochastic differential equations (3.1) with another pair of Wiener processes

$$\begin{aligned} dX_t &= d\tilde{W}^1_t, \\ dY_t &= A\left(\frac{X_t}{\sqrt{t}}, \frac{Y_t}{\sqrt{t}}\right) d\tilde{W}^1_t + B\left(\frac{X_t}{\sqrt{t}}, \frac{Y_t}{\sqrt{t}}\right) d\tilde{W}^2_t. \end{aligned}$$

with initial values $(\tau^{-1/2} X_\tau, \tau^{-1/2} Y_\tau)$. □

Due to the uniqueness of the weak solutions of the SDE (see [5]) Proposition 4.1 implies Proposition 4.2.

Proposition 4.2. *We assume **A1**, **A2** and **A3**. Then the solution $(X'_t, Y'_t)_{t\geq 1}$ of SDE (3.1) with initial values law equivalent to $(X_\tau/\sqrt{\tau}, Y_\tau/\sqrt{\tau})$, $\tau > 1$,*

$$(X'_1, Y'_1) \sim \left(\frac{X_\tau}{\sqrt{\tau}}, \frac{Y_\tau}{\sqrt{\tau}}\right),$$

is law equivalent to the τ-rescaled process

$$(X'_t, Y'_t)_{t\geq 1} \sim \left(\frac{X_{\tau t}}{\sqrt{\tau}}, \frac{Y_{\tau t}}{\sqrt{\tau}}\right)_{t\geq 1}.$$

4.3 Existence

We assume **A1**, **A2** and **A3** throughout this section.

First we establish some estimates which imply the continuity of solutions of equations (3.1) in L^2 norm.

In more details, since both $(X_t)_{t\geq 1}$ and $(Y_t)_{t\geq 1}$ coincide with Wiener processes we have for $s, t \geq 1$

$$\begin{aligned} ||X_t - X_s||^2_{L^2} &= \mathbb{E}((X_t - X_s)^2) = |t-s|, \\ ||Y_t - Y_s||^2_{L^2} &= \mathbb{E}((Y_t - Y_s)^2) = |t-s|. \end{aligned} \tag{4.2}$$

Let us consider two processes $(X'_t, Y'_t)_{t\geq 1}$ and $(X''_t, Y''_t)_{t\geq 1}$ which solve equation (3.1) with different initial values:

$$X'_t = X'_1 + W^1_t - W^1_1, \tag{4.3}$$

$$X''_t = X''_1 + W^1_t - W^1_1, \tag{4.4}$$

$$Y'_t = Y'_1 + \int_1^t A\left(\frac{X'_s}{\sqrt{s}}, \frac{Y'_s}{\sqrt{s}}\right) dW^1_s + \int_1^t B\left(\frac{X'_s}{\sqrt{s}}, \frac{Y'_s}{\sqrt{s}}\right) dW^2_s, \tag{4.5}$$

$$Y''_t = Y''_1 + \int_1^t A\left(\frac{X''_s}{\sqrt{s}}, \frac{Y''_s}{\sqrt{s}}\right) dW^1_s + \int_1^t B\left(\frac{X''_s}{\sqrt{s}}, \frac{Y''_s}{\sqrt{s}}\right) dW^2_s, \tag{4.6}$$

where $B(x,y) = \sqrt{1 - A(x,y)^2}$.

By the following lemma the distances between $(X'_t)_{t\geq 1}$ and $(X''_t)_{t\geq 1}$ and also between $(Y'_t)_{t\geq 1}$ and $(Y''_t)_{t\geq 1}$ are linear functions of the distances between the initial values.

Lemma 4.3. *For any $T \geq 1$:*

$$||X'_T - X''_T||_{L^2} = ||X'_1 - X''_1||_{L^2}, \tag{4.7}$$

$$\begin{aligned} \left|\left|Y'_T - Y''_T\right|\right|^2_{L^2} &\leq \left(1 + T^L \ln(T)\right) \left|\left|Y'_1 - Y''_1\right|\right|^2_{L^2} \\ &+ L\left(\ln(T) + \frac{T^L}{2}\ln^2(T)\right) \left|\left|X'_1 - X''_1\right|\right|^2_{L^2}, \end{aligned} \tag{4.8}$$

where L is a sum of squares of Lipschitz coefficients for A and B,
$\forall (x', y'), (x'', y'') \in \mathbb{R}^2$

$$|A(x',y') - A(x'',y'')|^2 + |B(x',y') - B(x'',y'')|^2 \leq L\left(|x' - x''|^2 + |y' - y''|^2\right).$$

Proof.

Equality (4.7) follows readily since

$$X'_t - X''_t = X'_1 - X''_1. \tag{4.9}$$

For the proof of (4.8), we apply subsequently the Ito formula, Lipschitz inequality and (4.9). Finally we get for $t \in [0, T-1]$

$$\begin{aligned} \mathbb{E}\left|Y'_{1+t} - Y''_{1+t}\right|^2 &= \mathbb{E}\left|Y'_1 - Y''_1\right|^2 \\ &+ \mathbb{E}\int_1^{1+t} \left(\left(A\left(\frac{X'_s}{\sqrt{s}}, \frac{Y'_s}{\sqrt{s}}\right) - A\left(\frac{X''_s}{\sqrt{s}}, \frac{Y''_s}{\sqrt{s}}\right)\right)^2\right. \\ &\qquad \left. + \left(B\left(\frac{X'_s}{\sqrt{s}}, \frac{Y'_s}{\sqrt{s}}\right) - B\left(\frac{X''_s}{\sqrt{s}}, \frac{Y''_s}{\sqrt{s}}\right)\right)^2\right) ds \\ &\leq \mathbb{E}\left|Y'_1 - Y''_1\right|^2 + \int_1^{1+t} \frac{L}{s}\left(\mathbb{E}\left|Y'_s - Y''_s\right|^2 + \mathbb{E}\left|X'_s - X''_s\right|^2\right) ds \\ &\leq \mathbb{E}\left|Y'_1 - Y''_1\right|^2 + L\ln(1+t)\mathbb{E}\left|X'_1 - X''_1\right|^2 \\ &\quad + \int_1^{1+t} \frac{L}{s}\mathbb{E}\left|Y'_s - Y''_s\right|^2 ds. \end{aligned}$$

Hence

$$\begin{aligned} \frac{1}{1+t}\mathbb{E}\left|Y'_{1+t} - Y''_{1+t}\right|^2 &\leq \frac{1}{1+t}\mathbb{E}\left|Y'_1 - Y''_1\right|^2 + \frac{L\ln(1+t)}{1+t}\mathbb{E}\left|X'_1 - X''_1\right|^2 \\ &\quad + \frac{L}{1+t}\int_1^{1+t} \frac{1}{s}\mathbb{E}\left|Y'_s - Y''_s\right|^2 ds. \end{aligned} \tag{4.10}$$

The Gronwall Lemma (see [11] Appendix B.j) applied for a function

$$\eta(t) = \int_1^{1+t} \frac{1}{s}\mathbb{E}\left|Y'_s - Y''_s\right|^2 ds \tag{4.11}$$

yields for $t \in [0, T-1]$

$$\eta(t) \leq (1+t)^L \left(\ln(1+t)\mathbb{E}\left|Y'_1 - Y''_1\right|^2 + \frac{L}{2}\ln^2(1+t)\mathbb{E}\left|X'_1 - X''_1\right|^2\right). \tag{4.12}$$

Subsequently

$$\begin{aligned} ||Y'_{1+t} - Y''_{1+t}||^2_{L^2} &\leq \left(1 + (1+t)^L \ln(1+t)\right) ||Y'_1 - Y''_1||^2_{L^2} \\ &+ L\left(\ln(1+t) + \frac{(1+t)^L}{2}\ln^2(1+t)\right) ||X'_1 - X''_1||^2_{L^2}. \end{aligned} \tag{4.13}$$

Having substituted $t = T - 1$ we conclude the proof of Lemma. □

To prove the existence of self-similar solutions, we follow the approach of Khasminskii (§2.2 [21]) and construct a self-similar solution basing on the averages of a process.

In more details, let C be any copula. We shall consider the solutions $(X_t, Y_t)_{t\geq 1}$ of the set of equations (3.1) with initial values (X_1, Y_1), for which we assume that both X_1 and Y_1 are standard normal $N(0, 1)$ and the copula of their joint distribution is C. By $H(t, C)$ we denote the copula of (X_t, Y_t). Since X_t and Y_t have normal distribution with mean 0 and variance equal t ($N(0, t)$), the joint cumulative distribution of the pair (X_t, Y_t) is given by

$$F_t(x, y) = H(t, C)\left(\Phi(\frac{x}{\sqrt{t}}), \Phi(\frac{y}{\sqrt{t}})\right). \tag{4.14}$$

Since for any copula C there exists a pair of random variables with joint distribution function $C(\Phi(x), \Phi(y))$, the function H is well defined on $[1, \infty) \times \mathcal{C}^2$ (compare [21] Th. 3.4).

For the solution $(X_t, Y_t)_{t\geq 1}$ of (3.1) to be a 1/2-self-similar process we need to show that $H(t, C) = C$ for all $t \geq 1$, therefore we are in fact interested in the existence of fixed points of H.

In the following proposition we list the basic properties of H.

Proposition 4.4. *1. H defines a representation of the multiplicative semigroup* $([1, \infty), \cdot, 1)$. *For any* $t, s \geq 1$ *and copula C*

$$H(t, H(s, C)) = H(ts, C).$$

2. H is continuous with respect to the metric

$$d((t, C), (s, D)) = |t - s| + \sup\{|C(u, v) - D(u, v)| : (u, v) \in [0, 1]^2\},$$

where C and D are copulas and s and t are real numbers.
3. H commutes with the mixture of copulas, ie. if $C_\theta(u, v)$, $\theta \in \Theta \subset \mathbb{R}^k$ *is a measurable family of copulas then for any probabilistic measure* μ *on* Θ *and any* $t \geq 1$

$$H\left(t, \int_\Theta C_\theta d\mu(\theta)\right) = \int_\Theta H(t, C_\theta) d\mu(\theta).$$

Proof.

Point 1. Follows directly from Proposition 4.1.

Point 2. Follows from the continuity of solutions of equations (3.1) in L^2 norm. For the proof we shall apply Skorohod representation Theorem (see [3] Th. 25.6).

Let $\{C_n\}_{n\in\mathbb{N}_+}$ be any sequence of copulas convergent to copula C_∞. Then, by the Skorohod Theorem, there exists a probability space $(\Omega_1, \mathcal{F}_1, \mathbb{P}_1)$, a sequence of random pairs $\{(Z_n^1, Z_n^2)\}_{n\in\mathbb{N}_+}$ and a random vector (Z_∞^1, Z_∞^2) (all defined on $(\Omega_1, \mathcal{F}_1, \mathbb{P}_1)$), such that:
1. $C_n(\Phi(x), \Phi(y))$ is a distribution function of (Z_n^1, Z_n^2), $n = 1, 2, \ldots, \infty$;
2. (Z_n^1, Z_n^2) almost surely converges to (Z_∞^1, Z_∞^2).

Let us consider the following product of probability spaces

$$(\Omega', \mathcal{F}', \mathbb{P}') = (\Omega \times \Omega_1, \mathcal{F} \times \mathcal{F}_1, \mathbb{P} \times \mathbb{P}_1).$$

Since Wiener processes W^1, W^2 have been defined on $(\Omega, \mathcal{F}, \mathbb{P})$ and the initial values Z_n^i on $(\Omega_1, \mathcal{F}_1, \mathbb{P}_1)$ we extend them onto the product space by putting

$$W_t^i(\omega, \omega_1) = W_t^i(\omega), \quad Z_n^i(\omega, \omega_1) = Z_n^i(\omega_1).$$

Next we analyze the stochastic equations (3.1) on the previously defined product space, denoting by $(X_{n,t}, Y_{n,t})$ their solutions with initial values (Z_n^1, Z_n^2), for $n = 1, 2, \dots, \infty$.

Obviously $H(t, C_n)$ are copulas of $(X_{n,t}, Y_{n,t})$.

Note that due to Lebesgue dominated convergence Theorem the sequence $\{(Z_n^1, Z_n^2)\}_{n\in\mathbb{N}_+}$ converges in L^2 norm

$$\mathbb{E}\Big((Z_n^1 - Z_\infty^1)^2 + (Z_n^2 - Z_\infty^2)^2\Big) \overset{n\to\infty}{\longrightarrow} 0.$$

Therefore, by Lemma 4.3 and equalities (4.2), for any convergent sequence of indices $\{s_n\}_{n\in\mathbb{N}_+} \subset [1, \infty)$ the sequence of random pairs $\{(X_{n,s_n}, Y_{n,s_n})\}_{n\in\mathbb{N}_+}$ converges in L^2 to $(X_{\infty,s_\infty}, Y_{\infty,s_\infty})$, where s_∞ denotes the limit of s_n.

By the following lemma, convergence in L^2 and convergence in distribution are closely related. To simplify the notation we put for two random pairs $V_i = (X_i, Y_i)$, $i = 1, 2$

$$||V_2 - V_1||_{L^2}^2 = \mathbb{E}\Big((X_2 - X_1)^2\Big) + \mathbb{E}\Big((Y_2 - Y_1)^2\Big). \tag{4.15}$$

Lemma 4.5. *Let V_1 and V_2 be two random pairs with standard normal margins and copulas respectively C^1 and C^2, defined on the same probability space. Then*

$$||C^2 - C^1||_\infty \le \frac{3}{\sqrt[3]{2\pi}} ||V_1 - V_2||_{L^2}^{2/3}.$$

Proof.
Let (x, y) be any point of the real plane and ε a positive constant. Applying the elementary set theory and Markov inequality (see [3] formula (21.12)) we get the following estimate

$$\begin{aligned}
&C^2(\Phi(x), \Phi(y)) - C^1(\Phi(x+\varepsilon), \Phi(y+\varepsilon)) \\
&\quad = \mathbb{P}(X_1 \le x, Y_1 \le y) - \mathbb{P}(X_2 \le x+\varepsilon, Y_2 \le y+\varepsilon) \\
&\quad \le \mathbb{P}\Big(\{\omega : X_1(\omega) \le x \wedge Y_1(\omega) \le y\} \setminus \{\omega : X_2(\omega) \le x+\varepsilon \wedge Y_2(\omega) \le y+\varepsilon\}\Big) \\
&\quad = \mathbb{P}\Big(\{\omega : X_1(\omega) \le x \wedge Y_1(\omega) \le y \wedge \Big(X_2(\omega) > x+\varepsilon \vee Y_2(\omega) > y+\varepsilon\Big)\}\Big) \\
&\quad = \mathbb{P}\Big(\{\omega : \Big(X_1(\omega) \le x \wedge Y_1(\omega) \le y \wedge X_2(\omega) > x+\varepsilon\Big) \\
&\qquad \vee \Big(X_1(\omega) \le x \wedge Y_1(\omega) \le y \wedge Y_2(\omega) > y+\varepsilon\Big)\}\Big) \\
&\quad \le \mathbb{P}\Big(\{\omega : X_1(\omega) \le x \wedge X_2(\omega) > x+\varepsilon\}\Big) + \mathbb{P}\Big(\{\omega : Y_1(\omega) \le y \wedge Y_2(\omega) > y+\varepsilon\}\Big) \\
&\quad \le \mathbb{P}\Big(|X_1 - X_2| \ge \varepsilon\Big) + \mathbb{P}\Big(|Y_1 - Y_2| \ge \varepsilon\Big) \\
&\quad \le \frac{1}{\varepsilon^2}\Big(\mathbb{E}\Big(|X_1 - X_2|^2\Big) + \mathbb{E}\Big(|Y_1 - Y_2|^2\Big)\Big) = \frac{1}{\varepsilon^2}||V_1 - V_2||_{L^2}^2.
\end{aligned} \tag{4.16}$$

Since copulas are Lipschitz functions (with Lipschitz constant 1) and Φ is also Lipschitz function (with Lipschitz constant $\varphi(0) = \frac{1}{\sqrt{2\pi}}$) we further estimate

$$\begin{aligned}
&C^2(\Phi(x), \Phi(y)) - C^1(\Phi(x), \Phi(y)) \\
&\quad \le C^1(\Phi(x+\varepsilon), \Phi(y+\varepsilon)) - C^1(\Phi(x), \Phi(y)) + \frac{1}{\varepsilon^2}||V_1 - V_2||_{L^2}^2 \\
&\quad = \frac{\sqrt{2}}{\sqrt{\pi}}\varepsilon + \frac{1}{\varepsilon^2}||V_1 - V_2||_{L^2}^2.
\end{aligned} \tag{4.17}$$

Since the bound is valid for all positive ε, we substitute

$$\varepsilon = \sqrt[6]{2\pi}||V_1 - V_2||_{L^2}^{2/3},$$

which minimizes the estimate.

$$C^2(\Phi(x),\Phi(y)) - C^1(\Phi(x),\Phi(y)) \le 3\frac{||V_1 - V_2||_{L^2}^{2/3}}{\sqrt[3]{2\pi}}.$$

The above bound remains valid when we replace C^2 and C^1. Moreover it is valid for all points $(x, y) \in \mathbb{R}^2$. Therefore

$$\begin{aligned} ||C^2 - C^1||_\infty &= \sup\{|C^2(u,v) - C^1(u,v)| : (u,v) \in [0,1]^2\} \\ &= \sup\{|C^2(\Phi(x),\Phi(y)) - C^1(\Phi(x),\Phi(y))| : (x,y) \in \mathbb{R}^2\} \\ &\le \frac{3}{\sqrt[3]{2\pi}}||V_1 - V_2||_{L^2}^{2/3}. \end{aligned} \tag{4.18}$$

□

Thus convergence in L^2 implies convergence in distribution and the joint cumulative distribution functions of (X_{n,s_n}, Y_{n,s_n}) are converging to the cumulative distribution functions of $(X_{\infty,s_\infty}, Y_{\infty,s_\infty})$

$$\lim_{n\to\infty} H(s_n, C_n)(\Phi(x),\Phi(y)) = H(s_\infty, C_\infty)(\Phi(x),\Phi(y)). \tag{4.19}$$

Hence the copulas $H(s_n, C_n)$ converge to $H(s_\infty, C_\infty)$

$$\lim_{n\to\infty} H(s_n, C_n) = H\left(\lim_{n\to\infty} s_n, \lim_{n\to\infty} C_n\right). \tag{4.20}$$

□

Point 3. follows from the fact that a solution of equations (3.1) with random initial values is a mixture of solutions with deterministic initial values (compare [21] Theorem 3.4 point 2). Thus we have to show the associativity of iterated mixtures.

We recall that a random pair Z_μ is a mixture of random pairs Z_θ, $\theta \in \Theta$, with respect to the probabilistic measure μ on Θ, when for every bounded Borel function f on $\mathbb{R}^2$

$$E(f(Z_\mu)) = \int_\Theta E(f(Z_\theta))d\mu. \tag{4.21}$$

We assume that the random pairs Z_μ and Z_θ, $\theta \in \Theta$, and the two-dimensional Wiener process $(W_t^1 - W_1^1, W_t^2 - W_1^2)_{t\ge1}$ are independent. We denote by (X_t^z, Y_t^z) (respectively by (X_t^θ, Y_t^θ) and by (X_t^μ, Y_t^μ)) a solution of (3.1) with deterministic initial values $(X_1, Y_1) = z$, $z \in \mathbb{R}^2$ (respectively with random initial values $(X_1, Y_1) = Z_\theta$ and $(X_1, Y_1) = Z_\mu$). Thanks to the assumptions **A1, A2** and **A3** such objects exist and are unique (Theorem 3.4 [21]). Let f be a bounded Borel function on $\mathbb{R}^2$. We put

$$u(t,z) = E(f(X_t^z, Y_t^z)). \tag{4.22}$$

We get

$$E(f(X_t^\theta, Y_t^\theta)) = E(E(f(X_t^\theta, Y_t^\theta)|X_1^\theta, Y_1^\theta) = E(u(t, Z^\theta)) \tag{4.23}$$

and similarly

$$E(f(X_t^\mu, Y_t^\mu)) = E(u(t, Z^\mu)). \tag{4.24}$$

Applying (4.21) to the right side, we obtain

$$E(f(X_t^\mu, Y_t^\mu)) = \int_\Theta E(u(t, Z^\theta))d\mu = \int_\Theta E(f(X_t^\theta, Y_t^\theta))d\mu. \tag{4.25}$$

Thus for each $t \ge 1$ (X_t^μ, Y_t^μ) is a mixture of (X_t^θ, Y_t^θ).

To conclude the proof of point 3 it is enough to select

$$Z^\theta \sim C_\theta(\Phi(x), \Phi(y)), \text{ and } Z^\mu \sim C_\mu(\Phi(x), \Phi(y)),$$

where

$$C_\mu(x, y) = \int\limits_\Theta C_\theta(x, y) d\mu(\theta).$$

□

Next we show that the set of fixed points of the semigroup H is not empty. We select a copula C and as a fixed point take the generalized average of the trajectory of C. In more details.

Proposition 4.6. *For any copula C the set of cluster points of the sequence $(C^k)_{k\geq 1}$, given by*

$$C^k(u, v) = \int\limits_1^{e^k} \frac{1}{ks} H(s, C)(u, v) ds$$

is not empty. Furthermore any such cluster point is a fixed point of H.

Proof.
Due to Ascoli theorem there exists a subsequence of C^k having a limit, i.e. a cluster point of the sequence. We denote this limit copula by $C^\star$

$$C^{k_n} \longrightarrow C^\star.$$

Since H is continuous and commutes with the mixtures we get

$$\begin{aligned} & H(t, C^\star) \\ = \; & H(t, \lim_{n\to\infty} C^{k_n}) = \lim_{n\to\infty} H(t, C^{k_n}) = \lim_{n\to\infty} H(t, \int\limits_1^{e^{k_n}} \frac{1}{k_n s} H(s, C) ds) \\ = \; & \lim_{n\to\infty} \int\limits_1^{e^{k_n}} \frac{1}{k_n s} H(t, H(s, C)) ds = \lim_{n\to\infty} \int\limits_1^{e^{k_n}} \frac{1}{k_n s} H(ts, C) ds \\ = \; & \lim_{n\to\infty} \int\limits_t^{te^{k_n}} \frac{1}{k_n s} H(s, C) ds. \end{aligned} \tag{4.26}$$

Therefore, since copulas are bounded by 1, we get for $t \geq 1$

$$\begin{aligned} & |H(t, C^\star)(u, v) - C^\star(u, v)| \\ = \; & \lim_{n\to\infty} |H(t, C^{k_n})(u, v) - C^{k_n}(u, v)| \\ = \; & \lim_{n\to\infty} \left| \int\limits_t^{te^{k_n}} \frac{1}{k_n s} H(s, C)(u, v) ds - \int\limits_1^{e^{k_n}} \frac{1}{k_n s} H(s, C)(u, v) ds \right| \\ = \; & \lim_{n\to\infty} \left| \int\limits_{e^{k_n}}^{te^{k_n}} \frac{1}{k_n s} H(s, C)(u, v) ds - \int\limits_1^{t} \frac{1}{k_n s} H(s, C)(u, v) ds \right| \\ \leq \; & \lim_{n\to\infty} \left(\int\limits_{e^{k_n}}^{te^{k_n}} \frac{1}{k_n s} ds + \int\limits_1^{t} \frac{1}{k_n s} ds \right) \\ = \; & \lim_{n\to\infty} \frac{1}{k_n} \left(\ln(t) + \ln(e^{k_n}) - \ln(e^{k_n}) + \ln(t) - \ln(1) \right) = \lim_{n\to\infty} \frac{2\ln(t)}{k_n} = 0. \end{aligned} \tag{4.27}$$

□

The following corollary concludes the proof of Theorem 3.2.

Corollary 4.7. *Let $C^\star$ be a fixed point of H. Then the solution of equations (3.1) with initial values*

$$X_1 = W_1^1, \quad Y_1 = \Phi^{-1}\Psi(\Phi(W_1^1), \Phi(W_1^2)),$$

where Ψ is a generalized inverse of the weak derivative of $C^\star$

$$\Psi(u, v) = \inf\{w : D^{1,0}C^\star(u, w) \geq v\}.$$

is (1/2)-self-similar.

Proof.
Let $F_t(x, y)$ be a cumulative distribution function of the solution (X_t, Y_t). Since

$$F_1(x, y) = C^\star(\Phi(x), \Phi(y)),$$

we get from Proposition 4.4

$$\begin{aligned} F_t(x, y) &= H(t, C^\star)\left(\Phi\left(\frac{x}{\sqrt{t}}\right), \Phi\left(\frac{y}{\sqrt{t}}\right)\right) = C^\star\left(\Phi\left(\frac{x}{\sqrt{t}}\right), \Phi\left(\frac{y}{\sqrt{t}}\right)\right) \\ &= F_1\left(\frac{x}{\sqrt{t}}, \frac{y}{\sqrt{t}}\right). \end{aligned}$$

□

4.4 Proof of Proposition 3.3

In this section we assume that the sum of squares of Lipschitz coefficients of A and B, denoted by L, is smaller than 1. Let $C^\star$ be a fixed point of the semigroup H from Proposition 4.6 and $(X_t, Y_t)_{t\geq 1}$ be a 1/2-self-similar solution of equations (3.1) from Corollary 4.7. Its distribution functions are equal to

$$F_t(x, y) = C^\star\left(\Phi\left(\frac{x}{\sqrt{t}}\right), \Phi\left(\frac{y}{\sqrt{t}}\right)\right).$$

We compare it with an arbitrary solution of equations (3.1) $(X'_t, Y'_t)_{t\geq 1}$. Let

$$F'_t(x, y) = C_t\left(\Phi\left(\frac{x}{\sqrt{t}}\right), \Phi\left(\frac{y}{\sqrt{t}}\right)\right).$$

be its distribution functions. Due to Lemmas 4.3 and 4.5 we have for every $t \geq 1$

$$\begin{aligned} ||C_\star - C_t||_\infty^3 &\leq \frac{3^3}{2\pi}||\frac{1}{\sqrt{t}}(X_t - X'_t, Y_t - Y'_t)||_{L^2}^2 \\ &= \frac{27}{2\pi}\frac{1}{t}\left(||X_t - X'_t||_{L^2}^2 + ||Y_t - Y'_t||_{L^2}^2\right) \\ &\leq \frac{27}{2\pi}\frac{1}{t}\left(\left(1 + t^L\ln(t)\right)||Y'_1 - Y''_1||_{L^2}^2\right. \\ &\quad + \left.\left(1 + L\ln(t) + L\frac{t^L}{2}\ln^2(t)\right)||X'_1 - X''_1||_{L^2}^2\right), \end{aligned} \tag{4.28}$$

Since X_1, X'_1, Y_1 and Y'_1 have mean 0 and variance 1 we may bound

$$||X_t - X'_t||_{L^2}^2 \leq 4, \quad ||Y_t - Y'_t||_{L^2}^2 \leq 4.$$

Hence

$$||C_\star - C_t||_\infty^3 \leq \frac{54}{\pi}\left(\frac{2+L\ln(t)}{t} + t^{L-1}\left(\ln(t) + \frac{L}{2}\ln^2(t)\right)\right). \tag{4.29}$$

Since L is smaller than 1,

$$\lim_{t\to\infty} ||C_t - C_\star||_\infty = 0. \tag{4.30}$$

Which concludes the proof of Proposition 3.3.

4.5 Generalized solutions of PDEs

In this section we begin the proof of Theorem 3.4.

We show that the copula process C_t is a "weak generalized" solution of PDE (3.12) (we follow the naming used in [13] and [11], see also [15]). We assume **A1** – **A4**. By H^1 we denote the Hilbert space of weakly differentiable functions, which together with their derivatives are square integrable.

Proposition 4.8. *For any $h \in H^1(\mathbb{R}^2)$ and $t \geq 0$*

$$\frac{d}{dt}\int_{[0,1]^2} h(u,v)C_t(u,v)dudv = -B_t(h, C_t), \tag{4.31}$$

where for fixed $t \geq 0$ and fixed copula C, $B_t(\cdot, C)$ is a continuous linear functional on $H^1(\mathbb{R}^2)$, given by the formula

$$B_t(h,C) = \frac{1}{2t}\int_{[0,1]^2} D^{1,0}\left(h(u,v)\varphi(\Phi^{-1}(u))^2\right)D^{1,0}C(u,v)dudv \tag{4.32}$$

$$+ \frac{1}{2t}\int_{[0,1]^2} D^{0,1}\left(h(u,v)\varphi(\Phi^{-1}(v))^2\right)D^{0,1}C(u,v)dudv$$

$$+ \frac{1}{t}\int_{[0,1]^2} D^{0,1}\left(h(u,v)\varphi(\Phi^{-1}(u))\varphi(\Phi^{-1}(v))A(\Phi^{-1}(u),\Phi^{-1}(v))\right)D^{1,0}C(u,v)dudv.$$

Proof.
We base on results from [15]. We shift time t by 1, put $\sigma_i = 1$ and $\mu_i = 0$ and apply Theorem 4.1 from [15]. □

Since for selfsimilar processes the copula process C_t is constant, say $C_t = C$ for $t \geq 1$, we get

$$B_1(h, C) = 0. \tag{4.33}$$

To continue the proof of theorem 3.4 we have to improve the regularity of C. We apply Theorem 8.8 of [13]. In the notation used in [13], the divergence form of B_1 looks like

$$\begin{aligned} B_1(h,C) &= \int_{[0,1]^2} a^{1,1}D^{1,0}CD^{1,0}h + a^{2,2}D^{0,1}CD^{0,1}h + a^{2,1}D^{1,0}CD^{0,1}h \\ &+ \left(c^1D^{1,0}C + c^2D^{0,1}C\right)hdudv, \end{aligned} \tag{4.34}$$

where

$$a^{1,1}(u,v) = \frac{1}{2}\varphi(\Phi^{-1}(u))^2, \tag{4.35}$$

$$a^{2,2}(u,v) = \frac{1}{2}\varphi(\Phi^{-1}(v))^2, \tag{4.36}$$

$$a^{2,1}(u,v) = \varphi(\Phi^{-1}(u))\varphi(\Phi^{-1}(v))A(\Phi^{-1}(u),\Phi^{-1}(v)), \tag{4.37}$$

$$\begin{aligned} c^1(u,v) &= -\varphi(\Phi^{-1}(u))\phi^{-1}(u) - \varphi(\Phi^{-1}(u))\Phi^{-1}(v)A(\Phi^{-1}(u),\Phi^{-1}(v)) \\ &+ \varphi(\Phi^{-1}(u))D^{0,1}A(\Phi^{-1}(u),\Phi^{-1}(v)), \end{aligned} \tag{4.38}$$

$$c^2(u,v) = -\varphi(\Phi^{-1}(u))\phi^{-1}(v). \tag{4.39}$$

We restrict the domain to the smaller square $\Omega_r = (\Phi(-r), \Phi(r))^2$, $r > 0$. We observe that, for any $r > 0$, on Ω_r the coefficients $a^{1,1}$, $a^{2,2}$, $a^{2,1}$, c^1 and c^2 are bounded. Furthermore $a^{1,1}$, $a^{2,2}$, $a^{2,1}$ are Lipschitz and B is strongly elliptic

$$a^{1,1}(u,v)a^{2,2}(u,v) - \frac{1}{4}a^{2,1}(u,v)^2 = \frac{1}{4}\varphi(\Phi^{-1}(u))^2\varphi(\Phi^{-1}(v))^2(1 - A(\Phi^{-1}(u), \Phi^{-1}(v))^2) > 0. \tag{4.40}$$

Therefore Theorem 8.8 of [13] implies that $C(u, v)$ belongs to $H^2(\Omega_r)$ for any $r > 0$, thus it is weakly twice differentiable on $(0, 1)^2$ and fulfills almost everywhere equation (3.11).

$$\begin{aligned} &D^{2,0}C(u,v)\varphi(\Phi^{-1}(u))^2 + D^{0,2}C(u,v)\varphi(\Phi^{-1}(v))^2 \\ &+2A(\Phi^{-1}(u), \Phi^{-1}(v))D^{1,1}C(u,v)\varphi(\Phi^{-1}(u))\varphi(\Phi^{-1}(v)) = 0. \end{aligned} \tag{4.41}$$

When C is weakly twice differentiable so is the cumulative distribution function $F(x, y) = C(\Phi(x), \Phi(y))$. Furthermore

$$\begin{aligned} D^{1,0}F(x,y) &= D^{1,0}C(\Phi(x), \Phi(y))\varphi(x), \\ D^{0,1}F(x,y) &= D^{0,1}C(\Phi(x), \Phi(y))\varphi(y), \\ D^{1,1}F(x,y) &= D^{1,1}C(\Phi(x), \Phi(y))\varphi(x)\varphi(y), \\ D^{2,0}F(x,y) &= D^{2,0}C(\Phi(x), \Phi(y))\varphi(x)^2 - D^{1,0}C(\Phi(x), \Phi(y))x\varphi(x), \\ D^{0,2}F(x,y) &= D^{0,2}C(\Phi(x), \Phi(y))\varphi(y)^2 - D^{0.1}C(\Phi(x), \Phi(y))y\varphi(y). \end{aligned} \tag{4.42}$$

If C fulfills almost everywhere equation (4.41), then F fulfills almost everywhere on $\mathbb{R}^2$ the equation (3.12)

$$D^{2,0}F + D^{0,2}F + 2A(x,y)D^{1,1}F + xD^{1,0}F + yD^{0,1}F = 0 \tag{4.43}$$

In such a way we conclude the proof of Theorem 3.4.

Remark 4.1. *1. Equation (3.11) can be derived from the Fokker-Planck equation (see [7]) But this requires the existence and twice-differentiability of the density of the process* (X_t, Y_t).
2. Equations (3.11) and (3.12) can be derived without the assumption ***A4****, see [22].*

4.6 Dominance and uniqueness

Theorem 4.9. *Let* A_i, $i = 1, 2$, *be bounded Lipschitz functions,* $|A_i| < 1$. *If copulas* C_1 *and* C_2 *fulfill PDE (3.11) with* A *equal respectively to* A_1 *and* A_2 *and*

$$\forall (x,y) \in \mathbb{R}^2 \quad A_1(x,y) \geq A_2(x,y),$$

then copula C_1 *dominates in concordance ordering*

$$\forall (u,v) \in [0,1]^2 \quad C_1(u,v) \geq C_2(u,v),$$

The proof of Theorem 3.5 follows from the maximum principle for elliptic partial differential equations, see [13] or [11]. In more details: If copulas C_1 and C_2 fulfill PDE (3.11) with A equals respectively to A_1 and A_2 and A_1 dominates

$$\forall (x,y) \in \mathbb{R}^2 \quad A_1(x,y) \geq A_2(x,y),$$

then a function $U(u, v) = C_1(u, v) - C_2(u, v)$ is a sub-solution of the equation (3.11) with A equals to A_2 with zero boundary condition. Indeed, since C_1 is a copula its mixed derivative is almost everywhere nonnegative and for almost all $(u, v) \in (o, 1)^2$ we get

$$\begin{aligned} &D^{2,0}U(u,v)\varphi(\Phi^{-1}(u))^2 + D^{0,2}U(u,v)\varphi(\Phi^{-1}(v))^2 \\ &+2A_2(\Phi^{-1}(u), \Phi^{-1}(v))D^{1,1}U(u,v)\varphi(\Phi^{-1}(u))\varphi(\Phi^{-1}(v)) \\ = \; &D^{2,0}C_1(u,v)\varphi(\Phi^{-1}(u))^2 + D^{0,2}C_1(u,v)\varphi(\Phi^{-1}(v))^2 \\ &+2A_2(\Phi^{-1}(u), \Phi^{-1}(v))D^{1,1}C_1(u,v)\varphi(\Phi^{-1}(u))\varphi(\Phi^{-1}(v)) \\ = \; &2\big(A_2(\Phi^{-1}(u), \Phi^{-1}(v)) - A_1(\Phi^{-1}(u), \Phi^{-1}(v))\big)D^{1,1}C_1(u,v)\varphi(\Phi^{-1}(u))\varphi(\Phi^{-1}(v)) \leq 0. \end{aligned} \tag{4.44}$$

Let δ be a minimum of U and let it be attained at the point (u_0, v_0)

$$\delta = \min\{U(u, v) : (u, v) \in [0, 1]^2\} = U(u_0, v_0).$$

We show that δ must be 0. If it happened that $\delta < 0$ then knowing that U is a Lipschitz function vanishing on the border of the unit square we would be able to select an enough big r that:
1. the point (u_0, v_0) would belong to $\Omega_r = (\Phi(-r), \Phi(r))^2$;
2. the infimum of U on the complement of Ω_r would be greater than $\delta/2$.
Next basing on the same arguments as in the previous section, we would apply Theorem 8.1 from [13] and would get

$$0 > \delta = U(u_0, v_0) = \inf\{U(u, v) : (u, v) \in \Omega_r \geq \inf\{U(u, v) : (u, v) \in \partial\Omega_r\} > \frac{\delta}{2},$$

Dividing both sides by negative δ we would get $1 < 1/2$, a contradiction.

Hence $U = C_1 - C_2$ must be nonnegative, i.e. copula C_1 dominates in concordance ordering.

This concludes the proof of Theorem 3.5.

To prove Corollary 3.6 we consider two copulas $C(u, v)$ and $D(u, v)$ which fulfill PDE (3.11) with the same A. Due to proved above Theorem 3.5 C dominates D and D dominates C and they have to be equal each other.

Theorem 3.8 follows from the "existence" Theorem 3.2, "PDE characterization" Theorem 3.4 and the "uniqueness" Corollary 3.6.

Let a copula $C(x, y)$ be twice differentiable in a weak sense and fulfill the equation (3.11) with $A(x, y)$ bounded, $|A| < 1$, A and $\sqrt{1 - A^2}$ Lipschitz and A differentiable with respect to the second variable. Due to Theorem 3.2 there exits a $1/2$-self-similar solution $(X_t, Y_t)_{t\geq 1}$ of eq. (3.1) with coefficients A and $B = \sqrt{1 - A^2}$. Let D be the copula of (X_t, Y_t) for all $t \geq 1$. Due to Theorem 3.4 copula D is a solution of the equation (3.11) with coefficient $A(x, y)$. Hence Corollary 3.6 implies that $C = D$. Which means that C is a copula of the mentioned above a $1/2$-self-similar process $(X_t, Y_t)_{t\geq 1}$.

Theorem 3.9 follows from Proposition 4.6, Corollary 4.7, "PDE characterization" Theorem 3.4 and the "uniqueness" Corollary 3.6.
Proposition 4.6 and Corollary 4.7 imply that every cluster point (cluster copula) of the sequence

$$C^n = \frac{1}{n} \int_1^{e^n} \frac{1}{s} C_s^C, \quad n = 1, 2, \dots$$

is a copula of a 1/2-selfsimilar process. If we assume condition **A4** then the cluster copulas fulfill the same PDE. Hence due to Corollary 3.6 they coincide. Which implies that the limit of the sequence C_n exists.

5 Examples

5.1 Gaussian copula

Let us recall that the Gaussian copula is defined as follows:

$$Ga_\rho(u, v) = \Phi_\rho(\Phi^{-1}(u), \Phi^{-1}(v)), \quad \text{for } \rho \in (-1, 1)$$

where Φ_ρ is the joint distribution function of a bi-dimensional standard normal vector, with linear correlation coefficient ρ. For more details see [25] or [20] §4.3.1.

Taking derivatives of Ga_ρ yields

$$\begin{aligned}
\frac{\partial^2 Ga_\rho}{\partial u^2}(u,v) &= \frac{-\rho}{\sqrt{1-\rho^2}}\phi\left(\frac{\Phi^{-1}(v)-\rho\Phi^{-1}(u)}{\sqrt{1-\rho^2}}\right)\frac{1}{\phi(\Phi^{-1}(u))},\\
\frac{\partial^2 Ga_\rho}{\partial v^2}(u,v) &= \frac{-\rho}{\sqrt{1-\rho^2}}\phi\left(\frac{\Phi^{-1}(u)-\rho\Phi^{-1}(v)}{\sqrt{1-\rho^2}}\right)\frac{1}{\phi(\Phi^{-1}(v))},\\
\frac{\partial^2 Ga_\rho}{\partial u\partial v}(u,v) &= \frac{1}{\sqrt{1-\rho^2}}\exp\left(\frac{\Phi^{-1}(u)^2+\Phi^{-1}(v)^2}{2}\right)\times\\
&\quad\times\exp\left(\frac{2\rho\Phi^{-1}(u)\Phi^{-1}(v)-\Phi^{-1}(u)^2-\Phi^{-1}(v)^2}{2(1-\rho^2)}\right).
\end{aligned}$$

We may substitute into eq. (3.11) and obtain

$$\begin{aligned}
&\frac{\partial^2 Ga_\rho}{\partial u^2}(u,v)\varphi(\Phi^{-1}(u))^2+\frac{\partial^2 Ga_\rho}{\partial v^2}(u,v)\varphi(\Phi^{-1}(v))^2\\
&+2\rho\frac{\partial^2 Ga_\rho}{\partial u\partial v}(u,v)\varphi(\Phi^{-1}(u))\varphi(\Phi^{-1}(v))=0
\end{aligned}$$

thus $A(x,y)=\rho$.

By the use of Theorem 3.8 we conclude that for any initial conditions X_1 and Y_1 having standard normal distribution and independent with the increments of the underlying Wiener process there exists a $1/2$-self-similar solution of eq. (3.1), with coefficients $A(x,y)=\rho, B(x,y)=\sqrt{1-\rho^2}$, where $|\rho|<1$, such that the copula of its initial values is Gaussian with linear correlation coefficient ρ.

Since Gaussian copulas are solutions of PDE (3.11) with constant coefficient A equal to ρ, Theorem 3.5 implies that when the quadratic covariation A is bounded from 1 or -1 than the corresponding copula C is dominated by a Gaussian copula or dominates such. The above implies Corollary 3.7.

To conclude the Gaussian example, we consider the solutions of eq. (3.1), with coefficients $A(x,y)=\rho, B(x,y)=\sqrt{1-\rho^2}$ when the initial values (X_1,Y_1) having bivariate normal distribution with standard margins and independent with the increments of the underlying Wiener process, have the linear correlation coefficient ρ_1 different than ρ, $|\rho|<1$ and $|\rho_1|<1$.
We have

$$\begin{aligned}
X_t &= X_1+W_t^1-W_1^1,\\
Y_t &= Y_1+\rho(W_t^1-W_1^1)+\sqrt{1-\rho^2}(W_t^2-W_1^2).
\end{aligned}$$

The joint distribution of the pair (X_t,Y_t) is normal and the linear correlation coefficient equals

$$Corr(X_t,Y_t)=\rho+\frac{1}{t}(\rho_1-\rho).$$

For τ rescaled solution we get

$$Corr(\frac{1}{\sqrt{\tau}}X_{t\tau},\frac{1}{\sqrt{\tau}}Y_{t\tau})=\rho+\frac{1}{t\tau}(\rho_1-\rho)\stackrel{\tau\to\infty}{\longrightarrow}\rho,$$

which implies that the rescaled processes when $\tau\to\infty$ are converging in distribution to the self-similar solution.
Basing on [25], formula 2.6, we get the following uniform bound for corresponding copulas

$$\begin{aligned}
\left|Ga_{\rho+(\rho_1-\rho)/t/\tau}(u,v)-Ga_\rho(u,v)\right| &= \left|\int_\rho^{\rho+(\rho_1-\rho)/t/\tau}\phi_2(\Phi^{-1}(u),\Phi^{-1}(v),r)dr\right|\\
&\le \frac{1}{2\pi}\max\left(\frac{1}{\sqrt{1-\rho^2}},\frac{1}{\sqrt{1-\rho_1^2}}\right)\frac{1}{t\tau}|\rho_1-\rho|\\
&\le \frac{1}{\tau}\frac{1}{2\pi}\frac{1}{\sqrt{1-\max(\rho^2,\rho_1^2)}}|\rho_1-\rho|,
\end{aligned}$$

where $t \geq 1$ and $\phi_2(x, y, r)$ is a density of a bivariate normal distribution with correlation r and standard margins. Note that ϕ_2 is bounded. For all $(x, y) \in \mathbb{R}^2$ and $r \in (-1, 1)$

$$\phi_2(x, y, r) \leq \frac{1}{2\pi}\frac{1}{\sqrt{1 - r^2}}.$$

5.2 FGM copula

The Farlie-Gumbel-Morgenstern copula is defined as

$$C(u, v) = uv\left(1 + a(1 - u)(1 - v)\right), \quad \text{for } a \in [-1, 1]$$

The partial derivatives of C are given by

$$\begin{aligned} \frac{\partial^2 C}{\partial u^2}(u, v) &= 2av(v - 1), \qquad \frac{\partial^2 C}{\partial v^2}(u, v) = 2au(u - 1), \\ \frac{\partial^2 C}{\partial u \partial v}(u, v) &= 1 + a(1 - 2u)(1 - 2v). \end{aligned}$$

After substituting into eq. (3.11) we obtain

$$\begin{aligned} &2av(v - 1)\varphi(\Phi^{-1}(u))^2 + 2au(u - 1)\varphi(\Phi^{-1}(v))^2 \\ &+2A(\Phi^{-1}(u), \Phi^{-1}(v))(1 + a(1 - 2u)(1 - 2v))\varphi(\Phi^{-1}(u))\varphi(\Phi^{-1}(v)) = 0 \end{aligned}$$

where

$$A(x, y) = \frac{a\Phi(y)(1 - \Phi(y))e^{\frac{y^2 - x^2}{2}} + a\Phi(x)(1 - \Phi(x))e^{\frac{x^2 - y^2}{2}}}{1 + a(1 - 2\Phi(x))(1 - 2\Phi(y))}. \tag{5.1}$$

Note that

$$A(0, 0) = \frac{2a \cdot \frac{1}{4}}{1 + a \cdot 0} = \frac{a}{2}.$$

We show that $|A|$ is bounded by $|a|/2$ for any $a \in [-1, 1]$, so the maximum of $|A|$ is attained to the point $(0, 0)$ (compare [16] Section 3.4). Indeed, we have an estimate for any $x, y \in \mathbb{R}$ and $a \in [-1, 1]$

$$|A(x, y)| \leq |a|\frac{\Phi(y)(1 - \Phi(y))e^{\frac{y^2 - x^2}{2}} + \Phi(x)(1 - \Phi(x))e^{\frac{x^2 - y^2}{2}}}{1 - |(1 - 2\Phi(x))(1 - 2\Phi(y))|}. \tag{5.2}$$

Since the right side of (5.2) is invariant with respect to change of sign of x, change of sign of y and symmetry $(x, y) \to (y, x)$ it is enough to show that for $y \geq x \geq 0$, a function

$$\chi(x, y) = 2\Phi(y)\Phi(-y)e^{\frac{1}{2}(y^2 - x^2)} + 2\Phi(x)\Phi(-x)e^{\frac{1}{2}(x^2 - y^2)} + (1 - 2\Phi(x))(1 - 2\Phi(y))$$

is bounded by 1. We observe that the directional derivative

$$y\frac{\partial \chi(x, y)}{\partial x} + x\frac{\partial \chi(x, y)}{\partial y} = 2(y - x)\left[\phi(x)\left(2\Phi(y) - 1\right) - \phi(y)\left(2\Phi(x) - 1\right)\right]$$

is nonnegative for $y \geq x \geq 0$. Hence χ is nondecreasing along the hyperbola

$$y^2 = r^2 + x^2, \quad y \geq x \geq 0.$$

Since moreover $\chi(x, y)$ is positive for $y \geq x \geq 0$, we get

$$\begin{aligned} 0 \leq \chi(x, y) &\leq \sup\{\chi(\xi, \sqrt{r^2 + \xi^2}) : \xi > x\} \\ &= \lim_{\xi \to +\infty} \chi(\xi, \sqrt{r^2 + \xi^2}) = 1 \end{aligned} \tag{5.3}$$

Furthermore, since both nominator and denominator have bounded derivatives and for $|a| < 1$ denominator is bounded from zero, A is differentiable with bounded derivatives. Hence A and $B(x, y) = \sqrt{1 - A(x, y)^2}$ are Lipschitz. Therefore by Theorem 3.4 point 3 there exists a $1/2$-self-similar solution of eq. (3.1), with coefficients A as above and $B(x, y) = \sqrt{1 - A(x, y)^2}$, such that the copula of its initial values is FGM with parameter $a \in (-1, 1)$.

5.3 Archimedean copulas

We recall that by a strict Archimedean copula we mean a copula that takes the following form

$$C_g(u,v) = g^{-1}(g(u)+g(v)),$$

where a function $g\colon [0,1] \to [0,+\infty]$ is continuous, strictly decreasing, convex and

$$g(1) = 1, \quad g(0) = +\infty.$$

The function g is called a generator of the copula C_g. For more details concerning Archimedean copulas, both strict and nonstrict, the reader is referred to [9, 19, 20, 26].

Due to convexity, g is differentiable at all but at most countably many points and the derivative may have jumps. Therefore, for simplicity, we restrict ourselves to twice differentiable generators g. Then we have for $t \in (0,1)$

$$g'(t) < 0, \quad g''(t) \geq 0. \tag{5.4}$$

We start with calculating the partial derivatives of C. In order to simplify the notation let us introduce auxiliary variable $z = g^{-1}(g(u)+g(v))$.

$$\begin{aligned}
\frac{\partial C(u,v)}{\partial u} &= \frac{g'(u)}{g'(z)}, &(5.5)\\
\frac{\partial C(u,v)}{\partial v} &= \frac{g'(v)}{g'(z)}, &(5.6)\\
\frac{\partial^2 C(u,v)}{\partial u^2} &= \frac{g''(u)g'(z) - g'(u)g''(z)\frac{g'(u)}{g'(z)}}{\left(g'(z)\right)^2} \\
&= \frac{g''(u)}{g'(z)} - \left(g'(u)\right)^2 \frac{g''(z)}{\left(g'(z)\right)^3}, &(5.7)\\
\frac{\partial^2 C(u,v)}{\partial v^2} &= \frac{g''(v)g'(z) - g'(v)g''(z)\frac{g'(v)}{g'(z)}}{\left(g'(z)\right)^2} \\
&= \frac{g''(v)}{g'(z)} - \left(g'(v)\right)^2 \frac{g''(z)}{\left(g'(z)\right)^3}, &(5.8)\\
\frac{\partial C(u,v)}{\partial u \partial v} &= g'(u)\cdot(-1)\cdot\left(g'(z)\right)^{-2} g''(z)\frac{g'(v)}{g'(z)} \\
&= -\frac{g''(z)g'(u)g'(v)}{\left(g'(z)\right)^3}. &(5.9)
\end{aligned}$$

If we further assume that $g''(t) \neq$ for all $t \in (0,1)$, then substituting into eq. (3.11) we obtain

$$\begin{aligned}
A(x,y) \;=\; &-\frac{1}{2}\left[\frac{g'(\Phi(x))}{g'(\Phi(y))} - \frac{g''(\Phi(x))\left(g'(C_g(\Phi(x),\Phi(y)))\right)^2}{g''(C_g(\Phi(x),\Phi(y)))g'(\Phi(x))g'(\Phi(y))}\right] e^{\frac{y^2-x^2}{2}} \\
&-\frac{1}{2}\left[\frac{g'(\Phi(y))}{g'(\Phi(x))} - \frac{g''(\Phi(y))\left(g'(C_g(\Phi(x),\Phi(y)))\right)^2}{g''(C_g(\Phi(x),\Phi(y)))g'(\Phi(x))g'(\Phi(y))}\right] e^{\frac{x^2-y^2}{2}}.
\end{aligned}$$

As we show in next subsections for some generators A fulfill assumptions **A2** and **A3**, for some not.

5.3.1 Clayton copula

As a specific example let us consider the following generator

$$g(t) = \frac{1}{\alpha}\left(t^{-\alpha} - 1\right), \qquad \alpha > 0. \tag{5.10}$$

This way we in fact obtain a member of the Clayton family of copulas (compare [9, 20, 26]),

$$C_g(u, v) = \left(u^{-\alpha} + v^{-\alpha} - 1\right)^{-\frac{1}{\alpha}}, \quad \alpha > 0. \tag{5.11}$$

Since

$$g'(t) = -t^{-\alpha-1} \quad \text{and} \quad g''(t) = (\alpha + 1)t^{-\alpha-2},$$

we get

$$A(x, y) = \frac{1}{2}\left((1 - \Phi(y)^{\alpha})\frac{\Phi(y)\phi(x)}{\Phi(x)\phi(y)} + (1 - \Phi(x)^{\alpha})\frac{\Phi(x)\phi(y)}{\Phi(y)\phi(x)}\right). \tag{5.12}$$

We show that A is not bounded by 1. We apply the limit

$$\lim_{z\to-\infty} \frac{|z|\Phi(z)}{\phi(z)} = 1, \tag{5.13}$$

which follows from the de l'Hospital rule:

$$\lim_{z\to-\infty} \frac{\Phi(z)}{z^{-1}\phi(z)} \overset{H}{=} \lim_{z\to-\infty} \frac{\phi(z)}{-\phi(z) - z^{-2}\phi(z)} = -1. \tag{5.14}$$

We substitute $x = z$ and $y = 2z$ and take the limit as z tends to $-\infty$.

$$\begin{aligned} &\lim_{z\to-\infty} A(z, 2z) \\ = &\lim_{z\to-\infty} \frac{1}{2}\left((1 - \Phi(2z)^{\alpha})\frac{1}{2}\frac{2|z|\Phi(2z)}{\phi(2z)}\frac{\phi(z)}{|z|\Phi(z)} + (1 - \Phi(z)^{\alpha}) \cdot 2\frac{|z|\Phi(z)}{\phi(z)}\frac{\phi(2z)}{2|z|\Phi(2z)}\right) \\ = &\frac{1}{2}\left(\frac{1}{2} + 2\right) = 1{,}25. \end{aligned} \tag{5.15}$$

5.3.2 Frank copula

As a next specific example let us consider the following generator

$$g(t) = -\ln\frac{e^{-\alpha t} - 1}{e^{-\alpha} - 1}, \qquad \alpha \neq 0. \tag{5.16}$$

This way we in fact obtain a member of the Frank family of copulas (compare [9, 20, 26]),

$$C_g(u, v) = -\frac{1}{\alpha}\ln\left(1 + \frac{(e^{-\alpha u} - 1)(e^{-\alpha v} - 1)}{e^{-\alpha} - 1}\right), \quad \alpha \neq 0. \tag{5.17}$$

We have

$$g'(t) = \frac{\alpha e^{-\alpha t}}{e^{-\alpha t} - 1} = \frac{\alpha}{1 - e^{\alpha t}}, \tag{5.18}$$

$$g''(t) = \frac{\alpha^2 e^{-\alpha t}}{(e^{-\alpha t} - 1)^2} = \frac{\alpha g'(t)}{e^{-\alpha t} - 1} = e^{\alpha t}(g'(t))^2 \tag{5.19}$$

and

$$\begin{aligned}\frac{g'(u)}{g'(v)} - \frac{g''(u)g'(z)^2}{g''(z)g'(u)g'(v)} &= \frac{g'(u)}{g'(v)} - \frac{e^{\alpha u}g'(u)}{e^{\alpha z}g'(v)} \\ &= \frac{g'(u)}{g'(v)}\left(1 - e^{-\alpha z}e^{\alpha u}\right) \\ &= \frac{g'(u)}{g'(v)}\left(1 - \left(1 + \frac{(e^{-\alpha u}-1)(e^{-\alpha v}-1)}{e^{-\alpha}-1}\right)e^{\alpha u}\right) \\ &= \frac{1-e^{\alpha v}}{1-e^{\alpha u}} \\ &\times \frac{(e^{-\alpha}-1)(1-e^{\alpha u}) - (e^{-\alpha u}-1)(e^{-\alpha v}-1)e^{\alpha u}}{e^{-\alpha}-1} \\ &= \frac{e^{-\alpha}-e^{-\alpha v}}{e^{-\alpha}-1}\cdot\left(1-e^{\alpha v}\right) \\ &= -\frac{1+e^{-\alpha}-e^{-\alpha v}-e^{-\alpha(1-v)}}{1-e^{-\alpha}}\end{aligned} \tag{5.20}$$

Finally the quadratic covariation is given by

$$A_\alpha(x,y) = \frac{1}{2}\left(\chi_\alpha(\Phi(y))\frac{\phi(x)}{\phi(y)} + \chi_\alpha(\Phi(x))\frac{\phi(y)}{\phi(x)}\right), \tag{5.21}$$

where

$$\begin{aligned}\chi_\alpha(t) &= \frac{1-e^{-\alpha t}+e^{-\alpha}-e^{-\alpha(1-t)}}{1-e^{-\alpha}} \\ &= \frac{1-e^{\alpha t}+e^{\alpha}-e^{\alpha(1-t)}}{e^{\alpha}-1}.\end{aligned} \tag{5.22}$$

Since for any $t \in (0,1)$

$$\lim_{\alpha\to\pm\infty}\chi_\alpha(t) = \pm 1,$$

we get for $x^2 \neq y^2$

$$\lim_{\alpha\to\pm\infty}|A_\alpha(x,y)| = \frac{1}{2}\left(\frac{\phi(x)}{\phi(y)} + \frac{\phi(y)}{\phi(x)}\right) > 1.$$

Hence for sufficiently big $|\alpha|$ $|A_\alpha|$ is not bounded by 1.
But for sufficiently small $|\alpha|$ $|A_\alpha|$ is bounded by 1. Specifically it is bounded for $|\alpha| \leq 2$. Indeed, since the functions $|\chi_\alpha(t)|$ and $\phi(\Phi^{-1}(t))$ are concave, we get estimates for $t \in (0,1)$

$$|\chi_\alpha(t)| < |\alpha|\min(t, 1-t) \leq \frac{|\alpha|}{2\phi(0)}\phi(\Phi^{-1}(t)). \tag{5.23}$$

Hence for all $x, y \in \mathbb{R}$

$$\begin{aligned}|A_\alpha(x,y)| &\leq \frac{1}{2}\left(\frac{|\chi_\alpha(\Phi(y))|}{\phi(y)}\phi(x) + \frac{|\chi_\alpha(\Phi(x))|}{\phi(x)}\phi(y)\right) \\ &< \frac{1}{2}\left(\frac{|\alpha|}{2\phi(0)}\phi(0) + \frac{|\alpha|}{2\phi(0)}\phi(0)\right) = \frac{|\alpha|}{2}.\end{aligned} \tag{5.24}$$

Moreover $A(x,y)$ is Lipschitz continuous since derivatives of the ratio $\chi_\alpha(\Phi(\xi))/\phi(\xi)$ are bounded. Hence for $|\alpha| < 2$ Theorem 3.4 point 3 is applicable.

Conflict of interest statement: Authors state no conflict of interest.

References

[1] Baudoin, F. (2014). *Diffusion Processes and Stochastic Calculus*. European Mathematical Society, Zürich.

[2] Bernardi M., F. Durante and P. Jaworski (2017). CoVaR of families of copulas. *Statist. Probab. Lett. 120*, 8–17.
[3] Billingsley, P. (1979). *Probability and Measure*. Wiley, Chichester.
[4] Brezis, H. (2011). *Functional Analysis, Sobolev Spaces and Partial Differential Equations.* Springer, New York.
[5] Cherny, A. S. (2002). On the uniqueness in law and the pathwise uniqueness for stochastic differential equations. *Theory Probab. Appl. 46*(3), 406–419.
[6] Cherubini U., E. Luciano and W. Vecchiato (2004). *Copula Methods in Finance.* Wiley, Chichester.
[7] Choe H. J., C. Ahn, B. J. Kim and Y. Ma (2013). Copulas from the Fokker-Planck equation. *J. Math. Anal. Appl. 406*(2), 519–530.
[8] Durante F. and P. Jaworski (2010). A new characterization of bivariate copulas. *Comm. Statist. Theory Methods 39*(16), 2901–2912.
[9] Durante F. and C. Sempi (2016). *Principles of Copula Theory.* CRS Press, Boca Raton FL.
[10] Embrechts P. and M. Maejima (2002). *Selfsimilar Processes*. Princeton University Press.
[11] Evans, L. (1998). *Partial Differential Equations.* American Mathematical Society, Providence RI.
[12] Gihman Ĭ. Ī. and A. V. Skorohod (1972). *Stochastic Differential Equations.* Springer, New York.
[13] Gilbarg, D. and N. S. Trudinger (2001). *Elliptic Partial Differential Equations of Second Order.* Springer, Berlin.
[14] Jaworski, P. (2017). On the Conditional Value at Risk (CoVaR) for tail-dependent copulas. *Depend. Model. 5*, 1–19.
[15] Jaworski, P. (2019). On Copula-Ito processes. *Depend. Model. 7*, 322–347.
[16] Jaworski P. and M. Krzywda (2013). Coupling of Wiener processes by using copulas. *Statist. Probab. Lett. 83*(9), 2027–2033.
[17] Jaworski P., F. Durante and W. Härdle (2013). *Copulae in Mathematical and Quantitative Finance.* Springer, Heidelberg.
[18] Jaworski P., F. Durante, W. Härdle and T. Rychlik (2010). *Copula Theory and its Applications.* Springer, Heidelberg.
[19] Joe, H. (1997). *Multivariate Models and Dependence Concepts.* Chapman & Hall, London.
[20] Joe, H. (2015). *Dependence Modeling with Copulas.* CRS Press, Boca Raton FL.
[21] Khasminskii, R. (2012). *Stochastic Stability of Differential Equations*. Springer, Heidelberg.
[22] Krzywda, M. (2020). *Copulas of Self-Similar Itô Diffusion Processes and Their Applications in Financial Mathematics.* Phd Thesis, Jagiellonian University, Poland (in Polish).
[23] Ladyzhenskaya, O. A. and N. N. Ural'tseva (1968). *Linear and Quasilinear Elliptic Equations*. Academic Press, New York.
[24] Łojasiewicz, S. (1988). *An Introduction to the Theory of Real Functions.* Third edition. Wiley, Chichester.
[25] Meyer, C. (2013). The bivariate normal copula. *Comm. Statist. Theory Methods 42*(13), 2402–2422.
[26] Nelsen, R. B. (2006). *An Introduction to Copulas.* Second edition. Springer, New York.
[27] Oksendal, B. (2003). *Stochastic Differential Equations. An Introduction with Applications.* Sixth edition. Springer, Berlin.
[28] Rudin, W. (1966). *Real and Complex Analysis*. McGraw-Hill, New York.
[29] Schuss, Z. (2010). *Theory and Applications of Stochastic Processes. An Analytical Approach.* Springer, New York.
[30] Schwartz, L. (1967). *Analyse Mathématique*. Hermann, Paris.
[31] Sempi, C. (2010). Coupled Brownian motion. In C. Borgelt, G. Gonzáles-Rodríguez, W. Trutschnig, M. A. Lubiano, M. Ángeles Gil, P. Grzegorzewski and O. Hryniewicz (Eds.), *Combining Soft Computing and Statistical Methods in Data Analysis*, pp. 569–574. Springer, Berlin.
[32] Sempi, C. (2016). Copulae of processes related to the Brownian motion: a brief survey. In S. Saminger-Platz and R. Mesiar (Eds.), *On Logical, Algebraic, and Probabilistic Aspects of Fuzzy Set Theory*, pp. 173–180. Springer, Cham.
[33] Taqqu, M. S. (2003). Fractional Brownian motion and long-range dependence. In P. Doukhan, G. Oppenheim and M. S. Taqqu (Eds.), *Theory and Applications of Long Range-Dependence*, pp. 5–38. Birkhauser, Boston.

Research Article
Special Issue in memory of Abe Sklar

Open Access

Jonathan Ansari* and Ludger Rüschendorf

Sklar's theorem, copula products, and ordering results in factor models

https://doi.org/10.1515/demo-2021-0113
Received January 19, 2021; accepted July 16, 2021

Abstract: We consider a completely specified factor model for a risk vector $X = (X_1, \dots, X_d)$, where the joint distributions of the components of X with a risk factor Z and the conditional distributions of X given Z are specified. We extend the notion of $*$-product of d-copulas as introduced for $d = 2$ and continuous factor distribution in Darsow et al. [6] and Durante et al. [8] to the multivariate and discontinuous case. We give a Sklar-type representation theorem for factor models showing that these $*$-products determine the copula of a completely specified factor model. We investigate in detail approximation, transformation, and ordering properties of $*$-products and, based on them, derive general orthant ordering results for completely specified factor models in dependence on their specifications. The paper generalizes previously known ordering results for the worst case partially specified risk factor models to some general classes of positive or negative dependent risk factor models. In particular, it develops some tools to derive sharp worst case dependence bounds in subclasses of completely specified factor models.

Keywords: componentwise convex copulas, concordance order, conditional distribution function, conditional independence, factor model, product of copulas

MSC: 60E15, 60E05, 28A50

1 Introduction

A relevant class of distributions for modeling dependencies are factor models where each component of the underlying random vector $X = (X_1, \dots, X_d)$ is supposed to depend on some common random factor Z through

$$X_i = f_i(Z, \varepsilon_i), \quad 1 \le i \le d,$$

for some functions f_i and a random vector $(\varepsilon_1, \dots, \varepsilon_d)$ that is independent of Z . In this paper, we consider the case where Z is a real-valued random variable. If the bivariate distribution of (X_i, Z) is specified and the distribution of $X|Z = z$ is known for all i and z , then the distribution of X is fully specified. We denote this setting a completely specified factor model (CSFM).

For applications to risk modeling, partially specified factor models (PSFMs) are introduced in Bernard et al. [5]. In these models, the distributions of (X_i, Z) are specified. The joint distribution of $(\varepsilon_1, \dots, \varepsilon_d)$ is, however, not prescribed. This means, that only the distributions of X_i and Z as well as the copula $D^i = C_{X_i,Z}$ of (X_i, Z) are given. Then, the worst case distribution in the PSFM is determined by the conditionally comonotonic random vector $X_Z^c = (F^{-1}_{X_1|Z}(U), \dots, F^{-1}_{X_d|Z}(U))$, where $U \sim U(0, 1)$ is independent of Z , assuming generally a non-atomic underlying probability space $(\Omega, \mathcal{A}, P)$. If Z has a continuous distribution, the copula of X_Z^c is given by the upper product of the bivariate copulas D^i , see [2].

***Corresponding Author: Jonathan Ansari:** Department of Quantitative Finance, University of Freiburg, Germany, E-mail: jonathan.ansari@finance.uni-freiburg.de
Ludger Rüschendorf: Department of Mathematical Stochastics, University of Freiburg, Germany, E-mail: ruschen@stochastik.uni-freiburg.de

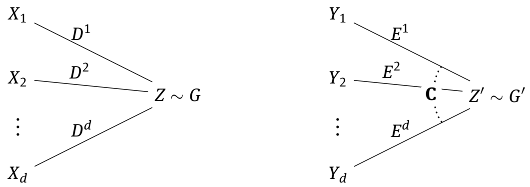

Figure 1 On the left: a partially specified factor model with dependence specifications $D^1, \dots, D^d$ and risk factor distribution function G . On the right: a completely specified factor model with dependence specifications $E^1, \dots, E^d$, conditional copula family $\mathbf{C}$ and factor distribution function G' .

In standard factor models, the individual factors $\varepsilon_1, \dots, \varepsilon_d$ are assumed to be independent. Then, the distribution of X is completely specified and the components of X are conditionally independent given $Z = z$ for all z . Further, the copula of X is then given by the conditional independence product of the bivariate specifications D^i , which is an extension of the bivariate copula product introduced in Darsow et al. [6] to arbitrary dimension, see [15].

In this paper, we introduce and study the $*$-product of copulas as an extension of the bivariate copula product considered in [8] to the multivariate case and to general factor distribution functions in order to model the copula of $X = (f_i(Z, \varepsilon_i))_i$ for general dependence structures among $(\varepsilon_1, \dots, \varepsilon_d)$ and also discontinuous Z . We provide a simple representation of a conditional distribution function by the corresponding univariate distribution functions and a generalized derivative of the associated copula. Then, we derive a Sklar-type theorem implying that the dependence structure of X is determined by the $*$-product of the dependence specifications in the CSFM. Further, we establish a general continuity result for the $*$-product in dependence on all its arguments which is useful for corresponding approximation results. We study transformation properties of the $*$-product and introduce, as a counterpart of the upper product, the lower product of bivariate copulas in the two- and three-dimensional case.

In Section 3, we derive general lower and upper orthant ordering results for the $*$-product in dependence on the copula specifications. This requires the consideration of integral inequalities like the rearrangement results of Lorentz [16] and Fan and Lorentz [11]. We extend and strengthen several recent results on the lower and upper orthant ordering of upper products to general $*$-products. In particular, we show that componentwise convexity of the conditional copulas plays an important role for the ordering of the $*$-products. We introduce the $\leq_{\partial_2 S}$-ordering on the set of bivariate copulas based on the Schur-ordering of copula derivatives allowing to derive a meaningful comparison criterion. We show that many well-known copula families satisfy this ordering.

Finally, in Section 4, we combine the $*$-product ordering results with the ordering of marginal distributions and obtain several general ordering results in CSFMs. As a consequence, this yields maximum elements and, thus, sharp bounds w.r.t. the lower and upper orthant ordering for classes of PSFMs as well as for classes of CSFMs with the conditional independence assumption.

2 The $*$-product of copulas in completely specified factor models

A d-copula is a distribution function $C\colon [0, 1]^d \to [0, 1]$ with uniform univariate marginal distribution functions. Due to Sklar's theorem, every d-dimensional distribution function F can be decomposed into a composition of a d-copula C and the univariate marginal distribution functions $F_1, \dots, F_d$ of F , i.e.,

$$F(x) = C\left(F_1(x_1), \dots, F_d(x_d)\right) \tag{1}$$

for all $x = (x_1, \dots, x_d) \in \mathbb{R}^d$. The copula C is uniquely determined on the Cartesian product $\times_{i=1}^d \mathsf{Ran}(F_i)$ of the ranges of F_i, $1 \le i \le d$. Further, for every d-copula and for all distribution functions $F_1, \dots, F_d$, the right-hand side in (1) defines a d-variate distribution function, see the original papers of Sklar [29] and Schweizer and Sklar [27], see also Nelsen [22], Rüschendorf [25], and Durante and Sempi [10]. Denote by $\mathcal{C}_d$ the set of all d-copulas and by $\mathcal{F}^d$ ($\mathcal{F}_c^d$) the set of (continuous) d-dimensional distribution functions.

In the setting of a completely specified factor model, the distribution function F of $X = (X_1, \dots, X_d)$ can be decomposed into

$$F(x_1, \dots, x_d) = \int_{-\infty}^{\infty} F_z(x_1, \dots, x_d)\, \mathrm{d}G(z) = \int_{-\infty}^{\infty} C_z\left(F_{1|z}(x_1), \dots, F_{d|z}(x_d)\right) \mathrm{d}G(z),$$

where F_z is the conditional distribution function of $(X_1, \dots, X_d)|Z = z$ with univariate marginal conditional distribution functions $F_{i|z}$ and conditional copula $C_z \in \mathcal{C}_d$. Each $F_{i|z}$ depends via

$$D^i\left(F_i(x), G(z)\right) = F_{X_i,Z}(x, z) = \int_{-\infty}^{z} F_{i|s}(x)\, \mathrm{d}G(s)$$

only on the dependence specification $D^i = C_{X_i,Z}$ and the marginal distribution functions F_i and G, where $G = F_Z$ denotes the distribution function of Z.

Altogether, this motivates to introduce the $*$-product of copulas as a product of the specifications $D^1, \dots, D^d \in \mathcal{C}_2$, of the conditional copulas $(C_z)_z$, $C_z \in \mathcal{C}_d$, and of the risk factor distribution function $G \in \mathcal{F}^1$. In a Sklar-type theorem, we show that the $*$-product is a copula that describes the dependence structure of the risk vector X in the CSFM. We give the basic properties of the $*$-products that are used in the following sections to develop several ordering results for $*$-products and, thus, ordering results for CSFMs.

Our results extend the bivariate $*$-product considered in Durante et al. [8] and the bivariate conditional independence product introduced in Darsow et al. [6]. A discussion of some properties of bivariate $*$-products is given in Durante and Sempi [10, Section 5.5]. An important particular case of the $*$-product in the present paper is the multivariate conditional independence product which describes the dependence structure of the commonly used factor models with conditional independence assumption, compare Krupskii and Joe [15]. The particular case of upper products that corresponds to upper risk bounds in partially specified factor models has been investigated in [2]. As a counterpart of upper products, we introduce the lower product of bivariate copulas that describes best case bounds in the two-, respectively, three-dimensional PSFM.

2.1 Definition of $*$-products

The consideration of general factor distributions needs the following notion of generalized differentiation. For $G \in \mathcal{F}^1$ denote by

$$\iota_G \colon [0,1] \to \mathsf{Ran}(G), \qquad t \mapsto G \circ G^{-1}(t),$$
$$\iota_G^- \colon [0,1] \to \mathsf{Ran}(G^-), \qquad t \mapsto G^- \circ G^{-1}(t),$$

the transformation of the identity w.r.t. to G, resp. G^-, where $G^{-1} \colon [0,1] \to \mathbb{R} \cup \{\pm\infty\}$ given by $G^{-1}(u) := \inf\{x \mid G(x) \ge u\}$, $\inf \emptyset = \infty$, is the generalized inverse of G, and G^- is the left-continuous version of G. Several properties of the transformations ι_G and ι_G^- are given in Lemma A.1 in the Appendix, see also Figure 2.

Define for a function $f \colon [0,1] \to \mathbb{R}$ the generalized differential operator ∂^G by the left-hand limit

$$\partial^G f(t_0) := \lim_{t \nearrow t_0} \frac{f(\iota_G(t_0)) - f(\iota_G^-(t))}{\iota_G(t_0) - \iota_G^-(t)}, \tag{2}$$

$t_0 \in (0,1]$, if the limit exists. As usual, denote by ∂_i^G the operator ∂^G which is applied to the i-th coordinate of a function of several arguments.

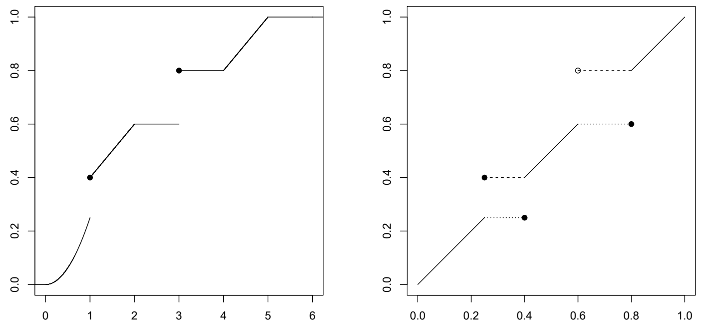

Figure 2 On the left: a distribution function G ; on the right: its corresponding transformations ι_G (dashed and solid line) and ι_G^- (dotted and solid line) which both coincide with the identity function on the interior of the range of G . Note that ι_G^- is left-continuous, and ι_G is neither left- nor right-continuous, see Lemma A.1.

Remark 2.1. *(a) The denominator in* (2) *is positive for all* $0 \le t < t_0 \le 1$ *because* $\iota_G(t_0) \ge t_0 > t \ge \iota_G^-(t)$ *by Lemma A.1*(iv).

(b) If f is left-continuous and if the (ordinary) left-hand derivative $f'_-(t_0) := \lim_{t \nearrow t_0} \frac{f(t_0)-f(t)}{t_0-t}$ *exists, then* $\partial^G f(t_0)$ *exists for all* $G \in \mathcal{F}^1$ *. To see this, we know by* (a) *that* $\iota_G(t_0) \ge t_0 \ge \lim_{t \nearrow t_0} \iota_G^-(t)$ *. Hence, if* $\iota_G(t_0) = \lim_{t \nearrow t_0} \iota_G^-(t) = t_0$ *, then* $\partial^G f(t_0) = f'_-(t_0)$ *. If* $\iota_G(t_0) > \lim_{t \nearrow t_0} \iota_G^-(t)$ *, then* $\partial^G f(t_0)$ *exists since f and ι_G are left-continuous, see Lemma A.1*(vi).

(c) A useful transformation property of ∂^G is that

$$\partial^G f(t) = \partial^G f(\iota_G(t)) = \partial^G f(G(x)) \quad \text{for all } G \in \mathcal{F}^1 \text{ and for Lebesgue-almost all } t\,, \tag{3}$$

where $x = G^{-1}(t)$ *. This is a consequence of Lemma A.1*(v) *considering the cases where G is continuous at x or has a jump discontinuity at x , compare equations* (38) *and* (39) *in the proof of Theorem 2.2.*

The following result gives the representation of a conditional distribution function by the univariate marginals and the generalized partial derivative of the corresponding copula.

Theorem 2.2 (Representation of conditional distribution functions)**.**
For $F, G \in \mathcal{F}^1$ *, let* $X \sim F$ *and* $Z \sim G$ *be real random variables with copula* $C \in \mathcal{C}_2$ *, i.e.,* $C = C_{X,Z}$ *. Then, the following statements hold true:*

(i) For all $x \in \mathbb{R}$ *, there exists a G-null set N_x such that the conditional distribution function of X given $Z = z$ evaluated at x is represented by*

$$F_{X|Z=z}(x) = \lim_{h \searrow 0} \frac{C(F(x), G(z)) - C(F(x), G(z-h))}{G(z) - G(z-h)} = \partial_2^G C(F(x), G(z)) \tag{4}$$

for all $z \in N_x^c$ *.*

(ii) There exists a G-null set N such that

$$F_{X|Z=z}(x) = \lim_{w \searrow x} \partial_2^G C(F(w), G(z)) \tag{5}$$

for all $x \in \mathbb{R}$ *and for all* $z \in N^c$ *.*

The proof is given in the Appendix.

Remark 2.3. *(a) For the representation of the conditional distribution function in* (4) *and* (5)*, we make use of the left-hand limit in the definition of the generalized differential operator given by* (2)*. If G has a discontinuity at z, then the operator ∂_2^G is the difference quotient operator w.r.t. the second component of C between $G(z)$ and $G^-(z)$. If G is continuous at z, the operator ∂_2^G reduces to the ∂_2^--operator denoting the left-hand partial derivative with respect to the second variable. Hence, if G is continuous for all z, then it holds that $\partial_2^G = \partial_2^-$. Denote by ∂_2 the operator which takes the partial derivative w.r.t. the second component of a multivariate function. Since copulas are almost surely partially differentiable, see Nelsen [22, Theorem 2.2.7], it holds for all u, that $\partial_2^- C(u,v) = \partial_2 C(u,v)$ for almost all v.*

(b) We point out that the right-hand expression in (4) *is not necessarily right-continuous in x, and, thus, it does not generally define a distribution function in x. However, in the following definition of the $*$-product as well as in most results of the paper, we integrate over the conditioning variable and, then, this representation of the conditional distribution function is appropriate.*

In the following definition, we extend the $*$-product introduced by Darsow et al. [6] for Markov structures, and, for arbitrary conditional dependencies, by Durante et al. [8] (for $d = 2$) and [2] (for $d \geq 2$) to $G \in \mathcal{F}^1$ allowing also discontinuous factor distribution functions.

We need a measurability assumption which is implicitly assumed in the above mentioned literature by the definition of the corresponding integrals. We call a family $\mathbf{B} = (B_t)_{t\in[0,1]}$ of d-copulas *measurable* if the mapping $(t,u) \mapsto B_t(u)$, $(t,u) \in [0,1]\times[0,1]^d$, is measurable.

The $*$-product of bivariate copulas is defined in dependence on a measurable family $\mathbf{B} = (B_t)_{t\in[0,1]}$ of d-dimensional copulas and on a distribution function $G \in \mathcal{F}^1$.

Definition 2.4 ($*$-product of copulas)**.**

(i) Let $\mathbf{B} := (B_t)_{t\in[0,1]}$ be measurable, $B_t \in \mathcal{C}_d$, $0 \leq t \leq 1$, and $G \in \mathcal{F}^1$. For bivariate copulas $D^1, \ldots, D^d \in \mathcal{C}_2$, the ($d$-dimensional) $*$-product *of $D^1, \ldots, D^d$ w.r.t. $\mathbf{B}$ and G is defined by*

$$\mathop{\ast}\nolimits_{\mathbf{B},G} D^i\,(u) := \mathop{\ast}\nolimits_{i=1,\mathbf{B},G}^{d} D^i\,(u) := \int_0^1 B_t^G\left(\partial_2^G D^1(u_1,t),\ldots,\partial_2^G D^d(u_d,t)\right)\,\mathrm{d}t \tag{6}$$

for $u = (u_1,\ldots,u_d) \in [0,1]^d$ where B_t^G is defined by

$$B_t^G := \begin{cases} B_t\,, & \text{if } \iota_G^-(t) = \iota_G(t)\,,\\ \frac{1}{\iota_G(t)-\iota_G^-(t)}\int_{\iota_G^-(t)}^{\iota_G(t)} B_s\,\mathrm{d}s\,, & \text{if } \iota_G^-(t) \neq \iota_G(t)\,. \end{cases} \tag{7}$$

(ii) If there exists a copula $B \in \mathcal{C}_d$ such that $B_t^G = B$ for almost all t, then we use the notion $\mathop{\ast}\nolimits_{B,G} D^i := \mathop{\ast}\nolimits_{\mathbf{B},G} D^i$ and call it simplified $*$-product *of $D^1, \ldots, D^d$ w.r.t. B and G.*

(iii) If G is continuous, then the (simplified) $*$-product is abbreviated by $\mathop{\ast}\nolimits_{\mathbf{B}} D^i := \mathop{\ast}\nolimits_{\mathbf{B},G} D^i$ and $\mathop{\ast}\nolimits_{B} D^i := \mathop{\ast}\nolimits_{B,G} D^i$, respectively.

Note that the number d of bivariate copulas is typically clear from the context and therefore the simplified notation is used. We also sometimes use the notation $D^1 *_{\mathbf{B},G} \cdots *_{\mathbf{B},G} D^d := \mathop{\ast}\nolimits_{\mathbf{B},G} D^i$ for the $*$-product of d bivariate copulas $D^1 \ldots, D^d$ w.r.t. to $\mathbf{B}$ and G.

Note that for fixed $u_1, \ldots, u_d \in [0,1]$ the integrand in (6) is well-defined as a consequence of Remark 2.1(b) because copulas are Lipschitz-continuous. The justification for the simplified notation in (iii) of the above definition is due to Proposition 2.14.

As usual, we denote by Π^d, M^d, and W^d, defined by

$$\Pi^d(u) := u_1 \cdots u_d\,, \qquad M^d(u) := \min_{1\leq i\leq d}\{u_i\}\,, \qquad W^d(u) := \max_{1\leq i\leq d}\left\{\sum_{i=1}^d u_i - d + 1\,, 0\right\}\,,$$

the *product copula*, the *upper Fréchet copula*, and the *lower Fréchet bound*, respectively, where W^d is a copula only for $d \leq 2$. As special $*$-products, we consider the following simplified products of bivariate copulas.

Definition 2.5 (Some specific simplified $*$-products)**.**
(i) *The* conditional independence product *is defined as* $\Pi_G D^i := \ast_{\Pi^d,G} D^i$.
(ii) *The* upper product *is defined as* $\bigvee_G D^i := D^1 \vee_G \cdots \vee_G D^d := \ast_{M^d,G} D^i$.
(iii) *The* lower product *is defined as* $\bigwedge_G D^i := D^1 \wedge_G \cdots \wedge_G D^d := \ast_{W^d,G} D^i$.

If G is continuous, we abbreviate the independence product by $\Pi D^i = \Pi_{i=1}^d D^i$, the upper product by $\bigvee D^i = \bigvee_{i=1}^d D^i = D^1 \vee \cdots \vee D^d$, and the lower product by $\bigwedge D^i = \bigwedge_{i=1}^d D^i = D^1 \wedge \cdots \wedge D^d$.

Since W^d is a copula only if $d \leq 2$, we clarify that for $d \geq 3$, the lower product is defined in the sense of (6).

The following result shows that the $*$-product is a copula. It extends [2, Proposition 2.1] from continuous to general factor distribution functions.

Proposition 2.6. *For all measurable $\mathbf{B} = (B_t)_{t \in [0,1]}$, $B_t \in \mathcal{C}_d$ for all t , for all $G \in \mathcal{F}^1$, and for all $D^1, \ldots, D^d \in \mathcal{C}_2$, the $*$-product $\ast_{\mathbf{B},G} D^i$ is a d-copula.*

Proof. Due to Theorem 2.2, the functions H_z^i , $1 \leq i \leq d$, defined by $H_z^i(u) := \lim_{v \searrow u} \partial_2^G D^i(v, G(z))$ for $u \in [0,1)$ and $H_z^i(1) := \partial_2^G D^i(1, G(z)) = 1$ are univariate distribution functions for G-almost all $z \in \mathbb{R}$. Then, by Sklar's Theorem, F_z defined by

$$F_z(u_1, \ldots, u_d) := B_{G(z)}^G(H_z^1(u_1), \ldots, H_z^d(u_d)), \quad (u_1, \ldots, u_d) \in [0,1]^d ,$$

is a d-dimensional distribution function on $[0,1]^d$, where (B_t^G) is defined by (7). It follows that

$$\begin{aligned} \ast_{\mathbf{B},G} D^i (u_1, \ldots, u_d) &= \int_0^1 B_t^G \left(\partial_2^G D^1(u_1, t), \ldots, \partial_2^G D^d(u_d, t) \right) \mathrm{d}t \\ &= \int_0^1 B_{\iota_G(t)}^G \left(\partial_2^G D^1(u_1, \iota_G(t)), \ldots, \partial_2^G D^d(u_d, \iota_G(t)) \right) \mathrm{d}t \\ &= \int_{\mathbb{R}} B_{G(z)}^G \left(\partial_2^G D^1(u_1, G(z)), \ldots, \partial_2^G D^d(u_d, G(z)) \right) \mathrm{d}G(z) \\ &= \int_{\mathbb{R}} B_{G(z)}^G \left(H_z^1(u_1), \ldots, H_z^d(u_d) \right) \mathrm{d}G(z) = \int_{\mathbb{R}} F_z(u_1, \ldots, u_d) \, \mathrm{d}G(z) , \end{aligned} \tag{8}$$

where we apply (3) for the second equality and use that $B_t^G = B_{\iota_G(t)}^G$ which follows from Lemma A.1(v). The third equality follows from the transformation formula, see, e.g., [32, Theorem 2]. For the fourth equality, we use for fixed $(u_1, \ldots, u_d) \in [0,1]^d$ that $H_z^i(u_i) = \partial_2^G D^i(u_i, G(z))$, $1 \leq i \leq d$, for G-almost all z , see Theorem 2.2(i). Since the last integral is a mixture of distribution functions, the product $\ast_{\mathbf{B},G} D^i$ is a distribution function. The measurability of $F_z(u_1, \ldots, u_d)$ in z is a consequence of the measurability of $\mathbf{B}$ and, by (4), of $t \mapsto \partial_2^G D^i(u_i, t)$ for all $u_i \in [0,1]$, $1 \leq i \leq d$.

It remains to show that $\ast_{\mathbf{B},G} D^i$ has uniform marginals. For $i \in \{1, \ldots, d\}$, let $v = (v_1, \ldots, v_d)$ with $v_i \in [0,1]$ and $v_j = 1$ for all $j \neq i$. Since $\partial_2^G D^j(v_j, t) = 1$ for all t and $j \neq i$, it follows that

$$\ast_{\mathbf{B},G} D^j (v_1, \ldots, v_d) = \int_0^1 \partial_2^G D^i(v_i, t) \, \mathrm{d}t = \int_{\mathbb{R}} \partial_2^G D^i(v_i, G(z)) \, \mathrm{d}G(z) = v_i ,$$

where the first equality holds due to the uniform marginals of the copula B_t^G , the second one is a consequence of the transformation formula and (3), and the last equality is given by Theorem 2.2 and the disintegration theorem. $\square$

2.2 Sklar-type theorem for factor models

The following theorem describes the meaning of the notion of $*$-products. It is a version of Sklar's Theorem for completely specified factor models and states that the dependence structure of a random vector $(X_i)_{1\le i\le d}$ that follows a completely specified factor model, $X_i = f_i(Z, \varepsilon_i)$, is given by a $*$-product of the specifications $G = F_Z$, $C^i = C_{X_i,Z}$, and $B_t^G = C_{X_1,\dots,X_d|Z=G^{-1}(t)}$, $t \in [0, 1]$.

Theorem 2.7 (Sklar's Theorem for completely specified factor models)**.**
Let $F_{1,\dots,d+1} \in \mathcal{F}^{d+1}$ be a $(d+1)$-dimensional distribution function with univariate marginal distribution functions $F_1, \dots, F_{d+1}$. Denote by $F_{i,d+1}$ the bivariate marginal distribution function of its $(i, d+1)$-marginal, by $F_{1,\dots,d}$ the distribution function of its first d components, and by $F_{1,\dots,d|F_{d+1}^{-1}(t)}$ the conditional distribution function of its first d components given that the $(d+1)$-st component equals $F_{d+1}^{-1}(t)$. Then, there exist bivariate copulas $C^1, \dots, C^d \in \mathcal{C}_2$ and a measurable family $\mathbf{B} = (B_t)_{t\in[0,1]}$ of d-copulas such that

$$F_{i,d+1}(x_i, x_{d+1}) = C^i(F_i(x_i), F_{d+1}(x_{d+1})) \qquad \text{for } i = 1, \dots, d\,, \tag{9}$$

$$F_{1,\dots,d|F_{d+1}^{-1}(t)}(x_1, \dots, x_d) = B_t^{F_{d+1}}\left(\left(\lim_{w_i\searrow x_i} \partial_2^{F_{d+1}} C^i(F_i(w_i), t)\right)_{1\le i\le d}\right) \qquad \text{for almost all } t \in [0, 1]\,, \tag{10}$$

$$F_{1,\dots,d}\,(x_1, \dots, x_d) = *_{\mathbf{B},F_{d+1}} C^i\,(F_1(x_1), \dots, F_d(x_d)) \tag{11}$$

for all $(x_1, \dots, x_{d+1}) \in \mathbb{R}^{d+1}$.
Conversely, for distribution functions $F_1, \dots, F_{d+1} \in \mathcal{F}^1$, bivariate copulas $C^1, \dots, C^d \in \mathcal{C}_2$ and a measurable family $\mathbf{B} = (B_t)_{t\in[0,1]}$ of d-copulas, the family $\left(F_{1,\dots,d|F_{d+1}^{-1}(t)}\right)_{t\in[0,1]}$ in (10) defines a $(d+1)$-dimensional distribution function $F_{1,\dots,d+1}$ with bivariate marginal distribution functions $F_{i,d+1}$ given by (9) and d-variate distribution function $F_{1,\dots,d}$ given by (11).

Further, for $1 \le i \le d$, the copula C^i is uniquely determined on $\mathrm{Ran}(F_i) \times \mathrm{Ran}(F_{d+1})$, and $B_t^{F_{d+1}}$ is uniquely determined on $\bigtimes_{i=1}^d \mathrm{Ran}\left(\lim_{w_i\searrow\cdot} \partial_2^{F_{d+1}} C^i(F_i(w_i), t)\right)$ for almost all $t \in [0, 1]$.

Proof. Due to Sklar's Theorem in the bivariate case, there exist $C^1, \dots, C^d \in \mathcal{C}_2$ such that (9) holds for all $(x_1, \dots, x_{d+1}) \in \mathbb{R}^{d+1}$. The univariate marginal distribution functions of $F_{1,\dots,d|F_{d+1}^{-1}(t)}$ are given by

$$F_{i|F_{d+1}^{-1}(t)}(x) = \lim_{w\searrow x} \partial_2^{F_{d+1}} C^i(F_i(w), t)\,, \quad \text{for all } x \in \mathbb{R} \text{ and for almost all } t \in [0, 1]\,, \tag{12}$$

$1 \le i \le d$, see Theorem 2.2(ii). Due to Sklar's Theorem in the d-variate case, $B_t \in \mathcal{C}_d$, $t \in [0, 1]$, with

$$B_t(u) = F_{1\dots d|F_{d+1}^{-1}(t)}\left(F^{-1}_{1|F_{d+1}^{-1}(t)}(u_1), \dots, F^{-1}_{d|F_{d+1}^{-1}(t)}(u_d)\right)\,, \quad u = (u_1, \dots, u_d) \in [0, 1]^d\,,$$

for almost all t defines a family $\mathbf{B} = (B_t)_{t\in[0,1]}$ of d-copulas such that (10) holds true. Note that $\mathbf{B}$ is measurable because the mappings $[0, 1]\times\mathbb{R}^d \ni (t, x) \mapsto F_{1\dots d|F_{d+1}^{-1}(t)}(x)$ and $[0, 1]\times[0, 1] \ni (t, u_i) \mapsto F^{-1}_{i|F_{d+1}^{-1}(t)}(u_i)$, $1 \le i \le d$, are measurable.

To show (11), we apply the disintegration theorem and obtain for all $x = (x_1, \dots, x_d) \in \mathbb{R}^d$ that

$$F_{1,\dots,d}(x) = \int_0^1 F_{1,\dots,d|F_{d+1}^{-1}(t)}(x)\,\mathrm{d}t = \int_0^1 B_t^{F_{d+1}}\left(\left(\partial_2^{F_{d+1}} C^i(F_i(x_i), t)\right)_{1\le i\le d}\right)\,\mathrm{d}t = *_{\mathbf{B},F_{d+1}} C^i\,(F_i(x_i))\,,$$

where for the second equality we use the representation in (10) and that $\lim_{w_i\searrow x_i} \partial_2^{F_{d+1}} C^i(F_i(w_i), t) = \partial_2^{F_{d+1}} C^i(F_i(x_i), t)$ for all t outside a Lebesgue-null set $N_x \subset [0, 1]$, see Theorem 2.2.

For the converse direction, let $F_1, \dots, F_{d+1} \in \mathcal{F}_1$, $C^1, \dots, C^d \in \mathcal{C}_2$ and $\mathbf{B} = (B_t)_{t\in[0,1]}$ be measurable, $B_t \in \mathcal{C}_d$ for all t. Then, by Theorem 2.2 and Sklar's Theorem, $F_{1,\dots,d|F_{d+1}^{-1}(t)}$ given by (10) defines a d-variate distribution function for almost all $t \in [0, 1]$. As a consequence of the measurability of $\mathbf{B}$, the mapping $t \mapsto F_{1,\dots,d|F_{d+1}^{-1}(t)}(x)$ is measurable for all $x \in \mathbb{R}^d$, compare (8). Hence, $F_{1,\dots,d+1}$ defined by

$$F_{1,\dots,d+1}(x, z) = \int_0^{F_{d+1}(z)} F_{1,\dots,d|F_{d+1}^{-1}(t)}(x_1, \dots, x_d)\,\mathrm{d}t\,,$$

$x = (x_1, \dots, x_d) \in \mathbb{R}^d$, $z \in \mathbb{R}$, is a $(d+1)$-dimensional distribution function with marginal distribution function of the first d components given by

$$\begin{aligned} F_{1,\dots,d}(x) = F_{1,\dots,d+1}(x,\infty) &= \int_0^1 B_t^{F_{d+1}}\left(\left(\partial_2^{F_{d+1}} C^i(F_i(x_i),t)\right)_{1\le i\le d}\right) \mathrm{d}t \\ &= \int_0^1 B_t^{F_{d+1}}\left(\left(\lim_{w_i \searrow x_i} \partial_2^{F_{d+1}} C^i(F_i(w_i),t)\right)_{1\le i\le d}\right) \mathrm{d}t = \ast_{\mathbf{B},F_{d+1}} C^i\left(F_1(x_1),\dots,F_d(x_d)\right) \end{aligned}$$

and bivariate marginal distribution functions w.r.t. to the i-th and $(d+1)$-st component given by

$$F_{i,d+1}(x_i,z) = F_{1,\dots,d+1}(\infty,\dots,\infty,x_i,\infty,\dots,\infty,z) = \int_0^{F_{d+1}(z)} \partial_2^{F_{d+1}} C^i\left(F_i(x_i),t\right) \mathrm{d}t = C^i(F_i(x_i),F_{d+1}(z))\,.$$

The uniqueness properties follow directly from the uniqueness properties in Sklar's Theorem. □

Remark 2.8. *(a) For $F_1,\dots,F_d,G \in \mathcal{F}^1$, let $(X_1,\dots,X_d,Z)$ be a $(d+1)$-dimensional random vector with $X_i \sim F_i$, $1 \le i \le d$ and $Z \sim G$. Then, from Theorem 2.7 it follows that*

$$(X_1,\dots,X_d) \sim \ast_{\mathbf{B},G} D^i(F_1,\dots,F_d)\,,$$

for $D^i = C_{X_i,Z}$ and $\mathbf{B} = (B_t)_{t\in[0,1]}$ measurable such that $B_t^G = C_{X_1,\dots,X_d|Z=G^{-1}(t)}$ is the conditional copula of $(X_1,\dots,X_d)$ given $Z = G^{-1}(t)$.

(b) As a weakening of (10), *there exists for all $x = (x_1,\dots,x_d) \in \mathbb{R}^d$ a Lebesgue-null set N_x such that*

$$F_{1,\dots,d|F_{d+1}^{-1}(t)}(x) = B_t^{F_{d+1}}\left(\left(\partial_2^{F_{d+1}} C^i(F_i(x_i),t)\right)_{1\le i\le d}\right) \qquad \text{for all } t \in N_x^c\,,$$

compare Theorem 2.2.

As a consequence of Sklar's theorem 2.7 for factor models, the conditional independence product, the upper product, and the lower product is characterized by conditional independence, conditional comonotonicity, and conditional countermonotonicity, respectively.

Corollary 2.9. *For $1 \le i \le d$ and $F_i \in \mathcal{F}^1$, let $X_i \sim F_i$ be random variables on a non-atomic probability space. Then, for $G \in \mathcal{F}^1$ and $D^1,\dots,D^d \in \mathcal{C}_2$, the following statements hold true.*

(i) $(X_1,\dots,X_d) \sim \Pi_G D^i(F_1,\dots,F_d)$ if and only if there exists a random variable $Z \sim G$ such that $(X_1,\dots,X_d)|Z = z$ is independent for G-almost all z.
(ii) $(X_1,\dots,X_d) \sim \bigvee_G D^i(F_1,\dots,F_d)$ if and only if there exists a random variable $Z \sim G$ such that $(X_1,\dots,X_d)|Z = z$ is comonotonic for G-almost all z.
(iii) $(X_1,X_2) \sim D^1 \wedge_G D^2(F_1,F_2)$ if and only if there exists a random variable $Z \sim G$ such that $(X_1,X_2)|Z = z$ is countermonotonic for G-almost all z.

Throughout the following sections, the copula families $\mathbf{B}$ and $\mathbf{C}$ are assumed to be measurable.

2.3 Basic properties of $\ast$-products

For a d-copula C, denote by $\bar{C}$ the corresponding survival function and by $\widehat{C}$ its survival copula. Then, the survival function and the survival copula of the $\ast$-product are determined as follows.

Proposition 2.10 (Survival function and survival copula)**.**
The survival function $\overline{\ast_{\mathbf{B},G} D^i}$ *and the survival copula* $\widehat{\ast_{\mathbf{B},G} D^i}$ *of the* $\ast$*-product* $\ast_{\mathbf{B},G} D^i$ *are given by*

$$\overline{\ast_{\mathbf{B},G} D^i}(u) = \int_0^1 \widehat{B}_t^G \left(1 - \partial_2^G D^1(u_1, t), \dots, 1 - \partial_2^G D^d(u_i, t)\right) \mathrm{d}t\,, \tag{13}$$

$$\widehat{\ast_{\mathbf{B},G} D^i}(u) = \int_0^1 \widehat{B}_t^G \left(1 - \partial_2^G D^1(1 - u_1, t), \dots, 1 - \partial_2^G D^d(1 - u_i, t)\right) \mathrm{d}t\,,$$

for $u = (u_1, \dots, u_d) \in [0,1]^d$ *, where* $\widehat{B}_t^G$ *is the survival copula of* B_t^G *.*

Proof. Let $(U_1, \dots, U_d, Z)$ be a random vector such that U_i is uniformly distributed on $(0,1)$, $Z \sim G$, and

$$(U_1, \dots, U_d) | Z = G^{-1}(t) \sim B_t \left(\lim_{w_1 \searrow \cdot} \partial_2^G D^1(w_1, t), \dots, \lim_{w_d \searrow \cdot} \partial_2^G D^d(w_d, t)\right)$$

for almost all $t \in (0,1)$ and $C_{U_i,Z} = D^i$ for all $1 \le i \le d$, compare Remark 2.8(a). Then, it holds by (11) that $\ast_{\mathbf{B},G} D^i(u) = P(U_i \le u_i\,,\ 1 \le i \le d)$. Further, we obtain

$$\begin{aligned}
\overline{\ast_{\mathbf{B},G} D^i}\,(u) = P(U_i > u_i\ \forall i) &= \int_0^1 P\left(U_i > u_i\ \forall i \mid Z = G^{-1}(t)\right) \mathrm{d}t \\
&= \int_0^1 \widehat{B}_t^G \left(1 - \lim_{w_1 \searrow u_1} \partial_2^G D^1(w_1, t), \dots, 1 - \lim_{w_d \searrow u_d} \partial_2^G D^d(w_d, t)\right) \mathrm{d}t \\
&= \int_0^1 \widehat{B}_t^G \left(1 - \partial_2^G D^1(u_1, t), \dots, 1 - \partial_2^G D^d(u_i, t)\right) \mathrm{d}t\,,
\end{aligned}$$

where the third equality follows by the application of Sklar's Theorem for survival functions to the conditional survival function in the integrand, see, e.g., Georges et al. [12, Theorems 1 and 2] using that the i-th conditional marginal survival function is given by $\bar{F}_{U_i|Z=G^{-1}(t)}(u_i) = 1 - F_{U_i|Z=G^{-1}(t)}(u_i) = 1 - \lim_{w_i \searrow u_i} \partial_2^G D^i(w_i, t)$. The fourth equality is a consequence of Theorem 2.2.
The second statement follows from the relationship $\widehat{C}(u_1, \dots, u_d) = \bar{C}(1 - u_1, \dots, 1 - u_d)$, $(u_1, \dots, u_d) \in [0,1]^d$, between the survival copula $\widehat{C}$ and the survival function $\bar{C}$ of a copula $C \in \mathcal{C}_d$. □

For some particular specifications, the $\ast$-products simplify as follows.

Proposition 2.11 (Particular specifications)**.**
For all $u = (u_1, \dots, u_d) \in [0,1]^d$ *, the following statements hold true.*

(i) If $D^j = M^2$ *for all* $j \neq i$ *then* $\ast_{\mathbf{B}} D^k(u) = D^i(u_i, \min_{j \neq i}\{u_j\})$ *.*
(ii) If $D^j = W^2$ *for all* $j \neq i$ *then* $\ast_{\mathbf{B}} D^k(u) = D^i_*(u_i, \min_{j \neq i}\{u_j\})$ *, where* $D_*(v_1, v_2) := v_1 - D(v_1, 1 - v_2)$ *.*
(iii) If $D^i = \Pi^2$ *for all* i *, then* $\ast_{\mathbf{B},G} D^i(u) = \int_0^1 B_t^G(u)\,\mathrm{d}t$ *.*
(iv) Marginalization property: For $J \subset \{1, \dots, d\}$ *, the* J*-marginal of* $\ast_{\mathbf{B},G} D^i$ *is given by* $\ast_{\mathbf{B}',G} D^j$ *with bivariate copulas* $(D^j)_{j \in J}$ *and conditional copulas* $\mathbf{B}' = (B'_t)_t$ *where* $B'_t \in \mathcal{C}_{|J|}$ *is the* J*-marginal of* B_t *.*
(v) Identifiability property: If $D^j = M^2$ *, then the* (i,j)*-marginal of* $\ast_{\mathbf{B}} D^k$ *is given by* D^i *.*

Proof. To show statement (i), observe that $\partial_2 M^2(v,t) = \mathbb{1}_{\{t<v\}}$ for almost all t. This yields

$$\begin{aligned}\ast_{\mathbf{B}} D^k(u) &= \int_0^1 B_t\left(\mathbb{1}_{\{t<u_1\}},\dots,\mathbb{1}_{\{t<u_{i-1}\}}, \partial_2 D^i(u_i,t), \mathbb{1}_{\{t<u_{i+1}\}},\dots,\mathbb{1}_{\{t<u_d\}}\right)\,\mathrm{d}t \\ &= \int_0^{\min_{j\neq i}\{u_j\}} \partial_2 D^i(u_i,t)\,\mathrm{d}t = D^i(u_i, \min_{j\neq i}\{u_j\})\end{aligned}$$

for $u = (u_1,\dots,u_d) \in [0,1]^d$, where the second equality follows because all B_t have uniform univariate marginals and are grounded.
Statement (ii) follows similarly with $\partial_2 W^2(v,t) = \mathbb{1}_{\{t>1-v\}}$ for almost all t, and statement (iii) follows from $\partial_2^G \Pi^2(v,t) = v$.
(iv): For $u = (u_1,\dots,u_d)$ with $u_i = 1$ for $i \notin J$, it follows that $\ast_{\mathbf{B},G} D^i(u) = \ast_{\mathbf{B}',G} D^j(u_J)$, where $u_J = (u_j)_{j\in J}$.
Statement (v) is a consequence of (i) setting $u_k = 1$ for all $k \in \{1,\dots,d\}\setminus\{i,j\}$. □

Note that statements (i), (ii), and (v) in the above result are formulated w.r.t. continuous risk factor distribution functions and cannot be generalized to arbitrary $G \in \mathcal{F}^1$. A counterexample can be constructed from the following example.

Example 2.12. *Let $D^i = M^2$ for all i and $G = \mathbb{1}_{[0,\infty)}$ be the distribution function of the Dirac distribution in 0. Then, it holds that $\Pi_G D^i = \Pi^d \neq M^d$ using that $\iota_G(t) = 1$ and $\iota_G^-(t) = 0$ for all $t \in (0,1)$. In fact, for $Z \sim G$, it holds that $P(Z=0) = 1$, and, thus, the dependence specifications $C_{X_i,Z} = D^i = M^2$ do not yield any information on the X_i and cannot force comonotonicity of $(X_1,\dots,X_d)$.*

Next, we study the product $\ast_{\mathbf{B},G} D^i$ in the case where $D^i = M^2$ for all i. We make use of ordinal sums defined as follows.

Let $J \subset \mathbb{N}$ be a finite or countable subset of the natural numbers. Let $(a_k,b_k)_{k\in J}$ be a family of pairwise disjoint, open subinterval of $[0,1]$ and let $(C_k)_{k\in J}$ be a family of d-copulas. Then, the ordinal sum $(\langle a_k,b_k,C_k\rangle)_{k\in J}$ of $(C_k)_{k\in J}$ w.r.t. $(a_k,b_k)_{k\in J}$ is defined by

$$(\langle a_k,b_k,C_k\rangle)_{k\in J}(u) = \begin{cases} a_k + (b_k-a_k)C_k & \left(\frac{\min\{u_1,b_k\}-a_k}{b_k-a_k},\dots,\frac{\min\{u_d,b_k\}-a_k}{b_k-a_k}\right), \\ & \text{if } \min\{u_1,\dots,u_d\} \in (a_k,b_k) \text{ for some } k \in J \\ \min\{u_1,\dots,u_d\} & \text{else}, \end{cases}$$

where $u = (u_1,\dots,u_d) \in [0,1]^d$, see, e.g., Mesiar and Sempi [17].

The following proposition characterizes ordinal sums by $\ast$-products.

Proposition 2.13 (Ordinal sums)**.**
For $G \in \mathcal{F}^1$, for a measurable family $\mathbf{B} = (B_t)_{t\in[0,1]}$ and a sequence $(C_k)_{k\in J}$ of d-copulas, and for pairwise disjoint open subintervals $(a_k,b_k)_{k\in J}$ of $(0,1)$, the following statements are equivalent:

(i) $\ast_{\mathbf{B},G} M^2 = (\langle a_k,b_k,C_k\rangle)_{k\in J}$
(ii) $(a_k,b_k)_{k\in J} = \{(\iota_G^-(t),\iota_G(t)) \mid \iota_G^-(t) \neq \iota_G(t),\ t \in (0,1)\}$ *and* $C_k = B_t^G$ *for* $t \in (a_k,b_k) = (\iota_G^-(t),\iota_G(t))$.

Proof. For $u = (u_1,\dots,u_d) \in [0,1]^d$, let $v := \min\{u_i\}$. Then, we have that

$$\begin{aligned}\ast_{\mathbf{B},G} M^2(u) &= \int_0^1 B_t^G\left((\partial_2 \mathbb{1}_{\{t\le u_i\}})_{1\le i\le d}\right)\,\mathrm{d}t \\ &= \begin{cases} v & \text{if } \iota_G^-(v) = \iota_G(v) \\ \iota_G^-(v) + (\iota_G(v)-\iota_G^-(v))\,B_v^G\left(\left(\frac{\min\{u_i,\iota_G(v)\}-\iota_G^-(v)}{\iota_G(v)-\iota_G^-(v)}\right)_{1\le i\le d}\right) & \text{if } \iota_G^-(v) \neq \iota_G(v), \end{cases}\end{aligned}$$

which implies the assertion. Note that B_t^G is constant for $t \in (\iota_G^-(t), \iota_G(t))$. □

Denote by $\overline{A}$ the closure of a set $A \subset \mathbb{R}$. The following result justifies the simplified notation for the $*$-products where the argument G is omitted in the case that G is continuous, see Definition 2.4(iii). The proof is given in the Appendix.

Proposition 2.14.
Let $d \geq 2$. Then, the following statements are equivalent:

(i) $*_{\mathbf{B},G_1} D^i = *_{\mathbf{B},G_2} D^i$ *for all measurable families* $\mathbf{B} = (B_t)_{0\leq t\leq 1}$ *of d-copulas and for all* $D^i \in \mathcal{C}_2$, $1 \leq i \leq d$,
(ii) $\overline{\mathsf{Ran}(G_1)} = \overline{\mathsf{Ran}(G_2)}$.
(iii) $\iota_{G_1}(t) = \iota_{G_2}(t)$ *for Lebesgue-almost all* $t \in [0, 1]$.

As a consequence of the above result, the $*$-product depends only on the closure of the range of the risk factor distribution G. Thus, the copula of a completely specified factor model is invariant under strictly increasing transformations of the factor variable.

The following result shows in which relevant cases the $*$-product attains the upper Fréchet copula.

Proposition 2.15 (Maximality)**.**
For the $*$-product, the following statements hold true.

(i) If $D^i = M^2$ *for all* i, *then* $*_{\mathbf{B}} D^i = M^d$.
(ii) If $D^i = W^2$ *for all* i, *then* $*_{\mathbf{B}} D^i = M^d$.
(iii) $\bigvee_G D^i = M^d$ *if and only if* $D^j = D^k$ *on* $[0, 1] \times \mathsf{Ran}(G)$ *for all* $j \neq k$.

Proof. Statements (i) and (ii) follow from Proposition 2.11(i) and (ii).
Statement (iii) is an extension of [2, Proposition 2.4(v)] to arbitrary $G \in \mathcal{F}^1$. We give the proof in the Appendix. □

The definition of the $*$-product also yields an invariance property under Lebesgue-measure preserving transformations.

Let λ be the Lebesgue measure on $\mathcal{B}([0, 1])$. Denote by $\mathcal{T}$ the set of measurable transformations $T\colon ([0, 1], \mathcal{B}([0, 1]), \lambda) \to ([0, 1], \mathcal{B}([0, 1]), \lambda)$ that are *measure preserving*, i.e. $\lambda^T = \lambda$, where $\lambda^T(A) := \lambda(T^{-1}(A))$ for all $A \in \mathcal{B}([0, 1])$ denotes the distribution of T w.r.t. λ. Let $\mathcal{T}_P$ be the set of all $T \in \mathcal{T}$ such that T is bijective. Then, elements of $\mathcal{T}_P$ are called *shuffles*, see [9].

The following statement shows that simplified $*$-products are invariant under joint shuffles of the factor variable Z assuming a continuous distribution function.

Proposition 2.16 (Invariance under shuffles)**.** *For all $T \in \mathcal{T}_P$ and $C \in \mathcal{C}_2$, the function $\mathcal{S}_T(C)\colon [0, 1]^2 \to [0, 1]$ given through*

$$\mathcal{S}_T(C)(u, v) := \int_0^v \partial_2 C(u, T(t))\, \mathrm{d}t$$

is a bivariate copula. Furthermore, for simplified $*$-products with continuous factor distribution function and $B \in \mathcal{C}_d$ holds

$$*_B C^i = *_B \mathcal{S}_T(C^i).$$

The proof is given in the Appendix.

2.4 Continuity results for $*$-products

In this section, we derive continuity properties of the $*$-product w.r.t. to all its specifications.

For the approximation of $*$-products w.r.t. the factor distribution, we need the following lemma. The proof is given in the Appendix.

Lemma 2.17. *For $G_n, G \in \mathcal{F}^1$, $n \in \mathbb{N}$, the following statements hold true.*

(i) ι_G determines ι_G^- uniquely by $\iota_G^-(t) = \inf\{s \mid \iota_G(s) \geq t\}$.
(ii) If $\iota_{G_n} \to \iota_G$, then $\iota_{G_n}^- \to \iota_G^-$, where each convergence is almost surely pointwise.

In the following example, we consider some typical approximations of distribution functions for which the corresponding transformations ι converge almost surely pointwise.

Example 2.18. *(a) Denote by $\mathcal{F}_0^1$ the set of distribution functions with compact support. For $G \in \mathcal{F}_0^1$, consider the discretization $(G_n)_n$ given by $G_n(x) := \frac{\lceil nG(x) \rceil}{n}$. Then, G_n is a distribution function for all n with $\mathrm{Ran}(G_n) \subseteq \{0, \frac{1}{n}, \frac{2}{n}, \ldots, 1\}$. For $t \in (0,1)$ such that G^{-1} is continuous at t, it can be verified that $\iota_{G_n}(t) \to \iota_G(t)$. Thus, ι_{G_n} converges to ι_G almost surely pointwise.*
(b) For $G \in \mathcal{F}^1$, consider the discretization $(G_n)_n \in \mathcal{F}^1$ given by

$$G_n(x) := \begin{cases} \sup\{\frac{k}{n} \mid G(x) \geq \frac{k}{n}, k \in \mathbb{N}_0\}, & \text{if } G(x) < \frac{1}{2}, \\ \frac{1}{2} & \text{if } G(x) = \frac{1}{2}, \\ \inf\{\frac{k}{n} \mid G(x) \leq \frac{k}{n}, k \in \mathbb{N}_0\}, & \text{if } G(x) > \frac{1}{2}. \end{cases}$$

Similarly to the above example, it holds that $\mathrm{Ran}(G_n) \subseteq \{0, \frac{1}{n}, \frac{2}{n}, \ldots, 1\}$ and $\iota_{G_n} \to \iota_G$ almost surely pointwise.

The following two counterexamples show that, in general, neither convergence in distribution (denoted by $\xrightarrow{\mathcal{D}}$) implies almost surely pointwise convergence of the corresponding transformations ι nor that the converse holds true.

Example 2.19 ($G_n \xrightarrow{\mathcal{D}} G \not\Rightarrow \iota_{G_n} \to \iota_G$).
Let $G_n = F_{N(0,1/n)}$ be the distribution function of the normal distribution with mean 0 and variance $\frac{1}{n}$. Then, $G_n \xrightarrow{\mathcal{D}} G = \mathbb{1}_{[0,\infty)}$. But $\iota_{G_n} \not\to \iota_G$ almost surely pointwise because for all $t \in (0,1)$ it holds that $\iota_{G_n}(t) = t \neq \mathbb{1}_{(0,1)}(t) = \iota_G(t)$ for all $n \in \mathbb{N}$.

Example 2.20 ($\iota_{G_n} \to \iota_G \not\Rightarrow G_n \xrightarrow{\mathcal{D}} G$).
Let $G, G' \in \mathcal{F}_c^1$ be continuous distribution functions with $G \neq G'$. Let $(G_n)_n$ be the approximation of G given by Example 2.18(b). *Then, $\iota_{G_n} \to \iota_{G'}$ almost surely pointwise because $\iota_{G_n}(t) \to \iota_G(t) = t = \iota_{G'}(t)$ for all $t \in (0,1)$. But $G_n \overset{\mathcal{D}}{\not\to} G'$ because $G_n \xrightarrow{\mathcal{D}} G$ and $G' \neq G$.*

For a continuity result of the $*$-product $*_{\mathbf{B},G} D^i$ w.r.t. the bivariate dependence specifications D^i, we consider as slightly generalized version of the ∂-convergence for bivariate copulas in Mikusiński and Taylor [18].

Definition 2.21 (∂_2-convergence).
Let $D_n, D \in \mathcal{C}_2$ be bivariate copulas for all $n \in \mathbb{N}$. Then, the ∂_2-convergence $D_n \xrightarrow{\partial_2} D$ is defined by

$$\int_0^1 |\partial_2 D_n(x,t) - \partial_2 D(x,t)| \, \mathrm{d}t \xrightarrow{n\to\infty} 0 \quad \text{for all } x \in [0,1].$$

Remark 2.22. *a) Some typical approximations of copulas are the checkerboard, the checkmin and the Bernstein approximation, respectively. All these approximations are w.r.t. the ∂-convergence, see Mikusiński and Taylor [18], and, thus, also w.r.t. the ∂_2-convergence. In contrast, the ∂_2-convergence does not generally hold for the shuffle-of-min approximation, see Mikusiński and Taylor [18, Example 4].*

b) For a bivariate copula D, denote by D^T with $D^T(u,v) := D(v,u)$, $(u,v) \in [0,1]^2$, the transposed copula of D, and by K_D the associated Markov kernel defined by $K_D(t,[0,v]) := \lim_{u \searrow v} \partial_1 D(t,u)$ for all $u \in [0,1]$ and for Lebesgue-almost all $t \in [0,1]$. Then, the ∂_2-convergence is metrizable with a metric r given by $r(A,B) = \mathrm{D}_1(A^T, B^T)$, where D_1 denotes the metric defined by

$$\mathrm{D}_1(A,B) := \int_0^1 \int_0^1 \left| K_A(t,[0,v]) - K_B(t,[0,v]) \right| \,\mathrm{d}t\,\mathrm{d}v \tag{14}$$

for $A, B \in \mathcal{C}_2$, see [30]. Note that for all $v \in [0,1]$, there exists a Lebesgue-null set N_v such that $\partial_1 A(t,v) = K_A(t,[0,v])$ and $\partial_1 B(t,v) = K_B(t,[0,v])$ for all $t \in N_v^c$, see, e.g., Theorem 2.2(i).

As a main result, we give sufficient conditions for the continuity of the $*$-product w.r.t. all its arguments.

Theorem 2.23 (Continuity of $*$-products)**.**
Let $D_n^i, D^i \in \mathcal{C}_2$ be bivariate copulas, $1 \le i \le d$, $\mathbf{B^n} = (B_t^n)_{t\in[0,1]}$, $\mathbf{B} = (B_t)_{t\in[0,1]}$ be measurable families of d-copulas, and $G_n, G \in \mathcal{F}^1$ be distribution functions for all $n \in \mathbb{N}$. If

(i) $D_n^i \xrightarrow{\partial_2} D^i$ for all $1 \le i \le d$,
(ii) $B_t^n \xrightarrow{\mathcal{D}} B_t$ for Lebesgue-almost all $t \in [0,1]$, and
(iii) $\iota_{G_n}(t) \to \iota_G(t)$ for Lebesgue-almost all $t \in [0,1]$,

then it holds true that

$$*_{\mathbf{B^n},G_n} D_n^i \;\to\; *_{\mathbf{B},G} D^i \quad \textit{uniformly}\,.$$

Proof. We show for $u = (u_1,\ldots,u_d) \in [0,1]^d$ that

$$*_{\mathbf{B^m},G_k} D_n^i\,(u) \xrightarrow{n\to\infty} *_{\mathbf{B^m},G_k} D^i\,(u) \quad \text{for all } k, m \in \mathbb{N}\,, \tag{15}$$

$$*_{\mathbf{B^m},G_k} D_n^i\,(u) \xrightarrow{m\to\infty} *_{\mathbf{B},G_k} D_n^i\,(u) \quad \text{for all } k, n \in \mathbb{N}\,, \tag{16}$$

$$*_{\mathbf{B^m},G_k} D_n^i\,(u) \xrightarrow{k\to\infty} *_{\mathbf{B^m},G} D_n^i\,(u) \quad \text{for all } n, m \in \mathbb{N}\,. \tag{17}$$

Due to the equicontinuity of copulas, the above $*$-products converge uniformly using Arzelà-Ascoli's theorem. Thus, the statement follows from the exchangeability of applying the limits and, again, from Arzelà-Ascoli's theorem.

First, we show (17). Assume w.l.g. that $D_n^i = D^i$ and $\mathbf{B^m} = \mathbf{B}$ for all $n, m \in \mathbb{N}$. From Lemma 2.17(ii) we obtain that $\iota_{G_k} \to \iota_G$ a.s. implies that $\iota_{G_k}^-(t) \to \iota_G^-(t)$ for all $t \in N_0^c \cap [0,1]$ outside a Lebesgue-null set N_0. Fix $t \in N_0^c \cap [0,1]$.

If $\iota_G^-(t) < \iota_G(t)$, then there exists $R \in \mathbb{N}$ such that for all $k \ge R$ it holds that $\iota_{G_k}^-(t) < \iota_{G_k}(t)$ and, thus,

$$B_t^{G_k}(u) = \frac{1}{\iota_{G_k}(t) - \iota_{G_k}^-(t)} \int_{\iota_{G_k}^-(t)}^{\iota_{G_k}(t)} B_s(u)\,\mathrm{d}s \;\xrightarrow{k\to\infty}\; \frac{1}{\iota_G(t) - \iota_G^-(t)} \int_{\iota_G^-(t)}^{\iota_G(t)} B_s(u)\,\mathrm{d}s = B_t^G(u)$$

and

$$\partial_2^{G_k} D^i(u_i,t) = \frac{D^i(u_i,\iota_{G_k}(t)) - D^i(u_i,\iota_{G_k}^-(t))}{\iota_{G_k}(t) - \iota_{G_k}^-(t)} \;\xrightarrow{k\to\infty}\; \frac{D^i(u_i,\iota_G(t)) - D^i(u_i,\iota_G^-(t))}{\iota_G(t) - \iota_G^-(t)} = \partial_2^G D^i(u_i,t)$$

for $i = 1, \dots, d$.

If $\iota_G^-(t) = \iota_G(t)$ and $\iota_{G_k}^-(t) = \iota_{G_k}(t)$ for all k, then it follows that

$$B_t^{G_k}(u) = B_t(u) = B_t^G(u) \quad \text{and} \quad \partial_2^{G_k} D^i(u_i, t) = \partial_2 D^i(u_i, t) = \partial_2^G D^i(u_i, t) \quad \text{for all } k\,.$$

If $\iota_G^-(t) = \iota_G(t)$ and $\iota_{G_{k_l}}^-(t) < \iota_{G_{k_l}}(t)$ for a subsequence $(k_l)_{l\in\mathbb{N}}$, then it follows from Lebesgue's differential theorem, see, e.g., [4, Theorem 8.4.6], that

$$B_t^{G_{k_l}} = \frac{1}{\iota_{G_{k_l}}(t) - \iota_{G_{k_l}}^-(t)} \int\limits_{\iota_{G_{k_l}}^-(t)}^{\iota_{G_{k_l}}(t)} B_s(u)\,\mathrm{d}s \xrightarrow{l\to\infty} B_t(u) = B_t^G(u)$$

and, since the partial derivative of a copula exists almost surely, that

$$\partial_2^{G_{k_l}} D^i(u_i, t) = \frac{D^i(u_i, \iota_{G_{k_l}}(t)) - D^i(u_i, \iota_{G_{k_l}}^-(t))}{\iota_{G_{k_l}}(t) - \iota_{G_{k_l}}^-(t)} \xrightarrow{l\to\infty} \partial_2 D^i(u_i, t) = \partial_2^G D^i(u_i, t)$$

if $t \in N_1^c \cap [0,1]$ is outside of a Lebesgue-null set $N_1 \supseteq N_0$.

Altogether, this yields

$$B_s^{G_k}\left(\left(\partial_2^{G_k} D^i(u_i, s)\right)_{1\le i\le d}\right) \xrightarrow{k\to\infty} B_s^G\left(\left(\partial_2^G D^i(u_i, s)\right)_{1\le i\le d}\right)$$

for all $s \in N_1^c \cap [0,1]$ using that $B_s^{G_k} \in \mathcal{C}_d$ is equicontinuous for all s. This implies

$$*_{\mathbf{B},G_k} D^i\,(u) = \int\limits_0^1 B_s^{G_k}\left(\left(\partial_2^{G_k} D^i(u_i, s)\right)_{1\le i\le d}\right) \mathrm{d}s \longrightarrow \int\limits_0^1 B_s^G\left(\left(\partial_2^G D^i(u_i, s)\right)_{1\le i\le d}\right) \mathrm{d}s = *_{\mathbf{B},G} D^i\,(u)$$

as $k \to \infty$, where we apply the dominated convergence theorem.

To show (15), let $j \in \{1, \dots, d\}$ and choose w.l.g. $G_k = G$, $\mathbf{B^m} = \mathbf{B}$, and $D_n^i = D_n$ for all $k, m, n \in \mathbb{N}$ and $i \neq j$. Let $(G^l)_{l\in\mathbb{N}}$ be the discrete approximation of G given in Example 2.18(b). Then, the family $(B_t^{G^l})_t$ is constant in t on the intervals $(\frac{\kappa-1}{l}, \frac{\kappa}{l})$, $1 \le \kappa \le l$, and each $B_t^{G^l}$ is Lipschitz continuous with Lipschitz constant 1.

Thus, for the Lebesgue measure λ on $[0,1]$, it holds that

$$\begin{aligned}
&\lambda\left(\{t\colon |B_t^{G^l}((\partial_2^G D_n^i(u_i,t))_{1\le i\le d}) - B_t^{G^l}((\partial_2^{G^l} D_n^i(u_i,t))_{1\le i\le d})| > \varepsilon\} \cap (\tfrac{\kappa-1}{l}, \tfrac{\kappa}{l})\right)\\
&\le \lambda\left(\{t\colon |\partial_2^{G^l} D_n^j(u_j,t) - \partial_2^{G^l} D_n^j(u_j,t)| > \varepsilon\} \cap (\tfrac{\kappa-1}{l}, \tfrac{\kappa}{l})\right) \xrightarrow{n\to\infty} 0
\end{aligned}$$

for all $\varepsilon > 0$ and $1 \le \kappa \le l$, where the convergence follows from the assumption that $D_n^j \xrightarrow{\partial_2} D^j$. Then

$$\begin{aligned}
&\int\limits_0^1 |B_t^{G^l}((\partial_2^{G^l} D_n^i(u_i,t))_{1\le i\le d}) - B_t^{G^l}(\partial_2^{G^l}(D^i(u_i,t))_{1\le i\le d})|\,\mathrm{d}t\\
&= \sum_{\kappa=1}^{l} \int\limits_{\frac{\kappa-1}{l}}^{\frac{\kappa}{l}} |B_t^{G^l}((\partial_2^{G^l} D_n^i(u_i,t))_i) - B_t^{G^l}((\partial_2^{G^l} D^i(u_i,t))_{1\le i\le d})|\,\mathrm{d}t \xrightarrow{n\to\infty} 0\,,
\end{aligned}$$

which implies that $*_{\mathbf{B},G^l} D_n^i(u) \to *_{\mathbf{B},G^l} D^i(u)$ as $n \to \infty$ for all l. Thus, the statement follows from $*_{\mathbf{B},G^l} D_n^i \xrightarrow{l\to\infty} *_{\mathbf{B},G} D_n^i$ uniformly, see (17).

Statement (16) follows with the dominated convergence theorem. □

In the following remark, we note that a weak approximation of the bivariate dependence specifications or a weak approximation of the factor distribution does not guarantee the convergence of the corresponding $*$-products.

Remark 2.24. *(a) In general, the $*$-product $\ast_{\mathbf{B},G} D^i$ is not continuous in D^i w.r.t. weak convergence. A counterexample is given for the upper product and $G \in \mathcal{F}_c^1$ in [2, Example 2.7].*

(b) In general, the $*$-product is not continuous in the factor distribution w.r.t. weak convergence, i.e. $G_n \xrightarrow{\mathcal{D}} G \not\Rightarrow \ast_{\mathbf{B},G_n} D^i \xrightarrow{\mathcal{D}} \ast_{\mathbf{B},G} D^i$.

For a counterexample, let $(G_n)_n$ be the approximation of G given by Example 2.19. Then, $G_n \xrightarrow{\mathcal{D}} G = \mathbb{1}_{[0,\infty)}$. If the D^i do not coincide for all i, then the $*$-products do not necessarily converge because, e.g., for the upper products, it holds that

$$\bigvee_{G_n} D^i = \bigvee D^i \neq M^d = \bigvee_{G} D^i ,$$

where the first equality holds due to the continuity of G_n for all n, and the inequality is true because of the maximality property of the upper product, see Proposition 2.15(iii)*. The last equality follows from*

$$\min\{u_1, \dots, u_d\} = \int_0^1 \min\left\{\partial_2^G D^1(u_1, t), \dots, \partial_2^G D^d(u_d, t)\right\} \, \mathrm{d}t$$

since

$$\partial_2^G D^i(u_i, t) = \frac{D^i\left(u_i, \iota_G(t)\right) - D^i\left(u_i, \iota_G^-(t)\right)}{\iota_G(t) - \iota_G^-(t)} = u_i$$

for all $u_i \in [0,1]$ and $1 \le i \le d$ because $\iota_G(t) = 1$ and $\iota_G^-(t) = 0$ for all $t \in (0,1)$.

2.5 The lower product of bivariate copulas

In the following proposition, we provide basic properties for the lower product of bivariate copulas which are parallel to some results in [2] for the upper product.

For a bivariate copula $D \in \mathcal{C}_2$, define the reflected copulas D_* and D^* by

$$D_*(u,v) := v - D(1-u, v), \quad \text{and} \quad D^*(u,v) := u - D(u, 1-v), \tag{18}$$

respectively, for $(u,v) \in [0,1]^2$. Remember that the transposed copula D^T is defined by $D^T(u,v) := D(v,u)$, $(u,v) \in [0,1]^2$.

Proposition 2.25. *For $D, E \in \mathcal{C}_2$ and for a random vector (U_1, U_2, U_3) the following statements hold true:*

(i) Minimality property: $D \wedge_G E = W^2$ if and only if $D = E_*$ on $[0,1] \times \mathrm{Ran}(G)$.
(ii) $M^2 \wedge_G D \wedge_G E$ is a 3-copula if and only if $G \in \mathcal{F}_c^1$.
(iii) $(U_1, U_2, U_3) \sim M^2 \wedge D \wedge E \;\Leftrightarrow\; C_{U_1,U_2} = D^*$, $C_{U_1,U_3} = E^*$ and $(U_2, U_3)|U_1 = t$ is countermonotonic for almost all t.
(iv) $D \wedge_G M^2 = D$ on $[0,1] \times \mathrm{Ran}(G)$ and $M^2 \wedge_G D = D^T$ on $\mathrm{Ran}(G) \times [0,1]$.
(v) $D \wedge_G W^2 = D^*$ on $[0,1] \times \mathrm{Ran}(G)$ and $W^2 \wedge_G D = (D^*)^T$ on $\mathrm{Ran}(G) \times [0,1]$.
(vi) In general, the lower product is neither commutative nor associative.

The proof is given in the Appendix.

3 Ordering results for $*$-products

In this section, we establish lower and upper orthant ordering results for the $*$-product $\ast_{\mathbf{B},G} D^i$ w.r.t. the conditional copulas $\mathbf{B}$ and the bivariate specifications D^i. By the Sklar-representation theorem (Theorem

2.7) these results imply corresponding dependence ordering results for CSFM w.r.t. their specifications.

Definition 3.1 (Stochastic orderings)**.**
Let ξ, ξ' be d-dimensional random vectors. Then, define the

(i) lower orthant ordering $\xi \leq_{lo} \xi'$ *if for the corresponding distributions holds that* $F_\xi(x) \leq F_{\xi'}(x)$ *for all* $x \in \mathbb{R}^d$,
(ii) upper orthant ordering $\xi \leq_{uo} \xi'$ *if for the corresponding survival functions holds that* $\overline{F}_\xi(x) \leq \overline{F}_{\xi'}(x)$ *for all* $x \in \mathbb{R}^d$,
(iii) concordance ordering $\xi \leq_c \xi'$ *if it holds that* $\xi \leq_{lo} \xi'$ *and* $\xi \leq_{uo} \xi'$,

Note that all these orderings depend only on the distributions and, thus, are also defined for the corresponding distribution functions. A comparison w.r.t. the concordance ordering requires that the corresponding univariate marginal distributions are equal, i.e., $(\xi_1, \ldots, \xi_d) \leq_c (\xi_1, \ldots, \xi_d)$ implies $\xi_i \stackrel{\mathrm{d}}{=} \xi_i'$ for all i . Further, if $d = 2$ and $\xi_i \stackrel{\mathrm{d}}{=} \xi_i'$, the orderings $\leq_{lo}$, $\leq_{uo}$, and $\leq_c$ are equivalent and we denote them as the *standard bivariate dependence orderings*.

For an overview of stochastic orderings, see Müller and Stoyan [21], Shaked and Shanthikumar [28] and Rüschendorf [26].

In comparison to the ordering of $\ast_{\mathbf{B},G} D^i$ w.r.t. the specifications D^i , an ordering w.r.t. the copula family $\mathbf{B}$ is a simple task and given by the following proposition which extends Durante et al. [8, Proposition 3].

Proposition 3.2 (Ordering w.r.t. conditional copulas)**.**
Let $\mathbf{B} = (B_t)_{0 \leq t \leq 1}$, $\mathbf{C} = (C_t)_{0 \leq t \leq 1}$ *be measurable families of d-copulas. If* $B_t \prec C_t$ *for almost all* t , *where* $\prec$ *is one of the orders* $\leq_{lo}$, $\leq_{uo}$, *and* $\leq_c$, *respectively, then it holds true that*

$$\ast_{\mathbf{B},G} D^i \prec \ast_{\mathbf{C},G} D^i$$

for all $G \in \mathcal{F}^1$ *and for all copulas* $D^i \in \mathcal{C}_2$, $1 \leq i \leq d$.

Proof. The statement follows from the closure of these orders under mixtures (see Shaked and Shanthikumar [28, Theorems 6.G.3.(d)]). □

In the sequel, we are interested in ordering conditions for $\ast_{\mathbf{B},G} D^i$ w.r.t. the specifications D^i .

Intuitively, if the D^i increase in the standard bivariate dependence orderings, then the product $\ast_{\mathbf{B},G} D^i$ should increase due to the following reason: If all the D^i get closer to the upper Fréchet bound M^2 , then each $X_i = f_i(Z, \varepsilon_i)$ depends more strongly positively on Z . Thus, the copula $C_{X_1,\ldots,X_d}$ of $(X_1, \ldots, X_d)$ should be closer to the upper Fréchet bound M^d . But it turns out that ordering criteria on D^i are more complicated. One can also couple each X_i more strongly negatively with Z which also leads to a stronger positive dependence among the X_i . Further, as we see in Theorem 3.7, general ordering conditions for $\ast_{\mathbf{B},G} D^i$ in D^i for fixed D^j , $j \neq i$, restrict the choice of the conditional copula family $\mathbf{B}$.

Another difficulty is that, for fixed $i \in \{1, \ldots, d\}$, ordering results for $\ast_{j=1,\mathbf{B},G}^d D^j$ w.r.t. D^i always involve integral inequalities because

$$\ast_{j=1,\mathbf{B},G}^d D^j\,(u) = \int_0^1 B_t^G \left((\partial_2^G D^j(u_j, t))_{1 \leq j \leq d} \right) \,\mathrm{d}t$$

depends on D^i through the (generalized) partial derivative $\partial_2^G D^i$ of D^i . More precisely, a pointwise ordering of the integrands w.r.t. D^i and E^i , i.e., $B_t^G \left((\partial_2^G D^j(u_j, t))_{1 \leq j \leq d} \right) \leq B_t^G \left((\partial_2^G E^j(u_j, t))_{1 \leq j \leq d} \right)$ for all $(u_1, \ldots, u_d) \in$

$[0,1]^d$ and $t \in (0,1)$, is not possible: If we set $u_j = 1$ for all $j \neq i$, then

$$\partial_2^G D^i(u_i,t) = B_t^G\left((\partial_2^G D^j(u_j,t))_{1\le j\le d}\right) \le B_t^G\left((\partial_2^G E^j(u_j,t))_{1\le j\le d}\right) = \partial_2^G E^i(u_i,t))$$

for all t implies $D^i = E^i$ on $[0,1]\times \mathrm{Ran}(G)$ and, thus, $\ast_{j=1,\mathbf{B},G}^d D^j = \ast_{j=1,\mathbf{B},G}^d E^j$.

In the remaining part of this section, we derive several lower and upper orthant ordering results for $\ast_{j=1,\mathbf{B},G}^d D^j$ w.r.t. the D^i verifying integral inequalities based on the Schur-ordering, the sign-change ordering, and the lower orthant ordering, respectively.

3.1 Ordering results for componentwise convex conditional copulas

Denote by $\prec_S$ the Schur-ordering for functions, i.e., for integrable functions $f,g\colon [0,1]\to\mathbb{R}$, the relation $f \prec_S g$ is defined by $\int_0^x f^*(t)\,\mathrm{d}t \le \int_0^x g^*(t)\,\mathrm{d}t$ for all $x\in(0,1)$ and $\int_0^1 f(t)\,\mathrm{d}t = \int_0^1 g(t)\,\mathrm{d}t$. Here h^* denotes the decreasing rearrangement of an integrable function h, i.e., the (essentially w.r.t. the Lebesgue measure λ) uniquely determined decreasing function h^* such that $\lambda(h^*\le t) = \lambda(h\le t)$ for all $t\in\mathbb{R}$.

We say that a family $(\Phi_t)_{t\in[0,1]}$ of functions $\Phi_t\colon \Theta\to\mathbb{R}$ is continuous, $\Theta = \mathbb{R}^d$ or $\Theta = [0,1]^d$, if the mapping $(t,x)\mapsto \Phi_t(x)$ is continuous for all $(t,x)\in[0,1]\times\Theta$. As a basic integral inequality result, we make use of the following theorem on rearrangements from Fan and Lorentz [11, Theorem 1].

Theorem 3.3 (Ky Fan–Lorentz Theorem).
Let $\Phi_t\colon \mathbb{R}^d\to\mathbb{R}$, $t\in[0,1]$, be a family of continuous functions. Then, the following statements are equivalent:

(i) For all bounded and decreasing functions f_i, g_i on $[0,1]$ with $f_i \prec_S g_i$, it holds that

$$\int_0^1 \Phi_t(f_1,\dots,f_d)\,\mathrm{d}t \le \int_0^1 \Phi_t(g_1,\dots,g_d)\,\mathrm{d}t\,. \tag{19}$$

(ii) Φ with $\Phi(t,\cdot) := \Phi_t(\cdot)$ satisfies the following conditions for all $0\le t\le 1$, $0\le a\le 1-2\delta$, $\delta>0$, $u_k\ge 0$, $k=1,\dots,d$, $h\ge 0$ and $i\neq j$ where those arguments are omitted which are the same in each expression:

$$\Phi(u_i+h,u_j+h) - \Phi(u_i+h,u_j) - \Phi(u_i,u_j+h) + \Phi(u_i,u_j) \ge 0\,, \tag{20}$$

$$\Phi(u_i+h) - 2\Phi(u_i) + \Phi(u_i-h) \ge 0\,, \tag{21}$$

$$\int_0^\delta \left(\Phi_{a+\delta+s}(u_i) - \Phi_{a+\delta+s}(u_i+h) + \Phi_{a+s}(u_i+h) - \Phi_{a+s}(u_i)\right)\,\mathrm{d}s \ge 0\,. \tag{22}$$

For a function $f\colon\mathbb{R}^d\to\mathbb{R}$, let $\triangle_i^\varepsilon f(x) := f(x+\varepsilon e_i) - f(x)$ be the difference operator where $\varepsilon>0$ and where e_i denotes the i-th unit vector w.r.t. the canonical base in $\mathbb{R}^d$. Then, f is said to be *supermodular*, respectively, *directionally convex* if $\triangle_i^{\varepsilon_i}\triangle_j^{\varepsilon_j} f(x)\ge 0$ for all $x\in\mathbb{R}^d$, for all $\varepsilon_i,\varepsilon_j>0$, and for all $1\le i<j\le d$, respectively, $1\le i\le j\le d$. Note that in the literature, directionally convex functions are also called ultramodular or Wright convex.

Here, Condition (20) is supermodularity of Φ_t for all t, condition (21) is convexity of Φ_t in each component for all t. Functions that fulfill both conditions (20) and (21) are directionally convex. Motivated by Theorem 3.3, we consider the class $\mathcal{C}_d^{ccx}$ of componentwise convex d-copulas which is identical to the class of directionally convex copulas since copulas are supermodular.

Remark 3.4. *(a) As a consequence of the transformation formula, Theorem 3.3 also holds true if "decreasing" in* (i) *is substituted by "increasing" and the inequality in* (22) *is reversed, i.e.,*

$$\int_0^{\delta} \big(\Phi_{a+\delta+s}(u_i) - \Phi_{a+\delta+s}(u_i + h) + \Phi_{a+s}(u_i + h) - \Phi_{a+s}(u_i)\big)\, \mathrm{d}s \le 0\,. \tag{23}$$

for all $0 \le a \le 1 - 2\delta$, $\delta > 0$, $u_k \ge 0$, $k = 1, \dots, d$, $h \ge 0$.

(b) If Φ has continuous second partial derivatives w.r.t. all variables, then conditions (20), (21), (22), *and* (23), *respectively, are equivalent to*

$$\frac{\partial^2\Phi}{\partial u_i \partial u_j} \ge 0 \quad \forall i \ne j\,, \qquad \frac{\partial^2\Phi}{\partial u_i^2} \ge 0 \quad \forall i\,, \qquad \frac{\partial^2\Phi}{\partial t \partial u_i} \le 0 \quad \forall i\,, \qquad \text{and} \quad \frac{\partial^2\Phi}{\partial t \partial u_i} \ge 0 \quad \forall i\,,$$

respectively, see Lorentz [16].

In order to apply the Ky Fan–Lorentz Theorem to $*$-products, we consider an important class of bivariate copulas which are convex or concave in the second variable.

Definition 3.5 (CI/CIS/CDS copula)**.** *A bivariate copula $D \in \mathcal{C}_2$ is said to be* conditionally increasing in sequence *(CIS), respectively,* conditionally decreasing in sequence *(CDS) if $\partial_2 D(u, t)$ is decreasing, respectively, increasing in t for all $u \in [0, 1]$.*
Further, D is conditionally increasing *(CI) if D and D^T are CIS, i.e., D is concave in both components.*

In the literature, the CIS property is often defined by the partial derivative w.r.t. the first component. However, we define it in this way because the $*$-product depends on the derivatives of the bivariate copulas w.r.t. the second component.

Note that for a random vector $(U_1, U_2) \sim D$ with $D \in \mathcal{C}_2$ CIS, the conditional distribution $U_1 \mid U_2 = v$ is stochastically increasing in v. This explains the denomination of conditional increasingness.

For the next theorem, we need the following lemma. The proof is given in the Appendix.

Lemma 3.6. *For $G \in \mathcal{F}^1$, conditions* (21), (22), *and* (23), *respectively, transfer from measurable $\mathbf{B} = (B_t)_{t\in[0,1]}$, $B_t \in \mathcal{C}_d$, to the mixtures $\mathbf{B}^G = (B_t^G)_{t\in[0,1]}$.*

As a consequence of the Ky Fan–Lorentz Theorem 3.3, we obtain that general $\le_{lo}$-ordering results for $*_{\mathbf{B},G} D^i$ w.r.t. D^i require convexity of B_t in each component for all t.

Theorem 3.7 ($\le_{\mathbf{lo}}$-ordering of componentwise convex $*$-products)**.**
Assume that $\mathbf{B} = (B_t)_{t\in[0,1]}$ is a continuous family of d-copulas. Then, the following statements are equivalent:

(i) For all $G \in \mathcal{F}^1$ and for all CIS copulas $D^i, E^i \in \mathcal{C}_2$ with $D^i \le_{lo} E^i$, $1 \le i \le d$, it holds

$$*_{\mathbf{B},G} D^i \le_{lo} *_{\mathbf{B},G} E^i\,.$$

(ii) **B** *fulfills conditions* (21) *and* (22).

Proof. Assume (ii). Let $G \in \mathcal{F}^1$ and $D^i, E^i \in \mathcal{C}_2$ be CIS. For $(u_1, \dots, u_d) \in [0, 1]^d$, define $f_i(t) := \partial_2^G D^i(u_i, t)$ and $g_i(t) := \partial_2^G E^i(u_i, t)$, for almost all $t \in (0, 1)$. For $v \in (0, 1]$, we obtain from $D^i \le_{lo} E^i$ that

$$\begin{aligned}\int_0^v f_i(t)\,\mathrm{d}t &= D^i(u_i, \iota_G^-(v)) + (v - \iota_G^-(v))\, \partial_2^G D^i(u_i, v)\\ &\le E^i(u_i, \iota_G^-(v)) + (v - \iota_G^-(v))\, \partial_2^G E^i(u_i, v) = \int_0^v g_i(t)\,\mathrm{d}t\end{aligned}$$

with equality if $v = 1$. Since D^i and E^i are CIS, the functions f_i and g_i are decreasing; this yields $f_i \prec_S g_i$. Together with the boundedness of f_i and g_i it follows from the Ky Fan–Lorentz Theorem 3.3 that

$$\ast_{\mathbf{B},G} D^i(u) = \int_0^1 B_t^G\left(f_1(t),\dots,f_d(t)\right)\,\mathrm{d}t \le \int_0^1 B_t^G\left(g_1(t),\dots,g_d(t)\right)\,\mathrm{d}t = \ast_{\mathbf{B},G} E^i(u),$$

because $(B_t^G)_t$ fulfills conditions (21) and (22), see Lemma 3.6. This proves (i).

The reverse direction follows in the same way as in the proof of the Ky Fan–Lorentz Theorem 3.3 (see Fan and Lorentz [11, Theorem 1]) because for all decreasing functions $f_i, g_i\colon [0,1] \to [0,1]$ with $f_i \prec_S g_i$, there exist $(u_1,\dots,u_d) \in [0,1]^d$ and copulas $D^i, E^i \in \mathcal{C}_2$ with $D^i \le_{lo} E^i$ such that $f_i(t) = \partial_2 D^i(u_i,t)$ and $g_i(t) = \partial_2 E^i(u_i,t)$ holds. □

A similar result holds true w.r.t. the upper orthant ordering as follows.

Theorem 3.8 ($\le_{\mathbf{uo}}$-ordering of componentwise convex $\ast$-products)**.**
Assume that $\mathbf{B} = (B_t)_{t\in[0,1]}$ is a continuous family of d-copulas. Then, the following statements are equivalent:

(i) For all $G \in \mathcal{F}^1$ and for all CIS copulas $D^i, E^i \in \mathcal{C}_2$ with $D^i \le_{lo} E^i$, $1 \le i \le d$, holds

$$\ast_{\mathbf{B},G} D^i \le_{uo} \ast_{\mathbf{B},G} E^i .$$

(ii) The survival copulas $\widehat{\mathbf{B}} = (\widehat{B}_t)_{t\in(0,1)}$ fulfill conditions (21) *and* (23).

Proof. The proof is similar to the proof of Theorem 3.7 applying the Ky Fan–Lorentz Theorem to the survival functions of the $\ast$-products given by (13). Since $1 - \partial_2^G D^i(u_i,t)$ is increasing in t for all u_i and i, the survival copulas $\widehat{\mathbf{B}}$ have to fulfill condition (23), see Remark 3.4(a). □

Remark 3.9. *(a) If $\mathbf{B}$ and the associated survival copulas $\widehat{\mathbf{B}}$ fulfill the convexity condition* (21) *as well as condition* (22) *and* (23)*, respectively, then it holds $\ast_{\mathbf{B},G} D^i \le_c \ast_{\mathbf{B},G} E^i$ for all CIS copulas D^i, E^i with $D^i \le_{lo} E^i$, $1 \le i \le d$. For simplified $\ast$-products, condition* (22) *and* (23) *are trivially fulfilled. If $d = 2$, then B_t is componentwise convex if and only if $\widehat{B}_t$ is componentwise convex.*

(b) The componentwise convexity condition (21) *for each B_t implies negative lower orthant dependence for all the bivariate marginals of B_t. To see this, let $i \ne j$ and $u = (u_1,\dots,u_d)$ with $u_k = 1$ for all $k \ne i,j$. Then, it holds true that*

$$B_t(u) = \int_0^{u_i} \partial_i B_t(u_1,\dots,u_{i-1},s,u_{i+1},\dots,u_d)\,\mathrm{d}s \le \int_0^{u_i} u_j\,\mathrm{d}t = \Pi^d(u)$$

using the uniform marginal condition $\int_0^1 \partial_i B_t(u)\,\mathrm{d}u_i = u_j$ and that $\partial_i B_t(u)$ is increasing in u_i. For a discussion of componentwise convex copulas, see, e.g., Klement et al. [14] and Klement et al. [13].

If B_t is not componentwise convex for some t outside a null set, then general lower orthant ordering results for $\ast_{\mathbf{B}} D^k$ w.r.t. D^i depend on D^j, $j \ne i$.

For example, the conditional copula M^d corresponding to the upper product $\bigvee = \ast_{M^d}$ is component-wise concave (and not convex). Due to the maximality property of the upper product, general ordering conditions for $\bigvee D^k$ w.r.t. D^i depend on D^j, see Proposition 2.15(iii)*.*

(c) The ordering results for comonotonic random vectors in Rüschendorf [24, Corollary 3(b)] and for random vectors with common CI copula in Müller and Scarsini [19, Theorem 4.5], respectively, are based on the application of the Ky Fan–Lorentz Theorem 3.3 to (conditional) quantile functions. In contrast, Theorem 3.7 follows from the Ky Fan–Lorentz Theorem 3.3 comparing conditional distribution functions w.r.t. the conditioning variable.

We make use of another integral inequality due to Lorentz [16] as follows.

Theorem 3.10 (Lorentz)**.** *Let $\Phi\colon [0,1]\times\mathbb{R}^d \to \mathbb{R}$ be continuous. The following statements are equivalent:*

(i) For all positive bounded measurable functions f_k on $[0,1]$, $1 \le k \le d$, holds

$$\int_0^1 \Phi(t, f_1(t), \dots, f_d(t))\,\mathrm{d}t \le \int_0^1 \Phi(t, f_1^*(t), \dots, f_d^*(t))\,\mathrm{d}t\,.$$

(ii) Φ satisfies conditions (20) *and* (22)*.*

Note that the above result also holds true if we replace the decreasing rearrangements f_i^* by the increasing rearrangements f_{i*} of f_i and condition (22) by (23).

As a consequence of the Lorentz Theorem 3.10, we obtain for continuous factor distribution functions $G \in \mathcal{F}_c^1$ the following result concerning shuffles.

Proposition 3.11. *Let $D^1, \dots, D^d \in \mathcal{C}_2$ be CIS copulas.*

(i) If $\mathbf{B} = (B_t)_{t\in[0,1]}$ is a continuous family of d-copulas that fulfills condition (22)*, then it holds true that*

$$\ast_{\mathbf{B}} \mathcal{S}_{T_i}(D^i) \le_{lo} \ast_{\mathbf{B}} D^i$$

for all shuffles $T_i \in \mathcal{T}_P$ of D^i.

(ii) For $B \in \mathcal{C}_d$, the simplified products satisfy

$$\ast_B \mathcal{S}_{T_i}(D^i) \le_{lo} \ast_B \mathcal{S}_T(D^i) \tag{24}$$

for all shuffles $T_i, T \in \mathcal{T}_P$.

Proof. For $u = (u_1, \dots, u_d) \in [0,1]^d$, define $g_{i,u_i}(t) := \partial_2 \mathcal{S}_{T_i}(D^i)(u_i, t)$. Since D^i is conditionally increasing, the decreasing rearrangement is given by $g_{i,u_i}^*(t) = \partial_2 D^i(u_i, t)$ for almost all t. Hence, Theorem 3.10 implies

$$\ast_{\mathbf{B}} \mathcal{S}_{T_i}(D^i)(u) = \int_0^1 B_t\left((g_{i,u_i}(t))_{1\le i\le d}\right)\,\mathrm{d}t \le \int_0^1 B_t\left((g_{i,u_i}^*(t))_{1\le i\le d}\right)\,\mathrm{d}t = \ast_{\mathbf{B}} D^i(u)\,.$$

The second statement follows from the first one with Proposition 2.16. □

Remark 3.12. *(a) Note that the specifications on the right side of* (24) *are jointly shuffled.*

(b) A similar result to Proposition 3.11 holds true w.r.t. the upper orthant ordering. A generalization to arbitrary factor distribution functions $G \in \mathcal{F}^1$ is not possible because, in general, $g_{i,u_i} = \partial_2^G \mathcal{S}_{T_i}(D^i)(u_i,\cdot) \not\prec_S \partial_2^G D^i(u_i,\cdot)$ and, thus, $g_{i,u_i}^* \ne \partial_2^G D^i(u_i,\cdot)$, see also Example 3.15.

To apply Lorentz's Theorem 3.10 to the ordering of $\ast_{\mathbf{B},G} D^k$ w.r.t. D^i, we introduce and study the orderings $\le_{\partial_2 S,G}$ and $\le_{\partial_2 S}$ on the set $\mathcal{C}_2$ of bivariate copulas.

Definition 3.13 ($\le_{\partial_2 S}$, Schur order for copula derivatives)**.**
For $G \in \mathcal{F}^1$ and $D, E \in \mathcal{C}_2$, define the Schur order for the partial copula derivative *(w.r.t. the second variable) by*

$$D \le_{\partial_2 S,G} E \quad \text{if } \partial_2^G D(v,\cdot) \prec_S \partial_2^G E(v,\cdot) \quad \text{for all } v \in [0,1]\,.$$

For $G \in \mathcal{F}_c^1$, we abbreviate $\le_{\partial_2 S,G}$ by $\le_{\partial_2 S}$.

The least element in $\mathcal{C}_2$ w.r.t to the $\le_{\partial_2 S}$-order is given by the independence copula Π^2, i.e., it holds that $\Pi^2 \le_{\partial_2 S} C$ for all $C \in \mathcal{C}_2$. In contrast, a greatest element does not exist. However, M^2 and W^2 as well as every shuffle of these copulas are maximal elements.

Let $\zeta_1 \colon \mathcal{C}_2 \to [0, 1]$ be the dependence measure defined by

$$\zeta_1(A) := 3\,\mathrm{D}_1(A, \Pi^2)\,,$$

see [31]. By the following result, ζ_1 is increasing w.r.t. the $\le_{\partial_2 S}$-ordering, compare Figure 4.

Proposition 3.14. *Let D and E be bivariate copulas. Then $D \le_{\partial_2 S} E$ implies $\zeta_1(D^T) \le \zeta_1(E^T)$.*

Proof. By definition of the D_1-metric in (14) and by the transpose of a copula, we have that

$$\begin{aligned}
\zeta_1(D^T) = \mathrm{D}_1(D^T, (\Pi^2)^T) &= \int_0^1 \int_0^1 \left| K_{D^T}(t, [0, v]) - K_{(\Pi^2)^T}(t, [0, v]) \right| \,\mathrm{d}t\,\mathrm{d}v \\
&= \int_0^1 \int_0^1 \left| \partial_1 D^T(t, v) - \partial_1 (\Pi^2)^T(t, v) \right| \,\mathrm{d}t\,\mathrm{d}v = \int_0^1 \int_0^1 \left| \partial_2 D(v, t) - v \right| \,\mathrm{d}t\,\mathrm{d}v \\
&\le \int_0^1 \int_0^1 \left| \partial_2 E(v, t) - v \right| \,\mathrm{d}t\,\mathrm{d}v = \cdots = \zeta_1(E^T)\,,
\end{aligned}$$

where the inequality follows from the Hardy-Littlewood-Polya theorem which states that $f \le_S g$ is equivalent to $\int_0^1 \varphi(f(t))\,\mathrm{d}t \le \int_0^1 \varphi(g(t))\,\mathrm{d}t$ for all convex functions $\varphi \colon \mathbb{R} \to \mathbb{R}$ such that the expectations exist, see, e.g., [26, Theorem 3.21]. □

In general, $D \le_{\partial_2 S} E$ does not imply $D \le_{\partial_2 S, G} E$ even if E is a CIS copula, which is shown by the following counterexample.

Example 3.15 ($\le_{\partial_2 S} \not\Rightarrow \le_{\partial_2 S, G}$).
Let $D = (\langle a_k, b_k, C_k \rangle)_{k \in \{1,2,3\}}$ be the ordinal sum of $C_1 = C_2 = \Pi^2$ and $C_3 = M^2$ w.r.t. the intervals $(a_1, b_1) = (0, \frac{1}{4})$, $(a_2, b_2) = (\frac{1}{4}, \frac{1}{2})$, and $a_3, b_3) = (\frac{1}{2}, 1)$. Consider the symmetric copula D^* of D defined by (18). It can easily be seen that $D^* \le_{\partial_2 S} M^2$ and $D^* \ne_{\partial_2 S} M^2$. Let $G \in \mathcal{F}^1$ be given by $G(x) = \frac{1}{2}(1 + (x \wedge 1))\mathbb{1}_{\{x \ge 0\}}$. Then, it holds that $\mathrm{Ran}(G) = \{0\} \cup [\frac{1}{2}, 1]$, $\iota_G(t) = \frac{1}{2}\mathbb{1}_{(0,1/2]}(t) + t \cdot \mathbb{1}_{(1/2,1]}(t)$, and $\iota_G^-(t) = t \cdot \mathbb{1}_{(1/2,1]}(t)$.
For $u \le \frac{1}{4}$, it holds that

$$\begin{aligned}
\partial_2^G M^2(u, t) &= \begin{cases} \frac{\min\{u, \iota_G(t)\} - \min\{u, \iota_G^-(t)\}}{\iota_G(t) - \iota_G^-(t)} = \frac{u - 0}{\mathrm{fk}\frac{1}{2} - 0} = 2u\,, & \text{if } t \in (0, \frac{1}{2}]\,, \\ \lim_{s \nearrow t} \frac{\min\{u, t\} - \min\{u, s\}}{t - s} = 0\,, & \text{if } t \in (\frac{1}{2}, 1]\,, \end{cases} \\
&= 2u \cdot \mathbb{1}_{[0,1/2]}(t)\,, \quad \text{and} \\
\partial_2^G D^*(u, t) &= 4u \cdot \mathbb{1}_{(3/4,1]}(t)\,.
\end{aligned}$$

Hence, we obtain for $u \in (0, \frac{1}{4}]$ that $\partial_2^G D^*(u, \cdot) \succ_S \partial_2^G M^2(u, \cdot)$ and $\partial_2^G D^*(u, \cdot) \ne \partial_2^G M^2(u, \cdot)$. But this means that $D^* \not\le_{\partial_2 S, G} M^2$.

However, if both D and E are CIS (or CDS), then it can easily be verified that $D \le_{\partial_2 S} E$ yields $D \le_{\partial_2 S, G} E$.

A relation of the $\le_{\partial_2 S}$-ordering to the lower orthant ordering is given as follows. Note that we obtain from the definition of the reflected copula E^* of E in (18) that $E^* =_{\partial_2 S} E$, where, as usual, $=_{\partial_2 S}$ holds if $\le_{\partial_2 S}$ and $\ge_{\partial_2 S}$ is fulfilled.

Lemma 3.16. *For $D, E \in \mathcal{C}_2$, the following statements hold true.*

(i) If E is CIS, then $D \le_{\partial_2 S} E$ implies $E^* \le_{lo} D \le_{lo} E$.

(ii) If D and E are CIS, then $D \leq_{\partial_2 S} E$ and $D \leq_{lo} E$ are equivalent.

Proof. (i): Let $(u, v) \in [0, 1]^d$. For the decreasing rearrangement g_u^* of $\partial_2 D(u, \cdot)$, it follows that

$$D(u, v) = \int_0^v \partial_2 D(u, t)\,\mathrm{d}t \leq \int_0^v g_u^*(t)\,\mathrm{d}t \leq \int_0^v \partial_2 E(u, t)\,\mathrm{d}t = E(u, v)\,.$$

For the increasing rearrangement g_*^u of $\partial_2 D(u, t)$, it similarly holds that

$$E^*(u, v) = \int_0^v \partial_2 E^*(u, t)\,\mathrm{d}t = \int_0^v \partial_2 E(u, 1-t)\,\mathrm{d}t \leq \int_0^v g_*^u(t)\,\mathrm{d}t \leq \int_0^v \partial_2 D(u, t)\,\mathrm{d}t = D(u, v)\,.$$

(ii): If $D \leq_{lo} E$, then $D \leq_{\partial_2 S} E$ follows from

$$\int_0^v \partial_2 D(u, t)\,\mathrm{d}t = D(u, v) \leq E(u, v) = \int_0^v \partial_2 E(u, t)\,\mathrm{d}t$$

for all $u, v \in [0, 1]$, using that D and E are CIS. The reverse direction is given by (i). □

Consider the class

$$\mathcal{C}_2^E = \left\{ D \in \mathcal{C}_2 \mid D \leq_{\partial_2 S} E \right\}\,,$$

of bivariate copulas that are closer than E to the independence copula or equal to E w.r.t. the $\leq_{\partial_2 S}$-ordering. Due to the following result, the class $\mathcal{C}_2^E$ has a least and a greatest element w.r.t. the lower orthant ordering given by a CDS and a CIS copula.

Proposition 3.17. *There exist a unique CDS copula $E_\downarrow \in \mathcal{C}_2^E$ and a unique CIS copula $E_\uparrow \in \mathcal{C}_2^E$ such that*

$$E_\downarrow =_{\partial_2 S} E =_{\partial_2 S} E_\uparrow\,. \tag{25}$$

It holds that $E_\downarrow = E_\uparrow^*$, where $E_\uparrow^*$ is defined by (18), *and*

$$E_\downarrow \leq_{lo} D \leq_{lo} E_\uparrow \quad \text{for all } D \in \mathcal{C}_2^E\,. \tag{26}$$

Proof. To show (25), let $u \in [0, 1]$ and denote by $f_u\colon (0, 1) \to [0, 1]$ the essentially (w.r.t. the Lebesgue measure) unique decreasing rearrangement of $\partial_2 E(u, \cdot)$. For $(u, v) \in [0, 1]^2$, define

$$E_\uparrow(u, v) := \int_0^v f_u(t)\,\mathrm{d}t\,.$$

Then, $E_\uparrow$ is a bivariate copula, where the property of 2-increasingness follows for $(u_1, v_1) \leq (u_2, v_2)$ from

$$E_\uparrow(u_1, v_1) + E_\uparrow(u_2, v_2) - E_\uparrow(u_1, v_2) - E_\uparrow(u_2, v_1) = \int_{v_1}^{v_2} \underbrace{f_{u_2}(t) - f_{u_1}(t)}_{\geq 0}\,\mathrm{d}t \geq 0$$

because $\partial_2 E(u_2, t) \geq \partial_2 E(u_1, t)$ for all t.
Since $\partial_2 E_\uparrow(u, \cdot)$ is a rearrangement of $\partial_2 E(u, \cdot)$, it holds that $E =_{\partial_2 S} E_\uparrow$. Since $\partial_2 E_\uparrow(u, t) = f_u(t)$ for almost all t and f_u is the essentially uniquely determined decreasing rearrangement of $\partial_2 E(u, \cdot)$, it follows that $E_\uparrow$ is the uniquely determined CIS copula with $E =_{\partial_2 S} E_\uparrow$.

For the lower bound $E_\downarrow$, given by $E_\downarrow(u, v) := \int_{1-v}^1 f_u(t)\,\mathrm{d}t$, $(u, v) \in [0, 1]^2$, the statement follows similarly, so (25) is proved. Since $\int_0^1 f_u(t)\,\mathrm{d}t = u$ for all $u \in [0, 1]$, it follows that

$$E_\downarrow(u, v) = u - \int_0^{1-v} f_u(t)\,\mathrm{d}t = u - E_\uparrow(u, 1 - v) = E_\uparrow^*(u, v)$$

for all $(u, v) \in [0, 1]^2$. Statement (26) follows with Lemma 3.16 (i). □

In the following, we give some examples of $\le_{\partial_2 S}$-ordered copula families.

Example 3.18 (Elliptical copulas).
Let $(D^\rho)_{\rho\in[-1,1]}$ be a family of bivariate elliptical copulas with correlation parameter ρ . If $D^{|\rho_1|}$ and $D^{|\rho_2|}$ are CI, then

$$|\rho_1| \le |\rho_2| \quad\Rightarrow\quad D^{\rho_1} \le_{\partial_2 S} D^{\rho_2}\,. \tag{27}$$

A sufficient condition for $D^{|\rho_1|}$ to be CI is given by Abdous et al. [1, Proposition 1.2]. Then also $D^{|\rho_2|}$ is CI. Note that only in the Gaussian case, D^0 is CI, compare Rüschendorf [23, Theorem 2].

To show (27), let $0 \le \rho_1 \le \rho_2$. Since elliptical distributions are increasing w.r.t. the lower orthant ordering in the (generalized) correlation parameter, see Das Gupta et al. [7, Theorem 5.1], it follows that $D^{\rho_1} \le_{lo} D^{\rho_2}$. Then, Lemma 3.16(ii) *implies $D^{\rho_1} \le_{\partial_2 S} D^{\rho_2}$ using that D^{ρ_1} and D^{ρ_2} are CI. The general case follows from the symmetry $(D^\rho)^* = D^{-\rho}$ of elliptical copulas.*

Example 3.19 (Archimedean copulas).
Let C_ψ defined by $C_\psi(u,v) = \psi(\psi^{-1}(u)+\psi^{-1}(v))$ be the bivariate Archimedean copula with (strict) generator $\psi\colon \mathbb{R}_+ \to [0,1]$. The CI-property of C_ψ is characterized by the log-convexity of $-\psi'$, see Müller and Scarsini [20, Theorem 2.8]. Further, it holds that $C_{\psi_1} \le_{lo} C_{\psi_2}$ if and only if $\psi_1^{-1}\circ\psi_2$ is subadditive, see Nelsen [22, Theorem 4.4.2]. Thus, we obtain from Lemma 3.16(ii) *the following $\le_{\partial_2 S}$-criterion for Archimedean copulas:*
If $-\psi_i'$ is log-convex for $i = 1,2$, then it holds that $C_{\psi_1} \le_{\partial_2 S} C_{\psi_2}$ if and only if $\psi_1^{-1}\circ\psi_2$ is subadditive.

Sufficient conditions for the subadditivity are given in [22, Section 4.4]. We give some illustrating examples of $\le_{\partial_2 S}$-increasing Archimedean copula families. The log-convexity of $-\psi'$ can be verified straightforwardly.

(a) The Clayton family $(C_{\psi_\theta})_{\theta\in[-1,\infty)}$ with (inverse) generator $\varphi_\theta(t) = \psi_\theta^{-1}(t) = \frac{1}{\theta}\left(t^{-\theta}-1\right)$ for $\theta \ne 0$ and $\varphi_0(t) = \psi_0^{-1}(t) = -\ln(t)$ is $\le_{lo}$-increasing, see Nelsen [22, Example 4.14]. Since $-\psi_\theta'$ is log-convex for $\theta \ge 0$, it follows that $(C_{\psi_\theta})_{\theta\ge 0}$ is $\le_{\partial_2 S}$-increasing in θ .
(b) The Gumbel-Hougaard family $(C_{\psi_\theta})_{\theta\in[1,\infty)}$ with (inverse) generator $\varphi_\theta(t) = \psi_\theta^{-1}(t) = (-\ln(t))^\theta$ is $\le_{lo}$-increasing, see Nelsen [22, Example 4.12]. Since $-\psi_\theta'$ is log-convex for all θ , it follows that $(C_{\psi_\theta})_{\theta\ge 1}$ is $\le_{\partial_2 S}$-increasing in θ .
(c) The Frank family $(C_{\psi_\theta})_{\theta\in\mathbb{R}}$ with (inverse) generator $\varphi_\theta(t) = \psi_\theta^{-1}(t) = -\ln\left(\frac{\exp(-\theta t)-1}{\exp(-\theta)-1}\right)$ for $\theta \ne 0$ and $\varphi_0(t) = \psi_0^{-1}(t) = -\ln(t)$ is $\le_{lo}$-increasing, see Nelsen [22, p. 150]. Since $-\psi_\theta'$ is log-convex for $\theta \ge 0$ and $C_{\psi_\theta}^* = C_{\psi_{-\theta}}$, see Nelsen [22, p. 133], it follows that $|\theta| \le |\theta'|$ implies $C_{\psi_\theta} \le_{\partial_2 S} C_{\psi_{\theta'}}$.

Combining the Ky Fan–Lorentz Theorem 3.3 and Lorentz's Theorem 3.10, we get the following main result.

Theorem 3.20 ($\le_{\partial_2 S}$-ordering criterion).
Let $G \in \mathcal{F}^1$ and let $D^i, E^i \in \mathcal{C}_2$ be bivariate copulas with E^i CIS and $D^i \le_{\partial_2 S, G} E^i$ for all $1 \le i \le d$. Assume that $\mathbf{B} = (B_t)_{t\in[0,1]}$ is continuous, $B_t \in \mathcal{C}_d$ for all t .

(i) If $\mathbf{B}$ fulfills condition (22) *and if $B_t \in \mathcal{C}_d^{ccx}$ for all t , then*

$$\ast_{\mathbf{B},G} D^i \le_{lo} \ast_{\mathbf{B},G} E^i\,.$$

(ii) If $\widehat{\mathbf{B}} = (\widehat{B}_t)_{t\in[0,1]}$ fulfills condition (23) *and if $\widehat{B}_t \in \mathcal{C}_d^{ccx}$ for all t , then*

$$\ast_{\mathbf{B},G} D^i \le_{uo} \ast_{\mathbf{B},G} E^i\,.$$

Proof. To show (i), define for $u = (u_1,\dots,u_d) \in [0,1]^d$ the function $g_{i,u_i}(t) := \partial_2^G D^i(u_i,t)$ for almost all $t \in (0,1)$ and for $i = 1,\dots,d$. Since E^i is CIS, it holds that $\partial_2^G E^i(u_i,\cdot)$ is decreasing. From the assumption that $D^i \le_{\partial_2 S, G} E^i$, we obtain for the decreasing rearrangement g_{i,u_i}^* of g_{i,u_i} that $g_{i,u_i}^* \prec_S \partial_2^G E^i(u_i,\cdot)$. This yields the integral inequalities

$$\begin{aligned}
\ast_{\mathbf{B},G} D^i(u) &= \int_0^1 B_t^G\left(\partial_2^G D^1(u_1,t),\dots,\partial_2^G D^d(u_d,t)\right)\,\mathrm{d}t \\
&\le \int_0^1 B_t^G\left(g^*_{1,u_1}(t),\dots,g^*_{d,u_d}(t)\right)\,\mathrm{d}t \\
&\le \int_0^1 B_t^G\left(\partial_2^G E^1(u_1,t),\dots,\partial_2^G E^d(u_d,t)\right)\,\mathrm{d}t = \ast_{\mathbf{B},G} E^i(u)
\end{aligned}$$

where we apply Theorems 3.10 and 3.3 using that also the copulas $(B_t^G)_t$ are componentwise convex and fulfill condition (22), see Lemma 3.6.
Statement (ii) follows similarly to (i) applying formula (13) for the survival function of the $\ast$-product. □

Since the independence copula coincides with its survival copula and is componentwise convex, we obtain the following result as a consequence of Theorem 3.20.

Corollary 3.21 (Ordering the conditional independence product)**.**
If $G \in \mathcal{F}^1$ and $D^i, E^i \in \mathcal{C}_2$ such that E^i is CIS and $D^i \le_{\partial_2 S,G} E^i$ for all $1 \le i \le d$, then

$$\Pi_G D^i \le_c \Pi_G E^i\,.$$

Remark 3.22. *(a) For simplified $\ast$-products, condition* (22) *of Proposition 3.11 and Theorem 3.20 are trivially fulfilled. In Proposition 3.11 there is no convexity condition w.r.t.* **B** *and B, respectively. The statement in Theorem 3.20 also holds true if the E^i are conditionally decreasing, i.e. if $\partial_2 E^i(u_i,\cdot)$ is increasing for all u_i.*
(b) Corollary 3.21 extends [15, Proposition 1] to arbitrary dimension and general factor distribution $G \in \mathcal{F}^1$ where the authors show that $\Pi_{i=1}^2 E^i(u,v) \ge \Pi^2(u,v) = uv$ for $u,v \in [0,1]$ and CIS copulas $E^1, E^2 \in \mathcal{C}_2$.
(c) The intuition why Theorem 3.20 is true can be seen in the following explanation. The condition $D^i \le_{\partial_2 S} E^i$ indicates that D^i is closer to the independence copula than E^i. Since additionally E^i is CIS, E^i is closer to the upper Fréchet bound than D^i, and $(E^i)^*$ is closer to the lower Fréchet bound than D^i. This yields a stronger positive dependence among the X_i in the factor model with specifications E^i. However, in general, such a statement cannot be obtained if the conditional copulas are not componentwise convex.

The following counterexample shows that the assumption $D^i \le_{\partial_2 S,G} E^i$ in Theorem 3.20 cannot be simplified to $D^i \le_{\partial_2 S} E^i$, and that, in general, $\ast_{\mathbf{B}} D^i \le_{lo} \ast_{\mathbf{B}} E^i$ does not imply $\ast_{\mathbf{B},G} D^i \le_{lo} \ast_{\mathbf{B},G} E^i$ for $G \in \mathcal{F}^1 \setminus \mathcal{F}_c^1$.

Example 3.23. *For $d \ge 2$, let $D^i = D^*$ and $E^i = M^2$ with D^* and $G \in \mathcal{F}^1$ given by Example 3.15, $i = 1,\dots,d$. As shown there, it holds that $D^i \le_{\partial_2 S} E^i$ as well as $\partial_2^G D^i(u,\cdot) \succ_S \partial_2^G E^i(u,\cdot)$ for $u \le \frac{1}{4}$. This yields by Theorem 3.20 for all continuous $\mathbf{B} = (B_t)_{t\in[0,1]}$, $B_t \in \mathcal{C}_d^{ccx}$, satisfying condition* (22) *that $\ast_{\mathbf{B}} D^i \le_{lo} \ast_{\mathbf{B}} E^i$ and, as a consequence of Theorems 3.3 and 3.10, that $\ast_{\mathbf{B},G} D^i(u,\dots,u) \ge \ast_{\mathbf{B},G} E^i(u,\dots,u)$. So, for general $G \in \mathcal{F}^1 \setminus \mathcal{F}_c^1$ and for a general continuous family* **B** *of componentwise convex d-copulas which fulfills condition* (22)*, we have the following diagram:*

$$\begin{array}{ccc}
D^i \le_{\partial_2 S} E^i & \not\Rightarrow & D^i \le_{\partial_2 S,G} E^i \\
\Downarrow & \not\nearrow & \Downarrow \\
\ast_{\mathbf{B}} D^i \le_{lo} \ast_{\mathbf{B}} E^i & \not\Rightarrow & \ast_{\mathbf{B},G} D^i \le_{lo} \ast_{\mathbf{B},G} E^i\,.
\end{array}$$

3.2 Upper product ordering results

To derive ordering results for upper and lower products of bivariate copulas, consider on $\mathcal{C}_2$ the sign change ordering and the symmetric sign change ordering defined as follows.

For bivariate copulas $D, E \in \mathcal{C}_2$, define the function $f_{u,v}\colon [0,1] \to [-1,1]$ by

$$f_{u,v}(t) = \partial_2 E(u,t) - \partial_2 D(v,t)$$

for almost all $t \in (0,1)$ as the difference of the partial derivatives of E and D w.r.t. the second variable for fixed first components $u, v \in [0,1]$.

Definition 3.24 (Sign change orderings)**.**
The sign change ordering $D \leq_{\partial\Delta} E$ *, respectively, the* symmetric sign change ordering $D \leq_{s\partial\Delta} E$ *is defined via the property that for all* u, v *, respectively, for all* $u = v$ *, the function* $f_{u,v}$ *has no* $(-,+)$*-sign change.*

The sign change orderings strengthen the standard bivariate dependence orderings. It holds true that

$$D \leq_{\partial\Delta} E \quad \Rightarrow \quad D \leq_{s\partial\Delta} E \quad \Rightarrow \quad D \leq_c E \quad \Leftrightarrow \quad D \leq_{lo} E\,,$$

see [2, Proposition 3.4]. Note that the lower and upper Fréchet copula are the least and greatest element, respectively, w.r.t. the $\leq_{\partial\Delta}$-ordering, i.e., it holds that $W^2 \leq_{\partial\Delta} D \leq_{\partial\Delta} M^2$ for all $D \in \mathcal{C}_2$. Examples of $\leq_{\partial\Delta}$-ordered copula families are elliptical copulas and some families of Archimedean copulas, see [2].

Each of both conditions

$$D^j = E^j \leq_{\partial\Delta} D^d, E^d\,, \quad D^d \leq_{s\partial\Delta} E^d\,, \quad \forall\, 1 \leq j \leq d-1\,, \tag{28}$$

$$\text{and} \quad D^j = E^j \geq_{\partial\Delta} D^d, E^d\,, \quad D^d \geq_{s\partial\Delta} E^d\,, \quad \forall\, 1 \leq j \leq d-1\,, \tag{29}$$

implies $\bigvee D^i \geq_c \bigvee E^i$, see [2, Proposition 3.6]. We generalize this result to arbitrary factor distributions as follows.

Theorem 3.25 (Sign-change ordering criterion for upper products)**.**
Let $G \in \mathcal{F}^1$ *be a distribution function and let* $D^i, E^i \in \mathcal{C}_2$ *,* $1 \leq i \leq d$ *, be bivariate copulas. If either* (28) *or* (29) *holds, then it follows that*

$$\bigvee_G D^i \geq_c \bigvee_G E^i\,.$$

Proof. Assume (28). For $1 \leq i \leq d-1$ and $u_i, v \in [0,1]$, the functions $f_i, g_i, h\colon (0,1) \to [-1,1]$ given by

$$\begin{aligned} f_i(t) &= \partial_2 E^d(v,t) - \partial_2 D^i(u_i,t)\,, \\ g_i(t) &= \partial_2 D^d(v,t) - \partial_2 D^i(u_i,t)\,, \quad \text{and} \\ h(t) &= f_i(t) - g_i(t) = \partial_2 E^d(v,t) - \partial_2 D^d(v,t) \end{aligned}$$

have a.s. no $(-,+)$-sign change. Then, also the piecewise averaged functions f_i^G , g_i^G , $h^G\colon (0,1) \to [-1,1]$ given by

$$\begin{aligned} f_i^G(t) &= \partial_2^G E^d(v,t) - \partial_2^G D^i(u_i,t)\,, \\ g_i^G(t) &= \partial_2^G D^d(v,t) - \partial_2^G D^i(u_i,t)\,, \quad \text{and} \\ h^G(t) &= f_i^G(t) - g_i^G(t) = \partial_2^G E^d(v,t) - \partial_2^G D^d(v,t) \end{aligned}$$

have a.s. no $(-,+)$-sign change. Thus, the assertion follows in the same way as the proof of [2, Proposition 3.6].
Under the assumption of (29), the statement follows similarly with [2, Lemma 3.2], using that the functions f_i^G, g_i^G, and h_i^G have a.s. no $(+,-)$-sign change. □

Since we make use of it later on, we cite another concordance ordering criterion for upper products, based on the lower orthant ordering of the arguments.

Proposition 3.26 ($\le_{lo}$-ordering criterion for upper products)**.**
For $D^2, \dots, D^d, E \in \mathcal{C}_2$, the following statements are equivalent:

(i) $D^i \le_{lo} E$ *for all* $1 \le i \le d$,
(ii) $M^2 \vee D^2 \vee \dots \vee D^d \le_c M^2 \vee \underbrace{E \vee \dots \vee E}_{(d-1)\text{-times}}$.

The result of Proposition 3.26 is given by [3, Theorem 1] even for the tighter supermodular ordering.

3.3 Lower product ordering results

An ordering criterion similar to the sign change criterion for upper products in Theorem 3.25 holds true for lower products. Remember that, in general, the lower products $M^2 \wedge_G D \wedge_G E$ and $W^2 \wedge_G D \wedge_G E$ are 3-copulas only for continuous G. The symmetric copula D_* associated with $D \in \mathcal{C}_2$ is defined in (18).

Theorem 3.27 (Sign-change ordering criterion for lower products)**.**
For bivariate copulas $D^1, D^2, D^3 \in \mathcal{C}_2$ and $G \in \mathcal{F}^1$, the following statements hold true:

(i) *If $D^1_* \le_{\partial\Delta} D^2, D^3$ and $D^2 \le_{s\partial\Delta} D^3$, then*
$$M^2 \wedge D^1 \wedge D^2 \le_{lo} M^2 \wedge D^1 \wedge D^3 \quad \text{and} \quad D^1 \wedge_G D^2 \le_{lo} D^1 \wedge_G D^3 .$$
(ii) *If $D^1_\star \ge_{\partial\Delta} D^2, D^3$ and $D^2 \ge_{s\partial\Delta} D^3$, then*
$$W^2 \wedge D^1 \wedge D^2 \le_{lo} W^2 \wedge D^1 \wedge D^3 \quad \text{and} \quad D^1 \wedge_G D^2 \le_{lo} D^1 \wedge_G D^3 .$$

Proof. To show the lower orthant ordering in (i), let $u = (u_1, u_2, u_3) \in [0,1]^3$. In the case that $G \in \mathcal{F}^1 \setminus \mathcal{F}^1_c$ is discontinuous, set $u_1 = 1$. Consider the functions $f, g, h\colon [0,1] \to [-1,1]$ defined by

$$\begin{aligned}
f(t) &:= \partial_2^G D^2(1-u_3, t) - \partial_2^G D^1_*(u_2, t)\,,\\
g(t) &:= \partial_2^G D^3(1-u_3, t) - \partial_2^G D^1_*(u_2, t)\,,\\
h(t) &:= \partial_2^G D^3(1-u_3, t) - \partial_2^G D^2(1-u_3, t) = g(t) - f(t)\,.
\end{aligned}$$

Then f, g, h have no $(-,+)$-sign change and it holds that $\int f(t)\,\mathrm{d}t = \int g(t)\,\mathrm{d}t$. This yields the integral inequality
$$\int_0^{u_1} \max\{f(t), 0\}\,\mathrm{d}t \le \int_0^{u_1} \max\{g(t), 0\}\,\mathrm{d}t\,,$$
compare [2, Lemma 3.2]. Thus, we obtain

$$\begin{aligned}
M^2 \wedge_G D^1 \wedge_G D^2(u) &= \int_0^{u_1} \max\left\{\partial_2^G D^1(u_2, t) + \partial_2^G D^2(1-u_3, t) - 1\,, 0\right\}\,\mathrm{d}t\\
&= \int_0^{u_1} \max\left\{\partial_2^G D^2(1-u_3, t) - \partial_2^G D^1_*(u_2, t)\,, 0\right\}\,\mathrm{d}t\\
&= \int_0^{u_1} \max\left\{f(t)\,, 0\right\}\,\mathrm{d}t\\
&\le \int_0^{u_1} \max\left\{g(t)\,, 0\right\}\,\mathrm{d}t = \dots = M^2 \wedge_G D^1 \wedge_G D^3(u)\,,
\end{aligned}$$

where the first equality follows from $\partial_2^G M^2(u_1, t) = \mathbb{1}_{\{u_1>t\}}$ for almost all t and for arbitrary $u_1 \in [0,1]$ in the case that G is continuous, respectively, for $u_1 = 1$ if G is discontinuous. This yields $M^2 \wedge D^1 \wedge D^2 \leq_{lo} M^2 \wedge D^1 \wedge D^3$ in the continuous case and $D^1 \wedge_G D^2 \leq_{lo} D^1 \wedge_G D^3$ for arbitrary G.
For the upper orthant ordering in (i), we obtain analogously that

$$\begin{aligned}\overline{M^2 \wedge_G D^1 \wedge_G D^2}(u) &= \int_{u_1}^{1} \max\left\{1 - \partial_2^G D^1(u_2, t) - \partial_2^G D^2(1-u_3, t), 0\right\} \mathrm{d}t \\ &\leq \int_{u_1}^{1} \max\left\{1 - \partial_2^G D^1(u_2, t) - \partial_2^G D^3(1-u_3, t), 0\right\} \mathrm{d}t = \overline{M^2 \wedge_G D^1 \wedge_G D^3}(u).\end{aligned}$$

Statement (ii) follows analogously. □

Similarly to the $\leq_{lo}$-ordering criterion for the concordance ordering of upper products in Proposition 3.26, we obtain a concordance-ordering result for lower products based on a $\leq_{lo}$-ordering criterion for the bivariate dependence specifications.

Theorem 3.28 ($\leq_{lo}$-ordering criterion for lower products)**.**
Let $D, E^1, E^2 \in \mathcal{C}_2$ be bivariate copulas. Then, the following statements are equivalent:

(i) $D \leq_{lo} E^1$ *and* $D_* \leq_{lo} E^2$,
(ii) $M^2 \wedge D \wedge D_* \leq_c M^2 \wedge E^1 \wedge E^2$.

Proof. Assume (i). To show the lower orthant ordering, let $u = (u_1, u_2, u_3) \in [0,1]^3$. Then, it holds that

$$\begin{aligned}M^2 \wedge D \wedge D_*(u) &= \int_0^{u_1} \max\{\partial_2 D(u_2, t) + \partial_2 D_*(u_3, t) - 1, 0\} \mathrm{d}t = \int_0^{u_1} \max\{\partial_2 D(u_2, t) - \partial_2 D(1-u_3, t), 0\} \mathrm{d}t \\ &= \int_0^{u_1} \max\{\partial_2 D(u_2, t), \partial_2 D(1-u_3, t)\} \mathrm{d}t - D(1-u_3, u_1) \\ &= \max\{D(u_2, u_1), D(1-u_3, u_1)\} - D(1-u_3, u_1) \\ &\leq \max\left\{E^1(u_2, u_1), E^2_*(1-u_3, u_1)\right\} - E^2_*(1-u_3, u_1) \\ &\leq \int_0^{u_1} \max\left\{\partial_2 E^1(u_2, t), \partial_2 E^2_*(1-u_3, t)\right\} \mathrm{d}t - E^2_*(1-u_3, u_1) = M^2 \wedge E^1 \wedge E^2(u),\end{aligned}$$

where the first inequality follows from the assumptions using that $D_* \leq_{lo} E^2$ if and only if $D \geq_{lo} E^2_*$. The second inequality holds due to Jensen's inequality.
For the upper orthant ordering, we similarly obtain

$$\begin{aligned}\overline{M^2 \wedge D \wedge D_*}(u) &= \int_{u_1}^{1} \max\{\partial_2 D(u_2, t) - \partial_2 D_*(u_3, t), 0\} \mathrm{d}t \\ &\leq \int_{u_1}^{1} \max\left\{\partial_2 E^1(u_2, t) - \partial_2 E^2(u_3, t), 0\right\} \mathrm{d}t = \overline{M^2 \wedge E^1 \wedge E^2}(u).\end{aligned}$$

Assume (ii). Then, (i) follows from the closure of the lower orthant ordering under marginalization and from the marginalization property of $*$-products, see Proposition 2.11(iv). □

3.4 Ordering results for convex combinations

In Section 3.1, we have established that general lower orthant ordering results for $\ast_{\mathbf{B},G} D^i$ in D^i for fixed D^j, $i \neq j$, are only possible if the conditional copulas $\mathbf{B} = (B_t)_t$ fulfill the convexity condition (21). Remember that this convexity condition implies negative dependence of the bivariate marginals of B_t.

Motivated by Theorem 3.20 for componentwise convex conditional copulas and by Proposition 3.26 concerning a $\leq_{lo}$-ordering criterion for the upper product, the question arises for which $\ast$-products ordering results of the form

$$D^i \prec_{\partial_2 S} E\,,\ \forall i\,, E \text{ CIS} \quad \Rightarrow \quad \ast_{\mathbf{B}} D^i \leq_{lo} \ast_{\mathbf{B}} E \tag{30}$$

hold true. Note that E is assumed to be a joint upper bound for the D^i.

To partly answer this question, we generalize the necessary integral ordering condition in the Ky Fan–Lorentz Theorem 3.3 under an additional ordering assumption on the upper bound.

Proposition 3.29. *If for all decreasing and bounded functions f_i, g_i on $[0,1]$ with $g_{i_1} \leq \ldots \leq g_{i_d}$, $i_1, \ldots, i_d \in \{1, \ldots, d\}$, such that $f_i \prec_S g_i$ the integral inequality* (19) *holds true, then Φ fulfills the milder convexity condition*

$$\Phi(u_i + h) - 2\phi(u_i) + \Phi(u_i - h) \geq 0 \quad \textit{for all } i\,,\ u_i\,,\ u_j \neq u_i\,,\ h \leq \min_{j \neq i} |u_i - u_j| \tag{31}$$

where those components are omitted which are the same in each expression.

Proof. We modify the proof of Fan and Lorentz [11, Theorem 1]: Let $0 \leq a < a + 2\delta \leq 1$, and $u_1, \ldots, u_d \in [0,1]$. For some fixed i, let $h \in [0, \min_{j \neq i} |u_j - u_i|]$. Assume w.l.g. that $u_j \neq u_i$ for all $j \neq i$. Define

$$f_i(t) = \begin{cases} u_i + h\,, & 0 \leq t \leq a\,, \\ u_i\,, & a < t \leq a + 2\delta\,, \\ u_i - h\,, & a + 2\delta < t \leq 1\,, \end{cases} \qquad g_i(t) = \begin{cases} u_i + h\,, & 0 \leq t \leq a + \delta\,, \\ u_i - h\,, & a + \delta < t \leq 1\,, \end{cases}$$
$$f_j(t) = g_j(t) = u_j\,, \quad j \neq i\,.$$

Then, it holds true that $f_k \prec_S g_k$, $k = 1, \ldots, d$. Further, there exist $i_1, \ldots, i_d \in \{1, \ldots, d\}$ pairwise different such that $g_{i_1}(t) \leq \ldots \leq g_{i_d}(t)$ for all t. The inequality (19) reduces in this case to

$$\int_0^\delta \left\{\Phi(a + t, u_i + h) - \Phi(a + t, u_i) + \Phi(a + \delta + t, u_i - h_i) - \Phi(a + \delta + t, u_i)\right\} \mathrm{d}t \geq 0\,.$$

Dividing by δ and making $\delta \to 0$, yields (31). □

As a consequence, we obtain that lower orthant ordering results for $\ast$-products with a joint upper bound for all copulas also restrict the choice of conditional copulas.

Corollary 3.30. *If for all CIS copulas $D^i, E \in \mathcal{C}_2$ with $D^i \prec_{\partial_2 S} E$ the inequality*

$$\ast_{\mathbf{B}} D^i \leq_{lo} \ast_{\mathbf{B}} E \tag{32}$$

holds true, then $\mathbf{B}$ *fulfills the milder convexity condition* (31).

Proof. Let f_i, g_i be decreasing and bounded such that $f_i \prec_S g_i$ and $g_{i_1} \leq \ldots \leq g_{i_d}$, $i_1, \ldots, i_d \in \{1, \ldots, d\}$. Assume w.l.g. that $0 \leq f_i, g_i \leq 1$. Then, there exist $u_1, \ldots, u_d \in [0,1]$ and CIS copulas $D^i, E \in \mathcal{C}_2$ with $D^i \prec_{\partial_2 S} E$ such that $f_i(t) = \partial_2 D^i(u_i, t)$ and $g_i(t) = \partial_2 E(u_i, t)$. Thus, the statement follows from Proposition 3.29. □

Remark 3.31. *(a) Due to Corollary 3.30, ordering results of the form* (30) *can not be obtained for all continuous families* $\mathbf{B} = (B_t)_{t\in[0,1]}$ *of d-copulas.*

(b) The upper Fréchet copula M^d *fulfills the milder convexity condition* (31). *In this case, inequality* (32) *is trivially fulfilled because* $\bigvee E^i = M^d$ *whenever* $E^i = E$ *for all* i *. Note that for the upper product the non-trivial generalized inequality*

$$M^2 \vee D^2 \vee \cdots \vee D^d \leq_c M^2 \vee \underbrace{E \vee \cdots \vee E}_{(d-1)\text{-times}} \tag{33}$$

holds true whenever $D^i \leq_{lo} E$ *(see Proposition 3.26).*

Denote by $co(M^d, \mathcal{C}_d^{ccx})$ the set of convex combinations of M^d with elements of $\mathcal{C}_d^{ccx}$. Then, we obtain the following result.

Theorem 3.32. *Let* $D^1 = E^1 = M^2$ *and* $D^i \in \mathcal{C}_2$ *such that for a CIS copula* $E \in \mathcal{C}_2$ *holds* $D^i \prec_{\partial_2 S} E = E^i$ *,* $2 \leq i \leq d$ *. Let* $B \in co(M^d, \mathcal{C}_d^{ccx})$ *. Then, for the simplified* $*$*-products, it holds true that*

$$\ast_B D^i \leq_{lo} \ast_B E^i .$$

Proof. The copula B is of the form $B = aM^d + (1-a)C$, for some $a \in [0,1]$, where $C \in \mathcal{C}_d^{ccx}$ fulfills the convexity condition (21). Thus, the statement follows from Theorem 3.20 and from (33) using that $D^i \prec_{\partial_2 S} E$ implies $D^i \leq_{lo} E$, see Lemma 3.16. □

Note that in the above result, $E^i = E$ for $i \in \{2,\dots,d\}$ is a joint upper bound for the copulas $D^2,\dots,D^d$.

4 Ordering results for completely specified factor models

In this section, we combine the ordering results on $*$-products in Section 3 with the ordering of the univariate marginal distributions. This leads to lower and upper orthant as well as concordance ordering results for CSFMs and, thus, to bounds w.r.t. these orderings in classes of CSFMs and PSFMs, respectively.

Suppose that $X = (X_1,\dots,X_d)$ with $X_i = f_i(Z,\varepsilon_i)$ and $Y = (Y_1,\dots,Y_d)$ with $Y_i = g_i(Z',\varepsilon_i')$ are d-dimensional random vectors that follow a completely specified factor model with factor distribution function $G = F_Z$ and $G' = F_{Z'}$, respectively, such that $\overline{\mathrm{Ran}(G)} = \overline{\mathrm{Ran}(G')}$. Then the corresponding copulas are given by the $*$-products

$$C_X = \ast_{\mathbf{B},G} D^i \quad \text{and} \quad C_Y = \ast_{\mathbf{C},G'} E^i ,$$

respectively, where $D^i = C_{X_i,Z}$, $E^i = C_{Y_i,Z'}$, $B_t^G = C_{X|Z=G^{-1}(t)}$, and $C_t^{G'} = C_{Y|Z'=G'^{-1}(t)}$, see Theorem 2.7. Further, by Sklar's Theorem, the corresponding distribution functions are given by

$$F_X = \ast_{\mathbf{B},G} D^i\,(F_{X_1},\dots,F_{X_d}) \quad \text{and} \quad F_Y = \ast_{\mathbf{C},G} E^i\,(F_{Y_1},\dots,F_{Y_d}) ,$$

using that $\overline{\mathrm{Ran}(G)} = \overline{\mathrm{Ran}(G')}$, see Proposition 2.14.

We establish conditions on the conditional copula families **B** and **C** assumed generally to be measurable, on the dependence specifications D^i and E^i , and on the distributions of the components X_i and Y_i to infer lower orthant, upper orthant and concordance comparison results for X and Y .

The following proposition compares CSFMs where the bivariate dependence specifications D^i and E^i coincide.

Proposition 4.1 (Ordering conditional copulas).
Assume that $D^i = E^i$ *for all* i *. Then, the following statements hold true.*

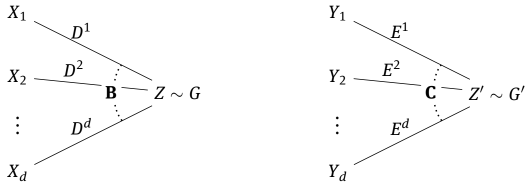

Figure 3 The setting in Section 4: completely specified factor models with dependence specifications D^i and $\mathbf{B} = (B_t)_t$ as well as E^i and $\mathbf{C} = (C_t)_t$, $1 \le i \le d$, and with factor distribution function G and G' , respectively, such that $\overline{\mathrm{Ran}(G)} = \overline{\mathrm{Ran}(G')}$.

(i) *If* $\mathbf{B} \le_{lo} \mathbf{C}$ *and* $X_i \le_{lo} Y_i$ *then* $X \le_{lo} Y$.
(ii) *If* $\mathbf{B} \le_{uo} \mathbf{C}$ *and* $X_i \le_{uo} Y_i$ *then* $X \le_{uo} Y$.
(iii) *If* $\mathbf{B} \le_{c} \mathbf{C}$ *and* $X_i \stackrel{\mathrm{d}}{=} Y_i$ *then* $X \le_{c} Y$.

Proof. The statements follow from Proposition 3.2 for fixed marginal distributions together with Sklar's Theorem (respectively, Sklar's Theorem for survival functions) for fixed conditional copulas using that $X_i \le_{lo} Y_i$ (respectively, $X_i \le_{uo} Y_i$) implies $F_{X_i}(x) \le F_{Y_i}(x)$ (respectively, $F_{X_i}(x) \ge F_{Y_i}(x)$) for all $x \in \mathbb{R}$ and $1 \le i \le d$. □

In the remaining part of this section, we also establish ordering conditions w.r.t. the dependence specifications D^i and E^i .

For the following theorem, we need a family of componentwise convex conditional (survival) copulas that lies between $\mathbf{B}$ and $\mathbf{C}$. Then, we obtain a general ordering condition in dependence on the bivariate specifications, the conditional copulas and the marginal distributions.

Theorem 4.2. *Let* $\mathbf{B}' = (B'_t)_{t\in[0,1]} \subset \mathcal{C}_d$ *be continuous. Assume that all* E^i *are CIS.*

(i) *If* $\mathbf{B}'$ *satisfies condition* (22) *and if* $B'_t \in \mathcal{C}_d^{ccx}$ *for all* t *, then*

$$\mathbf{B} \le_{lo} \mathbf{B}' \le_{lo} \mathbf{C}\,,\ D^i \le_{\partial_2 S,G} E^i\,,\ \textit{and } X_i \le_{lo} Y_i \textit{ for all } i \ \Rightarrow\ X \le_{lo} Y\,.$$

(ii) *If* $\widehat{\mathbf{B}}'$ *satisfies condition* (23) *and if* $\widehat{B}_t \in \mathcal{C}_d^{ccx}$ *for all* t *, then*

$$\mathbf{B} \le_{uo} \mathbf{B}' \le_{uo} \mathbf{C}\,,\ D^i \le_{\partial_2 S,G} E^i\,,\ \textit{and } X_i \le_{uo} Y_i \textit{ for all } i \ \Rightarrow\ X \le_{uo} Y\,.$$

(iii) *If* $\mathbf{B}'$ *and* $\widehat{\mathbf{B}}'$ *satisfy condition* (22) *and* (23)*, respectively, and if* $B_t, \widehat{B}_t \in \mathcal{C}_d^{ccx}$ *for all* t *, then*

$$\mathbf{B} \le_{c} \mathbf{B}' \le_{c} \mathbf{C}\,,\ D^i \le_{\partial_2 S,G} E^i\,,\ \textit{and } X_i \stackrel{\mathrm{d}}{=} Y_i \textit{ for all } i \ \Rightarrow\ X \le_{c} Y\,.$$

Proof. To show (i), we obtain from Proposition 3.2 and Theorem 3.20 that

$$\ast_{\mathbf{B},G} D^i \le_{lo} \ast_{\mathbf{B}',G} D^i \le_{lo} \ast_{\mathbf{B}',G} E^i \le_{lo} \ast_{\mathbf{C},G} E^i\,.$$

Then, the statement follows with Sklar's Theorem. Statements (ii) and (iii) follow analogously. □

Since the independence copula and its associated survival copula are componentwise convex, we obtain as a consequence of the above theorem ordering results for the standard factor model.

Corollary 4.3 (Ordering results for standard factor models)**.**
Assume that $\mathbf{B} = \mathbf{C} = \Pi^{\mathbf{d}} = (\Pi^d)$ *. If* $D^i \le_{\partial_2 S,G} E^i$ *and if* E^i *is CIS for all* i *, then*

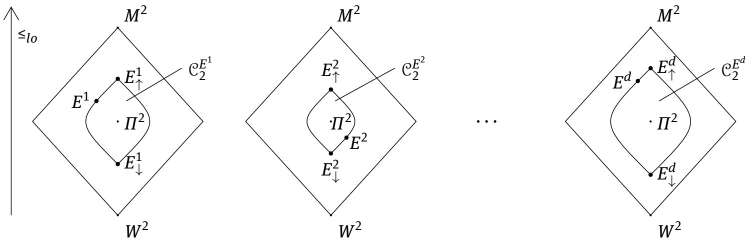

Figure 4 Classes $\mathcal{C}_2^{E^i} = \{C \in \mathcal{C}_2 \mid C \leq_{\partial_2 S} E^i\}$ of bivariate copulas generated by the copulas $E^i \in \mathcal{C}_2$, $i = 1, \ldots, d$, via the $\leq_{\partial_2 S}$-ordering. Note that M^2, Π^2, and W^2 denote the upper Fréchet copula, the independence copula, and the lower Fréchet copula, respectively. The copulas $E^i_\uparrow$ and $E^i_\downarrow$ are the uniquely determined copulas that are CIS and CDS, respectively, such that $E^i_\uparrow =_{\partial_2 S} E^i =_{\partial_2 S} E^i_\downarrow$, see Proposition 3.17. As a consequence of Proposition 3.14, it holds for all $D \in \mathcal{C}_2^{E^i}$ that $\zeta_1(D^T) \leq \zeta_1((E^i)^T)$.

(i) $X_i \leq_{lo} Y_i$ *for all* i *implies* $X \leq_{lo} Y$,
(ii) $X_i \leq_{uo} Y_i$ *for all* i *implies* $X \leq_{uo} Y$,
(iii) $X_i \overset{\mathrm{d}}{=} Y_i$ *for all* i *implies* $X \leq_c Y$.

In the following remark, we determine sharp bounds for some relevant classes of CSFMs including classes of standard factor models with bounded bivariate specification sets.

Remark 4.4. *Let* $F_i \in \mathcal{F}^1$ *for all* i. *Denote by* $\prec$ *one of the orderings* $\leq_{lo}$ *and* $\leq_{uo}$. *For* $E^i \in \mathcal{C}_2$, *denote by* $\mathcal{C}_2^{E^i} := \{C \in \mathcal{C}_2 \mid C \leq_{\partial_2 S} E^i\}$ *the class of bivariate copulas that is upper bounded w.r.t.* $\leq_{\partial_2 S}$ *by* E^i, $1 \leq i \leq d$. *For a risk factor* $Z \sim G$, $G \in \mathcal{F}_c^1$, *consider the class*

$$\mathcal{X}^f = \left\{\xi = (\xi_1, \ldots, \xi_d) \mid C_{\xi_i, Z} \in \mathcal{C}_2^{E^i},\ F_{\xi_i} \prec F_i \text{ for all } i,\ C_{\xi|Z=z} \prec \Pi^d \text{ for all } z\right\}$$

of d*-variate random vectors that are conditionally on* $Z = z$ *negatively dependent w.r.t.* $\prec$, *have marginal distributions upper bounded by* F_i, *and have dependence specifications* $C_{\xi_i, Z} \in \mathcal{C}_2^{E^i}$, *see Figure 4. Then, for all* $\xi \in \mathcal{X}^f$, *it holds that*

$$F_\xi \prec \Pi_{i=1}^d E^i_\uparrow (F_1, \ldots, F_d), \tag{34}$$

where $E^i_\uparrow$ *is the uniquely determined CIS copula such that* $E^i_\uparrow =_{\partial_2 S} E^i$, *see Proposition 3.17. Further, a vector* $\xi \in \mathcal{X}^f$ *such that* $\xi \sim \Pi_{i=1}^d E^i_\uparrow (F_1, \ldots, F_d)$ *can explicitly be determined which implies that the bound in* (34) *is attained, compare Corollary 2.9.*

In PSFMs, the conditional copulas are not specified. For the comparison of upper bounds in classes of PSFMs, we note that the worst case distribution in a PSFM w.r.t. the orthant orders is obtained when the conditional copula specifications attain the upper Fréchet copula.

Theorem 4.5 (Upper bounds in classes of PSFMs)**.**
Assume that $\mathbf{C} = \mathbf{M^d} = (M^d)$. *If* $D^j = E^j \leq_{\partial\Delta} D^d, E^d$ *for* $j = 1, \ldots, d-1$ *and* $D^d \leq_{s\partial\Delta} E^d$, *then*

(i) $X_i \leq_{lo} Y_i$ *for all* i *implies* $X \leq_{lo} Y$,
(ii) $X_i \leq_{uo} Y_i$ *for all* i *implies* $X \leq_{uo} Y$,
(iii) $X_i \overset{\mathrm{d}}{=} Y_i$ *for all* i *implies* $X \leq_c Y$.

Proof. From Proposition 3.2 and Theorem 3.25 we obtain that

$$\ast_{\mathbf{B},G} D^i \leq_{lo} \bigvee_G D^i \leq_{lo} \bigvee_G E^i = \ast_{\mathbf{C},G} E^i .$$

Then (i) follows with Sklar's Theorem. Statements (ii) and (iii) follow analogously. □

Similarly, we obtain for lower bounds in the two- and three-dimensional case the following result.

Theorem 4.6 (Lower bounds in classes of PSFMs, $d = 3$)**.**
Assume that $\mathbf{B} = \mathbf{W^3} = (W^3)$ *and* $D^1 = E^1 = M^2$ *. If* $D^2_* = E^2_* \leq_{\partial\Delta} D^3, E^3$ *and* $D^3 \leq_{s\partial\Delta} E^3$ *, Then*

(i) $X_i \leq_{lo} Y_i$ *implies* $(X_2, X_3) \leq_{lo} (Y_2, Y_3)$ *, and if* $G \in \mathcal{F}_c^1$ *then* $(X_1, X_2, X_3) \leq_{lo} (Y_1, Y_2, Y_3)$ *,*
(ii) $X_i \leq_{uo} Y_i$ *implies* $(X_2, X_3) \leq_{uo} (Y_2, Y_3)$ *, and if* $G \in \mathcal{F}_c^1$ *then* $(X_1, X_2, X_3) \leq_{uo} (Y_1, Y_2, Y_3)$ *,*
(iii) $X_i \stackrel{d}{=} Y_i$ *implies* $(X_2, X_3) \leq_c (Y_2, Y_3)$ *, and if* $G \in \mathcal{F}_c^1$ *then* $(X_1, X_2, X_3) \leq_c (Y_1, Y_2, Y_3)$ *,*

Proof. For $G \in \mathcal{F}_c^1$, we obtain from Theorem 3.27 and Proposition 3.2 that

$$\ast_{\mathbf{B}} D^i = M^2 \wedge D^2 \wedge D^3 \leq_{lo} M^2 \wedge E^2 \wedge E^3 \leq_{lo} \ast_{\mathbf{C}} E^i .$$

Then, $(X_1, X_2, X_3) \leq_{lo} (Y_1, Y_2, Y_3)$ follows with Sklar's Theorem.
For general $G \in \mathcal{F}^1$, denote by $\mathbf{C^{23}} = (C_t^{23})_{t\in[0,1]}$ the bivariate $(2, 3)$-marginal copulas of $\mathbf{C}$, i.e., $C_t^{23}(u, v) = C_t(1, u, v)$ for all $t, u, v \in [0, 1]$. Similarly, $\mathbf{B^{23}} = \mathbf{W^2} = (W^2)$. Then, we obtain that

$$D^2 \ast_{\mathbf{B^{23}},G} D^3 = D^2 \wedge_G D^3 \leq_{lo} E^2 \wedge_G E^3 \leq_{lo} E^2 \ast_{\mathbf{C^{23}},G} E^3 ,$$

and, thus, $(X_2, X_3) \leq_{lo} (Y_2, Y_3)$. For the upper orthant and concordance ordering, the statements follow analogously. □

Note that the same results hold true if the inequality signs $\leq_{\partial\Delta}$ and $\leq_{s\partial\Delta}$ in Theorem 4.5 and Theorem 4.6 (with $D^1 = E^1 = W^2$) are reversed.

For classes of *partially specified internal factor models* (PSIFMs) where the first component of the risk vector in the PSFM coincides with (an increasing function of) the factor variable, see [3], we obtain the following results. Note that in this class, the first bivariate dependence specification is given by the upper Fréchet copula M^2 .

Theorem 4.7 (Upper bounds in classes of PSIFMs)**.**
Assume that $G \in \mathcal{F}_c^1$ *and* $\mathbf{C} = \mathbf{M^d} = (M^d)$ *. If* $D^1 = E^1 = M^2$ *and* $D^j \leq_{lo} E^j = E$ *for* $j = 2, \dots, d$ *and* $E \in \mathcal{C}_2$ *, then*

(i) $X_i \leq_{lo} Y_i$ *for all i implies* $X \leq_{lo} Y$ *,*
(ii) $X_i \leq_{uo} Y_i$ *for all i implies* $X \leq_{uo} Y$ *,*
(iii) $X_i \stackrel{d}{=} Y_i$ *for all i implies* $X \leq_c Y$ *.*

Proof. From Proposition 3.2 and Theorem 3.26, we obtain that

$$\ast_{\mathbf{B}} D^i \leq_{lo} \bigvee D^i \leq_{lo} \bigvee E^i = \ast_{\mathbf{C}} E^i .$$

Thus, the statement follows with Sklar's Theorem. Statements (ii) and (iii) follow analogously. □

For lower bounds in the three-dimensional case, we obtain the following result.

Theorem 4.8 (Lower bounds in classes of PSIFMs, $d = 3$)**.**
Assume that $G \in \mathcal{F}_c^1$ *and* $\mathbf{B} = \mathbf{W^3} = (W^3)$ *. If* $D^1 = E^1 = M^2$ *and* $D^2 = D^3_* \leq_{lo} E^j$ *for* $j = 2, 3$ *, then*

(i) $X_i \leq_{lo} Y_i$ *for all i implies* $X \leq_{lo} Y$,
(ii) $X_i \leq_{uo} Y_i$ *for all i implies* $X \leq_{uo} Y$,
(iii) $X_i \stackrel{\mathrm{d}}{=} Y_i$ *for all i implies* $X \leq_c Y$.

Proof. From Theorem 3.28 and Proposition 3.2, we obtain

$$\ast_{\mathbf{B}} D^i = D^1 \wedge D^2 \wedge D^3 \leq_{lo} E^1 \wedge E^2 \wedge E^3 \leq_{lo} \ast_{\mathbf{C}} E^i .$$

Then, $(X_1, X_2, X_3) \leq_{lo} (Y_1, Y_2, Y_3)$ follows with Sklar's Theorem. Statements (ii) and (iii) follow analogously. □

Conclusion

In this paper, we obtain some general ordering results for factor models w.r.t. the specifications of the joint distributions of the components with the risk factor variable. The results generalize the upper product ordering results in [2, 3] to general conditional dependence structures and are based essentially on a version of Sklar‘s theorem as well as on classical ordering results based on rearrangements. The results in this paper allow to determine worst case distributions w.r.t. the orthant orderings for classes of CSFMs as well as in subclasses of PSFMs for any $d \geq 2$ and, similarly, of best case distributions for $d = 2, 3$. Related ordering results w.r.t. the stronger supermodular and the directionally convex ordering need different techniques and are the subject of a subsequent study.

Conflict of interest statement: Authors state no conflict of interest.

Acknowledgements: The authors thank both reviewers for their valuable comments which improved the manuscript. In particular, the counterexample G in (35) is due to one of the referees. We also thank Wolfgang Trutschnig for the hint to Remark 2.22(b).

A Appendix

Lemma A.1. *For* $G \in \mathcal{F}^1$ *, the following statements hold true.*

(i) ι_G *and* ι_G^- *are non-decreasing and Lebesgue-almost surely continuous.*
(ii) $G^{-1}(\iota_G(t)) = G^{-1}(t)$ *and* $\iota_G(G(x)) = G(x)$ *for all* $t \in [0, 1]$ *and* $x \in \mathbb{R}$.
(iii) *If* $G(x - \varepsilon) < G(x)$ *for all* $\varepsilon > 0$ *, then* $\iota_G^-(G(x)) = G^-(x)$.
(iv) $\iota_G^-(t) \leq t \leq \iota_G(t)\, \forall t \in (0, 1)$, $\iota_G^-(0) = 0 = \iota_G(0)$ *and* $\iota_G^-(1) \leq 1 = \iota_G(1)$.
(v) $\iota_G \circ \iota_G = \iota_G$ *and* $\iota_G^- \circ \iota_G = \iota_G^-$.
(vi) ι_G^- *is left-continuous.*
(vii) *For all* $y \in \mathbb{R}$, ι_G *is left-continuous at* $G(y)$ *and* ι_G^- *is continuous at* $G^-(y)$.
(viii) *In general,* ι_G *is neither left-continuous nor right-continuous.*
(ix) $\iota_G^-(t) = t = \iota_G(t)$ *if and only if G is continuous at* $G^{-1}(t)$.
(x) $\iota_G^-(t) = t = \iota_G(t)$ *for all* $t \in [0, 1]$ *if and only if* $G \in \mathcal{F}_c^1$.
(xi) *If* $\iota_G(t) > t$ *, then* $\iota_G(t + \varepsilon) = \iota_G(t)$ *and* $\iota_G^-(t + \varepsilon) = \iota_G^-(t)$ *for all* $0 < \varepsilon \leq \iota_G(t) - t$.
(xii) *If* $\iota_G^-(t) < t$ *, then* $\iota_G(t - \varepsilon) = \iota_G(t)$ *and* $\iota_G^-(t - \varepsilon) = \iota_G^-(t)$ *for all* $0 < \varepsilon < t - \iota_G^-(t)$.

Proof. (i): The non-decreasingness is clear. Since ι_G and ι_G^- can only have an at most countable number of jumps, the set of discontinuity points is a Lebesgue-null set.

(ii), (iii) and (iv) follow from the definition of G^{-1} and G^{-}, respectively, considering the cases where G is discontinuous and constant around x, respectively.
(v) is a consequence of (ii).
(vi): This follows from the left-continuity of G^{-} and G^{-1}.
(vii): To show the left-continuity of ι_G at $G(y)$, let $(t_n)_{n\in\mathbb{N}}$ be strictly increasing in $[0,1]$ with limit $G(y) > 0$. Then, we have

$$G(y) = \iota_G(G(y)) \geq \iota_G(t_n) \geq t_n \to G(y)$$

as $n \to \infty$ applying (ii), (i), and (iv). To show the right-continuity of ι_G^{-} at $G^{-}(y)$, let $(t_n)_{n\in\mathbb{N}}$ be strictly decreasing in $[0,1]$ with limit $G^{-}(y) < 1$. Then, we obtain similarly that

$$G^{-}(y) = \iota_G^{-}(G^{-}(y)) \leq \iota_G^{-}(t_n) \leq t_n \to G^{-}(y)\,.$$

(viii): Consider the distribution functions G and H defined by

$$G(x) = \begin{cases} 0 & \text{if } x < 0\,, \\ \frac{2}{3}x & \text{if } x \in [0, \frac{1}{2})\,, \\ \frac{2}{3}x + \frac{1}{3} & \text{if } x \in [\frac{1}{2}, 1]\,, \\ 1 & \text{if } x > 1\,. \end{cases} \qquad H(x) = \begin{cases} 0 & \text{if } x < 0\,, \\ x & \text{if } x \in [0, \frac{1}{2})\,, \\ \frac{1}{2} & \text{if } x \in [\frac{1}{2}, 1)\,, \\ 1 & \text{if } x \geq 1\,. \end{cases} \tag{35}$$

Then ι_G and ι_H are given by

$$\iota_G(t) = \begin{cases} t & \text{if } t \in [0, \frac{1}{3}) \cup (\frac{2}{3}, 1]\,, \\ \frac{2}{3} & \text{if } t \in [\frac{1}{3}, \frac{2}{3}]\,, \end{cases} \qquad \iota_H(t) = \begin{cases} t & \text{if } t \in [0, \frac{1}{2}]\,, \\ 1 & \text{if } t > \frac{1}{2}\,. \end{cases}$$

So, ι_G is not left-continuous at $t = \frac{1}{3}$ and ι_H is not right-continuous at $t = \frac{1}{2}$.
(ix): $\iota_G(t) = t = \iota_G^{-}(t)$ holds if and only if $G(G^{-1}(t)) = G^{-}(G^{-1}(t))$, which is equivalent to the continuity of G at $G^{-1}(t)$.
(x): If $G \in \mathcal{F}_c^1$, the statement follows from (ix). For the reverse direction, assume that $G \in \mathcal{F}^1 \setminus \mathcal{F}_c^1$ is discontinuous. Then there exists x such that $G(x) > G^{-}(x)$. For $t \in [G^{-}(x), G(x))$, it follows that $\iota_G(t) = G(G^{-1}(t)) = G(x) > t$.
(xi): Let $\varepsilon \in (0, \iota_G(t) - t]$. Then, the non-decreasingness of ι_G and (v) imply $\iota_G(t) \leq \iota_G(t+\varepsilon) \leq \iota_G(\iota_G(t)) = \iota_G(t)$ and $\iota_G^{-}(t) \leq \iota_G^{-}(t+\varepsilon) \leq \iota_G^{-}(\iota_G(t)) = \iota_G^{-}(t)$.
(xii): Let $x = G^{-1}(t)$. Then, the statement follows from $G^{-}(x) < t - \varepsilon \leq G(x)$. □

Proof of Theorem 2.2.
Consider the set $\mathcal{I}_c := \{(z_0, z_1) \mid z_0 < z_1\,,\ G \text{ is continuous on } (z_0, z_1)\}$ of open intervals on which G is continuous, and denote by $\mathcal{I}_s := \{\{z\} \mid z \in \mathbb{R}\}$ the set of one-point sets. Note that each element of $\mathcal{I}_c$ is the intersection of an open interval in $\mathbb{R}$ and the preimage of $(0,1)$ under G. We show that

$$\int_{(z_0,z_1)} F_{X|Z=z}(x)\,\mathrm{d}G(z) = \int_{(z_0,z_1)} \partial_2^G C(F(x), G(z))\,\mathrm{d}G(z)\,, \quad \text{and} \tag{36}$$

$$\int_{\{z\}} F_{X|Z=z}(x)\,\mathrm{d}G(z) = \int_{\{z\}} \partial_2^G C(F(x), G(z))\,\mathrm{d}G(z) \tag{37}$$

for all $(z_0, z_1) \in \mathcal{I}_c$ and $\{z\} \in \mathcal{I}_s$. Since G has at most countably many jump discontinuities, every open interval $(y_0, y_1) \subset \mathbb{R}$ can be written as a disjoint union of at most countably many elements of $\mathcal{I}_c$ and $\mathcal{I}_s$. Then, (36) and (37) imply

$$\int_{(y_0,y_1)} F_{X|Z=z}(x)\,\mathrm{d}G(z) = \int_{(y_0,y_1)} \partial_2^G C(F(x), G(z))\,\mathrm{d}G(z)$$

for all open intervals $(y_0, y_1) \subset \mathbb{R}$. Hence, the integrands coincide for G-almost all z, which yields (i).

To show (36), let $(z_0, z_1) \in \mathcal{I}_c$. Assume w.l.g. that $t_0 := G(z_0) < G^-(z_1) =: t_1$. Then we obtain from the disintegration theorem and Sklar's Theorem that

$$
\begin{aligned}
\int\limits_{(z_0,z_1)} F_{X|Z=z}(x)\,\mathrm{d}G(z) &= \lim_{z \nearrow z_1} F(x,z) - F(x,z_0) = C(F(x), G^-(z_1)) - C(F(x), G(z_0)) \\
&= \int\limits_{(t_0,t_1)} \partial_2 C(F(x), s)\,\mathrm{d}s = \int\limits_{(t_0,t_1)} \lim_{\varepsilon \searrow 0} \frac{C(F(x), s) - C(F(x), s-\varepsilon)}{\varepsilon}\,\mathrm{d}s \\
&= \int\limits_{(t_0,t_1)} \lim_{\varepsilon \searrow 0} \frac{C(F(x), \iota_G(\iota_G(s))) - C(F(x), \iota_G^-(\iota_G(s)-\varepsilon))}{\iota_G(\iota_G(s)) - \iota_G^-(\iota_G(s)-\varepsilon)}\,\mathrm{d}s \\
&= \int\limits_{(t_0,t_1)} \partial_2^G C(F(x), \iota_G(s))\,\mathrm{d}s = \int\limits_{(z_0,z_1)} \partial_2^G C(F(x), G(z))\,\mathrm{d}G(z)\,,
\end{aligned}
\tag{38}
$$

where the third equality follows from the disintegration theorem applied on copulas. For the fourth equality, we use that the left-hand derivative and the derivative of the copula w.r.t. the second component coincide for Lebesgue-almost all s. The fifth equality follows from $\iota_G(s) = s = \iota_G^-(s)$ and $\iota_G^-(s-\varepsilon) = s - \varepsilon$ for all $s \in (t_0, t_1)$ and $\varepsilon \in (0, s - t_0)$ because G is continuous at $G^{-1}(s)$ and $G^{-1}(s-\varepsilon)$, respectively, see Lemma A.1(ix). The sixth equality holds by definition of the differential operator in (2), and the last equality is a consequence of the transformation formula.

To show (37), assume for $z \in \mathbb{R}$ w.l.g. that $G(z) > G^-(z)$. Then we obtain

$$
\begin{aligned}
\int\limits_{\{z\}} F_{X|Z=y}(x)\,\mathrm{d}G(y) &= F(x,z) - \lim_{w \nearrow z} F(x,w) = C(F(x), G(z)) - C(F(x), G^-(z)) \\
&= \frac{C(F(x), \iota_G(G(z))) - C(F(x), \iota_G^-(G(z)))}{\iota_G(G(z)) - \iota_G^-(G(z))} \cdot (G(z) - G^-(z)) \\
&= \lim_{\varepsilon \searrow 0} \frac{C(F(x), \iota_G(G(z))) - C(F(x), \iota_G^-(G(z)-\varepsilon))}{\iota_G(G(z)) - \iota_G^-(G(z)-\varepsilon)} \cdot (G(z) - G^-(z)) \\
&= \int\limits_{\{z\}} \partial_2^G C(F(x), G(z))\,\mathrm{d}G(z)\,,
\end{aligned}
\tag{39}
$$

where we use $G(z) > G^-(z)$ and apply Lemma A.1(v) for the third equality. For the fourth equality, we use the left-continuity of ι_G^-, see Lemma A.1(vi). The last equality follows with the definition of the operator ∂_2^G in (2).

To show statement (ii) of Theorem 2.2, denote by $\mathbb{Q}$ the rational numbers. Due to part (i) it holds that

$$F_{X|Z=z}(x) = \partial_2^G C(F(x), G(z))$$

for all $x \in \mathbb{Q}$ and for all z outside the G-null set $N := \bigcup_{x \in \mathbb{Q}} N_x$. Then we obtain for $x \in \mathbb{R}$ that

$$F_{X|Z=z}(x) = \lim_{\substack{w \searrow x \\ w \in \mathbb{Q}}} F_{X|Z=z}(w) = \lim_{\substack{w \searrow x \\ w \in \mathbb{Q}}} \partial_2^G C(F(w), G(z)) =: H_z(x)$$

for all $z \in N^c$. For $z \in N^c$, the function H_z is by definition right-continuous. Since C is a 2-copula and thus 2-increasing, H_z is non-decreasing. Further, $H_z(-\infty) = 0$ and $H_z(\infty) = 1$. Hence, $H_z(x) = \lim_{w \searrow x} \partial_2^G C(F(w), G(z))$ coincides with $F_{X|Z=z}(x)$ for all $x \in \mathbb{R}$ and for all $z \in N^c$. This proves the assertion. □

Proof of Proposition 2.14.
(i) ⇒ (ii): Assume that $\overline{\mathsf{Ran}(G_1)} \neq \overline{\mathsf{Ran}(G_2)}$. As a consequence of Proposition 2.13, the d-variate products $\Pi_{G_1} M^2$ and $\Pi_{G_2} M^2$ do not coincide because for $G \in \mathcal{F}^1$, $\Pi_G M^2$ defines an ordinal sum with intervals $\{(\iota_G^-(t), \iota_G(t)) \mid \iota_G^-(t) \neq \iota_G(t)\,,\, t \in (0,1)\}$ which are different for $G = G_1$ and $G = G_2$ unless $\overline{\mathsf{Ran}(G_1)} = \overline{\mathsf{Ran}(G_2)}$.

(ii) ⇒ (iii): First assume that $\mathrm{Ran}(G_1) = \mathrm{Ran}(G_2)$. Then, for all $t \in (0,1)$, it holds that

$$\begin{aligned}\iota_{G_1}(t) &= G_1\left(\inf\{x \mid G_1(x) \geq t\}\right) = \inf\{u \in \mathrm{Ran}(G_1) | u \geq t\}\\ &= \inf\{u \in \mathrm{Ran}(G_2) | u \geq t\} = G_2\left(\inf\{x \mid G_2(x) \geq t\}\right) = \iota_{G_2}(t)\,.\end{aligned}$$

In the general case that $\overline{\mathrm{Ran}(G_1)} = \overline{\mathrm{Ran}(G_2)}$, it holds that $\mathrm{Ran}(G_1)$ and $\mathrm{Ran}(G_2)$ only differ in a Lebesgue-null set because distribution functions have at most countably many jump discontinuities. Hence, the first part implies that $\iota_{G_1}(t) = \iota_{G_2}(t)$ for Lebesgue-almost all t .
(iii) ⇒ (i): This follows from the definition of the $*$-product because $\ast_{\mathbf{B},G} D^i$ depends on G only through $\iota_G(t)$ for Lebesgue-almost all t . □

Proof of Proposition 2.15(iii). Assume that $D^i = D^j$ for all $i \neq j$. Then, for $u = (u_1, \dots, u_d) \in [0,1]^d$, it holds true that

$$\bigvee_G D^i(u) = \int_0^1 \min_i \left\{\partial_2^G D^1(u_i, t)\right\} \mathrm{d}t = \int_0^1 \partial_2^G D^1(\min_i\{u_i\}, t)\, \mathrm{d}t = \int_0^1 \partial_2^G D^1(\min_i\{u_i\}, G(y))\, \mathrm{d}G(y) = \min_i\{u_i\}\,,$$

where the second equality holds because $\partial_2^G D^1(\cdot, t)$ is increasing for all t , the third equality follows from (3) and the transformation formula, see, e.g., [32, Theorem 2], and the last equality is a consequence of Theorem 2.2 and the disintegration theorem.

For the reverse direction, assume w.l.g. that $d = 2$ and $D^1(w_1, w_2) > D^2(w_1, w_2)$ for some $(w_1, w_2) \in [0,1] \times \mathrm{Ran}(G)$. Then, there exist $(u, v) \in (0,1) \times \mathrm{Ran}(G)$ and an ε-ball $B_\varepsilon(u, v) \subset (0,1)^2$ such that

$$\partial_2^G D^1(x,t) > \partial_2^G D^2(x,t) \quad \text{for almost all } (x,t) \in B_\varepsilon((u,v))\,, \tag{40}$$

because, otherwise, it would hold that

$$D^1(w_1, w_2) = \int_0^{w_2} \partial_2^G D^1(w_1, t)\, \mathrm{d}t \leq \int_0^{w_2} \partial_2^G D^2(w_1, t)\, \mathrm{d}t = D^2(w_1, w_2)\,.$$

which is a contradiction to $D^1(w_1, w_2) > D^2(w_1, w_2)$. As a consequence of (40), we obtain that

$$M^2(u,u) = u = \int_0^1 \partial_2^G D^1(u,t)\, \mathrm{d}t > \int_0^1 \min\left\{\partial_2^G D^1(u,t), \partial_2^G D^2(u,t)\right\} \mathrm{d}t = D^1 \vee_G D^2(u,u)\,.$$

This yields $D^1 \vee_G D^2 \neq M^2$. □

Proof of Proposition 2.16.
The first statement is a consequence of [31, Proposition 3].

Since $\mathcal{S}_T(C^i) \in \mathcal{C}_2$ for all i , the product $\ast_B \mathcal{S}_T(C^i)$ is well-defined. Hence, the second statement follows from

$$\begin{aligned}\ast_B \mathcal{S}_T(C^i)\,(u_1, \dots, u_d) &= \int_0^1 B\left((\partial_2 \mathcal{S}_T(C^i)(u_i, t))_{1 \leq i \leq d}\right) \mathrm{d}t = \int_0^1 B\left((\partial_2 C^i(u_i, T(t)))_{1 \leq i \leq d}\right) \mathrm{d}t\\ &= \int_{[0,1]} B\left((\partial_2 C^i(u_i, s))_{1 \leq i \leq d}\right) \mathrm{d}\lambda^T(s) = \int_{[0,1]} B\left((\partial_2 C^i(u_i, s))_{1 \leq i \leq d}\right) \mathrm{d}\lambda(s)\\ &= \ast_B C^i\,(u_1, \dots, u_d)\end{aligned}$$

for all $(u_1, \dots, u_d) \in [0,1]^d$, using that $\partial_2 \mathcal{S}_T(C)(u,t) = \partial_2 C(u, T(t))$ for λ-almost all t . □

Proof of Lemma 2.17. (i): Let $t \in (0,1)$. Due to Lemma A.1, we consider three cases.
In the first case, assume that $\iota_G(t) = t$ and $\iota_G(t-\varepsilon) = t$ for some $\varepsilon > 0$. Define

$$t_0 := \inf\{s \mid \iota_G(s) = \iota_G(t)\}\,. \tag{41}$$

Then, Lemma A.1(xi) implies that ι_G^- in constant on $(t_0, t]$. We show that $\iota_G^-(t) = t_0$. Suppose that $\iota_G^-(t) > t_0$. Let $\eta = \iota_G^-(t) - t_0$. Then, $\iota_G^-(t_0+\delta) = t_0 + \eta$ for some $\delta \in (0,\eta)$. But this is a contradiction to Lemma A.1(iv). Suppose that $\iota_G^-(t) < t_0$. Then, Lemma A.1(xii) implies that ι_G is constant on $(\iota_G^-(t), t]$. But this is a contradiction to (41).
In the second case, assume that $\iota_G(t) = t$ and $\iota_G(t-\delta) = t-\delta$ for all $0 < \delta < \varepsilon$ for some $\varepsilon > 0$. Then, Lemma A.1(xii) implies that $\iota_G^-(t) = t$.
In the third case, assume that $\iota_G(t) \neq t$. Then, Lemma A.1(iv) implies that $\iota_G(t) > t$. Lemma A.1(xi) implies that ι_G^- is constant on $(t_0, \iota_G(t)]$ for t_0 defined by (41). We show that $\iota_G^-(t_0+\delta) = t_0$ for all $0 < \delta < \iota_G(t) - t_0$ and, thus, $\iota_G^-(t) = t_0$. Suppose that $\iota_G^-(t_0+\delta) > t_0$ for some $\delta \in (0, \iota_G(t) - t_0)$. Then, there is a contradiction to Lemma A.1(iv). Suppose that $\iota_G^-(t_0+\delta) < t_0$ for some $\delta \in (0, \iota_G(t) - t_0)$. Then, Lemma A.1(xii) yields a contradiction to the minimality of t_0.
All of the three above considered cases imply that $\iota_G^-(t) = \inf\{s \mid \iota_G(s) \geq \iota_G(t)\}$. It remains to show that $\iota_G(s) \geq \iota_G(t) \Leftrightarrow \iota_G(s) \geq t$. From Lemma A.1(iv), we obtain that $\iota_G(t) \geq t$, which implies the direction from left to right. For the reverse direction, we obtain from Lemma A.1(v) and (i) that $\iota_G(s) = \iota_G(\iota_G(s)) \geq \iota_G(t)$.

(ii): Consider the functions $F_n, F\colon \mathbb{R} \to [0,1]$, $n \in \mathbb{N}$, defined by

$$F_n(t) = \begin{cases} 0 & \text{if } t < 0\,,\\ \lim_{s \searrow t} \iota_{G_n}(s) & \text{if } t \in [0,1]\,,\\ 1 & \text{if } t > 1\,, \end{cases} \qquad F(t) = \begin{cases} 0 & \text{if } t < 0\,,\\ \lim_{s \searrow t} \iota_{G}(s) & \text{if } t \in [0,1]\,,\\ 1 & \text{if } t > 1\,, \end{cases}$$

$$F_n^-(t) = \begin{cases} 0 & \text{if } t < 0\,,\\ \lim_{s \nearrow t} \iota_{G_n}(s) & \text{if } t \in [0,1]\,,\\ 1 & \text{if } t > 1\,, \end{cases} \qquad F^-(t) = \begin{cases} 0 & \text{if } t < 0\,,\\ \lim_{s \nearrow t} \iota_{G}(s) & \text{if } t \in [0,1]\,,\\ 1 & \text{if } t > 1\,. \end{cases}$$

Then, F_n and F are distribution functions with left-continuous version F_n^- and F^-, respectively. Since by assumption $\iota_{G_n} \to \iota_G$ almost surely pointwise, we obtain that $F_n(t) \to F(t)$ for all t at which F is continuous. This implies that the generalized inverse distribution functions converge almost surely, i.e.,

$$F_n^{-1}(t) \to F^{-1}(t) \quad \text{for almost all } t \in [0,1]\,. \tag{42}$$

Since $F^{-1}(t) = \inf\{s \mid F(s) \geq t\} = \inf\{s \mid F^-(s) \geq t\}$, it holds by construction of F and F^- and by (i) that $F^{-1}(t) = \inf\{s \mid \iota_G(s) \geq t\} = \iota_G^-(t)$. Similarly, we obtain that $F_n^{-1}(t) = \iota_{G_n}^-(t)$. Hence, (42) implies that $\iota_{G_n}^-(t) \to \iota_G^-(t)$ for almost all $t \in [0,1]$. □

Proof of Proposition 2.25. (i): Let $D = E_*$ on $[0,1] \times \mathrm{Ran}(G)$. Then, for $(u,v) \in [0,1]^2$, it holds that

$$\begin{aligned} D \wedge_G E(u,v) &= \int_0^1 \max\left\{\partial_2^G D(u,t) + \partial_2^G E(v,t) - 1, 0\right\} = \int_0^1 \max\left\{\partial_2^G D(u,t) - \partial_2^G E_*(1-v,t), 0\right\} \mathrm{d}t \\ &= \int_0^1 \max\left\{\partial_2^G D(u,t), \partial_2^G D(1-v,t)\right\} \mathrm{d}t - 1 + v = \int_0^1 \partial_2^G D(\max\{u, 1-v\}, t)\,\mathrm{d}t - 1 + v \\ &= \max\{u, 1-v\} - 1 + v = W^2(u,v)\,. \end{aligned}$$

For the reverse direction, assume w.l.g. that $D(w_1,w_2) < E_*(w_1,w_2)$ for some $(w_1,w_2) \in [0,1] \times \mathrm{Ran}(G)$. Then, there exist $(u,v) \in (0,1) \times \mathrm{Ran}(G)$ and an ε-ball $B_\varepsilon(u,v) \subset (0,1)^2$ such that

$$\partial_2^G D(x,t) < \partial_2^G E_*(x,t) \quad \text{for almost all } (x,t) \in B_\varepsilon((u,v))\,, \tag{43}$$

because, otherwise, it would hold that

$$D(w_1, w_2) = \int_0^{w_2} \partial_2^G D(w_1, t)\,\mathrm{d}t \geq \int_0^{w_2} \partial_2^G E_*(w_1, t)\,\mathrm{d}t = E_*(w_1, w_2)\,,$$

which is a contradiction to $D^1(w_1, w_2) < E_*(w_1, w_2)$. As a consequence of (43), we obtain that

$$\begin{aligned} W^2(u, 1-u) = 0 = &\int_0^1 \partial_2^G D(u, t)\,\mathrm{d}t - u \\ &< \int_0^1 \max\left\{\partial_2^G D(u, t), \partial_2^G E_*(u, t)\right\}\,\mathrm{d}t - u = \int_0^1 \max\left\{\partial_2^G D(u, t), 1 - \partial_2^G E(1-u, t)\right\}\,\mathrm{d}t - u \\ &= \int_0^1 \max\left\{\partial_2^G D(u, t) + \partial_2^G E(1-u, t) - 1, 0\right\}\,\mathrm{d}t = D \wedge_G E\,(u, 1-u)\,. \end{aligned}$$

This yields $D \wedge_G E \neq W^2$.
(ii): If $G \in \mathcal{F}_c^1$ is continuous, then it holds that

$$\begin{aligned} M^2 \wedge_G D \wedge_G E(u) &= \int_0^1 \max\left\{\mathbb{1}_{\{u_1 \geq t\}} + \partial_2 D(u_2, t) + \partial_2 E(u_3, t) - 2\,, 0\right\}\,\mathrm{d}t \\ &= \int_0^{u_1} \max\left\{\partial_2 D(u_2, t) + \partial_2 E(u_3, t) - 1\,, 0\right\}\,\mathrm{d}t \end{aligned}$$

for $u = (u_1, u_2, u_3) \in [0, 1]^3$. This defines a 3-copula, compare Durante et al. [8, Proposition 2].
For the reverse direction, assume that $G \in \mathcal{F}^1 \backslash \mathcal{F}_c^1$ is discontinuous and that $M^2 \wedge_G D \wedge_G E$ is a 3-copula. Then, Theorem 2.7 implies that there exists a random vector (U_1, U_2, U_3, Z) (under an extension of the probability space if necessary) such that $Z \sim G$, $C_{U_1,Z} = M^2$, $C_{U_2,Z} = D$, and $C_{U_3,Z} = E$ as well as

$$P\left((U_1, U_2, U_3) \leq u \mid Z = z\right) = W^3\left(\partial_2^G M^2(u_1, t), \partial_2^G D(u_2, t), \partial_2^G E(u_3, t)\right)$$

for all $z = G^{-1}(t)$, $t \in (0, 1)$, and for all $u = (u_1, u_2, u_3) \in [0, 1]^3$. Since G is discontinuous, there exists $t_0 \in (0, 1)$ such that $\iota_G^-(t_0) < \iota_G(t_0)$. This implies that the conditional distribution functions $\partial_2^G M^2(\cdot, t_0)$, $\partial_2^G D(\cdot, t_0)$, and $\partial_2^G E(\cdot, t_0)$ are continuous. Now, choose $u = (u_1, u_2, u_3) \in [0, 1]^3$ and $v = (v_1, v_2, v_3) = (1, 1, 1)$ such that

$$\begin{aligned} \partial_2^G D(u_1, t_0) = \partial_2^G E(u_2, t_0) = \partial_2^G M^2(u_3, t_0) = 0.5\,, \\ \partial_2^G D(v_1, t_0) = \partial_2^G E(v_2, t_0) = \partial_2^G M^2(v_3, t_0) = 1\,. \end{aligned}$$

Then, it follows that

$$P\left((U_1, U_2, U_3) \in [u, v] \mid Z = G^{-1}(t_0)\right) = V_{W^3}\left([\tfrac{1}{2}, 1]^3\right) = -0.5 < 0\,,$$

where $V_{W^3}([\frac{1}{2}, 1]^3)$ denotes the W^3-volume of the box $[\frac{1}{2}, 1]^3 \subset [0, 1]^3$, see Nelsen [22, Exercise 2.36]. This yields a contradiction and, thus, $M^2 \wedge_G D \wedge_G E$ is not a copula.
(iii) is a consequence of Theorem 2.7 and Remark 2.8.
(iv) and (v): For $(u, v) \in [0, 1] \times \mathrm{Ran}(G)$, it holds that

$$D \wedge_G M^2(u, v) = \int_0^1 \max\left\{\partial_2^G(u, t) + \partial_2^G M^2(v, t) - 1, 0\right\}\,\mathrm{d}t = \int_0^v \partial_2^G D(u, t)\,\mathrm{d}t = D(u, v)\,,$$

where the second equality holds true because $\partial_2^G M^2(v,t) = \mathbb{1}_{\{v>t\}}$ using that $v \in \mathrm{Ran}(G)$. The third equality follows from Theorem 2.2.
The other statements follow similarly.
(vi): As a consequence of (iv), $\bigwedge$ is not commutative if D is not symmetric. For a counterexample for associativity, let $D^i \in \mathcal{C}_2$ be a Gaussian copula with correlation $\rho_i \in (-1,1)$, $i = 1,2,3$. Then, $C^i \wedge C^j$ is a Gaussian copula with correlation $m(\rho_i,\rho_j) = \rho_i\rho_j - \sqrt{1-\rho_i^2}\sqrt{1-\rho_j^2}$. Obviously, in general, it holds that $m(\rho_1, m(\rho_2,\rho_3)) \neq m(m(\rho_1,\rho_2),\rho_3)$.

□

Proof of Lemma 3.6. For condition (21), the statement is trivial.
For condition (22), we need to show that

$$\int_0^\delta \left[B_{a+\delta+s}(u) - B_{a+\delta+s}(v) + B_{a+s}(v) - B_{a+s}(u)\right] \,\mathrm{d}s \geq 0\,, \quad \forall\, 0 \leq a \leq 1-2\delta\,,\ \forall \delta > 0\,, \tag{44}$$

implies

$$\int_0^\delta \left[B^G_{a+\delta+s}(u) - B^G_{a+\delta+s}(v) + B^G_{a+s}(v) - B^G_{a+s}(u)\right] \,\mathrm{d}s \geq 0\,, \quad \forall\, 0 \leq a \leq 1-2\delta\,,\ \forall \delta > 0\,, \tag{45}$$

where $u = (u_k)$, $v = (v_k) \in [0,1]^d$ such that for some $i \in \{1,\dots,d\}$ and $u_i \leq v_i$, $u_j = v_j$ for all $j \neq i$.

Consider the function $f\colon [0,1] \to [-1,0]$ given by

$$f(t) = B_t(u) - B_t(v)\,.$$

Then, condition (44) is equivalent to

$$\int_0^\delta f(a+\delta+s)\,\mathrm{d}s \geq \int_0^\delta f(a+s)\,\mathrm{d}s \quad \text{for all } 0 \leq a \leq 1-2\delta \text{ and } \delta > 0\,,$$

which means that f is increasing. Thus, the smoothed function $f^G\colon [0,1] \to [-1,0]$ given by

$$\begin{aligned} f^G(t) &= \begin{cases} f(t)\,, & \text{if } \iota_G^-(t) = \iota_G(t)\,, \\ \frac{1}{\iota_G(t)-\iota_G^-(t)} \int_{\iota_G^-(t)}^{\iota_G(t)} f(s)\,\mathrm{d}s\,, & \text{if } \iota_G^-(t) \neq \iota_G(t) \end{cases} \\ &= B_t^G(u_i) - B_t^G(u_i + h) \end{aligned}$$

is also increasing. But this is equivalent to (45).

For condition (23), the statement follows analogously. □

Conflict of interest statement: Authors state no conflict of interest.

References

[1] Abdous, B., C. Genest, and B. Rémillard (2005). Dependence properties of meta-elliptical distributions. In P. Duchesne and B. Rèmillard (Eds.), *Statistical Modeling and Analysis for Complex Data Problems*, pp. 1–15. Springer, New York.
[2] Ansari, J. and L. Rüschendorf (2018). Ordering risk bounds in factor models. *Depend. Model. 6*, 259–287.
[3] Ansari, J. and L. Rüschendorf (2020). Upper risk bounds in internal factor models with constrained specification sets. *Probab. Uncertain. Quant. Risk 5*, no. 3, 30 pp.
[4] Benedetto, J. J. and W. Czaja (2009). *Integration and Modern Analysis*. Birkhäuser, Boston.
[5] Bernard, C., L. Rüschendorf, S. Vanduffel, and R. Wang (2017). Risk bounds for factor models. *Finance Stoch. 21*(3), 631–659.
[6] Darsow, W. F., B. Nguyen, and E. T. Olsen (1992). Copulas and Markov processes. *Illinois J. Math. 36*(4), 600–642.

[7] Das Gupta, S., I. Olkin, L. J. Savage, M. L. Eaton, M. Perlman, and M. Sobel (1972). Inequalities on the probability content of convex regions for elliptically contoured distributions. *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, pp. 241–265.

[8] Durante, F., E. P. Klement, J. J. Quesada-Molina, and P. Sarkoci (2007). Remarks on two product-like constructions for copulas. *Kybernetika 43*(2), 235–244.

[9] Durante, F., P. Sarkoci, and C. Sempi (2009). Shuffles of copulas. *J. Math. Anal. Appl. 352*(2), 914–921.

[10] Durante, F. and C. Sempi (2016). *Principles of Copula Theory*. CRC Press, Boca Raton FL.

[11] Fan, K. and G. G. Lorentz (1954). An integral inequality. *Amer. Math. Monthly 61*, 626–631.

[12] Georges, P., A.-G. Lamy, E. Nicolas, G. Quibel, and T. Roncalli (2001). Multivariate survival modelling: a unified approach with copulas. Available at https://ssrn.com/abstract=1032559.

[13] Klement, E. P., A. Kolesárová, R. Mesiar, and S. Saminger-Platz (2017). Copula constructions using ultramodularity. In M. Ùbeda Flores, E. de Amo Artero, F. Durante and J. Fernàndez Sànchez (Eds.), *Copulas and Dependence Models with Applications*, pp. 135–156. Springer, Cham.

[14] Klement, E. P., M. Manzi, and R. Mesiar (2014). Ultramodularity and copulas. *Rocky Mountain J. Math. 44*(1), 189–202.

[15] Krupskii, P. and H. Joe (2013). Factor copula models for multivariate data. *J. Multivariate Anal. 120*, 85–101.

[16] Lorentz, G. G. (1953). An inequality for rearrangements. *Amer. Math. Monthly 60*, 176–179.

[17] Mesiar, R. and C. Sempi (2010). Ordinal sums and idempotents of copulas. *Aequationes Math. 79*(1-2), 39–52.

[18] Mikusiński, P. and M. D. Taylor (2010). Some approximations of n-copulas. *Metrika 72*(3), 385–414.

[19] Müller, A. and M. Scarsini (2001). Stochastic comparison of random vectors with a common copula. *Math. Oper. Res. 26*(4), 723–740.

[20] Müller, A. and M. Scarsini (2005). Archimedean copulae and positive dependence. *J. Multivariate Anal. 93*(2), 434–445.

[21] Müller, A. and D. Stoyan (2002). *Comparison Methods for Stochastic Models and Risks*. Wiley, Chichester.

[22] Nelsen, R. B. (2006). *An Introduction to Copulas. Second Edition*. Springer, New York.

[23] Rüschendorf, L. (1981). Characterization of dependence concepts in normal distributions. *Ann. Inst. Statist. Math. 33*(3), 347–359.

[24] Rüschendorf, L. (1983). Solution of a statistical optimization problem by rearrangement methods. *Metrika 30*, 55–61.

[25] Rüschendorf, L. (2009). On the distributional transform, Sklar's theorem, and the empirical copula process. *J. Statist. Plann. Inference 139*(11), 3921–3927.

[26] Rüschendorf, L. (2013). *Mathematical Risk Analysis*. Springer, Heidelberg.

[27] Schweizer, B. and A. Sklar (1974). Operations on distribution functions not derivable from operations on random variables. *Studia Math. 52*, 43–52.

[28] Shaked, M. and J. G. Shanthikumar (2007). *Stochastic Orders*. Springer, New York.

[29] Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris 8*, 229–231.

[30] Trutschnig, W. (2011). On a strong metric on the space of copulas and its induced dependence measure. *J. Math. Anal. Appl. 384*(2), 690–705.

[31] Trutschnig, W. and J. F. Sánchez (2013). Some results on shuffles of two-dimensional copulas. *J. Statist. Plann. Inference 143*(2), 251–260.

[32] Winter, B. B. (1997). Transformations of Lebesgue-Stieltjes integrals. *J. Math. Anal. Appl. 205*(2), 471–484.

DE GRUYTER Depend. Model. 2021; 9:307–326

Research Article Open Access
Special Issue in memory of Abe Sklar

Thimo M. Kasper, Sebastian Fuchs, and Wolfgang Trutschnig*

On convergence of associative copulas and related results

https://doi.org/10.1515/demo-2021-0114
Received May 27, 2021; accepted July 22, 2021

Abstract: Triggered by a recent article establishing the surprising result that within the class of bivariate Archimedean copulas $\mathcal{C}_{\mathrm{ar}}$ different notions of convergence - standard uniform convergence, convergence with respect to the metric D_1, and so-called weak conditional convergence - coincide, in the current contribution we tackle the natural question, whether the obtained equivalence also holds in the larger class of associative copulas $\mathcal{C}_{\mathrm{a}}$. Building upon the fact that each associative copula can be expressed as (finite or countably infinite) ordinal sum of Archimedean copulas and the minimum copula M we show that standard uniform convergence and convergence with respect to D_1 are indeed equivalent in $\mathcal{C}_{\mathrm{a}}$. It remains an open question whether the equivalence also extends to weak conditional convergence. As by-products of some preliminary steps needed for the proof of the main result we answer two conjectures going back to Durante et al. and show that, in the language of Baire categories, when working with D_1 a typical associative copula is Archimedean and a typical Archimedean copula is strict.

Keywords: associative copulas, Archimedean copulas, weak convergence, Baire category

MSC: 62H05, 60E05, 54E52

1 Introduction

Various different notions of convergence in the family of bivariate copulas $\mathcal{C}$ have been considered in the literature: The standard uniform metric d_∞ is probably the most common choice; since, however, d_∞ is not capable of distinguishing independence and complete dependence (or, in the words of [24], d_∞ does not 'distinguish between different types of statistical dependence') the stronger metric D_1 was introduced in [36]. Letting $K_A(\cdot,\cdot), K_B(\cdot,\cdot)$ denote Markov kernels (regular conditional distributions) of the copulas A and B, respectively, the metric D_1 is defined by

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

According to [36], D_1 is a metrization of the strong operator topology (of the Markov operators induced by the copulas, see [24, 29] for more information on the strong operator topology and the one-to-one correspondence between bivariate copulas and Markov operators), convergence with respect to D_1 implies convergence with respect to d_∞ (but not vice versa) and, setting $\zeta_1(A) := 3D_1(A,\Pi)$ yields a dependence measure ζ_1 fulfilling that $\zeta_1(A)$ is maximal if and only if A is completely dependent.

Thimo M. Kasper: Department for Mathematics, University of Salzburg, E-mail: thimo.kasper@sbg.ac.at
Sebastian Fuchs: Department for Mathematics, University of Salzburg, E-mail: sebastian.fuchs@sbg.ac.at
***Corresponding Author: Wolfgang Trutschnig:** Department for Mathematics, University of Salzburg, E-mail: wolfgang.trutschnig@sbg.ac.at

Sticking to Markov kernels/conditional distributions and considering weak convergence of almost all conditional distributions of copulas results in the notion of weak conditional convergence which was introduced recently in [17]. Weak conditional convergence implies convergence with respect to D_1, and it is straightforward to construct examples illustrating that the reverse implication is wrong (again see [17]). When working within the subclass of absolutely continuous copulas other natural notions of convergence are the total variation distance TV (which, in the absolutely continuous setting coincides with the L_1-distance of the densities) and the Kullback-Leibler divergence KL (see, e.g., [35] and the references therein). In case the copula densities are greater than 0 almost everywhere then according to [26] we have the following interrelation, where $a \Rightarrow b$ indicates the convergence with respect to a implies convergence with respect to b:

$$KL \Rightarrow TV \Rightarrow D_1 \Rightarrow d_\infty$$

As mentioned above, the topologies induced by d_∞, D_1 and weak conditional convergence on $\mathcal{C}$ do not coincide: the topology induced by weak conditional convergence is strictly finer than the one induced by D_1, which, in turn is strictly finer than the one induced by d_∞. According to [17], however, on large classes like the family of Archimedean copulas as well as on the family of Extreme Value copulas, all three topologies do coincide. Considering that the family of Archimedean copulas is dense in the family of associative copulas with respect to d_∞ (see [18, 19]) and that every associative copula admits a so-called ordinal sum representation (see [23]) in terms of finitely or countably infinitely many Archimedean copulas and M it is natural to ask, whether the three notions of convergence are also equivalent on the family $\mathcal{C}_a$ of all associative copulas. In the current contribution we provide a partial affirmative answer and show that on $\mathcal{C}_a$ convergence with respect to d_∞ and D_1 are indeed equivalent. The question, whether weak conditional convergence is equivalent too, remains open - we have neither been able to prove equivalence nor to construct a counterexample.

Although the main result of the paper is the equivalence of d_∞ and D_1 on $\mathcal{C}_a$ several auxiliary results on denseness of the class of strict and the class of non-strict Archimedean copulas in $(\mathcal{C}_a, D_1)$ produce nice by-products in so far as we are able to answer two conjectures going back to Durante et al. in [6] and show that, in the language of Baire categories, when working with D_1 a 'typical' associative copula is Archimedean and a 'typical' Archimedean copula is strict.

The remainder of this paper is organized as follows: Section 2 contains notation and preliminaries which will be used in the sequel. In Section 3 we first prove the fact that the family of all strict and the family of all non-strict Archimedean copulas are dense in $(\mathcal{C}_a, D_1)$ by following the procedure studied in [19] for d_∞, and then show the afore-mentioned Baire category results. Finally, Section 4 focuses on convergence of associative copulas and establishes the main result in several steps. Some examples and graphics illustrate the chosen approach.

2 Notation and preliminaries

In the sequel we will let $\mathcal{C}$ denote the family of all bivariate copulas. For each copula C the corresponding doubly stochastic measure will be denoted by μ_C, i.e. $\mu_C([0, x] \times [0, y]) = C(x, y)$ for all $x, y \in [0, 1]$. For more background on copulas and doubly stochastic measures we refer to [8, 27].

As copulas can be seen as binary operations on $[0, 1]^2$ associativity is of particular interest, that is, $C \in \mathcal{C}$ is called *associative* if for all $x, y, z \in [0, 1]$ we have

$$C(C(x, y), z) = C(x, C(y, z)).$$

$\mathcal{C}_a$ denotes the class of all associative copulas. Associative copulas are closely related to *triangular norms*. In fact, according to [25] a copula C is a triangular norm if and only if $C \in \mathcal{C}_a$ whereas a triangular norm T is a copula if and only if it is Lipschitz continuous with Lipschitz constant 1 (see, for example, [1, 18]).

The standard *uniform metric* d_∞ on $\mathcal{C}$ is defined by

$$d_\infty(C_1, C_2) := \max_{(x,y)\in[0,1]^2} |C_1(x, y) - C_2(x, y)|. \tag{2.1}$$

It is well known that the metric space $(\mathcal{C}, d_\infty)$ is compact and that pointwise and uniform convergence of a sequence of copulas $(C_n)_{n\in\mathbb{N}}$ are equivalent (see [8, 37]).

For every metric space (S, d) the Borel σ-field on S will be denoted by $\mathcal{B}(S)$. In what follows Markov kernels will play a prominent role. A *Markov kernel* from $\mathbb{R}$ to $\mathbb{R}$ is a mapping $K : \mathbb{R} \times \mathcal{B}(\mathbb{R}) \to [0, 1]$ such that for every fixed $E \in \mathcal{B}(\mathbb{R})$ the mapping $x \mapsto K(x, E)$ is (Borel-)measurable and for every fixed $x \in \mathbb{R}$ the mapping $E \mapsto K(x, E)$ is a probability measure. Given two real-valued random variables X, Y on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and letting $\mathbb{1}_E$ denote the characteristic function of a set $E \subseteq \mathbb{R}$ we say that a Markov kernel K is a *regular conditional distribution* of Y given X if $K(X(\omega), E) = \mathbb{E}(\mathbb{1}_E \circ Y|X)(\omega)$ holds $\mathbb{P}$-almost surely for every $E \in \mathcal{B}(\mathbb{R})$. It is well-known that for X, Y as above, a regular conditional distribution of Y given X always exists and is unique for $\mathbb{P}^X$-a.e. $x \in \mathbb{R}$ whereby $\mathbb{P}^X$ denotes the push-forward of $\mathbb{P}$ via X, i.e., $\mathbb{P}^X(E) = \mathbb{P}(X^{-1}(E))$ for every $E \in \mathcal{B}(\mathbb{R})$ (see, e.g., [15, 20]). In case (X, Y) has distribution function $C \in \mathcal{C}$ we will let $K_C : [0, 1] \times \mathcal{B}([0, 1]) \to [0, 1]$ denote (a version of) the regular conditional distribution of Y given X and refer to it as *Markov kernel of C*. Defining the x-section of a set $G \in \mathcal{B}([0, 1]^2)$ by $G_x := \{y \in [0, 1] : (x, y) \in G\}$ we have the following *disintegration formula* (see [15, 20])

$$\int_{[0,1]} K_C(x, G_x)\, \mathrm{d}\mathbb{P}^X(x) = \mu_C(G) \tag{2.2}$$

hence, in particular,

$$\int_{[0,1]} K_C(x, E)\, \mathrm{d}\mathbb{P}^X(x) = \int_{[0,1]} K_C(x, E)\, \mathrm{d}\lambda(x) = \lambda(E) \tag{2.3}$$

for every $E \in \mathcal{B}([0, 1])$, whereby λ denotes the Lebesgue measure on $\mathbb{R}$. For more information on conditional expectation and general disintegration we refer to [15, 20].

Following [8] the rightside upper Dini derivative $\bar{\partial}_1^+ C : [0, 1) \times [0, 1] \to [0, 1]$ of a copula C with respect to the first coordinate is defined by

$$\bar{\partial}_1^+ C(x, y) := \limsup_{h\to 0+} \frac{C(x + h, y) - C(x, y)}{h} \in [0, 1].$$

As shown in [7], the map $y \mapsto \bar{\partial}_1^+ C(x, y)$ is non-decreasing which we will use subsequently to clarify the relation between Markov kernel and rightside upper Dini derivative:

Lemma 2.1. *Let C be an arbitrary copula. Then*

$$\bar{\partial}_1^+ C(x, y) = K_C(x, [0, y])$$

for λ_2-a.e. $(x, y) \in (0, 1)^2$.

Proof. Letting ∂_i denote the partial derivative with respect to the i-th coordinate it is well-known (see, e.g., [8, 33]) that for every $y \in [0, 1]$ there exists a set $\Lambda_y \in \mathcal{B}([0, 1])$ of full measure such that for all $x \in \Lambda_y$ we have that $\partial_1 C(x, y)$ exists and

$$\bar{\partial}_1^+ C(x, y) = \partial_1 C(x, y) = K_C(x, [0, y])$$

holds. Then $\Lambda := \bigcap_{y\in(0,1)\cap\mathbb{Q}} \Lambda_y \in \mathcal{B}([0, 1])$ is a set of full measure and for all $x \in \Lambda$, $y \in (0, 1) \cap \mathbb{Q}$ we have

$$\bar{\partial}_1^+ C(x, y) = K_C(x, [0, y]).$$

As for $x \in \Lambda$, $y \in (0, 1)$ we have $K_C(x, [0, y]) \geq \bar{\partial}_1^+ C(x, y)$ it suffices to show that

$$\Gamma := \{(x, y) \in [0, 1]^2 : x \in \Lambda, K_C(x, [0, y]) > \bar{\partial}_1^+ C(x, y)\}$$

is a λ_2-null set which can be done as follows: First, note that $(x, y) \mapsto K_C(x, [0, y])$ as well as $(x, y) \mapsto K_C(x, [0, y))$ are measurable functions. Second, for $x \in \Lambda$, $y \in [0, 1]$ we have $K_C(x, [0, y]) > \bar{\partial}_1^+ C(x, y) \Leftrightarrow$

$K_C(x,[0,y]) > K_C(x,[0,y))$ whence

$$\begin{aligned}\Gamma &= \{(x,y) \in \Lambda \times [0,1] : K_C(x,[0,y)) < K_C(x,[0,y])\}\\ &= \{(x,y) \in [0,1]^2 : K_C(x,[0,y)) < K_C(x,[0,y])\} \cap (\Lambda \times [0,1])\end{aligned}$$

which is measurable. As for $x \in \Lambda$ we have Γ_x is at most countably infinite, applying disintegration yields $\lambda_2(\Gamma) = 0$. □

The following consequence of Lemma 2.1 is immediate:

Corollary 2.2. *Let $C_1, C_2 \in \mathcal{C}$ and $U, V \subseteq [0,1]$ with U open. If C_1 is identical to C_2 on $U \times V \subseteq [0,1]^2$ then*

$$K_{C_1}(x,[0,y]) = K_{C_2}(x,[0,y])$$

for λ_2-a.e. $(x,y) \in U \times V$.

A copula C is called *completely dependent* if there exists a λ-preserving transformation $h : [0,1] \to [0,1]$ (i.e., a transformation fulfilling $\lambda(h^{-1}(E)) = \lambda(E)$ for every $E \in \mathcal{B}([0,1])$) such that $K(x,E) := \mathbb{1}_E(h(x))$ is a Markov kernel of C. The class of all completely dependent copulas will be denoted by $\mathcal{C}_{\mathrm{cd}}$.
For two copulas C_1, C_2 the so-called *Markov product* or *star product* $C_1 * C_2$ is defined by

$$(C_1 * C_2)(x,y) := \int_{[0,1]} \partial_2 C_1(x,t) \cdot \partial_1 C_2(t,y)\, \mathrm{d}\lambda(t).$$

The star-product is always a copula (see [5]) and according to results of [38] a Markov kernel of $C_1 * C_2$ is given by

$$(K_{C_1} \circ K_{C_2})(x,E) := \int_{[0,1]} K_{C_2}(y,E) K_{C_1}(x,\mathrm{d}y).$$

Subsequently we will make use of the following properties of the star product: First, $C \in \mathcal{C}$ is completely dependent if and only if there exists a copula B such that the identity $B * C = M$ holds, and second, if $B \in \mathcal{C}$ and if the sequence of copulas $(C_n)_{n\in\mathbb{N}}$ converges uniformly to another copula C then $(B * C_n)_{n\in\mathbb{N}}$ converges uniformly to $B * C$. For more properties of the class $\mathcal{C}_{\mathrm{cd}}$ and the star product we refer to [22, 36] and [5, 14], respectively.

Considering Markov kernels allows to define stronger metrics than the standard uniform one:

$$D_1(C_1,C_2) := \int_{[0,1]^2} |K_{C_1}(x,[0,y]) - K_{C_2}(x,[0,y])|\, \mathrm{d}\lambda_2(x,y), \tag{2.4}$$

$$D_2^2(C_1,C_2) := \int_{[0,1]^2} (K_{C_1}(x,[0,y]) - K_{C_2}(x,[0,y]))^2\, \mathrm{d}\lambda_2(x,y),$$

$$D_\infty(C_1,C_2) := \sup_{y\in[0,1]} \int_{[0,1]} |K_{C_1}(x,[0,y]) - K_{C_2}(x,[0,y])|\, \mathrm{d}\lambda(x).$$

In [36] it was shown that D_1, D_2, D_∞ are three metrics generating the same topology on $\mathcal{C}$. In what follows we will primarily work with D_1 and refer to [9, 36] for more information on D_2 and D_∞. The metric space $(\mathcal{C}, D_1)$ is complete and separable but not compact (again see [36]).
An even stronger notion of convergence involving Markov kernels is introduced and studied in [17]: Let $C, C_1, C_2, \ldots$ be copulas with corresponding Markov kernels $K_C, K_{C_1}, K_{C_2}, \ldots$. Then $(C_n)_{n\in\mathbb{N}}$ is said to converge *weakly conditional* to C if and only if for λ-almost every $x \in [0,1]$ 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 $K_C(x,\cdot)$ (in short, $C_n \xrightarrow{wcc} C$, where 'wcc' stands for 'weak conditional convergence'). The following interrelation of the afore-mentioned modes of convergence holds:

$$\text{weak conditional convergence} \Rightarrow \text{convergence w.r.t. } D_1 \Rightarrow \text{convergence w.r.t. } d_\infty.$$

For counterexamples for the reverse implications again see [17] and [36].

3 The interrelation of Archimedean and associative copulas with respect to D_1

Triggered by [19], where the authors prove that the class of Archimedean copulas is dense in the class of associative copulas with respect to the uniform metric d_∞, in what follows we show the result w.r.t. D_1 and, as a by-product, answer two open questions posed in [6] regarding Baire category results for Archimedean copulas. In Section 4 we then exploit the proven denseness to derive surprising convergence results of associative copulas with respect to D_1.

An *Archimedean copula* $A = A_\varphi$ is a copula induced by a convex and strictly decreasing function $\varphi : [0,1] \to [0,\infty]$ with $\varphi(1) = 0$, called *generator*, via

$$A(x,y) = \varphi^-(\varphi(x) + \varphi(y)), \quad x, y \in [0,1].$$

Thereby $\varphi^- : [0,\infty] \to [0,1]$ denotes the pseudoinverse of φ, defined by

$$\varphi^-(x) := \begin{cases} \varphi^{-1}(x) & \text{if } x \in [0, \varphi(0+)) \\ 0 & \text{if } x \geq \varphi(0+) \end{cases}, \tag{3.1}$$

where $\varphi(x\pm) := \lim_{t\to x^\pm} \varphi(t)$ is the respective one-sided limit. We refer to A as the Archimedean copula induced by φ; in case of $\varphi(0+) = \infty$, A is called *strict* and *non-strict* otherwise. In the sequel $\mathcal{C}_{\text{ar}}$ denotes the class of all Archimedean copulas, $\mathcal{C}_{\text{ar}}^{\text{s}}$ the subclass of strict Archimedean copulas and $\mathcal{C}_{\text{ar}}^{\neg\text{s}}$ the subclass of non-strict Archimedean copulas. Since generators are only unique up to a multiplicative constant we will from now on also assume (without explicit reference) that the generator is normalized in the sense that $\varphi(\frac{1}{2}) = 1$ holds. With this convention there is a one-to-one correspondence between generators and induced copulas.

In what follows we let $D^+\varphi(x)$ ($D^-\varphi(x)$) denote the right-hand (left-hand) derivative of φ at $x \in (0,1)$. By convexity, φ is differentiable outside a countable subset of $(0,1)$, i.e. $D^+\varphi(x) = D^-\varphi(x)$ holds for all but at most countably many $x \in (0,1)$ and $D^+\varphi$ is non-decreasing and right-continuous. According to [16, 31] we additionally have $D^-\varphi(x) = D^+\varphi(x-)$ for every $x \in (0,1)$. Letting $\mathrm{Cont}(D^+\varphi) \subseteq (0,1)$ denote the set of all continuity points of $D^+\varphi$ in $(0,1)$ it follows that $[0,1] \setminus \mathrm{Cont}(D^+\varphi)$ is at most countably infinite and thus has Lebesgue measure 0. Setting $D^+\varphi(0) = -\infty$ in case of strict φ as well as $D^+\varphi(1) = 0$ (for strict and non-strict φ) allows to view $D^+\varphi$ as non-decreasing and right-continuous function on the full unit interval.
The Kendall distribution function of an Archimedean copula A with generator φ is given by (see, e.g., [13])

$$F_A^K(x) = x - \frac{\varphi(x)}{D^+\varphi(x)}. \tag{3.2}$$

Following [10, 27] for every $t \in [0,1]$ we define the t-level function $f^t : [t,1] \to [0,1]$ by $f^t(x) := \varphi^{-1}(\varphi(t) - \varphi(x))$ so that for $t > 0$

$$\mathrm{graph}(f^t) = \{(x, f^t(x)) : x \in [t,1]\} = \{(x,y) \in [0,1]^2 : C(x,y) = t\} =: L_t$$

holds since C is strictly increasing coordinatewise above $C^{-1}(\{0\})$. If φ is strict then according to [10]

$$K_A(x,[0,y]) = \begin{cases} \frac{D^+\varphi(x)}{D^+\varphi(A(x,y))} & \text{if } x \in (0,1) \\ 1 & \text{if } x \in \{0,1\} \end{cases} \tag{3.3}$$

is (a version of) the Markov kernel of A, for non-strict φ a version is given by

$$K_A(x,[0,y]) = \begin{cases} 0 & \text{if } x \in (0,1), y < f^0(x) \\ \frac{D^+\varphi(x)}{D^+\varphi(A(x,y))} & \text{if } x \in (0,1), y \geq f^0(x) \\ 1 & \text{if } x \in \{0,1\} \end{cases}. \tag{3.4}$$

According to [17] all three afore-mentioned modes of convergence coincide within $\mathcal{C}_{\text{ar}}$, that is, a sequence of Archimedean copulas $(A_n)_{n\in\mathbb{N}}$ converges uniformly to some $A \in \mathcal{C}_{\text{ar}}$ if and only if $(A_n)_{n\in\mathbb{N}}$ converges to A with respect to D_1 if and only if $(A_n)_{n\in\mathbb{N}}$ converges to A weakly conditional (as $n \to \infty$).

Archimedean copulas play a central role in our study since, together with the Fréchet-Hoeffding upper bound M, they form the building blocks of associative copulas: According to [23] (see also [1]) every associative copula C admits an ordinal sum representation in the following sense: there exists some index set $I \subseteq \mathbb{N}$, a partition $\{[a_i, b_i]\}_{i\in I}$ of $[0, 1]$ and $(A_i)_{i\in I}$ a sequence of copulas such that $A_i \in \mathcal{C}_{\text{ar}} \cup \{M\}$ and

$$C(x, y) = \begin{cases} a_i + (b_i - a_i)A_i\left(\frac{x-a_i}{b_i-a_i}, \frac{y-a_i}{b_i-a_i}\right) & \text{if } (x, y) \in (a_i, b_i)^2 \\ M(x, y) & \text{otherwise} \end{cases}.$$

In this case we write $C = (\langle a_i, b_i, A_i\rangle)_{i\in I}$. Notice that, contrary to literature, the above definition explicitly exposes (shrunk versions of) M along diagonal blocks. Defining the affine transformation $T_i : [a_i, b_i] \to [0, 1]$, $x \mapsto \frac{x-a_i}{b_i-a_i}$ and its inverse $T_i^{-1} : [0, 1] \to [a_i, b_i]$, $x \mapsto a_i + (b_i - a_i)x$ allows to write

$$C(x, y) = T_i^{-1}(A_i(T_i(x), T_i(y)))$$

for every $(x, y) \in (a_i, b_i)^2$. We will also work with the pre-image of T_i and also denote it by T_i^{-1} since no confusion will arise.
In this paper a *finite* associative copula is by definition an associative copula with finite index set I. Moreover, a subinterval (a_i, b_i) appearing in the ordinal sum representation of an associative copula C will also be simply called *subinterval of C*.

We start with the following Lemma stating that any non-strict Archimedean copula can be approximated arbitrarily well by strict Archimedean copulas with respect to D_1 and vice versa:

Theorem 3.1. *$\mathcal{C}_{ar}^{s}$ and $\mathcal{C}_{ar}^{\neg s}$ are dense in $(\mathcal{C}_{ar}, D_1)$.*

Proof. We only prove denseness of strict Archimedean copulas, the non-strict case follows in a similar manner. Fix $\varepsilon > 0$ and let $A_{\neg s}$ have non-strict generator $\varphi_{\neg s}$. In [19] it is shown that for every $n \in \mathbb{N}$ there exists a strict generator φ_s^n with $\varphi_s^n = \varphi_{\neg s}$ on $(\frac{1}{n}, 1]$ such that the corresponding strict Archimedean copula A_s^n fulfills $d_\infty(A_{\neg s}, A_s^n) \le \frac{1}{n}$. Letting $n \to \infty$ we obtain pointwise convergence of $\varphi_s^n \to \varphi_{\neg s}$ on $(0, 1]$ and using [17, Theorem 4.2] we even have weak conditional convergence of $(A_s^n)_{n\in\mathbb{N}}$ implying $D_1(A_s^n, A_{\neg s}) \to 0$ as $n \to \infty$. □

Next we show that an associative copula can be approximated arbitrarily well w.r.t. D_1 by a *finite* ordinal sum consisting only of strict Archimedean copulas or M.

Lemma 3.2. *Let C be an associative copula. Then for every $\varepsilon > 0$ there exist a finite index set J, a sequence $(A_j)_{j\in J}$ in $\mathcal{C}_{ar}^{s} \cup \{M\}$ and a partition $\{[a_j, e_j]\}_{j\in J}$ such that*

$$D_1(C, (\langle a_j, e_j, A_j\rangle)_{j\in J}) \le \varepsilon.$$

Proof. Fix $\varepsilon > 0$ and let $C = (\langle a_i, b_i, A_i\rangle)_{i\in I}$, where $I \subseteq \mathbb{N}$ is some finite or countably infinite index set. Choose a finite set $J \subseteq I$ such that $\sum_{j\in J} \lambda([a_j, e_j]) \ge 1 - \eta$ for some $\eta > 0$. Then $O := (\langle a_j, e_j, A_j\rangle)_{j\in J}$ is a finite ordinal sum of Archimedean copulas with

$$\begin{aligned} D_1(C, O) &= \int_{[0,1]^2} |K_C(x, [0, y]) - K_O(x, [0, y])| \, d\lambda_2(x, y) \\ &= \int_{\bigcup_{j\in J}[a_j,e_j]^2} |K_C(x, [0, y]) - K_O(x, [0, y])| \, d\lambda_2(x, y) \\ &\quad + \int_{\bigcup_{j\in I\setminus J}[a_j,e_j]^2} |K_C(x, [0, y]) - K_O(x, [0, y])| \, d\lambda_2(x, y) \\ &\le 0 + \eta. \end{aligned}$$

Choosing $\eta \in (0, \varepsilon)$ we obtain the assertion for arbitrary Archimedean copulas. Moreover, applying Theorem 3.1 to all those Archimedean components that are non-strict completes the proof. □

Loosely speaking, we found that selecting only finitely many 'large' components of an associative copula yields a finite associative copula that is close to the original one with respect to D_1. However, we do not necessarily have a finite ordinal sum purely consisting of Archimedean components as there can still exist subintervals (a_i, b_i) of C containing M in our approximation. The following lemma resolves this issue:

Lemma 3.3. *For every strict generator φ setting $\varphi_n := \varphi^n$ and letting A_n denote the Archimedean copula with generator φ_n yields* $\lim_{n\to\infty} D_1(A_n, M) = 0$.

Proof. We show that for generators of the form $\varphi_n(x) = \varphi(x)^n$ for some (strict and) continuously differentiable Archimedean generator φ we have convergence of the induced sequence $(A_n)_{n\in\mathbb{N}}$ to M w.r.t. D_1 (see Remark 3.4 for the proof based on a general (strict) generator φ):
Notice that $\varphi_n^{-1}(x) = \varphi^{-1}(\sqrt[n]{x})$, $\varphi_n'(x) = n\varphi(x)^{n-1}\varphi'(x) = n\varphi_{n-1}(x)\varphi'(x)$ and the corresponding copula is given by $A_n(x, y) = \varphi^{-1}(\sqrt[n]{\varphi(x)^n + \varphi(y)^n})$. It is well-known (see, e.g. [18, Proposition 8.5]) that the sequence $(A_n)_{n\in\mathbb{N}}$ converges uniformly to M. In fact, this follows from either of the following two relations:

$$A_n(x, x) = \varphi^{-1}(\sqrt[n]{2} \cdot \varphi(x)) \xrightarrow{n\to\infty} x,$$
$$F^K_{A_n}(x) = x - \frac{1}{n}\frac{\varphi(x)}{\varphi'(x)} \xrightarrow{n\to\infty} x.$$

The Markov kernel K_n corresponding to A_n is given by

$$K_n(x, [0, y]) = \begin{cases} \frac{\varphi'(x)}{\varphi'(A_n(x,y))} \cdot \frac{\varphi_{n-1}(x)}{\varphi_{n-1}(A_n(x,y))} & \text{if } x \in (0, 1) \\ 1 & \text{if } x \in \{0, 1\} \end{cases}.$$

First suppose $0 < x < y < 1$. As φ is continuously differentiable $\frac{\varphi'(x)}{\varphi'(A_n(x,y))}$ tends to 1 as $n \to \infty$. Concerning the second factor we have

$$\begin{aligned}
\frac{\varphi_{n-1}(x)}{\varphi_{n-1}(A_n(x, y))} &= \left(\frac{\varphi(x)}{\varphi(A_n(x, y))}\right)^{n-1} \cdot \frac{\varphi(x)}{\varphi(A_n(x, y))} \cdot \frac{\varphi(A_n(x, y))}{\varphi(x)} \\
&= \frac{\varphi_n(x)}{\varphi_n(A_n(x, y))} \cdot \frac{\varphi(A_n(x, y))}{\varphi(x)} \\
&= \frac{\varphi_n(x)}{\varphi_n(x) + \varphi_n(y)} \cdot \frac{\varphi(A_n(x, y))}{\varphi(x)} \\
&= \left(1 + \left(\frac{\varphi(y)}{\varphi(x)}\right)^n\right)^{-1} \cdot \sqrt[n]{1 + \left(\frac{\varphi(y)}{\varphi(x)}\right)^n} \xrightarrow{n\to\infty} 1 \cdot 1 = 1.
\end{aligned}$$

In case of $x > y$ we have $A_n(x, y) \to y$ for $n \to \infty$ and the first factor of the Markov kernel converges to the constant $\frac{\varphi'(x)}{\varphi'(y)}$ whereas for the second factor we have $\frac{\varphi(x)}{\varphi(A_n(x,y))} < 1$. Indeed, choose $\delta > 0$ such that $0 < y < x - \delta$ then there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ we have $A_n(x, y) < x - \delta$ which directly yields that for these n

$$\frac{\varphi(x)}{\varphi(A_n(x, y))} < \frac{\varphi(x)}{\varphi(x - \delta)} < 1$$

holds. As consequence the second factor converges to 0 as $n \to \infty$ which shows that for $x \in (0, 1)$ we have weak convergence of $K_n(x, \cdot) \xrightarrow{n\to\infty} K_M(x, \cdot)$. Hence $(A_n)_{n\in\mathbb{N}}$ even converges weakly conditional to M which implies converges w.r.t. D_1 (again see [36]). □

Remark 3.4.

1. Notice that the assumption of φ being continuously differentiable is not necessary for Lemma 3.3 to hold. In fact, for some general (strict) Archimedean generator we obtain $\varphi_n^{-1}(x) = \varphi^{-1}(\sqrt[n]{x})$ and $D^+\varphi_n(x) = n\varphi(x)^{n-1}D^+\varphi(x) = n\varphi_{n-1}(x)D^+\varphi(x)$. We have to show that there exists some $\Lambda \in \mathcal{B}([0, 1])$ of full measure

such that for every $x \in \Lambda$ and $y \in U_x \subseteq [0,1]$ with dense set U_x we have pointwise convergence of $(K_n(x,[0,y]))_{n\in\mathbb{N}}$ to $K_M(x,[0,y])$.
To this end, setting $\Lambda := \mathrm{Cont}(D^+\varphi)$ yields $\lambda(\Lambda) = 1$, and for $x \in \Lambda$ define $U_x := \bigcap_{n\in\mathbb{N}}\{u \in [0,1] : A_n(x,u) \in \mathrm{Cont}(D^+\varphi)\}$ it follows that U_x is the countable intersection of sets of full measure and as such has full measure itself. The rest of the proof is identical to the one of Lemma 3.3 above.

2. Notice that by Lemma 4.3 for $x > y$ we even have weak conditional convergence of general sequences of Archimedean copulas uniformly converging to M.

Example 3.5. The (normalized) generator $\varphi(x) = 1/x - 1$ produces the copula commonly abbreviated by $A_\varphi = \frac{\Pi}{\Sigma - \Pi}$ which is a member of numerous families of Archimedean copulas. Moreover, the generators $\varphi_n(x) = \varphi(x)^n = (1/x-1)^n$ define the family (4.2.12) in Nelsen [27] and the sequence of induced copulas $(A_n)_{n\in\mathbb{N}}$ not only converges uniformly to M, but also weakly conditional by Lemma 3.3.

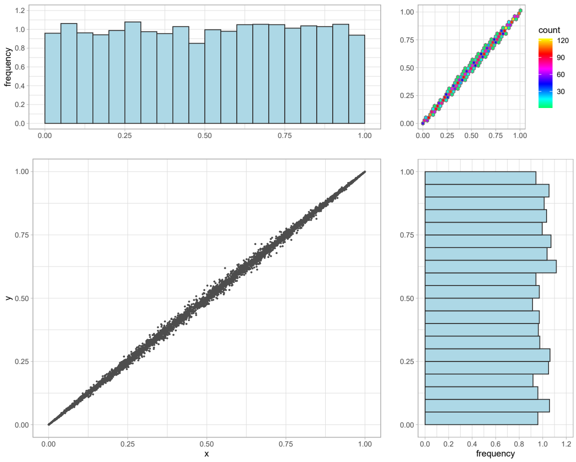

Figure 1: Sample of size 5000 from the copula A_n with $n = 40$ as considered in Example 3.5 as well as marginal histograms and bivariate histogram.

Theorem 3.6. *Let C be an associative copula. Then for every $\varepsilon > 0$ there is a finite ordinal sum A of strict Archimedean copulas such that*

$$D_1(C,A) \leq \varepsilon$$

holds.

Proof. Fix $\varepsilon > 0$. By Lemma 3.2 there exists a finite ordinal sum O (with strict Archimedean components) such that $D_1(C,O) < \frac{\varepsilon}{2}$ holds. Applying Lemma 3.3 we can find some strict $A_\varepsilon \in \mathcal{C}_{ar}$ with $D_1(A_\varepsilon, M) < \frac{\varepsilon}{2^{k+1}}$, where k denotes the size of the partition corresponding to O. Replacing every segment (along the diagonal) of O containing a shrunk version of M by a shrunk version of A_ε yields another finite ordinal sum A consisting only of (shrunk versions of strict) Archimedean copulas. Altogether we get $D_1(C,A) \leq D_1(C,O) + D_1(O,A) \leq \varepsilon$ which completes the proof. $\square$

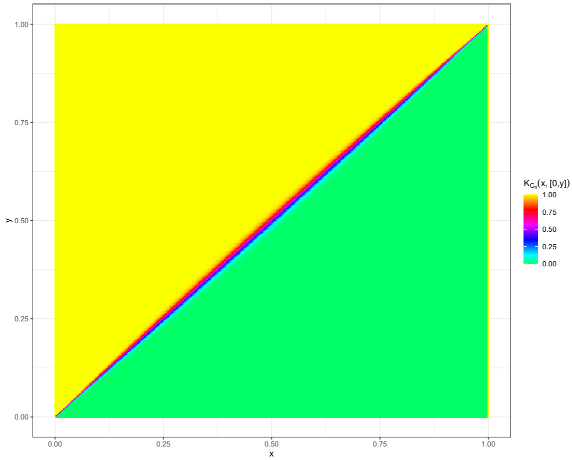

Figure 2: Heatmap of $(x, y) \mapsto K_{A_n}(x, [0, y])$ of the copula A_n for $n = 40$ as considered in Example 3.5.

As next step we show that the ordinal sum $C = (\langle 0, c, A_1\rangle, \langle c, 1, A_2\rangle)$ of two strict Archimedean copulas can be approximated arbitrarily well with respect to D_1 by *one* strict Archimedean copula. In [19] the authors define for every $\varepsilon > 0$ a strict Archimedean copula A_ε having the property that $d_\infty(A_\varepsilon, C) < \varepsilon$. We prove that for this particular glueing construction even $\lim_{\varepsilon\to 0} D_1(A_\varepsilon, C) \to 0$ holds. For each $\varepsilon > 0$ and $c \in (0, 1)$ setting

$$\varphi_\varepsilon(x) = \begin{cases} \varphi_1(T_1(x)) & \text{if } x \in [0, c - \frac{\varepsilon}{2}], \\ g(x) & \text{if } x \in (c - \frac{\varepsilon}{2}, x_0], \\ \frac{g(x_0)}{\varphi_2(T_2(x_0))} \cdot \varphi_2(T_2(x)) & \text{if } x \in (x_0, 1] \end{cases}$$

defines the generator of a strict Archimedean copula A_ε fulfilling $d_\infty(C, A_\varepsilon) < \varepsilon$. Here, φ_1, φ_2 are the generators of A_1, A_2 and T_1, T_2 are the transformations from $[0, c]$ and $[c, 1]$ into $[0, 1]$, respectively. The linear function g defined on $(c - \frac{\varepsilon}{2}, x_0]$ for some $x_0 < c + \frac{\varepsilon}{2}$ connects the two (appropriately transformed) generators of A_1, A_2 in a way such that the resulting map is again an Archimedean generator. We will make use of this construction, however, the upcoming proof is based on the order of the copulas C and A_ε in different sections of $[0, 1]^2$ rather than on the formula of φ_ε.

Example 3.7. Let $\varphi_1(x) = -\log(x)/\log(2)$ and $\varphi_2(x) = (1/x - 1)^{40}$ then the induced strict Archimedean copulas are Π and A_{40} from Example 3.5, respectively. Choosing the partition $\{[0, 0.4], [0.4, 1]\}$ we obtain an associative copula $C = (\langle 0, 0.4, A_1\rangle, \langle 0.4, 1, A_2\rangle)$. Following the construction in [19] and considering $\varepsilon > 0$ yields a strict Archimedean copula A_ε fulfilling $d_\infty(C, A_\varepsilon) < \varepsilon$.

Lemma 3.8. *Let $C = (\langle 0, c, A_1\rangle, \langle c, 1, A_2\rangle)$ be the ordinal sum of the two strict Archimedean copulas A_1, A_2 with respect to some $c \in (0, 1)$ and, for $\varepsilon > 0$, let A_ε be the strict Archimedean copula with the afore-mentioned generator φ_ε. Then*

$$D_1(C, A_\varepsilon) \le 4\varepsilon.$$

Proof. In order to simplify notation, within integrals we will (sometimes) set $F_C^x(y) = K_C(x, [0, y])$ and $F_\varepsilon^x(y) = K_\varepsilon(x, [0, y])$ for the Markov kernels of C and A_ε, respectively.

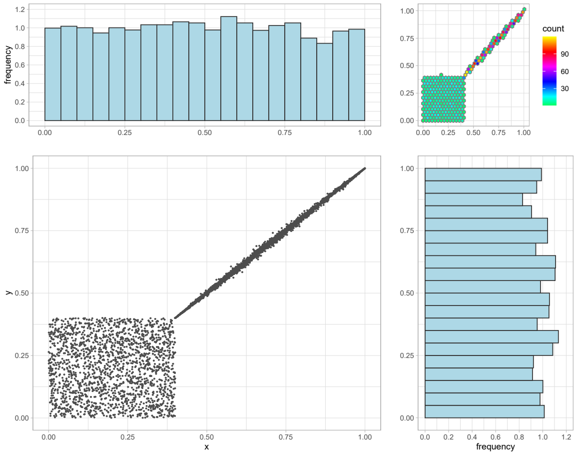

Figure 3: Sample of size 5000 from the Archimedean copula A_ε for $\varepsilon = 1/1000$ as considered in Example 3.7 as well as marginal histograms and two dimensional histogram.

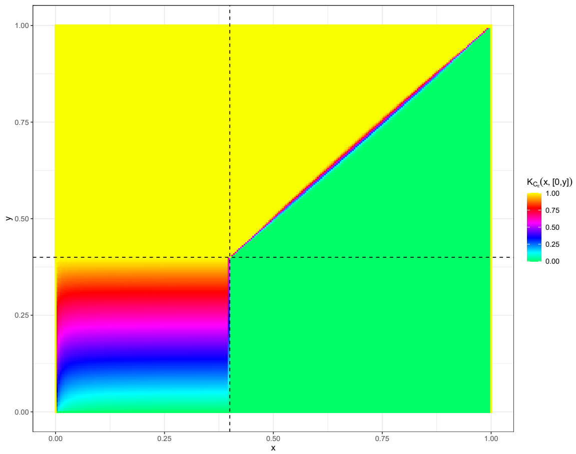

Figure 4: Heatmap of the Markov kernel $K_{A_\varepsilon}(x, [0, y])$ of the copula A_ε for $\varepsilon = 1/1000$ as considered in Example 3.7.

We start with the following segmentation of $[0,1]^2$ and estimate the absolute value within the central horizontal and vertical ε-strips very roughly:

$$\begin{aligned} D_1(C, A_\varepsilon) &= \int_{[0,1]^2} |K_C(x,[0,y]) - K_\varepsilon(x,[0,y])| \,\mathrm{d}\lambda_2(x,y) \\ &\leq \underbrace{\int_{[0,c-\frac{\varepsilon}{2}]^2} |F_C^x(y) - F_\varepsilon^x(y)| \,\mathrm{d}\lambda_2(x,y)}_{\mathrm{I}} + \underbrace{\int_{[c+\frac{\varepsilon}{2},1]^2} |F_C^x(y) - F_\varepsilon^x(y)| \,\mathrm{d}\lambda_2(x,y)}_{\mathrm{II}} \\ &\quad + \underbrace{\int_{[0,c-\frac{\varepsilon}{2}]\times[c+\frac{\varepsilon}{2},1]} |F_C^x(y) - F_\varepsilon^x(y)| \,\mathrm{d}\lambda_2(x,y)}_{\mathrm{III}} + \underbrace{\int_{[c+\frac{\varepsilon}{2},1]\times[0,c-\frac{\varepsilon}{2}]} |F_C^x(y) - F_\varepsilon^x(y)| \,\mathrm{d}\lambda_2(x,y)}_{\mathrm{IV}} \\ &\quad + \underbrace{\int_{(c-\frac{\varepsilon}{2},c+\frac{\varepsilon}{2})\times[0,1]} 1 \,\mathrm{d}\lambda_2(x,y)}_{=\varepsilon} + \underbrace{\int_{[0,1]\times(c-\frac{\varepsilon}{2},c+\frac{\varepsilon}{2})} 1 \,\mathrm{d}\lambda_2(x,y)}_{=\varepsilon}. \end{aligned}$$

The integrals I and II are identical to 0 as in the considered squares $A_\varepsilon = C$ and therefore $K_\varepsilon = K_C$ by Corollary 2.2. Regarding III, for every $(x,y) \in [0, c-\frac{\varepsilon}{2}]\times[c+\frac{\varepsilon}{2}, 1]$, we have $1 = K_C(x,[0,y]) \geq K_\varepsilon(x,[0,y])$ and $A_\varepsilon(x,y) \in (x-\varepsilon, x]$ so that $x = C(x,y) \geq A_\varepsilon(x,y)$. Thus,

$$\begin{aligned} \mathrm{III} &= \int_{[0,c-\frac{\varepsilon}{2}]\times[c+\frac{\varepsilon}{2},1]} K_C(x,[0,y]) - K_\varepsilon(x,[0,y]) \,\mathrm{d}\lambda_2(x,y) \\ &= \int_{[c+\frac{\varepsilon}{2},1]} C\big(c-\tfrac{\varepsilon}{2}, y\big) - A_\varepsilon\big(c-\tfrac{\varepsilon}{2}, y\big) \,\mathrm{d}\lambda(y) \ \leq \varepsilon. \end{aligned}$$

Regarding IV, for every $(x,y) \in [c+\frac{\varepsilon}{2}, 1]\times[0, c-\frac{\varepsilon}{2}]$, we have $0 = K_C(x,[0,y]) \leq K_\varepsilon(x,[0,y])$ and $A_\varepsilon(x,y) \in (y-\varepsilon, y]$ so that $y = C(x,y) \geq A_\varepsilon(x,y)$. As direct consequence,

$$\begin{aligned} \mathrm{IV} &= \int_{[c+\frac{\varepsilon}{2},1]\times[0,c-\frac{\varepsilon}{2}]} |K_C(x,[0,y]) - K_\varepsilon(x,[0,y])| \,\mathrm{d}\lambda_2(x,y) \\ &= \int_{(c+\frac{\varepsilon}{2},1]\times[0,c-\frac{\varepsilon}{2}]} K_\varepsilon(x,[0,y]) \,\mathrm{d}\lambda_2(x,y) \\ &= \int_{[0,c-\frac{\varepsilon}{2}]} y - A_\varepsilon\big(c+\tfrac{\varepsilon}{2}, y\big) \,\mathrm{d}\lambda(y) \ \leq \varepsilon. \end{aligned}$$

Summing up, we have shown $D_1(C, A_\varepsilon) \leq \varepsilon + \varepsilon + 2\varepsilon = 4\varepsilon$, which completes the proof. □

Combining Theorem 3.6 and Lemma 3.8 and using induction on the number of segments shows the following result:

Theorem 3.9. *$\mathcal{C}_{ar}$ is dense in $(\mathcal{C}_a, D_1)$.*

We conclude this section by answering two open problems posed in [6] regarding Baire category results for Archimedean copulas with respect to D_1. Following [6] or [30, Chapters 1 & 9] a subset S of a metric space (Ω, d) is called *nowhere dense* if the closure of S in (Ω, d) does not contain any non-empty open ball of (Ω, d). Moreover, a subset S of (Ω, d) is said to be *meager* or of *first category* if it is a countable union of nowhere dense sets. S is called *co-meager* if its complement $\Omega \setminus S$ is meager. Additionally, S is of *second category* if it is not meager. For more background on Baire categories we refer to [30].

In [6] the authors considered the metric spaces $(\mathcal{C}, d_\infty)$ and $(\mathcal{C}, D_1)$ and derived several Baire category results for different subclasses of $\mathcal{C}$. In particular, they showed that $\mathcal{C}_{\mathrm{ar}}^s$ is co-meager in $(\mathcal{C}_{\mathrm{ar}}, d_\infty)$ as well as that $(\mathcal{C}_{\mathrm{ar}}, d_\infty)$ is co-meager in $(\mathcal{C}_{\mathrm{a}}, d_\infty)$, and then conjectured that the same results should hold with respect to the stronger metric D_1. The proof of the first conjecture is a direct consequence of results from [17] whereas denseness of $\mathcal{C}_{\mathrm{ar}}$ in $(\mathcal{C}_{\mathrm{a}}, D_1)$ (as established above) is the key to confirm the second conjecture:

Corollary 3.10. *$\mathcal{C}^{s}_{ar}$ is co-meager in $(\mathcal{C}_{ar}, D_1)$.*

Proof. Since in $\mathcal{C}_{\mathrm{ar}}$ convergence with respect to d_∞ is equivalent to D_1-convergence both metrics induce the same topology on $\mathcal{C}_{\mathrm{ar}}$. It therefore follows immediately that $(\mathcal{C}_{\mathrm{ar}}, D_1)$ contains the same co-meager subsets as $(\mathcal{C}_{\mathrm{ar}}, d_\infty)$. In particular, $\mathcal{C}^{\mathrm{s}}_{\mathrm{ar}}$ is co-meager in $(\mathcal{C}_{\mathrm{ar}}, D_1)$. ☐

Corollary 3.11. *$\mathcal{C}_{ar}$ is co-meager (hence of second category) in $(\mathcal{C}_a, D_1)$.*

Proof. As mentioned before, denseness of $\mathcal{C}_{\mathrm{ar}}$ in $(\mathcal{C}_{\mathrm{a}}, D_1)$ is the missing piece to derive the desired result. For the sake of completeness we present the adapted proof (cf. [6, Theorem 2.4]) of the statement: We have to show that $\mathcal{C}_{\mathrm{a}} \setminus \mathcal{C}_{\mathrm{ar}}$ is the countable union of nowhere dense subsets. Setting for $k \in \mathbb{N}$

$$\mathcal{A}_k := \left\{ C \in \mathcal{C}_{\mathrm{a}} : \exists x \in \left[\frac{1}{k}, 1 - \frac{1}{k}\right] \text{ with } C(x, x) = x \right\}$$

we obviously have $\mathcal{C}_{\mathrm{a}} \setminus \mathcal{C}_{\mathrm{ar}} = \bigcup_{k\in\mathbb{N}} \mathcal{A}_k$. Notice that $\mathcal{A}_k$ is closed in $(C_{\mathrm{a}}, d_\infty)$ hence also in $(\mathcal{C}_{\mathrm{a}}, D_1)$. Considering that $\mathcal{C}_{\mathrm{ar}}$ is dense in $\mathcal{C}_{\mathrm{a}}$ w.r.t. D_1, $\mathcal{A}_k$ cannot contain any non-empty open ball of $\mathcal{C}_{\mathrm{a}}$, i.e. $\mathcal{A}_k$ is nowhere dense in $\mathcal{C}_{\mathrm{a}}$. As direct consequence $\mathcal{C}_{\mathrm{a}} \setminus \mathcal{C}_{\mathrm{ar}}$ is the union of nowhere dense subsets and the proof is complete. ☐

In this sense (see [3]), with respect to D_1 a typical associative copula is Archimedean and a typical Archimedean copula is strict.

4 Convergence in the class of associative copulas

In this section we prove the surprising result that within the class of associative copulas, convergence w.r.t. d_∞ is equivalent to convergence w.r.t. D_1.

As already mentioned for $C \in \mathcal{C}_{\mathrm{a}}$ there exist some finite or countably infinite index set $I \subseteq \mathbb{N}$, $\{(a_i, b_i)\}_{i\in I}$ a collection of subintervals forming a 'partition' of $[0, 1]$ and a sequence of copulas $(A_i)_{i\in I}$ with $A_i \in \mathcal{C}_{\mathrm{ar}}$ or $A_i = M$ such that

$$C(x, y) = T_i^{-1}(A_i(T_i(x), T_i(y)))$$

whenever $(x, y) \in (a_i, b_i)^2$ and $C(x, y) = M(x, y)$ otherwise; see Section 3. It follows that a Markov kernel of C is given by

$$K_C(x, [0, y]) = \begin{cases} K_{A_i}(T_i(x), [0, T_i(y)]) & \text{if } (x, y) \in (a_i, b_i)^2 \\ K_M(x, [0, y]) & \text{otherwise} \end{cases}.$$

Moreover, for $t \in (a_i, b_i)$ with $A_i \in \mathcal{C}_{\mathrm{ar}}$ the t-level curve of C restricted to $(a_i, b_i)^2$ is given by

$$f^t(x) = T_i^{-1}(f_i^{T_i(t)}(T_i(x))).$$

Note that the level curves of associative copulas are convex.

In the sequel we always assume that our points of interest lie in some subinterval (a_i, b_i) with $A_i \in \mathcal{C}_{\mathrm{ar}}$ and handle the subintervals of C containing M separately later on. Following [10] and defining the set $E^C_{s,t}$ for fixed $s, t \in [0, 1]$ by

$$E^C_{s,t} := \{(x, y) \in [0, 1]^2 : x \le s, C(x, y) \le t\}$$

it follows that for $s, t \in (a_i, b_i)$

$$\mu_C(E^C_{s,t}) = a_i + (b_i - a_i) \cdot \mu_{A_i}(E^{A_i}_{T_i(s),T_i(t)}) \tag{4.1}$$

holds and therefore the following formulas follow immediately: For $s, t \in (a_i, b_i)$ with $t \le s$ we have

$$\mu_C(E^C_{s,t}) = t + (b_i - a_i) \cdot \frac{\varphi_i(T_i(s)) - \varphi_i(T_i(t))}{D^+\varphi_i(T_i(t))}$$

whereas for $s > b_i > t > a_i$ we have

$$\mu_C(E_{s,t}) = t - (b_i - a_i) \cdot \frac{\varphi_i(T_i(t))}{D^+\varphi_i(T_i(t))}.$$

Denoting by L_t the t-level set we further have

$$\mu_C(L_t) = (b_i - a_i) \cdot \left[\frac{-\varphi_i(T_i(t))}{D^+\varphi_i(T_i(t))} + \frac{\varphi_i(T_i(t))}{D^-\varphi_i(T_i(t))}\right].$$

Moreover, setting $s = b_i$ immediately yields a representation for the Kendall distribution function

$$F^K_C(t) = T_i^{-1}(F^K_{A_i}(T_i(t)))$$

for $t \in (a_i, b_i)$ which will be of use in the sequel.

We now focus on convergence to M, or more generally $C \in \mathcal{C}_{\mathrm{cd}}$, and show the following surprising result:

Theorem 4.1. *Let $C_1, C_2, \ldots$ be copulas, $C \in \mathcal{C}_{cd}$ and suppose that $(C_n)_{n\in\mathbb{N}}$ converges uniformly to C. Then the following assertions hold:*

1. *$(C_n)_{n\in\mathbb{N}}$ converges to C w.r.t. D_1.*
2. *$(C_n^t * C_n)_{n\in\mathbb{N}}$ converges to M w.r.t. d_∞.*

Proof. Considering the D_2-distance and the definition of the star product (see Section 2) for arbitrary $A, B \in \mathcal{C}$ we have

$$\begin{aligned} D_2^2(A,B) &= \int_{[0,1]^2} (K_A(x,[0,y]) - K_B(x,[0,y]))^2 \, \mathrm{d}\lambda_2(x,y) \\ &= \int_{[0,1]} A^t * A(y,y) + B^t * B(y,y) - 2 \cdot A^t * B(y,y) \, \mathrm{d}\lambda(y). \end{aligned}$$

Considering that for $C \in \mathcal{C}_{\mathrm{cd}}$ the identity $C^t * C = M$ holds and that the star product is continuous in each argument with respect to d_∞ (see [8]) it follows that

$$\begin{aligned} 0 \le \limsup_{n\to\infty} D_2^2(C_n, C) &\le \limsup_{n\to\infty} \int_{[0,1]} C_n^t * C_n(s,s) \, \mathrm{d}\lambda(s) + \int_{[0,1]} M(s,s) \, \mathrm{d}\lambda(s) \\ &\quad + \limsup_{n\to\infty} \Big(-2 \cdot \int_{[0,1]} C_n^t * C(s,s) \, \mathrm{d}\lambda(s) \Big) \\ &= \limsup_{n\to\infty} \int_{[0,1]} \underbrace{C_n^t * C_n(s,s)}_{\le M(s,s)} \, \mathrm{d}\lambda(s) - \frac{1}{2} \le 0 \end{aligned}$$

implying $\lim_{n\to\infty} D_2^2(C_n, C) = 0$.
From the first part we already know that $\lim_{n\to\infty} \int_{[0,1]} C_n^t * C_n(s,s) \, \mathrm{d}\lambda(s) = \frac{1}{2}$. Letting δ_n denote the diagonal of $C_n^t * C$ and considering Lipschitz continuity of diagonals it follows that δ_n converges pointwise to the identity function $\mathrm{id}_{[0,1]}$. Letting B_{δ_n} denote the Bertino copula with diagonal δ_n (see [11, 12]) yields $B_{\delta_n} \le C_n^t * C_n$ as B_{δ_n} is the smallest symmetric copula with diagonal δ_n. Using the fact that $\delta_n \to \mathrm{id}_{[0,1]}$ we have $B_{\delta_n} \to M$ for $n \to \infty$. Hence considering $C_n^t * C_n \ge B_{\delta_n}$ it follows that $C_n^t * C_n \to M$ uniformly and the proof is complete. □

The second assertion is surprising in so far that the $*$-product is not jointly continuous w.r.t. d_∞ (but w.r.t. D_1).

The following Corollary, which is immediate from Theorem 4.1, states the promised equivalence of d_∞- and D_1-convergence whenever the limiting associative copula is the Fréchet-Hoeffding upper bound M:

Corollary 4.2. *Let $C_1, C_2, \ldots$ be associative copulas. Then the following assertions are equivalent:*

(a) $\lim_{n\to\infty} d_\infty(C_n, M) = 0$.
(b) $\lim_{n\to\infty} D_1(C_n, M) = 0$.

Lemma 4.3. *Let $A_1, A_2, \ldots$ be Archimedean copulas. If $(A_n)_{n\in\mathbb{N}}$ converges uniformly to M then for $x > y$ we have*

$$\lim_{n\to\infty} K_{A_n}(x, [0, y]) = K_M(x, [0, y]).$$

Proof. Recall that

$$\lim_{n\to\infty} \frac{\varphi_n(t)}{D^+\varphi_n(s)} = 0 \tag{4.2}$$

whenever $s, t \in (0, 1)$ with $s \le t$; see, e.g. [4, Proposition 5].
Now, fix $x > y$ and suppose that $\varepsilon > 0$ fulfills $y+\varepsilon < x$. Convexity of φ_n yields $0 \ge \varphi_n(y)-\varphi_n(x) \ge D^+\varphi_n(x)\cdot(y-x)$. Taking the absolute value and dividing by $|D^+\varphi_n(y)|$ we obtain

$$\frac{\varphi_n(y)}{|D^+\varphi_n(y)|} - \frac{\varphi_n(x)}{|D^+\varphi_n(y)|} \ge (x-y)\cdot\left|\frac{D^+\varphi_n(x)}{D^+\varphi_n(y)}\right|.$$

As direct consequence of (4.2) we have $\lim_{n\to\infty} \frac{D^+\varphi_n(x)}{D^+\varphi_n(y)} = 0$. Considering $\lim_{n\to\infty} A_n(x, y) = y$ there exists some $n_0 \in \mathbb{N}$ such that for all $n \ge n_0$ we have $A_n(x, y) \le y + \varepsilon$, hence

$$0 \le \lim_{n\to\infty} K_{A_n}(x, [0, y]) = \lim_{n\to\infty} \frac{D^+\varphi_n(x)}{D^+\varphi_n(A_n(x, y))} \le \lim_{n\to\infty} \frac{D^+\varphi_n(x)}{D^+\varphi_n(y+\varepsilon)} = 0.$$

□

Remark 4.4. As result of the above Lemma, in case of Archimedean copulas converging uniformly to M, we have weak conditional convergence whenever $x > y$ in $[0, 1]$. It remains an open question whether the result also holds above the diagonal (i.e. for $x < y$) which would be key for proving that within $\mathcal{C}_a$ we even have equivalence of d_∞-convergence, D_1-convergence *and* weak conditional convergence.

As next step we consider a sequence $(A_n)_{n\in\mathbb{N}}$ of Archimedean copulas uniformly converging to some associative copula C whereby we focus on studying convergence within Archimedean blocks of C. That is, we focus on $(x, y) \in (a_i, b_i)^2$, for some $i \in I$, with corresponding copula $A_i \in \mathcal{C}_{ar}$ in the 'partition' $\{[a_i, b_i]\}_{i\in I}$ underlying our associative limit C. However, since it is not necessary that A_i appears in the sequence $(A_n)_{n\in\mathbb{N}}$ we deviate from this notation; subsequently we denote by $(a_\star, e_\star)$ the subinterval of interest corresponding to its Archimedean content $A_\star$ with generator $\varphi_\star$ in the ordinal sum representation of C.

Lemma 4.5. *Let $A_1, A_2, \ldots$ be Archimedean copulas with generators $\varphi_1, \varphi_2, \ldots$, respectively and let $C \in \mathcal{C}$ be an associative copula such that $(A_n)_{n\in\mathbb{N}}$ converges uniformly to C. Then for every subinterval $(a_\star, b_\star)$ of C containing an Archimedean copula the following assertions hold:*

1. $\lim_{n\to\infty} F^K_{A_n}(x) = F^K_C(x)$ *for all* $x \in (a_\star, b_\star) \cap T_\star^{-1}(Cont(D^+\varphi_\star))$,
2. $\lim_{n\to\infty} \frac{\varphi_n(x)}{D^+\varphi_n(x)} = (b_\star - a_\star)\cdot \frac{\varphi_\star(T_\star(x))}{D^+\varphi_\star(T_\star(x))}$ *for all* $x \in (a_\star, b_\star) \cap T_\star^{-1}(Cont(D^+\varphi_\star))$,
3. $\lim_{n\to\infty} \frac{\varphi_n(x)}{\varphi_n(y)} = \frac{\varphi_\star(T_\star(x))}{\varphi_\star(T_\star(y))}$ *for all* $x, y \in (a_\star, b_\star)$,

4. $\lim_{n\to\infty} \frac{\varphi_n(x)}{D^+\varphi_n(y)} = (b_\star - a_\star) \cdot \frac{\varphi_\star(T_\star(x))}{D^+\varphi_\star(T_\star(y))}$ *for all* $x \in (a_\star, b_\star)$, $y \in (a_\star, b_\star) \cap T_\star^{-1}(Cont(D^+\varphi_\star))$,
5. $\lim_{n\to\infty} \frac{D^+\varphi_n(x)}{D^+\varphi_n(y)} = \frac{D^+\varphi_\star(T_\star(x))}{D^+\varphi_\star(T_\star(y))}$ *for all* $x, y \in (a_\star, b_\star) \cap T_\star^{-1}(Cont(D^+\varphi_\star))$.

Proof. It is well-known that $\lim_{n\to\infty} d_\infty(A_n, C) = 0$ implies $F^K_{A_n}(x) \to F^K_C(x)$ pointwise in continuity points of F^K_C. Hence, (1) is a direct consequence of the before established fact that $F^K_C(x) = T_\star^{-1}(F^K_{A_\star}(T_\star(x)))$ and (2) follows immediately. According to [13], an Archimedean generator φ fulfills

$$\varphi(x) = \varphi(y) \cdot \exp\left(\operatorname{sign}(x-y) \int\limits_{\min(x,y)}^{\max(x,y)} \frac{D^+\varphi(t)}{\varphi(t)} \mathrm{d}t\right)$$

for $x \in [0, 1]$ and $y \in (0, 1)$. Therefore,

$$\begin{aligned}\lim_{n\to\infty} \frac{\varphi_n(x)}{\varphi_n(y)} &= \lim_{n\to\infty} \exp\left(\operatorname{sign}(x-y) \int\limits_{\min(x,y)}^{\max(x,y)} \frac{D^+\varphi_n(t)}{\varphi_n(t)} \mathrm{d}t\right) \\ &= \exp\left(\operatorname{sign}(x-y) \int\limits_{\min(x,y)}^{\max(x,y)} \lim_{n\to\infty} \frac{D^+\varphi_n(t)}{\varphi_n(t)} \mathrm{d}t\right),\end{aligned}$$

where the exchange of limit and integral is justified in [4, Proposition 2]. Considering $x > y$, property (2) and using change of coordinates yields

$$\lim_{n\to\infty} \frac{\varphi_n(x)}{\varphi_n(y)} = \exp\left(\int\limits_{[y,x]} \frac{1}{b_\star - a_\star} \cdot \frac{D^+\varphi_\star(T_\star(t))}{\varphi_\star(T_\star(t))} \mathrm{d}t\right) = \frac{\varphi_\star(T_\star(x))}{\varphi_\star(T_\star(y))},$$

the case $x < y$ follows similarly. This shows (4) and (5) is a consequence of properties (2) and (4). □

We now show that for a sequence of Archimedean copulas converging to an associative limit C, within the Archimedean blocks of C we have weak convergence of the corresponding Markov kernels on a dense set above the level curve $f^{a_\star}$ of C. Notice that $f^{a_\star}(x) = T_\star^{-1}(f^0_\star(T_\star(x)))$, $x \in (a_\star, b_\star)$ is the $a_\star$-level curve of C. Hence the following Theorem specifically includes the case of strict Archimedean blocks in C whereas non-strict blocks are tackled in Theorem 4.9.

Theorem 4.6. *Let $A_1, A_2, \dots$ be Archimedean copulas with generators $\varphi_1, \varphi_2, \dots$ and let C be an associative copula such that $(A_n)_{n\in\mathbb{N}}$ converges uniformly to C. Then for every subinterval $(a_\star, b_\star)$ of C containing an Archimedean copula $A_\star \in \mathcal{C}_{ar}$ there exists some $\Lambda \in \mathcal{B}((a_\star, b_\star))$ of full measure and for $x \in \Lambda$ some set $U_x \subseteq (f^{a_\star}(x), b_\star)$ which is dense in $(f^{a_\star}(x), b_\star)$ such that for every $x \in \Lambda$ and $y \in U_x$ we have*

$$\lim_{n\to\infty} K_{A_n}(x, [0, y]) = K_{A_\star}(T_\star(x), [0, T_\star(y)]) = K_C(x, [0, y]).$$

Proof. Suppose that $x \in (a_\star, b_\star)$. Then considering $y = T_\star^{-1}(1/2) \in (a_\star, b_\star)$ property (3) of Lemma 4.5 yields

$$\lim_{n\to\infty} \frac{\varphi_n(x)}{\varphi_n(T_\star^{-1}(\frac{1}{2}))} = \varphi_\star(T_\star(x)).$$

Set $\Lambda := T_\star^{-1}(\mathrm{Cont}(D^+\varphi_\star))$ then Λ is of full measure in $(a_\star, b_\star)$ and for $x \in \Lambda$ convexity (of $\varphi_\star(T_\star(x))$) implies (see, e.g., [21, 32])

$$\lim_{n\to\infty} \frac{1}{\varphi_n(T_\star^{-1}(\frac{1}{2}))} D^+\varphi_n(x) = T'_\star(x) D^+\varphi_\star(T_\star(x)),$$

whereby the chain rule is applicable since T is differentiable and increasing. For $x \in \Lambda$ the set $U_x := \{v \in (f^{a_\star}(x), b_\star) : T_\star(C(x, v)) \in \mathrm{Cont}(D^+\varphi_\star)\}$ is of full measure. In fact,

$$\begin{aligned}U_x &= \{v \in (f^{a_\star}(x), b_\star) : A_\star(T_\star(x), T_\star(v)) \in \mathrm{Cont}(D^+\varphi_\star)\} \\ &= (h^x_\star)^{-1}(\mathrm{Cont}(D^+\varphi_\star)),\end{aligned}$$

where $h^x_\star(v) := A_\star(T_\star(x), T_\star(v))$ is an increasing homeomorphism from $(f^{a_\star}(x), b_\star)$ to $[0, 1]$ (cf. [27]). Thus, $(h^x_\star)^{-1}(\mathrm{Cont}(D^+\varphi_\star)^\complement)$ is at most countably infinite whence $\lambda((h^x_\star)^{-1}(\mathrm{Cont}(D^+\varphi_\star))) = b_\star - f^{a_\star}(x)$.
Letting $x \in \Lambda$, $y \in U_x$ and invoking continuous convergence of the right-hand derivatives yields

$$\lim_{n\to\infty} \frac{1}{\varphi_n(T_\star^{-1}(\frac{1}{2}))} D^+\varphi_n(A_n(x, y)) = T'_\star(C(x, y))D^+\varphi_\star(T_\star(C(x, y)))$$

and we directly obtain

$$\begin{aligned}\lim_{n\to\infty} K_{A_n}(x, [0, y]) &= \lim_{n\to\infty} \frac{D^+\varphi_n(x)}{D^+\varphi_n(A_n(x, y))} \\ &= \lim_{n\to\infty} \frac{\frac{D^+\varphi_n(x)}{\varphi_n(T_\star^{-1}(\frac{1}{2}))}}{\frac{D^+\varphi_n(A_n(x,y))}{\varphi_n(T_\star^{-1}(\frac{1}{2}))}} \\ &= \frac{T'_\star(x)D^+\varphi_\star(T_\star(x))}{T'_\star(C(x, y))D^+\varphi_\star(T_\star(C(x, y)))} \\ &= \frac{D^+\varphi_\star(T_\star(x))}{D^+\varphi_\star(A_\star(T_\star(x), T_\star(y)))} = K_C(x, [0, y])\end{aligned}$$

since obviously $T'_\star(x) = T'_\star(C(x, y)) = \frac{1}{b_\star - a_\star}$. □

According to results of [36] convergence with respect to D_1 (within $(a_\star, e_\star)$) follows immediately.
Since every Archimedean copula is an ordinal sum with only trivial idempotent elements, the convergence results from Lemma 4.5 and Theorem 4.6 can be seen as generalizations of [17, Theorem 4.2]:

Theorem 4.7. *Let $A, A_1, A_2, \ldots$ be Archimedean copulas with generators $\varphi, \varphi_1, \varphi_2, \ldots$ such that $(A_n)_{n\in\mathbb{N}}$ converges uniformly to A. Then there exists some $\Lambda \in \mathcal{B}((0, 1))$ of full measure and some set $U_x \subseteq (f^0(x), 1)$ which is dense in $(f^0(x), 1)$ such that for every $x \in \Lambda$ and $y \in U_x$ we have*

$$\lim_{n\to\infty} K_{A_n}(x, [0, y]) = K_A(x, [0, y]).$$

Remark 4.8. At this point we have a very strong convergence property for the following special case of convergence within $\mathcal{C}_\mathrm{a}$: Let $A_1, A_2, \ldots$ be a sequence of Archimedean copulas converging uniformly to some finite associative copula C having no sections containing M, i.e. a finite ordinal sum purely consisting of Archimedean copulas. Then $(A_n)_{n\in\mathbb{N}}$ converges weakly conditional to C. This includes the frequently presented case of Archimedean copulas converging to ordinal sums of W (see, e.g., [4],[19]), also called “prototypes” in [34].

Using the afore-mentioned results we can now finally prove the equivalence of convergence with respect to d_∞ and convergence with respect to D_1 within the class of associative copulas:

Theorem 4.9. *Let $C, C_1, C_2, \ldots$ be associative copulas. If the sequence $(C_n)_{n\in\mathbb{N}}$ converges uniformly to C then the sequence even converges with respect to D_1.*

Proof. As proved in Section 3 the class $\mathcal{C}^\mathrm{s}_\mathrm{ar}$ of strict Archimedean copulas is dense in $(\mathcal{C}_\mathrm{a}, D_1)$. Consequently, for every C_n there exists some $A_n \in \mathcal{C}^\mathrm{s}_\mathrm{ar}$ such that $D_1(C_n, A_n) < \frac{1}{n}$. Thus,

$$D_1(C_n, C) \le D_1(C_n, A_n) + D_1(A_n, C) < \frac{1}{n} + D_1(A_n, C)$$

and it suffices to show that $D_1(A_n, C) \to 0$ for $n \to \infty$ which can be done as follows: According to [6, Lemma 3] $d_\infty(A_n, C) \le 2 \cdot \sqrt{D_1(A_n, C_n)} + d_\infty(C_n, C)$ and the latter converges to 0 for $n \to \infty$ whence $(A_n)_{n\in\mathbb{N}}$ converges uniformly to C. Write $C = (\langle a_i, b_i, B_i\rangle)_{i\in I}$ for some finite or countably infinite index set I. Fix $\varepsilon > 0$ then as in Lemma 3.2 we can choose a finite set $J \subseteq I$ such that $\sum_{j\in J} \lambda((a_j, b_j)) \ge 1 - \varepsilon$ and consider the following partition of $[0, 1]^2$ (as before we set $F^x_n(y) = K_n(x, [0, y]) = K_{A_n}(x, [0, y])$ and $\mu_n = \mu_{A_n}$, similarly for the limit C):

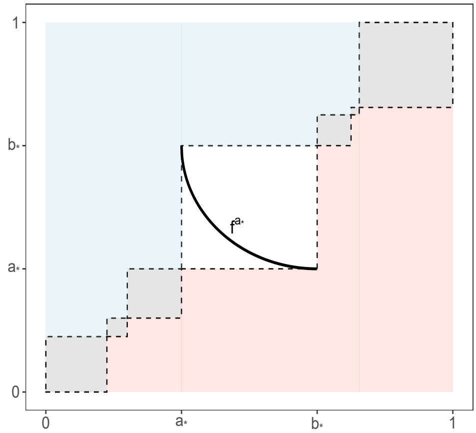

Figure 5: Illustration of the (support of the) associative limit C with $a_\star$-level curve $f^{a_\star}(x)$ for $x \in (a_\star, b_\star)$. The regions U and L considered in the proof of Theorem 4.9 are colorized in light blue and light red, respectively.

$$D_1(A_n, C) = \underbrace{\int\limits_{\bigcup_{j\in J}(a_j,b_j)^2} |F_n^x(y) - F_C^x(y)| \mathrm{d}\lambda_2(x,y)}_{\mathrm{I}} + \underbrace{\int\limits_{\bigcup_{i\in I\setminus J}(a_i,b_i)^2} |F_n^x(y) - F_C^x(y)| \mathrm{d}\lambda_2(x,y)}_{\mathrm{II}}$$
$$+ \underbrace{\int\limits_{[0,1]^2\setminus\bigcup_{i\in I}(a_i,b_i)^2} |F_n^x(y) - F_C^x(y)| \mathrm{d}\lambda_2(x,y)}_{\mathrm{III}}.$$

Clearly II $\le \varepsilon$; for calculating III divide the integration area in the sets L and U denoting the lower and upper part, respectively, i.e. $U = \bigcup_{i\in I}(a_i, b_i) \times (b_i, 1)$ and $L = \bigcup_{i\in I}(a_i, b_i) \times (0, a_i)$ (as depicted in Figure 5). Using disintegration we get

$$\begin{aligned}\int\limits_U |K_n(x,[0,y]) - K_C(x,[0,y])| \mathrm{d}\lambda_2(x,y) &= \int\limits_U 1 - K_n(x,[0,y])\, \mathrm{d}\lambda_2(x,y)\\
&= \int\limits_U K_n(x,(y,1])\, \mathrm{d}\lambda_2(x,y)\\
&= \sum_{i\in I} \int\limits_{(b_i,1)} \underbrace{\mu_n((a_i,b_i)\times(y,1])}_{\le \mu_n((a_i,b_i)\times(b_i,1))}\, \mathrm{d}\lambda(y)\\
&\le \sum_{i\in I} \mu_n((a_i,b_i)\times(b_i,1)) \cdot (1 - b_i)\\
&\le \sum_{i\in I} \mu_n((a_i,b_i)\times(b_i,1)) = \mu_n(U)\end{aligned}$$

and as U is a μ_C-continuity set it follows that $\lim_{n\to\infty} \mu_n(U) = \mu_C(U) = 0$. The lower area L follows analogously whence there is $n_0 \in \mathbb{N}$ such that III $\le \varepsilon$ for all $n \ge n_0$.
It remains to show that the same holds true for I: We split the index set J into J_1 denoting the subset of indices

corresponding to blocks containing an Archimedean copula and J_2 corresponding to blocks containing M, i.e.

$$\mathrm{I} = \underbrace{\int\limits_{\bigcup_{j\in J_1}(a_j,b_j)^2} |F_n^x(y) - F_C^x(y)| \, \mathrm{d}\lambda_2(x,y)}_{\mathrm{I.I}} + \underbrace{\int\limits_{\bigcup_{j\in J_2}(a_j,b_j)^2} |F_n^x(y) - F_M^x(y)| \, \mathrm{d}\lambda_2(x,y)}_{\mathrm{I.II}}.$$

We begin with the Archimedean case I.I: To facilitate reading we focus on a single subinterval $(a_\star, b_\star) := (a_j, b_j)$ for some $j \in J_1$ containing an Archimedean copula. Then

$$\begin{aligned}
&\int\limits_{(a_\star,b_\star)^2} |F_n^x(y) - F_C^x(y)| \, \mathrm{d}\lambda(x)\mathrm{d}\lambda(y) \\
&= \int\limits_{(a_\star,b_\star)} \int\limits_{\{x<f^{a_\star}(y)\}} |F_n^x(y) - F_C^x(y)| \, \mathrm{d}\lambda(x)\mathrm{d}\lambda(y) + \int\limits_{(a_\star,b_\star)} \int\limits_{\{x\geq f^{a_\star}(y)\}} |F_n^x(y) - F_C^x(y)| \, \mathrm{d}\lambda_2(x,y),
\end{aligned}$$

where for fixed $y \in (a_\star, b_\star)$ the set $\{x < f^{a_\star}(y)\} = \{x \in (a_\star, b_\star) : x < f^{a_\star}(y)\} = (a_\star, f^{a_\star}(y))$. The first integrand fulfills $F_C^x(y) = K_C(x, [0, y]) = 0$ and we get

$$\int\limits_{(a_\star,b_\star)} \int\limits_{\{x<f^{a_\star}(y)\}} K_n(x, [0, y]) \, \mathrm{d}\lambda_2(x,y) = \int\limits_{(a_\star,b_\star)} \mu_n((a_\star, f^{a_\star}(y)) \times [0, y]) \, \mathrm{d}\lambda(y)$$

which converges to $\int_{(a_\star,b_\star)} \mu_C((a_\star, f^{a_\star}(y)) \times [0, y]) \, \mathrm{d}\lambda(y) = 0$ for $n \to \infty$ by Dominated Convergence. For the second summand we have

$$\int\limits_{(a_\star,b_\star)} \int\limits_{\{x\geq f^{a_\star}(y)\}} |F_n^x(y) - F_C^x(y)| \, \mathrm{d}\lambda_2(x,y) = \int\limits_{(a_\star,b_\star)} \int\limits_{\{y\geq f^{a_\star}(x)\}} |F_n^x(y) - F_C^x(y)| \, \mathrm{d}\lambda_2(y,x).$$

Observe that now we are in the setting of Theorem 4.6 and hence also have convergence to 0 for $n \to \infty$.
To calculate I.II we proceed similarly, fix a subinterval $(\bar{a}_\star, \bar{b}_\star) := (a_j, b_j)$ for some $j \in J_2$ of C containing a copy of M and consider

$$\begin{aligned}
&\int\limits_{(\bar{a}_\star,\bar{b}_\star)^2} |K_n(x, [0, y]) - K_M(x, [0, y])| \, \mathrm{d}\lambda_2(x,y) \\
&\qquad = \int\limits_{(\bar{a}_\star,\bar{b}_\star)} \Big\{ \int\limits_{(\bar{a}_\star,y]} 1 - K_n(x, [0, y]) \, \mathrm{d}\lambda(x) + \int\limits_{[y,\bar{b}_\star)} K_n(x, [0, y])\mathrm{d}\lambda(x) \Big\} \mathrm{d}\lambda(y) \\
&\qquad = \int\limits_{(\bar{a}_\star,\bar{b}_\star)} \mu_n((\bar{a}_\star, y] \times (y, 1]) + \mu_n([y, \bar{b}_\star) \times [0, y]) \, \mathrm{d}\lambda(y) \\
&\qquad \to \int\limits_{(\bar{a}_\star,\bar{b}_\star)} \mu_C((\bar{a}_\star, y] \times (y, 1]) + \mu_C([y, \bar{b}_\star) \times [0, y]) \, \mathrm{d}\lambda(y) = 0
\end{aligned}$$

for $n \to \infty$ (again using Dominated Convergence). It directly follows that there exists $n_1 \in \mathbb{N}$ such that for all $n \geq n_1$ we have $\mathrm{I} \leq \varepsilon$.
Altogether, there exists $n_2 \in \mathbb{N}$ such that for all $n \geq n_2 = \max(n_0, n_1)$ we have

$$D_1(A_n, C) = \mathrm{I} + \mathrm{II} + \mathrm{III} \leq 3 \cdot \varepsilon,$$

which completes the proof. □

Acknowledgements
The first author gratefully acknowledges the financial support from Porsche Holding Austria and Land Salzburg within the WISS 2025 project KFZ (P1900123). Moreover, the second and the third author gratefully acknowledge the support of the WISS 2025 project 'IDA-lab Salzburg' (20204-WISS/225/197-2019 and 0102-F1901166-KZP).

Conflict of interest statement: Authors state no conflict of interest.

References

[1] Alsina, C., B. Schweizer, and M. Frank (2006). *Associative Functions*. World Scientific Publishing, Singapore.
[2] Billingsley, P. (1999). *Convergence of Probability Measures*. Second edition. John Wiley & Sons, New York.
[3] Bruckner, A. M, J. B. Bruckner, and B. S. Thomson (2008). *Real Analysis*. Second edition. Available at http://classicalrealanalysis.info/com/Real-Analysis.php.
[4] Charpentier, A. and J. Segers (2008). Convergence of Archimedean copulas. *Statist. Probab. Lett. 78*(4), 412–419.
[5] Darsow, W. F., B. Nguyen, and E. T. Olsen (1992). Copulas and Markov processes. *Illinois J. Math. 36*(4), 600–642.
[6] Durante, F., J. Fernández-Sánchez, and W. Trutschnig (2016). Baire category results for exchangeable copulas. *Fuzzy Sets Syst. 284*, 146–151.
[7] Durante, F. and P. Jaworski (2010). A New Characterization of Bivariate Copulas. *Comm. Statist. Theory Methods 39*, 2901–2912.
[8] Durante, F. and C. Sempi (2016). *Principles of Copula Theory*. CRC Press, Boca Raton FL.
[9] Fernández-Sánchez, J. and W. Trutschnig (2015a). Conditioning based metrics on the space of multivariate copulas and their interrelation with uniform and levelwise convergence and Iterated Function Systems. *J. Theoret. Probab. 28*, 1311–1336.
[10] Fernández-Sánchez, J. and W. Trutschnig (2015b). Singularity aspects of Archimedean copulas. *J. Math. Anal. Appl. 432*(1), 103–113.
[11] Fernández-Sánchez, J. and W. Trutschnig (2016). Some members of the class of (quasi-)copulas with given diagonal from the Markov kernel perspective. *Comm. Statist. Theory Methods 45*(5), 1508–1526.
[12] Fredricks, G. A. and R. B. Nelsen (2002). The Bertino family of copulas. In Cuadras C.M., Fortiana J., Rodriguez-Lallena J.A. (Eds.), *Distributions With Given Marginals and Statistical Modelling*, pp. 81–91. Springer, Dordrecht.
[13] Genest, C. and L.-P. Rivest (1993). Statistical inference procedures for bivariate Archimedean copulas. *J. Amer. Statist. Assoc. 88*(423), 1034–1043.
[14] Griessenberger, F., J. Fernández-Sánchez, and W. Trutschnig (2021). Some properties of double shuffles of bivariate copulas and (extreme) copulas invariant with respect to Lüroth double shuffles. *Fuzzy Sets Syst.*, to appear. Available at https://doi.org/10.1016/j.fss.2021.02.014.
[15] Kallenberg, O. (2002). *Foundations of Modern Probability*. Springer, New York.
[16] Kannan, R. and C. K. Krueger (1996). *Advanced Analysis*. Springer, New York.
[17] Kasper, T., S. Fuchs, and W. Trutschnig (2020). On weak conditional convergence of bivariate Archimedean and extreme value copulas, and consequences to nonparametric estimation. *Bernoulli*, to appear. Available at http://www.bernoulli-society.org/index.php/publications/bernoulli-journal/bernoulli-journal-papers.
[18] Klement, E. P., R. Mesiar, and E. Pap (2000). *Triangular Norms*. Springer, Dordrecht.
[19] Klement, E. P., R. Mesiar, and E. Pap (2001). Uniform approximation of associative copulas by strict and non-strict copulas. *Illinois J. Math. 45*(4), 1393–1400.
[20] Klenke, A. (2008). *Wahrscheinlichkeitstheorie*. Springer, Berlin.
[21] Kosmol, P. (2010). *Optimierung und Approximation*. De Gruyter, Berlin.
[22] Lancaster, H. O. (1963). Correlation and complete dependence of random variables. *Ann. Math. Statist. 34*(4), 1315–1321.
[23] Ling, C.-H. (1965). Representation of associative functions. *Publ. Math. Debrecen 12*, 189–212.
[24] Li, X., P. Mikusinski, M.D. Taylor (1998). Strong approximation of copulas. *J. Math. Anal. Appl. 255*(2), 608-623
[25] Moynihan, R. (1978). On τ_T semigroups of probability distribution functions II. *Aequationes Math. 17*, 19–40.
[26] Mroz, T., S. Fuchs, W. Trutschnig (2021). How simplifying and flexible is the simplifying assumption in pair-copula constructions - analytic answers in dimension three and a glimpse beyond. *Electron. J. Stat. 15*, 1951-1992
[27] Nelsen, R. B. (2006). *An Introduction to Copulas*. Second edition. Springer, New York.
[28] Nelsen, R. B., J. J. Quesada-Molina, J. A. Rodríguez-Lallena, and M. Úbeda Flores (2009). Kendall distribution functions and associative copulas. *Fuzzy Sets Syst. 160*(1), 52–57.
[29] Olsen, E.T., W.F. Darsow, B. Nguyen (1996). Copulas and Markov operators, in L. Rüschendorf, B. Schweizer, and M.D. Taylor (Eds.), *Distributions with Fixed Marginals and Related Topics*, pp. 244-259. Institute od Mathematical Statistics, Hayward CA.

[30] Oxtoby, J. C. (1980). *Measure and Category*. Springer, New York.
[31] Pollard, D. (2001). *A User's Guide to Measure Theoretic Probability*. Cambridge University Press.
[32] Rockafeller, R.T. (1972). *Convex Analysis*. Princeton University Press.
[33] Rudin, W. (1976). *Principles of Mathematical Analysis*. McGraw-Hill, Auckland.
[34] Schreyer, M., R. Paulin, and W. Trutschnig (2017). On the exact region determined by Kendall's τ and Spearman's ρ. *J. R. Stat. Soc. Ser. B Stat. Methodol. 79*(2), 613–633.
[35] Spanhel, F. and M. Kurz (2019). Simplified vine copula models: Approximations based on the simplifying assumption. *Electron. J. Stat. 13*, 1254-1291.
[36] Trutschnig, W. (2011). On a strong metric on the space of copulas and its induced dependence measure. *J. Math. Anal. Appl. 384*(2), 690–705.
[37] Trutschnig, W. (2012). Some results on the convergence of (quasi-) copulas. *Fuzzy Sets Syst. 191*, 113–121.
[38] Trutschnig, W. and J. Fernández-Sánchez (2012). Idempotent and multivariate copulas with fractal support. *J. Statist. Plann. Inference 142*(12), 3086–3096.

DE GRUYTER Depend. Model. 2021; 9:327–346

Research Article **Open Access**
Special Issue in memory of Abe Sklar

Dietmar Pfeifer* and Olena Ragulina

Generating unfavourable VaR scenarios under Solvency II with patchwork copulas

https://doi.org/10.1515/demo-2021-0115
Received November 12, 2020; accepted August 23, 2021

Abstract: The central idea of the paper is to present a general simple patchwork construction principle for multivariate copulas that create unfavourable VaR (i.e. Value at Risk) scenarios while maintaining given marginal distributions. This is of particular interest for the construction of Internal Models in the insurance industry under Solvency II in the European Union. Besides this, the Delegated Regulation by the European Commission requires all insurance companies under supervision to consider different risk scenarios in their risk management system for the company's own risk assessment. Since it is unreasonable to assume that the potential worst case scenario will materialize in the company, we think that a modelling of various unfavourable scenarios as described in this paper is likewise appropriate. Our explicit copula approach can be considered as a special case of ordinal sums, which in two dimensions even leads to the technically worst VaR scenario.

Keywords: Solvency II, copulas, patchwork copulas, Bernstein copulas, Monte Carlo methods

MSC: 62H05, 62H12, 62H17, 11K45

1 Introduction

Reasonable VaR-estimates from original data or suitable scenarios for risk management within so-called Internal Models are of particular interest in the insurance industry under Solvency II (see, e.g., [1, 4, 6–8, 21, 31]). In this paper, we propose a simple stochastic Monte Carlo algorithm on patchwork copulas for the generation of VaR scenarios that are suitable for comparison purposes in Internal Models for the calculation of Solvency Capital Requirements (SCR), in particular for the Non-Life Module. Note that in the Standard Formula of Solvency II, there is a formula for the calculation of the non-life premium and reserve risk SCR given by the volume factor

$$\rho_{1-\alpha}(\sigma)_{\mathrm{VaR}} = \frac{\exp\left(k_{1-\alpha}\sqrt{\ln\left(1+\sigma^2\right)}\right)}{\sqrt{1+\sigma^2}} - 1$$

applied to the volume measure (i.e. premium income) of the year considered (see, e.g., [31, p. 324, relation (21.9b)]; cf. also [16, p. 329 ff.]). Here α denotes the risk level (i.e. 0.5% in Solvency II) and $k_{1-\alpha}$ the corresponding $1-\alpha$ quantile of the standard normal distribution. Further, σ denotes the standard deviation of the underlying risk, i.e. the ultimate combined loss ratio, which is assumed to be lognormally distributed with expectation 1=100% (which is the limit towards certain ruin according to the law of large numbers). However, this formula is questionable from a scientific point of view (see [14]). Note also that this formula was simplified in the Commission Delegated Regulation of the EU [12, Article 115]:

$$\rho_{1-\alpha}(\sigma)_{\mathrm{VaR}} \approx 3\sigma \quad \text{for} \quad \alpha = 0.005.$$

***Corresponding Author: Dietmar Pfeifer:** Carl von Ossietzky Universität Oldenburg, Germany,
E-mail: dietmar.pfeifer@uni-oldenburg.de
Olena Ragulina: Taras Shevchenko National University of Kyiv, Ukraine, E-mail: ragulina.olena@gmail.com

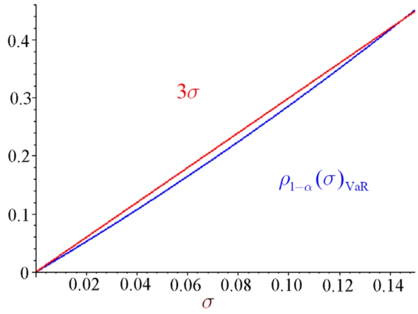

Figure 1: Plot of the Non-Life SCR volume factor $\rho_{1-\alpha}(\sigma)_{\mathrm{VaR}}$ vs. its simplification 3σ.

This is a reasonable conservative approximation as long as $\rho < 0.15$ (see Figure 1).

Another questionable point here is the aggregation to the overall SCR from different module SCR's by correlations in Solvency II (see, e.g., [31]). This has been discussed in detail, e.g., in [24, 29].

Note that no official legislative paper on Solvency II contains a strict mathematical definition of the underlying risk measure Value at Risk, cf. [11, Article 104, L 335/52, No. 4] or the Commission Delegated Regulation of the EU [12, L 12/20 (53)]. The wording used in these documents, however, suggests that "the Value-at-Risk measure with a 99.5% confidence level" is the corresponding lower quantile of the risk distribution.

Note also that the above-mentioned Commission Delegated Regulation [12] concerning the implementation of Solvency II in the EU requires the consideration of risk scenarios in several Articles, in particular in Article 259, L 12/161 on Risk Management Systems saying that insurance and reinsurance undertakings shall, where appropriate, include performance of stress tests and scenario analyses with regard to all relevant risks faced by the undertaking, in their risk-management system. The results of such analyses also have to be reported in the ORSA (Own Risk and Solvency Assessment, see, e.g., [23]) as described in Article 306 of the Commission Delegated Regulation of the EU [12]. In the light of the outlined structural problems with the standard formula above, the ORSA is probably a better instrument to rate the enterprise's risks in a more reliable way. The problem is, however, that the Commission Delegated Regulation does not make any clear statements on how such stress tests or scenario analyses have to be performed.

Article 1 of the Commission Delegated Regulation of the EU [12, L 12/20, No. 2] defines a "scenario analysis" as an analysis of the impact of a combination of adverse events. The Monte Carlo simulation algorithm developed in this paper allows for a mathematically rigorous description how such scenarios can be generated, being flexible enough to cover also extreme situations.

In what follows, we shall focus mainly on the Non-Life Modules under Solvency II. Therefore, we only consider continuous risk distributions. In this case, VaR is simply a lower quantile of the cumulative risk distribution function. For corresponding considerations for the Life and Capital Asset Modules under Solvency II, we refer to [3, 32].

Besides Solvency II aspects, the method proposed in this paper might also be of interest for reinsurance companies for the risk assessment of statistically dependent natural perils like windstorm, hail or flooding triggered by adverse climate conditions.

2 Unfavourable patchwork copulas

Patchwork copulas in the context of risk management have been treated in detail in [1, 5, 15, 24–26, 30], among others. In several of the cited papers the question of an unfavourable, i.e. superadditive VaR estimate for a

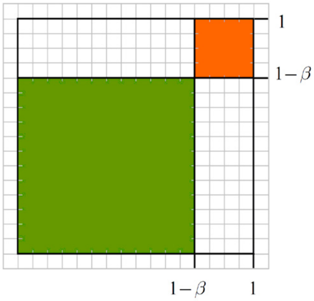

Figure 2: Shape of the support of the underlying patchwork copula **W**.

portfolio of aggregated risks was in particular emphasized, see also [27]. However, the construction of worst VaR scenarios in this context is quite complicated; note that a worst VaR is a supremum of VaR's over the Fréchet class of all possible joint distributions with given marginals (see, e.g., [10, Sections 1.1 and 1.3]). The situation is more simple in the two-dimensional case with identical margins (see [10, Section 2]). A numerical approach to a constructive solution to the general problem is given, e.g., by the rearrangement algorithm (see, e.g., [1, 10, 20]). From a practical point of view, simpler and yet explicit constructions for unfavourable but not necessarily worst VaR estimates by appropriate copula constructions seem to be a useful alternative. In this paper, we describe how such a construction could be performed. We start with an explicit approach in two dimensions, which is later extended to arbitrary dimensions. For better readability, all proofs are shifted to the Appendix.

Lemma 1. *Let, for $d \geq 2$, $d \in \mathbb{N}$, $\mathbf{U} = (U_1, \dots, U_d)$ and $\mathbf{V} = (V_1, \dots, V_d)$ be d-dimensional random vectors over $[0,1]^d$ with continuous uniform margins (i.e.,* **U** *and* **V** *represent d-dimensional copulas). Let further I denote a binomially distributed random variable, independent of* **U** *and* **V**, *with $\mathbb{P}(I = 1) = p \in (0,1)$. Then the random vector* **W** *with components $W_i := IpU_i + (1-I)[p + (1-p)V_i]$ for $1 \leq i \leq d$ also has continuous uniform margins, i.e.* **W** *represents a d-dimensional copula.*

Note that **W** can be considered as a special case of ordinal sums (cf. [22, Chapter 3.3.2] for the two-dimensional case, and [18, relation (4.31)], [19, Definition 2.1] and [9, Example 2.2.10 and Chapter 3.8] for the multivariate case).

Suppose now that a portfolio of d insurance risks is considered where a mutual probabilistic dependence structure is assumed to be described by **U**. If the d (for simplicity assumed continuous) marginal risk distribution functions are denoted by $F_1, \dots, F_d$ and by $Q_1, \dots, Q_d$ their pseudo-inverses (quantile functions), then both random vectors $\big(Q_1(U_1), \dots, Q_d(U_d)\big)$ and $\big(Q_1(W_1), \dots, Q_d(W_d)\big)$ represent a risk vector $\mathbf{X} = (X_1, \dots, X_d)$ with the given marginal distributions. However, w.r.t. to risk aggregation, $\mathbf{X} := \big(Q_1(W_1), \dots, Q_d(W_d)\big)$ creates in general an unfavourable VaR scenario for $S = \sum_{i=1}^d X_i$, even if p is close to 1 and therefore **U** and **W** differ only marginally. The graph in Figure 2 shows the corresponding support of **W** in two dimensions.

In the sequel, put $p := 1 - \beta$ for $0 < \beta < 1$. Then $\mathbf{W} = I(1-\beta)\mathbf{U} + (1-I)(1-\beta+\beta\mathbf{V})$.

We start with some further preliminary Lemmata.

Lemma 2. *Let $\mathbf{W}_1 := (1-\beta)\mathbf{U}$, $\mathbf{W}_2 := 1 - \beta + \beta\mathbf{V}$, $Z_{1i} := Q_i(W_{1i})$ and $Z_{2i} := Q_i(W_{2i})$, $i = 1, 2$. Then there hold*

$$F_{Z_{1i}}(x,\beta) = \begin{cases} \dfrac{F_i(x)}{1-\beta}, & 0 \leq x \leq Q_i(1-\beta), \\ 1, & x \geq Q_i(1-\beta), \end{cases} \quad \text{and} \quad F_{Z_{2i}}(x,\beta) = \begin{cases} 0, & 0 \leq x \leq Q_i(1-\beta), \\ \dfrac{F_i(x)+\beta-1}{\beta}, & x \geq Q_i(1-\beta). \end{cases}$$

Lemma 3. *Assume that f and g are Lebesgue densities of independent random variables X and Y, concentrated on the same finite interval* $[0, M]$ *with* $M > 0$. *Then* $S := X + Y$ *has the density* h_1 *given by*

$$h_1(x) = \begin{cases} \int_0^x f(x-y)g(y)\,dy, & 0 \le x \le M, \\ \int_{x-M}^M f(x-y)g(y)\,dy, & M \le x \le 2M. \end{cases}$$

If f and g are concentrated on the same infinite interval $[M, \infty)$ *with* $M \ge 0$, *then* $S := X + Y$ *has the density* h_2 *given by*

$$h_2(x) = \int_M^{x-M} f(x-y)g(y)\,dy, \quad x \ge 2M.$$

In particular, if F and G are the corresponding cdf's pertaining to f and g, respectively, then in either case, $\frac{d}{dx} F \star G(x)\big|_{x=2M} = 0$, *where* $\star$ *means convolution.*

Lemma 4. *Assume that all* $F_i \equiv F$ *being equal with quantile function Q, and that* **U** *and* **V** *have independent components each. Denote*

$$\underline{F}(x,\beta) := \begin{cases} \frac{F(x)}{1-\beta}, & x \le Q(1-\beta), \\ 1, & x \ge Q(1-\beta), \end{cases} \quad \text{and} \quad \overline{F}(x,\beta) := \frac{F\big(x + Q(1-\beta)\big) + \beta - 1}{\beta}, \quad x \ge 0.$$

Let further denote $X_i := Q(W_i)$ *and* $S = \sum_{i=1}^d X_i$. *Then we can conclude that*

$$F_S(x,\beta) = \begin{cases} (1-\beta)\underline{F}^{d\star}(x,\beta), & x \le dQ(1-\beta), \\ (1-\beta) + \beta\overline{F}^{d\star}\big(x - dQ(1-\beta), \beta\big), & x > dQ(1-\beta), \end{cases}$$

where $\star$ *again means convolution. If F has a density f, then correspondingly*

$$\underline{f}(x,\beta) := \begin{cases} \frac{f(x)}{1-\beta}, & x \le Q(1-\beta), \\ 1, & x \ge Q(1-\beta), \end{cases} \quad \text{and} \quad \overline{f}(x,\beta) := \frac{f\big(x + Q(1-\beta)\big)}{\beta}, \quad x \ge 0,$$

and

$$f_S(x,\beta) = \begin{cases} (1-\beta)\underline{f}^{d\star}(x,\beta), & x \le dQ(1-\beta), \\ (1-\beta) + \beta\overline{f}^{d\star}\big(x - dQ(1-\beta), \beta\big), & x > dQ(1-\beta). \end{cases}$$

The following examples show the effect of a risk aggregation with an unfavourable VaR scenario for two dimensions in detail.

Example 1 (exponential distributions)**.** Assume that

$$F_1 = F_2 = \begin{cases} 0, & x < 0, \\ 1 - e^{-x}, & x \ge 0. \end{cases}$$

Then

$$F_{Z_{1i}}(x,\beta) = \frac{1-e^{-x}}{1-\beta}, \quad 0 \le x \le -\ln(\beta), \quad \text{and} \quad F_{Z_{2i}}(x,\beta) = \frac{\beta - e^{-x}}{\beta} = 1 - e^{-x-\ln(\beta)}, \quad x \ge -\ln(\beta), \quad i = 1, 2.$$

For the corresponding densities, we obtain by differentiation

$$f_{Z_{1i}}(x,\beta) = \begin{cases} \frac{e^{-x}}{1-\beta}, & 0 \le x \le -\ln(\beta), \\ 0, & x > -\ln(\beta), \end{cases} \quad \text{and} \quad f_{Z_{2i}}(x,\beta) = \begin{cases} 0, & x < -\ln(\beta), \\ e^{-x-\ln(\beta)}, & x \ge -\ln(\beta), \end{cases} \quad i = 1, 2,$$

and

$$\underline{f}(x,\beta) = \begin{cases} \frac{e^{-x}}{1-\beta}, & 0 \le x \le -\ln(\beta), \\ 0, & x > -\ln(\beta), \end{cases} \quad \text{and} \quad \overline{f}(x,\beta) = \begin{cases} 0, & x < 0, \\ e^{-x}, & x \ge 0. \end{cases}$$

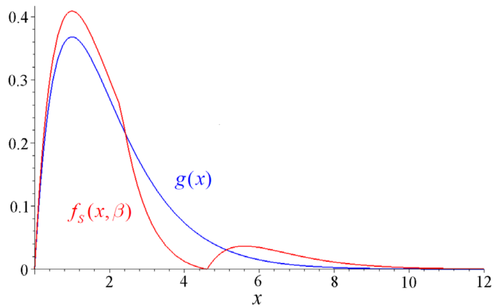

Figure 3: Plots of the densities $f_S(x, \beta)$ for $\beta = 0.1$ and $g(x)$ in Example 1.

By Lemma 4, we obtain the following density f_S of the aggregated risk S:

$$f_S(x,\beta) = \begin{cases} \dfrac{xe^{-x}}{1-\beta}, & 0 \le x \le -\ln(\beta), \\ \dfrac{(-2\ln(\beta) - x)e^{-x}}{1-\beta}, & -\ln(\beta) \le x \le -2\ln(\beta), \\ \dfrac{(x + 2\ln(\beta))e^{-x}}{\beta}, & x \ge -2\ln(\beta), \end{cases}$$

with the corresponding cdf F_S:

$$F_S(x,\beta) = \begin{cases} \dfrac{1-(1+x)e^{-x}}{1-\beta}, & 0 \le x \le -\ln(\beta), \\ \dfrac{1-2\beta+2e^{-x}\ln(\beta)+(1+x)e^{-x}}{1-\beta}, & -\ln(\beta) \le x \le -2\ln(\beta), \\ \dfrac{\beta-2e^{-x}\ln(\beta)-(1+x)e^{-x}}{\beta}, & x \ge -2\ln(\beta). \end{cases}$$

For the graph in Figure 3, let g denote the density of $T := Q_1(U_1) + Q_2(U_2)$ (independent summands, Gamma distribution). In what follows, let G denote the cdf of $T := Q_1(U_1)+Q_2(U_2)$ (independent summands, Gamma distribution) and H be the cdf of S under the worst VaR scenario (see the graph in Figure 4), i.e. the distribution of $\mathbf{V}$ corresponds to the lower Fréchet bound or countermonotonicity copula (see, e.g., Remark 3 and the comments after Figure 3 in [10], or [24]). In this case we have

$$H(x,\beta) = \begin{cases} F_S(x), & x \le -2\ln(\beta), \\ 1-\beta, & -2\ln(\beta) \le x \le -2\ln(\beta/2), \\ 1-\beta+\sqrt{\beta^2-4e^{-x}}, & x \ge -2\ln(\beta/2). \end{cases}$$

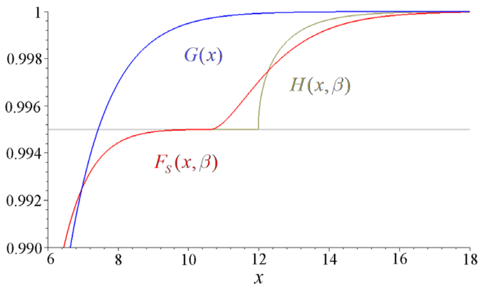

Figure 4: Plots of the cdf's $F_S(x,\beta)$, $G(x)$ and $H(x,\beta)$ for $\beta = 0.005$ in Example 1.

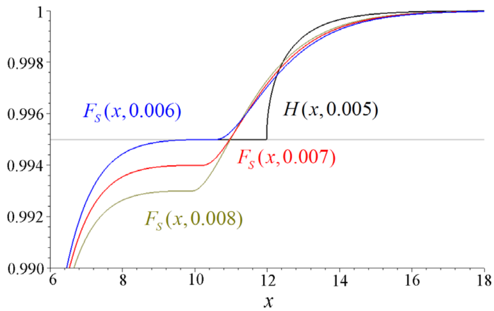

Figure 5: Plots of the cdf's $F_S(x, 0.005 + \varepsilon)$ for $\varepsilon \in \{0.001,\ 0.002,\ 0.003\}$ and $H(x, 0.005)$ in Example 1.

Note that with the Solvency II standard $\alpha = 0.005$, we get here, for $\beta = \alpha$, $\mathrm{VaR}_\alpha(S) = -2\ln(\beta) = 10.5914 > \mathrm{VaR}_\alpha(T) = 7.4301$, where $\mathrm{VaR}_\alpha(T)$ is the numerical solution to the equation $(1 + x)e^{-x} = \alpha$. For the worst VaR scenario, however, we get $\mathrm{wVaR}_\alpha(S) = -2\ln\left(\frac{\beta}{2}\right) = 11.9829$ with $10.5966 = \mathrm{SVaR}_\alpha := \mathrm{VaR}_\alpha(X_1) + \mathrm{VaR}_\alpha(X_2) > \mathrm{VaR}_\alpha(S) = 10.5914$. This means that even with the construction for S with $\beta = \alpha$, we still have a (quite small) diversification effect, but not in the worst VaR scenario. This changes, however, if we look at $\mathrm{VaR}_\alpha(S) = 10.9630$ when we replace β by $\alpha + \varepsilon$ in the definition of $\mathbf{W}$ for e.g. $\varepsilon = 0.001$.

The graph in Figure 5 shows the cdf's in the tails for several choices of ε. The graph in Figure 6 shows the values of $Q_S(0.995, \beta) = F_S^{-1}(0.995, \beta)$ in the range $0.0062 \le \beta \le 0.0076$. A numerical calculation shows that for $\alpha = 0.005$ the worst $\mathrm{VaR}_\alpha(S) = 10.9829$ is attained for $\beta = 0.0068$, i.e. $\varepsilon = 0.0018$. Table 1 summarizes the results found for $\alpha = 0.005$.

Table 1: Summarized results for Example 1.

β	0.0050	0.0060	0.0068	0.0070	0.0080
$\mathrm{VaR}_\alpha(S)$	10.5914	10.9630	10.9829	10.9821	10.9618
$\mathrm{VaR}_\alpha(T)$	7.4301	7.4301	7.4301	7.4301	7.4301
$\mathrm{wVaR}_\alpha(S)$	11.9829	11.9829	11.9829	11.9829	11.9829
SVaR_α	10.5966	10.5966	10.5966	10.5966	10.5966

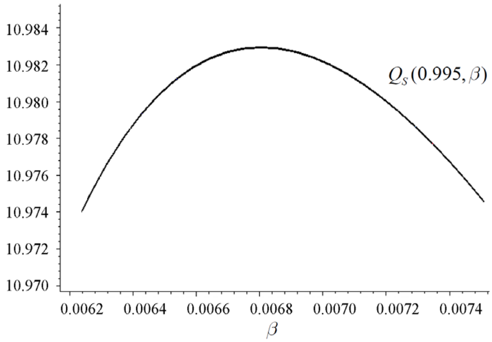

Figure 6: Plot of the parametrized quantile function $Q_S(0.995, \beta) = F_S^{-1}(0.995, \beta)$ in Example 1.

Example 2 (uniform distributions). Assume that

$$F_1 = F_2 = \begin{cases} 0, & x \le 0, \\ x, & 0 \le x \le 1, \\ 1, & x \ge 1. \end{cases}$$

Then

$$F_{Z_{1i}}(x,\beta) = \frac{x}{1-\beta}, \quad 0 \le x \le 1-\beta, \quad \text{and} \quad F_{Z_{2i}}(x,\beta) = \frac{x+\beta-1}{\beta}, \quad x \ge 1-\beta, \quad i = 1, 2.$$

By Lemma 4, we obtain the following density f_S of the aggregated risk S:

$$f_S(x,\beta) = \begin{cases} \frac{x}{1-\beta}, & x \le 1-\beta, \\ \frac{2-2\beta-x}{1-\beta}, & 1-\beta \le x \le 2-2\beta, \\ \frac{x-2+2\beta}{\beta}, & 2-2\beta \le x \le 2-\beta, \\ \frac{2-x}{\beta}, & 2-\beta \le x \le 2, \end{cases}$$

with the corresponding cdf F_S:

$$F_S(x,\beta) = \begin{cases} \frac{x^2}{2(1-\beta)}, & x \le 1-\beta, \\ \frac{4x(1-\beta)-x^2-2(1-\beta)^2}{2(1-\beta)}, & 1-\beta \le x \le 2-2\beta, \\ \frac{4(1-\beta)(1-x)+x^2-2\beta+2\beta^2}{2\beta}, & 2-2\beta \le x \le 2-\beta, \\ \frac{2\beta-4(1-x)-x^2}{2\beta}, & 2-\beta \le x \le 2. \end{cases}$$

In what follows, g is the density of $T := Q_1(U_1)+Q_2(U_2)$ (independent summands, triangle distribution), see the graph in Figure 7. In the graph in Figure 8, G is the cdf for $T := Q_1(U_1) + Q_2(U_2)$ (independent summands, triangle distribution) and H is the cdf for S under the worst VaR scenario, i.e. the distribution of $\mathbf{V}$ corresponds to the lower Fréchet bound. In this case we have

$$H(x,\beta) = \begin{cases} F_S(x), & x \le 2-2\beta, \\ 1-\beta, & 2-2\beta \le x < 2-\beta, \\ 1, & x \ge 2-\beta. \end{cases}$$

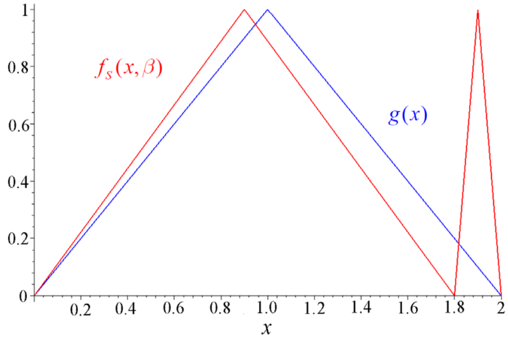

Figure 7: Plots of the densities $f_S(x,\beta)$ for $\beta = 0.1$ and $g(x)$ in Example 2.

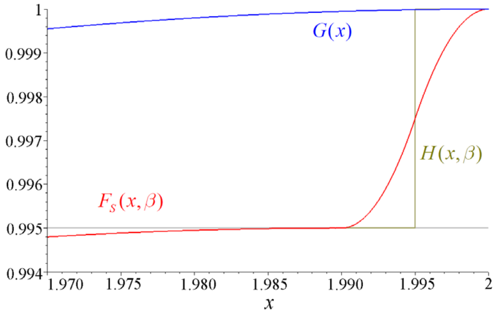

Figure 8: Plots of the cdf's $F_S(x,\beta)$, $G(x)$ and $H(x,\beta)$ for $\beta = 0.005$ in Example 2.

Note that with the Solvency II standard $\alpha = 0.005$, we have here, for $\beta = \alpha$, $\mathrm{VaR}_\alpha(S) = 2 - 2\alpha = 1.9900 = \mathrm{VaR}_\alpha(T)$. For the worst VaR scenario, however, we get here $\mathrm{wVaR}_\alpha(S) = 2 - \alpha = 1.9950 > 1.9900 = \mathrm{VaR}_\alpha(X_1) + \mathrm{VaR}_\alpha(X_2) = \mathrm{SVaR}_\alpha = \mathrm{VaR}_\alpha(S)$. This means that with the construction for S we have no true diversification effect, in contrast to the worst VaR scenario. This changes, however, if we look at $\mathrm{VaR}_\alpha(S) = 1.9910$ when we replace β by $\alpha + \varepsilon$ in the definition of $\mathbf{W}$ for e.g. $\varepsilon = 0.001$.

The graph in Figure 9 shows the cdf's in the tails for several choices of ε. The graph in Figure 10 shows the values of $Q_S(0.995,\beta) = F_S^{-1}(0.995,\beta)$ in the range $0.0054 \le \beta \le 0.0070$. A numerical calculation shows that for $\alpha = 0.005$ the maximal $\mathrm{VaR}_\alpha(S) = 1.9915$ is attained for $\beta = 0.0060$, i.e. $\varepsilon = 0.0010$.

Note that in this example a closed-form representation for $Q_S(u,\beta)$ is given by

$$Q_S(u,\beta) = 2 - 2\beta + \sqrt{2\beta(\beta + u - 1)}, \quad 1 - \beta \le u \le 1 - \frac{\beta}{2}.$$

This implies

$$Q_S(1-\alpha,\beta) = 2 - 2\beta + \sqrt{2\beta(\beta - \alpha)}, \quad \alpha \le \beta \le 2\alpha,$$

with its maximum being attained for $\beta_0 = \frac{1+\sqrt{2}}{2}\alpha$ with value $Q_S(1-\alpha,\beta_0) = 2 - \left(1 + \frac{\sqrt{2}}{2}\right)\alpha$. Note that in contrast the worst VaR here is $\mathrm{wVaR}_\alpha(S^*) = 2 - \alpha$. Table 2 summarizes the results found for $\alpha = 0.005$.

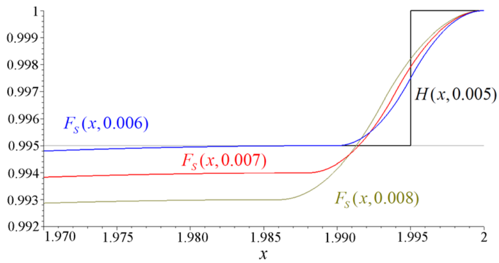

Figure 9: Plots of $F_S(x, 0.005 + \varepsilon)$ for $\varepsilon \in \{0.001, 0.002, 0.003\}$ and $H(x, 0.005)$ in Example 2.

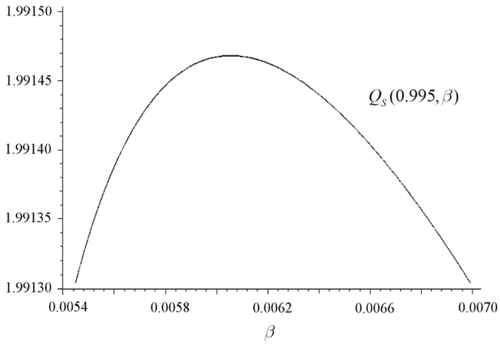

Figure 10: Plot of the parametrized quantile function $Q_S(0.995, \beta) = F_S^{-1}(0.995, \beta)$ in Example 2.

Table 2: Summarized results for Example 2.

β	0.0050	0.0055	0.0060	0.0065	0.0070
$\mathrm{VaR}_\alpha(S)$	1.9900	1.9130	1.9915	1.9914	1.9913
$\mathrm{VaR}_\alpha(T)$	1.9900	1.9900	1.9900	1.9900	1.9900
$\mathrm{wVaR}_\alpha(S)$	1.9950	1.9950	1.9950	1.9950	1.9950
SVaR_α	1.9900	1.9900	1.9900	1.9900	1.9900

Example 3 (Pareto distributions). Assume that

$$F_1 = F_2 = \begin{cases} 0, & x \le 0, \\ \dfrac{x}{1+x}, & x > 0. \end{cases}$$

Then

$$F_{Z_{1i}}(x,\beta) = \frac{x}{(1-\beta)(1+x)}, \quad 0 \le x \le \frac{1}{\beta} - 1, \quad \text{and} \quad F_{Z_{2i}}(x,\beta) = 1 - \frac{1}{\beta(1+x)}, \quad x \ge \frac{1}{\beta} - 1, \quad i = 1, 2.$$

For the corresponding densities, we obtain by differentiation

$$f_{Z_{1i}}(x,\beta) = \begin{cases} \dfrac{1}{(1-\beta)(1+x)^2}, & 0 \le x \le \dfrac{1}{\beta} - 1, \\ 0, & x > \dfrac{1}{\beta} - 1, \end{cases} \quad \text{and} \quad f_{Z_{2i}}(x,\beta) = \begin{cases} 0, & x < \dfrac{1}{\beta} - 1, \\ \dfrac{1}{\beta(1+x)^2}, & x \ge \dfrac{1}{\beta} - 1, \end{cases}$$

and

$$\underline{f}(x,\beta) = \begin{cases} \dfrac{1}{(1-\beta)(1+x)^2}, & 0 \le x \le \dfrac{1}{\beta} - 1, \\ 0, & x > \dfrac{1}{\beta} - 1, \end{cases} \quad \text{and} \quad \overline{f}(x,\beta) = \begin{cases} 0, & x < 0, \\ \dfrac{\beta}{(1+\beta x)^2}, & x \ge 0. \end{cases}$$

In order to calculate the density f_S of the aggregated risk S, we need a suitable partial fraction representation of $\underline{f}(x-y)\underline{f}(y)$ and $\overline{f}(x-y)\overline{f}(y)$. Note that in general, we have

$$\frac{1}{(1+x-y)(1+y)} = \frac{1}{2+x}\left(\frac{1}{1+x-y} + \frac{1}{1+y}\right)$$

and

$$\begin{aligned} \frac{1}{(1+x-y)^2(1+y)^2} &= \frac{1}{(2+x)^2}\left(\frac{1}{1+x-y} + \frac{1}{1+y}\right)^2 \\ &= \frac{1}{(2+x)^2}\left(\frac{1}{(1+x-y)^2} + \frac{1}{(1+y)^2} + \frac{2}{2+x}\left(\frac{1}{1+x-y} + \frac{1}{1+y}\right)\right), \end{aligned}$$

from which we obtain, by Lemma 4,

$$F_S(x,\beta) = \begin{cases} \dfrac{x^2 + 2x - 2\ln(1+x)}{(2+x)^2(1-\beta)}, & 0 \le x \le \dfrac{1}{\beta} - 1, \\ \dfrac{(1-2\beta)x^2 + (4-6\beta)x - 4\beta + 4 + 2\ln(\beta x + 2\beta - 1)}{(2+x)^2(1-\beta)}, & \dfrac{1}{\beta} - 1 \le x \le 2\left(\dfrac{1}{\beta} - 1\right), \\ \dfrac{x^2 - 2x + \frac{2}{\beta}\ln(\beta x + 2\beta - 1)}{(2+x)^2}, & x \ge 2\left(\dfrac{1}{\beta} - 1\right). \end{cases}$$

The density $f_S(x)$ follows by differentiation.

In the following, g denotes the density of $T := Q_1(U_1) + Q_2(U_2)$ (independent summands), see the graph in Figure 11. In the graph in Figure 12, G is the cdf of $T := Q_1(U_1) + Q_2(U_2)$ (independent summands) and H is the cdf of S under the worst VaR scenario, i.e. the distribution of $\mathbf{V}$ corresponds again to the lower Fréchet bound. In this case we have

$$H(x,\beta) = \begin{cases} F_S(x), & x \le \dfrac{2}{\beta} - 2, \\ 1 - \beta, & \dfrac{2}{\beta} - 2 \le x \le \dfrac{4}{\beta} - 2, \\ 1 - \beta + \sqrt{\beta^2 - \dfrac{4\beta}{2+x}}, & x \ge \dfrac{4}{\beta} - 2. \end{cases}$$

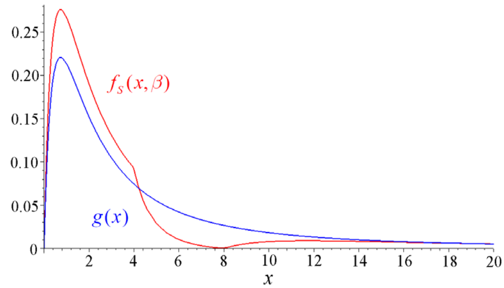

Figure 11: Plots of the densities $f_S(x,\beta)$ for $\beta = 0.1$ and $g(x)$ in Example 3.

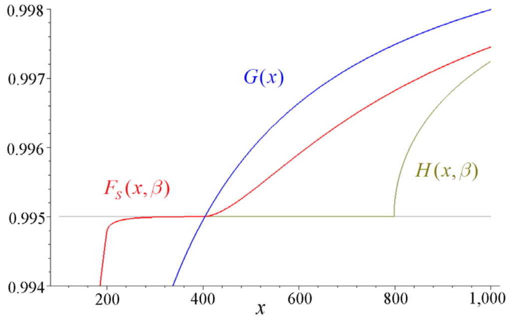

Figure 12: Plots of the cdf's $F_S(x,\beta)$, $G(x)$ and $H(x,\beta)$ for $\beta = 0.005$ in Example 3.

Note that with the Solvency II standard $\alpha = 0.005$, we have here, for $\beta = \alpha$, $\mathrm{VaR}_\alpha(S) = 397.3168 < \mathrm{VaR}_\alpha(T) = 403.9161$. For the worst VaR scenario, however, we get $\mathrm{wVaR}_\alpha(S) = \frac{4}{\beta} - 2 = 798 > 398 = \mathrm{VaR}_\alpha(X_1) + \mathrm{VaR}_\alpha(X_2) = \mathrm{SVaR}_\alpha > \mathrm{VaR}_\alpha(S) = 397.3168$. This means that even with the construction for S we still have a (quite small) diversification effect, but not in the worst VaR scenario, as expected. This changes, however, if we look at $\mathrm{VaR}_\alpha(S) = 488.2116$ when we replace β by $\beta+\varepsilon$ in the definition of $\mathbf{W}$ for, e.g., $\varepsilon = 0.001$.

The graph in Figure 13 shows the cdf's in the tail for several choices of ε. The graph in Figure 14 shows the values of $Q_S(0.995,\beta) = F_S^{-1}(0.995,\beta)$ in the range $0.007 \le \beta \le 0.012$. A numerical calculation shows that for $\alpha = 0.005$ the maximum $\mathrm{VaR}_\alpha(S) = 509.3804$ is attained for $\beta = 0.0089$, i.e. $\varepsilon = 0.0039$. Table 3 summarizes the results found for $\alpha = 0.005$.

Table 3: Summarized results for Example 3.

β	0.0050	0.0070	0.0089	0.0100	0.0110
$\mathrm{VaR}_\alpha(S)$	397.3168	503.2848	509.3804	508.6489	507.0076
$\mathrm{VaR}_\alpha(T)$	403.9161	403.9161	403.9161	403.9161	403.9161
$\mathrm{wVaR}_\alpha(S)$	798.0000	798.0000	798.0000	798.0000	798.0000
SVaR_α	398.0000	398.0000	398.0000	398.0000	398.0000

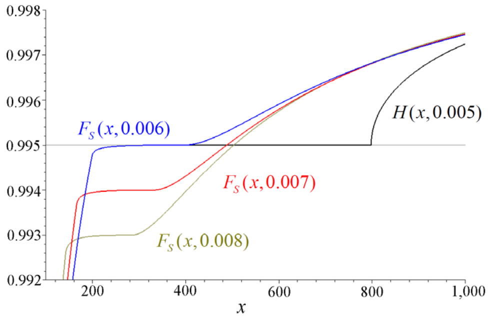

Figure 13: Plots of the cdf's $F_S(x, 0.005 + \varepsilon)$ for $\varepsilon \in \{0.001,\ 0.002,\ 0.003\}$ and $H(x, 0.005)$ in Example 3.

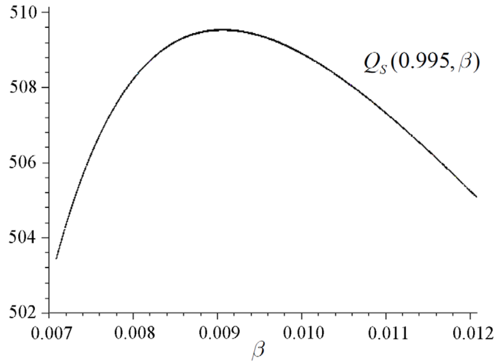

Figure 14: Plot of the parametrized quantile function $Q_S(0.995, \beta) = F_S^{-1}(0.995, \beta)$ in Example 3.

Examples 1–3 show that it is generally possible to obtain unfavourable VaR scenarios by a suitable choice of $\beta = \alpha + \varepsilon$ in the definition of $\mathbf{W}$, i.e. scenarios which lead to an opposite diversification effect in the portfolio and which are sometimes close to the worst VaR scenario.

We continue with a particular construction of $\mathbf{W}$ which allows in general for an unfavourable VaR scenario.

Lemma 5. *For $d \in \mathbb{N}$, $d > 1$, let $\mathbf{I}_d$ denote the d-dimensional unit matrix, $\mathbf{e}_d = (1, \dots, 1)$ the d-dimensional row vector consisting of just ones, and $\mathbf{E}_d = \mathbf{e}_d^{tr}\mathbf{e}_d$ the $d \times d$ matrix with all entries equal to unity. Then $\Sigma_d = (1-r)\mathbf{I}_d + r\mathbf{E}_d$ is a correlation matrix iff $-\frac{1}{d-1} \le r \le 1$. In the general case, the latent roots λ_i of Σ_d are given by $\lambda_1 = 1 + (d-1)r$ and $\lambda_i = 1 - r$, $i = 2, \dots, d$. An orthonormal basis $T_1, \dots, T_d$ of corresponding latent vectors is given by $T_1 = \dfrac{1}{\sqrt{d}}\mathbf{e}_d^{tr}$ and $T_j = (t_{1j}, \dots, t_{dj})^{tr}$ for $2 \le j \le d$, where*

$$t_{ij} = \begin{cases} -\dfrac{1}{\sqrt{j(j-1)}}, & 1 \le i < j, \\ \sqrt{\dfrac{j-1}{j}}, & j = i, \\ 0, & i > j. \end{cases}$$

Hence Σ_d possesses the spectral decomposition $\Sigma_d = \mathbf{A}\mathbf{A}^{tr}$ with $\mathbf{A} = \mathbf{T}\sqrt{\Delta}$, where $\mathbf{T} = [T_1, \dots, T_d]$ and $\Delta = \operatorname{diag}(\lambda_1, \dots, \lambda_d)$.

Note that there is also an alternative possibility to represent latent roots $\lambda_j^\star$ and normalized latent vectors $T_j^\star = (t_{1j}^\star, \dots, t_{dj}^\star)^{tr}$, $j = 1, \dots, d$, of Σ_d since Σ_d is a particular symmetric Toeplitz matrix for which the latent roots and normalized latent vectors can be expressed via trigonometric functions (see [2, relations (5.89) and (5.90)]). In particular, we can choose

$$\lambda_j^\star = 1 + r\sum_{i=1}^{d-1}\cos\left(\frac{2\pi ij}{d}\right) = \begin{cases} 1 - r, & j = 1, \dots, d-1, \\ 1 + (d-1)r, & j = d, \end{cases}$$

and

$$t_{ij}^\star = \frac{\cos\left(\frac{2\pi ij}{d}\right) + \sin\left(\frac{2\pi ij}{d}\right)}{\sqrt{d}}, \quad 1 \le i, j \le d.$$

This is due to the fact that the latent roots have multiplicities, hence the linear space spanned by the corresponding latent vectors is $(d-1)$-dimensional, allowing for different representations of the corresponding linear basis. However, for our purposes, the representation in Lemma 5 is more suited.

In what follows we will call a Gaussian copula derived from the correlation matrix $\Sigma_d = \frac{d}{d-1}\mathbf{I}_d - \frac{1}{d-1}\mathbf{E}_d$ for $r = -\frac{1}{d-1}$ a *minimal correlation Gaussian copula*.

Note that the corresponding multivariate normal distribution is degenerated since Σ_d is singular, i.e. a random vector $\mathbf{X}$ with zero mean and correlation matrix Σ_d has the representation $\mathbf{X} = A\mathbf{Y}$, where $\mathbf{Y}$ has a standard multivariate normal distribution with mean zero and variance-covariance matrix $\mathbf{I}_d$. For $d = 2$, the minimal correlation Gaussian copula is identical to the lower Fréchet bound or countermonotonicity copula.

3 A case study

The following example shows the effects of such an approach for the 19-dimensional data set discussed in [26]. Table 4 contains insurance losses from a non-life portfolio of natural perils in $d = 19$ areas in central Europe over a time period of 20 years. The losses are given in million monetary units (MMU).

A statistical analysis of the data shows a good fit to lognormal $\mathcal{LN}(\mu, \sigma)$-distributions for the losses per Area k, $k = 1, \dots, 19$. The parameters μ_k and σ_k for Area k shown in Table 5 were hence estimated from the log data by calculating means and standard deviations.

Table 4: Insurance losses from a Nat Cat portfolio in central Europe.

Year	Area 1	Area 2	Area 3	Area 4	Area 5	Area 6	Area 7	Area 8	Area 9	Area 10
1	23.664	154.664	40.569	14.504	10.468	7.464	22.202	17.682	12.395	18.551
2	1.080	59.545	3.297	1.344	1.859	0.477	6.107	7.196	1.436	3.720
3	21.731	31.049	55.973	5.816	14.869	20.771	3.580	14.509	17.175	87.307
4	28.990	31.052	30.328	4.709	0.717	3.530	6.032	6.512	0.682	3.115
5	53.616	62.027	57.639	1.804	2.073	4.361	46.018	22.612	1.581	11.179
6	29.950	41.722	12.964	1.127	1.063	4.873	6.571	11.966	15.676	24.263
7	3.474	14.429	10.869	0.945	2.198	1.484	4.547	2.556	0.456	1.137
8	10.020	31.283	21.116	1.663	2.153	0.932	25.163	3.222	1.581	5.477
9	5.816	14.804	128.072	0.523	0.324	0.477	3.049	7.791	4.079	7.002
10	170.725	576.767	108.361	41.599	20.253	35.412	126.698	71.079	21.762	64.582
11	21.423	50.595	4.360	0.327	1.566	64.621	5.650	1.258	0.626	3.556
12	6.380	28.316	3.740	0.442	0.736	0.470	3.406	7.859	0.894	3.591
13	124.665	33.359	14.712	0.321	0.975	2.005	3.981	4.769	2.006	1.973
14	20.165	49.948	17.658	0.595	0.548	29.350	6.782	4.873	2.921	6.394
15	78.106	41.681	13.753	0.585	0.259	0.765	7.013	9.426	2.180	3.769
16	11.067	444.712	365.351	99.366	8.856	28.654	10.589	13.621	9.589	19.485
17	6.704	81.895	14.266	0.972	0.519	0.644	8.057	18.071	5.515	13.163
18	15.550	277.643	26.564	0.788	0.225	1.230	26.800	64.538	2.637	80.711
19	10.099	18.815	9.352	2.051	1.089	6.102	2.678	4.064	2.373	2.057
20	8.492	138.708	46.708	3.680	1.132	1.698	165.600	7.926	2.972	5.237

Year	Area 11	Area 12	Area 13	Area 14	Area 15	Area 16	Area 17	Area 18	Area 19
1	1.842	4.100	46.135	14.698	44.441	7.981	35.833	10.689	7.299
2	0.429	1.026	7.469	7.058	4.512	0.762	14.474	9.337	0.740
3	0.209	2.344	22.651	4.117	26.586	3.920	13.804	2.683	3.026
4	0.521	0.696	31.126	1.878	29.423	6.394	18.064	1.201	0.894
5	2.715	1.327	40.156	4.655	104.691	28.579	17.832	1.618	3.402
6	4.832	0.701	16.712	11.852	29.234	7.098	17.866	5.206	5.664
7	0.268	0.580	11.851	2.057	11.605	0.282	16.925	2.082	1.008
8	0.741	0.369	3.814	1.869	8.126	1.032	14.985	1.390	1.703
9	0.524	6.554	5.459	3.007	8.528	1.920	5.638	2.149	2.908
10	9.882	6.401	106.197	44.912	191.809	90.559	154.492	36.626	36.276
11	1.052	8.277	22.564	8.961	19.817	16.437	25.990	2.364	6.434
12	0.136	0.364	28.000	7.574	3.213	1.749	12.735	1.744	0.558
13	1.990	15.176	57.235	23.686	110.035	17.373	7.276	2.494	0.525
14	0.630	0.762	25.897	3.439	8.161	3.327	24.733	2.807	1.618
15	0.770	15.024	36.068	1.613	6.127	8.103	12.596	4.894	0.822
16	0.287	0.464	24.211	38.616	51.889	1.316	173.080	3.557	11.627
17	0.590	2.745	16.124	2.398	20.997	2.515	5.161	2.840	3.002
18	0.245	0.217	12.416	4.972	59.417	3.762	24.603	7.404	19.107
19	0.415	0.351	10.707	2.468	10.673	1.743	27.266	1.368	0.644
20	0.566	0.708	22.646	6.652	14.437	63.692	113.231	7.218	2.548

Table 5: Distributional parameters for fitted lognormal loss distributions

Parameter	Area 1	Area 2	Area 3	Area 4	Area 5	Area 6	Area 7	Area 8	Area 9	Area 10
μ_k	2.806	4.072	3.141	0.638	0.398	1.223	2.321	2.212	1.078	2.106
σ_k	1.216	1.052	1.211	1.569	1.300	1.599	1.198	0.988	1.145	1.253

Parameter	Area 11	Area 12	Area 13	Area 14	Area 15	Area 16	Area 17	Area 18	Area 19
μ_k	−0.323	0.382	3.020	1.749	3.041	1.550	3.070	1.244	0.938
σ_k	1.088	1.335	0.803	1.003	1.122	1.477	0.962	0.858	1.214

As is to be expected, insurance losses in locally adjacent areas show a high degree of stochastic dependence, which can also be seen from the correlation Tables 6 and 7. For a better readability, only two decimal places are displayed.

Table 6: Empirical correlations between original losses in adjacent areas.

Parameter	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	A14	A15	A16	A17	A18	A19
A1	1	0.46	0.03	0.16	0.47	0.20	0.35	0.49	0.41	0.24	0.78	0.64	0.91	0.63	0.85	0.66	0.30	0.67	0.56
A2	0.46	1	0.64	0.78	0.67	0.36	0.51	0.76	0.57	0.51	0.58	−0.04	0.59	0.84	0.68	0.58	0.87	0.77	0.90
A3	0.03	0.64	1	0.93	0.41	0.26	0.11	0.16	0.33	0.16	0.08	−0.09	0.13	0.64	0.25	0.10	0.74	0.14	0.35
A4	0.16	0.78	0.93	1	0.54	0.36	0.16	0.25	0.43	0.19	0.22	−0.10	0.30	0.79	0.36	0.19	0.84	0.32	0.49
A5	0.47	0.67	0.41	0.54	1	0.41	0.35	0.51	0.84	0.63	0.59	0.02	0.64	0.67	0.59	0.50	0.58	0.71	0.67
A6	0.20	0.36	0.26	0.36	0.41	1	0.07	0.11	0.28	0.19	0.28	0.14	0.31	0.42	0.24	0.27	0.39	0.27	0.40
A7	0.35	0.51	0.11	0.16	0.35	0.07	1	0.44	0.27	0.19	0.48	−0.07	0.46	0.35	0.45	0.91	0.64	0.61	0.49
A8	0.49	0.76	0.16	0.25	0.51	0.11	0.44	1	0.50	0.75	0.61	−0.03	0.54	0.47	0.71	0.53	0.40	0.75	0.90
A9	0.41	0.57	0.33	0.43	0.84	0.28	0.27	0.50	1	0.66	0.68	−0.01	0.52	0.60	0.50	0.41	0.46	0.65	0.63
A10	0.24	0.51	0.16	0.19	0.63	0.19	0.19	0.75	0.66	1	0.33	−0.12	0.27	0.28	0.43	0.24	0.23	0.45	0.65
A11	0.78	0.58	0.08	0.22	0.59	0.28	0.48	0.61	0.68	0.33	1	0.19	0.79	0.65	0.80	0.73	0.43	0.84	0.74
A12	0.64	−0.04	−0.09	−0.10	0.02	0.14	−0.07	−0.03	−0.01	−0.12	0.19	1	0.44	0.21	0.28	0.17	−0.12	0.13	0.03
A13	0.91	0.59	0.13	0.30	0.64	0.31	0.46	0.54	0.52	0.27	0.79	0.44	1	0.71	0.86	0.74	0.47	0.76	0.65
A14	0.63	0.84	0.64	0.79	0.67	0.42	0.35	0.47	0.60	0.28	0.65	0.21	0.71	1	0.74	0.54	0.79	0.68	0.72
A15	0.85	0.68	0.25	0.36	0.59	0.24	0.45	0.71	0.50	0.43	0.80	0.28	0.86	0.74	1	0.69	0.47	0.71	0.75
A16	0.66	0.58	0.10	0.19	0.50	0.27	0.91	0.53	0.41	0.24	0.73	0.17	0.74	0.54	0.69	1	0.63	0.77	0.64
A17	0.30	0.87	0.74	0.84	0.58	0.39	0.64	0.40	0.46	0.23	0.43	−0.12	0.47	0.79	0.47	0.63	1	0.59	0.64
A18	0.67	0.77	0.14	0.32	0.71	0.27	0.61	0.75	0.65	0.45	0.84	0.13	0.76	0.68	0.71	0.77	0.59	1	0.86
A19	0.56	0.90	0.35	0.49	0.67	0.40	0.49	0.90	0.63	0.65	0.74	0.03	0.65	0.72	0.75	0.64	0.64	0.86	1

Table 7: Empirical correlations between log losses in adjacent areas.

Parameter	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	A14	A15	A16	A17	A18	A19
A1	1	0.27	0.30	0.16	0.17	0.45	0.28	0.32	0.32	0.29	0.67	0.51	0.76	0.34	0.67	0.74	0.18	0.21	0.29
A2	0.27	1	0.48	0.66	0.39	0.37	0.71	0.69	0.52	0.64	0.30	−0.02	0.45	0.66	0.58	0.45	0.73	0.74	0.78
A3	0.30	0.48	1	0.70	0.40	0.31	0.42	0.51	0.58	0.53	0.18	0.07	0.21	0.32	0.54	0.26	0.47	0.21	0.57
A4	0.16	0.66	0.70	1	0.77	0.47	0.46	0.47	0.59	0.49	0.18	−0.13	0.33	0.50	0.47	0.18	0.76	0.43	0.54
A5	0.17	0.39	0.40	0.77	1	0.59	0.30	0.20	0.49	0.39	0.28	0.08	0.35	0.56	0.44	0.16	0.55	0.36	0.41
A6	0.45	0.37	0.31	0.47	0.59	1	0.14	0.01	0.36	0.34	0.33	0.12	0.48	0.46	0.48	0.37	0.59	0.17	0.50
A7	0.28	0.71	0.42	0.46	0.30	0.14	1	0.52	0.27	0.40	0.45	−0.07	0.31	0.31	0.46	0.62	0.63	0.58	0.57
A8	0.32	0.69	0.51	0.47	0.20	0.01	0.52	1	0.64	0.81	0.27	−0.02	0.38	0.35	0.56	0.35	0.28	0.62	0.63
A9	0.32	0.52	0.58	0.59	0.49	0.36	0.27	0.64	1	0.78	0.40	0.19	0.27	0.50	0.44	0.30	0.33	0.57	0.61
A10	0.29	0.64	0.53	0.49	0.39	0.34	0.40	0.81	0.78	1	0.21	−0.02	0.21	0.37	0.52	0.30	0.31	0.53	0.81
A11	0.67	0.30	0.18	0.18	0.28	0.33	0.45	0.27	0.40	0.21	1	0.47	0.49	0.45	0.60	0.67	0.20	0.45	0.39
A12	0.51	−0.02	0.07	−0.13	0.08	0.12	−0.07	−0.02	0.19	−0.02	0.47	1	0.44	0.21	0.24	0.46	−0.23	0.25	0.05
A13	0.76	0.45	0.21	0.33	0.35	0.48	0.31	0.38	0.27	0.21	0.49	0.44	1	0.55	0.60	0.71	0.37	0.39	0.24
A14	0.34	0.66	0.32	0.50	0.56	0.46	0.31	0.35	0.50	0.37	0.45	0.21	0.55	1	0.59	0.43	0.57	0.58	0.53
A15	0.67	0.58	0.54	0.47	0.44	0.48	0.46	0.56	0.44	0.52	0.60	0.24	0.60	0.59	1	0.59	0.36	0.35	0.63
A16	0.74	0.45	0.26	0.18	0.16	0.37	0.62	0.35	0.30	0.30	0.67	0.46	0.71	0.43	0.59	1	0.38	0.43	0.39
A17	0.18	0.73	0.47	0.76	0.55	0.59	0.63	0.28	0.33	0.31	0.20	−0.23	0.37	0.57	0.36	0.38	1	0.52	0.56
A18	0.21	0.74	0.21	0.43	0.36	0.17	0.58	0.62	0.57	0.53	0.45	0.25	0.39	0.58	0.35	0.43	0.52	1	0.60
A19	0.29	0.78	0.57	0.54	0.41	0.50	0.57	0.63	0.61	0.81	0.39	0.05	0.24	0.53	0.63	0.39	0.56	0.60	1

The graph in Figure 15 shows estimated cdf's on a basis of 100,000 Monte Carlo simulations for the aggregated loss using lognormal margins with the parameters from Table 5 with a Bernstein copula representing $\mathbf{U}$ and a minimal correlation Gaussian copula representing $\mathbf{V}$, for various values of p. For comparison purposes, we have also added an estimated cdf for the aggregated loss for a Bernstein copula representing $\mathbf{U}$ and an upper Fréchet (or comonotonicity) copula representing $\mathbf{V}$. Note that the Bernstein copula is here constructed according to [5] on the basis of the ranks of the risk vectors (see also [28, Section 3]).

The plots in Figure 15 for the tail cdf's correspond to a Bernstein copula $\mathbf{U}$ with a minimal correlation Gaussian copula $\mathbf{V}$: $p = 1$ ($F_1(x)$), $p = 0.99$ ($F_2(x)$), $p = 0.994$ ($F_3(x)$), as well as to a Bernstein copula $\mathbf{U}$ with $p = 0.994$ but different copulas $\mathbf{V}$: upper Fréchet bound or comonotonicity copula ($F_4(x)$) and independence copula ($F_5(x)$).

Table 8 shows the estimated risk measures VaR_α for $\alpha = 0.005$ (Solvency II standard) for the various values of p and different types of $\mathbf{V}$.

Table 8: Survey over VaR-estimates under different copula models with lognormal margins, in MMU.

p	0.99	0.994	0.994	0.994	1
$\mathbf{V}$	minimal correlation Gaussian	minimal correlation Gaussian	upper Fréchet	independence	–
VaR_α	4,647	5,272	3,976	5,018	2,229

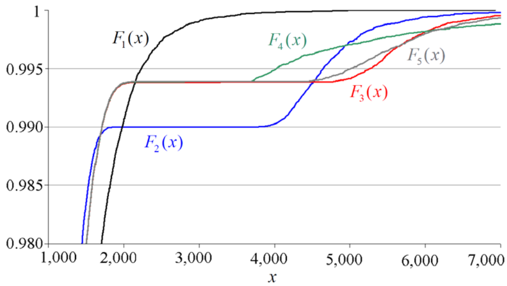

Figure 15: Plots of the estimated cdf's $F_i(x)$, $i = 1, \ldots, 5$, in the tail.

As can clearly be seen, the patchwork construction with the minimal correlation Gaussian copula representing **V** with no tail dependence gives the largest VaR estimate here and is typically larger than the upper Fréchet copula, which has a positive tail dependence. Note that the sum of individual VaR's is given by 3,976 MMU, which means that using the Bernstein copula alone would lead to a diversified portfolio while all other copula models do not.

Finally, it should be pointed out that the effects described here are independent of the particular copula chosen for **U**, i.e. the magnitude of the estimated VaR's under the patchwork construction would remain roughly equal also under an elliptical, an Archimedean, a vine or an adapted Bernstein copula approach for **U** ([28], cf. also the comments after Figure 3 in [10]).

4 Concluding remarks

The patchwork copula construction presented in this paper allows for a simple but yet effective and well-defined way to generate unfavourable VaR scenarios, i.e. scenarios with opposite diversification effects in particular for applications in Solvency II. Such scenario considerations are prescribed by legislative guidelines as, e.g., specified in the Commission Delegated Regulation of the EU [12] (p. L12/6 (16), L12/9 (49), L12/12 (75) or (77), just to mention some). Besides Solvency II, such unfavourable VaR scenario generations could also be of interest in the Basel III framework (e.g., economic scenario generators) or in the reinsurance industry, in particular w.r.t. extreme natural perils.

Although there is theoretically also a method to create worst VaR scenarios by means of the rearrangement algorithm, the latter approach easily becomes numerically cumbersome in high-dimensional portfolios as in our case study, especially, if the risk distributions are not identical (see [10, Section 2.2]). Hence a sub-optimal but easy to implement alternative is of value, in particular, since it seems unlikely that the worst VaR scenario would actually occur in real life portfolios.

The approach discussed in this paper seems, at a first glance, to be related to the recent paper [27]. The essential difference is, however, that the latter paper is not based on an observation-free copula construction for the tails as in the present paper. The algorithm proposed there leads only to stochastic approximations of the underlying distributions by a marginal-wise backwards transformation of the simulated multivariate distribution with the quantile functions of the originally estimated marginal cdf's. This emphasizes the fact that unfavourable VaR estimates cannot perhaps be characterized by the copula structure alone but that the interplay between the dependence structure and the marginal distributions is also essential (see the discussion in [17]). Such a kind of interplay could potentially also be considered in the present approach, allowing non-constant negative pairwise correlations in the matrix Σ_d for the Gaussian copula in Lemma 5.

Note that Value at Risk is not the only risk measure that is used for calculating capital requirements in Europe. For instance, the Swiss Solvency Test uses the Expected Shortfall (ES) $\mathrm{ES}_\alpha(X)$ of risks X as the underlying risk measure (cf. [13]). In accordance with our terminology and under the assumption of a continuous risk distribution, it is defined as

$$\mathrm{ES}_\alpha(X) = \mathbb{E}\big(X \mid X > \mathrm{VaR}_\alpha(X)\big), \quad 0 < \alpha < 1.$$

Unfortunately, it is impossible to generate true unfavourable ES scenarios since ES is a coherent (i.e. subadditive) risk measure which in the worst case generates additive risk scenarios if the risks involved follow a comonotone dependence structure (see, e.g., [21, Chapter 7.2]). Note, however, that it is sufficient for the generation of additive ES scenarios to use a dependence structure as in Lemma 1 with the upper Fréchet bound for $\mathbb{V}$, which is a copula in any dimension.

Acknowledgements: We would like to thank the referees for several helpful comments that improved the presentation of the paper essentially.

Conflict of interest statement: Authors state no conflict of interest.

Appendix: Proofs of Lemmata

Proof of Lemma 2. We have

$$\begin{aligned} F_{Z_{1i}}(x,\beta) &= \mathbb{P}\big(Q_i\big((1-\beta)U_i\big) \le x\big) = \mathbb{P}\big((1-\beta)U_i \le F_i(x)\big) \\ &= \mathbb{P}\left(U_i \le \frac{F_i(x)}{1-\beta}\right) = \frac{F_i(x)}{1-\beta}, \quad 0 \le x \le Q_i(1-\beta), \end{aligned}$$

and

$$\begin{aligned} F_{Z_{2i}}(x,\beta) &= \mathbb{P}\big(Q_i\big(1-\beta+\beta V_i\big) \le x\big) = \mathbb{P}\big(1-\beta+\beta V_i \le F_i(x)\big) \\ &= \mathbb{P}\left(V_i \le \frac{F_i(x)+\beta-1}{\beta}\right) = \frac{F_i(x)+\beta-1}{\beta}, \quad x \ge Q_i(1-\beta), \quad i = 1, 2. \end{aligned}$$

□

Proof of Lemma 3. In the finite interval case, we have, by the usual convolution formula,

$$h_1(x) = \int\limits_{\substack{0 \le y \le M \\ 0 \le x-y \le M}} f(x-y)g(y)\,\mathrm{d}y = \int\limits_{\max(0,x-M) \le y \le \min(x,M)} f(x-y)g(y)\,\mathrm{d}y.$$

Now for $0 \le x \le M$, we have $\max(0, x-M) = 0$, $\min(x, M) = x$, from which the upper formula in brackets for $h_1(x)$ follows. For $M \le x \le 2M$, we have $\max(0, x-M) = x-M$, $\min(x, M) = M$, from which the lower formula in brackets for $h_1(x)$ follows.

The proof for the infinite interval case is analogous, observing that for $x \ge 2M$, we have

$$h_2(x) = \int\limits_{\substack{M \le y \le x \\ M \le x-y}} f(x-y)g(y)\,\mathrm{d}y = \int\limits_{M \le y \le x-M} f(x-y)g(y)\,\mathrm{d}y.$$

Further, under the conditions made, we have, in either case,

$$\frac{\mathrm{d}}{\mathrm{d}x} F \star G(x)\Big|_{x=2M} = h_1(2M) = h_2(2M) = \int_M^M f(x-y)g(y)\,\mathrm{d}y = 0,$$

as stated. □

Proof of Lemma 4. Let ξ_i and ζ_i be independent random variables with the cdf's $\underline{F}(\bullet,\beta)$ and $\overline{F}(\bullet,\beta)$, respectively. Then $I\xi_i + (1-I)\big(Q(1-\beta)+\zeta_i\big)$ is a stochastic representation of X_i, $i = 1,\dots,d$, where again I is a binomial random variable with $\mathbb{P}(I=1) = 1-\beta$ and $\mathbb{P}(I=0)=\beta$, independent of $(\mathbf{U},\mathbf{V})$ according to Lemma 2. Hence

$$I\sum_{i=1}^{d}\xi_i + (1-I)\sum_{i=1}^{d}\big(Q(1-\beta)+\zeta_i\big) = I\sum_{i=1}^{d}\xi_i + (1-I)\left(dQ(1-\beta)+\sum_{i=1}^{d}\zeta_i\right)$$

is a stochastic representation of S. Note that the cdf of $\sum_{i=1}^{d}\xi_i$ is $\underline{F}^{d\star}(\bullet,\beta)$ and that of $\sum_{i=1}^{d}\zeta_i$ is $\overline{F}^{d\star}(\bullet,\beta)$, from which the assertion follows. □

Proof of Lemma 5. The proof relies on the following two relations:

a) $\displaystyle\sum_{k=2}^{d}\frac{1}{k(k-1)} = \frac{d-1}{d}$ for all $d \geq 2$;

b) $\displaystyle\frac{i-1}{i} + \sum_{k=1}^{d-i}\frac{1}{(i+k)(i+k-1)} = \frac{d-1}{d}$ for all $d \geq 2$ and $1 \leq i \leq d$.

Clearly a) follows easily by induction. Relation b) follows immediately from a) since

$$\frac{i-1}{i} = \sum_{k=2}^{i}\frac{1}{k(k-1)} \quad\text{and}\quad \sum_{k=1}^{d-i}\frac{1}{(i+k)(i+k-1)} = \sum_{k=i+1}^{d}\frac{1}{k(k-1)}.$$

To prove Lemma 5, we first show that $\mathbf{T}\mathbf{T}^{tr} = \mathbf{I}_d = \mathbf{T}^{tr}\mathbf{T}$. Let $\mathbf{T}\mathbf{T}^{tr} = [b_{ij}]_{i,j=1,\dots,d}$. For $1 \leq i \leq d$, we obtain, by relation b) above,

$$b_{ii} = \frac{1}{d} + \frac{i-1}{i} + \sum_{k=1}^{d-i}\frac{1}{(i+k)(i+k-1)} = 1.$$

For $1 \leq i,j \leq d$ with $i \neq j$ we get, with $i \vee j := \max(i,j)$, following again relation b),

$$\begin{aligned} b_{ij} &= \frac{1}{d} - \frac{1}{i\vee j} + \sum_{k=i\vee j+1}^{d}\frac{1}{k(k-1)} = \frac{1}{d} - \frac{1}{i\vee j} + \sum_{k=1}^{d-i\vee j}\frac{1}{(k+i\vee j)(k+i\vee j-1)} \\ &= \frac{1}{d} - \frac{1}{i\vee j} + \frac{d-1}{d} - \frac{i\vee j-1}{i\vee j} = 1-1 = 0. \end{aligned}$$

This proves $\mathbf{T}\mathbf{T}^{tr} = \mathbf{I}_d$. On the other hand, let $\mathbf{T}^{tr}\mathbf{T} = [c_{ij}]_{i,j=1,\dots,d}$. It is obvious that $c_{11} = \frac{1}{d}d = 1$ and for all $2 \leq i \leq d$, $c_{ii} = \frac{1}{i(i-1)}(i-1) + \frac{i-1}{i} = 1$. Next, for all $2 \leq j \leq d$, we obtain

$$c_{1j} = \frac{1}{\sqrt{d}}\left(-\frac{1}{\sqrt{j(j-1)}}(j-1) + \sqrt{\frac{j-1}{j}}\right) = 0,$$

and for all $2 \leq i \leq d$, we get

$$c_{i1} = \frac{1}{\sqrt{d}}\left(-\frac{1}{\sqrt{i(i-1)}}(i-1) + \sqrt{\frac{i-1}{i}}\right) = 0.$$

Finally, for $2 \leq i,j \leq d$ with $i \neq j$, we get

$$c_{ij} = -\frac{1}{\sqrt{(i\vee j)(i\vee j-1)}}\left(-\frac{1}{\sqrt{(i\vee j)(i\vee j-1)}}(i\vee j-1) + \sqrt{\frac{i\vee j-1}{i\vee j}}\right) = 0.$$

This proves $\mathbf{T}^{tr}\mathbf{T} = \mathbf{I}_d$.

Now let $\lambda_1 = 1+(d-1)r$, $\lambda_i = 1-r$, $i = 2,\dots,d$, and $\Delta_t = \operatorname{diag}(\lambda_1 - t,\dots,\lambda_d - t)$. A standard computation yields, for $t \in \mathbb{R}$, $\mathbf{T}\Delta_t = [s_{ij}]_{i,j=1,\dots,d}$, where

$$s_{ij} = \begin{cases} \frac{1+(d-1)r-t}{\sqrt{d}}, & j = 1, \\ -\frac{1-r-t}{\sqrt{j(j-1)}}, & 1 \leq i < j, \\ \sqrt{\frac{i-1}{j}}\,(1-r-t), & 1 < i = j, \\ 0, & \text{otherwise.} \end{cases}$$

Let $\mathbf{T}\Delta_t\mathbf{T}^{tr} = [d_{ij}]_{i,j=1,\dots,d}$. From relation a) above it follows that

$$d_{11} = \frac{1+(d-1)r-t}{d} + (1-r-t)\sum_{k=2}^{d}\frac{1}{k(k-1)} = \frac{1+(d-1)r-t}{d} + (1-r-t)\,\frac{d-1}{d} = 1-t,$$

and for $2 \le i \le d$, relation b) gives

$$\begin{aligned} d_{ii} &= \frac{1+(d-1)r-t}{d} + (1-r-t)\left(\frac{i-1}{i} + \sum_{k=1}^{d-i}\frac{1}{(i+k)(i+k-1)}\right) \\ &= \frac{1+(d-1)r-t}{d} + (1-r-t)\,\frac{d-1}{d} = 1-t. \end{aligned}$$

Next, for $2 \le i, j \le d$ with $i \ne j$ we obtain from relation b) above that

$$\begin{aligned} d_{ij} &= \frac{1+(d-1)r-t}{d} - \frac{1-r-t}{i\vee j} + (1-r-t)\left(\sum_{k=1}^{d-i\vee j}\frac{1}{(i\vee j+k)(i\vee j+k-1)}\right) \\ &= \frac{1+(d-1)r-t}{d} - \frac{1-r-t}{i\vee j} + (1-r-t)\left(\frac{d-1}{d} - \frac{i\vee j-1}{i\vee j}\right) = r. \end{aligned}$$

This in turn means $\mathbf{T}\Delta_t\mathbf{T}^{tr} = \Sigma_d - t\mathbf{I}_d$. Consequently, the characteristic polynomial for Σ_d is given by

$$\begin{aligned} \varphi_{\Sigma_d}(t) &= \det(\Sigma_d - t\mathbf{I}_d) = \det(\mathbf{T}\Delta_t\mathbf{T}^{tr}) = \det(\mathbf{T})\,\det(\Delta_t)\,\det(\mathbf{T}^{tr}) \\ &= \det(\mathbf{T})\,\det(\Delta_t)\,\det(\mathbf{T}^{-1}) = \det(\Delta_t) = \prod_{i=1}^{d}(\lambda_i - t). \end{aligned}$$

Hence λ_i, $1 \le i \le d$, are the latent roots of Σ_d. Therefore, Σ_d is a correlation matrix, i.e. positive semidefinite iff $\lambda_i \ge 0$ for all $1 \le i \le d$, i.e. $-\frac{1}{d-1} \le r \le 1$. Thus Lemma 5 is proved. □

Conflict of interest statement: Authors state no conflict of interest.

References

[1] Arbenz, P., C. Hummel, and G. Mainik (2012). Copula based hierarchical risk aggregation through sample reordering. *Insurance Math. Econ. 51*(1), 122–133.

[2] Basilevsky, A. (1983). *Applied Matrix Algebra in the Statistical Sciences.* North-Holland, New York.

[3] Boonen, T. J. (2017). Solvency II solvency capital requirement for life insurance companies based on expected shortfall. *Eur. Actuar. J. 7*, 405–434.

[4] Cadoni, P. (Ed.) (2014). *Internal Models and Solvency II: From Regulation to Implementation.* RISK Books, London.

[5] Cottin, C. and D. Pfeifer (2014). From Bernstein polynomials to Bernstein copulas. *J. Appl. Funct. Anal. 9*(3-4), 277–288.

[6] Cruz, M. (Ed.) (2009). *The Solvency II Handbook: Developing ERM Frameworks in Insurance and Reinsurance Companies.* RISK Books, London.

[7] Doff, R. (2011). *Risk Management for Insurers: Risk Control, Economic Capital and Solvency II,*. Second edition. RISK Books, London.

[8] Doff, R. (Ed.) (2014). *The Solvency II Handbook: Practical Approaches to Implementation.* RISK Books, London.

[9] Durante, F. and C. Sempi (2016). *Principles of Copula Theory.* CRC Press, Boca Raton FL.

[10] Embrechts, P., G. Puccetti, and L. Rüschendorf (2013). Model uncertainty and VaR aggregation. *J. Bank. Financ. 37*(8), 2750–2764.

[11] European Parliament and the Council (2009). Directive 2009/138/EC of the European Parliament and of the Council of 25 November 2009 on the taking-up and pursuit of the business of Insurance and Reinsurance (Solvency II) (recast). *Official Journal of the European Union* 17.12.2009: L335/1–L335/155.

[12] European Parliament and the Council (2015). Commission Delegated Regulation EU 2015/35 of 10 October 2014 supplementing Directive 2009/138/EC of the European Parliament and of the Council on the taking-up and pursuit of the business of Insurance and Reinsurance (Solvency II). *Official Journal of the European Union* 17.1.2015: L12/1–L12/797.

[13] Federal Office of Private Insurance (FINMA) (2006). Technical document on the Swiss Solvency Test. Available at https://www.finma.ch/FinmaArchiv/bpv/download/e/SST_techDok_061002_E_wo_Li_20070118.pdf.

[14] Hampel, M. and D. Pfeifer (2011). Proposal for correction of the SCR calculation bias in Solvency II. *ZVersWiss 100*, 733–743.
[15] Hummel, C. (2018). Shaping tail dependencies by nesting box copulas. Available at https://arxiv.org/abs/0906.4853v2.
[16] Hürliman, W. (2009). Optimisation of the non-life insurance risk diversification in Solvency II. In M. Cruz (Ed.), *The Solvency II Handbook: Developing ERM Frameworks in Insurance and Reinsurance Companies,* Ch. 12, RISK Books, London.
[17] Ibragimov, R. and A. Prokhorov (2017). *Heavy Tails and Copulas: Topics in Dependence Modelling in Economics and Finance.* World Scientific, Singapore.
[18] Jaworski, P. and T. Rychlik (2008). On distributions of order statistics for absolutely continuous copulas with applications to reliability. *Kybernetika 44*(6), 757–776.
[19] Jaworski, P. (2009). On copulas and their diagonals. *Inform. Sci. 179*(17), 2863–2871.
[20] Mainik, G. (2015). Risk aggregation with empirical margins: Latin hypercubes, empirical copulas, and convergence of sum distributions. *J. Multivariate Anal. 141*, 197–216.
[21] McNeil, A. J., R. Frey, and P. Embrechts (2015). *Quantitative Risk Management: Concepts, Techniques and Tools.* Revised edition. Princeton University Press.
[22] Nelsen, R. B. (2006). *An Introduction to Copulas.* Second edition. Springer, New York.
[23] Ozdemir, B. (2015). *ORSA: Design and Implementation.* RISK Books, London.
[24] Pfeifer, D. (2013). Correlation, tail dependence and diversification. In C. Becker, R. Fried, and S. Kuhnt (Eds.), *Robustness and Complex Data Structures: Festschrift in Honour of Ursula Gather,* pp. 301–314. Springer, Berlin.
[25] Pfeifer, D., A. Mändle, and O. Ragulina (2017). New copulas based on general partitions-of-unity and their applications to risk management (part II). *Depend. Model. 5*, 246–255.
[26] Pfeifer, D., A. Mändle, O. Ragulina, and C. Girschig (2019). New copulas based on general partitions-of-unity (part III) – the continuous case. *Depend. Model. 7*, 181–201.
[27] Pfeifer, D. and O. Ragulina (2018). Generating VaR scenarios under Solvency II with product beta distributions. *Risks 6*(4), Article ID 122, 15 pages.
[28] Pfeifer, D. and O. Ragulina (2020). Adaptive Bernstein copulas and risk management. *Mathematics 8*(12), Article ID 2221, 22 pages.
[29] Pfeifer, D. and D. Strassburger (2008). Solvency II: stability problems with the SCR aggregation formula. *Scand. Actuar. J. 2008*(1), 61–77.
[30] Pfeifer, D., H. A. Tsatedem, A. Mändle, and C. Girschig (2016). New copulas based on general partitions-of-unity and their applications to risk management. *Depend. Model. 4*, 123–140.
[31] Sandström, A. (2011). *Handbook of Solvency for Actuaries and Risk Managers: Theory and Practice.* CRC Press, London.
[32] Varnell, E. M. (2011). Economic scenario generators and Solvency II. *British Actuar. J. 16*(1), 121–159.

DE GRUYTER Depend. Model. 2021; 9:347–373

Research Article **Open Access**
Special Issue in memory of Abe Sklar

Susanne Saminger-Platz, Anna Kolesárová, Adam Šeliga, Radko Mesiar, and Erich Peter Klement*

New results on perturbation-based copulas

https://doi.org/10.1515/demo-2021-0116
Received May 30, 2021; accepted September 20, 2021

Abstract: A prominent example of a perturbation of the bivariate product copula (which characterizes stochastic independence) is the parametric family of Eyraud-Farlie-Gumbel-Morgenstern copulas which allows small dependencies to be modeled. We introduce and discuss several perturbations, some of them perturbing the product copula, while others perturb general copulas. A particularly interesting case is the perturbation of the product based on two functions in one variable where we highlight several special phenomena, e.g., extremal perturbed copulas. The constructions of the perturbations in this paper include three different types of ordinal sums as well as flippings and the survival copula. Some particular relationships to the Markov product and several dependence parameters for the perturbed copulas considered here are also given.

Keywords: copula, dependence parameter, Eyraud-Farlie-Gumbel-Morgenstern copula, ordinal sum, perturbation

MSC: 60E05, 62H05, 62H20

Dedicated to the memory of Professor Abe Sklar: we would like to pay homage to a great mathematician who made deep and innovative contributions to probabilistic metric spaces and who founded the theory of copulas.

1 Introduction

The earliest use of what is now called *perturbation theory* was to deal with otherwise unsolvable mathematical problems of celestial mechanics. When Kepler published his first law "*the orbit of every planet is an ellipse with the Sun at one of the two foci*," in [66] and [67, book 5, part 1, III. De Figura Orbitæ]) at the beginning of the 17th century, he provided an analytical solution of a classical *two-body problem*, the two bodies being the planet under consideration and the Sun — in this scenario no perturbation occurred. Many decades later, when *three-body problems* were studied, e.g., the system Moon–Earth–Sun [59], one observed that the Moon (which has a much smaller mass than both Sun and Earth) does not move along a simple ellipse à la Kepler because of the competing gravitation of the Earth and the Sun, i.e., the constants describing the motion of a planet around the Sun are also influenced by the motion of other planets and may vary in time. In addition, in the second half of the 19th century the increasing accuracy of astronomical observations also required a

Susanne Saminger-Platz: Institute for Mathematical Methods in Medicine and Data Based Modeling, Johannes Kepler University, Linz, Austria, E-mail: susanne.saminger-platz@jku.at, ORCID 0000-0002-9606-5751
Anna Kolesárová: Institute of Information Engineering, Automation and Mathematics, Faculty of Chemical and Food Technology, Slovak University of Technology, Bratislava, Slovakia, E-mail: anna.kolesarova@stuba.sk, ORCID 0000-0001-6438-0039
Adam Šeliga: Department of Mathematics and Descriptive Geometry, Faculty of Civil Engineering, Slovak University of Technology, Bratislava, Slovakia, E-mail: adam.seliga@stuba.sk, ORCID 0000-0002-3215-6195
Radko Mesiar: Department of Mathematics and Descriptive Geometry, Faculty of Civil Engineering, Slovak University of Technology, Bratislava, Slovakia, and Institute for Research and Applications of Fuzzy Modeling, NSC IT4Innovations, University of Ostrava, Czech Republic, E-mail: radko.mesiar@stuba.sk, ORCID 0000-0001-6503-949X
***Corresponding Author: Erich Peter Klement:** Institute for Mathematical Methods in Medicine and Data Based Modeling, Johannes Kepler University, Linz, Austria, E-mail: ep.klement@jku.at, ORCID 0000-0002-2054-3286

higher accuracy of the solutions to Newton's gravitational equations [95] and motivated mathematicians such as Lagrange [80] and Laplace [81] to develop and study fundamental methods of perturbation theory (for a classical survey see, e.g., [9]).

In mathematics, physics, and chemistry, perturbation theory deals with mathematical methods for finding an approximate solution to a problem, by starting from the exact solution of a related, simpler problem. A critical feature of the technique is a middle step that breaks the problem into "solvable" and "perturbative" parts. Perturbation theory is widely used when the problem at hand does not have a known exact solution, but can be expressed as a "small" change to a known solvable problem.

In quantum mechanics (see [32] and [12]), perturbation theory can be used to describe a complicated unsolved system using a simple, solvable system. Starting with such a simple system whose mathematical solution is known, one may add a weak disturbance to the system (a so-called "Hamiltonian" [7]). If this disturbance is not too large, physical quantities related to the perturbed system (e.g., its energy levels) can be seen as "corrections" to those of the simple system. If these corrections are small compared to the size of the quantities themselves, approximate methods such as asymptotic series can help to calculate them.

Now let us turn to the concept of bivariate copulas (introduced in [119], see also [39, 64, 93]), i.e., to mathematical objects which capture the dependence structure among random variables and which are the topic of our current research. In this context, perturbation usually means that a (small) bivariate function (often called the perturbation factor) is added to a given copula, and one is interested to find out under which conditions the result is again a copula [36, 98]. A prominent example of such a perturbation is the family of Eyraud-Farlie-Gumbel-Morgenstern (or EFGM) copulas given in (2.7), where a parameterized family of perturbation factors is added to the product copula Π (for more details about EFGM copulas and for other families of copulas which can be considered as perturbations see [39, 68, 93]).

In this paper, we investigate two different types of perturbations of general copulas which will be introduced in Definitions 4.1 and 5.8 (where we can identify an interesting class of extremal elements), respectively.

In these constructions we make use of several techniques which are well-known in the theory of copulas, such as x- and y-flippings, the construction of the survival copula, and three variants of the concept of ordinal sums, to name just a few. We study in detail a number of mathematical properties of these perturbations and some interrelations between them. Some of the perturbations discussed here induce interesting extensions of the family of Eyraud-Farlie-Gumbel-Morgenstern copulas (compare [104]). Finally, some relationships to the Markov product of copulas are emphasized, and the values of four dependence parameters (Spearman's rho, Kendall's tau, Blomqvist's beta, and Gini's gamma) of our perturbations are given.

2 Preliminaries

Copulas are mathematical objects capturing the dependence structure among random variables. The name "copula" for functions linking an n-dimensional distribution and its one-dimensional marginals goes back to Sklar's paper [119] (compare also [120]), where he proved (for the case $n = 2$) a result which is now referred to as Sklar's Theorem. However, links between multivariate distributions and their one-dimensional marginals have been studied before, e.g., by Hoeffding [60, 61], Fréchet [50], Dall'Aglio [23–25], and Féron [48], and also later on without any reference to the concept of copulas (see, e.g., [108, 123]).

Formally, a (bivariate) copula [119] is a function $C\colon [0,1]^2 \to [0,1]$ which satisfies, for each $x \in [0,1]$, the boundary conditions

$$C(x,0) = C(0,x) = 0 \qquad \text{and} \qquad C(x,1) = C(1,x) = x, \tag{C1}$$

and which is 2-increasing, i.e., for all $(x,y), (x^\star, y^\star) \in [0,1]$ with $(x,y) \le (x^\star, y^\star)$

$$C(x,y) + C(x^\star, y^\star) - C(x, y^\star) - C(x^\star, y) \ge 0. \tag{C2}$$

The set of all bivariate copulas will be denoted by $\mathscr{C}$.

There are infinitely many elements in $\mathscr{C}$: in the books [63, 93] and, more recently, [39] one finds plenty of examples of parametric families (usually with one or two parameters) of copulas, on the one hand, and classes of copulas which can be constructed and characterized by functions in one variable (e.g., by additive and/or multiplicative generators [1, 83, 111, 112] in the case of Archimedean copulas), on the other hand.

Sklar's Theorem establishing bivariate copulas as the link between bivariate probability distributions and their marginals was given in [119] (see also [39, 64, 93]).

In brief, Sklar's Theorem states that, whenever (X, Y) is a random vector with its two marginal distributions $F_X, F_Y\colon \mathbb{R} \to [0,1]$, then there exists a copula $C_{X,Y} \in \mathscr{C}$ (which is uniquely determined if and only if X and Y are continuous) such that the joint distribution $F_{X,Y}\colon \mathbb{R}^2 \to [0,1]$ is given by

$$F_{X,Y}(u,v) = C_{X,Y}(F_X(u), F_Y(v)). \tag{2.1}$$

Conversely, for each copula $C \in \mathscr{C}$ the function $F\colon \mathbb{R}^2 \to [0,1]$ given by

$$F(u,v) = C(F_X(u), F_Y(v))$$

is a two-dimensional probability distribution of the random vector (X, Y) such that $C_{X,Y} = C$.

Sklar's Theorem also holds for the general case of n-dimensional probability distributions (proofs for this general case and alternative proofs for the bivariate case can be found in [5, 33–35, 43, 46, 96, 100]).

In the language of Sklar's Theorem the following three basic copulas describe a pair of independent or comonotone dependent or countermonotone dependent random variables X and Y, respectively: the product copula $\Pi\colon [0,1]^2 \to [0,1]$ and the Fréchet-Hoeffding lower and upper bounds $W\colon [0,1]^2 \to [0,1]$ and $M\colon [0,1]^2 \to [0,1]$ given by, respectively, $\Pi(x,y) = xy$, $W(x,y) = \max(x+y-1, 0)$, and $M(x,y) = \min(x,y)$.

If $C\colon [0,1]^2 \to [0,1]$ is an arbitrary copula then several related copulas can be considered, e.g., the x-flipping $C^{x\text{flip}}\colon [0,1]^2 \to [0,1]$ and the y-flipping $C^{y\text{flip}}\colon [0,1]^2 \to [0,1]$ of C, and the survival copula $C^{\text{surv}}\colon [0,1]^2 \to [0,1]$, which are defined (see [30, 93]) by, respectively,

$$C^{x\text{flip}}(x,y) = y - C(1-x,y), \qquad C^{y\text{flip}}(x,y) = x - C(x,1-y), \qquad C^{\text{surv}}(x,y) = x+y-1+C(1-x,1-y). \tag{2.2}$$

If X and Y are two continuous random variables and if $C_{X,Y} \in \mathscr{C}$ is the (unique) copula satisfying (2.1) in Sklar's Theorem then the following stochastic interpretation is an immediate consequence of [93, Theorem 2.4.4]:

$$(C_{X,Y})^{x\text{flip}} = C_{-X,Y}, \qquad (C_{X,Y})^{y\text{flip}} = C_{X,-Y}, \qquad (C_{X,Y})^{\text{surv}} = C_{-X,-Y}. \tag{2.3}$$

As an immediate consequence of (2.3), the x-flipping, the y-flipping and the construction of the survival copula (the latter being the composition of x-flipping and y-flipping) are involutive operations on $\mathscr{C}$, i.e., for each copula $C \in \mathscr{C}$ we have

$$C^{\text{surv}} = (C^{x\text{flip}})^{y\text{flip}} = (C^{y\text{flip}})^{x\text{flip}} \qquad \text{and} \qquad (C^{x\text{flip}})^{x\text{flip}} = (C^{y\text{flip}})^{y\text{flip}} = (C^{\text{surv}})^{\text{surv}} = C. \tag{2.4}$$

For the three basic copulas W, Π, and M the following relationships can be verified easily:

$$W^{x\text{flip}} = W^{y\text{flip}} = M, \qquad \Pi^{x\text{flip}} = \Pi^{y\text{flip}} = \Pi, \qquad M^{x\text{flip}} = M^{y\text{flip}} = W.$$

A copula $C \in \mathscr{C}$ which is invariant with respect to the construction of survival copulas, i.e., which satisfies $C^{\text{surv}} = C$, is also called radially symmetric [55] (see also [6, 8, 19]). The three basic copulas W, Π and M are trivial examples of radially symmetric copulas.

Given a copula $C \in \mathscr{C}$, we sometimes will work with a distinguished section of it, the so-called opposite diagonal section $\omega_C\colon [0,1] \to \mathbb{R}$ defined by

$$\omega_C(x) = C(x, 1-x). \tag{2.5}$$

There is an axiomatization for a function $\omega\colon [0,1] \to [0,1]$ to be an opposite diagonal section of copulas, and for each such ω there exists at least one $C \in \mathscr{C}$ such that $\omega_C = \omega$ (see, e.g., [31, 45, 52, 53]).

In Section 6, we shall be concerned with several dependence parameters of a copula $C \in \mathscr{C}$, in particular with Spearman's rho [121], Kendall's tau [65], Blomqvist's beta [11], and Gini's gamma [56] which can be defined for each copula $C \in \mathscr{C}$ and assume their values in the interval $[-1, 1]$. The corresponding functions $\varrho, \tau, \beta, \gamma\colon \mathscr{C} \to [-1, 1]$ are given by (see, e.g., [93]), respectively:

$$\varrho(C) = 12 \iint_{[0,1]^2} C(x, y)\,dx\,dy - 3, \qquad \tau(C) = 4 \iint_{[0,1]^2} C(x, y)\,dC(x, y) - 1,$$
$$\beta(C) = 4\,C(\tfrac{1}{2}, \tfrac{1}{2}) - 1, \qquad \gamma(C) = 4 \int_0^1 (C(x, x) + C(x, 1 - x))\,dx - 2. \tag{2.6}$$

For the three basic copulas W, Π and M and for each function $\xi\colon \mathscr{C} \to [-1, 1]$ such that $\xi \in \{\varrho, \tau, \beta, \gamma\}$ we obtain the special values $\xi(W) = -1$, $\xi(\Pi) = 0$, and $\xi(M) = 1$.

From Theorems 5.1.1 and 5.1.9 in [93] (see also Definition 5.1.7 in this monograph) it follows that for each copula $C \in \mathscr{C}$ and for each function $\xi\colon \mathscr{C} \to [-1, 1]$ such that $\xi \in \{\varrho, \tau, \beta, \gamma\}$ we get for the x-flipping $C^{x\text{flip}}$ and the y-flipping $C^{y\text{flip}}$ of C and for the survival copula C^{surv} given by (2.2):

$$\xi\big(C^{x\text{flip}}\big) = \xi\big(C^{y\text{flip}}\big) = -\xi\big(C^{\text{surv}}\big) = -\xi(C).$$

A particularly interesting and important family of copulas, which is used quite often when the weak dependence of exchangeable random variables should be modeled, is $(C_\theta^{\mathbf{EFGM}})_{\theta\in[-1,1]}$, where each $C_\theta^{\mathbf{EFGM}}\colon [0, 1]^2 \to [0, 1]$ is defined by

$$C_\theta^{\mathbf{EFGM}}(x, y) = xy + \theta xy(1 - x)(1 - y). \tag{2.7}$$

This family was usually referred to as the family of *Farlie-Gumbel-Morgenstern copulas* [42, 58, 92, 93]. In [39] (see also [13, 14, 91]) it was pointed out that the corresponding distributions had already been studied in the earlier and, for many years, forgotten work by Eyraud [41]. In recognition of that we will consistently use the name *Eyraud-Farlie-Gumbel-Morgenstern copulas* (*EFGM copulas* for short) in this paper.

Recall the concept of an M-ordinal sum which was introduced first for triangular norms (see [109–111]) in [112] and then also for copulas [49]. This construction was based on earlier results for partially ordered sets [10, 40] and abstract semigroups [15–18] (compare also [3, 72, 74, 83, 113]).

In an analogous way, another type of ordinal sums of copulas based on the lower Fréchet-Hoeffding bound W was suggested in [90] (see also [39]).

A third type of ordinal sum of copulas, the so-called (vertical) Π-ordinal sums, will be used in Section 5. They were originally introduced in [76] as a generalization of some patchwork techniques [28, 37, 38] and of the gluing of copulas proposed in [118].

In each of the three cases we start with an arbitrary family $(]a_k, b_k[)_{k\in K}$ of non-empty, pairwise disjoint open subintervals of $[0, 1]$ and with an arbitrary family $(C_k)_{k\in K}$ of copulas. Then each of the three functions $C^{M\text{os}}, C^{W\text{os}}, C^{\Pi\text{os}}\colon [0, 1]^2 \to [0, 1]$ defined by, respectively,

$$C^{M\text{os}}(x, y) = \begin{cases} a_k + (b_k - a_k)C_k\big(\frac{x-a_k}{b_k-a_k}, \frac{y-a_k}{b_k-a_k}\big) & \text{if } (x, y) \in [a_k, b_k]^2, \\ M(x, y) & \text{otherwise,} \end{cases} \tag{2.8}$$

$$C^{W\text{os}}(x, y) = \begin{cases} (b_k - a_k)C_k\big(\frac{x-a_k}{b_k-a_k}, \frac{y+b_k-1}{b_k-a_k}\big) & \text{if } (x, y) \in [a_k, b_k] \times [1 - b_k, 1 - a_k], \\ W(x, y) & \text{otherwise,} \end{cases} \tag{2.9}$$

$$C^{\Pi\text{os}}(x, y) = \begin{cases} a_k y + (b_k - a_k)C_k\big(\frac{x-a_k}{b_k-a_k}, y\big) & \text{if } x \in [a_k, b_k], \\ \Pi(x, y) & \text{otherwise,} \end{cases} \tag{2.10}$$

is a well-defined copula.

We call $C^{M\text{os}}$ the *M-ordinal sum*, $C^{W\text{os}}$ the *W-ordinal sum*, and $C^{\Pi\text{os}}$ the (*vertical*) *Π-ordinal sum* of the summands $(]a_k, b_k[, C_k)_{k\in K}$, and we shall write

$$C^{M\text{os}} = M\text{-}(\langle a_k, b_k, C_k\rangle)_{k\in K}, \quad C^{W\text{os}} = W\text{-}(\langle a_k, b_k, C_k\rangle)_{k\in K}, \quad C^{\Pi\text{os}} = \Pi\text{-}(\langle a_k, b_k, C_k\rangle)_{k\in K}.$$

An ordinal sum $M\text{-}(\langle a_k, b_k, C_k\rangle)_{k\in K}$, $W\text{-}(\langle a_k, b_k, C_k\rangle)_{k\in K}$ or $\Pi\text{-}(\langle a_k, b_k, C_k\rangle)_{k\in K}$ is said to be non-trivial if the family of open intervals $(]a_k, b_k[)_{k\in K}$ does not consist of $]0, 1[$ only, i.e., if $\{]a_k, b_k[\mid k \in K\} \neq \{]0, 1[\}$.

Another extremal case which is covered by (2.8)–(2.10) is that of an empty index set: the empty ordinal sums $M\text{-}(\langle a_k, b_k, C_k\rangle)_{k\in\emptyset}$, $W\text{-}(\langle a_k, b_k, C_k\rangle)_{k\in\emptyset}$ and $\Pi\text{-}(\langle a_k, b_k, C_k\rangle)_{k\in\emptyset}$ coincide with M, W and Π, respectively.

Some deeper investigations of these and other types of ordinal sums can be found in [102] (see also, e.g., [38, 40, 69, 70, 74, 101]).

3 Some known results on perturbations of copulas

In a number of papers, various perturbations of copulas were introduced and studied from different points of view. In most cases, the authors fixed a copula $D \in \mathscr{C}$ and a suitable bivariate function $H\colon [0, 1]^2 \to \mathbb{R}$, and looked for conditions on H (and D) guaranteeing that also the function $C\colon [0, 1]^2 \to \mathbb{R}$ given by

$$C(x, y) = D(x, y) + H(x, y) \tag{3.1}$$

was a copula.

Remark 3.1. Several interesting variants and special cases of this general approach were investigated in detail. If not stated otherwise, $D\colon [0, 1]^2 \to [0, 1]$ always denotes a given copula.

(i) In [87] the following functional equation was studied:

$$C_\theta(x, y) = D(x, y) + \theta(x - D(x, y))(y - D(x, y)). \tag{3.2}$$

Observe that for $D = \Pi$ and $\theta \in [-1, 1]$ we obtain the Eyraud-Farlie-Gumbel-Morgenstern copula $C_\theta = C_\theta^{\mathbf{EFGM}}$ as given in (2.7).

(ii) Put $D = M$ and $\theta \in [0, 1]$. Then it was mentioned in [87] that

$$C_\theta(x, y) = \max((1 - \theta)M(x, y) + \theta(x + y - 1), 0) \tag{3.3}$$

yields a copula.

(iii) As a variant of (ii), for a function $H\colon [0, 1]^2 \to \mathbb{R}$ also the functional equation

$$C(x, y) = \max(D(x, y) + H(x, y), 0) \tag{3.4}$$

was studied in [87] (compare also [77–79, 86]).

(iv) Let $f, g\colon [0, 1] \to \mathbb{R}$ be suitable functions and put

$$C(x, y) = D(x, y) + f(\max(x, y))g(\min(x, y)). \tag{3.5}$$

This situation was investigated in [36], while the special case $D = \Pi$ had already been discussed in [4].

(v) As a special case of (iv), a necessary and sufficient condition for the function

$$C(x, y) = xy + f(x)g(y), \tag{3.6}$$

to be a copula was given in [99, Theorem 2.3], and a simpler sufficient condition for being an absolutely continuous copula can be found in Theorem 2.5 in the same paper. Probably the simplest necessary and sufficient condition for C being a copula (see [39, Example 1.6.10]) is that the functions f and g are Lipschitz, vanish in 0 and 1 and satisfy $f'(x)g'(y) \geq -1$ for all $(x, y) \in [0, 1]^2$ for which the derivatives exist (compare also [39, Theorem 1.6.9]). In Section 5, we shall present some new results related to this type of perturbations.

(vi) As a generalization of (v), the parameterized function

$$C_\theta(x, y) = xy + \theta f(x)g(y), \tag{3.7}$$

was considered for the first time in the context of copulas in [105] and later in [82, 98, 117]. A generalized form of (3.7), where in the first summand on the right-hand side the product was replaced by an arbitrary copula, was discussed in [68]. For a survey and other generalizations of the cases (3.5)–(3.7) see [4].

(vii) For an interesting generalization of (vi), namely,

$$C(x,y) = \Pi(x,y) + \sum_{i=1}^{n} \theta_i f_i(x) g_i(y), \tag{3.8}$$

a necessary and sufficient condition for $\theta_i \in \mathbb{R}$ and $f_i, g_i \colon [0,1] \to \mathbb{R}$ with $i = 1, 2, \ldots, n$ was given in [88, Theorem 2.2] (also several sufficient conditions which are easier to check, together with a number of examples, were provided).

(viii) Let $N_1, N_2 \colon [0,1] \to [0,1]$ be two convex strong negations on $[0,1]$ (see, e.g., [72]). The problem, under which conditions on θ and N_1 and N_2 the function

$$C_\theta(x,y) = \max(xy + \theta N_1(x) N_2(y), 0) \tag{3.9}$$

is a copula was solved in [47] (see also [86]).

Remark 3.2. Other approaches to perturbations are based on weighted arithmetic and geometric means. For example, the weighted arithmetic mean of two arbitrary copulas is again a copula, and thus we can perturb a given copula D by means of some (arbitrary) copula C putting $E = (1-\varepsilon)D + \varepsilon C$, where $\varepsilon \in\,]0,1[$. In the case of the weighted geometric mean, it is known that the set of extreme-value copulas [57] is closed under this averaging operator. For some deeper study of this kind of problems see [20]. For some other types of perturbation, in particular for those connected with the diagonal expansion of a copula, we refer to [21]. More recently, other perturbations of copulas which are related to random noise were described and investigated in [89, 115, 116].

The family of EFGM copulas $\left(C_\theta^{\mathbf{EFGM}}\right)_{\theta \in [-1,1]}$ has many nice properties (see, e.g., [104]), but also some drawbacks. One of them is that the values of the dependence functions of the four dependence parameters given in (2.6) are bounded by $-\frac{2}{9}$ and $\frac{2}{9}$ for each EFGM copula, so only weak dependencies can be modeled by these copulas.

A number of extensions of the family of EFGM copulas has been presented in the literature in order to overcome this constraint. A comprehensive extension (i.e., containing the three basic copulas W, Π and M) was introduced in [62], other approaches were based on quadratic constructions of copulas [75] or on some forms of convexity (such as ultramodularity [85] and Schur concavity [107]), see [103]. A very interesting and natural extension of EFGM copulas are polynomial copulas [122] (in [114] polynomial copulas of degree five are studied in detail) which can also be seen as special cases of the general perturbation (3.1). Also some of the perturbations investigated in [36, 68, 77, 78, 87, 98, 105] turn out to be extensions of the family of EFGM copulas.

4 Copula-based perturbations of copulas

We now consider perturbations of an arbitrary copula $C \in \mathscr{C}$ by means of the opposite diagonal sections given by (2.5) of two copulas $C_1, C_2 \in \mathscr{C}$. The study of these perturbations was motivated by the investigations in [36] or [86, 87]. Again, in some special cases we obtain extensions of the family of EFGM copulas.

Definition 4.1. Let C, C_1 and C_2 be three arbitrary copulas and $\theta \in \mathbb{R}$. Then we consider the function $[C, C_1, C_2]_\theta \colon [0,1]^2 \to [0,1]$ given by

$$[C, C_1, C_2]_\theta(x,y) = C(x,y) + \theta\, C_1(x, 1-x)\, C_2(y, 1-y). \tag{4.1}$$

Note that (4.1) is a variant of (3.7) in Remark 3.1(vi): in the first summand on the right-hand side, the product is replaced by an arbitrary copula C, and in the second summand the functions f and g equal the opposite diagonals (given by (2.5)) ω_{C_1} and ω_{C_2}, respectively, of two copulas C_1 and C_2.

In order to find out under which conditions the function $[C, C_1, C_2]_\theta$ is a copula, we have to recall some measure theoretic facts about absolutely continuous and singular parts of copulas (compare [93, Section 2.4]).

If $\lambda\colon \mathscr{B}([0,1]^2) \to [0,1]$ denotes the Lebesgue measure on the σ-algebra $\mathscr{B}([0,1]^2)$ of Borel subsets of $[0,1]^2$ then, as a consequence of [93, Theorem 2.2.7], for each copula $C\colon [0,1]^2 \to [0,1]$ the mixed derivative $\frac{\partial^2 C(x,y)}{\partial x\,\partial y}$ exists almost everywhere on $[0,1]^2$ (with respect to λ). Therefore, the function $\psi_{A_C}\colon [0,1]^2 \to \mathbb{R}$ defined by

$$\psi_{A_C}(x,y) = \begin{cases} \frac{\partial^2 C(x,y)}{\partial x\,\partial y} & \text{if } \frac{\partial^2 C(x,y)}{\partial x\,\partial y} \text{ exists,} \\ 0 & \text{otherwise,} \end{cases}$$

is integrable over $[0,1]^2$, and the function $A_C\colon [0,1]^2 \to [0,1]$ defined by

$$A_C(x,y) = \int_0^x \int_0^y \psi_{A_C}(u,v)\,dv\,du$$

is absolutely continuous. This function A_C is called the absolutely continuous part of the copula C, and the function $S_C\colon [0,1]^2 \to [0,1]$ given by $S_C = C - A_C$ the singular part of the copula C (see [93, (2.4.1)]). Finally, put

$$\alpha_C = \operatorname{essinf}\{\psi_{A_C}(x,y) \mid (x,y) \in [0,1]^2\} = \sup\big\{b \in \mathbb{R} \bigm| \lambda(\{(x,y) \in [0,1]^2 \mid \psi_{A_C}(x,y) < b\}) = 0\big\}. \tag{4.2}$$

The following result provides a complete solution of the problem under which conditions the function $[C, C_1, C_2]_\theta$ is a copula.

Theorem 4.2. *Let $C\colon [0,1]^2 \to [0,1]$ be a copula, $A_C\colon [0,1]^2 \to [0,1]$ the absolutely continuous part of C, and α_C as given by (4.2). Then the following are equivalent:*
(i) *for all copulas $C_1, C_2\colon [0,1]^2 \to [0,1]$ the function $[C, C_1, C_2]_\theta\colon [0,1]^2 \to [0,1]$ defined by (4.1) is a copula;*
(ii) *$\theta \in [-\alpha_C, \alpha_C]$.*

Proof. Fix an arbitrary copula C and some $\theta \in \mathbb{R}$, and assume that condition (i) holds. Define the two functions $f, h\colon [0,1] \to [0,1]$ by $f(x) = \min(x, 1-x)$ and $h(x) = 2x - \lfloor 2x \rfloor$, where $\lfloor u \rfloor$ denotes the floor of the real number u, and the sequence of functions $(g_n\colon [0,1] \to [0,1))_{n\in\mathbb{N}}$ given inductively by $g_1 = f$ and $g_{n+1} = \frac{1}{2}(g_n \circ h)$. Observe that, for each $n \in \mathbb{N}$, the functions f and g_n are non-negative and continuous, and that they vanish at the boundaries of $[0,1]$, i.e., $f(0) = g_n(0) = f(1) = g_n(1) = 0$. Moreover, we have $\{f'(x), g_n'(y)\} \subseteq \{-1, 1\}$ for each point $(x,y) \in [0,1]^2$ where these derivatives exist. Then each of these functions is the opposite diagonal section of some copula (see, e.g., [31, 45]), i.e., there exists a sequence of copulas $(C_n)_{n\in\mathbb{N}}$ such that $C_1(x, 1-x) = f(x)$ and $C_n(x, 1-x) = g_n(x)$ for each $x \in [0,1]$ and all $n \in \mathbb{N} \setminus \{1\}$. For each $n \in \mathbb{N}$ consider the sets

$$A_n = \{(x,y) \in [0,1]^2 \mid f'(x) g_n'(y) = 1\}, \qquad B_n = \{(x,y) \in [0,1]^2 \mid f'(x) g_n'(y) = -1\}$$

which may be written as

$$\begin{aligned} A_n &= \bigcup_{i=0}^{2^{n-1}-1} \left(\left(\left]0, \frac{1}{2}\right[\times \left]\frac{2i}{2^n}, \frac{2i+1}{2^n}\right[\right) \cup \left(\left]\frac{1}{2}, 1\right[\times \left]\frac{2i+1}{2^n}, \frac{2i+2}{2^n}\right[\right) \right), \\ B_n &= \bigcup_{i=0}^{2^{n-1}-1} \left(\left(\left]0, \frac{1}{2}\right[\times \left]\frac{2i+1}{2^n}, \frac{2i+2}{2^n}\right[\right) \cup \left(\left]\frac{1}{2}, 1\right[\times \left]\frac{2i}{2^n}, \frac{2i+1}{2^n}\right[\right) \right), \end{aligned}$$

and observe that for each $n \in \mathbb{N}$

$$\lambda(A_n) = \lambda(B_n) = \frac{1}{2} \qquad \text{and} \qquad \lambda(A_n \cup B_n) = \lambda\Big(\bigcup_{j\in\mathbb{N}} A_j\Big) = \lambda\Big(\bigcup_{j\in\mathbb{N}} B_j\Big) = 1.$$

If, for some $\alpha > \alpha_C$ and for $H_{C,\alpha} = \{(x,y) \in [0,1]^2 \mid \psi_{A_C}(x,y) < \alpha\}$ we have $\lambda(H_{C,\alpha}) > 0$, then there exist $m, n \in \mathbb{N}$ such that $\lambda(A_m \cap H_{C,\alpha}) > 0$ and $\lambda(B_n \cap H_{C,\alpha}) > 0$. Then, for any $(x,y) \in B_n \cap H_{C,\alpha}$, the mixed derivative of C as well as the derivatives of f and g_n exist such that

$$\frac{\partial^2 [C, C_1, C_n]_\theta(x,y)}{\partial x\,\partial y} = \psi_{A_C}(x,y) + \theta f'(x) g_n'(y) < \alpha - \theta.$$

Clearly, if $\alpha - \theta < 0$ then $[C, C_1, C_n]_\theta$ has a negative density on a Borel subset of $[0,1]^2$ with positive Lebesgue measure and, therefore, cannot be a copula. Thus, necessarily, $\theta \le \alpha$. Similarly, $[C, C_1, C_m]_\theta$ cannot be a copula whenever $\alpha + \theta < 0$, and thus $\theta \ge -\alpha$. Since $\alpha > \alpha_C$ was chosen arbitrarily, these two inequalities imply $\theta \in [-\alpha_C, \alpha_C]$, showing that (i) implies (ii).

Conversely, fix a copula C and some $\theta \in [-\alpha_C, \alpha_C]$, and note that for all copulas $C_1, C_2 \in \mathscr{C}$ the two functions $f, g\colon [0,1] \to [0,1]$ given by $f(x) = C_1(x, 1-x)$ and $g(x) = C_2(x, 1-x)$ are 1-Lipschitz and that they satisfy the property $f(0) = f(1) = g(0) = g(1) = 0$. Also, the inequality

$$\frac{\partial^2 [C, C_1, C_2]_\theta(x,y)}{\partial x\,\partial y} = \psi_{A_C}(x,y) + \theta f'(x) g'(y) \ge \alpha_C - |\theta| \ge 0$$

holds on the set of all $(x,y) \in [0,1]^2$ where $\frac{\partial^2 [C,C_1,C_2]_\theta(x,y)}{\partial x\,\partial y}$, $f'(x)$ and $g'(y)$ exist (this set has Lebesgue measure 1) and where the mixed derivative is bounded from below by α_C. Moreover, define the functions $f_0, g_0\colon [0,1] \to \mathbb{R}$ by

$$f_0(x) = \begin{cases} f'(x) & \text{if } f'(x) \text{ exists,} \\ 0 & \text{otherwise,} \end{cases} \qquad g_0(x) = \begin{cases} g'(x) & \text{if } g'(x) \text{ exists,} \\ 0 & \text{otherwise,} \end{cases}$$

and observe that the following sum of integrals

$$\iint\limits_{[0,1]^2} \psi_{A_C}(x,y)\,dx\,dy + \iint\limits_{[0,1]^2} f_0(x) g_0(y)\,dx\,dy = \iint\limits_{[0,1]^2} \psi_{A_C}(x,y)\,dx\,dy + \int_0^1 f_0(x)\,dx \int_0^1 g_0(y)\,dy = \iint\limits_{[0,1]^2} \psi_{A_C}(x,y)\,dx\,dy$$

equals the mass of the absolutely continuous part of the copula C, showing that the function $[C, C_1, C_2]_\theta$ is also a copula. □

Theorem 4.2 clarifies, for a given copula C, the minimal range for the parameter θ ensuring that $[C, C_1, C_2]_\theta$ is a copula for an arbitrary choice of C_1 and C_2. In case that C_1 and C_2 are fixed beforehand, too, also for values of θ outside of the interval $[-\alpha_C, \alpha_C]$ the function $[C, C_1, C_2]_\theta$ may be a copula (compare Remark 4.4(i) below as well as Remark 5.16(ii) for some particular examples).

Example 4.3. Consider the copula $C^{\mathbf{CA}}_{0.5} = \sqrt{M \cdot \Pi}$ which is not absolutely continuous and which belongs to the family of Cuadras-Augé copulas first discussed in [22]. The formula (4.2) yields $\alpha_{C^{\mathbf{CA}}_{0.5}} = 0.5$, so Theorem 4.2 tells us that, for all copulas $C_1, C_2 \in \mathscr{C}$ and for each $\theta \in [-0.5, 0.5]$, the triplet $[C^{\mathbf{CA}}_{0.5}, C_1, C_2]_\theta$ given in (4.1) is a copula. The greatest perturbation of $C^{\mathbf{CA}}_{0.5}$ of this type is obtained if we choose $\widehat{C_1} = \widehat{C_2} = M$ and $\theta = 0.5$, in which case we get for all copulas $C_1, C_2 \in \mathscr{C}$, for each $\theta \in [-0.5, 0.5]$ and for each $(x,y) \in [0,1]^2$

$$\left| \left[C^{\mathbf{CA}}_{0.5}, C_1, C_2\right]_\theta(x,y) - C^{\mathbf{CA}}_{0.5}(x,y) \right| \le \left[C^{\mathbf{CA}}_{0.5}, M, M\right]_{0.5}(0.5, 0.5) - C^{\mathbf{CA}}_{0.5}(0.5, 0.5) = 0.125.$$

Remark 4.4. Some properties of the three basic copulas W, M and Π follow directly from Theorem 4.2:

(i) For all $C, C_1, C_2 \in \mathscr{C}$ we have $[C, C_1, C_2]_0 = C$ and, if $W \in \{C_1, C_2\}$, then we get $[C, C_1, C_2]_\theta = C$ for each $\theta \in \mathbb{R} \supseteq [-\alpha_C, \alpha_C]$.

(ii) Observe that $\psi_{A_W}(x,y) = \psi_{A_M}(x,y) = 0$ for each $(x,y) \in [0,1]^2$ and, therefore, $\alpha_W = \alpha_M = 0$. As a consequence, we get $[W, C_1, C_2]_0 = W$ and $[M, C_1, C_2]_0 = M$ for all $C_1, C_2 \in \mathscr{C}$, and neither $[W, C_1, C_2]_\theta$ nor $[M, C_1, C_2]_\theta$ is a copula if $\theta \ne 0$.

(iii) On the other hand, we have $\psi_{A_\Pi}(x,y) = 1$ for each $(x,y) \in [0,1]^2$ and, therefore, $\alpha_\Pi = 1$, i.e., $[\Pi, C_1, C_2]_\theta$ is a copula for all $C_1, C_2 \in \mathscr{C}$ and for all $\theta \in [-1, 1]$.

The following monotonicity properties are an immediate consequence of (4.1):

Corollary 4.5. *Let C, C_1, C_2 and D_1, D_2 be copulas and $\eta, \theta \in [-\alpha_C, \alpha_C]$. Then we have:*
(i) *if $\eta \le \theta$ then $[C, C_1, C_2]_\eta \le [C, C_1, C_2]_\theta$;*
(ii) *if $C_1 \le D_1$ and $C_2 \le D_2$ then*

$$[C, D_1, D_2]_\theta \le [C, C_1, C_2]_\theta \qquad \text{whenever } \theta \in [-\alpha_C, 0[\,,$$
$$[C, C_1, C_2]_\theta \le [C, D_1, D_2]_\theta \qquad \text{whenever } \theta \in [0, \alpha_C]\,;$$

(iii) $[C, M, M]_{-\alpha_C} \le [C, C_1, C_2]_\theta \le [C, M, M]_{\alpha_C}$.

Remark 4.6. Recall that $[\Pi, \Pi, \Pi]_\theta = C_\theta^{\mathbf{EFGM}}$ for each $\theta \in [-1, 1]$. The family

$$([C, C_1, C_2]_\theta)_{(C,C_1,C_2,\theta)\in\mathscr{C}^3\times[-\alpha_C,\alpha_C]}$$

can be considered as an extension of the family $\left(C_\theta^{\mathbf{EFGM}}\right)_{\theta\in[-1,1]}$ of EFGM copulas, as well as several subfamilies thereof, e.g.,

$$([C, \Pi, \Pi]_\theta)_{(C,\theta)\in\mathscr{C}\times[-\alpha_C,\alpha_C]} \qquad \text{and} \qquad ([\Pi, C_1, C_2]_\theta)_{(C_1,C_2,\theta)\in\mathscr{C}^2\times[-1,1]}.$$

As a matter of fact, a copula C is not necessarily symmetric, i.e., we may have $C(x, y) \neq C(y, x)$ for some $(x, y) \in [0, 1]^2$. In the literature, there are several concepts to measure the asymmetry of a (quasi-)copula (see, for instance, [29] for a recent approach to study the asymmetry of a (quasi-)copula with respect to a curve).

Using the Chebyshev norm, in [71, 94] the degree of asymmetry of a copula $C \in \mathscr{C}$ was defined as

$$\operatorname{asymm}(C) = \sup\left\{\left|C(x, y) - C(y, x)\right| \,\middle|\, (x, y) \in [0, 1]^2\right\}. \tag{4.3}$$

Obviously, a copula C is symmetric if and only if $\operatorname{asymm}(C) = 0$, and we always have $\operatorname{asymm}(C) \le \frac{1}{3}$, the maximal value $\frac{1}{3}$ being attained, e.g., by the copula $C^{[\frac{1}{3}]}$ given by $C^{[\frac{1}{3}]}(x, y) = \max\left(M\left(x, y - \frac{1}{3}\right), W(x, y)\right)$.

Example 4.7. In special cases we can say something about the degree of asymmetry of copulas $[C, C_1, C_2]_\theta$ constructed as in (4.1), in particular about the relationship between $\operatorname{asymm}(C)$ and $\operatorname{asymm}([C, C_1, C_2]_\theta)$.
(i) If C, C_1 and C_2 are arbitrary copulas such that the opposite diagonal sections of C_1 and C_2 coincide, i.e., $\omega_{C_1} = \omega_{C_2}$, then for each $\theta \in [-a_C, a_C]$ we get

$$\operatorname{asymm}([C, C_1, C_2]_\theta) = \operatorname{asymm}(C).$$

In other words, if $\omega_{C_1} = \omega_{C_2}$ and if $[C, C_1, C_2]_\theta$ is a copula then the construction (4.1) preserves the degree of asymmetry of the copula C.
(ii) If C, C_1 and C_2 are symmetric copulas and $\theta \in [-a_C, a_C]$ then $[C, C_1, C_2]_\theta$ may be asymmetric: put $C = \Pi$, $C_1 = M$ and $C_2 = \Pi$ then we have

$$0 = \operatorname{asymm}(C) < \operatorname{asymm}([C, C_1, C_2]_1) = \frac{1}{32}.$$

Writing $D = [C, C_1, C_2]_1$ we see that $[D, M, \Pi]_{-1} = C$ and, therefore,

$$\frac{1}{32} = \operatorname{asymm}(D) > \operatorname{asymm}([D, C_1, C_2]_{-1}) = 0.$$

(iii) Fix the two W-ordinal sums $\widehat{C_1} = W\text{-}(\langle\frac{1}{2}, 1, M\rangle)$ and $\widehat{C_2} = W\text{-}(\langle 0, \frac{1}{2}, M\rangle)$ which are asymmetric copulas ($\operatorname{asymm}(\widehat{C_1}) = \operatorname{asymm}(\widehat{C_2}) = \frac{1}{4}$) and whose opposite diagonal sections $\omega_{\widehat{C_1}}, \omega_{\widehat{C_2}} \colon [0, 1] \to [0, 1]$ are given by

$$\omega_{\widehat{C_1}}(x) = \max\left(\min\left(1 - x, x - \frac{1}{2}\right), 0\right) \qquad \text{and} \qquad \omega_{\widehat{C_2}}(x) = \omega_{\widehat{C_1}}(1 - x). \tag{4.4}$$

Then $\omega_{\widehat{C_1}}(x) \cdot \omega_{\widehat{C_2}}(y) \neq 0$ implies $(x, y) \in \left]\frac{1}{2}, 1\right] \times \left[0, \frac{1}{2}\right[$, in which case we have $\omega_{\widehat{C_1}}(y) \cdot \omega_{\widehat{C_2}}(x) = 0$. Thus the maximal absolute difference between the values $\omega_{\widehat{C_1}}(x) \cdot \omega_{\widehat{C_2}}(y)$ and $\omega_{\widehat{C_1}}(y) \cdot \omega_{\widehat{C_2}}(x)$ for $(x, y) \in [0, 1]^2$

equals $\frac{1}{16}$ and it is attained at the point $\left(\frac{3}{4}, \frac{1}{4}\right)$, i.e., for each $\theta \in [-1, 1]$ we get for the degree of asymmetry of the copula $\left[\Pi, \widehat{C_1}, \widehat{C_2}\right]_\theta$:

$$\begin{aligned}
\operatorname{asymm}\left(\left[\Pi, \widehat{C_1}, \widehat{C_2}\right]_\theta\right) &= \sup\left\{\left|\left[\Pi, \widehat{C_1}, \widehat{C_2}\right]_\theta(x, y) - \left[\Pi, \widehat{C_1}, \widehat{C_2}\right]_\theta(y, x)\right| \;\middle|\; (x, y) \in [0, 1]^2\right\} \\
&= |\theta| \cdot \sup\left\{\left|\omega_{\widehat{C_1}}(x) \cdot \omega_{\widehat{C_2}}(y) - \omega_{\widehat{C_1}}(y) \cdot \omega_{\widehat{C_2}}(x)\right| \;\middle|\; (x, y) \in [0, 1]^2\right\} \\
&= |\theta| \cdot \left|\omega_{\widehat{C_1}}(\tfrac{3}{4}) \cdot \omega_{\widehat{C_2}}(\tfrac{1}{4}) - \omega_{\widehat{C_1}}(\tfrac{1}{4}) \cdot \omega_{\widehat{C_2}}(\tfrac{3}{4})\right| \\
&= \frac{1}{16}\,|\theta|\,.
\end{aligned} \tag{4.5}$$

Here the relationship between the parameter θ and the degree of asymmetry of $\left[\Pi, \widehat{C_1}, \widehat{C_2}\right]_\theta$ is very simple, as it is given by a linear function of $|\theta|$. In Remarks 5.13 and 5.17 we present two families of copulas whose degrees of asymmetry depend on the respective parameters in a more complex way.

(iv) If D is an arbitrary copula and if D_1 and D_2 are asymmetric copulas having the same opposite diagonal sections as the two copulas $\widehat{C_1}$ and $\widehat{C_2}$, respectively, which we considered in (iii), i.e., if ω_{D_1} and ω_{D_2} coincide with $\omega_{\widehat{C_1}}$ and $\omega_{\widehat{C_2}}$ as given in (4.4), respectively, then for each $\theta \in [-\alpha_D, \alpha_D]$ the degree of asymmetry of $[D, D_1, D_2]_\theta$ is bounded from above by the sum of the degrees of asymmetry of D and $\left[\Pi, \widehat{C_1}, \widehat{C_2}\right]_\theta$. In other words, for each $\theta \in [-\alpha_D, \alpha_D]$ we have

$$\operatorname{asymm}([D, D_1, D_2]_\theta) \le \operatorname{asymm}(D) + \operatorname{asymm}\left(\left[\Pi, \widehat{C_1}, \widehat{C_2}\right]_\theta\right) = \operatorname{asymm}(D) + \frac{1}{16}\,|\theta|\,. \tag{4.6}$$

If the copula D is symmetric then in (4.6) the equality holds.

5 Perturbations of the product copula and ordinal sums

For the copulas considered in this section, recall the constructions of M-ordinal sums and W-ordinal sums in (2.8) and (2.9), respectively, and the x-flipping, the y-flipping and the survival copula of a given ordinal sum defined in (2.2). The verification of the equalities in Lemma 5.1 is a matter of tedious computations.

Lemma 5.1. *The relationship between M-ordinal sums $M\text{-}(\langle a_k, b_k, C_k\rangle)_{k\in K}$ and W-ordinal sums $W\text{-}(\langle a_k, b_k, C_k\rangle)_{k\in K}$, on the one hand, and the x-flipping, the y-flipping and the survival copula of a given copula, on the other hand, can be formalized as follows:*

$$\begin{aligned}
\left(M\text{-}(\langle a_k, b_k, C_k\rangle)_{k\in K}\right)^{\text{xflip}} &= W\text{-}\left(\left\langle 1 - b_k, 1 - a_k, (C_k)^{\text{xflip}}\right\rangle\right)_{k\in K}, \\
\left(W\text{-}(\langle a_k, b_k, C_k\rangle)_{k\in K}\right)^{\text{xflip}} &= M\text{-}\left(\left\langle 1 - b_k, 1 - a_k, (C_k)^{\text{xflip}}\right\rangle\right)_{k\in K}, \\
\left(M\text{-}(\langle a_k, b_k, C_k\rangle)_{k\in K}\right)^{\text{yflip}} &= W\text{-}\left(\left\langle a_k, b_k, (C_k)^{\text{yflip}}\right\rangle\right)_{k\in K}, \\
\left(W\text{-}(\langle a_k, b_k, C_k\rangle)_{k\in K}\right)^{\text{yflip}} &= M\text{-}\left(\left\langle a_k, b_k, (C_k)^{\text{yflip}}\right\rangle\right)_{k\in K}, \\
\left(M\text{-}(\langle a_k, b_k, C_k\rangle)_{k\in K}\right)^{\text{surv}} &= M\text{-}\left(\left\langle 1 - b_k, 1 - a_k, (C_k)^{\text{surv}}\right\rangle\right)_{k\in K}, \\
\left(W\text{-}(\langle a_k, b_k, C_k\rangle)_{k\in K}\right)^{\text{surv}} &= W\text{-}\left(\left\langle 1 - b_k, 1 - a_k, (C_k)^{\text{surv}}\right\rangle\right)_{k\in K}.
\end{aligned} \tag{5.1}$$

Let us start with a special class of M-ordinal sums constructed by means of two copies of the product copula Π.

Definition 5.2. For each $r \in\,]0, \infty[$ consider the copula $\Pi^{\langle r\rangle} \colon [0, 1]^2 \to [0, 1]$, defined as M-ordinal sum as follows:

$$\Pi^{\langle r\rangle} = M\text{-}\left(\left\langle 0, \frac{1}{r+1}, \Pi\right\rangle, \left\langle \frac{1}{r+1}, 1, \Pi\right\rangle\right). \tag{5.2}$$

If in Definition 5.2 we switch to the functional expression for M-ordinal sums given in (2.8), then for each $r \in \,]0, \infty[$ we obtain the following explicit formula for the copula $\Pi^{\langle r\rangle}\colon [0,1]^2 \to [0,1]$:

$$\Pi^{\langle r\rangle}(x,y) = \begin{cases} (r+1)xy & \text{if } (x,y) \in \left[0, \frac{1}{r+1}\right]^2, \\ \frac{(r+1)xy-x-y+1}{r} & \text{if } (x,y) \in \left]\frac{1}{r+1}, 1\right]^2, \\ M(x,y) & \text{otherwise.} \end{cases}$$

If we take into account that $\Pi^{\mathrm{xflip}} = \Pi^{\mathrm{yflip}} = \Pi^{\mathrm{surv}} = \Pi$, then the next three equalities follow in a straightforward way from (5.1):

Corollary 5.3. *For each $r \in \,]0, \infty[$ we have*

$$\begin{aligned} \left(\Pi^{\langle r\rangle}\right)^{\mathrm{xflip}} &= W\text{-}\left(\left\langle 0, \frac{r}{r+1}, \Pi\right\rangle, \left\langle \frac{r}{r+1}, 1, \Pi\right\rangle\right), \\ \left(\Pi^{\langle r\rangle}\right)^{\mathrm{yflip}} &= W\text{-}\left(\left\langle 0, \frac{1}{r+1}, \Pi\right\rangle, \left\langle \frac{1}{r+1}, 1, \Pi\right\rangle\right), \\ \left(\Pi^{\langle r\rangle}\right)^{\mathrm{surv}} &= M\text{-}\left(\left\langle 0, \frac{r}{r+1}, \Pi\right\rangle, \left\langle \frac{r}{r+1}, 1, \Pi\right\rangle\right). \end{aligned} \tag{5.3}$$

Comparing formulas (5.2) and (5.3), we see that the two copulas $\Pi^{\langle r\rangle}\colon [0,1]^2 \to [0,1]$ and $\left(\Pi^{\langle r\rangle}\right)^{\mathrm{yflip}}$ share the same pair of summands, namely, $(\langle 0, \frac{1}{r+1}, \Pi\rangle, \langle \frac{1}{r+1}, 1, \Pi\rangle)$, but they are ordinal sums of different types: the former is an M-ordinal sum, while the latter is a W-ordinal sum. For the sake of simplicity of the notations, let us fix a shortcut for the y-flipping of the copula $\Pi^{\langle r\rangle}$:

Definition 5.4. For each $r \in \,]0, \infty[$, define the copula $\Pi^{\langle -r\rangle}\colon [0,1]^2 \to [0,1]$ by

$$\Pi^{\langle -r\rangle} = \left(\Pi^{\langle r\rangle}\right)^{\mathrm{yflip}}. \tag{5.4}$$

Clearly, for each constant $r \in \,]-\infty, 0[\, \cup \,]0, \infty[$, we have $\frac{z}{z+1} = \frac{1}{r+1}$ if and only if $z = \frac{1}{r}$. This allows us to derive the notions $\Pi^{\langle 1/r\rangle}$ and $\Pi^{\langle -1/r\rangle}$ directly, replacing the variable $r \in \mathbb{R} \setminus \{0\}$ in the formulas (5.2) and (5.4) by its reciprocal value $\frac{1}{r} \in \mathbb{R} \setminus \{0\}$.

Corollary 5.5. *For each $r \in \,]-\infty, 0[\, \cup \,]0, \infty[$ we have*

$$\Pi^{\langle -1/r\rangle} = \left(\Pi^{\langle r\rangle}\right)^{\mathrm{xflip}} \qquad \textit{and} \qquad \Pi^{\langle 1/r\rangle} = \left(\Pi^{\langle r\rangle}\right)^{\mathrm{surv}}. \tag{5.5}$$

Definition 5.6. Denote the set of all copulas $\Pi^{\langle r\rangle}$ and $\Pi^{\langle -r\rangle}$ given in (5.2) and (5.4), respectively, by $\mathscr{C}_\Pi^\star$, i.e.,

$$\mathscr{C}_\Pi^\star = \left\{\Pi^{\langle r\rangle} \;\middle|\; r \in \,]-\infty, 0[\, \cup \,]0, \infty[\right\}. \tag{5.6}$$

Six copulas $\Pi^{\langle r\rangle} \in \mathscr{C}_\Pi^\star$ as given in (5.2) and (5.4)–(5.5) are shown in Figure 1, as well as, in Figure 3, some contour plots of such copulas.

Remark 5.7. If we take into account (2.4), we immediately see that the set $\mathscr{C}_\Pi^\star$ and the corresponding operations (x-flipping, y-flipping and the construction of the survival copula) induce a commutative diagram which is shown in the left-hand part of Figure 2 (compare also Figure 3).

Obviously, this commutative diagram is isomorphic to the commutative diagram in the right-hand part of Figure 2, where we “calculate” only with the parameters of the copulas $\Pi^{\langle r\rangle}$, due to (5.4)–(5.5). A canonical isomorphism $\varphi\colon \mathbb{R} \setminus \{0\} \to \mathscr{C}_\Pi^\star$ between these two commutative diagrams is given by $\varphi(r) = \Pi^{\langle r\rangle}$.

However, the copulas $\Pi^{\langle r\rangle} \in \mathscr{C}_\Pi^\star$ have some additional properties. In particular, they are extremal elements (with respect to the usual partial order $\leq$ for functions from $[0,1]^2$ to $[0,1]$) of a special class of copulas which has been studied in a number of papers (see, e.g., [4, 36, 99, 105]).

From Example 1.6.10 in [39] we know that the function $C\colon [0,1]^2 \to [0,1]$ defined by $C(x,y) = xy + f(x)g(y)$ (see (3.6) in Remark 3.1(v)) is a copula if and only if the functions $f, g\colon [0,1] \to \mathbb{R}$ are Lipschitz, vanish at the border of the unit interval $[0,1]$ and satisfy $f'(x)g'(y) \geq -1$ for all $(x,y) \in [0,1]^2$ where $f'(x)g'(y)$ exists.

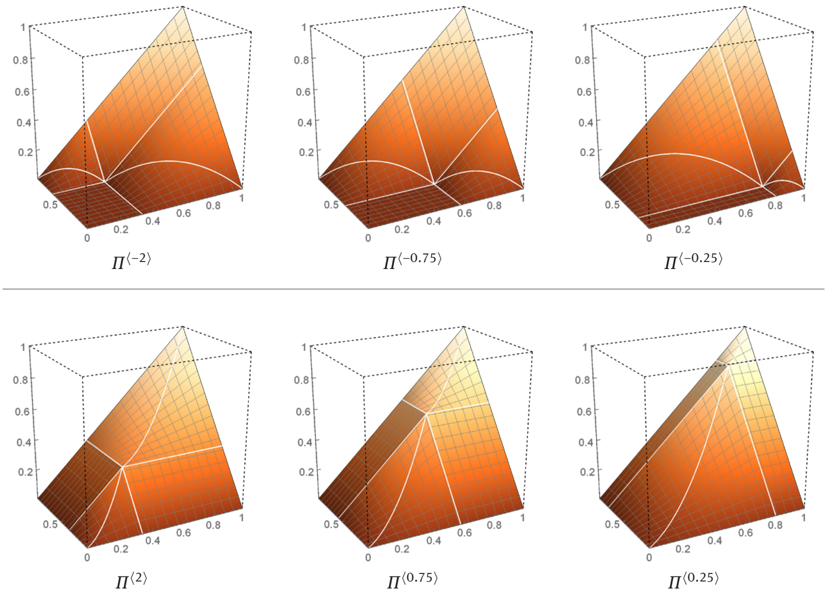

Figure 1: 3D plots of three minimal elements given by (5.4) (top) and three maximal elements given by (5.2) (bottom) of the set $\mathscr{C}_{\Pi}^{\star}$ (see Theorem 5.9).

Definition 5.8. Denote the set of Lipschitz functions on the unit interval vanishing at the boundaries of $[0, 1]$ by $\mathscr{L}_{\mathrm{b}}$:

$$\mathscr{L}_{\mathrm{b}} = \{f\colon [0,1] \to \mathbb{R} \mid f \text{ is Lipschitz and } f(0) = f(1) = 0\}.$$

Furthermore, define the set $\mathscr{F}_{\Pi}$ by

$$\mathscr{F}_{\Pi} = \{(f,g) \in \mathscr{L}_{\mathrm{b}} \times \mathscr{L}_{\mathrm{b}} \mid f'(x)g'(y) \geq -1 \text{ whenever } f'(x)g'(y) \text{ exists}\}$$

and, whenever $(f,g) \in \mathscr{F}_{\Pi}$, the function $\Pi_{[f,g]}\colon [0,1]^2 \to \mathbb{R}$ by

$$\Pi_{[f,g]}(x,y) = \Pi(x,y) + \Pi(f(x), g(y)).$$

Finally, since [39, Example 1.6.10] tells us that each function $\Pi_{[f,g]}$ with $(f,g) \in \mathscr{F}_{\Pi}$ is a copula, we can consider the partially ordered set $\mathscr{C}^{\mathrm{pert},\Pi}$ of copulas given by

$$\mathscr{C}^{\mathrm{pert},\Pi} = \{\Pi_{[f,g]} \mid (f,g) \in \mathscr{F}_{\Pi}\}. \tag{5.7}$$

With these notations in mind, we are ready to show that each of the copulas $\Pi^{\langle r\rangle}$ with $r \in\,]{-\infty}, 0[\, \cup\,]0, \infty[$ is an extremal element of the partially ordered set of copulas $\mathscr{C}^{\mathrm{pert},\Pi}$.

Theorem 5.9. *For each $r \in\,]0, \infty[$ we have:*

(i) *the copula $\Pi^{\langle r\rangle}$ is a maximal element of the set $\mathscr{C}^{\mathrm{pert},\Pi}$;*

(ii) *the copula $\Pi^{\langle -r\rangle}$ is a minimal element of the set $\mathscr{C}^{\mathrm{pert},\Pi}$.*

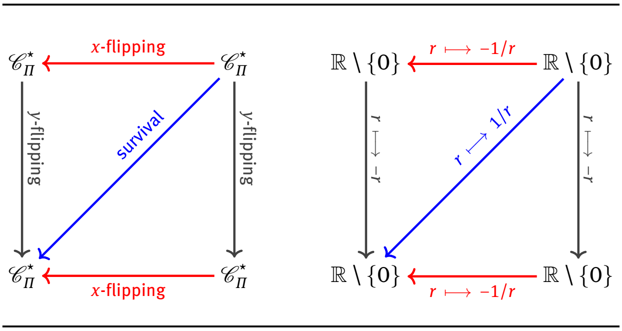

Figure 2: Two isomorphic commutative diagrams discussed in Remark 5.7: the partially ordered sets $(\mathscr{C}^\star_\Pi, \subseteq)$ and $(\mathbb{R} \setminus \{0\}, \le)$ (compare also Figure 3).

Proof. Fix $a, b \in \mathbb{R}$ with $\min(a, b) > 0$ and a non-constant function $f_{\langle a,b\rangle} \in \mathscr{L}_\mathrm{b}$ such that

$$a = \sup\{-f'_{\langle a,b\rangle}(x) \mid x \in [0,1] \text{ and } f'_{\langle a,b\rangle}(x) \text{ exists}\},$$
$$b = \sup\{f'_{\langle a,b\rangle}(x) \mid x \in [0,1] \text{ and } f'_{\langle a,b\rangle}(x) \text{ exists}\}.$$

It follows that $f_{\langle a,b\rangle}(x) \le bx$ and $f_{\langle a,b\rangle}(x) \le a(1-x)$ for all $x \in [0,1]$, i.e., the function $f^\star_{\langle a,b\rangle} : [0,1] \to \mathbb{R}$ given by

$$f^\star_{\langle a,b\rangle}(x) = \min(bx, a(1-x))$$

is the greatest function in $\mathscr{L}_\mathrm{b}$ such that the range of its derivative is a subset of $[-a, b]$. Taking into account that $\Pi_{[f_{\langle a,b\rangle}, g]} \in \mathscr{C}^{\mathrm{pert},\Pi}$ implies $f'_{\langle a,b\rangle}(x) g'(y) \ge -1$ for all $(x,y) \in [0,1]^2$ where the derivatives $f'_{\langle a,b\rangle}(x)$ and $g'(y)$ exist, we see that the function $g^\star_{\langle 1/b,1/a\rangle} : [0,1] \to \mathbb{R}$ given by

$$g^\star_{\langle 1/b,1/a\rangle}(x) = \min\left(\frac{x}{a}, \frac{1-x}{b}\right)$$

is the greatest function in $\mathscr{L}_\mathrm{b}$ such that the range of its derivative is a subset of $\left[-\frac{1}{b}, \frac{1}{a}\right]$. If we put $r = \frac{b}{a} > 0$, this implies that for each $g \in \mathscr{L}_\mathrm{b}$ with $(f_{\langle a,b\rangle}, g) \in \mathscr{F}_\Pi$ we obtain

$$\Pi_{[f_{\langle a,b\rangle}, g]} \le \Pi_{\left[f^\star_{\langle a,b\rangle}, g\right]} \le \Pi_{\left[f^\star_{\langle a,b\rangle}, g^\star_{\langle 1/b,1/a\rangle}\right]} = \Pi^{\langle r\rangle},$$

where $\Pi^{\langle r\rangle} : [0,1]^2 \to [0,1]$ is given by (5.2). Since for any $r_1, r_2 > 0$ with $r_1 \ne r_2$ the copulas $\Pi^{\langle r_1\rangle}$ and $\Pi^{\langle r_2\rangle}$ are incomparable because of their M-ordinal sum structure, we see that $\Pi^{\langle r\rangle}$ is a maximal element of $\mathscr{C}^{\mathrm{pert},\Pi}$ for each $r \in]0, \infty[$, showing that (i) holds.

In order to show (ii), choose an arbitrary $r \in]0, \infty[$. Note first that if $g \in \mathscr{L}_\mathrm{b}$ then for the flipped function $g^- : [0,1] \to [0,1]$ given by $g^-(x) = -g(1-x)$ we also have $g^- \in \mathscr{L}_\mathrm{b}$ and, thus, $\left(\Pi_{[f,g]}\right)^{\mathrm{yflip}} = \Pi_{[f,g^-]}$ for each $\Pi_{[f,g]} \in \mathscr{C}^{\mathrm{pert},\Pi}$, i.e., the set $\mathscr{C}^{\mathrm{pert},\Pi}$ is closed under y-flipping. Since the y-flipping reverses the order and preserves the incomparability of copulas, a copula is maximal if and only if its y-flipping is minimal, showing that for each $r \in]0, \infty[$ the copula $\Pi^{\langle -r\rangle} = \left(\Pi^{\langle r\rangle}\right)^{\mathrm{yflip}}$ (compare also (5.2)) is a minimal element of $\mathscr{C}^{\mathrm{pert},\Pi}$. □

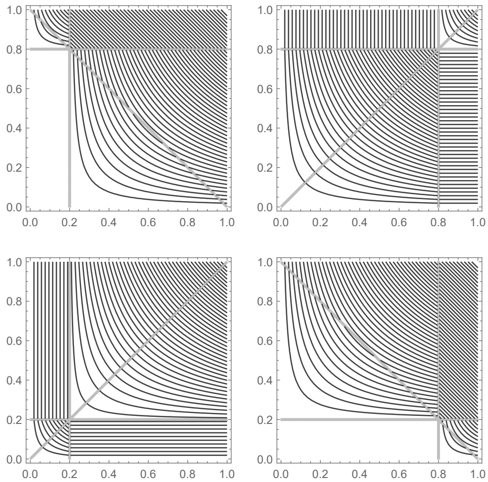

Figure 3: Contour plots of some extremal copulas in $\mathscr{C}_\Pi^\star$ (compare the two commutative diagrams in Figure 2): $\Pi^{\langle 0.25\rangle}$, $\Pi^{\langle -0.25\rangle}$, $\Pi^{\langle 4\rangle}$, and $\Pi^{\langle -4\rangle}$ (clockwise, from top right).

We can even show more: the copulas $\Pi^{\langle r\rangle}$ with $r \in \,]0, \infty[$ are the only elements of $\mathscr{C}^{\mathrm{pert},\Pi}$ which have a non-trivial idempotent element. Recall that $u \in [0, 1]$ is called a non-trivial idempotent element of a copula $C\colon [0, 1]^2 \to [0, 1]$ if $0 < u < 1$ and $C(u, u) = u$.

Proposition 5.10. *Let $(f, g) \in \mathscr{F}_\Pi$ be a pair of functions such that the copula $\Pi_{[f,g]} \in \mathscr{C}^{\mathrm{pert},\Pi}$ has a non-trivial idempotent element. Then there exists an $r \in \,]0, \infty[$ such that $\Pi_{[f,g]} = \Pi^{\langle r\rangle}$.*

Proof. Suppose that $\hat{u} \in \,]0, 1[$ is a non-trivial idempotent element of $\Pi_{[f,g]}$. Then there exist two copulas C_1 and C_2 such that $\Pi_{[f,g]}$ is an M-ordinal sum:

$$\Pi_{[f,g]} = M\text{-}(\langle 0, \hat{u}, C_1\rangle, \langle \hat{u}, 1, C_2\rangle).$$

This means $\Pi_{[f,g]}(x, y) = x$ whenever $\hat{u} \in [x, y]$, and $\Pi_{[f,g]}(x, y) = y$ whenever $\hat{u} \in [y, x]$.

Therefore, if $\hat{u} \in [x, y]$ then $\Pi_{[f,g]}(x, y) = xy + f(x)g(y) = x$ and, as a consequence, $f(x)g(y) = x(1 - y)$, i.e., there is a constant $b > 0$ such that $f(x) = bx$ for all $x \in [0, \hat{u}]$ and $g(x) = \frac{1-x}{b}$ for all $x \in [\hat{u}, 1]$. Similarly, if $\hat{u} \in [y, x]$ then we get $f(x)g(y) = (1 - x)y$, implying that there is a constant $a > 0$ such that $f(x) = a(1 - x)$ whenever $x \in [\hat{u}, 1]$, and $g(x) = \frac{x}{a}$ whenever $x \in [0, \hat{u}]$. Moreover, we have $f(\hat{u}) = b\hat{u} = a(1 - \hat{u})$ and $g(\hat{u}) = \frac{1-\hat{u}}{b} = \frac{\hat{u}}{a}$, both implying $a = \frac{b\hat{u}}{1-\hat{u}}$. Putting $r = \frac{1-\hat{u}}{\hat{u}} > 0$ we have $a = \frac{b}{r}$.

Summarizing, we have identified a solution $(f_r, g_r) \in \mathscr{F}_\Pi$ of the functional equation $\Pi_{[f,g]}(x, y) = xy + f(x)g(y)$ which is unique up to pairs of positive multiplicative constants $(a, \frac{1}{a})$ and which is given by

$$f_r(x) = \min(rx, 1 - x) \qquad \text{and} \qquad g_r(x) = \min\left(x, \frac{1 - x}{r}\right), \tag{5.8}$$

Now it is not difficult to check that $\Pi_{[f_r,g_r]} = \Pi^{\langle r\rangle}$. ☐

Remark 5.11. Let $\Pi_{[f,g]}$ be a copula in $\mathscr{C}^{\mathrm{pert},\Pi}$ with $(f, g) \in \mathscr{F}_\Pi$.

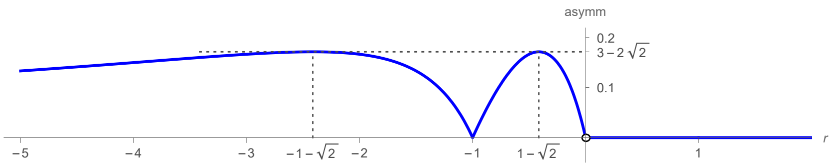

Figure 4: Asymmetry in $\mathscr{C}_{\Pi}^{\star}$: the degree of asymmetry $\operatorname{asymm}(\Pi^{\langle r\rangle})$ (see Remark 5.13) for $r \in [-5,0[\,\cup\,]0,2]$. Each copula $\Pi^{\langle r\rangle}$ with $r>0$ is symmetric, and so is $\Pi^{\langle -1\rangle}$. The maximal degree of asymmetry in $\mathscr{C}_{\Pi}^{\star}$, namely, $3-2\sqrt{2}\approx 0.171573$, is obtained for $r\in\{-1-\sqrt{2},1-\sqrt{2}\}$.

(i) As follows from the proof of Proposition 5.10, $\Pi_{[f,g]}$ can have at most one non-trivial idempotent element $\hat{u}\in\,]0,1[$, which uniquely determines the corresponding copula $\Pi^{\langle(1-\hat{u})/\hat{u}\rangle}=\Pi_{[f,g]}$.

(ii) Similarly, one can show that there can exist at most one element $\hat{u}\in\,]0,1[$ such that for $\Pi_{[f,g]}(\hat{u},1-\hat{u})=0$, which uniquely determines the corresponding copula $\Pi^{\langle(\hat{u}-1)/\hat{u}\rangle}=\Pi_{[f,g]}$ (which, because of $\frac{\hat{u}-1}{\hat{u}}<0$, necessarily is a minimal element of $\mathscr{C}^{\mathrm{pert},\Pi}$).

In Proposition 5.10 we have seen that, for $r\in\,]0,\infty[$, the functions f_r and g_r given by (5.8) generate the copula $\Pi^{\langle r\rangle}$ in a canonical way: $\Pi^{\langle r\rangle}=\Pi_{[f_r,g_r]}$. Also the copulas $\Pi^{\langle -r\rangle}$, $\Pi^{\langle -1/r\rangle}$ and $\Pi^{\langle 1/r\rangle}$ can be obtained in a similar way.

Remark 5.12. As a by-product of the second part of the proof of Theorem 5.9, it follows that the set of copulas $\mathscr{C}^{\mathrm{pert},\Pi}$ is closed under y-flipping. However, the set $\mathscr{C}^{\mathrm{pert},\Pi}$ is also closed under x-flipping and the construction of the survival copula: fix an arbitrary $r\in\,]0,\infty[$ and start with the pair of functions $(f_r,g_r)\in\mathscr{F}_{\Pi}$ given by (5.8). Define the functions $f_r^-,g_r^-\colon[0,1]\to[0,1]$ in the same way as in the proof of Theorem 5.9 by

$$f_r^-(x)=-f_r(1-x)=-\min(r(1-x),x)\qquad\text{and}\qquad g_r^-(x)=-g_r(1-x)=-\min\left(1-x,\frac{x}{r}\right),$$

and observe that $\{f_r^-,g_r^-\}\subseteq\mathscr{L}_{\mathrm{b}}$. Then it is not difficult to check that

$$\left(\Pi^{\langle r\rangle}\right)^{x\text{-flip}}=\Pi_{[f_r^-,g_r]}=\Pi^{\langle -1/r\rangle}\qquad\text{and}\qquad\left(\Pi^{\langle r\rangle}\right)^{\mathrm{surv}}=\Pi_{[f_r^-,g_r^-]}=\Pi^{\langle 1/r\rangle}.$$

Finally, note that we also have the following limit properties:

$$\lim_{r\to-\infty}\Pi^{\langle r\rangle}=\lim_{r\uparrow 0}\Pi^{\langle r\rangle}=\lim_{r\downarrow 0}\Pi^{\langle r\rangle}=\lim_{r\to\infty}\Pi^{\langle r\rangle}=\Pi.$$

Remark 5.13. Figures 1 and 3 indicate clearly that the set $\mathscr{C}_{\Pi}^{\star}$ contains symmetric and asymmetric copulas. More precisely, for each $\theta\in\mathbb{R}\setminus\{0\}$ the degree of asymmetry $\operatorname{asymm}(\Pi^{\langle r\rangle})$ of the copula $\Pi^{\langle r\rangle}$ defined in (4.3) (see [71, 94]) is given by (the relationship between the parameter r and $\operatorname{asymm}(\Pi^{\langle r\rangle})$ is illustrated in Figure 4)

$$\operatorname{asymm}(\Pi^{\langle r\rangle})=\begin{cases}\frac{1+r}{r(1-r)} & \text{if } r\in\,]-\infty,-1[\,,\\ \frac{r(r+1)}{r-1} & \text{if } r\in[-1,0[\,,\\ 0 & \text{if } r\in\,]0,\infty[\,.\end{cases}$$

For the following family of copulas we shall make use of the so-called (vertical) Π-ordinal sums of copulas given in (2.10) (for more details see [102]).

Definition 5.14. For each $r\in\,]0,\infty[\,\setminus\{1\}$ define the copula $\Pi\text{-}(W,M)_{\langle r\rangle}\colon[0,1]^2\to[0,1]$ by

$$\Pi\text{-}(W,M)_{\langle r\rangle}=\begin{cases}\Pi\text{-}(\langle 0,1-r,W\rangle,\langle 1-r,1,M\rangle) & \text{if } r\in\,]0,1[\,,\\ \Pi\text{-}\left(\left\langle 0,\frac{1}{r},M\right\rangle,\left\langle\frac{1}{r},1,W\right\rangle\right) & \text{if } r\in\,]1,\infty[\,.\end{cases}\tag{5.9}$$

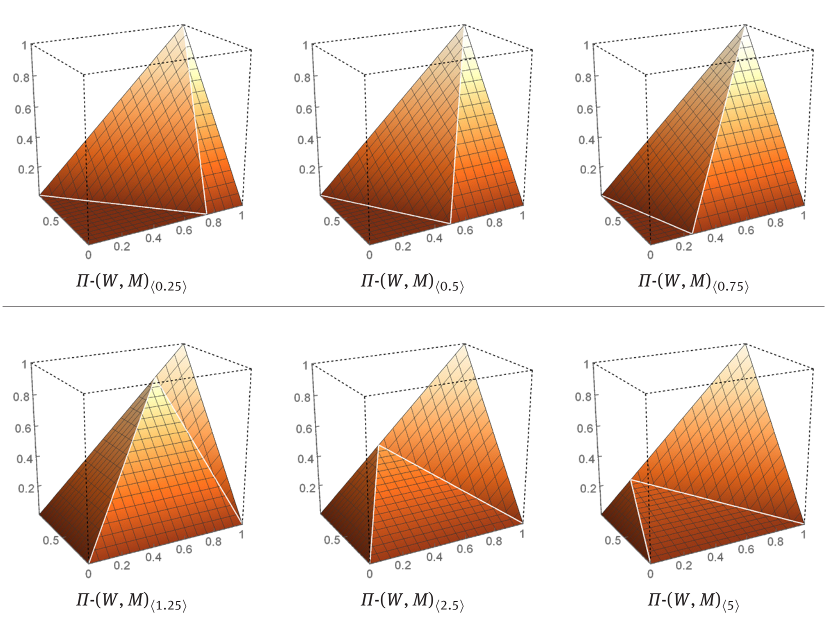

Figure 5: 3D plots of several Π-ordinal sums $\Pi\text{-}(W,M)_{\langle r\rangle}$ with $r \in \,]0,\infty[\, \setminus \{1\}$, as given in Definition 5.14.

If in (5.9) we switch to the functional expression for Π-ordinal sums as given in (2.10), then we get the following explicit formulas for the copula $\Pi\text{-}(W,M)_{\langle r\rangle}\colon [0,1]^2 \to [0,1]$ in the two cases, $r \in \,]0,1[$ and $r \in \,]1,\infty[$, respectively:

(i) in the case $r \in \,]0,1[$ we have

$$\Pi\text{-}(W,M)_{\langle r\rangle}(x,y) = \begin{cases} 0 & \text{if } y \in \left[0, \frac{-x+1-r}{1-r}\right], \\ y & \text{if } y \in \left[0, \frac{x+r-1}{r}\right], \\ x+(y-1)(1-r) & \text{otherwise;} \end{cases}$$

(ii) in the case $r \in \,]1,\infty[$ we have

$$\Pi\text{-}(W,M)_{\langle r\rangle}(x,y) = \begin{cases} x & \text{if } y \in [rx, 1], \\ x+y-1 & \text{if } y \in \left[\frac{rx-r}{1-r}, 1\right], \\ \frac{y}{r} & \text{otherwise.} \end{cases}$$

Remark 5.15. Recall the support of a copula $C \in \mathscr{C}$ which is defined as the complement of the union of all open subsets of $[0,1]^2$ with C-measure zero, where the C-measure is the probability measure on $[0,1]^2$ which is induced by the copula C (for more details see [93, Section 2.4]).

It is not difficult to see that, for $r \in \,]0,1[$, the support of the singular copula $\Pi\text{-}(W,M)_{\langle r\rangle} \in \mathscr{C}$ consists of the line segments connecting the point $(1-r,0)$ with $(0,1)$ and $(1,1)$ (blue in Figure 6), respectively. For $r \in \,]1,\infty[$, the support of $\Pi\text{-}(W,M)_{\langle r\rangle} \in \mathscr{C}$ is given by the line segments connecting the point $(\frac{1}{r},1)$ with $(0,0)$ and $(1,0)$ (red in Figure 6).

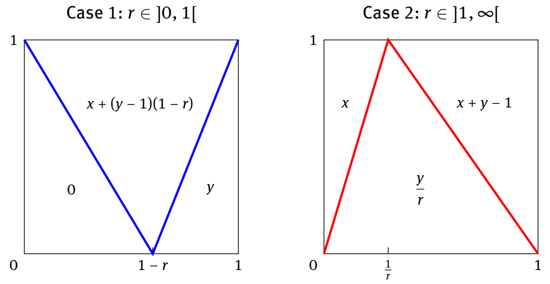

Figure 6: The singular copulas $\Pi\text{-}(W,M)_{\langle r\rangle}$ with $r \in]0,\infty[$ given in (5.9): Case 1, i.e., $r \in]0,1[$ (left, support shown in blue) and Case 2, i.e., $r \in]1,\infty[$ (right, support shown in red).

Remark 5.16. The family of copulas $\left(\Pi\text{-}(W,M)_{\langle r\rangle}\right)_{r\in]0,\infty[\setminus\{1\}}$ has a number of properties which can be derived directly from Definition 5.14.

(i) Because of

$$\lim_{r\uparrow 1}\Pi\text{-}(W,M)_{\langle r\rangle} = \lim_{r\downarrow 1}\Pi\text{-}(W,M)_{\langle r\rangle} = M \qquad \text{and} \qquad \lim_{r\to 0}\Pi\text{-}(W,M)_{\langle r\rangle} = \lim_{r\to\infty}\Pi\text{-}(W,M)_{\langle r\rangle} = W,$$

it makes sense to define $\Pi\text{-}(W,M)_{\langle 1\rangle} = M$ and $\Pi\text{-}(W,M)_{\langle 0\rangle} = \Pi\text{-}(W,M)_{\langle\infty\rangle} = W$. Then the family of copulas $\left(\Pi\text{-}(W,M)_{\langle r\rangle}\right)_{r\in[0,\infty]}$ is continuous with respect to the parameter r, and non-decreasing on $[0,1]$ and non-increasing on $[1,\infty]$ (also with respect to the parameter r). Six 3D plots of copulas of the form $\Pi\text{-}(W,M)_{\langle r\rangle}$ with $r \in]0,\infty[$ in Figure 7 illustrate this construction.

(ii) Consider the opposite diagonal section $\omega_{\Pi\text{-}(W,M)_{\langle r\rangle}}: [0,1] \to [0,1]$ of $\Pi\text{-}(W,M)_{\langle r\rangle}$ given by (2.5) and observe that we get $\omega_{\Pi\text{-}(W,M)_{\langle r\rangle}}(x) = \min(rx, 1-x)$ whenever $r \in]0,1[$, and $\omega_{\Pi\text{-}(W,M)_{\langle r\rangle}}(x) = \min\left(x, \frac{1-x}{r}\right)$ whenever $r \in]0,\infty[$. Recalling the pair of functions (f_r, g_r) defined in (5.8), we see that

$$\omega_{\Pi\text{-}(W,M)_{\langle r\rangle}} = \begin{cases} f_r & \text{if } r \in]0,1[\,, \\ g_r & \text{if } r \in]1,\infty[\,. \end{cases}$$

(iii) As a consequence, the copulas $\Pi\text{-}(W,M)_{\langle r\rangle}$ studied here provide an interesting bridge between the copulas $\Pi^{\langle r\rangle}$ considered in (5.2) and the copulas $[C, C_1, C_2]_\theta$ defined in (4.1):

$$\Pi^{\langle r\rangle} = \begin{cases} \left[\Pi, \Pi\text{-}(W,M)_{\langle r\rangle}, \Pi\text{-}(W,M)_{\langle r\rangle}\right]_{\frac{1}{r}} & \text{if } r \in]0,1]\,, \\ \left[\Pi, \Pi\text{-}(W,M)_{\langle r\rangle}, \Pi\text{-}(W,M)_{\langle r\rangle}\right]_{r} & \text{if } r \in]1,\infty[\,. \end{cases} \tag{5.10}$$

Note that, with the exception of $r = 1$ where we have $\Pi^{\langle 1\rangle} = [\Pi, M, M]_1$, in all other cases the parameters $\frac{1}{r}$ (if $r < 1$) and r (if $r > 1$) of the expressions in formula (5.10) are not contained in the interval $[-1,1] = [-\alpha_\Pi, \alpha_\Pi]$.

(iv) Defining, for an arbitrary copula $C \in \mathscr{C}$, the copula $\overline{C}: [0,1]^2 \to [0,1]$ by $\overline{C}(x,y) = C(y,x)$, we can extend the range of the parameter r of $\Pi^{\langle r\rangle}$ from $[0,\infty]$ to $[-\infty,\infty]$ by introducing negative parameters:

$$\Pi^{\langle r\rangle} = \begin{cases} \left[\Pi, \Pi\text{-}(W,M)_{\langle -r\rangle}, \overline{\Pi\text{-}(W,M)_{\langle -r\rangle}}\right]_{\frac{1}{r}} & \text{if } r \in]-1,0[\,, \\ [\Pi, M, M]_{-1} & \text{if } r = -1, \\ \left[\Pi, \Pi\text{-}(W,M)_{\langle -r\rangle}, \overline{\Pi\text{-}(W,M)_{\langle -r\rangle}}\right]_{r} & \text{if } r \in]-\infty,-1[\,. \end{cases}$$

Observe that the definition $\Pi^{\langle -1\rangle} = [\Pi, M, M]_{-1}$ makes sense because of $\lim_{r\uparrow -1} \Pi^{\langle r\rangle} = [\Pi, M, M]_{-1} = \lim_{r\downarrow -1} \Pi^{\langle r\rangle}$.

Remark 5.17. The 3D plots in Figure 5 suggest that the family $\big(\Pi\text{-}(W, M)_{\langle r\rangle}\big)_{r\in]0,\infty[\setminus\{1\}}$ contains asymmetric copulas. Indeed for each $r \in\,]0, \infty[\setminus \{1\}$ the degree of asymmetry $\operatorname{asymm}(\Pi\text{-}(W, M)_{\langle r\rangle})$ of the copula $\Pi\text{-}(W, M)_{\langle r\rangle}$ defined in (4.3) (see [71, 94]) is given by (for a visualization of the dependence of $\operatorname{asymm}(\Pi\text{-}(W, M)_{\langle r\rangle})$ on the parameter r see Figure 7)

$$\operatorname{asymm}(\Pi\text{-}(W, M)_{\langle r\rangle}) = \begin{cases} \frac{r(r-1)}{r^2-r-1} & \text{if } r \in\,]0, 1[\,, \\ \frac{r-1}{r^2+r-1} & \text{if } r \in\,]1, \infty[\,. \end{cases}$$

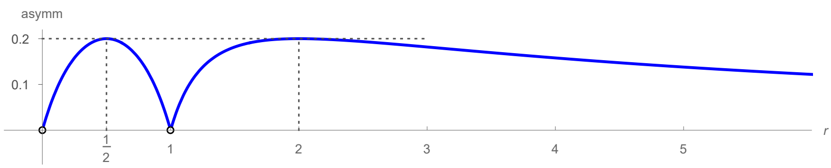

Figure 7: The degree of asymmetry $\operatorname{asymm}(\Pi\text{-}(W, M)_{\langle r\rangle})$ (see Remark 5.13) for $r \in\,]0, 1[\, \cup\,]1, 6[$. No member of the family of copulas $\big(\Pi\text{-}(W, M)_{\langle r\rangle}\big)_{r\in]0,\infty[\setminus\{1\}}$ is symmetric, and the maximal degree of asymmetry for this special set of Π-ordinal sums, namely, $\frac{1}{5}$, is obtained for $r \in \{\frac{1}{2}, 2\}$.

6 Markov product and dependence parameters

Several properties of copulas in the classes $\mathscr{C}_\Pi^\star$ given by (5.6) and $\mathscr{C}^{\text{pert},\Pi}$ given by (5.7) can be nicely related to some results on the Markov product of copulas. Finally, we shall discuss the dependence parameters of different perturbation-based copulas.

6.1 Perturbations which are idempotent with respect to the Markov product

Given two copulas $C_1, C_2 \in \mathscr{C}$, it is well-known that the partial derivatives $\frac{\partial C_1}{\partial y}$ and $\frac{\partial C_2}{\partial x}$ exist almost everywhere. Therefore the Markov product $*\colon \mathscr{C} \times \mathscr{C} \to \mathscr{C}$ (which was introduced as $*$ product in [26], see also [39, 73, 97] and [93, Section 6.4]) is well-defined, and the copula $C_1 * C_2\colon [0, 1]^2 \to [0, 1]$ is given by

$$(C_1 * C_2)(x, y) = \int_0^1 \frac{\partial C_1(x, t)}{\partial t} \cdot \frac{\partial C_2(t, y)}{\partial t}\, dt. \tag{6.1}$$

It is easy to see that the upper Fréchet-Hoeffding bound M and the product copula Π are the neutral element and the annihilator, respectively, of $\mathscr{C}$ with respect to the Markov product, i.e., for all $C \in \mathscr{C}$ we have

$$M * C = C * M = C \qquad \text{and} \qquad \Pi * C = C * \Pi = \Pi.$$

Moreover, for each $C \in \mathscr{C}$ the Markov product of C and the lower Fréchet-Hoeffding bound W are related to the x- and y-flipping and the survival copula of C as follows (see also [93]):

$$W * C = C^{\text{xflip}}, \qquad C * W = C^{\text{yflip}}, \qquad \text{and} \qquad W * C = C^{\text{surv}} * W.$$

Since the Markov product is associative (see [26, Theorem 2.4]), the pair $(\mathscr{C}, *)$ is a monoid, i.e., a semigroup with neutral element:

Corollary 6.1. *The pair* $(\mathscr{C}^{\text{pert},\Pi}, *)$ *is a sub-semigroup, and* $(\mathscr{C}^{\text{pert},\Pi} \cup \{M\}, *)$ *is a sub-monoid of* $(\mathscr{C}, *)$.

Proof. We only have to show that the set $\mathscr{C}^{\text{pert},\Pi}$ is closed under the Markov product $*$. Choose two arbitrary copulas $C_1, C_2 \in \mathscr{C}^{\text{pert},\Pi}$ such that $C_1 = \Pi_{[f_1,g_1]}$ and $C_2 = \Pi_{[f_2,g_2]}$ for some pairs of functions $(f_1, g_1), (f_2, g_2) \in \mathscr{F}_\Pi$. If we put $c_{(g_1,f_2)} = \int_0^1 g_1'(t)f_2'(t)\,dt$ then straightforward calculations yield for each $(x, y) \in [0, 1]^2$

$$C_1 * C_2(x, y) = xy + f_1(x) \cdot g_2(y) \cdot \int_0^1 g_1'(t) \cdot f_2'(t)\,dt = \Pi_{\left[c_{(g_1,f_2)} \cdot f_1, g_2\right]}(x, y).$$

Note that the constant $c_{(g_1,f_2)}$ may be split arbitrarily into two positive multiplicative factors, acting on f_1 and g_2, respectively. □

It has already been discussed in [93] that the family of Eyraud-Farlie-Gumbel-Morgenstern copula $C_\theta^{\mathbf{EFGM}}$ is also closed under the Markov product, i.e., for all $\alpha, \beta \in [-1, 1]$

$$C_\alpha^{\mathbf{EFGM}} * C_\beta^{\mathbf{EFGM}} = C_{\alpha\beta/3}^{\mathbf{EFGM}}.$$

Also the family of Fréchet copulas [51, 93] (and, therefore, the family of Mardia copulas [84, 93]) is closed under the Markov product $*$. Note that the family of Fréchet copulas $C_{\alpha,\beta}^{\mathbf{Fréchet}}$ is defined, for all $\alpha, \beta \in [0, 1]$ with $\alpha + \beta \le 1$, and for all $(x, y) \in [0, 1]^2$ by

$$C_{\alpha,\beta}^{\mathbf{Fréchet}}(x, y) = \alpha M(x, y) + (1 - \alpha - \beta)\Pi(x, y) + \beta W(x, y)$$

and can be seen as another type of perturbation of the product copula Π. For the Markov product the following holds (compare [93], but also [2]):

$$C_{\alpha_1,\beta_1}^{\mathbf{Fréchet}} * C_{\alpha_2,\beta_2}^{\mathbf{Fréchet}} = C_{\alpha_1\alpha_2+\beta_1\beta_2,\alpha_1\beta_2+\alpha_2\beta_1}^{\mathbf{Fréchet}}.$$

Remark 6.2. In [2, Remark 3.2] (see also [27, 44]) copulas C which are idempotent with respect to the Markov product, i.e., which satisfy $C * C = C$, were studied. Several classes of idempotent copulas were identified, among them:

(i) the class of M-ordinal sums M-$(\langle a_k, b_k, C_k\rangle)_{k\in K}$ with $C_k = \Pi$ for each $k \in K$;

(ii) the class of copulas of the form $\Pi_{[f,f]}(x, y) = xy+f(x)f(y)$, where the function $f\colon [0, 1] \to [0, 1]$ is Lipschitz with $f(0) = f(1) = 0$ and $f'(x) \in \{-1, 1\}$ wherever $f'(x)$ exists.

Remark 6.3. The set $\mathscr{C}^{\text{pert},\Pi}$ given in (5.7) has also some nice properties related to the Markov product given in (6.1).

(i) Based on Corollary 6.1, a copula $\Pi_{[f,g]} \ne \Pi$ is idempotent with respect to the Markov product if and only if $\Pi_{[f,g]} = \Pi_{[f,g]} * \Pi_{[f,g]} = \Pi_{\left[c_{(g,f)} \cdot f, g\right]}$, which holds if and only if $c_{(g,f)} = 1$, i.e., if $\int_0^1 g'(t)f'(t)\,dt = 1$.

(ii) From [2, Remark 3.2] (see also Remark 6.2) it follows that for each $r \in\,]0, \infty[$ the copula $\Pi^{\langle r\rangle}$ is idempotent with respect to the Markov product given in (6.1), i.e., $\Pi^{\langle r\rangle} * \Pi^{\langle r\rangle} = \Pi^{\langle r\rangle}$.
Because of (5.8) we have $\Pi^{\langle r\rangle} = \Pi_{[f_r,g_r]} = \Pi_{\left[af_r, \frac{1}{a}g_r\right]}$ for each $a \in\,]0, \infty[$. Choosing $a = \frac{1}{\sqrt{r}}$ we obtain $\frac{1}{\sqrt{r}}f_r = g_r\sqrt{r}$. Defining the function $h\colon [0, 1] \to \mathbb{R}$ by $h_r = \frac{1}{\sqrt{r}}f_r = g_r\sqrt{r}$, i.e.,

$$h_r(x) = \min\left(x\sqrt{r}, \frac{1 - x}{\sqrt{r}}\right),$$

then we have $(h_r, h_r) \in \mathscr{F}_\Pi$ and $\Pi^{\langle r\rangle} = \Pi_{[h_r,h_r]}$.
If we want to check directly that $\Pi^{\langle r\rangle}$ is idempotent with respect to $*$, we see that $\int_0^1 (h_r'(t))^2\,dt = 1$ because of

$$h_r'(t) = \begin{cases} \sqrt{r} & \text{if } t \in \left]0, \frac{1}{r+1}\right[, \\ -\frac{1}{\sqrt{r}} & \text{if } t \in \left]\frac{1}{r+1}, 1\right[. \end{cases}$$

6.2 Dependence parameters

Finally, we shall look for the dependence parameters (Spearman's rho, Kendall's tau, Blomqvist's beta, and Gini's gamma) given in (2.6) of the perturbation-based copulas studied in this paper.

First, we will compute the dependence parameters of the perturbations discussed in Section 4 (for the exact formula for $[C, C_1, C_2]_\theta$ see Definition 4.1).

Proposition 6.4. *Let C, C_1 and C_2 be copulas and $\theta \in \mathbb{R}$ such that the function $[C, C_1, C_2]_\theta$ defined by* (4.1) *is a copula. For the dependence parameters given in* (2.6) *we obtain the following formulas:*

$$\varrho([C, C_1, C_2]_\theta) = \varrho(C) + 12 \cdot \theta \int_0^1 C_1(x, 1-x)\,dx \cdot \int_0^1 C_2(y, 1-y)\,dy,$$

$$\tau([C, C_1, C_2]_\theta) = \tau(C) + 8\theta \int_0^1 \int_0^1 C(x,y) \tfrac{d}{dx} C_1(x, 1-x) \tfrac{d}{dy} C_2(y, 1-y)\,dy\,dx$$

$$\beta([C, C_1, C_2]_\theta) = \beta(C) + 4 \cdot \theta \cdot C_1(\tfrac{1}{2}, \tfrac{1}{2}) \cdot C_2(\tfrac{1}{2}, \tfrac{1}{2}) = \beta(C) + \tfrac{1}{4}\theta(\beta(C_1) + 1)(\beta(C_2) + 1),$$

$$\gamma([C, C_1, C_2]_\theta) = \gamma(C) + 4\theta \int_0^1 C_1(x, 1-x)[C_2(x, 1-x) + C_2(1-x, x)]\,dx.$$

Proof. The formulas for $\varrho([C, C_1, C_2]_\theta)$, $\beta([C, C_1, C_2]_\theta)$ and $\gamma([C, C_1, C_2]_\theta)$ are straightforward. Concerning Kendall's tau, it follows from formula (5.1.12) and Theorem 5.1.5 in [93] that $\tau([C, C_1, C_2]_\theta)$ can be expressed in terms of the partial derivatives of $[C, C_1, C_2]_\theta$ with respect to x and y. Then partial integration together with the boundary conditions (C1) of C_1 and C_2 yields

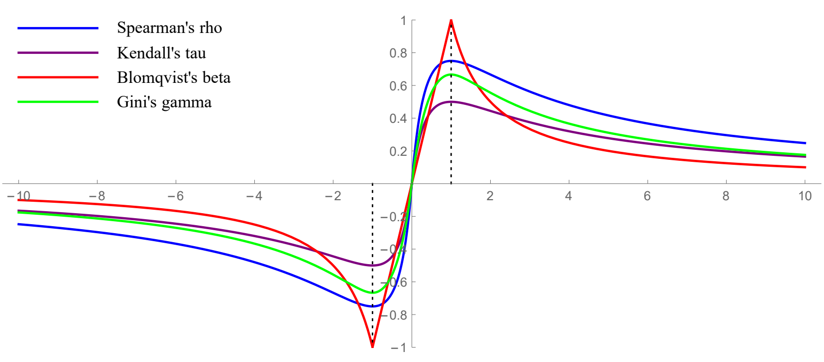

Figure 8: Spearman's rho, Kendall's tau, Blomqvist's beta, and Gini's gamma for the extremal copulas $\Pi^{\langle r \rangle} \in \mathscr{C}_\Pi^\star$ with $r \in [-10, 10] \setminus \{0\}$.

$$\begin{aligned}
\tau\big([C,C_1,C_2]_\theta\big) &= 1-4\iint\limits_{[0,1]^2} \tfrac{\partial}{\partial x}[C,C_1,C_2]_\theta(x,y)\tfrac{\partial}{\partial y}[C,C_1,C_2]_\theta(x,y)\,dx\,dy\\
&= 1-4\iint\limits_{[0,1]^2} \tfrac{\partial}{\partial x}C(x,y)\cdot\tfrac{\partial}{\partial y}C(x,y)\,dx\,dy - 4\theta\iint\limits_{[0,1]^2} C_1(x,1-x)\cdot\tfrac{\partial}{\partial x}C(x,y)\cdot\tfrac{d}{dy}C_2(y,1-y)\,dx\,dy\\
&\quad -4\theta\iint\limits_{[0,1]^2} C_2(y,1-y)\cdot\tfrac{\partial}{\partial y}C(x,y)\cdot\tfrac{d}{dx}C_1(x,1-x)\,dx\,dy\\
&\quad -4\theta^2\iint\limits_{[0,1]^2} C_1(x,1-x)\cdot\tfrac{d}{dx}C_1(x,1-x)\cdot C_2(y,1-y)\cdot\tfrac{d}{dy}C_2(y,1-y)\,dx\,dy\\
&= \tau(C)-4\theta\int_0^1 \tfrac{d}{dy}C_2(y,1-y)\Big(\int_0^1 C_1(x,1-x)\cdot\tfrac{\partial}{\partial x}C(x,y)\,dx\Big)\,dy\\
&\quad -4\theta\int_0^1 \tfrac{d}{dx}C_1(x,1-x)\Big(\int_0^1 C_2(y,1-y)\cdot\tfrac{\partial}{\partial y}C(x,y)\Big)\,dy\,dx\\
&\quad -4\theta^2\int_0^1 C_1(x,1-x)\cdot\tfrac{d}{dx}C_1(x,1-x)\,dx\cdot\int_0^1 C_2(y,1-y)\cdot\tfrac{d}{dy}C_2(y,1-y)\,dy\\
&= \tau(C)+8\theta\int_0^1\int_0^1 C(x,y)\tfrac{d}{dx}C_1(x,1-x)\tfrac{d}{dy}C_2(y,1-y)\,dy\,dx,
\end{aligned}$$

which completes the proof. □

Remark 6.5. There are several immediate consequences of Proposition 6.4 for special choices of either C or C_1 and C_2:

(i) For an arbitrary copula C and a parameter θ such that $[C,M,M]_\theta\in\mathscr{C}$ and $[C,\Pi,\Pi]_\theta\in\mathscr{C}$ we have the following relationships for the dependence parameters ϱ, β, and γ given in (2.6):

$$\begin{aligned}
\varrho\big([C,M,M]_\theta\big) &= \varrho(C)+\frac{3}{4}\theta, & \varrho\big([C,\Pi,\Pi]_\theta\big) &= \varrho(C)+\frac{1}{3}\theta,\\
\beta\big([C,M,M]_\theta\big) &= \beta(C)+\theta, & \beta\big([C,\Pi,\Pi]_\theta\big) &= \beta(C)+\frac{1}{4}\theta,\\
\gamma\big([C,M,M]_\theta\big) &= \gamma(C)+\frac{2}{3}\theta, & \gamma\big([C,\Pi,\Pi]_\theta\big) &= \gamma(C)+\frac{4}{15}\theta.
\end{aligned}$$

(ii) In particular, for each $\theta\in[-1,1]$ and for each of the dependence parameters given in (2.6) we obtain for the copula $[\Pi,M,M]_\theta\colon[0,1]^2\to[0,1]$ given by (4.1):

$$\varrho\big([\Pi,M,M]_\theta\big)=\frac{3}{4}\theta,\qquad \tau\big([\Pi,M,M]_\theta\big)=\frac{1}{2}\theta,\qquad \beta\big([\Pi,M,M]_\theta\big)=\theta,\qquad \gamma\big([\Pi,M,M]_\theta\big)=\frac{2}{3}\theta.$$

(iii) If for some $C\in\mathscr{C}$ the equalities $\varrho(C)=\tau(C)=0$ hold then we obtain for all copulas $C_1,C_2\in\mathscr{C}$ satisfying $\varrho\big([C,C_1,C_2]_\theta\big)\cdot\tau\big([C,C_1,C_2]_\theta\big)\neq 0$

$$\varrho\big([C,C_1,C_2]_\theta\big):\tau\big([C,C_1,C_2]_\theta\big)=3:2.$$

We can also give some properties of the four dependence parameters related to perturbations discussed in Section 5, in particular to the copulas $\Pi^{\langle r\rangle}\in\mathscr{C}_\Pi^\star$ (see (5.2), (5.4) and Definition 5.6) and $\Pi_{[f,g]}\in\mathscr{C}^{\mathrm{pert},\Pi}$ (see Definition 5.8).

Remark 6.6. For the dependence parameters given in (2.6) we obtain the following formulas (for a visualization see Figure 8):

(i) For each $r \in \mathbb{R} \setminus \{0\}$ we get

$$\varrho\big(\Pi^{\langle r\rangle}\big) = \operatorname{sign}(r)\frac{3|r|}{(|r|+1)^2}, \qquad \tau\big(\Pi^{\langle r\rangle}\big) = \operatorname{sign}(r)\frac{2|r|}{(|r|+1)^2},$$

$$\beta\big(\Pi^{\langle r\rangle}\big) = \operatorname{sign}(r)\min\Big(|r|, \frac{1}{|r|}\Big), \qquad \gamma\big(\Pi^{\langle r\rangle}\big) = \begin{cases} \operatorname{sign}(r)\frac{2|r|(3-|r|)}{3(|r|+1)} & \text{if } |r| \le 1, \\ \operatorname{sign}(r)\frac{2(3|r|-1)}{3|r|(|r|+1)} & \text{otherwise.} \end{cases}$$

(ii) In each case, the extremal values are attained for $r = \pm 1$:

$$\varrho\big(\Pi^{\langle \pm 1\rangle}\big) = \pm\frac{3}{4}, \qquad \tau\big(\Pi^{\langle \pm 1\rangle}\big) = \pm\frac{1}{2}, \qquad \beta\big(\Pi^{\langle \pm 1\rangle}\big) = \pm 1, \qquad \gamma\big(\Pi^{\langle \pm 1\rangle}\big) = \pm\frac{2}{3}.$$

(iii) Consider an open interval $\Theta \subseteq \mathbb{R}$ with $0 \in \Theta$ and a family of copulas $(C_\theta)_{\theta\in\Theta}$ which is continuous with respect to the parameter θ. In [54, Theorem 3.1] (compare also [106]), the authors have shown that, under mild regularity conditions, we have

$$\lim_{\theta\to 0}\frac{\varrho(C_\theta)}{\tau(C_\theta)} = \frac{3}{2}.$$

Putting $\Pi^{\langle 0\rangle} = \Pi$ we see that the family of copulas $(\Pi^{\langle r\rangle})_{r\in\mathbb{R}}$ is continuous with respect to the parameter r, and it illustrates the result above because of

$$\varrho\big(\Pi^{\langle r\rangle}\big) : \tau\big(\Pi^{\langle r\rangle}\big) = 3 : 2$$

for each $r \neq 0$. This is the same situation as for the family of Eyraud-Farlie-Gumbel-Morgenstern copulas $(C_\theta^{\mathbf{EFGM}})_{\theta\in[-1,1]}$ since for each $\theta \neq 0$ we have

$$\varrho\big(C_\theta^{\mathbf{EFGM}}\big) : \tau\big(C_\theta^{\mathbf{EFGM}}\big) = 3 : 2.$$

Proposition 6.7. *If $\xi\colon \mathscr{C} \to [-1,1]$ is one of the dependence parameters ϱ, τ, β, and γ given in* (2.6)*, then for each $\Pi_{[f,g]} \in \mathscr{C}^{\mathrm{pert},\Pi}$ we have*

$$\xi\big([\Pi, M, M]_{-1}\big) \le \xi\big(\Pi_{[f,g]}\big) \le \xi\big([\Pi, M, M]_1\big).$$

Proof. Note that, whenever for some pairs of functions $(f_1, g_1), (f_2, g_2) \in \mathscr{F}_\Pi$ we have

$$\int_0^1 f_1(x)\,dx \int_0^1 g_1(x)\,dx \le \int_0^1 f_2(x)\,dx \int_0^1 g_2(x)\,dx$$

then also $\xi\big(\Pi_{[f_1,g_1]}\big) \le \xi\big(\Pi_{[f_2,g_2]}\big)$ for each $\xi \in \{\varrho, \tau, \beta, \gamma\}$. Together with Corollary 4.5 and Remark 6.6 this shows that our claim holds. □

Still keeping the notations of Proposition 6.7, we can identify the ranges of the dependence parameters for the copulas $\Pi_{[f,g]} \in \mathscr{C}^{\mathrm{pert},\Pi}$:

Corollary 6.8. *For each $\Pi_{[f,g]} \in \mathscr{C}^{\mathrm{pert},\Pi}$ we obtain*

$$\big|\varrho\big(\Pi_{[f,g]}\big)\big| \le \frac{3}{4}, \qquad \big|\tau\big(\Pi_{[f,g]}\big)\big| \le \frac{1}{2},$$

$$\big|\beta\big(\Pi_{[f,g]}\big)\big| \le 1, \qquad \big|\gamma\big(\Pi_{[f,g]}\big)\big| \le \frac{2}{3}.$$

7 Concluding remarks

Perturbation of functions has a long and multifaceted tradition — also for copulas starting from a copula and perturbing it either by some other function(s) or some other copulas or some derived copulas has been of interest to the scientific community for different reasons. Mostly, the question whether or not or under which conditions the newly built function is again a copula, has been in the focus of investigations.

Also in this contribution, we discussed classes of perturbed copulas — starting with some copula C being perturbed by the product of the opposite diagonal sections of two (possibly different) copulas C_1 and C_2. We showed that the search for largest flexibility with respect to the choice of the copulas C_1, C_2 leads to a (possibly rather restricted) interval of possible values for the parameter. However, we also worked with different ordinal sums of copulas involving the three basic copulas, showing, e.g., that some particular ordinal sums with the independence copula as its only summand play an important role: they give rise to infinitely many, maximal and minimal elements of a set of perturbation copulas based on the independence copula and two one-place functions of a particular type.

Summarizing we may stress that an emphasis in this contribution has been the identification of new structural elements and properties of particular classes of perturbation copulas — revealing new properties, but also revealing new insights by connecting different types of perturbation copulas and by connecting our results to results known from different areas of copula theory.

For our further research on perturbation of copulas, we think, on the one hand, about the perturbation of copulas of higher dimensions. On the other hand, when considering random variables affected by some random noise, we see that the original copula, describing the stochastic structure of a random vector, is perturbed into a new copula, and our aim would be an analytical description of the perturbed copula. For some preliminary results in this direction see, e.g., [89].

Acknowledgements: Funded by the Johannes Kepler Open Access Publishing Fund. Anna Kolesárová is thankful for the support by the grant VEGA 1/0267/21 and, together with Radko Mesiar and Adam Šeliga, for the support by the grant APVV-18-0052. Radko Mesiar and Adam Šeliga also acknowledge the support by the grant APVV-17-0066.

The authors are also grateful to the two anonymous referees for their valuable comments and suggestions which eventually led to an improvement of an earlier version of this paper.

Conflict of interest statement: The authors declare that there is no conflict of interest.

References

[1] Abel, N. H. (1826). Untersuchung der Functionen zweier unabhängig veränderlichen Größen x und y, wie $f(x, y)$, welche die Eigenschaft haben, daß $f(z, f(x, y))$ eine symmetrische Function von z, x und y ist. *J. Reine Angew. Math. 1*, 11–15.

[2] Albanese A. and C. Sempi (2016). Idempotent copulæ: Ordinal sums and Archimedean copulæ. *J. Math. Anal. Appl. 438*(2), 1055–1065.

[3] Alsina, C., M. J. Frank, and B. Schweizer (2006). *Associative Functions.* World Scientific Publishing, Singapore.

[4] Amblard, C. and S. Girard (2009). A new extension of bivariate FGM copulas. *Metrika 70*, 1–17.

[5] Anakkamatee, W., S. Dhompongsa, and S. Tasena (2014). A constructive proof of the Sklar's theorem on copulas. *J. Nonlinear Convex Anal. 15*(6), 1137–1145.

[6] Arias-García, J. J., H. De Meyer, and B. De Baets (2018). On the construction of radially symmetric copulas in higher dimensions. *Fuzzy Set. Syst. 335*, 30–47.

[7] Arnold, V. I (1989). *Mathematical Methods of Classical Mechanics.* Second edition. Springer, New York.

[8] Bahraoui, T. and J.-F. Quessy (2017). Tests of radial symmetry for multivariate copulas based on the copula characteristic function. *Electron. J. Stat. 11*(1), 2066–2096.

[9] Bellman, R. (1964). *Perturbation Techniques in Mathematics, Physics, and Engineering*. Holt, Rinehart and Winston, New York.

[10] Birkhoff, G. (1973). *Lattice Theory*. Third edition. American Mathematical Society, Providence RI.

[11] Blomqvist, N. (1950). On a measure of dependence between two random variables. *Ann. Math. Statist. 21*(4), 593–600.

[12] Bransden, B. H. and C. J. Joachain (2000). *Quantum Mechanics*. Second edition. Pearson Education, Harlow.

[13] Cambanis, S. (1977). Some properties and generalizations of multivariate Eyraud-Gumbel-Morgenstern distributions. *J. Multivariate Anal. 7*(4), 551–559.

[14] Cambanis, S. (1991). On Eyraud-Farlie-Gumbel-Morgenstern random processes. In G. Dall'Aglio, S. Kotz, and G. Salinetti (Eds.), *Advances in Probability Distributions with Given Marginals*, pp. 207–222. Kluwer Academic Publishers, Dordrecht.

[15] Clifford, A. H. (1954). Naturally totally ordered commutative semigroups. *Amer. J. Math. 76*(3), 631–646.

[16] Clifford, A. H. (1958). Connected ordered topological semigroups with idempotent endpoints. I. *Trans. Amer. Math. Soc. 88*(1), 80–98.

[17] Clifford, A. H. (1958). Totally ordered commutative semigroups. *Bull. Amer. Math. Soc. 64*(6), 305–316.

[18] Clifford, A. H. (1959). Connected ordered topological semigroups with idempotent endpoints. II. *Trans. Amer. Math. Soc. 91*(2), 193–208.

[19] Cooray, K. (2019). A new extension of the FGM copula for negative association. *Comm. Statist. Theory Methods 48*(8), 1902–1919.

[20] Cuadras, C. M. (2009). Constructing copula functions with weighted geometric means. *J. Statist. Plann. Inference 139*(11), 3766–3772.

[21] Cuadras, C. M. (2015). Contributions to the diagonal expansion of a bivariate copula with continuous extensions. *J. Multivariate Anal. 139*, 28–44.

[22] Cuadras, C. M. and J. Augé (1981). A continuous general multivariate distribution and its properties. *Comm. Statist. Theory Methods 10*(4), 339–353.

[23] Dall'Aglio, G. (1956). Sugli estremi dei momenti delle funzioni di ripartizione doppia. *Ann. Sc. Normale Super. Pisa Cl. Sci. (3)10*, 35–74.

[24] Dall'Aglio, G. (1959). Sulla compatibilità delle funzioni di ripartizione doppia. *Rend. Mat. Appl. (3-4)18*, 385–413.

[25] Dall'Aglio, G. (1960). Les fonctions extrêmes de la classe de Fréchet à 3 dimensions. *Publ. Inst. Statist. Univ. Paris 9*, 175–188.

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

[27] Darsow, W. F. and E. T. Olsen (2010). Characterization of idempotent 2-copulas. *Note Mat. 30*(1), 147–177.

[28] De Baets, B. and H. De Meyer (2007). Orthogonal grid constructions of copulas. *IEEE Trans. Fuzzy Syst. 15*(6), 1053–1062.

[29] De Baets, B., H. De Meyer, and T. Jwaid (2019). On the degree of asymmetry of a quasi-copula with respect to a curve. *Fuzzy Set. Syst. 354*, 84–103.

[30] De Baets, B., H. De Meyer, J. Kalická, and R. Mesiar (2009). Flipping and cyclic shifting of binary aggregation functions. *Fuzzy Set. Syst. 160*(6), 752–765.

[31] De Baets, B., H. De Meyer, and M. Úbeda-Flores (2009). Opposite diagonal sections of quasi-copulas and copulas. *Internat. J. Uncertain. Fuzz. 17*(4), 481–490.

[32] Dirac, P. A. M. (1927). The quantum theory of the emission and absorption of radiation. *Proc. Roy. Soc. London Ser. A 114*(767), 243–265.

[33] Durante, F., J. Fernández-Sánchez, and C. Sempi (2012). Sklar's theorem obtained via regularization techniques. *Nonlinear Anal. 75*(2), 769–774.

[34] Durante, F., J. Fernández-Sánchez, and C. Sempi (2013). A topological proof of Sklar's theorem. *Appl. Math. Lett. 26*(9), 945–948.

[35] Durante, F., J. Fernández-Sánchez, and C. Sempi (2013). How to prove Sklar's theorem. In H. Bustince, J. Fernandez, R. Mesiar, and T. Calvo (Eds.), *Aggregation Functions in Theory and in Practise*, pp. 85–90. Springer, Heidelberg.

[36] Durante, F., J. Fernández-Sánchez, and M. Úbeda-Flores (2013). Bivariate copulas generated by perturbations. *Fuzzy Set. Syst. 228*, 137–144.

[37] Durante, F., S. Saminger-Platz, and P. Sarkoci (2008). On representations of 2-increasing binary aggregation functions. *Inform. Sci. 178*(23), 4534–4541.

[38] Durante, F., S. Saminger-Platz, and P. Sarkoci (2009). Rectangular patchwork for bivariate copulas and tail dependence. *Comm. Statist. Theory Methods 38*(15), 2515–2527.

[39] Durante, F. and C. Sempi (2016). *Principles of Copula Theory*. CRC Press, Boca Raton FL.

[40] Dvořák, A., M. Holčapek, and J. Paseka (2021). On ordinal sums of partially ordered monoids: A unified approach to ordinal sum constructions of t-norms, t-conorms and uninorms. *Fuzzy Set. Syst.*, to appear. Available at https://doi.org/10.1016/j.fss.2021.04.008.

[41] Eyraud, H. (1936). Les principes de la mesure des corrélations. *Ann. Univ. Lyon, Sect. A 1*, 30–47.

[42] Farlie, D. J. G. (1960). The performance of some correlation coefficients for a general bivariate distribution. *Biometrika 47*(3-4), 307–323.

[43] Faugeras, O. P. (2013). Sklar's theorem derived using probabilistic continuation and two consistency results. *J. Multivariate Anal. 122*, 271–277.

[44] Fernández-Sánchez, J., W. Trutschnig, and M. Tschimpke (2021). Markov product invariance in classes of bivariate copulas characterized by univariate functions. *J. Math. Anal. Appl. 501*(2), Article ID 125184, 15 pages.

[45] Fernández-Sánchez, J. and M. Úbeda-Flores (2018). Constructions of copulas with given diagonal (and opposite diagonal) sections and some generalizations. *Depend. Model. 6*, 139–155.
[46] Fernández-Sánchez, J. and M. Úbeda-Flores (2018). Proving Sklar's theorem via Zorn's lemma. *Internat. J. Uncertain. Fuzz. 26*(1), 81–85.
[47] Fernández-Sánchez, J. and M. Úbeda-Flores (2019). Solution to two open problems on perturbations of the product copula. *Fuzzy Set. Syst. 354*, 116–122.
[48] Féron, R. (1956). Sur les tableaux de corrélation dont les marges sont données. Cas de l'espace à trois dimensions. *Publ. Inst. Statist. Univ. Paris 5*, 3–12.
[49] Frank, M. J. (1979). On the simultaneous associativity of $F(x, y)$ and $x + y - F(x, y)$. *Aequationes Math. 19*, 194–226.
[50] Fréchet, M. (1951). Sur les tableaux de corrélation dont les marges sont données. *Ann. Univ. Lyon, Sect. A(3) 14*, 53–77.
[51] Fréchet, M. (1958). Remarques au sujet de la note précédente. *C. R. Acad. Sci. Paris 246*, 2719–2720.
[52] Fredricks, G. A. and R. B. Nelsen (1997). Copulas constructed from diagonal sections. In V. Beneš and J. Štěpán (Eds.), *Distributions with Given Marginals and Moment Problems*, pp. 129–136. Springer, Dordrecht.
[53] Fredricks, G. A. and R. B. Nelsen (2002). The Bertino family of copulas. In C. M. Cuadras, J. Fortiana, and J. A. Rodríguez-Lallena (Eds.), *Distributions with Given Marginals and Statistical Modelling*, pp. 81–91. Springer, Dordrecht.
[54] Fredricks, G. A. and R. B. Nelsen (2007). On the relationship between Spearman's rho and Kendall's tau for pairs of continuous random variables. *J. Statist. Plann. Inference 137*(7), 2143–2150.
[55] Genest, C. and J. G. Nešlehová (2014). On tests of radial symmetry for bivariate copulas. *Statist. Papers 55*, 1107–1119.
[56] Gini, C. (1955). Variabilità e mutabilità. In E. Pizetti and T. Salvemini (Eds.), *Memorie di Metodologica Statistica*. Libreria Eredi Virgilio Veschi, Roma.
[57] Gudendorf, G. and J. Segers (2010). Extreme-value copulas. In P. Jaworski, F. Durante, W. K. Härdle, and T. Rychlik (Eds.), *Copula Theory and Its Applications*, pp. 127–145. Springer, Berlin.
[58] Gumbel, E. J. (1958). Distributions à plusieurs variables dont les marges sont données. *C. R. Acad. Sci. Paris 246*, 2717–2719.
[59] Gutzwiller, M. C. (1998). Moon-Earth-Sun: The oldest three-body problem. *Rev. Modern Phys. 70*, 589–639.
[60] Hoeffding, W. (1940) Maßstabinvariante Korrelationstheorie. *Schr. Math. Inst. Inst. Angew. Math. Univ. Berlin 5*, 181–233. Also in Hoeffding, W. (1994). Scale-invariant correlation theory. In N. I. Fisher and P. K. Sen (Eds.), *The Collected Works of Wassily Hoeffding*, pp. 57–107. Springer, New York.
[61] Hoeffding, W. (1941). Maßstabinvariante Korrelationsmaße für diskontinuierliche Verteilungen. *Arch. Math. Wirtsch.-Sozialforschg. 7*, 49–70. Also in Hoeffding, W. (1994). Scale-invariant correlation theory. In N. I. Fisher and P. K. Sen (Eds.), *The Collected Works of Wassily Hoeffding*, pp. 109–133. Springer, New York.
[62] Hürlimann, W. (2017). A comprehensive extension of the FGM copula. *Statist. Papers 58*, 373–392.
[63] Joe, H. (1997). *Multivariate Models and Dependence Concepts*. Chapman & Hall, London.
[64] Joe, H. (2015). *Dependence Modeling with Copulas*. CRC Press, Boca Raton FL.
[65] Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika 30*(1-2), 81–93.
[66] Kepler, J. (1609). *Astronomia Nova*. Prague.
[67] Kepler, J. (1618-1621). *Epitome Astronomiae Copernicanae*. Linz.
[68] Kim, J.-M., E. A. Sungur, T. Choi, and T.-Y. Heo (2011). Generalized bivariate copulas and their properties. *Model Assist. Stat. Appl. 6*(2), 127–136.
[69] Klement, E. P., A. Kolesárová, R. Mesiar, and S. Saminger-Platz (2017). On the role of ultramodularity and Schur concavity in the construction of binary copulas. *J. Math. Inequal. 11*(2), 361–381.
[70] Klement, E. P., A. Kolesárová, R. Mesiar, and C. Sempi (2007). Copulas constructed from horizontal sections. *Comm. Statist. Theory Methods 36*(11), 2901–2911.
[71] Klement, E. P. and R. Mesiar (2006). How non-symmetric can a copula be? *Comment. Math. Univ. Carolin. 47*(1), 141–148.
[72] Klement, E. P., R. Mesiar, and E. Pap (2000). *Triangular Norms*. Springer, Dordrecht.
[73] Klement, E. P., R. Mesiar, and E. Pap (2001). Uniform approximation of associative copulas by strict and non-strict copulas. *Illinois J. Math. 45*(4), 1393–1400.
[74] Klement, E. P., R. Mesiar, and E. Pap (2002). Triangular norms as ordinal sums of semigroups in the sense of A. H. Clifford. *Semigroup Forum 65*, 71–82.
[75] Kolesárová, A., G. Mayor, and R. Mesiar (2015). Quadratic constructions of copulas. *Inform. Sci. 310*, 69–76.
[76] Kolesárová, A., R. Mesiar, and J. Kalická (2013). On a new construction of 1-Lipschitz aggregation functions, quasi-copulas and copulas. *Fuzzy Set. Syst. 226*, 19–31.
[77] Komorník, J., M. Komorníková, and J. Kalická (2017). Dependence measures for perturbations of copulas. *Fuzzy Set. Syst. 324*, 100–116.
[78] Komorník, J., M. Komorníková, and J. Kalická (2018). Families of perturbation copulas generalizing the FGM family and their relations to dependence measures. In V. Torra, R. Mesiar, and B. De Baets (Eds.), *Aggregation Functions in Theory and in Practice*, pp. 53–63. Springer, Cham.
[79] Komorník, J., M. Komorníková, J. Kalická, and C. Nguyen (2016). Tail dependence of perturbed copulas. *J. Stat. Theory Appl. 15*(2), 153–160.
[80] Lagrange, J.-L. (1788). *Méchanique Analitique*. La Veuve Desaint, Paris.

[81] Laplace, P.-S. (1798–1825). *Traité de Mécanique Céleste*. L'Imprimerie de Crapelet, Paris.

[82] Lee, M.-L. T. (1996). Properties and applications of the Sarmanov family of bivariate distributions. *Comm. Statist. Theory Methods 25*(6), 1207–1222.

[83] Ling, C.-H. (1965). Representation of associative functions. *Publ. Math. Debrecen 12*, 189–212.

[84] Mardia, K. V. (1970). *Families of Bivariate Distributions*. Charles Griffin, London.

[85] Marinacci, M. and L. Montrucchio (2005). Ultramodular functions. *Math. Oper. Res. 30*(2), 311–332.

[86] Mesiar, R., J. Komorník, and M. Komorníková (2013). On some construction methods for bivariate copulas. In H. Bustince, J. Fernandez, R. Mesiar, and T. Calvo (Eds.), *Aggregation Functions in Theory and in Practise*, pp. 39–45. Springer, Heidelberg.

[87] Mesiar, R., M. Komorníková, and J. Komorník (2015). Perturbation of bivariate copulas. *Fuzzy Set. Syst. 268*, 127–140.

[88] Mesiar, R. and V. Najjari (2014). New families of symmetric/asymmetric copulas. *Fuzzy Set. Syst. 252*, 99–110.

[89] Mesiar, R., A. Sheikhi, and M. Komorníková (2019). Random noise and perturbation of copulas. *Kybernetika 55*(2), 422–434.

[90] Mesiar, R. and J. Szolgay (2004). W-ordinal sums of copulas and quasi-copulas. *Proceedings MAGIA & UWPM 2004*, pp. 78–83. Publishing House of Slovak University of Technology, Bratislava.

[91] Mikami, T. (1997). Large deviations and central limit theorems for Eyraud-Farlie-Gumbel-Morgenstern processes. *Statist. Probab. Lett. 35*(1), 73–78.

[92] Morgenstern, D. (1956). Einfache Beispiele zweidimensionaler Verteilungen. *Mitteilungsbl. Math. Statist. 8*, 234–235.

[93] Nelsen, R. B. (2006). *An Introduction to Copulas*. Second edition. Springer, New York.

[94] Nelsen, R. B. (2007). Extremes of nonexchangeability. *Statist. Papers 48*, 329–336.

[95] Newton, I. (1687). *Philosophiæ Naturalis Principia Mathematica*. Jussu Societatis Regiae ac typis Josephi Streater, London.

[96] Oertel, F. (2015). An analysis of the Rüschendorf transform - with a view towards Sklar's Theorem. *Depend. Model. 3*, 113–125.

[97] Olsen, E. T., W. F. Darsow, and B. Nguyen (1996). Copulas and Markov operators. In L. Rüschendorf, B. Schweizer, and M. D. Taylor (Eds.), *Distributions with Fixed Marginals and Related Topics*, pp. 244–259. Institute of Mathematical Statistics, Hayward CA.

[98] Rodríguez-Lallena, J. A. and M. Úbeda-Flores (2004). A new class of bivariate copulas. *Statist. Probab. Lett. 66*(3), 315–325.

[99] Rodríguez-Lallena, J. A. and M. Úbeda-Flores (2005). Best-possible bounds on sets of multivariate distribution functions. *Comm. Statist. Theory Methods 33*(4), 805–820.

[100] Rüschendorf, L. (2009). On the distributional transform, Sklar's theorem, and the empirical copula process. *J. Statist. Plann. Inference 139*(11), 3921–3927.

[101] Saminger, S. (2006). On ordinal sums of triangular norms on bounded lattices. *Fuzzy Set. Syst. 157*(10), 1403–1416.

[102] Saminger-Platz, S., M. Dibala, E. P. Klement, and R. Mesiar (2017). Ordinal sums of binary conjunctive operations based on the product. *Publ. Math. Debrecen 91*(1-2), 63–80.

[103] Saminger-Platz, S., A. Kolesárová, R. Mesiar, and E. P. Klement (2020). The key role of convexity in some copula constructions. *Eur. J. Math. 6*, 533–560.

[104] Saminger-Platz, S., A. Kolesárová, A. Šeliga, R. Mesiar, and E. P. Klement (2021). The impact on the properties of the EFGM copulas when extending this family. *Fuzzy Set. Syst. 415*, 1–26.

[105] Sarmanov, O. V. (1966). Generalized normal correlation and two-dimensional Fréchet classes. *Dokl. Akad. Nauk SSSR 168*(1), 32–35.

[106] Schreyer, M., R. Paulin, and W. Trutschnig (2017). On the exact region determined by Kendall's τ and Spearman's ϱ. *J. R. Stat. Soc. Ser. B. Stat. Methodol. 79*(2), 613–633.

[107] Schur, I. (1923). Über eine Klasse von Mittelbildungen mit Anwendungen auf die Determinantentheorie. *S.-B. Berlin. Math. Ges. 22*, 9–20.

[108] Schwarz, G. (1985). Multivariate distributions with uniformly distributed projections. *Ann. Probab. 13*(4), 1371–1372.

[109] Schweizer, B. and A. Sklar (1958). Espaces métriques aléatoires. *C. R. Acad. Sci. Paris 247*, 2092–2094.

[110] Schweizer, B. and A. Sklar (1960). Statistical metric spaces. *Pacific J. Math. 10*(1), 313–334.

[111] Schweizer, B. and A. Sklar (1961). Associative functions and statistical triangle inequalities. *Publ. Math. Debrecen 8*, 169–186.

[112] Schweizer, B. and A. Sklar (1963). Associative functions and abstract semigroups. *Publ. Math. Debrecen 10*, 69–81.

[113] Schweizer, B. and A. Sklar (1983). *Probabilistic Metric Spaces*. North-Holland, New York.

[114] Šeliga, A., M. Kauers, S. Saminger-Platz, R. Mesiar, A. Kolesárová, and E. P. Klement (2021). Polynomial bivariate copulas of degree five: characterization and some particular inequalities. *Depend. Model. 9*, 13–42.

[115] Sheikhi, A., V. Amirzadeh, and R. Mesiar (2021). A comprehensive family of copulas to model bivariate random noise and perturbation. *Fuzzy Set. Syst. 415*, 27–36.

[116] Sheikhi, A., F. Arad, R. Mesiar, and L. Vavríková (2020). Random noise and perturbation of copula with a copula induced noise. *Int. J. Gen. Syst. 49*(8), 856–871.

[117] Shubina, M. and M.-L. T. Lee (2004). On maximum attainable correlation and other measures of dependence for the Sarmanov family of bivariate distributions. *Comm. Statist. Theory Methods 33*(5), 1031–1052.

[118] Siburg, K. F. and P. A. Stoimenov (2008). Gluing copulas. *Comm. Statist. Theory Methods 37*(19), 3124–3134.

[119] Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris 8*, 229–231.
[120] Sklar, A. (1973). Random variables, joint distribution functions, and copulas. *Kybernetika 9*(6), 449–460.
[121] Spearman, C. (1904). The proof and measurement of association between two things. *Am. J. Psychol. 15*, 72–101.
[122] Sriboonchitta, S. and V. Kreinovich (2018). Why are FGM copulas successful? A simple explanation. *Adv. Fuzzy Syst. 2018*, Article ID 5872195, 5 pages.
[123] Tchen, A. H. (1980). Inequalities for distributions with given marginals. *Ann. Probab. 8*(4), 814–827.

Depend. Model. 2021; 9:374–384 DE GRUYTER

Research Article Open Access
Special Issue in memory of Abe Sklar

Barry C. Arnold and Matthew Arvanitis*

On a general class of gamma based copulas

https://doi.org/10.1515/demo-2021-0117
Received May 25, 2021; accepted September 29, 2021

Abstract: A large family of copulas with gamma components is examined, and interesting submodels are defined and analyzed. Parameter estimation is demonstrated for some of these submodels. A brief discussion of higher-dimensional versions is included.

Keywords: copula, gamma components, bivariate beta, tail dependence, likelihood-free estimation, approximate Bayesian computation

MSC: 62H05, 62E10

1 Introduction

Arnold and Ng [3] introduced a bivariate second kind beta, or beta(2), distribution involving 8 independent random variables with gamma distributions (subsequently such random variables will be referred to as gamma components). It was identified as the most general bivariate model whose marginals are ratios of sums of independent gamma variables. The model involves 8 independent components $U_1, U_2, \ldots, U_8$ with $U_j \sim \Gamma(\alpha_j, 1),\ \ j = 1, 2, \ldots, 8$. The two-dimensional random vector (X, Y) is then defined by

$$\begin{aligned} X &= \frac{U_1 + U_5 + U_7}{U_3 + U_6 + U_8}, \\ Y &= \frac{U_2 + U_5 + U_8}{U_4 + U_6 + U_7}. \end{aligned} \tag{1}$$

This defines an 8-parameter family of bivariate distributions with beta(2) marginal distributions. If (X, Y) is defined as in (1) then we write: $(X, Y) \sim BB(2)(\underline{\alpha})$, to be read as (X, Y) has a bivariate second kind beta distribution with parameter vector $\underline{\alpha}$. Each α_i can assume any positive real value, so that the parameter space of the model is $(0, \infty)^8$. A corresponding family of bivariate distributions with beta marginals of the first or usual kind is obtained from (1) by defining

$$(V_1, V_2) = (X/(1 + X), Y/(1 + Y)).$$

Why there are 8 U_j's, and where they are located in the model (1), may require some explanation. There are four locations where a particular U_j may be placed. (1) In the numerator of X. (2) In the denominator of X. (3) In the numerator of Y and (4) In the denominator of Y. The variables U_1, U_2, U_3 and U_4 appear only once, and each one of them appears in only one of the four possible locations. A variable U_j cannot appear in both the numerator and denominator of X, nor of Y, since otherwise the independence of numerators and denominators, required for beta(2) marginals, would be destroyed. U_5 appears in the numerator of both X and Y. U_6 appears in the denominator of both X and Y. U_7 appears in the numerator of X and in the denominator of Y, while U_8 appears in the denominator of X and in the numerator of Y No U_j can appear in 3 or in 4 of

Barry C. Arnold: Department of Statistics, University of California, Riverside, USA, E-mail: barnold@ucr.edu
***Corresponding Author: Matthew Arvanitis:** USDA Forest Products Laboratory, Madison, Wisconsin, USA, E-mail: matthew.arvanitis@usda.gov

the possible locations, since that would destroy the required independence of at least one numerator and its corresponding denominator. If an additional independent gamma variable is introduced in one or two permissible locations in (1) then it can be combined with one of the existing 8 U_j's and no enrichment of the model will result. Thus for example, if U_9 is added to both numerators, then $U_5 + U_9$ will continue to play the role of U_5 with an adjusted shape parameter $\alpha_5 + \alpha_9$.

We adopt the convention that a random variable with a $\Gamma(\alpha, 1)$ distribution with $\alpha = 0$ will be defined to be a random variable that is degenerate at 0. By setting some of the α_j's in the Arnold-Ng model (1) equal to zero, simplified submodels (some of which have been discussed in the literature) will be obtained. Note that after setting certain α_j's equal to zero, we must retain $\alpha_1 + \alpha_5 + \alpha_7 > 0$, $\alpha_3 + \alpha_6 + \alpha_8 > 0$, $\alpha_2 + \alpha_5 + \alpha_8 > 0$, and $\alpha_4 + \alpha_6 + \alpha_7 > 0$, in order to continue to have beta(2) marginal distributions.

While this general bivariate beta model, and particularly its submodels, have demonstrated flexibility and usefulness in practice, the focus of this paper is on a specific class of submodels of the beta(1) containing only copulas, that is, distributions with uniform marginals. This paper is organized as follows. In Section 2, we construct this specific class of copulas by limiting the parameter space and discuss how further limitations of the parameter space can produce several familiar copula models discussed in the literature. In Section 3, natural symmetries of the copulas are considered. Section 4 discusses higher-dimensional versions. We demonstrate some examples of parameter estimation in Section 5, and finish with concluding remarks in Section 6.

2 Copulas

Our gamma based copulas are obtained by setting the values of the parameters in the Arnold-Ng(8) bivariate beta distribution so that the marginals are $Uniform(0, 1)$.

Recall that the Arnold-Ng 8-parameter bivariate beta model is of the form

$$(V_1, V_2) = \left(\frac{U_1 + U_5 + U_7}{U_1 + U_5 + U_7 + U_3 + U_6 + U_8}, \frac{U_2 + U_5 + U_8}{U_2 + U_5 + U_8 + U_4 + U_6 + U_7} \right), \tag{2}$$

where $U_1, U_2, U_3, \ldots, U_8$ are independent random variables with $U_i \sim \Gamma(\alpha_i, 1)$, $\; i = 1, 2, \ldots, 8$.

In order to have uniform marginals the α_i's must satisfy:

$$\alpha_1 + \alpha_5 + \alpha_7 = 1, \tag{3}$$
$$\alpha_3 + \alpha_6 + \alpha_8 = 1, \tag{4}$$
$$\alpha_2 + \alpha_5 + \alpha_8 = 1, \tag{5}$$

and

$$\alpha_4 + \alpha_6 + \alpha_7 = 1, \tag{6}$$

where all the α_i's are non-negative (and necessarily ≤ 1).

We may rewrite these constraints in the following form:

$$\begin{aligned} \alpha_1 &= 1 - \alpha_5 - \alpha_7; \\ \alpha_2 &= 1 - \alpha_5 - \alpha_8; \\ \alpha_3 &= 1 - \alpha_6 - \alpha_8; \\ \alpha_4 &= 1 - \alpha_6 - \alpha_7; \end{aligned} \tag{7}$$

so that the parameter space, the set of permissible values for $(\alpha_5, \alpha_6, \alpha_7, \alpha_8)$, may be defined by:

$$\begin{aligned} \alpha_5 &\in [0, 1]; \\ \alpha_6 &\in [0, 1]; \\ \alpha_7 &\in [0, 1 - \max\{\alpha_5, \alpha_6\}]; \end{aligned} \tag{8}$$

$$\alpha_8 \in [0, 1 - \max\{\alpha_5, \alpha_6\}].$$

It is convenient to use notation BB($i_1, i_2, \ldots, i_k$) to denote the model obtained from (2) by setting certain α_i's equal to 0 and retaining only $\alpha_{i_1}, \alpha_{i_2}, \ldots, \alpha_{i_k}$ in the model. So, for $\alpha_5 = \alpha_6 = 1$ and $\alpha_7 = \alpha_8 = 0$, this model reduces to the co-monotone copula (BB(5,6)), that is, $V_1 = V_2$. Likewise, if $\alpha_5 = \alpha_6 = 0$ and $\alpha_7 = \alpha_8 = 1$, this model reduces to the counter-monotone copula (BB(7,8)), that is, $V_1 = 1 - V_2$. Also, if $\alpha_5 = \alpha_6 = \alpha_7 = \alpha_8 = 0$ (BB(1,2,3,4)), then V_1 is independent of V_2. Therefore, this model encompasses a continuous space of distributions with the co-monotone and counter-monotone copulas, also known as the Fréchet–Hoeffding bounds, and the independent copula as boundary points. Throughout this paper, correlations for this model or its submodels are expressed in terms of Spearman's Rank Correlation and are obtained by simulation when they cannot be determined analytically.

Submodels of this family are of particular interest. For example, the simplest is given when $\alpha_5 = \alpha$, and $\alpha_6 = \alpha_7 = \alpha_8 = 0$, so that it takes the form:

$$(V_1, V_2) = \left(\frac{U_1 + U_5}{U_1 + U_5 + U_3}, \frac{U_2 + U_5}{U_2 + U_5 + U_4} \right), \tag{9}$$

This family has positive correlations ranging from 0 to 0.478 (discussed in more detail later in the section), and the correlation has a monotone relationship with the parameter α. At $\alpha = 0$, this model reduces to the independent case. A larger submodel, the Magnussen model, is obtained by setting only $\alpha_7 = \alpha_8 = 0$ (see [5]). It thus is of the form:

$$(V_1, V_2) = \left(\frac{U_1 + U_5}{U_1 + U_5 + U_3 + U_6}, \frac{U_2 + U_5}{U_2 + U_5 + U_4 + U_6} \right), \tag{10}$$

where

$$\alpha_1 + \alpha_5 = 1, \tag{11}$$

$$\alpha_3 + \alpha_6 = 1, \tag{12}$$

$$\alpha_2 + \alpha_5 = 1, \tag{13}$$

and

$$\alpha_4 + \alpha_6 = 1, \tag{14}$$

where all the α_i's are non-negative (and necessarily ≤ 1). This is a two parameter family of copulas parameterized by $\alpha_5, \alpha_6 \in (0, 1)$. The other four α_i's are determined by (11)-(14). This model exhibits a full range of positive correlations, with the independent copula and the co-monotone copula appearing at the two extremes of the parameter space.

The Arnold-Ng 5 parameter model (introduced in [3]) is obtained by setting $\alpha_3 = \alpha_4 = \alpha_5 = 0$, and is therefore denoted by BB(1,2,6,7,8). The corresponding family of copulas includes those of the form:

$$(V_1, V_2) = \left(\frac{U_1 + U_7}{U_1 + U_7 + U_6 + U_8}, \frac{U_2 + U_8}{U_2 + U_8 + U_6 + U_7} \right), \tag{15}$$

where

$$\alpha_1 + \alpha_7 = 1, \tag{16}$$

$$\alpha_6 + \alpha_8 = 1, \tag{17}$$

$$\alpha_2 + \alpha_8 = 1, \tag{18}$$

and

$$\alpha_6 + \alpha_7 = 1, \tag{19}$$

where all the α_i's are non-negative (and necessarily ≤ 1). This is a one parameter family of copulas parameterized by $\alpha_1 \in (0, 1)$. The other four α_i's are determined by (16)-(19).

Note that if $(V_1, V_2) \sim BB(1, 2, 6, 7, 8)$ then

- $(1 - V_1, 1 - V_2) \sim BB(3, 4, 5, 7, 8)$
- $(1 - V_1, V_2) \sim BB(2, 3, 5, 6, 7)$
- $(V_1, 1 - V_2) \sim BB(1, 4, 5, 6, 8)$

We will denote these models by

AN(5A)=BB(1,2,6,7,8),
AN(5B)=BB(3,4,5,7,8),
AN(5C)=BB(2,3,5,6,7), and
AN(5D)=BB(1,4,5,6,8).

For these to be copulas we must impose the following constraints:

For AN(5A):

$$\alpha_1 + \alpha_7 = 1, \tag{20}$$
$$\alpha_6 + \alpha_8 = 1, \tag{21}$$
$$\alpha_2 + \alpha_8 = 1, \tag{22}$$
$$\alpha_6 + \alpha_7 = 1. \tag{23}$$

For AN(5B)

$$\alpha_5 + \alpha_7 = 1, \tag{24}$$
$$\alpha_3 + \alpha_8 = 1, \tag{25}$$
$$\alpha_5 + \alpha_8 = 1, \tag{26}$$
$$\alpha_4 + \alpha_7 = 1. \tag{27}$$

For AN(5C)

$$\alpha_5 + \alpha_7 = 1, \tag{28}$$
$$\alpha_3 + \alpha_6 = 1, \tag{29}$$
$$\alpha_2 + \alpha_5 = 1, \tag{30}$$
$$\alpha_6 + \alpha_7 = 1. \tag{31}$$

For AN(5D)

$$\alpha_1 + \alpha_5 = 1, \tag{32}$$
$$\alpha_6 + \alpha_8 = 1, \tag{33}$$
$$\alpha_5 + \alpha_8 = 1, \tag{34}$$
$$\alpha_4 + \alpha_6 = 1. \tag{35}$$

So we have four one parameter families of copulas.

But there are 45 more one parameter submodels that can be constructed by setting three of the five parameters to zero and applying the appropriate constraints. Four of the $\binom{8}{5} = 56$ choices must be excluded as they would produce zeros in one of the numerators or denominators in (1), and four additional choices all represent the same family as the following example shows. Take the BB(4,5,6,7,8) model for which we need to impose the following constraints:

$$\alpha_5 + \alpha_7 = 1, \tag{36}$$
$$\alpha_6 + \alpha_8 = 1, \tag{37}$$
$$\alpha_5 + \alpha_8 = 1, \tag{38}$$
$$\alpha_4 + \alpha_6 + \alpha_7 = 1. \tag{39}$$

We can take $\alpha_8 = \alpha \in (0, 1)$. Then $\alpha_5 = 1 - \alpha,\ \alpha_6 = 1 - \alpha,\ \alpha_7 = \alpha$, and, necessarily, $\alpha_4 = 0$, so our model is equivalent to the BB(5,6,7,8) with constraints. But how can one pick among all these models, and what properties do they have?

Let us look at the Olkin-Liu BB [7] model:

$$(V_1, V_2) = \left(\frac{U_1}{U_1 + U_6}, \frac{U_2}{U_2 + U_6} \right), \tag{40}$$

For this to be a copula, we need $\alpha_1 = \alpha_2 = \alpha_6 = 1$. Thus we have a 0-dimensional family of copulas. Three related copulas are obtainable by reflection about $1/2$. Clearly, when $\alpha_5 = 1$ for the simple family (9), it reduces to a reflected version of (40), and when the Magnussen parameters satisfy $\alpha_5 = 1 - \alpha_6$, it also reduces to (a possibly reflected version of) this model. As a copula, the Olkin-Liu BB model is also known as the Ali-Mikhail-Haq copula [1], a distribution that tends to appear in a vast array of copula families, including many of the Archimedean type (It should be noted that while this paper refers to the "Ali-Mikhail-Haq" copula multiple times, the copula that is being referenced is only one specific member of an entire family of copulas which is not a subset of the present family.) [6].

Let us return to consider the Arnold-Ng(5A) model of the form:

$$(V_1, V_2) = \left(\frac{U_1 + U_7}{U_1 + U_7 + U_6 + U_8}, \frac{U_2 + U_8}{U_2 + U_8 + U_6 + U_7} \right), \tag{41}$$

where

$$\alpha_1 + \alpha_7 = 1, \tag{42}$$

$$\alpha_6 + \alpha_8 = 1, \tag{43}$$

$$\alpha_2 + \alpha_8 = 1, \text{ and} \tag{44}$$

$$\alpha_6 + \alpha_7 = 1, \tag{45}$$

where all the α_i's are non-negative (and necessarily ≤ 1). This is a one parameter family of copulas parameterized by $\alpha_1 \in (0, 1)$. The other four α_i's are determined by (42)-(45). So we have $\alpha_1 = \alpha_2 = \alpha_6 = \alpha \in (0, 1)$ and $\alpha_7 = \alpha_8 = 1 - \alpha$. A plot of the correlation of the AN(5A) model, as a function of α, is given in Figure 1.

One immediate observation about the AN(5A) model is that its range of correlations is not full, that is, it is a proper subset of $[-1, 1]$. While this may at first thought appear to be a disappointing range of correlations, it should be noted that this family exhibits a different feature that may be desirable for many applications: it is a continuous collection of distributions ranging from the counter-monotone copula (having a correlation of -1) to an Ali-Mikhail-Haq copula, which favors strong upper tail (positive) dependence only, which coincides with the 0.478 correlation. Contour plots of the densities for various values of α are shown in Figure 2. It is applicable to phenomena that do not exhibit significant lower tail dependence, but do exhibit a trade-off between upper tail dependence and an overall negative correlation. For example, if X_1 represents a citizen's government entitlements, and X_2 represents the same citizen's overall wealth, then applying the AN(5A) to $V_1 = F_{X_1}(X_1)$ and $V_2 = F_{X_2}(X_2)$, α could be a measure of the government's plutocratic tendencies, since we might observe α nearer to 0 for socially democratic governments, while α may be much closer to 1 for highly plutocratic governments.

Remark Among the 4 avatars of the copula based on the Arnold-Ng 5 parameter bivariate beta model, two, namely AN(5A) and AN(5B), contain the counter-monotone copula but not the co-monotone copula. In contrast AN(5C) and AN(5D) contain the co-monotone copula but not the counter-monotone copula.

3 Symmetry considerations

There are three symmetry conditions that are on occasion deemed to be desirable properties for modeling purposes. It is therefore of interest to identify which members of our general copula family exhibit such symmetry features. A copula will be said to be marginally symmetric if $V_1 \stackrel{d}{=} 1 - V_1$ and $V_2 \stackrel{d}{=} 1 - V_2$, where $\stackrel{d}{=}$

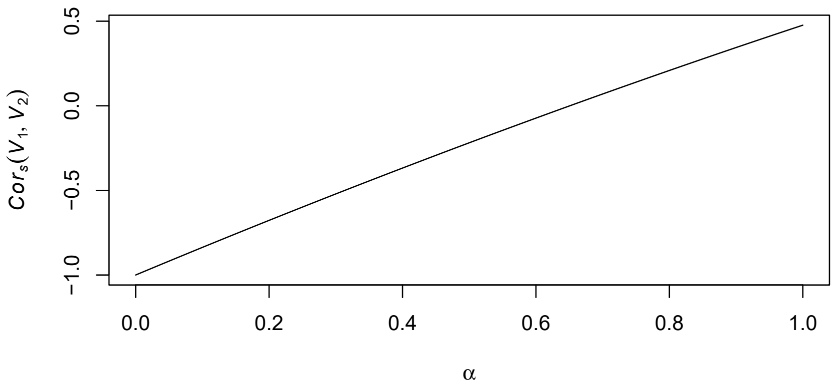

Figure 1: AN(5A) (Spearman) correlation as a function of α.

signifies equality of the distribution functions associated with the random variables. Clearly any BB copula with parameter vector $\underline{\alpha}$ satisfying the constraints (3)-(6) will exhibit marginal symmetry since V_1 and V_2 have $Uniform(0, 1)$ distributions.

Radial symmetry requires more restrictions. In order for a copula to be radially symmetric it must be the case that $(V_1, V_2) \stackrel{d}{=} (1 - V_1, 1 - V_2)$. For this to be true the parameter vector $\underline{\alpha}$ must be of the form

$$\underline{\alpha} = (\alpha, \alpha, \alpha, \alpha, \beta, \beta, 1 - \alpha - \beta, 1 - \alpha - \beta),$$

where $0 \le \alpha + \beta \le 1$. We thus have a two parameter subfamily of radially symmetric copulas.

For joint symmetry, even more is required, namely $(V_1, V_2) \stackrel{d}{=} (V_1, 1-V_2) \stackrel{d}{=} (1-V_1, V_2) \stackrel{d}{=} (1-V_1, 1-V_2)$. For this to occur we must have a copula with parameter vector of the form

$$\underline{\alpha} = (\alpha, \alpha, \alpha, \alpha, (1 - \alpha)/2, (1 - \alpha)/2, (1 - \alpha)/2, (1 - \alpha)/2).$$

where $0 \le \alpha \le 1$. We thus have a one parameter subfamily of jointly symmetric copulas.

There is another "symmetry" condition that can be considered. We will say that a copula is exchangeable if $(V_1, V_2) \stackrel{d}{=} (V_2, V_1)$. For this to be the case, the parameter vector of the copula must be of the form

$$(\alpha, \alpha, \beta, \beta, \gamma, \gamma, 1 - \alpha - \gamma, 1 - \beta - \gamma).$$

This is a three parameter family of copulas where the parameters satisfy the following constraints

$$0 \le \gamma \le 1, \quad 0 \le \alpha \le 1 - \gamma, \text{ and } \quad 0 \le \beta \le 1 - \gamma.$$

As examples, the AN(5C) and AN(5D) (among many others) are exchangeable.

4 On multivariate gamma based copulas

k-dimensional versions of the Arnold-Ng beta(2) distribution were mentioned in Arnold and Ghosh [2], in a context of copula models. First we consider the three dimensional case. It will then be evident how to deal with higher dimensions.

A three dimensional beta(2) distribution will be one whose structure is of a form which involves 26 independent gamma distributed U_j's. Thus, there are a total of 26 parameters in the model where U_j, j =

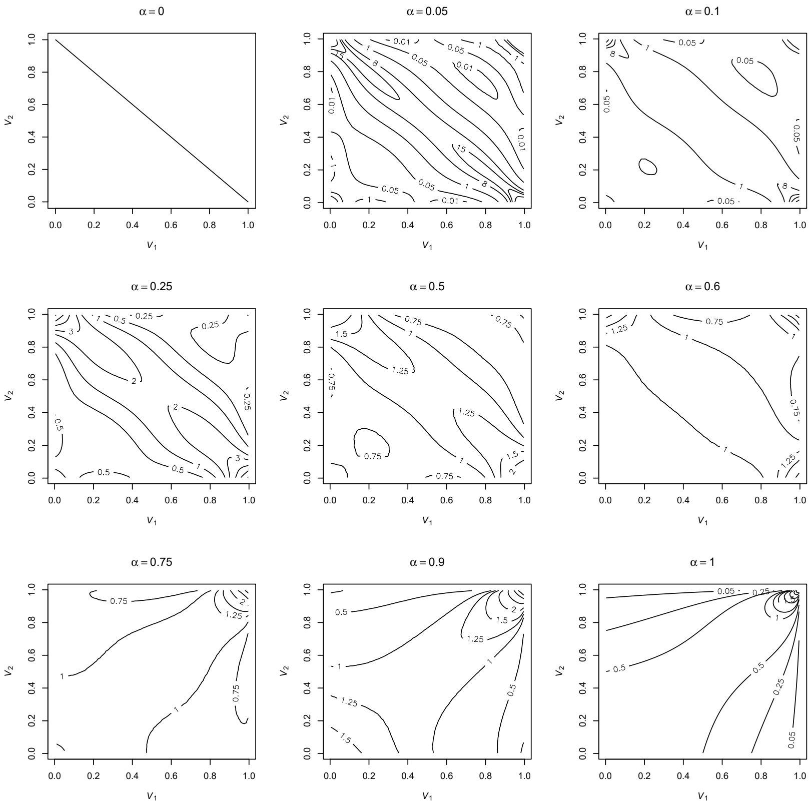

Figure 2: AN(5A) densities for various values of α.

1, 2, ..., 26 are independent variables with $U_j \sim \Gamma(\alpha_j, 1)$ for each j. The model can then be expressed in the following form.

$$X = \frac{U_1 + U_7 + U_8 + U_9 + U_{10} + U_{19} + U_{20} + U_{21} + U_{22}}{U_4 + U_{11} + U_{12} + U_{13} + U_{14} + U_{23} + U_{24} + U_{25} + U_{26}}, \tag{46}$$

$$Y = \frac{U_2 + U_7 + U_{11} + U_{15} + U_{16} + U_{19} + U_{20} + U_{23} + U_{24}}{U_5 + U_9 + U_{13} + U_{17} + U_{18} + U_{21} + U_{22} + U_{25} + U_{26}}, \tag{47}$$

$$Z = \frac{U_3 + U_8 + U_{12} + U_{15} + U_{17} + U_{19} + U_{21} + U_{23} + U_{25}}{U_6 + U_{10} + U_{14} + U_{16} + U_{18} + U_{20} + U_{22} + U_{24} + U_{26}}. \tag{48}$$

We must then impose 6 constraints to ensure that the associated trivariate beta distribution has uniform marginals.

The pattern for the dimensions of the parameter spaces of the multivariate copula models can now be recognized. The univariate model involves 2 U's with two constraints, i.e., a $3^1 - 1 - 2 = 0$ parameter model. The bivariate model involves 8 U's with 4 constraints, i.e., a $3^2 - 1 - 4 = 4$ parameter model. The trivariate case involves 26 U's with 6 constraints, i.e., a $3^3 - 1 - 6 = 20$ parameter model, and, in general, the k-dimensional model involves $3^k - 1$ U's with $2k$ constraints.

Use of fully parameterized k-dimensional copulas would almost never be recommended. Instead simplified sub-models, obtained by setting many of the α's equal to zero, can be expected to be adequate for many data sets.

5 Parameter estimation

In general, densities are not available in analytic form for the gamma-based copulas under discussion, except for special cases, such as the Ali-Mikhail-Haq case. So, if a sample is available from the bivariate gamma based copula then maximum likelihood is unavailable for parameter estimation. What we do have available are relatively simple expressions for mixed moments of the coordinate random variables. In principle, then, we could choose several sample mixed moments, equate them to their expectations and solve the resulting equations for the α_i parameters. This, in many cases, will prove to be a non-trivial exercise. However, it will typically be feasible for simplified sub-models involving only a few of the α_i's. With fewer equations to deal with, the method of moments approach is often not difficult to implement.

Example 1. *As a very simple example consider a model in which $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$. The model is thus of the form*

$$\begin{aligned} V_1 &= (U_5 + U_7)/(U_5 + U_7 + U_6 + U_8), \textit{ and} \\ V_2 &= (U_5 + U_8)/(U_5 + U_8 + U_6 + U_7). \end{aligned} \tag{49}$$

In this one parameter family of copulas we can denote α_5 by α and then we have $\alpha_6 = \alpha$ and $\alpha_7 = \alpha_8 = 1 - \alpha$. In this case, we can verify that

$$E(V_1 V_2) = (\alpha + 1)/6.$$

If we define $M_{V_1 V_2} = (1/n) \sum_{i=1}^n V_{1i} V_{2i}$, then a moment based estimate of α will be

$$\widetilde{\alpha} = 6M_{V_{1i}V_{2i}} - 1.$$

Example 2. *Suppose that (V_1, V_2) is of the AN(5A) type. While it may be tempting to make use of the simple relationship between the Spearman correlation and the parameter α shown in Figure 1 in a method of moments approach, it is not the case that Spearman is the best option. Figure 3 depicts the three most well-known correlation measures for the AN(5A) model. In addition to the correlations, the figure shows 95% confidence ranges of values for the correlation measures, given observed samples of the shown sizes. In this light, Kendall's Tau appears to be the superior choice (at least among these three choices). But even using Kendall's Tau, it is clear that samples would need to be large in order to obtain reasonably accurate estimates of α. For example, at $n = 50$, a 95% confidence interval for α would have a width of about 0.4, while for $n = 250$, that width would reduce to about 0.15, depending on the actual value of α.*

Examples in which more of the α's are non-zero will generally require iterative solutions of the moment equations, but except for that, they can be expected to yield reasonable estimates of the parameters provided that sample sizes are adequate, and enough moments are included in the estimation procedure. More complex examples exist for which a simple method of moments approach may be inadequate.

Example 3. *Consider the two-parameter case where $\alpha, \beta \in [0, 1]$, $\alpha_5 = \alpha\beta$, $\alpha_6 = \alpha$, and $\alpha_7 = (1 - \alpha)\beta$, and $\alpha_8 = (1-\alpha\beta)(1-\alpha)$, and a sample, $\mathbf{v}$, of size n, is available. This is a two-parameter subfamily which contains both the co-monotone copula and the counter-monotone copula, and therefore can exhibit the full range of correlations. It also, unsurprisingly, includes two rotated versions of the Ali-Mikhail-Haq copula. Approximate Bayesian Computation, or ABC, can be applied, where the assumed prior for (α, β) can be any distribution deemed appropriate. Here, we will assume it to be a bivariate distribution with independent $Uniform(0, 1)$ marginals.*

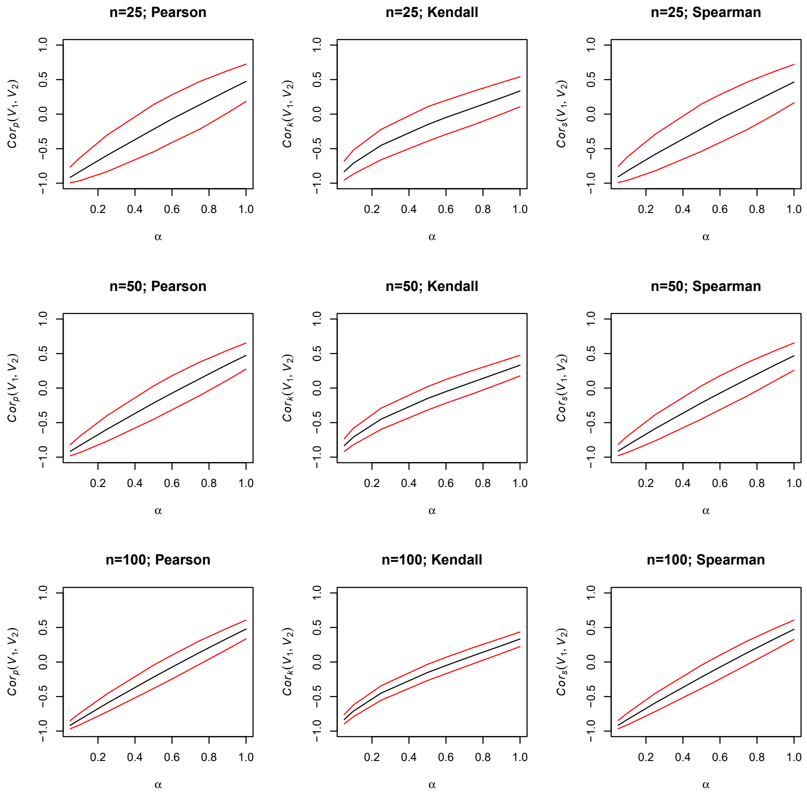

Figure 3: AN(5A) correlations, Pearson, Kendall's Tau, and Spearman, as a function of α, with 95% confidence bounds, obtained from simulation.

Now, given the behavior of this particular subfamily of distributions based on (α, β), we choose eight specific measures. The first four are simply proportions; partitioning the unit square into four equally-sized squares, we record the total number of observations in each, and divide by n. For the other four, we compute the four rotated Ali-Mikhail-Haq log-liklihood functions on the data contained in the same four regions (normalized).

$$S_1(\boldsymbol{v}) = \frac{n_{11}}{n}, \textit{ where } n_{11} = \sum_{k=1}^{n} I\left\{v_{1k} \le \frac{1}{2} \& v_{2k} \le \frac{1}{2}\right\},$$

$$S_2(\boldsymbol{v}) = \frac{n_{12}}{n}, \textit{ where } n_{12} = \sum_{k=1}^{n} I\left\{v_{1k} \le \frac{1}{2} \& v_{2k} > \frac{1}{2}\right\},$$

$$S_3(\boldsymbol{v}) = \frac{n_{21}}{n}, \textit{ where } n_{21} = \sum_{k=1}^{n} I\left\{v_{1k} > \frac{1}{2} \& v_{2k} \le \frac{1}{2}\right\},$$

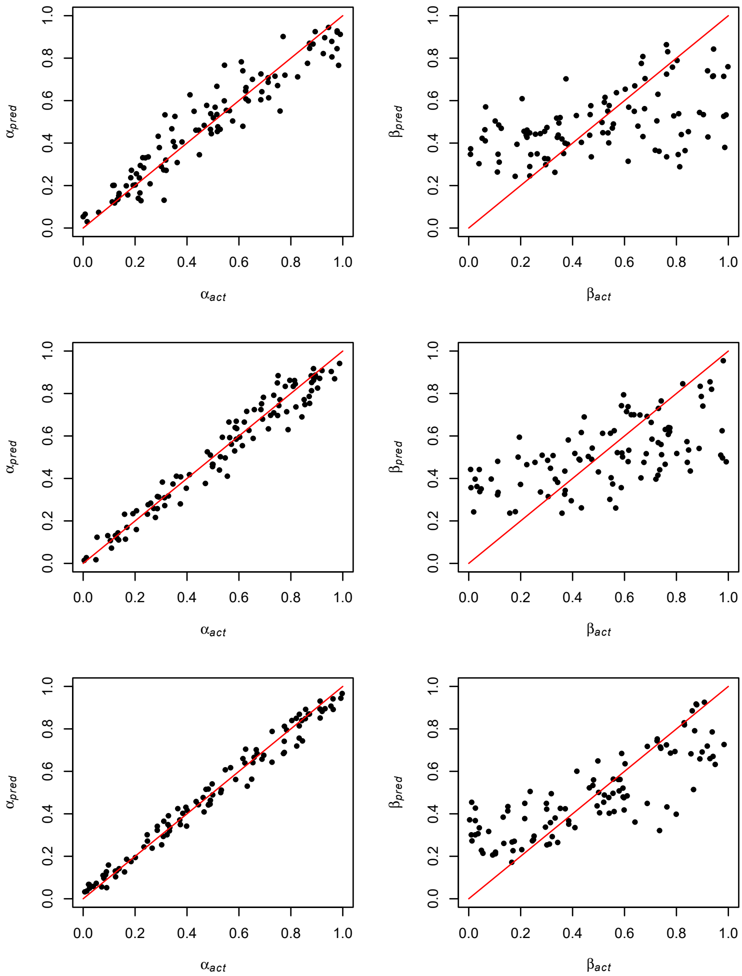

Figure 4: Comparisons of actual values of α and β with the predicted values by the ABC procedure. The three rows of plots represent sample sizes of 100, 250, and 1000, respectively.

$$S_4(\boldsymbol{v}) = \frac{n_{22}}{n}, \textit{where } n_{22} = \sum_{k=1}^{n} I\left\{v_{1k} > \frac{1}{2} \& v_{2k} > \frac{1}{2}\right\}, \tag{50}$$

$$S_5(\boldsymbol{v}) = \frac{1}{n_{11}} \sum_{k: v_{1k} \le \frac{1}{2} \& v_{2k} \le \frac{1}{2}} \log\left[\frac{(1+v_{1k})(1+v_{2k}) - 2 + (1-v_{1k})(1-v_{2k})}{(1-(1-v_{1k})(1-v_{2k}))^3}\right],$$

$$S_6(\boldsymbol{v}) = \frac{1}{n_{12}} \sum_{k: v_{1k} \le \frac{1}{2} \& v_{2k} > \frac{1}{2}} \log\left[\frac{(1+v_{1k})(2-v_{2k}) - 2 + (1-v_{1k})v_{2k}}{(1-(1-v_{1k})v_{2k})^3}\right],$$

$$S_7(\boldsymbol{v}) = \frac{1}{n_{21}} \sum_{k: v_{1k} > \frac{1}{2} \& v_{2k} \le \frac{1}{2}} \log\left[\frac{(2-v_{1k})(1+v_{2k}) - 2 + v_{1k}(1-v_{2k})}{(1-v_{1k}(1-v_{2k}))^3}\right],$$

$$S_8(\boldsymbol{v}) = \frac{1}{n_{22}} \sum_{k: v_{1k} > \frac{1}{2} \& v_{2k} > \frac{1}{2}} \log\left[\frac{(2-v_{1k})(2-v_{2k}) - 2 + v_{1k}v_{2k}}{(1-v_{1k}v_{2k})^3}\right],$$

We tested the ABC procedure for this family by arbitrarily selecting 100 values of (α, β) *from a* $Uniform\left((0,1)\times(0,1)\right)$ *distribution, simulating a sample for each, and applying the ABC procedure, repeating for three different sample sizes:* 100, 250, *and* 1000. *Figure 4 depicts the results. It can be observed that there is much more difficulty in estimating* β *than* α. *This is a common issue, and oftentimes with various subfamilies such as this, reasonable estimation may be assured with only very large sample sizes [4].*

6 Concluding Remarks

The present paper has provided an introduction to a broad spectrum of gamma based copula models and sub-models which can potentially be useful additions to the modeler's tool kit. These new flexible copula models can be expected to find application in cases in which the simpler well-known copula models prove to be inadequate to adapt to particular data sets, or where "big" data is encountered. It will be unlikely that use of the full 4 parameter model will be frequently deemed appropriate.

Acknowledgements: We are grateful for the helpful suggestions and dialogue provided by the referees which have resulted in an improved presentation.

Conflict of interest statement: Authors state no conflict of interest.

References

[1] Ali, M. M., N. Mikhail, and M. S. Haq (1978). A class of bivariate distributions including the bivariate logistic. *J. Multivariate Anal. 8*(3), 405–412.

[2] Arnold, B. C. and I. Ghosh (2014). Bivariate Kumaraswamy models involving use of Arnold-Ng copulas. *J. Appl. Stat. Sci. 22*(3/4), 227–240.

[3] Arnold, B. C. and H. K. T. Ng (2011). Flexible bivariate beta distributions. *J. Multivariate Anal. 102*(8), 1194–1202.

[4] Arvanitis, M. (2018). *Likelihood-free Estimation for Some Flexible Families of Distributions*. Ph.D. thesis, UC Riverside.

[5] Magnussen, S. (2004). An algorithm for generating positively correlated beta-distributed random variables with known marginal distributions and a specified correlation. *Comput. Statist. Data Anal. 46*(2), 397–406.

[6] Nelsen, R. B. (2006). *An Introduction to Copulas*. Second edition. Springer, New York.

[7] Olkin, I. and R. Liu (2003). A bivariate beta distribution. *Statist. Probab. Lett. 62*(4), 407–412.

DE GRUYTER Depend. Model. 2021; 9:385–393

Research Article Open Access
Special Issue in memory of Abe Sklar

Mhamed Mesfioui* and Julien Trufin

Dispersive order comparisons on extreme order statistics from homogeneous dependent random vectors

https://doi.org/10.1515/demo-2021-0118
Received May 31, 2021; accepted November 10, 2021

Abstract: In this paper, we investigate sufficient conditions for preservation property of the dispersive order for the smallest and largest order statistics of homogeneous dependent random vectors. Moreover, we establish sufficient conditions for ordering with the dispersive order the largest order statistics from dependent homogeneous samples of different sizes.

Keywords: Dispersive order, copulas, Archimedean copulas, extreme order statistics

MSC: 60E15, 62G30, 62H05

1 Introduction

Order statistics play an important role in statistics, risk management, auction theory, reliability and many other theoretical and applied probability areas. They have received a lot of attention from many researchers. For comprehensive references one may refer to Balakrishnan and Rao ([1], [2]) and David and Nagaraja [3].

For a random vector $\boldsymbol{X} = (X_1, \ldots, X_n)$, denote as $X_{i:n}$ the corresponding ith order statistic, $i = 1, \ldots, n$. Most of the research on stochastic comparisons between order statistics has been dedicated to the case of independent and identically distributed (i.i.d.) random variables $X_1, \ldots, X_n$. We can quote Boland et al. [4] who prove stochastic comparisons between order statistics with the hazard rate order and the likelihood ratio order, Raqab and Amin [5] who further prove comparison with the likelihood ratio order between order statistics from samples of different sizes, Kochar [6] who proves comparisons with the dispersive order between order statistics for decreasing failure rate (DFR) distributions or Khaledi and Kochar [7] who study comparisons with the dispersive order between order statistics from samples of different sizes for DFR distributions.

In practical situations, the observations are usually not i.i.d. During the last three decades, the case of independent but not necessarily identically distributed random variables has also got the attention of researchers. We refer the reader to Kochar [8] and Balakrishnan and Zhao [9] for comprehensive references. To mention a few with the dispersive order, Dykstra et al. [10] study comparisons of the largest order statistics with the dispersive order for independent exponential random variables. Khaledi and Kochar [11] extend the latter result from the exponential case to the proportional hazard PH) sample. More recently, Fang and Zhang [12] have obtained the dispersive order between maximums of one heterogeneous- and one homogeneous-independent samples for Weibull random variables sharing a common shape parameter.

In the last decade, some papers have been devoted to the study of the ordering properties of the order statistics from dependent samples. For example, Li and Fang [13] generalize the result in Fang and Zhang

***Corresponding Author: Mhamed Mesfioui:** Département de mathématiques et d'informatique, Université du Québec à Trois-Rivières, Trois-Rivières (Québec), Canada G9A 5H7
Julien Trufin: Department of Mathematics, Université Libre de Bruxelles (ULB), Bruxelles, Belgium

[12] to the dependent case where two proportional hazards samples have a common Archimedean copula and one has heterogeneous hazards and the other has the homogeneous arithmetic average hazards. For two Weibull samples having a common Archimedean survival copula, Li and Li [14] further prove dispersive order inequalities between minimums of one heterogeneous and one homogeneous samples. Fang et al. [15] derive the usual stochastic order, the dispersive order and the star order of order statistics from the PH sample with Archimedean survival copulas and the proportional reversed hazards (PRH) sample with Archimedean copulas. Li et al. [16] investigate order statistics from random variables following the scale model and obtain in the presence of the Archimedean copula or survival copula the usual stochastic order of the sample extremes and the second smallest order statistic, the dispersive order and the star order of the sample extremes.

In this paper, we consider homogeneous dependent samples and we establish sufficient conditions on the copulas or survival copulas and the marginal distribution functions in order to preserve the dispersive order for the smallest and largest order statistics. Moreover, we obtain sufficient conditions on the copulas and the marginal distribution functions for ordering with the dispersive order the largest order statistics from dependent homogeneous samples of different sizes.

The paper is structured as follows. In Section 2, we introduce the useful concepts that will be used in the rest of the paper. Section 3 is devoted to homogeneous dependent samples with different copulas and marginal distributions, and we obtain sufficient conditions for preserving the dispersive order for the smallest and largest order statistics. Finally, in Section 4, we derive sufficient conditions for the dispersive order of the sample extremes in the case of dependent homogeneous samples of different sizes with common copulas and marginal distributions.

2 Preliminaries

In this section, we recall the concepts that are important in the following.

Let X and Y be two random variables with their respective distribution functions F and G, survival functions $\bar{F}$ and $\bar{G}$ and right continuous inverses F^{-1} and G^{-1}.

Definition 2.1. X is said to be smaller than Y in the dispersive order, denoted $X \preceq_{disp} Y$, if

$$F^{-1}(\beta) - F^{-1}(\alpha) \le G^{-1}(\beta) - G^{-1}(\alpha) \quad \text{for all } 0 < \alpha \le \beta < 1,$$

or equivalently if $G^{-1}(F(x)) - x$ is increasing in x.

For more details on stochastic orders, we refer the reader to Shaked and Shanthikumar [17].

Let $\boldsymbol{X} = (X_1, \dots, X_n)$ be a random vector with joint distribution function $\boldsymbol{F}$, joint survival function $\bar{\boldsymbol{F}}$, univariate marginal distribution functions $F_1, \dots, F_n$, survival functions $\bar{F}_1, \dots, \bar{F}_n$ and right continuous inverses $F_1^{-1}, \dots, F_n^{-1}$. If there exists $C : [0,1]^n \to [0,1]$ and $\hat{C} : [0,1]^n \to [0,1]$ such that

$$\begin{aligned} C(u_1, \dots, u_n) &= \boldsymbol{F}\left(F_1^{-1}(u_1), \dots, F_n^{-1}(u_n)\right) \\ \hat{C}(u_1, \dots, u_n) &= \bar{\boldsymbol{F}}\left(\bar{F}_1^{-1}(u_1), \dots, \bar{F}_n^{-1}(u_n)\right) \end{aligned}$$

for all $(u_1, \dots, u_n) \in [0,1]^n$, then C and $\hat{C}$ are called the copula and the survival copula of $\boldsymbol{X}$, respectively. The functions $\delta_C(u) = C(u, \dots, u)$ and $\delta_{\hat{C}}(u) = \bar{C}(u, \dots, u)$ are known as the diagonal sections of C and $\hat{C}$.

The Archimedean copulas form an important class of copulas.

Definition 2.2. Let $\phi : [0, +\infty) \to [0,1]$ with $\phi(0) = 1$ and $\phi(+\infty) = 0$. A n-dimensional copula C_ϕ is said to be an Archimedean copula with generator ϕ if, for all $(u_1, \dots, u_n) \in [0,1]^n$,

$$C_\phi(u_1, \dots, u_n) = \phi\left(\psi(u_1) + \dots + \psi(u_n)\right),$$

where we denote $\psi = \phi^{-1}$ the right continuous inverse of ϕ for convenience.

For $(n-2)$th differentiable ϕ, we know from McNeil and Nešlehová [18] that the function $C_\phi : [0,1]^n \to [0,1]$ is a n-dimensional copula if, and only if, the generator ϕ is a n-monotone function. Denoting $\phi^{(i)}(x)$ the ith derivative of ϕ, recall that ϕ is said to be n-monotone if $(-1)^i \phi^{(i)}(x) \geq 0$ for $i = 1, \dots, n-2$ and $(-1)^{n-2}\phi^{(n-2)}(x)$ is non-increasing and convex.

3 Comparison of the smallest and largest order statistics with the dispersive order

Let $\boldsymbol{X} = (X_1, \dots, X_n)$ and $\boldsymbol{Y} = (Y_1, \dots, Y_n)$ be two homogeneous random vectors with copulas $C_{\boldsymbol{X}}$ and $C_{\boldsymbol{Y}}$ and univariate marginal distribution functions F and G, respectively. Under the conditions that both $C_{\boldsymbol{X}}$ and $C_{\boldsymbol{Y}}$ are the independence copula, Theorem 3.B.26 in Shaked and Shanthikumar [17] states that if $X \preceq_{disp} Y$, then $X_{i:n} \preceq_{disp} Y_{i:n}$ for $i = 1, \dots, n$, where $X_{i:n}$ and $Y_{i:n}$ are the ith order statistics of $\boldsymbol{X}$ and $\boldsymbol{Y}$, respectively. As mentioned in Shaked and Shanthikumar [17], this preservation property of the dispersive order is useful in reliability theory and in nonparametric statistics. Actually, the latter result can be extended for $\boldsymbol{X}$ and $\boldsymbol{Y}$ having a common copula, as shown next.

Proposition 3.1. *If $C_{\boldsymbol{X}} = C_{\boldsymbol{Y}}$ and $X \preceq_{disp} Y$, then $X_{i:n} \preceq_{disp} Y_{i:n}$ for $i = 1, \dots, n$.*

Proof. Since $\boldsymbol{X}$ and $\boldsymbol{Y}$ have a common copula, their exists a uniform random vector $(U_1, \dots, U_n)$ with distribution $C_{\boldsymbol{X}}$ such that $X_i = F^{-1}(U_i)$ and $Y_i = G^{-1}(U_i)$, so that we have $X_{i:n} = F^{-1}(U_{i:n})$ and $Y_{i:n} = G^{-1}(U_{i:n})$. This ensures that the distributions of $X_{i:n}$ and $Y_{i:n}$ are of the form $F_{i:n} = H_{i:n} \circ F$ and $G_{i:n} = H_{i:n} \circ G$, respectively, where $H_{i:n}$ denotes the distribution of $U_{i:n}$. The announced result then follows from $G_{i:n}^{-1}(F_{i:n}(x)) - x = G^{-1}(F(x)) - x$ for all x. □

Proposition 3.1 considers samples $\boldsymbol{X}$ and $\boldsymbol{Y}$ with a common copula. The next result show that the dispersive order can still be preserved for the smallest and largest order statistics when we consider different copulas for $\boldsymbol{X}$ and $\boldsymbol{Y}$ at the cost of requiring additional conditions on the marginal distribution G and the diagonal sections $\delta_{C_{\boldsymbol{X}}}$ and $\delta_{C_{\boldsymbol{Y}}}$ of the copulas $C_{\boldsymbol{X}}$ and $C_{\boldsymbol{Y}}$ for the largest order statistics, and on the survival function $\bar{G}$ and the diagonal sections $\delta_{\hat{C}_{\boldsymbol{X}}}$ and $\delta_{\hat{C}_{\boldsymbol{Y}}}$ of the survival copulas $\hat{C}_{\boldsymbol{X}}$ and $\hat{C}_{\boldsymbol{Y}}$ for the smallest order statistics.

Proposition 3.2. *For F and G twice differentiable,*
(i) if $\delta^{-1}_{C_{\boldsymbol{X}}}(\alpha)/\delta^{-1}_{C_{\boldsymbol{Y}}}(\alpha)$ is decreasing in $\alpha \in [0,1]$, G is log-convex and $X \preceq_{disp} Y$, then $X_{n:n} \preceq_{disp} Y_{n:n}$;
(ii) if $\delta^{-1}_{\hat{C}_{\boldsymbol{X}}}(\alpha)/\delta^{-1}_{\hat{C}_{\boldsymbol{Y}}}(\alpha)$ is decreasing in $\alpha \in [0,1]$, $\bar{G}$ is log-convex and $X \preceq_{disp} Y$, then $X_{1:n} \preceq_{disp} Y_{1:n}$.

Proof. (i) Obviously, the quantile functions of $X_{n:n}$ and $Y_{n:n}$ are $F^{-1}\{\delta^{-1}_{C_{\boldsymbol{X}}}(\alpha)\}$ and $G^{-1}\{\delta^{-1}_{C_{\boldsymbol{Y}}}(\alpha)\}$, $\alpha \in [0,1]$, respectively. Thus, $X_{n:n} \preceq_{disp} Y_{n:n}$ if, and only if, $F^{-1}\{\delta^{-1}_{C_{\boldsymbol{X}}}(\alpha)\} - G^{-1}\{\delta^{-1}_{C_{\boldsymbol{Y}}}(\alpha)\}$ is decreasing in α. This latter difference can be rewritten as the sum of

$$A(\alpha) = F^{-1}\{\delta^{-1}_{C_{\boldsymbol{X}}}(\alpha)\} - G^{-1}\{\delta^{-1}_{C_{\boldsymbol{X}}}(\alpha)\}$$

and

$$B(\alpha) = G^{-1}\{\delta^{-1}_{C_{\boldsymbol{X}}}(\alpha)\} - G^{-1}\{\delta^{-1}_{C_{\boldsymbol{Y}}}(\alpha)\}.$$

Now, since $X \preceq_{disp} Y$ implies that $A(\alpha)$ is decreasing in α (since $\delta^{-1}_{C_{\boldsymbol{X}}}(\alpha)$ is an increasing function), it suffices to prove that $B(\alpha)$ is decreasing in α as well. Simple calculations yield

$$\begin{aligned} B'(\alpha) &= (G^{-1})'\{\delta^{-1}_{C_{\boldsymbol{X}}}(\alpha)\}(\delta^{-1}_{C_{\boldsymbol{X}}})'(\alpha) - (G^{-1})'\{\delta^{-1}_{C_{\boldsymbol{Y}}}(\alpha)\}(\delta^{-1}_{C_{\boldsymbol{Y}}})'(\alpha) \\ &= \frac{\delta^{-1}_{C_{\boldsymbol{X}}}(\alpha)}{G'(G^{-1}(\delta^{-1}_{C_{\boldsymbol{X}}}(\alpha)))}\Big(\ln\{\delta^{-1}_{C_{\boldsymbol{X}}}(\alpha)\}\Big)' - \frac{\delta^{-1}_{C_{\boldsymbol{X}}}(\alpha)}{G'(G^{-1}(\delta^{-1}_{C_{\boldsymbol{Y}}}(\alpha)))}\Big(\ln\{\delta^{-1}_{C_{\boldsymbol{Y}}}(\alpha)\}\Big)' \\ &= h\{G^{-1}(\delta^{-1}_{C_{\boldsymbol{X}}}(\alpha))\}\Big(\ln\{\delta^{-1}_{C_{\boldsymbol{X}}}(\alpha)\}\Big)' - h\{G^{-1}(\delta^{-1}_{C_{\boldsymbol{Y}}}(\alpha))\}\Big(\ln\{\delta^{-1}_{C_{\boldsymbol{Y}}}(\alpha)\}\Big)', \end{aligned} \tag{3.1}$$

where $h(t) = G(t)/G'(t)$. As $\delta_{C_X}^{-1}(\alpha)/\delta_{C_Y}^{-1}(\alpha)$ is decreasing in $\alpha \in [0,1]$ and $\delta_{C_X}^{-1}(1)/\delta_{C_Y}^{-1}(1) = 1$, one has $0 \le \left(\ln\{\delta_{C_X}^{-1}(\alpha)\}\right)' \le \left(\ln\{\delta_{C_Y}^{-1}(\alpha)\}\right)'$ and $\delta_{C_X}^{-1}(\alpha) \ge \delta_{C_Y}^{-1}(\alpha)$. Hence, from (3.1), we have

$$B'(\alpha) \le \left[h\{G^{-1}(\delta_{C_X}^{-1}(\alpha))\} - h\{G^{-1}(\delta_{C_Y}^{-1}(\alpha))\}\right]\left(\ln\{\delta_{C_X}^{-1}(\alpha)\}\right)'. \tag{3.2}$$

Finally, since G is log-convex, it is plain that h is a decreasing function so that the right-hand side in (3.2) is negative.

(ii) The quantile function of $X_{1:n}$ is clearly given by $F_{1:n}^{-1}(\alpha) = \bar{F}^{-1}(\delta_{\hat{C}_X}^{-1}(1-\alpha))$, $\alpha \in [0,1]$. So, $X_{1:n} \preceq_{disp} Y_{1:n}$ if, and only if, $\bar{F}^{-1}\{\delta_{\hat{C}_X}^{-1}(\alpha)\} - \bar{G}^{-1}\{\delta_{\hat{C}_Y}^{-1}(\alpha)\}$ is increasing in α, which is the case when $\hat{C}(\alpha) = \bar{F}^{-1}\{\delta_{\hat{C}_Y}^{-1}(\alpha)\} - \bar{G}^{-1}\{\delta_{\hat{C}_Y}^{-1}(\alpha)\}$ is increasing in α since $X \preceq_{disp} Y$. Proceeding in a similar manner than in (3.1) and using the fact that $\delta_{\hat{C}_X}^{-1}(\alpha)/\delta_{\hat{C}_Y}^{-1}(\alpha)$ is decreasing in α and $\delta_{\hat{C}_X}^{-1}(1)/\delta_{\hat{C}_Y}^{-1}(1) = 1$, one has

$$C'(\alpha) \ge \left[\bar{h}\{\bar{G}^{-1}(\delta_{\hat{C}_X}^{-1}(\alpha))\} - \bar{h}\{\bar{G}^{-1}(\delta_{\hat{C}_Y}^{-1}(\alpha))\}\right]\left(\ln\{\delta_{\hat{C}_X}^{-1}(\alpha)\}\right)', \tag{3.3}$$

where $\bar{h}(t) = \bar{G}(t)/\bar{G}'(t)$. Consequently, as $\bar{G}$ is log-convex, it is plain that $\bar{h}$ is a decreasing function so that the right-hand side in (3.3) is positive. □

In particular, for Archimedean copulas, we directly get the following result.

Corollary 3.3. *(i) For $C_{\boldsymbol{X}} = C_{\phi_1}$ and $C_{\boldsymbol{Y}} = C_{\phi_2}$, where C_{ϕ_1} and C_{ϕ_2} are two Archimedean copulas, if $\phi_1(\psi_1(\alpha)/n))/\phi_2(\psi_2(\alpha)/n))$ is decreasing in $\alpha \in [0,1]$, G is log-convex and $X \preceq_{disp} Y$, then $X_{n:n} \preceq_{disp} Y_{n:n}$.*

(ii) For $\hat{C}_{\boldsymbol{X}} = \hat{C}_{\phi_1}$ and $\hat{C}_{\boldsymbol{Y}} = \hat{C}_{\phi_2}$, where $\hat{C}_{\phi_1}$ and $\hat{C}_{\phi_2}$ are two Archimedean copulas, if $\phi_1(\psi_1(\alpha)/n))/\phi_2(\psi_2(\alpha)/n))$ is decreasing in $\alpha \in [0,1]$, $\bar{G}$ is log-convex and $X \preceq_{disp} Y$, then $X_{1:n} \preceq_{disp} Y_{1:n}$.

Proof. It suffices to notice that the diagonal section of an Archimedean copula with generator ϕ is the form $\phi(\psi(\alpha)/n))$. Then, the result immediately follows from Proposition 3.2. □

We illustrate Corollary 3.3 by an example where $\phi_1(\psi_1(\alpha)/n))/\phi_2(\psi_2(\alpha)/n))$ is decreasing in α.

Example 3.4. (i) Consider

$$\phi_1(t) = (\theta_1 t + 1)^{-1/\theta_1}$$

and

$$\phi_2(t) = (\theta_2 t + 1)^{-1/\theta_2}$$

with $0 < \theta_1 \le \theta_2$, so that C_{ϕ_1} and C_{ϕ_2} are Clayton copulas. Simple calculations yield

$$\begin{aligned} k_1(\alpha) &= \phi_1(\psi_1(\alpha)/n))/\phi_2(\psi_2(\alpha)/n)) \\ &= n^{1/\theta_1 - 1/\theta_2} \frac{\left(\alpha^{-\theta_1} - 1 + n\right)^{-1/\theta_1}}{\left(\alpha^{-\theta_2} - 1 + n\right)^{-1/\theta_2}}, \end{aligned}$$

so that $k_1(\alpha)$ is decreasing in α when $0 < \theta_1 \le \theta_2$, as illustrated in Figure 3.1 for $n = 3$, $\theta_1 = 1$ and $\theta_2 = 2$. As a result, from Corollary 3.3 (i), we know that if G is log-convex and $X \preceq_{disp} Y$, then $X_{n:n} \preceq_{disp} Y_{n:n}$. A log-convex distribution function G whose support is $]-\infty, a]$ with $a \in \mathbb{R}$ can be expressed as

$$G(x) = (e^{h(x)} - 1)\mathrm{I}[x < a] + 1, \quad x \in \mathbb{R}, \tag{3.4}$$

where h is an increasing and convex function such that $h(-\infty) = -\infty$ and $h(a) = 0$, and where $\mathrm{I}[\cdot]$ denotes the indicator function, equal to 1 if the event appearing in the brackets is realized and to 0 otherwise. As an example, distribution functions of the form

$$G(x) = (e^{-|x|^\alpha} - 1)\mathrm{I}[x < 0] + 1, \quad x \in \mathbb{R},$$

with $\alpha \le 1$ are log-convex. In particular, for two homogeneous samples with the same marginal distribution functions, that is when $F = G$, Corollary 3.3 (i) tells us that the variability (in terms of the dispersive order) of the largest statistics is increasing in the dependence parameter θ.

(ii) Likewise, if we consider two Clayton copulas $\hat{C}_{\phi_1}$ and $\hat{C}_{\phi_2}$ for the survival copulas with parameters $0 < \theta_1 \le \theta_2$, respectively, then we know from Corollary 3.3 (ii) that if $\bar{G}$ is log-convex and $X \preceq_{disp} Y$, then $X_{1:n} \preceq_{disp} Y_{1:n}$. Well-known distributions have log-convex survival functions $\bar{G}$, as shown in [19]. Let us mention the Pareto distribution, the Gamma distribution with density function $f(x) = \frac{x^{c-1}\exp(-c)}{\Gamma(c)}$ and $0 < c < 1$ as well as the Weibull distribution with density function $f(x) = cx^{c-1}\exp(-x^c)$ and $0 < c < 1$.

The log-convexity condition on the marginal distribution function G in Proposition 3.2 and Corollary 3.3 is necessary to establish dispersive order inequalities among the largest order statistics, as illustrated in the following example.

Example 3.5. Consider $\boldsymbol{U} = (U_1, \dots, U_n)$ and $\boldsymbol{V} = (V_1, \dots, V_n)$ two homogeneous uniform random vectors distributed as C_{ϕ_1} and C_{ϕ_2}, respectively. Clearly, $U_{n:n} \preceq_{disp} V_{n:n}$ if, and only if $\phi_1(\psi_1(\alpha)/n)) - \phi_2(\psi_2(\alpha)/n))$ is decreasing in α, which is actually not fulfilled for most of the Archimedean copulas. In particular, if we consider the Clayton copula, we get

$$\begin{aligned} k_2(\alpha) &= \phi_1(\psi_1(\alpha)/n)) - \phi_2(\psi_2(\alpha)/n) \\ &= n^{1/\theta_1}\left(\alpha^{-\theta_1} - 1 + n\right)^{-1/\theta_1} - n^{1/\theta_2}\left(\alpha^{-\theta_2} - 1 + n\right)^{-1/\theta_2}. \end{aligned}$$

As a result, $k_2(\alpha)$ is not monotone, as illustrated in Figure 3.2 for $n = 3$, $\theta_1 = 1$ and $\theta_2 = 2$. Therefore, $U_{n:n}$ and $V_{n:n}$ cannot be compared with the dispersive order, the reason is that the uniform distribution is log-concave and not log-convex.

It is interesting to notice that $\delta^{-1}_{C_X}(\alpha)/\delta^{-1}_{C_Y}(\alpha)$ is decreasing in $\alpha \in [0,1]$ if, and only if, $\ln \delta^{-1}_{C_X}(\alpha) - \ln \delta^{-1}_{C_Y}(\alpha)$ is decreasing in $\alpha \in [0,1]$, which amounts to requiring $\ln U_{n:n} \preceq_{disp} \ln V_{n:n}$. Therefore the condition in Proposition 3.2 that $\delta^{-1}_{C_X}(\alpha)/\delta^{-1}_{C_Y}(\alpha)$ is decreasing in $\alpha \in [0,1]$ can be rewritten as $\ln U_{n:n} \preceq_{disp} \ln V_{n:n}$.

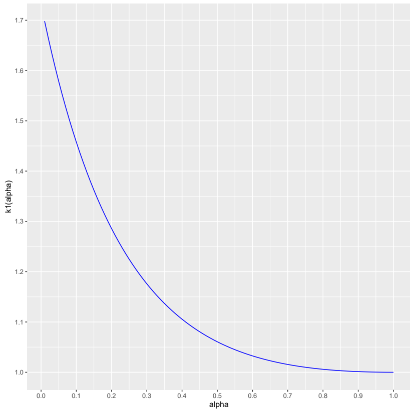

Figure 3.1: $k_1(\alpha) = \phi_1(\psi_1(\alpha)/n))/\phi_2(\psi_2(\alpha)/n))$ for $\alpha \in [0,1]$.

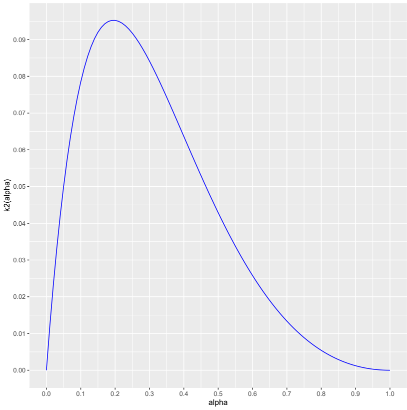

Figure 3.2: $k_2(\alpha) = \phi_1\left(\psi_1(\alpha)/n\right) - \phi_2\left(\psi_2(\alpha)/n\right)$ for $\alpha \in [0,1]$.

4 Comparison of the largest order statistics with the dispersive order for different sample sizes

Let $\boldsymbol{X}_n = (X_1, \ldots, X_n)$ be an homogeneous random vector with common distribution function F assumed to be twice differentiable and let $\boldsymbol{X}_{n-1} = (X_1, \ldots, X_{n-1})$. In this section, we aim to compare the largest order statistics of samples $\boldsymbol{X}_n$ and $\boldsymbol{X}_{n-1}$ with the dispersive order. We denote by δ_n and δ_{n-1} the diagonal sections of the copulas of $\boldsymbol{X}_n$ and $\boldsymbol{X}_{n-1}$, respectively.

Proposition 4.1. *If F is log-convex and $\delta_n^{-1}(\alpha)/\delta_{n-1}^{-1}(\alpha)$ is decreasing in $\alpha \in [0,1]$, then*

$$X_{n:n} \preceq_{disp} X_{n-1:n-1}.$$

Proof. First, we know that $F_{n:n}^{-1}(\alpha) = F^{-1}\{\delta_n^{-1}(\alpha)\}$ and $F_{n-1:n-1}^{-1}(\alpha) = F^{-1}\{\delta_{n-1}^{-1}(\alpha)\}$. Therefore, we clearly have that $X_{n:n} \preceq_{disp} X_{n-1:n-1}$ if, and only if, $D(\alpha) = F^{-1}\{\delta_n^{-1}(\alpha)\} - F^{-1}\{\delta_{n-1}^{-1}(\alpha)\}$ is a decreasing function. Similar calculations than in the proof of Proposition 3.2 yield

$$D'(\alpha) = k\{F^{-1}(\delta_n^{-1}(\alpha))\}\left(\ln\{\delta_n^{-1}(\alpha)\}\right)' - k\{F^{-1}(\delta_{n-1}^{-1}(\alpha))\}\left(\ln\{\delta_{n-1}^{-1}(\alpha)\}\right)'$$

where $k(t) = F(t)/F'(t)$. Since, $\delta_n^{-1}(\alpha)/\delta_{n-1}^{-1}(\alpha)$ is decreasing in $\alpha \in [0,1]$, one has

$$D'(\alpha) \le \left(k\{F^{-1}(\delta_n^{-1}(\alpha))\} - k\{F^{-1}(\delta_{n-1}^{-1}(\alpha))\}\right)\left(\ln\{\delta_n^{-1}(\alpha)\}\right)'.$$

Using the fact that $\delta_{n-1}^{-1}(\alpha) \le \delta_n^{-1}(\alpha)$, it follows that $D'(\alpha) \le 0$ if $k = F/F'$ is a decreasing function, that is, if F is log-convex. □

For Archimedean copulas, Proposition 4.1 directly leads to the next result.

Corollary 4.2. *For $C_{\boldsymbol{X}} = C_\phi$ where C_ϕ is an Archimedean copula with generator ϕ, if F is log-convex and $t\phi'(t)/\phi(t)$ is decreasing in $t > 0$, then*

$$X_{n:n} \preceq_{disp} X_{n-1:n-1}.$$

Proof. Since the copula C_ϕ is Archimedean with generator ϕ, we have $\delta_n^{-1}(\alpha) = \phi\left(\frac{\psi(\alpha)}{n}\right)$ and $\delta_{n-1}^{-1}(\alpha) = \phi\left(\frac{\psi(\alpha)}{n-1}\right)$. Consequently, $\delta_n^{-1}(\alpha)/\delta_{n-1}^{-1}(\alpha)$ is a decreasing function, if, and only if,

$$\frac{\phi'\left(\frac{\psi(\alpha)}{n}\right)}{\phi\left(\frac{\psi(\alpha)}{n}\right)}\frac{\psi(\alpha)}{n} \geq \frac{\phi'\left(\frac{\psi(\alpha)}{n-1}\right)}{\phi\left(\frac{\psi(\alpha)}{n-1}\right)}\frac{\psi(\alpha)}{n-1}$$

for all $\alpha \in (0,1)$. The latter inequality is fulfilled for all $\alpha \in (0,1)$ since $t\phi'(t)/\phi(t)$ is decreasing in $t > 0$. □

The following example illustrates the condition on ϕ involved in Corollary 4.2. As mentioned in Li and Fang [20] and Mesfioui et al. [21], this condition is satisfied for most of the Archimedean copulas, such as Clayton, Frank and Gumbel copulas.

Example 4.3. Consider the generator $\phi(t) = (\theta t + 1)^{-1/\theta}$ with $\theta > 0$ of a Clayton copula. It verifies

$$\left(\frac{t\phi'(t)}{\phi(t)}\right)' = \frac{-1}{(\theta t + 1)^2} < 0,$$

so that for an homogeneous random vector $\boldsymbol{X} = (X_1, \ldots, X_n)$ with such a dependence structure and a log-convex marginal distribution function F, we know from Corollary 4.2 that we have

$$X_{n:n} \preceq_{disp} X_{n-1:n-1} \preceq_{disp} \cdots \preceq_{disp} X_{1:1}.$$

As already discussed in the previous section for Proposition 3.2 and Corollary 3.3, the log-convexity of F required in Proposition 4.1 and Corollary 4.2 is also crucial here to ensure the ordering of $X_{n:n}$ and $X_{n-1:n-1}$ with the dispersive order, as revealed by the next example.

Example 4.4. Consider the homogeneous uniform random vector $\boldsymbol{U} = (U_1, \ldots, U_n)$ distributed as the independence copula C_ϕ. The generator of the independence copula, that is $\phi(t) = e^{-t}$, satisfies $t\phi'(t)/\phi(t) = -t$, so that it is well decreasing in $t > 0$, as required in Corollary 4.2. However, it is easy to see that the uniform distribution F is not log-convex but log-concave. In this case, the quantile functions of $X_{n-1:n-1}$ and $X_{n:n}$ are $F_{n-1:n-1}^{-1}(\alpha) = \alpha^{\frac{1}{n-1}}$ and $F_{n:n}^{-1}(\alpha) = \alpha^{\frac{1}{n}}$, $\alpha \in [0,1]$, respectively, which implies that the function $g(\alpha) = F_{n:n}^{-1}(\alpha) - F_{n-1:n-1}^{-1}(\alpha)$ is not monotone, as illustrated in Figure 4.1 for $g(\alpha) = F_{3:3}^{-1}(\alpha) - F_{2:2}^{-1}(\alpha)$. As a result, $X_{n-1:n-1}$ and $X_{n:n}$ cannot be compared with the dispersive order.

Actually, for a uniform random vector $\boldsymbol{U} = (U_1, \ldots, U_n)$ distributed as C_ϕ, a necessary and sufficient condition on the generator ϕ to get $U_{n:n} \preceq_{disp} U_{n-1:n-1}$ can be obtained, as shown next.

Proposition 4.5. *For an homogeneous uniform random vector $\boldsymbol{U} = (U_1, \ldots, U_n)$ distributed as C_ϕ, we have that*

$$U_{n:n} \preceq_{disp} U_{n-1:n-1} \iff t\phi'(t) \text{ is a decreasing function in } t > 0.$$

Proof. Since the quantile function of $U_{n:n}$ is $F_{n:n}^{-1}(x) = \phi\left(\frac{\psi(u)}{n}\right)$, we get

$$\begin{aligned}
U_{n:n} \preceq_{disp} U_{n-1:n-1} &\iff \phi\left(\frac{\psi(u)}{n}\right) - \phi\left(\frac{\psi(u)}{n-1}\right) \text{ is a decreasing function in } u \in [0,1] \\
&\iff \phi\left(\frac{x}{n}\right) - \phi\left(\frac{x}{n-1}\right) \text{ is an increasing function in } x > 0 \\
&\iff \frac{x}{n-1}\phi'\left(\frac{x}{n-1}\right) \leq \frac{x}{n}\phi'\left(\frac{x}{n}\right) \text{ for all } x > 0 \\
&\iff t\phi'(t) \text{ is a decreasing function in } t > 0.
\end{aligned}$$

□

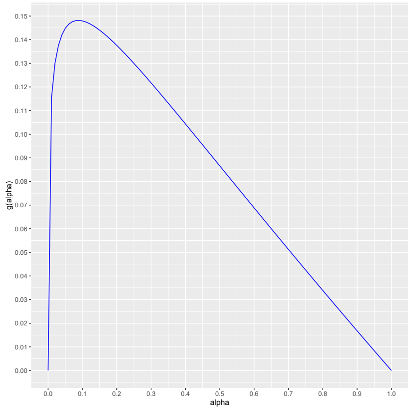

Figure 4.1: $g(\alpha) = F^{-1}_{3:3}(\alpha) - F^{-1}_{2:2}(\alpha)$ for $\alpha \in [0, 1]$.

Unfortunately, the function $t\phi'(t)$ is not monotone for most of the Archimedean copulas, including the independence, Clayton, Frank and Gumbel copulas. Note that this is the reason why an additional condition on F is needed for Proposition 4.1 to get the dispersive order for $X_{n-1:n-1}$ and $X_{n:n}$.

Conflict of interest statement: Authors state no conflict of interest.

Acknowledgments: Mhamed Mesfioui acknowledges the financial support of the Natural Sciences and Engineering Research Council of Canada No 06536-2018.

References

[1] Balakrishnan N. and C.R. Rao (1998). *Order Statistics: Theory and Methods*. Elsevier, New York.
[2] Balakrishnan N. and C.R. Rao (1998). *Order Statistics: Applications*. Elsevier, New York.
[3] David H.A. and H.N. Nagaraja (2003). *Order Statistics*. Third edition. Wiley, New York.
[4] Boland P.J., E. El-Neweihi and F. Proschan (1994). Applications of the hazard rate ordering in reliability and order statistics. *J. Appl. Probab. 31*(1), 180–192.
[5] J. Appl. Probab. Raqab M.Z. and W.A. Amin (1996). Some ordering results on order statistics and record values. *IAPQR Transactions 21*, 1–8.
[6] S. C. Kochar (1996). Dispersive ordering of order statistics. *Statist. Probab. Lett. 27*(3), 271–274.
[7] Khaledi B.-E.and S. C. Kochar (2000). On dispersive ordering between order statistics in one-sample and two-sample problems. *Statist. Probab. Lett. 46*(3), 257–261.
[8] S. C. Kochar (2012). Stochastic comparisons of order statistics and spacings: a review. *ISRN Probab. Statist. 2012*, Article ID 839473, 47 pages.
[9] Balakrishnan N. and P. Zhao (2013). Ordering properties of order statistics from heterogeneous populations: a review with an emphasis on some recent developments. *Probab. Eng. Inf. Sci. 27*(4), 403–443.
[10] Dykstra R., S. C. Kochar and J. Rojo (1997). Stochastic comparisons of parallel systems of heterogeneous exponential components. *J. Statist. Plann. Inference 65*(2), 203–211.
[11] Khaledi B.-E. and S. C. Kochar (2000). Some new results on stochastic comparisons of parallel systems. *J. Appl. Probab. 37*(4), 1123–1128.
[12] Fang L. and X. Zhang (2013). Stochastic comparisons of series systems with heterogeneous Weibull components. *Statist. Probab. Lett. 83*(7), 1649–1653.

[13] Li X. and, R. Fang (2015). Ordering properties of order statistics from random variables of Archimedean copula with applications. *J. Multivariate Anal. 133*, 304–320.

[14] Li C. and X. Li (2015). Likelihood ratio order of sample minimum from heterogeneous Weibull random variables. *Statist. Probab. Lett. 97*, 46–53.

[15] Fang R., X. Li and C. Li (2016). Stochastic comparisons on sample extremes of dependent and heterogenous observations. *Statistics 50*(4), 930-955.

[16] Li C., R. Fang and X. Li (2016). Stochastic comparisons of order statistics from scaled and interdependent random variables. *Metrika 79*, 553-578.

[17] Shaked M. and J. G. Shanthikumar (2007). *Stochastic Orders.* Springer, New York.

[18] McNeil A.J. and J. Nešlehová (2009). Multivariate Archimedean copulas, D-monotone functions and l_1-norm symmetric distributions. *Ann. Statist. 37*(5B), 3059-3097.

[19] Bagnoli, M. and T. Bergstrom (2005). Log-concave probability and its applications. *Econom. Theory 26*, 445-469.

[20] Li X. and R. Fang (2015). Ordering properties of order statistics from random variables of Archimedean copulas with applications. *J. Multivariate Anal. 133*, 304-320.

[21] Mesfioui M., M. Kayid and S. Izadkhah (2017). Stochastic comparisons of order statistics from heterogeneous random variables with Archimedean copula. *Metrika 80*, 749-760.

Depend. Model. 2021; 9:394–423 **DE GRUYTER**

Research Article **Open Access**
Special Issue in memory of Abe Sklar

Rachele Foschi, Giovanna Nappo*, and Fabio L. Spizzichino

Diagonal sections of copulas, multivariate conditional hazard rates and distributions of order statistics for minimally stable lifetimes

https://doi.org/10.1515/demo-2021-0119
Received June 10, 2021; accepted November 11, 2021

Abstract: As a motivating problem, we aim to study some special aspects of the marginal distributions of the order statistics for exchangeable and (more generally) for *minimally stable* non-negative random variables $T_1, \ldots, T_r$. In any case, we assume that $T_1, \ldots, T_r$ are identically distributed, with a common survival function $\overline{G}$ and their survival copula is denoted by K. The diagonal sections of K, along with $\overline{G}$, are possible tools to describe the information needed to recover the laws of order statistics.
When attention is restricted to the absolutely continuous case, such a joint distribution can be described in terms of the associated multivariate conditional hazard rate (m.c.h.r.) functions. We then study the distributions of the order statistics of $T_1, \ldots, T_r$ also in terms of the system of the m.c.h.r. functions. We compare and, in a sense, we combine the two different approaches in order to obtain different detailed formulas and to analyze some probabilistic aspects for the distributions of interest. This study also leads us to compare the two cases of exchangeable and minimally stable variables both in terms of copulas and of m.c.h.r. functions. The paper concludes with the analysis of two remarkable special cases of stochastic dependence, namely Archimedean copulas and load sharing models. This analysis will allow us to provide some illustrative examples, and some discussion about peculiar aspects of our results.

Keywords: Minimally stable random vectors, diagonal sections of survival copulas, diagonal dependence, t-exchangeability, absolute continuity, Archimedean copulas, load-sharing models

MSC: 60E05, 62H05, 62G30, 60K10, 62N05

1 Introduction

Concerning the basic role of the concept of copula and of the Sklar's theorem in the analysis of stochastic dependence, a main issue is the study of the distributions of the order statistics $X_{1:r}, \ldots, X_{r:r}$ for a set of interdependent random variables $X_1, \ldots, X_r$. On the one hand, the condition of exchangeability is specially relevant (see in particular Galambos [11]) in such a study. On the other hand, the marginal distributions of $X_{1:r}, \ldots, X_{r:r}$ are strictly related to the diagonal sections of copulas (see, e.g., Jaworski [12], Durante and Sempi [9]). For these reasons, in the theory of order statistics, the study of diagonal sections of copulas has been mainly concentrated on the case of exchangeable random variables.

Rachele Foschi: University of Pisa, Department of Economics and Management, Via C. Ridolfi, 10, 56124 Pisa, Italy, E-mail: rachele.foschi@unipi.it
***Corresponding Author: Giovanna Nappo:** University of Rome La Sapienza, Department of Mathematics,Piazzale Aldo Moro, 5, 00185, Rome, Italy, E-mail: nappo@mat.uniroma1.it
Fabio L. Spizzichino: University of Rome La Sapienza, Piazzale Aldo Moro, 5, 00185, Rome, Italy, E-mail: fabio.spizzichino@fondazione.uniroma1.it

Really, in such a study, the assumption of exchangeability can at any rate be replaced by the more general condition that, for $d = 2, \dots, r-1$, all the diagonal sections of the d-dimensional marginal copulas do coincide. Such a condition has been attracting more and more interest in the recent literature, where it has been however designated by means of different terminologies. In fact, such a condition can actually manifest under different mathematical forms, as we will discuss in details. For our purposes it is specially convenient to look at it as the condition that $X_1, \dots, X_r$ are minimally stable (see Definition 3 below).

In this note we concentrate attention on the case of non-negative, minimally stable, random variables which we denote by $T_1, \dots, T_r$.

Generally, concerning with non-negative random variables, stochastic dependence can also be conveniently described in terms of stochastic intensities of related counting processes. See in particular Arjas [1], Bremaud [3], Arjas and Norros [2]. Such a description, in particular, can be based on the knowledge of the so-called *multivariate conditional hazard rates (m.c.h.r.)* functions, when attention is restricted to the absolutely continuous case (see in particular the papers by Shaked and Shanthikumar [26–28]). In such a case the family of those functions gives rise to a method to describe a joint distribution, which is alternative to the one based on copulas and marginal distributions or on the joint density function.

From an analytical view-point the two methods are actually equivalent: on the one hand the family of the m.c.h.r. functions can be obtained in terms of the joint density function, on the other hand the joint density can be recovered when the m.c.h.r. functions are known. As a matter of fact, however, the corresponding formulas are not easily handleable in general cases. The two methods, furthermore, are respectively apt to explain completely different aspects of stochastic dependence.

In this paper we aim to establish a bridge between the two different approaches. Maintaining the attention focused on the minimally stable case, then, we are primarily interested in the relations tying the system of the diagonal sections with the system formed by the m.c.h.r. functions. Such relations will allow us to detect, both in terms of copulas and in terms of the m.c.h.r. functions, which are the minimal sets of functions able to convey sufficient information to recover the family of the marginal distributions of the order statistics $T_{1:r}, \dots, T_{r:r}$.

In such a framework, interesting questions also concern with understanding the real difference between the cases when $T_1, \dots, T_r$ are exchangeable and when they are minimally stable. On this purpose, the differences between the two properties will be detailed both using the language of copulas and the language of the m.c.h.r. functions. Still by using and combining the two approaches, we will also face the problem of constructing examples of random variables $T_1, \dots, T_r$ which are minimally stable but not exchangeable.

More in details, the plan of this paper goes as follows.

In Section 2 we introduce some needed notation and then we review basic facts about distributions of order statistics, about diagonal sections of copulas, and about the relations tying these two families of objects. We also show in details the equivalence among different forms under which one can represent the condition that $T_1, \dots, T_r$ are minimally stable. Some relevant remarks are given and an example is presented concerning the construction of random variables which are minimally stable but not exchangeable.

In Section 3 we will first recall, in general, the definition and some basic aspects of the family of the multivariate conditional hazard rate functions. We will then show special features of the cases where the lifetimes $T_1, \dots, T_r$ are exchangeable or minimally stable. In this frame, the results of Section 2 will emerge as natural tools to obtain, in Section 4, the relations existing among diagonal sections of copulas, the distributions of order statistics, and a special subclass $\mathcal{C}$ of multivariate conditional hazard rates (see (37) and (39)), corresponding to the (unconditional) one-dimensional hazard rates of $\min(T_1, \dots, T_\ell)$, $\ell = 1, \dots, r$. See in particular Propositions 18 and 19.

In order to demonstrate some special aspects of the results presented in the Sections 3 and 4, Section 5 will be devoted to a detailed discussion of the remarkable cases of Archimedean copulas and of minimally stable time homogeneous load-sharing models. Some more general examples will be presented in the Appendix.

Often, along the paper, the term *lifetime* will be used as a short-hand for "non-negative random variable".

Notation: For any natural number n, we set $[n] := \{1, 2, \dots, n\}$.

For any subset $J \subseteq [n]$, we denote by $|J|$ the cardinality of J, and as usual we denote by J^c the complementary set of J, i.e., the set of indices $[n] \setminus J$. Furthermore, for any $k \leq |J|$ we denote by

$$\Pi_k(J) = \left\{(j_1, \dots, j_k) : j_\ell \in J, \; \forall \ell = 1, .., k, \; j_\ell \neq j_h, \; \forall \ell \neq h\right\},$$

the set of k-permutations of J. When $k = |J|$ we drop the index k and write simply $\Pi(J)$. The symbol

$$(n)_k := n(n-1)\cdots(n-(k-1))) = |\Pi_k([n])|$$

denotes the number of k-permutations in $\Pi_k([n])$.

For any subset $A = \{j_1, \dots, j_\ell\} \subset [n]$ we denote by $\mathbf{e}_A$ the vector whose i-th component is equal to 1 if $i \in A$, and is equal to 0, otherwise.

2 Diagonal sections and distributions of order statistics

Let $T_1, \dots, T_r$ denote r non-negative random variables, defined on a same probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with joint survival function

$$\overline{F}(t_1, \dots, t_r) := \mathbb{P}\left(T_1 > t_1, \dots, T_r > t_r\right),$$

and survival copula $K : [0, 1]^r \to [0, 1]$.

All over the paper we generally assume the following conditions, unless specified otherwise,

(H1) the random variables $T_1, \dots, T_r$ are identically distributed with common one-dimensional marginal survival function $\overline{G}$, i.e.,

$$\overline{G}(t) := \mathbb{P}\left(T_j > t\right), \quad \text{for } j = 1, \dots, r, \text{ and for } t > 0.$$

(H2) $\overline{G}(t)$ is continuous, strictly positive, and strictly decreasing on $(0, \infty)$.

(H3) the random variables $T_1, \dots, T_r$ are no-ties, i.e., $\mathbb{P}(T_i = T_j) = 0$, for $i \neq j$.

Since $T_1, \dots, T_r$ are non-negative, condition **(H2)** implies that $\overline{G}(0) = 1$ and that $\overline{G}(t)$ is invertible. Though **(H2)** is not strictly necessary (as, for example, in Proposition 8), we assume it for simplicity's sake (for example we use **(H2)** in Proposition 5, Remark 6 and Corollary 9, within this section and somewhere else, within the other sections). Condition **(H3)** allows the order statistics $T_{1:r}, \dots, T_{r:r}$ of $(T_1, \dots, T_r)$ to be defined without ambiguity. We denote by

$$\overline{G}_{1:r}(t) := \mathbb{P}\left(T_{1:r} > t\right), \dots, \overline{G}_{r:r}(t) := \mathbb{P}\left(T_{r:r} > t\right) \tag{1}$$

the corresponding marginal survival functions.

Note that the order statistics $T_{1:r}, \dots, T_{r:r}$ may be considered as the jump times of the counting process

$$N(t) := \sum_{i=1}^{r} \mathbf{1}_{\{T_i \leq t\}}$$

i.e., the process such that $N(t) = k$ for $t \in [T_{k:r}, T_{k+1:r})$, where we have set $T_{0:r} = 0$ and $T_{r+1:r} = \infty$.

Before continuing we recall the following definition.

Definition 1. *For a r-dimensional copula C the* **diagonal section** *is the function*

$$\delta^C : \; [0, 1] \to [0, 1]; \quad u \mapsto \delta^C(u) = C(u, u, \dots, u)$$

Furthermore, for any $A \subset [r]$, by δ_A^C we denote the diagonal section of the marginal copula, corresponding to the A-components, i.e., the function

$$\delta_A^C : \; [0, 1] \to [0, 1]; \quad u \mapsto \delta_A^C(u) := C(u\mathbf{e}_A + \mathbf{e}_{A^c}).$$

In particular we refer to the functions

$$\delta^C_{[\ell]}(u) = C(\overbrace{u,\dots,u}^{\ell\ times},\overbrace{1,\dots,1}^{r-\ell\ times}), \quad 2 \le \ell \le r$$

as the **diagonal sections associated to** C.

For the functions $\delta^C_{[\ell]}$ we will also use the shorter notation δ^C_ℓ, namely

$$\delta^C_\ell(u) := \delta^C_{[\ell]}(u), \quad 2 \le \ell \le r,$$

and the shorter term ***diagonal sections*** of C. Such terminology turns out to be convenient for the ensuing arguments. It is clear that, for $A \subsetneq [r]$, with $|A| = \ell$ and $A \neq [\ell]$, the two functions δ^C_A and δ^C_ℓ are generally different.
It is also clear that $\delta^C_\ell(u)$ is an increasing function and that $\delta^C_2(u) \ge \delta^C_3(u) \ge \dots \ge \delta^C_r(u)$. Conditions, for a function $\delta : [0,1] \to [0,1]$ to be the diagonal section of a copula, are given, in particular, in Jaworski [12], and Durante and Sempi [9].

In what follows, when dealing with the diagonal sections associated to the survival copula K of $T_1,\dots,T_r$ we drop the superscript, i.e., we set

$$\delta_\ell(u) := \delta^K_\ell(u), \quad 2 \le \ell \le r.$$

Assume for the moment that the joint survival function $\overline{F}(t_1,\dots,t_r)$ is exchangeable, namely

$$\overline{F}(t_1,\dots,t_r) = K\left(\overline{G}(t_1),\dots,\overline{G}(t_r)\right), \quad \text{for } t_1,\dots,t_r > 0,$$

with K permutation-invariant.

As well-known, a direct relationship can be established between δ_r and the probability law of the minimal order statistics $T_{1:r}$, in fact one immediately obtains, for $t > 0$,

$$\overline{G}_{1:r}(t) = \mathbb{P}(T_1 > t,\dots,T_r > t) = \delta_r(\overline{G}(t)). \tag{2}$$

By taking into account exchangeability of $T_1,\dots,T_r$ one can similarly write

$$\begin{aligned}\mathbb{P}\left(T_{j_1} > t,\dots,T_{j_d} > t\right) &= \mathbb{P}(T_1 > t,\dots,T_d > t),\\ &= \overline{F}(t,\dots,t,0,\dots,0) = K\left(\overline{G}(t),\dots,\overline{G}(t),1,\dots,1\right) = \delta_d(\overline{G}(t)),\end{aligned} \tag{3}$$

for $d = 1,2,\dots,r-1$ and for any subset of indices $J = \{j_1,..,j_d\} \subset [r]$ of cardinality d. Whence one can write

$$\overline{G}_{\ell:r}(t) = \sum_{h=r-\ell+1}^{r} (-1)^{h-r-1+\ell}\binom{r}{h}\binom{h-1}{r-\ell}\delta_h(\overline{G}(t)), \quad \ell = 1,\dots,r. \tag{4}$$

In fact, by using (3), the latter formula is readily obtained from the formula expressing the survival functions of the order statistics of exchangeable variables in terms of the survival functions of the minima within subsets of the same variables (see in particular David and Nagaraja, p. 46 [5], Jaworski and Rychlik [13], Rychlik [22]).

As we will see in Proposition 7, formula (4) for $\overline{G}_{\ell:r}(t)$ is still valid when the joint distribution of $T_1,\dots,T_r$ satisfy the specific symmetry conditions recalled in Definitions 2 and 3, below. Such conditions are actually weaker than exchangeability, and turn out to be equivalent each other (see Proposition 5 below).

Definition 2. *We will say that the random variables $T_1,\dots,T_r$ are* t**-Exchangeable** *if for every $t \ge 0$, the binary random variables $X_i(t) = \mathbf{1}_{\{T_i>t\}}$, $i = 1,\dots,r$, are exchangeable, or equivalently the events $\{T_i > t\}$, $i = 1,\dots,r$, are exchangeable.*

We will briefly refer to the previous property as t-EX.

Definition 3. *The random variables $T_1, \ldots, T_r$ are said* **minimally stable**, *when, for any $\ell = 1, \ldots, r$ and for any subset $A = \{j_1, \ldots, j_\ell\} \subseteq [r]$*

$$\mathbb{P}\left(T_{j_1} > t, \ldots, T_{j_\ell} > t\right) = \mathbb{P}\left(T_1 > t, \ldots, T_\ell > t\right), \quad \forall\, t > 0, \tag{5}$$

namely $\mathbb{P}\left(T_{j_1} > t, \ldots, T_{j_\ell} > t\right) = \overline{F}(\underbrace{t, \ldots, t}_{\ell \text{ times}}, \underbrace{0, \ldots, 0}_{r-\ell \text{ times}}),\ \forall\, t > 0.$

Finally we recall the strictly related concept of diagonal dependent copulas (see Navarro and Fernandez-Sanchez [17]). Such a concept can be obtained as a special case of the one of k-diagonal dependence, for $k \le r$, as introduced by Okolewski in [21].

Definition 4. *Let C be an r-dimensional copula C. The copula C is said to be a* **k-diagonal dependent copula**, *with $k \le r$, if for any subsets $A, B \subset [r]$, with $|A| = |B| \le k$*

$$\delta_A^C(u) = \delta_B^C(u), \quad \forall\, u \in [0, 1]. \tag{6}$$

When $k = r$, the copula C is said **diagonal dependent**.

As in Navarro and Fernandez-Sanchez [17] we briefly refer to the property of diagonal dependence as DD.

The following result can be obtained by taking into account basic and well known properties of exchangeable binary random variables originally obtained by de Finetti (see [6]). See also Navarro et al. [18].

Proposition 5. *Under the conditions* **(H1)–(H3)** *the following properties are equivalent*
(i) *The random variables $T_1, \ldots, T_r$, are t-Exchangeable;*
(ii) *For all $H, H' \subseteq \{1, 2, \ldots, r\}$, with $|H| = |H'|$*

$$\mathbb{P}\big(T_j > t, \forall j \in H, T_i \le t, \forall i \notin H\big) = \mathbb{P}\big(T_j > t, \forall j \in H', T_i \le t, \forall i \notin H'\big). \tag{7}$$

(iii) *The random variables $T_1, \ldots, T_r$ are minimally stable;*
(iv) *The random variables $T_1, \ldots, T_r$ are identically distributed and their survival copula K is diagonal dependent.*

Proof. Properties ***(i)*** and ***(ii)*** are clearly equivalent: indeed

$$\mathbb{P}\big(T_j > t, \forall j \in H, T_i \le t, \forall i \notin H\big) = \mathbb{P}\big(X_j(t) = 1, \forall j \in H, X_i(t) = 0, \forall i \notin H\big).$$

Similarly properties ***(iii)*** and ***(iv)*** are equivalent: indeed if $T_1, \ldots, T_r$ are minimally stable, then by taking $\ell = 1$ in (5), they are identically distributed, and therefore for all $A \subset [r]$ with $|A| = \ell$

$$\mathbb{P}(T_i > t,\ \forall i \in A) = K(\overline{G}(t)\mathbf{e}_A + \mathbf{e}_{A^c}) = \delta_\ell(\overline{G}(t)), \quad \forall\, t \ge 0,$$

and $\overline{G}(t)$ is invertible, in view of the regularity condition **(H2)**.
Finally ***(iv)*** is equivalent to ***(ii)***, in view of the inclusion-exclusion formula.

□

Remark 6. *The previous result (Proposition 5) holds true also without the regularity assumption* **(H2)** *on $\overline{G}$, but in the general case an extension of the notion of Diagonal Dependence is needed (see Navarro et al. [18]).*

The interest for the properties ***(i)*** and ***(iv)*** had independently emerged in the two papers Marichal et al. [15] and Navarro and Fernandez-Sanchez [17] with reference to the field of systems' reliability. Still in the same framework, furthermore, the study of conditions for the equivalence between ***(i)*** and ***(iv)*** has been developed in Navarro et al. [18]. See Remark 14 below for details about the connections with system reliability.

We are now in a position to establish the following result.

Proposition 7. *Assume* **(H1)–(H3)** *and any condition among* ***(i)–(iv)***. *Then the equations* (4) *hold.*

Actually the validity of (4) hinges on the Eq. (3), which only requires the DD property of the survival copula and the identical distribution of T_i, $i = 1, \dots, r$.

From now on we make the following further assumption

(H4) the random variables $T_1, \dots, T_r$ are minimally stable.

Namely we assume the condition ***(iii)*** of Proposition 5 and, at a time, we highlight that such a result just ensures the validity of the equivalence among all the conditions ***(i)–(iv)***, under our standing hypotheses **(H1)–(H3)**.

Under the assumptions **(H1)–(H4)** we thus proceed to establish detailed results concerning the relations between the following families of functions

$$\mathcal{A} := \{\overline{G}; \delta_2, \dots, \delta_r\}, \quad \mathcal{B} := \{\overline{G}_{1:r}, \dots, \overline{G}_{r:r}\}. \tag{8}$$

Since the marginal survival functions $\overline{G}_{1:r}, \dots, \overline{G}_{r:r}$ are determined by the knowledge of the joint distribution of $T_1, \dots, T_r$ then, in principle, the family $\mathcal{B}$ should depend on the survival copula K and the common survival marginal $\overline{G}(t)$. Actually the full knowledge of K is not necessary, since the knowledge of the associated diagonal sections is sufficient as shown by the formula (4). More precisely the families $\mathcal{A}$ and $\mathcal{B}$ convey the same amount of information concerning the joint distribution of $T_1, \dots, T_r$, as we point out in details in what follows and summarize in Proposition 10.

To this end we start by recalling that when $\overline{G}_{1:r}, \dots, \overline{G}_{r:r}$ are known, we can easily recover the common marginal survival function $\overline{G}(t)$. Indeed, the random variables $T_1,\dots,T_r$ are identically distributed and therefore

$$\overline{G}(t) = \frac{1}{r}\sum_{k=1}^{r} \overline{G}_{k:r}(t), \tag{9}$$

as immediately follows by observing that $\sum_{h=1}^{r} \mathbf{1}_{\{T_h>t\}} = \sum_{k=1}^{r} \mathbf{1}_{\{T_{k:r}>t\}}$.

Furthermore the same formula (4) would permit to recover, step-by-step, the functions $\delta_2, \dots, \delta_r$. Here we follow a different path and the detailed formula is given in the Corollary 9 of the following result.

Proposition 8. *Under the conditions* **(H1)–(H4)**, *for every $d \in [r]$, and $J \subseteq [r]$, with $|J| = d$*

$$\begin{aligned}\mathbb{P}\big(T_j > t,\ \forall j \in J\big) &= \sum_{h=d}^{r} \frac{(h)_d}{(r)_d}\Big(\overline{G}_{r-h+1:r}(t) - \overline{G}_{r-h:r}(t)\Big) &(10)\\ &= \frac{d}{(r)_d}\sum_{k=1}^{r-d+1}(r-k)_{d-1}\,\overline{G}_{k:r}(t), \quad t > 0, &(11)\end{aligned}$$

where by convention $\overline{G}_{0:r}(t) = 0$, for $t > 0$.

Also for what concerns the proof of Proposition 8, similarly to what we mentioned for Proposition 5, one could apply well-known and simple results (see, e.g., de Finetti [6]) about exchangeable binary random variables. For the ease of the reader we give a self-contained, and detailed, proof at the end of this section. Here we only point out that the most important ingredient of the proof amounts to the validity of the following identity for any subset $J \subset [r]$:

$$\mathbb{P}\big(T_j > t,\ \forall j \in J\big) = \sum_{K:\, K \subseteq J^c} \mathbb{P}\big(T_j > t,\ \forall j \in J \cup K,\ T_i \le t,\ \forall i \notin J \cup K\big) \tag{12}$$

or equivalently

$$\mathbb{P}\big(T_j > t,\ \forall j \in J\big) = \sum_{H:\, H \supseteq J} \mathbb{P}\big(T_j > t,\ \forall j \in H,\ T_i \le t,\ \forall i \notin H\big). \tag{13}$$

From Proposition 8 we get the expression of δ_d in terms of the marginal survival functions $\overline{G}_{k:r}(t)$, $k = 1, \dots, r$.

Corollary 9. *Under the conditions* **(H1)–(H4)**, *for every* $d \in \{1, 2, \dots, r\}$, *the following equalities hold*

$$\delta_d(u) = \sum_{h=d}^{r} \frac{(h)_d}{(r)_d} \left(\overline{G}_{r-h+1:r}\big(\overline{G}^{-1}(u)\big) - \overline{G}_{r-h:r}\big(\overline{G}^{-1}(u)\big) \right) \tag{14}$$

$$= \frac{d}{(r)_d} \sum_{h=1}^{r-d+1} (r-h)_{d-1}\, \overline{G}_{h:r}\big(\overline{G}^{-1}(u)\big), \quad u \in [0, 1]. \tag{15}$$

Indeed, this result is a consequence of equations (10), (11), and (3), since, as already observed the condition appearing in (3) also holds for minimally stable variables.

The above results, concerning the relations tying the families $\mathcal{A}$ and $\mathcal{B}$, will be now summarized by means of the following proposition.

Proposition 10. *Under the conditions* **(H1)–(H4)**, *the family of the survival functions* $\mathcal{B} = \{\overline{G}_{1:r}, \dots, \overline{G}_{r:r}\}$ *is determined by the family* $\mathcal{A} = \{\overline{G}, \delta_2, \dots, \delta_r\}$ *by means of formula* (4). *Viceversa the family* $\mathcal{A}$ *is determined by the family* $\mathcal{B}$ *by means of formula* (14) *(or* (15)*) and formula* (9).

Remark 11. *For non-exchangeable, but minimally stable lifetimes* $T_1, \dots, T_r$ *there still exist exchangeable lifetimes* $\widetilde{T}_1, \dots, \widetilde{T}_r$, *such that* $\mathbb{P}(\widetilde{T}_j > t) = \mathbb{P}(T_j > t) = \overline{G}(t)$, *and the diagonal sections* $\widetilde{\delta}_\ell(u)$ *associated to their survival copula* $\widetilde{K}$ *coincide with the diagonal sections* $\delta_\ell(u)$ *associated to the survival copula* K. *Indeed* $\widetilde{K}$ *may be constructed by symmetrizing* K*:*

$$\widetilde{K}(u_1, \dots, u_r) = \frac{1}{r!} \sum_{\sigma \in \Pi([r])} K(u_{\sigma_1}, \dots, u_{\sigma_r}).$$

The above construction can be of help in obtaining the explicit form of $\mathbb{P}(T_j > t, \forall j \in A)$ *in some special cases (see in particular Subsection 5.2).*

Remark 12. *The problem of constructing examples of vectors which are not exchangeable, but still minimally stable, naturally arises. Since minimally stable variables* $T_1, \dots, T_r$ *are identically distributed, constructing such examples is equivalent to constructing diagonal-dependent copulas, which are not exchangeable. In the following Example 13 we present a simple path to such a construction. Other examples can be found in Navarro, Fernandez-Sanchez [17], and Navarro et al. [18]. See also Example 31 in the Appendix.*

Example 13. *First of all we notice that, when* $r = 2$, *then any pair* (T_1, T_2) *of identical distributed random variables is minimally stable, but in general is not exchangeable. Similarly, and trivially, any* 2*-dimensional copula* C *is minimally stable. Indeed* $\delta_2^C(u) = C(u, u)$, *and* $\delta_1^C(u) = C(u, 1) = C(1, u) = u$. *Starting from the copula* C *one may define two* 3*-dimensional copulas as follows*

$$C_{(1,2,3)}(u_1, u_2, u_3) := \frac{1}{3}\left[C(u_1, u_2)u_3 + C(u_2, u_3)u_1 + C(u_3, u_1)u_2\right]$$

$$C_{(3,2,1)}(u_1, u_2, u_3) := \frac{1}{3}\left[C(u_3, u_2)u_1 + C(u_2, u_1)u_3 + C(u_1, u_3)u_2\right]$$

respectively obtained as the symmetric mixture over the cyclic permutations of $(1, 2, 3)$ *and the cyclic permutations of* $(3, 2, 1)$. *Notice that when* C *is non-exchangeable, then* $C_{(1,2,3)}$ *and* $C_{(3,2,1)}$ *are non-exchangeable: indeed if* $u, v \in (0, 1)$ *are such that* $C(u, v) \neq C(v, u)$ *then, for example,*

$$C_{(1,2,3)}(u, v, 1) := \frac{1}{3}\left[C(u, v) + C(v, 1)u + C(1, u)v\right] = \frac{1}{3}\, C(u, v) + \frac{2}{3}\, uv,$$

which is clearly different from

$$C_{(1,2,3)}(v, u, 1) := \frac{1}{3}\left[C(v, u) + C(u, 1)v + C(1, v)u\right] = \frac{1}{3}\, C(v, u) + \frac{2}{3}\, uv,$$

thus $C_{(1,2,3)}$ is non-exchangeable, though the 2-dimensional marginal distributions are all equal, namely

$$C_{(1,2,3)}(u, v, 1) = C_{(1,2,3)}(v, 1, u) = C_{(1,2,3)}(1, u, v) = \frac{1}{3}\, C(v, u) + \frac{2}{3}\, uv.$$

By iterating the above construction, one can also obtain a DD, non-exchangeable, copula which is n-dimensional. Details can be found in the Appendix (see Example 30).

Remark 14. *As already mentioned, the topics developed in the papers Marichal et al. [15], Navarro and Fernandez-Sanchez [17], and Navarro et al. [18] are motivated by questions arising in the field of systems' reliability. More precisely these papers deal with the so-called signature representation for the survival function $R_S^{(\varphi)}(t)$ of the lifetime T_S of a binary coherent system S, made with r binary components, with structure function φ, and for which the random variables $T_1, \ldots, T_r$ have the meaning of the components' lifetimes. The signature of S is a probability distribution $\mathbf{s}^{(\varphi)} := (s_1^{(\varphi)}, \ldots, s_r^{(\varphi)})$ over $[r]$ which is a combinatorial invariant associated to φ (see in particular Samaniego [24]). The signature representation means that the equation*

$$R_S^{(\varphi)}(t) = \sum_{h=1}^{r} s_h^{(\varphi)} \overline{G}_{h:r}(t) \tag{16}$$

holds for any $t > 0$.
Under our standing hypotheses **(H1)**—**(H4)***, the properties* ***(i)*** *and* ***(iv)*** *are equivalent and also imply the signature representation* (16) *for the survival function $R_S^{(\varphi)}(t)$. When the functions $\overline{G}_{1:r}, \ldots, \overline{G}_{r:r}$ are known, one can then recover from* (16) *the function $R_S^{(\varphi)}(t)$, relatively to any structure φ for which the signature $\mathbf{s}^{(\varphi)}$ is known. At the same time the family of all the functions $R_S^{(\varphi)}(t)$ in particular contains the family $\mathcal{B} = \{\overline{G}_{1:r}, \ldots, \overline{G}_{r:r}\}$. In fact, the survival functions $\overline{G}_{k:r}$ can be seen as the reliability functions of the coherent systems of the type k-out-of-r, for $k = 1, \ldots, r$.*

We end this section with the afore announced proof of Proposition 8.

Proof of Proposition 8. We start by recalling that $N(t) = \sum_{i=1}^{r} \mathbf{1}_{\{T_i \le t\}}$ and observing that

$$\mathbb{P}(N(t) = r - h) = \sum_{J:|J|=h} \mathbb{P}(T_j > t,\ \forall j \in J,\ \ T_i \le t,\ \forall\, i \notin J)$$

so that, thanks to Eq. (7)

$$\mathbb{P}(N(t) = r - h) = \binom{r}{h} \mathbb{P}(T_j > t,\ \forall j \in \{1, 2, .., h\},\ \ T_i \le t,\ \ \forall i \in \{h + 1, \ldots, r\}),$$

or equivalently, for any $H \subset [r]$, with $|H| = h$

$$\mathbb{P}\big(T_j > t,\ \forall j \in H,\ \ T_i \le t,\ \forall i \notin H\big) = \frac{1}{\binom{r}{h}}\, \mathbb{P}(N(t) = r - h). \tag{17}$$

On the other hand, we observe that

$$\begin{aligned} &\mathbb{P}(N(t) = r - h) = \mathbb{P}(T_{r-h:r} \le t < T_{r-h+1:r}) \\ =&\mathbb{P}(T_{r-h+1:r} > t) - \mathbb{P}(T_{r-h:r} > t) = \overline{G}_{r-h+1:r}(t) - \overline{G}_{r-h:r}(t). \end{aligned} \tag{18}$$

Then the thesis follows immediately: indeed, for every $J \subset \{1, 2, \ldots, r\}$ with $|J| = d$, Eq.s (17) and (18), together with (12), imply (with the convention that $\binom{0}{0} = 1$)

$$\mathbb{P}\big(T_j > t,\ \forall j \in J\big) = \sum_{h=d}^{r} \binom{r-d}{h-d} \frac{1}{\binom{r}{h}} \left(\overline{G}_{r-h+1:r}(t) - \overline{G}_{r-h:r}(t)\right),$$

and formula (10) follows by observing that

$$\frac{\binom{r-d}{h-d}}{\binom{r}{h}} = \frac{(h)_d}{(r)_d}.$$

Finally, from (10), taking into account the convention that $\overline{G}_{0:r}(t) = 0$, one obtains

$$\begin{aligned}(r)_d\,\mathbb{P}\big(T_j > t,\ \forall j \in J\big) &= \sum_{h=d}^{r}(h)_d\overline{G}_{r-(h-1):r}(t) - \sum_{h=d}^{r-1}(h)_d\overline{G}_{r-h:r}(t)\\ &= \sum_{k=d-1}^{r-1}(k+1)_d\overline{G}_{r-k:r}(t) - \sum_{h=d}^{r-1}(h)_d\overline{G}_{r-h:r}(t)\\ &= d!\,\overline{G}_{r-(d-1):r}(t) + \sum_{k=d}^{r-1}\big[(k+1)_d - (k)_d\big]\,\overline{G}_{r-k:r}(t).\end{aligned}$$

Therefore, by observing that

$$(k+1)_d - (k)_d = (k)_{d-1}\,[k+1-(k-(d-1))] = (k)_{d-1}\,d,$$

one gets

$$\mathbb{P}\big(T_j > t,\ \forall j \in J\big) = \frac{d}{(r)_d}\sum_{k=d-1}^{r-1}(k)_{d-1}\overline{G}_{r-k:r}(t),$$

Then formula (11) follows by setting $h = r - k$ in the last sum. □

3 The use of multivariate conditional hazard rates

In this section attention will be restricted to random vectors of lifetimes with absolutely continuous joint distributions, so that their probabilistic properties can be alternatively described in terms of the multivariate conditional hazard rate (m.c.h.r.) functions. In a first part of this section we recall the definition of the m.c.h.r. functions associated to generic random variables $\{V_j, j \in [n]\}$, reviewing related properties and providing some references. Furthermore, for $A \subset [n]$, we focus attention on the joint distributions of $\{V_j,\ j \in A\}$. In particular we analyze the probability distributions of their minima by means of the m.c.h.r. functions associated to $\{V_j,\ j \in A\}$. In the second part, coming back to our lifetimes $T_1, \dots, T_r$, and adding absolute continuity condition to our standing hypotheses, we characterize both the exchangeability and the minimal stability conditions by means of the m.c.h.r. functions associated to $\{T_j, j \in [r]\}$.

3.1 Multivariate conditional hazard rates and distribution of minima

In this subsection we briefly recall some definitions and basic properties of multivariate conditional hazard rate functions for n non-negative random variables $V_1, \dots, V_n$ with an absolutely continuous joint distribution whose joint density function is denoted by $f_{\mathbf{V}}$. For simplicity's sake we will assume moreover that there exists a version of the joint density which is strictly positive on $\mathbb{R}^n_+$, i.e., $f_{\mathbf{V}}(v_1, .., v_n) > 0$, when $v_i > 0$, for all $i = 1, 2, \dots, n$.

For $k = 1, \dots, n-1$, and for any k-permutation $\mathbf{j} = (j_1, \dots, j_k) \in \Pi_k([n])$, the symbol $\mathbf{V_j}$ denotes the vector of lifetimes $(V_{j_1}, \cdots, V_{j_k})$; for any subset $J \subseteq [n]$ we denote

$$V_{1:J} := \min_{j \in J} V_j; \tag{19}$$

furthermore, if $\mathbf{j} \in \Pi(J)$, for $0 < v_1 < \cdots < v_k \le v$ the symbol

$$\mathbf{V_j} = \mathbf{v}; \quad V_{1:J^c} > v \tag{20}$$

briefly denotes the observation

$$V_{j_1} = v_1, \cdots, V_{j_k} = v_k, \quad \min_{j \in J^c} V_j > v.$$

The observation $\mathbf{V_j} = \mathbf{v};\ V_{1:J^c} > v$ in (20) is often called a “dynamic history".

For the given $k \ge 1$, $J \subset [n]$, with $|J| = k$, $\mathbf{j} = (j_1, \ldots, j_k) \in \Pi(J)$, $v > 0$, $0 < v_1 < \cdots < v_k \le v$, $j \notin J$, the multivariate conditional hazard rates (m.c.h.r.) function $v \mapsto \lambda_{j|\mathbf{j}}(v|v_1, \cdots, v_k)$ is defined by the limit (where it exists)

$$\lambda_{j|j_1,\ldots,j_k}(v|v_1, \cdots, v_k) := \lim_{\Delta \to 0^+} \frac{\mathbb{P}\big(V_j \le v + \Delta | \mathbf{V_j} = \mathbf{v}; V_{1:J^c} > v\big)}{\Delta}, \quad a.e. \tag{21}$$

Note that the above m.c.h.r. functions $\lambda_{j|j_1,\ldots,j_k}$ depend also on the version of the conditional probability.

Furthermore, for any $j \in [n]$, the m.c.h.r. function $\lambda_{j|\emptyset}(v) : [0, \infty) \to [0, \infty)$ is defined by the limit (where it exists)

$$\lambda_{j|\emptyset}(v) := \lim_{\Delta \to 0^+} \frac{\mathbb{P}\big(V_j \le v + \Delta | V_{1:n} > v\big)}{\Delta}, \quad a.e. \tag{22}$$

In the sequel we will use the convention

$$\lambda_{j|j_1,\ldots,j_k}(v|v_1, \cdots, v_k) = \lambda_{j|\emptyset}(v), \quad \text{when } k = 0. \tag{23}$$

The above limits make sense in view of the assumption of absolute continuity of the joint distribution of $V_1, \ldots, V_n$ and the m.c.h.r. functions can be seen as direct extensions of the common concept of hazard rate function of a univariate non-negative random variable.

For the random vector $\mathbf{V} \equiv (V_1, \ldots, V_n)$, the system of the m.c.h.r. functions in (21) and (22) can be computed in terms of the joint density $f_\mathbf{V}$. It is remarkable the circumstance that the function f_V can be obtained from the knowledge of the set of all the m.c.h.r. functions in terms of a formula, that we are going to recall next. Preliminarily, we first notice an obvious difference between f_V and the functions $\lambda_{j|j_1,\ldots j_k}(v|v_1, \ldots, v_k)$: while the arguments $v_1, \ldots, v_n$ of f_V are generally not ordered, the arguments $v_1, \ldots, v_k$ of the functions $\lambda_{j|j_1,\ldots j_k}(v|v_1, \ldots, v_k)$ are necessarily listed in increasing order by definition. Furthermore, for given non-ordered values $v_1, \ldots, v_n$, we denote by $v_{1:n}, \ldots, v_{n:n}$ the same values rearranged in increasing order. Then the following formula holds: for $(v_1, \ldots, v_n)$, let $\mathbf{j} = (j_1, \ldots, j_n)$ a permutation in $\Pi([n])$ such that $v_{1:n} = v_{j_1} \le v_{2:n} = v_{j_2} \le \cdots \le v_{n:n} = v_{j_n}$,

$$\begin{aligned} f_\mathbf{V}(v_1, \ldots, v_n) = \Pi_{k=1}^n \lambda_{j_k|j_1,\ldots,j_{k-1}}(v_{j_k}|v_{j_1}, \ldots, v_{j_{k-1}}) \cdot \\ \cdot\, e^{-\int_{v_{j_{k-1}}}^{v_{j_k}} \Lambda_{j_1,\ldots,j_{k-1}}(u|v_{j_1},\ldots,v_{j_{k-1}})\, du}, \end{aligned} \tag{24}$$

where we have set $v_{j_0} = 0$,

$$\Lambda_{j_1,\ldots.,j_{k-1}}(u|v_{j_1}, \ldots, v_{j_{k-1}}) = \sum_{j \notin \{j_1,\ldots,j_{k-1}\}} \lambda_{j|j_1,\ldots,j_{k-1}}(u|v_{j_1}, \ldots, v_{j_{k-1}}), \tag{25}$$

and we have used the convention that, when $k = 1$,

$$\begin{aligned} \lambda_{j|j_1,\ldots,j_{k-1}}(u|v_{j_1}, \ldots, v_{j_{k-1}}) = \lambda_{j|\emptyset}(u), \\ \Lambda_{j_1,\ldots.,j_{k-1}}(u|v_{j_1}, \ldots, v_{j_{k-1}})\Big|_{k=1} = \Lambda_\emptyset(u) = \sum_{j \in [n]} \lambda_{j|\emptyset}(u). \end{aligned} \tag{26}$$

For proofs, details, and for general aspects of the m.c.h.r. functions see Shaked, Shantikumar [26, 27]. See also the reviews contained within the more recent papers Shaked, Shantikumar [28], Spizzichino [30].

For any subset of indices $A = \{h_1, \ldots, h_{|A|}\} \subset [n]$, one can also consider joint density of the random vector $(V_{h_1}, \ldots V_{h_{|A|}})$. Such a density may be defined by means of a different family of m.c.h.r. functions, related to the choice of the set A. Namely for any $k \le |A|$ and for any k-permutation $(j_1, \dots, j_k) \in \Pi_k(A)$, $j \in A \setminus \{j_1, \dots, j_k\}$, $0 < v_1 < \dots < v_k \le v$, we can consider the m.c.h.r. functions defined as follows

$$\lambda^A_{j|j_1,\dots,j_k}(v|v_1, \cdots, v_k) := \lim_{\Delta \to 0^+} \frac{\mathbb{P}\big(V_j \le v + \Delta | \mathbf{V_j} = \mathbf{v}; V_{1:A\setminus\{j_1,\dots,j_k\}} > v\big)}{\Delta}, \quad a.e. \tag{27}$$

and, for $j \in A$, $v > 0$

$$\lambda^A_{j|\emptyset}(v) := \lim_{\Delta \to 0^+} \frac{\mathbb{P}\big(V_j \le v + \Delta | V_{1:A} > v\big)}{\Delta}, \quad a.e. \tag{28}$$

In view of the characterization of minimal stability, we are interested in the distributions of minima over different subsets of $[r]$. It is therefore relevant to highlight that, in particular, the functions $\lambda_{j|\emptyset}(v)$, for $j \in [n]$, are strictly related to the marginal law of the minimal order statistic $V_{1:n} \equiv V_{1:[n]} = \min_{j=1,\dots,n} V_j$. In this respect the following identity holds:

$$\mathbb{P}(V_{1:n} > v) = \exp\Big\{ - \int_0^v \sum_{j=1}^n \lambda_{j|\emptyset}(s) ds \Big\} = \exp\Big\{ - \int_0^v \Lambda_\emptyset(s) ds \Big\}. \tag{29}$$

(See, e.g., De Santis et al. [7], where a more detailed description of the probabilistic behavior of $V_{i:n}$ in terms of $\lambda_{j|\emptyset}(v)$, for $j \in [n]$, is pointed out). Similarly, we can also consider the survival function of $V_{1:A}$, the minimal order statistic among the variables V_j, with $j \in A$, for $A \subset [r]$. With the notation introduced so far, one can write

$$\mathbb{P}(V_{1:A} > v) = \exp\Big\{ - \int_0^v \sum_{j \in A} \lambda^A_{j|\emptyset}(s) ds \Big\} = \exp\left\{ - \int_0^v \Lambda^A_\emptyset(s)\, ds \right\}, \tag{30}$$

where

$$\Lambda^A_\emptyset(t) := \sum_{j \in A} \lambda^A_{j|\emptyset}(t). \tag{31}$$

Notice that therefore the functions $\Lambda_\emptyset(t)$ and $\Lambda^A_\emptyset(t)$ can be respectively interpreted as the "usual" hazard rate functions for the random variables $V_{1:n}$ and $V_{1:A}$. This observation will be a key point for the discussion in Section 4.

3.2 The m.c.h.r. functions and characterizations of Exchangeability and of Minimal Stability

Here we come back to our lifetimes $T_1, \dots, T_r$. We maintain the condition **(H1)**, whereas the conditions **(H2)** and **(H3)** are replaced by the following stronger condition:

(H5) The joint distribution of $T_1, \dots, T_r$ is absolutely continuous, with the joint density such that

$$f_{\mathbf{T}}(t_1, .., t_r) > 0, \quad a.e. \text{ in } \mathbb{R}^r_+$$

We start with the exchangeable case, noticing that such a case leads to a remarkable simplification of notation, technical results, and conceptual aspects concerning the m.c.h.r. functions.

First notice that the symmetry conditions among the different random variables, as requested by exchangeability, imply a specially simple form for the m.c.h.r. functions. More precisely, the functions $\lambda_{j|i_1,\dots,i_k}(t|t_{i_1}, \dots, t_{i_k})$ cannot depend on the index $j \notin \{i_1, \dots, i_k\}$. Furthermore all the k-permutations $(i_1, \dots, i_k)$ are to be considered as similar one another and thus the dependence of a m.c.h.r. function w.r.t. to $(i_1, \dots, i_k)$ is encoded in the number k. For the present exchangeable case, for any $t > 0$, $0 < t_1 < \dots < t_k \le t$, we then introduce the symbols $\mu(t|k; t_1, \dots, t_k)$ and $\mu(t|0)$ with the following meaning: for any $\mathbf{j} = (j_1, \dots, j_k) \in \Pi(I)$

$$\lambda_{j|j_1,\dots,j_k}(t|t_1, \dots, t_k) = \mu(t|k; t_1, \dots, t_k), \qquad \lambda_{j|\emptyset}(t) = \mu(t|0). \tag{32}$$

Thus, for any $(t_1, \ldots, t_r) \in \mathbb{R}_+^r$, and denoting by $t_{1:r}, \ldots, t_{r:r}$ the values $t_1, \ldots, t_r$ rearranged in an increasing order, by Eq. (24), $f_{\mathbf{T}}$ takes the form

$$f_{\mathbf{T}}(t_1, \ldots, t_r) = \Pi_{k=0}^{r-1} \mu(t_{k+1:r}|k; t_{1:r}, \ldots, t_{k:r}) e^{-(r-k)\int_{t_{k:r}}^{t_{k+1:r}} \mu(s|k; t_{1:r}, \ldots, t_{k:r})\, ds}, \tag{33}$$

where we used the further convention that $t_{0:k} = 0$.
On the other hand, when the form (32) is assumed for the family of the m.c.h.r. functions, the consequent formula (33) shows that $f_{\mathbf{T}}$ actually depends on the arguments $t_1, \ldots, t_r$ only through the ordered values $t_{1:r}, \ldots, t_{r:r}$ and thus it is necessarily exchangeable. In conclusion, the following characterization of exchangeability holds (see also Spizzichino [29], chap.2).

Proposition 15. *Non-negative random variables $T_1, \ldots, T_r$ with a strictly positive joint density are exchangeable if and only if the corresponding m.c.h.r. functions are of the form* (32).

As far as minimal stability is concerned, one can find natural conditions involving the hazard rates functions $\Lambda_\emptyset^A(t)$ of the minima $T_{1:A}$, for $A \subset [r]$. See in particular Lemma 17 in the next Section 4, where this topic is dealt with in some details. Here we point out that the functions $\Lambda_\emptyset^A(t)$ can be recovered once the m.c.h.r. functions $\lambda_{j|\emptyset}^A(t)$, associated to the random variables $\{T_j,\, j \in A\}$, are known (see (31)). However in general, when the distribution of $T_1, \ldots, T_r$ is specified in terms of the associated m.c.h.r. functions $\lambda_{j|j_1,\ldots,j_k}(t|t_1, \ldots, t_k)$, it is not easy to recover the m.c.h.r. functions $\lambda_{j|\emptyset}^A(t)$ associated to $\{T_j,\, j \in A\}$. Therefore it is useful to find conditions for minimal stability expressed directly in terms of the m.c.h.r. functions $\lambda_{j|j_1,\ldots,j_k}(t|t_1, \ldots, t_k)$. On this purpose, taking into account the equivalence between condition (7) and minimal stability (i.e., between conditions ***(ii)*** and ***(iii)*** of Proposition 5), it is relevant to express

$$\mathbb{P}\big(T_j > t, \forall j \in A, T_i \le t, \forall i \in [r] \setminus A\big)$$

in terms of the m.c.h.r. functions. To this end for any d-permutation $\mathbf{j} = (j_1, \ldots, j_d)$, we set

$$\begin{aligned}\Psi(t;[r],\mathbf{j}) :=& \mathbb{P}\big(T_{j_1} < T_{j_2} < \cdots < T_{j_d} \le t, T_i > t\, \forall i \notin \{j_1, \ldots, j_d\}\big)\\ =& \int_0^t ds_d \int_0^{s_d} ds_{d-1} \cdots \int_0^{s_2} ds_1\, e^{-\int_{s_d}^t \Lambda_{j_1,\ldots,j_d}(\tau|s_1,\ldots,s_d)d\tau} \cdot \\ & \cdot \prod_{\ell=1}^d \lambda_{j_\ell|j_1,\ldots,j_{\ell-1}}(s_\ell|s_1, \ldots, s_{\ell-1}) e^{-\int_{s_{\ell-1}}^{s_\ell} \Lambda_{j_1,\ldots,j_{\ell-1}}(\tau|s_1,\ldots,s_{\ell-1})d\tau}.\end{aligned} \tag{34}$$

Then, for any subset $A \subset [r]$, we can write

$$\mathbb{P}\big(T_j > t, \forall j \in A, T_i \le t, \forall i \in [r] \setminus A\big) = \sum_{\mathbf{j} \in \Pi([r]\setminus A)} \Psi(t;[r],\mathbf{j}). \tag{35}$$

The above Eq. (35), together with Eq. (34), and Proposition 5 can be used to get the following characterization of minimal stability.

Proposition 16. *Non-negative random variables $T_1, \ldots, T_r$ with a strictly positive joint density are minimally stable if and only if the corresponding m.c.h.r. satisfy the condition that whenever $A, B \subset [r]$, with $|A| = |B| \le r-1$, then*

$$\sum_{\mathbf{j} \in \Pi(A)} \Psi(t;[r],\mathbf{j}) = \sum_{\mathbf{j}' \in \Pi(B)} \Psi(t;[r],\mathbf{j}'), \quad t > 0, \tag{36}$$

where we have used the notation (34).

Observe that the exchangeability condition implies the identity $\Psi(t;[r],\mathbf{j}) = \Psi(t;[r],\mathbf{j}')$ for any pair of d-permutations $\mathbf{j}, \mathbf{j}'$, so that condition (36) is trivially satisfied.
The characterization of Proposition 16, besides its conceptual meaning, reveals to be effective in some special cases. In particular we will use it when dealing with a subclass of Load Sharing models (see Subsection 5.2).

4 Relations among diagonal sections of DD copulas, distributions of order statistics, and hazard rates of minima

Concerning with the joint distribution of minimally stable lifetimes $T_1, \dots, T_r$, it has been pointed out in Proposition 10 that the two systems of functions $\mathcal{A} = \{\overline{G}; \delta_2, \dots, \delta_r\}$ and $\mathcal{B} = \{\overline{G}_{1:r}, \dots, \overline{G}_{r:r}\}$ convey the same information about the joint distribution of $T_1,\dots,T_r$ and that they can be then recovered one from the other. Like in Subsection 3.2 we assume **(H1)** and **(H5)**, where the joint distribution of $T_1,\dots,T_r$ can be described in terms of the corresponding m.c.h.r. functions. In terms of those functions, we aim to single out characteristics of the joint distribution, whose knowledge may be equivalent, under the condition **(H4)**, to that of the systems of functions $\mathcal{A}$ and $\mathcal{B}$ defined in (8). It will emerge that the information contained in the systems $\mathcal{A}$ and $\mathcal{B}$ is equivalent to the knowledge embedded in the systems of functions defined by

$$\mathcal{C} =: \{\Lambda_\emptyset^{[1]}, \dots, \Lambda_\emptyset^{[r]}\}. \tag{37}$$

Such equivalence is demonstrated by the relations tying $\mathcal{C}$ with $\mathcal{A}$ and $\mathcal{B}$. Such relations will be detailed below by means of the following Propositions 18 and 19. More precisely, in Proposition 18 we express each of the families $\mathcal{A}$ and $\mathcal{B}$ in terms of $\mathcal{C}$, whereas in Proposition 19 the family $\mathcal{C}$ is expressed in terms of $\mathcal{A}$ and in terms of $\mathcal{B}$.

We start by giving a characterization of minimal stability in terms of the hazard rates of minima. As already observed (see Eq.s (28) and (30)) the m.c.h.r. functions $\lambda_{j|\emptyset}^A(t)$, $j \in A$, are related to the law of the minimum on an arbitrary set $A \subset [r]$, indeed

$$\mathbb{P}(T_{1:A} > t) = \exp\left\{-\int_0^t \Lambda_\emptyset^A(s)\,ds\right\}, \quad t > 0, \tag{38}$$

where $\Lambda_\emptyset^A(t) = \sum_{j\in A} \lambda_{j|\emptyset}^A(t)$ is the one-dimensional failure rate of $T_{1:A}$. Concerning with this notation, observe that the failure rate $\Lambda_\emptyset^{[r]}(t)$ coincides with $\Lambda_\emptyset(t) = \sum_{j=1}^r \lambda_{j|\emptyset}(t)$, so that $\overline{G}_{1:r}(t) = \mathbb{P}(T_{1:r} > t) = \exp\left\{-\int_0^t \Lambda_\emptyset(s)\,ds\right\}$, (see also Eq. (29)). Equation (38) leads immediately to the following simple characterization of minimal stability.

Lemma 17. *Assume condition* **(H5)** *and assume that the hazard rates $\Lambda_\emptyset^A(t)$ of the minima $T_{1:A}$ are known, for every non-empty subset $A \subset [r]$. Then each of the following conditions is necessary and sufficient for the minimal stability condition* **(H4)***:*

$$\forall A \subset [r], \quad \Lambda_\emptyset^A(t) = \Lambda_\emptyset^{[d]}(t) \quad a.e.,\ t > 0 \quad \text{where } d = |A|, \tag{39}$$

and

$$\forall A \subset [r], \quad \mathbb{P}(T_{1:A} > t) = \exp\left\{-\int_0^t \Lambda_\emptyset^{[d]}(s)\,ds\right\}, \quad t > 0, \quad \text{where } d = |A|. \tag{40}$$

Proof. Conditions (39) and (40) are clearly equivalent to each other, and if (40) holds then **(H4)** is immediate. Viceversa, by the assumption of minimal stability we may restrict attention on the subset of variables $T_1, \dots, T_d$, which are minimally stable, as well. Indeed, since $\mathbb{P}(T_{1:A} > t) = \mathbb{P}(T_{1:B} > t)$ for any $A, B \subseteq [r]$ such that $|A| = |B|$, then (40) follows, and therefore also (39) holds. □

Clearly when the m.c.h.r. functions $\lambda_{j|\emptyset}^A(t)$ are known, for every non-empty subset $A \subset [r]$, then (39) is equivalent to

$$\forall A \subset [r], \quad \sum_{j=1}^d \lambda_{j|\emptyset}^A(t) = \sum_{j=1}^d \lambda_{j|\emptyset}^{[d]}(t), \quad a.e.,\ t > 0, \quad \text{where } d = |A|, \tag{41}$$

Proposition 18. *Assume the minimal stability condition* **(H4)***, and the joint absolute continuity condition* **(H5)***. Assume furthermore that the family $\mathcal{C}$ of the hazard rate functions $\Lambda_\emptyset^{[d]}(t)$ is known. Then*

(i) *the family $\mathcal{A}$ is given by:*

$$\overline{G}(t) = \exp\left\{-\int_0^t \Lambda_\emptyset^{[1]}(s)\,ds\right\}, \quad \textit{for any } t > 0, \tag{42}$$

and for $d = 2, \ldots, r$ and for any $u \in [0,1]$,

$$\delta_d(u) = \exp\left\{-\int_0^{\overline{G}^{-1}(u)} \Lambda_\emptyset^{[d]}(s)\,ds\right\}; \tag{43}$$

(ii) *the family $\mathcal{B}$ is given by:*
for $\ell = 1, \ldots, r$ and for any $t > 0$,

$$\overline{G}_{\ell:r}(t) = \sum_{h=r-\ell+1}^{r} (-1)^{h-r-1+\ell} \binom{r}{h}\binom{h-1}{r-\ell} \exp\left\{-\int_0^t \Lambda_\emptyset^{[h]}(s)\,ds\right\}. \tag{44}$$

*Proof of **(i)**.* Due to minimal stability the random variables T_i share the marginal survival function with T_1, i.e.,

$$\exp\left\{-\int_0^t \lambda_{i|\emptyset}(s)\,ds\right\} = \exp\left\{-\int_0^t \lambda_{1|\emptyset}(s)\,ds\right\}, \quad t > 0.$$

Therefore $\lambda_{i|\emptyset}(t) = \lambda_{1|\emptyset}(t) \equiv \Lambda_\emptyset^{[1]}(t)$, for any $t > 0$, and (42) follows.

As already observed, we may concentrate attention on the random variables $T_1, \ldots, T_d$. Therefore to prove (43), on the one hand one has

$$\mathbb{P}(\min_{i=1,\ldots,d} T_i > t) = \exp\left\{-\int_0^t \Lambda_\emptyset^{[d]}(s)\,ds\right\}.$$

On the other hand, taking into account (3) we can also write

$$\mathbb{P}(\min_{i=1,\ldots,d} T_i > t) = \delta_d(\overline{G}(t)).$$

Thus Eq. (43) is immediately achieved by comparing the preceding two formulas and recalling that condition **(H5)** implies condition **(H2)**, which in its turn implies that $\overline{G}$ is invertible.
*Proof of **(ii)**.* Taking into account Proposition 7, Eq. (44) is immediately achieved by combining Eq. (43) with Eq. (4). □

Proposition 19. *Assume the minimal stability condition* **(H4)**, *and the joint absolute continuity condition* **(H5)**.
(i) *If the family $\mathcal{A} = \{\overline{G}, \delta_2, \ldots, \delta_r\}$ is known, then, setting $\delta_1(u) = u$,*

$$\Lambda_\emptyset^{[\ell]}(t) = -\frac{d}{dt}\log\left[\delta_\ell(\overline{G}(t))\right], \quad a.e.\ t > 0, \quad \ell = 1, 2, \ldots, r. \tag{45}$$

(ii) *If the family $\mathcal{B} = \{\overline{G}_{1:r}, \ldots, \overline{G}_{r:r}\}$ is known, then, respectively denoting by $g_{1:r}, \ldots, g_{r:r}$ the probability density of the order statistics $T_{1:r}, \ldots, T_{r:r}$,*

$$\begin{aligned}\Lambda_\emptyset^{[\ell]}(t) &= \frac{\sum_{h=\ell}^r (h)_\ell \left(g_{r-h+1:r}(t) - g_{r-h:r}(t)\right)}{\sum_{h=\ell}^r (h)_\ell \left(\overline{G}_{r-h+1:r}(t) - \overline{G}_{r-h:r}(t)\right)} \\ &= \frac{\sum_{k=1}^{r-(\ell-1)} (r-k)_{\ell-1}\, g_{k:r}(t)}{\sum_{k=1}^{r-(\ell-1)} (r-k)_{\ell-1}\, \overline{G}_{k:r}(t)}, \qquad a.e.\ t > 0, \quad \ell = 1, 2, \ldots, r.\end{aligned} \tag{46}$$

Proof of ***(i)***. For $\ell = 1$ Eq. (45) follows immediately by Eq. (42). For $\ell = 2, \ldots, r$, Eq. (45) is immediately achieved by inverting the Eq. (43).

Proof of ***(ii)***. Eq. (46) is obtained by resorting to Eq. (45) and Proposition 8, together with the circumstance that $\mathbb{P}\left(T_{1:A} > t\right) = \delta_d\left(\overline{G}(t)\right)$. □

These results can be further specialized to the case of exchangeable times $T_1, \ldots, T_r$. For any $d = 1, \ldots, r$, the exchangeable random lifetimes $T_1, \ldots, T_d$, are characterized by the m.c.h.r. functions $\mu^{[d]}(t|k; t_1, .., t_k)$ for $k = 0, \ldots, d-1$, $0 < t_1 < \cdots < t_k \le t$, and therefore (see (32) and Proposition 15 for $r = d$)

$$\Lambda_{\emptyset}^{[d]}(t) = d\mu^{[d]}(t|0). \tag{47}$$

In view of this remark, the set of functions $\mathcal{C}$ is equivalent to the set of functions

$$\mathcal{C}' := \{\mu^{[1]}(t|0), \ldots, \mu^{[r]}(t|0)\}, \tag{48}$$

which is therefore equivalent also to the systems of functions $\mathcal{A}$ and $\mathcal{B}$.

More precisely, when $T_1, \ldots, T_r$ are exchangeable and satisfy **(H5)**, then, with the above notation, Eq.s (42), (43) and (44) can be rewritten as

$$\overline{G}(t) = \exp\left\{-\int_0^t \mu^{[1]}(s|0)\,ds\right\}, \quad t > 0, \tag{49}$$

$$\delta_d(u) = \exp\left\{-d\int_0^{\overline{G}^{-1}(u)} \mu^{[d]}(s|0)\,ds\right\}, \quad u \in [0,1],\ d = 2, \ldots, r, \tag{50}$$

and

$$\overline{G}_{\ell:r}(t) = \sum_{h=r-\ell+1}^{r} (-1)^{h-r-1+\ell} \binom{r}{h}\binom{h-1}{r-\ell} \exp\left\{-h\int_0^t \mu^{[h]}(s|0)\,ds\right\}. \tag{51}$$

Similarly Eq.s (45) and (46) take the form

$$\mu^{[\ell]}(t|0) = -\frac{1}{\ell}\frac{d}{dt}\log\left[\delta_\ell(\overline{G}(t))\right], \quad a.e.\ t > 0, \quad \ell = 1, 2, \ldots, r, \tag{52}$$

$$\begin{aligned}\mu^{[\ell]}(t|0) &= \frac{1}{\ell}\frac{\sum_{h=\ell}^{r}(h)_\ell\left(g_{r-h+1:r}(t) - g_{r-h:r}(t)\right)}{\sum_{h=\ell}^{r}(h)_\ell\left(\overline{G}_{r-h+1:r}(t) - \overline{G}_{r-h:r}(t)\right)}\\ &= \frac{1}{\ell}\frac{\sum_{k=1}^{r-(\ell-1)}(r-k)_{\ell-1}\,g_{k:r}(t)}{\sum_{k=1}^{r-(\ell-1)}(r-k)_{\ell-1}\,\overline{G}_{k:r}(t)}, \quad a.e.\ t > 0, \quad \ell = 1, \ldots, r.\end{aligned} \tag{53}$$

5 Special cases

The arguments developed in the previous sections will now be illustrated by considering in Subsections 5.1 and 5.2 the two remarkable classes of models respectively defined by Archimedean copulas and by multivariate conditional hazard rate functions satisfying the load-sharing condition. These choices in a sense correspond to the simplest possible forms admitted in the two types of descriptions of a joint distribution for lifetimes, respectively.

In particular, the analysis of these classes will allow us to obtain some examples of application for some of the results derived so far, by showing the special form taken by related formulas. Subsection 5.1 is tailored to illustrate basic aspects of the exchangeable case. On the other hand the arguments in Subsection 5.2 permit to pave the way for a better understanding of the differences between exchangeability and minimal stability, and to present some heuristic ideas at the basis of the construction of minimally stable, but non-exchangeable, multivariate models.

5.1 Archimedean Copulas

Let us consider the case when the survival copula K of $T_1, \ldots, T_r$ is Archimedean with generator ψ. For simplicity's sake we assume that K is a strict Archimedean copula, i.e., the generator ψ is a strictly decreasing, continuous and convex function, such that $\psi(1) = 0$ and $\psi(0^+) = \infty$, so that $K = C_\psi$, with

$$C_\psi(u_1, \ldots, u_r) := \psi^{-1}\left(\psi(u_1) + \cdots + \psi(u_r)\right).$$

It is important to recall that the function C_ψ is an r-dimensional copula if and only if the inverse function ψ^{-1} is r-monotonic (see Theorem 6.3.6 in Schweizer and Sklar [25], Nelsen [20], see also Mc Neil and Nešlehová [16]). By Definition 1 it is immediately seen that the diagonal sections associated to $K = C_\psi$ assume the form

$$\delta_\ell(u) = \psi^{-1}(\ell\psi(u)), \quad 2 \le \ell \le r.$$

Furthermore, the survival copula K being symmetric, the model $T_1, \ldots, T_r$ is exchangeable when the lifetimes share the same common marginal survival function $\overline{G}$, i.e., under condition **(H1)**. Therefore, once ψ and $\overline{G}$ are given, then the family $\mathcal{A}$ in (8) coincides with the family $\{\overline{G}(t), \psi^{-1}(2\psi(u)), \ldots, \psi^{-1}(r\psi(u))\}$. From Eq. (4) (see also Proposition 10) we know how the family $\mathcal{B}$ is generally obtained from $\mathcal{A}$. In the present case, one can more precisely write

$$\overline{G}_{\ell:r(t)} = \sum_{h=r-\ell+1}^{r} (-1)^{h-r-1+\ell} \binom{r}{h} \binom{h-1}{r-\ell} \psi^{-1}(h\psi(\overline{G}(t))), \quad 1 \le \ell \le r.$$

Moreover, under the further regularity condition **(H5)**, by Eq. (52) one can get the m.c.h.r. functions

$$\begin{aligned} \mu^{[\ell]}(t|0) &= -\frac{1}{\ell}\frac{d}{dt}\log\left[\psi^{-1}\left(\ell\psi(\overline{G}(t))\right)\right], \quad a.e. \\ &= \frac{1}{\psi'\left(\psi^{-1}\left(\ell\psi(\overline{G}(t))\right)\right)}\psi'\left(\overline{G}(t)\right)g(t), \quad a.e. \end{aligned} \tag{54}$$

Conversely, when the family $\mathcal{B}$ of the survival distribution functions $\overline{G}_{k:r}$ is given, then we can recover the family $\mathcal{A}$. In this respect we stress that, even if the explicit expression of the generator ψ is not known, still by Corollary 9 (see also Proposition 10), from $\mathcal{B}$ we immediately get the diagonal sections $\delta_d(u)$, $d = 2, \ldots, r$. Then the following question naturally arises:

Does the knowledge of $\mathcal{B}$ allow us to identify the generator ψ?

In other words we wonder whether the identity $\delta_d(u) = \psi^{-1}(d\psi(u))$ for any $d = 2, \ldots, r$ is sufficient to identify ψ. We briefly discuss about this problem in Remark 20 below. To this end it is useful to write down explicitly the case $d = r$:

$$\psi^{-1}(r\psi(u)) = \delta_r(u) = \overline{G}_{1:r}\left(\overline{G}^{-1}(u)\right),$$

and the case $d = r - 1$:

$$\psi^{-1}((r-1)\psi(u)) = \delta_{r-1}(u) = \frac{1}{r}\overline{G}_{2:r}\left(\overline{G}^{-1}(u)\right) + \left(1 - \frac{1}{r}\right)\delta_r(u).$$

Remark 20. *It is interesting to point out (see Jaworski [12]) that when $r > 2$ the generator ψ is uniquely determined (up to a multiplicative constant) by the pair δ_r and δ_{r-1}, though the proof of the latter claim is not* *constructive. On the other hand in general one cannot uniquely determine ψ only from δ_r, when $r > 2$, since there may exist infinite generators with the same diagonal section δ_r (see again [12]). However, when the diagonal δ_r satisfy the condition $\delta_r'(1^-) = r$, then Erderly et al. [10] show that ψ is uniquely determined (up to a multiplicative constant) by δ_r and, furthermore, ψ can be approximated thanks to the following formula:*

$$\psi(u) \propto \lim_{m\to\infty} r^m \big(1 - \delta_r^{-m}(u)\big),$$

where δ_r^{-m} is the composition of δ_r^{-1} with itself m times.

A further particular case within Archimedean models is the Schur-constant case, i.e., when

$$\mathbb{P}(T_1 > t_1, \ldots, T_r > t_r) = \overline{G}(t_1 + \cdots + t_r), \tag{55}$$

corresponding to the choice $\overline{G} = \psi^{-1}$ (see, e.g., [29]). In this case both the diagonal sections δ_h and the survival functions $\overline{G}_{\ell:r}$ are determined by the marginal survival function $\overline{G}(t)$:

$$\delta_h(\overline{G}(t)) = \psi^{-1}(h\psi(\overline{G}(t))) = \overline{G}(ht) \tag{56}$$

$$\overline{G}_{\ell:r}(t) = \sum_{h=r-\ell+1}^{r} (-1)^{h-r-1+\ell} \binom{r}{h} \binom{h-1}{r-\ell} \overline{G}(ht), \quad 1 \le \ell \le r.$$

Furthermore, in view of the particularly simple form (56) of the diagonal sections, Eq. (52) can be rewritten as

$$\mu^{[\ell]}(t|0) \quad = \quad -\frac{1}{\ell}\frac{d}{dt}\log\left[\overline{G}(\ell t)\right] = \frac{g(\ell t)}{\overline{G}(\ell t)}, \quad a.e. \tag{57}$$

Note that for Schur-constant models one could get the above results also directly, taking into account that (55) implies

$$\mathbb{P}(T_1 > t, \ldots, T_\ell > t) = \mathbb{P}(T_1 > t, \ldots, T_\ell > t, T_{\ell+1} > 0, \ldots, T_r > 0) = \overline{G}(\ell\, t).$$

Example 21. *In order to illustrate how to compute the m.c.h.r. functions for Archimedean models, in this example we consider two special cases within the class of Archimedean models sharing the same generator*

$$\psi(u) = (u^{-\alpha} - 1)^{\frac{1}{\beta}}, \quad \alpha > 0,\ \beta \ge 1.$$

Note that the inverse function $\psi^{-1}(t) = \frac{1}{(t^\beta+1)^\alpha}$ is completely monotonic, so that C_ψ is a copula for any $r \ge 2$. The first case is the Archimedean model with $\overline{G}(t) = e^{-t}$. Then, for any $A \subset [r]$ with $|A| = \ell$, one has

$$\begin{aligned}\mathbb{P}(T_{1:A} > t) &= \delta_\ell(\overline{G}(t)) = \psi^{-1}(\ell\psi(\overline{G}(t)))\\ &= \frac{1}{\left(\left(\ell(e^{t\alpha} - 1)^{\frac{1}{\beta}}\right)^\beta + 1\right)^\alpha} = \frac{1}{\left(\ell^\beta e^{t\alpha} - \ell^\beta + 1\right)^\alpha}.\end{aligned}$$

Therefore by (54) *and taking into account that*

$$\psi'(u) = -\frac{\alpha}{\beta}\, u^{-(\alpha+1)}\, \left(u^{-\alpha} - 1\right)^{\frac{1}{\beta}-1},$$

we get

$$\mu^{[\ell]}(t|0) = \frac{e^{\alpha t}\left(e^{\alpha t} - 1\right)^{\frac{1}{\beta}-1}}{\left(\ell^\beta e^{t\alpha} - \ell^\beta + 1\right)^{\alpha^2+\alpha}\left(\left(\ell^\beta e^{t\alpha} - \ell^\beta + 1\right)^{\alpha^2} - 1\right)^{\frac{1}{\beta}-1}}.$$

As a second case, we consider the Schur-constant model with the same generator ψ, which corresponds to the choice

$$\overline{G}(t) = \psi^{-1}(t) = \frac{1}{(t^\beta + 1)^\alpha}, \quad g(t) = \alpha\,\beta\frac{t^{\beta-1}}{(t^\beta + 1)^{\alpha+1}}.$$

Then by Eq.s (56) *and* (57) *we get*

$$\mathbb{P}(T_{1:A} > t) = \frac{1}{\left((\ell t)^\beta + 1\right)^\alpha}, \quad \mu^{[\ell]}(t|0) = \alpha\,\beta\frac{(\ell t)^{\beta-1}}{(\ell t)^\beta + 1}.$$

5.2 Time homogeneous load-sharing models

Load sharing models are characterized by the condition that the m.c.h.r. functions depend on current time and on the set of failed components at the current time, but do not depend on the failure times. Load sharing models are well known and recurrently studied in the reliability literature (see, e.g., Spizzichino [30], Rychlik and Spizzichino [23], and the references cited therein). In particular the joint and marginal distributions of the order statistics have been studied in some details. For what concerns the special case of exchangeability see also Kamps [14].
In the literature it is also assumed that the m.c.h.r. functions do not depend on the order of failures, however it is interesting here to extend such a definition to a generalized class of models in which instead also the order of failure times may influence the m.c.h.r. functions. Actually some of the existing results on load sharing models can be easily extended to this class.

Definition 22. *The joint distribution of the random variables $T_1, \dots, T_r$ is an* **Order Dependent Load Sharing model (ODLS)** *if it is absolutely continuous, and the m.c.h.r. functions do not depend on the failure times, i.e., for any $k = 0, 1, \dots, r-1$, there exist $\binom{r}{k}k!(r-k)$ functions $v \mapsto \lambda_{j|j_1,\dots,j_k}(v)$ such that, for any $0 < v_1 < \dots < v_k < v$, and $(j_1, \dots, j_k) \in \Pi_k([r])$*

$$\lambda_{j|j_1,\dots,j_k}(v|v_1, \cdots, v_k) = \lambda_{j|j_1,\dots,j_k}(v).$$

The model is said simply **Load Sharing (LS) model** *when the m.c.h.r. functions depend neither on the failure times nor on the order of failures, i.e., for any $k = 0, 1, \dots, r\text{-}1$, there exist $\binom{r}{k}(r-k)$ functions $v \mapsto \lambda_{j|\{j_1,\dots,j_k\}}(v)$ such that, for any $0 < v_1 < \dots < v_k < v$, and $(j_1, \dots, j_k) \in \Pi_k([r])$*

$$\lambda_{j|j_1,\dots,j_k}(v|v_1, \cdots, v_k) = \lambda_{j|\{j_1,\dots,j_k\}}(v).$$

When furthermore the functions $v \mapsto \lambda_{j|j_1,\dots,j_k}(v) = \lambda_{j|j_1,\dots,j_k}$ (respectively, the functions $v \mapsto \lambda_{j|\{j_1,\dots,j_k\}}(v)$) are constant w.r.t. time v, then the model $T_1, \dots, T_r$ is said an **Order Dependent Time Homogeneous Load Sharing model** *(ODTHLS) (respectively, a* **Time Homogeneous Load Sharing model** *(THLS)).*

Clearly a Load Sharing model is also an Order Dependent Load Sharing model. To distinguish between the two cases, we will sometimes say that a model is a ***strictly Order Dependent Load Sharing model*** when the m.c.h.r. functions do depend on the order. From an engineering-oriented viewpoint, strictly ODLS models do not seem very significant for applications in the field of reliability. However models with this property may emerge in different fields, as shown in De Santis, Spizzichino [8] in the analysis of aggregation paradoxes. See also Example 29 below, where the condition of load sharing must be limited to strictly ODTHLS models on the purpose of finding among uniform frailty models (see (76)) those which are minimally stable, without falling in the exchangeable case.

Concerning the analysis of the minimal stability, it is important to stress that, in general, d-marginal models of load sharing models are not load sharing. This fact entails that the m.c.h.r. functions $\lambda^A_{j|\emptyset}$ are in general not easy to compute, even in the exchangeable case. In the latter case however we will be able to compute such functions explicitly by using the results of Section 4. On the contrary in the non-exchangeable case, we will give minimal stability conditions in terms of the m.c.h.r. functions $\lambda_{j|j_1,\dots,j_{k-1}}$.

Exchangeable THLS models. An exchangeable load sharing model clearly cannot be strictly order dependent, in that its m.c.h.r. functions are such that $\mu(t|k; t_1, \dots, t_k) = \mu(t|k)$. Furthermore it is time homogeneous if and only if for any $k = 0, 1, .., r-1$ there exists a constant $L(r-k)$ such that $\Lambda_{j_1,\dots,j_k}(t) = L(r-k)$ and

$$\mu(k) = \frac{L(r-k)}{r-k}. \tag{58}$$

In such a case it is easily seen that

$$\overline{G}_{k:r}(t) = \mathbb{P}\Big(\tfrac{X_0}{L(r)} + \tfrac{X_1}{L(r-1)} + \cdots + \tfrac{X_{k-1}}{L(r-(k-1))} > t\Big)$$

where $X_i \sim EXP(1)$, $i = 0, 1, 2\dots, r-1$, are independent random variables (see in particular Spizzichino [30], Kamps [14], Cramer and Kamps [4], and references therein). In other words, for any $k = 1, .., r$, the distribution of $T_{k:r}$ coincides with the distribution of the sum of k independent exponential distributions of parameters $\gamma_1 = L(r), \ldots., \gamma_k = L(r-(k-1))$. In the literature such a distribution is known as Generalized Erlang or Hypoexponential distribution: for a fixed vector $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_r) \in \mathbb{R}_+^r$,

$$\overline{G}_k^{\boldsymbol{\gamma}}(t) := \mathbb{P}\Big(\sum_{j=1}^{k} \tfrac{Y_j}{\gamma_j} > t\Big) \tag{59}$$

for $Y_1, \ldots, Y_r$, independent and standard exponential random variables.

When $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_r)$ is such that $\gamma_i \neq \gamma_j$ for all $i \neq j$, the above distribution is referred to as Hyperexponential, and furthermore (see, e.g., Cramer and Kamps [4] and references therein) the survival function and the probability density are respectively given by

$$\overline{G}_k^{\boldsymbol{\gamma}}(t) = \sum_{j=1}^{k} \left(\prod_{h \in [k]\setminus\{j\}} \frac{\gamma_h}{\gamma_h - \gamma_j} \right) e^{-\gamma_j t},$$

and

$$g_k^{\boldsymbol{\gamma}}(t) = \sum_{j=1}^{k} \left(\prod_{h \in [k]\setminus\{j\}} \frac{\gamma_h}{\gamma_h - \gamma_j} \right) \gamma_j e^{-\gamma_j t}.$$

Therefore, denoting by $\boldsymbol{L}$ the vector

$$\boldsymbol{L} := \big(L(r), L(r-1), \ldots, L(1)\big),$$

in the exchangeable case we can write

$$\overline{G}_{k:r}(t) = \overline{G}_k^{\boldsymbol{L}}(t). \tag{60}$$

Furthermore, on the one hand formula (9) takes the special form

$$\overline{G}(t) = \frac{1}{r} \sum_{k=1}^{r} \mathbb{P}\big(\tfrac{X_0}{L(r)} + \tfrac{X_1}{L(r-1)} + \cdots + \tfrac{X_{k-1}}{L(r-(k-1))} > t\big) = \frac{1}{r} \sum_{k=1}^{r} \overline{G}_k^{\boldsymbol{L}}(t). \tag{61}$$

On the other hand, taking into account (11), for any $A \subset [r]$, with $|A| = d$, one has

$$\begin{aligned} \mathbb{P}\big(T_{1:A} > t\big) = \delta_d\big(\overline{G}(t)\big) &= \frac{d}{(r)_d} \sum_{k=1}^{r-d+1} (r-k)_{d-1}\, \mathbb{P}\big(\tfrac{X_0}{L(r)} + \tfrac{X_1}{L(r-1)} + \cdots + \tfrac{X_{k-1}}{L(r-(k-1))} > t\big) \\ &= \frac{d}{(r)_d} \sum_{k=1}^{r-d+1} (r-k)_{d-1}\, \overline{G}_k^{\boldsymbol{L}}(t) \end{aligned} \tag{62}$$

and consequently, recalling the notation in (47), (53) becomes

$$\mu^{[d]}(t|0) = \frac{1}{d} \frac{\sum_{k=1}^{r-d+1} (r-k)_{d-1}\, g_k^{\boldsymbol{L}}(t)}{\sum_{k=1}^{r-d+1} (r-k)_{d-1}\, \overline{G}_k^{\boldsymbol{L}}(t)}. \tag{63}$$

In particular, assuming that $L(i) \neq L(j)$ for $i \neq j$, and setting

$$\vartheta_{\ell,k}^{\boldsymbol{L}} := \prod_{h \in \{0,\ldots,k-1\}\setminus\{\ell\}} \frac{L(r-h)}{L(r-h) - L(r-\ell)},$$

one has

$$\begin{aligned} \overline{G}(t) &= \frac{1}{r} \sum_{k=1}^{r} \sum_{j=1}^{k} \left(\prod_{h \in [k]\setminus\{j\}} \frac{L(r-(h-1))}{L(r-(h-1)) - L(r-(j-1))} \right) e^{-L(r-(j-1))t} \\ &= \frac{1}{r} \sum_{\ell=0}^{r-1} \left(\sum_{k=\ell+1}^{r} \vartheta_{\ell,k}^{\boldsymbol{L}} \right) e^{-L(r-\ell)t}, \\ \mathbb{P}\big(T_{1:A} > t\big) &= \frac{d}{(r)_d} \sum_{\ell=0}^{r-d} \left(\sum_{k=\ell+1}^{r-d+1} (r-k)_{d-1} \vartheta_{\ell,k}^{\boldsymbol{L}} \right) e^{-L(r-\ell)t} \end{aligned}$$

and

$$\mu^{[d]}(t|0) = \frac{1}{d}\,\frac{\sum_{\ell=0}^{r-d}\left(\sum_{k=\ell+1}^{r-d+1}(r-k)_{d-1}\,\vartheta^{\boldsymbol{L}}_{\ell,k}\right)\,L(r-\ell)\,e^{-L(r-\ell)t}}{\sum_{\ell=0}^{r-d}\left(\sum_{k=\ell+1}^{r-d+1}(r-k)_{d-1}\vartheta^{\boldsymbol{L}}_{\ell,k}\right)\,e^{-L(r-\ell)t}}$$

Note that, when $d = r$, then we obviously get that the function $t \mapsto \mu^{[r]}(t|0)$ is constant and $\mu^{[r]}(t|0) = \frac{1}{r}\,L(r)$, whereas for $d < r$ the function $t \mapsto \mu^{[d]}(t|0)$ is not constant.
This fact is somehow related to the afore-mentioned circumstance that the d-dimensional marginal distributions of a load sharing model is generally not load sharing.

Before passing to the non-exchangeable case, we observe that in the present THLS exchangeable case the function $\Psi(t;[r],\mathbf{j})$ defined in (34) can be explicitly computed:

$$\begin{aligned}\Psi(t;[r],\mathbf{j}) &= \mathbb{P}\big(T_{j_1} < T_{j_2} < \cdots < T_{j_d} \le t < T_i,\ \forall i \notin \{j_1,\dots,j_d\}\big)\\ &= \frac{1}{d!}\,\mathbb{P}\big(T_j \le t < T_i,\ \forall j \in \{j_1,\dots,j_d\},\ \forall\, i \notin \{j_1,\dots,j_d\}\big)\\ &= \frac{1}{d!}\,\frac{1}{\binom{r}{d}}\,\mathbb{P}\big(N(t)=d\big) = \frac{1}{(r)_d}\,\big[\overline{G}^{\boldsymbol{L}}_{d+1}(t) - \overline{G}^{\boldsymbol{L}}_{d}(t)\big],\end{aligned} \tag{64}$$

where we have used the notation introduced in (59). In view of Proposition 16, expression (64) turns out to be useful also in the analysis of minimal stability conditions. Indeed even for any ODTHLS model one has

$$\begin{aligned}\Psi(t;[r],\mathbf{j}) &= \mathbb{P}\big(T_{j_1} < T_{j_2} < \cdots < T_{j_d} \le t < T_i,\ \forall i \notin \{j_1,\dots,j_d\}\big)\\ &= \prod_{\ell=1}^{d}\lambda_{j_\ell|j_1,\dots,j_{\ell-1}}\cdot\int_0^t ds_d\int_0^{s_d} ds_{d-1}\cdots\int_0^{s_2} ds_1\\ &\qquad\qquad\left[e^{-(t-s_d)\Lambda_{j_1,\dots,j_d}}\prod_{\ell=1}^{d}e^{-(s_\ell-s_{\ell-1})\Lambda_{j_1,\dots,j_{\ell-1}}}\right]\end{aligned} \tag{65}$$

$$= \prod_{\ell=1}^{d}\frac{\lambda_{j_\ell|j_1,\dots,j_{\ell-1}}}{\Lambda_{j_1,\dots,j_{\ell-1}}}\,\big[\overline{G}^{\boldsymbol{\Lambda}}_{d+1}(t) - \overline{G}^{\boldsymbol{\Lambda}}_{d}(t)\big], \tag{66}$$

where $\boldsymbol{\Lambda} = \big(\Lambda_\emptyset, \Lambda_{j_1}, \dots, \Lambda_{j_1,\dots,j_k}, \dots, \Lambda_{j_1,\dots,j_{r-1}}\big)$.

In the following Example 23, for the case $r = 3$, we will analyze two different THLS models: an exchangeable THLS model and a minimally stable (non-exchangeable) THLS model which is trivially not strictly order dependent. On this purpose we will compute the survival functions of minima on sets of size $d = 1, 2, 3$ and the marginal survival functions of order statistics. We will focus both on the common features and on the differences between the two models.

Example 23. *We start by considering the non-exchangeable model: let $r = 3$ and let T_1, T_2, T_3 be lifetimes jointly distributed according to a THLS model with m.c.h.r. functions given as follows:*

$$\begin{aligned}&\lambda_{1|\emptyset}(t) = \lambda_{2|\emptyset}(t) = \lambda_{3|\emptyset}(t) = \frac{1}{3},\\ &\lambda_{3|1}(t) = \gamma,\ \lambda_{2|1}(t) = 1-\gamma, \quad \lambda_{1|2}(t) = \gamma,\ \lambda_{3|2} = 1-\gamma \quad \lambda_{2|3}(t) = \gamma,\ \lambda_{1|3}(t) = 1-\gamma,\end{aligned}$$

for a fixed value $\gamma \in \left(\frac{1}{2}, 1\right)$ and finally

$$\lambda_{1|2,3}(t) = \lambda_{1|3,2}(t) = \lambda_{2|1,3}(t) = \lambda_{2|3,1}(t) = \lambda_{3|1,2}(t) = \lambda_{3|2,1}(t) = 2.$$

For this model one has

(*i*) $\Lambda_\emptyset = 1, \quad \Lambda_{j_1} = 1, \quad \Lambda_{j_1,j_2} = 2,$ *for any $j_1 \neq j_2 \in \{1,2,3\}$.*

Furthermore we consider lifetimes $\widetilde{T}_1, \widetilde{T}_2, \widetilde{T}_3$ jointly distributed according to the exchangeable THLS model defined by

$$(ii) \quad \widetilde{\Lambda}_{\emptyset} = L(3) = 1,\ \widetilde{\Lambda}_{j_1} = L(2) = 1,\ \widetilde{\Lambda}_{j_1,j_2} = L(1) = 2,\ \textit{for any } j_1 \neq j_2 \in \{1,2,3\},$$

or, equivalently (recall (58) *with $r = 3$), such that $\mu(0) = \frac{1}{3},\ \mu(1) = \frac{1}{2},\ \mu(2) = 2$.*

As we are going to see below, the two models share the same marginal survival functions of the order statistics. Before checking this property, we point out a main difference between the two models: for the exchangeable model one obviously has

$$\mathbb{P}\big(\widetilde{T}_{j_1} < \widetilde{T}_{j_2} < \widetilde{T}_{j_3}\big) = \frac{1}{3!}, \quad \textit{for any } (j_1, j_2, j_3) \in \Pi([3]),$$

whereas, for instance, one has

$$\frac{1-\gamma}{3} = \mathbb{P}(T_1 < T_2 < T_3) < \mathbb{P}(T_1 < T_3 < T_2) = \frac{\gamma}{3}.$$

The above inequality is implied by the observation that for THLS models one has (see, e.g., Spizzichino [30])

$$\mathbb{P}(T_{j_1} < T_{j_2} < T_{j_3}) = \frac{\lambda_{j_1|\emptyset}}{\Lambda_\emptyset} \frac{\lambda_{j_2|j_1}}{\Lambda_{j_1}} \frac{\lambda_{j_3|j_1,j_2}}{\Lambda_{j_1,j_2}}.$$

The above inequality shows also that T_1, T_2, T_3 is a non-exchangeable THLS model. However the random variables T_1, T_2, T_3 are minimally stable. Indeed by the previous values in (i), and by Eq. (65)*, one has: for any $j_1 \in \{1,2,3\}$,*

$$\begin{aligned}\Psi\big(t;[3],j_1\big) &= \mathbb{P}\big(T_{j_1} \le t, T_j > t, j \neq j_1\big)\\ &= \int_0^t \lambda_{j_1|\emptyset} e^{-s\Lambda_\emptyset}\, e^{-(t-s)\Lambda_{j_1}}\, ds = \int_0^t \frac{1}{3}\, e^{-s}\, e^{-(t-s)}\, ds = \frac{1}{3}\, t\, e^{-t};\end{aligned} \tag{67}$$

for any $(j_1, j_2) \in \Pi_2([3])$,

$$\begin{aligned}\Psi\big(t;[3],(j_1,j_2)\big) &= \mathbb{P}\big(T_{j_1} \le T_{j_2} \le t, T_{j_3} > t\big)\\ &= \int_0^t ds \int_0^s ds'\, e^{-(t-s)\Lambda_{j_1,j_2}}\, \lambda_{j_1|\emptyset}\lambda_{j_2|j_1}\, e^{-s'\Lambda_\emptyset}\, e^{-(s-s')\Lambda_{j_1}}\, ds\\ &= \frac{1}{3}\, \lambda_{j_2|j_1}\, e^{-2t} \int_0^t s\, e^s\, ds = \frac{1}{3}\, \lambda_{j_2|j_1}\, \big(e^{-t} t - e^{-t} + e^{-2t}\big).\end{aligned}$$

Taking into account that, for any (j_1, j_2)

$$\lambda_{j_2|j_1} + \lambda_{j_1|j_2} = \gamma + (1-\gamma) = 1,$$

we may apply Proposition 16 to conclude that T_1, T_2, T_3 are minimally stable. Furthermore we get that, for any $(j_1, j_2, j_3) \in \Pi\big([3]\big)$,

$$\mathbb{P}\big(T_{j_1} \le t, T_{j_2} \le t, T_{j_3} > t\big) = \frac{1}{3}\, \big(e^{-t} t - e^{-t} + e^{-2t}\big). \tag{68}$$

From the previous computations, and in particular (67) *and* (68)*, we get explicitly the following survival functions*

$$\begin{aligned}\mathbb{P}(T_{1:\{1,2,3\}} > t) &= \mathbb{P}(T_1 > t, T_2 > t, T_3 > t) = e^{-t};\\ \mathbb{P}(T_{1:\{1,2\}} > t) &= \mathbb{P}(T_1 > t, T_2 > t) = \mathbb{P}(T_1 > t, T_2 > t, T_3 > t)\\ &\quad + \mathbb{P}(T_1 > t, T_2 > t, T_3 \le t) = e^{-t} + \frac{1}{3}\, t\, e^{-t} = e^{-t}\left(1 + \frac{t}{3}\right),\\ \overline{G}(t) = \mathbb{P}(T_1 > t) &= \mathbb{P}(T_1 > t, T_2 > t, T_3 > t) + \mathbb{P}(T_1 > t, T_2 > t, T_3 \le t)\\ &\quad + \mathbb{P}(T_1 > t, T_3 > t, T_2 \le t) + \mathbb{P}(T_1 > t, T_2 \le t, T_3 \le t)\\ &= e^{-t} + 2\,\frac{t}{3}\, e^{-t} + \frac{1}{3}\, \big(t\, e^{-t} - e^{-t} + e^{-2t}\big) = \frac{2}{3}\, e^{-t} + t\, e^{-t} + \frac{1}{3}\, e^{-2t}.\end{aligned}$$

To see that the families of the marginal survival functions of the order statistics coincide for the two models, we take into account that in general, even with $r \geq 3$,

$$\overline{G}_{1:r}(t) = e^{-\int_0^t \Lambda_\emptyset(s)\,ds} = \mathbb{P}(N(t) = 0), \quad \text{and} \quad \overline{G}_{\ell:r}(t) = \mathbb{P}(N(t) \leq \ell - 1).$$

We notice furthermore that, for minimally stable models, Eq. (17) *(with $h = r - k$ and $H = \{k+1, \ldots, r\}$) and Eq.* (35) *imply that*

$$\begin{aligned}\mathbb{P}(N(t) = k) &= \binom{r}{k} \mathbb{P}\big(T_i \leq t,\ \forall i \in \{1, 2, .., k\}, T_j > t,\ \forall j \in \{k+1, \ldots, r\}\big)\\ &= \binom{r}{k} \sum_{(j_1, \ldots, j_k) \in \Pi([k])} \Psi\big(t; [r], (j_1, j_2, .., j_k)\big), \quad 1 \leq k \leq r.\end{aligned}$$

Thus, comparing Eq. (66) *with Eq.* (64) *for minimal THLS models, and taking into account that $\boldsymbol{\Lambda} = \boldsymbol{L}$, i.e., $\Lambda_{j_1,\ldots,j_k} = \widetilde{\Lambda}_{j_1,\ldots,j_k} = L(r-k)$, we obtain the afore mentioned conclusion, i.e., that for both models one has*

$$\overline{G}_{1:3}(t) = e^{-\Lambda_\emptyset t} = e^{-t}, \quad \overline{G}_{2:3}(t) = e^{-t}\,(1+t), \quad \overline{G}_{3:3}(t) = 2\,e^{-t}\,t + e^{-2t}.$$

We will see moreover that even the respective joint distributions of the order statistics do coincide for the two models. Actually the latter circumstance is a consequence of the condition that the functions $(j_1, \ldots, j_k) \mapsto \Lambda_{j_1,\ldots,j_k}$ are constant, only depending on k (see Remark 27 below and condition (80) *in Example 32 in the Appendix).*

Minimally stable ODTHLS models. We start our discussion with a simple necessary condition for minimal stability of ODTHLS models.

Lemma 24. *Let $T_1, \ldots, T_r$ be an ODTHLS model. If $T_1, \ldots, T_r$ are minimally stable then necessarily*

$$\lambda_{i|\emptyset} = \frac{\Lambda_\emptyset}{r} \quad \text{and} \quad \Lambda_i = \Lambda_1, \quad \forall\, i \in [r]$$

Proof. By Proposition 5 we know that when $T_1, \ldots, T_r$ are minimally stable, for any $t > 0$ the probabilities $\mathbb{P}(T_i \leq t,\ T_j > t, \forall j \neq i)$ necessarily assume the same value for any $i \in [r]$. Taking into account that (see (65) and (66))

$$\mathbb{P}(T_i \leq t,\ T_j > t, \forall j \neq i) = \lambda_{i|\emptyset} \int_0^t e^{-\Lambda_\emptyset s}\, e^{-\Lambda_i (t-s)} ds$$

one immediately gets that

$$\mathbb{P}(T_i \leq t,\ T_j > t, \forall j \neq i) = \begin{cases} \lambda_{i|\emptyset}\, t\, e^{-\Lambda_\emptyset t} & \text{if } \Lambda_i = \Lambda_\emptyset. \\[2ex] \lambda_{i|\emptyset} \dfrac{e^{-\Lambda_i t} - e^{-\Lambda_\emptyset t}}{\Lambda_\emptyset - \Lambda_i} & \text{if } \Lambda_i \neq \Lambda_\emptyset, \end{cases}$$

If there exists $i_0 \in [r]$ such that $\Lambda_{i_0} = \Lambda_\emptyset$ then necessarily $\Lambda_i = \Lambda_\emptyset$, for any $i \in [r]$. Consequently, also $\lambda_{i|\emptyset} = \lambda_{i_0|\emptyset}$.

Viceversa if there exists $i_0 \in [r]$ such that $\Lambda_{i_0} \neq \Lambda_\emptyset$, then necessarily $\Lambda_i \neq \Lambda_\emptyset$, for any $i \in [r]$. Furthermore one necessarily has

$$\lambda_{i|\emptyset} \frac{e^{-\Lambda_i t} - e^{-\Lambda_\emptyset t}}{\Lambda_\emptyset - \Lambda_i} = \lambda_{1|\emptyset} \frac{e^{-\Lambda_1 t} - e^{-\Lambda_\emptyset t}}{\Lambda_\emptyset - \Lambda_1}.$$

Then the thesis follows by the linear independence of the functions $t \mapsto e^{at}$ for different values of $a \in \mathbb{R}$. □

In the next result (see Proposition 25 below) we show that the survival functions $\overline{G}_{h:r}$ of a minimally stable ODTHLS model is a mixture of Hypoexponential distributions.

Before stating formally our result we need to introduce some further notation. Let $T_1, \dots, T_r$ be an ODTHLS model. For any permutation $\mathbf{j} = (j_1, \dots, j_r) \in \Pi[r]$ we will write $\boldsymbol{\Lambda}_{j_1,\dots,j_r}$ to denote the vector

$$\boldsymbol{\Lambda}_{j_1,\dots,j_r} := (\Lambda_\emptyset, \Lambda_{j_1}, \dots, \Lambda_{j_1,\dots,j_k}, \dots, \Lambda_{j_1,\dots,j_{r-1}}).$$

We will also use the shorter notation $\boldsymbol{\Lambda}_\mathbf{j}$ instead of $\boldsymbol{\Lambda}_{j_1,\dots,j_r}$.

Then we consider the partition of $\Pi[r]$ generated by the equivalence relation

$$\mathbf{j} \sim \mathbf{j}' \quad \Leftrightarrow \quad \boldsymbol{\Lambda}_{j_1,\dots,j_r} = \boldsymbol{\Lambda}_{j'_1,\dots,j'_r}.$$

If we define

$$\mathcal{L} := \{\boldsymbol{L} : \exists\, \mathbf{j} \in \Pi([r]) \text{ with } \boldsymbol{\Lambda}_\mathbf{j} = \boldsymbol{L}\} \tag{69}$$

then the elements of the partition may be labeled by the vectors $\boldsymbol{L} \in \mathcal{L}$:

$$\Pi([r]) = \bigcup_{\boldsymbol{L}\in\mathcal{L}} \Pi([r]; \boldsymbol{L}),$$

where

$$\Pi([r]; \boldsymbol{L}) := \{\mathbf{j} \in \Pi([r]) \text{ such that } \boldsymbol{\Lambda}_\mathbf{j} = \boldsymbol{L}\}. \tag{70}$$

For our purposes it is convenient to label the coordinates of the vectors in $\mathcal{L}$ as follows:

$$\boldsymbol{L} = \big(L(r), L(r-1), \dots, L(1)\big).$$

When $T_1, \dots, T_r$ are minimally stable, then, in view of Lemma 24, $L(r)$ takes on the same value for any $\boldsymbol{L} \in \mathcal{L}$, and the same happens for $L(r-1)$. More precisely one has

$$L(r) = \Lambda_\emptyset, \quad L(r-1) = \Lambda_1 = \Lambda_i, \ \forall\, i \in [r].$$

Proposition 25. *Let $T_1, \dots, T_r$ be a minimally stable ODTHLS model. Then, with the notation introduced above, the survival functions $\overline{G}_{\ell:r}$, $\ell = 1, \dots, r$, can be obtained as the following mixture of Hypoexponential survival functions*

$$\overline{G}_{\ell:r}(t) = \sum_{\boldsymbol{L}\in\mathcal{L}} \frac{|\Pi([r]; \boldsymbol{L})|}{r!}\, \overline{G}^{\boldsymbol{L}}_\ell(t). \tag{71}$$

Before giving the proof of the previous proposition, it is convenient to present the following remarks.

Remark 26. *The previous expression* (71) *of $\overline{G}_{\ell:r}(t)$ is alternative to the expression* (44) *given in Proposition 18. The main difference is that in* (44) *we need to know explicitly the hazard rates $\Lambda^{[h]}_\emptyset$ of the minima, but we need not to know explicitly the m.c.h.r. functions $\lambda_{j|j_1,\dots,j_k}$, while, viceversa, in* (71) *we need to know explicitly the m.c.h.r. functions, but we need not to compute the hazard rates $\Lambda^{[h]}_\emptyset$ of the minima.*

Remark 27. *Note that (see* (60)*) the functions $\overline{G}^{\boldsymbol{L}}_k(t)$ are the survival functions of the order statistics of an exchangeable THLS model with $\mu(k) = L(r-k)/(r-k)$. Therefore the r.h.s. of (71) can be interpreted as a mixture of the survival functions of exchangeable models. In the model T_1, T_2, T_3 of Example 23 the mixture turns out to be degenerate, since the set $\mathcal{L}$ is the singleton $\{(1, 1, 2)\}$. The latter circumstance explains the reason why the two models in Example 23 share the same family $\{\overline{G}_{1:3}, \overline{G}_{2:3}, \overline{G}_{3:3}\}$.*

Proof of Proposition 25. Consider a random permutation $\sigma_1, \dots, \sigma_r$, uniformly distributed in $\Pi([r])$. Then $T_{\sigma_1}, \dots, T_{\sigma_r}$ is the symmetrized model of $T_1, \dots, T_r$, denoted by $\widetilde{T}_1, \dots, \widetilde{T}_r$ in Remark 11.

Clearly the two models share the same distributions for the order statistics, and the joint distribution of $\widetilde{T}_1, \dots, \widetilde{T}_r$ is the mixture over $\boldsymbol{L} \in \mathcal{L}$ of the exchangeable THLS models with m.h.c.r. functions

$$\mu(k) = \frac{L(r-k)}{r-k},$$

with mixture weights given by $\frac{|\Pi([r];\boldsymbol{L})|}{r!}$. For any $\boldsymbol{L} \in \mathcal{L}$, the survival functions of the order statistics of the above exchangeable THLS model is $\overline{G}^{\boldsymbol{L}}_{h:r}(t)$, and therefore the marginal survival functions of the order statistics of the model $\widetilde{T}_1,\dots,\widetilde{T}_r$ are given by

$$\mathbb{P}(\widetilde{T}_{h:r} > t) = \sum_{\boldsymbol{L}\in\mathcal{L}} \frac{|\Pi([r];\boldsymbol{L})|}{r!}\,\overline{G}^{\boldsymbol{L}}_{h:r}(t),$$

whence the thesis follows. □

Similarly to (71) one obtains that for minimally stable ODTHLS models the marginal survival function and the survival functions of the minima are mixtures over $k \in [r]$ and $\boldsymbol{L} \in \mathcal{L}$ of $\overline{G}^{\boldsymbol{L}}_k(t)$. More precisely by (9) one gets immediately that

$$\overline{G}(t) = \frac{1}{r}\sum_{k=1}^{r}\sum_{\boldsymbol{L}\in\mathcal{L}} \frac{|\Pi([r];\boldsymbol{L})|}{r!}\,\overline{G}^{\boldsymbol{L}}_k(t) \tag{72}$$

and, furthermore, by (11) one gets

$$\mathbb{P}\big(T_{1:A} > t\big) = \delta_d\big(\overline{G}(t)\big) = \frac{d}{(r)_d}\sum_{k=1}^{r-d+1}(r-k)_{d-1}\sum_{\boldsymbol{L}\in\mathcal{L}} \frac{|\Pi([r];\boldsymbol{L})|}{r!}\overline{G}^{\boldsymbol{L}}_k(t). \tag{73}$$

As a generalization of the arguments presented in Example 23, we now characterize the set of all minimally stable ODTHLS models T_1, T_2, T_3 in terms of the m.c.h.r. functions.

Example 28 (Minimally stable ODTHLS with $r = 3$)**.** *Let us consider the ODTHLS model with T_1, T_2, T_3 whose joint distribution is given in terms of the m.c.h.r. functions $\lambda_{j_1|\emptyset}$, $\lambda_{j_2|j_1}$, $\lambda_{j_3|j_1,j_2}$, $j_1, j_2, j_3 \in \{1,2,3\}$, $j_2 \neq j_1$, $j_3 \neq j_1, j_2$.*
We are going to prove that T_1, T_2, T_3 are minimally stable if and only if conditions **(A1)** *and* **(A2)** *below hold, together with either condition* **(A3)** *or condition* **(A3)′**, *where*

(A1) *there exists a value $L(3)$ such that*

$$\lambda_{1|\emptyset} = \lambda_{2|\emptyset} = \lambda_{3|\emptyset} = \frac{L(3)}{3};$$

(A2) *there exists a value $L(2)$ such that*

$$\Lambda_1 = \lambda_{2|1} + \lambda_{3|1} = \Lambda_2 = \lambda_{1|2} + \lambda_{3|2} = \Lambda_3 = \lambda_{1|3} + \lambda_{2|3} = L(2);$$

(A3) *there exists a value $L(1)$ such that*

$$\lambda_{j_3|j_1,j_2} = L(1), \quad \forall (j_1, j_2, j_3) \in \Pi([3]),$$

and there exist two values γ_1 and γ_2 (possibly equal) such that

$$\{\lambda_{j_2|j_1}, \lambda_{j_1|j_2}\} = \{\gamma_1, \gamma_2\}, \quad \textit{for any } \{j_1, j_2\}, \tag{74}$$

and

$$\gamma_1 + \gamma_2 = L(2); \tag{75}$$

(A3)′ *there exist two values $L'(1) \neq L''(1)$ such that*

$$\lambda_{j_3|j_1,j_2} \in \big\{L'(1),\ L''(1)\big\}, \quad \forall (j_1, j_2, j_3) \in \Pi([3]),$$

there exist two values γ_1 and γ_2 (possibly equal) such that (74) *and* (75) *hold, and furthermore*

$$\lambda_{j_2|j_1} = \gamma_1,\ \lambda_{j_3|j_1,j_2} = L'(1) \quad \textit{if and only if} \quad \lambda_{j_1|j_2} = \gamma_2,\ \lambda_{j_3|j_2,j_1} = L''(1).$$

Note that the model is strictly order dependent only under condition **(A3)**$'$.

Conditions **(A1)** *and* **(A2)** *are the necessary conditions of Lemma 24 with* $L(3) := \Lambda_\emptyset$ *and* $L(2) := \Lambda_1$ *and guarantee that the following probabilities (see* (65) *and* (66)*) take the same value for any* $j_1 \in \{1, 2, 3\}$,

$$\Psi\big(t; [3], j_1\big) = \mathbb{P}\big(T_{j_1} \le t, T_j > t, j \ne j_1\big) = \int_0^t \lambda_{j_1|\emptyset} e^{-s\Lambda_\emptyset}\, e^{-(t-s)\Lambda_{j_1}}\, ds.$$

Furthermore for any $(j_1, j_2) \in \Pi_2([3])$, *we get*

$$\begin{aligned}\Psi\big(t; [3], (j_1, j_2)\big) &= \mathbb{P}\big(T_{j_1} \le T_{j_2} \le t, T_{j_3} > t\big)\\ &= \int_0^t ds \int_0^s ds'\, e^{-(t-s)\Lambda_{j_1,j_2}}\, \lambda_{j_1|\emptyset}\lambda_{j_2|j_1}\, e^{-s'\Lambda_\emptyset}\, e^{-\Lambda_{j_1}(s-s')}\, ds\end{aligned}$$

whereas

$$\Psi\big(t; [3], (j_2, j_1)\big) = \int_0^t ds \int_0^s ds' \lambda_{j_2|\emptyset}\lambda_{j_1|j_2}\, e^{-\Lambda_{j_1,j_2}(t-s)}\, e^{-\Lambda_\emptyset s'}\, e^{-\Lambda_{j_2}(s-s')}.$$

Proposition 16 guarantees that T_1, T_2, T_3 *are minimally stable if and only if for any* $t > 0$ *the following probabilities take on the same value for any* $\{j_1, j_2\} \subset \{1, 2, 3\}$:

$$\mathbb{P}\big(T_{j_1} \le t, T_{j_2} \le t, T_{j_3} > t\big) = \Psi\big(t; [3], (j_1, j_2)\big) + \Psi\big(t; [3], (j_2, j_1)\big).$$

Taking into account the necessary conditions **(A1)** *and* **(A2)**, *and that when* $r = 3$, *then* $\Lambda_{j_1,j_2} = \lambda_{j_3|j_1,j_2}$ *the previous condition is equivalent to requiring that, for any* $t > 0$ *the following sums take on the same value for any* $\{j_1, j_2\} \subset \{1, 2, 3\}$:

$$\begin{aligned}&\lambda_{j_2|j_1} \int_0^t ds \int_0^s ds'\, e^{-\lambda_{j_3|j_1,j_2}(t-s)}\, e^{-L(3)s'}\, e^{-L(2)(s-s')}\\ &\quad + \lambda_{j_1|j_2} \int_0^t ds \int_0^s ds'\, e^{-\lambda_{j_3|j_2,j_1}(t-s)}\, e^{-L(3)s'}\, e^{-L(2)(s-s')}.\end{aligned}$$

In its turn the above requirement is equivalent to either condition **(A3)** *or* **(A3)**$'$.

As a generalization of the previous example, in Example 32 (see the Appendix), we characterize the minimal stability property for ODTHLS models whose $\mathcal{L}$ is a singleton.

Among load sharing models, an interesting subclass is the class of the so-called *uniform frailty models*, whose m.c.h.r. functions are such that, for any $k = 0, 1, 2, \ldots, r-1$,

$$\lambda_{j|j_1,\ldots,j_k} = \frac{\Lambda_{j_1,\ldots,j_k}}{r-k}, \quad \forall (j_1, \ldots, j_k) \in \Pi_k([r]). \tag{76}$$

We conclude this section by analyzing the property of minimal stability for the model of the previous Example 28, under the additional assumption of uniform frailty.

Example 29. *Let us consider the model* T_1, T_2, T_3 *of the previous Example 28. If besides minimal stability, we impose the uniform frailty condition, then the model* T_1, T_2, T_3 *turns out to be non-exchangeable only if it is strictly order dependent. Indeed when condition* **(A3)** *holds, then the additional uniform frailty condition implies that* $\gamma_1 = \gamma_2 = L(2)/2$ *and therefore the model is exchangeable: for any* $(j_1, j_2, j_3) \in \Pi([3])$

$$\lambda_{j|\emptyset} = \frac{L(3)}{3}, \quad \lambda_{j_2|j_1} = \frac{L(2)}{2}, \quad \lambda_{j_3|j_1,j_2} = L(1).$$

On the contrary, when condition **(A3)**$'$ *holds, the model is strictly order dependent. Then the uniform frailty and the minimal stability conditions together become: for any* $(j_1, j_2, j_3) \in \Pi([3])$

$$\lambda_{j_1|\emptyset} = \frac{L(3)}{3}, \quad \lambda_{j_2|j_1} = \frac{L(2)}{2}, \quad \{\lambda_{j_3|j_1,j_2}, \lambda_{j_3|j_2,j_1}\} = \{L'(1), L''(1)\}.$$

A More general examples

Example 30 (A procedure to construct DD, not exchangeable, n-dimensional copulas)**.** *Our aim is to prove a generalization of Example 13. We start by proving that, given a DD $n-1$-dimensional copula C_{n-1}, then it is possible to construct a n-dimensional copula C_n which is DD. Subsequently we show that by using this construction recursively, starting with a 2-dimensional not symmetric copula, the copulas C_n are not exchangeable.*

Given the DD copula C_{n-1}, we define a n-dimensional copula C_n as

$$\begin{aligned} &C_n(u_1, \dots, u_{n-1}, u_n) \\ :=&\frac{1}{n}\left[C_{n-1}(u_1, \dots, u_{n-1}) \cdot u_n + C_{n-1}(u_2, \dots, u_{n-1}, u_n) \cdot u_1 + \cdots \right. \\ &\left. \cdots + C_{n-1}(u_{n-1}, u_n, u_1, \dots, u_{n-3}) \cdot u_{n-2} + C_{n-1}(u_n, u_1, \dots, u_{n-2}) \cdot u_{n-1}\right], \end{aligned}$$

Namely C_n is obtained as the symmetric mixture of the copulas over the n cyclic permutations of $(1, 2, \dots, n)$

$$\sigma_1 = (1, 2, \dots, n), \ \textit{and } \sigma_k = (k, k+1, \dots, n, 1, \dots, k-1), \quad 2 \le k \le n,$$

$$C_n(u_1, \dots, u_{n-1}, u_n) := \frac{1}{n} \sum_{k=1}^{n} C_{n-1}(u_{\sigma_k(1)}, \dots, u_{\sigma_k(n-1)}) \cdot u_{\sigma_k(n)}. \tag{77}$$

It is easy to see that C_n is DD, with diagonal sections

$$\delta_n^{C_n}(u) = C_n(u, \dots, u, u) = \delta_{n-1}^{C_{n-1}}(u) \cdot u, \quad \delta_1^{C_n}(u) = u$$

and

$$\delta_d^{C_n}(u) = \frac{d}{n}\, \delta_{d-1}^{C_{n-1}}(u) \cdot u + \left(1 - \frac{d}{n}\right) \delta_d^{C_{n-1}}(u), \quad 2 \le d \le n-1. \tag{78}$$

Indeed when $(u_1, \dots, u_{n-1}, u_n) = (u\mathbf{e}_A + \mathbf{e}_{A^c})$, with $A \subset [n]$, and $|A| = d$, then, for any $k = 1, \dots, n$, one can write the $n-1$-dimensional vectors appearing in (77) *as*

$$(u_{\sigma_k(1)}, u_{\sigma_k(2)}, \dots, u_{\sigma_k(n-1)}) \equiv (u_k, u_{k+1}, \dots, u_n, \dots, u_{k-2}) = (u\mathbf{e}_B + \mathbf{e}_{B^c}),$$

where B is a suitable subset of $[n-1]$. Furthermore the cardinality $|B|$ takes on either the value d or the value $d-1$, depending on the value of u_{k-1}, namely

$$|B| = \begin{cases} d & \textit{if } u_{\sigma_k(n)} \equiv u_{k-1} = 1 \\ d-1 & \textit{if } u_{\sigma_k(n)} \equiv u_{k-1} = u, \end{cases}$$

where we have used the convention that $u_0 = u_n$.

Starting from (78) *one can easily prove that*

$$\delta_d^{C_n}(u) = \alpha_{n,d}\, u^d + (1 - \alpha_{n,d})\, C(u, u) u^{d-2}, \quad 1 \le d \le n,$$

with

$$\alpha_{n,1} = 1, \quad \alpha_{n,d} = \frac{d}{n} \alpha_{n-1,d-1} + \left(1 - \frac{d}{n}\right) \alpha_{n-1,d}.$$

Starting from a fixed permutation $(\pi(1), \dots, \pi(n)) \notin \{\sigma_k,\ k = 1, \dots, n\}$*, a similar procedure can be used to construct another DD n-dimensional copula (possibly different from* C_n*), by using the n cyclic permutations* $\sigma_k \in \Pi_{\mathcal{C}}(1, 2, \dots, n)$*:*

$$C_{n,\pi}(u_1, \dots, u_{n-1}, u_n) := \frac{1}{n} \sum_{k=1}^{n} C_{n-1}\left(u_{\pi(\sigma_k(1))}, u_{\pi(\sigma_k(2))}, \dots, u_{\pi(\sigma_k(n-1))}\right) \cdot u_{\pi(\sigma_k(n))}.$$

When the procedure is implemented recursively starting with a fixed permutation $\pi \in \Pi([n])$*, and with a not-symmetric copula* $C_2(u, v) = C(u, v)$*, as in Example 13, denote by* $U_1, \dots, U_n$ *the random variables associated to the copula of* $C_{n,\pi}$*,* $n \geq 3$*. Then for any fixed* $i = 1, \dots, n$ *the 2-dimensional marginals of* U_i, U_{i+1} *(with the convention that* $U_{n+1} = U_0$*) are obtained recursively as*

$$C_{n,\pi}(u, v, \overbrace{1, \dots, 1}^{n-2}) = \frac{1}{n}\left[(n-2)C_{n-1,\pi}(u, v, \overbrace{1, \dots, 1}^{n-3}) + 2uv\right], \quad u, v \in [0, 1],$$

and therefore are all equal, i.e.,

$$C_{n,\pi}\left(u\mathbf{e}_{\{i\}} + v\mathbf{e}_{\{i+1\}} + \mathbf{e}_{[n]\setminus\{i,i+1\}}\right) = C_{n,\pi}(u, v, 1, \dots., 1).$$

Notice that therefore

$$C_{n,\pi}(u, v, 1, \dots., 1) \neq C_{n,\pi}(v, u, 1, \dots., 1),$$

so that the copulas $C_{n,\pi}$ *are not symmetric.*

Since we are particularly interested in examples with absolutely continuous joint and marginal distributions, observe that if we start this procedure with a absolutely continuous copula we get absolutely continuous copulas.

Furthermore we recall the class of absolutely continuous examples given in Navarro and Fernandez-Sanchez [17] (see in particular Proposition 1 therein). The class in [17] may be seen as a particular case of the larger class considered in the next example.

Example 31 (Negative mixtures of DD copulas are DD)**.** *Suppose that* $D(u_1, \dots, u_r)$ *is an absolutely continuous exchangeable copula with probability density d such that*

$$0 < \underline{d}\,\rho(u_1, \dots, u_r) \leq d(u_1, \dots, u_r),$$

for some positive function ρ *and some positive constant* $\underline{d}$*.*
Let $C_i(u_1, \dots, u_r)$*,* $i = 1, 2$*, be two different copulas which are DD, but non-exchangeable, and absolutely continuous, with probability density* $c_i(u_1, \dots, u_r)$ *such that, for some positive constant* $\overline{c}$

$$0 \leq c_i(u_1, \dots, u_r) \leq \overline{c}\,\rho(u_1, \dots, u_r).$$

Assume also that the function $c_1(u_1, \dots, u_r) - c_2(u_1, \dots, u_r)$ *is not symmetric, and define*

$$K_\alpha(u_1, \dots, u_r) := D(u_1, \dots, u_r) + \alpha\left[C_1(u_1, \dots, u_r) - C_2(u_1, \dots, u_r)\right]. \tag{79}$$

If α *is strictly positive and sufficiently small, then* K_α *is an absolutely continuous DD copula, but not exchangeable.*

We now proceed with the proof of the previous statement.

The function K_α *defined in* (79) *has a density*

$$k_\alpha(u_1, \dots, u_r) = d(u_1, \dots, u_r) + \alpha\left[c_1(u_1, \dots, u_r) - c_2(u_1, \dots, u_r)\right],$$

such that the integral

$$\int_{[0,1]^n} k_\alpha(u_1,\dots,u_r)du_1\cdots du_r = 1.$$

Therefore k_α *is a probability density, if and only if* $k_\alpha(u_1,\dots,u_r) \geq 0$ *for any* $(u_1,\dots,u_r) \in (0,1)^r$. *The condition* $\overline{c}\alpha \leq \underline{d}$ *implies that*

$$k_\alpha(u_1,\dots,u_r) \geq \underline{d}\rho(u_1,\dots,u_r) + \alpha[0 - \overline{c}\,\rho(u_1,\dots,u_r)] \geq 0$$

The assumption on the densities may be weakened, for instance, it is clearly not necessary any assumption on the density of C_1. *Furthermore the example could be generalized to the case*

$$K_{\alpha_1,\dots,\alpha_m}(u_1,\dots,u_r) := D(u_1,\dots,u_r) + \sum_{k=1}^{m} \alpha_k \left[C_{1,k}(u_1,\dots,u_r) - C_{2,k}(u_1,\dots,u_r)\right],$$

with suitable conditions on α_i *and* $C_{i,k}$, *for* $k = 1,\dots,m$, $i = 1, 2$.

Finally it is interesting to note that K_α *is a negative mixture of copulas, and that negative mixture of i.i.d. random variable are linked to finite exchangeability, and the problem of extendibility.*

Example 32. *(Minimally stable ODTHLS models sharing the joint distribution of the order statistics with an exchangeable THLS model) Let us assume that* $T_1,..,T_r$, *with* $r \geq 3$, *is an ODTHLS model described by the m.c.h.r. functions* $\lambda_{j|j_1,\dots,j_{d-1}}$, $d = 1,\dots,r$, $(j_1,\dots,j_{d-1},j) \in \Pi([r])$ *with the usual convention that when* $d = 1$ *then* $\lambda_{j|j_1,\dots,j_{d-1}} = \lambda_{j|\emptyset}$. *We are going to characterize all the minimally stable ODTHLS models in the particular case when the set* $\mathcal{L}$ *is the singleton* $\{\boldsymbol{L} = (L(r), L(r-1),\dots,L(1))\}$, *i.e.,*

$$\Lambda_{j_1,\dots,j_k} = \sum_{j\notin\{j_1,\dots,j_k\}} \lambda_{j|j_1,\dots,j_k} = L(r-k), \quad \forall\, k = 0,1,\dots,r-1. \tag{80}$$

By Proposition 25, for all the above minimally stable ODTHLS models, one has $\overline{G}_{k:r} = \overline{G}_k^{\boldsymbol{L}}$, $k = 1,2,\dots,r$, *i.e., the same Hypoexponential marginal survival functions of the exchangeable THLS model* (58). *The characterization of the condition that* $T_1,\dots,T_r$ *are minimally stable is then a consequence of Proposition 16 together with the comparison between* (66) *and* (64): *the m.c.h.r. functions satisfy the following system of equations*

$$\begin{cases} \sum_{j\notin\{j_1,\dots,j_{d-1}\}} \lambda_{j|j_1,\dots,j_{d-1}} = L(r-(d-1)), & \{j_1,\dots,j_{d-1}\} \subset [r] \\ \sum_{\mathbf{j}\in\Pi(I)} \prod_{h=1}^{d} \lambda_{j_h|j_1,\dots,j_{h-1}} = \frac{1}{\binom{r}{d}} \prod_{\ell=1}^{d} L(r-(\ell-1)), & \textit{for any } I \subset [r], \textit{ with } |I| = d, \\ \qquad d = 1,\dots,r. \end{cases}$$

This characterization yields that there exist infinitely many minimally stable ODTHLS models satisfying condition (80). *This statement is a consequence of the observation that one can see the previous system as a family of nested linear systems:*

$$\begin{cases} \sum_{j\in[r]} \lambda_{j|\emptyset} = L(r), \\ \lambda_{j|\emptyset} = \frac{1}{r} L(r); & j \in [r] \end{cases}$$

once $\lambda_{j|\emptyset}$ *are given, then* $\lambda_{j_2|j_1}$ *are the solutions* x_{j_1,j_2}, $j_2 \neq j_1$, *of*

$$\begin{cases} \sum_{j\neq j_1} \lambda_{j_1|\emptyset} x_{j_1,j} = L(r-1), & \forall\, j_1 \in [r] \\ \lambda_{j_1|\emptyset}\, x_{j_1,j_2} + \lambda_{j_2|\emptyset}\, x_{j_2,j_1} = \frac{1}{\binom{r}{2}} L(r)L(r-1); & \forall\, \{j_1,j_2\} \subset [r]; \end{cases}$$

once $\lambda_{j|\emptyset}$ *and* $\lambda_{j|j_1}$ *are given, then* $\lambda_{j_3|j_1,j_2}$ *are the solutions* x_{j_1,j_2,j_3} *of*

$$\begin{cases} \sum_{j\notin\{j_1,j_2\}} x_{j_1,j_2,j} = L(r-2), & \forall\, \{j_1,j_2\} \subset [r], \\ \sum_{(j_1,j_2,j_3)\in\Pi(\{k_1,k_2,k_3\})} \lambda_{j_1|\emptyset}\lambda_{j_2|j_1} x_{j_1,j_2,j_3} \\ \qquad = \frac{1}{\binom{r}{3}} L(r)\,L(r-1)L(r-2); & \forall\, \{k_1,k_2,k_3\} \subset [r]; \end{cases}$$

and so on. Since the above nested systems always admit the solutions $x_{j_1,\ldots,j_k,j} = \lambda_{j|j_1,\ldots,j_k} = \frac{L(r-k)}{r-k}$, *then there are infinite solutions.*

We end this example by observing that under condition (80) *all the ODTHLS models (not necessarily minimally stable) share with the exchangeable THLS model* (58) *not only the marginal distributions but also the joint distribution of the order statistics. More precisely the joint distribution of* $(T_{1:r}, \ldots, T_{r:r})$ *coincides with the joint distribution of*

$$\left(\frac{Y_0}{L(r)}, \frac{Y_0}{L(r)} + \frac{Y_1}{L(r-1)}, \ldots, \frac{Y_0}{L(r)} + \frac{Y_1}{L(r-1)} + \cdots \frac{Y_{r-1}}{L(1)}\right), \tag{81}$$

where Y_k, $k = 0, 1, \ldots, r-1$, *are i.i.d. standard exponential. Indeed, one can easily extend Corollary 3 in Rychlik and Spizzichino [23] for THLS models, to ODTHLS ones: for any permutation* $(j_1, \ldots, j_r) \in \Pi([r])$, *the conditional joint distribution of* $(T_{1:r}, \ldots, T_{r:r})$ *given the event* $\{T_{j_1} < T_{j_2} < \cdots < T_{j_r}\}$, *coincides with the law of*

$$\left(\frac{Y_0}{\Lambda_\emptyset}, \frac{Y_0}{\Lambda_\emptyset} + \frac{Y_1}{\Lambda_{j_1}}, \ldots, \frac{Y_0}{\Lambda_\emptyset} + \frac{Y_1}{\Lambda_{j_1}} + \cdots + \frac{Y_{r-1}}{\Lambda_{j_1,\ldots,j_{r-1}}}\right), \tag{82}$$

and by (80), *the random vectors in Eq.'s* (81) *and* (82) *do coincide.*

Finally we observe that, as a consequence, these models satisfy also the following condition: for any permutation $(j_1, \ldots, j_r) \in \Pi([r])$,

$$\mathbb{P}\left(T_{k:r} > t | T_{j_1} < T_{j_2} < \cdots < T_{j_r}\right) = \mathbb{P}\left(T_{k:r} > t\right) \tag{83}$$

Indeed, in the case of ODTHLS models, we have just seen that condition (80) *implies an even stronger property:*

$$\mathbb{P}\left(T_{k:r} > t_k, k = 1, 2, .., r | T_{j_1} < T_{j_2} < \cdots < T_{j_r}\right) = \mathbb{P}\left(T_{k:r} > t_k, k = 1, 2, \ldots r\right). \tag{84}$$

Condition (83) *emerges in a natural way even in more general settings beyond load-sharing, as pointed out in Navarro et al. [19], where it has been referred to as a condition of* weak exchangeability *(see also Navarro et al. [18]).*
In the frame of load-sharing models, condition (80), *(and therefore also* (84) *and* (83)*) emerges in De Santis and Spizzichino [8], where it plays an important role for the special type of problems studied therein.*

Acknowledgements: The authors thank two anonymous Reviewers for several useful comments, which led to a significant improvement of the presentation.
Giovanna Nappo and Fabio Spizzichino acknowledge partial support of Ateneo Sapienza Project 2017 *Modelli stocastici con applicazioni nelle Scienze e nell'Ingegneria*. Rachele Foschi acknowledges partial support of PRIN 2017 Project *How good is your model? Empirical evaluation and validation of quantitative models in economics*.

Conflict of interest statement: Authors state no conflict of interest.

References

[1] Arjas, E. (1981). The failure and hazard processes in multivariate reliability systems. *Math. Oper. Res. 6*(4), 551–562.

[2] Arjas, E. and I. Norros (1991). Stochastic order and martingale dynamics in multivariate life length models: a review. In K. Mosler and M. Scarsini (Eds.), *Stochastic Orders and Decision Under Risk*, pp. 7–24. Inst. Math. Statist., Hayward CA.

[3] Brémaud, P. (1981). *Point Processes and Queues: Martingale Dynamics*. Springer, New York.

[4] Cramer, E. and U. Kamps (2003). Marginal distributions of sequential and generalized order statistics. *Metrika 58*(3), 293–310.

[5] David, H. A. and H. N. Nagaraja (2003). *Order Statistics*. Third edition. Wiley, Hoboken NJ.

[6] de Finetti, B. (1937). La prévision: ses lois logiques, ses sources subjectives. *Ann. Inst. H. Poincaré 7*(1), 1–68.

[7] De Santis, E., Y. Malinovsky, and F. Spizzichino (2021). Stochastic precedence and minima among dependent variables. *Methodol. Comput. Appl. Probab. 23*, 187–205.

[8] De Santis, E. and F. Spizzichino (2021). Construction of aggregation paradoxes through load-sharing dependence models. Available at https://arxiv.org/abs/2004.09417v2.

[9] Durante, F. and C. Sempi (2016). *Principles of Copula Theory*. CRC Press, Boca Raton FL.
[10] Erdely, A., J. M. González-Barrios, and M. M. Hernández-Cedillo (2014). Frank's condition for multivariate Archimedean copulas. *Fuzzy Set. Syst. 240*, 131–136.
[11] Galambos, J. (1982). The role of exchangeability in the theory of order statistics. In *Exchangeability in Probability and Statistics (Rome, 1981)*, pp. 75–86. North-Holland, Amsterdam.
[12] Jaworski, P. (2009). On copulas and their diagonals. *Inform. Sci. 179*(17), 2863–2871.
[13] Jaworski, P. and T. Rychlik (2008). On distributions of order statistics for absolutely continuous copulas with applications to reliability. *Kybernetika 44*(6), 757–776.
[14] Kamps, U. (1995). A concept of generalized order statistics. *J. Statist. Plann. Inference 48*(1), 1–23.
[15] Marichal, J.-L., P. Mathonet, and T. Waldhauser (2011). On signature-based expressions of system reliability. *J. Multivariate Anal. 102*(10), 1410–1416.
[16] McNeil, A. J. and J. Nešlehová (2009). Multivariate Archimedean copulas, d-monotone functions and l_1-norm symmetric distributions. *Ann. Statist. 37*(5B), 3059–3097.
[17] Navarro, J. and J. Fernández-Sánchez (2020). On the extension of signature-based representations for coherent systems with dependent non-exchangeable components. *J. Appl. Probab. 57*(2), 429–440.
[18] Navarro, J., T. Rychlik, and F. Spizzichino (2021). Conditions on marginals and copula of component lifetimes for signature representation of system lifetime. *Fuzzy Set. Syst. 415*, 99–117.
[19] Navarro, J., F. J. Samaniego, N. Balakrishnan, and D. Bhattacharya (2008). On the application and extension of system signatures in engineering reliability. *Naval Res. Logist. 55*(4), 313–327.
[20] Nelsen, R. B. (2006). *An Introduction to Copulas*. Second edition. Springer, New York.
[21] Okolewski, A. (2017). Distribution bounds for order statistics when each k-tuple has the same piecewise uniform copula. *Statistics 51*(5), 969–987.
[22] Rychlik, T. (2010). Copulae in reliability theory (order statistics, coherent systems). In P. Jaworski, F. Durante, W. K. Härdle, T. Rychlik (Eds.), *Copula Theory and its Applications*, pp. 187–208. Springer, Heidelberg.
[23] Rychlik, T. and F. Spizzichino (2021). Load-sharing reliability models with different component sensitivities to other components' working states. *Adv. Appl. Probab. 53*(1), 107–132.
[24] Samaniego, F. J. (2007). *System Signatures and their Applications in Engineering Reliability*. Springer, New York.
[25] Schweizer, B. and A. Sklar (1983). *Probabilistic Metric Spaces*. North-Holland, Amsterdam.
[26] Shaked, M. and J. G. Shanthikumar (1990). Multivariate stochastic orderings and positive dependence in reliability theory. *Math. Oper. Res. 15*(3), 545–552.
[27] Shaked, M. and J. G. Shanthikumar (2007). *Stochastic Orders*. Springer, New York.
[28] Shaked, M. and J. G. Shanthikumar (2015). Multivariate conditional hazard rate functions–an overview. *Appl. Stoch. Models Bus. Ind. 31*(3), 285–296.
[29] Spizzichino, F. (2001). *Subjective Probability Models for Lifetimes*. Chapman & Hall/CRC, Boca Raton FL.
[30] Spizzichino, F. L. (2019). Reliability, signature, and relative quality functions of systems under time-homogeneous load-sharing models. *Appl. Stoch. Models Bus. Ind. 35*(2), 158–176.