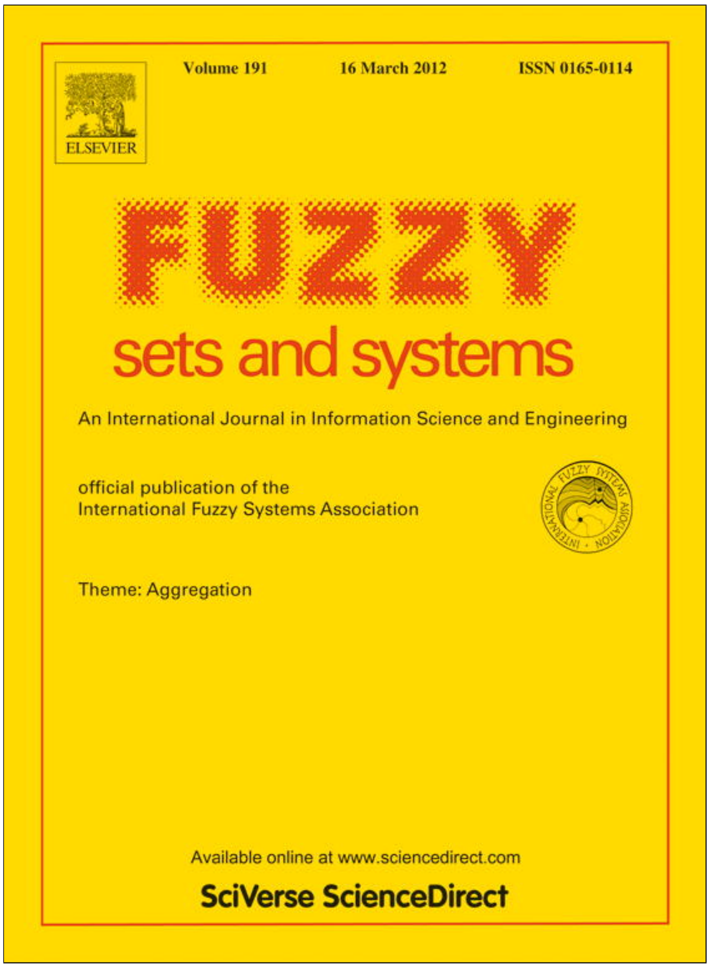
Volume 191
16 March 2012
ISSN 0165-0114
ELSEVIER
FUZZY
sets and systems
An International Journal in Information Science and Engineering
official publication of the
International Fuzzy Systems Association
Theme: Aggregation
Available online at www.sciencedirect.com
SciVerse ScienceDirect

Available online at www.sciencedirect.com

ScienceDirect

Fuzzy Sets and Systems 191 (2012) 103 – 112

FUZZY sets and systems

www.elsevier.com/locate/fss

Characterization of all copulas associated with non-continuous random variables

E. de Amo[a,*], M. Díaz Carrillo[b], J. Fernández-Sánchez[a]

[a] *University of Almería, Campus de La Cañada, 04120-Almería, Spain*
[b] *University of Granada, Campus Fuentenueva, 18071-Granada, Spain*

Received 15 June 2011; received in revised form 8 September 2011; accepted 4 October 2011
Available online 12 October 2011

Abstract

We introduce a constructive method, by means of a doubly stochastic measure, to describe all the copulas that, in view of Sklar's Theorem, are able to connect a bivariate distribution to its marginals. We use this to give the lower and upper optimal bounds for all the copulas that extend a given subcopula.

Keywords: Copula; Subcopula

1. Introduction

For any integer $n \geq 2$, a *multivariate* (or *n-dimensional*) *copula* is the restriction to the unit n-cube $[0, 1]^n$ of a multivariate cumulative distribution function whose marginals are uniform on $[0, 1]$. Copulas were introduced by Sklar in 1959 (see [1]), as the answer to a question posed by Fréchet, and they allow us to represent a joint distribution of random variables as a function of its marginal distributions. In fact, Sklar enunciated that if H is the joint distribution function of n random variables $X_1, \ldots, X_n$, and $F_1, \ldots, F_n$ are the distribution functions of $X_1, \ldots, X_n$, respectively, then there exists a multivariate copula C such that

$$H(x_1, \ldots, x_n) = C(F_1(x_1), \ldots, F_n(x_n))$$

for all $x_1, \ldots, x_n \in \mathbb{R}^n$. This C is uniquely determined on $\mathrm{Ran}(F_1) \times \cdots \times \mathrm{Ran}(F_n)$.

Nowadays, this result is known as Sklar's Theorem, and it has been one of the main tools in promoting the Theory of Copulas as one of the most up-to-date areas in Mathematics. The first proof of this theorem (in the bivariate case) was published in 1974 by Schweizer and Sklar [2]. (See [3], as well.) New proofs have been given since: [4–8], among others.

* Corresponding author. Tel.: +34 950015278; fax: +34 950015480.
E-mail addresses: edeamo@ual.es, enrideamo@gmail.com (E. de Amo), madiaz@ugr.es (M. Díaz Carrillo), juanfernandez@ual.es (J. Fernández-Sánchez).

0165-0114/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.fss.2011.10.005

The wide variety of different proofs of Sklar's Theorem are based on techniques that range from those which are purely probabilistic to others which are more analytic. In [2] the proof consists of a construction of the copula with the desired properties, and the method used is the extension from a subcopula. However, in the other cited papers, the authors only show the existence of, at least, one copula satisfying Sklar's Theorem.

In this paper we consider the bivariate case and, following the way of [2], we describe a method for finding all the copulas C that can be associated with a pair of random variables (Theorems 4 and 8).

The method we use, which we name the E-process, consists of finding suitable doubly stochastic measures in order to obtain these copulas C (Proposition 5). The procedure to obtain C is constructive and it is based on patchwork techniques. Several examples illustrate how our results can be applied to building copulas. To be specific, we obtain, as an application, the lower and upper bounds of copulas that extend a given subcopula, and that are copulas, as well (Theorems 11 and 12).

2. Preliminaries

Let $\mathbb{I} := [0, 1]$ be the closed unit interval and let $\mathbb{I}^2 := [0, 1]^2$ be the unit square. We use $\overline{A}$ to denote the closure of $A \subset \mathbb{I}$. For given sets A and B, we denote by A^B the Cartesian product of elements of A indexed in B, that is, the set of maps from B to A.

First, we give the definitions of subcopula and copula, and some of their elementary properties. For an overview, see for instance [3] or [9].

Definition 1. A bivariate subcopula (or a subcopula, for brevity) is a function $C^* : S_1 \times S_2 \longrightarrow \mathbb{I}$, where S_1 and S_2 are subsets of $\mathbb{I}$ containing 0 and 1, which satisfies the following:

1. $C^*(u, 0) = 0 = C^*(0, v)$, for all $u \in S_1, v \in S_2$;
2. C^* has uniform marginals, i.e. $C^*(u, 1) = u$, $C^*(1, v) = v$, for all $u \in S_1, v \in S_2$;
3. C^* is 2-increasing, i.e. C^*-volume V_{C^*} satisfies $V_{C^*}([u_1, u_2] \times [v_1, v_2]) = C^*(u_2, v_2) - C^*(u_2, v_1) - C^*(u_1, v_2) + C^*(u_1, v_1) \geq 0$, for all $u_1, u_2 \in S_1, v_1, v_2 \in S_2$.

A bivariate copula (or a copula, for brevity) is a subcopula C whose domain is $\mathbb{I}^2$. We denote by $\mathcal{C}$ the class of all copulas.

Well-known examples of copulas are the Fréchet–Hoeffding bounds $M(x, y) = \min\{x, y\}$, $W(x, y) = \max\{0, x + y - 1\}$, and the independence copula $\Pi(x, y) = xy$.

Each copula C induces a probability measure μ_C on $\mathbb{I}^2$ via the formula:

$$\mu_C([a, b] \times [c, d]) = V_C([a, b] \times [c, d]) = C(b, d) - C(b, c) - C(a, d) + C(a, c)$$

and, through standard measure-theoretic techniques, μ_C can be extended from the semi-ring of rectangles in $\mathbb{I}^2$ to the σ-algebra $\mathcal{B}(\mathbb{I}^2)$ of the Borel sets.

Therefore, we remark that there is a one-to-one correspondence between copulas and doubly stochastic measures defined in $\mathbb{I}^2$, that is, probability measures μ, such that for any measurable subset A of $\mathbb{I}$:

$$\mu(A \times \mathbb{I}) = \mu(\mathbb{I} \times A) = \lambda(A),$$

where λ denotes the standard Lebesgue measure on $\mathcal{B}(\mathbb{I})$.

Note that any distribution function H has an associated probability that we will denote by μ_H.

Finally, if C is a copula and $a \in \mathbb{I}$, then the functions $t \to C(t, a)$ (the *horizontal section* of C at a), and $t \to C(a, t)$ (the *vertical section* of C at a) are nondecreasing and 1-Lipschitz on $\mathbb{I}$, i.e. $|C(t_1, a) - C(t_2, a)| \leq |t_1 - t_2|$, for all $t_1, t_2 \in \mathbb{I}$.

Let us recall that Sklar's Theorem represents a bivariate distribution function H by means of the marginal distribution functions F and G, and the copula C. Both of them are connected by Eq. (1) below. Formally, we have the following:

If H is a joint distribution function in $[-\infty, +\infty]^2$ with marginals F and G in $[-\infty, +\infty]$, then there exists a copula C such that the following equation holds:

$$H(x, y) = C(F(x), G(y)), \quad \text{for all } x, y \in [-\infty, +\infty]. \tag{1}$$

If F and G are continuous, then C is unique; otherwise, C is uniquely determined on $\mathrm{Ran}(F) \times \mathrm{Ran}(G)$.

As can be seen in [10], the existing relation between C and H can corresponds to a wide variety of cases depending on the marginals F and G.

Besides, the reverse of (1) can be easily verified, that is:

Lemma 2. *If C is a copula and F and G are distribution functions, then the function H defined by* (1), *is a joint distribution function with marginals F and G.*

To prove Sklar's Theorem, one may define a subcopula in $\mathrm{Ran}(F) \times \mathrm{Ran}(G)$ by the equation given in (1) and, afterwards, extend it to its closure. For further considerations, it is convenient to rewrite this in the following way (see for example [3, Lemma 2.3.4]):

Lemma 3. *Let H be a joint distribution function with marginals F and G. Then, there exists a unique function*:

$$C^* : \overline{\mathrm{Ran}(F)} \times \overline{\mathrm{Ran}(G)} \longrightarrow \mathbb{I},$$

such that $C^*(F(x), G(y)) = H(x, y)$, for all x, y.

It is easy to verify that, when the restriction of a copula C_1 to $\overline{\mathrm{Ran}}(F) \times \overline{\mathrm{Ran}}(G)$ coincides with C^*, then $C_1(F(x), G(y)) = H(x, y)$, for all x, y.

3. The main result

The main result presented in this paper (Theorem 4) allows us to express all the copulas that extend a given subcopula. Equivalently, if X_1 and X_2 are random variables (non-necessarily continuous), with a joint distribution function H and marginals F and G, this result describes all the copulas that can represent H as a function of F and G. First, in order to make this statement, we introduce some notation.

Here and in what follows, we consider a bivariate joint distribution function $H : [-\infty, +\infty]^2 \longrightarrow \mathbb{I}$, with univariate marginals F and G.

For the distribution function $F : [-\infty, +\infty] \longrightarrow \mathbb{I}$, there exists an associated family S_1 of closed subintervals in $\mathbb{I}$, such that their pairwise intersections are empty. To check this, observe that the elements A of the projection of the graph of F on $\mathbb{I}$ are either an interval or a singleton. Let S_1 be the family constituted by the closures $\overline{A}$.

Now, we consider the class P_1 of elements in S_1 which are singletons, and set $D_1 := S_1 \backslash P_1$.

The complement in $\mathbb{I}$ of the union of elements of S_1 is a family of open intervals. We will denote by O_1 the class of all the closures of these (open) intervals. Finally, with $\mathcal{T}$ a index set, write $T := \{T_t = [a_t, b_t]; T_t \in D_1 \cup O_1\}_{t \in \mathcal{T}}$.

Similarly, for the distribution function G, there exist the corresponding sets S_2, P_2, D_2, O_2, and $J := \{J_j = [c_j, d_j]; J_j \in D_2 \cup O_2\}_{j \in \mathcal{J}}$, with $\mathcal{J}$ an index set.

Next, let us define auxiliary functions associated to the elements in the class O_1. In fact, for any $T_t \in O_1$ we select a family of distribution functions whose restriction to $\mathbb{I}$, $F_{tj} : \mathbb{I} \to \mathbb{I}$, satisfies

$$x = \frac{1}{b_t - a_t} \sum_j \beta_{tj} F_{tj}(x), \quad \forall x \in \mathbb{I}, \tag{2}$$

where

$$\beta_{tj} = C^*(b_t, d_j) + C^*(a_t, c_j) - C^*(b_t, c_j) - C^*(a_t, d_j).$$

Because $\sum_j \beta_{tj} = b_t - a_t$, let us note that it is possible to find functions F_{tj} satisfying (2). The easiest way to obtain this is setting $F_{tj}(x) = x$.

We proceed in a similar way to obtain functions $G_{tj} : \mathbb{I} \to \mathbb{I}$ that are associated to sets $J_j \in O_2$; here

$$x = \frac{1}{d_j - c_j} \sum_t \beta_{tj} G_{tj}(x). \tag{3}$$

There exist other auxiliary functions that are associated to rectangles in the form $T_t \times J_j \in D_1 \times O_2$ or $T_t \times J_j \in O_1 \times D_2$ if $\beta_{tj} \neq 0$.

In the case of $T_t \times J_j \in D_1 \times O_2$, we consider the distribution functions in the following way:

$$F_{tj}(x) = \frac{1}{\beta_{tj}}(C^*((b_t - a_t)x + a_t, d_j) + C^*(a_t, c_j) - (C^*(a_t, d_j) + C^*((b_t - a_t)x + a_t, c_j))). \tag{4}$$

If $T_t \times J_j \in O_1 \times D_2$, then we consider:

$$G_{tj}(x) = \frac{1}{\beta_{tj}}(C^*(b_t, (d_j - c_j)y + c_j) + C^*(a_t, c_j) - (C^*(a_t, (d_j - c_j)y + c_j) + C^*(b_t, c_j))). \tag{5}$$

With the above notations the main result of this paper can be presented as follows.

Theorem 4. *Let H be a bivariate distribution function in $[-\infty, +\infty]^2$, with given marginals F and G. Then, C is a copula satisfying the equation*:

$$C(F(x), G(y)) = H(x, y),$$

if and only if C can be expressed in the form:

$$C(x, y) = C^*(x, y), \quad if \ (x, y) \in \overline{\text{Ran}}(F) \times \overline{\text{Ran}}(G)$$

and

$$\begin{aligned} C(x, y) = {} & C^*(a_t, c_j) + \beta_{tj} C_{tj}\left(F_{tj}\left(\frac{x - a_t}{b_t - a_t}\right), G_{tj}\left(\frac{y - c_j}{d_j - c_j}\right)\right) \\ & + \sum_{t' \in S_t} \beta_{t'j} G_{t'j}\left(\frac{y - c_j}{d_j - c_j}\right) + \sum_{j' \in Z_j} \beta_{tj'} F_{tj'}\left(\frac{x - a_t}{b_t - a_t}\right), \end{aligned} \tag{6}$$

if $(x, y) \notin \overline{\text{Ran}}(F) \times \overline{\text{Ran}}(G)$ and $(x, y) \in T_t \times J_j$, where $C_{tj} \in \mathcal{C}$, F_{tj} and G_{tj} are distribution functions satisfying (2)–(5), *with $S_t = \{t' : a_{t'} < a_t\}$ and $Z_j = \{j' : c_{j'} < c_j\}$.*

3.1. The E-process

To show Theorem 4 we construct a measure μ on rectangles $T_t \times J_j$, $(t, j) \in \mathcal{T} \times \mathcal{J}$, and we prove that it is a doubly stochastic measure. This extension method will be called extension process (for short *E*-process).

We remark that, by Lemma 3, if H is a joint distribution function with marginals F and G, then there exists a unique function $C^* : \overline{\text{Ran}}(F) \times \overline{\text{Ran}}(G) \longrightarrow \mathbb{I}$, such that $C^*(F(x), G(y)) = H(x, y)$, for all $x, y \in [-\infty, +\infty]$.

Proposition 5. *Let H be a joint distribution function with marginals F and G. Then, there exists a doubly stochastic measure μ such that the restriction of its associated copula C_1 to $\overline{\text{Ran}}(F) \times \overline{\text{Ran}}(G)$ coincides with C^*.*

Proof. The method to produce the measure μ will be developed in three steps:

1. The construction of the continuous functions F_{tj} (resp. G_{tj}) associated to sets $T_t \in O_1$ (resp. O_2): If $T_t \in O_1$, then it is possible to choose a family of functions F_{tj} that satisfies (2), and functions G_{tj} associated to sets $J_j \in O_2$ satisfying (3).
2. Measure allocation using the joint distribution function with marginals F_{tj} and G_{tj}.
   There are four cases to consider:

   (a) $T_t \times J_j \in D_1 \times D_2$. If $[a, b] \times [c, d] \subseteq T_t \times J_j$, then

   $$\mu([a, b] \times [c, d]) = C^*(b, d) + C^*(a, c) - C^*(b, c) - C^*(a, d).$$

   And, the extension theorem allows us to extend the measure μ to every Borel set in the rectangle $D_1 \times D_2$.

(b) $T_t \times J_j \in D_1 \times O_2$. If $\beta_{tj} = 0$, then the measure of every Borel set in the rectangle is zero. On the other hand, if $\beta_{tj} \neq 0$, we consider the distribution function F_{ij} given by (4) and G_{tj} satisfying (3). It follows (from Lemma 2) that for F_{tj} and G_{tj}, and for any copula (which we denote by C_{tj}), a distribution function

$$H_{tj}(x, y) = C_{tj}(F_{tj}(x), G_{tj}(x))$$

exists. Now, the map $Q_{tj} : \mathbb{I}^2 \longrightarrow T_t \times J_j$ given by

$$Q_{tj}((x, y)) = ((b_t - a_t)x + a_t, (d_j - c_j)y + c_j),$$

allows us to move the mass distribution determined by H_{tj}, from $\mathbb{I}^2$ to $T_t \times J_j$. Hence, for each Borel set $A \subset T_t \times J_j$, the value of $\mu(A)$ is $\beta_{ij}\mu_{H_{tj}}(Q_{tj}^{-1}(A))$.

(c) $T_t \times J_j \in O_1 \times D_2$. It is analogous to (b).

(d) $T_t \times J_j \in O_1 \times O_2$. We proceed in a similar way as we did in (b). Here, the functions F_{tj} and G_{tj} were previously fixed (see (2) and (3)).

3. The probability measure μ is doubly stochastic:
We will restrict our attention to checking that $\mu([a, b] \times \mathbb{I}) = b - a$ in the case when $[a, b] \subset T_t$, for some $t \in \mathcal{T}$. We need to consider two subcases here.
If $T_t \in D_1$, then

$$\mu([a, b] \times \mathbb{I}) = \sum_j C^*(b, d_j) - C^*(a, d_j) - C^*(b, c_j) + C^*(a, c_j)$$
$$= C^*(b, 1) - C^*(a, 1) = b - a.$$

If $T_t \in O_1$, then

$$\mu([a, b] \times \mathbb{I}) = \frac{1}{b_t - a_t} \sum_j \beta_{tj}(F_{tj}(b) - F_{tj}(a)) = b - a.$$

Similar arguments apply to the case $\mathbb{I} \times [a, b]$.
Finally, this measure μ has an associated copula C_1. Let us note that, by construction, the restriction of C_1 to $\overline{\mathrm{Ran}}(F) \times \overline{\mathrm{Ran}}(G)$ coincides with C^*. Moreover, by Lemma 3, it follows that $C_1(F(x), G(y)) = H(x, y)$, which fulfils the statement. □

Note that, for a given distribution function H, the method in the Proposition 6 is associated with the copulas C_{tj} and the distribution functions F_{tj} and G_{tj}.

We conclude this subsection with three remarks.

1. If we set $C_{tj} = \Pi$ (that is, the independence or product copula), and $F_{tj}(x) = G_{tj}(x) = x$ in (2) and (3), then this is precisely the particular case given by Sklar and Schweizer in [2] for the proof of Sklar's Theorem.
2. Let us consider a copula C and a family $\{S_i\}_{i \in \mathfrak{I}}$ of closed and connected subsets of $\mathbb{I}^2$, with boundaries ∂S_i such that $S_i \cap S_j \subseteq \partial S_i \cap \partial S_j$ whenever $i \neq j$. Moreover, for every $i \in \mathfrak{I}$, let us consider an increasing continuous mapping $L_i : S_i \longrightarrow \mathbb{I}$, such that $C = L_i$ on ∂S_i. Then, the function $L : \mathbb{I}^2 \longrightarrow \mathbb{I}$ defined by

$$L(x, y) = \begin{cases} L_i(x, y), & (x, y) \in S_i, \\ C(x, y) & \text{otherwise,} \end{cases}$$

it said to be the *patchwork* (of $\{S_i\}_{i \in \mathfrak{I}}$) into the copula C. When sets S_i are rectangles, it is called a *rectangular patchwork* (see [11]).

According to the above definition, we can check the following result:

Proposition 6. *Let us denote by C_1 and C_2 two copulas obtained by two respective E-processes where the same distribution functions F_{tj} and G_{tj} have been considered. Then C_1 is a rectangular patchwork into C_2 (and vice versa).*

3. Following [12, Theorem 2], let $(]u_z, u'_z[)_{z \in Z}$ and $(]v_k, v'_k[)_{k \in K}$ be two families of nonempty, pairwise disjoint open subintervals of $]0, 1[$. Consider a copula C^b, called the *background copula*, a family $(C^f_{z,k})_{z \in Z, k \in K}$ of copulas,

called *foreground copulas*, and a family $(\lambda(u_z, u'_z, v_k, v'_k))_{z\in Z,k\in K}$ of positive multipliers. For any $z \in Z$ and $k \in K$, define the mapping $P^b_{z,k} : [u_z, u'_z] \times [v_k, v'_k] \to \mathbb{R}$ by

$$P^b_{z,k}(x, y) = C^b(x, y) - \lambda(u_z, u'_z, v_k, v'_k)C^b\left(\frac{x-u_z}{u'_z-u_z}, \frac{y-v_k}{v'_k-v_k}\right)$$

and the binary operation Q by

$$Q(x, y) = \begin{cases} P^b_{z,k}(x, y) + \lambda(u_z, u'_z, v_k, v'_k)C^f_{z,k}\left(\frac{x-u_z}{u'_z-u_z}, \frac{y-v_k}{v'_k-v_k}\right) & \text{if } (x, y) \in [u_z, u'_z] \times [v_k, v'_k], \\ C^b(x, y) & \text{otherwise.} \end{cases}$$

If for all $z \in Z$ and $k \in K$ it holds that $P^b_{z,k}$ is 2-increasing on $[u_z, u'_z] \times [v_k, v'_k]$, then Q is a copula.

Now, according to the above result, we can check the following result:

Proposition 7. *Let $C = C^b$ be a copula extending a subcopula C^*. If we choose the intervals $[u_z, u'_z]$ as elements in $O_1 \cup D_1$, intervals $[v_k, v'_k]$ as elements in $O_2 \cup D_2$, and multipliers $\lambda(u_z, u'_z, v_k, v'_k)$ such that $\lambda(u_z, u'_z, v_k, v'_k) = 0$ when $[u_z, u'_z] \in D_1$ and $[v_k, v'_k] \in D_2$, and such that $P^b_{z,k}$ is 2-increasing on $[u_z, u'_z] \times [v_k, v'_k]$, then Q is another copula that extends C^*.*

3.2. Description of all copulas associated with a pair of random variables

As we already mentioned in Introduction, the E-process given in this paper allows us to describe all the copulas that can be associated with a given distribution function H. In order to attain this goal, it is appropriate to introduce additional notation.

Let D'_1 denote the set of indices $t \in \mathcal{T}$ such that $T_t \in D_1$, and, in a similar way, we introduce the sets D'_2, O'_1 and O'_2. Let us denote by K the subset in $D'_1 \times O'_2 \cup O'_1 \times D'_2 \cup O'_1 \times O'_2$ of all indices such that $\beta_{tj} \neq 0$. We shall denote by Δ the class of all distribution functions, and let Δ_1 denote the subclass in $\Delta^{O'_1\times\mathcal{J}\cap K}$ such that $x = (1/(b_t - a_t))\sum_{j,(t,j)\in O'_1\times\mathcal{J}} \beta_{tj}F_{tj}(x)$. The subclass Δ_2 is defined analogously.

If we analyze the proof of Proposition 5, then we observe that, in fact, we have proved more than it is said in the statement. Precisely, for each element in $\mathcal{C}^K \times \Delta_1 \times \Delta_2$, its associated E-process gives rise a copula C such that $C(F(x), G(y)) = H(x, y)$.

On the other hand, for any copula C that extends a subcopula C^*, we can check that, if $\beta_{tj} \neq 0$ and we set

$$F_{tj}(x) = \frac{1}{\beta_{tj}}(C((b_t - a_t)x + a_t, d_j) + C(a_t, c_j) - (C(a_t, d_j) + C((b_t - a_t)x + a_t, c_j)))$$

and

$$G_{tj}(y) = \frac{1}{\beta_{tj}}(C((d_j - c_j)y + c_j, b_t) + C(a_t, c_j) - (C((d_j - c_j)y + c_j, a_t) + C(b_t, c_j))),$$

then the copula C is obtained as an E-process.

We are now in a position to prove our main result concerning the representation of copulas.

Theorem 8. *Let H be a bivariate distribution function with marginals F and G, and a subcopula C^* defined in $\overline{\mathrm{Ran}}(F) \times \overline{\mathrm{Ran}}(G)$ satisfying that $C^*(F(x), G(y)) = H(x, y)$. Then, the following statements hold*:

(a) *For each element in class $\mathcal{C}^K \times \Delta_1 \times \Delta_2$, the associated E-process gives rise to a copula C satisfying that $C(F(x), G(y)) = H(x, y)$. Moreover, the E-process is injective in the sense that for different elements in $\mathcal{C}^K \times \Delta_1 \times \Delta_2$, the corresponding copulas are different.*
(b) *Every copula C satisfying the equation $C(F(x), G(y)) = H(x, y)$, is generated by an E-process.*

We can now prove Theorem 4 as follows.

Proof of Theorem 4. By Theorem 8, because the copulas satisfying $C(F(x), G(y)) = H(x, y) = C^*(F(x), G(y))$ can be obtained by an E-process, if the copula C_{tj}, and the distribution functions F_{tj} and G_{tj} are those used in it, then C is the copula given in (6).

On the other hand, the constructing method and conditions (2)–(5) show that C is a copula. □

The two following examples illustrate Theorem 8.

Example 9. If $\text{Ran}(F) = \mathbb{I}$ and $\text{Ran}(G) = \{0, b, 1\}$ $(0 < b < 1)$, then the subcopula C^* can be essentially identified with the horizontal section $C^*(x, b) = h_b(x)$, where, the function $h_b : \mathbb{I} \longrightarrow [0, b]$ is an increasing bijection satisfying $|h_b(x) - h_b(y) \leq |x - y|$ (that is, it is 1-Lipschitz).

Conversely, if $h_b : \mathbb{I} \longrightarrow [0, b]$ is an increasing 1-Lipschitz bijection, then the map $C^* : \mathbb{I} \times \{0, b, 1\} \longrightarrow \mathbb{I}$ given by

$$C^*(x, y) = \begin{cases} 0 & \text{if } y = 0, \\ h_b(x) & \text{if } y = b, \\ x & \text{if } y = 1, \end{cases}$$

is a subcopula.

Here $T = \{T_0 = \mathbb{I}\}$, $J = \{J_0 = [0, b], J_1 = [b, 1]\}$, and $F_{0,0} = h_b(x)/b$, $F_{0,1} = (x - h_b(x))/(1 - b)$, $G_{0,0}(x) = G_{0,1}(x) = x$.

It is of interest to note that the above result shows a general expression for all the copulas with a horizontal section h_b. Precisely, it is given by

$$C(x, y) = \begin{cases} bC_1\left(\dfrac{h_b(x)}{b}, \dfrac{y}{b}\right) & \text{if } x \leq b, \\ (1 - b)C_2\left(\dfrac{x - h_b(x)}{1 - b}, \dfrac{y - b}{1 - b}\right) + h_b(x) & \text{otherwise,} \end{cases}$$

where C_1 and C_2 are copulas.

Note that we prove this fact by a different procedure from that in [11].

Finally, if we consider the Fréchet–Hoeffding boundary copulas in the following cases:

(a) $C_1(x, y) = C_2(x, y) = M(x, y) = \min\{x, y\}$, or
(b) $C_1(x, y) = C_2(x, y) = W(x, y) = \max\{0, x + y - 1\}$,
then, we obtain the upper and lower copulas of this family, respectively. This result corresponds to Theorems 3.1 and 3.2 in [13].

Example 10. If $\text{Ran}(F) = \{0, a, 1\}$ and $\text{Ran}(G) = \{0, b, 1\} (0 < a, b < 1)$, then the subcopula C^* only takes a non-trivial value $C^*(a, b) = \theta$, with θ such that $\max\{0, x + y - 1\} \leq \theta \leq \min\{x, y\}$.

A copula C extending C^* has the form:

$$C(x, y) = \begin{cases} \theta C_{00}\left(F_0\left(\dfrac{x}{a}\right), G_0\left(\dfrac{y}{b}\right)\right) & \text{if } (x, y) \in [0, a] \times [0, b] \\ (b - \theta)C_{10}\left(F_1\left(\dfrac{x - a}{1 - a}\right), G_0^+\left(\dfrac{y}{b}\right)\right) + \theta F_0\left(\dfrac{x}{a}\right) & \text{if } (x, y) \in [a, 1] \times [0, b] \\ (a - \theta)C_{01}\left(F_0^+\left(\dfrac{x}{a}\right), G_1\left(\dfrac{y - b}{1 - b}\right)\right) + \theta G_0\left(\dfrac{x}{a}\right) & \text{if } (x, y) \in [0, a] \times [b, 1] \\ \theta + (1 + \theta - a - b)C_{11}\left(F_1^+\left(\dfrac{x - a}{1 - a}\right), G_1^+\left(\dfrac{y - b}{1 - b}\right)\right) & \\ \quad +(b - \theta)F_1\left(\dfrac{x - a}{1 - a}\right) + (a - \theta)G_1\left(\dfrac{y - b}{1 - b}\right) & \text{if } (x, y) \in [a, 1] \times [b, 1] \end{cases},$$

where F_0 is an a/θ-Lipschitz distribution function, and F_1, G_0, and G_1 are, respectively, as well of parameters $(1-a)/(b-\theta)$, a/θ, and $(1-b)/(a-\theta)$. The function F_0^+ is determined by F_0 in the form $F_0^+(x) = (ax - \theta F_0(x))/(a-\theta)$, and in a similar way F_1^+, G_0^+, and G_1^+.

This example corresponds to the case of two random variables that follow Bernoulli distributions. In the case of independent variables, the expression becomes easier writing ab instead of θ.

4. Upper and lower bounds

For a given subcopula C^*, it is interesting to know the functions:

$$UC^*(x,y) = \sup\{C(x,y) : C \text{ is a copula that extends to } C^*\}$$

and

$$LC^*(x,y) = \inf\{C(x,y) : C \text{ is a copula that extends to } C^*\}.$$

We shall see that they are copulas, and we provide a method to produce them.

In fact, any interval T_t is divided into indexed subintervals (in $\mathcal{J}$) in such a way that the interval $T_t^j = [a_t^j, b_t^j] \subset T_t$ is an interval of length $V_{C^*}(T_t \times J_j)$, and its lower extreme is given by $a_t + \sum_{c_{j'} < c_j} V_{C^*}(T_t \times J_{j'})$. In the same manner, we can divide J_j into indexed subintervals J_j^t.

Now, we can define the function F_{tj}, as:

$$F_{tj}(x) = \begin{cases} 0 & \text{if } 0 \le x \le \dfrac{a_t^j - a_t}{b_t - a_t}, \\[2ex] \dfrac{b_t - a_t}{b_t^j - a_t^j}x + \dfrac{a_t - a_t^j}{b_t^j - a_t^j} & \text{if } \dfrac{a_t^j - a_t}{b_t - a_t} \le x \le \dfrac{b_t^j - a_t}{b_t - a_t}, \\[2ex] 1 & \text{if } \dfrac{b_t^j - a_t}{b_t - a_t} \le x \le 1 \end{cases} \tag{7}$$

and, similarly, for functions G_{tj}.

Theorem 11. *If we choose the functions F_{tj} and G_{tj} in the E-process as above (7), and the copula $C_{tj} = M$, then the associated copula with respect to the doubly stochastic measure is the copula UC^*.*

Proof. First, we can see that if $T_t \times J_j \in O_1 \times O_2$, then the mass distribution is uniformly distributed on the diagonal of the square $T_t^j \times J_j^t$ in this rectangle. If $T_t \times J_j \in D_1 \times O_2$, then the mass is distributed in the graph of an increasing bijection from T_t to J_j^t. If $T_t \times J_j \in O_1 \times D_2$, then the mass is also distributed in the graph of an increasing bijection from T_t^j to J_j.

To show that the copula associated to this measure is UC^*, three cases have to be considered. Here, the associated copula will be denoted by C':

1. If $(x,y) \in \overline{\text{Ran}}(F) \times \overline{\text{Ran}}(G)$, then it is obvious that C' is, in fact, $UC^*(x,y)$.
2. If $(x,y) \in T_t \times J_j \in O_1 \times O_2$, we distinguish some subcases:
   (a) If $x \le a_t^j$, $y \le c_j^t$, then

   $$C'(x,y) = C^*(a_t, c_j) + (x - a_t) + (y - c_j),$$

   and it is the maximum value obtainable;
   (b) If $x \ge b_t^j$, $y \ge d_j^t$, then $C'(x,y) = C^*(b_t, d_j)$, and that it is the maximum value obtainable, as well;
   (c) If $a_t^j \le x \le b_t^j$, $y \le c_j^t$, then $C'(x,y) = C^*(a_t, c_j) + \sum_{j'<j} V_{C^*}(T_t \times J_{j'}) + x - a_t$, and that it is newly the maximum value obtainable;
   (d) The same conclusion can be drawn to the case $x \le b_t^j$, $c_j^t \le y \le d_j^t$.

Similar arguments apply to the other cases:

(e) $x \geq b_t^j$, $y \leq c_j^t$;

(f) $x \leq a_t^j$, $y \geq d_j^t$;

(g) $a_t^j \leq x \leq b_t^j$, $y \geq d_j^t$;

(h) $x \geq b_t^j$, $c_j^t \leq y \leq d_j^t$; or

(i) $a_t^j \leq x \leq b_t^j$, $c_j^t \leq y \leq d_j^t$.

3. If $(x, y) \in T_t \times J_j \in D_1 \times O_2$ or $(x, y) \in T_t \times J_j \in O_1 \times D_2$, then the corresponding subcases can be treated as in case 2.

Observe that we obtain the lower bound in a similar way. In fact, the interval $T_t^j \subset T_t$ has length $V_{C^*}(T_t \times J_j)$, and its upper extreme is $b_t - \sum_{j'<j} V_{C^*}(T_t \times J_{j'})$. The interval J_j is divided into subintervals J_j^t, and the functions F_{tj} and G_{tj} are defined as above.

Theorem 12. *If we choose the functions F_{tj} and G_{tj} in the E-process as above (7), and copula $C_{tj} = W$, then the associated copula with respect to the doubly stochastic measure is the copula LC^*.*

Example 13. The last two theorems include, as a particular case, the result due to Carley in [14] (see [15], as well), when the sets Ran(F) and Ran(G) are finite.

5. Conclusions

We can describe all the elements of the set of copulas that extend a given subcopula (Theorem 4). We have called this technique we have used *E*-process. Furthermore, we describe the upper and lower bounds of this set that are copulas, as well.

If we transfer these ideas to the *n*-dimensional case, then we can obtain *n*-stochastic measures, but it does not seem easy to give an analogous to Theorem 4. Moreover, if $n \geq 3$, the Fréchet–Hoeffding lower bound is not a copula, and as a consequence it would be impossible to find a lower bound for the copulas that extend a given subcopula.

Acknowledgements

The authors are grateful to the support of the Ministerio de Ciencia e Innovación (Spain) under Research Project MTM2011-22394. They are also grateful to the anonymous referees for their useful suggestions that led to improvement of the paper.

References

[1] A. Sklar, Fonctions de répartition à n dimensions et leurs marges, Publ. Inst. Statist. Univ. Paris 8 (1959) 229–231.
[2] B. Schweizer, A. Sklar, Operations on distribution functions not derivable from operations on random variables, Stud. Math. 52 (1974) 43–52.
[3] R.B. Nelsen, An Introduction to Copulas, 2nd ed., Springer Series in Statistics. Springer, New York, 2006.
[4] A. Burchard, H. Hajaiej, Rearrangement inequalities for functionals with monotone integrands, J. Funct. Anal. 233 (2) (2006) 561–582.
[5] H. Carley, M.D. Taylor, A new proof of Sklar's theorem, in: C.M. Cuadras, J. Fostiana, J.A. Rodríguez-Lallena (Eds.), Proceedings of the Conference on Distributions with Given Marginals and Statistical Modelling, Barcelona, 2000, pp. 29–34.
[6] P. Deheuvels, Caractérisation complète des lois extrêmes multivariées et de la convergence des types extrêmes, Publ. Inst. Statist. Univ. Paris 23 (1978) 1–37.
[7] D.S. Moore, M.C. Spruill, Unified large-sample theory of general chi-squared statistics for tests of fit, Ann. Statist. 3 (1975) 599–616.
[8] L. Rüschendorf, On the distributional transform, Sklar's theorem, and the empirical copula process, J. Statist. Plan. Infer. 139 (11) (2009) 3921–3927.
[9] F. Durante, C. Sempi, Copula theory: an introduction, in: P. Jaworski, F. Durante, W. Härdle, T. Rychlik (Eds.), Workshop on Copula Theory and its Applications, Lecture Notes in Statistics – Proceedings, Springer, Dortrecht (NL), 2010.
[10] A.W. Marshall, Copulas, Marginals, and joint distributions. Distributions with fixed marginals and related topics (Seattle, WA, 1993), IMS Lecture Notes Monograph Series 28, Institute of Mathematical Statistics, Hayward, CA, 1996, pp. 213–222.

[11] F. Durante, S. Saminger-Platz, P. Sarkoci, Rectangular patchwork for bivariate copulas and tail dependence, Commun. Statist. Theory Methods 38 (15) (2009) 2515–2527.
[12] B. De Baets, H. De Meyer, Orthogonal grid constructions of copulas, IEEE Trans. Fuzzy Syst. 15 (6) (2007) 1053–1062.
[13] E.P. Klement, A. Kolesárová, R. Mesiar, C. Sempi, Copulas constructed from the horizontal section, Commun. Statist. Theory Methods 36 (2007) 2901–2911.
[14] H. Carley, Maximum and minimum extensions of finite subcopulas, Com. Statist. Theory Methods 31 (12) (2002) 2151–2166.
[15] C. Genest, J. Nešlehová, A primer on copulas for count data, ASTIN Bull. 37 (2007) 475–515.