

Paul Ressel\*

# Copulas, stable tail dependence functions, and multivariate monotonicity

<https://doi.org/10.1515/demo-2019-0013>

Received April 4, 2019; accepted June 11, 2019

**Abstract:** For functions of several variables there exist many notions of monotonicity, three of them being characteristic for resp. distribution, survival and co-survival functions. In each case the “degree” of monotonicity is just the basic one of a whole scale.

Copulas are special distribution functions, and stable tail dependence functions are special co-survival functions. It will turn out that for both classes the basic degree of monotonicity is the only one possible, apart from the (trivial) independence functions. As a consequence a “nesting” of such functions depends on particular circumstances. For nested Archimedean copulas the rather restrictive conditions known so far are considerably weakened.

**Keywords:** Copula, stable tail dependence function, multivariate monotonicity, nested Archimedean copula

**MSC:** 60E05, 62H05, 26A48, 26B40

## 1 Introduction

A copula is a multivariate distribution function (“d.f.”) defined on the unit cube  $[0, 1]^d$ , with uniform marginals. In particular it is fully  $d$ -increasing, a crucial property of  $d$ -dimensional d.f.s., and this is a special case of being  $\mathbf{n}$ -increasing, where  $\mathbf{n} = (n_1, \dots, n_d) \in \mathbf{N}^d$  is a specific index of multivariate monotonicity of a higher order, fully  $d$ -increasing corresponding to  $\mathbf{n} = \mathbf{1}_d := (1, \dots, 1)$ . For a copula it is enough to require it to be just  $d$ -increasing, since a copula is always grounded; cf. Theorem 2 in [11].

Now it has been of considerable interest to insert copulas into another one, the main problem in this connection being to assert the crucial monotonicity condition of this composed function. A general theorem characterizing functions  $f$  on  $[0, 1]^k$  which operate on  $k$  multivariate d.f.s., has been given in [14]. It requires  $f$  to be  $\mathbf{n}$ -increasing, with  $\mathbf{n} = (n_1, \dots, n_k)$  where  $n_i$  is the dimension of the  $i$ 'th d.f. to be inserted. Therefore if a copula was  $\mathbf{n}$ -increasing, other copulas of dimensions  $n_1, \dots, n_k$  could be inserted without further restrictions.

Our first result (Theorem 1) might appear “disappointing”: an  $\mathbf{n}$ -increasing copula with  $\mathbf{n} \geq \mathbf{2}_d := (2, \dots, 2)$  is necessarily the independence copula. Based on this result we shall prove an extension (from  $d = 2$  to any  $d \geq 2$ ) of an “impossibility theorem” due to Genest, Molina and Lallena [3]: a copula which operates on arbitrary bivariate copulas is again necessarily the independence copula (Theorem 2).

In Theorem 3 we give sufficient conditions for so-called *nested Archimedean copulas* to be “bona fide” copulas, with a new proof of a result of Rezapour, and these conditions are far less restrictive than those to be found in the earlier literature.

---

\*Corresponding Author: Paul Ressel: Kath. Universität Eichstätt-Ingolstadt, E-mail: paul.ressel@ku.de

Comparable results are then shown for stable tail dependence functions: if their intrinsic monotonicity property is sharpened, so that other such functions could be inserted, only the trivial independence function remains possible (Theorem 4), and this is also the only one which operates on arbitrary bivariate functions of this class (Theorem 6). For the proof of this last theorem a result of independent interest is needed (Theorem 5), establishing a (non-trivial) connection between  $\mathbf{n}$ -increasing and fully  $(\sum n_i)$ -increasing functions.

In this paper we use the following notations:

$\mathbf{R}_+ := [0, \infty[$ ,  $\mathbf{R}_- := ]-\infty, 0]$ ,  $\mathbf{N} := \{1, 2, 3, \dots\}$ ,  $\mathbf{N}_0 := \{0, 1, 2, \dots\}$   
 $x \odot y := (x_1y_1, x_2y_2, \dots)$  for two vectors of equal length,  
 $\frac{1}{x} := (\frac{1}{x_1}, \dots, \frac{1}{x_d})$  for  $x \in [0, \infty]^d$  with  $\frac{1}{0} := \infty$ ,  $\frac{1}{\infty} := 0$

$$|\mathbf{n}| := \sum_{i=1}^d n_i \text{ for } \mathbf{n} \in \mathbf{N}_0^d, \quad \mathbf{1}_d := (1, \dots, 1) \in \mathbf{N}^d, \quad \mathbf{2}_d := (2, \dots, 2),$$

(however mostly 0 instead of  $\mathbf{0}_d$ )

$$\binom{\mathbf{p}}{\mathbf{q}} := \prod_{i=1}^d \binom{p_i}{q_i} \text{ for } \mathbf{p}, \mathbf{q} \in \mathbf{N}_0^d \text{ (usually } \mathbf{q} \leq \mathbf{p})$$

$$(f \times g)(x, y) := (f(x), g(y)) \text{ for mappings } f, g$$

$$(f \otimes g)(x, y) := f(x) \cdot g(y) \text{ for real-valued } f, g$$

$$\sigma_k(x_1, \dots, x_k) := \sum_{i=1}^k x_i$$

$$\sigma_{\mathbf{n}} := \sigma_{n_1} \times \dots \times \sigma_{n_d} \text{ for } \mathbf{n} = (n_1, \dots, n_d) \in \mathbf{N}^d$$

$M_+^1(X)$  is the set of all Radon probability measures on a (Hausdorff) space  $X$ .

An interval in  $\mathbf{R}$  is any connected subset of  $\mathbf{R}$ ; it is non-degenerate if it contains more than one point.

A  $d$ -dimensional (non-degenerate) interval is a product of  $d$  one-dimensional (non-degenerate) intervals.

Examples are

$$[x, y] = \prod_{i=1}^d [x_j, y_j] \text{ for } x \leq y$$

and similarly  $]x, y]$  etc.

$e_1, \dots, e_d$  are the usual unit vectors in  $\mathbf{R}^d$

d.f. = distribution function

STDF = stable tail dependence function

## 2 Multivariate higher order monotonicity

Let  $I_1, \dots, I_d \subseteq \mathbf{R}$  be non-degenerate intervals,  $I := I_1 \times \dots \times I_d$ , and let  $f : I \rightarrow \mathbf{R}$  be any function. For  $s \in I, h \in \mathbf{R}_+^d$  such that also  $s + h \in I$  put

$$(E_h f)(s) := f(s + h)$$

and  $\Delta_h := E_h - E_0$ , i.e.  $(\Delta_h f)(s) := f(s + h) - f(s)$ , and  $\nabla_h := -\Delta_h \cdot E_h$  is often called a shift operator, and  $\Delta_h$  a difference operator.

Since  $\{E_h\}$  is commutative (where defined), so is  $\{\Delta_h\}$  and  $\{\nabla_h\}$ . In particular, with  $e_1, \dots, e_d$  denoting standard unit vectors in  $\mathbf{R}^d$ ,  $\Delta_{h_1 e_1}, \dots, \Delta_{h_d e_d}$  commute. As usual  $\Delta_h^0 f := f =: \nabla_h^0 f$ . For  $\mathbf{n} = (n_1, \dots, n_d) \in \mathbf{N}_0^d$  and  $h \in \mathbf{R}_+^d$  we put

$$\Delta_h^{\mathbf{n}} := \Delta_{h_1 e_1}^{n_1} \Delta_{h_2 e_2}^{n_2} \dots \Delta_{h_d e_d}^{n_d}$$

and similarly  $\nabla_h^n$ . Here  $\Delta_{h_i e_i}^{n_i}$  is the usual  $n_i$ 'th power of the difference operator  $\Delta_{h_i e_i}$  while  $\Delta_h^n$  is defined as the composition of the “partial” difference operators  $\Delta_{h_i e_i}^{n_i}$ . The multinomial theorem gives

$$(\Delta_h^n f)(s) = \sum_{0 \leq q \leq n} (-1)^{|n|-|q|} \binom{n}{q} f(s + q \odot h).$$

Note that  $\Delta_h^{1^d} \neq \Delta_h$  for  $d > 1$ .

**Definition.** Let  $n \in N_0^d \setminus \{0\}$ . A function  $f : I \rightarrow \mathbf{R}$  is called

(i)  **$n$ -increasing** (“ $n$ - $\uparrow$ ”) if

$$(\Delta_h^p f)(s) \geq 0$$

for all  $s \in I, h \in \mathbf{R}_+^d, p \in N_0^d, 0 \neq p \leq n$  such that  $s + p \odot h \in I$ .

(ii)  **$n$ -decreasing** (“ $n$ - $\downarrow$ ”) if instead  $(\nabla_h^p f)(s) \geq 0$ , and

(iii)  **$n$ -alternating** (“ $n$ - $\updownarrow$ ”) if instead  $(\nabla_h^p f)(s) \leq 0$ .

Several remarks are in order:

- (a) The notion “ $n$ -alternating” in this connection is new. In dimension one it is well-known and was introduced by Choquet [2], actually in a more general frame (which differs from our use for  $d \geq 2$ ).
- (b) The omission of  $p = 0$  is essential, since also functions with negative values will be considered.
- (c) Clearly the set of all  $n$ - $\uparrow$  functions (resp.  $n$ - $\downarrow$ , resp.  $n$ - $\updownarrow$ ) is a convex cone, closed under pointwise convergence.
- (d) With  $-I := \{-s \mid s \in I\}$  and  $f(-\cdot)(s) := f(-s)$  one has trivially

$$\{s, s + h\} \subseteq I \iff \{-s, -s - h\} \subseteq -I$$

and in this case

$$(\Delta_h f)(s) = (\nabla_h(f(-\cdot)))(-s - h).$$

More generally, for  $n \in N_0^d$

$$\{s, s + n \odot h\} \subseteq I \iff \{-s, -s - n \odot h\} \subseteq -I$$

and then

$$(\Delta_h^n f)(s) = (\nabla_h^n(f(-\cdot)))(-s - n \odot h).$$

As a result

$$f \text{ is } n\text{-}\downarrow \text{ on } I \iff f(-\cdot) \text{ is } n\text{-}\uparrow \text{ on } -I$$

and

$$f \text{ is } n\text{-}\updownarrow \text{ on } I \iff -f(-\cdot) \text{ is } n\text{-}\uparrow \text{ on } -I.$$

These “innocent” observations are really helpful, as we will see shortly.

- (e) A particularly important index of higher order monotonicity is  $n = (1, \dots, 1) =: \mathbf{1}_d$ , in which case the  $I_j$  need not be intervals – just non-empty subsets of  $\mathbf{R}$  (or even of  $\overline{\mathbf{R}}$ ). Functions which are  $\mathbf{1}_d$ - $\uparrow$  are also called “fully  $d$ -increasing”,  $\mathbf{1}_d$ - $\downarrow$  corresponds to “fully  $d$ -max increasing” (!), and finally  $\mathbf{1}_d$ - $\updownarrow$  is the same as “fully  $d$ -max decreasing”. We shall still occasionally use “fully  $d$ -increasing”. Note that  $(\Delta_h^{1^d} f)(s)$  is the “ $f$ -volume” of the  $d$ -dimensional rectangle  $]s, s + h[$ .
- (f) These notions make of course also perfect sense in one dimension. For univariate functions we write  $n$ - $\uparrow, n$ - $\downarrow$  or  $n$ - $\updownarrow, n \in \mathbf{N}$ . A function is  $1$ - $\uparrow$  ( $\downarrow$ ) iff  $f$  is increasing (decreasing) in the usual (weak) sense; and  $1$ - $\updownarrow$  is here the same as  $1$ - $\uparrow$ . For  $n = 2$  we have

$$f \text{ is } 2\text{-}\uparrow \iff f \text{ is increasing and convex}$$

$$f \text{ is } 2\text{-}\downarrow \iff f \text{ is decreasing and convex}$$

$$f \text{ is } 2\text{-}\updownarrow \iff f \text{ is increasing and concave}$$

In the literature  $n$ - $\downarrow$  functions are often called “ $n$ -monotone”, and  $n$ - $\uparrow$  ones “ $n$ -absolutely monotone”.

(g) If  $f : I \rightarrow \mathbf{R}$  is of the form (a “tensor sum”)

$$f(x_1, \dots, x_d) = g(x_1, \dots, x_k) + h(x_{k+1}, \dots, x_d)$$

(with  $1 \leq k < d$ ),  $\mathbf{m} \in \mathbf{N}^k$ ,  $\mathbf{n} \in \mathbf{N}^{d-k}$ , then

$$f \text{ is } (\mathbf{m}, \mathbf{n})\text{-}\uparrow(\downarrow\uparrow) \iff g \text{ is } \mathbf{m}\text{-}\uparrow(\downarrow\uparrow) \text{ and } h \text{ is } \mathbf{n}\text{-}\uparrow(\downarrow\uparrow).$$

In particular, for  $f(x) = \sum_{i=1}^d f_i(x_i)$  we see  $f$  is  $\mathbf{n}\text{-}\uparrow(\downarrow\uparrow) \iff f_i$  is  $n_i\text{-}\uparrow(\downarrow\uparrow) \forall i \leq d$ .

Note here the special case  $\mathbf{n} = \mathbf{1}_d$  :

$$f \text{ is } \mathbf{1}_d\text{-}\uparrow \iff f \text{ is } \mathbf{1}_d\text{-}\downarrow \iff \text{each } f_i \text{ is increasing.}$$

- (h) Up to (one-sided) continuity non-negative  $\mathbf{1}_d\text{-}\uparrow$  functions are  $d$ -dimensional distribution functions,  $\mathbf{1}_d\text{-}\downarrow$  characterising survival functions, and  $\mathbf{1}_d\text{-}\downarrow$  corresponding to so-called “co-survival functions”, cf. [11] and [13] for a detailed and precise exposition.
- (i) A natural fourth (and last) monotonicity condition for  $f : I \rightarrow \mathbf{R}$  would be  $\Delta_h^p f \leq 0$  in the above definition, to which no name has been assigned so far. This property would characterize (essentially) “co-distribution functions”, a notion not yet introduced, and not needed here.

Of fundamental importance is the following result about preservation of higher order monotonicity under appropriate conditions, cf. [14], Theorem 12.

**Proposition. Monotone Composition Theorem (“MCT”)**

Let  $d \in \mathbf{N}$ ,  $\mathbf{n} \in \mathbf{N}^d$ ,  $\mathbf{m}^{(i)} \in \mathbf{N}^{n_i}$ ,  $J_i \subseteq \mathbf{R}^{n_i}$  and  $I_i \subseteq \mathbf{R}$  be non-degenerate intervals,  $g_i : J_i \rightarrow I_i$ ,  $1 \leq i \leq d$ . Let further  $f : I_1 \times \dots \times I_d \rightarrow \mathbf{R}$ .

Then, if  $g_i$  is  $\mathbf{m}^{(i)}\text{-}\uparrow$  for each  $i \leq d$ , and  $f$  is  $(|\mathbf{m}^{(1)}|, \dots, |\mathbf{m}^{(d)}|)\text{-}\uparrow$ , the composed mapping  $f \circ (g_1 \times \dots \times g_d)$  is  $(\mathbf{m}^{(1)}, \mathbf{m}^{(2)}, \dots, \mathbf{m}^{(d)})\text{-}\uparrow$ .

(This is a function of  $|\mathbf{n}|$  variables.)

In connection with Remark (d) above the MCT is a highly applicable tool, even for  $d = 1$ . We give two examples:

**Example 1.** It is claimed that  $\varphi(t) := -\log(1 - e^{-t})$  is  $n\text{-}\downarrow$  on  $]0, \infty[$  for each  $n \in \mathbf{N}$  (i.e.  $\varphi$  is completely monotone). We must show that  $\varphi(\cdot)$  is  $n\text{-}\uparrow$  on  $] - \infty, 0[$ . Since  $\log(1 + \cdot)$  is  $n\text{-}\downarrow$  on  $\mathbf{R}_+$  (it is a well-known Bernstein function),  $-\log(1 - \cdot)$  is  $n\text{-}\uparrow$  on  $\mathbf{R}_-$ , and so is then

$$\varphi(\cdot) = [-\log(1 - \cdot)] \circ \exp,$$

the exponential function being  $n\text{-}\uparrow$  on  $\mathbf{R} \forall n$ .

In [4], Ex. 4 this is proved by using so-called Eulerian numbers, counting permutations with a certain number of ascents. Another proof has been given in [6], Corollary 3.2 (b).

**Example 2.** In [1], Theorem 2.1 it is shown that the composition  $\psi \circ \ell$  of a  $d\text{-}\downarrow$  function  $\psi$  on  $\mathbf{R}_+$  with a stable tail dependence function (“STDF”)  $\ell$  on  $\mathbf{R}_+^d$  is always a survival function, thereby solving an open problem. Now  $\ell$  is characterised by being homogeneous,  $\mathbf{1}_d\text{-}\downarrow$ , and normalized by  $\ell(e_i) = 1 \quad \forall i \leq d$ , where  $e_1, \dots, e_d$  are the usual unit vectors; cf. [13], Theorem 6.

With help of the MCT we have immediately the following, slightly more general result:

If  $f$  is a co-survival function, i.e. of the form  $f(x) = \mu([x, \infty]^c)$ , where  $\mu$  is a Radon measure on  $[0, \infty]^d \setminus \{\infty_d\}$ , and if  $\psi : \mathbf{R}_+ \rightarrow \mathbf{R}_+$  is  $d$ - $\downarrow$ ,  $\psi(0) = 1$ , then  $\psi \circ f$  is a survival function. (Note that  $f$  is a STDF iff  $\mu$  is homogeneous.)

The essential condition to check is that  $\psi \circ f$  is  $\mathbf{1}_d$ - $\downarrow$  (see [11], Theorem 10), equivalently that  $\psi \circ f(\cdot)$  is  $\mathbf{1}_d$ - $\uparrow$ . Now  $f$  is  $\mathbf{1}_d$ - $\uparrow$ , cf. [13], Theorem 3, and

$$\psi \circ f(\cdot) = \psi(\cdot) \circ (-f(\cdot)) \quad (\text{on } \mathbf{R}^d)$$

is indeed  $\mathbf{1}_d$ - $\uparrow$  by the MCT. Since  $f$  is left continuous, and  $0 \in \mathbf{R}_+$  is the only possible point of discontinuity for  $\psi$  (we have  $d \geq 2$ ), a moment's reflection gives left continuity of  $\psi \circ f$ . Together with  $\psi \circ f(0) = 1$  we see that  $\psi \circ f$  is a survival function.

### 3 Two “impossibility theorems” for copulas

In view of the MCT if a  $d$ -dimensional copula was  $\mathbf{n}$ - $\uparrow$ , then  $d$  other copulas of resp. dimensions  $n_1, \dots, n_d$  could be inserted without destroying the copula property. We shall see that this higher degree of monotonicity is not possible.

An important notion in the following is that of a *Bauer simplex*; this is a compact convex subset  $K$  of some locally convex Hausdorff space  $X$ , for which its set of extreme points  $\text{ex}(K)$  is closed (hence compact as well), and for which the integral representation given by the Krein-Milman theorem is unique. In other words, for each  $x \in K$  there is a unique  $\mu \in M_+^1(\text{ex}(K))$  such that

$$x = \int_{\text{ex}(K)} y \, d\mu(y)$$

in the sense of

$$\varphi(x) = \int_{\text{ex}(K)} \varphi(y) \, d\mu(y)$$

for each continuous linear functional  $\varphi : X \rightarrow \mathbf{R}$ . In our applications  $X$  will always be a suitable set of real-valued functions on some set  $T$ , with the topology of pointwise convergence. The point evaluations  $\varphi_t(x) := x(t)$  for some  $t \in T$  are then continuous linear functionals, and we obtain

$$x(t) = \int_{\text{ex}(K)} y(t) \, d\mu(y), \quad t \in T.$$

In the proof of Theorem 1 we'll need the integral representation of  $\mathbf{n}$ - $\uparrow$  functions on  $[0, 1]^d$ ,  $d \geq 2$ , as shown in [14], Theorem 5. We first define functions  $f_a$ ,  $0 \leq a \leq 1$ , on  $[0, 1]$  by

$$f_a(t) := \frac{(t - a)_+}{1 - a}, \quad 0 \leq a < 1, \text{ and } f_1 := 1_{\{1\}};$$

note that  $f_0(t) = t$ . For  $n \in \mathbf{N}$  let

$$K_n := \{f : [0, 1] \rightarrow \mathbf{R}_+ \mid f \text{ is } \mathbf{n}\text{-}\uparrow, f(1) = 1\}.$$

Then  $K_n$  is a Bauer simplex (for  $n \geq 2$ ) and

$$E_n := \text{ex}(K_n) = \{f_0^j \mid j = 0, 1, \dots, n - 2\} \cup \{f_a^{n-1} \mid 0 \leq a \leq 1\}.$$

For  $\mathbf{n} = (n_1, \dots, n_d) \in \mathbf{N}^d$  with  $\mathbf{n} \geq \mathbf{2}_d$  we put

$$E_{\mathbf{n}} := E_{n_1} \otimes \dots \otimes E_{n_d},$$

the set of all tensor products  $h_1 \otimes \dots \otimes h_d$  with  $h_i \in E_{n_i}$  for all  $i$ . And we consider

$$K_n := \{f : [0, 1]^d \rightarrow \mathbf{R}_+ \mid f \text{ is } \mathbf{n}\text{-}\uparrow, f(1_d) = 1\}.$$

Then  $K_n$  is again a Bauer simplex, with  $E_n$  as its set of extreme points.

**Theorem 1.** For  $d \geq 2$  and  $\mathbf{n} \geq \mathbf{2}_d$  the only  $\mathbf{n}\text{-}\uparrow$  copula is the independence copula. This is still true if  $\mathbf{n} \geq \mathbf{2}_d - e_i$  for one index  $i \leq d$ .

**Proof.** Let  $C$  be an  $\mathbf{n}\text{-}\uparrow$  copula,  $\mathbf{n} \geq \mathbf{2}_d$ . Since  $K_n$  is a Bauer simplex,  $C$  has the unique integral representation

$$C(s) = \int h_1(s_1) \cdot \dots \cdot h_d(s_d) d\mu(h_1, \dots, h_d), \quad s \in [0, 1]^d$$

where  $\mu \in M_+^1(E_n)$ . The copula condition  $C(1, \dots, 1, s_i, 1, \dots, 1) = s_i$  translates into

$$s_i = \int h_i(s_i) d\mu^i(h_i)$$

where  $\mu^i \in M_+^1(E_{n_i})$  is the  $i$ 'th projection of  $\mu$ . But  $s_i = f_0(s_i)$ , and the uniqueness of the integral representation (in dimension one) gives  $\mu^i = \varepsilon_{f_0}$ , the one-point measure on  $E_{n_i}$  in  $f_0$ . This being true for every  $i \leq d$ , we conclude that  $\mu = \varepsilon_{(f_0, \dots, f_0)}$ , that is  $C(s) = \prod_{i=1}^d s_i$ .

If  $C$  is  $(1, 2, \dots, 2)\text{-}\uparrow$  (without restriction), we may apply Theorem 7 in [14], and have a representation

$$C(s) = \int 1_{[a_1, 1]}(s_1) \cdot h_2(s_2) \cdot \dots \cdot h_d(s_d) d\mu(a_1, h_2, \dots, h_d)$$

with  $\mu \in M_+^1([0, 1] \times E_{\mathbf{2}_{d-1}})$ . Again the copula condition gives immediately  $\mu^i = \varepsilon_{f_0}$  for  $i = 2, \dots, d$ , or  $\mu = \mu^1 \otimes \varepsilon_{(f_0, \dots, f_0)}$ , with  $(f_0, \dots, f_0)$  of length  $d - 1$ . Therefore

$$\begin{aligned} C(s_1, 1, \dots, 1) = s_1 &= \int 1_{[a_1, 1]}(s_1) d\mu^1(a_1) \\ &= \int 1_{[0, s_1]}(a_1) d\mu^1(a_1) = \mu^1([0, s_1]), \end{aligned}$$

showing that  $\mu^1$  is the uniform distribution on  $[0, 1]$ , and finally

$$C(s) = \left( \int 1_{[a_1, 1]}(s_1) da_1 \right) \cdot s_2 \cdot \dots \cdot s_d = \prod_{i=1}^d s_i. \quad \square$$

**Remark.** If more than one component of  $\mathbf{n}$  is 1, the above result does no longer hold, since for any bivariate copula  $B$  the function  $B(s_1, s_2) \cdot s_3$  defines a trivariate copula which is  $(1, 1, n)\text{-}\uparrow$  for each  $n$ .

We will need the following result, cf. Theorem 11 in [14]:

Let  $f : [0, 1]^d \rightarrow \mathbf{R}_+$  and  $\mathbf{n} \in \mathbf{N}^d$  be given. Then there are equivalent:

- (i)  $f$  is  $\mathbf{n}\text{-}\uparrow$
- (ii) For any fully  $n_i$ -increasing  $\varphi_i : \{0, 1\}^{n_i} \rightarrow [0, 1]$  the composition  $f \circ (\varphi_1 \times \dots \times \varphi_d)$  is fully  $|\mathbf{n}|$ -increasing on  $\{0, 1\}^{|\mathbf{n}|}$ .

It is clear that in property (ii) we may consider instead  $\varphi_i : \{a, b\}^{n_i} \rightarrow [0, 1]$  for any  $a < b$ , and this will be used below.

**Lemma.** Let  $0 < a < b < 1$  and let  $\varphi : \{a, b\}^k \rightarrow [0, 1]$  be fully  $k$ -increasing. Then  $\varphi$  may be extended to a d.f. of some probability measure  $\mu$  on  $[0, 1]^k$ .

**Proof.** By Theorem 7 in [11] there is a finite measure  $\nu$  on  $\{a, b\}^k$  with d.f.  $\varphi$ . We extend  $\nu$  to a probability measure  $\mu$  on  $[0, 1]^k$  by  $\mu(\{\mathbf{1}_k\}) := 1 - \varphi(b, \dots, b)$ . Then the d.f. of  $\mu$  is obviously an extension of  $\varphi$ .  $\square$

The following “impossibility theorem” was proved for  $d = 2$  in [3]; it shows that general results on the preservation of the copula property cannot be expected.

**Theorem 2.** *Let  $C$  be a  $d$ -variate copula such that  $C \circ (C_1 \times \dots \times C_d)$  is again a copula for arbitrary bivariate copulas  $C_1, \dots, C_d$  (where even in one fixed component  $i$  the “univariate copula” is allowed). Then  $C$  is the independence copula.*

**Proof.** We choose  $0 < a < b < 1$ . Assuming  $i = 1$  to be the “exceptional” coordinate, let  $\varphi_2, \dots, \varphi_d : \{a, b\}^2 \rightarrow [0, 1]$  be fully 2-increasing. Applying the above Lemma we extend  $\varphi_2, \dots, \varphi_d$  to d.f.s  $\psi_2, \dots, \psi_d$  on  $[0, 1]^2$ , and each  $\psi_j$  can by Sklar’s theorem be written as

$$\psi_j = C_j \circ (f_j \times g_j)$$

with suitable bivariate copulas  $C_j$  and increasing functions  $f_j, g_j : [0, 1] \rightarrow [0, 1]$ . By assumption

$$C \circ (id \times C_2 \times \dots \times C_d)$$

is again a copula, hence fully  $1 + 2(d - 1)$ - $\uparrow$ , so is therefore

$$C \circ (id \times \psi_2 \times \dots \times \psi_d)$$

and finally also its restriction

$$C \circ (id \times \varphi_2 \times \dots \times \varphi_d).$$

We see that  $C$  is  $(1, 2, \dots, 2)$ - $\uparrow$ , and thus by Theorem 1  $C$  is the independence copula.  $\square$

## 4 Nested Archimedean copulas

We saw above that copulas in general do not operate on each other. Nonetheless the question under which additional conditions this might hold, was raised many times (cf. for ex. [10], p. 106). For so-called Archimedean copulas a rather satisfying result was shown by Rezapour [15], for which we give an alternative proof.

Let  $h : [0, \infty] \rightarrow [0, 1]$  be a decreasing surjection, decreasing strictly on  $\{h > 0\}$ . Then  $h^{-1}$  is well-defined on  $]0, 1]$ , and we complement it by  $h^{-1}(0) := \sup\{h > 0\}$ . Put

$$A_h(s) := h \left( \sum_{i=1}^d h^{-1}(s_i) \right), \quad s \in [0, 1]^d.$$

Then if  $A_h$  is a d.f., it is automatically a copula, and this holds iff  $h$  is  $d$ - $\downarrow$ , see [9], Theorem 3.5.  $A_h$  is then called *Archimedean copula with generator  $h$* , and  $h$  a  *$d$ -generator*.

Let us now consider the situation that other Archimedean copulas are inserted in  $A_h$ . That is, we look at the function

$$C(x) := h \left[ \sum_{i=1}^d h^{-1} (A_{h_i}(x_{\alpha_i})) \right]$$

where  $\alpha_1, \dots, \alpha_d$  is a partition of  $\{1, \dots, n_0\}$ , and  $h_i$  is an  $n_i$ -generator, with  $n_i := |\alpha_i|$ , the number of elements in  $\alpha_i$ .

More explicitly,

$$(*) \quad C(x) = h \left[ \sum_{i=1}^d (h^{-1} \circ h_i) \left( \sum_{j \in \alpha_i} h_i^{-1}(x_j) \right) \right].$$

and now the functions  $h^{-1} \circ h_i$  play a rôle.

**Theorem 3.** Let  $d \geq 2$ ,  $n_i \geq 2$  for  $i = 1, \dots, d$ , and  $n_0 := \sum_{i=1}^d n_i$ . Suppose  $g_i : \mathbf{R}_+ \rightarrow \mathbf{R}_+$  is  $n_i$ - $\downarrow$  and bijective,  $i = 1, \dots, d$ , and  $h$  is an  $n_0$ -generator. Put  $h_i := h \circ g_i$  for  $i \leq d$ . Then

- (i)  $h_i$  is  $n_i$ - $\downarrow$  and  $A_{h_i}$  is a copula  $\forall i \leq d$
- (ii)  $A_h \circ (A_{h_1} \times \dots \times A_{h_d})$  is also a copula.

**Proof.**

- (i) We have to show that  $h_i(-\cdot)$  is  $n_i$ - $\uparrow$ . Now

$$h_i(-\cdot) = h(-g_i(-\cdot))$$

is indeed  $n_i$ - $\uparrow$  as the composition of  $h(-\cdot)$  and  $-g_i(-\cdot)$ , by the MCT. Clearly  $h_i$  is then a generator and consequently  $A_{h_i}$  is a copula.

- (ii) We rewrite (\*) as

$$C(x) = h \left[ \sum_{i=1}^d g_i \left( \sum_{j \in \alpha_i} h_i^{-1}(x_j) \right) \right]$$

where  $-\sum_{j \in \alpha_i} h_i^{-1}(x_j)$  is  $\mathbf{1}_{n_i}$ - $\uparrow$ , cf. Remark (g) above, and  $-g_i(-\cdot)$  is  $n_i$ - $\uparrow$ , therefore, again by applying the MCT,

$$\varphi_i(x_{\alpha_i}) := -g_i \left( \sum_{j \in \alpha_i} h_i^{-1}(x_j) \right)$$

is also  $\mathbf{1}_{n_i}$ - $\uparrow$ , hence their sum  $\varphi := \sum_{i=1}^d \varphi_i$  is  $\mathbf{1}_{n_0}$ - $\uparrow$ , and finally  $C = h \circ (-\varphi) = h(-\cdot) \circ \varphi$  is also  $\mathbf{1}_{n_0}$ - $\uparrow$ , and then a copula.  $\square$

**Remark.** If  $g_i$  is  $n_i$  times differentiable, then

$$g_i \text{ is } n_i\text{-}\downarrow \iff g'_i \geq 0, g''_i \leq 0, g'''_i \geq 0, \dots, (-1)^{n_i} g^{(n_i)} \leq 0.$$

The conditions of Theorem 3 are thus a lot less restrictive than what has been assumed so far in the literature, namely requiring  $h, h_1, \dots, h_d$ , and  $g'_1, \dots, g'_d$  to be all completely monotone, cf. [5] and [8].

Let us look at the special case where  $n_i = 2$  for all  $i$  (and  $n_0 = 2d$ ): Then the functions  $g_i$  have to be just increasing concave bijections. If now  $h$  is, say completely monotone, a truly huge class of nested Archimedean copulas emerges!

## 5 Stable tail dependence functions

A stable tail dependence function (“STDF”)  $f : \mathbf{R}_+^d \rightarrow \mathbf{R}$  is a homogeneous  $\mathbf{1}_d$ -alternating function such that  $f(e_i) = 1$  for all  $i$ ; cf. [13], Theorem 6. This is equivalent with

$$C_f(s) := \exp [-f(-\log s_1, \dots, -\log s_d)], \quad s \in ]0, 1]^d$$

to be a so-called extreme value copula, i.e. fulfilling

$$C_f(s_1^t, \dots, s_d^t) = (C_f(s))^t \quad \forall s \in ]0, 1]^d, \quad \forall t > 0.$$



Note that the independence copula corresponds here to  $f(x) = \sum_{i=1}^d x_i$ , the “independence STDF”.

The following result is an analogue to Theorem 1, however with a completely different proof.

**Theorem 4.** *Let  $f : \mathbb{R}_+^d \rightarrow \mathbb{R}$  be a  $2_d$ -alternating STDF. Then  $f(x) = \sum_{i=1}^d x_i$  for all  $x$ .*

**Proof.** It follows from Corollary 1 in [13] that  $f$  is a seminorm. Furthermore  $f(e_i) = 1 \forall i$  (so that  $f$  is even a norm). Hence

$$f(x) = f\left(\sum_{i=1}^d x_i e_i\right) \leq \sum_{i=1}^d x_i \quad \forall x.$$

If  $f(x) = \sum_{i=1}^d x_i$  for some  $x$ , then also  $f(y) = \sum_{i=1}^d y_i \quad \forall 0 \leq y \leq x$ , because then

$$\begin{aligned} \sum_{i=1}^d x_i &= f(x) = f(x - y + y) \leq f(x - y) + f(y) \\ &\leq \sum_{i=1}^d (x_i - y_i) + \sum_{i=1}^d y_i = \sum_{i=1}^d x_i. \end{aligned}$$

$f$  being homogeneous, it suffices to find just one  $x \in ]0, \infty[^d$  with  $f(x) = \sum_{i=1}^d x_i$ , to have this equality everywhere.

We proceed by induction, starting with  $d = 2$ . The two inequalities

$$\begin{aligned} (\nabla_{e_1}^{(2,0)} f)(e_2) &= f(0, 1) - 2f(1, 1) + f(2, 1) \leq 0 \\ (\nabla_{e_2}^{(0,2)} f)(2e_1) &= f(2, 0) - 2f(2, 1) + f(2, 2) \leq 0 \end{aligned}$$

are added to give

$$3 - f(2, 1) \leq 0$$

i.e.  $f(2, 1) = 3$ , and thus  $f(x_1, x_2) = x_1 + x_2 \quad \forall x$ . Now we suppose the result to hold for some  $d \geq 2$ , and consider a  $2_{d+1}$ - $\downarrow$  STDF on  $\mathbb{R}_+^{d+1}$ . Then we know  $f(x) = \sum x_i$  whenever at least one  $x_i$  is zero. Adding the  $d + 1$  inequalities

$$\begin{aligned} f(0, \mathbf{1}_d) - 2f(\mathbf{1}_{d+1}) + f(2, \mathbf{1}_d) &\leq 0 \\ f(2, 0, \mathbf{1}_{d-1}) - 2f(2, \mathbf{1}_d) + f(2, 2, \mathbf{1}_{d-1}) &\leq 0 \\ f(2, 2, 0, \mathbf{1}_{d-2}) - 2f(2, 2, \mathbf{1}_{d-1}) + f(2, 2, 2, \mathbf{1}_{d-2}) &\leq 0 \\ \dots & \\ f(\mathbf{2}_{d-1}, 0, 1) - 2f(\mathbf{2}_{d-1}, 1, 1) + f(\mathbf{2}_d, 1) &\leq 0 \\ f(\mathbf{2}_d, 0) - 2f(\mathbf{2}_d, 1) + f(\mathbf{2}_{d+1}) &\leq 0 \end{aligned}$$

gives, since  $f(\mathbf{2}_{d+1}) = 2f(\mathbf{1}_{d+1})$ ,

$$\begin{aligned} d + (2 + d - 1) + (4 + d - 2) + \dots + (2(d - 1) + 1) + 2d \\ \leq f(2, \mathbf{1}_d) + f(2, 2, \mathbf{1}_{d-1}) + \dots + f(\mathbf{2}_d, 1) \\ \leq (2 + d) + (4 + d - 1) + \dots + (2(d - 1) + 2) + (2d + 1), \end{aligned}$$

i.e. the same value  $\frac{3}{2} \cdot d(d + 1)$  on both sides of this inequality, therefore (termwise) equality.  $\square$

**Remark.** The normalizing condition  $f(e_i) = 1$  for a STDF  $f$  is of course unimportant in the above result. If  $f$  is any homogeneous  $\mathbf{2}_d$ - $\uparrow$  function with  $a_i := f(e_i) > 0$  for all  $i$ , then

$$g(x) := f\left(\frac{x_1}{a_1}, \dots, \frac{x_d}{a_d}\right)$$

is a (normalized) STDF, still  $\mathbf{2}_d$ - $\uparrow$ , leading to  $f(x) = \sum_{i=1}^d a_i x_i$ , and this is of course also true if some  $a_i$  are 0; that is we have always

$$f(x) = \sum_{i=1}^d f(e_i) \cdot x_i \quad \forall x.$$

Notwithstanding Theorem 4 there may be many instances of STDFs which can be composed without leaving this class of functions. The best known subfamily in this regard is related to the so-called Gumbel-Hougaard copulas, and nothing else than the classical norms

$$\ell_\alpha(x_1, \dots, x_d) := \left(\sum_{i=1}^d x_i^\alpha\right)^{1/\alpha}, \quad \alpha \geq 1$$

where of course  $\ell_\infty(x_1, \dots, x_d) := \max_{i \leq d} x_i$ ; here we use only  $x \in \mathbf{R}_+^d$ .

If  $1 \leq \alpha \leq \alpha_i$  for each  $i$ , then the composed map

$$\ell_\alpha \circ (\ell_{\alpha_1} \times \dots \times \ell_{\alpha_d})$$

(with  $\ell_{\alpha_i}$  on  $\mathbf{R}_+^{n_i}$  for some  $n_i \geq 2$ ) is still a STDF; this is true because

$$\ell_\alpha(\ell_{\alpha_1}(x^{(1)}), \dots, \ell_{\alpha_d}(x^{(d)})) = \left[ \sum_i \left( \sum_{j_i \leq n_i} x_{j_i}^{(i)\alpha_i} \right)^{\alpha/\alpha_i} \right]^{1/\alpha}$$

and  $\alpha/\alpha_i \leq 1, 1/\alpha \leq 1$ , using the fact that  $t \mapsto t^\beta$  is completely alternating on  $\mathbf{R}_+ \forall \beta \in ]0, 1]$ , and the MCT.

The nested copula family corresponding to these STDF's is explained and constructed from a probabilistic modeling viewpoint in [7], Example 3.1.

Without the condition  $\alpha \leq \alpha_i$  the situation is different:

$$f(x_1, \dots, x_4) := \ell_2(\ell_1(x_1, x_2), \ell_1(x_3, x_4)) = \sqrt{(x_1 + x_2)^2 + (x_3 + x_4)^2}$$

is no longer a STDF, since

$$\begin{aligned} \left(\nabla_{(1,1,0,0)}^{(1,1,0,0)} f\right)(1, 1, 1, 1) &= f(1, 1, 1, 1) - f(2, 1, 1, 1) - f(1, 2, 1, 1) + f(2, 2, 1, 1) \\ &= \sqrt{8} - 2\sqrt{13} + \sqrt{20} = 0.089460\dots > 0. \end{aligned}$$

$f$  is not  $\mathbf{1}_4$ - $\uparrow$  and therefore not a STDF.

In fact also the second impossibility theorem for copulas (Theorem 2 above) has an analogue for STDFs. In order to prove it we need the following general result about multivariate monotone functions, certainly of intrinsic interest, too. For  $d = 1$  it was shown already in [12], Theorem 3. We'll use the notation  $\sigma_k : \mathbf{R}^k \rightarrow \mathbf{R}$  for the sum, i.e.  $\sigma_k(x_1, \dots, x_k) := x_1 + \dots + x_k$ .

**Theorem 5.** Let  $d \in \mathbf{N}$ ,  $\mathbf{n} = (n_1, \dots, n_d) \in \mathbf{N}^d$ , let  $J_i \subseteq \mathbf{R}^{n_i}$  be non-degenerate intervals,  $i \leq d$ ,  $I_i := \sigma_{n_i}(J_i)$ ,  $J := J_1 \times \dots \times J_d$ ,  $I := I_1 \times \dots \times I_d$  and  $\sigma_{\mathbf{n}} := \sigma_{n_1} \times \dots \times \sigma_{n_d}$ , so that  $I = \sigma_{\mathbf{n}}(J)$ . Then for any function  $f : I \rightarrow \mathbf{R}$  we have the equivalence

$$f \text{ is } \mathbf{n}\text{-}\uparrow \iff f \circ \sigma_{\mathbf{n}} \text{ is } \mathbf{1}_{|\mathbf{n}|}\text{-}\uparrow .$$

**Proof.**

„ $\Rightarrow$ “: This is an immediate consequence of the MCT.

„ $\Leftarrow$ “: It is clearly sufficient to consider compact intervals  $J_1, \dots, J_d$ . It is even no restriction to assume  $J_i = [0, 1]^{n_i}$  for each  $i \leq d$ : the formula

$$(\Delta_{\mathbf{h}}^{\mathbf{p}} f)(s) = \sum_{0 \leq \mathbf{q} \leq \mathbf{p}} (-1)^{|\mathbf{p}| - |\mathbf{q}|} \binom{\mathbf{p}}{\mathbf{q}} f(s + \mathbf{q} \odot \mathbf{h})$$

allows this transfer, since with  $g(s) := f(a \odot s + b)$ , where  $a \in ]0, \infty[^d$  and  $b \in \mathbf{R}^d$ , it gives

$$(\Delta_{\mathbf{h}}^{\mathbf{p}} g)(s) = \left( \Delta_{a \odot \mathbf{h}}^{\mathbf{p}} f \right) (a \odot s + b) ,$$

and for suitable  $a$  and  $b$  any non-degenerate  $k$ -dimensional compact interval may be mapped to  $[0, 1]^k$  via an affine bijection of the form  $s \mapsto a \odot s + b$ .

That is, we may assume  $J_i = [0, 1]^{n_i}$ ,  $I_i = [0, n_i]$ ,  $J = [0, 1]^{|\mathbf{n}|}$ , and  $I = [0, \mathbf{n}] \subseteq \mathbf{R}^d$ . Define  $g : [0, 1]^d \rightarrow \mathbf{R}$  by  $g(s) := f(\mathbf{n} \odot s)$ , and let  $M_i := \sigma_{n_i}/n_i$  denote the arithmetic mean on  $[0, 1]^{n_i}$ . Then

$$g \circ (M_1 \times \dots \times M_d) = f \circ \sigma_{\mathbf{n}}$$

is  $\mathbf{1}_{|\mathbf{n}|}\text{-}\uparrow$  on  $J = [0, 1]^{|\mathbf{n}|}$ , and we may apply Theorem 13 in [14] to infer that  $g$  is  $\mathbf{n}\text{-}\uparrow$ , or equivalently, that  $f$  is  $\mathbf{n}\text{-}\uparrow$ . (Note that the two conditions  $f \geq 0$  and  $f(\mathbf{1}_d) = 1$  in [14] are of no importance in the proof given there! Also  $\mathbf{n} \geq \mathbf{2}_d$  may be weakened to  $\mathbf{n} \geq \mathbf{1}_d$  without any change in the proof.)  $\square$

We are now ready to prove the following analogue to Theorem 2.

**Theorem 6.** Let  $\ell$  be a STDF on  $\mathbf{R}_+^d$  such that  $\ell \circ (\ell_1 \times \dots \times \ell_d)$  is again a STDF for arbitrary bivariate  $\ell_1, \dots, \ell_d$ . Then  $\ell(x) = \sum_{i=1}^d x_i \forall x$ .

**Proof.** Let  $\ell_1 = \dots = \ell_d$  be the bivariate independence STDF,  $\ell_i(s, t) = s + t$ . Then  $\ell(x_1 + x_2, x_3 + x_4, \dots)$  is  $\mathbf{1}_{2d}\text{-}\uparrow$  on  $\mathbf{R}_+^{2d}$ , equivalently

$$-\ell(-x_1 - x_2, -x_3 - x_4, \dots) \text{ is } \mathbf{1}_{2d}\text{-}\uparrow$$

on  $\mathbf{R}^{2d}$ , therefore by Theorem 5  $-\ell(\cdot)$  is  $\mathbf{2}_d\text{-}\uparrow$  on  $\mathbf{R}^d$ , which in turn means that  $\ell$  itself is  $\mathbf{2}_d\text{-}\downarrow$  on  $\mathbf{R}_+^d$ . But then  $\ell = \sigma_d$  because of Theorem 4.  $\square$

It is well-known (and easy to see) that any STDF  $\ell$  lies between  $\ell_1$  and  $\ell_\infty$ :

$$\max_{1 \leq i \leq d} x_i \leq \ell(x) \leq \sum_{i=1}^d x_i \quad \forall x .$$

Contrary to the sum, the max-STDF can be inserted in an arbitrary STDF  $\ell$ :  $\ell$  has the integral representation

$$\ell(x) = \int_C \max_{1 \leq i \leq d} (c_i x_i) d\mu(c) \quad , \quad x \in \mathbf{R}_+^d$$

where  $C := \{c \in \mathbf{R}_+^d \mid \max_{i \leq d} c_i = 1\}$ , and  $\mu \in M_+(C)$  is a finite (not probability) measure on  $C$  such that  $\int c_i d\mu(c) = 1$  for all  $i \leq d$ . Inserting  $\max_{j_1 \leq n_1} x_{j_1}^{(1)}, \dots, \max_{j_d \leq n_d} x_{j_d}^{(d)}$  at the place of  $x_1, \dots, x_d$  we get

$$\ell \left( \max_{j_1} x_{j_1}^{(1)}, \dots, \max_{j_d} x_{j_d}^{(d)} \right) = \int_C \max_{i, j_i} (c_i x_{j_i}^{(i)}) d\mu(c),$$

obviously a STDF on  $\mathbf{R}_+^{n_1 + \dots + n_d}$ .

## References

- [1] Charpentier, A., A.-L. Fougères, C. Genest and J.G. Nešlehová (2014). Multivariate Archimax copulas. *J. Multivariate Anal.* 126, 118–136.
- [2] Choquet, G. (1954). Theory of capacities. *Ann. Inst. Fourier* 5, 131–295.
- [3] Genest, C., J.J. Quesada Molina and J.A. Rodrigues Lallena (1995). De l'impossibilité de construire des lois à marges multidimensionnelles données à partir de copules. *C. R. Acad. Sci. Paris Sér. I Math.* 320(6), 723–726.
- [4] Genest, C., J.G. Nešlehová and L.-P. Rivest (2018). The class of multivariate max-id copulas with  $\ell_1$ -norm symmetric exponent measure. *Bernoulli* 24(4B), 3751–3790.
- [5] Hering, C., M. Hofert, J.-F. Mai and M. Scherer (2010). Constructing hierarchical Archimedean copulas with Lévy subordinators. *J. Multivariate Anal.* 101(6), 1428–1433.
- [6] Mai, J.-F. (2014). A note on the Galambos copula and its associated Bernstein function. *Depend. Model.* 2, 22–29.
- [7] Mai, J.-F. and M. Scherer (2012).  $H$ -extendible copulas. *J. Multivariate Anal.* 110, 151–160.
- [8] McNeil, A.J. (2008). Sampling nested Archimedean copulas. *J. Stat. Comput. Simul.* 78(6), 567–581.
- [9] Morillas, P. M. (2005). A characterization of absolutely monotonic ( $\Delta$ ) functions of a fixed order. *Publ. Inst. Math. (N.S.)* 78(92), 93–105.
- [10] Nelsen, R. B. (2006). *An Introduction to Copulas*. Second edition. Springer, New York.
- [11] 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.
- [12] Ressel, P. (2012). Functions operating on multivariate distribution and survival functions—With applications to classical mean-values and to copulas. *J. Multivariate Anal.* 105(1), 55–67.
- [13] Ressel, P. (2013). Homogeneous distributions—And a spectral representation of classical mean values and stable tail dependence functions. *J. Multivariate Anal.* 117, 246–256.
- [14] Ressel, P. (2014). Higher order monotonic functions of several variables. *Positivity* 18(2), 257–285.
- [15] Rezapour, M. (2015). On the construction of nested Archimedean copulas for  $d$ -monotone generators. *Statist. Probab. Lett.* 101, 21–32.