REVSTAT – Statistical Journal
Volume 0, Number 0, Month 0000, 000–000
https://doi.org/00.00000/revstat.v00i0.000

On entropy and divergence type measures of bivariate extreme value copulas

Authors: KONSTANTINOS ZOGRAFOS ✉
– Department of Mathematics, University of Ioannina, 451 10 Ioannina, Greece
kzograf@uoi.gr

Received: Month 0000 Revised: Month 0000 Accepted: Month 0000

Abstract:

- Pickands dependence function is the basis of extreme value copulas which formulate the extreme dependence between random variables. Exact forms of cumulative entropy and divergence type measures, which was discussed recently in Zografos (2023), are determined in the light of bivariate extreme value copulas and they are formulated in terms of the Pickands dependence function. Some fundamental properties of the presented entropy and divergence type indices are studied. The results presented here provide a link between statistical information theory and copula theory.

Keywords:

- *Cumulative entropy; Cumulative divergence; Extreme value copula; Pickands dependence function; Entropy of extreme value copula; Divergence of extreme value copula; Marshall-Olkin copula.*

AMS Subject Classification:

- 62B10, 62H05, 62G32, 60E05, 60G70.

✉ Corresponding author

1. INTRODUCTION

Statistical Information Theory is a branch of mathematics, probability and statistics, statistical mechanics, information engineering and other research fields which has many contacts with a lot of research areas in science and engineering, such as statistical inference, information transmission, source coding, cryptography, pattern recognition, among many others. This field provides with numerous measures that quantify the amount of information, where the word information is used in several conceptual frameworks and contexts. Several meanings of the word information, appeared in the early literature, was motivated Ferentinos and Papaioannou (1981), Papaioannou (1985, 2001) to classify the measures of information into three broad categories, namely, entropy, divergence and Fisher's type measures of information (cf. Zografos (2023) and references appeared therein for details). The characterization of the measures of information and their applications is also included in the main activities of the field of statistical information theory and the introduction in Zografos (2023) is gathering the respective literature.

The classical measures of information, like Shannon entropy, Kullback-Leibler divergence, Fisher's measure of information have been initially defined and applied by means of the probability density function that governs the experimental data. The last two decades, measures of information are defined in terms of the distribution function and they complement and strengthen the respective literature. This research activity has been developed on the basis of the seminal paper by Rao et al. (2004) and a timely elaboration of Rao's et al. measure by Zografos and Nadarajah (2005). The use of the cumulative distribution function or the survival function for the definition of measures of information is motivated by the fact that the distribution function is always existing while the density function is not always present. Moreover, the cumulative distribution function is easily estimated by the empirical respective model, in case where the distribution function that drives the data is unknown. The recent paper by Zografos (2023) provides with a review of the classical measures of information and their cumulative counterparts.

On the other hand, the notion of copula and the respective theory of copulas has received significant attention in the literature of the last decades. However, copulas are cumulative distribution functions themselves, defined on $[0, 1]^d$, with uniform marginals on the interval $[0, 1]$. Hence, measures of information defined by means of the cumulative distribution function can be formulated in terms of copulas and, then, to provide in this way entropy type measures of a copula or divergence or quasi-distance type measures between two copulas. This is exactly the point that motivates the present paper which aims to create a bridge between information theory and copula theory with an ultimate goal: information theoretic methods to be used for the formulation and solution of problems from the area of copula theory.

In the general framework, described above, this paper concentrates in the bivariate case and it focuses in a specific type of copulas, the bivariate extreme value copulas. The main aim is the derivation of the analytic expressions of cumulative entropies and divergences in the case of bivariate extreme value copulas and the study of some of the properties of the introduced here measures. Similar work is maybe of interest for other classes of copulas such as Archimedean or elliptical copulas and this line of work is open to the best our knowledge.

In the context described above, the next section concentrates on some preliminary concepts and the notation. More specifically it presents the definition of some entropy and divergence measures, defined in terms of distribution functions. The bivariate extreme value copula is also presented in section 2. Closed-form expressions for entropy type measures of bivariate extreme value copulas are obtained in section 3 and some fundamental properties of the said entropies are studied. In the same frame, section 4 is devoted to the analytic expressions for some divergence measures in the case of bivariate extreme value copulas. The paper is completed with some conclusions.

2. Preliminaries and notation

The first part of this section closely follows the exposition in Zografos (2023) and it is devoted to the presentation of a broad family of entropy type measures and a direct analog of Csiszár's type ϕ-divergences, in terms of the cumulative distribution functions. The cumulative counterpart of the density power divergence is also presented in the same setup. The said entropy and divergence type measures, introduced and defined in Zografos (2023), are reproduced here, not only for the sake of completeness but mainly because they will be the basis in the investigations of the subsequent sections.

Based on this last paper, in order to formulate entropy type and divergence type measures defined in terms of cumulative distribution functions, following standards arguments (cf. Billingsley (1986)), consider two d-dimensional random vectors $X = (X_1, ..., X_d)^t$ and $Y = (Y_1, ..., Y_d)^t$. Let F and G denote, respectively, the joint distribution functions of X and Y, defined by, $F(x_1, ..., x_d) = \Pr(X_1 \leq x_1, ..., X_d \leq x_d)$ and $G(y_1, ..., y_d) = \Pr(Y_1 \leq y_1, ..., Y_d \leq y_d)$, for $(x_1, ..., x_d) \in \mathbb{R}^d$ and $(y_1, ..., y_d) \in \mathbb{R}^d$. In this context and following the exposition in Zografos (2023), a broad family of entropy type measures has been defined in Chen et al. (2012), in a somewhat different setup, and some years later in the papers by Klein et al. (2016); Klein and Doll (2020) by

$$\mathcal{CE}_\varphi(F) = \int_{-\infty}^{+\infty} \varphi(F(x))dx, \tag{2.1}$$

where the entropy generating function φ is a non-negative and concave real valued function defined on $[0, 1]$. This measure is the direct analog of Burbea and Rao (1982) φ-entropy, which was defined by means of a probability density function, while for special choices of the concave function φ it leads to interesting particular entropy type measures, like that appeared in Table 3 of p. 13, in Klein and Doll (2020). The measure $\mathcal{CE}_\varphi(F)$, defined in (2.1) above, is an entropy type measure defined in terms of the cumulative distribution function F and this definition is immediately extended to the multivariate case. Shannon's type and Tsallis' type cumulative entropies are obtained from (2.1) for $\varphi(x) = -x \ln x$, $x \in [0, 1]$, by the convention $0 \ln 0 = 0$ and $\varphi(x) = \frac{x - x^\lambda}{\lambda - 1}, \lambda > 0, \lambda \neq 1$, $x \in [0, 1]$, respectively. Based on (Klein et al., 2016, p. 2), entropy type measures, like that for $\varphi(x) = -x \ln x$, $x \in [0, 1]$, "can rather be interpreted as measures of dispersion than as measures of information.". Moreover, following (Asadi et al., 2017, p. 1030), Tsallis' type cumulative entropy, obtained from (2.1) for $\varphi(x) = \frac{x - x^\lambda}{\lambda - 1}, \lambda > 0, \lambda \neq 1$, $x \in [0, 1]$, is a measure of concentration of the distribution. That is, Tsallis' type entropy is non-negative and equals zero if and only if the underline distribution is degenerate. The proof of this assertion is immediately obtained by a straightforward application of the proof in page 1030 in Asadi et al. (2017) and taking into account that $\varphi(x) = \frac{x-x^{\lambda}}{\lambda-1} = -x^{\lambda}L_{\lambda}(x), \lambda > 0, \lambda \neq 1,\ x \in (0,1]$, where $L_{\lambda}(x) = \frac{x^{1-\lambda}-1}{1-\lambda}$ is defined by (2.3), p. 1029 in Asadi et al. (2017).

In spite of entropy, a divergence type measure is defined in terms of two probability distribution functions and it usually serves as a measure of quasi-distance or statistical distance between the underlined probability distributions. A broad family of divergence type measures between two cumulative probability distributions was defined recently in Zografos (2023) as a direct analog of Csiszár's ϕ-divergences. The cumulative Csiszár's type ϕ-divergence between the distribution functions F and G is defined by

$$\mathcal{CD}_{\phi}(F,G) = \int\limits_{\mathbb{R}^d} G(x)\phi\left(\frac{F(x)}{G(x)}\right)dx - \left(\int\limits_{\mathbb{R}^d} G(x)dx\right)\phi\left(\frac{\int\limits_{\mathbb{R}^d} F(x)dx}{\int\limits_{\mathbb{R}^d} G(x)dx}\right), \tag{2.2}$$

where $\phi : (0,\infty) \to \mathbb{R}$ is a real valued convex function and $\phi \in \Phi$, defined by

$$\Phi = \left\{\phi : \phi \text{ is strictly convex at 1, with } \phi(1) = 0, 0\phi\left(\frac{0}{0}\right) = 0, 0\phi\left(\frac{u}{0}\right) = \lim_{v\to\infty}\frac{\phi(v)}{v}\right\}. \tag{2.3}$$

This measure satisfies the next non-negativity property (cf. Zografos (2023))

$$\mathcal{CD}_{\phi}(F,G) \geq 0 \text{ with equality if and only if } F(x) = G(x), \text{ on } \mathbb{R}^d,$$

which ensures the application of this measure as a quasi-distance in formulating and solving problems in probability and statistics. Specific choices of the convex function ϕ lead to particular cumulative type divergences, which are products of Csiszár's type cumulative ϕ-divergence, defined by (2.2) and (2.3). In this frame, Cressie and Read's type cumulative divergence (cf. Cressie and Read (1984); Read and Cressie (1988)) is directly obtained from (2.2) for $\phi(u) = \phi_{\lambda}(u) = \frac{u^{\lambda+1}-u-\lambda(u-1)}{\lambda(\lambda+1)}, \lambda \neq 0, -1, u > 0$, and it is defined as follows

$$\begin{aligned}&\mathcal{CD}_{\lambda}(F,G)\\ &= \tfrac{1}{\lambda(\lambda+1)}\left(\int\limits_{\mathbb{R}^d} G(x)\left(\tfrac{F(x)}{G(x)}\right)^{\lambda+1}dx - \left(\int\limits_{\mathbb{R}^d} G(x)dx\right)\left(\int\limits_{\mathbb{R}^d} F(x)dx/\int\limits_{\mathbb{R}^d} G(x)dx\right)^{\lambda+1}\right),\end{aligned} \tag{2.4}$$

for $\lambda \neq 0, -1$. Kullback-Leibler type cumulative divergence (cf. Zografos (2023))

$$\mathcal{CD}_{KL}(F,G) = \int\limits_{\mathbb{R}^d} F(x)\ln\left(\frac{F(x)}{G(x)}\right)dx - \left(\int\limits_{\mathbb{R}^d} F(x)dx\right)\ln\left(\int\limits_{\mathbb{R}^d} F(x)dx/\int\limits_{\mathbb{R}^d} G(x)dx\right), \tag{2.5}$$

is directly obtained from (2.2) for $\phi(u) = u\log u, u > 0$, or $\phi(u) = u\log u + u - 1, u > 0$. It is related with $\mathcal{CD}_{\lambda}(F,G)$ in the limiting sense, that follows,

$$\lim_{\lambda\to 0}\mathcal{CD}_{\lambda}(F,G) = \mathcal{CD}_{KL}(F,G) \text{ and } \lim_{\lambda\to -1}\mathcal{CD}_{\lambda}(F,G) = \mathcal{CD}_{KL}(G,F).$$

The cumulative analog of the density power divergence, is introduced in Zografos (2023) by,

$$\mathcal{C}d_a(F,G) = \int\limits_{\mathbb{R}^d}\left\{G(x)^{1+a} - \left(1+\frac{1}{a}\right)G(x)^a F(x) + \frac{1}{a}\,F(x)^{1+a}\right\}dx,\ a > 0, \tag{2.6}$$

and it is a straightforward extension of Basu et al. (1998) and Basu et al. (2011), Chapter 9, density power divergence measure. Following the steps of the proof of Theorem 9.1 of Basu et al. (2011), p. 301, it can be seen that $\mathcal{C}d_a(F,G)$ is non-negative, for all $a>0$ and it is equal to 0 if and only if the underlined cumulative distributions F and G are coincide. Hence, $\mathcal{C}d_a(F,G)$ can also serve as a quasi-distance between the underlined distributions F and G. It is seen that the case $a=0$ is excluded from the definition of $\mathcal{C}d_a(F,G)$. It can be easily shown that

$$\lim_{a\to 0}\mathcal{C}d_a(F,G)=CKL(F,G),$$

where

$$CKL(F,G)=\int\limits_{\mathbb{R}^d}F(x)\ln\left(\frac{F(x)}{G(x)}\right)dx+\int\limits_{\mathbb{R}^d}[G(x)-F(x)]dx, \tag{2.7}$$

has been defined, in the univariate case $d=1$, as an alternative cumulative type Kullback-Leibler divergence by Baratpour and Rad (2012) and Park et al. (2012). Moreover, motivated from the mutual information which was defined in terms of cumulative functions in Zografos (2023), pp. 308-310, a broad class of distance type measures between the joint distribution function and the respective one under the assumption of independence can be defined, in view of (2.6), if the distribution function G is used to represent the product of the marginal distribution functions $F_{X_i}(x_i)$, $i=1,...,d$, of the components of the random vector $X=(X_1,...,X_d)^t$, that is if $G(x)=\prod_{i=1}^d F_{X_i}(x_i)$. It is clear that such a measure can be exploited to define a broad class of measures of dependence and its empirical version can be also serve as a test statistic to develop tests of independence (cf. Zografos (2023) and references therein).

The cumulative type measures of entropy and divergence, defined above, are developed on the basis of cumulative distribution functions. Hence, it is clear that copulas can be used to replace the underline cumulative distribution functions of the above measures. To describe the way, let's concentrate, without loss of generality, in the bivariate case and let (X,Y) be a pair of continuous random variables with joint distribution function H and marginal distribution functions F and G. Then, Sklar's Theorem (cf. Nelsen (2006), pp. 18-22) proves that there is a unique copula C associated with H such that $H(x,y)=C(F(x),G(y))$, for x, y in $\mathbb{R}$. On the other hand, C is the distribution function of the pair of random variables $(U,V)=(F(X),G(Y))$, with margins uniform on the interval $(0,1)$ and $C(u,v)=H(F^{-1}(u),G^{-1}(v))$, $u,v\in(0,1)$. In this setting the entropy type measure (2.1) is formulated as follows,

$$\mathcal{CE}_\varphi(C)=\int\limits_0^1\int\limits_0^1\varphi\left(C(u,v)\right)dudv,$$

and all the measures presented above on the basis of distributions functions can be expressed in a similar way as information theoretic indices between copulas. For example, if C_1 and C_2 are two copula functions defined on $(0,1)^2$ and satisfying the standard conditions (cf., among many others, subsection 2.2 in Nelsen (2006), p. 10), then the cumulative analog of the density power divergence between copulas C_1 and C_2 is defined, in view of (2.6), by

$$\mathcal{C}d_a(C_1,C_2)=\int\limits_0^1\int\limits_0^1\left\{C_2^{1+a}(u,v)-\left(1+\frac{1}{a}\right)C_2^a(u,v)C_1(u,v)+\frac{1}{a}\,C_1^{1+a}(u,v)\right\}dudv,\ a>0, \tag{2.8}$$

and it can also serve as a quasi-distance between the underlined copulas C_1 and C_2. Based on the above discussion, (2.8) can be used to define a broad class of measures of dependence if C_1

denotes a copula function, say C, and C_2 is used to denote the independence copula, which is usually denoted by Π and it is defined by $\Pi(u,v) = uv$, $0 < u, v \le 1$, $(u,v) \neq (1,1)$. The empirical version of $\mathcal{C}d_a(C,\Pi)$ can be the basis of the test statistic for the development of a test of independence. Gathering the recent literature, Klein et al. (2016) provide with copula representations of correlations defined in terms of cumulative entropies, while Ma and Sun (2008) and De Keyser and Gijbels (2024) present dependence measures by means of copula densities. Recently, Nair and Sunoj (2023) propose survival copula entropy as alternative to various entropy measures.

Let now turn ourselves again to the bivariate case and consider the extreme value copula (cf. Beirlant et al. (2004), p. 273, 314, Nelsen (2006), p. 98, Drouet Mari and Kotz (2001), p. 95, Joe (2015), p. 148, 382, Ghoudi et al. (1998)). A bivariate copula is an extreme value copula if and only if

$$C_A(u,v) = \exp\left[\ln(uv) \cdot A\left\{\frac{\ln v}{\ln(uv)}\right\}\right], \quad 0 < u, v \le 1, \ (u,v) \neq (1,1), \tag{2.9}$$

where $A : [0,1] \to [1/2, 1]$ is a convex function which satisfies $A(0) = A(1) = 1$, $\max\{t, 1-t\} \le A(t) \le 1$, for each $t \in [0,1]$. The function $A(t)$, defined above, is the well known Pickands dependence function while the plot of $A(t) = \max\{t, 1-t\}$ (line) and the plot of $A(t) = 1$ (dash) are given in Figure 1,

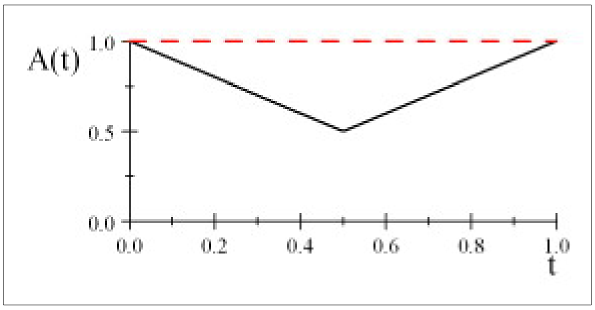

Figure 1: Pickands dependence function $A(t)$.

Based on Nelsen (2006) or Gudendorf and Segers (2010), p. 131, the value $A(t) = 1$ corresponds to independence, while the value $A(t) = \max\{t, 1-t\}$ refers to complete positive dependence, in the sense of comonotonicity, rather, instead of that of the functional relationship of the coordinates of the underlying random vectors.

3. Extreme Value Copula Entropies

Consider now the cumulative entropy of the extreme value copula C_A, which is defined, in view of (2.1), as follows

$$\mathcal{CE}_\varphi(C_A) = \int_{[0,1]^2} \varphi\left(C_A(u,v)\right) dudv, \tag{3.1}$$

in terms of the extreme value copula C_A, where the entropy generating function φ is a non-negative and concave real valued function defined on $[0,1]$ (cf. Klein et al. (2016)). $\mathcal{CE}_\varphi(C_A)$

above, will be called the extreme value copula φ-entropy. The next subsection focuses in particular cases of (3.1).

3.1. Shannon, Tsallis and generating type extreme value copula measures

The proposition that follows provides with analytic expressions for $\mathcal{CE}_\varphi(C_A)$ and special cases, obtained for particular choices of the concave function φ. The proof of the Proposition 3.1 is given in the Appendix A, subsection A.1.

Proposition 3.1. *Let the extreme value copula (2.9) and the extreme value copula φ-entropy (3.1). Then,*
(a) For the non-negative and concave real valued function φ defined on $[0,1]$,

$$\mathcal{CE}_\varphi(C_A) = -\int\limits_{[0,1]^2} v^{\frac{1-t}{t}} \frac{\ln v}{t^2} \varphi\left\{\exp\left[\frac{\ln v}{t} A(t)\right]\right\} dvdt.$$

(b) For $\varphi(x) = -x\ln x$, $x \in [0,1]$, $0\ln 0 = 0$, $\mathcal{CE}_\varphi$ leads to the Shannon's type extreme value copula entropy,

$$\mathcal{CE}_{Sh}(C_A) = -\int\limits_{[0,1]^2} C_A(u,v)\ln C_A(u,v)dudv = 2\int\limits_0^1 \frac{A(t)}{[1+A(t)]^3}dt. \tag{3.2}$$

(c) For $\varphi(x) = \frac{x-x^\lambda}{\lambda-1}, \lambda > 0, \lambda \neq 1$, $x \in [0,1]$, $\mathcal{CE}_\varphi$ leads to the Tsallis' type extreme value copula entropy,

$$\mathcal{CE}_{Ts,\lambda}(C_A) = \frac{1}{\lambda-1}\int\limits_{[0,1]^2}\left[C_A(u,v) - C_A^\lambda(u,v)\right]dudv, \ \lambda > 0, \lambda \neq 1, \tag{3.3}$$

where the generating type function, $\int_{[0,1]^2} C_A^\lambda(u,v)dudv$, $\lambda > 0$, is given by

$$\mathcal{CI}_\lambda(C_A) = \int\limits_{[0,1]^2} C_A^\lambda(u,v)dudv = \int\limits_0^1 \frac{1}{[1+\lambda A(t)]^2}dt, \ \lambda > 0. \tag{3.4}$$

Remark 1. (a) The expression for $\mathcal{CE}_\varphi(C_A)$ is quite general. However, for specific choices of φ, like that of cases (b) and (c) of the proposition, the respective expressions are quite neat and they depend on Pickands dependence function $A(t)$. Therefore $\mathcal{CE}_{Sh}(C_A)$ and $\mathcal{CE}_{Ts,\lambda}(C_A)$ can be thought as indices which are strongly related to the dependence in the bivariate extreme value case and more precisely they are more related to concordance measures, such as Spearman's rho, mentioned below.
(b) It is easy to see that extreme value copula Shannon and Tsallis type entropies are connected by a limiting behaviour, as follows,

$$\lim_{\lambda\to 1} \mathcal{CE}_{Ts,\lambda}(C_A) = \mathcal{CE}_{Sh}(C_A). \tag{3.5}$$

Tsallis type measure is motivated, in essence, by similar measures defined in Asadi et al. (2017), Rajesh and Sunoj (2019), Klein and Doll (2020), which are recently discussed in

Zografos (2023).
(c) For $\lambda = 1$, the generating type function $\mathcal{CI}_\lambda(C_A)$ is simplified as follows,

$$\mathcal{CI}_1(C_A) = \int_{[0,1]^2} C_A(u,v)dudv = \int_0^1 \frac{1}{[1+A(t)]^2}dt. \tag{3.6}$$

However, it is well known (cf. Nelsen (2006), p. 167) that Spearman's rho correlation coefficient is given by

$$\rho_S(A) = 12\int_{[0,1]^2} C_A(u,v)dudv - 3 = 12\mathcal{CI}_1(C_A) - 3 = 12\int_0^1 \frac{1}{[1+A(t)]^2}dt - 3, \tag{3.7}$$

and therefore

$$\mathcal{CI}_1(C_A) = \frac{\rho_S(A)+3}{12}. \tag{3.8}$$

(d) The measure $\mathcal{CI}_\lambda(C_A)$ is a type of generating function because its derivative, in respect to λ, at $\lambda = 1$, generates $\mathcal{CE}_{Sh}(C_A)$ in the sense that $(d/d\lambda)\mathcal{CI}_\lambda(C_A)|_{\lambda=1} = -\mathcal{CE}_{Sh}(C_A)$. This is motivated by the papers of Golomb (1966) and Guiasu and Reischer (1985).
(e) It is also the basis for the definition of Rényi's type extreme value copula entropy of the form

$$\mathcal{CE}_R(C_A) = (1-\lambda)^{-1}\log \mathcal{CI}_\lambda(C_A), \lambda > 0, \lambda \neq 1.$$

The example that follows concentrates on the bivariate Tawn extreme value copula and it illustrates on how the above defined measures are applied in this specific family of extreme value copulas.

Example 1. Asymmetric Tawn extreme value copula

Let consider, in this example, a broad family of extreme value copulas with Pickands dependence functions of the form

$$A(t) = (1-\psi_1)(1-t) + (1-\psi_2)t + \left[(\psi_1(1-t))^\theta + (\psi_2 t)^\theta\right]^{1/\theta}, \tag{3.9}$$

for $0 \le \psi_1, \psi_2 \le 1$, $\theta \ge 1$, by following the notation in p. 49 of Eschenburg (2013). This wide family of dependence functions has been introduced by equation (3.1) in Tawn (1990), where an asymmetric extension of the logistic model has been presented and studied. Based on Eschenburg (2013), p. 56, the symmetric version of Tawn copula is obtained from (3.9) if $\psi_1 = \psi_2 = \Psi$, with $0 \le \Psi \le 1$. In this case, the Pickands dependence function $A(t)$ in (3.9) is related with the respective function of the Gumbel copula by the formula,

$$A(t) = 1 + \Psi\left(A^G(t) - 1\right),$$

with $A^G(t)$ the Pickands dependence function of the Gumbel copula, given by $A^G(t) = \left[(1-t)^\theta + t^\theta\right]^{1/\theta}$, $\theta \ge 1$.

This example evaluates Shannon type extreme-value copula entropy (3.2) in case of Pickands dependence functions of the form (3.9) for various values of ψ_1, ψ_2 and $\theta = 20$.

Table 1 evaluates Shannon type measure (3.2) of Tawn copula with $A(t)$ defined in (3.9), with $\theta = 20$, $\psi_2 = 1$, for different values of ψ_1 while the figure that follows presents the associated Pickends function $A(t)$, in (3.9), with the same choices of the parameters. Figure 1 includes the plots of $A(t)$ for $\psi_1 = 0.1$ (black-solid), $\psi_1 = 0.2$ (blue-dash), $\psi_1 = 0.5$ (red-dots), $\psi_1 = 0.8$ (brown-dash) and $\psi_1 = 1$ (green-solid), the last one is corresponding to symmetry ($\psi_1 = \psi_2 = 1$, Gumbel copula).

	$\mathcal{CE}_{Sh}$ for $\theta = 20$, $\psi_2 = 1$ and different ψ_1					
ψ_1	0.001	0.1	0.2	0.5	0.8	1
$\mathcal{CE}_{Sh}(C_A)$	0.25006	0.25566	0.26032	0.26997	0.27549	0.27777

Table 1: Shannon type entropy of Tawn copula with $\theta = 20$, $\psi_2 = 1$, for different values of ψ_1.

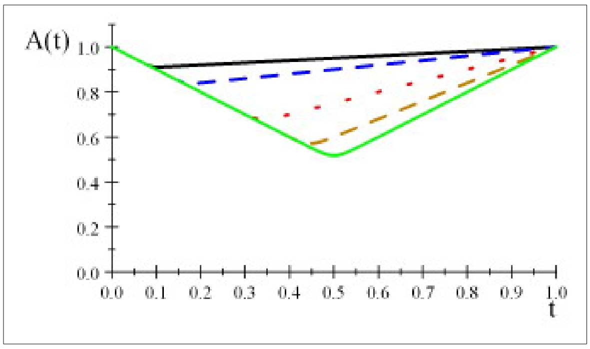

Figure 2: Pickands dependence function $A(t)$ in (3.9) with $\theta = 20$, $\psi_2 = 1$ and $\psi_1 = 0.1, 0.2, 0.5, 0.8, 1$.

We observe that for the specific values of the parameters $\theta = 20$ and $\psi_2 = 1$, Shannon's type entropy $\mathcal{CE}_{Sh}(C_A)$ increases as the value of the parameter ψ_1 increases and the Pickands dependence function is moving from the case of independence to the case of comonotonicity and symmetry.

Table 2 and Figure 3 that follows are the analogs of Table 1 and Figure 2 when the parameters of the model (3.9) are chosen to be $\theta = 20$, $\psi_1 = 1$, for different values of ψ_2.

	$\mathcal{CE}_{Sh}$ for $\theta = 20$, $\psi_1 = 1$ and different ψ_2					
ψ_2	0.001	0.1	0.2	0.5	0.8	1
$\mathcal{CE}_{Sh}(C_A)$	0.25006	0.25566	0.26032	0.26997	0.27549	0.27777

Table 2: Shannon type entropy of Tawn copula with $\theta = 20$, $\psi_1 = 1$, for different values of ψ_2.

The figure that follows includes the plots of $A(t)$ for $\psi_2 = 0.1$ (black-solid), $\psi_2 = 0.2$ (blue-dash), $\psi_2 = 0.5$ (red-dots), $\psi_2 = 0.8$ (brown-dash) and $\psi_2 = 1$ (green-solid), the last one is corresponding to symmetry ($\psi_1 = \psi_2 = 1$, Gumbel copula).

Again, for $\theta = 20$ and $\psi_1 = 1$, Shannon type entropy $\mathcal{CE}_{Sh}(C_A)$ increases as the value of the parameter ψ_2 increases and the Pickands dependence function is moving from the case of independence to the case of comonotonicity and symmetry.

We note at this point that Tables 1 and 2 are the same. This is due to the fact that the integral on the right hand side of (3.2), for $A(t)$ defined by (3.9), is invariant under the

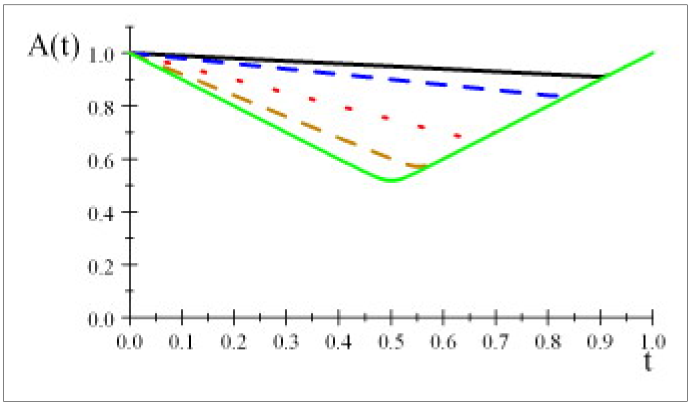

Figure 3: Pickands dependence function $A(t)$ in (3.9) with $\theta = 20$, $\psi_1 = 1$ and $\psi_2 = 0.1, 0.2, 0.5, 0.8, 1$.

transformation $z = 1 - t$, $0 < t < 1$, which reverses the role of the parameters ψ_1 and ψ_2 in (3.9), that is under a transformation which moves the dependence function $A(t)$, in (3.9), from right asymmetry to left asymmetry. Hence, a comparison of the tables and figures leads to the conclusion that Shannon type entropy of Tawn copula does not recognize the type of asymmetry of the said copula. Indeed, the values of $\mathcal{CE}_{Sh}$ for copulas of Figure 2 which correspond to right asymmetry are coincide with the similar ones of Figure 3 which correspond to left asymmetry. ■

We will now concentrate to the Marshall-Olkin bivariate extreme value copula and we will investigate the behaviour of Tsallis type entropy, defined by (3.3) and (3.4), for this family of extreme value copulas. Explicit expressions of the measures of Proposition 3.1 are derived for a specific case of Marshall-Olkin extreme value copula.

Example 2. Marshall-Olkin Extreme-Value Copula

The Pickands dependence function of the Marshall-Olkin extreme-value copula is defined by $A(t) = \max\{1-\alpha_1(1-t), 1-\alpha_2 t\}$, for $0 \leq \alpha_1, \alpha_2 \leq 1$. Based on Eschenburg (2013), p. 35, it is symmetric for $\alpha_1 = \alpha_2$ and it leads to the independence copula if α_1 or $\alpha_2 = 0$. Moreover, if $\alpha_1 = \alpha_2 = 1$, then $A(t) = \max\{t, 1-t\}$ and the respective extreme-value copula is $C_A(u,v) = u \wedge v = \min\{u,v\}$ (cf. Gudendorf and Segers (2010), p. 131).

Figure 4 includes the plots of $A(t)$ for $\alpha_1 = \frac{1}{2}, \alpha_2 = 1$ (blue-dash) which corresponds to right asymmetry, the plot for $\alpha_1 = 1, \alpha_2 = \frac{1}{2}$ (red-dots) which corresponds to left asymmetry and the plot for $\alpha_1 = \alpha_2 = 1$ (black–solid) which corresponds to symmetry of $A(t)$.

Table 3 evaluates Tsallis' type extreme-value copula entropy (3.3) in case of Pickands dependence functions $A(t) = \max\{1-\alpha_1(1-t), 1-\alpha_2 t\}$ of the Marshall-Olkin copula, for various values of α_1, α_2. This table includes also the approximate value of Shannon type extreme-value copula entropy, based on the limiting behaviour (3.5) of Tsallis' type measure. This approximation corresponds to the case $\lambda = 0.99$ of the table.

We observe that Tsallis' type entropy $\mathcal{CE}_{Ts,\lambda}(C_A)$ increases as the Pickands dependence function is moving from the case of independence (α_1 or $\alpha_2 = 0$) to the case of comonotonicity and symmetry ($\alpha_1 = \alpha_2 = 1$), except for $\lambda = 0.1$. Tsallis' type measure $\mathcal{CE}_{Ts,\lambda}$ decreases in respect to λ in case of right asymmetry ($\alpha_1 = 1/2$ and $\alpha_2 = 1$) and the same is happen in

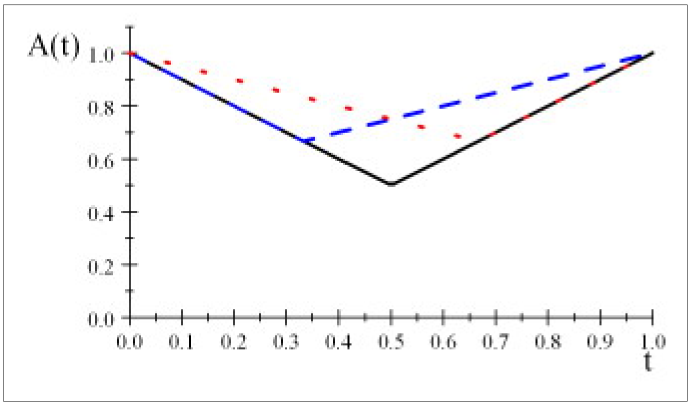

Figure 4: Pickand dependence function $A(t)$ of Marshall-Olkin copula with different values of α_1, α_2.

	$\alpha_1 = 0$				$\alpha_1 = 1/4$			
α_2	0	1/4	1/2	1	0	1/4	1/2	1
$\lambda = 0.1$	0.640 49	0.640 49	0.640 49	0.640 49	0.640 49	0.632 53	0.629 37	0.626 64
$\lambda = 0.5$	0.388 89	0.388 89	0.388 89	0.388 89	0.388 89	0.394 21	0.395 73	0.396 82
$\lambda = 0.99$	0.252	0.252	0.252	0.252	0.252	0.259	0.262	0.264
$\lambda = 1.5$	0.18	0.18	0.18	0.18	0.18	0.187 38	0.189 89	0.191 92
$\lambda = 2$	0.138 89	0.138 89	0.138 892	0.138 89	0.138 89	0.145 46	0.147 73	0.149 57

	$\alpha_1 = 1/2$				$\alpha_1 = 1$			
α_2	0	1/4	1/2	1	0	1/4	1/2	1
$\lambda = 0.1$	0.640 49	0.629 37	0.622 17	0.613 64	0.640 49	0.626 64	0.613 64	0.591 63
$\lambda = 0.5$	0.388 89	0.395 73	0.398 27	0.4	0.388 89	0.396 82	0.4	0.400 03
$\lambda = 0.99$	0.252	0.262	0.268	0.272	0.252	0.264	0.272	0.28
$\lambda = 1.5$	0.18	0.189 89	0.194 96	0.2	0.18	0.191 92	0.2	0.209 53
$\lambda = 2$	0.138 89	0.147 73	0.152 38	0.157 14	0.138 89	0.149 57	0.157 14	0.166 66

Table 3: Tsallis' type entropy of Marshall-Olkin copula for different values of α_1, α_2 and λ.

case of left asymmetry ($\alpha_1 = 1$ and $\alpha_2 = 1/2$).

Table 3 provides with numerical evaluation of Shannon and Tsallis type measures of Marshall-Olkin extreme-value copula with Pickands dependence function defined by $A(t) = \max\{1 - \alpha_1(1-t), 1 - \alpha_2 t\}$, for $0 \le \alpha_1, \alpha_2 \le 1$. In the sequel, explicit expressions of these measures are derived for the case $\alpha_1 = \alpha_2 = \alpha$, for $0 \le \alpha \le 1$. In this case,

$$A(t) = \max\{1 - \alpha(1-t), 1 - \alpha t\}, \text{ for } 0 \le \alpha \le 1.$$

Based on (3.4), (3.6), (3.3), (3.5) and detailed algebraic manipulations, given at the last section of the proofs of the statements, Appendix A.2, we obtain,

$$\begin{aligned}
\mathcal{CI}_\lambda(C_A) &= \frac{2}{(1+\lambda)[2-\lambda(\alpha-2)]}, \lambda>0,\\
\mathcal{CI}_1(C_A) &= \frac{1}{4-\alpha},\\
\mathcal{CE}_{Ts,\lambda}(C_A) &= \frac{1}{\lambda-1}\left[\frac{1}{4-\alpha}-\frac{2}{(1+\lambda)[2-\lambda(\alpha-2)]}\right], \lambda>0, \lambda\neq 1,\\
\mathcal{CE}_{Sh}(C_A) &= \frac{8-3\alpha}{2(4-\alpha)^2},
\end{aligned}$$

for $0 \leq \alpha \leq 1$, while for $\alpha = 0$,

$$\mathcal{CE}_{Ts,\lambda}(C_A) = \frac{\lambda+3}{4(\lambda+1)^2}.$$

The numerical results of Table 3 coincide with the similar ones obtained by an application of the previous explicit expressions of the measures for the specific values of $\alpha = \alpha_1 = \alpha_2$ and λ. To complete this example, we have to point out and to explain why several columns in Table 3 are equal. This is the case when at least one of the parameters α_1 or α_2 is equal to zero. However, in such a case, the Pickands dependence function of the Marshall-Olkin extreme value copula coincides with the independence copula $A(t) = 1$ and then, equations (3.3) and (3.4), entail that the respective Tsallis' type extreme value copula depends only on λ, $\lambda > 0, \lambda \neq 1$, something which is moreover confirmed by the last formula for $\mathcal{CE}_{Ts,\lambda}(C_A)$ above, for $\alpha = 0$. ■

3.2. Properties of the measures

The interest is now focused in the investigation of some of the properties of the measures which are presented in Proposition 3.1. In this context, a more systematic investigation of the presented entropy measures is attempted, in the sequel, by incorporating basic properties such as monotonicity and continuity, among others. The said properties support the conclusion that the measures introduced here are more related to concordance measures (like Spearman's rho) than to extent of information they contain.

The first property investigates the range of values of the measures $\mathcal{CE}_{Sh}(C_A)$, $\mathcal{CI}_\lambda(C_A)$ and $\mathcal{CI}_1(C_A)$ for Pickands dependence function A. Following the formulation of Kamnitui et al. (2019), p. 923, let $\mathcal{A}$ be the class of Pickands dependence functions A, defined above. Based on this paper, for $A_i \in \mathcal{A}, i = 1, 2$, we write $A_1 \leq A_2$ if $A_1(t) \leq A_2(t)$, for all $t \in [0, 1]$ and $A_1 < A_2$ if the inequality is strict in at least one point and hence, by continuity, on an interval. Taking into account this ordering, the proposition that follows states the range of values of the above mentioned measures. The proof is provided in the Appendix A.3 and it is shown there that $\mathcal{CE}_{Sh}(C_A)$ and $\mathcal{CI}_\lambda(C_A)$ are decreasing functions of A in the sense of Kamnitui et al. (2019), p. 923, mentioned above. Similar discussion is also provided in Theorem 3.4 in Ansari and Rockel (2023).

Proposition 3.2. *The measures $\mathcal{CE}_{Sh}(C_A)$, $\mathcal{CI}_\lambda(C_A)$ and $\mathcal{CI}_1(C_A)$ are maximized for $A(t) = \max\{t, 1-t\}$ and they are minimized for $A(t) = 1$, for $t \in [0, 1]$, while their range* *of values is as follows,*

$$\begin{aligned} 0.25 &\le \mathcal{CE}_{Sh}(C_A) \le \frac{5}{18}, \\ \frac{1}{(1+\lambda)^2} &\le \mathcal{CI}_\lambda(C_A) \le \frac{2}{(1+\lambda)(2+\lambda)}, \ \lambda > 0, \\ \frac{1}{4} &\le \mathcal{CI}_1(C_A) \le \frac{1}{3}. \end{aligned}$$

In the sequel, the interest is focused on the study of some properties of Shannon's and Tsallis' extreme value copula entropies or the generating type function $\mathcal{CI}_\lambda(C_A)$. The said properties are related to that of monotonicity with respect to the concordance ordering of copulas, continuity with respect to pointwise convergence of copulas, in the spirit of Definition 5.1.7 of Nelsen (2006), p. 168, or the Definition 2.8 in Joe (2015), p. 54. These properties are based on Scarsini (1984) axioms that should be obeyed by a measure of concordance, while similar properties of measures of concordance are also discussed in Nelsen (2002), Edwards et al. (2005), Liebscher (2014), Fuchs (2016) and references appeared therein.

The next proposition formulates monotonicity of the measures with respect to the concordance ordering of copulas. To proceed in this direction, let C_{A_1} and C_{A_2} are extreme value copulas with $C_{A_1} \prec C_{A_2}$ which means that $C_{A_1}(u,v) \le C_{A_2}(u,v)$, $0 < u,v \le 1$, $(u,v) \ne (1,1)$, in view of Definition 2.8.1 of Nelsen (2006), p. 39. In this setting, following the formulation of Kamnitui et al. (2019), p. 923, mentioned above, if $C_{A_1} \prec C_{A_2}$, that is if $C_{A_1}(u,v) \le C_{A_2}(u,v)$, $0 < u,v \le 1$, $(u,v) \ne (1,1)$, then $A_2 \le A_1$, taking into account that $\ln(uv) < 0$ in (2.9). Then, the monotonicity property is formulated as follows.

Proposition 3.3. *If C_{A_1} and C_{A_2} are extreme value copulas with $C_{A_1} \prec C_{A_2}$, then,*

$$\mathcal{CE}_{Sh}(C_{A_1}) \le \mathcal{CE}_{Sh}(C_{A_2}) \text{ and } \mathcal{CI}_\lambda(C_{A_1}) \le \mathcal{CI}_\lambda(C_{A_2}),\ \lambda > 0,\ \lambda \ne 1.$$

The proof is straightforward. The ordering $C_{A_1} \prec C_{A_2}$ leads to $A_2 \le A_1$ and the result follows immediately from Proposition 3.1, by taking into account the fact that the non-negative function $z_1(w) = \frac{2w}{(1+w)^3}$, is decreasing for $w \in [1/2, 1]$ and the same is happen for the function $z_2(w) = \frac{1}{(1+\lambda w)^2}$, for $w \ge 0$, $\lambda > 0, \lambda \ne 1$. Moreover, in view of (3.2) and (3.4), $\mathcal{CE}_{Sh}$ and $\mathcal{CI}_\lambda$ are related with the functions z_1 and z_2 as follows,

$$\mathcal{CE}_{Sh}(C_A) = \int_0^1 z_1(A(t))dt \ \text{ and } \ \mathcal{CI}_\lambda(C_A) = \int_0^1 z_2(A(t))dt,\ \lambda > 0,$$

which completes the proof.

The next property of the measures proposed here is related to their continuity with respect to pointwise convergence of copulas. In this frame, let $(X_n, Y_n), n = 1, 2, ...$ be a sequence of continuous random variables with extreme value copula C_{A_n} which converges pointwise to an extreme value copula C_A. In this setting,

$$C_{A_n}(u,v) = \exp\left[\ln(uv) \cdot A_n\left\{\frac{\ln v}{\ln(uv)}\right\}\right],\quad 0 < u,v \le 1,\ (u,v) \ne (1,1),$$

and

$$C_A(u,v) = \exp\left[\ln(uv) \cdot A\left\{\frac{\ln v}{\ln(uv)}\right\}\right],\quad 0 < u,v \le 1,\ (u,v) \ne (1,1),$$

for Pickands dependence functions A_n and A, where $A_n, A : [0,1] \to [1/2,1]$ are convex functions which satisfy $A_n(0) = A(0) = A_n(1) = A(1) = 1$, $\max\{t, 1-t\} \leq A_n(t), A(t) \leq 1$, for each $t \in [0,1]$. Given that copulas C_{A_n} converge pointwise to the copula C_A, it is immediate to see that

$$\lim_{n\to\infty} A_n(t) = A(t),\ t \in [0,1],$$

on the condition that the limits exist. Based on above formulation, the next proposition states that Shannon and Tsallis' type extreme value copula entropies are continuous with respect to pointwise convergence of copulas. The proof is outlined in the Appendix A.4.

Proposition 3.4. *If $(X_n, Y_n), n = 1, 2, ...$ is a sequence of continuous random variables with extreme value copula C_{A_n} which converges pointwise to an extreme value copula C_A, then*

$$\begin{aligned}
\lim_{n\to\infty} \mathcal{CE}_{Sh}(C_{A_n}) &= \mathcal{CE}_{Sh}(C_A), \\
\lim_{n\to\infty} \mathcal{CI}_\lambda(C_{A_n}) &= \mathcal{CI}_\lambda(C_A),\ \lambda \geq 1, \\
\lim_{n\to\infty} \mathcal{CE}_{Ts,\lambda}(C_{A_n}) &= \mathcal{CE}_{Ts,\lambda}(C_A),\ \lambda \geq 1.
\end{aligned}$$

It should be mentioned at this point that similar continuity property is obeyed by Spearman's rho (cf. Nelsen (2006), Theorem 5.1.9, p. 169) which is directly connected with $\mathcal{CI}_\lambda(C_{A_n})$ and $\mathcal{CI}_\lambda(C_A)$, for $\lambda = 1$, in view of Remark 1(c).

Invariance of copulas, under strictly increasing transformations of the corresponding continuous random variables, is an important property of copulas with particular usefulness in nonparametric statistics (cf. Nelsen (2006), p. 25). Based on Theorem 2.4.3, p. 25 in Nelsen (2006), if X and Y are two continuous random variables with copula C_{XY} and α, β are strictly increasing functions on the range of X and Y, respectively, then $C_{XY} = C_{\alpha(X)\beta(Y)}$, that is, C_{XY} is invariant under strictly increasing transformations of X and Y. The same is also happen in the case of extreme value bivariate copulas and the question which is raised at this point is related to the invariance of the corresponding measures of Proposition 3.1, namely the invariance of $\mathcal{CE}_{Sh}$, $\mathcal{CE}_{Ts,\lambda}$ and $\mathcal{CI}_\lambda$, under strictly increasing transformations of the underling random variables. Invariance is a characteristic property of a measure of concordance, stated in the paper by Scarsini (1984) and Theorem 5.1.8, p. 169 and Definition 2.8, p. 54 of Nelsen (2006) and Joe (2015), respectively. Spearman's rank correlation ρ_S, defined by (3.7), obeys invariance, in the sense described above, according to the previous citations. Motivated by Theorem 5.1.8, p. 169 in Nelsen (2006), next proposition formulates invariance of $\mathcal{CE}_{Sh}$, $\mathcal{CE}_{Ts,\lambda}$ and $\mathcal{CI}_\lambda$, under strictly increasing transformations of the underling random variables. The proof is immediately obtained in view of invariance of copulas, under strictly increasing transformations of the underlined continuous random variables.

Proposition 3.5. *If $\alpha(X)$ and $\beta(Y)$ are almost surely strictly monotone functions on range of X and Y respectively, then,*

$$\mathcal{CE}_{Sh}\left(C_A^{\alpha,\beta}\right) = \mathcal{CE}_{Sh}(C_A),\ \mathcal{CE}_{Ts,\lambda}\left(C_A^{\alpha,\beta}\right) = \mathcal{CE}_{Ts,\lambda}(C_A),\ \mathcal{CI}_\lambda\left(C_A^{\alpha,\beta}\right) = \mathcal{CI}_\lambda(C_A),$$

where the superscript α, β in the notation of $C_A^{\alpha,\beta}$ in the measures is used to denote the respective measure based on the extreme value copula of $\alpha(X)$ and $\beta(Y)$.

Remark 2. **(a)** Proposition 3.3 shows that Shannon type extreme value copula $\mathcal{CE}_{Sh}(C_A)$ and the generating type function $\mathcal{CI}_\lambda(C_A)$ are monotone with respect to the concordance ordering of the respective extreme value copulas. On the other hand, extreme value copulas are positively quadrant dependent (PQD) in view of Gudendorf and Segers (2010), p. 131. In this context, suppose that C_{A_2} is more concordant than C_{A_1} ($C_{A_1} \prec C_{A_2}$), or C_{A_2} is more PQD than C_{A_1}. Then, according to the Definition 5.7.2 on p. 223 in Nelsen (2006) and the discussion provided just before this definition, $\mathcal{CE}_{Sh}(C_{A_1}) \leq \mathcal{CE}_{Sh}(C_{A_2})$ and $\mathcal{CI}_\lambda(C_{A_1}) \leq \mathcal{CI}_\lambda(C_{A_2})$. Hence, small values of these measures correspond to less concordant or less PQD extreme value copulas. Rényi's type extreme value copula obeys a similar ordering because this measure is related with $\mathcal{CI}_\lambda$ by the formula $\mathcal{CE}_R(C_A) = (1-\lambda)^{-1}\log \mathcal{CI}_\lambda(C_A)$, $\lambda > 0, \lambda \neq 1$. In particular if the order λ of the measure is less than 1, then the ordering of $\mathcal{CE}_R$ is preserved. For $\lambda > 1$ the ordering of $\mathcal{CE}_R$ is reversed. Hence, the measures $\mathcal{CE}_{Sh}$, $\mathcal{CI}_\lambda$, $\lambda > 0, \lambda \neq 1$ and $\mathcal{CE}_R$, $0 < \lambda < 1$, quantify the degree of PQD of an extreme value copula with small values of these measures to correspond to less PQD copulas. It was not possible to derive a similar behavior of Tsallis' type entropy (3.3) because the function that generates this type of entropy, $\varphi(x) = \frac{x-x^\lambda}{\lambda-1} = -x^\lambda L_\lambda(x), \lambda > 0, \lambda \neq 1$, $x \in (0,1]$, is not decreasing in general, with $L_\lambda(x) = \frac{x^{1-\lambda}-1}{1-\lambda}$ to be the function which was defined by (2.3), p. 1029 in Asadi et al. (2017).

(b) In view of Proposition 3.2, $\mathcal{CE}_{Sh}(C_A)$ is maximized for $A(t) = \max\{t, 1-t\}$ and the maximum value of $\mathcal{CE}_{Sh}$ is

$$2\int_0^1 \frac{\max\{t,1-t\}}{[1+\max\{t,1-t\}]^3}dt = 0.277778.$$

That is, $\mathcal{CE}_{Sh}(C_A)$ is maximized in the case of complete dependence (comonotonicity) $C_A(u,v) = u \wedge v = \min\{u,v\}$ (cf. Gudendorf and Segers (2010), p. 131).

(c) In a similar manner, $\mathcal{CE}_{Sh}(C_A)$ is minimized for $A(t) = 1$, that is in the case of independence, with independence copula $\Pi(u,v) = uv$ and the minimum value of $\mathcal{CE}_{Sh}$ is $2\int_0^1 (1/8)dt = 0.25$. This is in harmony with part (a) of this remark, in view of the claim that if two random variables are PQD, then the graph of the copula of the said random variables "lies on or above the graph of the independence copula", cf. Nelsen (2006), p. 188. On the basis of this discussion, Shannon's type entropy of extreme value copula quantifies, in this sense, the degree of the PQD of the associated random variables. The smaller $\mathcal{CE}_{Sh}$ is, the less PQD the associated random variables would be.

(d) $\mathcal{CE}_{Sh}(C_A)$ does not recognize between right or left skew copulas, something which was observed in Example 1. It is also confirmed, for example, in the case $A(t) = \max\left\{\frac{x+1}{2}, 1-x\right\}$ which is right skew and the case $A(t) = \max\left\{x, 1-\frac{x}{2}\right\}$ which is left skew (cf. Eschenburg (2013), p. 76 and the cited there paper by Durante and Mesiar (2010)). In both cases the common value of $\mathcal{CE}_{Sh}(C_A)$ is 0.27.

(e) Proposition 3.5 formulates the invariance of the considered here measures under strictly increasing transformations of the underling random variables. Invariance is a characteristic property of a measure of concordance. Moreover, the definition and the study of a more general group of transformations on the collection of all bivariate copulas has been provided in the papers by Fuchs (2014) and Fuchs and Schmidt (2014).

4. Extreme Value Copula Divergences

The interest is focused, in this subsection, on the divergence between two bivariate extreme value copulas, that is on pseudo-distances of the form (2.2) or (2.4)-(2.7) between two bivariate extreme value copulas of the form (2.9). In this context, following the formulation of Kamnitui et al. (2019), p. 923, let $\mathcal{A}$ be the class of Pickands dependence functions, i.e., the set of functions A, defined above by (2.9). Let $A_i \in \mathcal{A}, i = 1, 2$, be two Pickands dependence functions with associated bivariate extreme value copulas

$$C_{A_i}(u,v) = \exp\left[\ln(uv) \cdot A_i\left\{\frac{\ln v}{\ln(uv)}\right\}\right], \quad 0 < u, v \leq 1, \ (u,v) \neq (1,1), \ i = 1, 2. \tag{4.1}$$

Then, it is of interest to define a distance type measure between C_{A_1} and C_{A_2}. In this frame, based on the measures (2.2) and (2.4)-(2.7) and on the fact that

$$\mathcal{CI}_1(C_{A_i}) = \int\limits_{[0,1]^2} C_{A_i}(u,v)dudv = \int\limits_0^1 \frac{1}{[1 + A_i(t)]^2}dt = \frac{\rho_S(A_i) + 3}{12}, i = 1, 2,$$

(cf. (3.6), (3.7) and (3.8)) we consider the following divergence type measures between C_{A_1} and C_{A_2}:

Csiszár type extreme value copulas ϕ-divergences,

$$\mathcal{CD}_\phi(C_{A_1}, C_{A_2}) = \int\limits_{[0,1]^2} C_{A_2}(u,v)\phi\left(\frac{C_{A_1}(u,v)}{C_{A_2}(u,v)}\right)dudv - \mathcal{CI}_1(C_{A_2})\phi\left(\frac{\mathcal{CI}_1(C_{A_1})}{\mathcal{CI}_1(C_{A_2})}\right), \tag{4.2}$$

where $\phi : (0, \infty) \to \mathbb{R}$ is a real valued convex function and $\phi \in \Phi$, defined by (2.3).

Kullback-Leibler type extreme value copulas divergence,

$$\mathcal{CD}_{KL}(C_{A_1}, C_{A_2}) = \int\limits_{[0,1]^2} C_{A_1}(u,v)\ln\left(\frac{C_{A_1}(u,v)}{C_{A_2}(u,v)}\right)dudv - \mathcal{CI}_1(C_{A_1})\ln\left(\frac{\mathcal{CI}_1(C_{A_1})}{\mathcal{CI}_1(C_{A_2})}\right), \tag{4.3}$$

and

Cressie-Read λ-power type extreme value copulas divergence,

$$\begin{aligned}&\mathcal{CD}_\lambda(C_{A_1}, C_{A_2}) \\ &= \tfrac{1}{\lambda(\lambda+1)}\left(\int\limits_{[0,1]^2} C_{A_2}(u,v)\left(\tfrac{C_{A_1}(u,v)}{C_{A_2}(u,v)}\right)^{\lambda+1}dudv - \mathcal{CI}_1(C_{A_2})\left(\tfrac{\mathcal{CI}_1(C_{A_1})}{\mathcal{CI}_1(C_{A_2})}\right)^{\lambda+1}\right),\end{aligned} \tag{4.4}$$

where $\lambda \in \mathbb{R}, \lambda \neq 0, -1$.

Moreover, let's define a copula type version of the density power divergence of Basu et al. (1998), namely, **the density power type extreme value copulas divergence**, which is defined, in view of (2.6) and (2.8), by,

$$\begin{aligned}&Cd_a(C_{A_1}, C_{A_2}) \\ &= \int\limits_{[0,1]^2}\left\{C_{A_2}^{1+a}(u,v) - \left(1 + \tfrac{1}{a}\right)C_{A_2}^a(u,v)C_{A_1}(u,v) + \tfrac{1}{a}\ C_{A_1}^{1+a}(u,v)\right\}dudv, \ a > 0.\end{aligned} \tag{4.5}$$

Cd_a is not obtained from (4.2) for a particular choice of the convex function ϕ while its limiting case, $\lim_{a\to 0} \mathcal{C}d_a(C_{A_1}, C_{A_2}) = CKL(C_{A_1}, C_{A_2})$, cf. (2.7), is defined by

$$(4.6) \qquad CKL(C_{A_1}, C_{A_2}) = \int\limits_{[0,1]^2} C_{A_1}(u,v) \ln\left(\frac{C_{A_1}(u,v)}{C_{A_2}(u,v)}\right) dudv + \mathcal{CI}_1(C_{A_2}) - \mathcal{CI}_1(C_{A_1}).$$

Next proposition provides with explicit expressions of the above divergence type measures.

Proposition 4.1. *Let the extreme value copulas (4.1). Then, based on (3.8) and (4.2)-(4.6),*
(a) For a real valued convex function $\phi : (0,\infty) \to \mathbb{R}$, with $\phi \in \Phi$, defined by (2.3),
(4.7)

$$\mathcal{CD}_\phi(C_{A_1}, C_{A_2}) = \int\limits_{[0,1]^2} \left(\ln\frac{1}{v}\right) v^{\frac{1-t+A_2(t)}{t}} \frac{1}{t^2} \phi\left(v^{\frac{A_1(t)-A_2(t)}{t}}\right) dvdt - \frac{\rho_S(A_2)+3}{12} \phi\left(\frac{\rho_S(A_1)+3}{\rho_S(A_2)+3}\right).$$

(b) The Kullback-Leibler type extreme value copulas divergence is given by,

$$(4.8) \qquad \mathcal{CD}_{KL}(C_{A_1}, C_{A_2}) = 2\int_0^1 \frac{A_2(t) - A_1(t)}{[1 + A_1(t)]^3} dt - \frac{\rho_S(A_1)+3}{12} \ln\left(\frac{\rho_S(A_1)+3}{\rho_S(A_2)+3}\right).$$

(c) For $\lambda \in \mathbb{R}, \lambda \neq 0, -1$, the Cressie-Read λ-power type extreme value copulas divergence is given by

$$(4.9) \qquad \begin{aligned} &\mathcal{CD}_\lambda(C_{A_1}, C_{A_2}) \\ &= \tfrac{1}{\lambda(\lambda+1)} \left(\int_0^1 \tfrac{1}{[1+A_1(t)+\lambda(A_1(t)-A_2(t))]^2} dt - \tfrac{\rho_S(A_2)+3}{12} \left(\tfrac{\rho_S(A_1)+3}{\rho_S(A_2)+3}\right)^{\lambda+1} \right). \end{aligned}$$

(d) For $a > 0$, the density power type extreme value copulas divergence is given by,

$$(4.10) \qquad \begin{aligned} \mathcal{C}d_a(C_{A_1}, C_{A_2}) = & \int_0^1 \frac{1}{[1+(1+a)A_2(t)]^2} dt + \frac{1}{a}\int_0^1 \frac{1}{[1+(1+a)A_1(t)]^2} dt \\ & - \left(1+\frac{1}{a}\right)\int_0^1 \frac{1}{[1+A_1(t)+aA_2(t)]^2} dt. \end{aligned}$$

(e) The $CKL(C_{A_1}, C_{A_2})$ extreme value copulas divergence is given by,

$$(4.11) \qquad CKL(C_{A_1}, C_{A_2}) = 2\int_0^1 \frac{A_2(t) - A_1(t)}{[1 + A_1(t)]^3} dt + \frac{1}{12}\left(\rho_S(A_2) - \rho_S(A_1)\right).$$

The proof of parts (c) and (d) of the proposition is outlined in subsection A.5, in the Appendix A. The proof of the remaining parts is quite similar and they are, therefore, omitted. ▲

Remark 3. **(a)** All the measures of the previous proposition obey the non-negativity and identity of indiscernibles property, a terminology which is conveyed by Weller-Fahy et al. (2015) and it means that the said measures are non-negative and they attain their minimum value if and only if the underline copulas are coincide. Hence, all these divergence type measures can by considered as quasi-distances or statistical distances between the underlined copulas.

(b) It is well known (cf. Nelsen (2006) or Gudendorf and Segers (2010), p. 131), that the value $A(t) = 1$ corresponds to independence. Then, based on (a), it is clear that if anyone of the five measures of the above proposition is applied for $A_2(t) = 1$, then they lead to a measure of the distance from independence. That is, they formulate on how close stands the copula C_{A_1} from the copula that expresses independence of the associated random variables. In this setting, they can be maybe applied to develop tests of independence.

(c) The cases (b) and (e) of the proposition concentrate to two different forms of Kullback-Leibler's type extreme value copulas divergences $\mathcal{CD}_{KL}(C_{A_1}, C_{A_2})$ and $CKL(C_{A_1}, C_{A_2})$. The logic of the construction of these two forms of Kullback-Leibler's type divergences between two extreme value copulas is discussed in detail in the first paragraph of p. 303 in Zografos (2023) for distribution functions in place of copulas. Based now in a well-known property of the logarithmic function, namely $x \ln(x/y) \geq x - y$, for $x, y > 0$, we can easily see that $\mathcal{CD}_{KL}(C_{A_1}, C_{A_2}) \leq CKL(C_{A_1}, C_{A_2})$, comparing (4.8) and (4.11). Following Zografos (2023), p. 306, it is clear, in view of the previous inequality, that the measure $CKL(C_{A_1}, C_{A_2})$ over estimate the divergence or the quasi-distance between the respective copulas in comparison with the respective Kullback-Leibler's type extreme value copulas divergence $\mathcal{CD}_{KL}(C_{A_1}, C_{A_2})$, defined by (4.8).

Example 3. Gumbel extreme value copula

Let's complete this section with an example which concentrates on Gumbel copula and its divergence, in the Kullback-Leibler sense of Proposition 4.1(b), from the copula that corresponds to independence. In this context, the Pickands dependence function of Gumbel copula is defined (cf. Nelsen (2006), p. 98, Example 3.23) by $A_1(t) = [t^\theta + (1-t)^\theta]^{1/\theta}$, $\theta \geq 1$, while the Pickands dependence function of the independence copula $\Pi(u, v) = uv$ is $A_2(t) = 1$. In this framework, Kullback-Leibler type extreme value copulas divergence (cf. Proposition 4.1(b)) is defined by,

$$\begin{aligned}
\mathcal{CD}_{KL}(C_{A_1}, \Pi) &= 2\int_0^1 \frac{1 - A_1(t)}{[1 + A_1(t)]^3} dt - \frac{\rho_S(A_1) + 3}{12} \ln\left(\frac{\rho_S(A_1) + 3}{\rho_S(A_2) + 3}\right) \\
&= 2\int_0^1 \frac{1 - [t^\theta + (1-t)^\theta]^{1/\theta}}{\left[1 + [t^\theta + (1-t)^\theta]^{1/\theta}\right]^3} dt - \frac{\rho_S(A_1) + 3}{12} \ln\left(\frac{\rho_S(A_1) + 3}{3}\right),
\end{aligned}$$

where

$$\rho_S(A_i) = 12 \int_{[0,1]^2} C_{A_i}(u, v) du dv - 3 = 12\mathcal{CI}_1(C_{A_i}) - 3 = 12 \int_0^1 \frac{1}{[1 + A_i(t)]^2} dt - 3, \; i = 1, 2,$$

with

$$\rho_S(A_2) = 0.$$

Table 4 presents the values of $\mathcal{CD}_{KL}$ for several values of the dependence parameter θ. We observe that the minimum value of $\mathcal{CD}_{KL}$, equal to 0, is attained in the case of independence, $\theta = 1$, which is expected in view of Remark 3(b).

	$\mathcal{CD}_{KL}$ for values of θ						
	$\theta = 1$	$\theta = 1.5$	$\theta = 2$	$\theta = 3$	$\theta = 8$	$\theta = 15$	$\theta = 50$
$\mathcal{CE}_{Sh}(C_A)$	0	0.0025	0.0056	0.0095	0.0141	0.0149	0.0152

Table 4: Kullback-Leibler divergence between Gumbel and independence copula.

5. Conclusions

The main aim is focused on cumulative entropies and cumulative divergences in the light of bivariate extreme value copulas. In this way it is developed a bridge between information theory and copula theory. On the other hand, based on the conclusions in the paper by Zografos (2023), divergences or statistical distances have been recently extensively used (cf. Pardo (2006) and Basu et al. (2011)) to present again and to rebuild robust statistical inference by means of distance type or divergence type measures. Hence, it is expected that the entropy and divergence measures between copulas, discussed above, can be used in this direction. This will maybe be the subject of a future work. Moreover, broad classes of copulas include the classes of elliptical or Archimedean copulas and it would be of interest to developed similar investigations and studies on the basis of these classes of copulas.

A. Appendix: Proofs of the statements

A.1. Proof of Proposition 3.1

(a) For the extreme value copula (2.9), the extreme value copula φ-entropy (3.1) is

$$\mathcal{CE}_{\varphi}(C_A) = \int\limits_{[01]^2} \varphi \left\{ \exp \left[\ln(uv) \cdot A \left\{ \frac{\ln v}{\ln(uv)} \right\} \right] \right\} dudv. \tag{1.1}$$

Let, $t = \frac{\ln v}{\ln(uv)}$, $(u, v) \in (0, 1]^2, (u, v) \neq (1, 1)$. Then, $t \in [0, 1)$ and

$$\ln(uv) = \frac{\ln v}{t}. \tag{1.2}$$

Moreover,

$$\frac{dt}{du} = -\frac{\ln v}{u(\ln(uv))^2} = -\frac{(\ln v)^2}{u(\ln(uv))^2 \ln v} \quad \text{and} \quad \frac{dt}{du} = -\frac{t^2}{u(\ln v)} \quad \text{or} \quad du = -\frac{u \ln v}{t^2} dt. \tag{1.3}$$

On the other hand,

$$t = \frac{\ln v}{\ln(uv)}, \text{ or } \ln u = \frac{(1-t)\ln v}{t}, \text{ or } u = v^{(1-t)/t},$$

and taking into account (1.3),

$$du = -v^{(1-t)/t}\frac{\ln v}{t^2}dt. \tag{1.4}$$

Then, equations (1.1), (1.2) and (1.4), lead to the general formula for $\mathcal{CE}_\varphi(C_A)$,

$$\mathcal{CE}_\varphi(C_A) = -\int\limits_{[0,1]^2} v^{\frac{1-t}{t}}\frac{\ln v}{t^2}\varphi\left\{\exp\left[\frac{\ln v}{t}A(t)\right]\right\}dvdt.$$

(b) Based on $u = v^{(1-t)/t}$, (1.2) and (1.4),

$$\begin{aligned}\mathcal{CE}_{Sh}(C_A) &= -\int\limits_{[0,1]^2} \exp\left[\ln(uv)\cdot A\left\{\frac{\ln v}{\ln(uv)}\right\}\right]\ln\left\{\exp\left[\ln(uv)\cdot A\left\{\frac{\ln v}{\ln(uv)}\right\}\right]\right\}dudv\\ &= -\int\limits_{[0,1]^2} \ln(uv)\cdot A\left\{\frac{\ln v}{\ln(uv)}\right\}\exp\left[\ln(uv)\cdot A\left\{\frac{\ln v}{\ln(uv)}\right\}\right]dudv\\ &= \int\limits_{[0,1]^2}\frac{1}{t^3}v^{\frac{1-t}{t}}(\ln v)^2 A(t)\exp\left\{\ln v^{A(t)/t}\right\}dvdt,\end{aligned}$$

and after standard manipulations,

$$\mathcal{CE}_{Sh}(C_A) = \int\limits_0^1 \frac{1}{t^3}A(t)\left\{\int\limits_0^1 v^{\frac{1+A(t)}{t}-1}(\ln v)^2 dv\right\}dt. \tag{1.5}$$

Based on formula 4.272-6, on p. 550 in Gradshteyn and Ryzhik (2007),

$$\int\limits_0^1\left(\log\frac{1}{x}\right)^{\mu-1}x^{\nu-1}dx = \frac{1}{\nu^\mu}\Gamma(\mu),\ Re\,\mu > 0,\ Re\,\nu > 0. \tag{1.6}$$

Applying this formula for $\mu = 3$ and $\nu = \frac{1+A(t)}{t}$, we obtain

$$\int\limits_0^1 v^{\frac{1+A(t)}{t}-1}(\ln v)^2 du = \frac{1}{\left(\frac{1+A(t)}{t}\right)^3}\Gamma(3) = \frac{2t^3}{[1+A(t)]^3}. \tag{1.7}$$

Then, (1.5) and (1.7) lead to the desired result,

$$\mathcal{CE}_{Sh}(C_A) = \int\limits_0^1 \frac{1}{t^3}A(t)\frac{2t^3}{[1+A(t)]^3}dt = 2\int\limits_0^1\frac{A(t)}{[1+A(t)]^3}dt.$$

(c) It is enough to obtain the expression for $\mathcal{CI}_\lambda(C_A)$, which is obtained by following exactly the same steps. For each $\lambda > 0$, based again on (1.2) and (1.4),

$$\begin{aligned}
\mathcal{CI}_\lambda(C_A) &= \int\limits_{[0,1]^2} C_A^\lambda(u,v)dudv = \int\limits_{[0,1]^2} \exp\left[\lambda \ln(uv) \cdot A\left\{\frac{\ln v}{\ln(uv)}\right\}\right] dudv \\
&= -\int\limits_{[0,1]^2} v^{\frac{1-t}{t}} \frac{\ln v}{t^2} \exp\left[\frac{\lambda A(t)}{t} \ln v\right] dvdt \\
&= -\int\limits_{[0,1]^2} v^{\frac{1-t}{t}} \frac{\ln v}{t^2} \exp\left[\ln v^{\frac{\lambda A(t)}{t}}\right] dvdt \\
&= -\int\limits_{[0,1]^2} v^{\frac{1-t}{t}} \frac{\ln v}{t^2} v^{\frac{\lambda A(t)}{t}} dvdt,
\end{aligned}$$

and therefore

$$\mathcal{CI}_\lambda(C_A) = \int\limits_0^1 \frac{1}{t^2} \left\{ \int\limits_0^1 \left(\ln \frac{1}{v}\right) v^{\frac{1+\lambda A(t)}{t} - 1} dv \right\} dt. \tag{1.8}$$

Applying again formula 4.272-6, on p. 550 in Gradshteyn and Ryzhik (2007), (1.6), for $\mu = 2$ and $\nu = \frac{1+\lambda A(t)}{t}$, we obtain

$$\int\limits_0^1 \left(\ln \frac{1}{v}\right) v^{\frac{1+\lambda A(t)}{t} - 1} dv = \frac{t^2}{[1 + \lambda A(t)]^2}. \tag{1.9}$$

Equations (1.8) and (1.9) lead to the desired result

$$\mathcal{CI}_\lambda(C_A) = \int\limits_0^1 \frac{1}{t^2} \frac{t^2}{[1+\lambda A(t)]^2} dt = \int\limits_0^1 \frac{1}{[1+\lambda A(t)]^2} dt, \quad \lambda > 0.$$

▲

A.2. Evaluation of the measures for Marshall-Olkin extreme value copula of Example 2

Based on (3.4)

$$\begin{aligned}
\mathcal{CI}_\lambda(C_A) &= \int\limits_0^1 \frac{1}{[1+\lambda A(t)]^2} dt = \int\limits_0^{1/2} \frac{1}{[1+\lambda A(t)]^2} dt + \int\limits_{1/2}^1 \frac{1}{[1+\lambda A(t)]^2} dt, \ \lambda > 0 \\
&= \int\limits_0^{1/2} \frac{1}{[1+\lambda(1-\alpha t)]^2} dt + \int\limits_{1/2}^1 \frac{1}{[1+\lambda(1-\alpha(1-t))]^2} dt \\
&= \int\limits_0^{1/2} \frac{1}{[1+\lambda-\lambda\alpha t)]^2} dt + \int\limits_{1/2}^1 \frac{1}{[1+\lambda-\lambda\alpha+\lambda\alpha t]^2} dt.
\end{aligned}$$

Based on formula 2.113, p. 69 in Gradshteyn and Ryzhik (2007), $\int \frac{1}{[a+bx]^2}dx = -\frac{1}{b[a+bx]}$, thus

$$\begin{aligned}\int_0^{1/2} \frac{1}{[1+\lambda-\lambda\alpha t)]^2}dt &= \left.\frac{1}{\lambda\alpha(1+\lambda-\lambda\alpha t)}\right|_0^{1/2} \\ &= -\frac{1}{\lambda\alpha(1+\lambda)} + \frac{1}{\lambda\alpha(1+\lambda-\lambda\alpha/2)} \\ &= \frac{1}{(1+\lambda)(2+2\lambda-\lambda\alpha)},\end{aligned}$$

and the same is the explicit expression of the second integral $\int_{1/2}^{1} \frac{1}{[1+\lambda-\lambda\alpha+\lambda\alpha t]^2}dt$. Hence,

$$\mathcal{CI}_\lambda(C_A) = \frac{2}{(1+\lambda)(2+2\lambda-\lambda\alpha)}.$$

For $\lambda = 1$ it is immediate to see that,

$$\mathcal{CI}_1(C_A) = \frac{1}{4-\alpha},$$

and then, based on (3.3),

$$\mathcal{CE}_{Ts,\lambda}(C_A) = \frac{1}{\lambda-1}\left[\frac{1}{4-\alpha} - \frac{2}{(1+\lambda)[2-\lambda(\alpha-2)]}\right].$$

Shannon type Marshall-Olkin extreme value copula is given, in view of (3.5), by

$$\mathcal{CE}_{Sh}(C_A) = \lim_{\lambda\to1}\mathcal{CE}_{Ts,\lambda}(C_A) = \frac{8-3\alpha}{2(4-\alpha)^2}.$$

▲

A.3. Proof of Proposition 3.2

The measures considered in this proposition are maximized for $A(t) = \max(t, 1-t)$ and they are minimized for $A(t) = 1$ because the non-negative function $z_1(w) = \frac{2w}{(1+w)^3}$, is decreasing for $w \in [1/2, 1]$ and the same is happen for the function $z_2(w) = \frac{1}{(1+\lambda w)^2}$, for $w \geq 0$, $\lambda > 0, \lambda \neq 1$, while $\mathcal{CE}_{Sh}$ and $\mathcal{CI}_\lambda$ are related with the functions z_1 and z_2 as follows,

$$\mathcal{CE}_{Sh}(C_{A_n}) = \int_0^1 z_1(A(t))dt \quad\text{and}\quad \mathcal{CI}_\lambda(C_{A_n}) = \int_0^1 z_2(A(t))dt,\ \lambda > 0,$$

for Pickands dependence function A with $\max(t, 1-t) \leq A(t) \leq 1$, for each $t \in [0, 1]$.

The range of values is outlined for $\mathcal{CI}_\lambda$. The proof for the other measures are derived in a similar manner. For $A(t) = \max(t, 1-t)$, the maximum value of $\mathcal{CI}_\lambda(C_A)$ is,

$$\mathcal{CI}_\lambda(C_{\max(t,1-t)}) = \int_0^1 \frac{1}{[1+\lambda\max(t,1-t)]^2}dt = \int_0^{1/2}\frac{1}{[1+\lambda(1-t)]^2}dt + \int_{1/2}^1 \frac{1}{[1+\lambda t]^2}dt,$$

and applying the transformation $y = 1 - t$ in the first integral

$$\mathcal{CI}_\lambda(C_{\max(t,1-t)}) = 2\int_{1/2}^{1} \frac{1}{[1+\lambda t]^2}dt.$$

Based on 2.113, p. 69 in Gradshteyn and Ryzhik (2007), $\int \frac{1}{[a+bx]^2}dx = -\frac{1}{b[a+bx]}$, thus

$$\mathcal{CI}_\lambda(C_{\max(t,1-t)}) = 2\int_{1/2}^{1} \frac{1}{[1+\lambda t]^2}dt = \frac{2}{(1+\lambda)(2+\lambda)}.$$

For $A(t) = 1$, the minimum value of $\mathcal{CI}_\lambda(C_A)$ is,

$$\mathcal{CI}_\lambda(C_1) = \int_0^1 \frac{1}{[1+\lambda]^2}dt = \frac{1}{(1+\lambda)^2},$$

and the proof is completed. ▲

A.4. Proof of Proposition 3.4

The proof follows the proof of Theorem 5.1.9 of Nelsen (2006), p. 169. Let's start with the $\lim_{n\to\infty}\mathcal{CI}_\lambda(C_{A_n})$. The main part of the proof shows that a positive power of an extreme value copula satisfies a Lipschitz type condition. In this direction, let $C_A(u,v) = \exp\left[\ln(uv)\cdot A\left\{\frac{\ln v}{\ln(uv)}\right\}\right]$, $0 < u, v \leq 1$, $(u,v) \neq (1,1)$, be an extreme value copula. Then, for $\lambda > 0$,

$$\begin{aligned} C_A^\lambda(u,v) &= \exp\left[\lambda\ln(uv)\cdot A\left\{\frac{\lambda\ln v}{\lambda\ln(uv)}\right\}\right] = \exp\left[\ln(u^\lambda v^\lambda)\cdot A\left\{\frac{\ln v^\lambda}{\ln(u^\lambda v^\lambda)}\right\}\right] \\ &= C_A(u^\lambda, v^\lambda). \end{aligned}$$

Theorem 2.2.4 of Nelsen (2006), p. 11, states that a subcopula, and hence a copula C, satisfies a Lipschitz condition, that is,

$$|C(u_2,v_2) - C(u_1,v_1)| \leq |u_2 - u_1| + |v_2 - v_1|.$$

Based on them,

$$\begin{aligned} \left|C_A^\lambda(u_2,v_2) - C_A^\lambda(u_1,v_1)\right| &= \left|C_A(u_2^\lambda,v_2^\lambda) - C_A(u_1^\lambda,v_1^\lambda)\right| \\ &\leq \left|u_2^\lambda - u_1^\lambda\right| + \left|v_2^\lambda - v_1^\lambda\right|, \end{aligned}$$

for $\lambda > 0$. Moreover, based on Lagrange mean value theorem, for each $x, y \in (0,1)$ with $x < y$, without loss of generality, $(x^\lambda - y^\lambda)/(x-y) = \lambda\xi^{\lambda-1}$, with $x < \xi < y$. However, for $\lambda \geq 1$ and $0 < x < \xi < y < 1$, $\xi^{\lambda-1} \leq 1$ and then $|x^\lambda - y^\lambda| \leq \lambda|x-y|$, $0 < x < y < 1$, $\lambda \geq 1$. Hence, the function $h(u) = u^\lambda$, $0 < u < 1$, satisfies a type of Lipschitz condition for $\lambda \geq 1$ of the form

$$|h(u_2) - h(u_1)| = \left|u_2^\lambda - u_1^\lambda\right| \leq \lambda\,|u_2 - u_1|,$$

and the same is valid for $\left|v_2^{\lambda}-v_1^{\lambda}\right| \leq \lambda\left|v_2-v_1\right|$. Both of them entail that

$$\left|C_A^{\lambda}(u_2,v_2)-C_A^{\lambda}(u_1,v_1)\right| \leq \lambda\left|u_2-u_1\right|+\lambda\left|v_2-v_1\right|,\ \lambda \geq 1.$$

Therefore, a positive power of an extreme value copula, C_A^{λ}, $\lambda \geq 1$, satisfies a Lipschitz type condition. In a similar manner to that of the proof of Theorem 5.1.9 of Nelsen (2006), p. 169, the above inequality means that the family C_A^{λ}, $\lambda \geq 1$, is equicontinuous, thus the convergence of $\left\{C_{A_n}^{\lambda}\right\}$ to C_A^{λ} is uniform, which entails that

$$\begin{aligned}\lim_{n\to\infty}\mathcal{CI}_{\lambda}(C_{A_n}) &= \lim_{n\to\infty}\int_{[0,1]^2} C_{A_n}^{\lambda}(u,v)dudv = \int_{[0,1]^2}\lim_{n\to\infty}C_{A_n}^{\lambda}(u,v)dudv \\ &= \int_{[0,1]^2} C_A^{\lambda}(u,v)dudv = \mathcal{CI}_{\lambda}(C_A),\end{aligned}$$

for $\lambda \geq 1$. The proof for Tsallis' type entropy is straightforward while the proof for Shannon's type entropy is immediate as, in view of (3.5),

$$\begin{aligned}\lim_{n\to\infty}\mathcal{CE}_{Sh}(C_{A_n}) &= \lim_{n\to\infty}\lim_{\lambda\to1}\mathcal{CE}_{Ts,\lambda}(C_{A_n}) = \lim_{\lambda\to1}\lim_{n\to\infty}\mathcal{CE}_{Ts,\lambda}(C_{A_n}) = \lim_{\lambda\to1}\mathcal{CE}_{Ts,\lambda}(C_A) \\ &= \mathcal{CE}_{Sh}(C_A),\end{aligned}$$

on the condition that the limits exist. ▲

A.5. Proof of Proposition 4.1(c) and 4.1(d)

(c) Based on (4.4), lets try to obtain the integral

$$\int_{[0,1]^2} C_{A_2}(u,v)\left(\frac{C_{A_1}(u,v)}{C_{A_2}(u,v)}\right)^{\lambda+1}dudv,$$

for

$$C_{A_i}(u,v)=\exp\left[\ln(uv)\cdot A_i\left\{\frac{\ln v}{\ln(uv)}\right\}\right],\quad 0<u,v\leq 1,\ (u,v)\neq(1,1),\ i=1,2.$$

Consider the transformation $t=\frac{\ln v}{\ln(uv)}$, $(u,v)\in(0,1]^2,(u,v)\neq(1,1)$. Then, $t\in[0,1)$ and taking into account (1.2) to (1.4),

$$\begin{aligned}\int_{[0,1]^2} C_{A_2}(u,v)\left(\frac{C_{A_1}(u,v)}{C_{A_2}(u,v)}\right)^{\lambda+1}dudv &= -\int_{[0,1]^2} v^{\frac{1-t}{t}}\frac{\ln v}{t^2}\exp\left(\ln v^{\frac{A_2(t)}{t}}\right) \\ &\quad\times\exp\left(\ln v^{\frac{(\lambda+1)[A_1(t)-A_2(t)]}{t}}\right)dvdt,\end{aligned}$$

or

$$\int_{[0,1]^2} C_{A_2}(u,v)\left(\frac{C_{A_1}(u,v)}{C_{A_2}(u,v)}\right)^{\lambda+1}dudv = \int_{[0,1]^2}\left(\ln\frac{1}{v}\right)\frac{1}{t^2}v^{\frac{1+A_1(t)+\lambda[A_1(t)-A_2(t)]}{t}-1}dvdt.$$

Using again (1.6), for $\mu = 2$ and $\nu = \frac{1+A_1(t)+\lambda[A_1(t)-A_2(t)]}{t}$, we obtain

$$\int\limits_{[0,1]^2} C_{A_2}(u,v)\left(\frac{C_{A_1}(u,v)}{C_{A_2}(u,v)}\right)^{\lambda+1} dudv = \int\limits_0^1 \frac{1}{t^2}\frac{t^2}{\left[1+A_1(t)+\lambda[A_1(t)-A_2(t)]\right]^2}dt,$$

which completes the proof in view of (4.4). ▲

(d) For

$$C_{A_i}(u,v) = \exp\left[\ln(uv)\cdot A_i\left\{\frac{\ln v}{\ln(uv)}\right\}\right], \quad 0<u,v\le 1,\ (u,v)\neq(1,1),\ i=1,2,$$

taking into account (4.5),

(1.10)
$$\mathcal{C}d_a(C_{A_1},C_{A_2}) = \int\limits_{[0,1]^2}\left\{C_{A_2}(u,v)^{1+a} - \left(1+\frac{1}{a}\right)C_{A_2}(u,v)^a C_{A_1}(u,v) + \frac{1}{a}C_{A_1}(u,v)^{1+a}\right\}dudv,$$

$a>0$. Based on (3.4),

$$\int\limits_{[0,1]^2} C_{A_i}(u,v)^{1+a}dudv = \int\limits_0^1 \frac{1}{[1+(1+a)A_i(t)]^2}dt,\ i=1,2. \tag{1.11}$$

Hence, it is enough to obtain, $\int\limits_{[0,1]^2} C_{A_2}(u,v)^a C_{A_1}(u,v)dudv$. To proceed with the evaluation of this integral, let's make the change of variables, $t = \frac{\ln v}{\ln(uv)}$, $(u,v)\in(0,1]^2$,$(u,v)\neq(1,1)$. Then, $t\in[0,1)$ and $\ln(uv) = \frac{\ln v}{t}$, $du = -v^{(1-t)/t}\frac{\ln v}{t^2}dt$, in view of (1.2) and (1.4). Based on them,

$$\begin{aligned}
\int\limits_{[0,1]^2} C_{A_2}(u,v)^a C_{A_1}(u,v)dudv &= \int\limits_{[0,1]^2} \exp\left\{\ln(uv)\cdot\left[A_1\left(\frac{\ln v}{\ln(uv)}\right) + aA_2\left(\frac{\ln v}{\ln(uv)}\right)\right]\right\}dudv \\
&= \int\limits_{[0,1]^2} \exp\left\{\frac{\ln v}{t}\cdot[A_1(t)+aA_2(t)]\right\}\left(-v^{(1-t)/t}\frac{\ln v}{t^2}\right)dvdt \\
&= -\int\limits_{[0,1]^2} \frac{1}{t^2}(\ln v)v^{(1-t)/t}\exp\left\{\ln v^{[A_1(t)+aA_2(t)]/t}\right\}dvdt \\
&= \int\limits_{[0,1]^2} \frac{1}{t^2}\left(\ln\frac{1}{v}\right)v^{\frac{1+A_1(t)+aA_2(t)}{t}-1}dvdt \\
&= \int\limits_0^1 \frac{1}{t^2}\left\{\int\limits_0^1\left(\ln\frac{1}{v}\right)v^{\frac{1+A_1(t)+aA_2(t)}{t}-1}dv\right\}dt. \qquad (1.12)
\end{aligned}$$

Applying again formula 4.272-6, on p. 550 in Gradshteyn and Ryzhik (2007),

$$\int\limits_0^1\left(\log\frac{1}{x}\right)^{\mu-1}x^{\nu-1}dx = \frac{1}{\nu^\mu}\Gamma(\mu),\ Re\,\mu>0,\ Re\,\nu>0,$$

for $\mu = 2$ and $\nu = \frac{1+A_1(t)+aA_2(t)}{t}$, we obtain

$$\int_0^1 \left(\ln \frac{1}{v}\right) v^{\frac{1+A_1(t)+aA_2(t)}{t}-1} dv = \frac{t^2}{\left[1+A_1(t)+aA_2(t)\right]^2}. \tag{1.13}$$

Then, (1.12) and (1.13) entail that

$$\int_{[0,1]^2} C_{A_2}(u,v)^a C_{A_1}(u,v) du dv = \int_0^1 \frac{1}{\left[1+A_1(t)+aA_2(t)\right]^2} dt, \tag{1.14}$$

and the proof of case (d) of Proposition 3.2 is integrated in view of (1.10), (1.11) and (1.14). ▲

ACKNOWLEDGMENTS

The author wishes to thank Prof. Christian Genest from McGill University, for fruitful discussions on copulas during his visit in the University of Ioannina, years ago, on May 2017. I also sincerely appreciate the Editor and two reviewers positive and constructive comments. The comments and suggestions of both reviewers have helped a lot to improve the presentation of the paper.

REFERENCES

Ansari, J. and Rockel, M. (2023). Dependence properties of bivariate copula families. *arXiv preprint arXiv:2310.17307*.

Asadi, M., Ebrahimi, N., and Soofi, E. S. (2017). Connections of gini, fisher, and shannon by bayes risk under proportional hazards. *Journal of Applied Probability*, 54(4):1027–1050.

Baratpour, S. and Rad, A. H. (2012). Testing goodness-of-fit for exponential distribution based on cumulative residual entropy. *Communications in Statistics-Theory and Methods*, 41(8):1387–1396.

Basu, A., Harris, I. R., Hjort, N. L., and Jones, M. (1998). Robust and efficient estimation by minimising a density power divergence. *Biometrika*, 85(3):549–559.

Basu, A., Shioya, H., and Park, C. (2011). *Statistical inference: the minimum distance approach*. CRC press, Boca Raton, FL.

Beirlant, J., Goegebeur, Y., Segers, J., and Teugels, J. L. (2004). *Statistics of extremes: theory and applications*. John Wiley & Sons, Chichester, UK.

Billingsley, P. (1986). *Probability and measure*. John Wiley & Sons, New York.

Burbea, J. and Rao, C. R. (1982). Entropy differential metric, distance and divergence measures in probability spaces: A unified approach. *Journal of Multivariate Analysis*, 12(4):575–596.

Chen, X., Kar, S., and Ralescu, D. A. (2012). Cross-entropy measure of uncertain variables. *Information Sciences*, 201:53–60.

Cressie, N. and Read, T. R. (1984). Multinomial goodness-of-fit tests. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 46(3):440–464.

De Keyser, S. and Gijbels, I. (2024). Hierarchical variable clustering via copula-based divergence measures between random vectors. *International Journal of Approximate Reasoning*, 165:109090.

Drouet Mari, D. and Kotz, S. (2001). *Correlation and Dependence*. Imperial College Press, London.

Durante, F. and Mesiar, R. (2010). L^∞-measure of non-exchangeability for bivariate extreme value and archimax copulas. *Journal of Mathematical Analysis and Applications*, 369(2):610–615.

Edwards, H., Mikusiński, P., and Taylor, M. (2005). Measures of concordance determined by d4–invariant measures on $(0,1)^2$. *Proceedings of the American Mathematical Society*, 133(5):1505–1513.

Eschenburg, P. (2013). Properties of extreme-value copulas. Master's thesis, Technische Universität München.

Ferentinos, K. and Papaioannou, T. (1981). New parametric measures of information. *Information and Control*, 51(3):193–208.

Fuchs, S. (2014). Multivariate copulas: Transformations, symmetry, order and measures of concordance. *Kybernetika*, 50(5):725–743.

Fuchs, S. (2016). Copula–induced measures of concordance. *Dependence Modeling*, 4:205–214.

Fuchs, S. and Schmidt, K. D. (2014). Bivariate copulas: transformations, asymmetry and measures of concordance. *Kybernetika*, 50(1):109–125.

Ghoudi, K., Khoudraji, A., and Rivest, E. L.-P. (1998). Propriétés statistiques des copules de valeurs extrêmes bidimensionnelles. *Canadian Journal of Statistics*, 26(1):187–197.

Golomb, S. (1966). The information generating function of a probability distribution (corresp.). *IEEE Transactions on Information Theory*, 12(1):75–77.

Gradshteyn, I. S. and Ryzhik, I. M. (2007). *Table of integrals, series, and products*. Elsevier/Academic Press, Amsterdam.

Gudendorf, G. and Segers, J. (2010). Extreme-value copulas. copula theory and its applications, 127–145. *Lect. Notes Stat. Proc*, 198.

Guiasu, S. and Reischer, C. (1985). The relative information generating function. *Information Sciences*, 35(3):235–241.

Joe, H. (2015). *Dependence modeling with copulas*. Chapman & Hall/CRC Press, Boca Raton, FL.

Kamnitui, N., Genest, C., Jaworski, P., and Trutschnig, W. (2019). On the size of the class of bivariate extreme-value copulas with a fixed value of Spearman's rho or Kendall's tau. *Journal of Mathematical Analysis and Applications*, 472(1):920–936.

Klein, I. and Doll, M. (2020). (Generalized) maximum cumulative direct, residual, and paired Φ entropy approach. *Entropy*, 22(1):91.

Klein, I., Mangold, B., and Doll, M. (2016). Cumulative paired φ-entropy. *Entropy*, 18(7):248.

Liebscher, E. (2014). Copula-based dependence measures. *Dependence Modeling*, 2(1):49–64.

Ma, J. and Sun, Z. (2008). Mutual information is copula entropy. *arXiv preprint arXiv:0808.0845, [cs.IT]*.

Nair, N. U. and Sunoj, S. (2023). Survival copula entropy and dependence in bivariate distributions. *REVSTAT-Statistical Journal (Accepted)*.

Nelsen, R. (2002). Concordance and copulas: a survey. In Cuadras, C. M., Fortiana, J., and Rodriguez-Lallena, J. A., editors, *Distributions with Given Marginals and Statistical Modelling*, pages 169–177. Kluwer Academic Publishers, Dordrecht.

Nelsen, R. B. (2006). *An introduction to copulas*. Springer, Berlin.

Papaioannou, T. (1985). Measures of information. In Kotz, S. and Johnson, N. L., editors, *Encyclopedia of Statistical Sciences*, pages 391–397. John Wiley & Sons, New York.

Papaioannou, T. (2001). On distances and measures of information: A case of diversity. In Charalambides, C., Koutras, M., and Balakrishnan, N., editors, *Probability and Statistical Models with Applications*, pages 503–515. Chapman & Hall, London.

Pardo, L. (2006). *Statistical Inference Based on Divergence Measures*. Chapman & Hall/CRC Press, Boca Raton, FL.

Park, S., Rao, M., and Shin, D. W. (2012). On cumulative residual kullback–leibler information. *Statistics & Probability Letters*, 82(11):2025–2032.

Rajesh, G. and Sunoj, S. (2019). Some properties of cumulative tsallis entropy of order α. *Statistical Papers*, 60(3):933–943.

Rao, M., Chen, Y., Vemuri, B. C., and Wang, F. (2004). Cumulative residual entropy: a new measure of information. *IEEE transactions on Information Theory*, 50(6):1220–1228.

Read, T. R. and Cressie, N. A. (1988). *Goodness-of-fit statistics for discrete multivariate data*. Springer Science & Business Media, New York.

Scarsini, M. (1984). On measures of concordance. *Stochastica*, 8(3):201–218.

Tawn, J. A. (1990). Modelling multivariate extreme value distributions. *Biometrika*, 77(2):245–253.

Weller-Fahy, D. J., Borghetti, B. J., and Sodemann, A. A. (2015). A survey of distance and similarity measures used within network intrusion anomaly detection. *IEEE Communications Surveys & Tutorials*, 17(1):70–91.

Zografos, K. (2023). On reconsidering entropies and divergences and their cumulative counterparts: Csiszár's, DPD's and Fisher's type cumulative and survival measures. *Probability in the Engineering and Informational Sciences*, 37(1):294–321.

Zografos, K. and Nadarajah, S. (2005). Survival exponential entropies. *IEEE Transactions on Information Theory*, 51(3):1239–1246.