Research Article **Open Access**

Carole Bernard and Alfred Müller

Dependence uncertainty bounds for the energy score and the multivariate Gini mean difference

https://doi.org/10.1515/demo-2020-0014
Received April 24, 2020; accepted September 5, 2020

Abstract: The energy distance and energy scores became important tools in multivariate statistics and multivariate probabilistic forecasting in recent years. They are both based on the expected distance of two independent samples. In this paper we study dependence uncertainty bounds for these quantities under the assumption that we know the marginals but do not know the dependence structure. We find some interesting sharp analytic bounds, where one of them is obtained for an unusual spherically symmetric copula. These results should help to better understand the sensitivity of these measures to misspecifications in the copula.

Keywords: dependence uncertainty bounds, energy score, Gini mean difference, spherically symmetric copula

MSC: 60E05, 60E15, 62C05, 62E10, 62E15

1 Introduction

In recent years the so-called *energy distance* became a famous tool in multivariate statistics used e.g., for goodness-of-fit tests and many other things. For a good overview over this topic we refer to [22]. Similar concepts have been suggested in the theory of multivariate probabilistic forecasting, where the so-called *energy score* has been suggested as a strictly proper scoring rule for multivariate distributions in the fundamental paper of [6]. Both concepts rely on functionals that are based on expected distances of independent copies of random vectors. This is related to the multivariate Gini mean difference, which has been studied in detail in [9]. In the univariate case the Gini mean difference is a well-known measure of spread of distributions or inequality in case of income distributions, see e.g., [24] for an overview.

In goodness-of-fit testing as well as in probabilistic forecasting one is interested in detecting misspecifications of stochastic models. Therefore it is an important question how sensitive the used functionals react to which kind of misspecification. [15] studied the discrimination ability of the energy score for the case of multivariate normal distributions. Based on simulation studies they conclude that the discrimination ability of the energy score may be limited when focusing on the dependence structure of multivariate probabilistic forecasts, but to the best of our knowledge there has been no general study of this problem so far for general distributions. Motivated from this study, [26] recently introduced the concept of a *copula energy score*. For their study it is important to find good lower bounds for the energy score of a copula.

In this paper we want to study this problem of so-called *dependence uncertainty bounds* for such quantities like the energy score and the Gini mean difference. By dependence uncertainty bounds we mean here

Carole Bernard: Department of Accounting, Law and Finance, Grenoble Ecole de Management, Grenoble, France, E-mail: carole.bernard@grenoble-em.com
Faculty of Economics, Vrije Universiteit Brussel, Brussels, Belgium
Alfred Müller: Department of Mathematics, University of Siegen, Siegen, Germany, E-mail: mueller@mathematik.uni-siegen.de

bounds for a functional of a multivariate distribution under the assumption that we only know the marginal distributions but do not know the dependence structure, i.e., we do not know the copula. The study of such uncertainty bounds has a long history going back to [7] and [4]. They considered this problem for correlation coefficients and for the value of cumulative distribution functions. In the meanwhile there is a vast literature on this topic for many kinds of functionals. For an overview see [18]. Very often the extremal positive dependence is given by the comonotone copula, in particular if the functional is an expectation of a supermodular function, as has been shown in [23] and [19]. It is typically more complicated to find the extremal negative dependence, even in the case of expectations of functions and thus linear functionals of the distributions, which is the case for most problems considered in the literature. An example of a non-linear problem is the case of finding the solution of an optimal stopping problem that was considered in [11]. In such a case of a non-linear problem the characterization of the optimizer can be very different from the case of a linear problem. Other non-linear problems include finding bounds on some risk measures.

The problem considered here also resembles the problems of optimal transport or optimal coupling. Our problem can be written in the form

$$\sup\left\{\int c(x_1,\dots,x_{2d})d\mu\right\}$$

where the supremum is taken over all $2d$-dimensional distributions μ with given marginals and the additional constraint that the first d components are independent of the other d. Without these additional constraints such problems have been considered e.g., in [20]. In recent years there is a lot of interest in such problems under an additional martingale constraint, see e.g., [5].

In this paper we also deal with a non-linear problem, but it will turn out that still the comonotone copula will typically lead to the extremal positive dependence. But for the extremal negative dependence we find in some cases a very interesting solution based on a spherical symmetric copula. This is an interesting copula, which does not seem to be well-known in the dependence modelling community.

The paper is organized as follows. In Section 2, we recall the definitions of the various concepts. We also introduce some important notation that will be used throughout the manuscript and present the problem that is considered in this paper. In Section 3, we focus on the expected distance between two multivariate distributions and its sensitivity to dependence uncertainty. Finally in Section 4, we provide a number of results on the dependence uncertainty bounds on the energy score. Section 5 concludes with a number of open questions that are left for future research.

2 Energy score and Gini mean difference

Throughout the manuscript $\boldsymbol{X}, \tilde{\boldsymbol{X}}$ shall be independent copies of a d-dimensional random vector with cumulative distribution function (cdf)

$$F(\boldsymbol{x}) = P(X_1 \le x_1, \dots, X_d \le x_d), \quad \boldsymbol{x} = (x_1, \dots, x_d) \in \mathbb{R}^d,$$

and let $\boldsymbol{Y}$ be a random vector with cdf G. We define the expected distance between two independent d-dimensional samples of F and G as

$$S(F, G) = \mathbb{E}\left(\|\boldsymbol{X} - \boldsymbol{Y}\|_2\right) = \int \|\boldsymbol{x} - \boldsymbol{y}\|_2 F(d\boldsymbol{x})G(d\boldsymbol{y}),$$

where we identify the cdfs F and G with the corresponding probability measures and denote as usual by

$$\|x\|_2 = \sqrt{\sum_{i=1}^{d} x_i^2}$$

the Euclidian distance. For an observation $\boldsymbol{y}$, we similarly define by identifying $\boldsymbol{y}$ with the one-point measure in $\boldsymbol{y}$

$$S(F, \boldsymbol{y}) = \mathbb{E}\left(\|\boldsymbol{X} - \boldsymbol{y}\|_2\right) = \int \|\boldsymbol{x} - \boldsymbol{y}\|_2 F(d\boldsymbol{x}).$$

The *energy distance* between two distributions F and G is defined as

$$\mathcal{E}(F, G) = 2S(F, G) - S(F, F) - S(G, G).$$

This is a distance between probability distributions, as it can be shown that $\mathcal{E}(F, G) \geq 0$ for all F, G and that $\mathcal{E}(F, G) = 0$ if and only if $F = G$. For details of this concept and applications we refer to the overview article of [22]. A strongly related concept is the so-called *energy score* for a distributional forecast F and an observation $\boldsymbol{y}$, which is given by

$$ES(F, \boldsymbol{y}) = S(F, \boldsymbol{y}) - \frac{1}{2}S(F, F).$$

This can be generalized by introducing a parameter $\beta \in (0, 2)$ as already considered in the fundamental paper of [6].

Definition 1. For $\beta \in (0, 2)$, the generalized expected distance between two independent d-dimensional samples of F and G is defined as

$$S_\beta(F, G) = \mathbb{E}\left(\|\boldsymbol{X} - \boldsymbol{Y}\|_2^\beta\right) = \int \|\boldsymbol{x} - \boldsymbol{y}\|_2^\beta F(d\boldsymbol{x})G(d\boldsymbol{y}),$$

and the *generalized energy score* as

$$ES_\beta(F, \boldsymbol{y}) = S_\beta(F, \boldsymbol{y}) - \frac{1}{2}S_\beta(F, F)$$

Similarly, the corresponding *generalized energy distance* is defined as

$$\mathcal{E}_\beta(F, G) = 2S_\beta(F, G) - S_\beta(F, F) - S_\beta(G, G). \tag{1}$$

Note that the limiting case $\beta = 2$ is excluded in the definition, as $\mathcal{E}_2(F, G)$ only depends on the marginal distributions of F and G and thus does not depend at all on the copula and in fact therefore is not a distance and does not lead to a proper scoring rule.

Remark 2. The function

$$S(F, F) = \mathbb{E}\left(\|\boldsymbol{X} - \tilde{\boldsymbol{X}}\|_2\right) = \int \|\boldsymbol{x} - \boldsymbol{y}\|_2 F(d\boldsymbol{x})F(d\boldsymbol{y})$$

is known (sometimes up to a constant $2d$) as *multivariate Gini mean difference* and has been studied in detail in [9]. To distinguish the univariate version of the Gini mean difference from the multivariate one we denote, from now on, the univariate version with a slight abuse of notation as

$$M(F) := M(F, F) = S(F, F) = \int |x - y|F(dx)F(dy) = 2\int (1 - F(x))F(x)dx$$

and its generalization for $\beta \in (0, 2)$ and different F and G similarly as

$$M_\beta(F, G) = \int |x - y|^\beta F(dx)G(dy).$$

Note that the bordering case of $\beta = 2$ yields for $X \sim F$ up to a factor of two the variance

$$M_2(F, F) = \int |x - y|^2 F(dx)F(dy) = 2var(X).$$

We will frequently consider the random variable $Z = |X - Y|$ for independent random variables X and Y. We use the following notation.

Definition 3. For independent X and Y with cdf F and G, respectively, we denote by $F\Diamond G$ the cdf of $Z = |X - Y|$ which is given by

$$F\Diamond G(x) = \int (F(y + x) - F(y - x))\, G(dy), \quad x \geq 0.$$

Notice that the Gini mean difference $M(F) = S(F, F)$ is the mean of $F \diamondsuit F$ and more general $M_\beta(F, G) = \mathbb{E}(Z^\beta)$ is the corresponding moment of order β.

Example 4. In case of standard uniform distributions U, V on $(0, 1)$ for $F = G$ we get for Z the density $f_Z(z) = 2 - 2z$ on $[0, 1]$ and thus

$$M_\beta(F) = \mathbb{E}(Z^\beta) = \int_0^1 z^\beta (2 - 2z) dz = \frac{2}{\beta + 1} - \frac{2}{\beta + 2} \tag{2}$$

with the special cases $S(F, F) = M(F) = \mathbb{E}Z = \frac{1}{3}$ and in the limiting case when $\beta = 2$ we get $\mathbb{E}Z^2 = \frac{1}{6} = 2var(U)$.

2.1 Dependence uncertainty bounds

We want to study how sensitive these quantities are with respect to the dependence information in the joint distribution. Therefore we investigate bounds for such expressions given that we only know the marginals of F and G. As usual we denote the marginals by

$$F_i(x) := P(X_i \le x), \quad x \in \mathbb{R}, \ i = 1, \dots, d.$$

By Sklar's theorem we can write the joint cumulative distribution F of $\boldsymbol{X}$ in the form

$$F(\boldsymbol{x}) = C(F_1(x_1), \dots, F_d(x_d)), \quad \boldsymbol{x} \in \mathbb{R}^d,$$

for some copula C, see e.g., [13]. We denote by $\mathcal{C}$ the set of all possible copulas and by $\mathcal{F} = \mathcal{F}(F_1, \dots, F_d)$ the so-called Fréchet class of all multivariate distributions with given marginals $F_1, \dots, F_d$. The well-known Fréchet bounds are denoted by

$$F^+(\boldsymbol{x}) = \min\{F_1(x_1), \dots, F_d(x_d)\}$$

and

$$F^-(\boldsymbol{x}) = \max\{F_1(x_1) + \dots + F_d(x_d) - d + 1, 0\}$$

and C^+ and C^- will be the corresponding copulas, typically called the comonotonic and countermonotonic copula, where one has to take into account that C^- is only a copula for $d = 2$.

Similarly, for a joint cumulative distribution function G of $\mathbf{X}$,

$$G(\boldsymbol{x}) = C(G_1(x_1), \dots, G_d(x_d)), \quad \boldsymbol{x} \in \mathbb{R}^d,$$

we denote by $\mathcal{G}$, the Fréchet class of all multivariate distributions with given marginals $G_1, \dots, G_d$.

We first study the corresponding dependence uncertainty bounds on the generalized expected distance between two independent d-dimensional samples from F and G. These bounds can then be written as

$$\inf_{F \in \mathcal{F}} S_\beta(F, F) \text{ and } \sup_{F \in \mathcal{F}} S_\beta(F, F), \tag{3}$$

when both samples come from the same distribution F, and by

$$\inf_{F \in \mathcal{F}, G \in \mathcal{G}} S_\beta(F, G) \text{ and } \sup_{F \in \mathcal{F}, G \in \mathcal{G}} S_\beta(F, G), \tag{4}$$

when the multivariate distributions F and G are not identical.

For a given observation $\boldsymbol{y}$, we then study dependence uncertainty bounds for its generalized energy score by considering

$$\inf_{F \in \mathcal{F}} ES_\beta(F, \boldsymbol{y}) \text{ and } \sup_{F \in \mathcal{F}} ES_\beta(F, \boldsymbol{y}) \tag{5}$$

The optimizations in (3), (4) and (5) are over the Fréchet classes $\mathcal{F}$ and $\mathcal{G}$. In fact, given that the marginal distributions are given, the uncertainty bounds can also be considered as solutions of optimization problems over the class $\mathcal{C}$ of all copulas.

3 Dependence uncertainty bounds for S_β

In this section, we provide analytic bounds for the expressions (3) and (4). To do so, we first look at some fundamental properties of S_β and ES_β.

It is well-known (see for instance [6]) that the generalized energy distance $\mathcal{E}_\beta$ is a distance for $\beta \in (0, 2)$, i.e., $\mathcal{E}_\beta(F, G) \geq 0$ for all F, G and therefore also $ES_\beta(F, \boldsymbol{y}) \geq 0$ for all F and $\boldsymbol{y}$. Moreover, this implies by definition (1) also that

$$2S_\beta(F, G) = S_\beta(F, G) + S_\beta(G, F) \geq S_\beta(F, F) + S_\beta(G, G). \tag{6}$$

It is also well-known that ES_β is a proper scoring rule, meaning that

$$ES_\beta(F, F) \leq ES_\beta(F, G) \quad \text{for all } F, G. \tag{7}$$

Note also that $ES_\beta(F, F) = \frac{1}{2}S_\beta(F, F)$. We can derive the following lemma on the concavity of $S_\beta(F, F)$. As already observed in section 2.1 of [6], this concavity holds for any proper scoring rule, but for completeness we add an elementary proof.

Lemma 5. *$F \to S_\beta(F, F)$ is concave.*

Proof. Indeed, S_β is linear in F and G and therefore we get for $\alpha \in (0, 1)$ from (6) that

$$\begin{aligned}
&S_\beta(\alpha F + (1-\alpha)G, \alpha F + (1-\alpha)G) \\
&\quad = \alpha^2 S_\beta(F, F) + \alpha(1-\alpha)(S_\beta(F, G) + S_\beta(G, F)) + (1-\alpha)^2 S_\beta(G, G) \\
&\quad \geq \alpha^2 S_\beta(F, F) + \alpha(1-\alpha)(S_\beta(F, F) + S_\beta(G, G)) + (1-\alpha)^2 S_\beta(G, G) \\
&\quad = \alpha S_\beta(F, F) + (1-\alpha) S_\beta(G, G).
\end{aligned}$$

□

3.1 Lower bound on $S_\beta(F, G)$

For finding a minimum value of $S_\beta(F, G)$ the following representation is going to be helpful.

$$S_\beta(F, G) = \mathbb{E}\left(\|\boldsymbol{X} - \boldsymbol{Y}\|_2^\beta\right) = \mathbb{E}\left(\left(\sum_{i=1}^d Z_i^2\right)^{\frac{\beta}{2}}\right), \tag{8}$$

where $Z_i = |X_i - Y_i| \sim F_i \diamondsuit G_i$ (see Definition 3). Thus we have a representation $S_\beta(F, G) = \mathbb{E}f(Z_1^2, \dots, Z_d^2)$ where we know the marginals of $(Z_1^2, \dots, Z_d^2)$ and the function f has the following properties: as a concave function of the sum it is submodular, i.e., $-f$ is supermodular. For the definition and properties of supermodular functions and their relevance for inequalities of expectations in case of distributions with given marginals we refer to Chapter 3 in [12].

From this we can derive the following lower bound.

Theorem 6. *For any random vector $\boldsymbol{X}$ and $\boldsymbol{Y}$ with cumulative cdfs F and G we get the following lower bound:*

$$S_\beta(F, G) \geq \mathbb{E}\left(Z^{\frac{\beta}{2}}\right)$$

for a random variable Z that is defined as

$$Z = \sum_{i=1}^d \left((F_i \diamondsuit G_i)^{-1}(U)\right)^2$$

for some standard uniform random variable U.

In case of identical marginals $F_1 = \ldots = F_d$ and $G_1 = \ldots = G_d$ this bound is sharp and is obtained for the upper Fréchet bounds $F = F^+$ and $G = G^+$. Furthermore, in this case it reduces to

$$S_\beta(F, G) \geq S_\beta\left(F^+, G^+\right) = d^{\frac{\beta}{2}} M_\beta(F_1, G_1).$$

Proof. According to (8) we can write $S(F, G) = \mathbb{E}f(Z_1^2, \ldots, Z_d^2)$ for a submodular function f. It follows from [23] that a lower bound for $\mathbb{E}f(Z_1^2, \ldots, Z_d^2)$ is obtained by assuming that the copula of $(Z_1^2, \ldots, Z_d^2)$ is given by the upper Fréchet bound C^+, or equivalently the copula of $(Z_1, \ldots, Z_d)$, as the copula is invariant under a strictly increasing transform of the components. This means that we can assume that $Z_i^2 = ((F_i \diamondsuit G_i)^{-1}(U))^2$ for some fixed uniform U and from this the first assertion immediately follows.

If each of F and G has identical marginals, then we have for F^+ and G^+ that $\boldsymbol{X} = (X_1, \ldots, X_1)$ and $\boldsymbol{Y} = (Y_1, \ldots, Y_1)$ and thus we get in (8) also the equality $Z_1 = \ldots = Z_d$ and therefore this lower bound is attained and reduces to

$$S_\beta\left(F^+, G^+\right) = \mathbb{E}\left((dZ_1^2)^{\frac{\beta}{2}}\right) = d^{\frac{\beta}{2}} M_\beta(F_1, G_1)$$

□

Note that $S_\beta(F^+, G^+)$ may not be a lower bound when the marginals of F and G are not identical. Consider $X = (4U_1, U_1)$ and $Y = (U_2, 4U_2)$ for two independent uniform U_1, U_2. Then Z_1 is large if U_1 is large and Z_2 is large if U_2 is large and they are far away from being comonotone. The support of (Z_1, Z_2) is given in the left panel of Figure 1 and does not correspond of the support of the upper Fréchet bound. The lower bound is obtained for something different from F^+ and G^+. In fact, if one takes for F the lower Fréchet bound $F = F^-$ and $G = G^+$ instead, i.e., $X = (4U_1, 1 - U_1)$ and $Y = (U_2, 4U_2)$, then we get more positively correlated Z_1 and Z_2 as depicted in the right panel of Figure 1. Indeed we get $S(F^-, G^+) \approx 2.48 < S(F^+, G^+) \approx 2.55$.

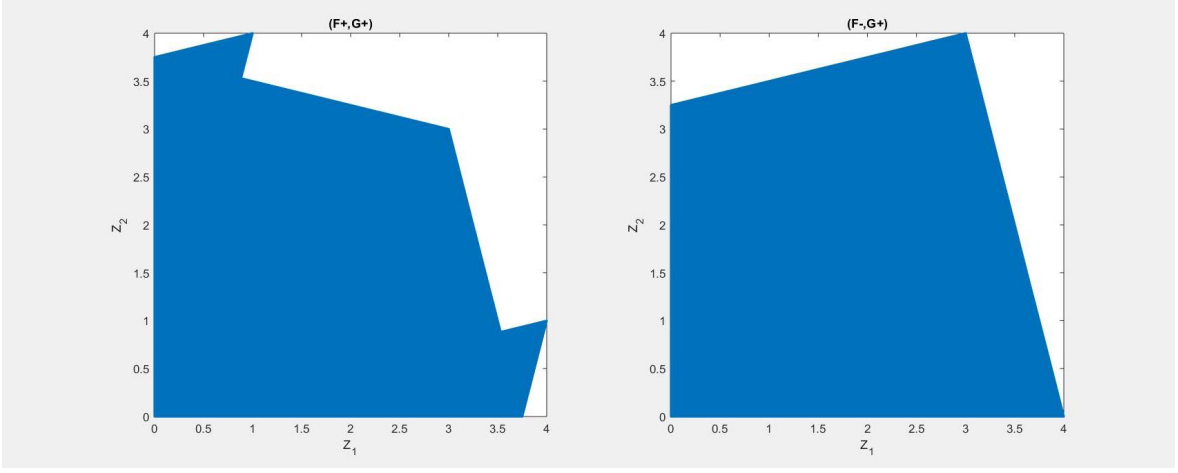

Figure 1: Support of (Z_1, Z_2). Left panel: $F = F^+$, $G = G^+$. Right panel: $F = F^-$, $G = G^+$.

3.2 Upper bound on $S_\beta(F, G)$

It seems to be much more difficult to find a sharp upper bound for $S_\beta(F, G)$, as it is a notoriously difficult problem to find a strongest possible negative dependence in the sense of maximizing the expression in (8). But we can easily derive an upper bound via Jensen's inequality.

Theorem 7. *For any random vector $\boldsymbol{X}$ and $\boldsymbol{Y}$ with cumulative cdfs F and G we get the following upper bound:*

$$S_\beta(F, G) \le \left(\sum_{i=1}^{d} M_2(F_i, G_i) \right)^{\frac{\beta}{2}}.$$

Proof. Applying Jensen's inequality to (8) we get that for any multivariate distributions F and G,

$$S_\beta(F, G) = \mathbb{E} \left(\sum_{i=1}^{d} Z_i^2 \right)^{\frac{\beta}{2}} \le \left(\sum_{i=1}^{d} \mathbb{E} Z_i^2 \right)^{\frac{\beta}{2}} = \left(\sum_{i=1}^{d} M_2(F_i, G_i) \right)^{\frac{\beta}{2}}.$$

□

The bound in Theorem 7 is obtained from Jensen's inequality. Therefore it is only tight if the expression $\|\boldsymbol{X} - \boldsymbol{Y}\|_2$ is constant. This will typically not be the case unless at least one of the distributions of $\boldsymbol{X}$ and $\boldsymbol{Y}$ is constant. However, this is not a characterization: a simple exception is given in the following example.

Example 8. Assume that $d = 2$ and that $F_1 = F_2 = G_1 = G_2$ are all Bernoulli with probability $1/2$. If F has the copula C^+ and G has the copula C^- then $\boldsymbol{X}$ has probability $1/2$ for the values $(0, 0)$ and $(1, 1)$, whereas $\boldsymbol{Y}$ has probability $1/2$ for the values $(0, 1)$ and $(1, 0)$, so that $\|\boldsymbol{X} - \boldsymbol{Y}\|_2 = 1$ almost sure, and hence, we get the sharp bound

$$S_\beta(F, G) \le S_\beta(C^+(F_1, F_2), C^-(G_1, G_2)) = 1.$$

3.3 Upper and lower bounds on $S_\beta(F, G)$ for copulas

In the special case of uniform marginal distributions, the class $\mathcal{F}$ is simply the class of all copulas and we are able to derive explicit expressions of the lower and upper bounds obtained in Theorems 6 and 7. Specifically, from Example 4, we have expression (2) and we get that $\mathbb{E} Z_i^2 = 1/6$. We can then immediately derive the following consequence.

Corollary 9. *For copulas C_1, C_2 we get the following bounds:*

$$S_\beta(C^+, C^+) = d^{\frac{\beta}{2}} \cdot \left(\frac{2}{\beta + 1} - \frac{2}{\beta + 2} \right) \le S_\beta(C_1, C_2) \le \left(\frac{d}{6} \right)^{\frac{\beta}{2}}$$

and in particular for $\beta = 1$,

$$S(C^+, C^+) = \frac{1}{3}\sqrt{d} \le S(C_1, C_2) \le \sqrt{\frac{d}{6}}.$$

Example 10. Similarly as in the case of Bernoulli variables considered in Example 8 one could conjecture that a sharp upper bound for copulas is obtained by $S_\beta(C^-, C^+)$. We will now give an explicit counterexample for the important case of $d = 2$ and $\beta = 1$. Using the invariance under rotation we can easily derive $S(C^-, C^+)$ in this case from the expected distance of two points on the axes.

$$S(C^-, C^+) = \frac{1}{\sqrt{2}} \int_0^1 \int_0^1 \sqrt{x^2 + y^2}\, dxdy = \frac{\sqrt{2} + \log(1 + \sqrt{2})}{3\sqrt{2}} \approx 0.541.$$

Now let us consider the copula $C^{\parallel}$, defined as the distribution of the following random vector (U_1, U_2) with $U_1 \sim U(0, 1)$ and

$$U_2 := \begin{cases} U_1 + \frac{1}{2}, & \text{if } U_1 \le \frac{1}{2}, \\ U_1 - \frac{1}{2}, & \text{if } U_1 > \frac{1}{2}. \end{cases} \tag{9}$$

Thus the support of the copula $C^{\parallel}$ consists of two parallel line segments as displayed in Figure 2.

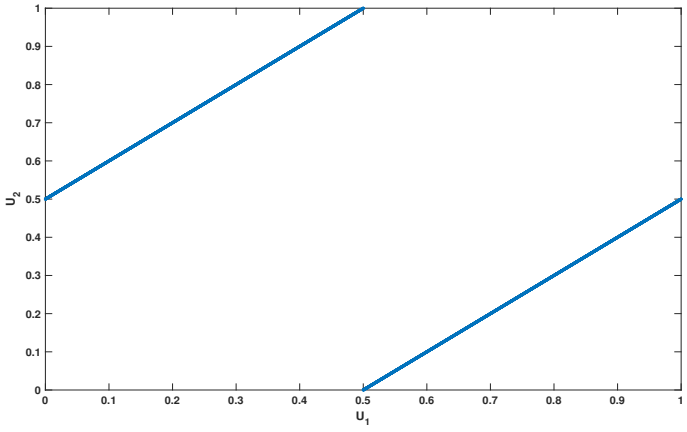

Figure 2: Support of the copula $C^{\parallel}$ defined in (9)

Then we get by a similar computation

$$S(C^+, C^{\parallel}) = \frac{1}{\sqrt{2}} \int_0^1 \int_0^1 \sqrt{\frac{1}{2} + \left(\frac{1}{2} + x - y\right)^2}\, dxdy \approx 0.549.$$

Thus $S(C^+, C^{\parallel}) > S(C^-, C^+)$. We notice, however, that this inequality is reversed for the energy distance. We get $\mathcal{E}(C^-, C^+) \approx 0.069 > 0.064 \approx \mathcal{E}(C^-, C^{\parallel})$ even though $S(C^-, C^+) < S(C^+, C^{\parallel})$, as $S(C^+, C^+) \approx 0.471$ is significantly smaller than $S(C^{\parallel}, C^{\parallel}) \approx 0.4985$. Therefore it is still an open problem whether for the energy distance $\mathcal{E}(C^-, C^+)$ maximizes $\mathcal{E}(C_1, C_2)$ among all copulas.

3.4 Upper bound on $S(F, F)$ for copulas

Under the assumption of equal copulas $C = C_1 = C_2$, we can improve the upper bounds in Corollary 9 and even find sharp bounds for $S(C, C)$ in the case of dimension $d = 2$ and $d = 3$ for some interesting copulas, which can be called spherical symmetric copulas. These do not seem to be very well-known in the community working on dependence modelling and copulas, but they have been considered from time to time in the statistics literature, e.g., in [2], [21] and [14]. Implicitly this result can also be found in section 4.9 of [8] where it is shown that uniform distributions can be marginal distributions of spherical symmetric distributions only in dimension $d = 2$ and $d = 3$. [14] show that in dimensions $d = 2$ and $d = 3$ there are unique spherical symmetric random vectors $\boldsymbol{X}$ whose marginals are uniformly distributed on $[-1, 1]$. In dimension $d = 2$, this distribution has the density

$$f(x, y) = \frac{1}{2\pi\sqrt{1 - x^2 - y^2}} \mathbf{1}_{[x^2+y^2<1]} \tag{10}$$

and in dimension $d = 3$ this is given by the uniform distribution on the sphere of a unit ball. The fact that the uniform distribution on the sphere of a unit ball in dimension $d = 3$ has uniform marginals is in fact a consequence of a famous theorem on spheres and cylinders going back to Archimedes. Notice that the bivariate case can be obtained from this by considering the two-dimensional marginals of the three-dimensional case. Transforming the marginals to uniform distributions on $[0, 1]$ via the transformation $X_i \mapsto (X_i + 1)/2$ we get copulas called *spherical symmetric copulas*, which we will denote by C°. In Figure 3 we show a discrete approximation of the bivariate spherical symmetric copula. Notice that the density is unbounded, going to infinity at the boundary of the support, and therefore in the discrete approximation there are many points there as one expects for a bivariate projection of points uniformly scattered on the sphere of a ball.

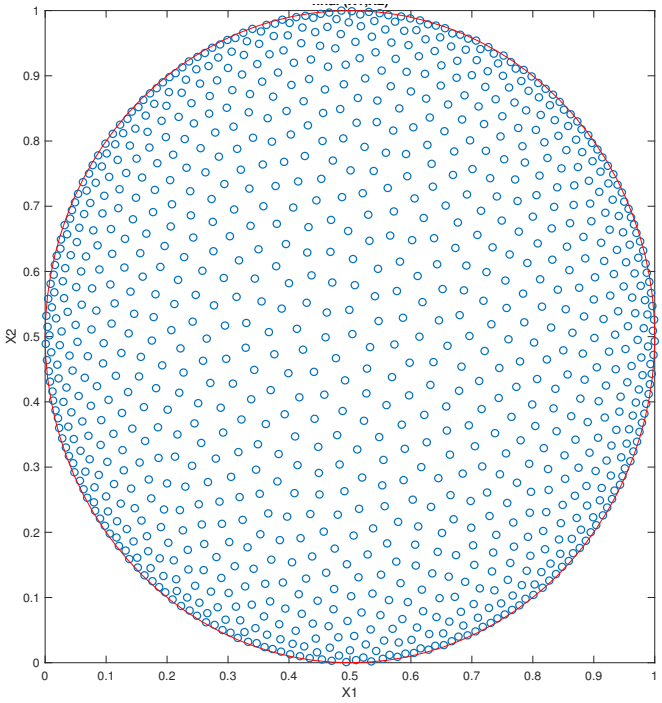

Figure 3: An illustration of the bivariate spherical symmetric copula C°

Moreover, we will use results that were obtained independently in [10], Theorem 2 and as main result in [1]. In our notation their results can be stated as follows.

Theorem 11. *The functional $S(F, F)$ is maximized among all random vectors $\boldsymbol{X} \sim F$ with $\mathbb{E}(\|\boldsymbol{X}\|^2) \le 1$ for $S(F^\star, F^\star)$, where $\boldsymbol{X}^\star \sim F^\star$ is given as follows.*
If $d \ge 3$ then $\boldsymbol{X}^\star$ is uniformly distributed on the sphere of a unit ball.
If $d = 2$ then $\sqrt{2/3} \cdot \boldsymbol{X}^\star$ has the density given in equation (10).

From this result we can derive the following improved bounds for copulas.

Theorem 12. *In dimension $d = 2$ it holds for any copula C that*

$$S(C, C) \le S(C^\circ, C^\circ) = \frac{\pi}{6}.$$

In dimension $d = 3$ it holds for any copula C that

$$S(C, C) \le S(C^\circ, C^\circ) = \frac{2}{3}.$$

Proof. We define the shift $T(X_1, \ldots, X_d) = (X_1 - 1/2, \ldots, X_d - 1/2)$ that transforms marginals from uniform on $(0, 1)$ to uniform on $(-1/2, 1/2)$. As obviously

$$\mathbb{E}\left(\|\boldsymbol{X} - \boldsymbol{Y}\|_2\right) = \mathbb{E}\left(\|T\boldsymbol{X} - T\boldsymbol{Y}\|_2\right)$$

this shift does not affect the functional S, and therefore we can replace the copulas by distributions F with uniform marginals on $(-1/2, 1/2)$. For any such random vector $\boldsymbol{X} \sim F$ we get

$$\mathbb{E}(\|\boldsymbol{X}\|^2) = \mathbb{E}(X_1^2 + \ldots + X_d^2) = \frac{d}{12} \text{ and hence } \mathbb{E}\left(\left\|\sqrt{\frac{12}{d}}\boldsymbol{X}\right\|^2\right) = 1.$$

For any copula C, and $\boldsymbol{X} \sim C$ and $\boldsymbol{Y} \sim C$ independent, we thus obtain that $\mathbb{E}\left(\left\| \sqrt{\frac{12}{d}} T\boldsymbol{X} \right\|^2\right) = 1$ and thus from Theorem 11,

$$\begin{aligned} S(C, C) &= \mathbb{E}\left(\|\boldsymbol{X} - \boldsymbol{Y}\|_2\right) = \mathbb{E}\left(\|T\boldsymbol{X} - T\boldsymbol{Y}\|_2\right) \\ &= \sqrt{\frac{d}{12}} \mathbb{E}\left(\left\| \sqrt{\frac{12}{d}} T\boldsymbol{X} - \sqrt{\frac{12}{d}} T\boldsymbol{Y} \right\|_2\right) \leq \sqrt{\frac{d}{12}} S(F^\star, F^\star), \end{aligned} \tag{11}$$

with $S(F^\star, F^\star)$ as described in Theorem 11. In case $d = 2$ and $d = 3$ we get equality for $C = C^\circ$ and the corresponding values for $S(F^\star, F^\star)$ are computed in [1], there denoted as M_d. They are given by

$$M_2 = \frac{\pi}{\sqrt{6}} \text{ and } M_3 = \frac{4}{3}.$$

Thus we derive for $d = 2$ that

$$S(C, C) \leq S(C^\circ, C^\circ) = \sqrt{\frac{2}{12}} \cdot M_2 = \frac{\pi}{6},$$

and for $d = 3$ that

$$S(C, C) \leq S(C^\circ, C^\circ) = \sqrt{\frac{3}{12}} \cdot M_3 = \frac{2}{3}.$$

□

We also get an improved bound for $d \geq 4$ from Theorem 11 but then it is not sharp, as the uniform distribution on the sphere no longer has uniform marginals. Indeed, [14] show that there cannot exist any spherical symmetric copula in dimension $d \geq 4$ and therefore we do not know, how the copula C that maximizes $C \mapsto S(C, C)$ looks like. We conjecture that it in some sense will be close to spherical symmetry. The improved bound that we can derive from (11) and Theorem 11 in case $d = 4$ is

$$S(C, C) \leq \sqrt{\frac{4}{12}} \cdot M_4 = \sqrt{\frac{1}{3}} \cdot \frac{64}{15\pi} \approx 0.784,$$

whereas the bound from Jensen's inequality for $d = 4$ given in Corollary 9 is $\sqrt{2/3} \approx 0.816$. For higher dimensions d the difference between the bounds derived from Jensen's inequality in Corollary 9 and the better ones derived from (11) and Theorem 11 become smaller and smaller as the latter ones are also approximately $\sqrt{d/6}$ for large d. A similar phenomenon has been observed for bounds on risk measures in [17] and [3].

Remark 13. We also notice that $S(C^\circ, C^\circ)$ is not an upper bound for the case that we allow the copulas to be different, as we have an explicit counterexample in Example 10 with

$$S(C^+, C^{\|}) = \frac{1}{\sqrt{2}} \int_0^1 \int_0^1 \sqrt{\frac{1}{2} + \left(\frac{1}{2} + x - y\right)^2}\, dxdy \approx 0.549 > \frac{\pi}{6} = S(C^\circ, C^\circ).$$

4 Bounds on the energy score

Let F be a distribution and $\boldsymbol{y}$ an observation. We now first study bounds on $S_\beta(F, \boldsymbol{y})$ in order to obtain bounds on the energy score $ES_\beta(F, \boldsymbol{y})$.

4.1 Bounds on $S_\beta(F, y)$

First, note that we cannot expect a general upper bound for $\boldsymbol{y} \mapsto S_\beta(F, \boldsymbol{y})$ as $S_\beta(F, \boldsymbol{y}) \to \infty$ for $\boldsymbol{y} \to \infty$. It is possible, however, to find general lower bounds and they are very important in applications like in the

definition of the copula energy score in [26]. We therefore concentrate on studying the lower bound in what follows.

It is clear that as a function of $\boldsymbol{y}$, the expression of $S_\beta(F, \boldsymbol{y})$ is small if $\boldsymbol{y}$ is in some sense near the center of the distribution. In the univariate case it is a well-known simple result that $y \mapsto \mathbb{E}|X - y|$ is minimized for $y^\star := F_X^{-1}(1/2)$ being the median of the distribution of X.

We can prove a similar lower bound in the multivariate case for the upper Fréchet bound F^+, if we assume that the marginals are symmetric and unimodal. Recall that a univariate distribution with cdf F is called symmetric and unimodal with respect to some $\mu \in \mathbb{R}$, if $F(\mu + t) + F(\mu - t) = 1$ for all $t > 0$ and F is convex on $(-\infty, \mu)$ and concave on (μ, ∞). Notice that unimodal distributions have a density if they are continuous in μ. We will need the following simple Lemma for such distributions that we state with a proof here, as we could not find it in the literature.

Lemma 14. *Assume that the random variable X has a continuous, unimodal and symmetric distribution with respect to μ. Then it holds for all $\mu \le x < y$ and all $\mu \ge x > y$ that*

$$|X - x| \le_{st} |X - y|.$$

Proof. Denote by F_x the cdf of $|X - x|$, and assume $\mu \le x < y$. We have to show that $F_x(t) \ge F_y(t)$ for all $t \ge 0$. A simple calculation shows $F_x(t) = F(x + t) - F(x - t)$. As F is continuous, unimodal and symmetric, it has a density f which is symmetric around μ and decreasing on $[\mu, \infty)$. Therefore

$$\frac{\partial}{\partial x} F_x(t) = f(x + t) - f(x - t) \le 0$$

as $|x + t - \mu| > |x - t - \mu|$. This implies the assertion for $\mu \le x < y$ and the case $\mu \ge x > y$ follows then by symmetry. □

Theorem 15. *Assume that the random vector $\boldsymbol{X}$ has a cdf F with marginals F_i that are continuous, unimodal and symmetric with respect to μ_i, $i = 1, \dots, d$. Then we have for all $\boldsymbol{y} \in \mathbb{R}^d$*

$$S_\beta(F, \boldsymbol{y}) \ge S_\beta(F^+, \boldsymbol{\mu}).$$

Proof. We have

$$S_\beta(F, \boldsymbol{y}) = \mathbb{E}\left(\sum_{i=1}^d Z_{i,y_i}^2\right)^{\frac{\beta}{2}}, \tag{12}$$

where $Z_{i,y_i}^2 = (X_i - y_i)^2$. It follows from Lemma 14 that

$$(X_i - y_i)^2 \ge_{st} (X_i - \mu_i)^2$$

for all i. Let us denote by $(Z_{i,y_i}^+)^2, i = 1, \dots, d$ comonotone random variables with the same distributions as Z_{i,y_i}^2. Notice that for general $\boldsymbol{y}$ the vector $(Z_{1,y_1}^2, \dots, Z_{d,y_d}^2)$ is typically not comonotone, even if F is the comonotonic upper Fréchet bound $F = F^+$. This is the case, however, if $y_i = \mu_i$ is the median for all $i = 1, \dots, d$. Therefore we get

$$S_\beta(F, \boldsymbol{y}) = \mathbb{E}\left(\sum_{i=1}^d Z_{i,y_i}^2\right)^{\frac{\beta}{2}} \ge \mathbb{E}\left(\sum_{i=1}^d (Z_{i,y_i}^+)^2\right)^{\frac{\beta}{2}} \tag{13}$$

$$\ge \mathbb{E}\left(\sum_{i=1}^d (Z_{i,\mu_i}^+)^2\right)^{\frac{\beta}{2}} = S_\beta(F^+, \boldsymbol{\mu}). \tag{14}$$

The first inequality follows as in the proof of Theorem 6. The second inequality follows from the fact that for random vectors $\boldsymbol{Z}$ and $\boldsymbol{Z}'$ with the same copula C^+ stochastic ordering $Z_h \le_{st} Z_h'$ of the marginals implies multivariate stochastic ordering and thus $\mathbb{E}f(\boldsymbol{Z}) \le \mathbb{E}f(\boldsymbol{Z}')$ for all increasing functions $f : \mathbb{R}^d \to \mathbb{R}$, see e.g., [12], Theorem 3.3.8. □

For copulas we get the following corollary. We denote here by $\mathbf{1} = (1, \dots, 1)$ a vector with all components being equal to one.

Corollary 16. *For all copulas C and all $\boldsymbol{y} \in [0,1]^d$ it holds*

$$S_\beta(C, \boldsymbol{y}) \geq S_\beta\left(C^+, \frac{1}{2}\mathbf{1}\right) = d^{\frac{\beta}{2}} \frac{\frac{1}{2}^\beta}{\beta + 1}.$$

Proof. This follows from Theorem 15, as uniform distributions U on $(0,1)$ are symmetric and unimodal with respect to $1/2$ and

$$\mathbb{E}\left(|U - 1/2|^\beta\right) = \frac{\frac{1}{2}^\beta}{\beta + 1}.$$

□

In the case of copulas we can also have a look at the case that the observation $\boldsymbol{y}$ is an extreme point, which can be assumed to be without loss of generality $\boldsymbol{y} = \mathbf{0}$.

Theorem 17. *For $\boldsymbol{y} = \mathbf{0}$ the function*

$$C \mapsto S_\beta(C, \mathbf{0})$$

attains its minimum for the upper Fréchet bound C^+.

Proof. The proof for the minimum of $C \mapsto S_\beta(C, \mathbf{0})$ follows the same lines as in Theorem 6. If $\boldsymbol{U}$ is a random vector with copula C, then

$$S_\beta(C, \mathbf{0}) = \mathbb{E} f(Z_1, \dots, Z_d)$$

where $Z_i = U_i^2,\ i = 1, \dots, d$, and f is a submodular function. As $\boldsymbol{Z} = (Z_1, \dots, Z_d)$ has the same copula as $\boldsymbol{U}$, we can conclude that $S_\beta(C, \mathbf{0}) \geq S_\beta(C^+, \mathbf{0})$. □

It is easy to see that it is not true in general that the upper Fréchet bound C^+ minimizes $C \mapsto S_\beta(C, \boldsymbol{y})$. Due to invariance under rotations the minimum is obtained for the lower Fréchet bound C^- if $\boldsymbol{y} = (0,1)$.

4.2 Bounds on $ES_\beta(F, y)$

Similarly to the above study of $S_\beta(F, \boldsymbol{y})$, one cannot expect a general upper bound for the energy score

$$ES_\beta(F, \boldsymbol{y}) = S_\beta(F, \boldsymbol{y}) - \frac{1}{2} S_\beta(F, F)$$

as the quantity tends to $+\infty$ for $\boldsymbol{y} \to \infty$. As we have to deal with a difference of two quantities, it is also in general more difficult to find sharp bounds, whereas one easily gets some bounds by bounding each of the two quantities using our previous results.

We now first consider a bounded domain for $\boldsymbol{y}$. We then characterize the copula that achieves the lower bound.

As $\boldsymbol{y} \mapsto \|\boldsymbol{x} - \boldsymbol{y}\|_2^\beta$ is convex for $\beta \geq 1$, we also get that $\boldsymbol{y} \mapsto S_\beta(F, \boldsymbol{y})$ is convex in this case. From Lemma 5, we thus can easily derive the following result.

Lemma 18. *For $\beta \in [1,2)$ the functions $\boldsymbol{y} \mapsto ES_\beta(F, \boldsymbol{y})$ and $F \mapsto ES_\beta(F, \boldsymbol{y})$ are convex.*

Considering a copula C and an observation $\boldsymbol{y} \in [0,1]^d$ we immediately get the following consequence.

Proposition 19. *For $\beta \in [1,2)$ the function*

$$\boldsymbol{y} \mapsto ES_\beta(C, \boldsymbol{y}), \quad \boldsymbol{y} \in [0,1]^d,$$

attains a maximum in the set $\{0,1\}^d$.

Proof. This immediately follows from the fact that a convex function on a compact convex domain attains a maximum in an extreme point. □

Following the results on the lower bound of $S_\beta(F, \boldsymbol{y})$ obtained for copulas in Section 4.1, it is natural to conjecture that the upper Fréchet bound C^+ is also the minimum of $C \mapsto ES_\beta(C, \mathbf{0})$. This is not true, however. We can show that in the bivariate case $ES(C^+, \mathbf{0}) > ES(\hat{C}, \mathbf{0})$ for the copula $\hat{C}$ of the following random vector $\hat{\boldsymbol{U}}$. Let $U, U' \sim U(0, 1)$ be independent uniformly distributed random variables and define $\hat{\boldsymbol{U}} = (\hat{U}_1, \hat{U}_2)$ as follows: $\hat{U}_1 = \hat{U}_2 = U$, if $U \le 1/2$. If $U > 1/2$ then $\hat{U}_1 = U$ and $\hat{U}_2 = (U' + 1)/2$.

We are not able to obtain an explicit lower bound but we can characterize some properties of the copula that achieves the minimum energy score.

Let $T : \mathbb{R}^d \to \mathbb{R}^d$ be a transformation that is an isometry and that preserves the marginal distributions in the following sense: if $\boldsymbol{X}$ is a random vector with distribution function F and marginals $F_1, \dots, F_d$, then $T(\boldsymbol{X})$ also has the same marginals and for any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$ we have

$$\|T(\boldsymbol{x}) - T(\boldsymbol{y})\|_2 = \|\boldsymbol{x} - \boldsymbol{y}\|_2.$$

It is easy to see that in the case of uniform marginals, i.e., for copulas, this holds for reflections of the form

$$T_i(\boldsymbol{x}) = (x_1, \dots, x_{i-1}, 1 - x_i, x_{i+1}, \dots, x_d) \tag{15}$$

and for permutations π of the coordinates

$$T_\pi(\boldsymbol{x}) = (x_{\pi_1}, \dots, x_{\pi_d}). \tag{16}$$

Let us denote by C_T the copula of the transformation $T(\boldsymbol{U})$, if $\boldsymbol{U}$ is a random vector with copula C, and let H be the finite group generated by all these isometric transformations of the hypercube that preserve the copula property. This symmetry group of the hypercube is also called *hyper-octahedral group* and is generated by the reflections and permutations described in (15) and (16), see e.g. [25].

We get the following theorem.

Theorem 20. *For $\beta \in [1, 2)$ and $\boldsymbol{y} = \frac{1}{2}\mathbf{1}$ the function*

$$C \mapsto ES_\beta(C, \boldsymbol{y})$$

attains a minimum for a copula that is invariant under H.

Proof. Let us define for a fixed C and $\boldsymbol{y} = \frac{1}{2}\mathbf{1}$

$$\hat{\boldsymbol{y}} := \frac{1}{|H|} \sum_{T \in H} T(\boldsymbol{y})$$

and

$$\hat{C} := \frac{1}{|H|} \sum_{T \in H} C_T.$$

Then $\hat{\boldsymbol{y}} = \boldsymbol{y}$, $T(\hat{\boldsymbol{y}}) = \hat{\boldsymbol{y}}$ for all $T \in H$ and $\hat{C}$ is invariant under H. Due to convexity of $C \mapsto ES_\beta(C, \boldsymbol{y})$ we get

$$\begin{aligned} ES_\beta(\hat{C}, \hat{\boldsymbol{y}}) &\le \frac{1}{|H|} \sum_{T \in H} ES_\beta(C_T, \hat{\boldsymbol{y}}) = \frac{1}{|H|} \sum_{T \in H} ES_\beta(C_T, T(\hat{\boldsymbol{y}})) \\ &= ES_\beta(C, \hat{\boldsymbol{y}}). \end{aligned}$$

□

With a similar argument we can show that the function $C \mapsto ES_\beta(C, \mathbf{0})$ attains a minimum for a copula that is invariant under permutations, but we are not able to derive an explicit solution for the moment.

5 Conclusions

We investigated dependence uncertainty bounds on the energy score and for related functionals. The obtained results indicate that indeed these functionals seem not to be very sensitive to the dependence structure as one can see e.g., from the inequality

$$\frac{1}{3}\sqrt{d} \le S(C_1, C_2) \le \sqrt{\frac{d}{6}}$$

in Corollary 9, which holds for all copulas C_1, C_2. Notice that we get the even closer sharp bounds

$$\frac{\sqrt{2}}{3} \le S(C, C) \le \frac{\pi}{6}$$

in Theorem 12 for the case $d = 2$ if we restrict to the case of equal copulas. Therefore our results support the corresponding claim in [15] which was based on a simulation study using multivariate normal distributions. However, many questions remain open and we hope to stimulate research on this topic that we consider as important. For example, we are not able to find explicitly the copula that achieves the lower bound of the energy score. We are only able to provide a partial characterization of it in Theorem 20. We are also working on the problem of finding the numerical solution of this optimization problem by using a variant of the swapping algorithm that was used in [16] for a related problem. First results indicate that the solution seems to be a copula with a very unusual shape, but these results will be reported in a forthcoming paper.

Acknowledgements: We thank Mohamed Barakat, Tilmann Gneiting and Marco Oesting as well as two anonymous referees for many helpful comments. C. Bernard gratefully acknowledges support from the Odysseus FWO research grant at Vrije Universiteit Brussel.

References

[1] Buja, A., B. F. Logan, J. A. Reeds, and L. A. Shepp (1994). Inequalities and positive-definite functions arising from a problem in multidimensional scaling. *Ann. Statist. 22*(1), 406–438.
[2] Eaton, M. L. (1981). On the projections of isotropic distributions. *Ann. Statist. 9*(2), 391–400.
[3] Embrechts, P., B. Wang, and R. Wang (2015). Aggregation-robustness and model uncertainty of regulatory risk measures. *Finance Stoch. 19*(4), 763–790.
[4] Fréchet, M. (1951). Sur les tableaux de corrélation dont les marges sont données. *Ann. Univ. Lyon Sect. A (3) 14*, 53–77.
[5] Ghoussoub, N., Y.-H. Kim, and T. Lim (2019). Structure of optimal martingale transport plans in general dimensions. *Ann. Probab. 47*(1), 109–164.
[6] Gneiting, T. and A. E. Raftery (2007). Strictly proper scoring rules, prediction, and estimation. *J. Amer. Statist. Assoc. 102*(477), 359–378.
[7] Hoeffding, W. (1940). Masstabinvariante Korrelationstheorie. *Schriftenreihe Math. Inst. Angew. Math. Univ. Berlin 5*, 179–233. English translation as Hoeffding, W. (1994). Scale-invariant correlation theory. In N. Fisher and P. Sen (Eds.), *The Collected Works of Wassily Hoeffding,* pp. 57–107. Springer, New York.
[8] Joe, H. (1997). *Multivariate Models and Dependence Concepts*. Chapman & Hall, London.
[9] Koshevoy, G. A. and K. Mosler (1997). Multivariate Gini indices. *J. Multivariate Anal. 60*(2), 252–276.
[10] Mattner, L. (1993). Extremal problems for probability distributions: a general method and some examples. In M. Shaked and Y. L. Tong (Eds.), *Stochastic Inequalities,* pp. 274–283. Institute of Mathematical Statistics, Hayward CA.
[11] Müller, A. and L. Rüschendorf (2001). On the optimal stopping values induced by general dependence structures. *J. Appl. Probab. 38*(3), 672–684.
[12] Müller, A. and D. Stoyan (2002). *Comparison Methods for Stochastic Models and Risks*. John Wiley & Sons, Chichester.
[13] Nelsen, R. B. (2006). *An Introduction to Copulas*. Second edition. Springer, New York.
[14] Perlman, M. D. and J. A. Wellner (2011). Squaring the circle and cubing the sphere: circular and spherical copulas. *Symmetry 3*(3), 574–599.
[15] Pinson, P. and J. Tastu (2013). Discrimination ability of the energy score. Available at https://orbit.dtu.dk/en/publications/discrimination-ability-of-the-energy-score.
[16] Puccetti, G. (2017). An algorithm to approximate the optimal expected inner product of two vectors with given marginals. *J. Math. Anal. Appl. 451*(1), 132–145.

[17] Puccetti, G. and L. Rüschendorf (2014). Asymptotic equivalence of conservative value-at-risk- and expected shortfall-based capital charges. *J. Risk 16*(3), 3–22.
[18] Puccetti, G. and R. Wang (2015). Extremal dependence concepts. *Statist. Sci. 30*(4), 485–517.
[19] Rüschendorf, L. (1983). Solution of a statistical optimization problem by rearrangement methods. *Metrika 30*(1), 55–61.
[20] Rüschendorf, L. and L. Uckelmann (1997). On optimal multivariate couplings. In V. Beneš and J. Štěpán (Eds.), *Distributions with Given Marginals and Moment Problems*, pp. 261–273. Springer, Dordrecht.
[21] Schwarz, G. (1985). Multivariate distributions with uniformly distributed projections. *Ann. Probab. 13*(4), 1371–1372.
[22] Székely, G. J. and M. L. Rizzo (2017). The energy of data. *Annu. Rev. Stat. Appl. 4*, 447–479.
[23] Tchen, A. H. (1980). Inequalities for distributions with given marginals. *Ann. Probab. 8*(4), 814–827.
[24] Yitzhaki, S. (2003). Gini's mean difference: a superior measure of variability for non-normal distributions. *Metron 61*(2), 285–316.
[25] Young, A. (1930). On quantitative substitutional analysis (fifth paper). *Proc. Lond. Math. Soc. (2) 31*, 273–288.
[26] Ziel, F. and K. Berk (2019). Multivariate forecasting evaluation: On sensitive and strictly proper scoring rules. Available at https://arxiv.org/abs/1910.07325.