Research Article **Open Access**

Helena Ferreira* and Marta Ferreira

Multivariate medial correlation with applications

https://doi.org/10.1515/demo-2020-0019
Received September 28, 2020; accepted November 23, 2020

Abstract: We define a multivariate medial correlation coefficient that extends the probabilistic interpretation and properties of Blomqvist's β coefficient, incorporates multivariate marginal dependencies and it preserves a partial ordering stronger than concordance relation. We illustrate the results in some models and provide an application on real datasets.

Keywords: Blomqvist β, multivariate medial correlation, multivariate stochastic order

MSC: 62H20

1 Introduction

Let us consider that $\mathbf{X} = (X_1, X_2)$ is a real random vector, over the probability space $(\Omega, \mathcal{A}, P)$, with continuous marginal distribution functions F_{X_i}, $i = 1, 2$, and let (U_1, U_2) represent the corresponding uniformized vector, that is, $U_i = F_{X_i}(X_i)$, $i = 1, 2$.

The medial correlation coefficient of (X_1, X_2), which we will represent by $\beta(X_1, X_2)$ or $\beta(\mathbf{X})$, is defined by

$$\beta(X_1, X_2) = P\left(\left(U_1 - \frac{1}{2}\right)\left(U_2 - \frac{1}{2}\right) > 0\right) - P\left(\left(U_1 - \frac{1}{2}\right)\left(U_2 - \frac{1}{2}\right) < 0\right). \tag{1}$$

The β coefficient introduced by Blomqvist ([1]), has its value in $[-1, 1]$ and compares the propensity for the margins of (X_1, X_2) to take both values above or both values below their respective medians, with the propensity for the occurrence of the contrary event.

Since

$$\beta(X_1, X_2) = 2\left(P\left(U_1 > \frac{1}{2}, U_2 > \frac{1}{2}\right) + P\left(U_1 < \frac{1}{2}, U_2 < \frac{1}{2}\right)\right) - 1, \tag{2}$$

and

$$\beta(X_1, X_2) = 4P\left(U_1 < \frac{1}{2}, U_2 < \frac{1}{2}\right) - 1, \tag{3}$$

if $C_{\mathbf{X}}(u_1, u_2)$ and $\hat{C}_{\mathbf{X}}(u_1, u_2)$, $(u_1, u_2) \in [0, 1]^2$, represent the copula and the survival copula of $\mathbf{X}$ (Nelsen [8]), respectively, we can say that

$$\beta(X_1, X_2) = 2\left(C_{\mathbf{X}}\left(\frac{1}{2}, \frac{1}{2}\right) + \hat{C}_{\mathbf{X}}\left(\frac{1}{2}, \frac{1}{2}\right)\right) - 1, \tag{4}$$

and

$$\beta(X_1, X_2) = 4C_{\mathbf{X}}\left(\frac{1}{2}, \frac{1}{2}\right) - 1. \tag{5}$$

***Corresponding Author: Helena Ferreira:** Universidade da Beira Interior, Centro de Matemática e Aplicações (CMA-UBI), Avenida Marquês d'Avila e Bolama, 6200-001 Covilhã, Portugal, E-mail: helenaf@ubi.pt
Marta Ferreira: Center of Mathematics of Minho University, Center for Computational and Stochastic Mathematics of University of Lisbon, Center of Statistics and Applications of University of Lisbon, Portugal, E-mail: msferreira@math.uminho.pt

The bivariate medial correlation coefficient $\beta(X_1, X_2)$, which can also be denoted by $\beta(C_{\mathbf{X}})$, enables to compare $C_{\mathbf{X}}(u_1, u_2)$ on $Q_L \cup Q_U = \left[0, \frac{1}{2}\right]^2 \cup \left]\frac{1}{2}, 1\right]^2$ with $C_{\mathbf{X}}(u_1, u_2)$ on $[0,1]^2 \setminus (Q_L \cup Q_U)$ or to compare $C_{\mathbf{X}}(u_1, u_2)$ on $Q_L = \left[0, \frac{1}{2}\right]^2$ with $C_{\mathbf{X}}(u_1, u_2)$ on $[0,1]^2 \setminus Q_L$.

The medial correlation coefficient can be related to other measures of global dependence in (X_1, X_2), or in $C_{\mathbf{X}}$, such as Spearman's ρ or Kendall's τ (Nelsen [8], Joe [3], Lebedev [6] and references therein).

Two bivariate vectors $\mathbf{X}$ and $\mathbf{Y}$, or their copulas, can be partially ordered by punctually comparing their copulas. We say that $\mathbf{X}$ is less concordant than $\mathbf{Y}$, and we write for that $\mathbf{X} \prec_c \mathbf{Y}$, if $C_{\mathbf{X}}(u_1, u_2) \le C_{\mathbf{Y}}(u_1, u_2)$, $(u_1, u_2) \in [0,1]^2$, or equivalent, if $\hat{C}_{\mathbf{X}}(u_1, u_2) \le \hat{C}_{\mathbf{Y}}(u_1, u_2)$, $(u_1, u_2) \in [0,1]^2$ (Nelsen [8]).

Thus, from the representations (4) or (5), we verify that

$$\text{if } \mathbf{X} \prec_c \mathbf{Y} \text{ then } \beta(\mathbf{X}) \le \beta(\mathbf{Y}). \tag{6}$$

In addition to the increasing with concordance ordering, the bivariate medial correlation coefficient β satisfies other properties that shape the definition of measure of concordance according to Scarsini ([9]).

Considering the countermonotonicity, independence and comonotonicity copulas, respectively, $C_W(u_1, u_2) = (u_1 + u_2 - 1) \vee 0$, $C_\Pi(u_1, u_2) = u_1 u_2$ and $C_M(u_1, u_2) = u_1 \wedge u_2$, $(u_1, u_2) \in [0,1]^2$, we have $C_W \prec_c C_{\mathbf{X}} \prec_c C_M$, $\beta(C_W) = -1$, $\beta(C_\Pi) = 0$, $\beta(C_M) = 1$ and we can also represent $\beta(X_1, X_2)$ by

$$\beta(X_1, X_2) = 2\left(C_{\mathbf{X}}\left(\frac{1}{2}, \frac{1}{2}\right) - C_\Pi\left(\frac{1}{2}, \frac{1}{2}\right) + \hat{C}_{\mathbf{X}}\left(\frac{1}{2}, \frac{1}{2}\right) - \hat{C}_\Pi\left(\frac{1}{2}, \frac{1}{2}\right)\right). \tag{7}$$

For a random vector $\mathbf{X} = (X_1, \ldots, X_d)$ with dimension $d > 2$, if we think about generalizing (1) to $P\left(\prod_{i=1}^{d}\left(U_i - \frac{1}{2}\right) > 0\right) - P\left(\prod_{i=1}^{d}\left(U_i - \frac{1}{2}\right) < 0\right)$ we definitely loose:
(i) interpretation as a measure of propensity for all margins to exceed their respective medians or all margins to be below their medians, and
(ii) information about the behaviour of $C_{\mathbf{X}}$ on $Q_k = \prod_{j=1}^{d} I_j$, $k = 1, \ldots, d-1$, where $I_j = \left[0, \frac{1}{2}\right]$ for k or $d-k$ values of j and $I_j = \left]\frac{1}{2}, 1\right]$ for the others.
On the other hand, any generalization of β in the multivariate context must preserve at least the property (i) and also verify
(iii) $\beta(C_\Pi) = 0$ and $\beta(C_M) = 1$.

The proposals of Nelsen ([7]), Úbeda-Flores ([13]) and Schmid and Schmidt ([10]) manage to keep (i) and (iii) above.

Starting from the multivariate version of (5), $4C_{\mathbf{X}}(\frac{1}{2}, \ldots, \frac{1}{2}) - 1$, rescaled by considering the quotient between its distance to the corresponding value for C_Π and the maximum value of that distance,

$$\begin{aligned} \beta'(X_1, \ldots, X_d) &= \frac{4C_{\mathbf{X}}\left(\frac{1}{2}, \ldots, \frac{1}{2}\right) - 1 - \left(4\left(\frac{1}{2}\right)^d - 1\right)}{4C_M\left(\frac{1}{2}, \ldots, \frac{1}{2}\right) - 1 - \left(4\left(\frac{1}{2}\right)^d - 1\right)} \\ &= \frac{2^d C_{\mathbf{X}}\left(\frac{1}{2}, \ldots, \frac{1}{2}\right) - 1}{2^{d-1} - 1}, \end{aligned} \tag{8}$$

we find Nelsen's generalization ([7]).

Úbeda-Flores ([13]) proposes the extension of (4) in

$$2\left(C_{\mathbf{X}}\left(\frac{1}{2}, \ldots, \frac{1}{2}\right) + \hat{C}_{\mathbf{X}}\left(\frac{1}{2}, \ldots, \frac{1}{2}\right)\right) - 1, \tag{9}$$

also rescaled by considering the quotient between its distance to the corresponding value for C_Π and the maximum value of that distance. In this way, we obtain the following generalization of β, which we will denote by β^* and where $\mathbf{\frac{1}{2}}$ represents the vector of suitable size and coordinates all equal to $\frac{1}{2}$:

$$\begin{aligned}\beta^*(X_1,\ldots,X_d) &= \frac{2\left(C_{\mathbf{X}}\left(\frac{1}{2},\ldots,\frac{1}{2}\right)+\hat{C}_{\mathbf{X}}\left(\frac{1}{2},\ldots,\frac{1}{2}\right)\right)-1-\left(\frac{1}{2^{d-2}}-1\right)}{2\left(C_M\left(\frac{1}{2},\ldots,\frac{1}{2}\right)+\hat{C}_M\left(\frac{1}{2},\ldots,\frac{1}{2}\right)\right)-1-\left(\frac{1}{2^{d-2}}-1\right)}\\ &= \frac{2^{d-1}\left(C_{\mathbf{X}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right)+\hat{C}_{\mathbf{X}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right)\right)-1}{2^{d-1}-1},\end{aligned} \tag{10}$$

which coincides with (8) when $C = \hat{C}$.

Reasoning in an equivalent way about (7), Schmid and Schmidt ([10]) propose

$$\frac{2\left(C_{\mathbf{X}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right)-C_{\Pi}\left(\frac{\mathbf{1}}{\mathbf{2}}\right)+\hat{C}_{\mathbf{X}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right)-\hat{C}_{\Pi}\left(\frac{\mathbf{1}}{\mathbf{2}}\right)\right)}{2\left(C_M\left(\frac{\mathbf{1}}{\mathbf{2}}\right)-C_{\Pi}\left(\frac{\mathbf{1}}{\mathbf{2}}\right)+\hat{C}_M\left(\frac{\mathbf{1}}{\mathbf{2}}\right)-\hat{C}_{\Pi}\left(\frac{\mathbf{1}}{\mathbf{2}}\right)\right)} = \frac{2^{d-1}\left(C_{\mathbf{X}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right)+\hat{C}_{\mathbf{X}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right)\right)-1}{2^{d-1}-1},$$

finding again the expression of Úbeda-Flores ([13]). In addition to this extension, Schmid and Schmidt ([10]) make a detailed study of a function resulting from a rescaling of $C_{\mathbf{X}}(\mathbf{u})+\hat{C}_{\mathbf{X}}(\mathbf{v})$, $\mathbf{u},\mathbf{v}\in[0,1]^d$, putting emphasis on the tail regions of the copula which determine the degree of large co-movements between the marginal random variables.

In order to keep (i), (ii) and (iii), we have Joe's sophisticated proposal ([4]) with an axiomatic on linear combinations of $C_{\sigma_{i_1}\sigma_{i_2}\ldots\sigma_{i_k}\mathbf{X}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right)$ and $\hat{C}_{\sigma_{i_1}\sigma_{i_2}\ldots\sigma_{i_k}\mathbf{X}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right)$, $1\le i_1<\ldots<i_K\le d$, $k=[\frac{d+1}{2}],\ldots,d$, where $\sigma_j\mathbf{X}$ denotes the j-th reflection of $\mathbf{X}$, that is, the vector $(X_1,\ldots X_{j-1},-X_j,X_{j+1},\ldots,X_d)$. Joe's axiomatic definition allows for various extensions of β, including those mentioned above and the arithmetic mean of $\beta(X_i,X_j)$, $1\le i<j\le d$.

The extensions referred for β increase with the multivariate concordance (Joe [5]). We say that $\mathbf{X}=(X_1,\ldots,X_d)$ is less concordant than $\mathbf{Y}=(Y_1,\ldots,Y_d)$, or $C_{\mathbf{X}}$ is less concordant than $C_{\mathbf{Y}}$, and in this case we write $\mathbf{X}\prec_c\mathbf{Y}$, when we have

$$C_{\mathbf{X}}(\mathbf{u})\le C_{\mathbf{Y}}(\mathbf{u}) \text{ and } \hat{C}_{\mathbf{X}}(\mathbf{u})\le\hat{C}_{\mathbf{Y}}(\mathbf{u}), \tag{11}$$

for $\mathbf{u}\in[\mathbf{0},\mathbf{1}]^d$. In the case of $d=2$ the two conditions are equivalent, as we have already mentioned.

The above proposed generalizations start from extensions of the representations of bivariate β in terms of copulas, considering the corresponding multivariate copulas.
The proposal that we will make, in the next section, for a multivariate correlation coefficient $\beta(\mathbf{X})$ starts from a generalization of the probabilistic interpretation of the definition (1) and satisfies almost all the desirable properties for a multivariate concordance measure (Taylor [11],[12]). It preserves a multivariate partial order relation that we introduce in section 4. We present several representations for $\beta(\mathbf{X})$, we demonstrate the main properties, relate it to the previously mentioned coefficients and illustrate with examples and applications.

2 Motivation for the multivariate medial correlation coefficient

For $d\ge 2$, $D=\{1,\ldots,d\}$, $I\subset D$, $\mathbf{X}=(X_1,\ldots,X_d)$ with continuous marginal distributions and $\mathbf{U}=(U_1,\ldots,U_d)=\left(F_{X_1}(X_1),\ldots,F_{X_d}(X_d)\right)$, we define

$$M(I)=\bigvee_{i\in I}U_i \text{ and } W(I)=\bigwedge_{i\in I}U_i, \tag{12}$$

where $\vee$ and $\wedge$ are the notations for the maximum and minimum operators, respectively.

When further clarification is needed, we write $M_{\mathbf{X}}(I)$ and $W_{\mathbf{X}}(I)$. Inequalities between vectors are understood by corresponding inequalities between homologous coordinates. By $\mathbf{X}_I$ we understand the subvector

of $\mathbf{X}$ with margins in I and $\mathcal{P}(D)$ represents the family of subsets of D. When $|I| = 1$, where $|A|$ denotes the cardinality of A, we consider $C_{\mathbf{X}_I}\left(\frac{\mathbf{1}}{\mathbf{2}}\right) = \hat{C}_{\mathbf{X}_I}\left(\frac{\mathbf{1}}{\mathbf{2}}\right) = \frac{1}{2}$.

Let's fix disjoint I and J in $\mathcal{P}(D)$. The propensity for margins of $\mathbf{X}_I$ and margins of $\mathbf{X}_J$ simultaneously taking values below the respective medians or simultaneously values above the respective medians is evaluated by $C_{\mathbf{X}_{I\cup J}}(\frac{\mathbf{1}}{\mathbf{2}}) + \hat{C}_{\mathbf{X}_{I\cup J}}(\frac{\mathbf{1}}{\mathbf{2}})$, that is, the probability of $\mathbf{U}_{I\cup J}$ taking values in $\left[0, \frac{1}{2}\right]^{|I\cup J|} \cup \left]\frac{1}{2}, 1\right]^{|I\cup J|}$. If we want to compare this probability with the probability of $\mathbf{U}_{I\cup J}$ taking values in $[0,1]^{|I\cup J|} \setminus \left(\left[0, \frac{1}{2}\right]^{|I\cup J|} \cup \left]\frac{1}{2}, 1\right]^{|I\cup J|}\right)$, we can do it briefly by calculating the coefficients

$$\begin{aligned} &\overline{\beta}\left(M(I), M(J)\right) := \\ :=\ & P\left(\left(M(I) - \tfrac{1}{2}\right)\left(M(J) - \tfrac{1}{2}\right) > 0\right) - P\left(\left(M(I) - \tfrac{1}{2}\right)\left(M(J) - \tfrac{1}{2}\right) < 0\right) \\ =\ & 2\left(P\left(M(I) > \tfrac{1}{2}, M(J) > \tfrac{1}{2}\right) + P\left(M(I) < \tfrac{1}{2}, M(J) < \tfrac{1}{2}\right)\right) - 1 \end{aligned} \tag{13}$$

and

$$\begin{aligned} &\underline{\beta}(W(I), W(J)) := \\ :=\ & P\left(\left(W(I) - \tfrac{1}{2}\right)\left(W(J) - \tfrac{1}{2}\right) > 0\right) - P\left(\left(W(I) - \tfrac{1}{2}\right)\left(W(J) - \tfrac{1}{2}\right) < 0\right) \\ =\ & 2\left(P\left(W(I) > \tfrac{1}{2}, W(J) > \tfrac{1}{2}\right) + P\left(W(I) < \tfrac{1}{2}, W(J) < \tfrac{1}{2}\right)\right) - 1. \end{aligned} \tag{14}$$

Let us make some comments about

$$\beta_{I,J}(\mathbf{X}) := \frac{\overline{\beta}(M(I), M(J)) + \underline{\beta}(W(I), W(J))}{2}. \tag{15}$$

(i) The expressions (13), (14) and (15) have $\beta(X_i, X_j)$ as a particular case, if we take $I = \{i\}$ and $J = \{j\}$. If $I = D$, $J = \emptyset$ and we consider that $M(\emptyset) = -\infty$ and $W(\emptyset) = +\infty$, then (15) is equal to $C_{\mathbf{X}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right) + \hat{C}_{\mathbf{X}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right) - 1$, which can be rescaled in order to obtain the proposal of Úbeda-Flores ([13]) and Schmid and Schmidt ([10]).

(ii) Despite the random variables $M(I)$ and $W(I)$, $I \subset D$, may fail to be uniformly distributed and therefore their medians may fail to be equal $1/2$, we can say from (13) and (14) that the value of $\beta_{I,J}(\mathbf{X})$ encompasses the behavior of all the margins of $\mathbf{X}_I$ and all the margins of $\mathbf{X}_J$ in relation to their medians, indicating a multivariate medial information about $\mathbf{X}$.

(iii) Since $\beta_{I,J}(\mathbf{X})$ is defined as an average of bivariate coefficients, it can be estimated by the methods available for the bivariate context (Blomqvist [1], Schmid and Schmidt [10] and references therein).

(iv) If $C_{\mathbf{X}} = C_M$ we have $\beta_{I,J}(\mathbf{X}) = 1$ and if $C_{\mathbf{X}} = C_{\Pi}$ then $\beta_{I,J}(\mathbf{X}) = 2^{2-|I|-|J|} - 2^{1-|I|} - 2^{1-|J|} + 1 = (2^{1-|I|} - 1)(2^{1-|J|} - 1)$. This value becomes null if and only if $|I| = 1$ or $|J| = 1$.

(v) A linear combination of $\beta_{\{i\},\{j\}}(\mathbf{X})$, $1 \le i < j \le d$, takes into account the bivariate dependencies in $\mathbf{X}$, but if we consider some function of the coefficients $\beta_{I,J}(\mathbf{X})$, with $I, J \in \mathcal{F}$, for some family $\mathcal{F} \subset \mathcal{P}(D)$ containing sets with more than one element, then we will be incorporating multivariate marginal dependencies.

The definition we propose, in the next section, for a multivariate medial correlation coefficient, will be based on the bivariate coefficients $\beta_{\{i\},D\setminus\{i\}}(\mathbf{X})$, $1 \le i \le d$, incorporating the dependency between each margin X_i and $\mathbf{X}_{D\setminus\{i\}}$, $1 \le i \le d$.
Our proposal contains, as a particular case, the Blomqvist bivariate coefficient, extends the probabilistic interpretation (1), takes values in $[-1, 1]$, becoming null naturally when $C_{\mathbf{X}} = C_{\Pi}$ and taking the maximum value when $C_{\mathbf{X}} = C_M$. The rest of the properties we proved allow us to consider it a measure for a multivariate concordance relation stronger than concordance order.

3 A multivariate medial correlation coefficient

We will propose to evaluate the multivariate medial correlation by comparing the propensity for all margins of $\mathbf{X}$ simultaneously taking values below the respective medians or all margins to exceed their respective

medians with the propensity of each margin X_i to contradicts this behavior. That is, we will take into account the coefficients $\beta_{I,J}$ with the particular choice of $I = \{i\}$ and $J = D \setminus \{i\}$, $i = 1, \ldots, d$.

Definition 3.1. *The multivariate medial correlation coefficient of the vector* $\mathbf{X}$ *with dimension* d*, or of its copula* $C_{\mathbf{X}}$*, is defined as*

$$\beta(\mathbf{X}) = \frac{1}{d}\sum_{i=1}^{d} \beta_{\{i\},D\setminus\{i\}}(\mathbf{X}), \tag{16}$$

where

$$\beta_{\{i\},D\setminus\{i\}}(\mathbf{X}) = \frac{\overline{\beta}\left(U_i, M(D\setminus\{i\})\right) + \underline{\beta}\left(U_i, W(D\setminus\{i\})\right)}{2}, \quad i = 1, \ldots, d. \tag{17}$$

We remark that, from comment (i), it can be concluded that $\beta(\mathbf{X})$ coincides with the Blomqvist coefficient when $d = 2$.

Below we present some representations of $\beta(\mathbf{X})$ that will be useful to clarify their properties and interpretation.

The following

$$\begin{aligned}\beta_{\{i\},D\setminus\{i\}}(\mathbf{X}) \quad &= 2\left(P\left(U_i < \tfrac{1}{2}, M\left(D\setminus\{i\}\right) < \tfrac{1}{2}\right) + P\left(U_i > \tfrac{1}{2}, W\left(D\setminus\{i\}\right) > \tfrac{1}{2}\right)\right)\\ &\quad -P\left(M\left(D\setminus\{i\}\right) < \tfrac{1}{2}\right) - P\left(W\left(D\setminus\{i\}\right) > \tfrac{1}{2}\right),\end{aligned} \tag{18}$$

holds, generalizing (2). We also have

$$\beta_{\{i\},D\setminus\{i\}}(\mathbf{X}) = 2\left(C_{\mathbf{X}}\left(\tfrac{1}{2}\right) + \hat{C}_{\mathbf{X}}\left(\tfrac{1}{2}\right)\right) - C_{\mathbf{X}_{\mathbf{D}\setminus\{\mathbf{i}\}}}\left(\tfrac{1}{2}\right) - \hat{C}_{\mathbf{X}_{\mathbf{D}\setminus\{\mathbf{i}\}}}\left(\tfrac{1}{2}\right), \tag{19}$$

generalizing (4). From the previous relation, it follows that

$$\beta_{\{i\},D\setminus\{i\}}(\mathbf{X}) = C_{\mathbf{X}}\left(\tfrac{1}{2}\right) + \hat{C}_{\mathbf{X}}\left(\tfrac{1}{2}\right) - C_{\sigma_i\mathbf{X}}\left(\tfrac{1}{2}\right) - \hat{C}_{\sigma_i\mathbf{X}}\left(\tfrac{1}{2}\right), \tag{20}$$

where $\sigma_i\mathbf{X}$ is the i-th reflection of $\mathbf{X}$, that is, $\sigma_i\mathbf{X} = (X_1, \ldots, X_{i-1}, -X_i, X_{i+1}, \ldots, X_d)$ and therefore $C_{\sigma_i\mathbf{X}}(\frac{\mathbf{1}}{\mathbf{2}}) = C_{(U_1,\ldots,U_{i-1},1-U_i,U_{i+1},\ldots,U_d)}(\frac{\mathbf{1}}{\mathbf{2}})$. We then obtain the following ways of representing the coefficient β.

Proposition 3.1. *The multivariate medial correlation coefficient of the vector* $\mathbf{X}$ *with dimension* d*, admits the following representations:*

$$\begin{aligned}\beta(\mathbf{X}) = \quad &2\left(P\left(\mathbf{U} \le \tfrac{\mathbf{1}}{\mathbf{2}}\right) + P\left(\mathbf{U} > \tfrac{\mathbf{1}}{\mathbf{2}}\right)\right)\\ &-\frac{1}{d}\sum_{i=1}^{d}\left(P\left(\mathbf{U}_{D\setminus\{i\}} \le \frac{\mathbf{1}}{\mathbf{2}}\right) + P\left(\mathbf{U}_{D\setminus\{i\}} > \frac{\mathbf{1}}{\mathbf{2}}\right)\right),\end{aligned} \tag{21}$$

$$\beta(\mathbf{X}) = 2\left(C_{\mathbf{X}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right) + \hat{C}_{\mathbf{X}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right)\right) - \frac{1}{d}\sum_{i=1}^{d}\left(C_{\mathbf{X}_{\mathbf{D}\setminus\{\mathbf{i}\}}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right) + \hat{C}_{\mathbf{X}_{\mathbf{D}\setminus\{\mathbf{i}\}}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right)\right), \tag{22}$$

$$\beta(\mathbf{X}) = C_{\mathbf{X}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right) + \hat{C}_{\mathbf{X}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right) - \frac{1}{d}\sum_{i=1}^{d}\left(C_{\sigma_i\mathbf{X}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right) + \hat{C}_{\sigma_i\mathbf{X}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right)\right). \tag{23}$$

The relation (23) rewritten in the form

$$\beta(\mathbf{X}) = \frac{1}{d}\sum_{i=1}^{d}\left(C_{\mathbf{X}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right) - C_{\sigma_i\mathbf{X}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right) + \hat{C}_{\mathbf{X}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right) - \hat{C}_{\sigma_i\mathbf{X}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right)\right), \tag{24}$$

reinforces the idea that $\beta(\mathbf{X})$ compares the propensity of each margin X_i to agree with the remaining margins together, $\mathbf{X}_{D\setminus\{i\}}$, and the propensity to disagree with them, when they are all above or all below their respective medians.

The above representations for β show that by considering β as a mapping on copulas it is linear with respect to convex combinations.

In the following, we establish relationships between $\beta(\mathbf{X})$ and the generalizations referred to in the introduction. By applying the definition (10) of $\beta^\star$, we conclude from the representation (23) that

$$\begin{aligned}\beta(\mathbf{X}) &= \frac{(2^{d-1}-1)\beta^\star(\mathbf{X})+1}{2^{d-1}} - \frac{1}{d}\sum_{i=1}^{d}\frac{(2^{d-1}-1)\beta^\star(\sigma_i\mathbf{X})+1}{2^{d-1}} \\ &= \frac{(2^{d-1}-1)}{2^{d-1}}\left(\beta^\star(\mathbf{X}) - \frac{1}{d}\sum_{i=1}^{d}\beta^\star(\sigma_i\mathbf{X})\right).\end{aligned} \tag{25}$$

By defining $\bar{N} = \sum_{i=1}^{d} \mathbf{1}_{\{U_i > \frac{1}{2}\}}$, the representation (23) of β leads to

$$\beta(\mathbf{X}) = P(\bar{N}=0) + P(\bar{N}=d) - \frac{1}{d}\left(P(\bar{N}=1) + P(\bar{N}=d-1)\right). \tag{26}$$

That fits Joe's representation (3.1.1) ([4]) with $w_d = 1$, $w_{d-1} = -\frac{1}{d}$ and the remaining weights w_i equal to zero.

Note that in the 3-dimensional case, the multivariate medial correlation coefficient β satisfies

$$\beta(\mathbf{X}) = \tfrac{4}{3}C_{\mathbf{X}}\left(\tfrac{\mathbf{1}}{\mathbf{2}}\right) + \tfrac{4}{3}\hat{C}_{\mathbf{X}}\left(\tfrac{\mathbf{1}}{\mathbf{2}}\right) - \tfrac{1}{3} = \beta^\star(\mathbf{X}) = \frac{\beta(X_1,X_2)+\beta(X_1,X_3)+\beta(X_2,X_3)}{3}.$$

Thus, in the 3-dimensional case β equals $\beta^\star$ and hence allows a different view on Blomqvist's β discussed in Úbeda-Flores ([13]).

We refer the properties of $\beta(\mathbf{X})$ in the next section and end this one with three examples.

Example 3.1. *Consider* $C_{\mathbf{X}}(u_1,\ldots,u_4) = \left(u_1^\delta \wedge u_2\right) u_1^{1-\delta}\left(u_3^\alpha \wedge u_4\right) u_3^{1-\alpha}$, *with* $0 \le \delta, \alpha \le 1$, *that is,* $C_{\mathbf{X}}$ *is the product of two Marshall-Olkin survival copulas ([5]). It holds that*

$$C_{\mathbf{X}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right) = \hat{C}_{\mathbf{X}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right) = \left(\frac{1}{2}\right)^{4-\delta-\alpha},$$

$$C_{\mathbf{X}_{D\setminus\{1\}}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right) = \hat{C}_{\mathbf{X}_{D\setminus\{1\}}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right) = C_{\mathbf{X}_{D\setminus\{2\}}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right) = \hat{C}_{\mathbf{X}_{D\setminus\{2\}}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right) = \left(\frac{1}{2}\right)^{3-\alpha},$$

$$C_{\mathbf{X}_{D\setminus\{3\}}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right) = \hat{C}_{\mathbf{X}_{D\setminus\{3\}}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right) = C_{\mathbf{X}_{D\setminus\{4\}}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right) = \hat{C}_{\mathbf{X}_{D\setminus\{4\}}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right) = \left(\frac{1}{2}\right)^{3-\delta}.$$

Therefore,

$$\beta(\mathbf{X}) = 2^{\delta+\alpha-2} - 2^{\alpha-3} - 2^{\delta-3}.$$

In the case of $\delta = \alpha = 0$ *the result agrees with what we expect, since in this case the margins of* $\mathbf{X}$ *are independent. The expression obtained can be related to* $\beta(X_1,X_2)$ *and* $\beta(X_3,X_4)$ *through*

$$\begin{aligned}\beta(\mathbf{X}) &= 2\times 2^{\delta+\alpha-3} - 2^{\alpha-3} - 2^{\delta-3} = \left(2^{\delta+\alpha-3} - 2^{\alpha-3}\right) + \left(2^{\delta+\alpha-3} - 2^{\delta-3}\right) \\ &= 2^{\alpha-3}\left(2^\delta - 1\right) + 2^{\delta-3}\left(2^\alpha - 1\right) \\ &= 2^{\alpha-3}\beta(X_1,X_2) + 2^{\delta-3}\beta(X_3,X_4),\end{aligned}$$

We verify that $\beta(\mathbf{X})$ *increases with* δ *and* α, *generalizing what we already knew to* $\beta(X_1,X_2)$ *and* $\beta(X_3,X_4)$. *Therefore* $\beta(\mathbf{X})$ *increases with the concordance of* $\mathbf{X}$.

Example 3.2. *Let us consider that* $\mathbf{X}$ *has a Gumbel copula*

$$C_{\mathbf{X}}(u_1, \ldots, u_d) = \exp\left\{ -\left(\sum_{i=1}^{d} (-\ln u_i)^{1/\delta} \right)^{\delta} \right\}, \tag{27}$$

with $0 < \delta \leq 1$. *For* $d = 3$ *we obtain* $\beta(\mathbf{X}) = 2^{2-2^{\delta}} - 1$, *coincident with* $\beta(X_i, X_j)$, $1 \leq i < j \leq 3$, *as expected, since in this case* $\beta(\mathbf{X}) = \beta^{\star}(\mathbf{X})$.

With simple calculations we can also conclude that

$$\beta(-X_1, X_2, X_3) = \frac{-2^{2-2^{\delta}} + 1}{3}$$

and that

$$\beta(X_1, X_2, X_3) + \beta(-X_1, X_2, X_3) = \frac{2}{2+1}\beta(X_2, X_3),$$

which corresponds to the verification in this example of a transition property that we present in the next section. Before we present the general expression of the multivariate correlation coefficient for a Gumbel distribution of dimension $d \geq 1$, *let's also calculate it specifically for* $d = 4$.

We have

$$C_{\mathbf{X}}\left(\tfrac{\mathbf{1}}{\mathbf{2}}\right) = 2^{-4^{\delta}}, \quad \hat{C}_{\mathbf{X}}\left(\tfrac{\mathbf{1}}{\mathbf{2}}\right) = -1 + 6 \times 2^{-2^{\delta}} - 4 \times 2^{-3^{\delta}} + 2^{-4^{\delta}},$$

and

$$C_{\mathbf{X}_{D\setminus\{i\}}}\left(\tfrac{\mathbf{1}}{\mathbf{2}}\right) = 2^{-3^{\delta}}, \quad \hat{C}_{\mathbf{X}_{D\setminus\{i\}}}\left(\tfrac{\mathbf{1}}{\mathbf{2}}\right) = 3 \times 2^{-2^{\delta}} - 2^{-3^{\delta}} - 2^{-1}, \quad \text{for } i = 1, 2, 3, 4.$$

Then

$$\beta(X_1, X_2, X_3, X_4) = 4 \times 2^{-4^{\delta}} - 8 \times 2^{-3^{\delta}} + 9 \times 2^{-2^{\delta}} - \frac{3}{2}.$$

These results for $d = 2, 3, 4$, *calculated directly, can also be obtained from the following general result.*

If d is even, we have

$$\beta(\mathbf{X}) = \frac{1-d}{2} + \sum_{k=1}^{d-2} \left(\binom{d-1}{k} + \binom{d}{k+1} \right) (-1)^{k+1} 2^{-(k+1)^{\delta}} + 4 \times 2^{-d^{\delta}} + (-1)^{d-1} 2^{-(d-1)^{\delta}},$$

(considering that a sum with the initial value of the counter greater than the final one is null) and if d is odd, we have

$$\beta(\mathbf{X}) = \frac{1-d}{2} + \sum_{k=1}^{d-2} \left(\binom{d-1}{k} + \binom{d}{k+1} \right) (-1)^{k+1} 2^{-(k+1)^{\delta}} - 2^{-(d-1)^{\delta}}.$$

Example 3.3. *Consider* $\mathbf{X}$ *of dimension* $d \geq 3$ *such that* $\mathbf{U} = (U, 1-U, U_3, \ldots, U_d)$. *Then*

$$\begin{aligned}
\beta(\mathbf{X}) = \quad & 2 \times (0 + 0) \\
& - \frac{1}{d}\left(C_{\mathbf{X}_{D\setminus\{1\}}}\left(\tfrac{\mathbf{1}}{\mathbf{2}}\right) + \hat{C}_{\mathbf{X}_{D\setminus\{1\}}}\left(\tfrac{\mathbf{1}}{\mathbf{2}}\right) + C_{\mathbf{X}_{D\setminus\{2\}}}\left(\tfrac{\mathbf{1}}{\mathbf{2}}\right) + \hat{C}_{\mathbf{X}_{D\setminus\{2\}}}\left(\tfrac{\mathbf{1}}{\mathbf{2}}\right) + 0 \right) \\
= \quad & - \frac{1}{d}\left(C_{\mathbf{X}_{D\setminus\{1\}}}\left(\tfrac{\mathbf{1}}{\mathbf{2}}\right) + C_{\mathbf{X}_{D\setminus\{2\}}}\left(\tfrac{\mathbf{1}}{\mathbf{2}}\right) + \hat{C}_{\mathbf{X}_{D\setminus\{1\}}}\left(\tfrac{\mathbf{1}}{\mathbf{2}}\right) + \hat{C}_{\mathbf{X}_{D\setminus\{2\}}}\left(\tfrac{\mathbf{1}}{\mathbf{2}}\right) \right) \\
= \quad & - \frac{1}{d}\left(C_{\mathbf{X}_{D\setminus\{1,2\}}}\left(\tfrac{\mathbf{1}}{\mathbf{2}}\right) + \hat{C}_{\mathbf{X}_{D\setminus\{1,2\}}}\left(\tfrac{\mathbf{1}}{\mathbf{2}}\right) \right).
\end{aligned}$$

It follows that, in this example we have $\beta(\mathbf{X}) \geq -\frac{1}{d}$ *and if, in particular* $(U_3, \ldots, U_d) = (V, \ldots, V)$, *then* $\beta(\mathbf{X}) = -\frac{1}{d}$.

4 Properties of the multivariate medial correlation coefficient

Since the coefficients $\beta_{\{i\},D\setminus\{i\}}(\mathbf{X})$, $i = 1, \ldots, d$, take values in $[-1, 1]$, the proposed coefficient takes values in the same range, being null for $C_{\mathbf{X}} = C_{\Pi}$. The maximum value is attainable when $C_{\mathbf{X}} = C_M = 1$.

The value of $\beta(\mathbf{X})$ may not increase with the concordance of $\mathbf{X}$. We can verify this with an example proposed by an anonymous referee.

Consider $\mathbf{X}$ and $\mathbf{Y}$ 4-dimensional vetors with copulas, respectively,

$$C_{\mathbf{X}}(u_1, u_2, u_3, u_4) = C_W(u_1, u_2)C_{\Pi}(u_3, u_4)$$

and

$$C_{\mathbf{Y}}(u_1, u_2, u_3, u_4) = C_W(u_1, u_2)C_M(u_3, u_4).$$

We have $\mathbf{X} \prec_c \mathbf{Y}$ and however $\beta(\mathbf{X}) = -\frac{1}{8} > -\frac{1}{4} = \beta(\mathbf{Y})$.

If $\mathbf{X} \prec_c \mathbf{Y}$ and, for each $i \in D$,

$$\begin{cases} C_{\sigma_i\mathbf{Y}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right) \leq C_{\sigma_i\mathbf{X}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right) \\ \hat{C}_{\sigma_i\mathbf{Y}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right) \leq \hat{C}_{\sigma_i\mathbf{X}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right), \end{cases} \tag{28}$$

then, from proposition 3.1, (23), we can conclude that $\beta(\mathbf{X}) \leq \beta(\mathbf{Y})$.

The verification of condition (28) together with $\mathbf{X} \prec_c \mathbf{Y}$, which can be illustrated with example 3.2, tells us that, in addition to the propensity for all margins to exceed their respective medians or all margins to be below their medians to be higher in $\mathbf{Y}$, also the propensity for each margin to disagree with the remaining, in this sense, is lower in $\mathbf{Y}$, reinforcing the relation $\mathbf{X} \prec_c \mathbf{Y}$.

When we have $\mathbf{X} \prec_c \mathbf{Y}$ and (28) we denote this type of relation by $\mathbf{X} \prec\prec \mathbf{Y}$.

The relation $\prec\prec$ is a point-wise partial ordering on the set of d-dimensional copulas that implies the concordance relation. For $d = 2$ both relations coincide. If $C_{\mathbf{Y}} = C_M$ then we have $\mathbf{X} \prec_c \mathbf{Y}$, $C_{\sigma_i\mathbf{Y}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right) = 0 \leq C_{\sigma_i\mathbf{X}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right)$ and $\hat{C}_{\sigma_i\mathbf{Y}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right) = 0 \leq \hat{C}_{\sigma_i\mathbf{X}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right)$. Therefore C_M is the maximal copula.
In particular copula classes, the relation $\prec\prec$ can induce a total order, as for example in the family of 3-dimensional copulas of example 3.2. In this class we can also see, from (27), that C_{Π} is the least element and $\prec\prec$ is a well order.
A weaker relation, although not so informative, could be considered in this work by replacing (28) with

$$\sum_{i\in D}\left(C_{\sigma_i\mathbf{Y}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right) + \hat{C}_{\sigma_i\mathbf{Y}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right)\right) \leq \sum_{i\in D}\left(C_{\sigma_i\mathbf{X}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right) + \hat{C}_{\sigma_i\mathbf{X}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right)\right).$$

The above properties on the values of the multivariate medial correlation coefficient are arranged in the following proposition.

Proposition 4.1. *The values of the multivariate medial correlation coefficient for vectors of dimension d satisfy the following properties:*
(i) If $\mathbf{X} \prec\prec \mathbf{Y}$ *then* $\beta(\mathbf{X}) \leq \beta(\mathbf{Y})$.
(ii) $\beta(\mathbf{X}) \in [-1, 1]$.
(iii) If $C_{\mathbf{X}} = C_{\Pi}$ *then* $\beta(\mathbf{X}) = 0$.
(iv) If $C_{\mathbf{X}} = C_M$ *then* $\beta(\mathbf{X}) = 1$.

Proof. Representation (23) leads to (i) and representation (21) leads to (ii).

Relations (iii) and (iv) may be obtained, for example, from (24). □

In the proposition below we present the properties of continuity, permutation invariance, duality, reflection symmetry and transition, which together with (i)-(iii) of the previous proposition and following Taylor [11], [12], justifies calling the proposed coefficient a measure for the relation $\prec\!\prec$.

Proposition 4.2. *The values of the multivariate medial correlation coefficient for vectors of dimension d satisfy the following properties:*
(i) If $\{C_{\mathbf{X}_n}\}_{n\geq 1}$ converges uniformly to $C_{\mathbf{X}}$, $n \to +\infty$, then $\lim_{n\to+\infty} \beta(\mathbf{X}_n) = \beta(\mathbf{X})$.
(ii) The value of $\beta(\mathbf{X})$ is invariant for permutations of the margins of $\mathbf{X}$.
(iii) $\beta(\mathbf{X}) = \beta(-\mathbf{X})$.
(iv) $\sum_{(\epsilon_1,\ldots,\epsilon_d)\in\{-1,1\}^d} \beta(\epsilon_1 X_1, \ldots, \epsilon_d X_d) = 0.$
(v) If $\mathbf{Y}$ is a $(d+1)$-dimensional random vector such that $C_{\mathbf{Y}}(u_1, \ldots, u_{i-1}, 1, u_{i+1}, \ldots, u_d) = C_{\mathbf{X}}(u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_d)$ then $\frac{d}{d+1}\beta(\mathbf{X}) = \beta(\mathbf{Y}) + \beta(\sigma_i\mathbf{Y})$.

Proof. The statement of (i) can be obtained, for example, from (22). From the representation (26) we can conclude (ii). The representation (23) leads to (iii) and (iv). Finally to obtain (v), let us note that, by (23), we have

$$
\begin{aligned}
&\beta(\mathbf{Y}) + \beta(\sigma_i\mathbf{Y}) \\
= \;& C_{\mathbf{Y}}\left(\tfrac{\mathbf{1}}{\mathbf{2}}\right) + C_{\sigma_i\mathbf{Y}}\left(\tfrac{\mathbf{1}}{\mathbf{2}}\right) + \hat{C}_{\mathbf{Y}}\left(\tfrac{\mathbf{1}}{\mathbf{2}}\right) + \hat{C}_{\sigma_i\mathbf{Y}}\left(\tfrac{\mathbf{1}}{\mathbf{2}}\right) \\
& -\frac{1}{d+1}\left(C_{\sigma_i\mathbf{Y}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right) + \hat{C}_{\sigma_i\mathbf{Y}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right) + C_{\sigma_i\sigma_i\mathbf{Y}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right) + \hat{C}_{\sigma_i\sigma_i\mathbf{Y}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right)\right) \\
& -\frac{1}{d+1}\sum_{j=1,j\neq i}^{d+1}\left(C_{\sigma_j\mathbf{Y}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right) + C_{\sigma_j\sigma_i\mathbf{Y}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right) + \hat{C}_{\sigma_j\mathbf{Y}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right) + \hat{C}_{\sigma_j\sigma_i\mathbf{Y}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right)\right) \\
= \;& C_{\mathbf{X}}\left(\tfrac{\mathbf{1}}{\mathbf{2}}\right) + \hat{C}_{\mathbf{X}}\left(\tfrac{\mathbf{1}}{\mathbf{2}}\right) - \frac{1}{d+1}\left(C_{\mathbf{X}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right) + \hat{C}_{\mathbf{X}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right)\right) \\
& -\frac{1}{d+1}\sum_{j=1}^{d}\left(C_{\sigma_j\mathbf{X}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right) + \hat{C}_{\sigma_j\mathbf{X}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right)\right) \\
= \;& \frac{d}{d+1}\left(C_{\mathbf{X}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right) + \hat{C}_{\mathbf{X}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right)\right) - \frac{d}{d+1}\frac{1}{d}\sum_{j=1}^{d}\left(C_{\sigma_j\mathbf{X}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right) + \hat{C}_{\sigma_j\mathbf{X}}\left(\frac{\mathbf{1}}{\mathbf{2}}\right)\right),
\end{aligned}
$$

that matches $\frac{d}{d+1}\beta(\mathbf{X})$, applying again (23). □

5 Application to real data

The multivariate medial correlation coefficient in (16) can be estimated through the bivariate coefficients in (17). Here we consider the respective empirical counterparts. This estimation procedure has already been addressed in literature (Blomqvist [1], Schmid and Schmidt [10] and references therein).

Let $(X_{1,j}, \ldots, X_{d,j})$, $j = 1, \ldots, n$, be a random sample generated from $(X_1, \ldots, X_d)$. Consider

$$
\hat{U}_{i,j} = \hat{F}_{X_i}(X_{i,j}) = \frac{1}{n+1}\sum_{l=1}^{n}\mathbb{1}_{\{X_{i,l}\leq X_{i,j}\}},\ i = 1, \ldots, d,\ j = 1, \ldots, n\,,
$$

as well as, $\hat{M}_j\left(D\setminus\{i\}\right) = \bigvee_{r\in D\setminus\{i\}} \hat{U}_{r,j}$ and $\hat{W}_j\left(D\setminus\{i\}\right) = \bigwedge_{r\in D\setminus\{i\}} \hat{U}_{r,j}$. Based on (16) we define

$$\hat{\beta} = \frac{1}{d}\sum_{i=1}^{d} \hat{\beta}_{\{i\},D\setminus\{i\}}, \tag{29}$$

where, according to (17), we take

$$\hat{\beta}_{\{i\},D\setminus\{i\}} = \frac{\hat{\overline{\beta}}\left(\hat{U}_i, \hat{M}\left(D\setminus\{i\}\right)\right) + \hat{\underline{\beta}}\left(\hat{U}_i, \hat{W}\left(D\setminus\{i\}\right)\right)}{2},$$

with

$$\begin{aligned}&\hat{\overline{\beta}}\left(\hat{U}_i, \hat{M}\left(D\setminus\{i\}\right)\right)\\ =\ & 2\left(\frac{1}{n}\sum_{j=1}^{n}\left(\mathbb{1}_{\{\hat{U}_{i,j}\le 1/2\}}\mathbb{1}_{\{\hat{M}_j(D\setminus\{i\})\le 1/2\}} + \mathbb{1}_{\{\hat{U}_{i,j}>1/2\}}\mathbb{1}_{\{\hat{M}_j(D\setminus\{i\})>1/2\}}\right)\right) - 1\end{aligned}$$

and

$$\begin{aligned}&\hat{\underline{\beta}}\left(\hat{U}_i, \hat{W}\left(D\setminus\{i\}\right)\right)\\ =\ & 2\left(\frac{1}{n}\sum_{j=1}^{n}\left(\mathbb{1}_{\{\hat{U}_{i,j}\le 1/2\}}\mathbb{1}_{\{\hat{W}_j(D\setminus\{i\})\le 1/2\}} + \mathbb{1}_{\{\hat{U}_{i,j}>1/2\}}\mathbb{1}_{\{\hat{W}_j(D\setminus\{i\})>1/2\}}\right)\right) - 1.\end{aligned}$$

We are going to apply the multivariate medial correlation coefficient estimator $\hat{\beta}$ in (29) on two datasets.

First, we consider the main GDP aggregates per capita in the European Union (EU), Germany and Portugal, available in https://ec.europa.eu/eurostat/data/database. We consider annual data from 2008 to 2019. The respective scatterplots are in Figure 1. Germany and EU seem the most correlated. The estimates of the bivariate coefficients $\beta_{\{i\},D\setminus\{i\}}$ and of the multivariate medial correlation coefficient β are in Table 1. The propensity of each country's annual GNP to agree with the remaining when compared to the propensity to disagree with them, in the sense of being all above or all below their respective medians, is estimated at 0.778. We can see that the bivariate medial correlation between Portugal and the remaining EU and Germany presents the lowest contribution to the estimated multivariate medial correlation.

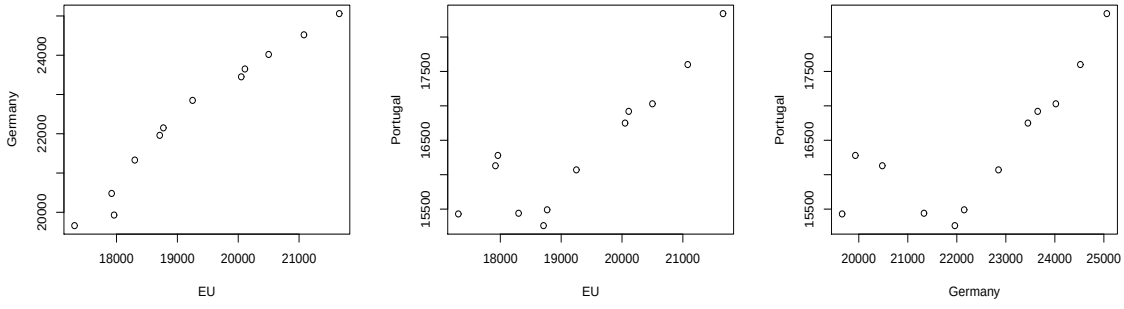

Figure 1: Annual main GDP aggregates per capita in the European Union versus Germany (left), European Union versus Portugal (center) and Germany versus Portugal (right).

Now we consider a dataset related to white variants of the Portuguese "Vinho Verde" wine, available in http://archive.ics.uci.edu/ml/datasets/Wine+Quality. See also Cortez *et al.* ([2]). Our analysis focuses on variables *residual sugar*, *total sulfur dioxide*, *density* and *alcohol*, whose respective scatterplots are plotted in Figure 2. It is visible some negative association with variable *alcohol*. The estimates of the bivariate coefficients $\beta_{\{i\},D\setminus\{i\}}$ and of the multivariate medial correlation coefficient β (Table 2) reflect this lack of concordance, with a larger negative bivariate coefficient between alcohol and the remaining variables. Indeed, the

Table 1: Estimates of the bivariate coefficients $\beta_{\{i\},D\setminus\{i\}}$ and of the multivariate medial correlation coefficient β of the annual main GDP aggregates per capita in the European Union, Germany and Portugal, from 2008 to 2019.

$\{i\}$	$D \setminus \{i\}$	$\hat{\beta}_{\{i\},D\setminus\{i\}}$	$\hat{\beta}$
{EU}	{Germany, Portugal}	0.833	
{Germany}	{EU, Portugal}	0.833	0.778
{Portugal}	{EU, Germany}	0.667	

propensity for all variables simultaneously taking values below the respective medians or all of them to exceed their respective medians in comparison with the propensity of each variable to contradict this behavior is estimated at −0.063, i.e., an almost null multivariate medial correlation coefficient.

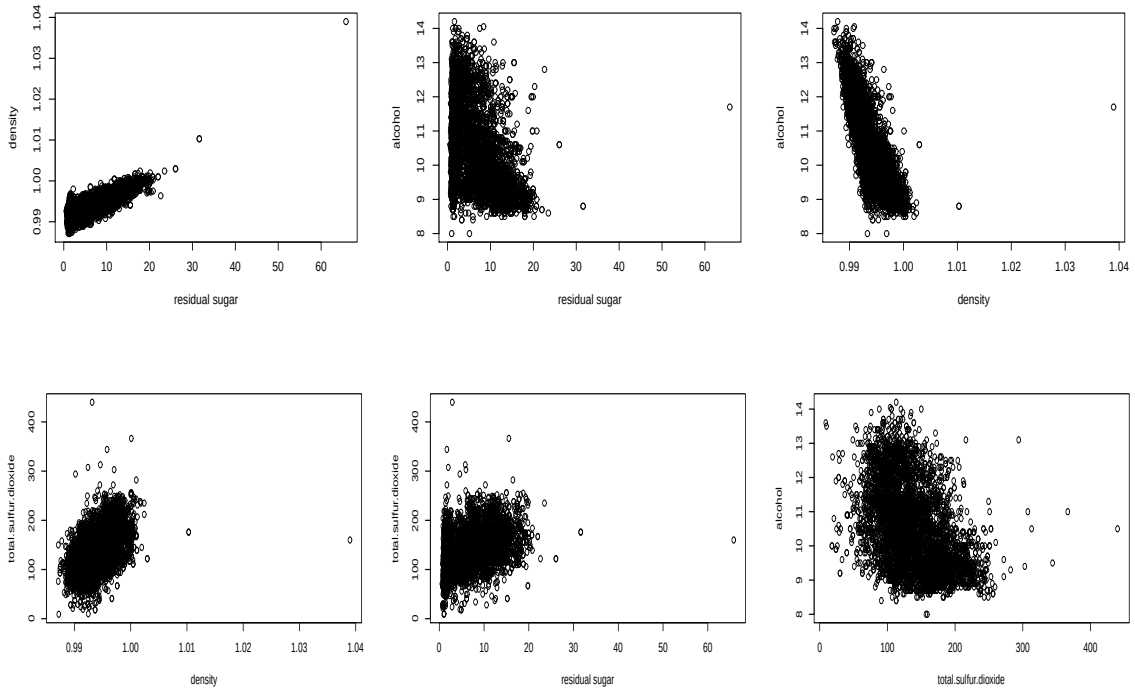

Figure 2: Scatterplots of the variables within the wine dataset: *residual sugar* versus *density* (top-left), *residual sugar* versus *alcohol* (top-center) and *density* versus *alcohol* (top-right); *density* versus *total sulfur dioxide* (bottom-left), *residual sugar* versus *total sulfur dioxide* (bottom-center) and *total sulfur dioxide* versus *alcohol* (bottom-right).

Table 2: Estimates of the bivariate coefficients $\beta_{\{i\},D\setminus\{i\}}$ and of the multivariate medial correlation coefficient β for the variables *residual sugar*, *total sulfur dioxide*, *density* and *alcohol* within the wine dataset.

$\{i\}$	$D \setminus \{i\}$	$\hat{\beta}_{\{i\},D\setminus\{i\}}$	$\hat{\beta}$
{residual sugar}	{total sulfur dioxide, density, alcohol}	0.088	-0.063
{total sulfur dioxide}	{residual sugar, density, alcohol}	0.027	
{density}	{residual sugar, total sulfur dioxide, alcohol}	0.046	
{alcohol}	{residual sugar, total sulfur dioxide, density}	-0.415	

6 Conclusion

The multivariate medial correlation coefficient that we propose extends the probabilistic interpretation and properties of the Blomqvist β coefficient, it is calculable from the copula, incorporates the dependence between each margin of the vector and the vector of the remaining margins and is a measure of a strong mode of multivariate concordance.

The estimation is addressed based on bivariate inferential methodology existing in literature and we illustrate its application using real data.

The adopted approach envisages the possibility of considering other functions of bivariate coefficients involving extremes of subvectors of $\mathbf{X}$, as well as the possibility of adapting the method to generalize other coefficients of bivariate dependence.

Acknowledgements: The authors thank the reviewers and the associated editor for the very important and valuable comments and corrections that contributed to the improvement of this work.
The first author was partially supported by the research unit Centre of Mathematics and Applications of University of Beira Interior UIDB/00212/2020 - FCT (Fundação para a Ciência e a Tecnologia). The second author was financed by Portuguese Funds through FCT - Fundação para a Ciência e a Tecnologia within the Projects UIDB/00013/2020 and UIDP/00013/2020 of Centre of Mathematics of the University of Minho, UIDB/00006/2020 of Centre of Statistics and its Applications of University of Lisbon and PTDC/MAT-STA/28243/2017.

References

[1] Blomqvist, N. (1950). On a measure of dependence between two random variables. *Ann. Math. Statist. 21*(4), 593–600.
[2] Cortez, P., A. Cerdeira, F. Almeida, T. Matos and J.Reis (2009). Modeling wine preferences by data mining from physicochemical properties. *Decis. Support Syst. 47*(4), 547–553.
[3] Joe, H. (1997). *Multivariate Models and Dependence Concepts*. Chapman and Hall, London.
[4] Joe, H. (1990). Multivariate Concordance. *J. Multivariate Anal. 35*(1), 12–30.
[5] Joe, H. (2015). *Dependence Modeling with Copulas*. CRC Press, Boca Raton FL.
[6] Lebedev, A.V. (2019). On the Interrelation between dependence coefficients of bivariate extreme value copulas. *Markov Process. Relat. 25*(4), 639–648.
[7] Nelsen, R.B. (2002). Concordance and Copulas: A Survey. In C.M. Cuadras, Fortiana J., Rodriguez-Lallena J.A. (Eds.) *Distributions With Given Marginals and Statistical Modelling*, pp 169–177. Springer, Dordrecht.
[8] Nelsen, R.B. (2006). *An Introduction to Copulas*. Second edition. Springer, New York.
[9] Scarsini, M. (1984). On Measures of Concordance. *Stochastica 8*(3), 201–218.
[10] Schmid, F. and R. Schmidt (2007). Nonparametric inference on multivariate versions of Blomqvist's beta and related measures of tail dependence. *Metrika 66*(3), 323–354.
[11] Taylor, M. D. (2007). Multivariate measures of concordance. *Ann. Inst. Statist. Math. 59*(4), 789–806.
[12] Taylor, M. D. (2016). Multivariate measures of concordance for copulas and their marginals. *Depend. Model. 4*, 224–236
[13] Úbeda-Flores, M. (2005) Multivariate versions of Blomqvist's beta and Spearman's footrule. *Ann. Inst. Statist. Math. 57*(4), 781–788.