Electronic Journal of Statistics
Vol. 10 (2016) 1064–1097
ISSN: 1935-7524
DOI: 10.1214/16-EJS1134

A general framework for testing homogeneity hypotheses about copulas

Jean-François Quessy

Université du Québec à Trois-Rivières
e-mail: jean-francois.quessy@uqtr.ca

Abstract: The dependence structure in a d-variate continuous random vector $\mathbf{X}$ is characterized by its unique copula. Starting from the fact that many copulas can be extracted from the *global* d-dimensional copula of $\mathbf{X}$, a very general framework is proposed here for testing that a given collection of induced p-dimensional copulas from a multivariate distribution are identical. Many hypotheses of interest in copula modeling fall into this category, including bivariate symmetry (diagonal, radial, joint), exchangeability, as well as various types of equality of copulas. Here, a broad class of test statistics is defined around a matrix representation of the null hypothesis and quadratic functionals including Cramér–von Mises and characteristic function mappings. Since the null hypotheses to be tested are composite by nature, the computation of P-values is achieved using multiplier bootstrap versions of the test statistics. The sample properties of the method are investigated when testing for several types of bivariate symmetry, exchangeability, equality of non-overlapping and overlapping copulas and equality of all bivariate copulas. The general conclusion is that the tests are good at keeping their nominal level and are powerful against a wide variety of alternatives, showing the relevance and reliability of the methodology for the modeling of multivariate datasets with the help of copulas.

Keywords and phrases: Copula, characteristic function, empirical copula process, multiplier bootstrap, quadratic functionals, symmetry hypotheses.

Received October 2014.

1. Introduction

Take a random vector $\mathbf{X} = (X_1, \ldots, X_d)$, $d \geq 2$, with joint distribution function $H(\mathbf{x}) = \mathrm{P}(\mathbf{X} \leq \mathbf{x})$, where $\mathbf{x} = (x_1, \ldots, x_d)$. If the marginal distributions $F_1, \ldots, F_d$ of $\mathbf{X}$ are continuous, then a unique copula $C : [0,1]^d \to [0,1]$ exists such that $H(\mathbf{x}) = C\{F_1(x_1), \ldots, F_d(x_d)\}$ for all $\mathbf{x} \in \mathbb{R}^d$; more details on the theoretical aspects of copulas can be found in the monographs by [13], [7] and [18]. For copula modeling, the testing of *shape hypotheses* has retained a lot of attention recently. These hypotheses only specify a general form for the underlying dependence structure C of $\mathbf{X}$. Among many contributions in this area, one can mention tests for extreme-value dependence ([15], [19]) and tests of Archimedeanity ([12], [3]). Because of the composite nature of the hypotheses of interest, the statistical procedures in these works typically exploit a special feature shared by all the members under the null, *e.g.*, the max-stability property of extreme-value copulas.

In this paper, a special attention is given to a very general class of shape hypotheses, namely those that can be written in terms of equality between copulas that one can extract from C. As will be seen, the class includes several hypotheses including many notions of symmetry ([11], [20], [10]) and the equality of two copulas ([21]). Formally, let $\mathcal{G}$ be the set of functions $g : [0,1] \to [0,1]$ such that $g = I$ or $g = 1 - I$, where I is the identity function. Given $g_1, \ldots, g_p \in \mathcal{G}$, let $\mathbf{g}(\mathbf{u}) = (g_1(u_1), \ldots, g_p(u_p))$, where $\mathbf{u} = (u_1, \ldots, u_p)$. Taking a p-tuple $A = (A_1, \ldots, A_p)$ of distinct elements of $\{1, \ldots, d\}$, consider the random subvector $\mathbf{U}_A = (U_{A_1}, \ldots, U_{A_p})$ of $\mathbf{U} = (U_1, \ldots, U_d) \sim C$ whose joint distribution

$$C_{A,\mathbf{g}}(\mathbf{v}) = \mathrm{P}\left\{\mathbf{g}\left(\mathbf{U}_A\right) \leq \mathbf{v}\right\}, \quad \mathbf{v} \in [0,1]^p, \tag{1}$$

is a p-variate copula extracted from C. The aim of this paper is to develop a general framework for testing hypotheses of the form

$$\mathcal{H}_0 : C_{A^{(1)},\mathbf{g}^{(1)}} = \cdots = C_{A^{(K)},\mathbf{g}^{(K)}}, \tag{2}$$

where $A^{(1)}, \ldots, A^{(K)}$ are p-tuples of distinct elements of $\{1, \ldots, d\}$ and $\mathbf{g}^{(1)}$, …, $\mathbf{g}^{(K)}$ are such that for each $k \in \{1, \ldots, K\}$, $\mathbf{g}^{(k)} = (g_1^{(k)}, \ldots, g_p^{(k)})$ with $g_1^{(k)}, \ldots, g_p^{(k)} \in \mathcal{G}$. The null hypothesis $\mathcal{H}_0$ corresponds to the equality in distribution of $\mathbf{g}^{(1)}(\mathbf{U}_{A^{(1)}})$, …, $\mathbf{g}^{(K)}(\mathbf{U}_{A^{(K)}})$. Although the methodology will be developed under a general setting, one should keep in mind the following special cases that are of a particular interest for copula modeling:

<u>Bivariate symmetry.</u> Let $d = p = 2$. If $A^{(1)} = (1,2)$, $A^{(2)} = (2,1)$ and $\mathbf{g}^{(1)} = \mathbf{g}^{(2)} = (I, I)$, then $\mathcal{H}_0$ is the hypothesis of diagonal symmetry $C(v_1, v_2) = C(v_2, v_1)$; if $A^{(1)} = A^{(2)} = (1,2)$, $\mathbf{g}^{(1)} = (I, I)$ and $\mathbf{g}^{(2)} = (1-I, 1-I)$, then $\mathcal{H}_0$ is the hypothesis of radial symmetry $C(v_1, v_2) = v_1 + v_2 - 1 + C(1-v_1, 1-v_2)$; if $A^{(1)} = A^{(2)} = A^{(3)} = (1,2)$, $\mathbf{g}^{(1)} = (I, I)$, $\mathbf{g}^{(2)} = (1-I, I)$ and $\mathbf{g}^{(3)} = (I, 1-I)$, then $\mathcal{H}_0$ is the hypothesis of joint symmetry.

<u>Exchangeability.</u> For $p = d > 2$, let $A^{(1)}, \ldots, A^{(K)}$ be the $K = d!$ sets of all possible permutations of $\{1, \ldots, d\}$ and put $\mathbf{g}^{(1)} = \cdots = \mathbf{g}^{(K)} = (I, \ldots, I)$; then $\mathcal{H}_0$ corresponds to the hypothesis of exchangeability.

<u>Equality of copulas.</u> For $2 \leq p < d$, consider K possibly overlapping p-tuples $A^{(1)}, \ldots, A^{(K)}$ of distinct elements of $\{1, \ldots, d\}$ and let $\mathbf{g}^{(1)} = \cdots = \mathbf{g}^{(K)} = (I, \ldots, I)$; then $\mathcal{H}_0$ is the hypothesis of equality of the p-variate copulas associated to $A^{(1)}, \ldots, A^{(K)}$. In particular, when $p = 2$ and if for $K = d(d-1)/2$, $A^{(1)}, \ldots, A^{(K)}$ are the sets of all possible pairs extracted from $\{1, \ldots, d\}$, then $\mathcal{H}_0$ is the hypothesis of equality of all the bivariate copulas.

In this article, a general statistical framework based on quadratic functionals is proposed in order to test any null hypothesis that can be written in the form $\mathcal{H}_0$. The broad class of nonparametric test statistics hence defined includes Cramér–von Mises, characteristic function and diagonal type statistics as special cases.

As will be seen, the fact that the test statistic are of a quadratic-type provides simple and easy-to-implement formulas; most importantly, it yields consistent tests in many situations of interest. Moreover, the test statistics recently considered by [16] appears as a special case of the general methodology developed here when the hypotheses to be tested are of the form (2).

The paper is organized as follows. In Section 2, the general class of quadratic test statistics is introduced and their asymptotic behavior is investigated; this is closely related to a very general result on the large-sample behavior of vectors of empirical copula processes that is of an independent interest. A method for the computation of P-values based on the multiplier bootstrap method is proposed and validated in Section 3. Many quadratic functionals on which the tests are based are described in Section 4. In Section 5, the sample properties of the tests are thoroughly investigated with the help of Monte–Carlo simulations. Concluding remarks are given in Section 6. All proofs and computational formulas are relegated to three appendices.

2. A general class of test statistics

2.1. Characterization of the null hypothesis

Starting from the null hypothesis $\mathcal{H}_0$ described in equation (2), consider the K-dimensional vector of p-variate copulas

$$\mathbf{C} = \left(C_{A^{(1)},\mathbf{g}^{(1)}}, \ldots, C_{A^{(K)},\mathbf{g}^{(K)}}\right)^\top. \tag{3}$$

The null hypothesis $\mathcal{H}_0$ entails that there exists a copula $D : [0,1]^p \to [0,1]$ such that $C_{A^{(k)},\mathbf{g}^{(k)}} = D$ for all $k \in \{1, \ldots, K\}$. This can be equivalently written

$$\mathbf{C}(\mathbf{v}) = D(\mathbf{v})\,\mathbf{1}_K \quad \text{for all } \mathbf{v} \in [0,1]^p, \tag{4}$$

where here and in the sequel, $\mathbf{1}_K = (1, \ldots, 1)^\top \in \mathbb{R}^K$. The alternative hypothesis specifies that there exists at least one pair $k \neq k' \in \{1, \ldots, K\}$ such that $C_{A^{(k)},\mathbf{g}^{(k)}} \neq C_{A^{(k')},\mathbf{g}^{(k')}}$. Next, consider a *combination* matrix $\mathcal{M} \in \mathbb{R}^{q\times K}$ such that for $\mathbf{z} \in \mathbb{R}^K$, one has $\mathcal{M}\,\mathbf{z} = \mathbf{0}_q$ if and only if $\mathbf{z} = r\,\mathbf{1}_K$ for some real number $r \neq 0$ and $\mathbf{0}_q = (0, \ldots, 0)^\top \in \mathbb{R}^q$; $\mathcal{M}$ can be taken as being of full rank without loss of generality. It follows from (4) that $\mathcal{M}\mathbf{C}(\mathbf{v}) = \mathbf{0}_q$ for all $\mathbf{v} \in [0,1]^p$ if and only if $\mathcal{H}_0$ is true. This characterization of the null hypothesis is at the heart of the upcoming developments.

2.2. Quadratic-type test statistics

In this subsection, a general class of test statistics will be built around quadratic functionals of $\mathcal{M}\mathbf{C}$. To this end, define for any $a, b \in \mathbb{N}$ the set $\mathcal{S}_{ab}$ of uniformly bounded functions $S : [0,1]^p \to \mathbb{R}^{a\times b}$. Then, let $\mathcal{F} : \mathcal{S}_{ab} \to \mathbb{R}^{b\times b}$ be a functional that is well defined for every $a, b \in \mathbb{N}$ and that is quadratic in the sense that for any $R \in \mathbb{R}^{b\times 1}$ and $S \in \mathcal{S}_{ab}$,

$$R^\top \mathcal{F}(S)\, R = \mathcal{F}(SR). \tag{5}$$

Property (5) entails $\mathcal{F}(\mathbf{0}_q) = 0$ since for any $r \in \mathbb{R}$, $\mathcal{F}(\mathbf{0}_q) = \mathcal{F}(\mathbf{0}_q\, r) = r^2 \mathcal{F}(\mathbf{0}_q)$. A measure of discrepancy from $\mathcal{H}_0$ can then be based on a quadratic functional $\mathcal{F}$ and a combination matrix $\mathcal{M} \in \mathbb{R}^{q\times K}$ via $T_{\mathcal{F},\mathcal{M}} : \mathcal{S}_{K1} \to \mathbb{R}$ such that $T_{\mathcal{F},\mathcal{M}}(S) = \mathcal{F}(\mathcal{M}S)$. The fact that $\mathcal{M}\mathbf{C} = \mathbf{0}_q$ if and only if $\mathcal{H}_0$ is true entails that under the null hypothesis,

$$T_{\mathcal{F},\mathcal{M}}(\mathbf{C}) = \mathcal{F}\left(\mathcal{M}\mathbf{C}\right) = \mathcal{F}\left(\mathbf{0}_q\right) = 0.$$

The null and alternative hypotheses may then be reformulated as

$$\mathcal{H}_0 : T_{\mathcal{F},\mathcal{M}}(\mathbf{C}) = 0 \quad \text{and} \quad \mathcal{H}_1 : T_{\mathcal{F},\mathcal{M}}(\mathbf{C}) > 0. \tag{6}$$

Now let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ be i.i.d. from a d-variate population with joint distribution H, continuous marginal distributions $F_1, \ldots, F_d$, and unique copula C. A nonparametric estimation of C first investigated by [22] is given by

$$C_n(\mathbf{u}) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}\left(\widehat{\mathbf{U}}_i \le \mathbf{u}\right), \quad \mathbf{u} = (u_1, \ldots, u_d) \in [0,1]^d,$$

where $\widehat{\mathbf{U}}_i = (F_{n1}(X_{i1}), \ldots, F_{nd}(X_{id}))$ and $F_{n1}, \ldots, F_{nd}$ are the univariate empirical distribution functions. This estimator is asymptotically equivalent to the empirical copula as defined, *e.g.*, by [26] and [25], and is employed here for computational convenience. In the sequel, C_n will be referred to as the empirical copula without confusion. A natural empirical version of the copula $C_{A,\mathbf{g}}$ defined in equation (1) is then given by

$$\widehat{C}_{A,\mathbf{g}}(\mathbf{v}) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}\left\{\mathbf{g}\left(\widehat{\mathbf{U}}_{iA}\right) \le \mathbf{v}\right\},$$

where for each $i \in \{1, \ldots, n\}$, $\widehat{\mathbf{U}}_{iA} = (\widehat{U}_{iA_1}, \ldots, \widehat{U}_{iA_p})$. An estimator of the vector of copulas $\mathbf{C}$ defined in (3) is therefore

$$\widehat{\mathbf{C}} = \left(\widehat{C}_{A^{(1)},\mathbf{g}^{(1)}}, \ldots, \widehat{C}_{A^{(K)},\mathbf{g}^{(K)}}\right)^\top.$$

In view of (6), a test would consist in rejecting $\mathcal{H}_0$ for large values of $T_{\mathcal{F},\mathcal{M}}(\widehat{\mathbf{C}})$. This statistic can be computed easily upon noting that $\widehat{\mathbf{C}} = \widehat{L}\,\mathbf{1}_n/n$, where the entries of $\widehat{L} \in \mathcal{S}_{Kn}$ are

$$\widehat{L}_{ki}(\mathbf{v}) = \mathbb{I}\left\{\mathbf{g}^{(k)}\left(\widehat{\mathbf{U}}_{iA^{(k)}}\right) \le \mathbf{v}\right\}.$$

The quadratic nature of $\mathcal{F}$ described in equation (5) entails

$$T_{\mathcal{F},\mathcal{M}}(\widehat{\mathbf{C}}) = \mathcal{F}\left(\mathcal{M}\widehat{L}\,\frac{\mathbf{1}_n}{n}\right) = \frac{1}{n^2}\,\mathbf{1}_n^\top\,\Lambda\,\mathbf{1}_n,$$

where $\Lambda = \mathcal{F}(\mathcal{M}\widehat{L}) \in \mathbb{R}^{n\times n}$. Simple and explicit formulas for computing the entries of Λ are given in Appendix B for the quadratic functionals that are described in Section 4.

2.3. Asymptotic behavior

The weak convergence of the empirical copula process $\mathbb{C}_n = \sqrt{n}(C_n - C)$ was obtained by [25]. Specifically, assuming that for each $\ell \in \{1,\ldots,d\}$, the partial derivative $C^{[\ell]}(\mathbf{u}) = \partial C(\mathbf{u})/\partial u_\ell$ exists and is continuous on the set $\{\mathbf{u} \in [0,1]^d : 0 < u_\ell < 1\}$, then $\mathbb{C}_n$ converges weakly in the space $\ell^\infty([0,1]^d)$ of bounded functions on $[0,1]^d$ to a process of the form

$$\mathbb{C}(\mathbf{u}) = \mathbb{B}(\mathbf{u}) - \sum_{\ell=1}^{d} C^{[\ell]}(\mathbf{u})\,\mathbb{B}(\mathbf{1}_{\ell-1}, u_\ell, \mathbf{1}_{d-\ell}), \tag{7}$$

where $\mathbb{B}$ is a tight centered Gaussian process such that for any $\mathbf{u}, \mathbf{u}' \in [0,1]^d$,

$$\mathrm{E}\{\mathbb{B}(\mathbf{u})\,\mathbb{B}(\mathbf{u}')\} = C(\mathbf{u} \wedge \mathbf{u}') - C(\mathbf{u})\,C(\mathbf{u}').$$

Here and in the sequel, weak convergence is understood in the meaning given by [26]. Note that $C^{[1]},\ldots,C^{[d]}$ must be defined properly on the frontier of $[0,1]^d$ in order that $\mathbb{C}$ exists and has continuous paths everywhere on $[0,1]^d$. To this end, one defines, for $\ell \in \{1,\ldots,d\}$,

$$C^{[\ell]}(\mathbf{u}) = \begin{cases} \displaystyle\limsup_{h\downarrow 0} \frac{C(\mathbf{u} + h\,\mathbf{e}_\ell)}{h}, & \text{if } u_\ell = 0; \\[2ex] \displaystyle\limsup_{h\downarrow 0} \frac{C(\mathbf{u}) - C(\mathbf{u} - h\,\mathbf{e}_\ell)}{h}, & \text{if } u_\ell = 1, \end{cases}$$

where $\mathbf{e}_1,\ldots,\mathbf{e}_d$ are the coordinate vectors in $\mathbb{R}^d$.

Proposition 2 establishes the asymptotic behavior under the null hypothesis of the K-dimensional vector of empirical copulas $\widehat{\mathbf{C}}$. Before stating it, a first step is to obtain the large-sample behavior of the process $\widehat{\mathbb{C}}_{A,\mathbf{g}} = \sqrt{n}(\widehat{C}_{A,\mathbf{g}} - C_{A,\mathbf{g}})$ for arbitrary A and $\mathbf{g} \in \mathcal{G}$.

Proposition 1. *Suppose that for each $j \in \{1,\ldots,p\}$, the partial derivative $C^{[j]}_{A,\mathbf{g}}(\mathbf{v}) = \partial C_{A,\mathbf{g}}(\mathbf{v})/\partial v_j$ exists and is continuous on the set $\{\mathbf{v} \in [0,1]^p : 0 < v_j < 1\}$. Then the empirical process $\widehat{\mathbb{C}}_{A,\mathbf{g}} = \sqrt{n}(\widehat{C}_{A,\mathbf{g}} - C_{A,\mathbf{g}})$ converges weakly in the space $\ell^\infty([0,1]^p)$ to a limit process of the form*

$$\mathbb{C}_{A,\mathbf{g}}(\mathbf{v}) = \mathbb{B}_{A,\mathbf{g}}(\mathbf{v}) - \sum_{j=1}^{p} C^{[j]}_{A,\mathbf{g}}(\mathbf{v})\,\mathbb{B}_{A,\mathbf{g}}(\mathbf{1}_{j-1}, v_j, \mathbf{1}_{p-j}),$$

where $\mathbb{B}_{A,\mathbf{g}}$ is a $C_{A,\mathbf{g}}$-Brownian bridge, i.e. a tight centered Gaussian process on $[0,1]^p$ such that $\mathrm{E}\{\mathbb{B}_{A,\mathbf{g}}(\mathbf{v})\,\mathbb{B}_{A,\mathbf{g}}(\mathbf{v}')\} = C_{A,\mathbf{g}}(\mathbf{v} \wedge \mathbf{v}') - C_{A,\mathbf{g}}(\mathbf{v})\,C_{A,\mathbf{g}}(\mathbf{v}')$.

The next proposition is the main theoretical result of this section. Before stating it, define for each $\mathbf{v}, \mathbf{v}' \in [0,1]^p$ the matrix $\gamma(\mathbf{v}, \mathbf{v}') \in \mathbb{R}^{K \times K}$ whose entries are

$$\gamma_{kk'}(\mathbf{v}, \mathbf{v}') = \mathrm{P}\left\{ \mathbf{g}^{(k)}(\mathbf{U}_{A^{(k)}}) \le \mathbf{v}, \mathbf{g}^{(k')}(\mathbf{U}_{A^{(k')}}) \le \mathbf{v}' \right\}.$$

Proposition 2. *Suppose $\mathbf{C} = D\,\mathbf{1}_K$, where D is a p-variate copula such that for each $j \in \{1, \ldots, p\}$, the partial derivative $D^{[j]}(\mathbf{v}) = \partial D(\mathbf{v})/\partial v_j$ exists and is continuous on the set $\{\mathbf{v} \in [0,1]^p : 0 < v_j < 1\}$. Then the vector of empirical processes $\widehat{\mathbb{V}} = \sqrt{n}(\widehat{\mathbf{C}} - D\,\mathbf{1}_K)$ converges weakly in the space $(\ell^\infty([0,1]^p))^{\otimes K}$ to*

$$\mathbb{V}(\mathbf{v}) = \mathbb{W}(\mathbf{v}) - \sum_{j=1}^{p} D^{[j]}(\mathbf{v})\,\mathbb{W}(\mathbf{1}_{j-1}, v_j, \mathbf{1}_{p-j}),$$

where $\mathbb{W}$ is a K-dimensional vector of centered Gaussian processes on $[0,1]^p$ such that $\mathrm{E}\left\{\mathbb{W}(\mathbf{v})\,\mathbb{W}(\mathbf{v}')^\top\right\} = \gamma(\mathbf{v}, \mathbf{v}') - D(\mathbf{v})\,D(\mathbf{v}')\,\mathbf{1}_K \mathbf{1}_K^\top$.

An immediate consequence of Proposition 2 is the characterization of the limit of the test statistic $T_{\mathcal{F},\mathcal{M}}(\widehat{\mathbf{C}})$ under the null hypothesis.

Corollary 1. *Let $\mathcal{M} \in \mathbb{R}^{q \times K}$ be a combination matrix such that $\mathcal{M}\,\mathbf{z} = \mathbf{0}_q$ if and only if $\mathbf{z} = r\,\mathbf{1}_K$, $r \in \mathbb{R} \setminus \{0\}$. Under the conditions of Proposition 2, $n\,T_{\mathcal{F},\mathcal{M}}(\widehat{\mathbf{C}})$ converges in distribution to a random variable having representation $\mathcal{F}(\mathbb{V}')$, where*

$$\mathbb{V}'(\mathbf{v}) = \mathbb{W}'(\mathbf{v}) - \sum_{j=1}^{p} D^{[j]}(\mathbf{v})\,\mathbb{W}'(\mathbf{1}_{j-1}, v_j, \mathbf{1}_{p-j})$$

and $\mathbb{W}'$ is a q-dimensional vector of centered Gaussian processes on $[0,1]^p$ with $\mathrm{E}\left\{\mathbb{W}'(\mathbf{v})\,\mathbb{W}'(\mathbf{v}')^\top\right\} = \mathcal{M}\,\gamma(\mathbf{v}, \mathbf{v}')\,\mathcal{M}^\top$.

The distribution of $\mathcal{F}(\mathbb{V}')$ in Corollary 1 has not a simple form in general. From the fact that $\mathcal{F}$ is quadratic in the meaning given in (5), one can conjecture that this limit distribution admits a representation in terms of a weighted sum of independent chi-squared random variables. Showing such a result is however out of the scope of this work.

As a complement, the asymptotic distribution of $T_{\mathcal{F},\mathcal{M}}(\widehat{\mathbf{C}})$ is now established under alternatives to $\mathcal{H}_0$, *i.e.* when $\mathbf{C}$ is such that $\mathcal{M}\mathbf{C} \neq \mathbf{0}_q$.

Proposition 3. *Suppose $\mathcal{F}$ is Hadamard differentiable with derivative at $S \in \mathcal{S}_{q1}$ given by $\mathcal{F}'_S$. If for each $\ell \in \{1, \ldots, d\}$, the partial derivative $C^{[\ell]}$ exists and is continuous on the set $\{\mathbf{u} \in [0,1]^d : 0 < u_\ell < 1\}$ and $\mathcal{M} \in \mathbb{R}^{q \times K}$ is such that $\mathcal{M}\mathbf{C} \neq \mathbf{0}_q$, then $\sqrt{n}\{T_{\mathcal{F},\mathcal{M}}(\widehat{\mathbf{C}}) - T_{\mathcal{F},\mathcal{M}}(\mathbf{C})\}$ is asymptotically normal with mean zero and variance $\sigma^2_{\mathcal{F},\mathcal{M}} = \mathrm{var}\,\{\mathcal{F}'_{\mathcal{M}\mathbf{C}}(\mathcal{M}\mathbb{V}^\star)\}$, where $\mathbb{V}^\star$ is the weak limit of $\sqrt{n}(\widehat{\mathbf{C}} - \mathbf{C})$.*

3. Multiplier bootstrap for the computation of P-values

3.1. The multiplier CLT for copula processes

For $\mathcal{H} \in \mathbb{N}$, consider the independent vectors $(\xi_1^{(1)}, \ldots, \xi_n^{(1)}), \ldots, (\xi_1^{(\mathcal{H})}, \ldots, \xi_n^{(\mathcal{H})})$ of independent and positive *multiplier* random variables from a probability distribution having mean one, variance one, and such that $\int_0^\infty \{\mathrm{P}(\xi_i^{(h)} > x)\}^{1/2} \mathrm{d}x < \infty$. For each $h \in \{1, \ldots, \mathcal{H}\}$, define a *multiplier* version of the process $\mathbb{B}$ appearing in equation (7) by

$$\widehat{\mathbb{B}}^{(h)}(\mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \gamma_i^{(h)} \, \mathbb{I}\left(\widehat{\mathbf{U}}_i \leq \mathbf{u}\right),$$

where $\gamma_i^{(h)} = (\xi_i^{(h)} - \bar{\xi}^{(h)})/\bar{\xi}^{(h)}$ and $\bar{\xi}^{(h)} = (\xi_1^{(h)} + \cdots + \xi_n^{(h)})/n$. This way of defining the multiplier random variables is known as the weighted bootstrap, which was introduced by [6]. The particular case of exponential multipliers is called bayesian bootstrap. The limiting representation of $\mathbb{C}_n$ in (7) suggests

$$\widehat{\mathbb{C}}^{(h)}(\mathbf{u}) = \widehat{\mathbb{B}}^{(h)}(\mathbf{u}) - \sum_{\ell=1}^{d} \widehat{C}^{[\ell]}(\mathbf{u}) \, \widehat{\mathbb{B}}^{(h)}(\mathbf{1}_{\ell-1}, u_\ell, \mathbf{1}_{d-\ell}) \tag{8}$$

as the multiplier empirical copula process, where for each $\ell \in \{1, \ldots, d\}$, $\widehat{C}^{[\ell]}$ is an estimator of the partial derivative $C^{[\ell]}$ such that for any $\epsilon \in (0, 1/2)$,

$$\sup_{\substack{\mathbf{u} \in [0,1]^d: \\ u_\ell \in [\epsilon, 1-\epsilon]}} \left| \widehat{C}^{[\ell]}(\mathbf{u}) - C^{[\ell]}(\mathbf{u}) \right| \xrightarrow{\mathrm{P}} 0.$$

One deduces from [25] that the vector of processes $(\mathbb{C}_n, \widehat{\mathbb{C}}^{(1)}, \ldots, \widehat{\mathbb{C}}^{(\mathcal{H})})$ converges weakly in the space $(\ell^\infty([0,1]^d))^{\otimes(1+\mathcal{H})}$ to $(\mathbb{C}, \mathbb{C}^{(1)}, \ldots, \mathbb{C}^{(\mathcal{H})})$, where $\mathbb{C}^{(1)}, \ldots, \mathbb{C}^{(\mathcal{H})}$ are independent copies of $\mathbb{C}$. This multiplier bootstrap method for empirical copula processes will now be adapted in order to mimic the asymptotic behavior of $\widehat{\mathbb{V}}$ under the null hypothesis.

3.2. Multiplier versions of the test statistics

The fact that the distribution of $n \, T_{\mathcal{F},\mathcal{M}}(\widehat{\mathbf{C}})$ under $\mathcal{H}_0$, both in finite samples and asymptotically, depends on the unknown value of the p-variate copula D such that $\mathbf{C} = D \, \mathbf{1}_K$ prevents from finding explicitly, or even numerically, a value Q_α such that $\mathrm{P}(T_{\mathcal{F},\mathcal{M}}(\mathbb{V}) > Q_\alpha) = \alpha$. A solution is to rely on bootstrap replicates of the limit $\mathcal{F}(\mathbb{V}')$ of $n \, T_{\mathcal{F},\mathcal{M}}(\widehat{\mathbf{C}})$ obtained from the multiplier method. To this end, for each $h \in \{1, \ldots, \mathcal{H}\}$, let

$$\widehat{\mathbb{W}}^{(h)} = \left(\widehat{\mathbb{B}}^{(h)}_{A^{(1)}, \mathbf{g}^{(1)}}, \ldots, \widehat{\mathbb{B}}^{(h)}_{A^{(K)}, \mathbf{g}^{(K)}}\right)^\top \in \mathcal{S}_{K1},$$

where for each $k \in \{1, \ldots, K\}$,

$$\widehat{\mathbb{B}}^{(h)}_{A^{(k)},\mathbf{g}^{(K)}}(\mathbf{v}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \gamma_i^{(h)} \, \mathbb{I}\left\{\mathbf{g}^{(k)}\left(\widehat{\mathbf{U}}_{iA^{(k)}}\right) \leq \mathbf{v}\right\}$$

are the multiplier versions of the limit process $\mathbb{B}_{A^{(k)},\mathbf{g}^{(k)}}$ appearing in the asymptotic representation of $\widehat{\mathbb{V}}$. Then, for uniformly bounded $\widehat{D}^{[1]}, \ldots, \widehat{D}^{[p]}$ that satisfy

$$\sup_{\substack{\mathbf{v} \in [0,1]^p: \\ v_j \in [\epsilon, 1-\epsilon]}} \left| \widehat{D}^{[j]}(\mathbf{v}) - D^{[j]}(\mathbf{v}) \right| \xrightarrow{\mathrm{P}} 0 \tag{9}$$

for any $\epsilon \in (0, 1/2)$, define

$$\widehat{\mathbb{V}}^{(h)}(\mathbf{v}) = \widehat{\mathbb{W}}^{(h)}(\mathbf{v}) - \sum_{j=1}^{p} \widehat{D}^{[j]}(\mathbf{v}) \, \widehat{\mathbb{W}}^{(h)}(\mathbf{1}_{j-1}, v_j, \mathbf{1}_{p-j}). \tag{10}$$

Estimators of $\widehat{D}^{[1]}, \ldots, \widehat{D}^{[p]}$ based on finite-differences are described in Section 5.

Proposition 4. *Under the conditions of Proposition 2, $(\widehat{\mathbb{V}}, \widehat{\mathbb{V}}^{(1)}, \ldots, \widehat{\mathbb{V}}^{(\mathcal{H})})$ converges weakly in the space $(\ell^\infty([0,1]^p))^{\otimes(1+\mathcal{H})}$ to $(\mathbb{V}, \mathbb{V}^{(1)}, \ldots, \mathbb{V}^{(\mathcal{H})})$, where $\mathbb{V}^{(1)}$, $\ldots$, $\mathbb{V}^{(\mathcal{H})}$ are independent copies of $\mathbb{V}$.*

Corollary 1 states that $n\, T_{\mathcal{F},\mathcal{M}}(\widehat{\mathbf{C}})$ converges in distribution to $\mathcal{F}(\mathbb{V}')$ under $\mathcal{H}_0$, where $\mathbb{V}' = \mathcal{M}\mathbb{V}$. Hence, multiplier versions of this test statistic are given by $\mathcal{F}(\mathcal{M}\widehat{\mathbb{V}}^{(1)}), \ldots, \mathcal{F}(\mathcal{M}\widehat{\mathbb{V}}^{(\mathcal{H})})$. Since $\mathcal{F}$ is continuous, an application of the continuous mapping theorem combined with the conclusion of Proposition 4 entails that the latter are asymptotically independent copies of $\mathcal{F}(\mathcal{M}\mathbb{V}) = \mathcal{F}(\mathbb{V}')$. A P-value is then given by

$$\widehat{p}_{\mathcal{F},\mathcal{M}} = \frac{1}{\mathcal{H}} \sum_{h=1}^{\mathcal{H}} \mathbb{I}\left\{\mathcal{F}\left(\mathcal{M}\widehat{\mathbb{V}}^{(h)}\right) > n\, T_{\mathcal{F},\mathcal{M}}(\widehat{\mathbf{C}})\right\}. \tag{11}$$

Note that the weak convergence result about the empirical process $\widehat{\mathbb{V}}$ and its multiplier versions still holds under *any* alternative such that for each $\ell \in \{1, \ldots, d\}$, the partial derivative $C^{[\ell]}$ of the underlying copula C exists and is continuous for $\mathbf{u} \in \{[0,1]^d : 0 < u_\ell < 1\}$. As a consequence, $(\widehat{\mathbb{V}}, \widehat{\mathbb{V}}^{(1)}, \ldots, \widehat{\mathbb{V}}^{(\mathcal{H})}) \rightsquigarrow (\mathbb{V}^\star, \mathbb{V}^{(1)}, \ldots, \mathbb{V}^{(\mathcal{H})})$, where $\mathbb{V}^{(1)}, \ldots, \mathbb{V}^{(\mathcal{H})}$ are independent copies of the limit $\mathbb{V}^\star$ of $\widehat{\mathbb{V}} = \sqrt{n}(\widehat{\mathbf{C}} - \mathbf{C})$, whether $\mathbf{C}$ equals $D\,\mathbf{1}_K$ or not. Thus, a test based on $n\, T_{\mathcal{F},\mathcal{M}}(\widehat{\mathbf{C}})$ is consistent whenever C is such that $\mathcal{F}(\mathcal{M}\mathbf{C}) > 0$. To see this, observe that since $\widehat{\mathbb{V}}$ is tight and $\mathcal{F}$ is assumed continuous,

$$T_{\mathcal{F},\mathcal{M}}\left(\widehat{\mathbf{C}}\right) = \mathcal{F}\left(\mathcal{M}\,\frac{\widehat{\mathbb{V}}}{\sqrt{n}} + \mathcal{M}\mathbf{C}\right) \xrightarrow{n\to\infty} \mathcal{F}\left(\mathcal{M}\mathbf{C}\right) > 0,$$

so that $n\, T_{\mathcal{F},\mathcal{M}}(\widehat{\mathbf{C}}) \overset{n\to\infty}{\to} \infty$ in probability. As a consequence, $\widehat{p}_{\mathcal{F},\mathcal{M}}$ defined in (11) is an asymptotically valid P-value for the test based on $n\, T_{\mathcal{F},\mathcal{M}}(\widehat{\mathbf{C}})$.

Simple formulas derive from the above multiplier bootstrap and the quadratic nature of the test statistics. First define for each $h \in \{1, \ldots, \mathcal{H}\}$ the vector $\boldsymbol{\gamma}^{(h)} = (\gamma_1^{(h)}, \ldots, \gamma_n^{(h)}) \in \mathbb{R}^n$. From the definition of $\widehat{\mathbb{V}}^{(h)}$ given in equation (10), one can write $\widehat{\mathbb{V}}^{(h)} = \widehat{P}\,\boldsymbol{\gamma}^{(h)}/\sqrt{n}$, where the entries of $\widehat{P} \in \mathcal{S}_{Kn}$ are

$$\widehat{P}_{ki} = \mathbb{I}\left\{\mathbf{g}^{(k)}\left(\widehat{\mathbf{U}}_{iA^{(k)}}\right) \leq \mathbf{v}\right\} - \sum_{j=1}^{p} \widehat{D}^{[j]}(\mathbf{v})\,\mathbb{I}\left\{g_j^{(k)}\left(\widehat{U}_{iA_j^{(k)}}\right) \leq v_j\right\}. \tag{12}$$

Using property (5) on $\mathcal{F}$, one readily obtains

$$\mathcal{F}\left(\mathcal{M}\widehat{\mathbb{V}}^{(h)}\right) = \frac{1}{n}\,(\boldsymbol{\gamma}^{(h)})^\top \widehat{\Lambda}\,\boldsymbol{\gamma}^{(h)},$$

where $\widehat{\Lambda} = \mathcal{F}(\mathcal{M}\widehat{P}) \in \mathbb{R}^{n\times n}$ needs to be computed only once from the data. As a consequence, the multiplier bootstrap replicates of the test statistic $n\,T_{\mathcal{F},\mathcal{M}}(\widehat{\mathbf{C}})$ obtains quickly. Approximation formulas for $\widehat{\Lambda}$ are given in Appendix C.

4. Some quadratic functionals

4.1. Cramér–von Mises and diagonal section

A Cramér–von Mises type functional for $S \in \mathcal{S}_{ab}$ is given by

$$\mathcal{F}_{\mathrm{CvM}}(S) = \int_{[0,1]^p} \{S(\mathbf{v})\}^\top S(\mathbf{v})\,\mathrm{d}\mathbf{v},$$

where here and in the sequel, the integration of a matrix of functions is understood to be taken componentwise. Note that if $\mathcal{M}\mathbf{C} \neq \mathbf{0}_q$, then $(\mathcal{M}\mathbf{C})^\top(\mathcal{M}\mathbf{C}) > 0$, so that $\mathcal{F}_{\mathrm{CvM}}(\mathcal{M}\mathbf{C}) > 0$. As a consequence, the test based on $T_{\mathcal{F},\mathcal{M}}(\widehat{\mathbf{C}})$ is consistent against any alternative to $\mathcal{H}_0$, whatever the form of $\mathcal{H}_0$. A version of $\mathcal{F}_{\mathrm{CvM}}$ that considers some sort of *dimension reduction* using the diagonal section instead is

$$\mathcal{F}_{\mathrm{Dia}}(S) = \int_0^1 \{S(v\,\mathbf{1}_p)\}^\top S(v\,\mathbf{1}_p)\,\mathrm{d}v.$$

However, unlike the test based on $\mathcal{F}_{\mathrm{CvM}}$, the test using $\mathcal{F}_{\mathrm{Dia}}$ is not consistent against all kind of alternatives to $\mathcal{H}_0$.

4.2. Characteristic function

The characteristic function of a p-variate random vector $\mathbf{X}$ with distribution function H is defined for $\mathbf{t} = (t_1, \ldots, t_p) \in \mathbb{R}^p$ by $\psi_H(\mathbf{t}) = \mathrm{E}(e^{i\mathbf{t}^\top \mathbf{X}})$, where $i = \sqrt{-1}$. A marginal-free version that depends only on the copula C of $\mathbf{X}$ is given by $\psi_C(\mathbf{t}) = \mathrm{E}(e^{i\mathbf{t}^\top \mathbf{U}})$, where $\mathbf{U} \sim C$. A mapping acting on this complex-valued function is $\int_{\mathbb{R}^p} |\psi_C(\mathbf{t})|^2 \mathrm{d}\omega(\mathbf{t})$, where $\mathrm{d}\omega > 0$ is a weight function. If C admits a density $\mathrm{d}C$, then one can equivalently write

$$\int_{\mathbb{R}^p} |\psi_C(\mathbf{t})|^2 \, \mathrm{d}\omega(\mathbf{t}) = \int_{\mathbb{R}^p} \left| \int_{[0,1]^p} e^{i\mathbf{t}^\top \mathbf{u}} \, \mathrm{d}C(\mathbf{u}) \right|^2 \mathrm{d}\omega(\mathbf{t}).$$

In order to generalize the above functional to $S \in \mathcal{S}_{ab}$, define

$$\mathcal{F}^{\omega}_{\mathrm{Cf}}(S) = \int_{\mathbb{R}^p} \left\{ \left(\int_{[0,1]^p} e^{i\mathbf{t}^\top \mathbf{v}} \, \mathrm{d}S(\mathbf{v}) \right)^\top \left(\int_{[0,1]^p} \overline{e^{i\mathbf{t}^\top \mathbf{u}} \, \mathrm{d}S(\mathbf{u})} \right) \right\} \mathrm{d}\omega(\mathbf{t}),$$

where $\overline{\mathbf{z}}$ is the componentwise complex conjugate of the complex matrix $\mathbf{z}$. In order that the integral with respect to S makes sense, it is assumed that S is of bounded variation on $[0,1]^p$ in the sense of Hardy–Krause. The following lemma provides a formula for $\mathcal{F}^{\omega}_{\mathrm{Cf}}$ that will prove useful for the computation of the related test statistics.

Lemma 1. *For any $S \in \mathcal{S}_{ab}$ of bounded variation,*

$$\mathcal{F}^{\omega}_{\mathrm{Cf}}(S) = \int_{[0,1]^{2p}} \beta^{\omega}(\mathbf{v} - \mathbf{u}) \, \mathrm{d}S^\top(\mathbf{v}) \, \mathrm{d}S(\mathbf{u}), \tag{13}$$

where for $\mathbf{r} = (r_1, \ldots, r_p) \in \mathbb{R}^p$,

$$\beta^{\omega}(\mathbf{r}) = \int_{\mathbb{R}^p} \cos\left(\mathbf{t}^\top \mathbf{r}\right) \mathrm{d}\omega(\mathbf{t}).$$

The following result provides an alternate formula for $\mathcal{F}^{\omega}_{\mathrm{Cf}}$. The proof is based on a multivariate integration by parts formula for Riemann-Stieltje integrals (see [8], for instance). The formula will prove useful for the computation of the multiplier versions of the test statistic based on $\mathcal{F}^{\omega}_{\mathrm{Cf}}$; see Appendix C for details.

Lemma 2. *For any $S \in \mathcal{S}_{ab}$ of bounded variation,*

$$\begin{aligned} \mathcal{F}^{\omega}_{\mathrm{Cf}}(S) \;\; = \;\; & \int_{[0,1]^{2p}} \sum_{\mathcal{I},\mathcal{J} \subseteq S_p} S^\top\left([\mathbf{v}^{\mathcal{I}}, \mathbf{1}]\right) S\left([\mathbf{u}^{\mathcal{J}}, \mathbf{1}]\right) \\ & \qquad \times \left(\frac{\partial}{\partial \mathbf{v}^{\mathcal{I}}} \frac{\partial}{\partial \mathbf{u}^{\mathcal{J}}} \beta^{\omega}\left(\mathbf{v}^{\mathcal{I}} - \mathbf{u}^{\mathcal{J}}\right) \right) \mathrm{d}\mathbf{v} \, \mathrm{d}\mathbf{u}, \end{aligned} \tag{14}$$

where $S_p = \{1, \ldots, p\}$, while $\mathbf{v}^{\mathcal{I}}_j = v_j$ if $j \in \mathcal{I}$ and $\mathbf{v}^{\mathcal{I}}_j = 0$ otherwise.

One can take $\mathcal{F}^{\omega}_{\mathrm{Cf}}$ in equation (14) as the definition for all $S \in \mathcal{S}_{ab}$. Because this functional is continuous, a consequence of Lemma 2 is that the asymptotic result stated in Corollary 1 holds for $\mathcal{F}^{\omega}_{\mathrm{Cf}}$.

4.3. Spearman's rank correlation and other linear measures of dependence

For some null hypotheses, it may be reasonable to only compare moments associated to $C_1, \ldots, C_K$. One such example is the use of Spearman's measure of association as considered by [9] for testing the hypothesis of an homogeneous correlation matrix. In the d-dimensional case, Spearman's rho can be defined as an affine transformation of

$$\eta_{\mathrm{Sp}}(C) = \int_{[0,1]^d} C(\mathbf{u})\,\mathrm{d}\mathbf{u}.$$

More generally, let $\eta : \ell^\infty([0,1]^d) \to \mathbb{R}$ be continuous and linear. For example, the medial correlation coefficient of [1] is an affine transformation of $\eta_{B\ell}(C) = C(\mathbf{1}_d/2)$. A quadratic functional based on η and acting on $S \in \mathcal{S}_{ab}$ may be defined by $\mathcal{F}_\eta(S) = \boldsymbol{\eta}(S)^\top \boldsymbol{\eta}(S)$, where the entries of $\boldsymbol{\eta}(S) \in \mathbb{R}^{a\times b}$ are $\boldsymbol{\eta}(S)_{ij} = \eta(S_{ij})$; see [17] and [24] for details on copula-based measures of association.

Note that a consequence of Corollary 1 applied to the current context yields the convergence in distribution of $n\,T_{\mathcal{F}_{\boldsymbol{\eta}},\mathcal{M}}(\widehat{\mathbf{C}})$ to $\boldsymbol{\eta}(\mathbb{V}')^\top \boldsymbol{\eta}(\mathbb{V}')$. Since $\boldsymbol{\eta}$ is continuous and linear and $\mathbb{V}'$ is a centered Gaussian process, Lemma 3.9.8 in [26] entails $\boldsymbol{\eta}(\mathbb{V}') \sim \mathbb{N}_q(\mathbf{0}_q, A_{\boldsymbol{\eta}})$, where $A_{\boldsymbol{\eta}} = \mathrm{E}\{\boldsymbol{\eta}(\mathbb{V}')\,\boldsymbol{\eta}(\mathbb{V}')^\top\}$. Classical results on the sums of squares of normal vectors that one can find in [23] entail that $n\,T_{\mathcal{F}_{\boldsymbol{\eta}},\mathcal{M}}(\widehat{\mathbf{C}})$ is asymptotically equivalent to $\lambda_1 Y_1^2 + \cdots + \lambda_r Y_r^2$, where $Y_1^2, \ldots, Y_r^2$ are independent chi-squared random variables with $\nu_1, \ldots, \nu_r$ degrees of freedom, respectively, and $\lambda_1, \ldots, \lambda_r$ are the r distinct eigenvalues of $A_{\boldsymbol{\eta}}$ with algebraic multiplicity $\nu_1, \ldots, \nu_r$. For Spearman's functional, one invokes Fubini's theorem to show that $A_{\boldsymbol{\eta}} = \int_{[0,1]^{2p}} \mathrm{E}\{\mathbb{V}'(\mathbf{v})\,\mathbb{V}'(\mathbf{v}')\}\,\mathrm{d}\mathbf{v}\,\mathrm{d}\mathbf{v}'$, while for Blomqvist's functional, $A_{\boldsymbol{\eta}} = \mathrm{E}\{\mathbb{V}'(\mathbf{1}_p/2)\,\mathbb{V}'(\mathbf{1}_p/2)\}$.

5. Investigation of the sample properties of the tests

5.1. General setup

The asymptotic behavior of the test statistic $n\,T^{\mathcal{F},\mathcal{M}}(\widehat{\mathbf{C}})$ under the null hypothesis in (2) has been established in Corollary 1, while the asymptotic validity of the multiplier bootstrap for the computation of P-values was formally obtained in Proposition 4. However, these limit results tell little about the behavior of the tests in small samples. For that reason, it is important to investigate the sample properties of the tests in terms of their ability to keep their nominal size under $\mathcal{H}_0$ and their power against selected alternative hypotheses.

The null hypotheses that were considered in this simulation study are the following: (i) three types of bivariate symmetries, namely diagonal, radial and joint symmetry, (ii) multivariate exchangeability, (iii) some variants of the equality among copulas and (iv) the equality of all the bivariate copulas in a d-variate vector. The main focus is put on the tests based on the Cramér–von Mises (CvM) statistic and on two characteristic function statistics (Cf_1, Cf_2)

corresponding respectively to the weight functions $d\omega_1(\mathbf{t}) = \prod_{j=1}^p e^{-t_j^2}$ and $d\omega_2(\mathbf{t}) = \prod_{j=1}^p e^{-|t_j|}$. In these cases, one can show that for $\mathbf{r} = (r_1, \ldots, r_p)$, $\beta^{\omega_1}(\mathbf{r}) = (\sqrt{\pi})^p \prod_{j=1}^p e^{-r_j^2/4}$ and $\beta^{\omega_2}(\mathbf{r}) = 2^p \prod_{j=1}^p (1+r_j^2)^{-1}$. These three tests are consistent under any alternative to $\mathcal{H}_0$. The tests based on the diagonal functional (Dia) was investigated only when testing for the equality of copulas, *i.e.* in the cases (iii) and (iv) above, because the latter vanishes in situations (i) and (ii); for the same reason, the Spearman and Blomqvist functionals (Sp, Bℓ) were only studied when testing for the equality of all bivariate copulas.

In Appendix C, the formulas for computing the multiplier versions of the test statistics are based on an approximation of $\widehat{P}$ on a grid of $[0,1]^p$ of size N^p; for the simulation results that are reported, $N = 20$ when $p = 2$ and $N = 10$ otherwise. The distribution of the multiplier random variables is the exponential with mean one and the number of multiplier bootstrap samples was $\mathcal{H} = 1\ 000$. All probabilities of rejection were estimated from 1 000 replicates. For each $j \in \{1, \ldots, p\}$, the estimation of the partial derivative $D^{[j]}$ appearing in the limiting representation of $\widehat{\mathbb{V}}$ under the null hypothesis will be based on a finite difference estimator. Namely, for a given A and $\mathbf{g}$, proceed as *e.g.* [25] and define for $h_n = n^{-1/2}$,

$$\widehat{C}^{[j]}_{A,\mathbf{g}}(\mathbf{v})$$

$$= \begin{cases} \dfrac{\widehat{C}_{A,\mathbf{g}}(\mathbf{v}_{j-1}, 2h_n, \mathbf{v}_{p-j})}{2h_n}, & v_j \in [0, h_n), \\[2ex] \dfrac{\widehat{C}_{A,\mathbf{g}}(\mathbf{v} + h_n\mathbf{e}_j) - \widehat{C}_{A,\mathbf{g}}(\mathbf{v} - h_n\mathbf{e}_j)}{2h_n}, & v_j \in [h_n, 1-h_n], \\[2ex] \dfrac{\widehat{C}_{A,\mathbf{g}}(\mathbf{v}_{j-1}, 1, \mathbf{v}_{p-j}) - \widehat{C}_{A,\mathbf{g}}(\mathbf{v}_{j-1}, 1-2h_n, \mathbf{v}_{p-j})}{2h_n}, & v_j \in (1-h_n, 1], \end{cases}$$

where $\mathbf{v}_{j-1} = (v_1, \ldots, v_{j-1})$ and $\mathbf{v}_{p-j} = (v_{j+1}, \ldots, v_p)$. Then, since $C^{[j]}_{A^{(k)},\mathbf{g}^{(k)}} = D^{[j]}$ for each $k \in \{1, \ldots, K\}$ under the null hypothesis, a combined estimator of $D^{[j]}$ is given by

$$\widehat{D}^{[j]}(\mathbf{v}) = \frac{1}{K} \sum_{k=1}^K \widehat{C}^{[j]}_{A^{(k)},\mathbf{g}^{(k)}}(\mathbf{v}).$$

The next proposition establishes that $\widehat{D}^{[1]}, \ldots, \widehat{D}^{[p]}$ satisfy the requirement of equation (9).

Proposition 5. *For each $j \in \{1, \ldots, p\}$, the estimator $\widehat{D}^{[j]}$ is uniformly bounded and for any $\epsilon \in (0, 1/2)$,*

$$\sup_{\mathbf{v} \in [0,1]^p : v_j \in [\epsilon, 1-\epsilon]} \left| \widehat{D}^{[j]}(\mathbf{v}) - D^{[j]}(\mathbf{v}) \right| \xrightarrow{\mathrm{P}} 0.$$

Many choices for the combination matrix are possible. For the upcoming simulations, one takes $\mathcal{M} = I_K - \mathbf{1}_K \mathbf{1}_K^\top / K$ so that $\mathcal{M}\mathbf{C} = \mathbf{C} - \bar{\mathbf{C}}$ for $\mathbf{C} \in \mathcal{S}_{K1}$, where $\bar{\mathbf{C}}$ is the componentwise mean of $\mathbf{C}$. This choice ensures that the tests remain as *omnibus* as possible. A version of $\mathcal{M}$ having linearly independent lines consists in defining $\widetilde{\mathcal{M}}$ as the first $K-1$ lines of $\mathcal{M}$, but $\mathcal{M}$ will be used in the sequel for convenience.

5.2. Bivariate symmetries

A bivariate copula C is symmetric with respect to the main diagonal of the unit square if $C(v_1, v_2) = C(v_2, v_1)$ for all $(v_1, v_2) \in [0,1]^2$. It corresponds to $A^{(1)} = (1,2)$, $A^{(2)} = (2,1)$ and $\mathbf{g}^{(1)} = \mathbf{g}^{(2)} = (I, I)$, because $\mathbf{g}^{(1)}(\mathbf{U}_{A^{(1)}}) = (U_1, U_2)$ and $\mathbf{g}^{(2)}(\mathbf{U}_{A^{(2)}}) = (U_2, U_1)$, so that the associated copulas are $C_{A^{(1)},\mathbf{g}^{(1)}}(v_1, v_2) = \mathrm{P}(U_1 \leq v_1, U_2 \leq v_2) = C(v_1, v_2)$ and

$$C_{A^{(2)},\mathbf{g}^{(2)}}(v_1, v_2) = \mathrm{P}(U_2 \leq v_1, U_1 \leq v_2) = C(v_2, v_1).$$

Test statistics for symmetry have been developed by [11] and [20]. In particular, one of the procedures suggested by [11] is based on the Cramér–von Mises functional described in Section 4.

In order to study the size of the introduced tests for diagonal symmetry, observations from the Clayton (Cℓ) and Frank (Fr) Archimedean copulas, as well as from the Normal copula (N) have been simulated; as is well known, these copulas are symmetric. The results about the estimated probability of rejection of the null hypothesis can be found in the top panel of Table 1 for the Cramér–von Mises and the two characteristic function statistics Cf_1 and Cf_2. Observations under $\mathcal{H}_1$ were generated using a particular case of an idea of [14] that allows to *asymmetrize* a symmetric copula C via $C^\star(v_1, v_2) = u_1^\delta C(v_1^{1-\delta}, v_2)$ for some $\delta \in (0,1)$. The special case $\delta = 1/2$ has been considered for the Clayton (C$\ell^{\mathcal{K}}$), Frank ($\mathrm{Fr}^{\mathcal{K}}$) and Normal ($\mathrm{N}^{\mathcal{K}}$) copulas.

A copula C is radial symmetric if $C(v_1, v_2) = v_1 + v_2 - 1 + C(1-v_1, 1-v_2)$ for all $(v_1, v_2) \in [0,1]^2$. It corresponds to $A^{(1)} = A^{(2)} = (1,2)$, $\mathbf{g}^{(1)} = (I, I)$ and $\mathbf{g}^{(2)} = (1-I, 1-I)$, since in that case, $\mathbf{g}^{(1)}(\mathbf{U}_{A^{(1)}}) = (U_1, U_2)$ and $\mathbf{g}^{(2)}(\mathbf{U}_{A^{(2)}}) = (1-U_1, 1-U_2)$, so that the corresponding copulas are $C_{A^{(1)},\mathbf{g}^{(1)}}(v_1, v_2) = \mathrm{P}(U_1 \leq v_1, U_2 \leq v_2) = C(v_1, v_2)$ and

$$\begin{aligned} C_{A^{(2)},\mathbf{g}^{(2)}}(v_1, v_2) &= \mathrm{P}(1-U_1 \leq v_1, 1-U_2 \leq v_2) \\ &= \mathrm{P}(U_1 \geq 1-v_1, U_2 \geq 1-v_2) \\ &= 1-(1-v_1)-(1-v_2)+C(1-v_1, 1-v_2) \\ &= v_1 + v_2 - 1 + C(1-v_1, 1-v_2). \end{aligned}$$

Test statistics for radial symmetry were investigated by [10]. For the results in the middle panel of Table 1, models considered under the null hypothesis were the Frank and Normal copulas. Note that Frank's model is the only member of the Archimedean family that possesses the radial symmetry property. Alternative hypotheses are provided by Clayton's copula. Finally note that a stronger

Table 1

Probability of rejection of the null hypotheses of diagonal (top panel), radial (middle panel) and joint symmetry (bottom panel), as estimated from 1 000 replicates from various scenarios under $\mathcal{H}_0$ and $\mathcal{H}_1$ when $n \in \{50, 100, 200\}$ for the Cramér–von Mises (CvM) and two characteristic function (Cf_1, Cf_2) statistics

	Copula model	τ	$n=50$ CvM	$n=50$ Cf_1	$n=50$ Cf_2	$n=100$ CvM	$n=100$ Cf_1	$n=100$ Cf_2	$n=200$ CvM	$n=200$ Cf_1	$n=200$ Cf_2
$\mathcal{H}_0$	Cℓ	.5	1.8	4.2	4.4	3.1	4.4	4.4	3.1	4.6	4.8
	N	.5	2.0	5.0	4.2	2.2	4.3	4.6	2.8	3.5	2.8
	Fr	.5	1.1	4.5	3.9	1.9	5.1	4.1	1.5	4.5	4.2
$\mathcal{H}_1$	C$\ell^{\mathcal{K}}$	.5	5.5	9.2	8.7	11.1	14.2	14.7	26.5	32.3	30.6
	N$^{\mathcal{K}}$	.5	9.2	13.6	13.5	15.6	23.6	24.2	38.6	51.5	50.8
	Fr$^{\mathcal{K}}$	.5	11.5	16.9	17.7	24.6	30.7	32.4	52.0	60.0	60.5
$\mathcal{H}_0$	Fr	.3	2.6	3.3	3.5	2.5	3.1	3.4	3.7	4.4	4.6
	Fr	.7	1.1	0.9	1.2	1.5	2.1	3.4	2.8	3.1	3.9
	N	.3	3.6	4.6	4.8	3.9	4.9	4.8	3.7	3.9	4.3
	N	.7	0.8	0.5	1.3	2.3	1.7	3.2	1.9	2.5	3.0
$\mathcal{H}_1$	Cℓ	.3	15.9	18.5	19.9	32.8	41.6	39.1	63.6	75.6	71.4
	Cℓ	.7	19.5	12.8	33.6	64.2	83.2	76.6	98.3	99.8	99.3
$\mathcal{H}_0$	Π	.0	5.0	5.9	5.2	4.3	4.9	4.6	3.8	4.0	3.8
	Fr$^{\mathcal{JS}}$	.3	4.6	5.2	5.0	4.1	4.3	4.9	3.2	3.3	3.4
	Fr$^{\mathcal{JS}}$	.7	2.6	2.6	2.3	5.2	5.2	5.5	5.2	4.8	4.9
	N$^{\mathcal{JS}}$	.3	4.8	5.9	5.5	4.0	3.6	4.2	4.3	4.0	3.9
	N$^{\mathcal{JS}}$	.7	4.8	5.2	4.4	3.6	2.7	3.4	5.4	4.8	5.3
$\mathcal{H}_1$	Cℓ	.1	16.9	18.3	18.2	29.9	31.1	31.4	50.4	52.0	51.4
	N	.1	15.1	18.2	17.0	26.6	29.8	28.6	46.3	52.4	50.0
	Fr	.1	19.6	21.6	21.6	28.5	30.9	30.5	52.1	53.8	53.9

hypothesis is joint symmetry; this occurs when $A^{(1)} = A^{(2)} = A^{(3)} = (1,2)$, $\mathbf{g}^{(1)} = (I, I)$, $\mathbf{g}^{(2)} = (1-I, I)$ and $\mathbf{g}^{(3)} = (I, 1-I)$. The independence copula $\Pi(v_1, v_2) = v_1 v_2$ is jointly symmetric has been considered as a model under the null hypothesis. In addition, joint symmetric versions of the Normal and Frank copulas arising from the mixture $(U,V) \stackrel{d}{=} (X,Y)$ or $(U,V) \stackrel{d}{=} (1-X,Y)$ with probability $1/2$, where (X,Y) follows a radial symmetric copula, have been considered; the models are referred to Fr$^{\mathcal{JS}}$ and N$^{\mathcal{JS}}$. Finally, alternative hypotheses based on the Clayton, Normal and Frank copulas have been considered. The results can be found in the bottom panel of Table 1.

The levels of dependence for the models in Table 1 are controlled via Kendall's tau whose expression for a copula C is $\tau(C) = 4\int_{[0,1]^2} C(u_1,u_2)\mathrm{d}C(u_1,u_2) - 1$. Looking globally at these results, one notes that the tests are good at keeping their nominal 5% level under the selected null hypotheses. An exception to that occurs for the Cramér–von Mises when $\tau(C) = .7$. As documented by [2], statistical tests using the multiplier method for empirical copulas often show some problems at keeping their size under high levels of dependence. A larger value of the approximation parameter N would possibly give more accurate results, but at the cost of much slower procedures. Overall, the tests are very powerful. Of course, for a given alternative, the estimated powers increase with the sample size. Interestingly, the characteristic function statistics are more powerful than the Cramér–von Mises statistic, with a mild superiority of Cf_1 over Cf_2. To the author's knowledge, it is the first time that rank-based characteristic functions are used for inference in copula models. Their good behavior here compared to the widely-used CvM statistic is very promising for other types of applications in semi-parametric inference.

5.3. Multivariate exchangeability

A generalization of diagonal symmetry for $d > 2$ dimensions is exchangeability. Specifically, a copula $C : [0,1]^d \to [0,1]$ is said to be exchangeable if $C(u_1, \ldots, u_d) = C(u_{\pi(1)}, \ldots, u_{\pi(d)})$ for any permutation $\pi(1), \ldots, \pi(d)$ of the first d integers. In that case, $A^{(1)}, \ldots, A^{(d!)}$ are the sets of all possible permutations of $\{1, \ldots, d\}$ and $\mathbf{g}^{(1)} = \cdots = \mathbf{g}^{(d!)} = (I, \ldots, I)$. All the members of the family of Archimedean copulas are exchangeable; this property follows easily from their representation $C(\mathbf{u}) = \phi^{-1}\{\phi(u_1) + \cdots + \phi(u_d)\}$, where ϕ is a univariate generator. Multivariate elliptical copulas with an equicorrelated covariance matrix are also exchangeable dependence models; this includes the so-called *equicorrelated* Normal copula.

Simulated data from two classes of models have been considered. The first is based on the d-variate Normal copula with correlation matrix $\Sigma \in \mathbb{R}^{d\times d}$. When Σ is such that $\theta = \Sigma_{12} = \Sigma_{21}$ and $\Sigma_{jj'} = 0.2$ for all $(j, j') \neq (1,2), (2,1)$, the model is noted $\mathrm{N}_1(\theta)$; if one takes Σ with $\Sigma_{jj'} = \theta^{|j-j'|}$ for some $\theta \in (0,1)$, the model is referred to $\mathrm{N}_2(\theta)$. Models $\mathrm{N}_1(.2)$ and $\mathrm{N}_2(.0)$ are thus situations when the null hypothesis of exchangeability holds. Another model construction will be based on the asymmetrization of the multivariate Clayton copula $C^{\mathrm{C}\ell}$ via $C^\star(\mathbf{u}) = \mathbf{u}^{\boldsymbol{\delta}} C^{\mathrm{C}\ell}(\mathbf{u}^{\mathbf{1}_d^\top - \boldsymbol{\delta}})$, where $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_d) \in [0,1]^d$, $\mathbf{u}^{\boldsymbol{\delta}} = (u_1^{\delta_1}, \ldots, u_d^{\delta_d})$ and $\mathbf{u}^{\mathbf{1}_d^\top - \boldsymbol{\delta}} = (u_1^{1-\delta_1}, \ldots, u_d^{1-\delta_d})$, where the level of dependence of $C^{\mathrm{C}\ell}$ is adjusted in such a way that Kendall's tau equals $1/2$. When $\delta_1 = \theta$ and $\delta_2 = \cdots = \delta_d = .4$, the model is referred to $\mathrm{C}\ell_1^{\mathcal{K}}(\theta)$; if $\delta_j = \theta^j$ for all $j \in \{1, \ldots, d\}$ and for some $\theta \in (0,1)$, this is noted $\mathrm{C}\ell_2^{\mathcal{K}}(\theta)$.

The results for the test statistics CvM, Cf_1 and Cf_2 are presented in Table 2 when $d = 3$. For the null hypotheses that were considered, *i.e.* $\mathrm{N}_1(.2)$, $\mathrm{N}_2(.0)$, $\mathrm{C}\ell_1^{\mathcal{K}}(.4)$ and $\mathrm{C}\ell_2^{\mathcal{K}}(.0)$, the test statistics Cf_1 and Cf_2 are good at keeping their nominal level, even for sample sizes as small as $n = 50$. However, the test based on the Cramér–von Mises functional has much more difficulty at keeping its size. This unwanted behavior seems to have a significant effect on the power of the test. Indeed, the power of the CvM statistic is much less than that of the two characteristic function statistics under all scenarios under $\mathcal{H}_1$.

TABLE 2

Probability of rejection of the null hypothesis of exchangeability in trivariate random vectors as estimated from 1 000 replicates from various scenarios under $\mathcal{H}_0$ and $\mathcal{H}_1$ when $n \in \{50, 100, 200\}$ for the Cramér–von Mises (CvM) and two characteristic function $(\mathrm{Cf}_1, \mathrm{Cf}_2)$ *statistics*

	Copula model	$n = 50$ CvM	Cf_1	Cf_2	$n = 100$ CvM	Cf_1	Cf_2	$n = 200$ CvM	Cf_1	Cf_2
$\mathcal{H}_0$	$\mathrm{N}_1(.2)$	1.7	4.4	3.8	2.4	4.9	5.2	2.2	5.6	4.5
	$\mathrm{N}_2(.0)$	2.2	6.4	5.5	2.0	6.8	6.4	3.2	7.4	6.1
$\mathcal{H}_0$	$\mathrm{C}\ell_1^{\mathcal{K}}(.4)$	1.6	4.7	4.1	2.2	6.0	5.8	1.9	4.6	4.3
	$\mathrm{C}\ell_2^{\mathcal{K}}(.0)$	0.8	4.5	3.9	0.7	5.3	5.2	1.5	3.4	2.3
$\mathcal{H}_1$	$\mathrm{N}_1(.4)$	5.2	16.2	15.2	9.3	33.7	30.7	22.9	65.6	60.0
	$\mathrm{N}_1(.6)$	20.8	65.1	59.6	51.1	96.1	94.0	89.9	100.0	99.7
	$\mathrm{N}_2(.6)$	9.0	36.9	33.1	29.5	76.5	73.3	75.5	98.9	97.7
	$\mathrm{N}_2(.8)$	6.9	48.5	39.9	31.2	89.5	84.8	91.7	99.8	99.7
$\mathcal{H}_1$	$\mathrm{C}\ell_1^{\mathcal{K}}(.6)$	2.3	8.1	7.3	2.6	10.0	9.8	5.1	17.3	16.7
	$\mathrm{C}\ell_1^{\mathcal{K}}(.8)$	4.6	16.1	14.5	8.9	29.3	28.6	22.3	53.4	51.2
	$\mathrm{C}\ell_2^{\mathcal{K}}(.6)$	2.4	12.4	11.2	5.2	23.9	23.1	15.0	43.6	42.4
	$\mathrm{C}\ell_2^{\mathcal{K}}(.8)$	3.0	6.7	5.9	3.0	8.0	8.5	5.5	13.5	11.7

5.4. Equality of copulas

The general framework developed in this paper allows easily for the testing of the equality of any subset of p-variate copulas that one can extract from the whole d-variate copula of a multivariate vector $\mathbf{X} = (X_1, \ldots, X_d)$. In the case when $K = 2$, it corresponds to the hypothesis of equality of two copulas as investigated by [21]. With the tools proposed here, it can easily be generalized to the equality of $K > 2$ copulas. In addition, one can consider copulas having one or more overlapping components, for example when testing for the equality of the dependence structures of (X_1, X_2) and (X_2, X_3).

All the simulation results that are presented in Table 3 used data generated from model $\mathrm{N}_1(\theta)$ described in subsection 5.3 and from a version $\mathrm{T}_3(\theta)$ based on the Student copula with $\nu = 3$ degrees of freedom. Only the results for $p = 2$ are presented. The values in the top panel of Table 3 correspond to situations where there are no overlapping components. Copulas with overlapping components were taken into account for the results in the bottom panel: one has $A^{(1)} = (1, 2)$, $A^{(2)} = (2, 3)$ when $K = 2$, $A^{(1)} = (1, 2)$, $A^{(2)} = (2, 3)$, $A^{(3)} = (3, 4)$ when $K = 3$, and $A^{(1)} = (1, 2)$, $A^{(2)} = (2, 3)$, $A^{(3)} = (3, 4)$, $A^{(4)} = (4, 5)$ when $K = 4$. The tests that were investigated are based on the Cramér–von Mises (CvM), diagonal (Dia) and the two characteristic function statistics (Cf_1, Cf_2).

Looking at the results in Table 3, one notes that the tests keep their nominal levels well under all scenarios under the null hypothesis, *i.e.*, when $\theta = .2$. One can also see that the four tests are very good at detecting departures from $\mathcal{H}_0$. As expected, the power increases as the sample size increases and as the departure from the null hypothesis gets larger. Probably due to their close

Table 3

Probability of rejection of the null hypothesis of equality of K bivariate copulas as estimated from 1 000 replicates from models $N_1(\theta)$ and $T_3(\theta)$ with $\theta = .2$ ($\mathcal{H}_0$) and $\theta \in \{.4, .6\}$ ($\mathcal{H}_1$) when $n \in \{50, 100\}$ for the Cramér–von Mises (CvM), diagonal (Dia) and two characteristic function $(\mathrm{Cf}_1, \mathrm{Cf}_2)$ statistics. Upper panel: no overlapping components; bottom panel: overlapping components

	Copula	$n = 50$				$n = 100$			
K	model	CvM	Dia	Cf_1	Cf_2	CvM	Dia	Cf_1	Cf_2
2	$\mathrm{N}_1(.2)$	5.4	5.0	5.9	6.0	4.7	4.9	5.8	5.4
	$\mathrm{N}_1(.4)$	16.8	15.5	21.3	19.8	22.5	21.8	28.5	26.4
	$\mathrm{N}_1(.6)$	51.3	48.9	62.2	58.1	84.1	81.7	90.1	87.7
	$\mathrm{T}_3(.2)$	4.9	5.0	5.9	6.1	4.8	6.1	5.1	4.9
	$\mathrm{T}_3(.4)$	44.2	48.1	48.9	48.1	78.0	75.4	82.2	81.2
	$\mathrm{T}_3(.6)$	88.6	92.2	90.8	90.7	99.9	99.9	99.9	99.9
3	$\mathrm{N}_1(.2)$	3.9	4.7	6.7	6.0	3.1	3.2	3.2	3.6
	$\mathrm{N}_1(.4)$	12.0	12.7	17.8	16.1	20.3	21.1	25.2	23.6
	$\mathrm{N}_1(.6)$	47.3	48.2	63.0	56.8	85.8	84.7	91.5	88.6
	$\mathrm{T}_3(.2)$	4.1	4.8	5.1	4.7	4.4	5.5	5.4	5.5
	$\mathrm{T}_3(.4)$	40.7	47.7	47.3	45.7	76.7	76.6	81.2	79.4
	$\mathrm{T}_3(.6)$	90.1	93.5	93.2	92.4	100.0	99.8	99.9	100.0
4	$\mathrm{N}_1(.2)$	3.3	5.5	5.4	5.4	2.9	3.9	3.2	3.1
	$\mathrm{N}_1(.4)$	9.8	12.0	16.1	14.1	19.6	20.6	27.1	25.0
	$\mathrm{N}_1(.6)$	43.4	45.8	57.2	53.5	81.5	80.3	88.9	86.4
	$\mathrm{T}_3(.2)$	2.8	4.4	3.7	3.6	4.6	4.6	6.0	5.5
	$\mathrm{T}_3(.4)$	34.6	41.7	43.9	42.3	74.4	75.5	78.8	77.6
	$\mathrm{T}_3(.6)$	86.3	91.7	91.0	90.5	99.8	99.8	100.0	100.0
2	$\mathrm{N}_1(.2)$	4.7	5.2	6.2	6.1	4.7	3.9	4.7	4.7
	$\mathrm{N}_1(.4)$	14.9	14.2	19.5	18.3	27.2	24.9	33.5	31.5
	$\mathrm{N}_1(.6)$	56.2	52.2	64.5	60.6	86.3	84.2	91.8	89.2
	$\mathrm{T}_3(.2)$	4.1	3.9	4.8	4.6	4.4	4.6	5.2	5.3
	$\mathrm{T}_3(.4)$	46.7	50.0	51.4	49.5	80.9	80.5	82.9	83.1
	$\mathrm{T}_3(.6)$	91.9	93.8	93.0	93.1	100.0	99.9	99.8	99.8
3	$\mathrm{N}_1(.2)$	3.3	4.8	5.7	4.8	2.9	4.1	3.3	3.8
	$\mathrm{N}_1(.4)$	10.3	12.8	16.1	14.8	21.0	21.7	29.0	26.9
	$\mathrm{N}_1(.6)$	50.3	51.8	64.7	59.2	87.4	85.1	92.9	90.4
	$\mathrm{T}_3(.2)$	3.0	4.4	3.9	3.7	4.0	5.1	4.3	4.5
	$\mathrm{T}_3(.4)$	40.8	48.1	48.7	46.9	80.5	81.9	83.3	82.8
	$\mathrm{T}_3(.6)$	90.9	93.3	94.5	93.0	99.8	100.0	99.8	99.8
4	$\mathrm{N}_1(.2)$	3.4	4.1	5.7	4.9	2.3	3.0	3.5	3.3
	$\mathrm{N}_1(.4)$	10.7	11.6	15.8	14.4	17.7	19.7	25.1	23.0
	$\mathrm{N}_1(.6)$	40.9	44.4	58.8	52.4	81.0	82.5	91.0	87.6
	$\mathrm{T}_3(.2)$	1.7	4.9	3.5	3.0	4.4	4.5	6.2	6.1
	$\mathrm{T}_3(.4)$	35.3	40.1	45.2	42.0	76.2	81.4	80.9	79.9
	$\mathrm{T}_3(.6)$	88.1	93.2	92.6	92.4	100.0	99.9	100.0	100.0

connection, the results for the Cramér–von Mises and the diagonal statistics are similar; this conclusion must be taken with care as the latter is not consistent against all types of alternatives. Here again, the characteristic function statistics are systematically more powerful than the CvM statistic. Also observe that for a given θ, the estimated powers are clearly higher for the Student copulas compared to alternatives based on the Normal model. Finally note that results not presented here indicate that the procedures based on the Spearman and Blomqvist functionals are quite powerful, too. However, these tests would have no power against alternatives where, *e.g.*, all Spearman measures are equal but where the copulas have different structures.

5.5. Distributional equality of all the pairs

The problem of testing for the equality of the $d(d-1)/2$ components of a Spearman matrix $\mathcal{R} \in \mathbb{R}^{d\times d}$, where $\mathcal{R}_{\ell\ell'}$ is the value of Spearman's rank correlation coefficient for the pair $(X_\ell, X_{\ell'})$, was considered by [9]. This hypothesis corresponds to $A^{(1)}, \ldots, A^{(d(d-1)/2)}$ being the $d(d-1)/2$ possible pairs in $\{1,\ldots,d\}$ and $\mathbf{g}^{(k)} = (I,I)$ for each $k \in \{1,\ldots,d(d-1)/2\}$. Statistics similar as those investigated by [9] arise when one takes the Spearman functional in the case $p = 2$. The problem can be generalized in many ways, *e.g.* for (i) testing the equality of all the pairwise dependence levels, as measured by an association measure like Blomqvist's index, (ii) testing the equality of p-variate dependence coefficients for a selection of p-variate copulas extracted from C when $p > 2$, and (iii) testing the equality of all bivariate copulas of C. The results in Table 4 concern item (iii), where it is assumed that all bivariate copulas are symmetric. In that case, there are $K = d(d-1)/2$ bivariate copulas that must be compared.

One can see from the results in Table 4 that the test statistics Cf_1, Cf_2, Sp and $\mathrm{B}\ell$ are overall good at keeping their nominal level. The Cramér–von Mises and diagonal statistics tend to be too conservative, especially under model $\mathrm{C}\ell_2^{\mathcal{K}}(.0)$, *i.e.* the Clayton copula. Under all scenarios of alternative hypotheses, the Spearman test statistic Sp considered by [9] is the most powerful; however, one has to keep in mind that this test is not consistent against alternatives where the bivariate copulas have the same level of dependence but different structures. Again, the characteristic function statistics perform very well, being superior to the Cramér–von Mises and diagonal statistics.

6. Concluding remarks

A general statistical methodology for testing a wide variety of composite hypotheses about copulas extracted from multivariate distributions has been introduced. The framework is very flexible, as it allows the testing of many hypotheses already considered in the literature, as well as many extensions and other ones that were never considered to date. The performance of the tests in small samples show the relevance of the introduced method.

The class of quadratic functionals from which the tests were developed enables for the investigation of several types of statistics including the well-known and widely-used Cramér–von Mises functional, as well as new rank-based diagonal and characteristic function mappings. From the empirical results presented in

Table 4

Probability of rejection of the null hypothesis of the equality in distribution of all pairs in a multivariate random vector as estimated from 1 000 replicates from various scenarios under $\mathcal{H}_0$ and $\mathcal{H}_1$ when $n \in \{50, 100\}$ for the Cramér–von Mises (CvM), diagonal (Dia), two characteristic function $(\mathrm{Cf}_1, \mathrm{Cf}_2)$, Spearman (Sp) and Blomqvist (Bℓ) statistics. Upper panel: $d = 3$; bottom panel: $d = 4$

	Copula model	$n=50$						$n=100$					
		CvM	Dia	Cf_1	Cf_2	Sp	Bℓ	CvM	Dia	Cf_1	Cf_2	Sp	Bℓ
$\mathcal{H}_0$	$\mathrm{N}_1(.2)$	2.9	5.2	6.0	4.8	6.0	6.4	4.9	4.6	5.0	5.4	5.4	5.0
	$\mathrm{N}_2(.0)$	3.0	4.0	4.9	4.4	5.3	5.7	4.5	4.9	5.1	4.6	5.6	5.8
	$\mathrm{C}\ell_1^{\mathcal{K}}(.4)$	2.3	4.0	4.7	3.8	5.1	4.8	4.4	5.7	6.4	5.9	6.2	5.9
	$\mathrm{C}\ell_2^{\mathcal{K}}(.0)$	1.3	4.0	4.6	3.2	4.5	4.9	1.7	3.6	5.6	4.6	5.5	5.4
$\mathcal{H}_1$	$\mathrm{N}_1(.4)$	12.6	13.7	18.7	16.6	20.3	11.3	25.5	25.3	34.7	31.3	36.6	19.5
	$\mathrm{N}_1(.6)$	53.1	55.2	69.2	64.5	72.3	36.3	89.3	87.4	94.7	93.2	95.4	65.7
	$\mathrm{N}_2(.4)$	18.8	18.5	30.3	27.3	32.8	14.6	36.1	35.6	50.6	45.9	54.4	24.9
	$\mathrm{N}_2(.6)$	22.8	27.4	45.8	38.9	48.6	19.1	61.0	57.6	78.9	73.2	82.7	33.1
$\mathcal{H}_1$	$\mathrm{C}\ell_1^{\mathcal{K}}(.6)$	4.1	6.0	6.0	6.0	7.4	6.6	6.8	8.7	9.4	8.9	10.5	7.8
	$\mathrm{C}\ell_1^{\mathcal{K}}(.8)$	12.1	14.3	17.3	16.4	19.0	11.6	20.8	24.5	24.9	24.0	27.4	18.4
	$\mathrm{C}\ell_2^{\mathcal{K}}(.5)$	11.5	15.5	18.2	15.8	20.2	11.3	20.8	26.4	27.6	25.3	31.4	17.7
	$\mathrm{C}\ell_2^{\mathcal{K}}(.7)$	6.7	8.2	11.0	10.5	12.2	9.8	8.7	10.6	11.4	10.7	13.7	8.0

	Copula model	$n=50$						$n=100$					
		CvM	Dia	Cf_1	Cf_2	Sp	Bℓ	CvM	Dia	Cf_1	Cf_2	Sp	Bℓ
$\mathcal{H}_0$	$\mathrm{N}_1(.2)$	1.6	2.7	3.7	3.2	4.4	5.6	2.3	3.3	4.2	4.3	4.1	4.9
	$\mathrm{N}_2(.0)$	1.7	3.0	4.1	3.6	4.3	5.3	3.2	4.7	5.5	4.7	5.7	4.8
	$\mathrm{C}\ell_1^{\mathcal{K}}(.4)$	1.0	3.4	3.2	3.2	3.3	4.8	1.4	3.5	3.3	3.0	4.0	3.7
	$\mathrm{C}\ell_2^{\mathcal{K}}(.0)$	0.5	1.4	2.8	1.5	2.8	4.2	1.1	2.6	3.8	3.7	4.0	5.0
$\mathcal{H}_1$	$\mathrm{N}_1(.4)$	3.9	8.2	11.2	9.0	13.7	9.1	12.3	16.1	20.2	17.8	23.3	13.2
	$\mathrm{N}_1(.6)$	29.6	39.6	51.6	45.1	57.0	27.5	70.2	77.2	87.1	82.7	89.9	54.1
	$\mathrm{N}_2(.4)$	22.4	27.3	42.8	36.2	48.4	19.6	61.3	60.6	78.6	72.3	82.5	41.4
	$\mathrm{N}_2(.6)$	46.6	52.4	76.2	67.3	79.4	35.6	88.6	89.7	97.9	95.8	98.7	61.7
$\mathcal{H}_1$	$\mathrm{C}\ell_1^{\mathcal{K}}(.6)$	3.2	5.5	6.2	5.3	8.1	7.0	5.6	10.3	9.7	9.6	11.7	10.3
	$\mathrm{C}\ell_1^{\mathcal{K}}(.8)$	9.8	15.4	17.5	15.2	20.4	13.9	28.2	35.5	36.5	35.2	40.8	25.3
	$\mathrm{C}\ell_2^{\mathcal{K}}(.5)$	19.2	29.7	34.7	30.0	39.7	21.9	47.0	61.3	61.0	57.8	66.4	37.8
	$\mathrm{C}\ell_2^{\mathcal{K}}(.7)$	8.6	13.6	16.1	14.7	18.9	12.8	21.1	29.4	29.9	29.3	34.8	21.6

Section 5, the latter are particularly appealing as they lead to consistent tests that seem to outperform the popular Cramér–von Mises statistics, at least for small sample sizes. A more formal approach considering the asymptotic power of these tests under contiguous alternatives could be an interesting avenue of future research. These characteristic function statistics could also provide competing procedures in many other inferential problems involving copulas, *e.g.*, goodness-of-fit tests to parametric families and tests of independence.

Another special case of the general class of quadratic functions that was introduced is the test statistics proposed by [16]. Their idea consists in considering a function $h \in \ell^\infty([0,1]^p)$ at a finite and pre-determined set of points $\Theta = (\theta_1, \ldots, \theta_L)$ and to define the quadratic functional $\mathcal{F}_{\mathrm{LG}}(h) = h_\Theta^\top h_\Theta$, where $h_\Theta = (h(\theta_1), \ldots, h(\theta_L))^\top$. In order to generalize to $S \in \mathcal{S}_{ab}$, define $S_{\theta_1}, \ldots, S_{\theta_L}$, where $S_{\theta_\ell} = S(\theta_\ell) \in \mathbb{R}^{a\times b}$, and let

$$\mathcal{F}_{\mathrm{LG}}(S) = \sum_{\ell=1}^{L} S_{\theta_\ell}^\top S_{\theta_\ell}.$$

Formulas for $\mathcal{F}_{\mathrm{LG}}(\mathcal{M}\widehat{L})$ and $\mathcal{F}_{\mathrm{LG}}(\mathcal{M}\widehat{P})$ would be easy to obtain.

Finally, note that the tools proposed in this work are not restricted to the i.i.d. case. Assuming a condition on the strong mixing coefficients of $(\mathbf{X}_i)_{i\in\mathbb{Z}}$, one deduces from [5] that the empirical copula process converges weakly to a Gaussian limit whose covariance function depends on the serial structure of the sequence. Re-sampling could be done from a serial version of the multiplier method [see 4]. As a consequence, the theoretical results derived here could be adapted in a straightforward manner to the context of multivariate time series.

Appendix A: Proofs

First consider the following auxiliary lemma.

Lemma 3. *Let $S_p = \{1,\ldots,p\}$ and $\widetilde{S}_p^{\mathbf{g}} = \{j \in S_p : g_j = 1 - I\}$, where I is the identity function. One has for $\mathbf{v} = (v_1,\ldots,v_p) \in [0,1]^p$ that*

$$C_{A,\mathbf{g}}(\mathbf{v}) = \sum_{B\subseteq \widetilde{S}_p^{\mathbf{g}}} (-1)^{|B|}\, C\left(\mathbf{v}^B\right) \quad \text{and} \quad \widehat{C}_{A,\mathbf{g}}(\mathbf{v}) = \sum_{B\subseteq \widetilde{S}_p^{\mathbf{g}}} (-1)^{|B|}\, C_n\left(\mathbf{v}^B\right),$$

where $|B|$ is the cardinality of the set B and for each $\ell \in \{1,\ldots,d\}$,

$$\mathbf{v}_\ell^B = \begin{cases} v_j, & \text{if } \ell = A_j \text{ for some } j \in S_p \setminus \widetilde{S}_p^{\mathbf{g}}; \\ 1 - v_j, & \text{if } \ell = A_j \text{ for some } j \in B; \\ 1, & \text{otherwise.} \end{cases}$$

Proof. One can write

$$\begin{aligned} C_{A,\mathbf{g}}(\mathbf{v}) &= \mathrm{E}\left\{ \prod_{j\in S_p} \mathbb{I}\left(g_j\left(U_{A_j}\right) \le v_j\right) \right\} \\ &= \mathrm{E}\left\{ \prod_{j\in S_p\setminus \widetilde{S}_p^{\mathbf{g}}} \mathbb{I}\left(U_{A_j} \le v_j\right) \prod_{j\in \widetilde{S}_p^{\mathbf{g}}} \mathbb{I}\left(1 - U_{A_j} \le v_j\right) \right\}. \end{aligned}$$

Then, note that

$$\begin{aligned} \prod_{j\in \widetilde{S}_p^{\mathbf{g}}} \mathbb{I}\left(1 - U_{A_j} \le v_j\right) &= \prod_{j\in \widetilde{S}_p^{\mathbf{g}}} \left\{1 - \mathbb{I}\left(U_{A_j} < 1 - v_j\right)\right\} \\ &= \sum_{B\subseteq \widetilde{S}_p^{\mathbf{g}}} (-1)^{|B|} \prod_{j\in B} \mathbb{I}\left(U_{A_j} < 1 - v_j\right). \end{aligned}$$

Hence,

$$\begin{aligned} C_{A,\mathbf{g}}(\mathbf{v}) &= \sum_{B\subseteq\widetilde{S}_p^{\mathbf{g}}} (-1)^{|B|}\,\mathrm{E}\left\{\prod_{j\in S_p\setminus\widetilde{S}_p^{\mathbf{g}}} \mathbb{I}\left(U_{A_j}\le v_j\right)\prod_{j\in B}\mathbb{I}\left(U_{A_j}<1-v_j\right)\right\} \\ &= \sum_{B\subseteq\widetilde{S}_p^{\mathbf{g}}} (-1)^{|B|}\,C\left(\mathbf{v}^B\right). \end{aligned}$$

The expression for $\widehat{C}_{A,\mathbf{g}}$ obtains similarly. □

A.1. Proof of Proposition 1

From Lemma 3, the empirical process $\widehat{\mathbb{C}}_{A,\mathbf{g}}$ can be written

$$\widehat{\mathbb{C}}_{A,\mathbf{g}}(\mathbf{v}) = \sum_{B\subseteq\widetilde{S}_p^{\mathbf{g}}} (-1)^{|B|}\,\mathbb{C}_n\left(\mathbf{v}^B\right),$$

where $\mathbb{C}_n = \sqrt{n}(C_n - C)$ is the empirical copula process. Then, again from Lemma 3,

$$C_{A,\mathbf{g}}^{[j]}(\mathbf{v}) = \sum_{B\subseteq\widetilde{S}_p^{\mathbf{g}}:j\in B\cup S_p\setminus\widetilde{S}_p^{\mathbf{g}}} (-1)^{|B|} C^{[A_j]}\left(\mathbf{v}^B\right)\,g_j'(v_j).$$

Hence the assumption on the partial derivatives of $C_{A,\mathbf{g}}$ ensures that for each $\ell = A_j$, $C^{[\ell]}(\mathbf{v}^B)$ exists and is continuous on $\{[0,1]^d : 0 < v_\ell < 1\}$ for any $B \subseteq \widetilde{S}_p^{\mathbf{g}}$. As a consequence, one has from [25] that $\mathbb{C}_n(\mathbf{v}^B) \rightsquigarrow \mathbb{C}(\mathbf{v}^B)$, where $\mathbb{C}$ is described in equation (7). Thus, in virtue of the continuous mapping Theorem, $\widehat{\mathbb{C}}_{A,\mathbf{g}}$ converges weakly in the space $\ell^\infty([0,1]^p)$ to a limit process of the form

$$\begin{aligned} &\mathbb{C}_{A,\mathbf{g}}(\mathbf{v}) \\ &\quad= \sum_{B\subseteq\widetilde{S}_p^{\mathbf{g}}} (-1)^{|B|}\,\mathbb{C}\left(\mathbf{v}^B\right) \\ &\quad= \mathbb{B}_{A,\mathbf{g}}(\mathbf{v}) + \sum_{B\subseteq\widetilde{S}_p^{\mathbf{g}}} (-1)^{|B|}\left\{-\sum_{\ell=1}^{d} C^{[\ell]}\left(\mathbf{v}^B\right)\,\mathbb{B}\left(\mathbf{1}_{\ell-1}, v_\ell^B, \mathbf{1}_{d-\ell}\right)\right\} \\ &\quad= \mathbb{B}_{A,\mathbf{g}}(\mathbf{v}) \\ &\qquad+ \sum_{B\subseteq\widetilde{S}_p^{\mathbf{g}}} (-1)^{|B|}\left\{-\sum_{j\in B\cup S_p\setminus\widetilde{S}_p^{\mathbf{g}}} C^{[A_j]}\left(\mathbf{v}^B\right)\,\mathbb{B}\left(\mathbf{1}_{A_j-1}, g_j(v_j), \mathbf{1}_{d-A_j}\right)\right\}, \end{aligned}$$

where

$$\mathbb{B}_{A,\mathbf{g}}(\mathbf{v}) = \sum_{B\subseteq\widetilde{S}_p^{\mathbf{g}}} (-1)^{|B|}\,\mathbb{B}\left(\mathbf{v}^B\right).$$

The last equality follows from the fact that $\mathbb{B}(\mathbf{1}_{\ell-1}, v_\ell^B, \mathbf{1}_{d-\ell})$ vanishes unless $\ell = A_j$ for some $j \in B \cup S_p \setminus \widetilde{S}_p^{\mathbf{g}}$. Observe that because $\mathbb{B}$ vanishes when one or more of its component is zero, all summands in the definition of $\mathbb{B}_{A,\mathbf{g}}$ vanish if $j \in S_p \setminus \widetilde{S}_p^{\mathbf{g}}$, except when $B = \emptyset$; then $\mathbb{B}_{A,\mathbf{g}}(\mathbf{1}_{j-1}, v_j, \mathbf{1}_{p-j}) = \mathbb{B}(\mathbf{1}_{A_j-1}, v_j, \mathbf{1}_{d-A_j})$ in that case. Otherwise, if $j \in \widetilde{S}_p^{\mathbf{g}}$, then the only nonzero term occurs when $B = \{j\}$, so that $\mathbb{B}_{A,\mathbf{g}}(\mathbf{1}_{j-1}, v_j, \mathbf{1}_{p-j}) = -\mathbb{B}(\mathbf{1}_{A_j-1}, 1 - v_j, \mathbf{1}_{d-A_j})$. Hence, $\mathbb{B}_{A,\mathbf{g}}(\mathbf{1}_{j-1}, v_j, \mathbf{1}_{p-j}) = g_j'(v_j)\, \mathbb{B}(\mathbf{1}_{A_j-1}, g_j(v_j), \mathbf{1}_{d-A_j})$ or equivalently,

$$\mathbb{B}(\mathbf{1}_{A_j-1}, g_j(v_j), \mathbf{1}_{d-A_j}) = g_j'(v_j)\, \mathbb{B}_{A,\mathbf{g}}(\mathbf{1}_{j-1}, v_j, \mathbf{1}_{p-j}).$$

It follows that

$$\begin{aligned}
&\sum_{B \subseteq \widetilde{S}_p^{\mathbf{g}}} (-1)^{|B|} \sum_{j \in B \cup S_p \setminus \widetilde{S}_p^{\mathbf{g}}} C^{[A_j]}\left(\mathbf{v}^B\right) \mathbb{B}\left(\mathbf{1}_{A_j-1}, g_j(v_j), \mathbf{1}_{d-A_j}\right) \\
&\quad = \sum_{B \subseteq \widetilde{S}_p^{\mathbf{g}}} (-1)^{|B|} \sum_{j \in B \cup S_p \setminus \widetilde{S}_p^{\mathbf{g}}} C^{[A_j]}\left(\mathbf{v}^B\right) g_j'(v_j)\, \mathbb{B}_{A,\mathbf{g}}\left(\mathbf{1}_{j-1}, v_j, \mathbf{1}_{p-j}\right) \\
&\quad = \sum_{j=1}^{p} \left\{ \sum_{B \subseteq \widetilde{S}_p^{\mathbf{g}} : j \in B \cup S_p \setminus \widetilde{S}_p^{\mathbf{g}}} (-1)^{|B|} C^{[A_j]}\left(\mathbf{v}^B\right) g_j'(v_j) \right\} \mathbb{B}_{A,\mathbf{g}}\left(\mathbf{1}_{j-1}, v_j, \mathbf{1}_{p-j}\right) \\
&\quad = \sum_{j=1}^{p} C_{A,\mathbf{g}}^{[j]}(\mathbf{v})\, \mathbb{B}_{A,\mathbf{g}}\left(\mathbf{1}_{j-1}, v_j, \mathbf{1}_{p-j}\right).
\end{aligned}$$

Hence, one can conclude that

$$\mathbb{C}_{A,\mathbf{g}}(\mathbf{v}) = \mathbb{B}_{A,\mathbf{g}}(\mathbf{v}) - \sum_{j=1}^{p} C_{A,\mathbf{g}}^{[j]}(\mathbf{v})\, \mathbb{B}_{A,\mathbf{g}}\left(\mathbf{1}_{j-1}, v_j, \mathbf{1}_{p-j}\right).$$

That $\mathbb{B}_{A,\mathbf{g}}$ is centered Gaussian is obvious from the fact that it is a linear combination of Gaussian processes. It remains to show that its covariance structure is that of a $C_{A,\mathbf{g}}$-Brownian bridge. To this end, note that $\mathbb{B}_{A,\mathbf{g}}$ is the limit of

$$\widehat{\mathbb{B}}_{A,\mathbf{g}}(\mathbf{v}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left\{ \mathbb{I}\left(\mathbf{g}\left(\mathbf{U}_{iA}\right) \leq \mathbf{v}\right) - C_{A,\mathbf{g}}(\mathbf{v}) \right\},$$

where $\mathbf{g}(\mathbf{U}_{1A}), \ldots, \mathbf{g}(\mathbf{U}_{nA})$ are i.i.d. $C_{A,\mathbf{g}}$. Hence the covariance structure of $\mathbb{B}_{A,\mathbf{g}}$ obtains from the finite-dimensional covariances. One obtains easily

$$\begin{aligned}
\mathrm{E}\left\{\widehat{\mathbb{B}}_{A,\mathbf{g}}(\mathbf{v})\, \widehat{\mathbb{B}}_{A,\mathbf{g}}(\mathbf{v}')\right\} &= \mathrm{E}\left\{\mathbb{I}\left(\mathbf{g}\left(\mathbf{U}_{1A}\right) \leq \mathbf{v}\right) \mathbb{I}\left(\mathbf{g}\left(\mathbf{U}_{1A}\right) \leq \mathbf{v}'\right)\right\} \\
&\quad - \mathrm{E}\left\{\mathbb{I}\left(\mathbf{g}\left(\mathbf{U}_{1A}\right) \leq \mathbf{v}\right)\right\} \mathrm{E}\left\{\mathbb{I}\left(\mathbf{g}\left(\mathbf{U}_{1A}\right) \leq \mathbf{v}\right)\right\} \\
&= C_{A,\mathbf{g}}(\mathbf{v} \wedge \mathbf{v}') - C_{A,\mathbf{g}}(\mathbf{v})\, C_{A,\mathbf{g}}(\mathbf{v}'),
\end{aligned}$$

finishing the proof.

A.2. Proof of Proposition 2

First note that

$$\widehat{\mathbb{V}} = \begin{pmatrix} \widehat{\mathbb{C}}_{A^{(1)},\mathbf{g}^{(1)}} \\ \vdots \\ \widehat{\mathbb{C}}_{A^{(K)},\mathbf{g}^{(K)}} \end{pmatrix}.$$

Hence, from Proposition 1 in the special case when $C_{A^{(k)},\mathbf{g}^{(k)}} = D$ for all $k \in \{1,\ldots,K\}$, one has

$$\widehat{\mathbb{V}}(\mathbf{v}) \rightsquigarrow \begin{pmatrix} \mathbb{B}_{A^{(1)},\mathbf{g}^{(1)}}(\mathbf{v}) \\ \vdots \\ \mathbb{B}_{A^{(K)},\mathbf{g}^{(K)}}(\mathbf{v}) \end{pmatrix} - \sum_{j=1}^{p} D^{[j]}(\mathbf{v}) \begin{pmatrix} \mathbb{B}_{A^{(1)},\mathbf{g}^{(1)}}(\mathbf{1}_{j-1}, v_j, \mathbf{1}_{p-j}) \\ \vdots \\ \mathbb{B}_{A^{(K)},\mathbf{g}^{(K)}}(\mathbf{1}_{j-1}, v_j, \mathbf{1}_{p-j}) \end{pmatrix},$$

where for each $k \in \{1,\ldots,K\}$,

$$\mathbb{B}_{A^{(k)},\mathbf{g}^{(k)}}(\mathbf{v}) = \sum_{B \subseteq \widetilde{S}_p^{(k)}} (-1)^{|B|}\, \mathbb{B}\left(\mathbf{v}^{B,(k)}\right),$$

with $\widetilde{S}_p^{(k)} = \{j \in S_p : g_j^{(k)} = 1 - I\}$ and for each $\ell \in \{1,\ldots,d\}$,

$$\mathbf{v}_\ell^{B,(k)} = \begin{cases} v_j, & \text{if } \ell = A_j^{(k)} \text{ for some } j \in S_p \setminus \widetilde{S}_p^{(k)}; \\ 1 - v_j, & \text{if } \ell = A_j^{(k)} \text{ for some } j \in B; \\ 1, & \text{otherwise.} \end{cases}$$

Then, let

$$\mathbb{W} = \begin{pmatrix} \mathbb{B}_{A^{(1)},\mathbf{g}^{(1)}} \\ \vdots \\ \mathbb{B}_{A^{(K)},\mathbf{g}^{(K)}} \end{pmatrix}$$

and note that this vector of processes appears as the tight limit of

$$\widehat{\mathbb{W}} = \begin{pmatrix} \widehat{\mathbb{B}}_{A^{(1)},\mathbf{g}^{(1)}} \\ \vdots \\ \widehat{\mathbb{B}}_{A^{(K)},\mathbf{g}^{(K)}} \end{pmatrix},$$

where

$$\widehat{\mathbb{B}}_{A^{(k)},\mathbf{g}^{(k)}}(\mathbf{v}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left\{ \mathbb{I}\left(\mathbf{g}^{(k)}\left(\mathbf{U}_{iA^{(k)}}\right) \le \mathbf{v}\right) - D(\mathbf{v}) \right\}.$$

The covariance structure of $\mathbb{W}$ then obtains easily from the finite-dimensional distributions, proceeding similarly as in the proof of Proposition 1. Specifically,

$$\begin{aligned} &\mathrm{E}\left\{\widehat{\mathbb{B}}_{A^{(k)},\mathbf{g}^{(k)}}(\mathbf{v})\,\widehat{\mathbb{B}}_{A^{(k')},\mathbf{g}^{(k')}}(\mathbf{v}')\right\} \\ &\quad = \mathrm{E}\left\{\mathbb{I}\left(\mathbf{g}^{(k)}\left(\mathbf{U}_{1A^{(k)}}\right) \le \mathbf{v}\right) \mathbb{I}\left(\mathbf{g}^{(k')}\left(\mathbf{U}_{1A^{(k')}}\right) \le \mathbf{v}'\right)\right\} \end{aligned}$$

$$
\begin{aligned}
& - \mathrm{E}\left\{\mathbb{I}\left(\mathbf{g}^{(k)}\left(\mathbf{U}_{1A^{(k)}}\right) \leq \mathbf{v}\right)\right\} \mathrm{E}\left\{\mathbb{I}\left(\mathbf{g}^{(k')}\left(\mathbf{U}_{1A^{(k')}}\right) \leq \mathbf{v}'\right)\right\} \\
= \;& \gamma_{kk'}(\mathbf{v}, \mathbf{v}') - D(\mathbf{v})\, D(\mathbf{v}').
\end{aligned}
$$

One can then write $\mathrm{E}\left\{\mathbb{W}(\mathbf{v})\, \mathbb{W}(\mathbf{v}')^\top\right\} = \gamma(\mathbf{v}, \mathbf{v}') - D(\mathbf{v})\, D(\mathbf{v}')\, \mathbf{1}_K \mathbf{1}_K^\top$.

A.3. Proof of Corollary 1

Using the fact that $\mathcal{M}\mathbf{C} = \mathbf{0}_q$ under $\mathcal{H}_0$ and from property (5), one can write

$$
n\, T_{\mathcal{F},\mathcal{M}}\left(\widehat{\mathbf{C}}\right) = n\, \mathcal{F}\left(\mathcal{M}\widehat{\mathbf{C}} - \mathcal{M}\mathbf{C}\right) = \mathcal{F}\left\{\mathcal{M}\left(\widehat{\mathbf{C}} - \mathbf{C}\right)\sqrt{n}\right\} = \mathcal{F}\left(\mathcal{M}\widehat{\mathbb{V}}\right).
$$

Since $\mathcal{F}$ is continuous, one can apply the continuous mapping theorem and conclude that $\mathcal{F}(\mathcal{M}\widehat{\mathbb{V}}) \rightsquigarrow \mathcal{F}(\mathbb{V}')$, where $\mathbb{V}' = \mathcal{M}\mathbb{V}$ and $\mathbb{V}$ is the weak limit of $\widehat{\mathbb{V}}$ given in Proposition 2. From the representation of $\mathbb{V}$, one can write

$$
\mathbb{V}'(\mathbf{v}) = \mathbb{W}'(\mathbf{v}) - \sum_{j=1}^{p} D^{[j]}(\mathbf{v})\, \mathbb{W}'(\mathbf{1}_{j-1}, v_j, \mathbf{1}_{p-j}),
$$

where $\mathbb{W}' = \mathcal{M}\mathbb{W}$ is a q-dimensional vector of centered Gaussian processes on $[0,1]^p$ with

$$
\begin{aligned}
\mathrm{E}\left\{\mathbb{W}'(\mathbf{v})\, \mathbb{W}'(\mathbf{v}')^\top\right\} &= \mathcal{M}\, \mathrm{E}\left\{\mathbb{W}(\mathbf{v})\, \mathbb{W}(\mathbf{v}')^\top\right\} \mathcal{M}^\top \\
&= \mathcal{M}\, \gamma(\mathbf{v}, \mathbf{v}')\, \mathcal{M}^\top - D(\mathbf{v})\, D(\mathbf{v}')\, \mathcal{M}\, \mathbf{1}_K (\mathcal{M}\, \mathbf{1}_K)^\top \\
&= \mathcal{M}\, \gamma(\mathbf{v}, \mathbf{v}')\, \mathcal{M}^\top.
\end{aligned}
$$

The last equality follows from $\mathcal{M}\, \mathbf{1}_K = \mathbf{0}_q$, by assumption on $\mathcal{M}$.

A.4. Proof of Proposition 3

Let $S \in \mathcal{S}_{q1}$ and $\Delta_n \to \Delta \in \mathcal{S}_{q1}$ as $n \to \infty$. By the Hadamard differentiability of $\mathcal{F}$, one has for $t_n \to 0 \in \mathbb{R}$ that

$$
\lim_{n\to\infty} \left| \frac{\mathcal{F}\left(S + \Delta_n\, t_n\right) - \mathcal{F}(S)}{t_n} - \mathcal{F}'_S(\Delta) \right| = 0.
$$

Now for $\mathcal{M} \in \mathbb{R}^{q\times K}$ and $\Delta_n \in S_{K1}$,

$$
\begin{aligned}
& \lim_{n\to\infty} \left| \frac{T_{\mathcal{F},\mathcal{M}}\left(S + \Delta_n\, t_n\right) - T_{\mathcal{F},\mathcal{M}}(S)}{t_n} - \mathcal{F}'_{\mathcal{M}S}(\mathcal{M}\Delta) \right| \\
&= \lim_{n\to\infty} \left| \frac{\mathcal{F}\left(\mathcal{M}S + \mathcal{M}\Delta_n\, t_n\right) - \mathcal{F}(\mathcal{M}S)}{t_n} - \mathcal{F}'_{\mathcal{M}S}(\mathcal{M}\Delta) \right| \\
&= 0,
\end{aligned}
$$

since $\mathcal{M}S \in \mathcal{S}_{q1}$ and $\mathcal{M}\Delta_n \to \mathcal{M}\Delta \in \mathcal{S}_{q1}$, and because of the Hadamard differentiability of $\mathcal{F}$. Thus $T_{\mathcal{F},\mathcal{M}}(S)$ is Hadamard differentiable with derivative at S given by $\mathcal{F}'_{\mathcal{M}S}(\mathcal{M}\Delta)$. The functional Delta method combined with

$\sqrt{n}(\widehat{\mathbf{C}} - \mathbf{C}) \rightsquigarrow \mathbb{V}^\star$ yields $\sqrt{n}\{T_{\mathcal{F},\mathcal{M}}(\widehat{\mathbf{C}}) - T_{\mathcal{F},\mathcal{M}}(\mathbf{C})\} \rightsquigarrow \mathcal{F}'_{\mathcal{M}\mathbf{C}}(\mathcal{M}\mathbb{V}^\star)$. Since $\mathcal{F}' : \mathcal{S}_{q1} \to \mathbb{R}$ is linear, Lemma 3.9.8 in [26] entails that $\mathcal{F}'_{\mathcal{M}\mathbf{C}}(\mathcal{M}\mathbb{V}^\star)$ is univariate Normal with mean $\mathrm{E}\{\mathcal{F}'_{\mathcal{M}\mathbf{C}}(\mathcal{M}\mathbb{V}^\star)\} = \mathcal{F}'_{\mathcal{M}\mathbf{C}}(\mathbf{0}_q) = 0$ and variance $\sigma^2_{\mathcal{F},\mathcal{M}} = \mathrm{var}\,\{\mathcal{F}'_{\mathcal{M}\mathbf{C}}\,(\mathcal{M}\mathbb{V}^\star)\}$.

A.5. Proof of Proposition 4

Using arguments similar as those in the proof of Lemma 3, one can show that

$$\widehat{\mathbb{B}}^{(h)}_{A^{(k)},\mathbf{g}^{(k)}}(\mathbf{v}) = \sum_{B\subseteq \widetilde{S}_p^{(k)}} (-1)^{|B|}\,\widehat{\mathbb{B}}^{(h)}\left(\mathbf{v}^{B,(k)}\right)$$

for each $k \in \{1,\ldots,K\}$, where $\widetilde{S}_p^{(k)}$ and $\mathbf{v}^{B,(k)}$ are defined in the proof of Proposition 2. If $j \in S_p \setminus \widetilde{S}_p^{(k)}$, then

$$\begin{aligned}
\widehat{\mathbb{B}}^{(h)}_{A^{(k)},\mathbf{g}^{(k)}}(\mathbf{1}_{j-1}, v_j, \mathbf{1}_{p-j}) &= \frac{1}{\sqrt{n}}\sum_{i=1}^n \gamma_i^{(h)}\,\mathbb{I}\left(\widehat{U}_{iA_j^{(k)}} \le v_j\right) \\
&= \widehat{\mathbb{B}}^{(h)}\left(\mathbf{1}_{A_j^{(k)}-1}, v_j, \mathbf{1}_{d-A_j^{(k)}}\right).
\end{aligned}$$

On the other side, if $j \in \widetilde{S}_p^{(k)}$,

$$\begin{aligned}
\widehat{\mathbb{B}}^{(h)}_{A^{(k)},\mathbf{g}^{(k)}}(\mathbf{1}_{j-1}, v_j, \mathbf{1}_{p-j}) &= \frac{1}{\sqrt{n}}\sum_{i=1}^n \gamma_i^{(h)}\,\mathbb{I}\left(\widehat{U}_{iA_j^{(k)}} \ge 1 - v_j\right) \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^n \gamma_i^{(h)}\left\{1 - \mathbb{I}\left(\widehat{U}_{iA_j^{(k)}} < 1 - v_j\right)\right\} \\
&= -\frac{1}{\sqrt{n}}\sum_{i=1}^n \gamma_i^{(h)}\,\mathbb{I}\left(\widehat{U}_{iA_j^{(k)}} < 1 - v_j\right) \\
&= -\widehat{\mathbb{B}}^{(h)}\left(\mathbf{1}_{A_j^{(k)}-1}, 1 - v_j, \mathbf{1}_{d-A_j^{(k)}}\right).
\end{aligned}$$

Thus, one can write

$$\widehat{\mathbb{B}}^{(h)}_{A^{(k)},\mathbf{g}^{(k)}}(\mathbf{1}_{j-1}, v_j, \mathbf{1}_{p-j}) = (g_j^{(k)})'(v_j)\,\widehat{\mathbb{B}}^{(h)}\left(\mathbf{1}_{A_j^{(k)}-1}, g_j^{(k)}(v_j), \mathbf{1}_{d-A_j^{(k)}}\right).$$

Hence, the k-th component of $\widehat{\mathbb{V}}^{(h)}$ is

$$\begin{aligned}
&\widehat{\mathbb{V}}_k^{(h)}(\mathbf{v}) \\
&= \sum_{B\subseteq \widetilde{S}_p^{(k)}} (-1)^{|B|}\left\{\widehat{\mathbb{B}}^{(h)}\left(\mathbf{v}^{B,(k)}\right) - \sum_{\ell=1}^d \widehat{D}^{[\ell]}\left(\mathbf{v}^{B,(k)}\right)\,\widehat{\mathbb{B}}^{(h)}\left(\mathbf{1}_{\ell-1}, v_\ell^B, \mathbf{1}_{d-\ell}\right)\right\}.
\end{aligned}$$

The expression inside the brackets is $\widehat{\mathbb{C}}^{(h)}$ defined in equation (8), but with $\widehat{C}^{[\ell]}$ replaced by $\widehat{D}^{[\ell]}$, so that one can write

$$\widehat{\mathbb{V}}_k^{(h)}(\mathbf{v}) = \sum_{B \subseteq \widetilde{S}_p^{(k)}} (-1)^{|B|}\, \widehat{\mathbb{C}}^{(h)}\left(\mathbf{v}^{B,(k)}\right).$$

From Proposition 3.2 in [25], one concludes that for each $k \in \{1, \ldots, K\}$, $(\widehat{\mathbb{V}}_k, \widehat{\mathbb{V}}_k^{(h)})$ converges weakly to $(\mathbb{V}_k, \mathbb{V}_k^{(h)})$, where $\mathbb{V}_k^{(h)}$ is an independent copy of

$$\sum_{B \subseteq \widetilde{S}_p^{(k)}} (-1)^{|B|}\, \mathbb{C}^{(h)}\left(\mathbf{v}^{B,(k)}\right) = \mathbb{C}_{A^{(k)},\mathbf{g}^{(k)}}(\mathbf{v}).$$

Hence, the vector $(\widehat{\mathbb{V}}, \widehat{\mathbb{V}}^{(1)}, \ldots, \widehat{\mathbb{V}}^{(M)})$ converges weakly to $(\mathbb{V}, \mathbb{V}^{(1)}, \ldots, \mathbb{V}^{(M)})$, where $\mathbb{V}^{(1)}, \ldots, \mathbb{V}^{(M)}$ are independent copies of $\mathbb{V}$.

A.6. Proof of Lemma 1

Because $e^{i\mathbf{t}^\top \mathbf{v}} = \cos(\mathbf{t}^\top \mathbf{v}) + i \sin(\mathbf{t}^\top \mathbf{v})$, one has

$$e^{i\mathbf{t}^\top \mathbf{v}}\, \overline{e^{i\mathbf{t}^\top \mathbf{u}}} = \cos(\mathbf{t}^\top \mathbf{v})\cos(\mathbf{t}^\top \mathbf{u}) + \sin(\mathbf{t}^\top \mathbf{v})\sin(\mathbf{t}^\top \mathbf{u}) + i\, r(\mathbf{u},\mathbf{v}),$$

where $r(\mathbf{u},\mathbf{v}) = \sin(\mathbf{t}^\top \mathbf{v})\cos(\mathbf{t}^\top \mathbf{u}) - \cos(\mathbf{t}^\top \mathbf{v})\sin(\mathbf{t}^\top \mathbf{u})$. Hence, one can write

$$\begin{aligned}
&\left(\int_{[0,1]^p} e^{i\mathbf{t}^\top \mathbf{v}}\, \mathrm{d}S(\mathbf{v})\right)^\top \left(\int_{[0,1]^p} \overline{e^{i\mathbf{t}^\top \mathbf{u}}\, \mathrm{d}S(\mathbf{u})}\right) \\
&\qquad = \int_{[0,1]^{2p}} e^{i\mathbf{t}^\top \mathbf{v}}\, \overline{e^{i\mathbf{t}^\top \mathbf{u}}}\, \mathrm{d}S^\top(\mathbf{v})\, \mathrm{d}S(\mathbf{u}) \\
&\qquad = \int_{[0,1]^{2p}} \left\{\cos(\mathbf{t}^\top \mathbf{v})\cos(\mathbf{t}^\top \mathbf{u}) + \sin(\mathbf{t}^\top \mathbf{v})\sin(\mathbf{t}^\top \mathbf{u})\right\} \mathrm{d}S^\top(\mathbf{v})\, \mathrm{d}S(\mathbf{u}), \quad (15)
\end{aligned}$$

where the last equality arises since $r(\mathbf{v},\mathbf{u}) = -r(\mathbf{u},\mathbf{v})$, so that

$$\begin{aligned}
\int_{[0,1]^{2p}} r(\mathbf{u},\mathbf{v})\, \mathrm{d}S^\top(\mathbf{v})\, \mathrm{d}S(\mathbf{u}) &= \int_{[0,1]^{2p}} r(\mathbf{v},\mathbf{u})\, \mathrm{d}S^\top(\mathbf{u})\, \mathrm{d}S(\mathbf{v}) \\
&= -\int_{[0,1]^{2p}} r(\mathbf{u},\mathbf{v})\, \mathrm{d}S^\top(\mathbf{u})\, \mathrm{d}S(\mathbf{v}) \\
&= -\int_{[0,1]^{2p}} r(\mathbf{u},\mathbf{v})\, \mathrm{d}S^\top(\mathbf{v})\, \mathrm{d}S(\mathbf{u}).
\end{aligned}$$

Using the trigonometric identity $\cos x \cos y + \sin x \sin y = \cos(x - y)$ and integrating over $\mathbb{R}^p$ with respect to the weight function $\mathrm{d}\omega$, one obtains using Fubini's theorem that

$$\mathcal{F}_{\mathrm{Cf}}^{\omega}(S) \quad = \quad \int_{\mathbb{R}^p} \left\{\int_{[0,1]^{2p}} \cos\left\{\mathbf{t}^\top(\mathbf{v}-\mathbf{u})\right\} \mathrm{d}S^\top(\mathbf{v})\, \mathrm{d}S(\mathbf{u})\right\} \mathrm{d}\omega(\mathbf{t})$$

$$
\begin{aligned}
&= \int_{[0,1]^{2p}} \left\{ \int_{\mathbb{R}^p} \cos\left\{ \mathbf{t}^\top (\mathbf{v} - \mathbf{u}) \right\} \mathrm{d}\omega(\mathbf{t}) \right\} \mathrm{d}S^\top(\mathbf{v})\, \mathrm{d}S(\mathbf{u}) \\
&= \int_{[0,1]^{2p}} \beta^\omega (\mathbf{v} - \mathbf{u})\, \mathrm{d}S^\top(\mathbf{v})\, \mathrm{d}S(\mathbf{u}).
\end{aligned}
$$

A.7. Proof of Lemma 2

From Proposition A.1 of [8], the fact that S is of bounded variation allows to deduce that for any function G of bounded variation on $[0,1]^p$,

$$
\int_{[0,1]^p} G(\mathbf{v})\, \mathrm{d}S(\mathbf{v}) = \int_{[0,1]^p} \sum_{\mathcal{I} \subseteq S_p} S\left([\mathbf{v}^{\mathcal{I}}, \mathbf{1}]\right) \left(\frac{\partial}{\partial \mathbf{v}^{\mathcal{I}}} G(\mathbf{v}^{\mathcal{I}}) \right) \mathrm{d}\mathbf{v},
$$

where $S_p = \{1, \ldots, p\}$. It then follows that

$$
\begin{aligned}
&\int_{[0,1]^{2p}} G(\mathbf{v})\, G(\mathbf{u})\, \mathrm{d}S^\top(\mathbf{v})\, \mathrm{d}S(\mathbf{u}) \\
&\quad = \int_{[0,1]^{2p}} \sum_{\mathcal{I}, \mathcal{J} \subseteq S_p} S^\top\left([\mathbf{v}^{\mathcal{I}}, \mathbf{1}]\right) S\left([\mathbf{u}^{\mathcal{J}}, \mathbf{1}]\right) \left(\frac{\partial}{\partial \mathbf{v}^{\mathcal{I}}} \frac{\partial}{\partial \mathbf{u}^{\mathcal{J}}} G(\mathbf{v}^{\mathcal{I}})\, G(\mathbf{u}^{\mathcal{J}}) \right) \mathrm{d}\mathbf{v}\, \mathrm{d}\mathbf{u}.
\end{aligned}
$$

Using the trigonometric identity $\cos x \cos y + \sin x \sin y = \cos(x - y)$, one deduces that formula (15) can be written

$$
\begin{aligned}
&\left(\int_{[0,1]^p} e^{i \mathbf{t}^\top \mathbf{v}}\, \mathrm{d}S(\mathbf{v}) \right)^\top \left(\int_{[0,1]^p} \overline{e^{i \mathbf{t}^\top \mathbf{u}}\, \mathrm{d}S(\mathbf{u})} \right) \\
&\qquad = \int_{[0,1]^{2p}} \sum_{\mathcal{I}, \mathcal{J} \subseteq S_p} S^\top\left([\mathbf{v}^{\mathcal{I}}, \mathbf{1}]\right) S\left([\mathbf{u}^{\mathcal{J}}, \mathbf{1}]\right) \\
&\qquad\qquad \times \left(\frac{\partial}{\partial \mathbf{v}^{\mathcal{I}}} \frac{\partial}{\partial \mathbf{u}^{\mathcal{J}}} \cos\left\{ \mathbf{t}^\top \left(\mathbf{v}^{\mathcal{I}} - \mathbf{u}^{\mathcal{J}} \right) \right\} \right) \mathrm{d}\mathbf{v}\, \mathrm{d}\mathbf{u}.
\end{aligned}
$$

Integrating this expression over $\mathbb{R}^p$ with respect to the weight function $\mathrm{d}\omega$ and making use of Fubini's theorem yields the result.

A.8. Proof of Proposition 5

From Lemma 3, one can show that for an arbitrary p-tuple A and a given $\mathbf{g} = (g_1, \ldots, g_p) \in \mathcal{G}$,

$$
\widehat{C}^{[j]}_{A,\mathbf{g}}(\mathbf{v}) = \sum_{\substack{B \subseteq \widetilde{S}^{\mathbf{g}}_p, \\ j \in B \cup S_p \setminus \widetilde{S}^{\mathbf{g}}_p}} (-1)^{|B|}\, \widehat{C}^{[A_j]}\left(\mathbf{v}^B\right) g_j'(v_j),
$$

where from [25], the estimator

$$\widehat{C}^{[\ell]}(\mathbf{u}) = \frac{C_n\left(\mathbf{u} + b_n \mathbf{e}_\ell\right) - C_n\left(\mathbf{u} - b_n \mathbf{e}_\ell\right)}{2\, b_n}$$

is such that for any $\epsilon \in (0, 1/2)$,

$$\sup_{\substack{\mathbf{u}\in[0,1]^d,\\ u_\ell \in[\epsilon,1-\epsilon]}} \left| \widehat{C}^{[\ell]}(\mathbf{u}) - C^{[\ell]}(\mathbf{u}) \right| \xrightarrow{\mathrm{P}} 0.$$

Hence,

$$\widehat{C}^{[j]}_{A,\mathbf{g}}(\mathbf{v}) - C^{[j]}_{A,\mathbf{g}}(\mathbf{v}) = \sum_{\substack{B\subseteq \widetilde{S}^{\mathbf{g}}_p,\\ j\in B\cup S_p\setminus \widetilde{S}^{\mathbf{g}}_p}} (-1)^{|B|} \left\{ \widehat{C}^{[A_j]}\left(\mathbf{v}^B\right) - C^{[A_j]}\left(\mathbf{v}^B\right) \right\} g_j'(v_j).$$

As a consequence, for any $\epsilon \in (0, 1/2)$,

$$\sup_{\substack{\mathbf{v}\in[0,1]^p,\\ v_j \in[\epsilon,1-\epsilon]}} \left| \widehat{C}^{[j]}_{A,\mathbf{g}}(\mathbf{v}) - C^{[j]}_{A,\mathbf{g}}(\mathbf{v}) \right| \le \sum_{\substack{B\subseteq \widetilde{S}^{\mathbf{g}}_p,\\ j\in B\cup S_p\setminus \widetilde{S}^{\mathbf{g}}_p}} \sup_{\substack{\mathbf{v}\in[0,1]^p,\\ v_j \in[\epsilon,1-\epsilon]}} \left| \widehat{C}^{[A_j]}\left(\mathbf{v}^B\right) - C^{[A_j]}\left(\mathbf{v}^B\right) \right|.$$

Since the expression on the righthand side of the last inequality converges in probability to zero, one finally has

$$\begin{aligned} &\sup_{\substack{\mathbf{v}\in[0,1]^p,\\ v_j \in[\epsilon,1-\epsilon]}} \left| \widehat{D}^{[j]}(\mathbf{v}) - D^{[j]}(\mathbf{v}) \right| \\ &\qquad \le \quad \frac{1}{K} \sum_{k=1}^K \sup_{\substack{\mathbf{v}\in[0,1]^p,\\ v_j \in[\epsilon,1-\epsilon]}} \left| \widehat{C}^{[j]}_{A^{(k)},\mathbf{g}^{(k)}}(\mathbf{v}) - C^{[j]}_{A^{(k)},\mathbf{g}^{(k)}}(\mathbf{v}) \right| \xrightarrow{\mathrm{P}} 0. \end{aligned}$$

Appendix B: Formulas for the test statistics

Formulas will be given for $\Lambda = \mathcal{F}(\mathcal{M}\widehat{L}) \in \mathbb{R}^{n\times n}$; this matrix appears in the computation of the test statistics. In the sequel, the combination matrix is $\mathcal{M} = I_K - \mathbf{1}_K\, \mathbf{1}_K^\top / K$. Also, the notation $\mathbf{G}_i^{(k)} = \mathbf{g}^{(k)}(\widehat{\mathbf{U}}_{iA^{(k)}})$ will be adopted throughout, so that $\widehat{L}_{ki}(\mathbf{v}) = \mathbb{I}(\mathbf{G}_i^{(k)} \le \mathbf{v})$.

B.1. Cramér–von Mises and diagonal section

Since $\mathcal{M}^\top\mathcal{M} = \mathcal{M}$ (*i.e.* $\mathcal{M}$ is idempotent), $(\mathcal{M}\,\widehat{L})^\top \mathcal{M}\,\widehat{L} = \widehat{L}^\top \mathcal{M}\,\widehat{L} = \widehat{M}^{(1)} - \widehat{M}^{(2)}$, where the entries of $\widehat{M}^{(1)} = \widehat{L}^\top \widehat{L}$ and $\widehat{M}^{(2)} = \widehat{L}^\top \mathbf{1}_K\, \mathbf{1}_K^\top \widehat{L}/K$ are

$$\widehat{M}^{(1)}_{ii'}(\mathbf{v}) = \sum_{k=1}^K \mathbb{I}\left(\mathbf{G}_i^{(k)} \vee \mathbf{G}_{i'}^{(k)} \le \mathbf{v} \right)$$

and

$$\widehat{M}^{(2)}_{ii'}(\mathbf{v}) = \frac{1}{K} \sum_{k,k'=1}^{K} \mathbb{I}\left(\mathbf{G}_i^{(k)} \vee \mathbf{G}_{i'}^{(k')} \leq \mathbf{v}\right),$$

with $\mathbf{r} \vee \mathbf{r}'$ being the componentwise maximum of $\mathbf{r}, \mathbf{r}' \in \mathbb{R}^p$. From straightforward computations, the entries of $\Lambda = \Lambda_{\mathrm{CvM}} = \mathcal{F}_{\mathrm{CvM}}(\mathcal{M}\widehat{L})$ are

$$\Lambda_{ii'} = \sum_{k=1}^{K} \prod_{j=1}^{p} \left\{1 - \left(G_{ij}^{(k)} \vee G_{i'j}^{(k)}\right)\right\} - \frac{1}{K} \sum_{k,k'=1}^{K} \prod_{j=1}^{p} \left\{1 - \left(G_{ij}^{(k)} \vee G_{i'j}^{(k')}\right)\right\}.$$

For the diagonal section statistic, note that

$$\widehat{M}^{(1)}_{ii'}(v\,\mathbf{1}_p) = \sum_{k=1}^{K} \mathbb{I}\left(\max_{1\leq j\leq p} G_{ij}^{(k)} \vee G_{i'j}^{(k)} \leq v\right)$$

and

$$\widehat{M}^{(2)}_{ii'}(v\,\mathbf{1}_p) = \frac{1}{K} \sum_{k,k'=1}^{K} \mathbb{I}\left(\max_{1\leq j\leq p} G_{ij}^{(k)} \vee G_{i'j}^{(k')} \leq v\right).$$

Hence, the entries of $\Lambda = \Lambda_{\mathrm{Dia}} = \mathcal{F}_{\mathrm{Dia}}(\mathcal{M}\widehat{L})$ are

$$\Lambda_{ii'} = \frac{1}{K} \sum_{k,k'=1}^{K} \left(\max_{1\leq j\leq p} G_{ij}^{(k)} \vee G_{i'j}^{(k')}\right) - \sum_{k=1}^{K} \left(\max_{1\leq j\leq p} G_{ij}^{(k)} \vee G_{i'j}^{(k)}\right).$$

B.2. Characteristic function

Letting $\delta_{\mathbf{v}}(\mathbf{r})$ give mass one at $\mathbf{v} = \mathbf{r}$ and zero otherwise, one can write

$$\mathrm{d}\widehat{C}_{A^{(k)},\mathbf{g}^{(k)}}(\mathbf{v}) = \frac{1}{n} \sum_{i=1}^{n} \delta_{\mathbf{v}}\left(\mathbf{G}_i^{(k)}\right).$$

If one defines $\mathcal{D} \in \mathcal{S}_{Kn}$ such that $\mathcal{D}_{ki}(\mathbf{v}) = \delta_{\mathbf{v}}(\mathbf{G}_i^{(k)})$, then $\mathrm{d}\widehat{L} = \mathcal{D}$; it follows that $(\mathcal{M}\mathcal{D}(\mathbf{v}))^\top \mathcal{M}\mathcal{D}(\mathbf{u}) = \mathcal{D}(\mathbf{v})^\top \mathcal{M}\mathcal{D}(\mathbf{u}) = \widehat{M}^{(1)}(\mathbf{v},\mathbf{u}) - \widehat{M}^{(2)}(\mathbf{v},\mathbf{u})$, where

$$\widehat{M}^{(1)}_{ii'}(\mathbf{v},\mathbf{u}) = \sum_{k=1}^{K} \delta_{\mathbf{v}}(\mathbf{G}_i^{(k)})\, \delta_{\mathbf{u}}(\mathbf{G}_{i'}^{(k)})$$

and

$$\widehat{M}^{(2)}_{ii'}(\mathbf{v},\mathbf{u}) = \frac{1}{K} \sum_{k,k'=1}^{K} \delta_{\mathbf{v}}(\mathbf{G}_i^{(k)})\, \delta_{\mathbf{u}}(\mathbf{G}_{i'}^{(k')}).$$

As a consequence, one has in view of (13) that the entries of $\Lambda = \mathcal{F}^{\omega}_{\mathrm{Cf}}(\mathcal{M}\widehat{L})$ are

$$\Lambda_{ii'} = \sum_{k=1}^{K} \beta^{\omega}\left(\mathbf{G}_i^{(k)} - \mathbf{G}_{i'}^{(k)}\right) - \frac{1}{K} \sum_{k,k'=1}^{K} \beta^{\omega}\left(\mathbf{G}_i^{(k)} - \mathbf{G}_{i'}^{(k')}\right).$$

B.3. Linear statistics

From the linearity of η and the fact that $\mathcal{M}^\top\mathcal{M} = \mathcal{M}$, one obtains

$$\Lambda = \mathcal{F}(\mathcal{M}\widehat{L}) = \boldsymbol{\eta}(\widehat{L})^\top\mathcal{M}^\top\mathcal{M}\,\boldsymbol{\eta}(\widehat{L}) = \boldsymbol{\eta}(\widehat{L})^\top\mathcal{M}\,\boldsymbol{\eta}(\widehat{L}).$$

Since $\boldsymbol{\eta}(\widehat{L})_{ki} = \eta(\widehat{L}_{ki})$, one has respectively for the Spearman and Blomqvist functionals that

$$\boldsymbol{\eta}_{\mathrm{Sp}}(\widehat{L})_{ki} = \prod_{j=1}^{p}\left(1 - G_{ij}^{(k)}\right) \quad \text{and} \quad \boldsymbol{\eta}_{\mathrm{B}\ell}(\widehat{L})_{ki} = \mathbb{I}\left(\max_{1\le j\le p} G_{ij}^{(k)} \le \frac{1}{2}\right).$$

Appendix C: Formulas for the multiplier versions

C.1. Approximation of $\widehat{P}$

Exactly as in Appendix B, adopt the notation $\mathbf{G}_i^{(k)} = \mathbf{g}^{(k)}(\widehat{\mathbf{U}}_{i,A^{(k)}})$, so that the entries of the function matrix $\widehat{P} \in \mathcal{S}_{Kn}$ defined in equation (12) are given by

$$\widehat{P}_{ki}(\mathbf{v}) = \mathbb{I}\left(\mathbf{G}_i^{(k)} \le \mathbf{v}\right) - \sum_{j=1}^{p} \widehat{D}^{[j]}(\mathbf{v})\,\mathbb{I}\left(\mathbf{G}_{ij}^{(k)} \le v_j\right).$$

It will be convenient to approximate $\widehat{P}$ on a grid of $[0,1]^p$ by choosing $N \in \mathbb{N}$ large and by considering

$$P_N(\mathbf{v}) = \sum_{\mathbf{s}\in\mathcal{B}^N} \widehat{P}\left(\frac{\mathbf{s} - 1/2}{N}\right)\mathbb{I}_{\Omega_{\mathbf{s}}}(\mathbf{v}),$$

where $\mathcal{B}^N = \{1,\dots,N\}^p$ and

$$\Omega_{\mathbf{s}} = \bigotimes_{j=1}^{p}\left(\frac{s_j - 1}{N}, \frac{s_j}{N}\right].$$

Since $\mathcal{F}$ is assumed continuous, the approximation of $\widehat{\Lambda} = \mathcal{F}(\mathcal{M}\widehat{P})$ with $\mathcal{F}(\mathcal{M}P_N)$ is valid.

C.2. Cramér–von Mises and diagonal section

For the Cramér–von Mises and diagonal functionals, one obtains from straightforward computations using the fact that $\mathcal{M}^\top\mathcal{M} = \mathcal{M}$ that

$$\mathcal{F}_{\mathrm{CvM}}(\mathcal{M}P_N) \quad = \quad \sum_{\mathbf{s},\mathbf{s}'\in\mathcal{B}^N} \widehat{P}\left(\frac{\mathbf{s} - 1/2}{N}\right)^\top \mathcal{M}\,\widehat{P}\left(\frac{\mathbf{s}' - 1/2}{N}\right)\int_{\Omega_{\mathbf{s}}\cap\Omega_{\mathbf{s}'}} \mathrm{d}\mathbf{v}$$

$$= \sum_{\mathbf{s}\in\mathcal{B}^N} \widehat{P}\left(\frac{\mathbf{s}-1/2}{N}\right)^\top \mathcal{M}\,\widehat{P}\left(\frac{\mathbf{s}-1/2}{N}\right)\int_{\Omega_\mathbf{s}} \mathrm{d}\mathbf{v}$$
$$= \frac{1}{N^p}\sum_{\mathbf{s}\in\mathcal{B}^N} \widehat{P}\left(\frac{\mathbf{s}-1/2}{N}\right)^\top \mathcal{M}\,\widehat{P}\left(\frac{\mathbf{s}-1/2}{N}\right).$$

Similar computations yield

$$\mathcal{F}_{\mathrm{Dia}}\left(\mathcal{M}P_N\right) = \frac{1}{N}\sum_{s=1}^{N} \widehat{P}\left\{\left(\frac{s-1/2}{N}\right)\mathbf{1}_p\right\}^\top \mathcal{M}\,\widehat{P}\left\{\left(\frac{s-1/2}{N}\right)\mathbf{1}_p\right\}.$$

C.3. Characteristic function

One can exploit the fact that $S(\mathbf{v}) = \mathcal{M}P_N(\mathbf{v})$ vanishes whenever $v_j = 0$ for some $j \in \{1,\ldots,p\}$, or when at least $p-1$ components of $\mathbf{v}$ are equal to one. In that case,

$$S([\mathbf{v}^{\mathcal{I}},\mathbf{1}]) = \sum_{\mathcal{K}\subseteq\mathcal{I}:|\mathcal{K}|>1} (-1)^{|\mathcal{K}|} S(\widetilde{\mathbf{v}}^{\mathcal{K}}),$$

where $\mathbf{v}_j^{\mathcal{K}} = v_j$ when $j\in\mathcal{K}$ and $v_j = 1$ otherwise. Plugging it into formula (14) and after some computations, one obtains

$$\mathcal{F}_{\mathrm{Cf}}^{\omega}(S) = \int_{[0,1]^{2p}} \sum_{\mathcal{K}\subseteq S_p:|\mathcal{K}|>1}\ \sum_{\mathcal{L}\subseteq S_p:|\mathcal{L}|>1} S(\widetilde{\mathbf{v}}^{\mathcal{K}})^\top S(\widetilde{\mathbf{u}}^{\mathcal{L}})\,\beta^{\omega}_{\mathcal{K},\mathcal{L}}(\mathbf{v},\mathbf{u})\,\mathrm{d}\mathbf{v}\,\mathrm{d}\mathbf{u},$$

where

$$\beta^{\omega}_{\mathcal{K},\mathcal{L}}(\mathbf{v},\mathbf{u}) = \sum_{\mathcal{I}:\mathcal{K}\subseteq\mathcal{I}}\ \sum_{\mathcal{J}:\mathcal{L}\subseteq\mathcal{J}} \frac{\partial}{\partial\mathbf{v}^{\mathcal{I}}}\frac{\partial}{\partial\mathbf{u}^{\mathcal{J}}}\,\beta^{\omega}\left(\mathbf{v}^{\mathcal{I}}-\mathbf{u}^{\mathcal{J}}\right).$$

When $p=2$, one can show that the formula simplifies to

$$\mathcal{F}_{\mathrm{Cf}}^{\omega}(S) = \int_{[0,1]^4} S(v_1,v_2)^\top S(u_1,u_2)\,\beta^{[2,2]}(\mathbf{v}-\mathbf{u})\,\mathrm{d}\mathbf{v}\,\mathrm{d}\mathbf{u},$$

where $\beta^{[2,2]}(r_1,r_2) = \partial^2\,\partial^2\,\beta^{\omega}(r_1,r_2)/\partial r_1^2\,\partial r_2^2$. One then has

$$\mathcal{F}_{\mathrm{Cf}}^{\omega}(\mathcal{M}P_N) = \frac{1}{N^4}\sum_{\mathbf{s},\mathbf{s}'\in\mathcal{B}^N} \widehat{P}\left(\frac{\mathbf{s}-1/2}{N}\right)^\top \mathcal{M}\,\widehat{P}\left(\frac{\mathbf{s}'-1/2}{N}\right)\beta^{[2,2]}\left(\frac{\mathbf{s}-\mathbf{s}'}{N}\right).$$

The formulas are explicit but cumbersome for $p>2$.

C.4. Linear statistics

The linearity of $\boldsymbol{\eta}$ and $\mathcal{M}^\top\mathcal{M} = \mathcal{M}$ entails $\mathcal{F}_{\boldsymbol{\eta}}(\mathcal{M}P_N) = \boldsymbol{\eta}\,(P_N)^\top\,\mathcal{M}\,\boldsymbol{\eta}\,(P_N)$, where

$$\boldsymbol{\eta}\,(P_N) = \sum_{\mathbf{s}\in\mathcal{B}^N} \widehat{P}\left(\frac{\mathbf{s}-1/2}{N}\right)\eta\left\{\mathbb{I}_{\Omega_\mathbf{s}}(\mathbf{v})\right\}.$$

Since

$$\mathbb{I}_{\Omega_{\mathbf{s}}}(\mathbf{v}) = \prod_{j=1}^{p} \mathbb{I}\left(\frac{s_j - 1}{N} < v_j \leq \frac{s_j}{N}\right),$$

one has respectively for the Spearman and Blomqvist functionals, $\eta_{\mathrm{Sp}}\{\mathbb{I}_{\Omega_{\mathbf{s}}}(\mathbf{v})\} = N^{-p}$ and

$$\eta_{\mathrm{B}\ell}\left\{\mathbb{I}_{\Omega_{\mathbf{s}}}(\mathbf{v})\right\} = \prod_{j=1}^{p} \mathbb{I}\left\{s_j \in \left[\frac{N}{2}, \frac{N}{2}+1\right)\right\} = \prod_{j=1}^{p} \mathbb{I}\left(s_j = \lfloor (N+1)/2 \rfloor\right).$$

As a consequence,

$$\boldsymbol{\eta}_{\mathrm{Sp}}(P_N) = \frac{1}{N^p} \sum_{\mathbf{s}\in\mathcal{B}^N} \widehat{P}\left(\frac{\mathbf{s}-1/2}{N}\right),$$

and because $\{\lfloor (N+1)/2 \rfloor - 1/2\}/N \approx 1/2$, one has $\boldsymbol{\eta}_{\mathrm{B}\ell}(P_N) \approx \widehat{P}(\mathbf{1}_p/2)$.

Acknowledgements

Thanks are due to the Editor, an Associate Editor and three referees for their careful reading of the manuscript and their valuable comments that led to an improvement of this paper. This research was supported in part by an individual grant from the Natural Sciences and Engineering Research Council of Canada (NSERC) and by the Canadian Statistical Science Institute (CANSSI).

References

[1] Blomqvist, N. (1950). On a measure of dependence between two random variables. *Ann. Math. Statistics* **21**, 593–600. MR0039190

[2] Bücher, A. & Dette, H. (2010). A note on bootstrap approximations for the empirical copula process. *Statistics and Probability Letters* **80**, 1925–1932. MR2734261

[3] Bücher, A., Dette, H. & Volgushev, S. (2012). A test for Archimedeanity in bivariate copula models. *J. Multivariate Anal.* **110**, 121–132. MR2927513

[4] Bücher, A. & Ruppert, M. (2013). Consistent testing for a constant copula under strong mixing based on the tapered block multiplier technique. *J. Multivariate Anal.* **116**, 208–229. MR3049901

[5] Bücher, A. & Volgushev, S. (2013). Empirical and sequential empirical copula processes under serial dependence. *J. Multivariate Anal.* **119**, 61–70. MR3061415

[6] Burke, M. D. (2000). Multivariate tests-of-fit and uniform confidence bands using a weighted bootstrap. *Statist. Probab. Lett.* **46**, 13–20. MR1730878

[7] Cherubini, U., Luciano, E. & Vecchiato, W. (2004). *Copula methods in finance.* Chichester: Wiley. MR2250804

[8] Fermanian, J.-D. (1998). *Contributions à l'analyse nonparamétrique des fonctions de hasard sur données multivariées et censurées*. Ph.D. thesis, Paris 6.

[9] Gaisser, S. & Schmid, F. (2010). On testing equality of pairwise rank correlations in a multivariate random vector. *J. Multivariate Anal.* **101**, 2598–2615. MR2719884

[10] Genest, C. & Nešlehová, J. G. (2013). On tests of radial symmetry for bivariate copulas. *Statistical Papers* , 1–13.

[11] Genest, C., Nešlehová, J. & Quessy, J.-F. (2012). Tests of symmetry for bivariate copulas. *Ann. Inst. Statist. Math.* **64**, 811–834. MR2927772

[12] Jaworski, P. (2010). Testing Archimedeanity. In *Combining Soft Computing and Statistical Methods in Data Analysis*. Springer, pp. 353–360.

[13] Joe, H. (1997). *Multivariate models and dependence concepts*, vol. 73 of *Monographs on Statistics and Applied Probability*. London: Chapman & Hall. MR1462613

[14] Khoudraji, A. (1995). *Contributions à l'étude des copules et à la modélisation de valeurs extrêmes bivariées*. Ph.D. thesis, Université Laval, Québec, Canada.

[15] Kojadinovic, I. & Yan, J. (2010). Nonparametric rank-based tests of bivariate extreme-value dependence. *J. Multivariate Anal.* **101**, 2234–2249. MR2671214

[16] Li, B. & Genton, M. G. (2013). Nonparametric identification of copula structures. *J. Amer. Statist. Assoc.* **108**, 666–675. MR3174650

[17] Nelsen, R. B. (1996). Nonparametric measures of multivariate association. In *Distributions with fixed marginals and related topics (Seattle, WA, 1993)*, vol. 28 of *IMS Lecture Notes Monogr. Ser.* Hayward, CA: Inst. Math. Statist., pp. 223–232. MR1485534

[18] Nelsen, R. B. (2006). *An introduction to copulas*. Springer Series in Statistics. New York: Springer, 2nd ed. MR2197664

[19] Quessy, J.-F. (2012). Testing for bivariate extreme dependence using Kendall's process. *Scandinavian Journal of Statistics* **39**, 398–417. MR2971634

[20] Quessy, J.-F. & Bahraoui, T. (2013). Graphical and formal statistical tools for the symmetry of bivariate copulas. *Canad. J. Statist.* **41**, 637–656. MR3146002

[21] Rémillard, B. & Scaillet, O. (2009). Testing for equality between two copulas. *J. Multivariate Anal.* **100**, 377–386. MR2483426

[22] Rüschendorf, L. (1976). Asymptotic distributions of multivariate rank order statistics. *Ann. Statist.* **4**, 912–923. MR0420794

[23] Scheffé, H. (1959). *The analysis of variance*. New York: John Wiley & Sons Inc. MR0116429

[24] Schmid, F., Schmidt, R., Blumentritt, T., Gaisser, S. & Ruppert, M. (2010). Copula-based measures of multivariate association. In *Copula theory and its applications*, vol. 198 of *Lect. Notes Stat. Proc.* Heidelberg: Springer, pp. 209–236. MR3051270

[25] SEGERS, J. (2012). Weak convergence of empirical copula processes under nonrestrictive smoothness assumptions. *Bernoulli* **18**, 764–782. MR2948900

[26] VAN DER VAART, A. W. & WELLNER, J. A. (1996). *Weak convergence and empirical processes*. Springer Series in Statistics. New York: Springer-Verlag. With applications to statistics. MR1385671