Sankhyā A : The Indian Journal of Statistics
https://doi.org/10.1007/s13171-019-00197-w
© 2020, Indian Statistical Institute

New Measure of the Bivariate Asymmetry

Tarik Bahraoui and Nikolai Kolev
University of São Paulo, São Paulo, Brazil

Abstract

A new measure of the bivariate asymmetry of a dependence structure between two random variables is introduced based on copula characteristic function. The proposed measure is represented as the discrepancy between the rank–based distance correlation computed over two complementary order-preserved sets. Generalproperties of the measure are established, as well as an explicit expression for the empirical version. It is shown that the proposed measure is asymptotically equivalent to a fourth–order degenerate V-statistics and that the limit distributions have representations in terms of weighted sums of an independent chi-square random variables. Under dependent random variables, the asymptotic behavior of bivariate distance covariance and variance process is demonstrated. Numerical examples illustrate the properties of the measures.

AMS (2000) *subject classification.* Primary 62H20, 62G30; Secondary 60E10.
Keywords and phrases. Copula characteristic function, Degenerate V-statistics, Distance correlation, Rank statistics, Dependence measure, Asymmetry

1 Introduction

Over the past twenty years, modeling of dependence among random variables using copulas has been developed rapidly. In parameter estimation and copula model selection, important progress has been made. For example, Dobrić and Schmid (2005) investigated a test of goodness-of-fit for parametric families of copulas with application to financial data set, see also Genest and Segers (2009) and Genest et al. (2011) for bivariate extreme-value copulas; testing the independence assumption based on the empirical copula process proposed by Genest and Rémillard (2004) and Kojadinovic and Yan (2011); testing for equality in law of dependent random variables, consult (Rémillard and Scaillet, 2009).

Otherwise, since there is a bijection between characteristic function and bivariate distribution, through its copula, one could alternatively work with the characteristic function associated with a copula. For a historical review of characteristic function we refer to Lukacs (1970) and Ushakov (1999). Some review methods for testing certain hypothesis based on the empirical characteristic function have been investigated, for example, goodness-of-fit tests procedures by Fan (1997); Meintanis (2004); Székely and Rizzo (2005) and Meintanis et al. (2016); for the problem of testing the symmetry see e.g., Feuerverger and Mureika (1977) and Henze et al. (2003); testing the concepts of independence consult Csörgő (1985); Kankainen and Ushakov (1998); Székely et al. (2007); Meintanis and Iliopoulos (2008); Fan et al. (2017) and the references therein. More recently, characteristic function associated to copula models have been investigated by Bahraoui et al. (2018) for goodness-of-fit testing in multivariate copula models, and used it for testing the structural assumptions, radial symmetry (Bahraoui and Quessy, 2017) and symmetry (Bahraoui et al., 2019).

In the current paper, we focus on the bivariate asymmetry measure (radial and diagonal symmetry), an important tool in many fields, e.g., in finance (Zhang and Shinki, 2007), insurance (Ang and Chen, 2002), and environmental science (Wang, 2016) and Salvadori et al. (2007). The literature review in measures of copula asymmetry has not been sufficiently set out, one of most classical measure is Skewness, see, e.g., Bücher et al. (2017) and others measures are based mainly on moments, Quantiles or treating upper and lower tail-weighted asymmetry. For a deep discussion on these measures, consult (Joe, 2015), Section 2.15, Rosco and Joe (2013); Krupskii and Joe (2015); Krupskii (2017) and for a very recent review see, e.g., Lee et al. (2018).

We suggest a new measure of asymmetry for bivariate copulas based on characteristic function in set $\gamma=\{\gamma^-,\gamma^+\}$, where γ^- and γ^+ will be called a complementary order-preserved sets $\gamma^+ = \{(t_1,t_2) : t_1 > t_2\}$ and $\gamma^- = \{(t_1, t_2) : t_1 < t_2\}$. The set γ is collection of sub-sets γ^+ and γ^- of $\mathbb{R}^2$. We show that the proposed measure is easily interpretable tool to detect a association between two variables, it belongs to the interval $[0,1]$ with lower bound zero characterize independence. In fact, the projection of the family of all two–dimensional copulas, namely $\mathscr{C}$ onto a class $\mathscr{B}$, where $\mathscr{B} \subseteq \mathbb{R}^2$ is the class of all bivariate copula characteristic function, provides many advantages:

(a) we overcome the problem of estimation of a partial derivative of the copulas;

(b) from an empirical point of view, the simple form of computation of the new measure;

(c) the measure discriminates well among the different copula models according to their tail behavior.

The paper is organized as follows. Section 2 contains a definition and some proprieties of measure of asymmetry. Section 3 is devoted to the empirical version of distance correlation/covariance. The asymptotic behavior of empirical bivariate distance correlation is investigated in Section 4 under independent and dependent random variables. Section 5 introduces a graphical interpretation of CCF in the set $\gamma = \{\gamma^-, \gamma^+\}$ with an estimation of the proposed measure of asymmetry for different bivariate copula models. A small application of a new measure of asymmetry to real-life data of a moderate sample size are presented in Section 6, and we conclude an additional discussion and possible future developments in Section 7. All the proofs are to be found in the Appendix.

2 A New Measure of Bivariate Asymmetry and its Properties

A new bivariate asymmetry measure is introduced in this section. We establish necessary and sufficient conditions that the asymmetry measure should fulfill and natural property of distance correlation. Then, some properties of dependence ordering and Fréchet-Hoeffding bound are derived in Subsection 2.2.

2.1. Definition and Proprieties

2.1.1. Measure of Bivariate Asymmetry. Let (X_1, X_2) be a pair of random variables with the joint distribution function $H(x_1, x_2) = \mathbb{P}(X_1 \leq x_1, X_2 \leq x_2)$, where $(x_1, x_2) \in \mathbb{R}^2$, $\mathbb{R} = (-\infty, +\infty)$ denote the ordinary real line. If the marginal distributions

$$F_1(x) = \mathbb{P}(X_1 \leq x_1) \quad \text{and} \quad F_2(x) = \mathbb{P}(X_2 \leq x_2)$$

are continuous, Sklar's Theorem ensures the existence of unique copula $C : [0,1]^2 \to [0,1]$ such that

$$H(x_1, x_2) = C\left\{F_1(x_1), F_2(x_2)\right\},$$

for each $(x_1, x_2) \in \mathbb{R}^2$. It is, in fact, the joint distribution of $\mathbf{U} = (U_1, U_2) = (F_1(X_1), F_2(X_2))$. Then, the copula is a distribution function on $[0,1]^2$ with uniform $\mathscr{U}(0,1)$ margins. For more details on the theory of copulas and their applications, we refer to Joe (2015); McNeil et al. (2015) and Nelsen (2006).

Let now Ψ_C be the bivariate copula characteristic function (CCF hereafter) of a copula $C \in \mathscr{C}$, defined for $i^2 = -1$ and $\mathbf{U} \sim C$ by

$$\Psi_C(\mathbf{t}) = \mathrm{E}\left\{e^{i(t_1 U_1 + t_2 U_2)}\right\} = \int_{[0,1]^2} e^{i(t_1 u_1 + t_2 u_2)} \mathrm{d}C(u_1, u_2), \quad \mathbf{t} = (t_1, t_2) \in \mathbb{R}^2,$$

Using the $\|.\|_w$-norm in the weighted L_2 space of functions, conditionally on set γ, we define a measure of dependence as quadratic distance between the empirical joint copula characteristic function and the product of its margins

$$\left\{\Lambda_\omega^{\gamma^+}, \Lambda_\omega^{\gamma^-}\right\}(C) = \Lambda_\omega^\gamma(C) = \int_\gamma |\Psi_C(\mathbf{t}) - \psi_1(t_1)\psi_2(t_2)|^2 \, \omega(\mathbf{t}) \mathrm{d}\,\mathbf{t}, \qquad (1)$$

where $w(\mathbf{t})$ is an arbitrary positive weight function for which the integral above exists and $\psi_j(t_j) = \mathrm{E}\left\{e^{i\,t_j\,U_j}\right\}$, $j = 1, 2$, denotes the marginal CCF.

As suggested by Székely et al. (2007), the specific choice $\omega(\mathbf{t})$=$(\pi|t_1||t_2|)^{-2}$ ensures that $\Lambda_\omega(C) = 0$ if and only if U_1 and U_2 are independent and where for $(U_1, U_2) \in \mathbb{R}^2$

$$\Lambda_\omega(C) = \int_{\mathbb{R}^2} |\Psi_C(t_1, t_2) - \psi_1(t_1)\psi_2(t_2)|^2 \, \omega_0(t_1, t_2)\, \mathrm{d}\, t_1 \, \mathrm{d}\, t_2.$$

We will consider further the measure $\Lambda_{\omega_0}(C)$, which can be interpreted as a distance covariance between U_1 and U_2 in arbitrary dimension by

$$\left\{\Lambda_{\omega_0}^{\gamma^+}, \Lambda_{\omega_0}^{\gamma^-}\right\}(C) = \Lambda_{\omega_0}^\gamma(C) = \int_\gamma |\Psi_C(\mathbf{t}) - \psi_1(t_1)\psi_2(t_2)|^2 \, \omega_0(\mathbf{t})\, \mathrm{d}\,\mathbf{t}, \qquad (2)$$

with non–integrable symmetry weighted function $\omega_0(\mathbf{t}) = (\pi|t_1||t_2|)^{-2}$ and where γ is collection of sub–sets γ^+ and γ^- of $\mathbb{R}^2$. Then, one can define the distance correlation by

$$R_{\omega_0}^\gamma(\mathbf{U}) = \Lambda_{\omega_0}^\gamma(C) \left\{\sigma_{\omega_0,\ell}^\gamma\right\}^{-\frac{1}{2}} \mathbb{I}\left(\sigma_{\omega_0,\ell}^\gamma > 0\right), \qquad (3)$$

where $\sigma_{\omega_0,\ell}^\gamma = \Lambda_{\omega_0,1}^\gamma \, \Lambda_{\omega_0,2}^\gamma$ is distance variance of $\mathbf{U}_\ell = (U_\ell, U_\ell) \sim C$, $\ell = 1, 2$.

Definition 1. A new measure of bivariate asymmetry is the mapping $\Delta_{\omega_0}(\Psi_C) : \mathscr{B} \to \mathbb{R}_+$ defined as follows

$$\Delta_{\omega_0}(\Psi_C) = \left|R_{\omega_0}^+(\mathbf{U}) - R_{\omega_0}^-(\mathbf{U})\right|, \qquad (4)$$

where the distance correlation is given in Eq. 3.

It is desirable to check the identifiability properties that the measure of bivariate asymmetry in a dependence structure should satisfy.

Lemma 1. *The measure $\Delta_{\omega_0} : \mathscr{B} \to \mathbb{R}_+$ of bivariate asymmetry, satisfies the following proprieties :*

(a) There exists $\varepsilon \in \mathbb{R}_+$ such that, $\Delta_{\omega_0}(\Psi_C) \leq \varepsilon$, for all $C \in \mathscr{C}$;

(b) $\Delta_{\omega_0}(\Psi_C) = 0$ if C is radially symmetric;

(c) $\Delta_{\omega_0}(\Psi_C) = \Delta_{\omega_0}(\Psi_S)$, *for any* C;

(d) If Ψ_n *converges uniformly to* Ψ_C, *as* $n \to \infty$ *and the fact that* $\int_\gamma \omega_0(\mathbf{t})\mathrm{d}\,\mathbf{t} < \infty$, *one has* $\Delta_{\omega_0}(\Psi_n) \to \Delta_{\omega_0}(\Psi_C)$ *(a.s), where for* $\mathbf{U} \sim C$, $S\,(u_1, u_2) = u_1 + u_2 - 1 + C\,(1 - u_1, 1 - u_1)$ *is a survival copula of* $C \in \mathscr{C}$ *and* Ψ_n *is empirical counterpart of* Ψ_C.

An overview description of these assumptions for bivariate copulas can be found in Durante et al. (2010); Dehgani et al. (2013) and for weaker version consult (Rosco and Joe, 2013).

2.1.2. Proprieties of Distance Correlation. As shown by Schweizer and Wolff (1981), several copula models based on dependence measures between two continuous random variables satisfy the axioms of Rényi (1959). Since the distance correlation defined in Eq. 10, is rank-based version of distance correlation $R^{\gamma}_{\omega_0}(\mathbf{U})$, all the axioms of Rényi (1959) are verified, specifically, the axiom of monotonic invariance. Likewise, distance correlation specified by Eq. 3 has properties of a true dependence measure and satisfies $R^{\gamma}_{\omega_0}(\mathbf{U}) \in [0, 1]$, $R^{\gamma}_{\omega_0}(\mathbf{U}) = 0$ characterizes independence between random variables U_1 and U_2, with marginal uniform distribution.

2.1.3. Diagonal Section Versus Distance Variance. For a given copula C, let $\delta_C : [0, 1] \to [0, 1]$ be the diagonal section of copula C. The function δ_C satisfies the following proprieties: $\delta_C(1) = 1$, $\delta_C(u) < u$, for all $u \in [0, 1]$ and $|\delta_C(u_1) - \delta_C(u_2)| < 2\,|u_1 - u_2|$, for all $(u_1, u_2) \in [0, 1]^2$. For more details about the diagonal section of the copula C, consult Alsina et al. (2006); Fredricks and Nelsen (1997) and Fernández-Sánchez and Úbeda Flores (2018).

Then, the distance variance can be written, for each $\mathbf{U}_\ell \sim C$, $\ell = 1, 2$ as

$$D^{\gamma}_{\delta,\omega_0,\ell} = \int_\gamma |\Psi_\delta(t_1 + t_2) - \psi_\delta(t_1)\,\psi_\delta(t_2)|^2\,\omega_0(\mathbf{t})\,\mathrm{d}\mathbf{t}, \tag{5}$$

where

$$\Psi_\delta(\mathbf{t}) = \Psi_\delta(t_1 + t_2) = \int_{[0,1]} e^{i\,(t_1+t_2)\,u}\mathrm{d}\delta_C(u).$$

Hence, if U_1 and U_2 are independent i.e., $C(u_1, u_2) = u_1\,u_2$, for all $(u_1, u_2) \in [0, 1]^2$, the diagonal section is then given by $\delta_C(u) = C(u, u) = u^2$. Then, after simple algebraic manipulations, one obtains the diagonal (CCF)

$$\Psi_\delta(t_1 + t_2) = \frac{2\,(\mathrm{e}^{\mathrm{i}(\mathrm{t}_1+\mathrm{t}_2)} - 1)}{i\,(t_1 + t_2)},$$

and the margins ψ_ℓ is a characteristic function of uniform distribution $\mathscr{U}(0, 1)$, that is equal to 1 if $t_\ell = 0$ and $\left(e^{i\,t_\ell} - 1\right)/it_\ell$, for each $\ell \in \{1, 2\}$.

As a consequence, $D^{\gamma}_{\delta,\omega_0,\ell}$ is concentrated to the diagonal section, depending only on (t_1, t_2) and bounded i.e., $D^{\gamma}_{\delta,\omega_0,\ell} < \infty$, for any $\ell = 1, 2$.

Moreover, the measure of bivariate asymmetry can be written as

$$\Delta_{\omega_0}(\Psi_C) = \sigma^{-1}_{\omega_0,\ell} \left| \Lambda^{+}_{\omega_0}(C) - \Lambda^{-}_{\omega_0}(C) \right|,$$

where $\Lambda^{\gamma}_{\omega_0}$ are given in Eq. 1. Thence, the asymptotic behavior of the measure of asymmetry relies on the asymptotic behavior of weighted distance covariance.

2.2. Stochastic Orders and Fréchet-Hoeffding Bound In this Subsection, an additional property of stochastic ordering is provided, the reader is referred to the monograph of Shaked and Shanthikumar (2007) for further details on this subject.

Let ψ_1 and ψ_2 be the characteristic function associated with C_1 and C_2, respectively. Hence, if a bivariate copula C_1 is stochastically dominated by another copula C_2 i.e., $C_1(u_1, u_2) \prec_{st} C_1(u_1, u_2)$ for all u_1, u_2 in $[0, 1]$ implies that the corresponding characteristic function also satisfies stochastically ordering $\psi_1(t_1, t_2) \prec_{st} \psi_2(t_1, t_2)$ in complementary order-preserved sets γ^+ and γ^-. A direct consequence of this property is the Fréchet-Hoeffding bound. We will present in the following Lemma, an important interpretation of the copula characteristic function associated to lower Fréchet-Hoeffding bound $W(u_1, u_2) = \max(u_1 + u_2 - 1, 0)$ and upper Fréchet-Hoeffding bound $M(u_1, u_2) = \min(u_1, u_2)$, for any u_1 and u_2 in $[0, 1]$.

Lemma 2. *For any copula $C \in \mathscr{C}$ and $\mathbf{t} = (t_1, t_2) \in \mathbb{R}^2$, one has*

$$\Psi_W(\mathbf{t}) \leq \Psi_C(\mathbf{t}) \leq \Psi_M(\mathbf{t}),$$

where the lower bound

$$\Psi_W(\mathbf{t}) = \frac{1}{t_1 t_2}\left(e^{it_1} - e^{i(t_1+t_2)}\right) - \frac{1}{t_2 - t_1}\left(e^{it_2} - e^{it_1}\right)$$

and the upper bound is product of the characteristic function of uniform distribution ψ_ℓ, for $\ell = 1, 2$,

$$\Psi_M(\mathbf{t}) = \psi_1(t_1)\,\psi_2(t_2) = \frac{\left(1 - e^{it_1}\right)\,\left(e^{it_2} - 1\right)}{t_1\, t_2}.$$

One can observe that the Fréchet-Hoeffding bounds of copulas characteristic function do not depend on arguments $u_1, u_2 \in [0, 1]$.

3 The Empirical Version

Here we will construct an empirical version of measure of asymmetry given by Eq. 4 based on ranks of observations and we give an explicit formula for computation the distance covariance. First, let $(X_{11}, X_{12}) \ldots, (X_{n1}, X_{n2})$ be independent copies of a random vector $\mathbf{X} = (X_1, X_2)$ with a copula C, whose marginal distributions F_1 and F_2 are continuous. In most of the applications, the marginal distributions are unknown, so that the uniform random vector $\mathbf{U} = (U_1, U_2) = (F_1(X_1), F_1(X_1))$ is unobservable. One can then work with the pairs of pseudo–observations $\widehat{\mathbf{U}}_1, \ldots, \widehat{\mathbf{U}}_n$ of $\mathbf{U} \sim C$, where $\widehat{\mathbf{U}}_j = (\widehat{U}_{j1}, \widehat{U}_{j2})$, with elements

$$\widehat{U}_{j\ell} = \frac{1}{n+1} \sum_{k=1}^{n} \mathbb{I}\left(X_{k\ell} \le X_{j\ell}\right), \quad \ell = 1, 2. \tag{6}$$

A natural empirical version of distance covariance $\varLambda^{\gamma}_{\omega_0}$ is given by

$$\widehat{\varLambda}^{\gamma}_{n,\omega_0} = \int_{\gamma} \left|\varPsi_n(\mathbf{t}) - \psi_{n1}(t_1)\, \psi_{n2}(t_2)\right|^2 \omega_0(\mathbf{t})\, \mathrm{d}\mathbf{t} \tag{7}$$

where $\varPsi_n$ is the empirical CCF specified by

$$\varPsi_n(\mathbf{t}) = \frac{1}{n} \sum_{j=1}^{n} \exp\left(i\, t_1\, \widehat{U}_{j1} + i\, t_2\, \widehat{U}_{j2}\right), \tag{8}$$

and the empirical marginal (CCF) are given by $\psi_{n1}(t_1) = \varPsi_n(t_1, 0)$ and $\psi_{n1}(t_2) = \varPsi_n(0, t_2)$. Hence following (4), the empirical version measure of bivariate asymmetry is defined as follows

$$\widehat{\varDelta}_{\omega_0}(\varPsi_n) = \left| R^{+}_{n,\omega_0}(\widehat{\mathbf{U}}_j) - R^{-}_{n,\omega_0}(\widehat{\mathbf{U}}_j) \right|, \tag{9}$$

where R^{γ}_{n} is the empirical distance correlation given by

$$R^{\gamma}_{n}(\widehat{\mathbf{U}}_j) = \widehat{\varLambda}^{\gamma}_{n,\omega_0} \left\{\widehat{\sigma}^{\gamma}_{n,\omega_0,\ell}\right\}^{-\frac{1}{2}} \mathbb{I}\left(\widehat{\sigma}^{\gamma}_{n,\omega_0,\ell} > 0\right). \tag{10}$$

Notice that $\widehat{\sigma}^{\gamma}_{n,\omega_0,\ell}$ is a product of the empirical bivariate distance variances i.e., $\widehat{\sigma}^{\gamma}_{n,\omega_0} = \widehat{\varLambda}^{\gamma}_{n,\omega_0,1}\, \widehat{\varLambda}^{\gamma}_{n,\omega_0,2}$, for each $\gamma = \{\gamma^+, \gamma^-\} \subset \mathbb{R}^2$.

In order to set an explicit formula for computation of distance covariance, let μ_{ω_0} be a measure defined by

$$\mu_{\omega_0}(\mathbf{a}, \mathbf{b}) = \int_{\gamma} \left(1 - \cos(\mathbf{a}\, t_1)\right) \left(1 - \cos(\mathbf{b}\, t_2)\right) \omega_0(\mathbf{t})\, \mathrm{d}\mathbf{t}. \tag{11}$$

for any $(\mathbf{a},\mathbf{b}) \in [0,1]^2$. By a tedious, but straightforward algebraic computation, we get

$$\mu_{\omega_0}(\mathbf{a},\mathbf{b}) = \frac{3}{2}\left(|\mathbf{a}|^2 + |\mathbf{b}|^2\right) + |\mathbf{a}|\,|\mathbf{b}|.$$

In the next Lemma, we obtain an explicit expression to compute the empirical distance covariance.

Lemma 3. *The empirical distance covariance $\Lambda^{\gamma}_{n,\omega_0}$ can be computed as follows*

$$\begin{aligned}
\widehat{\Lambda}^{\gamma}_{n,\omega_0} &= \frac{1}{n^2}\sum_{j,k=1}^{n}\mu_{\omega_0}\left(\widehat{U}_{j1}-\widehat{U}_{k1},\widehat{U}_{j2}-\widehat{U}_{k2}\right)\\
&\quad+\frac{1}{n^4}\sum_{j,k,s,l=1}^{n}\mu_{\omega_0}\left(\widehat{U}_{j1}-\widehat{U}_{k1},\widehat{U}_{s2}-\widehat{U}_{l2}\right)\\
&\quad-\frac{2}{n^3}\sum_{j,k,s=1}^{n}\mu_{\omega_0}\left(\widehat{U}_{j1}-\widehat{U}_{k1},\widehat{U}_{j2}-\widehat{U}_{s2}\right),
\end{aligned}$$

where $\widehat{U}_{j\ell}$, $\ell = 1,2$ are given in Eq. 6.

Remark 1. The evaluation of a measure $\mu_{\omega_0}(x,y)$ is different than those obtained in Lemma 1 of Székely et al. (2007), where for all $x \in \mathbb{R}$, one has

$$\int_{\mathbb{R}} \frac{1-\cos(s\,x)}{|s|^2}\,\mathrm{d}\,s = \pi\,|x|\,.$$

On the other hand, it is clear that the weighted function ω_0 is antisymmetric, that is, satisfies

$$\omega_0(\mathbf{t}) = \omega_0(-\mathbf{t}), \qquad \text{for all} \quad \mathbf{t}=(t_1,t_2)\in\mathbb{R}^2. \tag{12}$$

However, if the weight function ω is asymmetric, one can define the weight function $\omega^\star$ by

$$\omega^\star(\mathbf{t}) = \left(\omega(\mathbf{t})+\omega(-\mathbf{t})\right)/2,$$

wish satisfies Eq. 12, thus $\widehat{\Lambda}^{\gamma}_{n,\omega_0} = \widehat{\Lambda}^{\gamma}_{n,\omega^\star}$.

4 Asymptotic Behavior of Rank Degenerate V Statistic

In terms of stochastic processes approach, the technicalities need to derive the large sample distribution of the empirical copulas characteristic function process

$$\mathscr{Z}_n(\mathbf{t}) = \sqrt{n}\left\{\Psi_n(\mathbf{t}) - \psi_{n,1}(t_1)\psi_{n,2}(t_2)\right\}, \quad \text{for all } \mathbf{t}=(t_1,t_2)\in\mathbb{R}^2. \tag{13}$$

It seems to be a non–trivial problem in a complex-valued space. In this section, the asymptotic behavior of distance covariance $\Lambda^{\gamma}_{n,\omega}$ given in Eq. 7 is handled by theory of U and V-statistics in the context of ranks statistics. We refer the readers to the books of Lee (1990) and Koroljuk and Borovskich (1994). In the first step, we show that if the two random variables U_1 and U_2 are independent i.e., $\Psi_C(t_1,t_2) = \psi_1(t_1)\psi_2(t_2)$, the empirical distance covariance is represented as fourth order degenerate V-statistic, and its limit is weighted sums of independent chi-square variables.

Henceforth, we assume the following condition

$\mathscr{A}_1$. The weight function ω is such that

$$\begin{aligned} \omega(\mathbf{t}) &= \omega(-\mathbf{t}), \quad \text{and} \quad 0 < \int_{\gamma} \left(|t_1| + |t_2|\right)^4 \omega(\mathbf{t})\, \mathrm{d}\mathbf{t} < \infty, \\ \forall \mathbf{t} &= (t_1, t_2) \in \gamma. \end{aligned}$$

4.1. Degenerate V-Statistic Representation Among various measures of dependence between a of random variables (U_1, U_2) and a class of rank test procedures, can be written as the following linear form

$$\mathscr{M}_n = \frac{1}{n} \sum_{i_1=1}^{n} \dots \sum_{i_d=1}^{n} \varphi_n\left(R_{i_1}, \dots, R_{i_d}\right), \tag{14}$$

where φ_n is a symmetric kernel and R_i is the rank of X_i amongst $X_1, \dots, X_n$.

For example when the kernel φ_n is a Lipschitz function, one can extract from the representation (14) *Gini's mean difference*, *Wilcoxon's signed–rank test*, *Kendall rank correlation coefficient* or its projection into the family of linear rank statistic, the so-called *Spearman rank correlation.* For more details of the theory of rank statistics, see the excellent book of Hájek et al. (1999). An important step of representation (14) is that one can derive the asymptotic behavior of $\mathscr{M}_n$ and apply a multiplier bootstrap U and V-statistics technique to inference statistics. However, in many cases φ_n is simple to manipulate and it is robust concerning non–monotone dependence structure.

Since one uses the ranks of the observations, the distance measure representing as degenerate V-statistic and its limit distribution is a weighted sum of independent chi-square variables with some term, namely, "decentralization". Before we get that, if U_1 and U_2 are independent, that is, $\Lambda^{\gamma}_{\omega_0}(C) = 0$, for $\mathbf{U} \sim C$, we will represent the empirical distance covariance in the sets γ as a fourth–order degenerate V-statistic.

4.1.1. Independent Random Variables.
Let $\Upsilon_{\mathbf{t}}(\mathbf{u}_1, \mathbf{u}_2) = \left(\widetilde{\Upsilon}_{\mathbf{t}}(\mathbf{u}_1, \mathbf{u}_2) + \widetilde{\Upsilon}_{\mathbf{t}}(\mathbf{u}_2, \mathbf{u}_1)\right)/2$ be a symmetric kernel function, where for $\mathbf{u}_1 = (u_{11}, u_{12})$, $\mathbf{u}_2 = (u_{21}, u_{22})$ in $[0,1]^2$ and $\mathbf{t} = (t_1, t_2) \in \mathbb{R}^2$, the symmetric kernel $\widetilde{\Upsilon}_{\mathbf{t}}$ is given by

$$\widetilde{\Upsilon}_{\mathbf{t}}(\mathbf{u}_1, \mathbf{u}_2) = \exp\left(i\,\mathbf{t}\,\mathbf{u}_1^T\right) + i\,\mathbf{t}\,\{\mathbb{I}(\mathbf{u}_2 \leq \mathbf{u}_1) - \mathbf{u}_1\} - \kappa(\mathbf{t}),$$

where $\kappa(\mathbf{t}) = \psi_1(t_1)\,\psi_1(t_2) = \left(e^{i\,t_1} - 1\right)\left(e^{i\,t_2} - 1\right)/t_1 t_2$ denotes the product of uniform margins of copula characteristic function.

Proposition 1. *If U_1 and U_2 are independent and that Condition $\mathscr{A}_1$ holds, then for $\mathbf{U}_1, \ldots, \mathbf{U}_n$ i.i.d. generated from the copula C, the empirical bivariate distance covariance $\widehat{\Lambda}^{\gamma}_{n,\omega_0}$ given by* Eq. 7 *has the following representation*

$$\widehat{\Lambda}^{\gamma}_{n,\omega_0} = \frac{1}{n^3} \sum_{j,j',k,k'=1}^{n} K^{\gamma}_{\omega_0}\left(\mathbf{U}_j, \mathbf{U}_{j'}, \mathbf{U}_k, \mathbf{U}_{k'}\right),$$

where the functional symmetry kernel $12\,K^{\gamma}_{\omega_0}(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4)$ is defined by

$$\begin{aligned}
&\int_{\gamma} \{\Upsilon_{\mathbf{t}}(\mathbf{u}_1, \mathbf{u}_2) + \Upsilon_{\mathbf{t}}(\mathbf{u}_2, \mathbf{u}_1)\}\,\{\Upsilon_{-\mathbf{t}}(\mathbf{u}_3, \mathbf{u}_4) + \Upsilon_{-\mathbf{t}}(\mathbf{u}_4, \mathbf{u}_3)\}\,\omega_0(\mathbf{t})\,\mathrm{d}\mathbf{t} \\
&+ \int_{\gamma} \{\Upsilon_{\mathbf{t}}(\mathbf{u}_1, \mathbf{u}_3) + \Upsilon_{-\mathbf{t}}(\mathbf{u}_3, \mathbf{u}_1)\}\,\{\Upsilon_{-\mathbf{t}}(\mathbf{u}_2, \mathbf{u}_4,) + \Upsilon_{-\mathbf{t}}(\mathbf{u}_4, \mathbf{u}_2)\}\,\omega_0(\mathbf{t})\,\mathrm{d}\mathbf{t} \\
&+ \int_{\gamma} \{\Upsilon_{\mathbf{t}}(\mathbf{u}_1, \mathbf{u}_4) + \Upsilon_{\mathbf{t}}(\mathbf{u}_4, \mathbf{u}_1)\}\,\{\Upsilon_{-\mathbf{t}}(\mathbf{u}_2, \mathbf{u}_3) + \Upsilon_{\mathbf{t}}(\mathbf{u}_3, \mathbf{u}_2)\}\,\omega_0(\mathbf{t})\,\mathrm{d}\,\mathbf{t}.
\end{aligned}$$

First denote by

$$Q^{\gamma}_{\omega_0}(\mathbf{u}_1, \mathbf{u}_2) = \int_{\gamma} P_{\mathbf{t}}(\mathbf{u}_1)\,P_{-\mathbf{t}}(\mathbf{u}_2)\,\omega_0(\mathbf{t})\,\mathrm{d}\,\mathbf{t},$$

where $P_{\mathbf{t}}(\mathbf{u}) = \mathrm{E}\{\Upsilon_{\mathbf{t}}(\mathbf{u}, \mathbf{U}) + \Upsilon_{\mathbf{t}}(\mathbf{U}, \mathbf{u})\}$. Then, $\Lambda^{\gamma}_{n,\omega_0}$ shares the same limit as $\widetilde{\Lambda}^{\gamma}_{n,\omega_0}$ defined by

$$\widetilde{\Lambda}^{\gamma}_{n,\omega_0} = \frac{1}{n} \sum_{j,k=1}^{n} Q^{\gamma}_{\omega_0}\left(\mathbf{U}_j, \mathbf{U}_k\right) + o_{\mathbb{P}}(1).$$

The next statement is a consequence of the results of Bahraoui et al. (2018) where the Hájek projection method (e.g., Serfling (1980)) into two dimensional surface of V-statistics $\widehat{\Lambda}^{\gamma}_{n,\omega_0}$ is used and whose proof is omitted.

COROLLARY 1. *Following Corollary 1, p. 83 in* Lee (1990) *and if* U_1 *and* U_2 *are independent, one has for i.i.d.* $\mathbf{U}_1, \ldots, \mathbf{U}_n$ *generated from the copula* C*, that* $\widetilde{\Lambda}^{\gamma}_{n,\omega_0}$ *converges in distribution to*

$$\widetilde{\Lambda}^{\gamma}_{\omega_0} = \mathrm{E}\left\{Q^{\gamma}_{\omega_0}(\mathbf{U},\mathbf{U})\right\} + \sum_{\nu=1}^{\infty} \lambda_\nu \left(N_\nu^2 - 1\right),$$

where $\{N_\nu\}_{\nu=1}^{\infty}$ *is a sequence of i.i.d.* $\mathbb{N}(0,1)$ *random variables and* $\{\lambda_\nu\}_{\nu=1}^{\infty}$ *are the eigenvalues of* η*-operator, generated by the symmetry kernel* $Q^{\gamma}_{\omega_0}$ *as follows*

$$\eta : h \mapsto \mathrm{E}\{Q^{\gamma}_{\omega_0}(\mathbf{U},\mathbf{u})\, h(\mathbf{U})\}.$$

REMARK 2. It was shown in Section 2.1 that the bivariate distance variance is bounded under independent random variables U_1 and U_2. Then, one can represent the empirical measure of bivariate asymmetric $\widehat{\Delta}_{\omega_0}(\Psi_n)$ defined by Eq. 9 as a fourth–order degenerate V-statistic by

$$\widehat{\Delta}_{\omega_0}(\Psi_n) = \frac{1}{n^3} \sum_{j,j',k,k'=1}^{n} \Xi^{\gamma}_{\omega_0}\left(\mathbf{U}_j, \mathbf{U}_{j'}, \mathbf{U}_k, \mathbf{U}_{k'}\right),$$

where the symmetric kernel is given by $\Xi = \sigma^{-1}_{\omega_0}\left|K^{+}_{\omega_0} - K^{-}_{\omega_0}\right| \approx \left|K^{+}_{\omega_0} - K^{-}_{\omega_0}\right|$. Then as consequence of Corollary 1, one has that $\widehat{\Delta}_{\omega_0}(\Psi_n)$ converge in distribution to weighted sums of independent chi-square variables with $\{\lambda_\kappa\}_{\kappa=1}^{\infty}$ the eigenvalues of $\eta^\star$-operator, generated by the symmetry kernel $Q^{\gamma,\star}_{\omega_0}$ as follows

$$\eta^\star : h \mapsto \mathrm{E}\{Q^{\gamma,\star}_{\omega_0}(\mathbf{U},\mathbf{u})\, h(\mathbf{U})\},$$

where the symmetry kernel $Q^{\gamma,\star}_{\omega_0} = \left|Q^{+}_{\omega_0} - Q^{-}_{\omega_0}\right|$ is projection of kernel $\Xi^{\gamma}_{\omega_0}$.

4.1.2. Dependent Random Variables. Under dependent random variables U_1 and U_2, the asymptotic properties depend on unknown copula C and the marginal CCF. In this case, the empirical distance covariance is represented as sixth-order degenerate V-statistic and we will show that a limit has a normal distribution with mean zero and covariance given in Eq. 15. Let $\text{Ł}_{\mathbf{t}}$ be the symmetrized kernel function defined by

$$\text{Ł}_{\mathbf{t}}(\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_3) = (\widetilde{\text{Ł}}_{\mathbf{t}}(\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_3) + \widetilde{\text{Ł}}_{\mathbf{t}}(\mathbf{u}_2,\mathbf{u}_1,\mathbf{u}_3) + \widetilde{\text{Ł}}_{\mathbf{t}}(\mathbf{u}_3,\mathbf{u}_1,\mathbf{u}_2))/3,$$

where $\widetilde{Ł}$ is the kernel function given, for $\mathbf{u}_1 = (u_{11}, u_{12})$, $\mathbf{u}_2 = (u_{21}, u_{22})$, $\mathbf{u}_3 = (u_{31}, u_{32})$ in $[0,1]^2$ and $\mathbf{t} = (t_1, t_2) \in \mathbb{R}^2$, by

$$\begin{aligned}\widetilde{Ł}_{\mathbf{t}}(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3) &= e^{\left(i\,\mathbf{t}\mathbf{u}_1^T\right)} - e^{\left(i\,\mathbf{t}\mathbf{u}_2^T\right)} \\ &\quad + i\, e^{\left(i\,\mathbf{t}\mathbf{u}_1^T\right)} \left\{t_1\, \mathscr{I}(u_{11}, u_{31}) + t_2\, \mathscr{I}(u_{12}, u_{32})\right\} \\ &\quad - i\, e^{i\,(t_1\, u_{11} + t_2\, u_{22})} \left\{t_1 \mathscr{I}(u_{11}, u_{31}) + t_2 \mathscr{I}(u_{22}, u_{32})\right\},\end{aligned}$$

where $\mathscr{I}$ denotes $\mathscr{I}(a_1, a_2) = \mathbb{I}\,(a_2 \leq a_1) - a_2$ and $\mathbb{I}$ is indicator function.

The Proposition below states the connection between the empirical distance covariance and degenerate V-statistic, when U_1 and U_2 are dependent random variables and the margins are unknown.

PROPOSITION 2. *If Condition $\mathscr{A}_1$ and U_1 and U_2 are dependent variables are verified, then for $\mathbf{U}_1, \ldots, \mathbf{U}_n$ generated from the copula C, the empirical distance covariance $\widehat{\Lambda}^{\gamma}_{n,\omega_0}$ has the following representation*

$$\widehat{\Lambda}^{\gamma}_{n,\omega} = \frac{1}{n^5} \sum_{j,j',k,\ell,\ell',m=1}^{n} \Phi^{\gamma}_{\omega_0}\left(\mathbf{U}_j, \mathbf{U}_{j'}, \mathbf{U}_k, \mathbf{U}_\ell, \mathbf{U}_{\ell'}, \mathbf{U}_m\right),$$

where $72\, \Phi^{\gamma}_{\omega}(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4, \mathbf{u}_5, \mathbf{u}_6) =$

$$\begin{aligned}&\int_\gamma \{Ł_{\mathbf{t}}(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)\, Ł_{-\mathbf{t}}(\mathbf{u}_4, \mathbf{u}_5, \mathbf{u}_6) + Ł_{\mathbf{t}}(\mathbf{u}_1, \mathbf{u}_4, \mathbf{u}_5)\, Ł_{-\mathbf{t}}(\mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_6) \\ &+ Ł_{\mathbf{t}}(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_6)\, Ł_{-\mathbf{t}}(\mathbf{u}_4, \mathbf{u}_5, \mathbf{u}_3) + Ł_{\mathbf{t}}(\mathbf{u}_1, \mathbf{u}_5, \mathbf{u}_6)\, Ł_{-\mathbf{t}}(\mathbf{u}_4, \mathbf{u}_2, \mathbf{u}_3) \\ &+ Ł_{\mathbf{t}}(\mathbf{u}_1, \mathbf{u}_4, \mathbf{u}_6)\, Ł_{-\mathbf{t}}(\mathbf{u}_2, \mathbf{u}_5, \mathbf{u}_3) + Ł_{\mathbf{t}}(\mathbf{u}_1, \mathbf{u}_6, \mathbf{u}_3)\, Ł_{-\mathbf{t}}(\mathbf{u}_4, \mathbf{u}_5, \mathbf{u}_2) \\ &+ Ł_{\mathbf{t}}(\mathbf{u}_1, \mathbf{u}_5, \mathbf{u}_3) Ł_{-\mathbf{t}}(\mathbf{u}_4, \mathbf{u}_2, \mathbf{u}_6) + Ł_{\mathbf{t}}(\mathbf{u}_1, \mathbf{u}_4, \mathbf{u}_3) Ł_{-\mathbf{t}}(\mathbf{u}_2, \mathbf{u}_5, \mathbf{u}_6)\}\, \omega_0(\mathbf{t}) \mathrm{d}\mathbf{t}.\end{aligned}$$

Let now introduce the following quantities

$$\tau^2_{1,\omega_0} = \mathrm{Var}\left(\widetilde{\Phi}^{\gamma}_{\omega_0}(\mathbf{U}_1)\right) \quad \text{and} \quad \tau^2_{k,\omega_0} = \mathrm{Cov}\left(\widetilde{\Phi}^{\gamma}_{\omega_0}(\mathbf{U}_1), \widetilde{\Phi}^{\gamma}_{\omega_0}(\mathbf{U}_k)\right),$$

where $\widetilde{\Phi}^{\gamma}_{\omega_0}(\mathbf{u}) = \mathrm{E}\left\{\Phi^{\gamma}_{\omega_0}\left(\mathbf{u}, \mathbf{U}_2, \mathbf{U}_3, \mathbf{U}_4, \mathbf{U}_5, \mathbf{U}_6\right)\right\}$.

The next Proposition shows the large-sample behavior of the empirical distance covariance process

$$\mathbb{H}_{n,\omega_0}(\mathbf{U}) = n^{1/2}\left(\widehat{\Lambda}^{\gamma}_{n,\omega} - \theta(\Psi_C)(\mathbf{U})\right) \qquad \mathbf{U} = (U_1, U_2) \sim C$$

for dependent variables (U_1, U_2), where $\theta(\Psi_C)$ is an estimator of $\Lambda^{\gamma}_{n,\omega}$ given by

$$\begin{aligned}\theta(\Psi_C)(\mathbf{U}) &= \mathrm{E}\left\{\Phi^{\gamma}_{\omega_0}\left(\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3, \mathbf{U}_4, \mathbf{U}_5, \mathbf{U}_6\right)\right\} \\ &= \int_\gamma \left|\Psi_C(\mathbf{t}) - \psi_1(t_1)\psi_2(t_2)\right|^2 \omega_0(\mathbf{t}) \mathrm{d}\,\mathbf{t}.\end{aligned}$$

Proposition 3. *Assume Condition $\mathscr{A}_1$ holds. For $\mathbf{U}_1, \ldots, \mathbf{U}_n$ generated from the copula C, suppose that the covariance $\sum_{k=2}^{\infty} \tau_{k,\omega_0}^2$ is finite. Then, the empirical distance covariance process $\mathbb{H}_{n,\omega_0}$ converges weakly, as $n \to \infty$, to a centered independent Gaussian process $\mathbb{H}_{\omega_0}$ with covariance*

$$\widetilde{\tau}_{\omega_0}^2 = 6^2 \left(\tau_{1,\omega_0}^2 + 2 \sum_{k=2}^{\infty} \tau_{k,\omega_0}^2 \right).$$

Remark 3. Clearly, from previous Proposition 3, one can derive the large-sample theory of the empirical distance variance process $\mathbb{H}_{n,\omega_0}(\mathbf{U}_\ell)$, for every $\mathbf{U}_\ell = (U_\ell, U_\ell) \sim C$, $\ell = 1, 2$. The only terms that will change are the symmetry kernel $\Phi_{\omega_0}^{\gamma}$ and covariance function $\widetilde{\tau}_{\omega_0}^2$. In the sequel, we denote for each $\ell = 1, 2$, the symmetry kernel of distance variance by $\Phi_{\ell,\omega_0}^{\gamma}$ and covariance function $\widetilde{\rho}_{\ell,\omega_0}^2$. On the other hand, the stochastic representation of the empirical distance correlation $\widehat{R}_n^{\gamma}$ has the functional form

$$\widehat{R}_n^{\gamma}(\widehat{\mathbf{U}}_j) = \varphi\left(\widehat{\Lambda}_{n,\omega_0}^{\gamma}, \widehat{\sigma}_{n,\omega_0,\ell}^{\gamma} \right),$$

where $\varphi(x, y) = x/\sqrt{y}$ is continuous function. In fact, R_n^{γ} is V-functionals statistics with some symmetry ratio kernel. To derive the asymptotic behavior of R_n^{γ}, one can use the Functional Delta Method if the related symmetry ratio kernel is bounded, if not, the V-functionals statistics are not Hadamard differentiable. The methodologies introduced by Beutner and Zähle (2012) will be useful in this case. In the sequel, we suppose that the related symmetry kernel of R_n^{γ} is bounded.

The following Lemma shows the asymptotic behavior of $\widehat{R}_n^{\gamma}$ under dependent random variables U_1 and U_2.

Proposition 4. *Assume that the covariance $\sum_{k=2}^{\infty} \tau_{k,\omega_0}^2$ is finite and that Condition $\mathscr{A}_1$ holds. Then, the empirical distance correlation process*

$$\mathbb{R}_n^{\gamma}(\mathbf{U}) = \sqrt{n} \left\{ \widehat{R}_n^{\gamma}(\widehat{\mathbf{U}}_j) - R^{\gamma}(\mathbf{U}) \right\}, \qquad \mathbf{U} = (U_1, U_2) \sim C$$

converges in distribution, as $n \to \infty$, to $\mathbb{R}^{\gamma}(\mathbf{U})$ independent centered normal distribution with variance-covariance $\tau_{\omega_0}^2 = (\nabla\varphi)^T \Omega (\nabla\varphi)$, where $(\nabla\varphi) \in \mathbb{R}^{3\times 1}$ the gradient of φ and $\Omega \in \mathbb{R}^{3\times 3}$ is the covariance matrix.

5. Example in Copulas Models

In this section, we present a graphical visualization of characteristic function associated to a copula; computation of a measure of asymmetry $\Delta(\Psi_C)$

for some bivariate parametric copula models in the set γ^- and γ^+, and numerical example to illustrate the finite sample behavior. Six parametric bivariate copula families will be considered, the radially asymmetry copulas, namely, Clayton (Cℓ) and Gumbel (Gu) :

$$\begin{aligned} C_\theta^{\mathrm{C}\ell}(u_1,u_2) &= \left(u_1^{-\theta}+u_2^{-\theta}-1\right)^{-1/\theta}, \quad \theta>0,\\ C_\theta^{\mathrm{Gu}}(u_1,u_2) &= \exp\left\{-\left(|\ln u_1|^{1/(1-\theta)}+|\ln u_2|^{1/(1-\theta)}\right)^{1-\theta}\right\}, \quad \theta\in[0,1], \end{aligned}$$

the radially symmetry bivariate copula models, Frank (Fr), Plackett (Pℓ) and the Student T_ν, (with $\nu=6$ degree of freedom) copulas:

$$\begin{aligned} C_\theta^{\mathrm{Fr}}(u_1,u_2) &= -\frac{1}{\theta}\ln\left\{1+\frac{\left(e^{-\theta u_1}-1\right)\left(e^{-\theta u_2}-1\right)}{e^{-\theta}-1}\right\}, \quad \mathbb{R}\setminus\{0\}\\ C_\theta^{\mathrm{P}\ell}(u_1,u_2) &= \frac{\zeta_\theta(u_1,u_2)-\sqrt{\{\zeta_\theta(u_1,u_2)\}^2-4\theta(\theta-1)u_1u_2}}{2(\theta-1)}, \quad \theta\in[0,\infty)\\ C_\theta^{\mathrm{T}_\nu}(u_1,u_2) &= \int_{-\infty}^{t_\nu^{-1}(u_1)}\int_{-\infty}^{t_\nu^{-1}(u_2)}\varphi_{\nu,\theta}(x_1,x_2)\,\mathrm{d}x_2\,\mathrm{d}x_1, \quad \theta\in(-1,1), \end{aligned}$$

where $\varphi_{\nu,\theta}$ is the bivariate density of Student with ν degrees of freedom and $\zeta_\theta(u_1,u_2)=1+(\theta-1)(u_1+u_2)$. The last copula models that we consider is the radial symmetry Farlie-Gumbel-Morgenstern (FGM) copula introduced by Farlie (1960) and has the following representation:

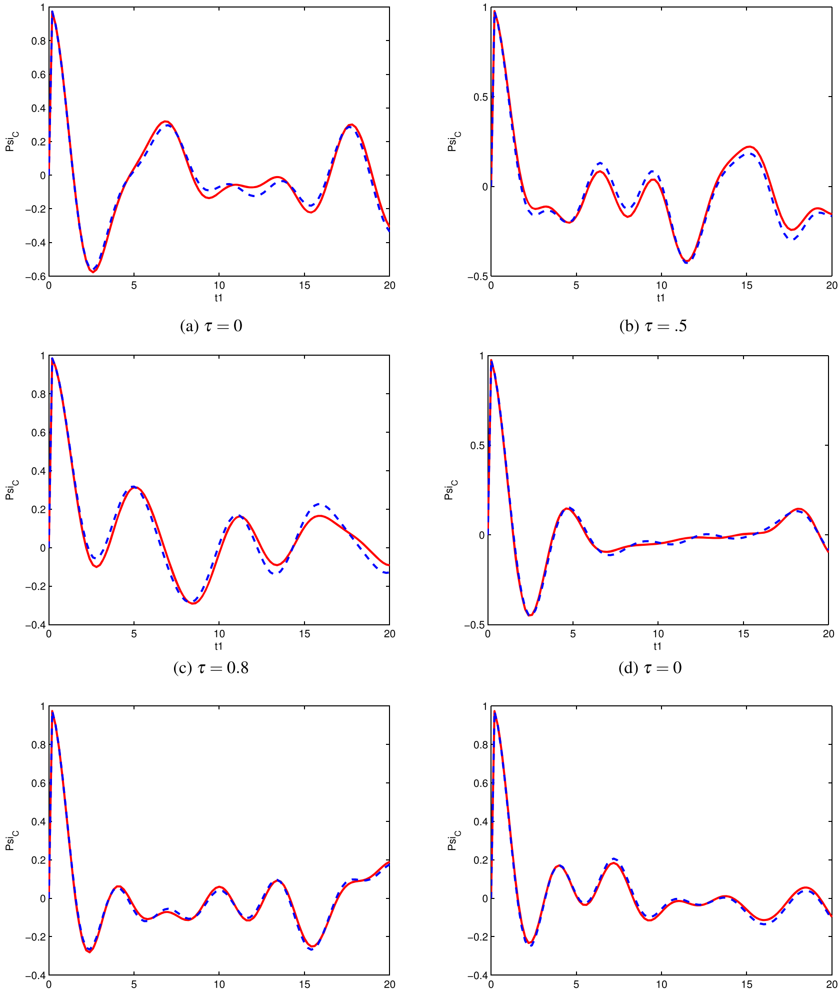
$$C_\theta^{\mathrm{FGM}}(u_1,u_2)=u_1\,u_2\,\{1-\theta\,(1-u_1)\,(1-u_2)\}\,,\quad -1<\theta<1.$$

The bivariate parametric copula models have been parameterized in terms of their associated Kendall's tau, defined for a given copula C by

$$\tau_C=4\int_0^1\int_0^1 C(u_1,u_2)\,\mathrm{d}C(u_1,u_2)-1.$$

Figure 1 shows the curves of the real part of Ψ_C associated with Gumbel(a-b-c) and Frank(d-e-f) for sample size $n=50$, where Ψ_C is given, for $\mathbf{U}\sim C$ and conditionally to the set $\gamma\subseteq\mathbb{R}^2$.

As one can see in Fig. 1, even with small or moderate sample sizes, the difference between curves of real part of Ψ_C in sets γ^- and γ^+ is significantly interesting when the strength of dependence increases and there exist more variation between (CCF) in a complementary order-preserved sets γ^+ and γ^-.



(a) $\tau = 0$ (b) $\tau = .5$ (c) $\tau = 0.8$ (d) $\tau = 0$ (e) $\tau = .5$ (f) $\tau = .8$

Figure 1: Curves of real part of $\Psi_C(t_1, t_2)$ for Gumbel (a-b-c) Frank (d-e-f) copula (in red) when $t_1 < t_2$ and (in blue) when $t_2 < t_1$ with $n = 50$

Table 1 reports the results of estimated measure of bivariate asymmetry for $n \in \{50, 100\}$ and $\tau \in \{0, 0.5.0.8\}$. One can observe the followings patterns: (a) under radial asymmetry copula, when τ increases, the estimated measure increases, near to one i.e., the strong relationship between variables, and under radial symmetry copula models, when sample size n increases and for small values of τ the measure of bivariate asymmetry decreases, the measure close to zero, this drop can be explained by the fact

Table 1: Measure of asymmetry $\widehat{\Delta}_{\omega_0}(\Psi_n)$ for Clayton (Cℓ), Gumbel (Gu), Farlie Gumbel Morgenstern (FGM), Plackett (Pℓ), Frank and Student (T_6) copulas

	τ_C	Cℓ	Gu	FGM	Pℓ	Fr	T_6
$n = 20$	0	.0842	.0792	.1056	.0812	.0795	.0729
	.5	.5317	.4086	.2607	.5266	.5287	.3398
	.8	.8152	.9292	.2636	.6045	.9763	.9427
$n = 50$	0	.0290	.0569	.0756	.0492	.0307	.0402
	.5	.5295	.4629	.1186	.3947	.7186	.3620
	.8	.8209	.9167	.1868	.6268	.9079	.7187
$n = 100$	0	.0160	.0159	.0186	.0135	.0119	.0260
	.5	.5184	.5319	.1553	.2949	.3356	.4058
	.8	.8937	.8988	.1502	.5562	.9126	.8763

that distance correlation $R^-_{\omega_0}$ is close to $R^+_{\omega_0}$ under radial symmetry copula models, and according to Fig. 1 one can see that the curve of real part of Ψ_C (C Frank copula) in the sets γ^+ and γ^- are near-identical. (b) in general and for all copulas under consideration, when $\tau = 0$ and sample size n increases, the measure of asymmetry decreases. As one can see, that the smallest value of measure appears in (FGM) copula, as a consequence, that the (FGM) copula has a small perturbation of independence copula. On the other side, the estimated measure of bivariate asymmetry $\widehat{\Delta}_{\omega_0}(\Psi_n)$ is related to the concept of bivariate tail dependence and tail order, see Joe (2015) for more details of concept on the tail of copula families.

6. Data Example

6.1. Nutrient Data The nutrient data consists of four-day measurements for the intake of (Calcium, iron, protein, vitamin A, vitamin C) from women aged 25 to 50 in the United States as part of the "Continuing Survey of Food Intakes of Individuals" program. The data has n=737 measurements collected from a cohort study of the Department of Agriculture (USDA) and it is available online from the University of Pennsylvania repository. The results of a measure of asymmetry $\widehat{\Delta}_{\omega_0}(\Psi_n)$ and the degree of asymmetry CCF are represented in Table 2. The data was used to illustrate the concepts of symmetry (Genest et al., 2012); Quessy and Bahraoui (2013) and to fit the multivariate data by Archimedean copula via Liouville generalization (McNeil and Nešlehová, 2010). The results of test of symmetry (Genest et al., 2012) find that there are six out of ten pairs bivariate margins for

Table 2: Measure of asymmetry $\widehat{\Delta}_{\omega_0}(\Psi_n)$ and degree of asymmetry between brackets ($\mu(\Psi_C)$) for the nutrient dataset, n=100

Nutrien	*Iron*	*Protein*	*VitaminA*	*VitaminC*
Calcium	.1939 (.1194)	.2253 (.1237)	.3165 (.1108)	.1069 (.1055)
Iron		.4143 (.1086)	.2378 (.1159)	.2209 (.1251)
Protein			.1340 (.1213)	.1489 (.1352)
vitamin A				.1957 (.1256)

which the null hypothesis of symmetry i.e., $\mathbb{H}_0 : C(u, v) - C(v, u) = 0$, for u and v in $[0, 1]^2$, is reject. For each pair of variables, we calculated the degree of asymmetry CCF defined by

$$\mu(\Psi_C) = \sup_{(t_1,t_2)\in\omega} |\Psi_C(t_1, t_2) - \Psi_C(t_2, t_1)|.$$

Table 2 resulted the asymmetry in the dependence of the pairs (Iron, Protein) and (Calcium, vitamin A) has a relatively high association compared with the other pairs; compared with the degree of asymmetry $\mu(\Psi_C)$, the measure $\widehat{\Delta}_{\omega_0}(\Psi_n)$ identified a strong asymmetry in the dependence when the degree of asymmetry CCF is important, maybe due to the existence of tails. Note that the degree of asymmetry calculated to CCF rather than copulas C, where for any copula, we have

$$\mu(\Psi_C) \leq \mu(C) = \sup_{(u,v)\in[0,1]^2} |C(u, v) - C(v, u)| \leq \frac{1}{3}.$$

7 Discussion

In this paper, we investigated a new measure of bivariate asymmetry based on the rank characteristic function associated with copula models and considered its properties. The theory of V-statistic seems to be more sophisticated than the approach based on the stochastic process to derive the asymptotic behavior of the empirical copula characteristic function process $\mathscr{Z}_n(\mathbf{t})$ given in Eq. 13. The proposed measure discriminates well the strength of dependence between different copulas according to their tail behavior i.e., tail dependence or independence and the geometrical structure radial asymmetry or symmetry.

It is worthwhile mentioning that the direction of asymmetry involves the direction of the association between the two random variables U_1 and U_2. In this sense, the properties of non–exchangeability partake some of the properties of independence. The direction of asymmetry is camouflaged or governed by an eigenvector related to an eigenvalue of symmetry kernel extracted from V-statistics. In other sense, there exists a relation between the procedure of test independence and test of non–exchangeability of the random variables.

In further research, it is hoped to extend the results of this paper in high dimensional, testing independence based on rank distance correlation of copula models, and measuring the asymmetry under weak dependent data, see e.g., Leucht (2012) and Bücher and Kojadinovic (2013).

Acknowledgements. The authors are partially supported by a *FAPESP* grants 2013/07375-0 and 2018/05262-7. The authors would like to thank the Editor and two referees for their valuable comments.

References

ALSINA, C., FRANK, M.J. and SCHWEIZER, B. (2006). Associative functions. World Scientific Publishing Co. Pte Ltd., Hackensack, NJ. Triangular norms and copulas.

ANG, A. and CHEN, J. (2002). Asymmetric correlations of equity portfolios. *J. Financial Econ.* **63**, 443–494.

BAHRAOUI, T., BOUEZMARNI, T. and QUESSY, J.-F. (2018). A family of goodness-of-fit tests for copulas based on characteristic functions. *Scand. J. Stat.* **45**, 301–323.

BAHRAOUI, T., BOUEZMARNI, T. and QUESSY, J.-F. (2019). Testing the symmetry of a dependence structure with a characteristic function. *Depend. Model.* **6**, 331–355.

BAHRAOUI, T. and QUESSY, J.-F. (2017). Tests of radial symmetry for multivariate copulas based on the copula characteristic function. *Electron. J. Stat.* **11**, 2066–2096.

BÜCHER, A., IRRESBERGER, F. and WEISS, G.N.F. (2017). Testing asymmetry in dependence with copula-coskewness. *North American Actuarial Journal* **21**, 267–280.

BEUTNER, E. and ZÄHLE, H. (2012). Deriving the asymptotic distribution of u- and v-statistics of dependent data using weighted empirical processes. *Bernoulli* **18**, 803–822.

BÜCHER, A. and KOJADINOVIC, I. (2013). A dependent multiplier bootstrap for the sequential empirical copula process under strong mixing. ArXiv e-prints.

CSÖRGŐ, S. (1985). Testing for independence by the empirical characteristic function. *J. Multivariate Anal.* **16**, 290–299.

DEHGANI, A., DOLATI, A. and ÚBEDA-FLORES, M. (2013). Measures of radial asymmetry for bivariate random vectors. *Statist. Papers* **54**, 271–286.

DOBRIĆ, J. and SCHMID, F. (2005). Testing goodness of fit for parametric families of copulas—application to financial data. *Comm. Statist. B—Simul. Comput.* **34**, 1053–1068.

DURANTE, F., KLEMENT, E.P., SEMPI, C. and ÚBEDA-FLORES, M. (2010). Measures of non-exchangeability for bivariate random vectors. *Statist. Papers* **51**, 687–699.

FAN, YANAN, LAFAYE DE MICHEAUX, PIERRE, PENEV, SPIRIDON, SALOPEK and DONNA (2017). Multivariate nonparametric test of independence. *J. Multivariate Anal.* **153**, 189–210. https://doi.org/10.1016/j.jmva.2016.09.014.

FAN, Y. (1997). Goodness-of-fit tests for a multivariate distribution by the empirical characteristic function. *J. Multivariate Anal.* **62**, 36–63.

FARLIE, D.J.G. (1960). The performance of some correlation coefficients for a general bivariate distribution. *Biometrika* **47**, 307–323.

Fernández-Sánchez, J. and Úbeda Flores, M. (2018). Constructions of copulas with given diagonal (and opposite diagonal) sections and some generalizations. *Depend. Model.* **6**, 139–155.

Feuerverger, A. and Mureika, R.A. (1977). The empirical characteristic function and its applications. *Ann. Statist.* **5**, 88–97.

Fredricks, G.A. and Nelsen, R.B. (1997). *Copulas constructed from diagonal sections.* Kluwer Acad. Publ., Dordrecht, p. 129–136.

Genest, C., Kojadinovic, I., Nešlehová, J. and Yan, J. (2011). A goodness-of-fit test for bivariate extreme-value copulas. *Bernoulli* **17**, 253–275.

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

Genest, C. and Rémillard, B. (2004). Tests of independence and randomness based on the empirical copula process. *Test* **13**, 335–370.

Genest, C. and Segers, J. (2009). Rank-based inference for bivariate extreme-value copulas. *Ann. Statist.* **37**, 2990–3022.

Hájek, J., Šidák, Z. and Sen, P.K. (1999). *Theory of Rank Tests*, 2nd edn. Probability and Mathematical Statistics. Academic Press, Inc, San Diego.

Henze, N., Klar, B. and Meintanis, S.G. (2003). Invariant tests for symmetry about an unspecified point based on the empirical characteristic function. *J. Multivariate Anal.* **87**, 275–297.

Herrndorf, N. (1984). An example on the central limit theorem for associated sequences. *Ann. Probab.* **12**, 912–917.

Joe, H. (2015). *Dependence Modeling with Copulas, Volume 134 of Monographs on Statistics and Applied Probability.* CRC Press, Boca Raton.

Kankainen, A. and Ushakov, N.G. (1998). A consistent modification of a test for independence based on the empirical characteristic function. *J. Math. Sci. (New York)* **89**, 1486–1494. Stability problems for stochastic models, Part 1(Moscow, 1996).

Kojadinovic, I. and Yan, J. (2011). Tests of serial independence for continuous multivariate time series based on a Möbius decomposition of the independence empirical copula process. *Ann. Inst. Statist. Math.* **63**, 347–373.

Koroljuk, V.S. and Borovskich, Y.V. (1994). Theory of U-statistics, Volume 273 of Mathematics and its Applications. Kluwer Academic Publishers Group, Dordrecht. Translated from the 1989 Russian original by P. V. Malyshev and D. V. Malyshev and revised by the authors.

Krupskii, P. (2017). Copula-based measures of reflection and permutation asymmetry and statistical tests. *Statist. Papers* **58**, 1165–1187.

Krupskii, P. and Joe, H. (2015). Tail-weighted measures of dependence. *J. Appl. Stat.* **42**, 614–629.

Lee, A.J. (1990). *U-Statistics, Volume 110 of Statistics: Textbooks and Monographs.* Marcel Dekker Inc. Theory and practice, New York.

Lee, D., Joe, H. and Krupskii, P. (2018). Tail-weighted dependence measures with limit being the tail dependence coefficient. *J. Nonparametr. Stat.* **30**, 262–290.

Leucht, A. (2012). Degenerate u- and v-statistics under weak dependence: asymptotic theory and bootstrap consistency. *Bernoulli* **18**, 552–585.

Lukacs, E. (1970). *Characteristic Functions*, 2nd edn. Hafner Publishing Co., New York. revised and enlarged.

McNeil, A.J., Frey, R. and Embrechts, P. (2015). Quantitative risk management (Revised ed.). Princeton Series in Finance. Princeton University Press, Princeton, NJ. Concepts, techniques and tools.

MCNEIL, A.J. and NEŠLEHOVÁ, J. (2010). From Archimedean to Liouville copulas. *J. Multivariate Anal.* **101**, 1772–1790.

MEINTANIS, S.G. (2004). A class of omnibus tests for the Laplace distribution based on the empirical characteristic function. *Comm. Statist. Theory Methods* **33**, 925–948.

MEINTANIS, S.G., ALLISON, J. and SANTANA, L. (2016). Goodness-of-fit tests for semiparametric and parametric hypotheses based on the probability weighted empirical characteristic function. *Statist. Papers* **57**, 957–976.

MEINTANIS, S.G. and ILIOPOULOS, G. (2008). Fourier methods for testing multivariate independence. *Comput. Statist. Data Anal.* **52**, 1884–1895.

NELSEN, R.B. (2006). *An Introduction to Copulas*, 2nd edn. Springer, New York. Springer Series in Statistics.

NEWMAN, C.M. (1980). Normal fluctuations and the fkg inequalities. *Comm. Math. Phys.* **74**, 119–128.

QUESSY, J.-F. and BAHRAOUI, T. (2013). Graphical and formal statistical tools for the symmetry of bivariate copulas. *Canad. J. Statist.* **41**, 637–656.

RÉMILLARD, B. and SCAILLET, O. (2009). Testing for equality between two copulas. *J. Multivariate Anal.* **100**, 377–386.

RÉNYI, A. (1959). On measures of dependence. *Acta Math. Acad. Sci. Hungar.* **10**, 441–451. unbound insert.

ROSCO, J.F. and JOE, H. (2013). Measures of tail asymmetry for bivariate copulas. *Statist. Papers* **54**, 709–726.

SALVADORI, G., DE MICHELE, C., KOTTEGODA, N.T. and ROSSO, R. (2007). *Extremes in Nature: an Approach using Copulas.* Springer, New York.

SCHWEIZER, B. and WOLFF, E.F. (1981). On nonparametric measures of dependence for random variables. *Ann. Statist.* **9**, 879–885.

SERFLING, R.J. (1980). *Approximation Theorems of Mathematical Statistics.* Wiley, New York. Wiley Series in Probability and Mathematical Statistics.

SHAKED, M. and SHANTHIKUMAR, J.G. (2007). *Stochastic Orders. Springer Series in Statistics.* Springer, New York.

SZÉKELY, G.J. and RIZZO, M.L. (2005). A new test for multivariate normality. *J. Multivariate Anal.* **93**, 58–80.

SZÉKELY, G.J., RIZZO, M.L. and BAKIROV, N.K. (2007). Measuring and testing dependence by correlation of distances. *Ann. Statist.* **35**, 2769–2794.

USHAKOV, N.G. (1999). Selected topics in characteristic functions. Modern Probability and Statistics. VSP, Utrecht.

WANG, F. (2016). A novel coefficient for detecting and quantifying asymmetry of California electricity market based on asymmetric detrended cross-correlation analysis. *Chaos* **26**, 063109, 10.

ZHANG, Z. and SHINKI, K. (2007). Extreme co-movements and extreme impacts in high frequency data in finance. *Journal of Banking & Finance* **31**, 1399–1415.

Publisher's Note. Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Appendix A: Proofs

A.1 Proof of Lemma 1

The property (a) will allow us to prove the boundedness of our measure. From Lemma 2, one has

$$\Delta_{\omega_0}(\Psi_W) \leq \Delta_{\omega_0}(\Psi_C) \leq \Delta_{\omega_0}(\Psi_M).$$

Using the fact that the Fréchet Hoeffding bounds of CCF are bounded, we can find ε_1 and ε_2 for which, for every $C \in \mathscr{C}$ and $\left|e^{i\,\mathbf{tu}}\right| \leq 1$, $\mathbf{u} = (u_1, u_2) \in [0,1]^2$, one has $|\Delta_{\omega_0}(\Psi_M)| \leq \varepsilon_1$ and $|\Delta_{\omega_0}(\Psi_W)| \leq \varepsilon_2$.

For property (b), if C is radially symmetric, one has $\Psi_C(t_1, t_2) = \Psi(-t_1, -t_2)$. In other words, it can be equivalently written as

$$\mathscr{L}_C(t_1, t_2) = \mathrm{E}\left\{\sin\left(t_1 W_1 + t_2 W_2\right)\right\} = 0, \tag{15}$$

where $(W_1, W_2) = (U_1 - 1/2, U_2 - 1/2)$. Then in this case, one has $R^+_{\omega_0}(\mathbf{U}) = R^-_{\omega_0}(\mathbf{U})$, for all $\mathbf{U} \sim C$.

The property (c) can be established directly.

For property (d), assume that

$$\lim_{n\to\infty} \|C_n - C\|_\infty = 0, \quad \text{for} \quad C \in \mathscr{C}.$$

Then the integrals $\int_{[0,1]^2} \phi(\mathbf{u})\mathrm{d}C(\mathbf{u})$ and $\int_{[0,1]^2} \phi(\mathbf{u})\mathrm{d}C_n(\mathbf{u})$ exists for all $\mathbf{u} = (u_1, u_2) \in [0,1]^2$, where $\phi(\mathbf{u}) = \exp(i\mathbf{tu})$. Hence, one has

$$\lim_{n\to\infty} \int_{[0,1]^2} \phi(\mathbf{u})\mathrm{d}C_n(\mathbf{u}) = \int_{[0,1]^2} \phi(\mathbf{u})\mathrm{d}C(\mathbf{u}).$$

A.2 Proof of Lemma 2

For the lower copula characteristic function bound Ψ_W associated to $W(u_1, u_2) = \max\left(u_1 + u_2 - 1, 0\right)$, one has

$$\begin{aligned}
\Psi_W(t_1, t_2) &= \int_{[0,1]^2} e^{i\,t_1\,u_1 + i\,t_2\,u_2}\,\mathrm{d}W(u_1, u_2) \\
&= \int_{[0,1]^2} e^{i\,t_1\,u_1 + i\,t_2\,u_2}\,\mathrm{d}\mathbb{I}\{u_1 \geq 1 - u_2\} \\
&= \int_0^1 \left\{\int_{1-u_2}^1 e^{i\,t_1\,u_1 + i\,t_2\,u_2}\,\mathrm{d}\,u_1\right\}\mathrm{d}\,u_2 \\
&= \int_0^1 \left\{\frac{1}{it_1}\left(e^{i\,t_1 + i\,t_2\,u_2} - e^{i\,t_1 + i\,(t_2 - t_1)\,u_2}\right)\right\}\mathrm{d}\,u_2 \\
&= \frac{1}{t_1 t_2}\left(e^{it_1} - e^{i(t_1+t_2)}\right) - \frac{1}{t_2 - t_1}\left(e^{it_2} - e^{it_1}\right).
\end{aligned}$$

For the upper copula characteristic function bound Ψ_M associated to $M(u_1, u_2) = \min(u_1, u_2)$, one has

$$\begin{aligned}
\Psi_M(t_1, t_2) &= \int_{[0,1]^2} e^{i\,t_1\,u_1 + i\,t_2\,u_2}\,\mathrm{d}M(u_1, u_2) \\
&= \int_{[0,1]^2} e^{i\,t_1\,u_1 + i\,t_2\,u_2}\,\mathrm{d}\left\{u_1 \mathbb{I}\{u_1 \le u_2\} + u_2 \mathbb{I}\{u_2 \le u_1\}\right\} \\
&= \int_0^1 \left\{ \int_0^{u_2} e^{i\,t_1\,u_1 + i\,t_2\,u_2}\,\mathrm{d}\,u_1 \right\} \mathrm{d}\,u_2 \\
&\quad + \int_0^1 \left\{ \int_0^{u_1} e^{i\,t_1\,u_1 + i\,t_2\,u_2}\,\mathrm{d}\,u_2 \right\} \mathrm{d}\,u_1 \\
&= \int_0^1 \left\{ \frac{1}{it_1} \left(e^{i\,(t_1+t_2)u_2} - e^{i\,t_2\,u_2} \right) \right\} \mathrm{d}\,u_2 \\
&\quad + \int_0^1 \left\{ \frac{1}{it_2} \left(e^{i\,(t_1+t_2)u_1} - e^{i\,t_1\,u_1} \right) \right\} \mathrm{d}\,u_1 \\
&= \frac{1}{t_1 t_2} \left(e^{it_1} + e^{it_2} - 2 \right) - \frac{2}{t_1\,t_2} \left(e^{i(t_1+t_2)} - 1 \right) \\
&= \frac{\left(e^{it_1} - 1\right)}{i\,t_1} \frac{\left(e^{it_2} - 1\right)}{i\,t_2} = \prod_{\ell=1}^{2} \psi_\ell(t_\ell),
\end{aligned}$$

where for each $\ell = 1, 2$, $\psi_\ell(t_\ell)$ is characteristic function of uniform distribution $\mathscr{U}(0, 1)$.

A.3 Proof of Lemma 3

First, write

$$\begin{aligned}
\left|\Psi_n(t_1, t_2) - \psi_{n1}(t_1)\,\psi_{n2}(t_2)\right|^2 &= \left|\Psi_n(t_1, t_2)\right|^2 + \left|\psi_{n1}(t_1)\psi_{n2}(t_2)\right|^2 \\
&\quad - 2\,\Re\left\{ \psi_{n1}(t_1)\psi_{n2}(t_2)\overline{\Psi_n(t_1, t_2)} \right\},
\end{aligned}$$

where $\overline{\Psi}_n$ is a complex conjugate of empirical CCF Ψ_n given by Eq. 8. For the first term, one has

$$\begin{aligned}
\left|\Psi_n(t_1, t_2)\right|^2 &= \Psi_n(t_1, t_2)\overline{\Psi_n(t_1, t_2)} \\
&= \frac{1}{n^2} \sum_{j,k=1}^{n} \Big\{ \cos\Big((\widehat{U}_{j1} - \widehat{U}_{k1})\,t_1 + (\widehat{U}_{j2} - \widehat{U}_{k2})\,t_2 \Big) \\
&\quad + i\,\sin\Big((\widehat{U}_{j1} - \widehat{U}_{k1})\,t_1 + (\widehat{U}_{j2} - \widehat{U}_{k2})\,t_2 \Big) \Big\}.
\end{aligned}$$

Using trigonometric identities $\cos(x+y) = \cos(x)\,\cos(y) - \sin(x)\,\sin(y)$, one has

$$|\Psi_n(t_1,t_2)|^2 = \frac{1}{n^2}\sum_{j,k=1}^{n} \cos\left((\widehat{U}_{j1} - \widehat{U}_{k1})\,t_1\right)\cos\left((\widehat{U}_{j2} - \widehat{U}_{k2})\,t_2\right) + \epsilon_1(t_1,t_2).$$

The same kind of arguments apply to the second term and one has the following representation

$$\begin{aligned} |\psi_{n1}(t_1)\psi_{n2}(t_2)|^2 &= \psi_{n1}(t_1)\psi_{n2}(t_2)\overline{\psi_{n1}(t_1)\psi_{n2}(t_2)} \\ &= \frac{1}{n^4}\sum_{j,k,s,l=1}^{n} \cos\left((\widehat{U}_{j1} - \widehat{U}_{k1})t_1\right)\cos\left((\widehat{U}_{s2} - \widehat{U}_{l2})t_2\right). \end{aligned}$$

Similarly, one can write the real part of the third term, with respect to the symmetric property of weighted function ω_0, as

$$\begin{aligned} \Re\left\{\psi_{n1}(t_1)\psi_{n2}(t_2)\overline{\Psi_n(t_1,t_2)}\right\} &= \frac{1}{n^3}\sum_{j,k,s=1}^{n} \cos\left((\widehat{U}_{j1} - \widehat{U}_{k1})t_1\right) \\ &\quad \times \cos\left((\widehat{U}_{j2} - \widehat{U}_{s2})t_2\right). \end{aligned}$$

The remainder imaginary part terms of $\Delta_n(t_1,t_2) = |\Psi_n(t_1,t_2) - \psi_{n1}(t_1)\,\psi_{n2}(t_2)|^2$ vanishes. Let $\Im\left\{\Delta_n(t_1,t_2)\right\} = \epsilon_1 + \epsilon_2 - 2\epsilon_3$, where

$$\epsilon_1(t_1,t_2) = -\frac{1}{n^2}\sum_{j,k=1}^{n} \sin\left((\widehat{U}_{j1} - \widehat{U}_{k1})\,t_1\right)\sin\left((\widehat{U}_{j2} - \widehat{U}_{k2})\,t_2\right).$$

$$\epsilon_2(t_1,t_2) = -\frac{1}{n^4}\sum_{j,k,s,\ell=1}^{n} \sin\left((\widehat{U}_{j1} - \widehat{U}_{k1})\,t_1\right)\sin\left((\widehat{U}_{s2} - \widehat{U}_{\ell 2})\,t_2\right).$$

and

$$\epsilon_3(t_1,t_2) = \frac{1}{n^3}\sum_{j,k,s=1}^{n} \sin\left((\widehat{U}_{j1} - \widehat{U}_{k1})\,t_1\right)\sin\left((\widehat{U}_{j2} - \widehat{U}_{s2})\,t_2\right).$$

Now, since the weighted function ω_0 is symmetric and the fact that $\epsilon_r(t_1,t_2) = -\epsilon_r(-t_1,t_2) = -\epsilon_r(t_1,-t_2)$ for any $r = 1,2,3$, one has

$$\int_\gamma \Im\left\{\Delta_n(t_1,t_2)\right\}\,\omega_0(t_1,t_2)\,\mathrm{d}t_1\,\mathrm{d}t_2 = 0.$$

Using now the identities

$$\cos(a)\cos(b) = 1 - (1-\cos(a)) - (1-\cos(b)) + (1-\cos(a))\ (1-\cos(b))\,,$$

one has $\widehat{\Lambda}^{\gamma}_{n,\omega_0} =$

$$\frac{1}{n^2}\sum_{j,k=1}^{n}\left\{1-\cos\left((\widehat{U}_{j1}-\widehat{U}_{k1})t_1\right)\right\}\left\{1-\cos\left((\widehat{U}_{j2}-\widehat{U}_{k2})t_2\right)\right\}$$
$$+\frac{1}{n^4}\sum_{j,k,s,l=1}^{n}\left\{1-\cos\left((\widehat{U}_{j1}-\widehat{U}_{k1})t_1\right)\right\}\left\{1-\cos\left((\widehat{U}_{s2}-\widehat{U}_{l2})t_2\right)\right\}$$
$$-\frac{2}{n^3}\sum_{j,k,s=1}^{n}\left\{1-\cos\left((\widehat{U}_{j1}-\widehat{U}_{k1})t_1\right)\right\}\left\{1-\cos\left((\widehat{U}_{j2}-\widehat{U}_{s2})t_2\right)\right\}.$$

Further, from the measure μ_{ω_0} given in Eq. 11, for any $(\mathbf{a},\mathbf{b}) \in [0,1]^2$, one gets expression given in the Lemma.2.

A.4 Proof of Proposition 2

First of all, let define for any $(u_1,u_2) \in [0,1]^2$ and $(t_1,t_2) \in \mathbb{R}^2$, the function $g(u_1,u_2) = \exp\{i\,t_1\,u_1 + i\,t_2\,u_2\}$ and its derivatives $g_k = \partial g(u_1,u_2)/\partial u_k$ at u_k and $g_{k,k'} = \partial^2 g(u_1,u_2)/\partial u_k \partial u_{k'}$, for all $k,k' = 1,2$. The mean-value theorem shows then that

$$\begin{aligned} g\left(\widehat{U}_{j1},\widehat{U}_{j2}\right) &= g(U_{j1},U_{j2}) + g_1\left(U_{j1},U_{j2}\right)\left(\widehat{U}_{j1}-U_{j1}\right) \\ &\quad + g_2\left(U_{j1},U_{j2}\right)\left(\widehat{U}_{j2}-U_{j2}\right) + e_{nj}, \end{aligned}$$

where for $U^{\star}_{j\ell}$ between $\widehat{U}_{j\ell}$ and $U_{j\ell}$, for $\ell = 1,2$, one has

$$\begin{aligned} e_{nj} &= \frac{1}{2}\left\{g_{1,1}(\widehat{U}^{\star}_{j1},\widehat{U}^{\star}_{j2})\left(\widehat{U}_{j1}-U_{j1}\right)^2 + g_{2,2}(\widehat{U}^{\star}_{j1})\left(\widehat{U}_{j2}-U_{j2}\right)^2\right. \\ &\quad \left. +2\,g_{1,2}\left(\widehat{U}^{\star}_{j1},\widehat{U}^{\star}_{j2}\right)\left(\widehat{U}_{j1}-U_{j1}\right)\left(\widehat{U}_{j2}-U_{j2}\right)\right\}. \end{aligned}$$

Since

$$\widehat{U}_{j1} = \frac{1}{n}\sum_{k=1}^{n}\mathbb{I}\left(U_{k1}\le U_{j1}\right) \quad \text{and} \quad \widehat{U}_{j2} = \frac{1}{n}\sum_{k=1}^{n}\mathbb{I}\left(U_{k2}\le U_{j2}\right).$$

The distance covariance $\Lambda^{\gamma}_{n,\omega_0}$ can be expressed as

$$\begin{aligned}
\widehat{\Lambda}^{\gamma}_{n,\omega_0} &= \int_{\gamma} \left|\Psi_n(t_1,t_2) - \psi_{n1}(t_1)\,\psi_{n2}(t_2)\right|^2 \omega_0(\mathbf{t})\,\mathrm{d}\mathbf{t} \\
&= \int_{\gamma} \left| \frac{1}{n}\sum_{j=1}^{n} e^{\left(it_1\widehat{U}_{j1}+it_2\widehat{U}_{j2}\right)} - \frac{1}{n^2}\sum_{j,j'=1}^{n} e^{\left(it_1\widehat{U}_{j1}+it_2\widehat{U}_{j'2}\right)} \right|^2 \omega_0(\mathbf{t})\mathrm{d}\mathbf{t} \\
&= \int_{\gamma} \left| \frac{1}{n^2}\sum_{j,j'=1}^{n} \left(e^{\left(it_1\widehat{U}_{j1}+it_2\widehat{U}_{j2}\right)} - e^{\left(it_1\widehat{U}_{j1}+i\,t_2\widehat{U}_{j'2}\right)} \right) \right|^2 \omega_0(\mathbf{t})\mathrm{d}\mathbf{t} \\
&= \int_{\gamma} \left| \frac{1}{n^2}\sum_{j,j'=1}^{n} \left\{ g\left(\widehat{U}_{j1},\widehat{U}_{j2}\right) - g\left(\widehat{U}_{j1},\widehat{U}_{j'2}\right) \right\} \right|^2 \omega_0(\mathbf{t})\,\mathrm{d}\mathbf{t} \\
&= \int_{\gamma} \left| \frac{1}{n^3}\sum_{j,j',k=1}^{n} L_{\mathbf{t}}\left(\mathbf{U}_j,\mathbf{U}_{j'},\mathbf{U}_k\right) + \frac{1}{n^2}\sum_{j,j'=1}^{n} \left(e_{nj}+e_{nj'}\right) \right|^2 \omega_0(\mathbf{t})\mathrm{d}\mathbf{t}.
\end{aligned}$$

Using now $|a+b|^2 = |a|^2+|b|^2+2\,\Re(a\,\bar{b})$, for any complex numbers a and b, it follows that $\Lambda^{\gamma}_{n,\omega_0} = V^{\gamma}_{n,\omega_0} + W^{\gamma}_{n1,\omega_0} + 2\,W^{\gamma}_{n2,\omega_0}$, for $\mathbf{U}=(U_1,U_2)\sim C$ and $\mathbf{t}=(t_1,t_2)\in\mathbb{R}^2$ where

$$V^{\gamma}_{n,\omega_0} = \frac{1}{n^6}\sum_{j,j',k,k',l,l',m=1}^{n} \int_{\gamma} L_{\mathbf{t}}\left(\mathbf{U}_j,\mathbf{U}_{j'},\mathbf{U}_k\right) L_{\mathbf{t}}\left(\mathbf{U}_{k'},\mathbf{U}_l,\mathbf{U}_{l'}\right) \omega_0(\mathbf{t})\,\mathrm{d}\mathbf{t},$$

$$\begin{aligned}
W^{\gamma}_{n1,\omega_0} &= \int_{\gamma} \left(\frac{1}{n^2}\sum_{j,j'=1}^{n} \left(e_{nj}+e_{nj'}\right) \right)^2 \omega_0(\mathbf{t})\,\mathrm{d}\mathbf{t} \\
&= \int_{\gamma} \left(\frac{1}{n^2}\sum_{j=1}^{n} e_{nj} \right)^2 \omega_0(\mathbf{t})\,\mathrm{d}\mathbf{t} + \int_{\gamma} \left(\frac{1}{n^2}\sum_{j'=1}^{n} e_{nj'} \right)^2 \omega_0(\mathbf{t})\,\mathrm{d}\mathbf{t} \\
&\quad +2\int_{\gamma} \left(\frac{1}{n^2}\sum_{j=1}^{n} e_{nj} \right)\left(\frac{1}{n^2}\sum_{j'=1}^{n} e_{nj'} \right) \omega_0(\mathbf{t})\,\mathrm{d}\mathbf{t} \\
&= W^{\gamma,(1)}_{n1} + W^{\gamma,(2)}_{n1} + 2W^{\gamma,(3)}_{n1}
\end{aligned}$$

and

$$W^{\gamma}_{n2,\omega_0} = 2\int_{\gamma} \Re\left\{ \left(\frac{1}{n^3}\sum_{j,j',k=1}^{n} L_{\mathbf{t}}\left(\mathbf{U}_j,\mathbf{U}_{j'},\mathbf{U}_k\right) \right) \overline{\left(\frac{1}{n^2}\sum_{j,j'=1}^{n} \left(e_{nj}+e_{nj'}\right) \right)} \right\} \omega_0(\mathbf{t})\mathrm{d}\mathbf{t}.$$

Using the fact that $g_{1,2}(u_1,u_2) = \partial g(u_1,u_2)/\partial u_1\, \partial u_2 = -t_1\, t_2\, g(u_1,u_2)$, one has

$$|e_{nj}(t_1,t_2)| \;\leq\; \frac{1}{2}\left\{|t_1|^2\left(\widehat{U}_{j1}-U_{j1}\right)^2 + |t_2|^2\left(\widehat{U}_{j2}-U_{j2}\right)^2 + 2|t_1||t_2|\left(\widehat{U}_{j1}-U_{j1}\right)\left(\widehat{U}_{j2}-U_{j2}\right)\right\}.$$

Since $\widehat{U}_{j\ell} - U_{j\ell} = O_{\mathbb{P}}(n^{-1/2})$, for any $\ell = 1,2$ and for each $j \in \{1,\ldots,n\}$, one can then conclude that

$$|e_{nj}(t_1,t_2)| = (|t_1|+|t_2|)^2\; O_{\mathbb{P}}(n^{-1}).$$

Then, since $\int_\gamma (|t_1|+|t_2|)^4\, \omega_0(\mathbf{t})\,\mathrm{d}\mathbf{t} < \infty$, one has

$$W_{n1}^{\gamma,(1)} = O_{\mathbb{P}}(n^{-1}) \int_\gamma (|t_1|+|t_2|)^4\, \omega_0(\mathbf{t})\,\mathrm{d}\mathbf{t} = o_{\mathbb{P}}(1).$$

The same arguments holds for $W_{n1}^{\gamma,(2)} = o_{\mathbb{P}}(1)$. For the last term, is a consequence of Cauchy-Schwartz inequality yields $W_{n1}^{\gamma,(3)} = o_{\mathbb{P}}(1)$. Therefore, $W_{n1,\omega_0}^{\gamma} = o_{\mathbb{P}}(1)$. Once again one apply a Cauchy-Schwartz inequality for square integrable complex functions f and g:

$$\left|\int_{\mathbb{R}^2} f(\mathbf{t})\,\overline{g(\mathbf{t})}\,\mathrm{d}\mathbf{t}\right| \leq \left(\int_{\mathbb{R}^2} |f(\mathbf{t})|^2\,\mathrm{d}\mathbf{t}\right)^{1/2} \left(\int_{\mathbb{R}^2} |g(\mathbf{t})|^2\,\mathrm{d}\mathbf{t}\right)^{1/2}.$$

That is

$$\left|W_{n2,\omega_0}^{\gamma}\right| \leq \sqrt{V_{n,\omega_0}^{\gamma}\, W_{n1,\omega_0}^{\gamma}},$$

and using the fact that $|\Re(z)| \leq |z|$ for any complex number $z = a + i\,b$, with combination to the last statement and that V_{n,ω_0}^{γ} is asymptotically equivalent to V-statistic, that is, converges in distribution, one obtains $W_{n2,\omega_0}^{\gamma} = o_{\mathbb{P}}(1)$, which complete the proof.

A.5 Proof of Proposition 3

Since $\widetilde{\Phi}_{\omega_0}^{\gamma}(\mathbf{u}) = \mathrm{E}\left\{\Phi_{\omega_0}^{\gamma}\left(\mathbf{u},\mathbf{U}_2,\mathbf{U}_3,\mathbf{U}_4,\mathbf{U}_5,\mathbf{U}_6\right)\right\} \neq 0$, the kernel $\Phi_{\omega_0}^{\gamma}$ is non–degenerate. Assume that $\sum_{k=2}^{\infty} \mathrm{Cov}\left(\widetilde{\Phi}_{\omega_0}^{\gamma}(\mathbf{U}_1), \widetilde{\Phi}_{\omega_0}^{\gamma}(\mathbf{U}_k)\right)$ are bounded. Then from Theorem 2 of Newman (1980) entails that the empirical distance covariance process converge weakly to an independent Gaussian random variables with mean zero and variance

$$\widetilde{\tau}_{\omega_0}^2 = 6^2\left(\mathrm{Var}\left(\widetilde{\Phi}_{\omega_0}^{\gamma}(\mathbf{U}_1)\right) + 2\sum_{k=2}^{\infty}\mathrm{Cov}\left(\widetilde{\Phi}_{\omega_0}^{\gamma}(\mathbf{U}_1), \widetilde{\Phi}_{\omega_0}^{\gamma}(\mathbf{U}_k)\right)\right).$$

As a conjecture, if the condition $\sum_{k=2}^{\infty} \mathrm{Cov}\left(\widetilde{\Phi}^{\gamma}_{\omega_0}(\mathbf{U}_1), \widetilde{\Phi}^{\gamma}_{\omega_0}(\mathbf{U}_k)\right)$ does not hold, there exist slowly varying function

$$L(z) \equiv 6^2 \left(\mathrm{Var}\left(\widetilde{\Phi}^{\gamma}_{\omega_0}(\mathbf{U}_1)\right) + 2\sum_{k=2}^{z} \mathrm{Cov}\left(\widetilde{\Phi}^{\gamma}_{\omega_0}(\mathbf{U}_1), \widetilde{\Phi}^{\gamma}_{\omega_0}(\mathbf{U}_k)\right)\right)$$

with

$$\frac{L(\alpha z)}{L(z)} \underset{z\to\infty}{\longrightarrow} 1, \quad \text{for any } \alpha > 0$$

such that the empirical distance covariance process

$$\mathbb{H}_{n,\omega_0}(\mathbf{U}) = \sqrt{n\,L(z)}\,\left(\widehat{\Lambda}^{\gamma}_{n,\omega_0} - \theta(\Psi_C)(\mathbf{U})\right), \qquad \mathbf{U} = (U_1, U_2) \sim C$$

does not converge in distribution to a Gaussian random variables. A detailed discussion can be found in Herrndorf (1984).

A.6 Proof of Proposition 4

First for $\mathbf{U} = (U_1, U_2) \sim C$, $\mathbf{U}_\ell = (\mathbf{U}_1, \mathbf{U}_2) = ((U_1, U_1), (U_2, U_2)) \sim C$ and as a consequence of Proposition 3 the empirical distance process

$$\begin{pmatrix} \mathbb{H}_{n,\omega_0}(\mathbf{U}) \\ \mathbb{H}_{n,\omega_0}(\mathbf{U}_1) \\ \mathbb{H}_{n,\omega_0}(\mathbf{U}_2) \end{pmatrix} \xrightarrow[n\to\infty]{d} N_3\left(0, 4\,\Omega\right)$$

a three-variate centered Gaussian distribution, where $\Omega \in \mathbb{R}^{3\times 3}$ is the covariance matrix of random vector $(\Phi_{\omega_0}, \Phi_{1,\omega_0}, \Phi_{2,\omega_0})^T$. Hence, by delta-method and we assume that the covariance $\sum_{k=2}^{\infty} \tau^2_{k,\omega_0}$ is finite, one has

$$\mathbb{R}^{\gamma}_n(\mathbf{U}) = \sqrt{n}\left\{\widehat{R}^{\gamma}_n(\widehat{\mathbf{U}}_j) - R^{\gamma}(\mathbf{U})\right\} \xrightarrow[n\to\infty]{d} N\left(0, \tau^2_{\omega_0}\right)$$

where $\tau^2_{\omega_0} = 4\,(\nabla\varphi)^T\,\Omega(\nabla\varphi)$ and $(\nabla\varphi) \in \mathbb{R}^{3\times 1}$ the gradient of φ at the point

$$\left(\Lambda^{\gamma}_{\omega_0}(C), \Lambda^{\gamma}_{\omega_0,1}, \Lambda^{\gamma}_{\omega_0,2}\right)^T.$$

Second, let

$$Z^{\gamma}_{n2} = \widetilde{Z}^{\gamma}_{n2,1}\,\widetilde{Z}^{\gamma}_{n2,2} = \frac{n\,\widehat{\sigma}^2_{n,\omega_0}}{\widetilde{\rho}_{\omega_0}},$$

where for every $\ell = 1, 2$, $\widetilde{Z}^{\gamma}_{n2,\ell} = \sqrt{n}\,\widehat{\Lambda}^{\gamma}_{n,\omega_0,\ell}/\widetilde{\rho}_{\ell,\omega_0}$ is the empirical distance variance process that converge in distribution, as n tends to infinity, to independent standard normal distribution. Then, one has Z^{γ}_{n2} converge in distribution, as n tends to infinity, to independent chi-squared random variables with one degree of freedom.

If Z_{n1}^{γ} and Z_{n2}^{γ} are independent i.e., $\mathbb{P}\,(Z_{n1}^{\gamma} = Z_{n2}^{\gamma}) = 1$, it follows by continuous mapping Theorem that the functional empirical distance covariance

$$R_n^{\gamma} = \varphi\,(Z_{n1}^{\gamma}, Z_{n2}^{\gamma}) = \frac{\sqrt{n}\,\widehat{\Lambda}_{n,\omega_0}^{\gamma}}{\widetilde{\tau}_{\omega_0}} \left\{ \frac{n\,\widehat{\sigma}_{n,\omega_0,\ell)}^{2}}{\widetilde{\rho}_{\omega_0}} \right\}^{-1/2}$$

converge in distribution, as $n \to \infty$, to limit $R = \varphi\,(Z_1, Z_2)$ where $\varphi(z_1, z_2) = z_1/\sqrt{z_2}$, that is, Student's t-distribution with one degree of freedom.

Tarik Bahraoui
Nikolai Kolev
Department of Statistics, University
of São Paulo, São Paulo, Brazil
E-mail: tarik.bahraoui@gmail.com
kolev.ime@gmail.com

Paper received: 25 October 2019.