JOURNAL OF COMPUTATIONAL AND GRAPHICAL STATISTICS
2025, VOL. 34, NO. 3, 1140–1152
https://doi.org/10.1080/10618600.2024.2432978

Visualization and Assessment of Copula Symmetry

Cristian F. Jiménez-Varón[a,b], Hao Lee[c], Marc G. Genton[d], and Ying Sun[d]

[a]Department of Mathematics, University of York, York, UK; [b]Department of Physics and Mathematics, Universidad Autónoma de Manizales, Manizales, Colombia; [c]Statistics & Data Science, Dietrich College of Humanities and Social Sciences, Carnegie Mellon University, Pittsburgh, PA; [d]Statistics Program, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia

ABSTRACT

Visualization and assessment of copula structures are crucial for accurately understanding and modeling the dependencies in multivariate data analysis. In this article, we introduce an innovative method that employs functional boxplots and rank-based testing procedures to evaluate copula symmetries. This approach is specifically designed to assess key characteristics such as reflection symmetry, radial symmetry, and joint symmetry. We first construct test functions for each specific property and then investigate the asymptotic properties of their empirical estimators. We demonstrate that the functional boxplot of these sample test functions serves as an informative visualization tool of a given copula structure, effectively measuring the departure from zero of the test function. Furthermore, we introduce a nonparametric testing procedure to assess the significance of deviations from symmetry, ensuring the accuracy and reliability of our visualization method. Through extensive simulation studies involving various copula models, we demonstrate the effectiveness of our testing approach. Finally, we apply our visualization and testing techniques to three real-world datasets: a nutritional habits survey with five variables, stock price data for the five top companies in the NASDAQ-100 stock index, and two major stock indices, the US S&P500 and German DAX. Supplementary materials for this article are available online.

ARTICLE HISTORY
Received December 2023
Accepted November 2024

KEYWORDS
Copula structure; Functional boxplot; Rank-based testing; Symmetry; Visualization

1. Introduction

Copula models have gained significant prominence in the field of statistics and data analysis due to their flexibility in modeling intricate dependence structures among random variables (Nelsen 2006; Joe 2014; Patton 2012). They have become indispensable tools for capturing and understanding various types of dependencies, such as tail dependence, asymmetry, and nonlinearity (Genest and Favre 2007; Cherubini, Luciano, and Vecchiato 2004). By decoupling the marginal distributions from the dependence structure, copula models offer a powerful framework for accurately characterizing complex multivariate relationships (Cherubini, Luciano, and Vecchiato 2004; Joe 2014). In addition to their theoretical significance, copula models have found extensive real-world applications in finance, insurance, and environmental sciences. In finance, for instance, copula models facilitate portfolio optimization, risk management, and pricing of complex financial derivatives by accurately modeling dependencies between financial assets (Cherubini, Luciano, and Vecchiato 2004; Patton 2012).

According to the representation theorem provided by Sklar (1959), every multivariate cumulative distribution function, F, of a continuous random vector $\mathbf{X} = (X_1, \ldots, X_d)^\top$ on $\mathbb{R}^d$, can be written as

$$\begin{aligned} F(x_1, \ldots, x_d) &= \mathbb{P}(X_1 \leq x_1, \ldots, X_d \leq x_d) \\ &= C\{F_1(x_1), \ldots, F_d(x_d)\}, \end{aligned} \tag{1}$$

where $F_l(x_l) = \mathbb{P}(X_l \leq x_l)$, $x_l \in \mathbb{R}$, are the continuous marginal distributions, $l = 1, \ldots, d$, and $C : [0,1]^d \to [0,1]$ is the unique copula that characterizes the dependence structure of the random vector $\mathbf{X}$ and can be obtained from

$$\begin{aligned} C(u_1, \ldots, u_d) &= \mathbb{P}(U_1 \leq u_1, \ldots, U_d \leq u_d) \\ &= F\{F_1^{-1}(u_1), \ldots, F_d^{-1}(u_d)\}, \end{aligned} \tag{2}$$

where $F_l^{-1}(u_l) = \inf\{x | F_l(x) \geq u_l\}$, $u_l \in [0,1]$, is the quantile function of F_l and $\mathbf{U} = (U_1, \ldots, U_d)^\top$ with $U_l = F_l(X_l)$. From (2), the copula function C serves as a cumulative distribution function for the random vector $\mathbf{U}$, residing within the d-dimensional unit hypercube and characterized by its marginal distributions. In practical applications, the representation provided in (1) enables the modeling of the dependence structure, given the knowledge of the marginal distributions. This can be achieved by selecting an appropriate parametric copula model from a wide range of options available in the literature (see e.g., Nelsen 2006; Joe 2014).

Choosing an appropriate copula model is a challenging task when quantifying dependence. In various practical applications, such as actuarial science, finance, and survival analysis (Nelsen 2006; Patton 2006; Aas et al. 2009), the common approach has been to rely on expert knowledge or choose a copula model based on mathematical convenience rather than its suitability for the specific data application. However, this approach can

CONTACT Cristian F. Jiménez-Varón cristian.jimenezvaron@york.ac.uk Department of Mathematics, University of York, York, UK.
Supplementary materials for this article are available online. Please go to www.tandfonline.com/r/JCGS.

introduce limitations and biases in the analysis (Mikosch 2006; Nelsen 2006; Aas et al. 2009; Joe 2014).

Several existing approaches in the literature for copula model selection are based on goodness-of-fit tests for copulas (Genest and Favre 2007; Genest, Rémillard, and Beaudoin 2009; Berg 2009). These methods typically treat the univariate marginal distribution as an infinite-dimensional nuisance parameter and replace the observations with maximally invariant statistics, such as ranks.

Understanding the properties and structure of copulas is crucial for capturing and interpreting the relationships between random variables. Various methods have been proposed in the literature to specify and test copula structures. Jaworski (2010) introduced a test for the associativity structure of copulas based on the asymptotic distribution of the pointwise copula estimator. However, this test only assesses associativity at a specific point rather than for the entire copula process, as discussed in Bücher, Dette, and Volgushev (2012). Bücher, Dette, and Volgushev (2012) derived Cramér-von Mises and Kolmogorov-Smirnov type test statistics for evaluating the characteristics of associativity. Additionally, they developed test statistics for Archimedean copulas (Bücher, Dette, and Volgushev 2012). Bücher, Dette, and Volgushev (2011) proposed a test for extreme value dependence based on the minimum weighted L^2-distance of extreme-value copulas. The bivariate symmetry test for copulas, based on Cramér-von Mises and Kolmogorov-Smirnov functionals of the rank-based empirical copula process, was introduced by Genest, Nešlehová, and Quessy (2012) and Genest and Nešlehová (2014).

Li and Genton (2013) proposed a nonparametric method for identifying copula symmetries using the asymptotic distribution of the empirical copula process. Quessy (2016) developed a statistical framework based on quadratic functionals to test the identity of copulas from a multivariate distribution. More recently, Jaser and Min (2021) proposed simpler nonparametric tests for the symmetry and radial symmetry of bivariate copulas. Their approach involves creating two bivariate samples by manipulating the underlying copula while preserving its dependence structure. The test statistics are based on the difference between the empirical Kendall's tau of both samples.

In this article, we present a new approach for visualizing and testing the structure of copula models, specifically focusing on the dependence properties such as symmetry, radial symmetry, and joint symmetry as defined in Nelsen (1993). Our approach complements existing goodness-of-fit tests for copula model selection. To visualize these copula structures, we employ the functional boxplot introduced by Sun and Genton (2011) as a visual tool to quantify the deviations from a given copula structure by measuring the departure from zero of sample test functions. We demonstrate that these visualizations offer insights into the extent to which specific copula structures are adhered to.

Additionally, we introduce a nonparametric testing procedure to assess the significance of deviations from symmetry. This testing procedure is motivated by the techniques proposed by Huang and Sun (2019) and Huang, Sun, and Genton (2023), which use a functional data framework to visualize and assess spatio-temporal covariance properties in both univariate and multivariate cases. We evaluate the effectiveness of our proposed testing approach through extensive simulation studies involving various copula models.

The article is structured as follows. In Section 2, we outline the copula symmetries of interest, including reflection symmetry, radial symmetry, and joint symmetry. We also provide details on the visualization and nonparametric testing procedures for each of these copula structures. Section 3 presents the simulation results regarding the size and power of our proposed nonparametric test. In Section 4, we apply our methods to three real-world datasets: a nutritional habits survey with five variables, stock price data for the five top companies in the NASDAQ-100 stock index, and two major stock indices, the US S&P500 and German DAX. Finally, the article concludes with a discussion in Section 5.

2. Methodology

In Section 2.1, we introduce copula symmetries. Section 2.2 covers the construction of test functions and provides asymptotic results for proper estimators. We present the visualization of test functions with functional boxplots in Section 2.3. Lastly, in Section 2.4, we describe a rank-based testing procedure for copula symmetries.

2.1. Copula Symmetries

Our discussion centers on the symmetry of bivariate copulas. Unlike the case of univariate functions, the concept of symmetry is not uniquely defined in a multivariate setting. Therefore, different notions of symmetry have been investigated in the context of copulas. Here we focus on the ones presented in Nelsen (1993).

Definition 1. A copula C is said to be *symmetric* if

$$C(u,v) - C(v,u) = 0, \forall (u,v) \in [0,1]^2. \tag{3}$$

Based on (3), one should notice that for any symmetric copula C, its distribution is symmetric with respect to the diagonal connecting the origin and the point $(1,1)$. Thus, the symmetry in Definition 1 is called *reflection symmetry* in some literature and we will employ this term to refer to this particular form of symmetry in the subsequent discussions.

Definition 2. A copula C is said to be *radially symmetric* if

$$C(u,v) - C(1-u,1-v) + 1 - u - v = 0, \forall (u,v) \in [0,1]^2. \tag{4}$$

Equivalently, one can state the radial symmetry property as $C(u,v) - C^*(u,v) = 0$ for all $(u,v) \in [0,1]^2$, where C^* stands for the survival copula associated with C; that is, for all $(u,v) \in [0,1]^2$, $C^*(u,v) = C(1-u,1-v) - 1 + u + v$. As pointed out by Nelsen (1993), there exist copulas that are reflection symmetric but not radially symmetric and, conversely, copulas that are radially symmetric but not reflection symmetric.

Definition 3. A copula C is said to be *jointly symmetric* if it satisfies

$$C(u,v) + C(u,1-v) - u = 0 \quad \text{and}$$
$$C(u,v) + C(1-u,v) - v = 0, \forall (u,v) \in [0,1]^2.$$

One can easily show that joint symmetry implies radial symmetry. However, there is no implication between reflection symmetry and joint symmetry. The Figure 1 of Li and Genton (2013) provides the interrelations among the three types of symmetry.

2.2. Test Functions

We propose to assess and visualize the three types of copula symmetries by the construction of test functions. First, we focus on the construction of test functions specifically for reflection symmetry (S). For any fixed $v \in [0,1]$, define the reflection symmetry test functions $f_v^S : [0,1] \to [-1,1]$ by

$$f_v^S(t) = C(t,v) - C(v,t)$$

for all $t \in [0,1]$. If C is reflection symmetric, we have $f_v^S = 0$; otherwise, the values of f_v^S vary with respect to t. For all $m \in \mathbb{N}$, any set $\{v_1,\dots,v_m,w_1,\dots,w_m\} \subseteq [0,1]$, we introduce $2m$ reflection symmetry test functions $f_{v_1}^S,\dots,f_{v_m}^S$ and $g_{w_1}^S,\dots,g_{w_m}^S$, where $g_{w_i} = -f_{w_i}$. This way, we create more test functions to better measure the departure from zero.

In practice, we need corresponding estimators $\widehat{f}_{v_1}^S,\dots,\widehat{f}_{v_m}^S$ and $\widehat{g}_{w_1}^S,\dots,\widehat{g}_{w_m}^S$ of the test functions $f_{v_1}^S,\dots,f_{v_m}^S$ and $g_{w_1}^S,\dots,g_{w_m}^S$. Intuitively, the estimators can be obtained through a linear combination of the corresponding empirical copulas. To provide a clearer definition of these estimators, we briefly summarize the important asymptotic results of the two-dimensional empirical copula process (see, e.g., Deheuvels 1979; Stute 1984; Tsukahara 2005). Additional recent results on convergence rate can be found in Genest and Segers (2010), Segers (2012), Swanepoel and Allison (2013), and the references therein.

Let C be a two-dimensional copula and $\boldsymbol{U} = (U_1,U_2)^\top$ be a random vector that is C-distributed. If one has direct access to a random sample $\boldsymbol{U}_1,\dots,\boldsymbol{U}_n$ of size n, then the copula C can be estimated by the empirical copula process defined by

$$\widehat{C}_n(u,v) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}\{U_{i1} \le u, U_{i2} \le v\}, \forall (u,v) \in [0,1]^2,$$

where $\mathbb{1}(\cdot)$ denotes the indicator function. As a well-known result, we have that for all $(u,v) \in [0,1]^2$,

$$\begin{aligned}\mathbb{C}_n(u,v) &:= \sqrt{n}\left\{\widehat{C}_n(u,v) - C(u,v)\right\} \\ &\xrightarrow{d} \mathbb{C}(u,v) \text{ in } \ell^\infty([0,1]^2), n \to \infty, \end{aligned} \tag{5}$$

where $\ell^\infty([0,1]^2)$ denotes the space of all the bounded functions over the compact set $[0,1]^2$ and $\mathbb{C}$ is a two-dimensional pinned C-Brownian sheet, that is, it is a zero-mean Gaussian random field with the covariance function given by

$$\begin{aligned}\operatorname{cov}\{\mathbb{C}(u,v),\mathbb{C}(x,y)\} = C(u\wedge x, v\wedge y) - C(u,v)C(x,y),\\ \forall u,v,x,y \in [0,1]\end{aligned}$$

where for all $a,b \in \mathbb{R}$, $a \wedge b = \min\{a,b\}$ (Genest, Nešlehová, and Quessy 2012). However, it is often the case that the n observations we have are generated from a random vector $\boldsymbol{X} = (X_1,X_2)^\top$ that is not necessarily uniformly distributed over the interval $[0,1]$. The representation theorem (Sklar 1959) states that it can be expressed as the composition of a copula C and marginals of $\boldsymbol{X}$. In this case, from every $\boldsymbol{X}_i$ in the random sample, one can estimate a $\boldsymbol{U}_i$ by the pseudo-observation $\widehat{\boldsymbol{U}}_i = (\widehat{U}_{i1},\widehat{U}_{i2})^\top$, where for all $s \in \{1,2\}$,

$$\widehat{U}_{is} = \frac{1}{n}\sum_{r=1}^{n} \mathbb{1}\{X_{rs} \le X_{is}\}.$$

With all of these $\widehat{\boldsymbol{U}}_i$'s, one can estimate C by

$$\widehat{D}_n(u,v) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}\{\widehat{U}_{i1} \le u, \widehat{U}_{i2} \le v\}, \forall (u,v) \in [0,1]^2.$$

It has been shown that when C is *regular* (see Definition 1 in Genest, Nešlehová, and Quessy (2012) for example), or loosely speaking, when C is differentiable with continuous partials, we have

$$\begin{aligned}\widehat{\mathbb{D}}_n(u,v) &:= \sqrt{n}\left\{\widehat{D}_n(u,v) - C(u,v)\right\} \\ &\xrightarrow{d} \mathbb{C}(u,v) - \dot{C}_1(u,v)\mathbb{C}(u,1) \\ &\quad -\dot{C}_2(u,v)\mathbb{C}(1,v) \text{ in } \ell^\infty([0,1]^2), n \to \infty, \end{aligned} \tag{6}$$

where $\dot{C}_1$ and $\dot{C}_2$ denote the partial derivatives of the copula C with respect to its first and second variables, respectively. In the sequel, if not otherwise stated, we always impose the regularity assumption on the underlying copula C. To sum up, the estimator of the reflection symmetry test function $\widehat{f}_v^S$ can be defined by either

$$\widehat{f}_v^S(t) = \widehat{C}_n(t,v) - \widehat{C}_n(v,t), \forall t \in [0,1],$$

or

$$\widehat{f}_v^S(t) = \widehat{D}_n(t,v) - \widehat{D}_n(v,t), \forall t \in [0,1],$$

depending on the type of the dataset we have, and we simply set $\widehat{g}_v^S(t) = -\widehat{f}_v^S(t)$.

Proposition 1. Given any fixed $v \in [0,1]$, the estimator $\widehat{f}_v^S(t)$ satisfies that for all $t \in [0,1]$,

$$\begin{aligned}&\sqrt{n}\left\{\widehat{f}_v^S(t) - f_v^S(t)\right\} \\ &\xrightarrow{d} \begin{cases} \mathbb{E}_v^S(t) & \text{if } \widehat{f}_v^S(t) = \widehat{C}_n(t,v) - \widehat{C}_n(v,t) \\ \widehat{\mathbb{E}}_v^S(t) & \text{if } \widehat{f}_v^S(t) = \widehat{D}_n(t,v) - \widehat{D}_n(v,t) \end{cases}\end{aligned}$$

in $\ell^\infty([0,1])$ as $n \to \infty$, where $\mathbb{E}_v^S$ and $\widehat{\mathbb{E}}_v^S$ are two zero-mean Gaussian random fields.

The proof of this Proposition can be found in the supplementary material in Section S3.

Similarly, we can construct radial symmetry (R) test functions, denoted as $\widehat{f}_v^R(t)$ and $\widehat{g}_v^R(t)$, based on Definition 2 to estimate $f_v^R(t)$ and $g_v^R(t)$, as described in Table 1. The convergence of these estimators is presented in Proposition 2.

Proposition 2. Given any fixed $v \in [0,1]$, the estimator $\widehat{f}_v^R(t)$ satisfies that for all $t \in [0,1]$,

Table 1. Definitions of the estimators of the test functions for different types of symmetry.

Type	Access	Test functions
S	(U_1, U_2)	$\widehat{f}^S_v(t) = \widehat{C}_n(t,v) - \widehat{C}_n(v,t), \widehat{g}^S_v(t) = -\widehat{f}^S_v(t)$
	(X_1, X_2)	$\widehat{f}^S_v(t) = \widehat{D}_n(t,v) - \widehat{D}_n(v,t) \widehat{g}^S_v(t) = -\widehat{f}^S_v(t)$
R	(U_1, U_2)	$\widehat{f}^R_v(t) = \widehat{C}_n(t,v) - \widehat{C}_n(1-t,1-v) + 1 - t - v, \widehat{g}^R_v(t) = -\widehat{f}^R_v(t)$
	(X_1, X_2)	$\widehat{f}^R_v(t) = \widehat{D}_n(t,v) - \widehat{D}_n(1-t,1-v) + 1 - t - v, \widehat{g}^R_v(t) = -\widehat{f}^R_v(t)$
J	(U_1, U_2)	$\widehat{f}^{J,1}_v(t) = \widehat{C}_n(t,v) + \widehat{C}_n(t,1-v) - t, \widehat{g}^{J,1}_v(t) = -\widehat{f}^{J,1}_v(t)$
		$\widehat{f}^{J,2}_v(t) = \widehat{C}_n(t,v) + \widehat{C}_n(1-t,v) - v, \widehat{g}^{J,2}_v(t) = -\widehat{f}^{J,2}_v(t)$
	(X_1, X_2)	$\widehat{f}^{J,1}_v(t) = \widehat{D}_n(t,v) + \widehat{D}_n(t,1-v) - t, \widehat{g}^{J,1}_v(t) = -\widehat{f}^{J,1}_v(t)$
		$\widehat{f}^{J,2}_v(t) = \widehat{D}_n(t,v) + \widehat{D}_n(t,1-v) - t, \widehat{g}^{J,2}_v(t) = -\widehat{f}^{J,2}_v(t)$

$$\sqrt{n}\left\{\widehat{f}^R_v(t) - f^R_v(t)\right\} \xrightarrow{d} \begin{cases} \mathbb{E}^R_v(t) & \text{if } \widehat{f}^R_v(t) = \widehat{C}_n(t,v) - \widehat{C}_n(1-t,1-v) + 1 - t - v \\ \widehat{\mathbb{E}}^R_v(t) & \text{if } \widehat{f}^R_v(t) = \widehat{D}_n(t,v) - \widehat{D}_n(1-t,1-v) + 1 - t - v \end{cases}$$

in $\ell^\infty([0,1])$ as $n \to \infty$, where $\mathbb{E}^R_v$ and $\widehat{\mathbb{E}}^R_v$ are two zero-mean Gaussian random fields.

The proof of this Proposition can be found in the supplementary material in Section S4.

As for the test functions for joint symmetry (J), one should investigate the two properties in Definition 3 separately. Hence, for any given $v \in [0,1]$, we construct four population joint symmetry test functions: $f^{J,1}_v, f^{J,2}_v, g^{J,1}_v, g^{J,2}_v : [0,1] \to [-1,1]$ that are given by

$$f^{J,1}_v(t) = C(t,v) + C(t,1-v) - t \text{ and } g^{J,1}_v(t) = -f^{J,1}_v(t)$$

as well as

$$f^{J,2}_v(t) = C(1-t,v) + C(t,1-v) - v \text{ and } g^{J,2}_v(t) = -f^{J,2}_v(t)$$

for all $t \in [0,1]$. As the cases of other symmetries, they can be estimated by the test functions involving empirical copulas (see Table 1 for their definitions), whose relevant asymptotic results are presented in Proposition 3.

Proposition 3. Given $v \in [0,1]$, the estimator $\widehat{f}^{J,1}_v(t)$ satisfies that for all $t \in [0,1]$,

$$\sqrt{n}\left\{\widehat{f}^{J,1}_v(t) - f^{J,1}_v(t)\right\} \xrightarrow{d} \begin{cases} \mathbb{E}^{J,1}_v(t) & \text{if } \widehat{f}^{J,1}_v(t) = \widehat{C}_n(t,v) + \widehat{C}_n(t,1-v) - t \\ \widehat{\mathbb{E}}^{J,1}_v(t) & \text{if } \widehat{f}^{J,1}_v(t) = \widehat{D}_n(t,v) + \widehat{D}_n(t,1-v) - t \end{cases}$$

in $\ell^\infty([0,1])$ as $n \to \infty$, where $\mathbb{E}^{J,1}_v$ and $\widehat{\mathbb{E}}^{J,1}_v$ are two centered Gaussian random fields.

Similarly, given $v \in [0,1]$, the estimator $\widehat{f}^{J,2}_v(t)$ satisfies that for all $t \in [0,1]$,

$$\sqrt{n}\left\{\widehat{f}^{J,2}_v(t) - f^{J,2}_v(t)\right\} \xrightarrow{d} \begin{cases} \mathbb{E}^{J,2}_v(t) & \text{if } \widehat{f}^{J,2}_v(t) = \widehat{C}_n(t,v) + \widehat{C}_n(1-v,t) - v \\ \widehat{\mathbb{E}}^{J,2}_v(t) & \text{if } \widehat{f}^{J,2}_v(t) = \widehat{D}_n(t,v) + \widehat{D}_n(1-v,t) - v \end{cases}$$

in $\ell^\infty([0,1])$ as $n \to \infty$, where $\mathbb{E}^{J,2}_v$ and $\widehat{\mathbb{E}}^{J,2}_v$ are two zero-mean Gaussian random fields.

The proof of this Proposition can be found in the supplementary material in Section S5.

2.3. Visualization

The visualization of test functions can be achieved using functional boxplots (Sun and Genton 2011). The functional boxplot is an effective and concise method for visualizing and summarizing functional data. Constructing a functional boxplot starts with data ordering. Since each observation is a real function, functional boxplots order the functional data using a notion of data depth, such as band depth or modified band depth. The whiskers of the functional boxplot extend to the most extreme observations outside the central region, providing additional insights into the distributional characteristics of the functional data. In the R programming language, the **fda** package offers the **fbplot** function, which allows users to create functional boxplots conveniently. To construct the functional boxplot of $\widehat{f}^S_{v_1}, \ldots, \widehat{f}^S_{v_m}, \widehat{g}^S_{w_1}, \ldots, \widehat{g}^S_{w_m}$, we need to discretize the interval $[0,1]$ into an evenly spaced range of p points for some $p \in \mathbb{N}$ and then evaluate the test functions $\widehat{f}^S_{v_1}, \ldots, \widehat{f}^S_{v_m}, \widehat{g}^S_{w_1}, \ldots, \widehat{g}^S_{w_m}$ over them. We denote the range of p points by $\{0 \le t_1 \le \cdots \le t_p \le 1\}$. According to Proposition 1, for every $j \in \{1, \ldots, m\}$, the joint distribution of the random vector $\left(\widehat{f}^S_{v_j}(t_1) - f^S_{v_j}(t_1), \ldots, \widehat{f}^S_{v_j}(t_p) - f^S_{v_j}(t_p)\right)^\top$ is close to a multivariate normal distribution with mean zero when the sample size n is sufficiently large, and so is the joint distribution of the random vector $\left(\widehat{g}^S_{w_j}(t_1) - g^S_{w_j}(t_1), \ldots, \widehat{g}^S_{w_j}(t_p) - g^S_{w_j}(t_p)\right)^\top$. Thus, the functional boxplot should be fairly centered around zero (even though its shape should not be expected as a horizontal band since the variance at each point $t_k, k \in \{1, \ldots, p\}$, can vary based on the underlying copula C). If the underlying copula C is reflection symmetric, it implies the equalities $f^S_{v_1} \equiv \cdots \equiv f^S_{v_m} \equiv g^S_{w_1} \equiv \cdots \equiv g^S_{w_m} \equiv 0$. Otherwise, the functional boxplot should be of an irregular shape that is visually deviated from zero. In the special case when $v_i \approx w_i$ for a great proportion of $i \in \{1, \ldots, n\}$, the functional boxplot may form an irregular envelope that is almost symmetric with respect to zero.

To assess radial symmetry, we analyze the functional boxplots of $\widehat{f}^R_{v_1}, \ldots, \widehat{f}^R_{v_m}, \widehat{g}^R_{w_1}, \ldots, \widehat{g}^R_{w_m}$. According to Proposition 2, the functional boxplots should be fairly centered around zero if and only if the underlying copula C is indeed radial symmetric. For joint symmetry, two types of test functions are defined, resulting in two functional boxplots per copula. Proposition 3 states that both functional boxplots should be centered around zero if the underlying copula is jointly symmetric.

To stress the interpretation, we demonstrate the functional boxplots based on various copulas models in Figure 1. In the plots, the central region is colored green if a specific copula model satisfies the corresponding type of symmetry. On the other hand, if the copula model does not meet the symmetry condition, the central region is colored red. The gradation of the colors used in the plots manifests the density of functional data: the darker the colors are, the more functional curves are located in that place. The corresponding structure(s) of each copula is flagged by the abbreviation(s) in bold: **S** (reflection symmetry), **R** (radial symmetry), **J** (joint symmetry) and **AS** (asymmetry) in the title of their respective functional boxplots.

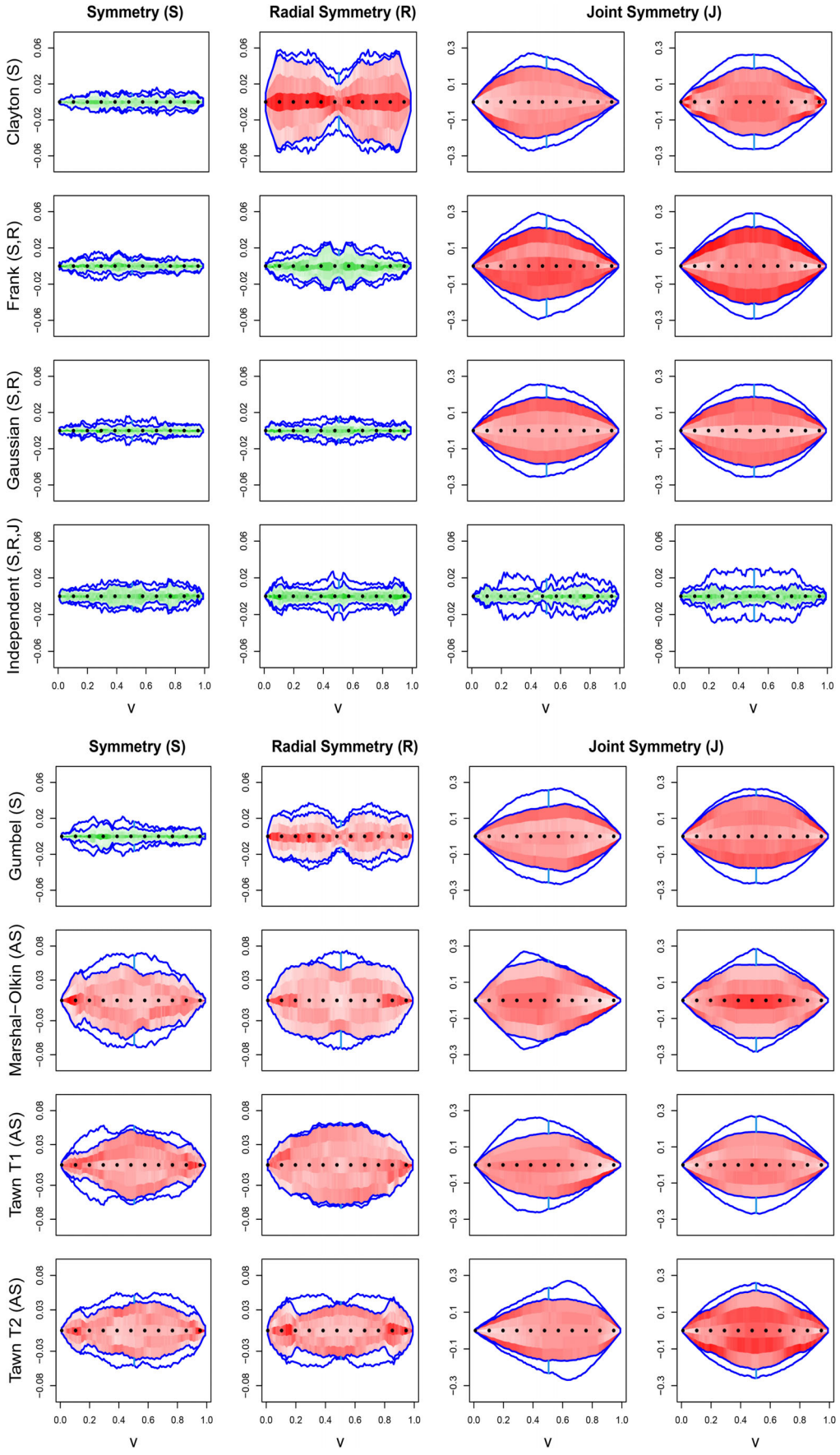

Figure 1. Functional boxplots for the visualization of copula symmetries. The parameter(s) of each copula is chosen to make its Kendall's tau equal/close to 0.5. Especially, the Marshall-Olkin copula takes the parameters (0.55, 0.85), and the Tawn copula, regardless of its types, takes the parameters (4.28, 0.60). The corresponding test functions are constructed with $n = 1000$, $m = 250$ and $p = 100$, and the values of v_j, w_j for all $j \in \{1, \ldots, m\}$ are randomly chosen from the unit interval $[0, 1]$.

2.4. Rank-based Hypothesis Testing Procedure

In Section 2.3, we explored the visualization of copula structures present in a given sample or dataset. We discussed the relevant asymptotic results and demonstrated how functional boxplots of selected test functions can provide intuitive indications of various symmetries. To thoroughly investigate and quantify these symmetries, we introduce a one-sample ranked-based hypothesis testing procedure. This procedure is a modification of the methods proposed by Huang and Sun (2019) and Huang, Sun, and Genton (2023) for testing the separability and symmetry of univariate and multivariate covariance functions. Our adapted approach allows us to assess the presence of symmetries in copula structures in a robust and statistically rigorous manner. Both of these methods can be considered as adaptations of the two-sample rank-based test proposed by López-Pintado and Romo (2009). The original test is designed to determine whether two sets of functional data are derived from the same distribution. In a similar manner, we modify this test to examine reflection symmetry as the initial step. By making specific adjustments, such as selecting appropriate estimators for the test functions and simulated distribution, the procedure can be extended to test for radial and joint symmetry.

Suppose that we have the observations $\boldsymbol{U}_i$'s or the pseudo-observations $\widehat{\boldsymbol{U}}_i$'s for $i = 1, \ldots, n$. As expected, for $v \in [0, 1]$ the null $\mathcal{H}_0$ and the alternative $\mathcal{H}_a$ are

- $\mathcal{H}_0 : f_v^S(t) = 0 \; \forall t \in [0, 1]$;
- $\mathcal{H}_a$: $\exists t \in [0, 1]$, s.t. $f_v^S(t) \neq 0$.

The details of the procedure are demonstrated as follows:

Step 1: Estimate the values of the reflection symmetry test functions $\widehat{f}^S_{v_1}, \ldots, \widehat{f}^S_{v_m}, \widehat{g}^S_{w_1}, \ldots, \widehat{g}^S_{w_m}$ over the evenly spaced points $\{0 \leq t_1, \leq \cdots \leq, t_p \leq 1\}$ as in Table 1 by $\boldsymbol{U}_i$'s, or $\widehat{\boldsymbol{U}}_i$'s, where for any $j \in \{1, \ldots, m\}$, v_j and w_j are randomly generated from the unit interval $[0, 1]$.

Step 2: Simulate from a reflection symmetric copula to obtain a set of n observations, $\boldsymbol{V}_i^{\mathcal{H}_0}, i \in \{1, \ldots, n\}$. Further details on how to simulate the observations $\boldsymbol{V}_i^{\mathcal{H}_0}$ are described in Section 2.4.1.

Step 3: Estimate the values of the reflection symmetry test functions $\widehat{f}^{\mathcal{H}_0}_{v_1}, \ldots, \widehat{f}^{\mathcal{H}_0}_{v_{m_0}}, \widehat{g}^{\mathcal{H}_0}_{w_1}, \ldots, \widehat{g}^{\mathcal{H}_0}_{w_{m_0}}$ over the evenly spaced points $\{0 \leq t_1 \leq \cdots \leq t_p \leq 1\}$ as in Table 1 by $\boldsymbol{V}_i^{\mathcal{H}_0}$'s, where for any $j^0 \in \{1, \ldots, m_0\}$, v_{j^0} and w_{j^0} are randomly generated from the unit interval $[0, 1]$.

Step 4: Combine the values of $\widehat{f}^S_{v_1}, \ldots, \widehat{f}^S_{v_m}, \widehat{g}^S_{w_1}, \ldots, \widehat{g}^S_{w_m}$ with those of $\widehat{f}^{\mathcal{H}_0}_{v_1}, \ldots, \widehat{f}^{\mathcal{H}_0}_{v_{m_0}}, \widehat{g}^{\mathcal{H}_0}_{w_1}, \ldots, \widehat{g}^{\mathcal{H}_0}_{w_{m_0}}$ to form a discretized functional dataset of size $2(m + m_0)$. Compute their modified band depths (see, e.g., López-Pintado and Romo 2009; Sun, Genton, and Nychka 2012).

Step 5: Rank the $2(m + m_0)$ test functions according to their depth values. In case of any ties, we assign distinct ordinal numbers at random to the test functions that compare equal in terms of the depth values. Suppose that $\widehat{f}^S_{v_1}, \ldots, \widehat{f}^S_{v_m}, \widehat{g}^S_{w_1}, \ldots, \widehat{g}^S_{w_m}$ are associated with the ranks $r_1, \ldots, r_{2m_0}$. Define the test statistic $W = \sum_{i=1}^{2m_0} r_i$.

The null hypothesis $\mathcal{H}_0$ is rejected when W is significantly small because it means that the test functions $\widehat{f}^S_{v_1}, \ldots, \widehat{f}^S_{v_m}, \widehat{g}^S_{w_1}, \ldots, \widehat{g}^S_{w_m}$ are more deviated from zero. The definition of the test statistic W here takes the essence of the Wilcoxon, or equivalently Mann-Whitney, test statistic. Hence, one can also deem the proposed hypothesis testing as a modification of the *two-sample* Wilcoxon rank-sum test for one functional dataset. The null distribution of W is estimated by N_b bootstrap samples of size n. More specifically, we generate N_b samples from the reflection symmetric copula in *Step 2*. For the bth sample, $b \in \{1, \ldots, N_b\}$, we regard it as a set of $\boldsymbol{U}_i$'s and follow the above procedure to calculate the corresponding test statistic W_b. Eventually, the null distribution of W is approximated by all of these test statistics: $W_1, \ldots, W_{N_b}$. For the details on how to carry out such a *one-sample* bootstrapping method in hypothesis testing, we refer to Section 16.4 in Efron and Tibshirani (1993).

Some features of the procedure are worth further discussion. First, the number of observations simulated from a reflection symmetric copula needs to be the same as the original sample size n. Otherwise, even if the test functions $\widehat{f}^S_{v_1}, \ldots, \widehat{f}^S_{v_m}, \widehat{g}^S_{w_1}, \ldots, \widehat{g}^S_{w_m}$ came from a reflection symmetric copula, they would still be more centered/deviated with respect to zero, thereby having greater depth values, compared to $\widehat{f}^{\mathcal{H}_0}_{v_1}, \ldots, \widehat{f}^{\mathcal{H}_0}_{v_{m_0}}, \widehat{g}^{\mathcal{H}_0}_{w_1}, \ldots, \widehat{g}^{\mathcal{H}_0}_{w_{m_0}}$. This is because larger or smaller sample sizes make empirical copulas better or worse approximations of the underlying true copulas.

Second, the values of p, m, and m_0 in *Steps 1* and *3* are hyperparameters that one can tune at will. Nonetheless, it is observed that setting $p = 100$ is sufficient, while m and m_0 have better to be large to guarantee the ideal empirical size and power of the test. Particularly, we notice that having m and m_0 greater than n yields better performances. Once both m and m_0 satisfy this property, their values tend to have little effect on its performance.

Third, the most notable difference between our procedure and those used in López-Pintado and Romo (2009), Huang and Sun (2019) and Huang, Sun, and Genton (2023) lies in the absence of a reference dataset. The ideas in their approaches originate with the hypothesis testing procedure designed by Liu and Singh (1993) that performs the detection of (dis)similarity in two multivariate distributions using *quality index*. Particularly, they introduced the treatment with a reference dataset to identify the change in population locations. In the context of functional data, it describes the scenario when the two samples of functional curves appear to form two separated bands, each of which comprises only curves from one group, as shown in Figure S1 in the supplementary material in Section S1.

If the reflection symmetry property is not satisfied, it is important to consider various scenarios when using these test functions to capture deviations from symmetry. In some cases, it may be challenging to quantify the departure from zero. For example, the test functions may exhibit two distinct bands, indicating a significant departure from zero for a substantial portion of the variable t. This is similar to the tests proposed by Liu and Singh (1993) and López-Pintado and Romo (2009) and illustrated in Figure S1 in the supplement. To address this situation, we introduce the testing functions $\widehat{g}^S_{w_1}, \ldots, \widehat{g}^S_{w_m}$ which are

rough reflections of the test functions $\widehat{f}^S_{v_1}, \ldots, \widehat{f}^S_{v_m}$ with respect to zero.

As one may see later, we actually generate the observations $\boldsymbol{V}_i^{\mathcal{H}_0}, i \in \{1, \ldots, n\}$ from a mixture (M) of empirical copulas in *Step 2* under the assumption that M is close to a truly reflection symmetric copula. Then, if we had to simulate a reference dataset, the empirical copula of the reference dataset might be more symmetric than that of the observations $\boldsymbol{V}_i^{\mathcal{H}_0}, i \in \{1, \ldots, n\}$ when the sample size n is not sufficiently large. This could lead to a not-so-small test statistic in the approach of Huang and Sun (2019) or Huang, Sun, and Genton (2023) as the test functions constructed from the reference dataset might tend to be slightly more centered around zero. Certainly, the absence of a reference dataset also saves memory storage.

2.4.1. Simulation under $\mathcal{H}_0$

Lastly, it is important to address the choice of the reflection symmetric copula in Step 2 and explain how to simulate from it. In order to maintain the nonparametric nature of the procedure, it is crucial to avoid selecting any parametric copula. In lieu, we propose to construct the estimated reflection symmetric copula $\widehat{C}^S$ given, for all (u, v), by either

$$\widehat{C}^S(u,v) = \frac{1}{2}\Big(\widehat{C}_n(u,v) + \widehat{C}_n(v,u)\Big)$$

or

$$\widehat{C}^S(u,v) = \frac{1}{2}\Big(\widehat{D}_n(u,v) + \widehat{D}_n(v,u)\Big),$$

depending on the types of dataset we have.

Proposition 4. Given any bivariate copula C, the copula $\tilde{C}$ defined, for all $(u, v) \in [0,1]^2$, by

$$\tilde{C}(u,v) = \frac{1}{2}\Big(C(u,v) + C(v,u)\Big)$$

is a reflection symmetric copula.

We prove this result in the supplementary material in Section S6.

As a mixture of two empirical copulas, one should have no difficulty simulating from $\widehat{C}^S$. Note that $\widehat{C}^S$ is not really a copula itself, but it serves as an appropriate estimation of the reflection symmetric copula $\tilde{C}$ (Rüschendorf 1976). Apart from the intuition suggested by Proposition 4, in the supplementary material in Section S2, the Figure S2 provides a visual assurance for the simulation method.

By closely following the aforementioned procedure and making appropriate adjustments, one can conduct tests for radial and joint symmetries. Specifically, in Step 2, a copula that exhibits radial or joint symmetry needs to be constructed. Table 2 provides a summary of the proposed estimated copulas for this purpose, and the following Propositions provide justification for these choices. To ensure clarity in our notation, we denote the survival copula of any copula C as C^*. Proofs of the Propositions 5 and 6 can be found in the supplementary material in Sections S7 and S8, respectively.

Proposition 5. Given any bivariate copula C, the copula $\tilde{C}$ defined, for all $(u, v) \in [0,1]^2$, by

$$\tilde{C}(u,v) = \frac{1}{2}\Big(C(u,v) + C^*(u,v)\Big)$$

is a radially symmetric copula.

Proposition 6. Given any copula C, the copula $\tilde{C}$ defined, for all $(u, v) \in [0,1]^2$, by

$$\tilde{C}(u,v) = \frac{1}{4}\Big(C(u,v)+u-C(u,1-v)+v-C(1-u,v)+C^*(u,v)\Big)$$

is a jointly symmetric copula.

3. Simulation Study

We performed simulations to evaluate the effectiveness of our testing procedures for the three copula structures described in Section 2.1. In each simulation, we estimated 1200 test functions, and the tests were conducted at a nominal level of 5%. We used $N_b = 1000$ bootstrap samples for our analysis. To assess the performance of our approach for testing reflection symmetry, we compared it to the tests proposed by Genest, Nešlehová, and Quessy (2012) and Li and Genton (2013). The results, including the sizes and powers of the test for reflection symmetry, are summarized in Table 3. To evaluate the power of our tests, we introduce asymmetry to the copula models using Khoudraji's device (Khoudraji 1995). Specifically, we used an asymmetric version of a copula $C(u, v)$, defined for $(u, v) \in [0,1]^2$, given by:

$$K_\delta(u,v) = u^\delta C(u^{1-\delta}, v),$$

where $\delta \in (0, 1)$. Previous studies (Genest, Nešlehová, and Quessy 2012) have shown that Khoudraji's device introduces minimal asymmetry when Kendall's $\tau \leq 0.5$. The maximum level of asymmetry is typically observed around $\delta = 0.5$. The R codes for implementing the visualization and hypotheses testing procedures are available at Visualization-and-Assessment-of–Copula-Symmetry.

Our results indicate that under small and moderate τ, the sizes converge to the nominal value as the sample size increases, but with larger τ, the sizes are somewhat below the nominal level. The powers of the tests increase as the sample size increases. When comparing our results to Table 3 in Genest, Nešlehová, and Quessy (2012), their approach generally achieves better powers for a small sample size of $n = 250$. In contrast, our approach demonstrates significantly improved powers as the sample size increases. This is expected as our rank-based test relies on the asymptotic distribution of the test functions and the empirical copula process.

When comparing our simulation results to Table 1 in Li and Genton (2013), we find that our results align with most of the reported powers and sizes. However, our approach achieves considerably higher powers even at smaller sample sizes, particularly for intermediate values of τ where the maximum asymmetry is expected.

For the tests of radial and joint symmetry, we selected five commonly used copulas. The sizes and powers of both testing procedures are presented in Table 4. To evaluate the performance of our approach in testing radial symmetry, we compared it to the methods proposed by Li and Genton (2013) and Genest and Nešlehová (2014). Additionally, we compared our approach for testing joint symmetry to the methods reported by Li and Genton (2013).

Table 2. Definitions of estimated copulas constructed to test different types of symmetry. $\widehat{C}_n^*$ and $\widehat{D}_n^*$ denote the estimator for the survival copulas associated with the underlying copula.

Type	Access	Simulated distribution
	(U_1, U_2)	$\widehat{C}^S(u,v) = \frac{1}{2}\left(\widehat{C}_n(u,v) + \widehat{C}_n(v,u)\right)$
S	(X_1, X_2)	$\widehat{C}^S(u,v) = \frac{1}{2}\left(\widehat{D}_n(u,v) + \widehat{D}_n(v,u)\right)$
	(U_1, U_2)	$\widehat{C}^R(u,v) = \frac{1}{2}\left(\widehat{C}_n(u,v) + \widehat{C}_n^*(u,v)\right)$
R	(X_1, X_2)	$\widehat{C}^R(u,v) = \frac{1}{2}\left(\widehat{D}_n(u,v) + \widehat{D}_n^*(u,v)\right)$
	(U_1, U_2)	$\widehat{C}^J(u,v) = \frac{1}{4}\left(\widehat{C}_n(u,v) + u - \widehat{C}_n(u,1-v) + v - \widehat{C}_n(1-u,v) + \widehat{C}_n^*(u,v)\right)$
J	(X_1, X_2)	$\widehat{C}^J(u,v) = \frac{1}{4}\left(\widehat{D}_n(u,v) + u - \widehat{D}_n(u,1-v) + v - \widehat{D}_n(1-u,v) + \widehat{D}_n^*(u,v)\right)$

Table 3. Empirical sizes and powers of the test of symmetry in the setting as Genest, Nešlehová, and Quessy (2012) and Li and Genton (2013).

			Clayton				Gaussian				Gumbel			
			n				n				n			
δ	τ		100	250	500	1000	100	250	500	1000	100	250	500	1000
	1/4	SIZE	0.044	0.048	0.059	0.062	0.053	0.052	0.046	0.066	0.050	0.063	0.059	0.063
0	1/2		0.021	0.043	0.051	0.053	0.031	0.039	0.048	0.064	0.038	0.044	0.055	0.053
	3/4		0.007	0.002	0.007	0.007	0.001	0.002	0.003	0.004	0.007	0.008	0.009	0.019
0	0*		0.045	0.060	0.059	0.060								
	0.5	POWER	0.121	0.297	0.547	0.796	0.089	0.267	0.552	0.850	0.119	0.312	0.609	0.903
1/4	0.7		0.374	0.814	0.981	0.999	0.336	0.860	0.995	1.000	0.365	0.886	0.994	1.000
	0.9		0.675	0.977	0.996	1.000	0.740	0.987	0.998	1.000	0.728	0.990	1.000	1.000
	0.5		0.149	0.358	0.630	0.872	0.187	0.511	0.826	0.990	0.264	0.690	0.929	0.998
1/2	0.7		0.483	0.926	0.998	1.000	0.664	0.986	1.000	1.000	0.736	0.992	1.000	1.000
	0.9		0.908	1.000	1.000	1.000	0.951	1.000	1.000	1.000	0.940	1.000	1.000	1.000
	0.5		0.087	0.210	0.334	0.557	0.178	0.450	0.743	0.959	0.301	0.656	0.930	0.995
3/4	0.7		0.254	0.629	0.897	0.996	0.500	0.929	0.999	1.000	0.605	0.957	0.998	1.000
	0.9		0.662	0.984	1.000	1.000	0.761	0.999	1.000	1.000	0.768	0.985	1.000	1.000

NOTE: The sizes for the independent copula are $\delta = \tau = 0$.

Table 4. Sizes and powers of the test of radial and joint symmetry in the setting as Li and Genton (2013) and Genest and Nešlehová (2014).

			Radial				Joint				
	τ	$n =$	100	250	500	1000	$n =$	100	250	500	1000
Π	0	SIZE	0.060	0.073	0.065	0.060	SIZE	0.030	0.029	0.024	0.048
	1/4		0.054	0.060	0.059	0.054	POWER	0.685	0.963	0.997	1.000
Frank	1/2		0.050	0.064	0.058	0.059		0.968	1.000	1.000	1.000
	3/4		0.016	0.041	0.037	0.046		0.975	0.999	1.000	1.000
	1/4		0.062	0.060	0.047	0.059		0.725	0.984	1.000	1.000
Gaussian	1/2		0.041	0.041	0.054	0.054		0.976	0.999	1.000	1.000
	3/4		0.009	0.022	0.014	0.031		0.973	1.000	1.000	1.000
	1/4		0.228	0.491	0.830	0.985		0.706	0.957	0.999	1.000
Clayton	1/2	P	0.516	0.833	0.985	1.000		0.966	0.999	1.000	1.000
	3/4	O	0.291	0.512	0.912	0.996		0.976	1.000	1.000	1.000
	1/4	W	0.110	0.333	0.510	0.695		0.696	0.974	0.998	1.000
Gumbel	1/2	E	0.166	0.486	0.723	0.909		0.974	1.000	1.000	1.000
	3/4	R	0.042	0.322	0.580	0.815		0.978	0.999	1.000	1.000

NOTE: One bootstrap sample is used to estimate the distribution of the test statistic W under the null.

The sizes of both tests closely align with the nominal level of 5% for values of τ equal to 1/4 and 1/2. However, for larger values of τ, the sizes are achieved at larger sample sizes. Specifically, in the case of the radial symmetry test, our results are consistent with those presented in Table 2 of Li and Genton (2013), and once again our testing approach achieves the nominal levels at smaller sample sizes.

In comparison to the results presented in Table 1 of Genest and Nešlehová (2014) for the Frank and Gaussian copula models they considered, our approach achieves sizes that are closer to the nominal level. Although their performance improves for larger sample sizes, our approach outperforms them even in scenarios with larger sample sizes.

Regarding the joint symmetry copula structure, our achieved powers are significantly higher compared to the powers obtained for the radial symmetry property, and similar results are presented in Table 2 of Li and Genton (2013).

4. Data Applications

To illustrate the procedures herein, we apply our visualization and hypothesis testing methodology to three real-world datasets. The nutritional habits survey data are in Section 4.1, stock price data for the five top companies in the NASDAQ-100 stock index are in Section 4.2, and two major stock indices, the US S&P500 and German DAX data, are in Section 4.3.

4.1. Nutritional Habits Survey Data

The dataset used here comes from a survey conducted by the U.S. Department of Agriculture in 1985. The survey aimed to investigate the dietary habits of 737 women aged between 25 and 50 years. Specifically, the survey collected daily intake measurements of five variables: calcium (mg), iron (mg), protein (g), vitamin A (mg), and vitamin C (mg).

In previous analyses, Genest, Nešlehová, and Quessy (2012) employed a Cramér-von Mises statistic to evaluate the bivariate reflection symmetry of the pairwise copulas. Similarly, Li and Genton (2013) used a nonparametric method based on the asymptotic distribution of the empirical copula process to assess reflection symmetry for this dataset. In our study, we conducted a test using our rank-based method and obtained corresponding p-values, which are presented in the top-right corner of each subplot in Figure 2.

The pattern of p-values obtained in our study exhibits similarities to the findings reported by Genest, Nešlehová, and Quessy (2012) and Li and Genton (2013), leading to similar conclusions for some pairs. However, there are notable differences. Unlike Genest, Nešlehová, and Quessy (2012), our test does not reject the reflection symmetry copula at a 5% significance level for the pairs (iron, vitamin A), (iron, vitamin C), and (protein, vitamin A). Conversely, we do reject the reflection symmetry copula for the pair (protein, vitamin C).

When comparing our results to those of Li and Genton (2013), we find much closer agreement. The only difference in conclusions arises for the two pairs: (protein, vitamin C), where we reject the reflection symmetry copula, and (iron, vitamin A), where we do not reject it. Notably, the pair (iron, vitamin A) differs from the conclusions of both Genest, Nešlehová, and Quessy (2012) and Li and Genton (2013). Based on our findings, we can make specific recommendations for appropriate copula models depending on the symmetry properties observed in the data. For pairs that exhibit reflection symmetry, such as calcium/vitamin C and iron/vitamin C, the Gumbel copula could be a suitable model. For pairs that present both reflection symmetry and radial symmetry, like iron/protein, iron/vitamin A, and vitamin A/vitamin C, the Gaussian and Frank copulas can be potential selections. However, in cases where all three symmetry structures (reflection, radial, and joint) are rejected, potential asymmetric copula models such as the Marshall-Olkin and Tawn's copulas may need to be considered.

4.2. Stock Prices Data

Assessing pairwise dependence between stock returns is crucial for portfolio diversification, risk management, and constructing optimal portfolios (Markowitz 1952). It provides insights into the interrelationships among stocks, aiding in the identification of risk sources and portfolio adjustments. Additionally, understanding pairwise dependence facilitates opportunities for hedging and risk mitigation, reducing overall portfolio risk and enhancing risk-adjusted returns (Campbell, Lo, and MacKinlay 1997). Moreover, it is valuable for financial modeling and forecasting, enabling the capture of joint movements and assessment of portfolio performance using models such as multivariate time series analysis, copula modeling, and factor models (Alexander 2008).

We analyze a time series dataset comprising the daily closing stock prices of five top companies within the NASDAQ-100 index: Amazon, Facebook, Google, Apple, and Microsoft. The data spans from January 1, 2018, to December 31, 2019, resulting in 502 observations. For each of the five assets, we have $T = 501$ log-return values. This dataset has previously been used in Lu and Ghosh (2023) for portfolio risk management, employing nonparametric estimation of multivariate copulas. The primary goal of our study is to assess the pairwise dependence among the equities under consideration and examine the potential presence of symmetry structures in the data. We use the visualization and testing procedures outlined in Sections 2.3 and 2.4 to evaluate potential symmetries within each bivariate time series.

Figure 3 presents the visualization and hypothesis testing results for the dataset under consideration, focusing on the three copula structures studied throughout this article. The results indicate that joint symmetry is rejected for all pairs of stocks. However, reflection symmetry is only dismissed in the Facebook-Google case, while radial symmetry is only rejected in the Amazon-Microsoft case. Our analysis reveals that a significant number of bivariate time series pass our tests for reflection symmetry and radial symmetry, suggesting that copula models exhibiting such symmetry properties could be potential models for the data. For cases where reflection and radial symmetry cannot be rejected, copula models such as the Gaussian and Frank copulas could be suitable choices. Deng, Smith, and Maneesoonthorn (2024) presents a comprehensive analysis of asymmetric dependencies in intraday equity returns through the application of large skew-t copula models. This study underscores the intricate and variable nature of dependency structures under differing market conditions and time resolutions. While our approach has concentrated on a specific time period and assumes a stable dependence structure, it could be extended to investigate temporal variations in dependency dynamics. Such an extension would contribute to a deeper understanding of asymmetry in financial markets and enhance the evaluation of symmetry within these evolving contexts.

4.3. S&P500 and DAX Return Data

This dataset consists of $n = 396$ observations of two major stock indices, the US S&P500 and the German DAX, during the years 2009 and 2010. The dataset was originally introduced by Brechmann and Schepsmeier (2013), where each time series was filtered using an ARMA(1,1)-GARCH(1,1) model with Student's t innovations, and the standardized residuals were transformed nonparametrically to copula data using the corresponding empirical distribution functions.

Subsequently, Li and Genton (2013) used this dataset to identify the underlying copula structure and propose a suitable parametric copula model for the dependence between the two stock indices. In Figure 4, we present the results of our proposed methodology for investigating the potential symmetry structures present in the two stock indices. Our analysis indicates that reflection symmetry cannot be rejected, however, we found

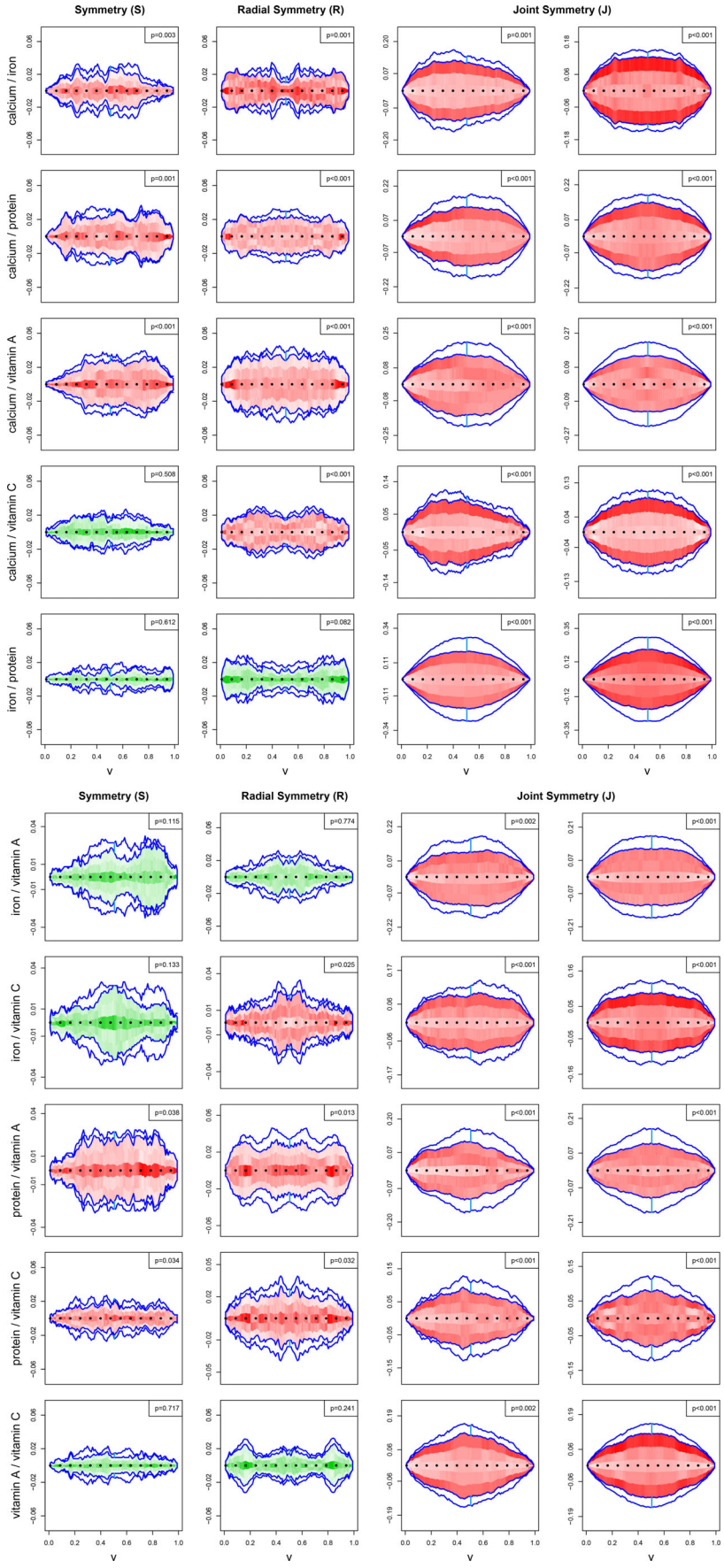

Figure 2. Visualizations and testing results for copula symmetries over the 5 random variables in the nutritional dataset.

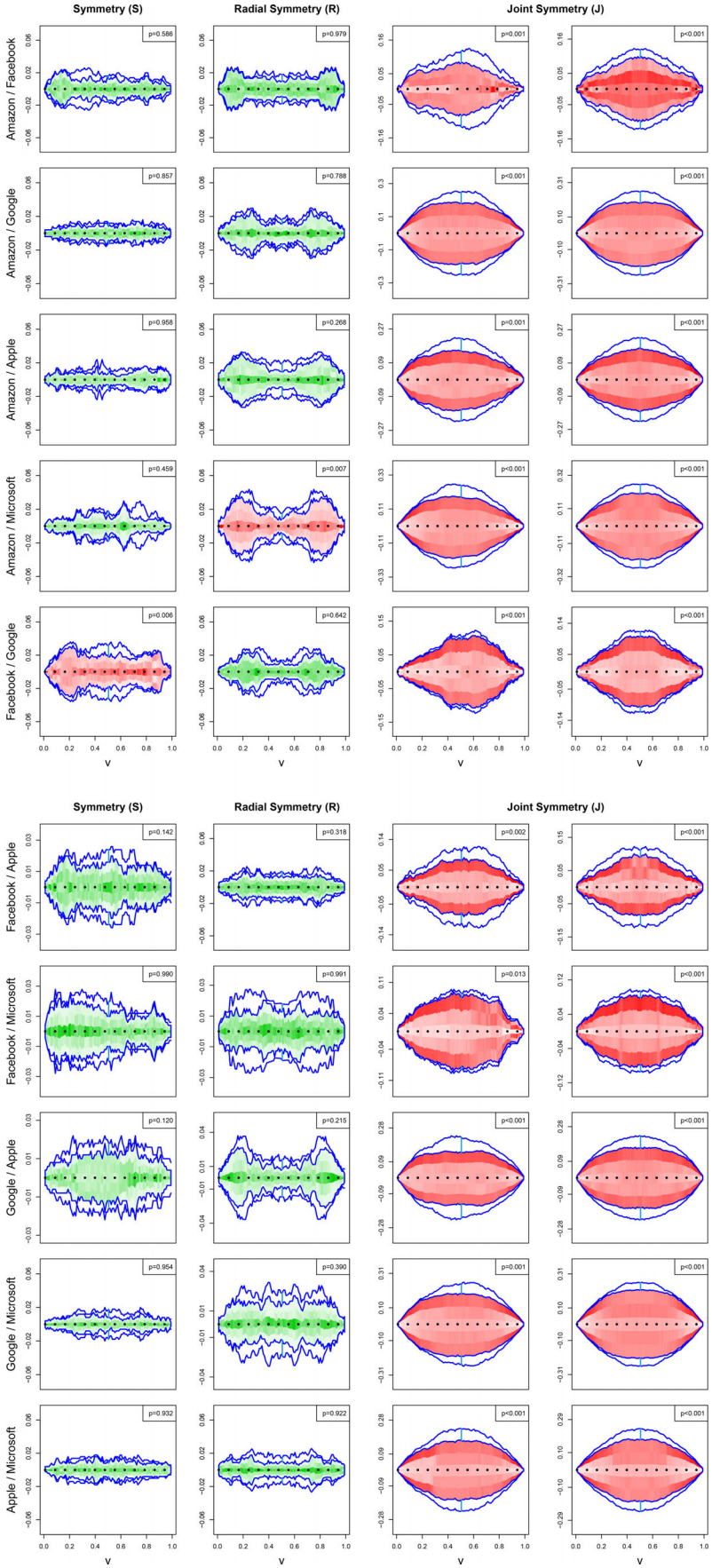

Figure 3. Visualizations and testing results for copula symmetries over the five top Nasdaq companies: Amazon, Facebook, Google, Apple, and Microsoft. The data spans from January 1, 2018, to December 31, 2019.

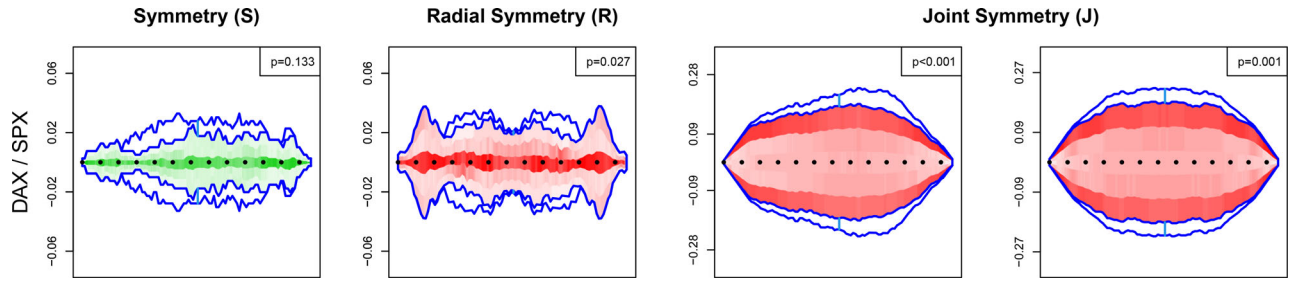

Figure 4. Visualization and testing results for copula symmetries over the two financial indexes S&P500 and DAX. The data spans the years 2009 and 2010.

sufficient evidence to reject the other two properties, radial symmetry and joint symmetry.

These findings differ from the results reported in Li and Genton (2013), where radial symmetry could not be rejected at the 5% significance level. In our analysis, however, radial symmetry is rejected with a p-value of 0.027.

Based on these insights, a potential parametric copula model that could be used to model the dependence between the S&P500 and DAX indices is the Gumbel copula, as illustrated in Figure 1.

5. Discussion

In this article, we have presented a comprehensive framework for visualizing and testing common assumptions regarding copula structures, such as reflection symmetry, radial symmetry, and joint symmetry. Our approach uses functional data analysis techniques to construct test functions based on bivariate copulas at specific discrete points. These test functions effectively summarize the copula structures and provide valuable insights into their adherence to specific structures.

To visually represent the copula structures, we employ functional boxplots, which depict the functional median and variability of the test functions. These visualizations allow us to assess the departure from zero and gain insights into the degree to which the copula structures conform to the desired assumptions.

Additionally, we have introduced a nonparametric testing procedure to evaluate the significance of deviations from symmetry. Through extensive simulation studies involving various copula models, we have demonstrated the reliability and power of our method, particularly for moderate to large sample sizes. The numerical experiments and data analyses conducted in this study have consistently shown robust testing results across different datasets, and the visualization technique has proven useful in extracting preliminary information directly from the data. It is worth mentioning that our functional data approach can be extended to test copula properties beyond symmetry as well.

It is important to note that our testing method relies on the asymptotic distribution of the estimators of the test functions, which are derived from empirical copula processes. As a result, the small sample properties of our test may not be as optimal as certain existing testing methods. The required sample size for our testing approach can vary depending on the specific copula structure under examination, but in general, larger sample sizes tend to enhance the size and power of the test. Thus, increasing the sample size is recommended to improve the overall performance of the test in terms of accuracy and sensitivity.

Finally, the proposed visualization and testing techniques were applied to three real-world datasets: a nutritional habits survey with five variables, five stock price data, and two major stock indices—the US S&P500 and the German DAX. These applications offered valuable insights into the underlying structures and patterns within the datasets, highlighting our approach's effectiveness in data exploration and analysis. To contextualize our findings within the literature, prior studies, such as Kole, Koedijk, and Verbeek (2007), have emphasized the importance of selecting appropriate copulas to model dependence structures in financial data. Our results suggest that the Gumbel copula is particularly suitable for capturing the dependence between the S&P500 and DAX indices, in line with research advocating for copulas that model tail dependence. Furthermore, our rejection of joint symmetry across all stock pairs underscores the limitations of symmetric copulas, such as the Gaussian and Frank, and reinforces the need for models that account for the asymmetric dependencies commonly observed in financial markets.

Supplementary Materials

- **Section S1: Interpretation of the Scenario Where Two Samples of Functional Curves Exhibit a Distinct Pattern** Analysis of cases where functional data shows two distinct samples of curves, such as forming separate bands with no overlap.
- **Section S2: Rank Plots for Different Copula Models and Symmetries** Visualization and analysis of rank plots for various copula models, including Clayton, Gaussian, and Gumbel, to evaluate symmetries like reflection and radial.
- **Sections S3 to S8: Proofs of Propositions 1 to 6 in the Main Manuscript** Detailed mathematical proofs for propositions, leveraging the Continuous Mapping Theorem and other tools to establish asymptotic properties of test functions.

Disclosure Statement

No potential conflict of interest was reported by the authors.

ORCID

Cristian F. Jiménez-Varón http://orcid.org/0000-0001-7471-3845
Marc G. Genton https://orcid.org/0000-0001-6467-2998
Ying Sun http://orcid.org/0000-0001-6703-4270

References

Aas, K., Czado, C., Frigessi, A., and Bakken, H. (2009), "Pair-Copula Constructions of Multiple Dependence," *Insurance: Mathematics and Economics*, 44, 182–198. [1140,1141]

Alexander, C. (2008), *Market Risk Analysis: Practical Financial Econometrics*, Chichester: Wiley. [1148]

Berg, D. (2009), "Copula Goodness-of-Fit Testing: An Overview and Power Comparison," *The European Journal of Finance*, 15, 675–701. [1141]

Brechmann, E. C., and Schepsmeier, U. (2013), "Modeling Dependence with c- and d-vine Copulas: The r Package cdvine," *Journal of Statistical Software*, 52, 1–27. [1148]

Bücher, A., Dette, H., and Volgushev, S. (2011), "New Estimators of the Pickands Dependence Function and a Test for Extreme-Value Dependence," *The Annals of Statistics*, 39, 1963–2006. [1141]

Bücher, A., Dette, H., and Volgushev, S. (2012), "A Test for Archimedeanity in Bivariate Copula Models," *Journal of Multivariate Analysis*, 110, 121–132. Special Issue on Copula Modeling and Dependence. [1141]

Campbell, J. Y., Lo, A., and MacKinlay, A. C. (1997), *The Econometrics of Financial Markets*, Princeton: Princeton University Press. [1148]

Cherubini, U., Luciano, E., and Vecchiato, W. (2004), *Copula Methods in Finance*, Chichester: Wiley. [1140]

Deheuvels, P. (1979), "Propriétés d'existence et propriétés topologiques des fonctions de dépendance avec applications à la convergence des types pour des lois multivariées," *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences. Séries A et B*, 288, A145–A148. [1142]

Deng, L., Smith, M. S., and Maneesoonthorn, W. (2024), "Large Skew-t Copula Models and Asymmetric Dependence in Intraday Equity Returns," *Journal of Business & Economic Statistics*, 1–17. [1148]

Efron, B., and Tibshirani, R. J. (1993), *An Introduction to the Bootstrap*, Volume 57 of Monographs on Statistics and Applied Probability. New York: Chapman and Hall. [1145]

Genest, C., and Favre, A.-C. (2007), "Everything You Always Wanted to Know About Copula Modeling but Were Afraid to Ask," *Journal of Hydrologic Engineering*, 12, 347–368. [1140,1141]

Genest, C., Nešlehová, J., and Quessy, J.-F. (2012), "Tests of Symmetry for Bivariate Copulas," *Annals of the Institute of Statistical Mathematics*, 64, 811–834. [1141,1142,1146,1147,1148]

Genest, C., and Nešlehová, D. (2014), "On Tests of Radial Symmetry for Bivariate Copulas," *Statistical Papers*, 55, 1107–1119. [1141,1146,1147]

Genest, C., Rémillard, B., and Beaudoin, D. (2009), "Goodness-of-Fit Tests for Copulas: A Review and a Power Study," *Insurance: Mathematics and Economics*, 44, 199–213. [1141]

Genest, C., and Segers, J. (2010), "On the Covariance of the Asymptotic Empirical Copula Process," *Journal of Multivariate Analysis*, 101, 1837–1845. [1142]

Huang, H., and Sun, Y. (2019), "Visualization and Assessment of Spatio-Temporal Covariance Properties," *Spatial Statistics*, 34, 100272, 18. [1141,1145,1146]

Huang, H., Sun, Y., and Genton, M. G. (2023), "Test and Visualization of Covariance Properties for Multivariate Spatio-Temporal Random Fields," *Journal of Computational and Graphical Statistics*, 32, 1545–1555. [1141,1145,1146]

Jaser, M., and Min, A. (2021), "On Tests for Symmetry and Radial Symmetry of Bivariate Copulas towards Testing for Ellipticity," *Computational Statistics*, 36, 1–26. [1141]

Jaworski, P. (2010), "Testing Archimedeanity," in *Combining Soft Computing and Statistical Methods in Data Analysis*, eds. C. Borgelt, G. González-Rodríguez, W. Trutschnig, M. A. Lubiano, M. Á. Gil, P. Grzegorzewski, and O. Hryniewicz, pp. 353–360, Berlin, Heidelberg: Springer. [1141]

Joe, H. (2014), *Dependence Modeling with Copulas*, New York: Chapman and Hall/CRC. [1140,1141]

Khoudraji, A. (1995), "Contributions à l'étude des copules et à la modélisation de valeurs extrêmes bivariées. Ph. D. thesis, National Library of Canada = Bibliothèquenationale du Canada Ottawa. [1146]

Kole, E., Koedijk, K., and Verbeek, M. (2007), "Selecting Copulas for Risk Management," *Journal of Banking & Finance*, 31, 2405–2423. [1151]

Li, B., and Genton, M. G. (2013), "Nonparametric Identification of Copula Structures," *Journal of the American Statistical Association*, 108, 666–675. [1141,1142,1146,1147,1148,1151]

Liu, R. Y., and Singh, K. (1993), "A Quality Index based on Data Depth and Multivariate Rank Tests," *Journal of the American Statistical Association*, 88, 252–260. [1145]

López-Pintado, S., and Romo, J. (2009), "On the Concept of Depth for Functional Data," *Journal of the American Statistical Association*, 104, 718–734. [1145]

Lu, L., and Ghosh, S. (2023), "Nonparametric Estimation of Multivariate Copula Using Empirical Bayes Methods," *Mathematics*, 11, 4383. [1148]

Markowitz, H. (1952), "Portfolio Selection," *The Journal of Finance*, 7, 77–91. [1148]

Mikosch, T. (2006), "Copulas: Tales and Facts—Rejoinder," *Extremes*, 9, 55–62. [1141]

Nelsen, R. B. (1993), "Some Concepts of Bivariate Symmetry," *Journal of Nonparametric Statistics*, 3, 95–101. [1141]

——— (2006), *An Introduction to Copulas*, New York: Springer. [1140,1141]

Patton, A. J. (2006), "Modelling Asymmetric Exchange Rate Dependence," *International Economic Review*, 47, 527–556. [1140]

——— (2012), "Copula-based Models for Financial Time Series," in *Handbook of Financial Time Series*, eds. T. G. Andersen, R. A. Davis, J.-P. Kreiß, and T. Mikosch, pp. 767–785, Berlin: Springer. [1140]

Quessy, J.-F. (2016), "A General Framework for Testing Homogeneity Hypotheses About Copulas," *Electronic Journal of Statistics*, 10, 1064–1097. [1141]

Rüschendorf, L. (1976), "Asymptotic Distributions of Multivariate Rank Order Statistics," *The Annals of Statistics*, 4, 912–923. [1146]

Segers, J. (2012), "Asymptotics of Empirical Copula Processes Under Non-Restrictive Smoothness Assumptions," *Bernoulli*, 18, 764–782. [1142]

Sklar, M. (1959), "Fonctions de répartition à *n* dimensions et leurs marges," *Publications de l'Institut de Statistique de l'Université de Paris*, 8, 229–231. [1140,1142]

Stute, W. (1984), "The Oscillation Behavior of Empirical Processes: The Multivariate Case," *The Annals of Probability*, 12, 361–379. [1142]

Sun, Y., and Genton, M. G. (2011), "Functional Boxplots," *Journal of Computational and Graphical Statistics*, 20, 316–334. [1141,1143]

Sun, Y., Genton, M. G., and Nychka, D. W. (2012), "Exact Fast Computation of Band Depth for Large Functional Datasets: How Quickly Can One Million Curves Be Ranked?" *Stat*, 1, 68–74. [1145]

Swanepoel, J. W. H., and Allison, J. S. (2013), "Some New Results on the Empirical Copula Estimator with Applications," *Statistics & Probability Letters*, 83, 1731–1739. [1142]

Tsukahara, H. (2005), "Semiparametric Estimation in Copula Models," *The Canadian Journal of Statistics*, 33, 357–375. [1142]