Fast large-sample goodness-of-fit tests for copulas

Ivan Kojadinovic[a], Jun Yan[b] and Mark Holmes[a]

[a]Department of Statistics
The University of Auckland, Private Bag 92019
Auckland 1142, New Zealand
{ivan,mholmes}@stat.auckland.ac.nz

[b]Department of Statistics
University of Connecticut, 215 Glenbrook Rd. U-4120
Storrs, CT 06269, USA
jun.yan@uconn.edu

Abstract

Goodness-of-fit tests are a fundamental element in the copula-based modeling of multivariate continuous distributions. Among the different procedures proposed in the literature, recent large scale simulations suggest that one of the most powerful tests is based on the empirical process comparing the empirical copula with a parametric estimate of the copula derived under the null hypothesis. As for most of the currently available goodness-of-fit procedures for copula models, the null distribution of the statistic for the latter test is obtained by means of a parametric bootstrap. The main inconvenience of this approach is its high computational cost, which, as the sample size increases, can be regarded as an obstacle to its application. In this work, fast large-sample tests for assessing goodness of fit are obtained by means of multiplier central limit theorems. The resulting procedures are shown to be asymptotically valid when based on popular method-of-moment estimators. Large scale Monte Carlo experiments, involving six frequently used parametric copula families and three different estimators of the copula parameter, confirm that the proposed procedures provide a valid, much faster alternative to the corresponding parametric bootstrap-based tests. An application of the derived tests to the modeling of a well-known insurance data set is presented. The use of the multiplier approach instead of the parametric bootstrap can reduce the computing time from about a day to minutes.

Key words and phrases: Empirical process; Multiplier central limit theorem; Pseudo-observation; Rank; Semiparametric model.

1 Introduction

The copula-based modeling of multivariate distributions is finding extensive applications in fields such as finance (Cherubini, Vecchiato, and Luciano, 2004; McNeil, Frey, and Embrechts, 2005), hydrology (Salvadori, Michele, Kottegoda, and Rosso, 2007), public health (Cui and Sun, 2004), and actuarial sciences (Frees and Valdez, 1998). The quite recent enthusiasm for the use of this modeling approach finds its origin in an elegant representation theorem due to Sklar (1959) that we present here in the bivariate case. Let (X, Y) be a random vector with continuous marginal cumulative distribution functions (c.d.f.s) F and G. A consequence of the work of Sklar (1959) is that the c.d.f. H of (X, Y) can be uniquely represented as

$$H(x, y) = C\{F(x), G(y)\}, \qquad x, y \in \mathbb{R},$$

where $C : [0, 1]^2 \to [0, 1]$, called a *copula*, is a bivariate c.d.f. with standard uniform margins. Given a random sample $(X_1, Y_1), \ldots, (X_n, Y_n)$ from c.d.f. H, this representation suggests breaking the multivariate model building into two independent parts: the fitting of the marginal c.d.f.s and the calibration of an appropriate parametric copula. The problem of estimating the parameters of the chosen copula has been extensively studied in the literature (see e.g. Genest and Favre, 2007; Genest, Ghoudi, and Rivest, 1995; Joe, 1997; Shih and Louis, 1995). Another very important issue that is currently drawing a lot of attention is whether the unknown copula C actually belongs to the chosen parametric copula family or not. More formally, one wants to test

$$H_0 : C \in \mathcal{C} = \{C_\theta : \theta \in \mathcal{O}\} \qquad \text{against} \qquad H_1 : C \notin \mathcal{C},$$

where $\mathcal{O}$ is an open subset of $\mathbb{R}^q$ for some integer $q \geq 1$.

A relatively large number of testing procedures have been proposed in the literature as can be concluded from the recent review of Genest, Rémillard, and Beaudoin (2009). Among the existing procedures, these authors advocate the use of so-called "blanket tests", i.e., those whose implementation requires neither an arbitrary categorization of the data, nor any strategic choice of smoothing parameter, weight function, kernel, window, etc. Among the tests in this last category, one approach that appears to perform particularly well according to recent large scale simulations (Berg, 2009; Genest et al., 2009) is based on the *empirical copula* (Deheuvels, 1981) of the data $(X_1, Y_1), \ldots, (X_n, Y_n)$, defined as

$$C_n(u, v) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}(U_{i,n} \leq u, V_{i,n} \leq v), \qquad u, v \in [0, 1], \tag{1}$$

where the random vectors $(U_{i,n}, V_{i,n})$ are *pseudo-observations* from C computed from the data by

$$U_{i,n} = \frac{1}{n+1} \sum_{j=1}^{n} \mathbf{1}(X_j \leq X_i), \qquad V_{i,n} = \frac{1}{n+1} \sum_{j=1}^{n} \mathbf{1}(Y_j \leq Y_i), \qquad i \in \{1, \ldots, n\}.$$

The empirical copula C_n is a consistent estimator of the unknown copula C, whether H_0 is true or not. Hence, as suggested in Fermanian (2005), Quessy (2005), and Genest and Rémillard (2008), a natural goodness-of-fit test consists of comparing C_n with an estimation C_{θ_n} of C obtained assuming that $C \in \mathcal{C}$ holds. Here, θ_n is an estimator of θ computed from the pseudo-observations $(U_{1,n}, V_{1,n}), \ldots, (U_{n,n}, V_{n,n})$. More precisely, these authors propose to base a test of goodness of fit on the empirical process

$$\sqrt{n}\{C_n(u,v) - C_{\theta_n}(u,v)\}, \qquad u, v \in [0,1]. \tag{2}$$

According to the large scale simulations carried out in Genest et al. (2009), the most powerful version of this procedure is based on the Cramér-von Mises statistic

$$S_n = \int_{[0,1]^2} n\{C_n(u,v) - C_{\theta_n}(u,v)\}^2 \mathrm{d}C_n(u,v) = \sum_{i=1}^{n} \{C_n(U_{i,n}, V_{i,n}) - C_{\theta_n}(U_{i,n}, V_{i,n})\}^2.$$

An approximate p-value for the test based on the above statistic is obtained by means of a parametric bootstrap whose validity was recently shown by Genest and Rémillard (2008). The large scale simulations carried out by Genest et al. (2009) and Berg (2009) suggest that, overall, this procedure yields the most powerful blanket goodness-of-fit test for copula models. The main inconvenience of this approach is its high computational cost as each parametric bootstrap iteration requires both random number generation from the fitted copula and estimation of the copula parameter. To fix ideas, Genest et al. (2009) mention the nearly exclusive use of 140 CPUs over a one-month period to estimate approximately 2000 powers and levels of goodness-of-fit tests based on parametric bootstrapping for $n \in \{50, 150\}$.

As the sample size increases, the application of parametric bootstrap-based goodness-of-fit tests becomes prohibitive. In order to circumvent this very high computational cost, we propose a fast large-sample testing procedure based on multiplier central limit theorems. Such techniques have already been used to simulate the null distributions of statistics in other types of tests based on empirical processes; see e.g. Lin, Fleming, and Wei (1994) and Fine, Yan, and Kosorok (2004) for applications in survival analysis or, more recently, Scaillet (2005), and Rémillard and Scaillet (2009) for applications in the copula modeling context. Starting from the seminal work of Rémillard and Scaillet (2009), we give two multiplier central limit theorems that suggest a fast asymptotically valid goodness-of-fit procedure. As we shall see in Section 5, for $n \approx 1500$, some computations that would typically require about a day when based on a parametric bootstrap can be performed in minutes when based on the multiplier approach.

The second section of the paper is devoted to the asymptotic behavior of the goodness-of-fit process $\sqrt{n}(C_n - C_{\theta_n})$ under the null hypothesis and to a brief presentation of the three most frequently used rank-based estimators of θ. In Section 3, we give two multiplier central limit theorems that are at the root of the proposed fast asymptotically valid goodness-of-fit procedures. Section 4 discusses extensive simulation results for $n = 75, 150$ and 300, and for three different rank-based estimators of the parameter θ. For $n = 150$, the resulting estimated powers and levels are compared with those obtained in Genest et al. (2009) using a parametric bootstrap. The proofs of the theorems and the details of the computations are relegated to the appendices. The penultimate section is devoted to an application of the proposed procedures to the modeling of dependence in well-known insurance data, while the last section contains methodological recommendations and concluding remarks.

Finally, note that all the tests studied in this work are implemented in the `copula` R package (Yan and Kojadinovic, 2010) available on the Comprehensive R Archive Network (R Development Core Team, 2009).

2 Asymptotic behavior of the goodness-of-fit process

In order to simplify the forthcoming expositions, we restrict our attention to bivariate one-parameter families of copulas, although many of the derivations to follow can be extended to the multivariate multiparameter case at the expense of higher complexity. Thus, let $\mathcal{O}$ be an open subset of $\mathbb{R}$, let $\mathcal{C} = \{C_\theta : \theta \in \mathcal{O}\}$ be an identifiable family of copulas, and assume that the true unknown copula belongs to the family $\mathcal{C}$, i.e., there exists a unique $\theta \in \mathcal{O}$ such that $C = C_\theta$.

The weak limit of the goodness-of-fit process (2) under the previous hypotheses partly follows from the following important result that characterizes the asymptotic behavior of the empirical copula (see e.g. Fermanian, Radulovic, and Wegkamp, 2004; Gänssler and Stute, 1987; Tsukahara, 2005).

Theorem 1. *Suppose that C_θ has continuous partial derivatives. Then, the empirical copula process $\sqrt{n}(C_n - C_\theta)$ converges weakly in $\ell^\infty([0,1]^2)$ to the tight centered Gaussian process*

$$\mathbb{C}_\theta(u,v) = \alpha_\theta(u,v) - C_\theta^{[1]}(u,v)\alpha_\theta(u,1) - C_\theta^{[2]}(u,v)\alpha_\theta(1,v), \qquad u,v \in [0,1],$$

where $C_\theta^{[1]} = \partial C_\theta/\partial u$, $C_\theta^{[2]} = \partial C_\theta/\partial v$ and α_θ is a C_θ-Brownian bridge, i.e., a tight centered Gaussian process on $[0,1]^2$ with covariance function $E[\alpha_\theta(u,v)\alpha_\theta(u',v')] = C_\theta(u \wedge u', v \wedge v') - C_\theta(u,v)C_\theta(u',v')$, $u,v,u',v' \in [0,1]$.

Let θ_n be an estimator of θ computed from $(U_{1,n}, V_{1,n}), \dots, (U_{n,n}, V_{n,n})$. The asymptotic behavior of the goodness-of-fit process (2) was studied in Quessy (2005) (see also Berg and Quessy, 2009) under the following three natural assumptions:

A1. For all $\theta \in \mathcal{O}$, the partial derivatives $C_\theta^{[1]}$ and $C_\theta^{[2]}$ are continuous.

A2. For all $\theta \in \mathcal{O}$, $\sqrt{n}(C_n - C_\theta)$ and $\sqrt{n}(\theta_n - \theta)$ jointly weakly converge to $(\mathbb{C}_\theta, \Theta)$ in $\ell^\infty([0,1]^2) \otimes \mathbb{R}$.

A3. For all $\theta \in \mathcal{O}$ and as $\varepsilon \downarrow 0$,

$$\sup_{|\theta^* - \theta| < \varepsilon} \sup_{u,v \in [0,1]} \left| \dot{C}_{\theta^*}(u,v) - \dot{C}_\theta(u,v) \right| \to 0,$$

where $\dot{C}_\theta = \partial C_\theta / \partial \theta$.

Proposition 1. *Under A1, A2 and A3, the goodness-of-fit process $\sqrt{n}(C_n - C_{\theta_n})$ converges weakly in $\ell^\infty([0,1]^2)$ to the tight centered Gaussian process*

$$\mathbb{C}_\theta(u,v) - \Theta \dot{C}_\theta(u,v), \qquad u,v \in [0,1]. \tag{3}$$

Assumption A1 is necessary to be able to apply Theorem 1. Assumption A3 is required to ensure that the process $\sqrt{n}(C_{\theta_n} - C_\theta)$ converges weakly to $\Theta\dot{C}_\theta(\mathbf{u})$ (see Berg and Quessy, 2009; Quessy, 2005, for more details). Assumption A2 then allows one to conclude that $\sqrt{n}(C_n - C_{\theta_n}) = \sqrt{n}(C_n - C_\theta) - \sqrt{n}(C_{\theta_n} - C_\theta)$ converges weakly to (3).

As can be seen, the weak limit (3) depends, through the random variable Θ, on the estimator θ_n chosen to estimate the parameter θ. Three rank-based estimation strategies are considered in this work. Two of the most popular involve the inversion of a consistent estimator of a moment of the copula. The two best-known moments are Spearman's rho and Kendall's tau. For a bivariate copula C_θ, these are, respectively,

$$\rho(\theta) = 12\int_{[0,1]^2} C_\theta(u,v)\mathrm{d}u\mathrm{d}v - 3 \quad \text{and} \quad \tau(\theta) = 4\int_{[0,1]^2} C_\theta(u,v)\mathrm{d}C_\theta(u,v) - 1.$$

Let $\mathcal{C}$ be a bivariate copula family such that the functions ρ and τ are one-to-one. Consistent estimators of θ are then given by $\theta_{n,\rho} = \rho^{-1}(\rho_n)$ and $\theta_{n,\tau} = \tau^{-1}(\tau_n)$, where ρ_n and τ_n are the sample versions of Spearman's rho and Kendall's tau, respectively.

A more general rank-based estimation method was studied by Genest et al. (1995) and Shih and Louis (1995) and consists of maximizing the log pseudo-likelihood

$$\log L_n(\theta) = \sum_{i=1}^n \log\{c_\theta(U_{i,n}, V_{i,n})\},$$

where c_θ denotes the density of the copula C_θ, assuming that it exists. As we continue, the resulting estimator is denoted by $\theta_{n,PL}$.

3 Multiplier goodness-of-fit tests

The proposed fast goodness-of-fit tests are based on multiplier central limit theorems that we state in the second subsection. These results partly rely on those obtained in Rémillard and Scaillet (2009) to test the equality between two copulas. However, an additional technical difficulty arises here from the fact that the estimation of θ is required. The resulting fast goodness-of-fit procedure is described in the last subsection.

3.1 Additional notation and setting

Let N be a large integer, and let $Z_i^{(k)}$, $i \in \{1,\dots,n\}$, $k \in \{1,\dots,N\}$, be i.i.d. random variables with mean 0 and variance 1 independent of the data $(X_1,Y_1),\dots,(X_n,Y_n)$. Moreover, for any $k \in \{1,\dots,N\}$, let

$$\begin{aligned}\alpha_n^{(k)}(u,v) &= \frac{1}{\sqrt{n}}\sum_{i=1}^n Z_i^{(k)}\left\{\mathbf{1}(U_{i,n} \le u, V_{i,n} \le v) - C_n(u,v)\right\} \qquad (4)\\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^n (Z_i^{(k)} - \bar{Z}^{(k)})\mathbf{1}(U_{i,n} \le u, V_{i,n} \le v), \qquad u,v \in [0,1].\end{aligned}$$

From Lemma A.1 in Rémillard and Scaillet (2009),

$$\left(\sqrt{n}(H_n - C_\theta), \alpha_n^{(1)}, \dots, \alpha_n^{(N)}\right) \rightsquigarrow \left(\alpha_\theta, \alpha_\theta^{(1)}, \dots, \alpha_\theta^{(N)}\right)$$

in $\ell^\infty([0,1]^2)^{\otimes(N+1)}$, where H_n is the empirical c.d.f. computed from the probability transformed data $(U_i, V_i) = (F(X_i), G(Y_i))$, and where $\alpha_\theta^{(1)}, \dots, \alpha_\theta^{(N)}$ are independent copies of α_θ. As consistent estimators of the partial derivatives $C_\theta^{[1]}$ and $C_\theta^{[2]}$, Rémillard and Scaillet (2009, Prop. A.2) suggest using

$$C_n^{[1]}(u,v) = \frac{1}{2n^{-1/2}} \left\{C_n(u + n^{-1/2}, v) - C_n(u - n^{-1/2}, v)\right\}$$

and

$$C_n^{[2]}(u,v) = \frac{1}{2n^{-1/2}} \left\{C_n(u, v + n^{-1/2}) - C_n(u, v - n^{-1/2})\right\},$$

respectively. From the proof of Theorem 2.1 in Rémillard and Scaillet (2009), it then follows that the empirical processes

$$\mathbb{C}_n^{(k)}(u,v) = \alpha_n^{(k)}(u,v) - C_n^{[1]}(u,v)\alpha_n^{(k)}(u,1) - C_n^{[2]}(u,v)\alpha_n^{(k)}(1,v) \tag{5}$$

can be regarded as approximate independent copies of the weak limit $\mathbb{C}_\theta$ defined in Theorem 1.

3.2 Multiplier central limit theorems

Before stating our first key result, we define a class of rank-based estimators of the copula parameter θ. As we continue, let $\Theta_n = \sqrt{n}(\theta_n - \theta)$.

Definition 1. *A rank-based estimator θ_n of θ is said to belong to class $\mathcal{R}_1$ if*

$$\Theta_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n J_\theta(U_{i,n}, V_{i,n}) + o_P(1), \tag{6}$$

where $J_\theta : [0,1]^2 \to \mathbb{R}$ is a score function that satisfies the following regularity conditions:

(a) for all $\theta \in \mathcal{O}$, J_θ is bounded on $[0,1]^2$ and centered, i.e., $\int_{[0,1]^2} J_\theta(u,v)\mathrm{d}C_\theta(u,v) = 0$,

(b) for all $\theta \in \mathcal{O}$, the partial derivatives $J_\theta^{[1]} = \partial J_\theta/\partial u$, $J_\theta^{[2]} = \partial J_\theta/\partial v$, $J_\theta^{[1,1]} = \partial^2 J_\theta/\partial u^2$, $J_\theta^{[1,2]} = \partial^2 J_\theta/\partial u \partial v$ and $J_\theta^{[2,2]} = \partial^2 J_\theta/\partial v^2$ exist and are bounded on $[0,1]^2$,

(c) for all $\theta \in \mathcal{O}$, the partial derivatives $\dot{J}_\theta = \partial J_\theta/\partial\theta$, $\dot{J}_\theta^{[1]} = \partial J_\theta^{[1]}/\partial\theta$ and $\dot{J}_\theta^{[2]} = \partial J_\theta^{[2]}/\partial\theta$ exist and are bounded on $[0,1]^2$,

(d) for all $\theta \in \mathcal{O}$, there exist $c_\theta > 0$ and $M_\theta > 0$ such that, if $|\theta' - \theta| < c_\theta$, then $|\dot{J}_{\theta'}| \le M_\theta$, $|\dot{J}_{\theta'}^{[1]}| \le M_\theta$, and $|\dot{J}_{\theta'}^{[2]}| \le M_\theta$.

It can be verified that, for the most popular bivariate one-parameter families of copulas, the method-of-moment estimator $\theta_{n,\rho}$ mentioned in the previous section belongs to class $\mathcal{R}_1$ (see e.g. Berg and Quessy, 2009, Section 3) as its score function is given by

$$J_{\theta,\rho}(u,v) = \frac{1}{\rho'(\theta)}\{12uv - 3 - \rho(\theta)\}, \qquad u,v \in [0,1].$$

The following result, proved in Appendix A, is instrumental for verifying the asymptotic validity of the fast goodness-of-fit procedure proposed in the next subsection when it is based on estimators from class $\mathcal{R}_1$.

Theorem 2. *Let θ_n be an estimator of θ belonging to class $\mathcal{R}_1$ and, for any $k \in \{1,\dots,N\}$, let*

$$\Theta_n^{(k)} = \frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i^{(k)} \mathcal{J}_{\theta_n,i,n}, \tag{7}$$

where

$$\mathcal{J}_{\theta_n,i,n} = J_{\theta_n}(U_{i,n}, V_{i,n}) + \frac{1}{n}\sum_{j=1}^n J_{\theta_n}^{[1]}(U_{j,n}, V_{j,n})\{\mathbf{1}(U_{i,n} \le U_{j,n}) - U_{j,n}\} \\ + \frac{1}{n}\sum_{j=1}^n J_{\theta_n}^{[2]}(U_{j,n}, V_{j,n})\{\mathbf{1}(V_{i,n} \le V_{j,n}) - V_{j,n}\}. \tag{8}$$

Then, under Assumptions A1 and A3,

$$\left(\sqrt{n}(C_n - C_{\theta_n}), \mathbb{C}_n^{(1)} - \Theta_n^{(1)}\dot{C}_{\theta_n}, \dots, \mathbb{C}_n^{(N)} - \Theta_n^{(N)}\dot{C}_{\theta_n}\right)$$

converges weakly to

$$\left(\mathbb{C}_\theta - \Theta\dot{C}_\theta, \mathbb{C}_\theta^{(1)} - \Theta^{(1)}\dot{C}_\theta, \dots, \mathbb{C}_\theta^{(N)} - \Theta^{(N)}\dot{C}_\theta\right) \tag{9}$$

in $\ell^\infty([0,1]^2)^{\otimes(N+1)}$, where Θ is the weak limit of $\Theta_n = \sqrt{n}(\theta_n - \theta)$ and $(\mathbb{C}_\theta^{(1)}, \Theta^{(1)}), \dots, (\mathbb{C}_\theta^{(N)}, \Theta^{(N)})$ are independent copies of $(\mathbb{C}_\theta, \Theta)$.

By replacing boundedness conditions in Definition 1 by more complex integrability conditions, Genest and Rémillard (2008) have defined a more general class of rank-based estimators of θ that also contains the maximum pseudo-likelihood estimator $\theta_{n,PL}$. For the latter, as shown in Genest et al. (1995), the asymptotic representation (6) holds with the score function

$$J_{\theta,PL}(u,v) = \left[\mathbb{E}\left\{\frac{\dot{c}_\theta(U,V)^2}{c_\theta(U,V)^2}\right\}\right]^{-1} \frac{\dot{c}_\theta(u,v)}{c_\theta(u,v)}, \qquad u,v \in (0,1),$$

where $\dot{c}_\theta = \partial c_\theta / \partial\theta$ and $(U,V) = (F(X), G(Y))$. An analogue of Theorem 2 remains to be proved for the above mentioned more general class of estimators.

We now define a second class of rank-based estimators of θ.

Definition 2. *A rank-based estimator θ_n of θ is said to belong to class $\mathcal{R}_2$ if*

$$\Theta_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} J_\theta(U_i, V_i) + o_P(1), \tag{10}$$

where $(U_i, V_i) = (F(X_i), G(Y_i))$ for all $i \in \{1, \ldots, n\}$, and $J_\theta : [0,1]^2 \to \mathbb{R}$ is a score function that satisfies the following regularity conditions:

(a) for all $\theta \in \mathcal{O}$, J_θ is bounded on $[0,1]^2$ and centered, i.e., $\int_{[0,1]^2} J_\theta(u,v)\mathrm{d}C_\theta(u,v) = 0$,

(b) for all $\theta \in \mathcal{O}$, the partial derivatives $J_\theta^{[1]} = \partial J_\theta/\partial u$ and $J_\theta^{[2]} = \partial J_\theta/\partial v$ exist and are bounded on $[0,1]^2$,

(c) for all $\theta \in \mathcal{O}$, the partial derivative $\dot{J}_\theta = \partial J_\theta/\partial\theta$ exists and is bounded on $[0,1]^2$,

(d) for all $\theta \in \mathcal{O}$, there exist $c_\theta > 0$ and $M_\theta > 0$ such that, if $|\theta' - \theta| < c_\theta$, then $|\dot{J}_{\theta'}| \leq M_\theta$.

An application of Hájek's projection method (see e.g. Berg and Quessy, 2009; Hájek, Šidák, and Sen, 1999) shows that the method-of-moment estimator $\theta_{n,\tau}$ can be expressed as in (10) with score function given by

$$J_{\theta,\tau}(u,v) = \frac{4}{\tau'(\theta)} \left\{ 2C_\theta(u,v) - u - v + \frac{1-\tau(\theta)}{2} \right\}, \qquad u, v \in [0,1].$$

It can be verified that, for the most popular bivariate one-parameter families of copulas, $\theta_{n,\tau}$ belongs to Class $\mathcal{R}_2$.

The following result, which is the analogue of Theorem 2 for Class $\mathcal{R}_2$, ensures the asymptotic validity of the fast goodness-of-fit procedure proposed in the next subsection when based on estimators from class $\mathcal{R}_2$. Its proof is omitted as it is very similar to and simpler than that of Theorem 2.

Theorem 3. *Let θ_n be an estimator of θ belonging to class $\mathcal{R}_2$ and, for any $k \in \{1, \ldots, N\}$, let*

$$\Theta_n^{(k)} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_i^{(k)} \mathcal{J}_{\theta_n,i,n} \qquad \text{where} \qquad \mathcal{J}_{\theta_n,i,n} = J_{\theta_n}(U_{i,n}, V_{i,n}). \tag{11}$$

Then, under Assumptions A1 and A3,

$$\left(\sqrt{n}(C_n - C_{\theta_n}), \mathbb{C}_n^{(1)} - \Theta_n^{(1)}\dot{C}_{\theta_n}, \ldots, \mathbb{C}_n^{(N)} - \Theta_n^{(N)}\dot{C}_{\theta_n}\right)$$

converges weakly to

$$\left(\mathbb{C}_\theta - \Theta\dot{C}_\theta, \mathbb{C}_\theta^{(1)} - \Theta^{(1)}\dot{C}_\theta, \ldots, \mathbb{C}_\theta^{(N)} - \Theta^{(N)}\dot{C}_\theta\right)$$

in $\ell^\infty([0,1]^2)^{\otimes(N+1)}$, where $(\mathbb{C}_\theta^{(1)}, \Theta^{(1)}), \ldots, (\mathbb{C}_\theta^{(N)}, \Theta^{(N)})$ are independent copies of $(\mathbb{C}_\theta, \Theta)$.

3.3 Goodness-of-fit procedure

Theorems 2 and 3 suggest adopting the following fast goodness-of-fit procedure:

1. Compute C_n from the pseudo-observations $(U_{1,n}, V_{1,n}), \dots, (U_{n,n}, V_{n,n})$ using (1), and estimate θ using an estimator from Class $\mathcal{R}_1$ or $\mathcal{R}_2$.
2. Compute the Cramér-von Mises statistic
$$S_n = \int_{[0,1]^2} n\{C_n(u,v) - C_{\theta_n}(u,v)\}^2 \mathrm{d}C_n(u,v) = \sum_{i=1}^{n} \{C_n(U_{i,n}, V_{i,n}) - C_{\theta_n}(U_{i,n}, V_{i,n})\}^2.$$
3. Then, for some large integer N, repeat the following steps for every $k \in \{1, \dots, N\}$:
    (a) Generate n i.i.d. random variates $Z_1^{(k)}, \dots, Z_n^{(k)}$ with expectation 0 and variance 1.
    (b) Form an approximate realization of the test statistic under H_0 by
$$\begin{aligned} S_n^{(k)} &= \int_{[0,1]^2} \left\{ \mathbb{C}_n^{(k)}(u,v) - \Theta_n^{(k)} \dot{C}_{\theta_n}(u,v) \right\}^2 \mathrm{d}C_n(u,v) \\ &= \frac{1}{n} \sum_{i=1}^{n} \left\{ \mathbb{C}_n^{(k)}(U_{i,n}, V_{i,n}) - \Theta_n^{(k)} \dot{C}_{\theta_n}(U_{i,n}, V_{i,n}) \right\}^2, \end{aligned}$$
    where $\mathbb{C}_n^{(k)}$ is defined by (5), and $\Theta_n^{(k)}$ by (7) or (11).
4. An approximate p-value for the test is finally given by $N^{-1} \sum_{k=1}^{N} \mathbf{1}(S_n^{(k)} \geq S_n)$.

The proposed procedure differs from the parametric bootstrap-based procedure considered in Genest and Rémillard (2008), Genest et al. (2009) and Berg (2009) only in Step 3. Instead of using multipliers, Step 3 of the parametric bootstrap-based procedure relies on random number generation from the fitted copula and estimation of the copula parameter to obtain approximate independent realizations of the test statistic under H_0. We detail Step 3 of the parametric bootstrap-based procedure hereafter:

3. For some large integer N, repeat the following steps for every $k \in \{1, \dots, N\}$:
    (a) Generate a random sample $(U_1^{(k)}, V_1^{(k)}), \dots, (U_n^{(k)}, V_n^{(k)})$ from copula C_{θ_n} and deduce the associated pseudo-observations $(U_{1,n}^{(k)}, V_{1,n}^{(k)}), \dots, (U_{n,n}^{(k)}, V_{n,n}^{(k)})$.
    (b) Let $C_n^{(k)}$ and $\theta_n^{(k)}$ stand for the versions of C_n and θ_n derived from the pseudo-observations $(U_{1,n}^{(k)}, V_{1,n}^{(k)}), \dots, (U_{n,n}^{(k)}, V_{n,n}^{(k)})$.
    (c) Form an approximate realization of the test statistic under H_0 as
$$S_n^{(k)} = \sum_{i=1}^{n} \{C_n^{(k)}(U_{i,n}^{(k)}, V_{i,n}^{(k)}) - C_{\theta_n^{(k)}}(U_{i,n}^{(k)}, V_{i,n}^{(k)})\}^2.$$

The multiplier procedure can be very rapidly implemented. Indeed, one first needs to compute the $n \times n$ matrix M_n whose elements are

$$M_n(i,j) = \mathbf{1}(U_{i,n} \leq U_{j,n}, V_{i,n} \leq V_{j,n}) - C_n(U_{j,n}, V_{j,n}) - C_n^{[1]}(U_{j,n}, V_{j,n})\{\mathbf{1}(U_{i,n} \leq U_{j,n}) - U_{j,n}\} \\ - C_n^{[2]}(U_{j,n}, V_{j,n})\{\mathbf{1}(V_{i,n} \leq V_{j,n}) - V_{j,n}\} - \mathcal{J}_{\theta_n,i,n}\dot{C}_{\theta_n}(U_{j,n}, V_{j,n}).$$

Then, to get one approximate realization of the test statistic under the null hypothesis, it suffices to generate n i.i.d. random variates $Z_1, \ldots, Z_n$ with expectation 0 and variance 1, and to perform simple arithmetic operations involving the Z_i's and the columns of matrix M_n.

Although an analogue of Theorem 2 and Theorem 3 remains to be proved for the maximum pseudo-likelihood estimator, we study the finite sample performance of the corresponding multiplier goodness-of-fit procedure in the next section. Its implementation (in a more general multivariate multiparameter context) is the subject of a companion paper (Kojadinovic and Yan, 2010a).

4 Finite-sample performance

The finite-sample performance of the proposed goodness-of-fit tests was assessed in a large-scale simulation study. The experimental design of the study was very similar to that considered in Genest et al. (2009). Six bivariate one-parameter families of copulas were considered, namely, the Clayton, Gumbel, Frank, normal, t (with $\nu = 4$ degrees of freedom), and Plackett (see Table 4 for more details). They are abbreviated by C, G, F, N, t_4 and P, respectively, in the forthcoming tables. Each one served both as true and hypothesized copula. The sample size $n = 150$ was considered in order to allow a comparison with the simulation results presented in Genest et al. (2009) obtained using the parametric bootstrap-based procedure described in the previous section in which estimation of the parameter was carried by inverting Kendall's tau. Note that we have not attempted to reproduce these results as they were obtained after an extensive use of high-performance computing grids. To make this empirical study more insightful, simulations were also carried out for $n = 75$ and 300. Three levels of dependence were considered, corresponding respectively to a Kendall's tau of 0.25, 0.5 and 0.75.

Three multiplier goodness-of-fit tests were compared, only differing according to the method used for estimating the unknown parameter θ of the hypothesized copula family. The three tests, based on Kendall's tau, Spearman's rho, and maximum pseudo-likelihood, are abbreviated by M-τ, M-ρ and M-PL. Similarly, the parametric bootstrap procedure based on Kendall's tau empirically studied in Genest et al. (2009) is denoted by PB-τ. In all executions of the multiplier-based tests, we used standard normal variates for the Z_i's in the procedure given in Subsection 3.3. As standard normal variates led to satisfactory results, we did not investigate the use of other types of multipliers. Also, the number of iterations N was fixed to 1000, which is equivalent to using 1000 bootstrap samples for PB-τ. For each testing scenario, 10 000 repetitions were performed to estimate the level or power of each of the three tests under consideration.

For $n = 75$ (results not reported), the empirical levels of the three tests appeared

Table 1: Percentage of rejection of the null hypothesis for sample size $n = 150$ obtained from 10 000 repetitions of the multiplier procedure with $N = 1000$. The results for PB-τ are taken from Genest et al. (2009).

		$\tau = 0.25$						$\tau = 0.5$						$\tau = 0.75$					
True	Meth	C	G	F	N	t	P	C	G	F	N	t	P	C	G	F	N	t	P
C	PB-τ	**4.6**	72.1	40.0	31.6	27.7	37.6	**5.3**	99.6	89.1	80.0	77.3	83.9	**5.4**	99.9	96.6	91.8	90.6	89.8
	M-τ	**6.2**	76.0	46.0	38.2	36.8	43.2	**5.0**	99.5	92.3	85.0	84.0	87.2	**3.6**	100.0	97.3	93.2	92.6	91.3
	M-PL	**7.0**	83.0	44.9	30.9	38.3	41.2	**7.1**	99.9	91.5	82.9	83.9	88.6	**7.8**	100.0	97.0	98.1	95.4	97.6
	M-ρ	**6.1**	75.5	46.4	38.8	33.5	42.7	**5.6**	99.5	92.8	87.6	81.0	85.6	**3.0**	99.5	96.7	91.7	85.1	76.1
F	PB-τ	86.1	**5.0**	33.4	23.8	19.1	30.4	99.9	**4.6**	63.0	38.3	33.9	48.4	100.0	**4.5**	81.9	38.5	33.9	45.8
	M-τ	85.4	**5.4**	34.2	25.5	22.9	30.8	99.9	**4.5**	68.5	41.7	37.6	51.1	100.0	**3.1**	84.4	38.4	34.3	45.3
	M-PL	90.8	**4.1**	33.8	18.6	23.1	28.5	100.0	**3.6**	69.8	37.0	41.7	56.3	100.0	**3.2**	88.6	46.4	47.9	82.7
	M-ρ	86.4	**6.0**	36.5	28.1	22.1	32.5	100.0	**5.1**	73.4	48.8	36.7	51.8	100.0	**4.1**	89.2	50.1	32.7	36.8
G	PB-τ	56.3	15.4	**5.3**	7.9	9.1	5.0	95.7	39.8	**4.8**	20.2	26.9	6.8	99.1	51.7	**4.7**	42.2	48.2	14.9
	M-τ	57.4	15.7	**5.1**	6.6	10.3	4.7	94.7	37.7	**4.8**	15.8	23.8	4.8	98.9	46.5	**3.4**	38.3	46.5	10.0
	M-PL	79.5	40.4	**5.2**	14.4	28.1	6.8	99.8	79.8	**5.3**	44.3	72.3	28.5	100.0	95.3	**4.5**	88.5	95.7	89.6
	M-ρ	57.2	15.2	**6.0**	7.8	7.7	5.3	94.2	27.5	**5.8**	13.7	12.8	3.6	95.4	20.7	**4.8**	25.7	19.2	2.3
N	PB-τ	50.2	10.1	4.8	**5.1**	4.9	6.8	93.7	18.3	19.9	**4.9**	5.2	9.8	99.8	12.3	40.9	**4.9**	4.1	7.7
	M-τ	51.7	11.7	8.7	**4.9**	5.9	7.4	93.9	19.1	24.3	**4.5**	4.6	10.9	99.7	10.6	45.8	**2.7**	2.6	6.6
	M-PL	65.9	21.1	9.3	**4.6**	14.6	8.6	98.4	32.7	26.0	**3.9**	16.9	23.8	100.0	24.1	53.9	**3.8**	13.8	70.6
	M-ρ	53.1	11.7	10.7	**6.4**	5.7	8.9	94.2	16.3	27.7	**5.1**	3.3	10.8	99.3	6.8	50.3	**3.7**	1.6	4.4
t_4	PB-τ	56.6	14.1	18.5	10.5	**4.8**	14.1	94.8	21.8	35.1	8.2	**5.0**	15.1	99.8	16.1	59.4	6.6	**4.9**	11.0
	M-τ	55.1	14.2	17.2	9.2	**5.0**	13.1	94.9	23.5	42.3	8.0	**4.6**	17.3	99.8	14.6	63.7	4.4	**3.0**	10.3
	M-PL	58.1	14.5	17.2	8.7	**4.5**	13.5	97.0	24.1	41.2	6.3	**4.0**	14.7	100.0	20.2	66.7	7.4	**4.2**	27.3
	M-ρ	59.8	16.9	22.8	15.0	**6.3**	16.7	95.7	27.2	50.4	17.3	**5.3**	18.5	99.6	16.2	75.1	15.0	**4.2**	8.7
P	PB-τ	56.0	14.3	5.7	7.9	7.7	**5.2**	95.8	29.8	8.4	13.2	13.8	**5.0**	99.5	25.8	20.6	16.5	15.7	**4.7**
	M-τ	55.8	16.3	6.0	7.3	9.0	**5.6**	95.3	29.5	10.3	10.5	11.8	**4.3**	99.5	24.6	23.6	13.1	13.3	**3.4**
	M-PL	78.3	34.9	5.4	14.3	21.9	**5.4**	99.7	61.1	9.1	39.6	31.0	**5.3**	100.0	65.1	18.6	78.4	28.8	**9.6**
	M-ρ	58.6	15.3	6.8	8.4	6.6	**5.7**	96.1	33.0	19.1	22.9	11.4	**5.4**	99.6	43.7	60.5	54.1	27.8	**5.3**

Table 2: Percentage of rejection of the null hypothesis for sample size $n = 300$ obtained from 10 000 repetitions of the multiplier procedure with $N = 1000$.

		$\tau = 0.25$						$\tau = 0.5$						$\tau = 0.75$					
True	Meth	C	G	F	N	t	P	C	G	F	N	t	P	C	G	F	N	t	P
C	M-τ	**4.9**	98.0	79.2	70.0	71.0	76.9	**5.0**	100.0	99.9	99.6	99.7	99.8	**2.9**	100.0	100.0	100.0	100.0	100.0
	M-PL	**5.5**	98.8	79.7	64.3	74.7	75.7	**5.5**	100.0	99.9	99.6	99.7	99.8	**6.2**	100.0	100.0	100.0	100.0	100.0
	M-ρ	**5.5**	98.2	80.2	72.3	70.3	77.5	**4.9**	100.0	99.9	99.7	99.5	99.8	**3.8**	100.0	100.0	100.0	100.0	100.0
F	M-τ	98.6	**4.5**	54.8	40.0	37.7	49.8	100.0	**4.3**	92.8	66.7	62.9	80.3	100.0	**2.8**	99.5	64.7	60.4	79.8
	M-PL	99.4	**4.5**	54.6	32.7	41.2	48.5	100.0	**4.1**	93.0	61.9	66.0	81.0	100.0	**3.3**	99.5	70.7	71.4	97.0
	M-ρ	98.8	**5.1**	57.2	43.8	37.4	51.8	100.0	**4.8**	93.9	72.7	61.4	80.8	100.0	**3.7**	99.7	76.0	57.6	79.0
G	M-τ	84.2	36.3	**5.0**	9.9	22.5	4.8	99.9	79.0	**4.4**	38.4	60.7	7.9	100.0	90.4	**2.8**	78.7	88.0	25.7
	M-PL	96.2	72.7	**4.7**	24.9	59.2	6.8	100.0	98.8	**4.1**	80.7	97.4	46.6	100.0	100.0	**3.7**	99.6	100.0	99.5
	M-ρ	84.9	34.7	**4.9**	9.4	16.3	4.4	99.9	72.3	**4.8**	34.4	42.7	5.4	100.0	75.1	**4.3**	69.8	66.9	12.0
N	M-τ	77.7	24.3	12.0	**4.8**	8.7	9.6	99.9	44.5	45.9	**4.0**	5.7	19.0	100.0	33.5	85.4	**2.7**	2.9	15.9
	M-PL	87.4	40.2	12.6	**4.0**	29.3	10.0	100.0	62.7	46.5	**4.2**	32.5	37.3	100.0	59.1	87.9	**3.0**	21.8	94.7
	M-ρ	77.9	24.7	12.3	**5.1**	7.1	10.3	99.8	41.5	46.6	**4.8**	4.5	21.2	100.0	26.2	86.8	**3.3**	1.5	18.3
t_4	M-τ	81.0	28.2	31.5	12.5	**4.8**	21.8	99.9	49.4	73.0	9.4	**4.3**	31.5	100.0	39.5	95.6	5.5	**3.1**	22.9
	M-PL	79.9	25.8	33.1	11.3	**4.4**	31.3	100.0	49.1	73.2	8.3	**4.3**	27.7	100.0	47.1	95.6	8.4	**3.8**	40.4
	M-ρ	82.1	31.6	35.0	18.8	**5.2**	24.1	99.9	54.7	78.0	22.3	**4.8**	32.6	100.0	44.7	97.8	20.1	**3.5**	22.0
P	M-τ	82.4	33.2	5.3	8.5	15.8	**4.6**	100.0	64.0	14.5	20.7	26.1	**4.4**	100.0	59.9	45.0	25.7	26.9	**3.2**
	M-PL	95.7	65.0	5.6	25.7	44.5	**5.0**	100.0	92.2	13.5	72.7	59.1	**5.0**	100.0	94.7	37.5	97.9	46.2	**6.6**
	M-ρ	83.9	33.2	6.3	10.5	12.0	**5.2**	100.0	68.9	25.2	40.5	22.6	**4.4**	100.0	85.0	86.9	86.7	54.4	**5.1**

overall to be too liberal. As n was increased to 150, the agreement between the empirical levels (in bold in Table 1) and the 5% nominal level seemed globally satisfactory, except for M-PL when data arise from the Clayton copula or from the Packett copula with $\tau = 0.75$. The empirical levels in Table 2 confirm that, as the sample size reaches 300, the multiplier approach provided, overall, an accurate approximation to the null distribution of the test statistics. In terms of power, as expected, the rejection percentages increased quite substantially when n went from 75 to 150, and then to 300.

From Theorems 2 and 3, and the work of Genest and Rémillard (2008), we know that the empirical processes involved in the multiplier and the parametric bootstrap-based tests are asymptotically equivalent if estimation is based on the inversion of Spearman's rho or Kendall's tau. The results presented in Table 1 for M-τ and PB-τ also suggest that their finite-sample performances are fairly close.

From the presented results, it also appears that the method chosen for estimating the parameter of the hypothesized copula family can greatly influence the power of the approach. For instance, from Tables 1 and 2, M-PL seems to perform particularly well when the hypothesized copula is the Plackett copula and the dependence is high. The procedure M-PL appears to give the best results overall. It is followed by M-τ.

5 Illustrative example

The insurance data of Frees and Valdez (1998) are frequently used for illustration purposes in copula modeling (see e.g. Ben Ghorbal, Genest, and Nešlehová, 2009; Genest, Quessy, and Rémillard, 2006; Klugman and Parsa, 1999). The two variables of interest are an indemnity payment and the corresponding allocated loss adjustment expense, and were observed for 1500 claims of an insurance company. Following Genest et al. (2006), we restrict ourselves to the 1466 uncensored claims.

The data under consideration contain a non negligible number of ties. As demonstrated in Kojadinovic and Yan (2010b), ignoring the ties, by using for instance mid-ranks in the computation of the pseudo-observations, may affect the results qualitatively. For these reasons, when computing the pseudo-observations, we assigned ranks at random in case of ties. This was done using the R function `rank` with its argument `ties.method` set to `"random"`. The random seed that we used is 1224. The approximate p-values of the multiplier goodness-of-fit tests and the corresponding parametric bootstrap-based procedures computed with $N = 10\,000$ are given in Table 3.

Among the six bivariate copulas, only the Gumbel family is not rejected at the 5% significance level, which might be attributed to the extreme value dependence in the data (Ben Ghorbal et al., 2009). This is in accordance with the results obtained e.g. in Chen and Fan (2005) using pseudo-likelihood ratio tests, or in Genest et al. (2006) using a goodness-of-fit procedure based on Kendall's transform. In addition to the p-values, the timings, performed on one 2.33 GHz processor, are provided. These are based on our mixed R and C implementation of the tests available in the `copula` R package. As one can notice, the use of the multiplier tests results in a very large computational gain while the conclusions remain the same.

Table 3: Approximate p-values computed with $N = 10\,000$ and timings of the goodness-of-fit tests M-τ, M-PL, M-ρ, PB-τ, PB-PL and PB-ρ for the insurance data. The timings were performed on one 2.33 GHz processor.

Estimation	Copula	M		PB	
method		p-value	time (min)	p-value	time (min)
τ	C	0.000	4.7	0.000	41.3
	G	0.246	4.7	0.236	41.7
	F	0.000	4.7	0.000	68.1
	N	0.000	4.8	0.000	166.4
	t_4	0.000	4.7	0.000	169.4
	P	0.000	4.7	0.000	41.6
	Total		28.3		528.5
PL	C	0.000	4.8	0.000	447.1
	G	0.179	4.7	0.169	107.8
	F	0.000	4.7	0.000	147.7
	N	0.000	4.8	0.000	824.5
	t_4	0.000	4.8	0.000	1053.7
	P	0.000	4.7	0.000	98.4
	Total		28.5		2679.3
ρ	C	0.000	4.7	0.000	6.4
	G	0.271	4.7	0.262	6.7
	F	0.000	4.7	0.000	46.9
	N	0.000	4.7	0.000	151.1
	t_4	0.000	4.7	0.000	155.1
	P	0.000	4.7	0.000	22.1
	Total		28.2		388.2

To ensure that the randomization does not affect the results qualitatively, the tests based on the pseudo-observations computed with `ties.method = "random"` were performed a large number of times in Kojadinovic and Yan (2010b). The numerical summaries presented in the latter study indicate that the randomization does not affect the conclusions qualitatively for these data.

6 Discussion

The previous illustrative example highlights the most important advantage of the studied procedures over their parametric bootstrap-based counterparts: the former can be much faster. From the simulation results presented in Section 4, we also expect the multiplier tests to be at least as powerful as the parametric bootstrap-based ones for larger n. In other words, the proposed multiplier procedures appear as appropriate large-sample substitutes to the parametric bootstrap-based goodness-of-fit tests used in Genest et al. (2009) and Berg (2009), and studied theoretically in Genest and Rémillard (2008).

From a practical perspective, the results presented in Section 4 indicate that the multiplier approach can be safely used even in the case of samples of size as small as 150 as long as estimation is based on Kendall's tau or Spearman's rho. As n reaches 300, all three tests, included the one based on the maximization of the pseudo-likelihood, appear to hold their nominal level. The latter version of the test also seems to be the most powerful in general.

From the timings presented in the previous section, we see that the computational gain resulting from the use of the multiplier approach appears to be much more pronounced when estimation is based on the maximization of the pseudo-likelihood. This is of particular practical importance as, in a general multivariate multiparameter context, the latter estimation method becomes the only possible choice. The finite-sample performance of the maximum pseudo-likelihood version of the multiplier test was recently studied in a companion paper for multiparameter copulas of dimension 3 and 4 (Kojadinovic and Yan, 2010a). The results of the large-scale Monte Carlo experiments reported therein confirm the satisfactory behavior of the multiplier approach in this higher dimensional context. For all these reasons, it would be important to obtain an analogue of Theorem 2 and Theorem 3 for the maximum pseudo-likelihood estimator.

A Proof of Theorem 2

Let $(U_i, V_i) = (F(X_i), G(Y_i))$ for all $i \in \{1, \dots, n\}$ and let F_n and G_n be the rescaled empirical c.d.f.s computed from the unobservable random samples $U_1, \dots, U_n$ and $V_1, \dots, V_n$ respectively, i.e.,

$$F_n(u) = \frac{1}{n+1} \sum_{i=1}^n \mathbf{1}(U_i \leq u), \qquad \text{and} \qquad G_n(v) = \frac{1}{n+1} \sum_{i=1}^n \mathbf{1}(V_i \leq v), \qquad u, v \in [0,1].$$

It is then easy to verify that, for any $i \in \{1, \dots, n\}$, $U_{i,n} = F_n(U_i)$ and $V_{i,n} = G_n(V_i)$.

The proof of Theorem 2 is based on five lemmas. In their proofs, we have sometimes delayed the use of the conditions stated in Definition 1 so that the reader can identify the difficulties associated with considering the more general class of estimators defined in Genest and Rémillard (2008, Definition 4).

Lemma 1. *Let θ_n be an estimator of θ belonging to class $\mathcal{R}_1$. Then,*

$$\Theta_n = \sqrt{n}(\theta_n - \theta) = \frac{1}{\sqrt{n}} \sum_{i=1}^n J_\theta(U_i, V_i) + \frac{1}{\sqrt{n}} \sum_{i=1}^n J_\theta^{[1]}(U_i, V_i)\{F_n(U_i) - U_i\} + \frac{1}{\sqrt{n}} \sum_{i=1}^n J_\theta^{[2]}(U_i, V_i)\{G_n(V_i) - V_i\} + o_P(1).$$

Proof. By the second-order mean value theorem, we have

$$J_\theta\{F_n(U_i), G_n(V_i)\} = J_\theta(U_i, V_i) + J_\theta^{[1]}(U_i, V_i)\{F_n(U_i) - U_i\} + J_\theta^{[2]}(U_i, V_i)\{G_n(V_i) - V_i\} + R_{i,n},$$

where

$$R_{i,n} = \frac{1}{2} J_\theta^{[1,1]}(U_{i,n}^{[1]}, V_{i,n}^{[1]})\{F_n(U_i) - U_i\}^2 + \frac{1}{2} J_\theta^{[2,2]}(U_{i,n}^{[2]}, V_{i,n}^{[2]})\{G_n(V_i) - V_i\}^2 \\ + J_\theta^{[1,2]}(U_{i,n}^{[3]}, V_{i,n}^{[3]})\{F_n(U_i) - U_i\}\{G_n(V_i) - V_i\},$$

and where, for any $k \in \{1,2,3\}$, $U_{i,n}^{[k]}$ is between U_i and $F_n(U_i)$, and $V_{i,n}^{[k]}$ is between V_i and $G_n(V_i)$. Then,

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n J_\theta\{F_n(U_i), G_n(V_i)\} = \frac{1}{\sqrt{n}} \sum_{i=1}^n J_\theta(U_i, V_i) + \frac{1}{\sqrt{n}} \sum_{i=1}^n J_\theta^{[1]}(U_i, V_i)\{F_n(U_i) - U_i\} \\ + \frac{1}{\sqrt{n}} \sum_{i=1}^n J_\theta^{[2]}(U_i, V_i)\{G_n(V_i) - V_i\} + \frac{1}{\sqrt{n}} \sum_{i=1}^n R_{i,n}.$$

Furthermore,

$$\left| \frac{1}{\sqrt{n}} \sum_{i=1}^n R_{i,n} \right| \le \sup_{u \in [0,1]} |F_n(u) - u| \times \sup_{u \in [0,1]} |\sqrt{n}\{F_n(u) - u\}| \times \frac{1}{2n} \sum_{i=1}^n |J_\theta^{[1,1]}(U_{i,n}^{[1]}, V_{i,n}^{[1]})| \\ + \sup_{v \in [0,1]} |G_n(v) - v| \times \sup_{v \in [0,1]} |\sqrt{n}\{G_n(v) - v\}| \times \frac{1}{2n} \sum_{i=1}^n |J_\theta^{[2,2]}(U_{i,n}^{[2]}, V_{i,n}^{[2]})| \\ + \sup_{u \in [0,1]} |F_n(u) - u| \times \sup_{v \in [0,1]} |\sqrt{n}\{G_n(v) - v\}| \times \frac{1}{n} \sum_{i=1}^n |J_\theta^{[1,2]}(U_{i,n}^{[3]}, V_{i,n}^{[3]})|.$$

The result follows from the fact that the second-order derivatives of J_θ are bounded on $[0,1]^2$, that $\sup_{u\in[0,1]} |\sqrt{n}\{F_n(u) - u\}|$ converges in distribution, and that $\sup_{u\in[0,1]} |F_n(u) - u| \xrightarrow{\mathbb{P}} 0$. $\square$

Lemma 2. *Let J_θ be the score function of an estimator of θ belonging to class $\mathcal{R}_1$, and, for any $i \in \{1, \dots, n\}$, let*

$$\tilde{\mathcal{J}}_{\theta,i,n} = J_\theta(U_i, V_i) + \frac{1}{n} \sum_{j=1}^n J_\theta^{[1]}(U_j, V_j)\{\mathbf{1}(U_i \le U_j) - U_j\} + \frac{1}{n} \sum_{j=1}^n J_\theta^{[2]}(U_j, V_j)\{\mathbf{1}(V_i \le V_j) - V_j\}. \tag{12}$$

Then, for any $k \in \{1, \dots, N\}$, we have

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i^{(k)} (\mathcal{J}_{\theta,i,n} - \tilde{\mathcal{J}}_{\theta,i,n}) \xrightarrow{\mathbb{P}} 0,$$

where $\mathcal{J}_{\theta,i,n}$ is defined as in (8) with θ_n replaced by θ.

Proof. For any $i \in \{1, \dots, n\}$, let $A_{\theta,i,n} = J_\theta\{F_n(U_i), G_n(V_i)\} - J_\theta(U_i, V_i)$, let

$$B_{\theta,i,n} = \frac{1}{n} \sum_{j=1}^n J_\theta^{[1]}\{F_n(U_j), G_n(V_j)\}[\mathbf{1}\{F_n(U_i) \le F_n(U_j)\} - F_n(U_j)] \\ - \frac{1}{n} \sum_{j=1}^n J_\theta^{[1]}(U_j, V_j)\{\mathbf{1}(U_i \le U_j) - U_j\},$$

and let

$$B'_{\theta,i,n} = \frac{1}{n}\sum_{j=1}^{n} J_\theta^{[2]}\{F_n(U_j), G_n(V_j)\}[\mathbf{1}\{G_n(V_i) \le G_n(V_j)\} - G_n(V_j)] \\ - \frac{1}{n}\sum_{j=1}^{n} J_\theta^{[2]}(U_j, V_j)\{\mathbf{1}(V_i \le V_j) - V_j\}.$$

Let $k \in \{1, \ldots, N\}$. Then, from (8) and (12), we obtain

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Z_i^{(k)}(\mathcal{J}_{\theta,i,n} - \tilde{\mathcal{J}}_{\theta,i,n}) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} Z_i^{(k)} A_{\theta,i,n} + \frac{1}{\sqrt{n}}\sum_{i=1}^{n} Z_i^{(k)} B_{\theta,i,n} + \frac{1}{\sqrt{n}}\sum_{i=1}^{n} Z_i^{(k)} B'_{\theta,i,n}.$$

By the mean value theorem, we have

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Z_i^{(k)} A_{\theta,i,n} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} Z_i^{(k)} J_\theta^{[1]}(U_{i,n}^{[1]}, V_{i,n}^{[1]})\{F_n(U_i) - U_i\} \\ + \frac{1}{\sqrt{n}}\sum_{i=1}^{n} Z_i^{(k)} J_\theta^{[2]}(U_{i,n}^{[2]}, V_{i,n}^{[2]})\{G_n(V_i) - V_i\},$$

where, for any $k = 1, 2$, $U_{i,n}^{[k]}$ (resp. $V_{i,n}^{[k]}$) is between U_i and $F_n(U_i)$ (resp. V_i and $G_n(V_i)$). Let $D_{\theta,n} = n^{-1/2}\sum_{i=1}^{n} Z_i^{(k)} J_\theta^{[1]}(U_{i,n}^{[1]}, V_{i,n}^{[1]})\{F_n(U_i) - U_i\}$. It is easy to check that $D_{\theta,n}$ has mean 0 and variance $\mathbb{E}(D_{\theta,n}^2) = n^{-1}\sum_{i=1}^{n} \mathbb{E}[\{J_\theta^{[1]}(U_{i,n}^{[1]}, V_{i,n}^{[1]})\}^2\{F_n(U_i) - U_i\}^2]$. Then,

$$\mathbb{E}(D_{\theta,n}^2) \le \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\left[\{J_\theta^{[1]}(U_{i,n}^{[1]}, V_{i,n}^{[1]})\}^2 \times \{\sup_{u\in[0,1]} |F_n(u) - u|\}^2\right] \\ \le \sup_{u,v\in[0,1]} \{J_\theta^{[1]}(u, v)\}^2 \times \mathbb{E}\left[\{\sup_{u\in[0,1]} |F_n(u) - u|\}^2\right].$$

The right-hand side tends to 0 as a consequence of the dominated convergence theorem. Hence, $D_{\theta,n} \xrightarrow{\mathbb{P}} 0$. Similarly, one has that $n^{-1/2}\sum_{i=1}^{n} Z_i^{(k)} J_\theta^{[2]}(U_{i,n}^{[2]}, V_{i,n}^{[2]})\{G_n(V_i) - V_i\} \xrightarrow{\mathbb{P}} 0$. It follows that $n^{-1/2}\sum_{i=1}^{n} Z_i^{(k)} A_{\theta,i,n} \xrightarrow{\mathbb{P}} 0$.

It remains to show that $n^{-1/2}\sum_{i=1}^{n} Z_i^{(k)} B_{\theta,i,n}$ and $n^{-1/2}\sum_{i=1}^{n} Z_i^{(k)} B'_{\theta,i,n}$ converge to zero in probability. First, using the fact that $\mathbf{1}\{F_n(U_i) \le F_n(U_j)\} = \mathbf{1}(U_i \le U_j)$, notice that $B_{\theta,i,n}$ can be expressed as

$$B_{\theta,i,n} = \frac{1}{n}\sum_{j=1}^{n} J_\theta^{[1]}\{F_n(U_j), G_n(V_j)\}\{\mathbf{1}(U_i \le U_j) - F_n(U_j)\} \\ - \frac{1}{n}\sum_{j=1}^{n} J_\theta^{[1]}(U_j, V_j)\{\mathbf{1}(U_i \le U_j) - F_n(U_j) + F_n(U_j) - U_j\},$$

which implies that $n^{-1/2}\sum_{i=1}^n Z_i^{(k)} B_{\theta,i,n}$ can be rewritten as the difference of

$$H_{\theta,n} = \frac{1}{\sqrt{n}}\sum_{i=1}^n Z_i^{(k)} \left(\frac{1}{n}\sum_{j=1}^n \left[J_\theta^{[1]}\{F_n(U_j), G_n(V_j)\} - J_\theta^{[1]}(U_j, V_j) \right] \{\mathbf{1}(U_i \le U_j) - F_n(U_j)\} \right)$$

and

$$H'_{\theta,n} = \left[\frac{1}{n}\sum_{j=1}^n J_\theta^{[1]}(U_j, V_j)\{F_n(U_j) - U_j\} \right] \left[\frac{1}{\sqrt{n}}\sum_{i=1}^n Z_i^{(k)} \right].$$

Now, $H_{\theta,n}$ can be rewritten as

$$H_{\theta,n} = \frac{1}{n}\sum_{j=1}^n \left[J_\theta^{[1]}\{F_n(U_j), G_n(V_j)\} - J_\theta^{[1]}(U_j, V_j) \right] \left[\frac{1}{\sqrt{n}}\sum_{i=1}^n Z_i^{(k)} \{\mathbf{1}(U_i \le U_j) - U_j\} \right.$$
$$\left. - \{F_n(U_j) - U_j\} \times \frac{1}{\sqrt{n}}\sum_{i=1}^n Z_i^{(k)} \right].$$

Therefore,

$$|H_{\theta,n}| \le \left[\sup_{u\in[0,1]} \left| \frac{1}{\sqrt{n}}\sum_{i=1}^n Z_i^{(k)} \{\mathbf{1}(U_i \le u) - u\} \right| + \sup_{u\in[0,1]} |F_n(u) - u| \times \left| \frac{1}{\sqrt{n}}\sum_{i=1}^n Z_i^{(k)} \right| \right]$$
$$\times \frac{1}{n}\sum_{j=1}^n \left| J_\theta^{[1]}\{F_n(U_j), G_n(V_j)\} - J_\theta^{[1]}(U_j, V_j) \right|.$$

From the multiplier central limit theorem (see e.g. Kosorok, 2008, Theorem 10.1) and the continuous mapping theorem, we have that

$$\sup_{u\in[0,1]} \left| \frac{1}{\sqrt{n}}\sum_{i=1}^n Z_i^{(k)} \{\mathbf{1}(U_i \le u) - u\} \right| \rightsquigarrow \sup_{u\in[0,1]} \alpha_\theta(u, 1).$$

Furthermore, from the mean value theorem, one can write

$$\frac{1}{n}\sum_{j=1}^n \left| J_\theta^{[1]}\{F_n(U_j), G_n(V_j)\} - J_\theta^{[1]}(U_j, V_j) \right| \le \frac{1}{n}\sum_{j=1}^n \left| J_\theta^{[1,1]}(U_{j,n}^{[1]}, V_{j,n}^{[1]}) \right| |F_n(U_j) - U_j|$$
$$+ \frac{1}{n}\sum_{j=1}^n \left| J_\theta^{[1,2]}(U_{j,n}^{[2]}, V_{j,n}^{[2]}) \right| |G_n(V_j) - V_j|,$$

which implies that

$$\frac{1}{n}\sum_{j=1}^n \left| J_\theta^{[1]}\{F_n(U_j), G_n(V_j)\} - J_\theta^{[1]}(U_j, V_j) \right| \le \sup_{u\in[0,1]} |F_n(u) - u| \times \frac{1}{n}\sum_{j=1}^n \left| J_\theta^{[1,1]}(U_{j,n}^{[1]}, V_{j,n}^{[1]}) \right|$$
$$+ \sup_{u\in[0,1]} |G_n(u) - u| \times \frac{1}{n}\sum_{j=1}^n \left| J_\theta^{[1,2]}(U_{j,n}^{[2]}, V_{j,n}^{[2]}) \right|.$$

Since the second-order derivatives of J_θ are bounded on $[0,1]^2$, the right-hand side converges to 0 in probability and hence $H_{\theta,n} \xrightarrow{\mathbb{P}} 0$. For $H'_{\theta,n}$, we can write

$$|H'_{\theta,n}| \le \left| \frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i^{(k)} \right| \times \sup_{u\in[0,1]} |F_n(u) - u| \times \frac{1}{n} \sum_{j=1}^n |J_\theta^{[1]}(U_j, V_j)|.$$

It follows that $H'_{\theta,n} \xrightarrow{\mathbb{P}} 0$. By symmetry, $n^{-1/2} \sum_{i=1}^n Z_i^{(k)} B'_{\theta,i,n} \xrightarrow{\mathbb{P}} 0$. □

Lemma 3. *Let θ_n be an estimator of θ belonging to class $\mathcal{R}_1$. Then, for any $k \in \{1, \ldots, N\}$, we have*

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i^{(k)} (\mathcal{J}_{\theta_n,i,n} - \mathcal{J}_{\theta,i,n}) \xrightarrow{\mathbb{P}} 0,$$

where $\mathcal{J}_{\theta_n,i,n}$ is defined in (8).

Proof. Let $k \in \{1, \ldots, N\}$. Starting from (8), for any $i \in \{1, \ldots, n\}$, one has

$$\begin{aligned}\mathcal{J}_{\theta_n,i,n} - \mathcal{J}_{\theta,i,n} &= J_{\theta_n}(U_{i,n}, V_{i,n}) - J_\theta(U_{i,n}, V_{i,n}) \\ &\quad + \frac{1}{n} \sum_{j=1}^n \{J_{\theta_n}^{[1]}(U_{j,n}, V_{j,n}) - J_\theta^{[1]}(U_{j,n}, V_{j,n})\}\{\mathbf{1}(U_i \le U_j) - U_{j,n}\} \\ &\quad + \frac{1}{n} \sum_{j=1}^n \{J_{\theta_n}^{[2]}(U_{j,n}, V_{j,n}) - J_\theta^{[2]}(U_{j,n}, V_{j,n})\}\{\mathbf{1}(V_i \le V_j) - V_{j,n}\}.\end{aligned}$$

From the mean-value theorem, for any $i \in \{1, \ldots, n\}$, there exist $\theta_{i,n}, \theta'_{i,n}$ and $\theta''_{i,n}$ between θ and θ_n such that

$$\begin{aligned}\mathcal{J}_{\theta_n,i,n} - \mathcal{J}_{\theta,i,n} &= \dot{J}_{\theta_{i,n}}(U_{i,n}, V_{i,n})(\theta_n - \theta) + \frac{1}{n} \sum_{j=1}^n \dot{J}_{\theta'_{j,n}}^{[1]}(U_{j,n}, V_{j,n})(\theta_n - \theta)\{\mathbf{1}(U_i \le U_j) - U_{j,n}\} \\ &\quad + \frac{1}{n} \sum_{j=1}^n \dot{J}_{\theta''_{j,n}}^{[2]}(U_{j,n}, V_{j,n})(\theta_n - \theta)\{\mathbf{1}(V_i \le V_j) - V_{j,n}\}.\end{aligned}$$

It follows that

$$\begin{aligned}\frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i^{(k)} (\mathcal{J}_{\theta_n,i,n} - \mathcal{J}_{\theta,i,n}) &= (\theta_n - \theta) \times \frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i^{(k)} \dot{J}_{\theta_{i,n}}(U_{i,n}, V_{i,n}) \\ &\quad + (\theta_n - \theta) \times \frac{1}{n} \sum_{j=1}^n \dot{J}_{\theta'_{j,n}}^{[1]}(U_{j,n}, V_{j,n}) \frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i^{(k)} \{\mathbf{1}(U_i \le U_j) - U_{j,n}\} \\ &\quad + (\theta_n - \theta) \times \frac{1}{n} \sum_{j=1}^n \dot{J}_{\theta''_{j,n}}^{[2]}(U_{j,n}, V_{j,n}) \frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i^{(k)} \{\mathbf{1}(V_i \le V_j) - V_{j,n}\}. \quad (13)\end{aligned}$$

Let $K_n = n^{-1/2} \sum_{i=1}^n Z_i^{(k)} \dot{J}_{\theta_{i,n}}(U_{i,n}, V_{i,n})$, and let us show that $L_n = (\theta_n - \theta) K_n \xrightarrow{\mathbb{P}} 0$. First, define

$$K'_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i^{(k)} \dot{J}_{\theta_{i,n}}(U_{i,n}, V_{i,n}) \mathbf{1}\{|\dot{J}_{\theta_{i,n}}(U_{i,n}, V_{i,n})| \le M_\theta\}$$

and

$$K_n'' = \frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i^{(k)} \dot{J}_{\theta_{i,n}}(U_{i,n}, V_{i,n}) \mathbf{1}\{|\dot{J}_{\theta_{i,n}}(U_{i,n}, V_{i,n})| > M_\theta\},$$

where M_θ is defined in Condition (d) of Definition 1. Clearly, $K_n = K_n' + K_n''$. Let us now show that $K_n'' \xrightarrow{\mathbb{P}} 0$. Let $\delta > 0$ be given. Then, for n sufficiently large, $\mathbb{P}(|\theta_n - \theta| < c_\theta) > 1 - \delta$. Hence,

$$1 - \delta < \mathbb{P}(|\theta_n - \theta| < c_\theta) \le \mathbb{P}\left(\bigcap_{i=1}^n \{|\theta_{i,n} - \theta| < c_\theta\} \right) \le \mathbb{P}\left(\bigcap_{i=1}^n \{|\dot{J}_{\theta_{i,n}}(U_{i,n}, V_{i,n})| < M_\theta\} \right),$$

which implies that $\mathbb{P}\left(\bigcup_{i=1}^n \{\mathbf{1}(|\dot{J}_{\theta_{i,n}}(U_{i,n}, V_{i,n})| > M_\theta) > 0\} \right) < \delta$, which in turn implies that $K_n'' \xrightarrow{\mathbb{P}} 0$.

To show that $L_n = (\theta_n - \theta) K_n' + (\theta_n - \theta) K_n'' \xrightarrow{\mathbb{P}} 0$, it therefore remains to show that $(\theta_n - \theta) K_n' \xrightarrow{\mathbb{P}} 0$. Let $\varepsilon, \delta > 0$ be given and choose M_δ such $M_\theta^2 / M_\delta^2 < \delta/2$. Then,

$$\mathbb{P}(|\theta_n - \theta||K_n'| > \varepsilon) \le \mathbb{P}(|\theta_n - \theta| > \varepsilon / M_\delta) + \mathbb{P}(|K_n'| > M_\delta).$$

Now, let n be sufficiently large so that $\mathbb{P}(|\theta_n - \theta| > \varepsilon/M_\delta) < \delta/2$. Furthermore, by Markov's inequality and the fact that

$$\mathbb{E}(K_n'^2) = \frac{1}{n} \sum_{i=1}^n \mathbb{E}\left[\{\dot{J}_{\theta_{i,n}}(U_{i,n}, V_{i,n})\}^2 \mathbf{1}\{|\dot{J}_{\theta_{i,n}}(U_{i,n}, V_{i,n})| \le M_\theta\} \right] \le M_\theta^2,$$

we have that

$$\mathbb{P}(|K_n'| > M_\delta) \le \frac{\mathbb{E}(K_n'^2)}{M_\delta^2} \le \frac{M_\theta^2}{M_\delta^2} < \frac{\delta}{2}.$$

We therefore obtain that $L_n \xrightarrow{\mathbb{P}} 0$.

To obtain the desired result, it remains to show that the second and third terms on the right side of (13) converge to 0 in probability. We shall only deal with the second term as the proof for the third term is similar. First, notice that

$$\begin{aligned}
&\left| \frac{1}{n} \sum_{j=1}^n \dot{J}_{\theta_{j,n}'}^{[1]}(U_{j,n}, V_{j,n}) \frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i^{(k)} \{\mathbf{1}(U_i \le U_j) - U_j + U_j - U_{j,n}\} \right| \\
&\qquad \le \frac{1}{n} \sum_{j=1}^n |\dot{J}_{\theta_{j,n}'}^{[1]}(U_{j,n}, V_{j,n})| \left[\sup_{u \in [0,1]} \left| \frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i^{(k)} \{\mathbf{1}(U_i \le u) - u\} \right| \right. \\
&\qquad\qquad \left. + \sup_{u \in [0,1]} |F_n(u) - u| \times \left| \frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i^{(k)} \right| \right].
\end{aligned}$$

It is easy to verify that the term between square brackets on the right side of the previous inequality converges in distribution. To obtain the desired result, it therefore suffices to show that $|\theta_n - \theta| \times n^{-1} \sum_{j=1}^n |\dot{J}^{[1]}_{\theta'_{j,n}}(U_{j,n}, V_{j,n})| \xrightarrow{\mathbb{P}} 0$. Let $\varepsilon > 0$ be given. Then,

$$\mathbb{P}\left(|\theta_n - \theta| \frac{1}{n} \sum_{j=1}^n |\dot{J}^{[1]}_{\theta'_{j,n}}(U_{j,n}, V_{j,n})| < \varepsilon\right)$$

$$\geq \mathbb{P}\left(|\theta_n - \theta| < \frac{\varepsilon}{M_\theta}, \frac{1}{n} \sum_{j=1}^n |\dot{J}^{[1]}_{\theta'_{j,n}}(U_{j,n}, V_{j,n})| \leq M_\theta\right)$$

$$\geq \mathbb{P}\left(|\theta_n - \theta| < \frac{\varepsilon}{M_\theta}, \bigcap_{j=1}^n \{|\theta'_{j,n} - \theta| < c_\theta\}\right) \to 1.$$

□

Lemma 4. *Let J_θ be the score function of an estimator of θ belonging to class $\mathcal{R}_1$. Then,*

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n \left[\frac{1}{n} \sum_{j=1}^n J^{[1]}_\theta(U_j, V_j)\{\mathbf{1}(U_i \leq U_j) - U_j\} - \int_{[0,1]^2} J^{[1]}_\theta(u, v)\{\mathbf{1}(U_i \leq u) - u\} \mathrm{d}C_\theta(u, v)\right] \xrightarrow{\mathbb{P}} 0$$

and

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n \left[\frac{1}{n} \sum_{j=1}^n J^{[2]}_\theta(U_j, V_j)\{\mathbf{1}(V_i \leq V_j) - V_j\} - \int_{[0,1]^2} J^{[2]}_\theta(u, v)\{\mathbf{1}(V_i \leq v) - v\} \mathrm{d}C_\theta(u, v)\right] \xrightarrow{\mathbb{P}} 0.$$

Proof. It can be verified that the first term and the second term have mean 0 and, that the variance of the first term and the variance of the second term tend to zero, which implies that the terms tend to zero in probability. □

Lemma 5. *Let J_θ be the score function of an estimator of θ belonging to class $\mathcal{R}_1$. For any $k \in \{1, \ldots, N\}$, we have*

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n Z^{(k)}_i \left[\frac{1}{n} \sum_{j=1}^n J^{[1]}_\theta(U_j, V_j)\{\mathbf{1}(U_i \leq U_j) - U_j\} - \int_{[0,1]^2} J^{[1]}_\theta(u, v)\{\mathbf{1}(U_i \leq u) - u\} \mathrm{d}C_\theta(u, v)\right] \xrightarrow{\mathbb{P}} 0$$

and

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n Z^{(k)}_i \left[\frac{1}{n} \sum_{j=1}^n J^{[2]}_\theta(U_j, V_j)\{\mathbf{1}(V_i \leq V_j) - V_j\} - \int_{[0,1]^2} J^{[2]}_\theta(u, v)\{\mathbf{1}(V_i \leq v) - v\} \mathrm{d}C_\theta(u, v)\right] \xrightarrow{\mathbb{P}} 0.$$

Proof. The proof is similar to that of Lemma 4. □

Proof of Theorem 2. Let $\tilde{\alpha}_n = \sqrt{n}(H_n - C_\theta)$, where H_n is the empirical c.d.f. computed from $(U_1, V_1), \ldots, (U_n, V_n)$, and, for any $k \in \{1, \ldots, N\}$, let

$$\tilde{\alpha}^{(k)}_n(u, v) = \frac{1}{\sqrt{n}} \sum_{i=1}^n Z^{(k)}_i \{\mathbf{1}(U_i \leq u, V_i \leq v) - C_\theta(u, v)\}, \qquad u, v \in [0, 1].$$

From the results given on page 383 of Rémillard and Scaillet (2009) and the multiplier central limit theorem (see e.g. Kosorok, 2008, Theorem 10.1), we have that

$$\left(\tilde{\alpha}_n, \tilde{\alpha}_n^{(1)}, \alpha_n^{(1)}, \ldots, \tilde{\alpha}_n^{(N)}, \alpha_n^{(N)}\right)$$

converges weakly to

$$\left(\alpha_\theta, \alpha_\theta^{(1)}, \alpha_\theta^{(1)}, \ldots, \alpha_\theta^{(N)}, \alpha_\theta^{(N)}\right)$$

in $(\ell^\infty([0,1]^2))^{\otimes 2N+1}$, where $\alpha_n^{(k)}$ is defined in (4), and $\alpha_\theta^{(1)}, \ldots, \alpha_\theta^{(N)}$ are independent copies of α_θ. Furthermore, from Lemma 1 and Lemma 4, we have that $\Theta_n = \tilde{\Theta}_n + o_P(1)$, where $\tilde{\Theta}_n = n^{-1/2}\sum_{i=1}^n \tilde{\mathcal{J}}_{\theta,i}$, and where

$$\tilde{\mathcal{J}}_{\theta,i} = J_\theta(U_i, V_i) + \int_{[0,1]^2} J_\theta^{[1]}(u,v)\{\mathbf{1}(U_i \le u) - u\}\mathrm{d}C_\theta(u,v) \\ + \int_{[0,1]^2} J_\theta^{[2]}(u,v)\{\mathbf{1}(V_i \le v) - v\}\mathrm{d}C_\theta(u,v).$$

Similarly, from Lemmas 2 and 3, we obtain that $\Theta_n^{(k)}$ and $n^{-1/2}\sum_{i=1}^n Z_i^{(k)}\tilde{\mathcal{J}}_{\theta,i,n}$ are asymptotically equivalent, where $\tilde{\mathcal{J}}_{\theta,i,n}$ is defined by (12). Then, from Lemma 5, it follows that $\Theta_n^{(k)}$ and $\tilde{\Theta}_n^{(k)} = n^{-1/2}\sum_{i=1}^n Z_i^{(k)}\tilde{\mathcal{J}}_{\theta,i}$ are asymptotically equivalent. Hence,

$$\begin{aligned}&\left(\tilde{\alpha}_n(u,v), \Theta_n, \alpha_n^{(1)}(u,v), \Theta_n^{(1)}, \ldots, \alpha_n^{(N)}(u,v), \Theta_n^{(N)}\right) \\ &\quad = \left(\tilde{\alpha}_n(u,v), \tilde{\Theta}_n, \tilde{\alpha}_n^{(1)}(u,v), \tilde{\Theta}_n^{(1)}, \ldots, \tilde{\alpha}_n^{(N)}(u,v), \tilde{\Theta}_n^{(N)}\right) + R_n(u,v), \quad (14)\end{aligned}$$

where $\sup_{(u,v)\in[0,1]^2}|R_n(u,v)| \xrightarrow{\mathbb{P}} 0$. Now, $\left(\tilde{\alpha}_n(u,v), \tilde{\Theta}_n, \tilde{\alpha}_n^{(1)}(u,v), \tilde{\Theta}_n^{(1)}, \ldots, \tilde{\alpha}_n^{(N)}(u,v), \tilde{\Theta}_n^{(N)}\right)$ can be written as the sum of i.i.d. random vectors

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n \Big(\{\mathbf{1}(U_i \le u, V_i \le v) - C_\theta(u,v)\}, \tilde{\mathcal{J}}_{\theta,i}, Z_i^{(1)}\{\mathbf{1}(U_i \le u, V_i \le v) - C_\theta(u,v)\}, Z_i^{(1)}\tilde{\mathcal{J}}_{\theta,i}, \ldots \\ \ldots, Z_i^{(N)}\{\mathbf{1}(U_i \le u, V_i \le v) - C_\theta(u,v)\}, Z_i^{(N)}\tilde{\mathcal{J}}_{\theta,i}\Big),$$

which, from the multivariate central limit theorem, converges in distribution to

$$\left(\alpha_\theta(u,v), \Theta, \alpha_\theta^{(1)}(u,v), \Theta^{(1)}, \ldots, \alpha_\theta^{(N)}(u,v), \Theta^{(N)}\right)$$

where $(\alpha_\theta^{(1)}(u,v), \Theta^{(1)}), \ldots, (\alpha_\theta^{(N)}(u,v), \Theta^{(N)})$ are independent copies of $(\alpha_\theta(u,v), \Theta)$. Similarly, for any finite collection of points $(u_1,v_1), \ldots, (u_k,v_k)$ in $[0,1]^2$,

$$\Big(\tilde{\alpha}_n(u_1,v_1), \ldots, \tilde{\alpha}_n(u_k,v_k), \tilde{\Theta}_n, \tilde{\alpha}_n^{(1)}(u_1,v_1), \ldots, \tilde{\alpha}_n^{(1)}(u_k,v_k), \tilde{\Theta}_n^{(1)}, \ldots \\ \ldots \tilde{\alpha}_n^{(N)}(u_1,v_1), \ldots, \tilde{\alpha}_n^{(N)}(u_k,v_k), \tilde{\Theta}_n^{(N)}\Big)$$

converges in distribution, i.e., we have convergence of the finite dimensional distributions. Tightness follows from the fact that $\tilde{\alpha}_n$ and $\tilde{\alpha}_n^{(k)}$ each converge weakly in $\ell^\infty([0,1]^2)$, so we obtain that

$$\left(\tilde{\alpha}_n, \tilde{\Theta}_n, \tilde{\alpha}_n^{(1)}, \tilde{\Theta}_n^{(1)}, \ldots, \tilde{\alpha}_n^{(N)}, \tilde{\Theta}_n^{(N)}\right) \rightsquigarrow \left(\alpha_\theta, \Theta, \alpha_\theta^{(1)}, \Theta^{(1)}, \ldots, \alpha_\theta^{(N)}, \Theta^{(N)}\right)$$

in $(\ell^\infty([0,1]^2)\otimes\mathbb{R})^{\otimes N+1}$. It follows from (14) that

$$\left(\tilde{\alpha}_n, \Theta_n, \alpha_n^{(1)}, \Theta_n^{(1)}, \ldots, \alpha_n^{(N)}, \Theta_n^{(N)}\right) \rightsquigarrow \left(\alpha_\theta, \Theta, \alpha_\theta^{(1)}, \Theta^{(1)}, \ldots, \alpha_\theta^{(N)}, \Theta^{(N)}\right)$$

in $(\ell^\infty([0,1]^2)\otimes\mathbb{R})^{\otimes N+1}$. Then, from the continuous mapping theorem,

$$\begin{aligned}
\Big(&\tilde{\alpha}_n(u,v) - C_\theta^{[1]}(u,v)\tilde{\alpha}_n(u,1) - C_\theta^{[2]}(u,v)\tilde{\alpha}_n(1,v) - \Theta_n\dot{C}_\theta(u,v),\\
&\alpha_n^{(1)}(u,v) - C_\theta^{[1]}(u,v)\alpha_n^{(1)}(u,1) - C_\theta^{[2]}(u,v)\alpha_n^{(1)}(1,v) - \Theta_n^{(1)}\dot{C}_\theta(u,v),\\
&\vdots\\
&\alpha_n^{(N)}(u,v) - C_\theta^{[1]}(u,v)\alpha_n^{(N)}(u,1) - C_\theta^{[2]}(u,v)\alpha_n^{(N)}(1,v) - \Theta_n^{(N)}\dot{C}_\theta(u,v)\Big)
\end{aligned}$$

converges weakly to (9) in $\ell^\infty([0,1]^2)^{\otimes(N+1)}$. Now, from the work of Stute (1984, page 371) (see also Tsukahara, 2005, Proposition 1), we have that

$$\sqrt{n}\{C_n(u,v) - C_\theta(u,v)\} = \tilde{\alpha}_n(u,v) - C_\theta^{[1]}(u,v)\tilde{\alpha}_n(u,1) - C_\theta^{[2]}(u,v)\tilde{\alpha}_n(1,v) + Q_n(u,v),$$

where $\sup_{(u,v)\in[0,1]^2}|Q_n(u,v)| \xrightarrow{\mathbb{P}} 0$. Furthermore, from the work of Quessy (2005, page 73) and under Assumption A3, we can write

$$\sqrt{n}\{C_{\theta_n}(u,v) - C_\theta(u,v)\} = \Theta_n\dot{C}_\theta(u,v) + T_n(u,v)$$

where $\sup_{(u,v)\in[0,1]^2}|T_n(u,v)| \xrightarrow{\mathbb{P}} 0$. It follows that

$$\begin{aligned}
\sqrt{n}\{C_n(u,v) - C_{\theta_n}(u,v)\} &= \sqrt{n}\{C_n(u,v) - C_\theta(u,v)\} - \sqrt{n}\{C_{\theta_n}(u,v) - C_\theta(u,v)\}\\
&= \tilde{\alpha}_n(u,v) - C_\theta^{[1]}(u,v)\tilde{\alpha}_n(u,1) - C_\theta^{[2]}(u,v)\tilde{\alpha}_n(1,v) - \Theta_n\dot{C}_\theta(u,v) + Q_n(u,v) - T_n(u,v).
\end{aligned}$$

Using the fact that $C_n^{[1]}$ and $C_n^{[2]}$ converge uniformly in probability to $C_\theta^{[1]}$ and $C_\theta^{[1]}$ respectively (Rémillard and Scaillet, 2009, Prop. A.2) and the fact that, from Assumption A3, $\dot{C}_{\theta_n}$ converges uniformly in probability to $\dot{C}_\theta$, we finally obtain that

$$\begin{aligned}
\Big(&\sqrt{n}\{C_n(u,v) - C_{\theta_n}(u,v)\},\\
&\alpha_n^{(1)}(u,v) - C_n^{[1]}(u,v)\alpha_n^{(1)}(u,1) - C_n^{[2]}(u,v)\alpha_n^{(1)}(1,v) - \Theta_n^{(1)}\dot{C}_{\theta_n}(u,v),\\
&\vdots\\
&\alpha_n^{(N)}(u,v) - C_n^{[1]}(u,v)\alpha_n^{(N)}(u,1) - C_n^{[2]}(u,v)\alpha_n^{(N)}(1,v) - \Theta_n^{(N)}\dot{C}_{\theta_n}(u,v)\Big)
\end{aligned}$$

also converges weakly to (9) in $\ell^\infty([0,1]^2)^{\otimes(N+1)}$. □

B Computational details

The computations presented in Sections 4 and 5 were performed using the R statistical system (R Development Core Team, 2009) and rely on C code for the most computationally demanding parts. Table 4, mainly taken from Frees and Valdez (1998), gives the c.d.f.s, Kendall's tau, and Spearman's rho for the one-parameter copula families considered in the simulation study. The functions D_1 and D_2 used in the expressions of Kendall's tau and Spearman's rho for the Frank copula are the so-called first and second *Debye functions* (see e.g. Genest, 1987).

Table 4: C.d.f.s, Kendall's tau and Spearman's rho for the one-parameter copula families considered in the simulation study. The functions D_1 and D_2 used in the expressions of Kendall's tau and Spearman's rho for the Frank copula are the so-called first and second *Debye functions* (see e.g. Genest, 1987).

Family	C.d.f. C_θ of the copula	Parameter range	Kendall's tau	Spearman's rho
Clayton	$\left(u^{-\theta}+v^{-\theta}-1\right)^{-1/\theta}$	$(0,+\infty)$	$\dfrac{\theta}{\theta+2}$	Numerical approximation
Gumbel	$\exp\left\{-\left((-\log u)^\theta+(-\log v)^\theta\right)^{1/\theta}\right\}$	$[1,+\infty)$	$1+\dfrac{1}{\theta}$	Numerical approximation
Frank	$-\dfrac{1}{\theta}\log\left(1+\dfrac{(\exp(-\theta u)-1)(\exp(-\theta v)-1)}{(\exp(-\theta)-1)}\right)$	$\mathbb{R}\setminus\{0\}$	$1-\dfrac{4}{\theta}\left[D_1(-\theta)-1\right]$	$1-\dfrac{12}{\theta}\left[D_2(-\theta)-D_1(-\theta)\right]$
Plackett	$\dfrac{1}{2(\theta-1)}\Big\{1+(\theta-1)(u+v)$ $-\left[(1+(\theta-1)(u+v))^2+4uv\theta(1-\theta)\right]^{1/2}\Big\}$	$[0,+\infty)$	Numerical approx.	$\dfrac{\theta+1}{\theta-1}-\dfrac{2\theta\log\theta}{(\theta-1)^2}$
Normal	$\Phi_\theta(\Phi^{-1}(u),\Phi^{-1}(v))$, where Φ_θ is the bivariate standard normal c.d.f. with correlation θ, and Φ is the c.d.f. of the univariate standard normal.	$[-1,1]$	$\dfrac{2}{\pi}\arcsin(\theta)$	$\dfrac{6}{\pi}\arcsin\left(\dfrac{\theta}{2}\right)$
t_ν	$t_{\nu,\theta}(t_\nu^{-1}(u),t_\nu^{-1}(v))$, where $t_{\nu,\theta}$ is the bivariate standard t c.d.f. with ν degrees of freedom and correlation θ, and t_ν is the c.d.f. of the univariate standard t with ν degrees of freedom.	$[-1,1]$	$\dfrac{2}{\pi}\arcsin(\theta)$	Numerical approximation

B.1 Expressions of τ^{-1}, ρ^{-1}, τ' and ρ'

Expressions of τ^{-1} and ρ^{-1} required for the moment estimation of the copula parameter can be obtained in most cases from the last two columns of Table 4. The same holds for the expressions of τ' and ρ' necessary for computing the score functions $J_{\theta,\tau}$ and $J_{\theta,\rho}$. The following cases are not straightforward:

- For the Frank copula, Spearman's rho and Kendall's tau were inverted numerically. Furthermore, in this case, the derivative of ρ is given by

$$\rho'(\theta) = \frac{12}{\theta(\exp(\theta) - 1)} - \frac{36}{\theta^2} D_2(\theta) + \frac{24}{\theta^2} D_1(\theta).$$

- For the Plackett copula, we proceeded as follows. First, using Monte Carlo integration, Kendall's tau was computed at a dense grid of θ values. Spline interpolation was then used to compute τ between the grid points. The values of τ at the grid points need only to be computed once. The derivative τ' was computed similarly. More details can be found in Kojadinovic and Yan (2009).

- For the Clayton, Gumbel and t_4 copulas, ρ and ρ' were computed using the same approach as for Kendall's tau for the Plackett copula.

B.2 Expressions of $\dot{C}_\theta$

Only the case of the two meta-elliptical copulas is not immediate. For the bivariate normal copula, the expression of $\dot{C}_\theta$ follows from the so-called *Plackett formula* (Plackett, 1954):

$$\frac{\partial \Phi_\theta(x, y)}{\partial \theta} = \frac{\exp\left(-\frac{x^2+y^2-2\theta xy}{2(1-\theta^2)}\right)}{2\pi\sqrt{1-\theta^2}},$$

where Φ_θ is the bivariate standard normal c.d.f. with correlation θ. The bivariate t generalization of the Plackett formula is given in Genz (2004):

$$\frac{\partial t_{\nu,\theta}(x, y)}{\partial \theta} = \frac{\left(1 + \frac{x^2+y^2-2\theta xy}{\nu(1-\theta^2)}\right)^{-\nu/2}}{2\pi\sqrt{1-\theta^2}},$$

where $t_{\nu,\theta}$ is the bivariate standard t c.d.f. with ν degrees of freedom and correlation θ.

References

Ben Ghorbal, M., Genest, C., and Nešlehová, J. (2009). On the test of Ghoudi, Khoudraji, and Rivest for extreme-value dependence. *The Canadian Journal of Statistics* **37**, 534–552.

Berg, D. (2009). Copula goodness-of-fit testing: An overview and power comparison. *The European Journal of Finance* **15**, 675–701.

Berg, D. and Quessy, J.-F. (2009). Local sensitivity analyses of goodness-of-fit tests for copulas. *Scandinavian Journal of Statistics* **36**, 389–412.

Chen, X. and Fan, Y. (2005). Pseudo-likelihood ratio tests for semiparametric multivariate copula model selection. *Canadian Journal of Statistics* **33**, 389–414.

Cherubini, G., Vecchiato, W., and Luciano, E. (2004). *Copula models in finance.* Wiley, New-York.

Cui, S. and Sun, Y. (2004). Checking for the gamma frailty distribution under the marginal proportional hazards frailty model. *Statistica Sinica* **14**, 249–267.

Deheuvels, P. (1981). A non parametric test for independence. *Publications de l'Institut de Statistique de l'Université de Paris* **26**, 29–50.

Fermanian, J.-D. (2005). Goodness-of-fit tests for copulas. *Journal of Multivariate Analysis* **95**, 119–152.

Fermanian, J.-D., Radulovic, D., and Wegkamp, M. (2004). Weak convergence of empirical copula processes. *Bernoulli* **10**, 847–860.

Fine, J., Yan, J., and Kosorok, M. (2004). Temporal process regression. *Biometrika* **91**, 683–703.

Frees, E. and Valdez, E. (1998). Understanding relationships using copulas. *North American Actuarial Journal* **2**, 1–25.

Gänssler, P. and Stute, W. (1987). *Seminar on empirical processes.* DMV Seminar 9. Birkhäuser, Basel.

Genest, C. (1987). Frank's family of bivariate distributions. *Biometrika* **74**, 549–555.

Genest, C. and Favre, A.-C. (2007). Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of Hydrological Engineering* **12**, 347–368.

Genest, C., Ghoudi, K., and Rivest, L.-P. (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika* **82**, 543–552.

Genest, C., Quessy, J.-F., and Rémillard, B. (2006). Goodness-of-fit procedures for copulas models based on the probability integral transformation. *Scandinavian Journal of Statistics* **33**, 337–366.

Genest, C. and Rémillard, B. (2008). Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models. *Annales de l'Institut Henri Poincaré: Probabilités et Statistiques* **44**, 1096–1127.

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.

Genz, A. (2004). Numerical computation of rectangular bivariate and trivariate normal and t probabilities. *Statistics and Computing* **14**, 251–260.

Hájek, J., Šidák, Z., and Sen, P. (1999). *Theory of rank tests, 2nd edition.* Academic Press.

Joe, H. (1997). *Multivariate models and dependence concepts.* Chapman and Hall, London.

Klugman, S. and Parsa, R. (1999). Fitting bivariate loss distributions with copulas. *Insurance: Mathematics and Economics* **24**, 139–148.

Kojadinovic, I. and Yan, J. (2009). Comparison of three semiparametric methods for estimating dependence parameters in copula models. *Insurance: Mathematics and Economics* **in press**.

Kojadinovic, I. and Yan, J. (2010a). A goodness-of-fit test for multivariate multiparameter copulas based on multiplier central limit theorems. *Statistics and Computing* **in press**.

Kojadinovic, I. and Yan, J. (2010b). Modeling multivariate distributions with continuous margins using the `copula` R package. *Journal of Statistical Software* **34**, 1–20.

Kosorok, M. (2008). *Introduction to empirical processes and semiparametric inference.* Springer, New York.

Lin, D., Fleming, T., and Wei, L. (1994). Confidence bands for survival curves under the proportional hazards model. *Biometrika* **81**, 73–81.

McNeil, A., Frey, R., and Embrechts, P. (2005). *Quantitative risk management.* Princeton University Press, New Jersey.

Plackett, R. (1954). A reduction formula for normal multivariate probabilities. *Biometrika* **41**, 351–369.

Quessy, J.-F. (2005). *Méthodologie et application des copules: tests d'adéquation, tests d'indépendance, et bornes sur la valeur-à-risque.* Ph.D. thesis, Université Laval, Québec, Canada.

R Development Core Team (2009). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Rémillard, B. and Scaillet, O. (2009). Testing for equality between two copulas. *Journal of Multivariate Analysis* **100**, 377–386.

Salvadori, G., Michele, C. D., Kottegoda, N., and Rosso, R. (2007). *Extremes in Nature: An Approach Using Copulas.* Water Science and Technology Library, Vol. 56. Springer.

Scaillet, O. (2005). A Kolmogorov-Smirnov type test for positive quadrant dependence. *Canadian Journal of Statistics* **33**, 415–427.

Shih, J. and Louis, T. (1995). Inferences on the association parameter in copula models for bivariate survival data. *Biometrics* **51**, 1384–1399.

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Université de Paris* **8**, 229–231.

Stute, W. (1984). The oscillation behavior of empirical processes: The multivariate case. *Annals of Probability* **12**, 361–379.

Tsukahara, H. (2005). Semiparametric estimation in copula models. *The Canadian Journal of Statistics* **33**, 357–375.

Yan, J. and Kojadinovic, I. (2010). *copula: Multivariate dependence with copulas.* R package version 0.9-5.