A TEST OF GOODNESS-OF-FIT FOR THE COPULA DENSITIES

Ghislaine GAYRAUD [a] and Karine TRIBOULEY [b]

[a] Université Rouen
76801 Saint-Étienne-du-Rouvray, France
and
LS-CREST
Timbre J340, 3 av. P. Larousse
92241 Malakoff Cedex, France
Ghislaine.Gayraud@univ-rouen.fr

[b] LPMA, Université Paris VII 175 rue du Chevaleret
75013 Paris, France
and
MODALX, Université Paris X, 200 avenue de la république
92001, Nanterre, France
karine.tribouley@u-paris10.fr

Abstract

We consider the problem of testing hypotheses on the copula density from n bi-dimensional observations. We wish to test the null hypothesis characterized by a parametric class against a composite nonparametric alternative. The density under the alternative are separated in the L_2-norm from any density lying in the null hypothesis. The copula densities under consideration are supposed to belong to a range of Besov balls. According to the minimax approach, the testing problem is solved in an adaptive framework: it leads to a log term loss in the minimax rate of testing in comparison with the non-adaptive case. A smoothness-free test statistic that achieves the minimax rate is proposed. The lower bound is also proved. The empirical behavior of the test procedure is also studied with both simulated and real data.

Index Terms — copulas, adaptive testing, composite null hypothesis, minimax rate of testing

1 Introduction

Copulas became a very popular and attractive tool in the recent literature for modeling multivariate observations. The nice feature of copulas is that they capture the structure dependence among the components of a multivariate observation without requiring the study of the univariate margins. More precisely, Sklar's Theorem ensures that any $d-$varied distribution function H may be expressed as

$$H(x^1, \ldots, x^d) = C\left(F^1(x^1), \ldots, F^d(x^d)\right),$$

where the F^p's are the margins and C is called the copula function. Sklar (1959) states the existence and the uniqueness of C as soon as the random variables with joint law H are continuous.

Modeling the dependence is a great challenge in statistics, specially in finance or assurance where (for instance) the identification of the dependence structure between assets is essential. Many authors proposed parametrical families of copulas $\{C_\lambda, \lambda \in \Lambda\}$, each of them being available to capture different dependence behavior. The elliptic family contains the Gaussian Copulas and the Student Copula which are often used in finance. For insurance purposes, heavy tails are needed and copulas coming from the archimedian family are used. Among others, the more common are the Gumbel Copula, the Clayton Copula or the Frank Copula. In view to illustrate the different behaviours of the tails of several copula densities, we present graphs corresponding to the models cited above. The parameters are chosen such a way that the associated Kendall's tau (i.e. the indicator of concordance/discordance) is identical in all illustrations.

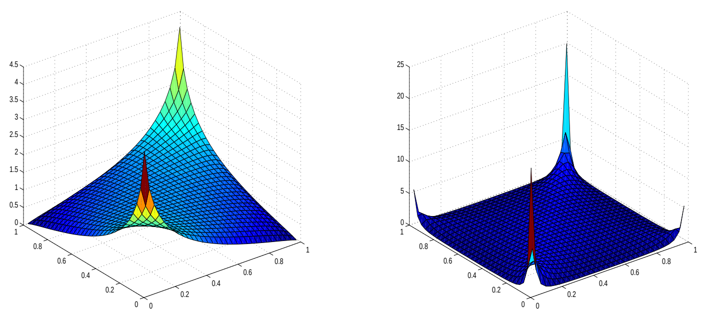

Figure 1: Kendall's tau$= 0.25$. Left: Bi-dimensional Gaussian copula density with parameter $\rho = 0.4$. Right: Bi-dimensional Student copula density with parameter $(\rho, \nu) = (0.4, 1)$.

Since many parametric copula models are now available, the crucial choice for the practitioner is to identify the model which is well-adapted

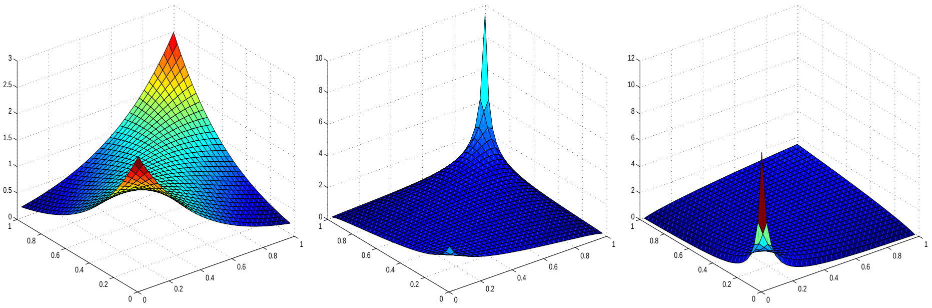

Figure 2: Kendall's tau= 0.25. Left: Bi-dimensional Frank copula density with parameter $\theta = 2.5$. Center: Bi-dimensional Gumbel copula density with parameter $\theta = 1.33$. Right: Bidimensional Clayton copula density with parameter $\theta = 0.66$.

to his data. Many goodness-of-fit tests are proposed in the literature. Genest et al. (2008b) give an excellent review and propose a detailed empirical study for different tests: we refer to this paper for any supplementary references. Roughly speaking, they study procedures based on empirical processes. Among others, they deal with rank-based versions of the Cramér-von-Mises and Kolmogorov-Smirnov statistics. They also consider test based on Kendall's transform. Basically, they restrict themselves to test statistics built from empirical distributions (empirical copula or transform of this latter). On a theoretical point of view, the asymptotic law under the null of the test statistic is stated in a number of papers (see by instance Deheuvels (1979), Deheuvels (1981a) and Deheuvels (1981b)). It allows in particular to derive the critical value but generally the alternative is unspecified and the properties on the power are empirically given from simulations.

In our paper, it is supposed that the copula C admits a density copula c with respect to the Lebesgue measure. To our knowledge, Fermanian (2005) was the first author to propose a goodness-of-fit test based on nonparametric kernel estimations of the density copula. In the same spirit as the papers cited above, he derived the asymptotic law of the test statistic under the null. His results are valid for bandwidths greater than $n^{-2/(8+d)}$ which correspond to enough smooth copula densities.

In this paper, we focus on the minimax theory framework: we define the test problem as initiated by Ingster (1982). One of the advantages of this point of view is to precisely define the alternative: it is then possible to quantify the risk associated to the test problem as the sum of the first type error and the second type of error. Since this risk measure provides a quality criterion, it is possible to compare the test procedures. Indeed, the alternative $H_1(v_n)$ is defined from a positive quantity v_n measuring the distance between the null and the latter. Obviously, the larger is this separating distance, the easier is the decision. The aim of the minimax theory is to determine the larger alternative for which the decision remains feasible. Solving **the lower bound problem** is equivalent to exhibit the faster separating rate v_n such that the risk is bounded from below by a given positive constant α: this rate is called **the minimax rate of testing**. Next, **the upper bound problem** has to be solved exhibiting a test procedure whose risk is bounded from above by a given α, that is, the statistic test allows to distinguish the null from $H_1(v_n)$, where v_n is the minimax rate.

In the white noise model or in the density model, the goodness-of-fit problem (stands as explained above) was solved for different regularity classes (Hölder or Sobolev or Besov) associated with various geometries: pointwise, quadratic and supremum norm. For fixed smoothness of the unknown density (**minimax context**), there is a rich literature summed-up in Ingster (1993) and in Ingster and Suslina (2002). Optimal test procedures include orthogonal projections, kernel estimates or χ^2 procedures. Goodness-of-fit tests with alternatives of variable smoothness into some given interval (**adaptive context**) were introduced by Spokoiny (1996) for L_2 distance in the Gaussian white noise model and generalized by Spokoiny (1998) to L_p distances. Ingster (2000) proved that a collection of χ^2 tests attains the adaptive rates of goodness-of-fit tests in L_2 distance as well as for the density model.

For sake of simplicity, we restrict ourselves to bi-dimensional data but there is no theoretical obstacle to generalize our results to higher dimensions. Suppose that we observe n i.i.d. copies $(X_i, Y_i)_{i\in\mathcal{I}}$ where $\mathcal{I} = \{1, \dots, n\}$ of (X, Y). The random vector (X, Y) is drawn from the distribution function H expressed through the copula C. Moreover, it is assumed that C has a copula density c with respect to the Lebesgue measure on $\mathbb{R}^2$ and F and G stand for the cdf's of X and Y respectively. From $(X_i, Y_i)_{i\in\mathcal{I}}$, we are interested in studying the goodness-of-fit problem when the null is a composite hypothesis $H_0 : c \in \mathcal{C}$ for a general class $\mathcal{C}$ of parametrical copula densities. Since the alternative is defined from the quadratic distance, we propose a goodness-of-fit test based on wavelet estimation of an integrated functional of the copula density. Indeed, Genest et al. (2008a) and Autin et al. (2008) show that the wavelet methods are an efficient tool to estimate the copula densities since these latter have very specifics behaviors. Unfortunately no direct observations $(F(X_i), G(Y_i))$ for $i \in \mathcal{I}$ are available since F and G are unknown, the test statistic is then built with pseudo-observations $(\widehat{F}(X_i), \widehat{G}(Y_i))_{i\in\mathcal{I}}$: as usual in the copula context, the quantities of interest are rank-based statistics. We provide an auto-driven test procedure and we produce its rate when the alternative contains a regular constraint: since the procedure is based on wavelet methods, the linked functional classes are the Besov classes $B_{s,p,q}$. We give results for $p \geq 2$ (**dense case**) and $s \geq 1/2$. The case $p < 2$ is called the **sparse case** and leads to different minimax rates (see Lepskii and Spokoiny (1999)); in this case, another test strategy must be constructed and it will be explored in a further work. The constraint $s \geq 1/2$ is due to the fact that pseudo-data are used and then a minimal regularity is required in order to pay no attention to substitute the direct data with the ranked data. Observe that Kerkyacharian and Picard (2004) have the same constraint in the univariate regression model when the design is random with unknown distribution. Next, we prove that our procedure is minimax (and adaptive) optimal by exhibiting the minimax adaptive rate. This one looks like the minimax rate but an extra $\log \log$ term appears: we prove that this loss is the price to paid for adaptivity. To our knowledge, the proof of the adaptive lower bound in the multivariate density model when the null is composite has never been clearly written.

Next, we allocate a large part in this paper to empirical studies. Simulation allows us to show that, when the theoretical framework is respected, the power qualities of our test procedures are good. We choose to make simulations starting from the parametrical copula families presented in the beginning of the introduction and which are the more common for applications. We compare our simulation results with those of Genest et al. (2008b). Then, we study a very well known sample of real life data of Frees and Valdez (1998) consisting of the indemnity payment (LOSS) and the allocated loss adjustment expense (ALAE) for 1500 general liability claims. The most popular model for the copula is a Gumbel copula model with parameter $\theta = 1.45$ (which may be estimated by inverting the Kendall's tau) given in Figure 3. Among other results, it is empirically shown that the Gumbel and the Gaussian copula models are acceptable while Student, Clayton or Frank models are rejected. Figure 3 gives a wavelet estimator of the copula density of $(LOSS, ALEA)$ by the method explained in Autin et al. (2008). Visually, fitting the unknown copula with the Gumbel model seems indeed to be the most appropriated.

The paper is organized as follows. In Section 2, we first provide a general description of orthonormal wavelet bases, focusing on the mathematical properties that are essential to the construction of the statistics we consider. In Section 3, we provide the inference procedures: first, we explain how to estimate the square L_2-norm of the copula density and next we derive the procedure of goodness-of-fit. The theoretical part is exposed in Section 4: first, we state very precisely the test problem under consideration; we define the criterion allowing to measure the quality of test procedures and define the separating minimax rate. In Section 5 the main results are stated: our test procedure is shown to be optimal in the sense defined in the previous section. Section 6 is devoted to practical results: first, we present an intensive simulation study and next a real-life data is exploited. We conclude these parts with a discussion in Section 7. The proof of the upper bound is given in Section 8 while the proof of the lower bound is given in Section 9. Finally, all technical or computational lemmas which are not essential to

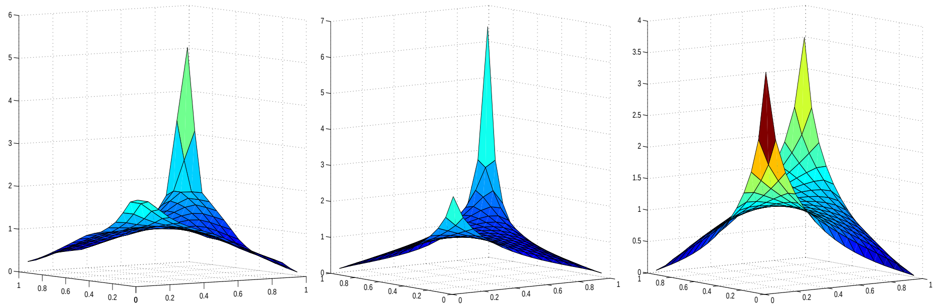

Figure 3: Left: Thresholded wavelet estimator for the copula density of $(LOSS, ALEA)$ as given in Autin et al. (2008). Center: Gumbel copula density with parameter $\theta = 1.45$. Right: Gaussien copula density with parameter $\rho = 0.48$.

understand the main proofs, are postponed in Appendix.

2 Wavelet Setting

2.1 Wavelet expansion

In the univariate case, let ϕ be a scaling function and ψ its associated wavelet function, which are chosen compactly supported on $[0, L]$, $L > 0$. Let j in $\mathbb{N}$, k_1 in $\mathbb{Z}$ and for any univariate function Φ, set $\Phi_{j,k_1}(\cdot) = 2^{j/2}\Phi(2^j \cdot -k_1)$. In the sequel, we use wavelet expansions for bivariate functions and we keep the same notation as for the univariate case. Then, a bivariate wavelet basis is built as follows:

$$\phi_{j,k}(x,y) = \phi_{j,k_1}(x)\phi_{j,k_2}(y), \qquad \psi^{(1)}_{j,k}(x,y) = \phi_{j,k_1}(x)\psi_{j,k_2}(y),$$
$$\psi^{(2)}_{j,k}(x,y) = \psi_{j,k_1}(x)\phi_{j,k_2}(y), \qquad \psi^{(3)}_{j,k}(x,y) = \psi_{j,k_1}(x)\psi_{j,k_2}(y),$$

where the subscript $k = (k_1, k_2)$ indicates the number of components of the functions $\phi_{j,k}$ and $\psi_{j,k}$. For a given $j \in \mathbb{N}$, the set

$$\{\phi_{j,k}, \psi^{\epsilon}_{\ell,k'},\ \ell \geq j, (k,k') \in \mathbb{Z}^2 \times \mathbb{Z}^2, \epsilon = 1,2,3\}$$

is an orthonormal basis of $L^2(\mathbb{R}^2)$ and the expansion of any real bivariate function Φ in $L^2(\mathbb{R}^2)$ is given by:

$$\Phi(x,y) = \sum_{k \in Z^2} A_{j,k}\phi_{j,k}(x,y) + \sum_{\ell=j}^{\infty} \sum_{k \in \mathbb{Z}^2} \sum_{\epsilon=1,2,3} B^{\epsilon}_{\ell,k}\psi^{\epsilon}_{\ell,k}(x,y),$$

where the scaling coefficients and the wavelet coefficients are

$$\forall j \in \mathbb{N}, \forall k \in \mathbb{Z}^2, \quad A_{j,k} = \int_{\mathbb{R}^2} \Phi \phi_{j,k}, \qquad B^{\epsilon}_{j,k} = \int_{\mathbb{R}^2} \Phi \psi^{\epsilon}_{j,k}.$$

The Parseval equality immediately leads to the expansion of the square L_2-norm of the function Φ:

$$\int \Phi^2 \quad = \quad T_j + B_j, \tag{1}$$

where the trend and the detail terms are respectively:

$$T_j = \sum_{k \in \mathbb{Z}^2} (A_{j,k})^2 \text{ and } B_j = \sum_{\ell=j}^{\infty} \sum_{k \in \mathbb{Z}^2} \sum_{\epsilon=1}^{3} (B^{\epsilon}_{\ell,k})^2. \tag{2}$$

Note that if Φ has a compact support on $[a_1, b_1] \times [a_2, b_2]$, the sum over the indices k is finite: there are no more than $(2^j(b_1 - a_1) + L)(2^j(b_2 - a_2) + L)$ terms in the sum (recall that L is the length of the support of ϕ); this is the case of the functions c under consideration since they are copula densities whose support is $[0, 1]^2$. In order to simplify the notations, the bounds of variation of k and ϵ in expansion of any Φ, are omitted in the sequel.

2.2 Besov Bodies and Besov spaces

Since we deal with wavelet expansions, it is natural to consider Besov bodies as functional spaces because they are characterized in term of wavelet coefficients as follows.

Definition 1. *For any $s > 0, p \geq 1$ and any radius $M > 0$, a function Φ defined on $\mathbb{R}^d$ belongs to the ball $b_{s,p,\infty}(M)$ of the Besov body $b_{s,p,\infty}$ if and only if its sequence of wavelet coefficients $B^{\epsilon}_{j,k}$ satisfies*

$$\forall j \in \mathbb{N}, \sum_{k \in \mathbb{Z}^2} \sum_{\epsilon=1}^{3} |B^{\epsilon}_{j,k}|^p < M 2^{-j(s+d/2-d/p)p}.$$

The Besov body $b_{s,p,\infty}$ coincides with the more standard Besov space $\mathcal{B}_{s,p,\infty}$ when there exists an integer N strictly larger than s and such that the $q-$th moment of the wavelet ψ vanishes for any $q = 0, \ldots, N-1$. It is possible to build univariate wavelets whose support is included in $[0, 2N-1]$ satisfying this property for any choice of N (see the Daubechies wavelets). An advantage of the Besov spaces $\mathcal{B}_{s,p,\infty}$ is to provide an useful tool to classify wavelet decomposed signals according to their regularity and sparsity properties (see for instance Donoho and Johnstone (1994a), Donoho and Johnstone (1994b)). Roughly speaking, the regularity increases with s, whereas the sparsity decreases with p; in particular, the spaces with indice $p < 2$ describe very wide classes of inhomogeneous but sparse functions. In this paper, only the case $p \geq 2$ is considered.

In the sequel, we need to bound the detail term B_j defined in (2). We use the following inequality

$$\forall j \in \mathbb{N}, \quad B_j \leq \sum_{\ell=j}^{\infty} \left(\sum_{k\in\mathbb{Z}^2} \sum_{\epsilon=1}^{3} |B_{\ell,k}^{\epsilon}|^p \right)^{2/p} \left(K\, 2^{2j}\right)^{1-2/p}$$

where K is a positive constant depending on the supports of Φ and ψ. Assuming that the function Φ belongs to $b_{s,p,\infty}(M)$ with s, p and M as in Definition 1, the following inequality holds

$$\forall j \in \mathbb{N}, \quad B_j \quad \leq \quad \tilde{K}\, 2^{-2js}, \tag{3}$$

where $\tilde{K}$ is a positive constant depending on the supports of Φ, ψ and on the radius M. When Φ is a copula density, $\tilde{K} = M^{2/p}\left(3(L+1)^2\right)^{1-2/p}$.

3 Statistical Procedures

Assuming that the copula density c belongs to L_2, we first explain the procedure to estimate the square L_2-norm of c

$$\theta = \|c\|^2 := \int_{[0,1]^2} c^2$$

which is used to define the alternative of the goodness-of-fit test. The statistical methods depend on parameters (the level j for the estimation procedure and j and the critical value t_j for the test procedure) which are discussed and determined in an optimal way in Section 5.

It is fundamental to note that, for any bivariate function Φ, one has

$$\mathbb{E}_c\left[\Phi(U,V)\right] \quad = \quad \mathbb{E}_h\left[\Phi(F(X),G(Y))\right], \tag{4}$$

where h stands for the joint density of (X,Y). This means in particular that the wavelet coefficients $\{c_{j,k}, c_{\ell,k}^{\epsilon},\ \ell \geq j, k \in \mathbb{Z}^2, \epsilon = 1,2,3\}$ of the copula density c on the wavelet basis

$$\{\phi_{j,k}, \psi_{\ell,k}^{\epsilon},\ \ell \geq j, k \in \mathbb{Z}^2, \epsilon = 1,2,3\}$$

are equal to the coefficients of the joint density h on the warped wavelet family

$$\{\phi_{j,k}(F(\cdot),G(\cdot)), \psi_{\ell,k}^{\epsilon}(F(\cdot),G(\cdot)),\ \ell \geq j, k \in \mathbb{Z}^2, \epsilon = 1,2,3\}.$$

The statistical procedures are based on the wavelet expansion of the copula density c, for which the wavelet coefficients have to be estimated.

3.1 Procedures to estimate θ

Let J be a subset of $\mathbb{N}$ and consider a given j in J. Motivated by the wavelet expansion (1), we propose to estimate θ with an estimator of the trend T_j omitting the detail term B_j. Using the orthonormality property of the wavelet basis, it leads to estimate the square of the coefficients of the copula density on the scaling function. As usual, a $U-$statistic associated to the empirical coefficients is used in order to remove the bias terms. Due to (4), we first consider the following family of statistics $\{\widehat{T}_j, j \in J\}$ defined by

$$\widehat{T}_j \;=\; \sum_k \widehat{\theta_{j,k}},$$

where $\widehat{\theta_{j,k}}$ is the following $U-$statistic

$$\widehat{\theta_{j,k}} \;=\; \frac{1}{n(n-1)} \sum_{\substack{i_1,i_2=1 \\ i_1 \neq i_2}}^{n} \phi_{j,k}\left(F(X_{i_1}), G(Y_{i_1})\right) \phi_{j,k}\left(F(X_{i_2}), G(Y_{i_2})\right).$$

Since no direct observation $(F(X_i), G(Y_i))$ is usually available, it is replaced in $\widehat{\theta_{j,k}}$ by the pseudo observation $(\widehat{F}(X_i), \widehat{G}(Y_i))$, where $\widehat{F}, \widehat{G}$ denote some estimator of the margins. To preserve the independence given by the observations, we split the initial sample $(X_i, Y_i)_{i \in \mathcal{I}}$ into disjoint samples $(X_i, Y_i)_{i \in \mathcal{I}_1}$ and $(X_i, Y_i)_{i \in \mathcal{I}_2}$ with $\mathcal{I}_2 \cup \mathcal{I}_1 = \mathcal{I}$, $\mathcal{I}_2 \cap \mathcal{I}_1 = \emptyset$, and whose size is n_1, n_2 respectively. The sub-sample with indices in $\mathcal{I}_1$ is used to estimate the marginal distributions and the second one with indices in $\mathcal{I}_2$ is devoted to the computation of the U-statistic. We consider the usual empirical distribution functions:

$$\widehat{F}(x) = \frac{1}{n_1} \sum_{i \in \mathcal{I}_1} \mathbb{I}_{\{X_i \leq x\}} \quad \text{and} \quad \widehat{G}(y) = \frac{1}{n_1} \sum_{i \in \mathcal{I}_1} \mathbb{I}_{\{Y_i \leq y\}}.$$

It leads to the family $\{\widetilde{T}_j, j \in J\}$ of estimators of θ

$$\widetilde{T}_j \;=\; \sum_k \widetilde{\theta_{j,k}},$$

with

$$\widetilde{\theta_{j,k}} \;=\; \frac{1}{n_2(n_2-1)} \sum_{\substack{i_1,i_2 \in \mathcal{I}_2 \\ i_1 \neq i_2}} \phi_{j,k}\left(\frac{R_{i_1}}{n_1}, \frac{S_{i_1}}{n_1}\right) \phi_{j,k}\left(\frac{R_{i_2}}{n_1}, \frac{S_{i_2}}{n_1}\right),$$

where $R_p = n_1 \widehat{F}(X_p)$ and $S_p = n_1 \widehat{G}(Y_p)$, $p \in \mathcal{I}_1$, could be viewed as estimates of the rank statistics of X_p and Y_p respectively.

3.2 Test Procedures

In this part, we consider a family of known bivariate copula densities $\mathcal{C}_\Lambda = \{c_\lambda, \lambda \in \Lambda\}$ indexed by a parameter λ varying in a given set $\Lambda \subset \mathbb{R}^{d_\Lambda}$, $d_\Lambda \in \mathbb{N}^*$. From the observations $(X_i, Y_i)_{i\in\mathcal{I}}$, our aim is to test the goodness-of-fit between any c_λ and a copula density c, which is *enough distant* in the L_2-norm, from the parametric family $\mathcal{C}_\Lambda$. Acting as in paragraph 3.1, we estimate the square L_2-norm between c and a fixed element c_λ lying in the family $\mathcal{C}_\Lambda$ by

$$\widetilde{T}_j(\lambda) \quad = \quad \sum_k \widetilde{\theta_{j,k}}(\lambda), \tag{5}$$

where

$$\widetilde{\theta_{j,k}}(\lambda) \quad = \quad \frac{1}{n_2(n_2-1)} \sum_{\substack{i_1,i_2\in\mathcal{I}_2\\ i_1\neq i_2}} \left(\phi_{j,k}\left(\frac{R_{i_1}}{n_1}, \frac{S_{i_1}}{n_1}\right) - c_{j,k}(\lambda)\right) \times \left(\phi_{j,k}\left(\frac{R_{i_2}}{n_1}, \frac{S_{i_2}}{n_1}\right) - c_{j,k}(\lambda)\right),$$

where $\{c_{j,k}(\lambda), k \in \mathbb{Z}^2, j \in \mathbb{N}\}$ denote the known scaling coefficients of the target copula density c_λ. Notice that, if the direct observations $(F(X_i), G(Y_i))_{i\in\mathcal{I}}$ would be available, the appropriate test statistic $\widehat{T}_j(\lambda)$ would be

$$\widehat{T}_j(\lambda) \quad = \quad \sum_k \widehat{\theta}_{j,k}(\lambda),$$

where

$$\widehat{\theta_{j,k}}(\lambda) \quad = \quad \frac{1}{n_2(n_2-1)} \sum_{\substack{i_1,i_2\in\mathcal{I}_2\\ i_1\neq i_2}} \left(\phi_{j,k}\left(F(X_{i_1}), G(Y_{i_1})\right) - c_{j,k}(\lambda)\right) \times \left(\phi_{j,k}\left(F(X_{i_2}), G(Y_{i_2})\right) - c_{j,k}(\lambda)\right).$$

Now we are ready to build the test procedures. Let us give a set of indices J and a set of critical values $\{t_j,\ j \in J\}$ and define $\{D_j^\Lambda, j \in J\}$, the family of test statistics

$$D_j^\Lambda \quad = \quad \mathbb{I}_{\inf_{\lambda\in\Lambda} \widetilde{T}_j(\lambda) > t_j} \quad ,$$

allowing to test if c belongs to the parametric family $\mathcal{C}_\Lambda = \{c_\lambda, \lambda \in \Lambda\}$. Note that $\Lambda = \{\lambda_0\}$ leads to the single null hypothesis $H_0 : c = c_{\lambda_0}$. We are also interested in building auto-driven procedures by considering all the tests in the family

$$D_\Lambda \quad = \quad \max_{j\in J} D_j^\Lambda = \mathbb{I}_{\max_{j\in J}(\inf_{\lambda\in\Lambda} \widetilde{T}_j(\lambda) - t_j) > 0} \,. \tag{6}$$

The sequence of parameters t_j of the method are determined in an optimal way in Section 5. We explain in Section 4 what "optimal way" means in giving a presentation of the minimax theory for our framework.

4 Minimax Theory

We adopt the minimax point of view to solve the problem of hypothesis testing, initiated by Ingster (1982) in Gaussian white noise. A review of results obtained in problems of minimax hypothesis testing is available in Ingster (1993) and Ingster and Suslina (2002). Let us describe this approach.

4.1 Minimax hypothesis testing Problem

As in the previous section, we consider $\mathcal{C}_\Lambda = \{c_\lambda, \lambda \in \Lambda\}$ a given functional class of copula densities. For any given $\tau = (s, p, M)$, with $s > 0, p \geq 1, M > 0$, the following statistical problem of hypothesis testing is considered,

$$H_0 : c = c_\lambda \in \mathcal{C}_\Lambda \qquad \text{against} \qquad H_1 : c \in \Gamma(v_n(\tau)), \tag{7}$$

with

$$\Gamma(v_n(\tau)) = b_{s,p,\infty}(M) \cap \left\{ c : \inf_{c_\lambda \in \mathcal{C}_\Lambda} \|c - c_\lambda\| \geq v_n(\tau) \right\},$$

where $b_{s,p,\infty}(M)$ is the ball of radius M of the Besov Body $b_{s,p,\infty}$ defined in Definition 1 and $v_n(\tau)$ is a sequence of positive numbers, depending on τ and decreasing to zero as n goes to infinity. Recall that $\|g\|$ denotes the L_2-norm of any function g in $L_2(\mathbb{R}^2)$. Observe that the functional class $\Gamma(v_n(\tau))$, which determines the alternative H_1, is characterized by three parameters: the regularity class $b_{s,p,\infty}$ where the copula density is supposed to belong, the L_2-norm which is the geometrical tool measuring the distance between both hypotheses, and the sequence $v_n(\tau)$.

According to the principle of the minimaxity, the regularity space and the loss function are chosen by the statistician. Notice that the parameter τ could be known or unknown. Obviously, our aim is to consider tests which are able to detect alternatives defined with sequences $v_n(\tau)$ as small as possible. It can be shown (Ingster (1993)) that $v_n(\tau)$ cannot be chosen in an arbitrary way: indeed, if $v_n(\tau)$ is too small, then H_0 and H_1 cannot be distinguished with a given error $\alpha \in (0, 1)$. Therefore, solving hypothesis testing problems via the minimax approach consists in determining the smallest sequence $v_n(\tau)$ for which such a test is still possible and to indicate the corresponding test functions. The smallest sequence $v_n(\tau)$ is called the minimax rate of testing. Let us denote D_n a *test statistic* that is an arbitrary function with possible values $0, 1$, measurable with respect to $(X_i, Y_i)_{i \in \mathcal{I}}$ and such that we accept H_0 if $D_n = 0$ and we reject it if $D_n = 1$.

Definition 2. *Assuming τ to be known, the sequence $v_n(\tau)$ is the minimax rate of testing H_0 versus H_1 if the relations (8) and (9) are fulfilled:*

- *for any given $\alpha_1 \in (0,1)$, there exists $a > 0$ such that*

$$\lim_{n\to+\infty} \inf_{D_n} \left(\sup_{c_\lambda \in \mathcal{C}_\Lambda} I\!P_\lambda(D_n = 1) + \sup_{c \in \Gamma(a\, v_n(\tau))} I\!P_c(D_n = 0) \right) \geq \alpha_1 \quad (8)$$

  where the infimum is taken over any test statistic D_n,

- *there exists a sequence of test statistics $(D_n^\star)_n$ for which for any given α_2 in $(0,1)$, it exists $A > 0$ such that*

$$\lim_{n\to+\infty} \left(\sup_{c_\lambda \in \mathcal{C}_\Lambda} I\!P_\lambda(D_n^\star = 1) + \sup_{c \in \Gamma(A\, v_n(\tau))} I\!P_c(D_n^\star = 0) \right) \leq \alpha_2, \quad (9)$$

where $I\!P_c$, respectively $I\!P_\lambda$ denotes the distribution function associated with the copula density c, respectively c_λ.

4.2 Adaptation

Nevertheless, since the copula function itself is unknown, the a priori knowledge on τ could appear unrealistic. Therefore, the purpose of this paper is to solve the previous problem of test in an adaptive framework i.e. in supposing that $\tau = (s, p, M)$ is unknown but varying in a known set $\mathcal{S}$. Comparing the adaptive case with the non-adaptive case, it has been proved in different frameworks that a loss of efficiency in the rate of testing is unavoidable (see for instance Spokoiny (1996), Gayraud and Pouet (2005)). This loss is expressed as t_n, a positive constant or a sequence of positive numbers increasing to infinity with n (as slow as possible), which appears in $v_{nt_n^{-1}}(\tau)$, the rate of testing. Similarly to the minimax rate of testing, we define the adaptive minimax rate of testing as follows

Definition 3. *The sequence $v_{nt_n^{-1}}(\tau)$ is the adaptive minimax rate of testing if the relations (10) and (11) are satisfied*

- *for any given $\alpha_1 \in (0,1)$, there exists $a > 0$ such that*

$$\lim_{n\to+\infty} \inf_{D_n} \left(\sup_{c_\lambda \in \mathcal{C}_\Lambda} I\!P_\lambda(D_n = 1) + \sup_{\tau \in \mathcal{S}} \sup_{c \in \Gamma(a\, v_{nt_n^{-1}}(\tau))} I\!P_c(D_n = 0) \right) \geq \alpha_1, \quad (10)$$

  where the infimum is taken over any test statistic D_n,

- *there exists a sequence of universal test statistics $D_n^\star$ (free of τ) such that, for any given α_2 in $(0,1)$, there exists $A > 0$ such that*

$$\lim_{n\to+\infty} \left(\sup_{c_\lambda \in \mathcal{C}_\Lambda} I\!P_\lambda(D_n^\star = 1) + \sup_{\tau \in \mathcal{S}} \sup_{c \in \Gamma(A\, v_{nt_n^{-1}}(\tau))} I\!P_c(D_n^\star = 0) \right) \leq \alpha_2 \quad (11)$$

where t_n is either a positive constant or a sequence of positive numbers increasing to infinity with n as slow as possible.

Note that relations (10) and (11) (instead of relations (8) and (9)) mean that the minimax rate of testing $v_n(\tau)$ is contaminated by the term t_n in the adaptive setting. Observe that the same phenomenon is observed in the estimation problem where an extra logarithm term $t_n = \log(n)$ has often (but not always) to be paid for the adaptation.

5 Main results

In this section, we focus on test problems for which the parametric family $\mathcal{C}_\Lambda$ is included in some $b_{s_\Lambda,p_\Lambda,\infty}(M_\Lambda)$ where $s_\Lambda > 0$, $p_\Lambda \geq 1$ and $M_\Lambda > 0$ are known.

Our theoretical results concern the minimax resolution of the problem of hypothesis testing defined in (7) in an adaptive framework. Theorem 2 states the result of the lower bound (see Relation (10)). Then Theorem 3 exhibits the rate achieved by the test procedure proposed in Section 3 (see Relation (11)). Comparing the rate of our procedure with the fastest rate given in Theorem 2 leads to the following result:

Theorem 1. *Let us set*

$$\mathcal{S} = \{\tau = (s,p,M), s \geq 1/2, p \geq 2, M > 0 : \ s - 2/p \leq s_\Lambda - 2/p_\Lambda, M_\Lambda \leq M\}. \quad (12)$$

Under the assumptions of Theorem 2 and Theorem 3, our test procedure defined by Relation (6) is adaptive optimal over the range of parameters $\tau \in \mathcal{S}$.

5.1 Lower Bound

Let us first state our assumptions.

- **AInf1:** there exists a parameter λ_0 in Λ such that

$$\forall (u,v) \in [0,1]^2, \quad c_{\lambda_0}(u,v) > m, \quad \text{for } m > 0.$$

- **AInf2:** $\text{card}(\Lambda) = o\left(\exp\left(n^{1/(4s_{max}+2)}\right)\right)$, where $s_{max} \geq 1/2$ appears in Theorem 2.

As it is usual for composite null hypotheses, the result of the lower bound requires the existence of a particular density $c_{\lambda_0} \in \mathcal{C}_\Lambda$ (Assumption **AInf1**) in order to construct the randomized class of functions which must be included in the alternatives. Moreover, a control of the complexity of $\mathcal{C}_\Lambda$ is needed (Assumption **AInf2**).

Theorem 2. *Suppose that $\mathcal{S}$ defined in (12) is nontrivial (see Spokoiny (1996)), which means that there exist $p \geq 2$, $M > 0$ and $0 < s_{min} < s_{max}$ such that*

$$\forall s \in [s_{min}, s_{max}], \quad (s, p, M) \in \mathcal{S}$$

and assume that **AInf1** *and* **AInf2** *hold. Set*

$$v_{nt_n^{-1}}(\tau) = (nt_n^{-1})^{-2s/(4s+2)} \text{ with } t_n = \sqrt{\log(\log(n))}.$$

Then, it exists a positive constant a such that

$$\lim_{n\to+\infty} \left(\inf_{D_n} \{ \sup_{\lambda\in\Lambda} I\!P_\lambda(D_n = 1) + \sup_{\tau\in\mathcal{S}} \sup_{c\in\Gamma(a\, v_{nt_n^{-1}}(\tau))} I\!P_c(D_n = 0) \} \right) = \tag{13}$$

where the infimum is taken over any test function D_n.

5.2 Upper Bound

Theorem 3 deals with Relation (11) which holds for the test statistic D_Λ defined by Relation (6) as soon as the parameters of the methods are chosen as follows. The set $J = \{\lfloor j_0 \rfloor, \ldots, \lfloor j_\infty \rfloor\}$ is determined by

$$2^{j_0} = \log(n_2)\log(n_1), \quad 2^{j_\infty} = \left(\frac{n_2}{\log(n_2)}\right)^{1/2} \wedge \left(\frac{n_1}{\log(n_1)}\right)^{1/2-1/2q} \tag{14}$$

where q is the order of differentiability of the scaling function ϕ. The critical values satisfy

$$\forall j \in J, \quad t_j = 3\mu\, \frac{2^j}{n_2} \sqrt{\log\log(n_2)}, \tag{15}$$

where μ is a positive constant such that $\mu > \sqrt{2K_g K_1}$, and K_g and K_1 are positive constants depending on $\|\phi\|_\infty$, $\|c\|_\infty$, $\|c_\lambda\|_\infty$ and the length of support of ϕ (see Lemma 3).

- **ASup:**

$$\text{card}(\Lambda) = o\left(\exp(\, n_2 \log(n_2) \log\log(n_2))^{1/4}\right).$$

Theorem 3. *Let us choose $n_1 = \pi\, n$ and $n_2 = (1-\pi)n$ for some π in $(0, 1)$. Assume that the scaling function ϕ is continuously $q-$differentiable for*

$$q \geq \left[1 - \frac{\log\left(\frac{n_2}{\log(n_2)}\right)}{\log\left(\frac{n_1}{\log(n_1)}\right)}\right]^{-1}.$$

Moreover assume that any density c under the alternatives or any c_λ under the null are uniformly bounded. Then, the test statistic D_Λ defined by (6) is such that

$$\lim_{n_1 \wedge n_2 \to +\infty} \sup_{c_\lambda \in \mathcal{C}_\Lambda} I\!P_\lambda(D_\Lambda = 1) = 0. \tag{16}$$

Assume that **ASup** *holds, then there exists a positive constant A such that*

$$\lim_{n_1 \wedge n_2 \to +\infty} \sup_{\tau \in \mathcal{S}} \sup_{c \in \Gamma(Av_{nt_n^{-1}}(\tau))} I\!P_c(D_\Lambda = 0) \quad = \quad 0, \tag{17}$$

where

$$v_{nt_n^{-1}}(\tau) = (n_2 t_{n_2}^{-1})^{-2s/(4s+2)} \text{ and } t_{n_2} = \sqrt{\log(\log(n_2))}.$$

Relation (11) of the upper bound holds since both relations (16) and (17) are satisfied. Note also that Relation (16) indicates that the test statistic D_Λ is asymptotically of any level in $(0, 1)$.

6 Practical results

The purpose of this section is to provide several examples to investigate the performances of the test procedure presented in Section 3. This part is not an illustration of the theoretical part. Instead of focussing on the separating rate between the alternative and the null hypothesis, our aim is to study the test procedure from a risk point of view. In the first part, we fix the test level $\alpha = 5\%$ and we study the empirical power function. In the second part, we present an application to some economical series.

6.1 Methodology

Contrary to the estimation problem, a smooth wavelet is not needed. The test statistic is then computed with the Haar wavelet since it has a small support and then it leads to a fast computation time. The critical value of the test is determined with bootstrap methods: the standard deviation of the test statistic is computed thanks to $N_{boot} = 20$ resampling. The size of the simulated samples is $n = 1024$ which is very reasonable for bi-dimensional problems in an asymptotic context. For the real life data example, the number of data is around $n = 4000$. We do not split the sample of data as it is indicated in the theoretical part. For the simulation part, the empirical level of the test is derived from $N_{MC} = 50$ replications for each test problem.

6.2 Simulations

We consider the usual parametrical families mentioned in Introduction. We focus on tests with a single density copula c_{λ_0} under the null hypothesis.

6.2.1 Goodness-of-fit when the data are issue from the parametric family

We first sample data drawn from densities c_λ lying in the same parametric family at which c_{λ_0} belongs. Table 2 and Table 3 give the empirical power function for two different dependence structures characterized by a different Kendall's tau value. Some comments on the tables are in order below.

Our procedure is empirically very conservative as it is the case in the theoretical part (see Relation (16)): $\hat{\alpha}$ is equal to zero provided that λ is sufficiently closed to λ_0. On a practical point of view, the estimation of the variance of the test statistic is not made accurately since a small N_{boot} is used in order to produce a tractable procedure. This implies that the critical value is not very sharp. When λ_0 and λ are enough closed and from a H_0-point of view the good issue is that H_1 is never chosen if H_0 is true whereas on a H_1-point of view, it is difficult to distinguish H_0 from H_1. It implies that our procedure doesn't guarantee to distinguish favorably c and c_{λ_0} even if the Euclidean distance between λ and λ_0 seems large enough. It is mainly due to the nonparametric side of our test statistic, which is applied for the simulations to a problem of test whose setting is obviously parametric. Moreover, it is secondly due to our grid choice of λ: we don't know the impact of the distance between two consecutive λ in the grid on the L_2-norm of the two corresponding c_λ. And obviously this relation depends on the parametric class of copulas. In particular the grid for the Frank family is not well-chosen. To reduce scale problems, the Kendall's tau seems to be a better indicator of the distance between the copula densities c_λ and c_{λ_0} since they produce confidence intervals of accepted fits more accurate.

The power is improved with a big dependence structure i.e. with a large Kendall's tau in absolute value. For example, observe that the empirical power when $\tau_0 = \tau \pm 0.1$ is at least 20%. It is interesting to note that the empirical power function is not a symmetric function of τ_0 with respect to τ: it is better when $\tau < \tau_0$. A possible explanation is that theoretical properties of our procedure are good as soon as the unknown copula density is sufficiently smooth. The smaller is τ, the smoother is the density.

The case of data issued from a Clayton with large parameter (e.g. $\lambda = 6$ which corresponds to $\tau = 0.75$) does not appear in the tables. Actually, our procedure fails: we never accept H_0 even if $c_{\lambda_0} = c_\lambda$. We think that this kind of copula density is not concerned with the theoretical part since it is very sparse (almost all the wavelet coefficients are zeros whereas a few ones are large). In other words, it does not belong to a dense Besov space $B_{s,p,\infty}$ with $p \geq 2$. Moreover, this copula density is not bounded. We emphasize that our procedure can not be apply when the data are issued from a sparse copula density but there is no difficulty to test a sparse or unbounded fit c_{λ_0}.

6.2.2 Goodness-of-fit for copula densities with similar Kendall's tau

Table 4 gives the empirical power level when the sample data is drawn from densities c_λ which do not belong the same family as c_{λ_0} does, but they have a similar Kendall's tau. Obviously, it is a more difficult case for the test procedure: the power is improved with a Kendall's tau associated with c_{λ_0} which is different from the Kendall's tau associated with the density sample. In order to compare our procedure with existing procedures, we choose the test based on rank-based versions of the familiar Cramér-von Mises statistics (see Genest et al. (2006)). In Genest et al. (2008b), an intensive empirical study is presented for a very different experimental design. Since we are concerning with asymptotic methods and minimax theory, we need enough data: here $n = 1024$ while in Genest et al. (2008b) $n = 150$. The aim of the latter paper is to study the accuracy of the goodness-of-fit tests with respect to a prescribed level. Therefore, they need to have a very good estimation of the variance of the test statistic. Then, they take $N_{boot} = 1000$ (while here $N_{boot} = 20$) in view to have a sharp critical value. In the same idea, they choose $N_{MC} = 10000$ to make the empirical probability very accurate. For our simulation, we take only $N_{boot} = 50$. Notice that the results by Genest et al. (2008b) required the nearly exclusive use of 140 CPUs over a one-month period while our aim is to provide fast and simple tests. Obviously, the results from Genest et al. (2008b) should be better than ours. Nevertheless, we give into the brackets in Table 4 their results. Several points are apparent from Table 4.

As previously (Table 2 and Table 3), our test is always degenerated: we always accept H_0 while the procedure of Genest et al. Genest et al. (2008b) produces an excellent estimation of the prescribed level α.

As it is mentioned in the previous part, our procedure fails when the data are issued from a Clayton copula density with a large Kendall's tau. But when $\tau = 0.25$, the procedure is excellent: as usual, the test admits $\hat{\alpha} = 0$ for empirical estimation of the first kind error. Moreover, when c_{λ_0} is Gumbel or Student, the empirical power is better than whose obtained by Genest et al. (for instance $\hat{\alpha} = 0.86$ instead of $\hat{\alpha} = 0.27$).

For small level of dependence ($\tau = 0.25$) our procedure is very competitive. For instance, when the fit c_{λ_0} is a Student(4), our procedure always produce a better power than the power given by the CvM test.

6.3 Real data

We present now an application to real data of our test procedure. The level of each test (with simple null hypothesis or multivariate null hypothesis) is $\alpha = 5\%$. To obtain the empirical level, $N = 50$ replications of our procedure computed with the half of the available data (chosen randomly) is used.

Table 1 gives the empirical probability to reject the null hypothesis and the final decision. "Yes" means that we accept that the structure of dependence belongs to the considered family and "No" that we reject the fitting.

We consider the data of Frees and Valdez (1998), which were also analyzed by Genest et al. (1998), Klugman and Parsa (1999), Chen and Fan (2005) and Genest et al. (2006), among others. The data consist of the indemnity payment (LOSS) and the allocated loss adjustment expense (ALAE) for 1466 general liability claims. The various authors who analyzed this data set concluded that the Gumbel copula provides an adequate representation of the underlying dependence structure. The Gumbel parametric family of extreme-value copulas captures the fact that almost all large indemnity payments generate important adjustment expenses (e.g., investigation and legal costs) while the effort invested in the treatment of a small claim is more variable. Accordingly, the copula exhibits positive but asymmetric dependence. Confirming this result, the adaptive method of estimation proposed by Autin et al. (2008) provides a benchmark (see Figure 3) for the copula density associated to the data.

We consider the following test problems:

$$H_0 : c \in \mathcal{C}_\Lambda$$

where the parametrical family $\mathcal{C}_\Lambda$ is described in Table 1. Since the Kendall's tau computed with the sample is $\tau = 0.31$, we choose an adapted grid of parameters for each parametrical family of copula densities. Next, assuming that the density copula of the data belongs to a fixed parametric family, we estimate the parameter λ

- with $\hat{\lambda}$ in inverting the Kendall's tau (third part of Table 1 where $H_0 : c = c_{\hat{\lambda}}$).
- with $\tilde{\lambda}$ in minimizing the average square error (ASE) computed thanks to the benchmark given in Figure 3 (fourth part of Table 1 where $H_0 : c = c_{\tilde{\lambda}}$). For information, we give the relative ASE computed with $c_{\tilde{\lambda}}$ into brackets.

7 Discussion

The paper is mainly devoted to construct an optimal procedure for solving a general nonparametric problem of test: both hypotheses are composite, very general parametric family could be considered under the null. Our procedure is proved asymptotically to be adaptive minimax and the minimax separating rate is exhibited over a range of Besov balls. Thanks to simulations and an application to real data, our procedure seems to be competitive on the power point of view even if the setting of test under consideration is clearly parametric. The copula model requires much more regularity (than the usual density model) since the approximation due to the rank-based statistics needs to be accurate enough.

Another very interesting point is that we focus in this paper on copulas densities belonging to **dense Besov spaces** (i.e. defined with parameters p larger than 2). Nevertheless, it seems that several copula densities among those exhibiting strong dependence structure belongs to **sparse Besov spaces** (i.e. defined with parameters p smaller than 2). Indeed, in the simulation part, we emphasize that our procedure fails when the Clayton copula with large parameters is concerned. In the white noise model and for testing the existence of the signal, Lepskii and Spokoiny (1999) prove that the minimax separating rate is not v_n and that it is possible to built adaptive minimax (non linear) procedures. The set of the copulas densities contains great examples of sparse functions and we will explore new procedures in a further work.

A very close problem is the sample comparison test (problem with two samples). It could be interesting to test if the structure of dependence between a couple of variables $V_1 = (X, Y)$ is the same as the structure of an other couple $V_2 = (Z, T)$. Here the problem of test becomes the following

$$H_0 : c_{V_1} = c_{V_2} \qquad \text{against} \qquad H_1 : (c_{V_1}, c_{V_2}) \in \Gamma(v_n(\tau)),$$

with

$$\Gamma(v_n) = \{c_{V_1} \in b_{s_1,p_1,\infty}(M_1)\} \cap \{c_{V_2} \in b_{s_2,p_2,\infty}(M_2)\} \\ \cap \{(c_{V_1}, c_{V_2}) : \|c_{V_1} - c_{V_2}\| \geq v_n.\}$$

where v_n is the separating rate of both hypotheses. In an analogous way with the test statistic used in this paper, the rule for the comparison test is

$$D = \mathbb{1}_{\max_{j\in J}(\sum_k \widetilde{\theta_{j,k}} - t_j) > 0}$$

where

$$\widetilde{\theta_{j,k}} = \frac{1}{n_2(n_2-1)} \sum_{\substack{i_1,i_2\in\mathcal{I}_2\\ i_1\neq i_2}} \left(\phi_{j,k}\left(\frac{R^X_{i_1}}{n_1}, \frac{R^Y_{i_1}}{n_1}\right) - \phi_{j,k}\left(\frac{R^Z_{i_1}}{n_1}, \frac{R^T_{i_1}}{n_1}\right)\right) \\ \times \left(\phi_{j,k}\left(\frac{R^X_{i_2}}{n_1}, \frac{R^Y_{i_2}}{n_1}\right) - \phi_{j,k}\left(\frac{R^Z_{i_2}}{n_1}, \frac{R^T_{i_2}}{n_1}\right)\right).$$

R^X, R^Y, R^Z, R^T are the rank statistics associated to X, Y, Z, T. Using the same tools as in Butucea and Tribouley (2006) where the homogeneity in law of the both samples is studied, it is possible to prove that this test is adaptive optimal and that the minimax separating rate is

$$v_n = \left(\frac{n}{\sqrt{\log(\log(n_2))}}\right)^{-2(s_1\wedge s_2)/(4(s_1\wedge s_2)+2)}.$$

Obviously, all these tests procedures could be used in the multivariate framework but as usual in the nonparametric context, the rates of testing become rapidly slow implying poor powers.

8 Proof of Theorem 3

Racall that for any given $\lambda \in \Lambda$, $\mathbb{P}_\lambda$ (respectively $\mathbb{P}_c$) denote the distribution associated with density c_λ, respectively c. In the same spirit, denote also $\mathbb{E}_\lambda$ and Var_λ (respectively $\mathbb{E}_c$ and Var_c) the expectation and the variance with respect to $\mathbb{P}_\lambda$, respectively to $\mathbb{P}_c$. When no index appears in $\mathbb{E}$ or in $\mathbb{P}$ it means that the underlying distribution is either $\mathbb{P}_c$ or $\mathbb{P}_\lambda$.

8.1 Expansion of the statistics of interest

Fix a level j in J. For the test problem, the statistic of interest $\widetilde{T}_j(\lambda)$ (for $\lambda \in \Lambda$) is defined in (5) and is an estimator of

$$T_j(\lambda) \;=\; \sum_k \theta_{j,k}(\lambda) = \sum_k \left(c_{j,k} - c_{j,k}(\lambda)\right)^2,$$

which is the quantity that we need to detect under the alternative. It would be useful to expand the statistic $\widetilde{T}_j(\lambda)$ as follows

$$\begin{aligned}\widetilde{T}_j(\lambda) &= 2T_j^{\diamondsuit}(\lambda) + T_j^{\heartsuit} + T_j^{\spadesuit} + 2T_j^{\clubsuit}(\lambda) + T_j(\lambda) \\ &= 2\sum_k \theta^{\diamondsuit}_{j,k}(\lambda) + \sum_k \theta^{\heartsuit}_{j,k} + \sum_k \theta^{\spadesuit}_{j,k} + 2\sum_k \theta^{\clubsuit}_{j,k}(\lambda) + \sum_k \theta_{j,k}(\lambda)\end{aligned} \tag{18}$$

where

$$\begin{aligned}
\theta^{\heartsuit}_{j,k} &= \frac{1}{n_2(n_2-1)} \sum_{\substack{i_1,i_2\in\mathcal{I}_2\\ i_1\neq i_2}} (\phi_{j,k}(F(X_{i_1}),G(Y_{i_1})) - c_{j,k}) \\
&\quad \times (\phi_{j,k}(F(X_{i_2}),G(Y_{i_2})) - c_{j,k}) \\
\theta^{\spadesuit}_{j,k} &= \frac{1}{n_2(n_2-1)} \sum_{\substack{i_1,i_2\in\mathcal{I}_2\\ i_1\neq i_2}} \left(\phi_{j,k}\left(\frac{R_{i_1}}{n_1},\frac{S_{i_1}}{n_1}\right) - \phi_{j,k}(F(X_{i_1}),G(Y_{i_1}))\right) \\
&\quad \times \left(\phi_{j,k}\left(\frac{R_{i_2}}{n_1},\frac{S_{i_2}}{n_1}\right) - \phi_{j,k}(F(X_{i_2}),G(Y_{i_2}))\right) \\
\theta^{\clubsuit}_{j,k}(\lambda) &= \frac{1}{n_2(n_2-1)} \sum_{\substack{i_1,i_2\in\mathcal{I}_2\\ i_1\neq i_2}} \left(\phi_{j,k}\left(\frac{R_{i_1}}{n_1},\frac{S_{i_1}}{n_1}\right) - \phi_{j,k}(F(X_{i_1}),G(Y_{i_1}))\right) \\
&\quad \times (\phi_{j,k}(F(X_{i_2}),G(Y_{i_2})) - c_{j,k}(\lambda)) \\
\theta^{\diamondsuit}_{j,k}(\lambda) &= \frac{1}{n_2} \sum_{i_1\in\mathcal{I}_2} (\phi_{j,k}(F(X_{i_1}),G(Y_{i_1})) - c_{j,k})(c_{j,k} - c_{j,k}(\lambda)).
\end{aligned}$$

The sequence $\{c_{j,k}\}_{j,k}$ denotes the unknown scaling coefficients of the unknown copula density c. Recall that

$$\widehat{T}_j(\lambda) = \sum_k \widehat{\theta_{j,k}}(\lambda)$$

with

$$\begin{aligned}\widehat{\theta_{j,k}}(\lambda) &= \frac{1}{n_2(n_2-1)} \sum_{\substack{i_1,i_2 \in \mathcal{I}_2 \\ i_1 \neq i_2}} (\phi_{j,k}(F(X_{i_1}), G(Y_{i_1})) - c_{j,k}(\lambda)) \\ &\quad \times (\phi_{j,k}(F(X_{i_2}), G(Y_{i_2})) - c_{j,k}(\lambda)) \,.\end{aligned}$$

The following lemma gives some evaluation for the first moments of each statistic of interest.

Lemma 1. *Let q be a positive integer and assume that ϕ is continuously $q-$differentiable. Let j be a level smaller than j_∞ defined in (14). Then, it exists some positive constant κ which may depend on ϕ, $\|c\|_\infty$, $\|c_\lambda\|_\infty$ and M such that*

$$\begin{aligned} \mathbb{E}\widehat{T}_j(\lambda) &= T_j(\lambda) \quad \text{and} \quad \mathrm{Var}\widehat{T}_j(\lambda) \leq \kappa \left(\left(\frac{2^j}{n_2}\right)^2 + \left(\frac{2^j}{n_2}\right) T_j(\lambda) \right) \\ \mathbb{E}|T_j^{\spadesuit}| &\leq \kappa \frac{\log(n_1)}{n_1} \\ \mathbb{E}_c|T_j^{\clubsuit}(\lambda)| &\leq \kappa \left(\frac{\log(n_1)}{n_1} T_j(\lambda) \right)^{1/2} \quad \text{and} \quad \mathbb{E}_\lambda (T_j^{\clubsuit}(\lambda))^2 \leq \kappa \, 2^j \left(\frac{\log(n_1)}{n_2 n_1} \right). \end{aligned}$$

Using the Bernstein Inequality, we establish the following bound for the deviation of the statistic $T_j^{\diamondsuit}(\lambda)$ under the alternative. The proof is postponed to Appendix B.

Lemma 2. *For any level j, for all $x > 0$*

$$\mathbb{P}_c \left(|T_j^{\diamondsuit}(\lambda)| \geq x \right) \leq \exp\left(-K \left(\frac{n_2^2 x^2}{2^{2j} T_j(\lambda) + n_2 x 2^j T_j(\lambda)^{1/2}} \right) \right)$$

where K is a positive constant depending on L, $\|\phi\|_\infty$ and $\|c\|_\infty$.

Using a result from Giné et al. (2000), we establish the following bound for the deviation of the U-statistics $\widehat{T}_j(\lambda)$ and $T_j^{\heartsuit}$. The proof is postponed to Appendix C.

Lemma 3. *For any level j, as soon as $x \geq 2^j \, n_2^{-1} \sqrt{\log(\log(n_2))}$, for all $\mu > 0$,*

$$\mathbb{P}_\lambda \left(|\widehat{T}_j(\lambda)| > \mu x \right) + \mathbb{P}_c \left(|T_j^{\heartsuit}| > \mu x \right) \quad \leq \quad K_g (\log(n_2))^{-\delta}$$

for any positive $\delta \leq \mu^2(K_g K_1)^{-1}$, where K_g is an universal positive constant given in Giné et al. (2000) and K_1 is a positive constant depending on $L, \|\phi\|_\infty$ and either $\|c_\lambda\|_\infty$ or $\|c\|_\infty$ depending on the underlying distribution i.e. either $I\!P_\lambda$ or $I\!P_c$.

8.2 Proof of Relation (16) (First type error)

Let us fix $\lambda \in \Lambda$ and set

$$p_\lambda = I\!P_\lambda \left(\max_{j \in J} \left[\inf_{\lambda' \in \Lambda} \widetilde{T}_j(\lambda') - t_j \right] > 0 \right).$$

Notice that under the null

$$T_j^\diamondsuit(\lambda) = T_j(\lambda) = 0 \text{ and } T_j^\heartsuit = \widehat{T}_j(\lambda).$$

Using the expansion (18), we get

$$\begin{aligned} p_\lambda &\leq \sum_{j \in J} I\!P_\lambda \left(\inf_{\lambda' \in \Lambda} \widetilde{T}_j(\lambda') > t_j \right) \\ &\leq \sum_{j \in J} I\!P_\lambda \left(\widetilde{T}_j(\lambda) > t_j \right) \\ &\leq \sum_{j \in J} \left\{ I\!P_\lambda \left(|\widehat{T}_j(\lambda)| > \frac{t_j}{3} \right) + I\!P_\lambda \left(|T_j^\spadesuit| > \frac{t_j}{3} \right) + I\!P_\lambda \left(|T_j^\clubsuit(\lambda)| > \frac{t_j}{3} \right) \right\} \end{aligned}$$

Due to Lemma 1 and using Markov Inequality, we obtain

$$\begin{aligned} p_\lambda &\leq \sum_{j \in J} I\!P_\lambda \left(|\widehat{T}_j(\lambda)| > \frac{t_j}{3} \right) + \sum_{j \in J} \left\{ \frac{I\!E_\lambda |T_j^\spadesuit|}{(t_j/3)} + \frac{I\!E_\lambda (T_j^\clubsuit(\lambda))^2}{(t_j/3)^2} \right\} \\ &\leq \sum_{j \in J} I\!P_\lambda \left(|\widehat{T}_j(\lambda)| > \frac{t_j}{3} \right) \\ &\quad + K \sum_{j \in J} \left\{ (t_j/3)^{-1} \frac{\log(n_1)}{n_1} + (t_j/3)^{-2} \left(\frac{2^j \log(n_1)}{n_1 n_2} \right) \right\}. \end{aligned}$$

Note that $\widehat{T}_j(\lambda)$ is centered under $I\!P_\lambda$, then applying Lemma 3, where t_j is $t_j = 3\mu\, 2^j\, n_2^{-1} \sqrt{\log(\log(n_2))}$, the constant μ is defined in (15) and since $\text{card}(J) \leq \log(n_2)$, one obtains

$$\begin{aligned} p_\lambda &\leq K_g \text{card}(J)\, (\log(n_2))^{-\delta} + K\text{card}(J) 2^{-j_0} \left(\frac{n_2 \log(n_1)}{n_1 \sqrt{\log\log(n_2)}} \right) \\ &\quad + K\text{card}(J) 2^{-j_0} \left(\frac{\log(n_1) n_2^2}{n_2 n_1 \sqrt{\log\log(n_2)}} \right) \\ &\leq K_g (\log(n_2))^{1-\delta} + K 2^{-j_0} \left(\frac{\log(n_1)\log(n_2)}{\sqrt{\log\log(n_2)}} \right), \end{aligned}$$

where the last inequality holds since δ satisfies $\delta \leq \mu^2(2K_gK_1)^{-1}$ (see Lemma 3). Since μ is such that $\mu > \sqrt{2K_gK_1}$, relation (16) is proved if one takes $\delta = \mu^2(2K_gK_1)^{-1}$.

8.3 Proof of Relation (17) (Second type error)

Let us fix $\tau \in \mathcal{S}$ and $c \in \Gamma(Av_{nt_n^{-1}}(\tau))$ and set

$$p_c = I\!\!P_c\left(\max_{j\in J}\inf_{\lambda\in\Lambda}\widetilde{T}_j(\lambda) - t_j \leq 0\right).$$

Using the expansion (18), we get, for any $j^\star \in J$

$$\begin{aligned}
p_c &\leq I\!\!P_c\left(\inf_\lambda\left\{2T^\diamondsuit_{j^\star}(\lambda) + T_j(\lambda) + T^\heartsuit_{j^\star} + T^\spadesuit_{j^\star} + 2T^\clubsuit_{j^\star}(\lambda)\right\} \leq t_{j^\star}\right)\\
&\leq I\!\!P_c\left(\inf_\lambda\left\{2T^\diamondsuit_{j^\star}(\lambda) + T_j(\lambda)\right\} \leq 2t_{j^\star}\right)\\
&\quad + I\!\!P_c\left(T^\heartsuit_{j^\star} + T^\spadesuit_{j^\star} + 2\inf_\lambda\left\{T^\clubsuit_{j^\star}(\lambda)\right\} \geq t_{j^\star}\right)\\
&\leq I\!\!P_c\left(\inf_\lambda\left\{2T^\diamondsuit_{j^\star}(\lambda) + T_j(\lambda)\right\} \leq 2t_{j^\star}\right) + I\!\!P_c\left(T^\heartsuit_{j^\star} \geq t_{j^\star}/3\right)\\
&\quad + I\!\!P_c\left(T^\spadesuit_{j^\star} \geq t_{j^\star}/3\right) + I\!\!P_c\left(\inf_\lambda\left\{T^\clubsuit_{j^\star}(\lambda)\right\} \geq t_{j^\star}/6\right)\\
&= p_{c1}(j^\star) + p_{c2}(j^\star) + p_{c3}(j^\star) + p_{c4}(j^\star). \qquad (19)
\end{aligned}$$

Let us explain how $j^\star$ is chosen. From the wavelet expansion (1) and Lemma 1, one has

$$I\!\!E_c\widehat{T_{j^\star}}(\lambda) = T_{j^\star}(\lambda) = \int (c - c_\lambda)^2 - B_{j^\star}(\lambda),$$

where $T_{j^\star}$, $B_{j^\star}$ are defined in (2) and $t^\star_j$ is the critical value given in (15). Since c is in $\Gamma(Av_{nt_n^{-1}}(\tau))$ and c_λ lies in $b_{s_\Lambda,p_\Lambda,\infty}(M_\Lambda) \subset b_{s,p,\infty}(M)$, the function $(c - c_\lambda)$ is in $b_{s,p,\infty}(M)$. We can choose $j^\star$ such that

$$2^{j^\star} = \left(\frac{K}{3\mu}\frac{n_2}{\sqrt{\log\log(n_2)}}\right)^{1/(2s+1)},$$

which is possible due to our choice of j_∞ and because $s \geq 1/2$. It implies that $B_{j^\star} \leq t_{j^\star}$ since $B_{j^\star} \leq \tilde{K}2^{-2j^\star s}$ (see Inequality (3)). Next, since $c \in \Gamma(Av_{nt_n^{-1}}(\tau))$, one has $\int(c - c_{\lambda'})^2 \geq A^2(v_{nt_n^{-1}}(\tau))^2$ for all $\lambda' \in \Lambda$. Focusing on rates $v_{nt_n^{-1}}(s)$ combined with positive constant A which satisfy $4t_{j^\star} \leq (Av_{nt_n^{-1}}(s))^2$, one obtains

$$\frac{t_{j^\star}}{I\!\!E_c\widehat{T_{j^\star}}(\lambda)} = \frac{t_{j^\star}}{T_{j^\star}(\lambda)} \leq 1/3, \qquad (20)$$

Coming back to the evaluation of the probability terms. First, by (20) and applying Lemma 2, we get

$$\begin{aligned}
p_{c1}(j^\star) &\leq \sum_{\lambda\in\Lambda} I\!P_c\left(2T^{\diamondsuit}_{j^\star}(\lambda)+T_{j^\star}(\lambda)\leq 2t_{j^\star}\right)\\
&\leq \sum_{\lambda\in\Lambda} I\!P_c\left(T^{\diamondsuit}_{j^\star}(\lambda)\leq -T_{j^\star}(\lambda)/6\right)\\
&\leq \sum_{\lambda\in\Lambda}\exp\left[-K\left(\left(\frac{n_2}{2^{j^\star}}\right)^2 T_{j^\star}(\lambda)\wedge\frac{n_2}{2^{j^\star}}T_{j^\star}(\lambda)^{1/2}\right)\right]\\
&\leq \sum_{\lambda\in\Lambda}\exp\left[-K\left(\left(\frac{n_2}{2^{j^\star}}\right)^2 t_{j^\star}\wedge\frac{n_2}{2^{j^\star}}t_{j^\star}^{1/2}\right)\right]\\
&\leq \sum_{\lambda\in\Lambda}\exp\left[-K\left(n_2 2^{-j_\infty}\sqrt{\log\log(n_2)}\right)^{1/2}\right] \qquad (21)
\end{aligned}$$

which is tending to zero as soon as $\mathrm{card}(\Lambda)=o\left(\exp\left[n_2\log(n_2)\log\log(n_2)\right]^{1/4}\right)$ which is ensured by assumption **ASup**.

Now, it remains to verify that $p_{c2}(j^\star), p_{c3}(j^\star)$ and $p_{c4}(j^\star)$ are going to zero as $n_1\wedge n_2$ goes to infinity. Using again the bound (20), Lemma 3 for some positive δ, Lemma 1 and the definition of the critical value (15), one gets

$$\begin{aligned}
&p_{c2}(j^\star)+p_{c3}(j^\star)+p_{c4}(j^\star)\\
&\leq I\!P_c\left(T^{\heartsuit}_{j^\star}\geq t_{j^\star}/3\right)++I\!P_c\left(T^{\spadesuit}_{j^\star}\geq t_{j^\star}/3\right)+\mathrm{card}(\Lambda)\,I\!P_c\left(T^{\clubsuit}_{j^\star}(\lambda)\geq t_{j^\star}/6\right)\\
&\leq K_g(\log(n_2))^{-\delta}+3\,\frac{I\!E_c|T^{\clubsuit}_{j^\star}(\lambda)|^2}{t^2_{j^\star}}+\mathrm{card}(\Lambda)\,6\,\frac{I\!E_c|T^{\spadesuit}_{j^\star}|}{t_{j^\star}}\\
&\leq K_g(\log(n_2))^{-\delta}+3\kappa\frac{2^{j^\star}\log(n_1)}{n_1n_2\,t^2_{j^\star}}+\mathrm{card}(\Lambda)\,6\kappa\frac{\log(n_1)}{n_1\,t_{j^\star}}\\
&\leq K_g(\log(n_2))^{-\delta}+6\kappa\,\mathrm{card}(\Lambda)\left(2^{-j^\star}\frac{n_2}{n_1}\frac{\log(n_1)}{\sqrt{\log\log(n_2)}}\right), \qquad (22)
\end{aligned}$$

which tends to zero with our choice of $j^\star$ as soon as $\mathrm{card}(\Lambda)=o(n_2^{1/(2s+1)})$ and where κ is the positive constant appearing in Lemma 1. Inequalities (21) and (22) entail that the right hand side of (19) is less than any $\alpha\in(0,1)$ as n is large enough. To finish the proof, observe that the choice of $v_{nt_n^{-1}}(s)$ is driven by the fact that it corresponds to the smallest sequence such that $4t^\star_j\leq(Av_{nt_n^{-1}}(s))^2$, which leads to

$$v_{nt_n^{-1}}(s)\geq\left(2^{j^\star}\mu\frac{\sqrt{\log(\log n_2)}}{n_2}\right)^{1/2}\geq\left(\frac{n_2}{\sqrt{\log\log(n_2)}}\right)^{2s/(4s+2)}.$$

9 Proof of Theorem 2

Without loss of generality, we suppose that the support of the scaling function ϕ and its associated wavelet function ψ is $[0,1]$. Moreover assume that $\int_0^1 \psi^\epsilon = 0$. Let us give some $a > 0$ which must be small enough.

9.1 Discretisation of $\mathcal{S}$

For any given $\tau = (s,p,M) \in \mathcal{S}$, denote the level $j(\tau)$ by

$$2^{j(\tau)} = (nt_n^{-1})^{2/(4s+2)}$$

and define s_j the solution of the equation $j = j(s_j,p,M)$ for any resolution level $j \in \tilde{J} = \{j_{s_{\max}}, \dots, j_{s_{\min}}\} \subset \{j_0, \dots, j_\infty\}$ with

$$j_{s_{\max}} = \lfloor j(s_{\max}, p, M) \rfloor \text{ and } j_{s_{\min}} = \lfloor j(s_{\min}, p, M) \rfloor.$$

Consider now the set $\mathcal{S}_n = \{\tau_j = (s_j,p,M), j \in \tilde{J}\}$ which appears as a discretisation version of a subset of $\mathcal{S}$ whose cardinality is of order $O(\log(n))$.

9.2 Prior and parametric family included in the alternatives

For any $s_j \in \mathcal{S}_n$, define a prior π_j which is concentrated on the class of the random functions

$$c_j(u,v) = c_{\lambda_0}(u,v) + \sum_k \sum_{\epsilon=1}^{3} \delta_k u_j(n) \psi^\epsilon_{j,k}(u,v),$$

where c_{λ_0} is defined in assumption **AInf1** and

$$P(\delta_k = 1) = P(\delta_k = -1) = 1/2 \quad \text{and} \quad u_j(n) = C_1 M (nt_n^{-1})^{-\frac{2(s_j+1)}{4s_j+2}}$$

for C_1 such that $3M^2C_1^2 = 2a^2$. Let j be any index in $\tilde{J}$. Since $\int \psi = 0$ and when a is small enough (to guarantee that $c_j \geq 0$), c_j is a density. Easy calculations imply that

$$\|c_j - c_{\lambda_0}\|^2 = M^2 C_1^2 \, (v_{nt_n^{-1}}(\tau_j))^2 > a^2 \, (v_{nt_n^{-1}})^2.$$

Moreover, if a is small enough, we have $3\, C_1^p < 1$ and

$$\begin{aligned} 2^{j(s_j+1-2/p)p} \sum_k \sum_\epsilon \left| \int c_j \psi^\epsilon_{jk} \right|^p &= 2^{j(s_j+1-2/p)p} \sum_k \sum_\epsilon |u_j(n)|^p \\ &= 3 C_1^p M^p \leq M^p, \end{aligned}$$

implying that $c_j \in b_{s_j,p,\infty}(M)$. Denote $\mathcal{A}_{j,n}(a)$ the set of densities

$$\{c \in b_{s_j,p,\infty}(M) : \inf_\lambda \|c - c_\lambda\|^2 > a^2 (v_{nt_n^{-1}}(\tau_j))^2\}.$$

and consider the variation between both distributions $I\!P_{\lambda_0}$ and $I\!P_\Pi$

$$Var(I\!P_{\lambda_0}, I\!P_\Pi) = \frac{1}{2}\int \left|\frac{dI\!P_\Pi}{dI\!P_{\lambda_0}} - 1\right| dI\!P_{\lambda_0},$$

where

$$\frac{dI\!P_\Pi}{dI\!P_{\lambda_0}} = \frac{1}{N_n}\sum_{j\in\tilde{J}} \frac{dI\!P_j}{dI\!P_{\lambda_0}} = \frac{1}{N_n}\sum_{j\in\tilde{J}} I\!E^{(n)}_{\pi_j}[\frac{c_j}{c_{\lambda_0}}].$$

Assuming that the following assertion holds

$$\lim_{n\to\infty}\inf_{j\in\tilde{J}} \pi_j(c\in\mathcal{A}_{j,n}(a)) = 1, \tag{23}$$

we deduce that the left hand side (LHS) of relation (13) without the limit is bounded from below by

$$\begin{aligned} LHS &\geq I\!P_{\lambda_0}(D_n = 1) + \sup_{s_j\in\mathcal{S}_n}\sup_{c\in\mathcal{A}_{j,n}(a)} I\!P_c(D_n = 0) \\ &\geq 1 - Var(I\!P_{\lambda_0}, I\!P_\Pi)(1 + o_n(1)), \end{aligned}$$

as n large enough. Since the supports of the functions c_j and $c_{j'}$ are disjoint for $j \neq j'$, one has

$$\begin{aligned} 1 - Var(I\!P_{\lambda_0}, I\!P_\Pi) &\geq 1 - \frac{1}{2}\frac{1}{N_n^2}\sum_{j\in\tilde{J}} I\!E_{\lambda_0}\left[\left(\int\prod_{i=1}^n \frac{c_j(U_i,V_i)}{c_{\lambda_0}(U_i,V_i)} d\pi_j(c_j)\right)^2 - 1\right] \\ &\geq 1 - o_n(1) \end{aligned}$$

provided that

$$\lim_{n\to\infty}\frac{1}{N_n^2}\sum_{j\in\tilde{J}} I\!E_{\lambda_0}\left[\left(\int\prod_{i=1}^n \frac{c_j(U_i,V_i)}{c_{\lambda_0}(U_i,V_i)} d\pi_j(c_j)\right)^2\right] = 0. \tag{24}$$

Relation (13) is thus proved if (23) and (24) are satisfied. The remaining proofs are given in the sequel.

9.3 Proof of Relation (23)

Let Λ' be a subsect of Λ. We have

$$\begin{aligned} \pi_j\left(\inf_{\lambda\in\Lambda}\|c_j - c_\lambda\|^2 \leq a^2(v_{nt_n^{-1}}(\tau_j))^2\right) &\leq \pi_j\left(\inf_{\lambda\in\Lambda/\Lambda'}\|c_j - c_\lambda\|^2 \leq a^2(v_{nt_n^{-1}}(\tau_j))^2\right) \\ &+ \pi_j\left(\inf_{\lambda\in\Lambda'}\|c_j - c_\lambda\|^2 \leq a^2(v_{nt_n^{-1}}(\tau_j))^2\right) \end{aligned} \tag{25}$$

Consider the particular subset Λ' defined by

$$\Lambda' = \{\lambda \in \Lambda : \ \|c_{\lambda_0} - c_\lambda\|^2 \le 6C_1^2 M^2 (v_{nt_n^{-1}}(\tau_j))^2\}.$$

Note that

$$\lambda \in \Lambda/\Lambda' \Longrightarrow \|c_\lambda - c_j\|^2 \ge a^2 (v_{nt_n^{-1}}(\tau_j))^2$$

due to the choice of C_1. It implies that the first term in the right hand side of (25) is null and then, it remains to prove that

$$\lim_{n\to\infty} \pi_j \left(\inf_{\lambda \in \Lambda'} \|c_j - c_\lambda\|^2 \le a^2 (v_{nt_n^{-1}}(\tau_j))^2 \right) = 0. \tag{26}$$

Since λ in Λ', we get

$$\begin{aligned} \|c_\lambda - c_j\|_2^2 &= \|c_{\lambda_0} - c_\lambda\|^2 + \sum_k \sum_\epsilon u_j(n)^2 + 2 \sum_k \sum_\epsilon \delta_k u_j(n) B_{j,k,\lambda,\lambda_0} \\ &\ge 3C_1^2 M^2 \, (v_{nt_n^{-1}}(\tau_j))^2 + 2 \sum_k \delta_k u_j(n) \sum_\epsilon B_{j,k,\lambda,\lambda_0}, \end{aligned}$$

where

$$B_{j,k,\lambda,\lambda_0} = \int \psi_{j,k}^\epsilon (c_{\lambda_0} - c_\lambda).$$

Therefore assertion (26) is equivalent to

$$\lim_{n\to\infty} \pi_j \left(\inf_{\lambda \in \Lambda'} 2 \sum_k \delta_k u_j(n) B_{j,k,\lambda,\lambda_0} \le -a^2 (v_{nt_n^{-1}}(\tau_j))^2 \right) = 0.$$

or

$$\lim_{n\to\infty} \pi_j \left(\sup_{\lambda \in \Lambda'} 2 \sum_k (-\delta_k) u_j(n) B_{j,k,\lambda,\lambda_0} \ge a^2 (v_{nt_n^{-1}}(\tau_j))^2 \right) = 0.$$

Finally, relation (26) is proved due to assumption **AInf2** and applying Bernstein inequality to

$$\pi_j \left(2 \sum_k (-\delta_k) u_j(n) B_{j,k,\lambda,\lambda_0} \ge a^2 (v_{nt_n^{-1}}(\tau_j))^2 \right),$$

with i.i.d. centered random variables $Z_k = -\delta_k B_{j,k,\lambda,\lambda_0}$. In particular, note that $|Z_k| < K_1 v_{nt_n^{-1}}(\tau_j)$ and $\sum_k \mathrm{Var}(Z_k) \le K_2 (v_{nt_n^{-1}}(\tau_j))^2$, where K_1 and K_2 are positive constants.

9.4 Proof of Relation (24)

Set

$$l_{n,\pi} = \int \prod_{i=1}^{n} \frac{c_j(U_i, V_i)}{c_{\lambda_0}(U_i, V_i)} d\pi_j(c_j).$$

Due to the fact that the functions $\psi_{j,k}^{\epsilon}$ have disjoint support, it is possible to rewrite c_j as follows

$$c_j = c_{\lambda_0} \prod_k (1 + \delta_k D_{j,k})$$

for

$$D_{j,k} = u_j(n) \sum_{\epsilon} \frac{\psi_{j,k}^{\epsilon}}{c_{\lambda_0}}.$$

Then,

$$\begin{aligned} l_{n,\pi} &= \prod_k \int \prod_{i=1}^{n} (1 + \delta_k D_{j,k}(U_i, V_i)) d\pi_j(\delta_k) \\ &= \prod_k \frac{1}{2} \left\{ \prod_{i=1}^{n} (1 + D_{j,k}(U_i, V_i)) + \prod_{i=1}^{n} (1 - D_{j,k}(U_i, V_i)) \right\}, \end{aligned}$$

and

$$\begin{aligned} l_{n,\pi}^2 &= \prod_k \frac{1}{4} \left\{ 2 \prod_{i=1}^{n} \left[1 + D_{j,k}^2(U_i, V_i) \right] + 2 \prod_{i=1}^{n} \left[1 - D_{j,k}^2(U_i, V_i) \right] \right. \\ &\quad \left. + H \left(D_{j,k}(U_i, V_i), \left(D_{j,k}^{b_t}(U_t, V_t) \right)_{t \in \{1, \dots, i-1, i+1, \dots, n\}} \right) \right\}, \end{aligned}$$

where b_t is either 0 or 2. Due to the independence of the data and acting as in Pouet (2000), it can be shown that

$$\mathbb{E}_{\lambda_0} \left[H \left(D_{j,k}(U_i, V_i), \left(D_{j,k}^{b_t}(U_t, V_t) \right)_{t \in \{1, \dots, i-1, i+1, \dots, n\}} \right) \right] = 0.$$

Therefore,

$$\begin{aligned} \mathbb{E}_{\lambda_0}[l_{n,\pi}^2(U_i, V_i)] &\leq \prod_k \left\{ \left(1 + \mathbb{E}_{\lambda_0} D_{j,k}^2(U_i, V_i) \right)^n + \left(1 - \mathbb{E}_{\lambda_0} D_{j,k}^2(U_i, V_i) \right)^n \right\} \\ &\leq \prod_k \cosh \left(n \mathbb{E}_{\lambda_0} D_{j,k}^2(U_i, V_i) \right). \end{aligned}$$

Using the inequality $\log(\cosh(u)) \leq Ku^2$ where K is a fixed constant and since c_{λ_0} is bounded from below by m, one obtains

$$\begin{aligned}
\frac{1}{N_n^2}\sum_{j\in\tilde{J}}\exp(\log(\mathbb{E}_{\lambda_0} l_{n,\pi})^2) &\leq \frac{1}{N_n^2}\sum_{j\in\tilde{J}}\exp\left\{Kn^2\sum_k\left(\mathbb{E}_{\lambda_0} D^2_{j,k}(U_i,V_i)\right)^2\right\} \\
&\leq \frac{1}{N_n^2}\sum_{j\in\tilde{J}}\exp\left\{\frac{3^2K}{m^2}n^2 2^{2j}u_j(n)^4\right\} \\
&\leq \frac{\log(n)^\kappa}{\log(n)(1+o_n(1))},
\end{aligned}$$

where $\kappa = K(3C_1^2M^2)^2m^{-2} = 4Ka^4m^{-2}$. Choosing a small enough and $\kappa < 1$, Relation (24) is then proved.

10 Appendix A: Proof of Lemma 1

In this part, κ denotes any positive constant which may depend on ϕ, M and on $\|c\|, \|c_\lambda\|$.

10.1 Notations and Preliminaries

Let us define or recall some notations that will be used below. For any $k \in \mathbb{Z}^2$, set

$$\begin{aligned}
\xi_k(X_i,Y_i) &= \phi_{j,k}\left(\widehat{F}(X_i),\widehat{G}(Y_i)\right) - \phi_{j,k}\left(F(X_i),G(Y_i)\right) \\
\omega^\lambda_{j,k}(X_i,Y_i) &= \phi_{j,k}(F(X_i),G(Y_i)) - c_{j,k}(\lambda) \\
\omega^\infty_{j,k}(X_i,Y_i) &= \phi_{j,k}(F(X_i),G(Y_i)) - c_{j,k},
\end{aligned}$$

where i is in $\mathcal{I}_2$. First, the localization property of the scaling function implies that only few $\xi_k(X_i,Y_i)$ will be used since the others are zero. Indeed, one has the following result

Lemma 4. *For any $k \in \mathbb{Z}^2$, let us denote*

$$N_j = \text{card} \quad \{i \in \mathcal{I}_2; \xi_k(X_i,Y_i) \neq 0\}.$$

Let $\delta > 0$. For any level j such that

$$2^j \leq \frac{2}{3\sqrt{\delta+1}}\left(\frac{n_2}{\log(n_2)}\right)^{1/2},$$

one has

$$\mathbb{P}(N_j > 2(2L+3)n_2 2^{-j})) \leq K(n_1^{-\delta} + n_2^{-\delta}).$$

We refer to Genest et al. (2008a) for the proof of this lemma since a similar result is established with an estimate $\widehat{F}$ built on the whole sample: it guarantees in particular that $\widehat{F}(X_{(i:n)}) = i/n$, where $X_{(i:n)}$ denotes the $i-$th (among n) order statistic. In our case, the situation is different since $\widehat{F}(X_{(i:n_2)})$ is based on the observations lying in the subsample whose indices are in $\mathcal{I}_1$ whereas it is calculated in an observation lying in the subsample whose indices are in $\mathcal{I}_2$; nevertheless, applying the Dvoretsky–Kiefer–Wolfovitz Inequality, the following deviation inequality holds. For any $\epsilon > 0$, $\mathbb{P}_{\widehat{F}} = \mathbb{P}\left(\left|\widehat{F}(X_{(i:n_2)}) - \frac{i}{n_2}\right| \geq 2\epsilon\right)$ is bounded from above by

$$\begin{aligned}
\mathbb{P}_{\widehat{F}} &\leq \mathbb{P}\left(\left|\widehat{F}(X_{(i:n_2)}) - F(X_{(i:n_2)})\right| \geq \epsilon\right) + \mathbb{P}\left(\left|F(X_{(i:n_2)}) - \widetilde{F}(X_{(i:n_2)})\right| \geq \epsilon\right) \\
&\leq \mathbb{P}\left(\|\widehat{F} - F\|_\infty \geq \epsilon\right) + \mathbb{P}\left(\|\widetilde{F} - F\|_\infty \geq \epsilon\right) \\
&\leq K\left(n_1^{-\delta} + n_2^{-\delta}\right),
\end{aligned}$$

as soon as we take $\epsilon = \sqrt{\delta \log(n_1)/(2n_1)} \vee \sqrt{\delta \log(n_2)/(2n_2)}$. Here $\widehat{F}$ represents the empirical margin computed with the subsample whose indices in $\mathcal{I}_1$ and $\widetilde{F}$, the empirical margin computed with the subsample whose indices in $\mathcal{I}_2$.

10.1.1 Study of $\widehat{T}_j(\lambda)$

Rewrite $\widehat{\theta_{j,k}}(\lambda)$ in $\widehat{T}_j(\lambda) = \sum_k \widehat{\theta_{j,k}}(\lambda)$ as follows

$$\widehat{\theta_{j,k}}(\lambda) = \frac{1}{n_2(n_2-1)} \sum_{\substack{i_1,i_2 \in \mathcal{I}_2 \\ i_1 \neq i_2}} \omega_{j,k}^\lambda(X_{i_1}, Y_{i_1}) \omega_{j,k}^\lambda(X_{i_2}, Y_{i_2}).$$

For all $i \in \mathcal{I}_2$, one has $\mathbb{E}(\omega_{j,k}^\lambda(X_i, Y_i)) = c_{j,k} - c_{j,k}(\lambda)$, which implies that

$$\mathbb{E}(\widehat{T}_j(\lambda)) = \sum_k \theta_{j,k}(\lambda) = T_j(\lambda).$$

Moreover for $p \neq k$, one obtains

$$\begin{aligned}
&\mathbb{E}(\widehat{\theta_{j,k}}(\lambda)\widehat{\theta_{j,p}}(\lambda)) \\
&= \frac{1}{(n_2(n_2-1))^2} \sum_{i_1 \neq i_2 \neq i_3 \neq i_4} \mathbb{E}\left[\omega_{j,k}^\lambda(X_{i_1}, Y_{i_1})\right] \mathbb{E}\left[\omega_{j,p}^\lambda(X_{i_3}, Y_{i_3})\right] \mathbb{E}\left[\omega_{j,k}^\lambda(X_{i_2}, Y_{i_2})\right] \mathbb{E}\left[\omega_{j,p}^\lambda(X_{i_4}, Y_{i_4})\right] \\
&\quad + 4\frac{1}{(n_2(n_2-1))^2} \sum_{i_1 \neq i_2 \neq i_3} \mathbb{E}\left[\omega_{j,k}^\lambda(X_{i_1}, Y_{i_1})\right] \mathbb{E}\left[\omega_{j,p}^\lambda(X_{i_3}, Y_{i_3})\right] \mathbb{E}\left[\omega_{j,k}^\lambda(X_{i_2}, Y_{i_2}) \omega_{j,p}^\lambda(X_{i_2}, Y_{i_2})\right] \\
&\quad + 2\frac{1}{(n_2(n_2-1))^2} \sum_{i_1 \neq i_2} \mathbb{E}\left[\omega_{j,k}^\lambda(X_{i_1}, Y_{i_1}) \omega_{j,p}^\lambda(X_{i_1}, Y_{i_1})\right] \mathbb{E}\left[\omega_{j,k}^\lambda(X_{i_2}, Y_{i_2}) \omega_{j,p}^\lambda(X_{i_2}, Y_{i_2})\right] \\
&\leq \theta_{j,k}(\lambda)\theta_{j,p}(\lambda) + \frac{4}{n_2} (c_{j,k} - c_{j,k}(\lambda))(c_{j,p} - c_{j,p}(\lambda)) \left[\int \left(\phi_{j,k} - \int \phi_{j,k} c_\lambda\right) \left(\phi_{j,p} - \int \phi_{j,p} c_\lambda\right) c\right] \\
&\quad + \frac{2}{n_2(n_2-1)} \left[\int \left(\phi_{j,k} - \int \phi_{j,k} c_\lambda\right) \left(\phi_{j,p} - \int \phi_{j,p} c_\lambda\right) c\right]^2,
\end{aligned}$$

which implies that

$$\begin{aligned}
\mathrm{Var}(\widehat{T_j}(\lambda)) &= I\!\!E\left(\sum_k \widehat{\theta_{j,k}}(\lambda)\right)^2 - \left(I\!\!E\sum_k \widehat{\theta_{j,k}}(\lambda)\right)^2 \\
&\leq \frac{4}{n_2}\sum_{k,p}(c_{j,k}-c_{j,k}(\lambda))(c_{j,p}-c_{j,p}(\lambda))\left[\int\left(\phi_{j,k}-\int\phi_{j,k}c_\lambda\right)\left(\phi_{j,p}-\int\phi_{j,p}c_\lambda\right)c\right] \\
&\quad + \frac{2}{n_2(n_2-1)}\sum_{k,p}\left[\int\left(\phi_{j,k}-\int\phi_{j,k}c_\lambda\right)\left(\phi_{j,p}-\int\phi_{j,p}c_\lambda\right)c\right]^2.
\end{aligned}$$

Applying the Hölder inequality and the consequence of the Parseval Equality, we get

$$\begin{aligned}
&\sum_{kp}\left[\int\left(\phi_{j,k}-\int\phi_{j,k}c_\lambda\right)\left(\phi_{j,p}-\int\phi_{j,p}c_\lambda\right)c\right]^2 \\
&\leq 2^2\left(\sum_{k,p}\left[\int\phi_{j,k}\phi_{j,p}c\right]^2 + 2\sum_{k,p}\left[\int\phi_{j,k}c_\lambda\int\phi_{j,p}c\right]^2 + \left(\sum_k\left[\int\phi_{j,k}c_\lambda\right]^2\right)^2\right) \\
&\leq 2^2\left(\left(\sum_k\int\phi_{j,k}^2c\right)^2 + 2\int c^2\int c_\lambda^2 + \left(\int c_\lambda^2\right)^2\right) \leq \kappa\, 2^{2j}.
\end{aligned}$$

We conclude that

$$\begin{aligned}
\mathrm{Var}(\widehat{T}_j(\lambda)) &\leq \kappa\left(\frac{4}{n_2}\left(\sum_{k,p}\theta_{j,k}(\lambda)\theta_{j,p}(\lambda)\right)^{1/2} 2^j + \frac{2^{2j}}{n_2(n_2-1)}\right) \\
&\leq \kappa\left(\frac{2^j}{n_2}T_j(\lambda) + \frac{2^{2j}}{n_2^2}\right).
\end{aligned}$$

which is the announced result for $\widehat{T}_j(\lambda)$.

10.1.2 Study of $T_j^\spadesuit$ and $T_j^\clubsuit(\lambda)$

Let us denote

$$\begin{aligned}
A_{i_1} &= [\xi_k(X_{i_1},Y_{i_1})]\,, \quad D_{i_1} = \sum_{k,p}\left(I\!\!E[\xi_k(X_{i_1},Y_{i_1})\xi_p(X_{i_1},Y_{i_1})]\right)^2 \\
B_{i_1,i_2} &= \sum_k[\xi_k(X_{i_1},Y_{i_1})\xi_k(X_{i_2},Y_{i_2}]\,,\ C_{i_1,i_2} = \xi_k(X_{i_1},Y_{i_1})\xi_p(X_{i_2},Y_{i_2}).
\end{aligned}$$

We need the following results which are stated in the lemma below

Lemma 5. *Assume that the scaling function is q-differentiable. For any level $j \leq j_\infty$, there exists some positive constant κ depending on ϕ, its derivatives and on $\|c\|_\infty$ (which might be $\|c_\lambda\|_\infty$ for some $\lambda \in \Lambda$) such that for any distinct indices i_1, i_2, one obtains*

$$I\!\!E|A_{i_1}| \leq \kappa\left(\frac{\log(n_1)}{n_1}\right)^{1/2} \qquad (27)$$

$$\begin{aligned} \mathbb{E}|B_{i_1,i_2}| &\leq \kappa 2^{2j}\left(\frac{\log(n_1)}{n_1}\right), \quad \mathbb{E}|C_{i_1,i_2}| \leq \kappa\left(\frac{\log(n_1)}{n_1}\right) \\ |D_{i_1}| &\leq 2^{6j}\left(\frac{\log(n_1)}{n_1}\right)^2. \end{aligned} \tag{28}$$

We prove Relation (27) in the next section, Relations (28) are proven in Genest et al. (2008a). We have

$$\mathbb{E}T_j^{\spadesuit} = \frac{1}{n_2(n_2-1)} \sum_{\substack{i_1,i_2\in\mathcal{I}_2\\ i_1\neq i_2}} \mathbb{E}[B_{i_1,i_2}].$$

Using Lemma 4 and Lemma 5, it follows

$$\mathbb{E}|T_j^{\spadesuit}| \leq \frac{1}{n_2(n_2-1)}(n_2 2^{-j})^2\, 2^{2j}\left(\frac{\log(n_1)}{n_1}\right) \leq \left(\frac{\log(n_1)}{n_1}\right).$$

Moreover, we get

$$T_j^{\clubsuit}(\lambda) = \frac{1}{n_2(n_2-1)} \sum_{\substack{i_1,i_2\in\mathcal{I}_2\\ i_1\neq i_2}} \sum_k \left[\xi_k(X_{i_1},Y_{i_1})\omega^{\lambda}_{j,k}(X_{i_2},Y_{i_2})\right].$$

By Hölder Inequality and from lemmas 4 and 5, one obtains

$$\mathbb{E}|T_j^{\clubsuit}(\lambda)| \leq \frac{1}{n_2(n_2-1)} \sum_{\substack{i_1,i_2\in\mathcal{I}_2\\ i_1\neq i_2}}^{n_2} \left(\sum_k (\mathbb{E}(A_i))^2 \sum_k (\mathbb{E}\omega^{\lambda}_{j,k}(X_{i_2},Y_{i_2}))^2\right)^{1/2}.$$

Remembering that $\mathbb{E}\omega^{\lambda}_{j,k}(X_{i_2},Y_{i_2}) = (c_{j,k}-c_{j,k}(\lambda))$ for any index i_2, we get

$$\begin{aligned} \mathbb{E}|T_j^{\clubsuit}(\lambda)| &\leq \frac{1}{n_2(n_2-1)}(n_2 2^{-j})n_2\left[2^{2j}\frac{\log(n_1)}{n_1}T_j(\lambda)\right]^{1/2} \\ &\leq K\left(\frac{\log(n_1)}{n_1}T_j(\lambda)\right)^{1/2}. \end{aligned}$$

Let us study the moments of $T_j^{\clubsuit}(\lambda)$ under $\mathbb{P}_\lambda$. Since $\mathbb{E}_\lambda\omega^{\lambda}_{j,k}(X_i,Y_i)=0$ for any k and i, we obviously have $\mathbb{E}_\lambda T_j^{\clubsuit}(\lambda)=0$ and

$$\mathbb{E}_\lambda(T_j^{\clubsuit}(\lambda))^2 = \left(\frac{1}{n_2(n_2-1)}\right)^2 \sum_{i_1\neq i_2} T_{i_1,i_2} + \left(\frac{1}{n_2(n_2-1)}\right)^2 \sum_{i_1\neq i_2\neq i_3} S_{i_1,i_2,i_3},$$

where

$$\begin{aligned} T_{i_1,i_2} &= \sum_{k,p}\left(\mathbb{E}_\lambda\left[\xi_k(X_{i_1},Y_{i_1})\xi_p(X_{i_1},Y_{i_1})\right]\mathbb{E}_\lambda\left[\omega^{\lambda}_{j,k}(X_{i_2},Y_{i_2})\omega^{\lambda}_{j,p}(X_{i_2},Y_{i_2})\right]\right), \\ S_{i_1,i_2,i_3} &= \sum_{k,p}\left(\mathbb{E}_\lambda\left[\xi_k(X_{i_1},Y_{i_1})\xi_p(X_{i_2},Y_{i_2})\right]\mathbb{E}_\lambda\left[\omega^{\lambda}_{j,k}(X_{i_3},Y_{i_3})\omega^{\lambda}_{j,p}(X_{i_3},Y_{i_3})\right]\right). \end{aligned}$$

By Hölder Inequality, we have

$$T_{i_1,i_2} = \sum_{k,p} I\!E_\lambda \left[\xi_k(X_{i_1},Y_{i_1})\xi_p(X_{i_1},Y_{i_1})\right] I\!E_\lambda \left[\omega^\lambda_{j,k}(X_{i_2},Y_{i_2})\omega^\lambda_{j,p}(X_{i_2},Y_{i_2})\right]$$
$$\leq D_{i_1}^{1/2}\left(\sum_{k,p}\left(I\!E_\lambda\left[\omega^\lambda_{j,k}(X_{i_2},Y_{i_2})\omega^\lambda_{j,p}(X_{i_2},Y_{i_2})\right]\right)^2\right)^{1/2}$$

With Parseval Equality, we get

$$\sum_{k,p}\left(I\!E_\lambda\left[\omega^\lambda_{j,k}(X_{i_2},Y_{i_2})\omega^\lambda_{j,p}(X_{i_2},Y_{i_2})\right]\right)^2 \leq \sum_{k,p}\left(\int \phi_{j,k}\phi_{j,p}c_\lambda\right)^2$$
$$\leq K\sum_k \int \phi^2_{j,k}c^2_\lambda \leq K\,2^{2j},$$

which combining with Lemma 5, implies that

$$T_{i_1,i_2} \leq K\left(2^{6j}\left(\frac{\log(n_1)}{n_1}\right)^2\right)^{1/2}\left(2^{2j}\right)^{1/2} \leq 2^{4j}\left(\frac{\log(n_1)}{n_1}\right).$$

In the same way,

$$S_{i_1,i_2,i_3} \leq K\left(\sum_{k,p}(I\!E_\lambda C_{i_1,i_2})^2\right)^{1/2}\left(\sum_{k,p}\left(I\!E\lambda\left[\omega^\lambda_{j,k}(X_{i_2},Y_{i_2})\omega^\lambda_{j,p}(X_{i_2},Y_{i_2})\right]\right)^2\right)^{1/2}$$
$$\leq \left(2^{2j}\right)^{1/2}\left(2^{4j}\left(\frac{\log(n_1)}{n_1}\right)^2\right)^{1/2} \leq 2^{3j}\left(\frac{\log(n_1)}{n_1}\right).$$

¿From Lemma 4, one has

$$I\!E_\lambda(T^\clubsuit_j(\lambda))^2 \leq K\frac{1}{n_2^2(n_2-1)^2}(n_2 2^{-j})n_2 2^{4j}\left(\frac{\log(n_1)}{n_1}\right)$$
$$+K\frac{1}{n_2^2(n_2-1)^2}(n_2 2^{-j})^2 n_2 2^{3j}\left(\frac{\log(n_1)}{n_1}\right)$$
$$\leq K\,2^j\left(\frac{\log(n_1)}{n_2 n_1}\right).$$

10.2 Proof of Lemma 5

The following expansion is crucial because it allows to reduce the study to univariate variables.

$$\xi_k(X_i,Y_i) = \xi_{k_1}(X_i)\xi_{k_2}(Y_i) + \xi_{k_1}(X_i)\phi_{jk_2}(G(Y_i)) + \xi_{k_2}(Y_i)\phi_{jk_1}(F(X_i)), \tag{29}$$

where the univariate statistics $\xi_{k_1}(X_i)$ and $\xi_{k_2}(Y_i)$ are defined as follows

$$\begin{aligned}\xi_{k_1}(X_i) &= \phi_{j,k_1}\left(\frac{\widehat{F}(X_i)}{n_1}\right) - \phi_{j,k_1}(F(X_i)) \\ \xi_{k_2}(Y_i) &= \phi_{j,k_2}\left(\frac{\widehat{G}(Y_i)}{n_1}\right) - \phi_{j,k_1}(G(Y_i)).\end{aligned}$$

Assuming that ϕ is continuously $q-$differentiable, we get

$$\xi_{k_1}(X_i) = \hat{z}_{k_1}(X_i) + \hat{w}_{k_1}(X_i),$$

where

$$\hat{z}_{k_1}(X_i) = \sum_{\ell=1}^{q-1} \frac{2^{j\ell}}{\ell!}(\widehat{F}(X_i) - F(X_i))^{\ell}\, \phi_{j,k_1}^{(\ell)}(F(X_i))$$

and

$$\hat{w}_{k_1}(X_i) = 2^{qj}\int_{\widehat{F}(X_i)}^{F(X_i)} \phi_{j,k_1}^{(q)}(t)\, (F(X_i)-t)^{q-1}dt.$$

A direct application of the Dvoretsky, Kiefer and Wolfovitz inequality leads to the following bound

$$P(\|\widehat{F} - F\|_\infty > \epsilon) \le K\, \exp(-2n_1\epsilon^2) \le K n_1^{-\delta},$$

as soon as $\epsilon = \sqrt{0.5\,\delta\log(n_1)/n_1}$. In the sequel, we take such an ϵ with δ large enough. Since $j \le j_\infty$ where j_∞ is defined in (14), observe that $2^j\epsilon \le 1$ and then we get

$$\begin{aligned}|\hat{z}_{k_1}(X_i)| &\le K\, 2^j\epsilon \max_{\ell=1,\ldots q-1} |\phi_{j,k_1}^{(\ell)}(F(X_i))|(1+o_P(1)) \\ |\hat{w}_{k_1}(X_i)| &\le K\, 2^{(q+1/2)j}\epsilon^q(1+o_P(1))\end{aligned}$$

which leads to the following bound

$$|\xi_{k_1}(X_i)| \le K\left(2^{(q+1/2)j}\epsilon^q + 2^j\epsilon \max_{\ell=1,\ldots q-1} |\phi_{j,k_1}^{(\ell)}(F(X_i))|\right)(1+o_P(1)).$$

The same kind of result obviously holds for $\xi_{k_2}(Y_i)$. In the sequel, we need the following evaluations (which also hold for any derivatives of ϕ). Using expansion (29), we get

$$\xi_k(X_i, Y_i) = S_1 + S_2,$$

where

$$\begin{aligned}S_1 &= \xi_{k_1}(X_i)\xi_{k_2}(Y_i), \\ S_2 &= \xi_{k_1}(X_i)\phi_{j,k_2}(G(Y_i)) + \xi_{k_2}(Y_i)\phi_{j,k_1}(F(X_i)).\end{aligned}$$

Using (30), we get

$$\begin{aligned} I\!E|S_1| &\leq K\left(2^{(2q+1)j}\epsilon^{2q} + 2^{(q+1)j}\epsilon^{q+1} + 2^j\epsilon^2\right), \\ I\!E|S_2| &\leq K\left(2^{qj}\epsilon^q + \epsilon\right). \end{aligned}$$

If $2^j \leq (n_1/\log(n_1))^{1/2-1/2q}$, we obtain $I\!E|\xi_k(X_i, Y_i)| \leq \epsilon$ which ends the proof.

11 Appendix B : Proof of Lemma 2

Applying Bernstein inequality leads to prove Lemma 2

$$I\!P_c\left(|T_j^\diamond(\lambda)| \geq x\right) \leq \exp -\left(\frac{n_2^2 x^2/2}{2^{2j}T_j(\lambda) + n_2 x 2^j T_j(\lambda)^{1/2}/3}\right)$$

provided that $T_j^\diamond(\lambda) = n_2^{-1}\sum_{i\in\mathcal{I}_2} Z_i$ where

$$\begin{aligned} Z_i &= \sum_k \left(\phi_{jk}(F(X_i), G(Y_i)) - c_{jk}\right)\left(c_{jk} - c_{jk}(\lambda)\right), \\ E_c Z_i &= 0, \\ |Z_i| &\leq \left(\sum_k \left(\phi_{jk}(F(X_i), G(Y_i)) - c_{jk}\right)^2 \sum_k \left(c_{jk} - c_{jk}(\lambda)\right)^2\right)^{1/2} \\ &\leq K\left(\left(2^j\right)^2 T_j(\lambda)\right)^{1/2} \leq K\, 2^j T_j(\lambda)^{1/2}, \\ V_c(Z_i) &\leq \sum_{k,p} E\left(\phi_{jk}(F(X_i), G(Y_i)) - c_{jk}\right)\left(\phi_{jp}(F(X_i), G(Y_i)) - c_{jp}\right) \\ &\qquad \times \left(c_{jk} - c_{jk}(\lambda)\right)\left(c_{jp} - c_{jp}(\lambda)\right) \\ &\leq \left(\sum_{k,p} E^2\left(\phi_{jk}(F(X_i), G(Y_i)) - c_{jk}\right)\left(\phi_{jp}(F(X_i), G(Y_i)) - c_{jp}\right)\right. \\ &\qquad \left. \times \sum_k \left(c_{jk} - c_{jk}(\lambda)\right)^2\right)^{1/2} \\ &\leq K\left(2^{4j}\left(2^j 2^j 2^{-2j}\right)^2\right)^{1/2} T_j(\lambda) \leq K\, 2^{2j} T_j(\lambda), \end{aligned}$$

where K is some positive constant depending on L, $\|\phi\|_\infty$ and $\|c\|_\infty$.

12 Appendix C: Proof of Lemma 3

12.1 U-Statistic

Let us first recall the result of Giné et al. (2000).

Proposition 1. (Theorem 3.3 p. 21 Giné et al. (2000))
It exists an universal positive constant $K_g < \infty$ such that, if Ω is a bounded canonical kernel of two variables for the i.i.d. Z_{i_1}, Z_{i_2}, i_1, $i_2 \in \{1, \ldots, \tilde{n}\}$, where $\tilde{n} \in \mathbb{N}$, for any $x > 0$, we have

$$\mathbb{P}\left(|\sum_{i_1,i_2} \Omega(Z_{i_1}, Z_{i_2})| > x \right) \leq K_g \exp\left(-\frac{1}{K_g} \min\left\{ \frac{x^2}{C^2}, \frac{x}{D}, \left(\frac{x}{B}\right)^{2/3}, \left(\frac{x}{A}\right)^{1/2} \right\} \right),$$

where

$$\begin{aligned} A &= \|\Omega(\cdot,\cdot)\|_\infty,\ B^2 = \tilde{n}\left[\|\mathbb{E}[\Omega^2(Z_1,\cdot)]\|_\infty + \|\mathbb{E}\Omega^2(\cdot, Z_2)\|_\infty \right], \\ C^2 &= \tilde{n}^2 \mathbb{E}[(\Omega(Z_1,Z_2))^2] \end{aligned}$$

and

$$D = \tilde{n} \sup_{\Omega_1,\Omega_2} \{\mathbb{E}[\Omega(Z_1,Z_2)\Omega_1(Z_1)\Omega_2(Z_2)] : \mathbb{E}[\Omega_1^2(Z_1)] \leq 1; \mathbb{E}[\Omega_2^2(Z_2)] \leq 1\}.$$

We apply this proposition for $Z_i = (F(X_i), G(Y_i))$, $\tilde{n} = n_2$ and the kernel

$$\Omega_{\tilde{c}}(Z_{i_1}, Z_{i_2}) = \sum_k \{\phi_{j,k}(Z_{i_1}) - \mathbb{E}_{\tilde{c}}[\phi_{j,k}(Z_{i_1})]\} \times \{\phi_{j,k}(Z_{i_2}) - \mathbb{E}_{\tilde{c}}[\phi_{j,k}(Z_{i_2})]\},$$

which is considered under the distribution $\mathbb{P}_{\tilde{c}}$ where $\tilde{c}$ is either c_λ or c. The quantities A, B, C and D are evaluated in the following lemma which is proved in the next section.

Lemma 6. *There exists some positive constant K_1 larger than either*

$$\left(12L^2\|\phi\|_\infty^2\right) \vee (2\|\tilde{c}\|_\infty) \vee \left(2L^2\|\phi\|_\infty^2\right) \vee \left(4\|\tilde{c}\|_\infty(\|\tilde{c}\|_\infty + 3L^4\|\phi\|_\infty^2)\right)$$

such that

$$A \leq K_1 2^{2j},\ B^2 \leq K_1 n_2 2^{2j},\ C^2 \leq K_1 n_2^2 2^{2j},\ D \leq K_1 n_2,$$

where $\tilde{c}$ is either c_λ or c.

Again define $\tilde{c}$ as c_λ or c, then applying both the result of Giné et al. (2000) and Lemma 6, for any level j and any $x \geq 2^j((n_2-1)n_2)^{-1/2}\sqrt{\log(\log(n_2))}$, it immediately follows that

$$\mathbb{P}_{\tilde{c}}\left(|\frac{1}{n_2(n_2-1)} \sum_{\substack{i_1,i_2 \in \mathcal{I}_2 \\ i_1 \neq i_2}} \Omega_{\tilde{c}}(Z_{i_1}, Z_{i_2})| > \mu x \right) \leq K_g \exp\left(-\delta \log(\log(n_2))\right).$$

which ends the proof of Lemma 3

12.2 Proof of Lemma 6

Let us denote $(U,V)=(F(X),G(Y))$ any pair of random variables whose marginal distribution are both uniform on $[0,1]$. Denote $\tilde{c}$ the copula density which is c_λ or c; in the same spirit, the coefficients $\tilde{c}_{j,k}$ stand for $c_{j,k}(\lambda)$ or $c_{j,k}$. Recall that

$$\begin{aligned} c_{j,k}(\lambda) &= \mathbb{E}_\lambda\left[\phi_{j,k}(F(X),G(Y))\right] = \int c_\lambda(u,v)\phi_{j,k}(u,v)dudv. \\ c_{j,k} &= \mathbb{E}\left[\phi_{j,k}(F(X),G(Y))\right] = \int c(u,v)\phi_{j,k}(u,v)dudv. \end{aligned}$$

Note that

$$\begin{aligned} \sum_{k,p} \mathbb{E}_{\tilde{c}}[\phi_{j,k}(U_{i_1},V_{i_1})\phi_{j,p}(U_{i_1},V_{i_1})] &\le 2^{2j}, \\ \sum_k (\mathbb{E}_{\tilde{c}}[\phi_{j,k}(U,V)])^2 &= \sum_k \tilde{c}^2_{j,k} \le \|\tilde{c}\|^2 \le M. \end{aligned}$$

We get

$$\begin{aligned} A &= \|\sum_k \left(\phi_{j,k}(u_1,v_1) - \mathbb{E}_{\tilde{c}}[\phi_{j,k}(U,V)]\right)\left(\phi_{j,k}(u_2,v_2) - \mathbb{E}_{\tilde{c}}[\phi_{j,k}(U,V)]\right)\|_\infty \\ &\le \|\sum_k \phi_{j,k}(u_1,v_1)\phi_{j,k}(u_2,v_2)\|_\infty + 2\|\sum_k \phi_{j,k}(u_1,v_1)\mathbb{E}_{\tilde{c}}[\phi_{j,k}(U,V)]\|_\infty \\ &\quad + \|\sum_k \left(\mathbb{E}_{\tilde{c}}[\phi_{j,k}(U,V)]\right)^2\|_\infty \\ &\le L^2\, 2^{2j}\|\phi\|^2_\infty + 2L^2\|\phi\|_\infty\|\tilde{c}\|_2 2^j + \|\tilde{c}\|_2^2 \le K 2^{2j}, \end{aligned}$$

where $K \ge 2L^2\|\phi\|^2_\infty$ and

$$\begin{aligned} B^2 &= 2n_2 \left\| \sum_{k,p} \mathbb{E}_{\tilde{c}}\left[\left(\phi_{j,k}(U_{i_1},V_{i_1}) - \mathbb{E}_{\tilde{c}}[\phi_{j,k}(U,V)]\right)\left(\phi_{j,p}(U_{i_1},V_{i_1}) - \mathbb{E}_{\tilde{c}}[\phi_{j,p}(U,V)]\right)\right] \right. \\ &\quad \left. \times \left(\phi_{j,k}(u_2,v_2) - \mathbb{E}_{\tilde{c}}[\phi_{j,k}(U,V)]\right)\left(\phi_{j,p}(u_2,v_2) - \mathbb{E}_{\tilde{c}}[\phi_{j,p}(U,V)]\right)\right\|_\infty \\ &\le 2n_2 \left\| \sum_{k,p} \left| \int \phi_{j,k}\phi_{j,p}\tilde{c} - \int \phi_{j,k}\,\tilde{c} \int \phi_{j,p}\,\tilde{c} \right| \left(\phi_{j,k}(u_2,v_2) - \mathbb{E}_{\tilde{c}}[\phi_{j,k}(U,V)]\right) \right. \\ &\quad \left. \times \left(\phi_{j,p}(u_2,v_2) - \mathbb{E}_{\tilde{c}}[\phi_{j,p}(U,V)]\right)\right\|_\infty \\ &\le 2n_2(2\|\tilde{c}\|_\infty) \left[\left\| \sum_{k,p} \phi_{j,k}(u_2,v_2)\phi_{j,p}(u_2,z_2) \right\|_\infty + 2\left\| \sum_{k,p} \phi_{j,k}(u_2,v_2)\mathbb{E}_{\tilde{c}}[\phi_{j,k}(U,V)] \right\|_\infty \right. \\ &\quad \left. + \left\| \sum_{k,p} \mathbb{E}_{\tilde{c}}[\phi_{j,k}(U,V)]\mathbb{E}_{\tilde{c}}[\phi_{j,p}(U,V)] \right\|_\infty \right] \\ &\le (4n_2\|\tilde{c}\|_\infty)\left(2^{2j}L^4 2\|\phi\|^2_\infty + 2L^2 2^j\|\phi\|_\infty + 2^{2j}\|\tilde{c}\|_\infty\right) \le K\, n_2 2^{2j} \end{aligned}$$

where $K \geq 4\|\tilde{c}\|_\infty(\|\tilde{c}\|_\infty + 3L^4\|\phi\|_\infty^2)$. Moreover,

$$\begin{aligned}
C^2 &= n_2^2 \sum_{k,p} \mathbb{E}_{\tilde{c}} \left[(\phi_{j,k}(U_{i_1}, V_{i_1}) - \mathbb{E}_{\tilde{c}}[\phi_{j,k}(U,V)]) (\phi_{j,p}(U_{i_1}, V_{i_1}) - \mathbb{E}_{\tilde{c}}[\phi_{j,p}(U,V)]) \right] \\
&\quad \times \mathbb{E}_{\tilde{c}} \left[(\phi_{j,k}(U_{i_2}, V_{i_2}) - \mathbb{E}_{\tilde{c}}[\phi_{j,k}(U,V)]) (\phi_{j,p}(U_{i_2}, V_{i_2}) - \mathbb{E}_{\tilde{c}}[\phi_{j,p}(U,V)]) \right] \\
&= n_2^2 \sum_{k,p} \left(\mathbb{E}_{\tilde{c}} \left[\phi_{j,k}(U_{i_1}, V_{i_1}) \phi_{j,p}(U_{i_1}, V_{i_1}) \right] - \mathbb{E}_{\tilde{c}} \left[\phi_{j,k}(U,V) \right] \mathbb{E}_{\tilde{c}} \left[\phi_{j,p}(U,V) \right] \right)^2 \\
&= n_2^2 \sum_{k,p} \left(\int \phi_{j,k}\phi_{j,p}\, \tilde{c} - \int \phi_{j,k}\, \tilde{c} \int \phi_{j,p}\, \tilde{c} \right)^2 \\
&\leq n_2^2 \sum_{k,p} \left(\int \phi_{j,k}\phi_{j,p}\, \tilde{c} \right)^2 + n_2^2 \left(\sum_k \left(\int \phi_{j,k}\, \tilde{c} \right)^2 \right)^2 \\
&\leq n_2^2 \sum_k \int \phi_{j,k}^2\, \tilde{c}^2 + n_2^2 \left(\int \tilde{c}^2 \right)^2 \\
&\leq \|\tilde{c}\|_\infty^2 n_2^2 2^{2j} + n_2^2 \|\tilde{c}\|_2^4 \leq K\, n_2^2 2^{2j},
\end{aligned}$$

where $K \geq 2\|\tilde{c}\|_\infty^2$. Denote $u_{\Omega_1,\Omega_2} = \mathbb{E}_{\tilde{c}}[\Omega_{\tilde{c}}(Z_1, Z_2)\Omega_{1,\tilde{c}}(Z_1)\Omega_{2,\tilde{c}}(Z_2)]$ and for $i = 1,\ 2$, put

$$c_i(k) = \int (\phi_{j,k} - \mathbb{E}\tilde{c}\phi_{j,k}(U,V))\Omega_{i,\tilde{c}}\, \tilde{c}.$$

By Hölder Inequality, we get

$$\begin{aligned}
u_{\Omega_1,\Omega_2} &= \sum_k \left(\int (\phi_{j,k} - \mathbb{E}_{\tilde{c}}\phi_{j,k}(U,V))\Omega_{1,\tilde{c}}\, \tilde{c} \right) \left(\int (\phi_{j,k} - \mathbb{E}_{\tilde{c}}\phi_{j,k}(U,V))\Omega_{2,\tilde{c}}\, \tilde{c} \right) \\
&\leq \sqrt{\sum_k (c_1(k))^2 \sum_k (c_2(k))^2}.
\end{aligned}$$

Applying again the inequality of Hölder to $\sum_k (c_1(k))^2$ (the same occurs for $c_2(k)$), one gets

$$\begin{aligned}
\sum_k (c_1(k))^2 \leq \sum_k (&\int (\phi_{j,k} - \mathbb{E}_{\tilde{c}}\phi_{j,k}(U,V))\Omega_{1,\tilde{c}}\, \tilde{c}\mathbb{1}_{[\frac{k_1}{2^j}, \frac{2N-1+k_1}{2^j}] \times [\frac{k_2}{2^j}, \frac{2N-1+k_2}{2^j}]})^2 \\
&\leq \sum_k \left(\int (\phi_{j,k} - \mathbb{E}_{\tilde{c}}\phi_{j,k}(U,V))^2 \tilde{c} \right) \times \\
&\qquad \left(\int (\Omega_{1,\tilde{c}})^2\, \mathbb{1}_{[\frac{k_1}{2^j}, \frac{2L-1+k_1}{2^j}] \times [\frac{k_2}{2^j}, \frac{2L-1+k_2}{2^j}]}\, \tilde{c} \right) \\
&\leq \|\tilde{c}\|_\infty \int (\Omega_{1,\tilde{c}})^2\, \tilde{c} \sum_k \mathbb{1}_{[\frac{k_1}{2^j}, \frac{2L-1+k_1}{2^j}] \times [\frac{k_2}{2^j}, \frac{2L-1+k_2}{2^j}]} \\
&\leq 12\|\phi\|_\infty^2 L^2 \qquad (30)
\end{aligned}$$

since $\mathbb{E}_{\tilde{c}}(\Omega_{1,\tilde{c}}(U))^2) \leq 1$. It follows that $D \leq K\, n_2$, where $K > 12L^2\|\phi\|_\infty^2$.

References

Autin, F., Le Pennec, E., and Tribouley, K. (2008). Thresholding methods to estimate the copula density. *Submitted available online at* http://www.cmi.univ-mrs.fr/∼autin/DONNEES/COPULAS.

Butucea, C. and Tribouley, K. (2006). Nonparametric homogeneity tests. *J. Statist. Plann. Inference*, 136:597–639.

Chen, X. and Fan, Y. (2005). Pseudo-likelihood ratio tests for semiparametric multivariate copula model selection. *Canad. J. Statist.*, 33:389–414.

Deheuvels, P. (1979). La fonction de dépendance empirique et ses propriétés: Un test non paramétrique d'indépendance. *Acad. Royal Bel., Bull. Class. Sci., 5e série*, 65:274–292.

Deheuvels, P. (1981a). A kolmogorov-smirnov type test for independence and multivariate samples. *Rev. Roum. Math. Pures et Appl.,*, 2:213–226.

Deheuvels, P. (1981b). A nonparametric test of independence. *Publications de l'ISUP,*, 26:29–50.

Donoho, D. and Johnstone, I. (1994a). Ideal spatial adaptation via wavelet shrinkage. *Biometrika*, 81:425–455.

Donoho, D. and Johnstone, I. (1994b). Minimax risk for l_q losses over l_p-balls. *Probab. Th. Rel. Fields*, 99:277–303.

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

Frees, E. W. and Valdez, E. A. (1998). Understanding relationships using copulas. *N. Am. Actuar. J.*, 2:1–25.

Gayraud, G. and Pouet, C. (2005). Adaptive minimax testing in the discrete regression scheme. *Probab. Th. Rel. Fields*, 4:531–558.

Genest, C., Ghoudi, K., and Rivest, L.-P. (1998). Comment on "Understanding relationships using copulas," by Edward w. Frees and Emiliano a. Valdez, January 1998. *N. Am. Actuar. J.*, 2:143–149.

Genest, C., Masiello, E., and Tribouley, K. (2008a). Estimating copula densities through wavelets.

Genest, C., Quessy, J.-F., and Rémillard, B. (2006). Goodness-of-fit procedures for copula models based on the integral probability transformation. *Scand. J. Statist.*, 33:337–366.

Genest, C., Rémillard, B., and Beaudoin, D. (2008b). Goodness-of-fit tests for copulas: A review and a power study. *Insurance Math. Econom.*, page in press.

Giné, E., Latala, R., and Zinn, J. (2000). Exponential and moment inequalities for u-statistics. *High Dimensional Probability II–Progress in Probability*, 47:13–38.

Ingster, Y. (1982). On minimax nonparametric detection of a signal in gaussian white noise. *Probl. Inf. Trans.*, 18:61–73.

Ingster, Y. (1993). Asymptotically minimax hypothesis testing for nonparametric alternatives, i, ii, iii. *Math. Methods Stat.*, 2:85–114; 171–189; 249–268.

Ingster, Y. (2000). Adaptive chi-square tests. *J. Math. Sciences*, 99:1110–1120.

Ingster, Y. and Suslina, I. (2002). *Nonparametric Goodness-of-Fit Testing Under Gaussian Models*, volume 169.

Kerkyacharian, G. and Picard, D. (2004). Regression in random design and warped wavelets. *Bernoulli*, 6:1053–1105.

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

Lepskii, O. and Spokoiny, V. (1999). Minimax nonparametric hypothesis testing: the case of inhomogeneous alternative. *Bernoulli*, 5:333–358.

Pouet, C. (2000). Tests minimax non-paramétriques : hypothèse nulle composite et constantes exactes. *These*.

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, 8:229–231.

Spokoiny, V. (1996). Adaptive hypothesis testing using wavelets. *Ann. Stat.*, 24:2477–2498.

Spokoiny, V. (1998). Adaptive ans spacially adaptive testing of a nonparametric hypothesis. *Math. Methods Stat.*, 7:245–273.

Family	parameter grid	Cardinal	$\hat{\alpha}$	Decision
Gumbel	1.05 : 0.1 : 1.95	10	0.00	Yes
Gaussian	0.0 : 0.1 : 0.9	10	0.04	Yes
Clayton	0.5 : 0.1 : 1.4	10	0.42	Yes
Frank	1.5 : 0.5 : 6.0	10	1.00	No
Gumbel	1.0 : 0.05 : 1.95	20	0.00	Yes
Gaussian	0.0 : 0.05 : 0.95	20	0.00	Yes
Clayton	0.5 : 0.05 : 1.45	20	0.54	No
Frank	1.25 : 0.25 : 6.0	20	1.00	No
Gumbel	1.45	1	0.10	Yes
Gaussien	0.48	1	0.12	Yes
Clayton	0.92	1	0.86	No
Frank	3.20	1	1.00	No
Gumbel	1.36 (1.07%)	1	0.02	Yes
Gaussien	0.45 (3.27%)	1	0.08	Yes
Clayton	0.41 (13.15%)	1	0.62	No
Frank	2.88 (3.93%)	1	1.00	No

Table 1: Empirical probability $\hat{\alpha}$ to reject the fit to a fixed parametrical family given in the first column and Decision at the prescribed level $\alpha = 5\%$. Multivariate null hypotheses (first and second part); $H_0 : c = c_{\hat{\lambda}}$, where $\hat{\lambda}$ is obtained by inversion of the empirical Kendall's tau (third part); $H_0 : c = c_{\tilde{\lambda}}$, where $\tilde{\lambda}$ is obtained by minimizing the ASE quantity which is given into brackets (fourth part).

Data Clayton $\lambda = 0.6, \tau = 0.23$			Data Gumbel $\lambda = 1.5, \tau = 0.33$			Data Gaussian $\lambda = 0.4, \tau = 0.26$			Data Frank $\lambda = 2.5, \tau = 0.25$		
λ_0	τ_0	$\hat{\alpha}$	λ_0	τ_0	$\hat{\alpha}$	λ_0	τ_0	$\hat{\alpha}$	λ_0	τ_0	$\hat{\alpha}$
0.05	0.02	1.00	1.05	0.04	1.00	0.05	0.03	1.00	0.50	0.06	0.98
0.10	0.04	0.96	1.10	0.09	1.00	0.10	0.06	0.98	1.00	0.10	0.86
0.15	0.07	0.98	1.15	0.13	1.00	0.15	0.09	0.92	1.50	0.14	0.32
0.20	0.09	0.76	1.20	0.16	0.94	0.20	0.12	0.62	2.00	0.22	0.00
0.25	0.11	0.54	1.25	0.20	0.68	0.25	0.16	0.22	2.50	0.25	0.00
0.30	0.13	0.20	1.30	0.23	0.22	0.30	0.19	0.00	3.00	0.30	0.00
0.35	0.14	0.16	1.35	0.25	0.08	0.35	0.22	0.00	3.50	0.35	0.20
0.40	0.16	0.06	1.40	0.28	0.00	0.40	0.26	0.00	4.00	0.38	0.58
0.45	0.18	0.00	1.45	0.31	0.00	0.45	0.29	0.00	4.50	0.41	1.00
0.50	0.20	0.00	1.50	0.33	0.00	0.50	0.33	0.02			
0.80	0.28	0.00	1.55	0.35	0.00	0.55	0.37	0.36			
0.85	0.29	0.06	1.60	0.37	0.00	0.60	0.40	0.86			
0.90	0.31	0.10	1.65	0.39	0.04	0.65	0.45	1.00			
0.95	0.32	0.20	1.70	0.41	0.16	0.70	0.49	1.00			
1.00	0.33	0.38	1.75	0.42	0.26						
1.05	0.34	0.50	1.80	0.44	0.58						
1.10	0.35	0.66	1.85	0.45	0.72						
1.15	0.36	0.78	1.90	0.47	0.88						
1.20	0.37	0.88	1.95	0.48	0.98						
1.25	0.38	0.96	2.00	0.50	0.98						
1.30	0.39	0.96	2.05	0.51	1.00						
1.35	0.40	1.00	2.10	0.52	1.00						

Table 2: Empirical power $\hat{\alpha}$ for the test of $H_0 : c = c_{\lambda_0}$ against $H_1 : c = c_\lambda$ (λ is given in the first line); c_{λ_0} and c_λ belong to the same parametrical family; $n = 1024$; prescribed level 5%; τ and τ_0 are the Kendall's tau.

Data Student (4) $\lambda = 0.95, \tau = 0.79$			Data Gumbel $\lambda = 4, \tau = 0.75$			Data Gaussian $\lambda = -0.9, \tau = -0.71$			Data Frank $\lambda = 14, \tau = 0.75$		
λ_0	τ_0	$\hat{\alpha}$	λ_0	τ_0	$\hat{\alpha}$	λ_0	τ_0	$\hat{\alpha}$	λ_0	τ_0	$\hat{\alpha}$
0.90	0.71	0.48	2.50	0.60	1.00	-0.95	-0.79	0.98	8.50	0.62	1.00
0.91	0.73	0.18	2.60	0.61	0.98	-0.90	-0.71	0.00	9.00	0.63	0.94
0.92	0.74	0.04	2.70	0.62	0.90	-0.85	-0.64	0.10	9.50	0.64	0.78
0.93	0.76	0.00	2.80	0.64	0.70	-0.80	-0.59	0.94	10.00	0.66	0.38
0.94	0.78	0.00	2.90	0.65	0.40	-0.75	-0.53	1.00	10.50	0.68	0.14
0.95	0.79	0.00	3.00	0.66	0.24	-0.70	-0.49	1.00	11.00	0.70	0.04
0.96	0.82	0.02	3.10	0.67	0.10				11.50	0.71	0.00
0.97	0.84	0.54	3.20	0.68	0.02				14.00	0.72	0.00
0.98	0.87	1.00	3.30	0.69	0.00				16.00	0.77	0.00
0.99	0.91	1.00	3.40	0.70	0.00				16.50	0.78	0.02
			3.50	0.71	0.00				17.00	0.78	0.02
			3.60	0.72	0.00				17.50	0.79	0.08
			3.70	0.72	0.00				18.00	0.79	0.12
			3.80	0.73	0.00				18.50	0.80	0.38
			3.90	0.74	0.00				19.00	0.81	0.38
			4.00	0.75	0.00				19.50	0.81	0.60
			4.10	0.75	0.02				20.00	0.81	0.76
			4.20	0.76	0.00				20.50	0.82	0.84
			4.30	0.76	0.90				21.00	0.82	0.94
			4.40	0.77	1.00				21.50	0.82	0.98
									22.00	0.83	0.99

Table 3: Empirical power $\hat{\alpha}$ for the test of $H_0 \ : \ c = c_{\lambda_0}$ against $H_1 \ : \ c = c_\lambda$ (λ is given in the first line); c_{λ_0} and c_λ belong to the same parametrical family; $n = 1024$; prescribed level 5%; τ and τ_0 are the Kendall's tau.

Copula for fit to c_{λ_0}	Data	$\tau = 0.25$	$\tau = 0.50$	$\tau = 0.75$
Gumbel	Clayton	0.98 (0.72)	1.00 (0.99)	1.00 (1.00)
	Gumbel	**0.00 (0.05)**	**0.00 (0.05)**	**0.00 (0.05)**
	Frank	0.12 (0.15)	0.50 (0.40)	0.94 (0.84)
	Normal	0.02 (0.10)	0.08 (0.18)	0.48 (0.61)
	Student(4)	0.24 (0.14)	0.10 (0.22)	0.28 (0.55)
Clayton	Clayton	**0.00 (0.05)**	**0.96** (0.05)	**1.00** (0.05)
	Gumbel	0.94 (0.86)	1.00 (1.00)	1.00 (1.00)
	Frank	0.38 (0.56)	1.00 (0.96)	1.00 (1.00)
	Normal	0.16 (0.50)	1.00 (0.93)	1.00 (1.00)
	Student(4)	0.48 (0.56)	1.00 (0.95)	1.00 (1.00)
Frank	Clayton	0.30 (0.40)	1.00 (0.89)	1.00 (0.97)
	Gumbel	0.02 (0.33)	0.32 (0.63)	0.24 (0.82)
	Frank	**0.00 (0.05)**	**0.00 (0.05)**	**0.00 (0.05)**
	Normal	0.00 (0.08)	0.02 (0.20)	0.02 (0.41)
	Student(4)	0.28 (0.18)	0.18 (0.08)	0.08 (0.06)
Normal	Clayton	0.20 (0.31)	1.00 (0.80)	1.00 (0.92)
	Gumbel	0.02 (0.24)	0.08 (0.38)	0.02 (0.38)
	Frank	0.00 (0.08)	0.02 (0.20)	0.14 (0.42)
	Normal	**0.00 (0.05)**	**0.00 (0.05)**	**0.00 (0.05)**
	Student(4)	0.24 (0.10)	0.02 (0.08)	0.00 (0.06)
Student(4)	Clayton	0.86 (0.27)	1.00 (0.77)	1.00 (0.93)
	Gumbel	0.42 (0.19)	0.08 (0.34)	0.02 (0.42)
	Frank	0.50 (0.09)	0.24 (0.27)	0.32 (0.41)
	Normal	0.16 (0.05)	0.00 (0.04)	0.00 (0.04)
	Student(4)	**0.00 (0.05)**	**0.00 (0.05)**	**0.00 (0.05)**

Table 4: Empirical power for the test of $H_0 : c = c_{\lambda_0}$ at the given level $\alpha = 5\%$ where c_{λ_0} is specified in the first column and the data are issue from a copula density specified in the second column. The parameter of each copula density is chosen such that the Kendall's tau is respectively $\tau = 0.25, 0.50, 0.75$.