INSTITUT NATIONAL DE LA STATISTIQUE ET DES ETUDES ECONOMIQUES
Série des Documents de Travail du CREST
(Centre de Recherche en Economie et Statistique)

n° 2003-34

Goodness of Fit Tests for Copulas

J.-D. FERMANIAN[1]

[1] CDC Ixis Capital Markets, 47 quai d'Austerlitz, 75648 Paris Cedex 13 and CREST, 15 boulevard Gabriel Péri, 92245 Malakoff Cedex, France. Mail : jdfermanian@cdcixis-cm.com

Goodness of Fit Tests for copulas

Jean-David Fermanian

CREST, 15 bd Gabriel Péri, 92245 Malakoff cedex, France and CDC Ixis Capital Markets, 47 quai d'Austerlitz, 75678 Paris cedex 13. jdfermanian@cdcixis-cm.com.

September, 30, 2003

abstract

This paper defines two distribution free goodness-of-fit test statistics for copulas and states their asymptotic distributions under some composite parametric assumptions. The results are stated formally in an independent identically distributed framework, and partially in time-dependent frameworks. We provide a simulation study and an empirical example by studying the dependency between several couples of equity indices.

Résumé

Nous définissons deux statistiques libres pour des tests d'adéquation de copules ; nous établissons leurs propriétés asymptotiques sous des hypothèses simples et composites. Les résultats sont valables formellement dans un cadre iid, et étendus partiellement dans un cadre de séries temporelles. Nous illustrons les résultats par des simulations et par l'étude de la dépendance entre plusieurs couples d'indices boursiers.

Key words: Copulas, GOF tests, Kernel, Time Series, Basket derivatives.

1 Introduction

In modern finance and insurance, to identify dependence structures between assets is becoming one of the main challenges we are faced with. Dependence must be evaluated for several purposes : pricing and hedging of credit sensitive instrument, particularly n-th to default and CDO's [1], basket derivatives and structured products [2], credit portfolio management [3], credit and market risk measures [4]. Copulas are recognized as key tools to analyze dependence structures. They are becoming more and more popular among academics and practitioners because it is well known the returns of financial assets are non gaussian and exhibit nonlinear features. Thus, multivariate gaussian random variables do not provide satisfying models [5].

The copula of a multivariate distribution can be considered as the part describing its dependence structure as a complement to the behavior of each of its margins. One attractive property of copulas is their invariance under strictly increasing transformations of the margins. Actually, the use of copulas allows to solve a difficult problem, namely to find a whole multivariate distribution, by performing two easier tasks. The first step starts by modelling every marginal distribution. The second step consists of estimating a copula, which summarizes all the dependencies between margins. However this second task is still in its infancy for most of multivariate financial series, partly because of the presence of temporal dependencies (serial autocorrelation, time varying heteroskedasticity,...) in returns of stock indices, credit spreads, interest rates of various maturities...

Estimation of copulas has essentially been spread out in the context of i.i.d. samples. If the true copula is assumed to belong to a parametric family $\mathcal{C} = \{C_\theta, \theta \in \Theta\}$, consistent and asymptotically normally distributed estimates of the parameter of interest can be obtained through maximum likelihood methods. There are mainly two ways to achieve this : a fully parametric method and a semiparametric method. The first method relies on the assumption of parametric marginal distributions. Each parametric margin is then plugged in the full

[1] see Li [45],who proposed a gaussian copula-based methodology for pricing of first-to default, Laurent and Gregory [44] and the references therein.

[2] For instance, Rosenberg [57] estimates copulas for the valuation of options on DAX30 and S&P500 indices. Cherubini and Luciano [12] calibrate Frank's copula for the pricing of digital options and options that are based on the minimum on some equity indices. Genest et al. [31] study the relation between multivariate option prices and several parametric copulas, by assuming a GARCH type model for individual returns.

[3] see e.g. Frey and McNeil [28]

[4] for instance Schönbucher [61], Schönbucher and Schubert [62], among others

[5] see Bouyé et al. [7] for a survey of financial applications, Frees and Valdez [27] for use in actuarial practice. Patton [49, 50] and Rockinger and Jondeau [56] introduce copulas for modelling conditional dependencies.

likelihood and this full likelihood is maximized with respects to θ. Alternatively and without parametric assumptions for margins, the marginal empirical cumulative distribution functions can be plugged in the likelihood. These two commonly used methods are detailed in Genest et al. [29] and Shi and Louis [63] [6].

Beside these two methods, it is also possible to estimate a copula by some nonparametric methods based on empirical distributions, following Deheuvels [14, 15, 16] [7]. The so-called empirical copulas look like usual multivariate empirical cumulative distribution functions. They are highly discontinuous (constant on some data-dependent pavements) and cannot be exploited as graphical device. Recently, smooth estimates of copulas in a time-dependent framework have been proposed in Fermanian and Scaillet [25]. They allow to guess which parametric copula family should be convenient. This intuition needs to be properly verified to be validated. In a statistical sense, it means to lead a goodness-of-fit test on the copula specification. This is our topic.

In section 2, we specify our notations and the GOF tests main results in a multidimensional framework. For i.i.d. data sets, we propose a first simple chi-square type test procedure in section 3. A more powerful and more sophisticated test statistics is described in section 4. The extension to time-dependent series are specified in section 5. The proofs are postponed in the appendix.

2 The framework

Consider an i.i.d. sample of d-dimensional vectors $(\mathbf{X}_i)_{i=1,\dots,n}$. Denote $\mathbf{X}_i = (X_{i1}, \dots, X_{id})$. and H, resp. C, the cumulative distribution function, resp. the copula, of $\mathbf{X}$. Usually, a GOF test tries to distinguish between two assumptions :

$$\mathcal{H}_0 : H = H_0, \text{ against } \mathcal{H}_a : H \neq H_0, \text{ when the zero-assumption is simple, or}$$

$$\mathcal{H}_0 : H \in \mathcal{F}, \text{ against } \mathcal{H}_a : H \notin \mathcal{F}, \text{ when the zero-assumption is composite.}$$

Here, H_0 denotes some known cdf, and $\mathcal{F} = \{H_\theta, \theta \in \Theta\}$ is some known parametric family of cdfs'.

Let us dealing first with simple assumptions. When $d = 1$, the problem is relatively simple. By considering the transformation of $\mathbf{X}$ by H, the empirical process tends weakly to

[6] see Cebrian, Denuit and Scaillet [9] for inference under misspecified copulas, Chen and Fan [10] for inference with β-mixing processes.

[7] see some weak convergence properties and extensions in Fermanian et al. [24].

a uniform Brownian bridge, under the null assumption. Thus, a lot of distribution free GOF statistics are available : Kolmogorov-Smirnov, Anderson-Darling...

To deal with the composite assumption, it is necessary to consider first some estimates of θ. The simplest solution is surely the chi-square test statistics, when θ is estimated by maximum likelihood over grouped data [8]. Actually, the one dimensional case has been deeply studied for a long time. A good reference is D'Agostino and Stephens [13].

In a multidimensional framework, it is more difficult to build distribution free GOF tests, particularly because the previous transformation $\mathbf{Y} = H(\mathbf{X})$ fails to work. More precisely, the law of $\mathbf{Y}$ is no longer distribution free. Thus, several authors have proposed some more or less satisfying solutions.

- Justel et al. [39] have proposed to use the transformation of Rosenblatt before testing a simple GOF assumption. But their method is not really distribution free, because it is necessary to calculate explicitly the transformation. The work is translated from the evaluation of a limiting asymptotic variance (dependent on H_0) towards a transformation of the data set (dependent on H_0). Moreover, the implementation is computationally intensive, because one needs to consider every permutations over $(1, \ldots, d)$ and the extension to composite assumptions is unknown.

- Several authors have tried to replace an evaluation over a d-dimensional space by a univariate function. Therefore, they consider some families of subsets in $\mathbb{R}^d$ indexed by a univariate parameter, for instance $A_\lambda = \{\mathbf{x} | H(\mathbf{x}) \leq \lambda\}$, $\lambda \in [0,1]$. Then, some Kolmogorov-Smirnov type test statistics are available. See Saunders and Laud [60], Foutz [26] or more recently Polonik [54]. The drawback is here to evaluate some subsets like A_λ and to choose the "best" ones. This latter point is the purpose of Polonik [54]. Moreover, extensions to composite assumptions seem to be difficult, even impossible.

- Khmaladze ([41, 42] and especially [43]) has transformed the usual empirical process into an asymptotically distribution free empirical process, for simple and composite assumptions. These techniques are the most conceptually satisfying. Nonetheless, it seems to be to difficult to adapt them in slightly different frameworks, such as copulas, because they are based on some intricate transformations of empirical processes.

[8] which often necessitates some ad-hoc estimation procedures.

Actually, the simplest way to build GOF composite tests for multivariate r.v. is to consider multidimensional chi-square tests, as in Pollard [53]. To do this, it is necessary to choose some subsets in $\mathbb{R}^d$. This choice produces some lack of power and the "best way" to choose these subsets is questionable.

Our goal is to find a technique to solve the similar GOF problem for copulas, say to distinguish between two assumptions :

$$\mathcal{H}_0 : C = C_0, \text{ against } \mathcal{H}_a : C \neq C_0, \text{ when the zero-assumption is simple, or}$$

$$\mathcal{H}_0 : C \in \mathcal{C}, \text{ against } \mathcal{H}_a : C \notin \mathcal{C}, \text{ when the zero-assumption is composite.}$$

Here, C_0 denotes some known copula, and $\mathcal{C} = \{C_\theta, \theta \in \Theta\}$ is some known parametric family of copulas. The copula is the cdf of $(F_1(X_1), \ldots, F_d(X_d))$. The difficulty is coming from the fact the marginal cdfs' F_j are unknown. Particularly, the chi square test procedures do not work anymore in general, when replacing marginal cdfs' by some usual estimates.

For all these reasons, the general problem of GOF test for copulas has not been dealt conveniently by authors. Some of them use the bootstrap procedure to evaluate the limiting distribution of the test statistic (Andersen et al. [3], e.g.). Genest and Rivest [29] solve the problem for the case of archimedian copulas, for which the problem can be reduced to a one dimensional one [9], for which some standard methods are available. For instance, Frees and Valdez [27] use Q-Q plots to fit the "best" archimedian copula. None of the authors have dealt the case of time dependent copulas [10]. Recently, some authors have applied Rosenblatt's transformation (cf [58]) to the original multivariate series, like in Justel et al. [39], before testing the copula specification: Breymann et al. [8], Chen et al. [11] [11]. Nonetheless, as we said previously, the use of Rosenblatt's transformation is a tedious preliminary, especially with high dimension variables, and it is model specific. Thus the test methodology is not really distribution-free.

Note that we could build some test procedures based on some estimates of $\mathbf{X}$'s cdf by modelizing the marginal distributions simultaneously. It seems to be a good idea, because some "more or less usual" tests are available to check the GOF of H itself. Nonetheless, it

[9]see the survey of de Matteis [17]

[10]with the exception of Patton [49, 50]. But the latter author tests all the joint specification and not only the copula itself.

[11]The latter authors compare the smoothed copula density of their transformed r.v. to the uniform density by means of a L^2 criterion, as in Hong and Li [40]. So their methodology is relatively closed to ours (see below the test statistics $\mathcal{T}$).

is not our point of view. Indeed, doing so produces tests for the whole specification - the copula and the margins - but not for the dependence structure itself - the copula only-. A slightly different point of view could be to test each marginal separately in a first step. If each marginal model is accepted, then a test of the whole multidimensional distribution can be led (by the previously cited methodologies). Nonetheless, such a procedure is heavy, and it is always necessary to deal with a multidimensional GOF test. Moreover, it is interesting to study dependence in depth first, independently of the modelization of margins. For instance, imagine the copula links the short not risky interest rate with some credit spreads. It should be useful to keep the possibility to switch from one model of the short rate to another (Vasicek, Longstaff and Schwartz, BGM...). Since this research area is very active, some new models appear regularly, others are forgotten, and current models are often improved. By choosing the copula independently of the marginal models, such evolutions are clearly very easy. Lastly, no universally accepted credit spreads model has emerged yet.

To build a GOF test, a natural way would be to use to asymptotic behavior of the empirical copula process. According to Fermanian et al. [24], we know that the bivariate empirical copula process $n^{1/2}(C_n - C_0)$ tends in law, under the null simple assumption, towards the gaussian process $\mathbb{G}_{C_0}$, where

$$\mathbb{G}_{C_0}(u,v) = \mathbb{B}_{C_0}(u,v) - \partial_1 C_0(u,v)\mathbb{B}_{C_0}(u,1) - \partial_2 C_0(u,v)\mathbb{B}_{C_0}(1,v), \quad (u,v) \in [0,1]^2.$$

We have denoted by $\mathbb{B}_C$ a brownian bridge on $[0,1]^2$, such that

$$E[\mathbb{B}_C(u,v)\mathbb{B}_C(u',v')] = C(u \wedge u', v \wedge v') - C(u,v)C(u',v'),$$

for every $u, u', v, v' \in [0,1]$. Here, $a \wedge b = \inf(a,b)$. Unfortunately, this limiting process is a lot more complicated than with the multidimensional brownian bridge $\mathbb{B}_{C_0}$. For instance, the covariance between $\mathbb{G}_{C_0}(u,v)$ and $\mathbb{G}_{C_0}(u',v')$ is the sum of 18 terms (while there are 2 terms in $\mathbb{B}_{C_0}$'s case). These terms involve C_0 and its partial derivatives. Thus, GOF tests based directly on empirical copula processes C_n seem to be unpractical, except by bootstrapping. Nonetheless, such procedures are computationally intensive. Even if the bootstrapped empirical copula process is weakly convergent (Fermanian et al. [24]), we prefer to propose a more usual test procedure.

3 A simple direct chi-square approach

There exists a simple direct way to circumvent the difficulty. Indeed, by smoothing the empirical copula process, we get an estimate of the copula density. The limit of this statistics is far simpler than $\mathbb{G}_{C_0}$. Let us consider first an i.i.d. framework.

For each index i, set the d-dimensional vectors

$$\mathbf{Y}_i = (F_1(X_{i,1}), \ldots, F_d(X_{i,d})), \text{ and}$$

$$\mathbf{Y}_{n,i} = (F_{n,1}(X_{i,1}), \ldots, F_{n,d}(X_{i,d})),$$

denoting by F_k and $F_{n,k}$ the true and the empirical k-th marginal cdf of $\mathbf{X}$. Obviously, the copula C is the cdf of $\mathbf{Y}_i$.

The empirical copula process we consider here is

$$C_n(\mathbf{u}) = \frac{1}{n} \sum_{i=1}^{n} \prod_{k=1}^{d} \mathbf{1}(F_{n,k}(X_{i,k}) \leq u_k),$$

instead of the "usual" copula process

$$C_n^*(\mathbf{u}) = F_n(F_{n,1}^{-}(u_1), \ldots, F_{n,d}^{-}(u_d)),$$

$$F_{n,k}^{-}(u) = \inf\{t | F_{n,k}(t) \geq u\}.$$

It is easy to verify these two empirical processes differ only by the small quantity n^{-1} at most. Thus, it would not be an hard task to adapt the proofs to C_n^*.

We will assume the law of the vectors $\mathbf{Y}_i$ has a density τ with respects to the Lebesgue measure. By definition the kernel estimator of a copula density τ at point $\mathbf{u}$ is

$$\tau_n(\mathbf{u}) = \frac{1}{h^d} \int K\left(\frac{\mathbf{u} - \mathbf{v}}{h}\right) C_n(d\mathbf{v}) = \frac{1}{nh^d} \sum_{i=1}^{n} K\left(\frac{\mathbf{u} - \mathbf{Y}_{n,i}}{h}\right), \tag{3.1}$$

where K is a d-dimensional kernel and $h = h(n)$ is a bandwidth sequence. More precisely, $\int K = 1$, $h(n) > 0$, and $h(n) \to 0$ when $n \to \infty$. As usual, we denote $K_h(\cdot) = K(\cdot/h)/h^d$. For convenience, we will assume

Assumption (K). The kernel K is the product of d univariate even compactly supported kernels K_r $r = 1, \ldots, d$. It is assumed p_K -times continuously differentiable.

These assumptions are far from minimal. Particularly, we could consider some multivariate kernels whose support is the whole space $\mathbb{R}^d$, if they tend to zero "sufficiently quickly" when their arguments tend to the infinity (for instance, at an exponential rate, like for the gaussian kernel). Since this speed depends on the behavior of τ, we are rather the simpler assumption (K).

As usual, the bandwidth needs to tend to zero not too quick.

Assumption (B0). When n tends to the infinity, $nh^d \to \infty$, $nh^{4+d} \to 0$ and

$$nh^{3+d/2}/(\ln_2 n)^{3/2} \longrightarrow \infty.$$

We have set $\ln_2 n = \ln(\ln n)$. Assumption (B0) can be weakened easily by assuming (K) with $p_K > 3$ (see details in the proofs).

Moreover, a certain amount of regularity of τ is necessary, for instance

Assumption (T0). $\tau(\mathbf{u}, \theta)$ and its first two derivatives with respects to $\mathbf{u}$ are uniformly continuous on $\mathcal{V}(\mathbf{u}_k) \times \mathcal{V}(\theta_0)$, for every vectors $\mathbf{u}_k$, $k = 1, \ldots, m$, denoting by $\mathcal{V}(\mathbf{u}_k)$ (resp. $\mathcal{V}(\theta_0)$) an open neighborhood of $\mathbf{u}_k$ (resp. θ_0).

In the appendix, we prove :

Theorem 1. *Under (K) with $p_K = 3$, (B0) and (T0), for every m and every vectors $\mathbf{u}_1, \ldots, \mathbf{u}_m$ in $]0,1[^d$, such that $\tau(\mathbf{u}_k) > 0$ for every k, we have*

$$(nh^d)^{1/2} \left((\tau_n - \tau)(\mathbf{u}_1), \ldots, (\tau_n - \tau)(\mathbf{u}_m) \right) \xrightarrow[T \to \infty]{law} \mathcal{N}(0, \Sigma),$$

where Σ is diagonal, and its k-th diagonal term is $\int K^2 . \tau^2(\mathbf{u}_k)$.

Now, imagine we want to build a procedure for a GOF test with some composite zero assumption. Under $\mathcal{H}_0$, the parametric family is $\mathcal{C} = \{C_\theta, \theta \in \Theta\}$. Assume we have estimated θ consistently by $\hat{\theta}$, and

$$\hat{\theta} - \theta_0 = O_P(n^{-1/2}). \tag{3.2}$$

We denote by $\tau(\cdot, \theta_0)$ (or simpler τ when there is no ambiguity) the "true" underlying copula density. Clearly, $\tau(\mathbf{u}, \hat{\theta}) - \tau(\mathbf{u}, \theta_0)$ tends to zero quicker than $(\tau_n - \tau)(\mathbf{u})$ under (T0) and equation (3.2). Thus, a simple GOF test could be

$$\mathcal{S} = \frac{nh^d}{\int K^2} \sum_{k=1}^{m} \frac{(\tau_n(\mathbf{u}_k) - \tau(\mathbf{u}_k, \hat{\theta}))^2}{\tau(\mathbf{u}_k, \hat{\theta})^2}.$$

Corollary 2. *Under the assumptions of theorem 1 and equation (3.2), if $\tau(\mathbf{u}, \theta)$ is continuously differentiable with respects to θ in a neighborhood of θ_0 for every $\mathbf{u} \in]0,1[^d$, then $\mathcal{S}$ tends in law towards a m-dimensional chi-square distribution under the composite zero-assumption $\mathcal{H}_0$.*

We could replace τ by the convolution of K and τ in theorem 1 and corollary 2. This allows to remove the assumption $nh^{4+d} \longrightarrow 0$. Indeed, this assumption prevents us from using the usual asymptotically optimal bandwidth that minimizes the asymptotic mean squared error.

The points $(\mathbf{u}_k)_{k=1,\dots,m}$ are chosen arbitrarily. They could be chosen in some particular areas of the d-dimensional square, where the user seeks a good fit. For instance, for risk management purposes, it would be fruitful to consider some dependencies in the tails [12]. For the particular copula family $\mathcal{C}$, it is necessary to specify these areas.

Clearly, the power of the $\mathcal{S}$ test depends strongly on the choice of the points $(\mathbf{u}_k)_{k=1,\dots,m}$. This is a bit the same drawback as the choice of cells in the usual GOF Chi-square test. Without a priori, it is always possible to choose a uniform grid of the type $(i_1/N, i_2/N, \dots, i_d/N)$, for every integers $i_1, \dots, i_d$, $1 \leq i_k \leq N-1$. Nonetheless, the number m will become very large when the dimension d increases.

More seriously, the power of the test will not be very large surely. Actually, the adequacy of the fit for a finite number of points is not a guarantee for a good adequacy of the whole copula. That is why we propose another test statistics. This statistics will take part of the whole underlying distribution potentially, and not only of a finite number of points.

4 The main test

This test is based on the proximity between the smoothed copula density and the estimated parametric density. Under $\mathcal{H}_0$, they will be near each other. To measure such a proximity, we will invoke the L^2 norm. To simplify, denote the estimated parametric $\tau(\cdot, \hat{\theta})$ density by $\hat{\tau}$. Consider the statistics

$$J_n = \int (\tau_n - K_h * \hat{\tau})^2(\mathbf{u})\omega(\mathbf{u})\, d\mathbf{u},$$

where ω is a weight function, viz a measurable function from $[0,1]^d$ towards $\mathbb{R}^+$. Note that we consider the convolution between the kernel K_h and $\hat{\tau}$ instead of $\hat{\tau}$ itself. This trick allows to remove a bias term in the limiting behavior of J_n (see Fan [19]). Note that the expectation of $\tau_n(u)$ is different from $K_h * \tau(u)$, contrary to the usual i.i.d. density case. This will complicate slightly the proof.

The minimization of the criterion J_n is known to produce consistent estimates in numerous situations. These ideas appear first in the seminal paper of Bickel and Rosenblatt [4]. They are applied in the usual density case for i.i.d. observations. Rosenblatt [59] extended the results in a two-dimensional framework and discusses consistency with respects to several

[12] Generally speaking, the tails are related to large values of each margins, so the $\mathbf{u}_k$ should chosen near the boundaries. Nonetheless, some particular directions could be exhibited (the main diagonal, for instance).

alternatives. Fan [19] extended these works to deal with every choices of the smoothing parameter. The comparison of some nonparametric statistics-especially nonparametric regressions- and their model-dependent equivalents has been formalized in a lot of papers in statistics and econometrics : Härdle and Mammen [37], Zheng [66], Fan and Li [20] [13], among others.

Similar results have been obtained for dependent processes more recently : Fan and Ullah [23], Hjellvik et al. [36], Gouriéroux and Tenreiro [33]... For instance, Aït-Sahalia [2] applies these techniques to find a convenient specification for the dynamics of the short interest rate. Recently, Gouriéroux and Gagliardini [34] use such a criterion to estimate possibly infinite dimensional parameters of a copula function [14]. Instead of for inference purposes, we will use J_n as a test statistics, like in Fan [19].

Let us assume that we have found a convenient estimator of θ.

Assumption (E). There exists $\hat{\theta} \in \mathbb{R}^q$ such that

$$\hat{\theta} - \theta_0 = n^{-1} A(\theta_0)^{-1} \sum_{i=1}^{n} B(\theta_0, \mathbf{Y}_i) + o_P(r_n), \tag{4.1}$$

and r_n tends to zero quicker than $n^{-1/2}(\ln_2 n)^{-1/2}$ when n tends to the infinity. Here, $A(\theta_0)$ denotes a $q \times q$ positive definite matrix and $B(\theta_0, \mathbf{Y})$ is a q-dimensional random vector. Moreover, $E[B(\theta_0, \mathbf{Y}_i)] = 0$ and $E[\|B(\theta_0, \mathbf{Y}_i)\|^2] < \infty$.

Particularly, under (E), $\hat{\theta} - \theta = O_P(n^{-1/2})$. Typically, $B(\theta, \cdot)$ is a score function. In section D in the appendix we prove these assumptions are satisfied particularly for the usual semiparametric maximum likelihood estimator whose theoretical properties are detailed in Genest et al. [29] and Chen and Fan [10]. But more general procedures can be used, like M-estimators.

Assumption (T). For some open neighborhood $\mathcal{V}(\theta_0)$ of θ_0,

- $\tau(\mathbf{u}, \theta)$ and its first two derivatives with respects to θ are uniformly continuous on $[0,1]^d \times \mathcal{V}(\theta_0)$, or
- $\tau(\mathbf{u}, \theta)$ and its first two derivatives with respects to θ are uniformly continuous on $[\varepsilon, 1-\varepsilon]^d \times \mathcal{V}(\theta_0)$, for some $\varepsilon > 0$, and the support of ω is included in $[\varepsilon_0, 1-\varepsilon_0]^d$, for some $\varepsilon_0 > \varepsilon$.

[13] see numerous references in Fan and Li [22]

[14] for instance, the univariate function defining an archimedian copula

When τ and its derivatives with respects of θ are uniformly bounded on $[0,1]^d \times \mathcal{V}(\theta_0)$, ω can be chosen arbitrarily. Unfortunately, it is not always the case. For instance, by choosing a bivariate gaussian copula density. To avoid technical troubles, we reduce the GOF test to a strict subsample of $[0,1]^d$, say ω's support.

Assumption (B). $nh^d \longrightarrow \infty$ and $nh^{4+d/2}/\ln_2^2 n \underset{n\to\infty}{\longrightarrow} \infty$.

Actually, the latter condition could be relaxed. It is sufficient to expand K up to higher order terms. We had chosen the order 4 so that condition (B) is not too strong. But, it is possible to exchange some degree of regularity of K against less constraints on the bandwidth.

Theorem 3. *Under assumptions $\mathcal{H}_0$, (T), (E), (B) and (K) with $p_K = 4$, we have*

$$nh^{d/2}\left(J_n - \frac{1}{nh^d}\int K^2(\mathbf{t}).(\tau\omega)(\mathbf{u}-h\mathbf{t})\,d\mathbf{t}\,d\mathbf{u} + \frac{1}{nh}\int \tau^2\omega.\sum_{r=1}^d \int K_r^2\right) \xrightarrow[n\to\infty]{law} \mathcal{N}(0,2\sigma^2),$$

$$\sigma^2 = \int \tau^2\omega \cdot \int \left\{\int K(\mathbf{u})K(\mathbf{u}+\mathbf{v})\,d\mathbf{u}\right\}^2 d\mathbf{v}.$$

Thus, a test statistics could be

$$\mathcal{T} = \frac{n^2h^d\left(J_n - (nh^d)^{-1}\int K^2(\mathbf{t}).(\hat\tau\omega)(\mathbf{u}-h\mathbf{t})\,d\mathbf{t}\,d\mathbf{u} + (nh)^{-1}\int \hat\tau^2\omega.\sum_{r=1}^d\int K_r^2\right)^2}{2\int \hat\tau^2\omega\cdot\int\left\{\int K(\mathbf{u})K(\mathbf{u}+\mathbf{v})\,d\mathbf{u}\right\}^2 d\mathbf{v}}.$$

Corollary 4. *Under the assumptions of theorem 3, the previous statistics $\mathcal{T}$ tends in law towards a chi-square distribution.*

See the proof in the appendix. Since the kernel K is even, we can replace the second term of the previous numerator by the simpler expression $-(nh^d)^{-1}\int K^2.\int\hat\tau\omega$. Moreover, the third term in the numerator can be replaced by

$$\frac{2}{nh}\sum_r \int K(\mathbf{t})K_r(t_r)\hat\tau(\mathbf{u}-h\mathbf{t})\tau(\mathbf{u})\omega(\mathbf{u})\,dt\,d\mathbf{u} - \frac{1}{nh}\sum_r\int\hat\tau^2\omega\cdot\int K_r^2.$$

This expression is a consequence of the proof, and offers surely a better approximation, even if it is a bit more complicated.

Moreover, under some additional regularity assumptions, we could replace $\hat\tau$ by τ_n inside $\mathcal{T}$. Indeed, it can be proved the kernel estimator of the density $\tau(\mathbf{u})$ converges uniformly with respects to $\mathbf{u}$ on ω's support at a convenient rate. The proof requires to control the uniform upper bound of the remainder terms $R_k(\mathbf{u})$, $k=1,2,3$ that are defined in the proof of theorem 1. This can be done by applying lemma B1 in Ai [1], e.g. The details are left to the reader.

Note that our test statistics differs from similar GOF test statistics in an i.i.d. framework with usual density functions (Fan [19], e.g.). Indeed, there is an additional term

$$(nh)^{-1} \int \hat{\tau}^2 \omega . \sum_{r=1}^{d} \int K_r^2,$$

in $\mathcal{T}$. This is the price to work with copulas, and to estimate the margins empirically. Nonetheless, when $d > 2$, this additional term is negligible with respects to

$$(nh^d)^{-1} \int K^2(\mathbf{t}).(\hat{\tau}\omega)(\mathbf{u} - h\mathbf{t})\, d\mathbf{t}\, d\mathbf{u}.$$

5 Extension to time-dependent series

Now, we would relax the assumption of independence between the random vectors $(\mathbf{X}_i)_{i\geq 0}$. Actually, most of financial series are dependent, sometimes strongly. This is particularly true for fixed income underlyings [15] [16].

Now, consider a $\mathbb{R}^d$-valued process $(\mathbf{X}_i)_{i\geq 1}$. It is assumed stationary. In practice, we are interested in the prediction of future values knowing the past. So, we should specify and test conditional copulas (see Patton [49, 50], e.g.). As in Chen and Fan [10], we restrict ourselves to the particular case when the margins are unconditional. For instance, if $(Z_i)_{i\geq 1}$ is an one-order Markov process, we can set $\mathbf{X}_i = (Z_i, Z_{i-1})$. Knowing the copula of $\mathbf{X}_i$ and the stationary distribution of Z_i is equivalent to specify the dynamic of process Z.

It seems to be relatively easy to extend theorem 1. To be specific assume

Assumption (M). $(\mathbf{X}_i)_{i\geq 1}$ is strictly stationary. Moreover, this process is β-mixing [17], viz

$$\beta_p \equiv \sup_{i\geq 1} E\left[\sup_{A\in\mathcal{M}_{i+p}^{\infty}} |P(A|\mathcal{M}_{-\infty}^{i}(\mathbf{X})) - P(A)| \right] \underset{p\to\infty}{\longrightarrow} 0,$$

denoting by $\mathcal{M}_a^b$, $a \leq b$, the sigma algebra generated by $(\mathbf{X}_a, \ldots, \mathbf{X}_b)$. Moreover, $\beta_p = O(\rho^p)$ for some $0 < \rho < 1$ and every integer p.

[15] think of the dynamics of the term structure of interest rates, for instance

[16] By differentiating series of indices or rates, dependence between successive observations is weakened significantly but not disappears. For instance, it is well-known the covariances of squared returns exhibit significant clustering. That is why a lot of GARCH-type models have been proposed to deal with these features, in the literature.

[17] or absolutely regular

This assumption is commonly used, even if it is far from the weakest one (see Doukhan [18] for some other mixing concepts). The geometric decay is encountered for a lot of processes. For instance, Mokkadem [48] and Pham and Tran [52] show that linear stationary ARMA processes satisfy assumption (M) under some conditions of regularity. It is the case for bilinear models (Pham [51]), nonlinear autoregressive models (Mokkadem [48]) and nonlinear ARCH models (Masry and Tjøstheim [46]).

Assumption (B1). $h \ln^2 n \to 0$ and $nh^{d+4}/\ln^4 n \longrightarrow \infty$.

Assumption (T1). The density $\tau_{i,j,k,l}$ of $(\mathbf{Y}_i, \mathbf{Y}_j, \mathbf{Y}_k, \mathbf{Y}_l)$ exists whenever $i < j < k < l$, is integrable and $\sup_{i<j<k<l} \int \tau_{i,j,k,l} < \infty$. Moreover, with obvious notations,

$$\sup_{i<j} \int |\tau_{i,j}(\mathbf{u}, \mathbf{v}) - \tau(\mathbf{u})\tau(\mathbf{v})| \, d\mathbf{u} \, d\mathbf{v} < \infty.$$

In the appendix, we prove :

Theorem 5. *Under (K) with $p_K = 2$, (B1), (M), (T0) and (T1), for every m and every vectors $\mathbf{u}_1, \ldots, \mathbf{u}_m$ in $]0,1[^d$ such that $\tau(\mathbf{u}_k) > 0$, we have*

$$(nh^d)^{1/2} \left((\tau_n - K_h * \tau)(\mathbf{u}_1), \ldots, (\tau_n - K_h * \tau)(\mathbf{u}_m) \right) \xrightarrow[T \to \infty]{law} \mathcal{N}(0, \Sigma),$$

where Σ is diagonal, and its k-th diagonal term is $\int K^2 . \tau^2(\mathbf{u}_k)$.

Thus, the test statistics $\mathcal{S}$ can be used exactly as in corollary 2. Notice we have used $K_h * \tau$ in the latter theorem instead of τ itself. Indeed, to remove the bias term $K_h * \tau - \tau$, we need to assume $nh^{d+4} \to 0$. But it can not be done in our case because of (B1). We get the test statistic

$$\tilde{\mathcal{S}} = \frac{nh^d}{\int K^2} \sum_{k=1}^{m} \frac{(\tau_n(\mathbf{u}_k) - K_h * \tau(\mathbf{u}_k, \hat{\theta}))^2}{\tau(\mathbf{u}_k, \hat{\theta})^2}.$$

A natural idea would be to extend the test statistic $\mathcal{T}$ for this dependent framework. Such an idea is surely possible. The tools we need is a Central Limit Theorem for degenerate U-Statistics of $\beta-$mixing processes. Fan and Li [21] have provided such a tool. Moreover, Fan and Ullah [23] have studied the asymptotic behavior of $\int (f_n - K_h * \hat{f})^2$ under the null hypothesis, when dealing with the marginal density f of a β-mixing process. Therefore, if we prove all the remainder terms that appear in the proof of theorem 3 are negligible, we can state the same results for dependent series.

By applying the same techniques as in theorem 5's proof, it seems to us we can achieve this goal. Particularly, we need to use Lemma 6 to control the distance between joint densities and products of marginal densities. Unfortunately, to state formally the result, it is necessary to investigate cautiously a lot of terms. Since it is very tedious, we are rather to conjecture:

Conjecture. Under (M) and some technical assumptions, theorem 3 and corollary 4 are true for β-mixing processes. Therefore, the test statistics $\mathcal{T}$ is still available.

6 A short simulation study

To asses the power of our test statistics, we have led a simple analysis by simulation. We generate some samples whose copula is the mixture of a bivariate frank's copula and an independent copula, viz

$$C_{\theta,\alpha}(u,v) = \alpha uv - \frac{(1-\alpha)}{\theta}\ln\left(1 + \frac{(\exp(-\theta u)-1)(\exp(-\theta v)-1)}{\exp(-\theta)-1}\right),\ \theta \neq 0,\ \alpha \in [0,1].$$

More precisely, we generate iid uniform samples $(U_{i,1},U_{i,2})_{i=1,\dots,200}$ on $[0,1]^2$. For every $i = 1,\dots,200$, we get

$$X_{i,1} = \Phi^{-1}(U_{i,1}) \text{ and } X_{i,2} = \Phi^{-1}(V_i),$$

where V_i satisfies the equality $\partial_1 C_{\theta,\alpha}(U_{i,1},V_i) = U_{i,2}$. Thus, the random vectors $(X_{i,1},X_{i,2})$ have the desired copula.

We compute the test statistics $\mathcal{S}$ and $\mathcal{T}$ with these data sets. Concerning $\mathcal{S}$, we choose 81 points on the uniform grid $(i/10, j/10)$, $i,j = 1,\dots,9$. We use the convolution between K and $\hat{\tau}$ instead of $\hat{\tau}$ itself. Concerning $\mathcal{S}$ and $\mathcal{T}$, the kernel is a sufficiently regular compactly supported kernel, say

$$K(\mathbf{u}) = \left(\frac{15}{16}\right)^2 \prod_{k=1}^{2}(1-u_k^2)^2\mathbf{1}(u_k \in [0,1]).$$

The bandwidths are chosen by the usual Silverman's rule (1986) :

$$\hat{h} = \frac{\sqrt{(\sigma_1^2+\sigma_2^2)/2}}{n^{1/6}},$$

denoting by σ_k^2 the empirical variance of $F_{n,k}$, $k = 1,2$. Note that these two variances are the same in our case because they depend on the sample size only. The weight function w

is chosen as $w(\mathbf{u}) = \mathbf{1}(\mathbf{u} \in [0.01, 0.99])$. The parameters of the copulas are estimated by the usual semiparametric maximum likelihood procedure (see Shi and Louis [63]).

For different values for α and θ, we compute the two test statistics $\mathcal{S}$ and $\mathcal{T}$. We have made 100 replications for 200 points samples: see table 1. When α is zero, we verify the asymptotic level 0.05 is underestimated by all but one case (thus the test is a bit too conservative). When α increases, the percentages of rejection grow, especially for the $\mathcal{T}$ test. The latter seems to be more powerful than $\mathcal{S}$, even if this advantage weakens when the copula is more and more far from the Frank's copula (viz when the zero assumption is more and more false). When α is near 1, note that the power is very weak. This is due to the fact the independent copula belongs to the boundary of the Frank's family (when leaving θ to tend to zero). Moreover, the reported powers are more and more high when θ is larger and larger, because the corresponding Frank's copula becomes far away from the independent one.

These partial results are very convincing. Particularly, with very small sample sizes, the power of the test $\mathcal{T}$ is far from ridiculous even when the proportion of perturbation is weak. Our results seem to be better than those reported by the test 1 proposed by Chen et al. [11]. In the latter case, the powers are near zero when the sample sizes is not greater than 500 for every level of perturbation but with a different model. Nonetheless, our results need to be completed by a more deep simulation study.

7 Empirical results

We have tried to find a convenient copula that would describe the dependence structure between the daily returns of several couples of equity indices: S&P500 and Nikkei225, Nasdaq Composite and Dow Jones Industrial, DAX30 and Swiss MI. These couples of indices are considered in commonly traded basket derivatives in the equity market. Each data set consists of $3,373$ observations from January 1st 1990 to April 12th 2002. The data have been downloaded from Datastream.

We will consider five families of bivariate copulas.

- gaussian copulas : $C(u, v) = \Phi_\rho(\Phi^{-1}(u), \Phi^{-1}(v))$, $0 \leq u, v \leq 1$, $|\rho| < 1$. Here, Φ_ρ denotes the bivariate cdf of a gaussian vector whose components have unit variance and correlation ρ.

% of noise	parameter	% of rejection (test $\mathcal{S}$)	% of rejection (test $\mathcal{T}$)
$\alpha = 0.0$	$\theta = 5$	0	2
	$\theta = 10$	0	0
	$\theta = 15$	0	1
	$\theta = 20$	0	1
	$\theta = 25$	0	8
$\alpha = 0.1$	$\theta = 5$	0	0
	$\theta = 10$	0	0
	$\theta = 15$	0	7
	$\theta = 20$	0	22
	$\theta = 25$	0	60
$\alpha = 0.2$	$\theta = 5$	1	1
	$\theta = 10$	1	5
	$\theta = 15$	3	36
	$\theta = 20$	17	80
	$\theta = 25$	31	95
$\alpha = 0.3$	$\theta = 5$	3	3
	$\theta = 10$	13	21
	$\theta = 15$	18	67
	$\theta = 20$	57	95
	$\theta = 25$	84	100
$\alpha = 0.6$	$\theta = 5$	12	8
	$\theta = 10$	21	8
	$\theta = 15$	56	50
	$\theta = 20$	86	84
	$\theta = 25$	99	98
$\alpha = 0.9$	$\theta = 5$	2	1
	$\theta = 10$	3	0
	$\theta = 15$	6	0
	$\theta = 20$	2	2
	$\theta = 25$	37	3

Table 1: Percentages of rejection at 5% level with $n = 200$ and 100 replications.

- Student copulas, whose density is

$$\tau(y_1, y_2) = \frac{\Gamma((\nu+2)/2))}{\Gamma(\nu/2)\sqrt{1-\rho^2}\pi\nu}\left(1 + (y_1^2 + y_2^2 - 2\rho y_1 y_2)/(\nu(1-\rho^2))\right)^{-(\nu+2)/2}, \ |\rho| < 1, \ \nu > 0.$$

- Frank's copulas :

$$C_\theta(u, v) = -\frac{1}{\theta}\ln\left(1 + \frac{(\exp(-\theta u) - 1)(\exp(-\theta v) - 1)}{\exp(-\theta) - 1}\right), \ \theta \neq 0.$$

- Clayton's copulas :

$$C_\theta(u, v) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}, \ \theta > 0.$$

- Gumbel's copulas :

$$C_\theta(u, v) = \exp\left(-[\ln(1/u)^{1/\theta} + \ln(1/v)^{1/\theta}]^\theta\right), \ \ \theta \in]0, 1].$$

Gaussian copulas are the most used in practice partly because their parameters can be estimated very easily by empirical means. Moreover, they extend the simple multidimensional gaussian framework naturally. Student copulas are of interest because they include the gaussian copulas family as a limit case (when ν tends to the infinity). But contrary to the latter, they exhibit dependence in the tails. As noticed in Hu [38], the three other families represent three degree of dependence. Gumbel (resp. Clayton) copulas allow to modelize dependence in the right (resp. left) tails. Finally, Frank's copulas are symmetric.

We do the assumption the sequence of daily returns are independent marginally [18]. By using the same specifications as in section 6, we get tables 2, 3 and 4.

The conclusions are the same for the two test statistics [19]. All these copulas are rejected by the test at the 1% level. Nonetheless, the Student copula seems to be the "less bad" one, especially for the couple (S&P500,Nikkei225). At the opposite, the gaussian copulas are strongly rejected. When Clayton's and Frank's copulas performances are closed, Gumbel's copulas provide a worse fit: The strong dependence between large negative returns can be explained partly by the former ones, contrary to the latter.

[18] even if it is a crude approximation: see ARCH, GARCH, stochastic volatility models...

[19] the critical values for $\mathcal{T}$ are 3.84 and 6.63 at the levels 5% and 1%. For $\mathcal{S}$, they are respectively 103.0 and 113.5.

copula family	$\mathcal{T}$	$\mathcal{S}$	estimated parameters
Student	6.30	91.4	$\hat{\nu} = 15.6,\ \hat{\rho} = 0.13$
Gaussian	38.3	430	$\hat{\rho} = 0.12$
Clayton	5.26	142.0	$\hat{\theta} = 0.16$
Frank	5.00	142.4	$\hat{\theta} = 0.80$
Gumbel	5.30	129.7	$\hat{\theta} = 0.93$

Table 2: Goodness-of-Fit test for the copula of the joint returns (S&P500, Nikkei225).

copula family	$\mathcal{T}$	$\mathcal{S}$	estimated parameters
Student	41.2	194	$\hat{\nu} = 4.15,\ \hat{\rho} = 0.68$
Gaussian	73.0	965	$\hat{\rho} = 0.67$
Clayton	35.8	481	$\hat{\theta} = 1.30$
Frank	27.0	974	$\hat{\theta} = 5.50$
Gumbel	78.0	7518	$\hat{\theta} = 0.55$

Table 3: Goodness-of-Fit test for the copula of the joint returns (Nasdaq, Dow Jones Ind).

copula family	$\mathcal{T}$	$\mathcal{S}$	estimated parameters
Student	42.8	250	$\hat{\nu} = 3.73,\ \hat{\rho} = 0.67$
Gaussian	69.3	1374	$\hat{\rho} = 0.67$
Clayton	27.5	604	$\hat{\theta} = 1.36$
Frank	24.0	873	$\hat{\theta} = 5.25$
Gumbel	67.8	6347	$\hat{\theta} = 0.56$

Table 4: Goodness-of-Fit test for the copula of the joint returns (Dax30, Swiss MI).

A Proof of theorem 1

We will prove that the behavior of $\tau_n(\mathbf{u})$ is the same as the behavior of

$$\tau_n^*(\mathbf{u}) = n^{-1}\sum_{i=1}^n K_h\left(\mathbf{u}-\mathbf{Y}_i\right),$$

for every $\mathbf{u}$. Indeed,

$$\begin{aligned}\tau_n(\mathbf{u}) &= \tau_n^*(\mathbf{u}) + \frac{(-1)}{nh}\sum_{i=1}^n (dK)_h(\mathbf{u}-\mathbf{Y}_i).(\mathbf{Y}_{ni}-\mathbf{Y}_i)\\ &+ \frac{1}{2nh^2}\sum_{i=1}^n (d^2K)_h(\mathbf{u}-\mathbf{Y}_i).(\mathbf{Y}_{ni}-\mathbf{Y}_i)^{(2)} + \frac{(-1)}{6nh^3}\sum_{i=1}^n (d^3K)_h(\mathbf{u}-\mathbf{Y}_{ni}^*).(\mathbf{Y}_{ni}-\mathbf{Y}_i)^{(3)}\\ &= \tau_n^*(\mathbf{u}) + R_1(\mathbf{u}) + R_2(\mathbf{u}) + R_3(\mathbf{u}),\end{aligned}$$

for some random vector $\mathbf{Y}_{ni}^*$ satisfying $\|\mathbf{Y}_{n,i}^* - \mathbf{Y}_i\| \leq \|\mathbf{Y}_{n,i}-\mathbf{Y}_i\|$ a.e.

Let us first study $R_1(\mathbf{u})$. Its expectation is $O(n^{-1}h^{-1})$. Moreover

$$\begin{aligned}E[R_1^2(\mathbf{u})] &= \frac{1}{n^2h^2}\sum_{i,j} E\left[(dK)_h(\mathbf{u}-\mathbf{Y}_i).(\mathbf{Y}_{ni}-\mathbf{Y}_i).(dK)_h(\mathbf{u}-\mathbf{Y}_j).(\mathbf{Y}_{nj}-\mathbf{Y}_j)\right]\\ &= \frac{1}{n^4h^2}\sum_{i,j}\sum_{k,l} E\left[(dK)_h(\mathbf{u}-\mathbf{Y}_i).(\mathbf{1}(\mathbf{Y}_k\leq\mathbf{Y}_i)-\mathbf{Y}_i).(dK)_h(\mathbf{u}-\mathbf{Y}_j).(\mathbf{1}(\mathbf{Y}_l\leq\mathbf{Y}_j)-\mathbf{Y}_j)\right].\end{aligned}$$

We will denote by $\mathbf{1}(\mathbf{y}\leq\mathbf{u})$ a d-dimensional vector whose k-th component is $\mathbf{1}(y_k\leq u_k)$. The expectations of the summands are zero, except if there are some equalities involving k and l. For instance, assume $k=l\neq i\neq j$. Let us note that

$$\begin{aligned}&E\left[(dK)_h(\mathbf{u}-\mathbf{Y}_j).(\mathbf{1}(\mathbf{Y}_i\leq\mathbf{Y}_j)-\mathbf{Y}_j)|\mathbf{Y}_i=\mathbf{y}_i\right] = \int (dK)_h(\mathbf{u}-\mathbf{v}).(\mathbf{1}(\mathbf{y}_i\leq\mathbf{v})-\mathbf{v})\tau(\mathbf{v})\,d\mathbf{v}\\ &= \sum_{r=1}^d\int(\partial_rK)_h(\mathbf{u}-\mathbf{v}).(\mathbf{1}(y_{i,r}\leq v_r)-v_r)\tau(\mathbf{v})\,d\mathbf{v}\\ &= \sum_{r=1}^d\int(\partial_rK)(\mathbf{v}).(\mathbf{1}(y_{i,r}\leq u_r-hv_r)-u_r+hv_r)\tau(\mathbf{u}-h\mathbf{v})\,d\mathbf{v}\\ &= \sum_{r=1}^d\int(\partial_rK)(\mathbf{v}).(\mathbf{1}(v_r\leq(u_r-y_{i,r})/h)-u_r+hv_r)\{\tau(\mathbf{u})+h\psi(\mathbf{u},\mathbf{v})\}\,d\mathbf{v},\end{aligned}$$

where ψ is a bounded compactly supported function, for n sufficiently large. Since we assume K is the product of some univariate kernels K_r, $r=1,\ldots,d$, we get for every couple (i,j) with $i\neq j$,

$$E\left[(dK)_h(\mathbf{u}-\mathbf{Y}_j).(\mathbf{1}(\mathbf{Y}_i\leq\mathbf{Y}_j)-\mathbf{Y}_j)|\mathbf{Y}_i=\mathbf{y}_i\right] = \tau(\mathbf{u})\sum_{r=1}^d K_r\left(\frac{u_r-y_{i,r}}{h}\right) + O(h).\phi(\mathbf{u}), \tag{A.1}$$

where ϕ is bounded, compactly supported and independent from $\mathbf{y}_i$.

Thus, the corresponding term in $E[R_1^2(\mathbf{u})]$ is

$$\frac{1}{nh^2}\int\left\{\tau(\mathbf{u})\sum_r K_r\left(\frac{u_r-y_r}{h}\right)+O(h)\phi(\mathbf{u})\right\}$$
$$\cdot\quad\left\{\tau(\mathbf{u})\sum_s K_s\left(\frac{u_s-y_s}{h}\right)+O(h)\phi(\mathbf{u})\right\}\tau(\mathbf{y}_1)\tau(\mathbf{y}_2)\,d\mathbf{y}$$
$$=\quad\frac{1}{nh^2}\int\tau^2(\mathbf{u})\left\{\sum_{r\neq s}K_r\left(\frac{u_r-y_r}{h}\right)K_s\left(\frac{u_s-y_s}{h}\right)\,dy_r\,dy_s\right.$$
$$+\quad\left.\sum_r K_r^2\left(\frac{u_r-y_r}{h}\right)\,dy_r\right\}+O(n^{-1}h^{-1})=O(\frac{1}{nh})=o(n^{-1}h^{-d}),$$

by some usual changes of variables with respects to y_r and y_s. The other equalities between i,j,k and l provide a similar conclusion. Thus, the variance of $R_1(\mathbf{u})$ is $o(n^{-1}h^{-d})$, and $R_1(\mathbf{u})=o_P(1/\sqrt{nh^d})$.

The study of $R_2(\mathbf{u})$ is similar. We get by the same method $E[R_2(\mathbf{u})]=O(n^{-1}h^{-2})$ and $E[R_2^2(\mathbf{u})]=O(n^{-2}h^{-4})$, hence $R_2(\mathbf{u})=o_P(1/\sqrt{nh^d})$. Since,

$$\|\mathbf{Y}_{n,i}-\mathbf{Y}_i\|_\infty=O_P\left(\left(\frac{\ln_2 n}{n}\right)^{1/2}\right),\tag{A.2}$$

we deduce directly $R_3(\mathbf{u})=O_P(h^{-3-d}n^{-3/2}.\ln_2^{3/2}n)$, which is $o_P(n^{-1/2}h^{-d/2})$ if $nh^{3+d/2}/\ln_2^{3/2}n$ tends to the infinity when $n\to\infty$. Thus, under our assumptions,

$$\tau_n(\mathbf{u})=\tau_n^*(\mathbf{u})+o_P\left(\frac{1}{\sqrt{nh^d}}\right).$$

Moreover, Bosq and Lecoutre's [5] theorem VIII.2 provides the asymptotic normality of the joint vector $(nh^d)^{1/2}((\tau_n^*-\tau)(\mathbf{u}_1),\ldots,(\tau_n^*-\tau)(\mathbf{u}_m))$. This concludes the proof. □.

B Proof of theorem 3

Clearly,

$$\begin{aligned}J_n&=\int(\tau_n-K_h*\hat{\tau})^2(\mathbf{u})\omega(\mathbf{u})\,d\mathbf{u}\\&=\int(\tau_n-E\tau_n)^2\omega+2\int(\tau_n-E\tau_n)(\mathbf{u}).(E\tau_n-K_h*\hat{\tau})(\mathbf{u})\omega(\mathbf{u})\,d\mathbf{u}\\&+\int(E\tau_n-K_h*\hat{\tau})^2\omega\\&\equiv\int(\tau_n-E\tau_n)^2\omega+2J_I+J_{II}.\end{aligned}\tag{B.1}$$

The main term of J_n will be

$$J_n^* = \int (\tau_n - E\tau_n)^2 \omega = \frac{1}{n} \int \left(\sum_{i=1}^{n} K_h(\mathbf{u} - \mathbf{Y}_{n,i}) - EK_h(\mathbf{u} - \mathbf{Y}_{ni}) \right)^2 \omega(\mathbf{u})\, d\mathbf{u}$$

$$= \frac{1}{n^2} \sum_{i,j=1}^{n} \int \left(K_h(\mathbf{u} - \mathbf{Y}_{n,i}) - EK_h(\mathbf{u} - \mathbf{Y}_{n,i}) \right) . \left(K_h(\mathbf{u} - \mathbf{Y}_{n,j}) - EK_h(\mathbf{u} - \mathbf{Y}_{n,j}) \right) \omega(\mathbf{u})\, d\mathbf{u}$$

Thus,

$$J_n^* = \frac{1}{n^2} \sum_{i=1}^{n} \int a_{n,i}^2 \omega + \frac{2}{n^2} \sum_{i<j} \int a_{n,i} a_{n,j} \omega \equiv J_{n,1}^* + J_{n,2}^* \,, \tag{B.2}$$

where we have set

$$a_{n,i}(\mathbf{u}) = K_h(\mathbf{u} - \mathbf{Y}_{n,i}) - EK_h(\mathbf{u} - \mathbf{Y}_{n,i}).$$

Intuitively, $a_{n,i}(\mathbf{u})$ is close to

$$a_i(\mathbf{u}) = K_h(\mathbf{u} - \mathbf{Y}_i) - EK_h(\mathbf{u} - \mathbf{Y}_i).$$

For technical reasons, we will need to expand the difference between the two latter terms up to the fourth order, viz

$$a_{n,i}(\mathbf{u}) - a_i(\mathbf{u}) = b_{n,i}(\mathbf{u}) + c_{n,i}(\mathbf{u}) + d_{n,i}(\mathbf{u}) + e_{n,i}(\mathbf{u}),$$

$$b_{n,i}(\mathbf{u}) = \frac{(-1)}{h} \left[(dK)_h(\mathbf{u} - \mathbf{Y}_i).(\mathbf{Y}_{n,i} - \mathbf{Y}_i) - E(dK)_h(\mathbf{u} - \mathbf{Y}_i).(\mathbf{Y}_{n,i} - \mathbf{Y}_i) \right],$$

$$c_{n,i}(\mathbf{u}) = \frac{1}{2h^2} \left[(d^2K)_h(\mathbf{u} - \mathbf{Y}_i).(\mathbf{Y}_{n,i} - \mathbf{Y}_i)^{(2)} - E(d^2K)_h(\mathbf{u} - \mathbf{Y}_i).(\mathbf{Y}_{n,i} - \mathbf{Y}_i)^{(2)} \right],$$

$$d_{n,i}(\mathbf{u}) = \frac{(-1)}{6h^3} \left[(d^3K)_h(\mathbf{u} - \mathbf{Y}_i).(\mathbf{Y}_{n,i} - \mathbf{Y}_i)^{(3)} - E(d^3K)_h(\mathbf{u} - \mathbf{Y}_i).(\mathbf{Y}_{n,i} - \mathbf{Y}_i)^{(3)} \right],$$

$$e_{n,i}(\mathbf{u}) = \frac{1}{24h^4} \left[(d^4K)_h(\mathbf{u} - \mathbf{Y}_{n,i}^*).(\mathbf{Y}_{n,i} - \mathbf{Y}_i)^{(4)} - E(d^4K)_h(\mathbf{u} - \mathbf{Y}_{n,i}^*).(\mathbf{Y}_{n,i} - \mathbf{Y}_i)^{(4)} \right],$$

for some $\mathbf{Y}_{n,i}^*$ that lies between $\mathbf{Y}_i$ and $\mathbf{Y}_{n,i}$ a.e. Most of the sums involving the previous terms will be negligible with respects to $1/(nh^{d/2})$.

B.1 Study of $J_{n,2}^*$

Now

$$J_{n,2}^* = \frac{2}{n^2} \sum_{i<j} \int \left[a_i + b_{n,i} + c_{n,i} + d_{n,i} + e_{n,i} \right] . \left[a_j + b_{n,j} + c_{n,j} + d_{n,j} + e_{n,j} \right] \omega$$

$$= \frac{2}{n^2} \sum_{i<j} \int a_i a_j \omega + \frac{2}{n^2} \sum_{i<j} \int (a_i b_{n,j} + a_j b_{n,i}) \omega + \dots$$

From Hall (1984), it is known that

$$\frac{nh^{d/2}}{2} \cdot \frac{1}{n^2} \sum_{i<j} \int a_i a_j \omega \xrightarrow{law} \frac{1}{2\sqrt{2}} \mathcal{N}(0, \sigma^2), \tag{B.3}$$

$$\sigma^2 = \int \tau^2 \omega \cdot \int \left[\int K(\mathbf{u}) K(\mathbf{u}+\mathbf{v})\, d\mathbf{u} \right]^2 d\mathbf{v}.$$

Therefore, the main term of $J^*_{n,2}$ seems to be of order $O(n^{-1}h^{-d/2})$. We will check it by studying the terms of the expansion of $J^*_{n,2}$ successively.

B.1.1 Study of $T_\alpha \equiv 2n^{-2}\sum_{i<j}\int a_i b_{n,j}\omega$

Note that the expectation of T_α is not zero, because some $\mathbf{Y}_i$ appears inside $b_{n,j}$, for every j. For convenience, set

$$b_{n,j}(\mathbf{u}) = \frac{(-1)}{nh}\sum_{k=1}^n b_{n,j,k}(\mathbf{u}), \text{ with}$$

$$b_{n,j,k}(\mathbf{u}) = (dK)_h(\mathbf{u}-\mathbf{Y}_j).(\mathbf{1}(\mathbf{Y}_k \le \mathbf{Y}_j) - \mathbf{Y}_j) - E\left[(dK)_h(\mathbf{u}-\mathbf{Y}_j).(\mathbf{1}(\mathbf{Y}_k \le \mathbf{Y}_j) - \mathbf{Y}_j)\right].$$

Moreover,

$$\begin{aligned} T_\alpha &= \left(\frac{-2}{n^3h}\right)\sum_{i<j}\sum_{k=1}^n \int a_i b_{n,j,k}\omega \\ &= \left(\frac{-2}{n^3h}\right)\left\{\sum_{i<j}\sum_{k\neq i,k\neq j}^n \int a_i b_{n,j,k}\omega + \sum_{i<j}\int a_i b_{n,j,i}\omega + \sum_{i<j}\int a_i b_{n,j,j}\omega\right\} \\ &\equiv T_\alpha^{(1)} + T_\alpha^{(2)} + T_\alpha^{(3)}. \end{aligned}$$

First, let us study $T_\alpha^{(3)}$. Its expectation is zero. Its variance is

$$\begin{aligned} E[(T_\alpha^{(3)})^2] &= \frac{4}{n^6h^2}\sum_{i_1<j_1}\sum_{i_2<j_2}\int E[a_{i_1}(\mathbf{u}_1)b_{n,j_1,j_1}(\mathbf{u}_1)a_{i_2}(\mathbf{u}_2)b_{n,j_2,j_2}(\mathbf{u}_2)]\omega(\mathbf{u}_1)\omega(\mathbf{u}_2)\,d\mathbf{u}_1\,d\mathbf{u}_2 \\ &= \frac{4}{n^6h^2}\left\{\sum_{i_1<j_1,i_2=i_1,j_2=j_1} + \sum_{i_1<j_1,i_2=j_1,j_2=i_1}\right\} \equiv V_{\alpha,1}^{(3)} + V_{\alpha,2}^{(3)}. \end{aligned}$$

The first of these terms is

$$\begin{aligned} V_{\alpha,1}^{(3)} &= \frac{4}{n^6h^2}\sum_{i<j}\int\left\{K_h(\mathbf{u}_1-\mathbf{y}_i) - \int K(\mathbf{v})\tau(\mathbf{u}_1-h\mathbf{v})\,d\mathbf{v}\right\}.\left\{K_h(\mathbf{u}_2-\mathbf{y}_i) - \int K(\mathbf{v})\tau(\mathbf{u}_2-h\mathbf{v})\,d\mathbf{v}\right\} \\ &\cdot \ \{(dK)_h(\mathbf{u}_1-\mathbf{y}_j).(1-\mathbf{y}_j) - E[(dK)_h(\mathbf{u}_1-\mathbf{Y}).(1-\mathbf{Y})]\}\cdot\{(dK)_h(\mathbf{u}_2-\mathbf{y}_j).(1-\mathbf{y}_j) \\ &- \ E[(dK)_h(\mathbf{u}_2-\mathbf{Y}).(1-\mathbf{Y})]\}\ \tau(\mathbf{y}_i)\tau(\mathbf{y}_j)\omega(\mathbf{u}_1)\omega(\mathbf{u}_2)\,d\mathbf{u}_1\,d\mathbf{u}_2\,d\mathbf{y}_i\,d\mathbf{y}_j. \end{aligned}$$

The "hardest" term among the latter ones is

$$\begin{aligned} &\frac{4}{n^6h^2}\sum_{i<j}\int K_h(\mathbf{u}_1-\mathbf{y}_i)K_h(\mathbf{u}_2-\mathbf{y}_i)(dK)_h(\mathbf{u}_1-\mathbf{y}_j).(1-\mathbf{y}_j).(dK)_h(\mathbf{u}_2-\mathbf{y}_j).(1-\mathbf{y}_j) \\ &\cdot \ \tau(\mathbf{y}_i)\tau(\mathbf{y}_j)\omega(\mathbf{u}_1)\omega(\mathbf{u}_2)\,d\mathbf{y}_i\,\mathbf{y}_j\,d\mathbf{u}_1\,d\mathbf{u}_2 \\ &= \ \frac{4}{n^6h^{2+d}}\sum_{i<j}\int K(\tilde{\mathbf{y}}_i)K(\tilde{\mathbf{u}}_2+\tilde{\mathbf{y}}_i)(dK)(\tilde{\mathbf{y}}_j).(1-\mathbf{u}_1+h\tilde{\mathbf{y}}_j) \\ &\cdot \ (dK)(\tilde{\mathbf{u}}_2+\tilde{\mathbf{y}}_j).(1-\mathbf{u}_1+h\tilde{\mathbf{y}}_j)\tau(\mathbf{u}_1-h\tilde{\mathbf{y}}_i)\tau(\mathbf{u}_1-h\tilde{\mathbf{y}}_j)\omega(\mathbf{u}_1)\omega(\mathbf{u}_1+h\tilde{\mathbf{u}}_2)\,d\tilde{\mathbf{y}}_i\,d\tilde{\mathbf{y}}_j\,d\mathbf{u}_1\,d\tilde{\mathbf{u}}_2. \end{aligned}$$

Since K is compactly supported, clearly, we can assume every variables belong to some compact real subset. Thus, the latter term is of order $n^{-4}h^{-2-d}$. It is $o(n^{-2}h^{-d})$ since nh tends to the infinity when n is large. The seven other terms of $V_{\alpha,1}^{(3)}$ can be dealt similarly. Actually, they are even of a weaker order (we win an extra factor h^d). Moreover,

$$V_{\alpha,2}^{(3)} = \frac{4}{n^6h^2}\sum_{i<j}\int\left\{K_h(\mathbf{u}_1-\mathbf{y}_i)-\int K(\mathbf{v})\tau(\mathbf{u}_1-h\mathbf{v})\,d\mathbf{v}\right\}$$
$$\cdot \quad \left\{K_h(\mathbf{u}_2-\mathbf{y}_j)-\int K(\mathbf{v})\tau(\mathbf{u}_2-h\mathbf{v})\,d\mathbf{v}\right\}$$
$$\cdot \quad \{(dK)_h(\mathbf{u}_1-\mathbf{y}_i).(1-\mathbf{y}_i)-E[(dK)_h(\mathbf{u}_1-\mathbf{Y}).(1-\mathbf{Y})]\}\cdot\{(dK)_h(\mathbf{u}_2-\mathbf{y}_j).(1-\mathbf{y}_j)$$
$$- \quad E[(dK)_h(\mathbf{u}_2-\mathbf{Y}).(1-\mathbf{Y})]\}\ \tau(\mathbf{y}_i)\tau(\mathbf{y}_j)\omega(\mathbf{u}_1)\omega(\mathbf{u}_2)\,d\mathbf{u}_1\,d\mathbf{u}_2\,d\mathbf{y}_i\,d\mathbf{y}_j.$$

Working exactly like $V_{\alpha,1}^{(3)}$, we can show $V_{\alpha,2}^{(3)} = O(n^{-4}h^{-2-d})$. Thus, we have proved that

$$T_\alpha^{(3)} = o_P\left(\frac{1}{nh^{d/2}}\right).$$

Second, let us study $T_\alpha^{(2)}$. Recall that

$$T_\alpha^{(2)} = \left(\frac{-2}{n^3h}\right)\sum_{i<j}\int a_i(\mathbf{u})(dK)_h(\mathbf{u}-\mathbf{Y}_j).(\mathbf{1}(\mathbf{Y}_i\leq\mathbf{Y}_j)-\mathbf{Y}_j)\omega(\mathbf{u})\,d\mathbf{u}.$$

The expectation of this term is not zero. By applying equation (A.1), we obtain

$$E[T_\alpha^{(2)}] = \frac{(-1)}{nh}(1-\frac{1}{n})\int E\Big[(K_h(\mathbf{u}-\mathbf{Y}_1)-E[K_h(\mathbf{u}-\mathbf{Y})])$$
$$\cdot \quad \left(\tau(\mathbf{u})\sum_{r=1}^d K_r\left(\frac{u_r-Y_{1,r}}{h}\right)+O(h)\phi(\mathbf{u})\right)\Big]\omega(\mathbf{u})\,du$$
$$= \quad \frac{(-1)}{nh}(1-\frac{1}{n})\sum_{r=1}^d\left\{\int(K_h(\mathbf{u}-\mathbf{y})-E[K_h(\mathbf{u}-\mathbf{Y})])\,K_r\left(\frac{u_r-y_r}{h}\right)\right.$$
$$\cdot \quad \tau(\mathbf{y})\tau(\mathbf{u})\omega(\mathbf{u})\,d\mathbf{u}\,d\mathbf{y}+O(h)\} = \frac{(-1)}{nh}\sum_{r=1}^d\int K_r^2.\int\tau^2\omega+O(n^{-1}).$$

Note we have used the fact that the density of Y_r is uniform on $[0,1]$.

The order of the expectation of $T_\alpha^{(2)}$ is then $(nh)^{-1}$. Unfortunately, it is not $o(1/nh^{d/2})$ when $d=2$. Nonetheless, its variance will be small enough so that we can consider this term is reduced to its expectation. Indeed,

$$
\begin{aligned}
Var(T_\alpha^{(2)}) &= \frac{4}{n^6h^2} \sum_{i_1<j_1,i_2<j_2} \int E\Big[a_{i_1}(\mathbf{u}_1).(dK)_h(\mathbf{u}_1-\mathbf{Y}_{j_1}).(\mathbf{1}(\mathbf{Y}_{i_1}\leq \mathbf{Y}_{j_1})-\mathbf{Y}_{j_1}) \\
&\cdot\quad a_{i_2}(\mathbf{u}_2).(dK)_h(\mathbf{u}_2-\mathbf{Y}_{j_2}).(\mathbf{1}(\mathbf{Y}_{i_2}\leq \mathbf{Y}_{j_2})-\mathbf{Y}_{j_2}) \\
&-\quad E[a_{i_1}(\mathbf{u}_1)b_{n,j_1,i_1}(\mathbf{u}_1)].E[a_{i_2}(\mathbf{u}_2)b_{n,j_2,i_2}(\mathbf{u}_2)]\Big]\omega(\mathbf{u}_1)\omega(\mathbf{u}_2)\,d\mathbf{u}_1\,d\mathbf{u}_2 \\
&= \frac{4}{n^6h^2}\left\{\sum_{i_1<j_1,i_2<j_2,i_1=i_2} + \sum_{i_1<j_1,i_2<j_2,i_1=j_2} + \sum_{i_1<j_1,i_2<j_2,j_1=i_2} + \sum_{i_1<j_1,i_2<j_2,j_1=j_2}\right\} \\
&\equiv V_{\alpha,1}^{(2)} + V_{\alpha,2}^{(2)} + V_{\alpha,3}^{(2)} + V_{\alpha,4}^{(2)}.
\end{aligned}
$$

Let us study the first of the previous terms.

$$
\begin{aligned}
V_{\alpha,1}^{(2)} &= \frac{4}{n^6h^2}\sum_{i_1<j_1,i_1<j_2}\int\left\{K_h(\mathbf{u}_1-\mathbf{y}_{i_1})-\int K(\mathbf{t})\tau(\mathbf{u}_1-h\mathbf{t})\,d\mathbf{t}\right\} \\
&\cdot\ \left\{K_h(\mathbf{u}_2-\mathbf{y}_{i_1})-\int K(\mathbf{t})\tau(\mathbf{u}_2-h\mathbf{t})\,d\mathbf{t}\right\}.\{(dK)(\tilde{\mathbf{y}}_{j_1}).(\mathbf{1}(\mathbf{y}_{i_1}\leq \mathbf{u}_1-h\tilde{\mathbf{y}}_{j_1})-(\mathbf{u}_1-h\tilde{\mathbf{y}}_{j_1}))\} \\
&\cdot\ \{(dK)(\tilde{\mathbf{y}}_{j_2}).(\mathbf{1}(\mathbf{y}_{i_1}\leq \mathbf{u}_2-h\tilde{\mathbf{y}}_{j_2})-(\mathbf{u}_2-h\tilde{\mathbf{y}}_{j_2}))\}\,\tau(\mathbf{y}_{i_1})\tau(\mathbf{u}_1-h\tilde{\mathbf{y}}_{j_1})\tau(\mathbf{u}_2-h\tilde{\mathbf{y}}_{j_2}) \\
&\cdot\ \omega(\mathbf{u}_1)\omega(\mathbf{u}_2)\,d\mathbf{u}_1\,d\mathbf{u}_2\,d\tilde{\mathbf{y}}_{j_1}\,d\tilde{\mathbf{y}}_{j_2} + O\left(\frac{1}{n^6h^2}\cdot\frac{n^2}{h^d}\right).
\end{aligned}
$$

The remainder term corresponds to the case $i_1=i_2$, $j_1=j_2$. The main previous term of $V_{\alpha,1}^{(2)}$ can be expressed as a sum of four terms. The first one involves the factor $K_h(\mathbf{u}_1-\mathbf{y}_{i_1}).K_h(\mathbf{u}_2-\mathbf{y}_{i_1})$. The second (resp. the third) one involves the factor $K_h(\mathbf{u}_1-\mathbf{y}_{i_1})$ (resp. $K_h(\mathbf{u}_2-\mathbf{y}_{i_1})$) only. The last one has no such factor (viz no more denominators h^{-d}). If necessary, we can set one or two changes of variables among $\tilde{\mathbf{y}}_{i_1}=(\mathbf{u}_1-\mathbf{y}_{i_1})/h$, $\tilde{\mathbf{y}}_{i_1}=(\mathbf{u}_2-\mathbf{y}_{i_1})/h$ or $\tilde{\mathbf{u}}_2=(\mathbf{u}_2-\mathbf{u}_1)/h$. It allows to clear all the factors h^{-d}. Thus we get easily,

$$
V_{\alpha,1}^{(2)} = O\left(\frac{1}{n^6h^2}\cdot n^3\right)+O\left(\frac{1}{n^4h^{2+d}}\right)=o\left(\frac{1}{n^2h^d}\right), \tag{B.4}
$$

since nh tends to the infinity. The three other terms $V_{\alpha,l}^{(2)}$, $l=2,3,4$ can be dealt similarly, because there exist always four free variables ($\mathbf{u}_1,\mathbf{u}_2$ and three ones among i_1,i_2,j_1,j_2) that can be used for some change of variables. Like previously, all the factors h^{-d} disappear. To conclude, $V_\alpha^{(2)}=O(1/(n^3h^2)+1/(n^4h^{2+d}))$, and

$$
T_\alpha^{(2)} = ET_\alpha^{(2)} + o_P\left(\frac{1}{nh^{d/2}}\right) = \frac{(-1)}{nh}\int\tau^2\omega.\sum_{r=1}^d\int K_r^2 + o_P\left(\frac{1}{nh^{d/2}}\right).
$$

Now, let us deal with $T_\alpha^{(1)}$. Recall that

$$
T_\alpha^{(1)} = \frac{(-2)}{n^3h}\sum_{i<j}\sum_{k,k\neq i,k\neq j}\int a_ib_{njk}\omega.
$$

Clearly, $T_\alpha^{(1)}$ is centered. Moreover, its variance is

$$E[\left(T_\alpha^{(1)}\right)^2] = \frac{4}{n^6h^2} \sum_{i_1<j_1} \sum_{i_2<j_2} \sum_{k_1\neq i_1,j_1} \sum_{k_2\neq i_2,j_2}$$
$$E\int (a_{i_1}b_{n,j_1k_1})(\mathbf{u}_1)\cdot(a_{i_2}b_{n,j_2k_2})(\mathbf{u}_2)\omega(\mathbf{u}_1)\omega(\mathbf{u}_2)\,d\mathbf{u}_1\,d\mathbf{u}_2.$$

A lot of the latter terms are zero. The only nonzero terms appear in the following cases :

- $k_1 = i_2$ and $k_2 = i_1$,
- $k_1 = k_2$ and $i_1 = i_2$,
- $k_1 = i_2$, $k_2 = j_1$ and $i_1 = j_2$,
- $k_1 = j_2$, $k_2 = i_1$ and $i_2 = j_1$,
- $k_1 = j_2$, $k_2 = j_1$ and $i_1 = i_2$,
- $k_1 = k_2$, $i_1 = j_2$ and $i_2 = j_1$.

Thus, the variance of $T_\alpha^{(1)}$ is the sum of six terms, denoted by $V_{\alpha,l}^{(1)}$, $l = 1,\ldots,6$. Assuming that there are no other equalities except $k_1 = i_2$ and $k_2 = i_1$, the first variance term is

$$V_{\alpha,1}^{(1)} = \frac{4}{n^4h^2} \sum_{i_1,i_2,i_1\neq i_2} \int$$
$$\left\{K_h(\mathbf{u}_1 - \mathbf{y}_{i_1}) - \int K(\mathbf{t})\tau(\mathbf{u}_1 - h\mathbf{t})\,d\mathbf{t}\right\}.\left\{K_h(\mathbf{u}_2 - \mathbf{y}_{i_2}) - \int K(\mathbf{t})\tau(\mathbf{u}_2 - h\mathbf{t})\,d\mathbf{t}\right\}$$
$$\cdot \quad \left\{\tau(\mathbf{u}_1)\sum_{r=1}^d K_r\left(\frac{u_{1r} - y_{i_2r}}{h}\right) + O(h)\phi(\mathbf{u}_1)\right\}\cdot\left\{\tau(\mathbf{u}_2)\sum_{s=1}^d K_r\left(\frac{u_{2s} - y_{i_1s}}{h}\right) + O(h)\phi(\mathbf{u}_2)\right\}$$
$$\cdot \quad \tau(\mathbf{y}_{i_2})\tau(\mathbf{y}_{i_1})\omega(\mathbf{u}_1)\omega(\mathbf{u}_2)\,d\mathbf{y}_{i_1}\,d\mathbf{y}_{i_2}\,d\mathbf{u}_1\,d\mathbf{u}_2.$$

This sum can be split into 16 other terms. The main one is

$$\frac{4}{n^4h^2}\sum_{r,s}\sum_{i_1,i_2}\int K_h(\mathbf{u}_1 - \mathbf{y}_{i_1})K_h(\mathbf{u}_2 - \mathbf{y}_{i_2})\tau(\mathbf{u}_1)K_r\left(\frac{u_{1r} - y_{i_2r}}{h}\right)$$
$$\cdot \quad \tau(\mathbf{u}_2)K_r\left(\frac{u_{2s} - y_{i_1s}}{h}\right)\tau(\mathbf{y}_{i_2})\tau(\mathbf{y}_{i_1})\omega(\mathbf{u}_1)\omega(\mathbf{u}_2)\,d\mathbf{y}_{i_1}\,d\mathbf{y}_{i_2}\,d\mathbf{u}_1\,d\mathbf{u}_2$$
$$= \quad \frac{4}{n^4h^2}\sum_{r,s}\sum_{i_1,i_2}\int K(\tilde{\mathbf{y}}_{i_1})K(\tilde{\mathbf{y}}_{i_2})\tau(\mathbf{u}_1)K_r\left(\frac{u_{1r} - u_{2r} + h\tilde{y}_{i_2r}}{h}\right)$$
$$\cdot \quad \tau(\mathbf{u}_2)K_r\left(\frac{u_{2s} - u_{1s} + h\tilde{y}_{i_1s}}{h}\right)\tau(\mathbf{u}_2 - h\tilde{\mathbf{y}}_{i_2})\tau(\mathbf{u}_1 - h\tilde{\mathbf{y}}_{i_1})\omega(\mathbf{u}_1)\omega(\mathbf{u}_2)\,d\tilde{\mathbf{y}}_{i_1}\,d\tilde{\mathbf{y}}_{i_2}\,d\mathbf{u}_1\,d\mathbf{u}_2$$

If $r \neq s$, set the change of variables $\tilde{u}_{1r} = (u_{1r} - u_{2r})/h$ and $\tilde{u}_{1s} = (u_{1s} - u_{2s})/h$ to get an extra factor h^2. If $r = s$, we obtain only one factor h. Thus, the previous variance term is $O(n^{-4}h^{-2}.n^2.h) = O(n^{-2}h^{-1})$. This is $o(n^{-2}h^{-d})$.

Imagine we have some other equalities between the indices i_1, i_2, j_1, j_2, k_1 and k_2 in $V_\alpha^{(1)}$. For instance $j_1 = j_2$. This would not be a problem because we gain a factor n and we can always remove the annoying factor h^{-d} by some change of variables with respects to $\mathbf{u}_1$, $\mathbf{u}_2$ and the variables $\mathbf{y}$. Thus, we get the order $O(n^{-6}h^{-2}.n^3) = o(n^{-2}h^{-d})$.

The 15 other terms that are coming from the expansion of $V_{\alpha,1}^{(1)}$ can be dealt similarly. Thus, $V_{\alpha,1}^{(1)} = o(n^{-2}h^{-d})$.

Another critical term should be

$$V_{\alpha,2}^{(1)} = \frac{4}{n^6h^2} \sum_{i<j_1} \sum_{i<j_2} \sum_{k,k\neq i,j_1,j_2} \int E\left[(a_i b_{n,j_1,k})(\mathbf{u}_1)(a_i b_{n,j_2,k})(\mathbf{u}_2)\right] \omega(\mathbf{u}_1)\omega(\mathbf{u}_2)\, d\mathbf{u}_1\, d\mathbf{u}_2.$$

Since k is different from all other indices, this equals

$$\begin{aligned}
&\frac{4}{n^6h^2} \sum_{i<j_1} \sum_{i<j_2} \sum_{k} \int \\
&\quad \left\{K_h(\mathbf{u}_1 - \mathbf{y}_i) - \int K(\mathbf{t})\tau(\mathbf{u}_1 - h\mathbf{t})\, d\mathbf{t}\right\} . \left\{K_h(\mathbf{u}_2 - \mathbf{y}_i) - \int K(\mathbf{t})\tau(\mathbf{u}_2 - h\mathbf{t})\, d\mathbf{t}\right\} \\
\cdot &\quad (dK)_h(\mathbf{u}_1 - \mathbf{y}_{j_1}).(\mathbf{1}(\mathbf{y}_k \leq \mathbf{y}_{j_1}) - \mathbf{y}_{j_1}) \cdot (dK)_h(\mathbf{u}_2 - \mathbf{y}_{j_2}).(\mathbf{1}(\mathbf{y}_k \leq \mathbf{y}_{j_2}) - \mathbf{y}_{j_2}) \\
\cdot &\quad \tau(\mathbf{y}_i)\tau(\mathbf{y}_{j_2})\tau(\mathbf{y}_{j_1})\tau(\mathbf{y}_k)\omega(\mathbf{u}_1)\omega(\mathbf{u}_2)\, d\mathbf{y}_i\, d\mathbf{y}_{j_1}\, d\mathbf{y}_{j_2}\, d\mathbf{y}_k\, d\mathbf{u}_1\, d\mathbf{u}_2 \\
= &\quad \frac{4}{n^4h^2} \sum_i \sum_k \int \left\{K(\tilde{\mathbf{y}}_i) - h^d \int K(\mathbf{t})\tau(\mathbf{u}_1 - h\mathbf{t})\, d\mathbf{t}\right\} . \left\{K_h(\mathbf{u}_2 - \mathbf{u}_1 + h\tilde{\mathbf{y}}_i) - \int K(\mathbf{t})\tau(\mathbf{u}_2 - h\mathbf{t})\, d\mathbf{t}\right\} \\
\cdot &\quad \left\{\tau(\mathbf{u}_1) \sum_{r=1}^d K_r\left(\frac{u_{1r} - y_{k,r}}{h}\right) + O(h)\phi(\mathbf{u}_1)\right\} \cdot \left\{\tau(\mathbf{u}_2) \sum_{s=1}^d K_s\left(\frac{u_{2s} - y_{k,s}}{h}\right) + O(h)\phi(\mathbf{u}_2)\right\} \\
\cdot &\quad \tau(\mathbf{u}_1 - h\tilde{\mathbf{y}}_i)\tau(\mathbf{y}_k)\omega(\mathbf{u}_1)\omega(\mathbf{u}_2)\, d\tilde{\mathbf{y}}_i\, d\mathbf{y}_k\, d\mathbf{u}_1\, d\mathbf{u}_2.
\end{aligned}$$

We have assumed there are no additional equalities between i, j_1, j_2. By setting $h\tilde{\mathbf{u}}_2 = \mathbf{u}_2 - \mathbf{u}_1$, we remove the factor h^{-d}. Moreover, by setting $h\tilde{u}_{1r} = u_{1r} - y_{kr}$, we get an extra factor h. Thus, the term if of order $O(n^{-4}h^{-2}.n^2.h = o(n^{-2}h^{-d})$. When there are some other equalities between the other indices i, j_1 and j_2, we gain a factor n even if we lose eventually a factor h^d. In every cases, the order of these terms is lower than $n^{-2}h^{-d}$. Therefore, $V_{\alpha,2}^{(1)} = o(n^{-2}h^{-d})$.

All the other terms $V_{\alpha,l}^{(1)}$, $l = 3, \ldots, 6$ are simpler. Indeed, with respects to $V_{\alpha,1}^{(1)}$, there is an additional equality between the indices. At the opposite, it should be harder to remove all the four terms h^{-d}. Actually, it can be done at least three times over four, because there are always two free variable $\mathbf{y}$ (at least), and we have $\mathbf{u}_1$ or $\mathbf{u}_2$ at our disposal too. Thus, all these terms are $O(n^{-6}h^{-2}.n^3h^{-d}) = o(n^{-2}h^{-d})$ since nh^2 tends to the infinity.

Therefore, the variance of $T_\alpha^{(1)}$ is negligible with respects to $n^{-2}h^{-d}$ and $T_\alpha^{(1)} = o_P(1/(nh^{d/2}))$. To conclude,

$$T_\alpha = \frac{(-1)}{nh}\int \tau^2\omega.\sum_{r=1}^d \int K_r^2 + o_P\left(\frac{1}{nh^{d/2}}\right). \tag{B.5}$$

B.1.2 Study of $T_\beta = 2n^{-2}\sum_{i<j} b_{n,i}b_{n,j}\omega$

Note that

$$T_\beta = \frac{2}{n^4h^2}\sum_{i<j}\sum_{k,k'}\int \{(dK)_h(\mathbf{u}-\mathbf{Y}_i).(\mathbf{1}(\mathbf{Y}_k \leq \mathbf{Y}_i) - \mathbf{Y}_i) - E[(dK)_h(\mathbf{u}-\mathbf{Y}_i).(\mathbf{1}(\mathbf{Y}_k \leq \mathbf{Y}_i) - \mathbf{Y}_i)]\}$$
$$\cdot \quad \{(dK)_h(\mathbf{u}-\mathbf{Y}_j).(\mathbf{1}(\mathbf{Y}_{k'} \leq \mathbf{Y}_j) - \mathbf{Y}_j) - E[(dK)_h(\mathbf{u}-\mathbf{Y}_j).(\mathbf{1}(\mathbf{Y}_{k'} \leq \mathbf{Y}_j) - \mathbf{Y}_j)]\}\,\omega(\mathbf{u})\,d\mathbf{u}.$$

The latter term needs to be considered with respects to the potential number of equalities between the indices i, j, k, k'.

No equalities between i, j, k, k' : $T_\beta^{(1)}$

Thus, the expectation of the corresponding term is zero. Moreover, its variance is

$$\frac{4}{n^8h^4}\sum_{i_1<j_1}\sum_{i_2<j_2}\sum_{k_1\neq k_1'\neq i_1,j_1}\sum_{k_2\neq k_2'\neq i_2,j_2} E\int (dK)_h(\mathbf{u}_1-\mathbf{Y}_{i_1}).(\mathbf{1}(\mathbf{Y}_{k_1} \leq \mathbf{Y}_{i_1}) - \mathbf{Y}_{i_1})$$
$$\cdot \quad (dK)_h(\mathbf{u}_1-\mathbf{Y}_{j_1}).(\mathbf{1}(\mathbf{Y}_{k_1'} \leq \mathbf{Y}_{j_1}) - \mathbf{Y}_{j_1}).(dK)_h(\mathbf{u}_2-\mathbf{Y}_{i_2}).(\mathbf{1}(\mathbf{Y}_{k_2} \leq \mathbf{Y}_{i_2}) - \mathbf{Y}_{i_2})$$
$$\cdot \quad (dK)_h(\mathbf{u}_2-\mathbf{Y}_{j_2}).(\mathbf{1}(\mathbf{Y}_{k_2'} \leq \mathbf{Y}_{j_2}) - \mathbf{Y}_{j_2})\omega(\mathbf{u}_1)\omega(\mathbf{u}_2)\,d\mathbf{u}_1\,d\mathbf{u}_2.$$

The expectations are zero, except if there are some equalities between our eight indices. More precisely, the equalities have to concern all the indices k_1, k_1', k_2, k_2', otherwise the corresponding term is zero. This provides the following cases :

- $k_1 = k_2$ and $k_1' = k_2'$,
- $k_1 = k_2'$ and $k_1' = k_2$,
- $k_1 = i_2$, $k_2 = i_1$, $k_1' = k_2'$, or their variations,
- $k_1 = i_2$, $k_1' = j_2$, $k_2 = i_1$, $k_2' = j_1$, or their variations.

The corresponding variances are called $V_{\beta,j}^{(1)}$, $j = 1, \ldots, 4$. Let us deal with the first configuration. It provides the "variance-type" term

$$
\begin{aligned}
&\frac{4}{n^8h^4}\sum_{i_1<j_1}\sum_{i_2<j_2}\sum_{k\neq k'} E\int (dK)_h(\mathbf{u}_1-\mathbf{Y}_{i_1}).(\mathbf{1}(\mathbf{Y}_k\leq \mathbf{Y}_{i_1})-\mathbf{Y}_{i_1})(dK)_h(\mathbf{u}_1-\mathbf{Y}_{j_1})\\
&\cdot\quad (\mathbf{1}(\mathbf{Y}_{k'}\leq \mathbf{Y}_{j_1})-\mathbf{Y}_{j_1})(dK)_h(\mathbf{u}_2-\mathbf{Y}_{i_2}).(\mathbf{1}(\mathbf{Y}_k\leq \mathbf{Y}_{i_2})-\mathbf{Y}_{i_2})\\
&\quad\ (dK)_h(\mathbf{u}_2-\mathbf{Y}_{j_2}).(\mathbf{1}(\mathbf{Y}_{k'}\leq \mathbf{Y}_{j_2})-\mathbf{Y}_{j_2})\omega(\mathbf{u}_1)\omega(\mathbf{u}_2)\,d\mathbf{u}_1\,d\mathbf{u}_2\\
&=\ \frac{4}{n^8h^4}\sum_{i_1<j_1}\sum_{i_2<j_2}\sum_{k\neq k'} E\int \left\{\tau(\mathbf{u}_1)\sum_{r=1}^d K_r\left(\frac{u_{1r}-Y_{kr}}{h}\right)+O(h)\phi(\mathbf{u}_1)\right\}\\
&\cdot\quad \left\{\tau(\mathbf{u}_1)\sum_{r'=1}^d K_{r'}\left(\frac{u_{1r'}-Y_{k'r'}}{h}\right)+O(h)\phi(\mathbf{u}_1)\right\}\cdot\left\{\tau(\mathbf{u}_2)\sum_{s=1}^d K_s\left(\frac{u_{2s}-Y_{ks}}{h}\right)+O(h)\phi(\mathbf{u}_2)\right\}\\
&\cdot\quad \left\{\tau(\mathbf{u}_2)\sum_{s'=1}^d K_{s'}\left(\frac{u_{1s'}-Y_{k's'}}{h}\right)+O(h)\phi(\mathbf{u}_2)\right\}\omega(\mathbf{u}_1)\omega(\mathbf{u}_2)\,d\mathbf{u}_1\,d\mathbf{u}_2.
\end{aligned}
$$

The main member of the previous expansion is

$$
\begin{aligned}
&\frac{4}{n^8h^4}\sum_{i_1<j_1}\sum_{i_2<j_2}\sum_{k\neq k'}\sum_{r,r',s,s'} E\int \tau^2(\mathbf{u}_1)K_r\left(\frac{u_{1r}-Y_{kr}}{h}\right)K_{r'}\left(\frac{u_{1r'}-Y_{k'r'}}{h}\right)\\
&\cdot\quad \tau^2(\mathbf{u}_2)K_s\left(\frac{u_{2s}-Y_{ks}}{h}\right)K_{s'}\left(\frac{u_{2s'}-Y_{k's'}}{h}\right)\omega(\mathbf{u}_1)\omega(\mathbf{u}_2)\,d\mathbf{u}_1\,d\mathbf{u}_2
\end{aligned}
$$

The "worse" situation occurs when $r=s$ and $r'=s'$. In this case, we get

$$
\begin{aligned}
&\frac{4}{n^8h^4}\sum_{i_1<j_1}\sum_{i_2<j_2}\sum_{k\neq k'}\int \tau^2(\mathbf{u}_1)K_r\left(\frac{u_{1r}-y_{kr}}{h}\right)K_{r'}\left(\frac{u_{1r'}-y_{k'r'}}{h}\right)\tau^2(\mathbf{u}_2)\\
&\cdot\quad K_r\left(\frac{u_{2r}-y_{kr}}{h}\right)K_{r'}\left(\frac{u_{2r'}-y_{k'r'}}{h}\right)\tau_r(y_{kr})\tau_{r'}(y_{k'r'})\omega(\mathbf{u}_1)\omega(\mathbf{u}_2)\,d\mathbf{u}_1\,d\mathbf{u}_2\,dy_{kr}\,dy_{k'r'}\\
&=\ \frac{4}{n^8h^2}\sum_{i_1<j_1}\sum_{i_2<j_2}\sum_{k\neq k'}\int \tau^2(\mathbf{u}_1)K_r(\tilde{y}_{kr})K_{r'}(\tilde{y}_{k'r'})\tau^2(\mathbf{u}_2)K_r\left(\frac{u_{2r}-u_{1r}+h\tilde{y}_{kr}}{h}\right)\\
&\cdot\quad K_{s'}\left(\frac{u_{2r'}-u_{1r'}+h\tilde{y}_{k'r'}}{h}\right)\tau_r(u_{1r}-h\tilde{y}_{kr})\tau_{r'}(u_{1r'}-h\tilde{y}_{k'r'})\omega(\mathbf{u}_1)\omega(\mathbf{u}_2)\,d\mathbf{u}_1\,d\mathbf{u}_2\,d\tilde{y}_{kr}\,d\tilde{y}_{k'r'}.
\end{aligned}
$$

By setting $h\tilde{u}_{2r}=u_{2r}-u_{1r}$, we get an extra factor h. The previous variance term is then $O(n^{-8}h^{-1}.n^6)=o(n^2h^{-d})$. Thus, $V^{(1)}_{\beta,1}=o(n^{-2}h^{-d})$.

The variance term $V^{(1)}_{\beta,2}$ corresponding to the case $k_1=k_2'$ and $k_1'=k_2$ can be dealt exactly as $V^{(1)}_{\beta,1}$. The third one, $V^{(1)}_{\beta,3}$, is

$$
\begin{aligned}
&\frac{4}{n^8h^4}\sum_{i_1<j_1}\sum_{i_2<j_2}\sum_{k} E\int (dK)_h(\mathbf{u}_1-\mathbf{Y}_{i_1}).(\mathbf{1}(\mathbf{Y}_{i_2}\leq \mathbf{Y}_{i_1})-\mathbf{Y}_{i_1})(dK)_h(\mathbf{u}_1-\mathbf{Y}_{j_1})\\
&\cdot\quad (\mathbf{1}(\mathbf{Y}_{k}\leq \mathbf{Y}_{j_1})-\mathbf{Y}_{j_1})(dK)_h(\mathbf{u}_2-\mathbf{Y}_{i_2}).(\mathbf{1}(\mathbf{Y}_{i_1}\leq \mathbf{Y}_{i_2})-\mathbf{Y}_{i_2})\\
&\cdot\quad (dK)_h(\mathbf{u}_2-\mathbf{Y}_{j_2}).(\mathbf{1}(\mathbf{Y}_{k}\leq \mathbf{Y}_{j_2})-\mathbf{Y}_{j_2})\omega(\mathbf{u}_1)\omega(\mathbf{u}_2)\,d\mathbf{u}_1\,d\mathbf{u}_2.
\end{aligned}
$$

It can be bounded easily : $V^{(1)}_{\beta,3}=O(1/(n^8h^4)\cdot n^5)=o(1/(n^2h^d))$, since nh^2 tends to the infinity when n is large.

$V^{(1)}_{\beta,4}$ and the other variance terms that are obtained by adding some equalities between the indices can be dealt similarly. All of them provide negligible terms. To conclude,

$$T^{(1)}_\beta = o_P\left(\frac{1}{nh^{d/2}}\right).$$

Only the equality $k=k'$: $T^{(2)}_\beta$

We get

$$T^{(2)}_\beta = \frac{2}{n^4h^2}\sum_{i<j}\sum_{k\neq i,j}\int (dK)_h(\mathbf{u}-\mathbf{Y}_i).(\mathbf{1}(\mathbf{Y}_k\leq \mathbf{Y}_i)-\mathbf{Y}_i)(dK)_h(\mathbf{u}-\mathbf{Y}_j).(\mathbf{1}(\mathbf{Y}_k\leq \mathbf{Y}_j)-\mathbf{Y}_j)\omega(\mathbf{u})\,d\mathbf{u}.$$

Its expectation is nonzero. More precisely,

$$\begin{aligned} ET^{(2)}_\beta &= \frac{2}{n^4h^2}\sum_{i<j}\sum_{k\neq i,j} E\int\left\{\tau(\mathbf{u})\sum_{r=1}^d K_r\left(\frac{u_r-Y_{kr}}{h}\right)+O(h)\phi(\mathbf{u})\right\} \\ &\cdot\ \left\{\tau(\mathbf{u})\sum_{s=1}^d K_s\left(\frac{u_s-Y_{ks}}{h}\right)+O(h)\phi(\mathbf{u})\right\}\omega(\mathbf{u})\,d\mathbf{u} \\ &= \frac{2}{n^4h^2}\left\{\sum_{r\neq s}\sum_{i<j}\sum_{k\neq i,j}+\sum_{r=s}\sum_{i<j}\sum_{k\neq i,j}\right\}+O(n^{-2}) \\ &\equiv E^{(2)}_{\beta,1}+E^{(2)}_{\beta,2}+O(n^{-2}). \end{aligned}$$

By setting $h\tilde{y}_{kr}=u_r-y_{kr}$ and $h\tilde{y}_{ks}=u_s-y_{ks}$, we get easily

$$E^{(2)}_{\beta,1} = O\left(\frac{1}{n^4h^2}\cdot n^3\cdot h^2\right) = O(n^{-2}).$$

Concerning $E^{(2)}_{\beta,1}$, one change of variables only is possible. It provides

$$\begin{aligned} E^{(2)}_{\beta,2} &= \frac{2}{n^4h^2}\sum_{i<j}\sum_{k\neq i,j}\sum_{r=1}^d\int \tau^2(\mathbf{u})K_r^2\left(\frac{u_r-y_{kr}}{h}\right)\omega(\mathbf{u})\mathbf{1}(y_{kr}\in[0,1])\,d\mathbf{u}\,dy_{kr} \\ &= \frac{2}{n^4h}\cdot\frac{n(n-1)}{2}\cdot(n-2)\left\{\sum_{r=1}^d\int K_r^2\cdot\int\tau^2\omega\right\} \\ &= \frac{1}{nh}\sum_r\int K_r^2\int\tau^2\omega+o(\frac{1}{nh}), \end{aligned}$$

for n sufficiently large. Therefore, the expectation of $T^{(2)}_\beta$ is not $o(n^{-1}h^{-d/2})$ (in the case $d=2$). Let us deal now with its variance. To lighten the notations, we set

$$e_0(\mathbf{u}) = E\left[(dK)_h(\mathbf{u}-\mathbf{Y}_1).(\mathbf{1}(\mathbf{Y}_3\leq\mathbf{Y}_1)-\mathbf{Y}_1).(dK)_h(\mathbf{u}-\mathbf{Y}_2).(\mathbf{1}(\mathbf{Y}_3\leq\mathbf{Y}_2)-\mathbf{Y}_2)\right].$$

Therefore,

$$
\begin{aligned}
Var(T_\beta^{(2)}) &= \frac{4}{n^8h^4}\sum_{i_1<j_1,i_2<j_2}\sum_{k_1\neq i_1,j_1}\sum_{k_2\neq i_2,j_2} E\int \{(dK)_h(\mathbf{u}_1-\mathbf{Y}_{i_1}).(\mathbf{1}(\mathbf{Y}_{k_1}\leq \mathbf{Y}_{i_1})-\mathbf{Y}_{i_1})\\
&\cdot\quad (dK)_h(\mathbf{u}_1-\mathbf{Y}_{j_1}).(\mathbf{1}(\mathbf{Y}_{k_1}\leq \mathbf{Y}_{j_1})-\mathbf{Y}_{j_1})-e_0(\mathbf{u}_1)\}\,.\{(dK)_h(\mathbf{u}_2-\mathbf{Y}_{i_2}).(\mathbf{1}(\mathbf{Y}_{k_2}\leq \mathbf{Y}_{i_2})-\mathbf{Y}_{i_2})\\
&\cdot\quad (dK)_h(\mathbf{u}_2-\mathbf{Y}_{j_2}).(\mathbf{1}(\mathbf{Y}_{k_2}\leq \mathbf{Y}_{j_2})-\mathbf{Y}_{j_2})-e_0(\mathbf{u}_2)\})\omega(\mathbf{u}_1)\omega(\mathbf{u}_2)\,d\mathbf{u}_1\,d\mathbf{u}_2.
\end{aligned}
$$

When there are no equalities between the indices $i_1, j_1, k_1, i_2, j_2, k_2$, the corresponding expectation is zero. At the opposite, there could be one, two or three equalities between them. In every case, it is always possible to make some changes of variables with respects to $\mathbf{y}_{i_1}$ and $\mathbf{y}_{j_1}$. Moreover, it is possible to set $h\tilde{\mathbf{u}}_2 = \mathbf{u}_2 - \mathbf{u}_1$, as previously. Thus, it is easy to verify that

$$
Var(T_\beta^{(2)}) = O\left(\frac{1}{n^8h^4}\cdot(n^5+n^4h^{-d})\right) = o(n^{-2}h^{-d}).
$$

Thus,

$$
T_\beta^{(2)} = \frac{1}{nh}\sum_r \int K_r^2 \int \tau^2\omega + o_P(n^{-1}h^{-d/2}).
$$

Only the equality $k=i$ or j (or $k'=i$ or j) : $T_\beta^{(3)}$

The expectation is zero and the variance can be dealt exactly as in the latter case.

Two equalities, or more, between the indices : $T_\beta^{(4)}$

To fix the ideas, imagine there are two equalities between our four indices. It means $i=k$ and $j=k'$, or the reverse. It is obvious to bound the expectation of $T_\beta^{(4)}$ by $O(n^{-4}h^{-2}.n^2) = o(n^{-1}h^{-d/2})$. Moreover, the variance is clearly $O(n^{-8}h^{-4}.n^4.h^{-d})$, by the same calculations as previously. Thus, $T_\beta^{(4)}$ is negligible with respects to $n^{-1}h^{-d/2}$, in probability.

To conclude,

$$
T_\beta = \frac{1}{nh}\sum_r \int K_r^2 \int \tau^2\omega + o_P(n^{-1}h^{-d/2}). \tag{B.6}
$$

B.1.3 Study of $2n^{-2}\sum_{i<j} a_i c_{n,j}\omega$ and $2n^{-2}\sum_{i<j} a_i d_{n,j}\omega$

To deal with these two terms simultaneously, denote

$$
\begin{aligned}
T_{\gamma,m} &= \frac{2}{n^{2+m}h^m}\sum_{i<j}\sum_{k_1,\ldots,k_m}\int \{K_h(\mathbf{u}-\mathbf{Y}_i) - EK_h(\mathbf{u}-\mathbf{Y}_i)\}\\
&\cdot\quad \{(d^mK)_h(\mathbf{u}-\mathbf{Y}_j).(\mathbf{1}(\mathbf{Y}_{k_1}\leq \mathbf{Y}_j)-\mathbf{Y}_j)\ldots(\mathbf{1}(\mathbf{Y}_{k_m}\leq \mathbf{Y}_j)-\mathbf{Y}_j)\\
&-\quad E[(d^mK)_h(\mathbf{u}-\mathbf{Y}_j).(\mathbf{1}(\mathbf{Y}_{k_1}\leq \mathbf{Y}_j)-\mathbf{Y}_j)\ldots(\mathbf{1}(\mathbf{Y}_{k_m}\leq \mathbf{Y}_j)]\}\,\omega(\mathbf{u})\,d\mathbf{u},
\end{aligned}
$$

for $m = 2, 3$. All the summands are centered, except when there are some equalities involving all the indices $k_1, \ldots, k_m$ and i (at least).

By splitting $T_{\gamma,m}$, we get several terms. If all the previous indices $i, j, k_1, \ldots, k_m$ are different from each other, the expectation is zero and the variance is

$$V_{\gamma,m}^{(1)} = \frac{4}{n^{4+2m}h^{2m}} \sum_{i_1<j_1} \sum_{i_2<j_2} \sum_{k_1,\ldots,k_m} \sum_{k'_1,\ldots,k'_m}$$
$$E \int \{K_h(\mathbf{u}_1 - \mathbf{Y}_{i_1}) - EK_h(\mathbf{u}_1 - \mathbf{Y}_{i_1})\} \cdot \{K_h(\mathbf{u}_2 - \mathbf{Y}_{i_2} - EK_h(\mathbf{u}_2 - \mathbf{Y}_{i_2})\}$$
$$\cdot \quad \{(d^m K)_h(\mathbf{u}_1 - \mathbf{Y}_{j_1}).(\mathbf{1}(\mathbf{Y}_{k_1} \leq \mathbf{Y}_{j_1}) - \mathbf{Y}_{j_1}) \ldots (\mathbf{1}(\mathbf{Y}_{k_m} \leq \mathbf{Y}_{j_1}) - \mathbf{Y}_{j_1})\}$$
$$\cdot \quad \left\{(d^m K)_h(\mathbf{u}_2 - \mathbf{Y}_{j_2}).(\mathbf{1}(\mathbf{Y}_{k'_1} \leq \mathbf{Y}_{j_2}) - \mathbf{Y}_{j_2}) \ldots (\mathbf{1}(\mathbf{Y}_{k'_m} \leq \mathbf{Y}_{j_2}) - \mathbf{Y}_{j_2})\right\} \omega(\mathbf{u}_1)\omega(\mathbf{u}_2)\, d\mathbf{u}_1\, d\mathbf{u}_2.$$

The corresponding terms are zero except when there are some equalities involving all the indices $k_1, \ldots, k_m, k'_1, \ldots, k'_m$ and i_1, i_2. There are at least $m + 1$ equalities. Moreover, there are always three "free" random variables at least, viz three integrations with respects to some $\mathbf{y}$ are available. It is possible to gain another factor h^d by the change of variables $h\tilde{\mathbf{u}}_2 = \mathbf{u}_2 - \mathbf{u}_1$. Thus, in every case,

$$V_{\gamma,m}^{(1)} = O\left(\frac{1}{n^{4+2m}h^{2m}} \cdot n^{4+2m-(m+1)}\right) = O\left(\frac{1}{n^{1+m}h^{2m}}\right).$$

This quantity is $o(n^{-2}h^{-d})$ when $m = 2, 3$ since nh^2 tends to the infinity when $n \to \infty$.

Imagine now there are some identities between the indices $i, j, k_1, \ldots, k_m$. The expectation of the corresponding term is zero, except if these equalities involve all i, $k_1, \ldots, k_m$. When $m = 2$ (resp. $m = 3$), two equalities at least are necessary. This implies the expectation is $O(n^{-2-m}h^{-m}n^m) = O(n^{-2}h^{-m}) = o(n^{-1}h^{-d/2})$. Moreover, its variance can be dealt exactly like $V_{\gamma,m}^{(1)}$. Thus, we have proved

$$T_{\gamma,m} = o_P\left(\frac{1}{nh^{d/2}}\right),$$

when $m = 2, 3$.

B.1.4 Study of $2n^{-2}\sum_{i<j} c_{n,i}c_{n,j}\omega$, $2n^{-2}\sum_{i<j} c_{n,i}b_{n,j}\omega$ and the other terms of the same type.

To deal with these terms simultaneously, denote

$$T_{\delta,m,p} = \frac{2}{n^{2+m+p}h^{m+p}} \sum_{i<j} \sum_{k_1,\ldots,k_m} \sum_{l_1,\ldots,l_p} \int \{(d^m K)_h(\mathbf{u} - \mathbf{Y}_i).(\mathbf{1}(\mathbf{Y}_{k_1} \leq \mathbf{Y}_i) - \mathbf{Y}_i)$$
$$\ldots \quad (\mathbf{1}(\mathbf{Y}_{k_m} \leq \mathbf{Y}_i) - \mathbf{Y}_i) - E[(d^m K)_h(\mathbf{u} - \mathbf{Y}_i).(\mathbf{1}(\mathbf{Y}_{k_1} \leq \mathbf{Y}_i) - \mathbf{Y}_i) \ldots (\mathbf{1}(\mathbf{Y}_{k_m} \leq \mathbf{Y}_j)]\}$$
$$\cdot \quad \left\{(d^p K)_h(\mathbf{u} - \mathbf{Y}_j).(\mathbf{1}(\mathbf{Y}_{l_1} \leq \mathbf{Y}_j) - \mathbf{Y}_j) \ldots (\mathbf{1}(\mathbf{Y}_{l_p} \leq \mathbf{Y}_j) - \mathbf{Y}_j)\right.$$
$$- \quad \left. E[(d^p K)_h(\mathbf{u} - \mathbf{Y}_j).(\mathbf{1}(\mathbf{Y}_{l_1} \leq \mathbf{Y}_j) - \mathbf{Y}_j) \ldots (\mathbf{1}(\mathbf{Y}_{l_p} \leq \mathbf{Y}_j)]\right\} \omega(\mathbf{u})\, d\mathbf{u},$$

for m and $l = 1, 2, 3$, $m + p \geq 3$. All the summands are centered, except when there are some equalities involving all the indices $k_1, \dots, k_m$ and $l_1, \dots, l_p$ (at least) .

Imagine we are dealing with all the terms of the previous sum corresponding to different indices. Thus the expectation is zero and the variance is a sum over $4 + 2(m + p)$ indices (denoted by $i_1, i_2, j_1, j_2,\ k_1, k_1', \dots, k_m, k_m',\ l_1, l_1', \dots, l_p, l_p'$ with obvious notations). Nonzero terms occurs when all the k, k', l and l' indices are matched. At least, this provides $m + p$ equalities. Moreover, there are always three opportunities to make some usual changes of variables and to remove the factors h^d. When this factor appears, it means we have an additional equality involving i or j indices. Thus, we win an extra factor n. Therefore, the variance is

$$O\left(\frac{1}{n^{4+2m+2p}h^{m+p}} \cdot (n^{4+m+p} + n^{3+m+p}h^{-d})\right).$$

In every case, this is $o(n^{-2}h^{-d})$.

Now, imagine there are some equalities between $i, j, k_1, \dots, k_m, l_1, \dots, l_p$. The variance of such a term can be dealt as previously. It is sufficient to verify that its expectation is negligible. This expectation is a sum of terms that are nonzero only if there are some equalities involving $k_1, \dots, k_m, l_1, \dots, l_p$. If $m + p$ is even, there are at least $(m + p)/2$ equalities. If $m + p$ is odd, there are at least $[(m + p)/2] + 1$ equalities. In every case, the factors h^d disappear by some changes of variables with respects to $\mathbf{y}_i$ and $\mathbf{y}_j$. To summarize, this expectation is $O(n^{-(m+p)/2}h^{-m-p})$ (resp. $O(n^{-[(m+p)/2]-1}h^{-m-p})$) if $m+p$ is even (resp. odd). These terms are $o(n^{-1}h^{-d/2})$ if $nh^3 \to \infty$.

Thus

$$T_{\delta,m,p} = o_P\left(\frac{1}{nh^{d/2}}\right).$$

B.1.5 Study of the remainder terms

These terms are like $2n^{-2}\sum_{i<j} a_i e_{n,j}\omega$. Actually, every term that involves $e_{n,j}$ is negligible. For instance

$$\left|2n^{-2}\sum_{i<j} a_i e_{n,j}\omega\right| \leq \frac{Cst}{n^2h^4} \cdot \frac{n^2}{h^d} \cdot \sup_j \|\mathbf{Y}_{n,j} - \mathbf{Y}_j\|_\infty^4 = O_P\left(\frac{\ln_2^2 n}{n^2h^{4+d}}\right).$$

This term is $o_P(n^{-1}h^{-d/2})$ under (B).

Thus, we have got

$$J^*_{n,2} = \frac{\sqrt{2}}{nh^{d/2}}\mathcal{N}_n + 2T_\alpha + T_\beta + o_P\left(\frac{1}{nh^{d/2}}\right)$$
$$= \frac{\sqrt{2}}{nh^{d/2}}\mathcal{N}_n + \frac{(-1)}{nh}\int \tau^2\omega . \sum_{r=1}^d \int K_r^2 + o_P\left(\frac{1}{nh^{d/2}}\right) \quad \text{(B.7)}$$

where $\mathcal{N}_n$ tends in law towards a gaussian r.v. $\mathcal{N}(0,\sigma^2)$.

B.2 Study of $J^*_{n,1}$

With the previous notations

$$J^*_{n,1} = \frac{1}{n^2}\sum_i \int a^2_{n,i}\omega = \frac{1}{n^2}\sum_i \int [a_i + b_{n,i} + c^*_{n,i}]^2\omega$$
$$= \frac{1}{n^2}\sum_i \int \left[a_i^2 + b^2_{n,i} + (c^*_{n,i})^2 + 2a_i b_{n,i} + 2b_{n,i}c^*_{n,i} + 2a_i c^*_{n,i}\right]\omega,$$

where the expansion of K has been stopped at the second order. We denote

$$c^*_{n,i}(\mathbf{u}) = \frac{1}{2h^2}\left\{(d^2K)_h(\mathbf{u}-\mathbf{Y}^*_{n,i}).(\mathbf{Y}_{n,i}-\mathbf{Y}_i)^{(2)} - E[(d^2K)_h(\mathbf{u}-\mathbf{Y}^*_{n,i}).(\mathbf{Y}_{n,i}-\mathbf{Y}_i)^{(2)}]\right\}$$
$$= O_P\left(\frac{1}{h^{d+2}}\sup_i \|\mathbf{Y}_{n,i}-\mathbf{Y}_i\|^2\right) = O_P\left(\frac{\ln_2 n}{nh^{d+2}}\right).$$

Therefore, it is easy to bound $\int a_i c^*_{n,i}\omega$, $\int b_{n,i}c^*_{n,i}\omega$, and $\int (c^*_{n,i})^2\omega$. All the corresponding terms in $J^*_{n,1}$ are negligible if

$$\frac{\ln_2 n}{n^2h^{d+2}} + \frac{\ln_2^{3/2} n}{n^{5/2}h^{d+3}} + \frac{\ln_2^2 n}{n^3h^{d+4}} << \frac{1}{nh^{d/2}}.$$

This is satisfied under condition (B). The main term of $J^*_{n,1}$ is provided by $\int a_i^2\omega$. Note that

$$E\frac{1}{n^2}\sum_i \int a_i^2\omega = \frac{1}{nh^d}\int K^2(\mathbf{t})(\tau\omega)(\mathbf{u}-h\mathbf{t})\,d\mathbf{t} + O(n^{-1})$$
$$= \frac{1}{nh^d}\int K^2 \int \tau\omega + O\left(\frac{h^2}{nh^d}\right),$$

since K is even. Moreover, the variance is

$$V \equiv \frac{1}{n^4}E\sum_{i,j}\int \Bigg[\left\{K_h(\mathbf{u}_1-\mathbf{Y}_i) - \int K(\mathbf{t})\tau(\mathbf{u}_1-h\mathbf{t})\,d\mathbf{t}\right\}^2$$
$$- E\left\{K_h(\mathbf{u}_1-\mathbf{Y}_i) - \int K(\mathbf{t})\tau(\mathbf{u}_1-h\mathbf{t})\,d\mathbf{t}\right\}^2\Bigg]\cdot\Bigg[\left\{K_h(\mathbf{u}_2-\mathbf{Y}_j) - \int K(\mathbf{t})\tau(\mathbf{u}_2-h\mathbf{t})\,d\mathbf{t}\right\}^2$$
$$- E\left\{K_h(\mathbf{u}_2-\mathbf{Y}_j) - \int K(\mathbf{t})\tau(\mathbf{u}_2-h\mathbf{t})\,d\mathbf{t}\right\}^2\Bigg]\,\omega(\mathbf{u}_1)\omega(\mathbf{u}_2)\,d\mathbf{u}_1\,d\mathbf{u}_2.$$

The non zero terms are obtained when $i = j$. By the change of variables $h\tilde{\mathbf{y}}_i = \mathbf{u}_1 - \mathbf{y}_i$ and $h\tilde{\mathbf{u}}_2 = \mathbf{u}_2 - \mathbf{u}_1$, it is easy to verify that $V = O(n^{-3} \cdot h^{-2d})$. Thus, since $nh^d \to \infty$, we get

$$\frac{1}{n^2}\sum_i \int a_i^2\omega = \frac{1}{nh^d}\int K^2(\mathbf{t})(\tau\omega)(\mathbf{u} - h\mathbf{t})\, d\mathbf{t} + o_P(n^{-1}h^{-d/2}).$$

Let us consider now $T \equiv n^{-2}\sum_i \int a_i b_{n,i}\omega$. Its expectation is

$$\begin{aligned} E[n^{-2}\sum_i \int a_i b_{n,i}\omega] &= n^{-1}\int E[a_1 b_{n,1}]\omega \\ &= \frac{(-1)}{n^2 h}\int \left\{ K_h(\mathbf{u}-\mathbf{y}) - \int K(\mathbf{t})\tau(\mathbf{u}-h\mathbf{t})\, d\mathbf{t} \right\} \\ &\cdot \{(dK)_h(\mathbf{u}-\mathbf{y}).(1-\mathbf{y}) - E[(dK)_h(\mathbf{u}-\mathbf{Y}).(1-\mathbf{Y})]\}\, \tau(\mathbf{y})\omega(\mathbf{u})\, d\mathbf{u}\, d\mathbf{y} \\ &= \frac{(-1)}{n^2 h^{1+d}}\int \left\{ K(\mathbf{v}) - \int K(\mathbf{t})\tau(\mathbf{u}-h\mathbf{t})\, d\mathbf{t} \right\} .(dK)(\mathbf{v}).(1-\mathbf{u}-h\mathbf{v}) \\ &\cdot \tau(\mathbf{u}-h\mathbf{v})\omega(\mathbf{u})\, d\mathbf{v}\, d\mathbf{u} + O(n^{-2}h^{-1}). \end{aligned}$$

Thus, this expectation is $o(n^{-1}h^{-d/2})$. Moreover, its variance is

$$\begin{aligned} Var(T) &= \frac{1}{n^4}E\sum_{i,j}\int a_i(\mathbf{u}_1)a_j(\mathbf{u}_2)b_{n,i}(\mathbf{u}_1)b_{n,j}(\mathbf{u}_2)\omega(\mathbf{u}_1)\omega(\mathbf{u}_2)\, d\mathbf{u}_1\, d\mathbf{u}_2 - E[T]^2 \\ &= \frac{1}{n^3}E\int a_1(\mathbf{u}_1)a_1(\mathbf{u}_2)b_{n,1}(\mathbf{u}_1)b_{n,1}(\mathbf{u}_2)\omega(\mathbf{u}_1)\omega(\mathbf{u}_2)\, d\mathbf{u}_1\, d\mathbf{u}_2 - E[T]^2 \\ &= \frac{1}{n^3h^2}E\int a_1(\mathbf{u}_1)a_1(\mathbf{u}_2)(dK)_h(\mathbf{u}_1 - \mathbf{Y}_1).(\mathbf{Y}_{n,1} - \mathbf{Y}_1) \\ &\cdot (dK)_h(\mathbf{u}_2 - \mathbf{Y}_1).(\mathbf{Y}_{n,1} - \mathbf{Y}_1)\omega(\mathbf{u}_1)\omega(\mathbf{u}_2)\, d\mathbf{u}_1\, d\mathbf{u}_2 + O(n^{-4}h^{-2-2d}) \end{aligned}$$

We get immediately, using an a.e. upper bound for the empirical process and three changes of variables,

$$\begin{aligned} Var(T) &\le \frac{1}{n^3h^2}E\int |a_1(\mathbf{u}_1)a_1(\mathbf{u}_2)|.\|(dK)_h\|_\infty^2(\mathbf{u}_1 - \mathbf{Y}_1).\|(dK)_h\|_\infty(\mathbf{u}_2 - \mathbf{Y}_1) \\ &\cdot \|\mathbf{Y}_{n,1} - \mathbf{Y}_1\|_\infty \omega(\mathbf{u}_1)\omega(\mathbf{u}_2)\, d\mathbf{u}_1\, d\mathbf{u}_2 + O(n^{-4}h^{-2-2d}) \\ &\le \frac{Cst}{n^3h^2}\cdot\frac{1}{h^d}\cdot\frac{\ln_2 n}{n} \end{aligned}$$

The latter upper bound is $o(n^{-2}h^{-d})$. Thus, we have proved $T = o_P(n^{-1}h^{-d/2})$.

It remains to deal with $n^{-2}\sum_i \int b_{n,i}^2\omega$. By a change of variable with respects to $\mathbf{u}$, we get directly the upper bound

$$n^{-2}\sum_i \int b_{n,i}^2\omega = O_P\left(\frac{1}{nh^2}\cdot\frac{1}{h^d}\cdot\frac{\ln_2 n}{n}\right) = o_P(n^{-1}h^{-d/2}),$$

if $nh^{2+d/2}/\ln_2 n \to \infty$. The latter condition could be relaxed by a more cautious analysis of the latter term, as done previously. It is useless, facing the set of technical assumptions we have already done.

To conclude,

$$J^*_{n,1} = \frac{1}{nh^d}\int K^2(\mathbf{t})(\tau\omega)(\mathbf{u}-h\mathbf{t})\,d\mathbf{t} + o_P\left(\frac{1}{nh^{d/2}}\right). \tag{B.8}$$

B.3 Study of J_I

Recall that

$$J_I = \int (\tau_n - E\tau_n).(E\tau_n - K_h * \hat\tau)\omega, \text{ and}$$

$$\tau_n(\mathbf{u}) - E\tau_n(\mathbf{u}) = n^{-1}\sum_i \int [a_i(\mathbf{u}) + b^*_{n,i}(\mathbf{u})]\omega(\mathbf{u})\,d\mathbf{u}, \text{ with}$$

$$b^*_{n,i}(\mathbf{u}) = \frac{(-1)}{nh}\sum_{i=1}^n \Big\{(dK)_h(\mathbf{u}-\mathbf{Y}^*_{n,i}).(\mathbf{Y}_{n,i}-\mathbf{Y}_i) - E[(dK)_h(\mathbf{u}-\mathbf{Y}^*_{n,i}).(\mathbf{Y}_{n,i}-\mathbf{Y}_i)]\Big\},$$

for some random variable $\mathbf{Y}^*_{n,i}$, $\|\mathbf{Y}^*_{n,i}-\mathbf{Y}_i\| \leq \|\mathbf{Y}_{n,i}-\mathbf{Y}_i\|$ a.e. Thus,

$$\begin{aligned}
J_I &= \int \left\{\frac{1}{n}\sum_i [a_i(\mathbf{u}) + b^*_{n,i}(\mathbf{u})]\right\}.\Big\{E\beta^*_{n,i}(\mathbf{u}) - K_h * (\hat\tau - \tau)(\mathbf{u})\Big\}\omega(\mathbf{u})\,d\mathbf{u} \\
&= \frac{1}{n}\sum_i \int a_i(\mathbf{u})K_h * (\tau - \hat\tau)(\mathbf{u})\omega(\mathbf{u})\,d\mathbf{u} + \frac{1}{n}\sum_i \int a_i(\mathbf{u})E\beta^*_{n,i}(\mathbf{u})\omega(\mathbf{u})\,d\mathbf{u} \\
&+ \frac{1}{n}\sum_i \int b^*_{n,i}(\mathbf{u})K_h * (\tau - \hat\tau)(\mathbf{u})\omega(\mathbf{u})\,d\mathbf{u} + \frac{1}{n}\sum_i \int b^*_{n,i}(\mathbf{u})E\beta^*_{n,i}(\mathbf{u})\omega(\mathbf{u})\,d\mathbf{u} \\
&\equiv J_I^{(0)} + J_I^{(1)} + J_I^{(2)} + J_I^{(3)},
\end{aligned}$$

by denoting

$$\beta^*_{n,i}(\mathbf{u}) = \frac{(-1)}{h}(dK)_h(\mathbf{u}-\mathbf{Y}^*_{n,i}).(\mathbf{Y}_{n,i}-\mathbf{Y}_i).$$

Clearly

$$\hat\tau(\mathbf{u}) - \tau(\mathbf{u}) = \partial_\theta\tau(\mathbf{u},\theta_0).(\hat\theta-\theta_0) + 2^{-1}\partial^2_\theta\tau(\mathbf{u},\tilde\theta).(\hat\theta-\theta_0)^{(2)}, \tag{B.9}$$

for some $\tilde\theta$, $\|\tilde\theta-\theta_0\| \leq \|\hat\theta-\theta_0\|$ a.e. Implicitly, $\tilde\theta$ depends on $\mathbf{u}$.

B.3.1 Study of $J_I^{(0)}$

Note that

$$\begin{aligned}
J_I^{(0)} &= n^{-1}\sum_{i=1}^n \int a_i(\mathbf{u})K(\mathbf{v})(\tau-\hat\tau)(\mathbf{u}-h\mathbf{v})\omega(\mathbf{u})\,d\mathbf{u}\,d\mathbf{v} \\
&= n^{-1}\sum_{i=1}^n \int a_i(\mathbf{u})K(\mathbf{v})\partial_\theta\tau(\mathbf{u}-h\mathbf{v},\theta_0).(\hat\theta-\theta_0)\omega(\mathbf{u})\,d\mathbf{u}\,d\mathbf{v} \\
&+ (2n)^{-1}\sum_{i=1}^n \int a_i(\mathbf{u})K(\mathbf{v})\partial^2_\theta\tau(\mathbf{u}-h\mathbf{v},\tilde\theta).(\hat\theta-\theta_0)^{(2)}\omega(\mathbf{u})\,d\mathbf{u}\,d\mathbf{v} \\
&= J_{I,1}^{(0)} + J_{I,2}^{(0)}.
\end{aligned}$$

Actually, the latter random quantity $\tilde{\theta}$ depends on $\mathbf{u}-h\mathbf{v}$. The first previous term $J_{I,1}^{(0)}$ can be dealt exactly as in Fan [19]. This author has assumed $\hat{\theta}$ is the maximum likelihood estimator of θ, which implies $B(\theta_0, \mathbf{Y}_i)$ is a score function. Actually, by reading carefully her proof, we notice we need only $B(\theta_0, \mathbf{Y}_i)$ is centered and belongs in L^2, viz our assumption (E). Thus, $J_{I,1}^{(0)} = O_P(n^{-1})$. Moreover, by some change of variables,

$$\|J_{I,2}^{(0)}\| \leq \frac{Cst}{n}\sum_{i=1}^n \int |K|(\tilde{\mathbf{u}})|K|(\mathbf{v})\|\partial^2_\theta \tau(\mathbf{Y}_i - h\tilde{\mathbf{u}} - h\mathbf{v}, \tilde{\theta})\|\omega(\mathbf{Y}_i - h\tilde{\mathbf{u}})\, d\tilde{\mathbf{u}}\, d\mathbf{v}.\|\hat{\theta} - \theta_0\|^2.$$

To bound the previous right hand side, we could assume

$$E\left[\sup_{\{(\mathbf{u},\mathbf{v},\theta)|\|\mathbf{u}\|+\|\mathbf{v}\|\leq 2h, \|\theta-\theta_0\|\leq\varepsilon\}} \|\partial^2_\theta \tau(\mathbf{Y}_i - \mathbf{u}, \theta)\| \cdot |\omega|(\mathbf{Y}_i - \mathbf{v})\right] < \infty. \tag{B.10}$$

This assumption is satisfied under the stronger condition (T), for n sufficiently large.

Thus, under (B.10), we get

$$J_{I,2}^{(0)} = O_P(\|\hat{\theta} - \theta_0\|^2) = O_P(n^{-1}).$$

B.3.2 Study of $J_I^{(1)}$

$$J_I^{(1)} = \frac{(-1)}{nh}\sum_i \int a_i(\mathbf{u}) E[(dK)_h(\mathbf{u} - \mathbf{Y}_i^*).(\mathbf{Y}_{n,i} - \mathbf{Y}_i)]\omega(\mathbf{u})\, d\mathbf{u}.$$

Clearly, this term is centered. Note that, by a limited expansion of K up to the p'th order, we can prove that

$$E[(dK)_h(\mathbf{u} - \mathbf{Y}_i^*).(\mathbf{Y}_{n,i} - \mathbf{Y}_i)] = O\left(\frac{1}{n} + \frac{1}{h^{p+d}} \cdot \left(\frac{\ln_2 n}{n}\right)^{(p+1)/2}\right). \tag{B.11}$$

The latter upper bound is uniform with respects to $\mathbf{u}$. Therefore, the variance of $J_I^{(1)}$ is

$$E[\left(J_I^{(1)}\right)^2] = \frac{1}{n^2h^2}\sum_i E\int a_i(\mathbf{u}_1)a_i(\mathbf{u}_2)E[(dK)_h(\mathbf{u}_1 - \mathbf{Y}_i^*).(\mathbf{Y}_{n,i} - \mathbf{Y}_i)]$$
$$\cdot \quad E[(dK)_h(\mathbf{u}_2 - \mathbf{Y}_i^*).(\mathbf{Y}_{n,i} - \mathbf{Y}_i)]\omega(\mathbf{u}_1)\omega(\mathbf{u}_2)\, d\mathbf{u}_2\, d\mathbf{u}_1$$
$$= \quad O\left(\frac{1}{nh^2}\cdot\frac{1}{n^2} + \frac{1}{nh^2}\cdot\frac{1}{h^{2p+2d}}\cdot\left(\frac{\ln_2 n}{n}\right)^{p+1}\right) = o\left(\frac{1}{n^2h^d}\right),$$

by a change of variables with respects to $\mathbf{y}$ and $\mathbf{u}_2$, and if $n^p h^{2+2p+d}/(\ln_2 n)^{p+1} \to \infty$. The latter condition is satisfied under our assumptions with $p = 2$.

B.3.3 Study of $J_I^{(2)}$

With obvious notations,

$$
\begin{aligned}
J_I^{(2)} &= n^{-1}\sum_{i=1}^n \int b_{n,i}^*(\mathbf{u})K(\mathbf{v})(\tau-\hat{\tau})(\mathbf{u}-h\mathbf{v})\omega(\mathbf{u})\,d\mathbf{u}\,d\mathbf{v} \\
&= n^{-1}\sum_{i=1}^n \int [b_{n,i}+c_{n,i}^*](\mathbf{u})K(\mathbf{v})\left[\partial_\theta\tau(\mathbf{u}-h\mathbf{v},\theta_0).(\hat{\theta}-\theta_0)\right. \\
&+ \left. 2^{-1}\partial_\theta^2\tau(\mathbf{u}-h\mathbf{v},\tilde{\theta}).(\hat{\theta}-\theta_0)^{(2)}\right]\omega(\mathbf{u})\,d\mathbf{u}\,d\mathbf{v} \\
&= n^{-1}\sum_{i=1}^n \int b_{n,i}(\mathbf{u})K(\mathbf{v})\left[\partial_\theta\tau(\mathbf{u}-h\mathbf{v},\theta_0).(\hat{\theta}-\theta_0)\right]\omega(\mathbf{u})\,d\mathbf{u}\,d\mathbf{v} \\
&+ O_P\left(\frac{\ln_2 n}{n}\cdot\frac{1}{h^2n^{1/2}}+\frac{1}{h}\cdot\left(\frac{\ln_2 n}{n}\right)^{1/2}\cdot\frac{1}{n}\right),
\end{aligned}
$$

under the condition (B.10). The main term of the latter expansion is

$$
T \equiv \frac{1}{n^2}\sum_{i,j}\int b_{n,i}(\mathbf{u})K(\mathbf{v})\partial_\theta\tau(\mathbf{u}-h\mathbf{v},\theta_0)A(\theta_0)^{-1}B(\theta_0,\mathbf{Y}_j)\omega(\mathbf{u})\,d\mathbf{u}.
$$

Thus, when $i\neq j$, the expectation of the summand is $O(n^{-1})$, and

$$
\begin{aligned}
E[T] &= \frac{1}{n^2}\sum_i E\left[b_{n,i}(\mathbf{u})K(\mathbf{v})\partial_\theta\tau(\mathbf{u}-h\mathbf{v},\theta_0)A(\theta_0)^{-1}B(\theta_0,\mathbf{Y}_i)\right]\omega(\mathbf{u})\,d\mathbf{u}+O(n^{-1}) \\
&= O\left(\frac{1}{nh}\cdot\left(\frac{\ln_2 n}{n}\right)^{1/2}+\frac{1}{n}\right)=o\left(\frac{1}{nh^{d/2}}\right).
\end{aligned}
$$

Moreover, by the same reasoning, its variance is

$$
Var(T)=O\left(\frac{1}{n^2h^2}\cdot\left(\frac{\ln_2 n}{n}\right)\right)=o(\frac{1}{n^2h^d}).
$$

Note that one remainder term is

$$
\frac{1}{n}\sum_i\int b_{n,i}(\mathbf{u})K(\mathbf{v})\partial_\theta\tau(\mathbf{u}-h\mathbf{v},\theta_0)\omega(\mathbf{u})\,d\mathbf{u}.o_P(r_n).
$$

The latter term is negligible if

$$
\left(\frac{\ln_2 n}{n}\right)^{1/2}\cdot\frac{r_n}{h} << \frac{1}{nh^{d/2}},
$$

viz if

$$
r_n = o\left(\frac{1}{\sqrt{n}\ln_2^{1/2} n}\cdot\frac{1}{h^{d/2-1}}\right).
$$

B.3.4 Study of $J_I^{(3)}$

Clearly, under the previous assumptions,

$$J_I^{(3)} = O_P\left(\frac{1}{h}\cdot\left(\frac{\ln_2 n}{n}\right)^{1/2}\cdot\frac{1}{nh}\right) = o_P\left(\frac{1}{nh^{d/2}}\right),$$

since $nh^2/\ln_2 n \to \infty$.

To conclude,

$$J_I = o_P\left(\frac{1}{nh^{d/2}}\right) \tag{B.12}$$

B.4 Study of J_{II}

With the previous notations,

$$\begin{aligned} J_{II} &= \int (E\tau_n - K_h * \hat{\tau})^2\omega \\ &= \int [K_h * (\hat{\tau}-\tau)]^2\omega + \int [E\beta_{ni}^*]^2\omega - 2\int K_h * (\hat{\tau}-\tau)E\beta_{ni}^*\omega. \end{aligned}$$

Applying equation (B.11) with $p = 2$, we get

$$E\beta_{ni}^*(\mathbf{u}) = O\left(\frac{1}{n} + \frac{1}{h^4}\cdot\left(\frac{\ln_2 n}{n}\right)^{3/2}\right),$$

uniformly with respects to $\mathbf{u}$. Thus, it is straightforward

$$\int [E\beta_{ni}^*]^2\omega = o\left(\frac{1}{nh^{d/2}}\right)$$

Moreover, under assumption (T) and by a limited expansion with respects to $\mathbf{u}$,

$$\int [K_h * (\hat{\tau}-\tau)]^2\omega = O_P\left(\frac{1}{n}\right).$$

By applying Schwartz's inequality, we obtain

$$J_{II} = o_P\left(\frac{1}{nh^{d/2}}\right). \tag{B.13}$$

The result of theorem 3 follows from equations (B.1), (B.2), (B.7), (B.8), (B.12) and (B.13).
□

C Proof of corollary 4

It is sufficient to prove that

$$\frac{1}{nh^d}\int K^2(\mathbf{t})((\hat\tau-\tau)\omega)(\mathbf{u}-h\mathbf{t})\,d\mathbf{t}\,d\mathbf{u} = o_P\left(\frac{1}{nh^{d/2}}\right), \text{ and} \tag{C.1}$$

$$\frac{1}{nh}\int K^2(\mathbf{t})(\hat\tau^2-\tau^2)\omega = o_P\left(\frac{1}{nh^{d/2}}\right). \tag{C.2}$$

Note that, under (T) and by a limited expansion with respects to θ, we have

$$\sup_{\mathbf{u}\in[\varepsilon,1-\varepsilon]^d}\|\hat\tau(\mathbf{u},\hat\theta)-\tau(\mathbf{u},\theta_0)\| = O_P(\|\hat\theta-\theta_0\|) = O_P(n^{-1/2}).$$

Thus, equation (C.1) and (C.2) are clearly satisfied because nh^d tends to the infinity when $n\to\infty$, proving the result. □

D The semiparametric estimator

Consider the parametric family $\mathcal{C} = \{\tau(\cdot,\theta),\theta\in\Theta\}$. The semiparametric estimator of θ satisfies, by definition,

$$\hat\theta = \arg\max_{\theta\in\Theta} Q_n(\theta), \text{ where}$$

$$Q_n(\theta) = n^{-1}\sum_{i=1}^n \ln\tau(\mathbf{Y}_{ni},\theta).$$

We prove that $\hat\theta$ satisfies condition (4.1). By a limited expansion, there exists some random vector θ^* such that

$$\partial^2_{\theta\theta}Q_n(\theta^*).(\hat\theta-\theta_0) = -\partial_\theta Q_n(\theta_0),$$

with $\|\theta^*-\theta_0\|\le\|\hat\theta-\theta_0\|$ a.e. First, with obvious notations,

$$\partial_\theta Q_n(\theta_0) = n^{-1}\sum_{i=1}^n\partial_\theta\ln\tau(\mathbf{Y}_i,\theta_0) + n^{-1}\sum_{i=1}^n\partial^2_{\mathbf{y},\theta}\ln\tau(\mathbf{Y}_i,\theta_0).(\mathbf{Y}_{n,i}-\mathbf{Y}_i)$$
$$+ \quad \frac{1}{2n}\sum_{i=1}^n\partial^3_{\mathbf{y}\mathbf{y}\theta}\ln\tau(\mathbf{Y}^*_{ni},\theta_0).(\mathbf{Y}_{n,i}-\mathbf{Y}_i)^{(2)} \equiv S_0+S_1+S_2.$$

We assume that

$$E\left[\|\partial_\theta\ln\tau(\mathbf{Y},\theta_0)\| + \|\partial^2_{\theta,\mathbf{y}}\ln\tau(\mathbf{Y},\theta_0)\| + \|\partial^3_{\theta,\mathbf{y},\mathbf{y}}\ln\tau(\mathbf{Y},\theta_0)\|\right] < \infty. \tag{D.1}$$

Obviously, S_0 is asymptotically normal. The expectation of S_1 is $O(n^{-1})$ and its variance is $O(n^{-2})$. Thus, S_1 is $O_P(n^{-1})$. Moreover,

$$\|S_2\| \le Cte\cdot\frac{1}{n}\sum_{i=1}^n\|\partial^3_{\mathbf{y}\mathbf{y}\theta}\ln\tau(\mathbf{Y}^*_{ni},\theta_0)\|.\|\mathbf{Y}_{ni}-\mathbf{Y}_i\|^2. \tag{D.2}$$

Assume the following conditions of regularity :

1. There exist some constants α et β such that, a.e.,

$$\|\partial^3_{\mathbf{y}\mathbf{y}\theta} \ln \tau(\mathbf{Y}^*_{ni}, \theta_0)\| \leq \alpha \|\partial^3_{\mathbf{y}\mathbf{y}\theta} \ln \tau(\mathbf{Y}_i, \theta_0)\| + \beta \|\partial^3_{\mathbf{y}\mathbf{y}\theta} \ln \tau(\mathbf{Y}_{ni}, \theta_0)\|, \text{ and}$$

2. For every $\mathbf{u} \in (0,1)^d$,

$$\|\partial^3_{\mathbf{y}\mathbf{y}\theta} \ln \tau(\mathbf{u}, \theta_0)\| \leq Cst.r(u_1)^{a_1} \ldots r(u_d)^{a_d},$$

   where $a_k = (-1+\delta)/p_k$, $1/p_1 + \ldots + 1/p_k = 1$, $\delta > 0$, and $r(t) = t(1-t)$.

The latter condition ensures the consistency of the empirical mean of $\|\partial^3_{\mathbf{y}\mathbf{y}\theta} \ln \tau(\mathbf{Y}_{ni}, \theta_0)\|$ (see Genest et al. [29], proposition A.1). Thus, we get $\|S_2\| = O_P(n^{-1} \ln_2 n)$. We have obtained

$$\partial_\theta Q_n(\theta_0) = n^{-1} \sum_{i=1}^n \partial_\theta \ln \tau(\mathbf{Y}_i, \theta) + O_P(\ln_2 n/n).$$

Moreover, with obvious notations,

$$\partial^2_\theta Q_n(\theta^*) = \partial^2_\theta Q_n(\theta_0) + n^{-1} \sum_{i=1}^n \partial^3_\theta \ln \tau(\mathbf{Y}_{ni}, \tilde{\theta}).(\theta^* - \theta_0)$$
$$= \lim_{n\infty} E[\partial^2_\theta Q_n(\theta_0)] + O_P(n^{-1/2}),$$

if $\partial^2_\theta Q_n(\theta)$ is asymptotically normal, and if

$$n^{-1} \sum_{i=1}^n \sup_{\theta \in \mathcal{V}(\theta_0)} \|\partial^3_\theta \ln \tau(\mathbf{Y}_{ni}, \theta)\| < \infty \text{ a.e.} \tag{D.3}$$

Here, $\mathcal{V}(\theta_0)$ denotes a neighborhood of θ_0. Applying proposition A.1 of Genest et al. [30], these two conditions can are ensured if :

1. For every $\mathbf{u} \in (0,1)^d$,

$$M(\mathbf{u}) \equiv \|\partial^2_\theta \ln \tau(\mathbf{u}, \theta_0)\| \leq Cst.r(u_1)^{b_1} \ldots r(u_d)^{b_d},$$

   where $b_k = (-0.5+\nu)/q_k$, $1/q_1 + \ldots + 1/q_k = 1$, $\nu > 0$. Moreover, $M(\mathbf{u})$ has continuous partial derivatives $M_k(\mathbf{u}) = \partial M(\mathbf{u})/\partial u_k$, such that

$$M_k(\mathbf{u}) \leq Cst.r(u_1)^{d_1^{(k)}} \ldots r(u_d)^{d_d^{(k)}},$$

   $d_k^{(k)} = b_k$, $d_j^{(k)} = b_j - 1$ if $j \neq k$.

2. For every $\mathbf{u} \in (0,1)^d$,

$$\sup_{\theta \in \mathcal{V}(\theta_0)} \|\partial^3_\theta \ln \tau(\mathbf{u}, \theta)\| \leq Cst.r(u_1)^{c_1} \ldots r(u_d)^{c_d},$$

   where $c_k = (-1+\eta)/p'_k$, $1/p'_1 + \ldots + 1/p'_k = 1$, $\eta > 0$.

Condition (1) ensures the asymptotic normality of the empirical mean of $M(\mathbf{Y}_{ni})$. Condition (2) ensures condition (D.3).

It can be verified that the previous conditions are satisfied by a large number of commonly used copula families. Particularly, it is the case for the gaussian copula.

Thus, under the previous conditions, we get

$$\sqrt{n}(\hat{\theta} - \theta_0) = \frac{1}{\sqrt{n}} A(\theta_0)^{-1} \cdot \sum_{i=1}^{n} \partial_\theta \ln \tau(\mathbf{Y}_i, \theta_0) + O_P\left(\frac{\ln_2 n}{n}\right),$$

$$A(\theta_0) = - \lim_{n\infty} E\left[\partial_\theta^2 Q_n(\theta)\right]$$

and (4.1) is satisfied. □

E Proof of theorem 5

We need a technical lemma to control the strength of the dependence between successive observations.

Let $(\xi_i)_{i\in\mathbb{Z}}$ be a d-dimensional strictly stationary β-mixing process. Let k and $i_1 < i_2 < \ldots < i_k$ be arbitrary integers. For any j, $1 \leq j \leq k-1$, put

$$P_j^{(k)}(E^{(j)} \times E^{(k-j)}) = P((\xi_{i_1}, \ldots, \xi_{i_j}) \in E^{(j)}) P((\xi_{i_{j+1}}, \ldots, \xi_{i_k}) \in E^{(k-j)}), \text{ and}$$

$$P_0^{(k)}(E^{(k)}) = P((\xi_{i_1}, \ldots, \xi_{i_k}) \in E^{(k)}),$$

where $E^{(i)}$ is a Borel set in $\mathbb{R}^{di}$. Lemma 1 in Yoshihara (1976) states

Lemma 6. *Let $h(\mathbf{x}_1, \ldots, \mathbf{x}_k)$ be a Borel measurable function such that*

$$\int |h(\mathbf{x}_1, \ldots, \mathbf{x}_k)|^{1+\delta} \, dP_j^{(k)} \leq M,$$

for some $\delta > 0$. Then

$$\left| \int h(\mathbf{x}_1, \ldots, \mathbf{x}_k) \, dP_0^{(k)} - \int h(\mathbf{x}_1, \ldots, \mathbf{x}_k) \, dP_j^{(k)} \right| \leq 4 M^{1/(1+\delta)} \beta_{i_{j+1}-i_j}^{\delta/(\delta+1)}.$$

We mimic the proof of theorem 1. Therefore, with the same notations,

$$\begin{aligned} \tau_n(\mathbf{u}) &= \tau_n^*(\mathbf{u}) + \frac{(-1)}{nh} \sum_{i=1}^{n} (dK)_h(\mathbf{u} - \mathbf{Y}_i).(\mathbf{Y}_{n,i} - \mathbf{Y}_i) \\ &+ \frac{1}{2nh^2} \sum_{i=1}^{n} (d^2K)_h(\mathbf{u} - \mathbf{Y}_{n,i}^*).(\mathbf{Y}_{n,i} - \mathbf{Y}_i)^{(2)} \\ &= \tau_n^*(\mathbf{u}) + \tilde{R}_1(\mathbf{u}) + \tilde{R}_2(\mathbf{u}), \end{aligned}$$

for some random vector $\mathbf{Y}^*_{n,i}$ satisfying $\|\mathbf{Y}^*_{n,i} - \mathbf{Y}_i\| \leq \|\mathbf{Y}_{n,i} - \mathbf{Y}_i\|$ a.e.

Let us first study $\tilde{R}_1(\mathbf{u})$. By applying lemma 6 and by noting that

$$\int (dK)_h(\mathbf{u} - \mathbf{y}_i).(\mathbf{1}(\mathbf{y}_k \leq \mathbf{y}_i) - \mathbf{y}_i)\tau(\mathbf{y}_i)\tau(\mathbf{y}_k)\, d\mathbf{y}_i\, d\mathbf{y}_k = 0, \quad k \neq i,$$

we get easily

$$\begin{aligned} |E\left[\tilde{R}_1(\mathbf{u})\right]| &\leq \frac{1}{n^2 h} \sum_{i,k} |E\left[(dK)_h(\mathbf{u} - \mathbf{Y}_i).(\mathbf{1}(\mathbf{Y}_k \leq \mathbf{Y}_i) - \mathbf{Y}_i)\right]| \\ &\leq \frac{Cst}{n^2 h^{d+1}} \left(1 + \sum_{i \neq k} \beta_{|k-i|}^{\delta/(1+\delta)}\right) \\ &= O\left(\frac{1}{nh^{1+d}}\right) << \frac{1}{\sqrt{nh^d}}, \end{aligned}$$

for every $\delta > 0$. Thus, the expectation of $\sqrt{nh^d}\tilde{R}_1(\mathbf{u})$ tends to zero, under our assumptions.

Moreover,

$$\begin{aligned} E[\tilde{R}_1^2(\mathbf{u})] &= \frac{1}{n^4 h^2} \sum_{i,j} \sum_{k,l} E\left[(dK)_h(\mathbf{u} - \mathbf{Y}_i).(\mathbf{1}(\mathbf{Y}_k \leq \mathbf{Y}_i) - \mathbf{Y}_i)\right. \\ &\cdot \left.(dK)_h(\mathbf{u} - \mathbf{Y}_j).(\mathbf{1}(\mathbf{Y}_l \leq \mathbf{Y}_j) - \mathbf{Y}_j)\right] \equiv \frac{1}{n^4 h^2} \sum_{i,j} \sum_{k,l} V_{i,j,k,l}. \end{aligned}$$

We have to consider successively every orders between the four indices i, j, k, l to apply lemma 6. Let $m = m(n)$ be a sequence of integers such that $m \leq n$, $m_n \to \infty$ and $m_n/n \to 0$. For convenience, we choose $m_n = (\ln n)^2$. Note that $n^\alpha \beta_{m_n}^\gamma$ tends to zero for every $\alpha > 0$ and $\gamma > 0$. Therefore,

- If $i \leq j \leq k \leq l$ and $|j - k| > m_n$: apply lemma 6 by setting $(i_1, i_2) = (i, j)$ and $(i_3, i_4) = (k, l)$ to get

  $$\begin{aligned} &\Big| V_{ijkl} - \int (dK)_h(\mathbf{u} - \mathbf{y}_i).(\mathbf{1}(\mathbf{y}_k \leq \mathbf{y}_i) - \mathbf{y}_i).\tau_{(\mathbf{Y}_i, \mathbf{Y}_j)}(\mathbf{y}_i, \mathbf{y}_j) \\ &\cdot \quad (dK)_h(\mathbf{u} - \mathbf{y}_j).(\mathbf{1}(\mathbf{y}_l \leq \mathbf{y}_j) - \mathbf{y}_j).\tau_{(\mathbf{Y}_k, \mathbf{Y}_l)}(\mathbf{y}_k, \mathbf{y}_l)\, d\mathbf{y}_i\, d\mathbf{y}_j\, d\mathbf{y}_k\, d\mathbf{y}_l \Big| \leq \frac{Cst}{h^{2d}} \beta_m^{\delta/(\delta+1)} \end{aligned} \quad (E.1)$$

  for every $\delta > 0$.

  If, moreover, $|k - l| > m_n$, apply lemma 6 again with $i_1 = k$ and $i_2 = l$. We obtain

  $$\begin{aligned} &\Big| \int (dK)_h(\mathbf{u} - \mathbf{y}_i).(\mathbf{1}(\mathbf{y}_k \leq \mathbf{y}_i) - \mathbf{y}_i).\tau_{(\mathbf{Y}_i, \mathbf{Y}_j)}(\mathbf{y}_i, \mathbf{y}_j) \\ &\cdot \quad (dK)_h(\mathbf{u} - \mathbf{y}_j).(\mathbf{1}(\mathbf{y}_l \leq \mathbf{y}_j) - \mathbf{y}_j).\tau_{(\mathbf{Y}_k, \mathbf{Y}_l)}(\mathbf{y}_k, \mathbf{y}_l)\, d\mathbf{y}_i\, d\mathbf{y}_j\, d\mathbf{y}_k\, d\mathbf{y}_l - 0 \Big| \leq \frac{Cst}{h^{2d}} \beta_m^{\delta/(\delta+1)}, \end{aligned}$$

for every $\delta > 0$.

Otherwise, $|k-l| \leq m$. Then there are two sub-cases. When $|i-j| > m$, apply lemma 6 again by setting $i_1 = i$ and $i_2 = j$ to get

$$\begin{aligned} & \left| \int (dK)_h(\mathbf{u}-\mathbf{y}_i).(\mathbf{1}(\mathbf{y}_k \leq \mathbf{y}_i) - \mathbf{y}_i).(dK)_h(\mathbf{u}-\mathbf{y}_j).(\mathbf{1}(\mathbf{y}_l \leq \mathbf{y}_j) - \mathbf{y}_j) \right. \\ & \cdot \quad \left. \left[\tau_{(\mathbf{Y}_i,\mathbf{Y}_j)}(\mathbf{y}_i,\mathbf{y}_j) - \tau(\mathbf{y}_i)\tau(\mathbf{y}_j) \right] \tau_{(\mathbf{Y}_k,\mathbf{Y}_l)}(\mathbf{y}_k,\mathbf{y}_l)\, d\mathbf{y}_i\, d\mathbf{y}_j\, d\mathbf{y}_k\, d\mathbf{y}_l \right| \\ & \leq \quad \frac{Cst}{h^{2d}} \beta_m^{\delta/(\delta+1)}, \end{aligned} \tag{E.2}$$

for every $\delta > 0$. Invoking equation (A.1), we have

$$\begin{aligned} & \int (dK)_h(\mathbf{u}-\mathbf{y}_i).(\mathbf{1}(\mathbf{y}_k \leq \mathbf{y}_i) - \mathbf{y}_i).(dK)_h(\mathbf{u}-\mathbf{y}_j).(\mathbf{1}(\mathbf{y}_l \leq \mathbf{y}_j) - \mathbf{y}_j)\tau(\mathbf{y}_i)\tau(\mathbf{y}_j) \\ & \quad \tau_{(\mathbf{Y}_k,\mathbf{Y}_l)}(\mathbf{y}_k,\mathbf{y}_l)\, d\mathbf{y}_i\, d\mathbf{y}_j\, d\mathbf{y}_k\, d\mathbf{y}_l = \int \left\{ \tau(\mathbf{u}) \sum_{r=1}^d K_r\left(\frac{u_r - y_{kr}}{h}\right) + O(h)\phi(\mathbf{u}) \right\} \\ & \cdot \quad \left\{ \tau(\mathbf{u}) \sum_{s=1}^d K_s\left(\frac{u_s - y_{ls}}{h}\right) + O(h)\phi(\mathbf{u}) \right\} \tau_{(\mathbf{Y}_k,\mathbf{Y}_l)}(\mathbf{y}_k,\mathbf{y}_l)\, d\mathbf{y}_k\, d\mathbf{y}_l = O(h). \end{aligned} \tag{E.3}$$

The number of terms of this type in the summation $\tilde{R}_1(\mathbf{u})$ is $O(mn^3)$. Recalling equations (E.1), (E.2) and (E.3), we get that the sum over all the considered indices (i,j,k,l) is

$$\frac{1}{n^4h^2} \sum_{i,j,k,l} V_{i,j,k,l} = O\left(\frac{mh}{nh^2} + \frac{1}{h^{2+2d}}\beta_m^{\delta/(\delta+1)}\right) << \frac{1}{nh^d}.$$

Else, if $|i-j| \leq m/2$, the number of terms of this type in $\tilde{R}_1(\mathbf{u})$ is $O(m^2n^2)$. Thus, applying equation (E.1), we get the corresponding sum is

$$\frac{1}{n^4h^2} \sum_{i,j,k,l} V_{i,j,k,l} = O\left(\frac{m^2}{n^2h^{2+2d}} + \frac{1}{h^{2+2d}}\beta_m^{\delta/(\delta+1)}\right) << \frac{1}{nh^d}.$$

- If $i \leq j \leq k \leq l$ and $|j-k| \leq m_n$: If $|l-k| > m$, apply lemma 6 by setting $(i_1,i_2,i_3) = (i,j,k)$ and $i_4 = l$. Else, the number of terms in the sum $\tilde{R}_1(\mathbf{u})$ is $O(m^2n^2)$. By summing them, we get an $O(m^2/(n^2h^{2+2d}))$.
- If $i \leq k \leq j \leq l$: apply lemma 6 by setting $(i_1,i_2,i_3) = (i,k,j)$ and $i_4 = l$, if $|j-l| > m$. Else, apply lemma 6 with $(i_1,i_2) = (i,k)$ and $(i_3,i_4) = (j,l)$. If, moreover, $|i-k| > m$, apply lemma 6 again with $i_1 = i$ and $i_2 = k$. Else, the number of corresponding terms is $O(m^2n^2)$.

- If $i \leq l \leq j \leq k$: If $|j-k| > m$, apply lemma 6 by setting $(i_1, i_2, i_3) = (i, l, j)$ and $i_4 = k$. Else, if $|j-k| \leq m$ and $|i-l| \leq m$, the number of such terms in the summation is $O(m^2 n^2)$.

  Otherwise, if $|j-k| \leq m$, $|i-l| > m$ and $|j-l| > m$, apply lemma 6 a first time with $(i_1, i_2) = (i, l)$ and $(i_3, i_4) = (j, k)$, and a second time with $i_1 = i$ and $i_2 = l$. If $|j-k| \leq m$, $|i-l| > m$ and $|j-l| \leq m$, the number of summands is $O(m^2 n^2)$. The upper bound follows by applying the previous techniques.

All the other orders between the four indices i, j, k, l can be dealt similarly. Thus, $E[\tilde{R}_1^2(\mathbf{u})] = o(n^{-1} h^{-d})$ and

$$\tilde{R}_1(\mathbf{u}) = o_P(n^{-1/2} h^{-d/2}).$$

Moreover, from proposition 7.1 in Rio [55], when the cdf F_k are continuous, we have

$$\sup_{\mathbf{u} \in \mathbb{R}^d} \sup_{k=1,\ldots,d} |F_{n,k}(\mathbf{u}) - F_k(\mathbf{u})| = O_P\left(\frac{\ln n}{\sqrt{n}}\right). \tag{E.4}$$

Thus, $\tilde{R}_2(\mathbf{u}) = O_P(h^{-2-d} \ln^2 n / n)$ and $\sqrt{nh^d} \tilde{R}_2(\mathbf{u})$ tends to zero in probability, under our assumptions.

At last, by revisiting Theorem 2.3 of Bosq [5] with our assumptions, we can prove easily the asymptotic normality of the joint vector $\sqrt{nh^d}((\tau_n^* - K_h * \tau)(\mathbf{u}_1), \ldots, (\tau_n^* - K_h * \tau)(\mathbf{u}_m))$. We omit the details. This concludes the proof. $\square$.

Acknowledgment. We thank P. Doukhan, C. Gouriéroux and M. Wegkamp for helpful discussions and comments.

References

[1] C. Ai, A semiparametric maximum likelihood estimator, Econometrica 65 (1997) 933 – 963.

[2] Y. Aït-Sahalia, Testing continuous-time models of the spot interest rate, Review Fin. Studies 9 (1996) 385 – 426.

[3] P.K. Andersen, C. Ekstrøm, J.P. Klein, Y. Shu and M.J. Zhang, Testing the Goodness-of-fit of a copula based on right censored data, mimeo (2002) available on the web.

[4] P. Bickel and M. Rosenblatt, On some global measures of the deviation of density function estimates, Ann. Statist. 1 (1973) 1071 – 1095.

[5] D. Bosq, Nonparametric statistics for stochastic processes. Lecture Notes in Statistics 110, Springer, 1998.

[6] D. Bosq and J-P. Lecoutre, Théorie de l'estimation fonctionnelle, Economica, Paris, 1987.

[7] E. Bouyé, V. Durrleman, A. Nikeghbali, G. Riboulet, and T. Roncalli, Copulas for finance. A reading guide and some applications, GRO, Crédit Lyonnais, 2000. Available on the web.

[8] W. Breyman, A. Dias and P. Embrechts, Dependence structures for multivariate high-frequency data in finance, Quantitative finance 3 (2003) 1-16.

[9] A. Cebrian, M. Denuit and O. Scaillet, Testing for Concordance Ordering, DP FAME 41 (2002).

[10] X. Chen and Y. Fan, Estimation of Copula-based semiparametric time series models, mimeo (2002).

[11] X. Chen, Y. Fan and A. Patton, Simple tests for models of dependence between financial time series: with applications to US equity returns and exchange rates, mimeo (2003).

[12] U. Cherubini and E. Luciano, Value at Risk trade-off and capital allocation with copulas, Economic notes 30, (2) (2000).

[13] R. B. D'Agostino and M. A. Stephens, Goodness-of-fit techniques, Marcel Dekker, 1986.

[14] P. Deheuvels, La fonction de dépendance empirique et ses propriétés. Un test non paramétrique d'indépendance. Acad. R. Belg., Bull. Cl. Sci., 5 Série 65 (1979) 274-292.

[15] P. Deheuvels, A Kolmogorov-Smirnov type test for independence and multivariate samples, Rev. Roum. Math. Pures et Appl., Tome XXVI, 2 (1981a) 213-226.

[16] P. Deheuvels, A Nonparametric Test of Independence, Publications de l'ISUP 26 (1981b) 29-50.

[17] R. de Matteis, Fitting copulas to data, Diploma thesis, Univ. Zurich (2001).

[18] P. Doukhan, Mixing, properties and examples, New-York, Springer, 1994.

[19] Y. Fan, Testing the goodness of fit of a parametric density function by kernel method. Econometric Theory 10 (1994) 316 – 356.

[20] Y. Fan and Q. Li, A general nonparametric model specification test, Manuscript, Departments of Economics, University of Windsor (1992).

[21] Y. Fan and Q. Li, Central limit theorem for degenerate U-statistics of absolutely regular processes with applications to model specification testing, J. Nonparametric Stat. 10 (1999) 245 – 271.

[22] Y. Fan and Q. Li, Consistent model specification tests, Econom. theory 16 (2000) 1016 – 1041.

[23] Y. Fan and A. Ullah, On goodness-of-fit tests for weakly dependent processes using kernel method, J. Nonparametric Statist. 11 (1999) 337 – 360.

[24] J-D. Fermanian, D. Radulovic, and M. Wegkamp, Weak convergence of empirical copula process, Working Paper CREST (2002) 2002 – 06.

[25] J-D. Fermanian and O. Scaillet, Nonparametric estimation of copulas for time series, Journal of Risk 5, Vol 4 (2002) 25-54.

[26] R. Foutz, A test for goodness-of-fit based on an empirical probability measure, Ann. Statist., 8 (1980) 989 – 1001.

[27] E. Frees and E. Valdez, Understanding Relationships Using Copulas, North American Actuarial J., 2 (1998) 1 – 25.

[28] R. Frey and A. McNeil, Modelling dependent defaults, working paper (2001).

[29] C. Genest and C. Rivest, Statistical inference procedures for bivariate archimedian copulas, J. Amer. Statist. Ass. 88 (1993) 1034 – 1043.

[30] C. Genest, K. Ghoudi and C. Rivest, A semiparametric estimation porcedure of dependence parameters in multivariate families of distributions, Biometrika 82 (1995) 543 – 552.

[31] C. Genest, R.W.J. van den Goorbergh and B. Werker, Multivariate option pricing using dynamic copula models, mimeo (2003).

[32] I. Gijbels and J. Mielniczuc, Estimating the density of a copula function, Comm. Stat. Theory Methods 19 No 2 (1990) 445 – 464.

[33] C. Gouriéroux and C. Tenreiro, Local power properties of kernel based goodness-of-fit tests, J. Multivariate Anal. 78 (2000) 161 – 190.

[34] C. Gouriéroux and P. Gagliardini, Constrained nonparametric copulas, Working paper CREST (2002).

[35] P. Hall, Central limit theorem for integrated square error of multivariate nonparametric density estimators, J. Multivariate Analysis, 14 (1984) 1 – 16.

[36] V. Hjellvik, Q. Yao and V. Tjøstheim, Linearity testing using local polynomial approximation. J. Statist. Plann. Inf. 68(1998) 295 – 321.

[37] W. Härdle and E. Mammen, Comparing nonparametric versus parametric regression fits, Ann. Statist., 21 (1993) 1926 – 1947.

[38] L. Hu, Dependence patterns accross financial markets : methods and evidence, mimeo (2002).

[39] A. Justel, D. Pena and R. Zamar, A multivariate Kolmogorov-Smirnov test of goodness of fit, Statist. Prob. Letters 35 (1997) 251 – 259.

[40] Y. Hong and H. Li, Nonparametric specification testing for continuous-time models with application to spot interest rates, mimeo (2002).

[41] E.V. Khmaladze, Martingale approach to the theory of goodness of fit tests, Theory Prob. Appl. 26 (1981) 240 – 257.

[42] E.V. Khmaladze, An innovation approach in goodness-of-fit tests in $\mathbb{R}^m$, Ann. Statist., 16 (1988) 1503 – 1516.

[43] E.V. Khmaladze, Goodness of fit problem and scanning innovation martingales, Ann. Statist. 21 (1993) 798 – 829.

[44] J-P. Laurent and J. Gregory, Basket default swaps, CDO's and factor copulas, mimeo (2002).

[45] D. Li, On default correlation : A copula function approach. Journal of Fixed Income 9 (2000) 43-54.

[46] E. Masry and D. Tjøstheim, Nonparametric estimation and identification of nonlinear ARCH time series : strong convergence and asymptotic normality, Econometric Theory, 11 (1995) 258 – 289.

[47] A. Mokkadem, Sur un modèle autoregressif non linéaire, ergodicité et ergodicité géométrique, J. Time Ser. Anal., 8 (1987) 195 – 204.

[48] A. Mokkadem, Mixing properties of ARMA processes, Stochastic. Process. Appl., 29 (1988) 309 – 315.

[49] A. Patton, Modelling Time-Varying Exchange Rate Dependence Using the Conditional Copula, UCSD WP 2001-09 (2001a).

[50] A. Patton, Estimation of Copula Models for Time Series of Possibly Different Lengths, UCSD WP 2001-17 (2001b).

[51] D. Pham, The mixing property of bilinear and generalized random coefficient autoregressive models, Stochastic Process. Appl., 23 (1986) 291 – 300.

[52] D. Pham and L. Tran, Some mixing properties of time series models, Stochastic Process. Appl., 19 (1985) 297 – 303.

[53] D. Pollard, General Chi-square goodness-of-fit tests with data-dependent cells, Z. Wahrsch. verw. Gebiete, 50 (1979) 317 – 331.

[54] W. Polonik, Concentration and goodness-of-fit in higher dimensions: (Asymptotically) distribution-free methods, Ann. Stat, 27 (1999) 1210-1229.

[55] E. Rio, Théorie asymptotique des processus aléatoires faiblement dépendants. SMAI, Mathématiques et applications, Springer, 2000.

[56] M. Rockinger and E. Jondeau, Conditional Dependency of Financial Series: An Application of Copulas. HEC Paris DP 723 (2001).

[57] J. Rosenberg, Nonparametric pricing of multivariate contingent claims, NYU, Stern School of Business (2001).

[58] M. Rosenblatt, Remarks on a Multivariate Transformation, Ann. Statist., 23 (1952) 470-472.

[59] M. Rosenblatt, Quadratic measure of deviation of two dimensional density estimates and a test of independence Ann. Statist. 3 (1975) 1 – 14.

[60] R. Saunders and P. Laud, The multidimensional Kolmogorov goodness-of-fit test, Biometrika 67 (1980) 237.

[61] P. Schönbucher, Factor models for portfolio credit risk, Univ. Bon, discussion paper 16/2001 (2001).

[62] P. Schönbucher and D. Schubert, Copula dependent default risk in intensity models, Univ. Bonn, mimeo (2001).

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

[64] B. Silverman, Density estimation for statistics and data analysis, Chapman & Hall, 1986.

[65] K-I. Yoshihara, Limiting behavior of U-statistics for stationary absolutely regular processes, Z. Wharscheinlichkeitstheorie verw. Gebiete 35 (1976) 237 – 252.

[66] J. Zheng, A consistent test of functional form via nonparametric estimation technique, J. Econometrics 75 (1996) 263 – 289.