An Empirical Central Limit Theorem with applications to copulas under weak dependence

Paul Doukhan (doukhan@ensae.fr)
ENSAE-CREST, J120, 3 Av. Pierre Larousse, 92245 Malakoff Cedex and SAMOS, Université Paris 1, Centre Pierre Mendès, 90 rue de Tolbiac, F-75634, Paris Cedex 13, France.

Jean-David Fermanian (fermania@ensae.fr)
CREST, J320, 15 bd Gabriel Péri, 92245 Malakoff Cedex, France.

Gabriel Lang (Corresponding author.) (lang@engref.fr)
GRESE, ENGREF, 19 av du Maine 75732 Paris Cedex 15, France.

Abstract. We state a multidimensional Functional Central Limit Theorem for weakly dependent random vectors. We apply this result to copulas. We get the weak convergence of the empirical copula process and of its smoothed version. The finite dimensional convergence of smoothed copula densities is also proved. A new definition and the theoretical analysis of conditional copulas and their empirical counterparts are provided.

Keywords: Copulas, multivariate FCLT, weak dependence.

AMS classification (2000): 62M10, 62G07, 60F17

1. Introduction

This paper is devoted to asymptotic results relative to the empirical process for weakly dependent sequences. Various definitions of weak dependence have been introduced in the literature. Among them, α-mixing and β-mixing have been developped, but these notions are not fully satisfactory, as they are defined with respect to filtrations and difficult to check in practice. Doukhan and Louhichi [11] introduce a definition of weak dependence that is easier to check on various examples of stationary processes (see Doukhan [8]). Various applications and developments of weak dependence are addressed in Ango Nze, Bühlmann and Doukhan[2] and Ango Nze and Doukhan [1].

The general notion of weak dependence corresponds to the following idea. Consider two finite samples with time indices P in the past and in the future F, separated by a gap r. The independence of P and F is equivalent to $\text{cov}(f(F), g(P)) = 0$ for a suitable class of measurable functions. A natural way to weaken this condition is to provide a precise control of these covariances as the gap r becomes larger, and to fix the rate of decrease of the control as r tends to infinity. Moreover the class of functions will be reduced to Lipschitz functions to make the weak dependence condition easy to check for a wide class of models.

Section 2 introduces the definition of weak dependence, provides examples and the functional central limit theorem for the multivariate empirical process. Section 3 is devoted to applications of the main theorem to copulas processes. The last section contains the proofs.

2. Definitions and main result

2.1. Weak dependence

We state here the definition of weak dependence that is used in the paper and a refinement of it. Define the Lipschitz modulus of a real function h on a space $\mathbb{R}^d$ as

$$\text{Lip }(h) = \sup_{\mathbf{x}\neq\mathbf{y}} \frac{|h(\mathbf{x}) - h(\mathbf{y})|}{\|\mathbf{x}-\mathbf{y}\|_1},$$

where $\|\mathbf{x}\|_1 = \|(x_1,\ldots,x_d)\|_1 = \sum_{i=1}^d |x_i|$. Define $\Lambda^{(1)}$ as the set of functions that are bounded by 1 and have a finite Lipschitz modulus. Declare two sequences of indices $i_1 \le \cdots \le i_u$ and $j_1 \le \cdots \le j_v$ as r-distant if $i_u \le j_1$ and $j_1 - i_u = r$.

Definition 1. *[Doukhan & Louhichi, 1999)] Let $\eta = (\eta_r)_{r\geq 0}$ (resp. $\theta = (\theta_r)_{r\geq 0}$) be a real positive sequence that tends to zero. We say that the d-dimensional process $(\xi_i)_{i\in\mathbb{Z}}$ is η-dependent (resp. θ-dependent) if, for any r-distant finite sequences $\mathbf{i} = (i_1,\ldots,i_u)$ and $\mathbf{j} = (j_1,\ldots,j_v)$, for any functions f and g in $\Lambda^{(1)}$ defined on $(\mathbb{R}^d)^u$ and $(\mathbb{R}^d)^v$ respectively, we have*

$$|\text{cov}(f(\xi_{i_1},\ldots,\xi_{i_u}), g(\xi_{j_1},\ldots,\xi_{j_v}))| \le (u\text{Lip } f + v\text{Lip } g)\eta_r, \qquad (1)$$

$$|\text{cov}(f(\xi_{i_1},\ldots,\xi_{i_u}), g(\xi_{j_1},\ldots,\xi_{j_v}))| \le v\text{Lip } g\theta_r. \qquad (2)$$

Remark 1. *The θ-dependence condition corresponds to causal processes and is more restrictive that η-dependence; note that $\eta_r \le \theta_r$. Mathematical advantages of θ-dependence are presented in Dedecker and Doukhan [5]. The forthcoming examples will make clear the differences between the two notions.*

Remark 2. *Note that if ξ is η-dependent and if f and g are bounded Lipschitz functions, the previous covariance is bounded by*

$$(u\text{Lip }(f)\|g\|_\infty + v\|f\|_\infty\text{Lip }(g))\,\eta_r.$$

2.2. Examples.

1. **Stable Markov processes.** Consider first stationary sequences satisfying a recurrence equation

$$X_n = F(X_{n-1}, \dots, X_{n-d}, \xi_n)$$

where the sequence (ξ_n) is iid. In this case $Y_n = (X_n, \dots, X_{n-d+1})$ is a Markov chain such that $Y_n = M(Y_{n-1}, \xi_n)$ with

$$M(x_1, \dots, x_d, \xi) = (F(x_1, \dots, x_d, \xi), x_1, \dots, x_{d-1}).$$

Consider a norm $\|\cdot\|$ on $\mathbb{R}^d$, then we set $\|Z\|_m = (\mathbb{E}\|Z\|^m)^{1/m}$ for $m \ge 1$ for any $\mathbb{R}^d$-valued random variable Z. Stationarity follows from Duflo's Theorem 1.IV.24 [16] if $\|F(0,\xi)\|_m < \infty$ and $\|F(x,\xi) - F(y,\xi)\|_m \le a\|x-y\|$ for some real $0 \le a < 1$ and $m \ge 1$. Here θ-dependence holds with $\theta_r = \mathcal{O}(a^{r/d})$(hence $\eta_r = \mathcal{O}(a^{r/d})$ too) (as $r \uparrow \infty$) for the following examples:

   – *Functional AR models*: $X_t = r(X_{t-1}, \dots, X_{t-d}) + \xi_t$ if $\|\xi_0\|_m < \infty$ and $|r(u_1, \dots, u_d) - r(v_1, \dots, v_d)| \le \sum_{i=1}^d a_i |u_i - v_i|$ for some $a_1, \dots, a_d \ge 0$ with $\sum_{i=1}^d a_i < 1$.

   – *Branching processes models.* Here $d = 1$, and $D \ge 2$. Set $\xi_t = \left(\xi_t^{(1)}, \dots, \xi_t^{(D)}\right)$. Let now $A_1(u), \dots, A_D(u)$ be Lipschitz functions (with $u \in \mathbb{R}$), and for $(u, z^{(1)}, \dots, z^{(D)}) \in \mathbb{R}^{D+1}$ let

$$M\left(u, \left(z^{(1)}, \dots, z^{(D)}\right)\right) = \sum_{j=1}^{D} A_j(u) z^{(j)}.$$

   $\mathbb{L}^m-$stationarity holds if $a = \sum_{j=1}^D \text{Lip}\,(A_j)\|\xi_0^{(j)}\|_m < 1$.

The following examples are not necessarily Markov models.

2. **Bernoulli shifts.** Let $H : \mathbb{R}^{\mathbb{Z}} \to \mathbb{R}^d$ be a measurable function. If the sequence $(\xi_n)_{n\in\mathbb{Z}}$ is independent and identically distributed on the real line, a Bernoulli shift with innovation process $(\xi_n)_{n\in\mathbb{Z}}$ is defined as

$$X_n = H\left((\xi_{n-i})_{i\in\mathbb{Z}}\right), \qquad n \in \mathbb{Z}.$$

A simple case of infinitely dependent Bernoulli shift is the moving average process, where the function H corresponds to a series. Assume that there exists a control of the functional dependence to the tail variables, *i.e.* a sequence δ_r decreasing to zero such that:

$$\mathbb{E}\left\|H\left(\xi_j, j \in \mathbb{Z}\right) - H\left(\xi_j \mathbf{1}_{|j|\le r}, j \in \mathbb{Z}\right)\right\| \le \delta_r. \tag{3}$$

where $\|\cdot\|$ is a norm on $\mathbb{R}^d$. Then the process is $\eta-$weakly dependent with $\eta_r \le 2\delta_{[r/2]}$, see Doukhan & Louhichi [11]. If $H\,(x_j, j \in \mathbb{Z})$ does not depend on the x_j's with $j < 0$, then the process is causal and θ-dependence holds with $\theta_r = \delta_r$.

A first example is a Volterra stationary process defined through a convergent Volterra expansion

$$X_t = v_0 + \sum_{k=1}^{\infty} V_{k;t}, \quad V_{k;t} = \sum_{-\infty<i_1<\cdots<i_k<\infty} a_{k;i_1,\ldots,i_k} \xi_{t-i_1} \cdots \xi_{t-k}$$

where v_0 denotes a constant and $(a_{k;i_1,\ldots,i_k})_{(i_1,\ldots,i_k)\in\mathbb{Z}^k}$ are real numbers for each $k \ge 1$. This expression converges in $\mathbb{L}^m$ for $m \ge 1$, provided that $\mathbb{E}|\xi_0|^m < \infty$ and $\sum_{k=1}^{\infty} \sum_{i_1<\cdots<i_k} |a_{k;i_1,\ldots,i_k}| < \infty$. Those models are η-dependent since (3) is satisfied, δ_r corresponding to the tail of the previous series.

The following examples illustrate this general class of models.

3. **LARCH(∞) models.** A vast literature is devoted to the study of conditionally heteroscedastic models. A simple equation in terms of a vector valued process allows a unified treatment of those models, see [15]. Let $(\xi_t)_{t\in\mathbb{Z}}$ be an iid sequence of random $d \times D$-matrices, $(A_j)_{j\in\mathbb{N}^*}$ be a sequence of $D \times d$ matrices, and a be a vector in $\mathbb{R}^D$. A vector valued LARCH(∞) model is a solution of the recurrence equation

$$X_t = \xi_t \left(a + \sum_{j=1}^{\infty} A_j X_{t-j} \right) \tag{4}$$

We provide below sufficient conditions for the following chaotic expansion

$$X_t = \xi_t \left(a + \sum_{k=1}^{\infty} \sum_{j_1,\ldots,j_k \ge 1} A_{j_1} \xi_{t-j_1} A_{j_2} \cdots A_{j_k} \xi_{t-j_1-\cdots-j_k} a \right) \tag{5}$$

Such LARCH(∞) models include a large variety of models, as

– Standard LARCH(∞) models, correspond to the case of real valued X_t and a_j.

– Bilinear model $X_t = \zeta_t \left(\alpha + \sum_{j=1}^{\infty} \alpha_j X_{t-j}\right) + \beta + \sum_{j=1}^{\infty} \beta_j X_{t-j}$ where the variables are real valued and ζ_t is the innovation. For this, we set $\xi_t = \begin{pmatrix} \zeta_t \\ 1 \end{pmatrix}$, $a = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ and $A_j = \begin{pmatrix} \alpha_j \\ \beta_j \end{pmatrix}$. Expansion (5) coincides with the chaotic expansion in [22].

- GARCH(p,q) models,

$$\begin{cases} r_t = \sigma_t \varepsilon_t \\ \sigma_t^2 = \sum_{j=1}^p \beta_j \sigma_{t-j}^2 + \gamma + \sum_{j=1}^q \gamma_j r_{t-j}^2 \end{cases}$$

  where $\gamma > 0$, $\gamma_i \geq 0$, $\beta_i \geq 0$ (and the variables ε are centered at expectation); this model is a special case of the bilinear model with $\alpha_0 = \frac{\gamma_0}{1-\sum \beta_i}$ et $\sum \alpha_i z^i = \frac{\sum \gamma_i z^i}{1-\sum \beta_i z^i}$ (see [22]).

- ARCH(∞) processes are given by equations,

$$\begin{cases} r_t = \sigma_t \varepsilon_t \\ \sigma_t^2 = \beta_0 + \sum_{j=1}^\infty \beta_j \sigma_{t-j}^2 \end{cases}$$

  One sets $\xi_t = \begin{pmatrix} \varepsilon_t & 1 \end{pmatrix}$, $a = \begin{pmatrix} \kappa\beta_0 \\ \lambda_1\beta_0 \end{pmatrix}$, $A_j = \begin{pmatrix} \kappa\beta_j \\ \lambda_1\beta_j \end{pmatrix}$ with $\lambda_1 = \mathbb{E}(\varepsilon_0^2)$, $\kappa^2 = \mathrm{Var}\ (\varepsilon_0^2)$.

Endow the sets of matrices with a norm $\|\cdot\|$ of algebra, derived from a norm for linear applications. Assume that $\Lambda = \|\xi_0\|_m \sum_{j\geq 1} \|A_j\| < 1$ then one stationary of solution of eqn. (4) in $\mathbb{L}^m$ is given as (5). The solution (5) of eqn. (4) is $\theta-$weakly dependent with

$$\theta_r = \left(\mathbb{E}\|\xi_0\| \sum_{k=1}^{r-1} k\Lambda^{k-1} R\left(\frac{r}{k}\right) + \frac{\Lambda^r}{1-\Lambda} \right) \mathbb{E}\|\xi_0\|\|a\|,$$

where $R(x) = \sum_{j\geq x} \|a_j\|$. There exists some constant $K>0$ and $b, C > 0$ such that

$$\theta_r \leq \begin{cases} K\frac{(\log(r))^{b\vee 1}}{r^b}, \text{ under Riemaniann decay } A(x) \leq Cx^{-b}, \\ K(q\vee\Lambda)^{\sqrt{r}}, \text{ under geometric decay } A(x) \leq Cq^x. \end{cases}$$

4. **Non-causal** *LARCH*(∞) **model.** Now A_j is defined for $j \neq 0$. Doukhan, Teyssière and Winant (2005) prove the same results of existence as for the previous causal case (replace summation for $j > 0$ by summation for $j \neq 0$) and the process is now η-weakly dependent with

$$\eta_r = \left(\|\xi_0\|_\infty \sum_{0\leq 2k<r} k\Lambda^{k-1} R\left(\frac{r}{2k}\right) + \frac{\Lambda^{r/2}}{1-\Lambda} \right) \mathbb{E}\|\xi_0\|\|a\|$$

where now

$$R(x) = \sum_{|j|\geq x} \|a_j\|, \qquad \Lambda = \|\xi_0\|_\infty \sum_{j\geq 1} \|A_j\| < 1.$$

Here we need the restrictive assumption that innovations are uniformly bounded.

2.3. Multivariate empirical central limit theorem

The main theoretical result of the paper is a functional central limit theorem for η-dependent vector-valued sequences $(\mathbf{X}_i)_{i\in\mathbb{Z}}$. It is an extension of the independent case, where the limit process in the space of càdlàg functions $D([0,1]^d)$ endowed with the Skorohod metric d_S is known to be a multivariate Brownian bridge $\mathbb{B}_0$, i.e. a Gaussian process with covariance function

$$\mathrm{cov}(\mathbb{B}_0(\mathbf{x}),\mathbb{B}_0(\mathbf{y})) = \mathbb{P}\left(\mathbf{X}_0 \le \mathbf{x}\wedge\mathbf{y}\right) - \mathbb{P}(\mathbf{X}_0 \le \mathbf{x})\mathbb{P}(\mathbf{X}_0 \le \mathbf{y}), \qquad (6)$$

for every vectors $\mathbf{x}$ and $\mathbf{y}$ in $[0,1]^d$. Here the order relation in $[0,1]^d$ is partial : $\mathbf{x}\le\mathbf{y}$ if it holds for every coordinates and $\mathbf{x}\wedge\mathbf{y} = (x_i \wedge y_i)_{i=1,\dots,d}$. In the case of weak dependence, the limiting distributions are not free of the distribution's process.

In this section, $\mathbf{Y}$ is a process with uniform marginal distributions and cdf F. We denote the empirical cdf:

$$F_n(\mathbf{x}) = n^{-1}\sum_{i=1}^n \mathbf{1}\{Y_{i,1}\le x_1,\dots,Y_{i,d}\le x_d\}, \qquad (7)$$

and define the normalized empirical process $B_n = \sqrt{n}(F_n - F)$ associated with $\mathbf{Y}$. Consider a centered Gaussian process $\mathbb{B}$ such that, for any vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^d$,

$$\mathrm{cov}(\mathbb{B}(\mathbf{u}),\mathbb{B}(\mathbf{v})) = \sum_{i\in\mathbb{Z}} \mathrm{cov}\left(\mathbf{1}\{\mathbf{Y}_0\le\mathbf{u}\},\mathbf{1}\{\mathbf{Y}_i\le\mathbf{v}\}\right). \qquad (8)$$

Note that the previous covariance structure depends on the joint distribution of $\mathbf{Y}_0$ and $\mathbf{Y}_i$, for every i. We consider a dependence relation based on the covariance of some indicator functions. The link with the weak dependence is given in Lemma 2.1.

Definition 2. *Let f be a function on $\mathbb{R}^u$, $\mathbf{i} = (i_1,\dots,i_u)$ be a sequence of elements in $\mathbb{Z}$ and $\mathbf{s} = (s_1,\dots,s_u)$ be a sequence of elements in $[0,1]^d$. With implicit reference to a process $\mathbf{Y}$, we define*

$$Z(f,\mathbf{i},\mathbf{s}) = f(\mathbf{1}\{\mathbf{Y}_{i_1}\le s_1\},\dots,\mathbf{1}\{\mathbf{Y}_{i_u}\le s_u\}).$$

Define $a_d^* := d + \sqrt{1+d^2}$. The main result of the paper is the following:

Theorem 1. *Assume that* $(\mathbf{Y}_i)_{i\in\mathbb{Z}}$ *is a centered process with uniform marginal distributions such that for any* r*-distant finite sequences* $\mathbf{i} = (i_1, \ldots, i_u)$ *and* $\mathbf{j} = (j_1, \ldots, j_v)$*, for any functions* f *and* g *in* $\Lambda^{(1)}$*, defined on* $\mathbb{R}^u$ *and* $\mathbb{R}^v$*:*

$$|\mathrm{cov}(Z(f, \mathbf{i}, \mathbf{s}), Z(g, \mathbf{j}, \mathbf{t}))| \leq (u\mathrm{Lip}\ f + v\mathrm{Lip}\ g)\eta_r. \tag{9}$$

Assume that there exist some constants $C > 0$ *and* $a > a_d^*$ *such that* $\eta_r \leq Cr^{-a}$*. Then* B_n *tends to* $\mathbb{B}$ *in distribution in* $D([0,1]^d, d_S)$*.*

See the proof in section 4. The following lemma is essential to apply theorem 1.

Lemma 2.1. *If* $(\mathbf{Y}_i)_{i\in\mathbb{Z}}$ *is* η*-dependent (with dependence coefficients* $\eta_{Y,r}$*), then condition (9) is satisfied with* $\eta_r = 3(\eta_{Y,r}d)^{\frac{1}{2}}$*.*

Proof. Define ϵ-approximations of $\mathbf{1}\{x \geq t\}$ by

$$h_{\epsilon,t}(x) = \prod_{p=1}^{d} \frac{(x^{(p)} - t^{(p)} + \epsilon)}{\epsilon} \mathbf{1}\{t^{(p)} - \epsilon < x^{(p)} < t^{(p)}\} + \mathbf{1}\{x \geq t\}.$$

Then $h_{\epsilon,t}(x)$ is $1/\epsilon$-Lipschitz, and $\mathbb{E}\|h_{\epsilon,t}(\mathbf{Y}_0) - \mathbf{1}\{\mathbf{Y}_0 \geq t\}\|_1 \leq \epsilon d$. Define the analogous approximation of $Z(f, \mathbf{i}, \mathbf{s})$ by

$$Z_\epsilon(f, \mathbf{i}, \mathbf{s}) = f(h_{\epsilon,s_1}(\mathbf{Y}_1), \ldots, h_{\epsilon,s_u}(\mathbf{Y}_u)).$$

Let f, g be in $\Lambda^{(1)}$ and set for short, $\psi = u\mathrm{Lip}\ f + v\mathrm{Lip}\ g$. Then

$$\begin{aligned}
&|\mathbb{E}\left(Z_\epsilon(f, \mathbf{i}, \mathbf{s})Z_\epsilon(g, \mathbf{j}, \mathbf{t})\right) - \mathbb{E}\left(Z(f, \mathbf{i}, \mathbf{s})Z(g, \mathbf{j}, \mathbf{t})\right)| \\
&\qquad \leq \|Z_\epsilon(f, \mathbf{i}, \mathbf{s})\|_\infty \mathbb{E}\left\|(Z_\epsilon(g, \mathbf{j}, \mathbf{t}) - Z(g, \mathbf{j}, \mathbf{t}))\right\|_1 \\
&\qquad + \|Z(g, \mathbf{j}, \mathbf{t})\|_\infty \mathbb{E}\left\|(Z_\epsilon(f, \mathbf{i}, \mathbf{s}) - Z(f, \mathbf{i}, \mathbf{s}))\right\|_1 \\
&\qquad\qquad \leq (v\mathrm{Lip}\ (g) + u\mathrm{Lip}\ (f))\epsilon d \leq \psi\epsilon d.
\end{aligned}$$

Similarly,

$$|\mathbb{E}\left(Z_\epsilon(f, \mathbf{i}, \mathbf{s})\right)\mathbb{E}\left(Z_\epsilon(g, \mathbf{j}, \mathbf{t})\right) - \mathbb{E}\left(Z(f, \mathbf{i}, \mathbf{s})\right)\mathbb{E}\left(Z(g, \mathbf{j}, \mathbf{t})\right)| \leq \psi\epsilon d. \tag{10}$$

As $\mathbf{Y}$ is η-weak dependent with dependence coefficients $\eta_{Y,r}$, for any r-distant sequences $\mathbf{i}$ and $\mathbf{j}$,

$$|\mathrm{cov}(Z_\epsilon(f, \mathbf{i}, \mathbf{s}), Z_\epsilon(g, \mathbf{j}, \mathbf{t}))| \leq \epsilon^{-1}\psi\eta_{Y,r}.$$

Choosing ϵ such that $\eta_{Y,r}\epsilon^{-1} = \epsilon d$, we get

$$|\mathrm{cov}(Z(f, \mathbf{i}, \mathbf{s}), Z(g, \mathbf{j}, \mathbf{t}))| \leq 3\psi \cdot (\eta_{Y,r}d)^{1/2}.$$

Several applications of theorem 1 are provided by a direct application of the functional delta-method. We shall consider below the empirical and the smoothed copula processes.

3. Applications to copula processes

3.1. Empirical copula processes

Copulas describe the dependence structure between some random vectors. They have been introduced a long time ago (Sklar [30]) and have been rediscovered recently, especially for their applications in finance and biostatistics. Briefly, a d-dimensional copula is a cdf on $[0,1]^d$ whose marginal distributions are uniform. It summarizes the dependence structure independently of the specification of the marginal distributions.

Consider a random vector $\mathbf{X} = (X_1, \ldots, X_d)$ whose joint cdf is F and whose marginal cdfs' are denoted by F_j, $j = 1, \ldots, d$. Then there exists a unique copula C defined on the product of the values taken by the r.v. $F_j(X_j)$, such that

$$C(F_1(x_1), \ldots, F_d(x_d)) = F(x_1, \ldots, x_d),$$

for any $\mathbf{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d$. C is called the copula associated with $\mathbf{X}$. When F is continuous, C is defined on $[0,1]^d$. If F is discontinuous, there are several choices to extend C to $[0,1]^d$ (see Nelsen [25] for a complete theory).

Let $(\mathbf{X}_i)_{i\in\mathbb{Z}}$ be a vector valued stationary process. The distribution of $\mathbf{X}_i$ is independent of i and we denote by C its copula which we shall estimate nonparametrically. For example, in the study a real valued stationary Markov sequence $(\mathbf{Z}_i)_{i\in\mathbb{Z}}$, one may consider the vector $\mathbf{X}_i = (\mathbf{Z}_i, \mathbf{Z}_{i+1}, \mathbf{Z}_{i+2})$, as in Chen and Fan [4].

The empirical copula is defined by

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

for every $u_1, \ldots, u_d$ in $[0,1]$. As usual, we denote the empirical cdfs'

$$F_{n,j}(x_j) = n^{-1} \sum_{i=1}^{n} \mathbf{1}\{X_{i,j} \le x_j\},\ j = 1, \ldots, d, \tag{11}$$

and we use the usual generalized inverse notations, for every univariate cdf G, $G^{-1}(u) = \inf\{t | G(t) \ge u\}$.

In the i.i.d. framework the consistency of C_n and the limiting behavior of $n^{1/2}(C_n - C)$ are obtained by Deheuvels ([6], [7]) under the strong assumption of independence between marginals; Gaensler and Stute [20] and Fermanian *et al.* [17] get rid of this restriction. Theorem 1 applied to $\mathbf{Y}_i = (F_1(X_{i,1}), \ldots, F_d(X_{i,d}))$ yields the extension to dependent data:

Theorem 2. *If* $(\mathbf{Y}_i)_{i\in\mathbb{Z}}$ *is* η*-dependent,* $\eta_n = O(n^{-a})$*,* $a > a_d^*$*, if* C *has continuous first partial derivatives, then* $n^{1/2}(C_n - C) \to \mathbb{G}$ *in* $(D([0,1]^d), d_S)$*; the Gaussian limit has continuous sample paths:*

$$\mathbb{G}(\mathbf{u}) = \mathbb{B}(\mathbf{u}) - \sum_{j=1}^{d} \frac{\partial C}{\partial u_j}(\mathbf{u})\mathbb{B}(\mathbf{v}_j), \tag{12}$$

here $\mathbf{v}_j \in [0,1]^d$ *is the vector with components equal to* 1 *excepted for the* j*-th, equal to* u_j*.*

The proof is based on our FCLT, theorem 1, for multivariate weakly dependent sequences. Note that the covariance structure of $n^{1/2}(C_n - C)$ relies on both (12) and (8).

Remark 3. *The same result applies for sequences such that multivariate FCLT holds. We thus quote that theorem 2 still holds under mixing conditions (see examples in Doukhan [9]):*

- *for stationary strongly mixing sequences, if* $\alpha_n = \mathcal{O}(n^{-a})$ *for some* $a > 1$*; we use Rio [27]'s empirical CLT for vector-valued sequences.*
- *in the absolutely regular case Doukhan, Massart and Rio [13]'s result yields assumption* $\beta_n = \mathcal{O}\left(n^{-1}\log^{-b} n\right)$ *for some* $b > 2$*.*

Other results yielding FCLT are recalled in [8].

In practice, smoothed copulas are prefered for graphical representation. Nonparametric estimation is often the first step before a parametric modelisation. For optimization purposes, estimates of the derivatives of underlying copulas are useful, *e.g* for portfolio optimization in a mean-variance framework (Markowitz [24]) or with respect to any other risk measure, estimation of the sensitivities of Value-at-Risk or Expected Shortfall with respect to notional amounts (Gouriéroux *et al.* [23] or Scaillet [29]). The smoothed empirical $\hat{F}_n$ the copula processes in d dimensions writes as:

$$\hat{F}_n(\mathbf{x}) = \int K((\mathbf{x} - \mathbf{v})/h)\, F_n(d\mathbf{v})$$

associated with the usual empirical process F_n (see equation 7), where K is the primitive function of a d-dimensional kernel k subject to the limit condition $\lim_{-\infty} K = 0$, and where $h = h_n$ is a bandwidth. More precisely, $\int k = 1$, $h_n > 0$, and $h_n \to 0$ when $n \to \infty$. Similarly, the j-th marginal cdf F_j is estimated nonparametrically by

$$\hat{F}_{n,j}(x_j) = \int K_j((x_j - v_j)/h)\, F_{n,j}(dv_j),$$

where K_j is the primitive function of a univariate kernel k_j. We assume for simplicity that the bandwidth h is the same for every marginal and that $k(u_1,\dots,u_d)=\prod_{j=1}^d k_j(u_j)$. Then, for every $\mathbf{u}\in[0,1]^d$, the smoothed empirical copula process writes as:

$$\hat{C}_n^{(1)}(\mathbf{u})=\hat{F}_n\left(\hat{F}_{n,1}^{-1}(u_1),\dots,\hat{F}_{n,d}^{-1}(u_d)\right),$$

or by smoothing directly the process C_n,

$$\hat{C}_n^{(2)}(\mathbf{u})=\int K((\mathbf{u}-\mathbf{v})/h)\,C_n(d\mathbf{v}).$$

As in the i.i.d. case, the uniform distance between empirical processes and smoothed empirical processes is $o_P(n^{-1/2})$ under some regularity conditions. To prove this result, we need some technical assumption on the kernels:

Assumption (K). Assume k is p times continuously differentiable, and:

- k is compactly supported, or
- there exists a sequence of positive real numbers a_n such that $h_n a_n$ tends to zero when $n\to\infty$, and

$$n^{1/2}\int_{\{\|\mathbf{v}\|>a_n\}}|k(\mathbf{v})|\,d\mathbf{v}\longrightarrow 0.$$

Moreover, we need:

Lemma 3.1. *Assume (K) and*

(i) the process $n^{1/2}(F_n-F)$ *is stochastically equicontinuous,*

(ii) $\|\mathbb{E}\hat{F}_n-F\|_\infty=o(n^{-1/2})$,

(iii) $nh^{2p}\to 0$.

Then $\|\hat{F}_n-F_n\|_\infty=o_P(n^{-1/2})$.

See the proof in section 4. Assumption (i) is satisfied when $\mathbf{X}$ is compactly supported, invoking theorem 1. We get assumption (ii) by assuming some regularity on F, e.g. F is p-times continuously differentiable. Therefore, folllowing the proof of theorem 10 in Fermanian *et al.* [17], we get:

Theorem 3. *Assume (K) and*

- *the process* $n^{1/2}(F_n-F)$ *is stochastically equicontinuous,*

- $(\mathbf{Y}_i)_{i\in\mathbb{Z}}$ *is* η*-dependent,* $\eta_n = O(n^{-a})$*,* $a > a_d^*$*,*
- F *is* p*-times continuously differentiable,*
- $nh^{2p} \to 0$*.*

Then $n^{1/2}(\hat{C}_n^{(1)} - C) \to \mathbb{G}$ *in* $(D([0,1]^d), d_S)$*.*

This result extends for weakly dependent processes the result on finite dimensional distributions in Fermanian and Scaillet [18]. Moreover, we can prove lemma 3.1 replacing F_n by C_n exactly by the same ways. Hence (see theorem 11 in [17]):

Theorem 4. *Assume (K) and*

- $(\mathbf{Y}_i)_{i\in\mathbb{Z}}$ *is* η*-dependent,* $\eta_n = O(n^{-a})$*,* $a > a_d^*$*,*
- C *is* p *times continuously differentiable,* $p \geq 1$*,*
- $nh_n^{2p} \to 0$*.*

Then $\|\hat{C}_n^{(2)} - C_n\|_\infty = o_P(n^{-1/2})$*. Hence* $n^{1/2}(\hat{C}_n^{(2)} - C) \to \mathbb{G}$ *in* $(D([0,1]^d), d_S)$*.*

3.2. Weak convergence of kernel copula densities

The limit of copulas is not distribution-free. This is why we also address the question of copulas densities. They are discussed in a semiparametric framework (section 3.2). In this case, limit laws of their finite distributions are asymptotically Gaussian and distribution-free, after a normalization. Assume each marginal law of the random vector $\mathbf{X}$, say the j-th, belongs to a parametric family $\{F_j(\cdot|\theta_j),\ \theta_j \in \Theta_j\}$, $j = 1,\ldots,d$. The true parameter is denoted by θ_j^0 and the true cdf by $F_j(\cdot|\theta_j^0)$ (or simpler F_j). Usually, marginal dsitributions are imposed by users, that like to put their commonly used univariate models into multivariate ones. Thus, we assume the parameters $\theta_1^0,\ldots,\theta_d^0$ are consistently estimated by $\hat{\theta}_1,\ldots,\hat{\theta}_d$. For convenience, denote $\hat{F}_j(\cdot) = F_j(\cdot|\hat{\theta}_j)$ (in this section θs only refer to parameters!). The semiparametric copula process is

$$\hat{C}(\mathbf{u}) = \frac{1}{n}\sum_{i=1}^{n}\prod_{k=1}^{d}\mathbf{1}\{F_k(X_{i,k}|\hat{\theta}_k) \leq u_k\}.$$

By smoothing this empirical copula process, we get an estimate of the copula density. The key point is that the asymptotic law of this statistics is far simpler than $\mathbb{G}$. For each index i the d-dimensional vectors we set

$$\mathbf{Y}_i = (F_1(X_{i,1}), \dots, F_d(X_{i,d})) \text{ and } \hat{\mathbf{Y}}_i = (\hat{F}_1(X_{i,1}), \dots, \hat{F}_d(X_{i,d})).$$

Assume that the law of the vectors $\mathbf{Y}_i$ has a density τ with respect to the Lebesgue measure on $\mathbb{R}^d$. The kernel estimator of a copula density τ at point $\mathbf{u}$ is thus

$$\hat{\tau}(\mathbf{u}) = \frac{1}{h^d} \int K\left(\frac{\mathbf{u}-\mathbf{v}}{h}\right) \hat{C}(d\mathbf{v}) = \frac{1}{nh^d} \sum_{i=1}^n K\left(\frac{\mathbf{u}-\hat{\mathbf{Y}}_i}{h}\right), \quad (13)$$

where K is a d-dimensional kernel and $h = h_n$ is a bandwidth sequence. As usual, we denote $K_h(\cdot) = K(\cdot/h)/h^d$. For convenience, we will assume

Assumption (K0). The kernel K is the product of d univariate even compactly supported kernels K_r, $r = 1, \dots, d$. It is assumed p_K-times continuously differentiable.

As previously, 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 argument tends to the infinity (for instance, at an exponential rate, like for the Gaussian kernel). As usual, the bandwidth sequence needs to tend to zero not too quickly.

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

Assumption (T0). Denoting by $\mathcal{V}(\theta_0)$ an open neighborhood of θ_0, for every $j = 1, \dots, d$, there exists a measurable function H_j s.t.

$$\sup_{\theta \in \mathcal{V}(\theta_0)} \|\partial^2_{\theta_j \theta_j} F_j(\mathbf{X}_j|\theta_j)\| < H_j(Y_j) \text{ a.e.}, \;\; \mathbb{E}[H_j(Y_j)] < \infty.$$

Moreover, τ and every density of $(\mathbf{Y}_0, \mathbf{Y}_k)$ are bounded in sup-norm, uniformly with respect to $k \in \mathbb{Z}$.

Assumption (E). For every $j = 1, \dots, d$,

$$\hat{\theta}_j - \theta_j^0 = n^{-1} A_j(\theta_j^0)^{-1} \sum_{i=1}^n B_j(\theta_j^0, Y_{i,j}) + o_P(r_n), \quad (14)$$

and r_n tends to zero quicker than $n^{-1/2}h^{1-d/2}$ when n tends to the infinity. Here, $A_j(\theta_j^0)$ denotes a positive definite non random matrix and $B_j(\theta_j^0, Y_j)$ is a random vector. Moreover, $\mathbb{E}[B_j(\theta_j^0, Y_j)] = 0$ and $\mathbb{E}[\|B_j(\theta_j^0, Y_j)\|^2] < \infty$. Typically, $B_j(\theta, \cdot)$ is a score function. It can be proved these assumptions are satisfied particularly for the usual maximum likelihood estimator, or more generally by M-estimators.

To invoke Doukhan and Coulhon-Prieur [14], who state the result for the usual kernel density estimates, we need the assumption:

Assumption (Y). The process $(\mathbf{Y}_i)_{i\in\mathbb{Z}}$ is stationary and η-dependent, with $\eta_n = O(n^{-a})$. The densities of the couples $(\mathbf{Y}_0, \mathbf{Y}_k)$ are uniformly bounded with respect to $k \geq 0$. Moreover the window width is assumed to satisfy $nh_n^{d\lambda} \to \infty$ as $n \to \infty$ and $a > 2 + \frac{1}{d} + \lambda$. Thus:

Theorem 5. *Under (K0) with $p_K = 2$, (B0), (T0), (E) and (Y), 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*

$$\sqrt{nh^d}\left((\hat{\tau} - K_h * \tau)(\mathbf{u}_1), \ldots, (\hat{\tau} - K_h * \tau)(\mathbf{u}_m)\right) \underset{n\to\infty}{\longrightarrow} \mathcal{N}(0, \Sigma),$$

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

Such a result can be used to prove some GOF tests, exactly as in Fermanian [19].

Remark 4. *We also derive the convergence*

$$\sqrt{nh^d}(\widehat{\tau}(\mathbf{x}) - \mathbb{E}\widehat{\tau}(\mathbf{x})) \underset{n\to\infty}{\longrightarrow} \mathcal{N}\left(0, \tau(\mathbf{x})\int K^2\right)$$

under the conditions $\theta_r = \mathcal{O}(r^{-a})$ for $a > 2 + \frac{1}{d}$ and $nh_n^{d\lambda} \to \infty$ as $n \to \infty$, as a corollary of theorem 1 in Coulon-Prieur and Doukhan [14]. The corresponding result also holds for finite dimensional distributions of this process (with independent limiting distributions).

3.3. Conditional copula processes

As previously, we consider stationary time series. Their conditional distributions with respect to past observations are often crucial to specify some underlying models. They are most of the time more useful than the joint or marginal unconditional distributions themselves. For instance, for a Markov process, the law of $\mathbf{X}_i$ conditionally on $\mathbf{X}_{i-1}$ defines the process itself. It can be written explicitly and sometimes simply, contrary to the joint law of $(\mathbf{X}_i, \ldots, \mathbf{X}_0)$. Dependence structures, copulas can be considered similarly. Patton [26] has introduced conditional copulas, namely copulas associated with conditional laws in a particular way. We first extend his definition.

Let $\mathbf{X}$ be a d-dimensional random vector. Consider some arbitrary sub σ-algebras $\mathcal{A}_1, \ldots, \mathcal{A}_d$ and $\mathcal{B}$, we denote $\mathcal{A} = (\mathcal{A}_1, \ldots, \mathcal{A}_d)$.

Assumption S. Let some d-vectors $\mathbf{x}$ and $\tilde{\mathbf{x}}$. For almost every $\omega \in \Omega$, $\mathbb{P}(X_j \leq x_j|\mathcal{A}_j)(\omega) = \mathbb{P}(X_j \leq \tilde{x}_j|\mathcal{A}_j)(\omega)$ for every $j = 1, \ldots, d$ implies $\mathbb{P}(\mathbf{X} \leq \mathbf{x}|\mathcal{B})(\omega) = \mathbb{P}(\mathbf{X} \leq \tilde{\mathbf{x}}|\mathcal{B})(\omega)$.

This technical assumption is satisfied particularly when every conditional cdfs' of $X_1, \ldots, X_d$ is strictly increasing. It is satisfied too when $\mathcal{A}_1 = \ldots = \mathcal{A}_d = \mathcal{B}$. Particularly, $\mathcal{B}$ may be the σ-algebra induced by the $\mathcal{A}_i$, $i = 1, \ldots, d$. We introduce

Definition 3. *A d-dimensional pseudo-copula is a function $C : [0,1]^d \longrightarrow [0,1]$ such that*

- *For every $\mathbf{u} \in [0,1]^d$, $C(\mathbf{u}) = 0$ when at least one coordinate of $\mathbf{u}$ is zero.*
- *$C(1, \ldots, 1) = 1$.*
- *For every $\mathbf{u}$ and $\mathbf{v}$ in $[0,1]^d$ such that $\mathbf{u} \leq \mathbf{v}$, the C-volume of $[\mathbf{u}, \mathbf{v}]$ (see Nelsen [25], definition 2.10.1) is positive.*

Thus, a pseudo-copula is "as a copula" except that the margins are not necessarily uniform. We get

Theorem 6. *For every random vector $\mathbf{X}$, there exists a random variable function $C : [0,1]^d \times \Omega \longrightarrow [0,1]$ such that*

$$\begin{aligned}\mathbb{P}(\mathbf{X} \leq \mathbf{u} | \mathcal{B})(\omega) &= C(\mathbb{P}(X_1 \leq u_1 | \mathcal{A}_1)(\omega), \ldots, \mathbb{P}(X_d \leq u_d | \mathcal{A}_d)(\omega), \omega) \\ &:= C(\mathbb{P}(X_1 \leq u_1 | \mathcal{A}_1), \ldots, \mathbb{P}(X_d \leq u_d | \mathcal{A}_d))(\omega),\end{aligned}$$

for every $\mathbf{u} \in [0,1]^d$ and almost every $\omega \in \Omega$. This function C is $\mathcal{B}([0,1]^d) \otimes \sigma(\mathcal{A}, \mathcal{B})$ measurable. For almost every $\omega \in \Omega$, $C(\cdot, \omega)$ is a pseudo-copula and is uniquely defined on the product of the values taken by $u_j \mapsto P(X_j \leq u_j | \mathcal{A}_j)(\omega)$, $j = 1, \ldots, d$.

When C is unique, it will be called the conditional $(\mathcal{A}, \mathcal{B})$-pseudo copula associated with $\mathbf{X}$. In general, it is not a copula, because of the difference between $\mathcal{B}$ and any $\mathcal{A}_i$ (in terms of information). The latter pseudo-copula is denoted by $C(\cdot | \mathcal{A}, \mathcal{B})$.

Typically, when we consider a d-dimensional process $(\mathbf{X}_n)_{n \in \mathbb{Z}}$, the previous sigma-algebras are indexed by n, namely they depend on the past values. For instance, $\mathcal{A}_{j,n} = \sigma(X_{j,n-1}, X_{j,n-2}, \ldots)$ and $\mathcal{B}_n = \sigma(\mathbf{X}_{n-1}, \ldots)$. Thus, conditional copulas depend on the index n and on the past values of $\mathbf{X}$, in general. Actually, we get sequences of copulas. When the process $\mathbf{X}$ is one-order Markov, conditional copulas depend only on the last observed value. In this paper, we consider two basic following cases:

(i) $\mathcal{A}_{j,n} = (X_{j,n-1} = x_j)$ for every $j = 1, \ldots, d$ and $\mathcal{B}_n = (\mathbf{X}_{n-1} = \mathbf{x})$,

(ii) $\mathcal{A}_{j,n} = (X_{j,n-1} \in [a_j, b_j])$, for some $a_j, b_j \in \bar{\mathbb{R}}$, $j = 1, \ldots, d$ and $\mathcal{B}_n = (\mathbf{X}_{n-1} \in [\mathbf{a}, \mathbf{b}])$.

It is particularly relevant to specify (i) and (ii) when the process $(\mathbf{X}_n)$ is Markov. Even if the process does not satisfy this property, we could consider the previous σ-algebras $\mathcal{A}_{j,n}$ and $\mathcal{B}_n$.

One key issue is to state wether these copulas depend really on the past values. This assumption is made most of the time in practice (Rosenberg [28] among others). Only a few papers try to modelize time dependent conditional copulas. For instance, to study the dependence between Yen-USD and Deutsche mark-USD exchange rates, Patton [26] assumes a bivariate Gaussian conditional copula whose correlation parameter follows a GARCH-type model. Alternatively, Genest *et al.* [21] postulate Kendall's tau is a function of current conditional univariate variances. Now, we try to estimate conditional copulas to test their constancy with respect to their conditioning subsets.

There exists a relation between copulas in the *(i)* and *(ii)* cases, denoted by $C_{(i)}$ and $C_{(ii)}$. More precisely, with obvious notations, we have

$$C_{(ii)}\left(F_{X_{1,n}}(x_1|X_{1,n-1}\in[a_1,b_1]),\ldots,F_{X_{d,n}}(x_d|X_{d,n-1}\in[a_d,b_d])\right)$$
$$= \int_{[\mathbf{a},\mathbf{b}]} \frac{C_{(i)}\left(F_1(x_1),\ldots,F_d(x_d)|\mathbf{X}_{n-1}=\mathbf{u}\right)\, d\mathbb{P}_{\mathbf{X}_{n-1}}(\mathbf{u})}{\mathbb{P}(\mathbf{X}_{n-1}\in[\mathbf{a},\mathbf{b}])},$$

by denoting $F_{X_{k,n}}(x_k|X_{k,n-1}=u_k) := F_k(x_k)$, $k=1,\ldots,d$.

Clearly, when the underlying distributions are continuous and when the diameter of the box $[\mathbf{a},\mathbf{b}]$ is "small", $F_{X_{i,n}}(x_i|X_{i,n-1}\in[a_i,b_i]) \simeq F_{X_{i,n}}(x_i|X_{i,n-1}=u_i)$ for every i and every $u_i\in[a_i,b_i]$. We deduce $C_{(i)}\simeq C_{(ii)}$ in this case. Thus, to test the constancy of $C_{(i)}(\cdot|\mathbf{X}_{n-1}=\mathbf{u})$ with respect to $\mathbf{u}$ is almost the same thing as to test the constancy of $C_{(ii)}(\cdot|\mathbf{X}_{n-1}\in[\mathbf{a},\mathbf{b}])$ with respect to "small" boxes $[\mathbf{a},\mathbf{b}]$. This intuitive argument justifies to test the zero assumption

$$\mathcal{H}_0 : C_{(ii)}(\cdot|\mathbf{X}_{n-1}\in[\mathbf{a},\mathbf{b}]) = C_0(\cdot) \text{ for every } \mathbf{a} \text{ and } \mathbf{b},$$

against its opposite. Actually, a direct test of a similar zero assumption with $C_{(i)}$ is more difficult because the marginal conditional cdfs' need to be estimated by some nonparametric techniques. At the opposite, we do not need such tools with $C_{(ii)}$, because the marginal conditioning probabilities can be easily estimated empirically.

Assume we observe a weakly dependent stationary sequence $(\mathbf{X}_i)_{0\leq i\leq n}$. Denoting by P_n the empirical measure, we see that $C_{(ii)}(\mathbf{u}|\mathbf{X}_0\in[\mathbf{a},\mathbf{b}])$ may be estimated by

$$C_{n,(ii)}(\mathbf{u}|[\mathbf{a},\mathbf{b}]) = \frac{P_n\left(X_{1,1}\leq x_1,\ldots,X_{d,1}\leq x_d,\mathbf{X}_0\in[\mathbf{a},\mathbf{b}]\right)}{P_n(\mathbf{X}_0\in[\mathbf{a},\mathbf{b}])}$$

where we set, for $j = 1, \ldots, d$ and $i \geq 1$:

$$\hat{F}_{X_{j,i}}(t|X_{j,i-1} \in [a_j, b_j]) = \frac{P_n\left(X_{j,m} \leq t, X_{j,i-1} \in [a_j, b_j]\right)}{P_n(X_{j,i-1} \in [a_j, b_j])}.$$
$$x_j = \hat{F}^{-1}_{X_{j,1}}(u_1|X_{j,0} \in [a_j, b_j]).$$

Note that the estimators $\hat{F}_{X_{j,i}}(\cdot|[a_j, b_j])$ and $C_{n,(ii)}(\cdot|[\mathbf{a}, \mathbf{b}])$ can be written as some regular functionals of the empirical cdf of $(\mathbf{X}_i, \mathbf{X}_{i-1})$.

By the same reasoning as in lemma 3 in [17], we check that the "copula" $C_{n,(ii)}(\cdot|[\mathbf{a}, \mathbf{b}])$ associated with the process $\mathbf{X}$ is the "copula" associated with the process $\mathbf{Y}$, but by replacing every a_j and b_j by $a'_j = F_j(a_j)$ and $b'_j = F_j(b_j)$, $j = 1, \ldots, d$. Thus, we could assume the underlying process has uniform marginals. By theorem 1 and the functional Delta method:

Theorem 7. *Assume $(\mathbf{Y}_i, \mathbf{Y}_{i-1})_{i\in\mathbb{Z}}$ is η-dependent, $\eta_n = O(n^{-a})$, $a > a^*_{2d}$, and that its copula has some continuous first partial derivatives. For every d-vectors $\mathbf{a}$ and $\mathbf{b}$, the process*

$$\sqrt{n}(C_{n,(ii)}(\cdot|[\mathbf{a}, \mathbf{b}]) - C_{(ii)}(\cdot|[\mathbf{a}, \mathbf{b}]))$$

converges to a Gaussian process in $D([0, 1]^d, d_S)$.

The proof is left to the reader. Thus, a test of $\mathcal{H}_0$ can be based on the limiting behavior of $\sqrt{n}(C_{n,(ii)}(\cdot|[\mathbf{a}, \mathbf{b}]) - C_0(\cdot))$. The covariance structure of the limiting process is particularly tedious. Thus, the critical values of such a test are obtained through Bootstrap procedures (see Fermanian *et al.* [17]).

4. Proofs

4.1. *Proof of theorem 1*

4.1.1. *CLT for the finite dimensional distributions of B_n*

Let $(s_1, \ldots, s_m)$ be a fixed sequence of elements in $[0, 1]^d$. Denote by $\mathbf{B}_n$ the vector-valued process

$$\mathbf{B}_n = (B_n(s_1), \ldots, B_n(s_m)).$$

To prove a CLT for the vector $\mathbf{B}_n$ is equivalent to prove the Gaussian convergence for any linear combination of its coordinates. Let $(\alpha_1, \ldots, \alpha_m)$ be a real vector such that $\sum_{j=1}^m \alpha_j \neq 0$. Define $Z_i = \sum_j \alpha_j(\mathbf{1}\{\mathbf{Y}_i \leq s_j\} - P(\mathbf{Y}_i \leq s_j))$.

Define also $\mathbf{S}_n = \frac{1}{\sqrt{n}} \sum_{1 \le i \le n} Z_i = \sum_{1 \le j \le m} \alpha_j B_n(s_j)$.

We use the Bernstein blocking technique, as described by [11]. Let $p(n)$ and $q(n)$ be sequences of integers such that $p(n) = o(n)$ and $q(n) = o(p(n))$. Assume that the Euclidean division of n by $(p+q)$ gives a quotient k. For $i = 1, \dots, k$, we define the interval $P_i = \{(p+q)(i-1)+q+1, \dots, (p+q)i\}$ and Q the set of indices that are not in one of the P_i. Note that the cardinal of Q is less than $(k+1)q$. For each block P_i and Q, we define the partial sums:

$$u_{i,n} = \frac{1}{\sqrt{n}} \sum_{j \in P_i} Z_j, \quad v_n = \frac{1}{\sqrt{n}} \sum_{j \in Q} Z_j.$$

We use lemma 11 of [11].

Lemma 4.1. *Let* $S_n = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} Z_k$ *be a sum of centered stationary r.v's, and set* $\sigma_n^2 = \mathrm{var}\ S_n$*. Assume that:*

$$\lim_{n \to \infty} \frac{1}{\sigma_n^2} \mathbb{E} v_n^2 = 0, \tag{15}$$

$$\sum_{j=2}^{k} \left| \mathrm{cov}\left(g\left(\frac{t}{\sigma_n} \sum_{i=1}^{j-1} u_{i,n} \right), h\left(\frac{t}{\sigma_n} u_{j,n} \right) \right) \right| \to 0, \ \textit{for all } t \in \mathbb{R}, \tag{16}$$

where h and g are one of the sine or the cosine function,

$$\lim_{n \to \infty} \frac{1}{\sigma_n^2} \sum_{i=1}^{k} \mathbb{E}\left[|u_{i,n}|^2 \mathbf{1}\{|u_{i,n}| \ge \epsilon \sigma_n\} \right] = 0 \ \ \textit{for all } \epsilon > 0, \tag{17}$$

$$\textit{and} \qquad \lim_{n \to \infty} \frac{1}{\sigma_n^2} \sum_{i=1}^{k} \mathbb{E}|u_{i,n}|^2 = 1. \tag{18}$$

Then S_n/σ_n converges in distribution to a Gaussian $\mathcal{N}(0,1)$-distribution.

Since the proof of this lemma is a direct adaptation of the proof of lemma 3.1 in [32], it is omitted. First note that

$$\sum_{i=0}^{n} \mathrm{cov}(Z_0, Z_i) < \infty \tag{19}$$

so that σ_n^2 tends to a constant, see [27]. If this constant is zero then the limit of S_n is 0. If it is not, we check the conditions of the preceding lemma for the sequence Z_j. To check (15), note that, with obvious notations,

$$\begin{aligned}
\mathbb{E}v_n^2 &\le \frac{1}{n}\sum_{i,j\in Q}\operatorname{cov}(Z_i,Z_j)\\
&\le \frac{2m^2\max_i\alpha_i^2}{n}\sum_{i\in Q}\sum_{j\in Q}\eta_{|j-i|}\\
&\le 4m^2\frac{(k+1)q}{n}\max_i\alpha_i^2\sum_{r=0}^{n-1}\eta_r\\
&= o(1).
\end{aligned}$$

Consider (16). Note that $g(\frac{t}{\sigma_n}\sum_{i=1}^{j-1}u_{i,n})$ is a function of at most $mp(k-1)$ indicator functions. Its Lipschitz modulus is less than $t\max_i\alpha_i/(\sqrt{n}\sigma_n)$. Similarly $h(\frac{t}{\sigma_n}u_{j,n})$ is a function of at most mp indicator functions whose Lipschitz modulus is less than $t\max_i\alpha_i/(\sqrt{n}\sigma_n)$. Invoking (9) we get

$$\begin{aligned}
\left|\operatorname{cov}\left(g\left(\frac{t}{\sigma_n}\sum_{i=1}^{j-1}u_{i,n}\right),h\left(\frac{t}{\sigma_n}u_{j,n}\right)\right)\right| &\le mpk\frac{t\max_i\alpha_i}{\sqrt{n}\sigma_n}\eta_q, \quad\text{hence}\\
\sum_{j=2}^{k}\left|\operatorname{cov}\left(g\left(\frac{t}{\sigma_n}\sum_{i=1}^{j-1}u_{i,n}\right),h\left(\frac{t}{\sigma_n}u_{j,n}\right)\right)\right| &\le mpk^2\frac{t\max_i\alpha_i}{\sqrt{n}\sigma_n}\eta_q\\
&= O\left(n^{3/2}p^{-1}q^{-a}\right).
\end{aligned}$$

Choosing $p=n^{5/6}$ and $q=n^{5/6a}$ gives a bound tending to 0.
To prove (17), it is sufficient to show that $\mathbb{E}|u_{i,n}|^4=O(k^{-2})$. But

$$\begin{aligned}
\mathbb{E}\left(\frac{1}{\sqrt{n}}\sum_{j\in P_i}Z_j\right)^4 &= \frac{p^2}{n^2}\mathbb{E}\left(\sum_{i=1}^{m}\alpha_iB_p(s_i)\right)^4\\
&\le \frac{p^2}{n^2}m^3\sum_{i=1}^{m}\alpha_i^4\mathbb{E}\left(B_p(s_i)-B_p(0)\right)^4,
\end{aligned}$$

and we conclude by applying proposition 1 for $l=2$ to the couples $(0,s_i)$:

$$\sup_i\mathbb{E}(B_p(s_i)-B_p(0))^4=O(1).$$

In order to prove (18), note that (15) implies that

$$\lim_{n\to\infty}\frac{1}{\sigma_n^2}\operatorname{var}\left(\sum_{i=1}^{k}u_{i,n}\right)=1.$$

Moreover, note that

$$\begin{aligned}\left|\operatorname{var}\left(\sum_{i=1}^{k} u_{i,n}\right) - \sum_{i=1}^{k}\mathbb{E}|u_{i,n}|^2\right| &\le 2\sum_{1\le i\ne j\le k}|\operatorname{cov}(u_{i,n},u_{j,n})| \\ &\le 2k\sum_{j=q}^{\infty}\eta_j \\ &= O(np^{-1}q^{-a+1}).\end{aligned}$$

Taking $p = n^{5/6}$ and $q = n^{5/6a}$, we get a bound tending to 0.

4.1.2. *Tightness of B_n*

As in [11], we prove a Rosenthal type inequality. This result is of independent interest.

Proposition 1. *Assume that $\mathbf{Y}$ has uniform marginals and is η-dependent with $\eta_r \le Cr^{-a}$. For every integer $l < (a+1)/2$ and s, t such that $s \le t$ and $\|t-s\|_1 < C$:*

$$\mathbb{E}(B_n(t) - B_n(s))^{2l} \le \frac{4(4l-2)!}{(2l-1)!}3^{2l} \left(\left(2k_l\left(\frac{\|t-s\|_1}{C}\right)^{1-1/a}\right)^l + (2l)!k_l n^{1-l}\left(\frac{\|t-s\|_1}{C}\right)^{1-(2l-1)/a}\right), \quad (20)$$

where $k_l = \left(C + \frac{C2^a}{a-2l+1}\right)$.

The same result may be easily proved when the marginal distributions have a bounded density (see [10]).

4.1.2.1. *Proof of proposition 1.* Let $s \le t$ be in $\mathbb{R}^d$. Denote $x_i(s,t) = \mathbf{1}\{\mathbf{Y}_i \le t\} - \mathbf{1}\{\mathbf{Y}_i \le s\} - F(t) + F(s)$. Because process $\mathbf{Y}$ has uniform margins, we get

$$|x_i(s,t)| \le 1, \quad (21)$$

and

$$|x_i(s,t)| \le \mathbf{1}\{\mathbf{Y}_i \le t\} - \mathbf{1}\{\mathbf{Y}_i \le s\} + F(t) - F(s),$$

so that

$$\begin{aligned}&\mathbb{E}|x_i(s,t)| \le \mathbb{E}(\mathbf{1}\{\mathbf{Y}_i \le t\} - \mathbf{1}\{\mathbf{Y}_i \le s\}) + F(t) - F(s) \\ &\le 2(F(t) - F(s)) \le 2\|t-s\|_1. \end{aligned} \quad (22)$$

For any multi-index $\mathbf{k}$ of $\mathbb{Z}$ denote $\Pi_{\mathbf{k}} = \prod_j x_{k_j}(s,t)$. Roughly,

$$|\mathrm{cov}\,(\Pi_{\mathbf{k}^1}, \Pi_{\mathbf{k}^2})| \leq 4\|t-s\|_1. \tag{23}$$

For any integer $q \geq 1$, set

$$A_q(n) = \sum_{\mathbf{k}\in\{1,\ldots,n\}^q} |\mathbb{E}\,(\Pi_{\mathbf{k}})|\,, \tag{24}$$

then

$$\mathbb{E}(B_n(s) - B_n(t))^{2l} \leq (2l)!n^{-l}A_{2l}(n). \tag{25}$$

For a finite sequence $\mathbf{k} = (k_1,\ldots,k_q)$ of elements of $\mathbb{Z}$, let $(k_{(1)},\ldots,k_{(q)})$ be the same sequence ordered from the smaller to the larger. The gap $r(\mathbf{k})$ in the sequence is defined as the max of the integers $k_{(i+1)} - k_{(i)}$, $j = 1,\ldots,q-1$. If $k_{(j+1)} - k_{(j)} = r$, define the two non-empty subsequences $\mathbf{k}^1 = (k_{(1)},\ldots,k_{(j)})$ and $\mathbf{k}^2 = (k_{(j+1)},\ldots,k_{(q)})$. Define the set $G_r(q,n) = \{\mathbf{k} \in \{1,\ldots,n\}^q;\ r(\mathbf{k}) = r\}$. Sorting the sequences of indices by their gaps, we get

$$A_q(n) \leq \sum_{k=1}^{n} \mathbb{E}|x_i(s,t)|^q + \sum_{r=1}^{n} \sum_{\mathbf{k}\in G_r(q,n)} |\mathrm{cov}\,(\Pi_{\mathbf{k}^1}, \Pi_{\mathbf{k}^2})| \tag{26}$$

$$+ \sum_{r=1}^{n} \sum_{\mathbf{k}\in G_r(q,n)} |\mathbb{E}\,(\Pi_{\mathbf{k}^1})\,\mathbb{E}\,(\Pi_{\mathbf{k}^2})|\,. \tag{27}$$

Define $V_q(n)$ as the right hand side of (26). In order to prove that the expression (27) is bounded by the product $\sum_m A_m(n)A_{q-m}(n)$, we make a first summation over the $\mathbf{k}$'s with $\#\mathbf{k}^1 = m$. Hence

$$A_q(n) \leq V_q(n) + \sum_{m=1}^{q-1} A_m(n)A_{q-m}(n). \tag{28}$$

To build a sequence $\mathbf{k}$ belonging to $G_r(q,n)$, we first fix one of the n points of $\{1,\ldots,n\}$. We choose a second point among the two points that are at distance r from the first point. The third point is in an interval of radius r centered on one of the preceding points, and so on. Thus

$$\#G_r(q,n) \leq n2(2r+1)\cdots(2(q-2)r+1) \leq n(q-1)!(3r)^{q-2}.$$

We use condition (9) (here $2q$ replaces $u\mathrm{Lip}\, f + v\mathrm{Lip}\, g$) and condition (23) to deduce

$$V_q(n) \leq 4n\left(\|t-s\|_1 + q!\sum_{r=1}^{2n}(3r)^{q-2}\min(\eta_r, \|t-s\|_1)\right).$$

Denote R the integer such that $R < (\|t-s\|_1/C)^{-1/a} \le R+1$. For any $2 \le q \le 2l$:

$$\begin{aligned} V_q(n) &\le 3^{(q-1)}4nq!\left(\|t-s\|_1\sum_{r=0}^{R-1} r^{q-2} + C\sum_{r=R}^{\infty} r^{q-2-a}\right) \\ &\le 3^{q-1}4nq!\left(\frac{\|t-s\|_1}{q-1}R^{q-1} + \frac{C}{(a-q+1)}R^{q-1-a}\right) \\ &\le 3^{q-1}4nq!\left(\frac{\|t-s\|_1}{C}\right)^{-(q-1)/a}\left(\frac{\|t-s\|_1}{q-1} + \frac{C}{(a-q+1)}R^{-a}\right). \end{aligned}$$

By assumption, $R \ge 1$, so that $(\|t-s\|_1/C)^{-1/a} \le 2R$, and

$$V_q(n) \le 3^{q-1}4nq!(\|t-s\|_1/C)^{1-(q-1)/a}\left(C + \frac{C2^a}{a-q+1}\right).$$

We find that:

$$V_q(n) \le 3^q 4nq! k_l(\|t-s\|_1/C)^{1-(q-1)/a}. \tag{29}$$

The rhs of equation (29) is a function of q that satisfies condition $(\mathcal{H}_0)$ of [11]:

$$\text{if } \ 2 \le p \le q,\ V_p^{q-2}(n) \le V_q^{p-2}(n)V_2^{q-p}(n).$$

Then, for $2 \le m \le q-1$,

$$(V_2^{m/2}(n) \vee V_m(n))(V_2^{(q-m)/2}(n) \vee V_{q-m}(n)) \le (V_2^{q/2}(n) \vee V_q(n)).$$

Defining $U_q = A_q(n)/(V_2^{q/2}(n) \wedge V_q(n))$, we see from (28) that

$$U_q \le \sum_{m=1}^{q-1} U_m U_{q-m} + 1.$$

Then, by invoking a lemma of [11] based on the Catalan's numbers property, we get that

$$U_q \le \frac{(2q-2)!}{q!(q-1)!}$$

and conclude that

$$\begin{aligned} A_{2l}(n) \le \frac{4(4l-2)!}{(2l)!(2l-1)!}3^{2l}\Bigg(\left(2k_l n\left(\frac{\|t-s\|_1}{C}\right)^{1-1/a}\right)^l \\ + (2l)!k_l n\left(\frac{\|t-s\|_1}{C}\right)^{1-(2l-1)/a}\Bigg), \end{aligned}$$

and (20) is proved.

Oscillation of the empirical process: We use this moment inequality and the techniques of [14] to compute the oscillations of the process. Let m be in $\mathbb{N}^d$, and (s,t) be two elements of $\mathbb{R}^d$, such that $s \leq t \leq s+m/n$. Let i be the element of $\mathbb{N}^d$ such that $s+i/n \leq t < s+i^+/n$, where $i^+ = (i_1+1,\ldots,i_d+1)$. Then

$$|B_n(t) - B_n(s)| \leq |B_n(t) - B_n(s+i/n)| + |B_n(s) - B_n(s+i/n)|.$$

Because B_n is the difference between two monotone functions, we get

$$\begin{aligned}
&|B_n(t) - B_n(s+i/n)| \\
&\leq \sqrt{n}|F_n(t) - F_n(s+i/n)| + \sqrt{n}|F(t) - F(s+i/n)| \\
&\leq \sqrt{n}|F_n(s+i^+/n) - F_n(s+i/n)| + \sqrt{n}|F(s+i^+/n) - F(s+i/n)| \\
&\leq |B_n(s+i^+/n) - B_n(s+i/n)| + 2\sqrt{n}|F(s+i^+/n) - F(s+i/n)| \\
&\leq |B_n(s+i^+/n) - B_n(s)| + |B_n(s+i/n) - B_n(s)| + 2d/\sqrt{n},
\end{aligned}$$

because the marginal distributions of F are uniform. Thus,

$$\sup_{s\leq t<s+m/n} |B_n(t) - B_n(s)| \leq 3 \max_{0\leq i\leq m} \left| B_n(s) - B_n\left(s+\frac{i}{n}\right)\right| + \frac{2d}{\sqrt{n}}. \tag{30}$$

For $s \in \mathbb{R}^d$ and $m \in \mathbb{N}^d$, define the "discrete" box $U = B(m,s) = \{s+i/n, 0 \leq i \leq m\}$. For such a box, $p_U^< = s$ and $p_U^> = s+m/n$ are opposite vertices of the box and we define

$$M(U) = \max_{t\in U}\left(\left|B_n(p_U^<) - B_n(t)\right| \wedge \left|B_n\left(p_U^>\right) - B_n(t)\right|\right).$$

Then

$$\max_{0\leq i\leq m}\left|B_n(s) - B_n\left(s+\frac{i}{n}\right)\right| \leq M(B(m,s)) + \left|B_n\left(s+\frac{m}{n}\right) - B_n(s)\right|. \tag{31}$$

Following [14], we use the moment inequality (20) to bound the distribution tail of $M(B(m,s))$:

Lemma 4.2. *Assume that p is an integer satisfying $a > 2p-1$ and $p(1-1/a) - d \geq 0$. Then*

$$\mathbb{P}\left(M(B(m,s)) \geq \lambda\right) \leq \frac{C_p}{K_p}\left(\frac{\|m\|_1}{n}\right)^{p(1-1/a)} \lambda^{-2p}, \tag{32}$$

with the constants $K_p = \frac{1}{2}\left(2^{(p(1-1/a)-d)/(2p+1)} - 1\right)^{2p+1}$ and C_p provided by proposition 1.

The first condition on p is needed to use the moment inequality. The second ensures that $K_p > 0$. The two constraints are satisfied particularly when $p = \left[\frac{1}{2}(1+d+\sqrt{1+d^2})\right]$. This induces the condition $a > a_d^*$.

Proof of the lemma. Note that (32) is true for $\|m\|_1 < 2$ and every s, because the box $B(m,s)$ contains at most two points so that $M(B(m,t)) = 0$. Let m be fixed, such that $\|m\|_1 \geq 2$ and for every $i < m$ and every t, the lemma is true for $M(B(i,t))$. Define $h = (s_1 + [m_1/2]/n, \ldots, s_d + [m_d/2]/n)$. Using h as a vertex, one defines a partition of 2^d sub-boxes of $B(m,s)$. Let $i \in B(m,s)$ and denote $U(i)$ the unique sub-box that contains i. Then

$$\begin{aligned}&\left|B_n(p^<_{B(m,s)}) - B_n(i)\right| \wedge \left|B_n\left(p^>_{B(m,s)}\right) - B_n(i)\right| \\ &\leq M(U(i)) + \left|B_n\left(p^>_{B(m,s)}\right) - B_n\left(p^>_{U(i)}\right)\right| \vee \left|B_n\left(p^<_{B(m,s)}\right) - B_n\left(p^<_{U(i)}\right)\right|.\end{aligned}$$

Because of the moment inequality and $\left\|p^>_{U(i)} - p^>_{B(m,s)}\right\|_1 \leq \|m\|_1/2n$,

$$\mathbb{P}\left(\left|B_n\left(p^>_{B(m,s)}\right) - B_n\left(p^>_{U(i)}\right)\right| \geq \lambda\right) \leq C_p \frac{\|m\|_1^{p(1-1/a)}}{(2n)^{p(1-1/a)}\lambda^{2p}},$$

and the same relation for the lower vertex yields

$$\begin{aligned}\mathbb{P}\left(\left|B_n\left(p^>_{B(m,s)}\right) - B_n\left(p^>_{U(i)}\right)\right| \vee \left|B_n\left(p^<_{B(m,s)}\right) - B_n\left(p^<_{U(i)}\right)\right| \geq \lambda\right) \\ \leq 2C_p \frac{\|m\|_1^{p(1-1/a)}}{(2n)^{p(1-1/a)}\lambda^{2p}}.\end{aligned}$$

From induction assumption and $\|p^>_{U(i)} - p^<_{U(i)}\|_1 \leq \|m\|_1/2n$:

$$\mathbb{P}\left(M(U(i)) \geq \lambda\right) \leq \frac{C_p}{K_p} \frac{\|m\|_1^{p(1-1/a)}}{(2n)^{p(1-1/a)}\lambda^{2p}}.$$

The following result may be found in [3] on p. 1661: assume that

$$\begin{aligned}&\mathbb{P}(A \geq \lambda) \leq a\lambda^{-2p} \text{ and } \mathbb{P}(B \geq \lambda) \leq b\lambda^{-2p} \text{ together imply} \\ &\qquad \mathbb{P}(A + B \geq \lambda) \leq (a^{1/(2p+1)} + b^{1/(2p+1)})^{2p+1}\lambda^{-2p}.\end{aligned}$$

We thus deduce by using the definition of K_p:

$$\begin{aligned}
&\mathbb{P}\Big(M(U(i)) + \Big|B_n\left(p^{>}_{B(m,s)}\right) - B_n\left(p^{>}_{U(i)}\right)\Big| \\
&\qquad \vee \Big|B_n\left(p^{<}_{B(m,s)}\right) - B_n\left(p^{<}_{U(i)}\right)\Big|\Big) \geq \lambda\Big) \\
&\leq C_p \frac{(2^{1/(2p+1)} + K_p^{-1/(2p+1)})^{2p+1}}{2^{p(1-1/a)}} \cdot \frac{\|m\|_1^{p(1-1/a)}}{n^{p(1-1/a)}\lambda^{2p}} \\
&\qquad \leq \frac{C_p 2^{-d}}{K_p} \frac{\|m\|_1^{p(1-1/a)}}{n^{p(1-1/a)}\lambda^{2p}}.
\end{aligned}$$

Now, using $\mathbb{P}(\max_{i=1,\dots k} A_i \geq \lambda) \leq \sum_{i=1,\dots k} \mathbb{P}(A_i \geq \lambda)$, we have

$$\mathbb{P}\left(M(B(m,s)) \geq \lambda\right) \leq \frac{C_p}{K_p} \frac{\|m\|_1^{p(1-1/a)}}{n^{p(1-1/a)}\lambda^{2p}},$$

so that (32) is proved for m. □

To prove the tightness of the sequence of processes B_n, we study the oscillations of B_n. Let $\epsilon > 0$. Let n be such that $2d/\sqrt{n} < \epsilon/8$. Let $\delta > 0$ and assume that $n\delta \geq 1$. Set $m = (2[n\delta] + 1, \dots, 2[n\delta] + 1)$. Because of relation (30), we have

$$\begin{aligned}
&\mathbb{P}\left(\sup_{\|s-t\|_1<\delta} |B_n(t) - B_n(s)| \geq \epsilon\right) \\
&\leq 2\mathbb{P}\left(\sup_{\|s-t\|_1<\delta} |B_n(t) - B_n(s\wedge t)| \geq \epsilon/2\right) \\
&\leq 2\mathbb{P}\left(\sup_{s\leq t<s+m/n} |B_n(t) - B_n(s)| \geq \epsilon/2\right) \\
&\leq 6\mathbb{P}\left(\max_{0\leq i\leq m}\left|B_n(s) - B_n\left(s + \frac{i}{n}\right)\right| \geq \epsilon/8\right).
\end{aligned}$$

Because of relation (31) and proposition 1, we obtain

$$\begin{aligned}
&\mathbb{P}\left(\max_{0\leq i\leq m}\left|B_n(s) - B_n\left(s + \frac{i}{n}\right)\right| \geq \epsilon/8\right) \\
&\leq \mathbb{P}\left(M(B(m,s))| \geq \epsilon/16\right) + \mathbb{P}\left(|B_n(s) - B_n(s + \frac{m}{n})| \geq \epsilon/16\right) \\
&\leq \frac{C_p}{K_p}\left(\frac{\|m\|_1}{n}\right)^{p(1-1/a)} \frac{16^{2p}}{\epsilon^{2p}} + C_p\left(\frac{\|m\|_1}{n}\right)^{p(1-1/a)} \frac{16^{2p}}{\epsilon^{2p}} \\
&\qquad \leq C_p\left(1 + K_p^{-1}\right)\frac{(2d(\delta + 1/n))^{p(1-1/a)}}{\epsilon^{2p}} 16^{2p},
\end{aligned}$$

so that B_n satisfies the tightness criteria for the multi-dimensional case, see [3]: for every $\epsilon > 0$,

$$\lim_{\delta\to 0}\limsup_{n\to\infty}\mathbb{P}\left(\sup_{\|s-t\|_1<\delta}|B_n(t)-B_n(s)|\geq\epsilon\right)=0,$$

proving the result. □

4.2. *Proof of theorem 2*

From lemma 3 in Fermanian *et al.* [17], it is enough to assume that the law of $\mathbf{X}$ is compactly supported on $[0,1]^d$ with uniform marginals. The argument relies on the functional delta method through the function ϕ defined on the Skorohod space $(D([0,1],d_S)$, $\phi : F_1 \mapsto F_1^{-1}$; from the compact support assumption, this application is now defined on $(l^\infty([0,1]),\|\cdot\|_\infty)$. As in [17], we apply theorem 3.9.23 in Van der Vaart and Wellner [31] to conclude. Note that, for any function $h\in C([0,1])$, the convergence of a sequence h_n to h in $(D([0,1]),d_S)$ is equivalent to the convergence in $(D([0,1]),\|\cdot\|_\infty)$. The result follows by applying theorem 3.9.4 in Van der Vaart and Wellner [31] and our theorem 1. □

4.3. *Proof of lemma 3.1*

First, let us assume that k is compactly supported. Then, by some integrations by parts, we get

$$\begin{aligned}\sqrt{n}(\hat{F}_n-F_n)(\mathbf{u}) &= \sqrt{n}\int[F_n(\mathbf{u}-h\mathbf{v})-F_n(\mathbf{u})]\,k(\mathbf{v})\,d\mathbf{v}\\ &= \int\sqrt{n}\,[(F_n-F)(\mathbf{u}-h\mathbf{v})-(F_n-F)(\mathbf{u})]\,k(\mathbf{v})\,d\mathbf{v}\\ &\quad + n^{1/2}\int(F(\mathbf{u}-h\mathbf{v})-F(\mathbf{u}))\,k(\mathbf{v})\,d\mathbf{v}.\end{aligned}$$

Since $\mathbf{v}$ belongs to a compact subset, $h\mathbf{v}$ is bounded above uniformly with respect to $\mathbf{v}$ and n. Equicontinuity of the process $\sqrt{n}(F_n-F)$ thus provides the result with our assumptions.

If k is not compactly supported, we lead the same reasoning. Now, for n sufficiently large,

$$\begin{aligned}&\mathbb{P}\left(|\sqrt{n}\int[(F_n-F)(\mathbf{u}-h\mathbf{v})-(F_n-F)(\mathbf{u})]\,k(\mathbf{v})\,d\mathbf{v}|>\eta\right)\\ &\leq\ \mathbb{P}\left(n^{1/2}\|k\|_{L_1}.\sup_{\|\mathbf{t}\|<\eta}|(F_n-F)(\mathbf{u}-\mathbf{t})-(F_n-F)(\mathbf{u})|>\eta/2\right)\\ &+\ \mathbb{P}\left(2n^{1/2}\int_{\{\|\mathbf{v}\|>a_n\}}|k(\mathbf{v})|\,d\mathbf{v}>\eta/2\right),\end{aligned}$$

which tends to zero under our assumptions. □

4.4. *Proof of theorem 5*

A Taylor expansion yields for every $\mathbf{u} \in [0,1]^d$,

$$\begin{aligned}\hat{\tau}(\mathbf{u}) &= \frac{1}{nh^d}\sum_{i=1}^n K\left(\frac{\mathbf{u}-\mathbf{Y}_i}{h}\right)\\ &- \frac{1}{nh}\sum_{i=1}^n dK\left(\frac{\mathbf{u}-\mathbf{Y}_i}{h}\right)(\hat{\mathbf{Y}}_i-\mathbf{Y}_i)\\ &+ \frac{1}{2nh^2}\sum_{i=1}^n d^2K\left(\frac{\mathbf{u}-\mathbf{Y}_i^*}{h}\right)(\hat{\mathbf{Y}}_i-\mathbf{Y}_i)^{\otimes 2}\\ &= \tau^*(\mathbf{u}) + R_1(\mathbf{u}) + R_2(\mathbf{u}),\end{aligned}$$

for some random vectors $\hat{\mathbf{Y}}_i^*$ satisfying $\|\hat{\mathbf{Y}}_i^* - \mathbf{Y}_i\| \leq \|\hat{\mathbf{Y}}_i - \mathbf{Y}_i\|$ a.e.

Note that τ^* is the kernel density estimator studied in Doukhan and Louhichi [12], when applied to the weakly dependent sequence $(\mathbf{Y}_i)_{i\in\mathbb{Z}}$, which is improved in the paper by Doukhan and Prieur-Coulon [14]. Thus we get fidi convergence of $\sqrt{nh^d}(\tau^*-\tau)$.

It remains to prove that $R_1(\mathbf{u})$ and $R_2(\mathbf{u})$ are negligible. Let us first study $R_1(\mathbf{u})$. Denote partial derivatives wrt u_j by ∂_j,

$$\begin{aligned}R_1(\mathbf{u}) &= -\frac{1}{n^2h}\sum_{i,k}\sum_{j=1}^d \partial_j K\left(\frac{\mathbf{u}-\mathbf{Y}_i}{h}\right)\partial_{\theta_j'}F_j(X_{i,j}|\theta_j^0)A_j^{-1}(\theta_j^0)B_j(\theta_j^0,Y_{k,j})\\ &+ O_P\left(\frac{1}{n^2h}\sum_{i,k}\sum_{j=1}^d\left|\partial_j K\left(\frac{\mathbf{u}-\mathbf{Y}_i}{h}\right)\right|\sup_{\theta_j^*}\|\partial^2_{\theta_j\theta_j'}F_j(X_{i,j}|\theta_j^*)\|\|\hat{\theta}_j-\theta_j^0\|^2\right)\\ &+ o_P\left(\frac{r_n}{h}\right),\end{aligned}$$

where θ_j^* belongs a.e. to a neighborhood of θ_j^0 for every j. Since the process $(\mathbf{Y}_i)_{i\in\mathbb{Z}}$ is weakly dependent, and since $B_j(\theta_j^0, Y_{k,j})$ is centered, we get

$$\mathbb{E}\left[\partial_j K\left(\frac{\mathbf{u}-\mathbf{Y}_i}{h}\right)\partial_{\theta_j'}F_j(X_{i,j}|\theta_j^0)A_j^{-1}(\theta_j^0)B_j(\theta_j^0,Y_{k,j})\right] = O\left(\frac{\eta_{|i-k|}}{h^{1+d}}\right),$$

and $\mathbb{E}R_1(\mathbf{u}) = O\left(\frac{1}{nh^{2+d}}\right) = o\left(\frac{1}{\sqrt{nh^d}}\right)\cdot$ Moreover, with obvious notations, the main term in the expansion of $\mathbb{E}R_1^2(\mathbf{u})$ is the expectation of

$$\frac{1}{n^4h^2}\sum_{i_1,i_2=1}^{n}\sum_{k_1,k_2=1}^{n}\sum_{j_1,j_2=1}^{d}\partial_{j_1}K\left(\frac{\mathbf{u}-\mathbf{Y}_i}{h}\right)\partial_{\theta'_{j_1}}F_{j_1}(X_{i_1,j_1}|\theta^0_{j_1})A^{-1}_{j_1}(\theta^0_{j_1})$$
$$B_{j_1}(\theta^0_{j_1},Y_{k_1,j_1})(\partial_{j_2}K)_h(\mathbf{u}-\mathbf{Y}_{i_2})\partial_{\theta'_{j_2}}F_{j_2}(X_{i_2,j_2}|\theta^0_{j_2})$$
$$A^{-1}_{j_2}(\theta^0_{j_2})B_{j_2}(\theta^0_{j_2},Y_{k_2,j_2})$$
$$:= \frac{1}{n^4h^2}\sum_{j_1,j_2}\sum_{i_1,i_2}\sum_{k_1,k_2}T_{i_1,j_1}\tilde{T}_{k_1,j_1}T_{i_2,j_2}\tilde{T}_{k_2,j_2}.$$

We consider every relative positions of the indices i_1, i_2, k_1, k_2 (j_1 and j_2 do not play any role). In each cases, weak dependence allows us to bound the expectation of $T_{i_1,j_1}\tilde{T}_{k_1,j_1}T_{i_2,j_2}\tilde{T}_{k_2,j_2}$. As in Theorem 5 of Fermanian [19], $\mathbb{E}R_1^2(\mathbf{u}) = o((nh^d)^{-1})$ and $R_1(\mathbf{u}) = o_P((nh^d)^{-1/2})$. The second term $R_2(\mathbf{u})$ is simpler because

$$R_2(\mathbf{u}) = O_P\left(\frac{1}{h^{2+d}}\cdot\frac{1}{n}\right) = o_P\left(\frac{1}{\sqrt{nh^d}}\right),$$

which proves the result.

References

1. Ango Nze, P. and Doukhan, P.: Weak dependence: models and applications, in *Dehling W. (ed.) et al., Empirical Processes Techniques for Dependent Data*, Birkhaüser, 117-137, 2002.
2. Ango Nze, P., Bühlmann, P. and Doukhan, P.: Weak dependence beyond mixing and asymptotics for non parametric regression, *Ann. Statist.* **30**, 397-430, 2002.
3. Bickel, P.J. and Wichura, M.J.: Convergence criteria for multiparameter stochastic and some applications, *Ann. Math. Statist.* **42**, 1656-1671, 1971.
4. Chen, X. and Fan, Y.: Estimation of Copula-based semiparametric time series models, J. of Econometrics 130, 307-335, 2006.
5. Dedecker, J., Doukhan, P.: A new covariance inequality and applications, *Stoch. Proc. Appl.* Vol. 106, n1, 63-80, 2003.
6. Deheuvels, P.: La fonction de dépendance empirique et ses propriétés, *Acad. Roy. Belg., Bull. C1 Sci. 5ième sér.* **65**, 274-292, 1979.
7. Deheuvels, P.: A Kolmogorov-Smirnov Type Test for independence and multivariate samples, *Rev. Roum. Math. Pures et Appl.* **26** N2, 213-226, 1981.
8. Doukhan, P.: *Mixing: Properties and Examples*, Lecture Notes in Statist. 85. Springer Verlag, New York, 1994.
9. Doukhan, P.: Limit theorems for stationary sequences, in *Theory and Applications of Long-range Dependence, Paul Doukhan, Georges Oppenheim and Murad S. Taqqu editors.* Birkhaüser, Boston, 43 – 100, 2002.
10. Doukhan, P. and Lang, G.: Rates in the empirical central limit theorem for stationary weakly dependent random fields, *Stat. Inf. Stoch. Proc.* **5**, 199-228, 2002.

11. Doukhan, P. and Louhichi, S.: A new weak dependence condition and application to moment inequalities, *Stoch. Proc. Appl.* **84**, 313-342, 1999.
12. Doukhan, P. and Louhichi, S.: Functional estimation for weakly dependent stationary time series, *Scand. J. Statist.* **28**, 325-342, 2001.
13. Doukhan, P., Massart, P. and Rio, E.: Invariance principle for absolutely regular empirical processes, *Ann. Inst. H. Poincaré* **31.2**, 393-427, 1995.
14. Doukhan, P. and Prieur-Coulon, C.: A CLT for triangular arrays of weakly dependent sequences, *Statist. Probab. Letters* **47**, 61-68, 2000.
15. Doukhan, P., Teyssière, G., Winant, P. A LARCH(∞) Vector Valued Process. To appear in statistics for times series ed. Soulier *et al.*, 2006.
16. Duflo, M. Algorithmes stochastiques *Math. & Appl. 23,* Springer, 1996.
17. Fermanian, J-D., Radulovic, D. and Wegkamp, M.: Weak convergence of empirical copula processes, *Bernoulli* **10**, n5 , 847860, 2004.
18. Fermanian, J-D. and Scaillet, O.: Nonparametric estimation of copulas for time series, *J. Risk* **5**, $25-54$, 2002.
19. Fermanian, J-D.: Goodness-of-fit tests for copulas, *J. Multivariate Anal.*, 95, 119-152, 2005.
20. Gaenssler, P. and Stute, W.: *Seminar on Empirical Processes.*, DMV Seminar, Band 9, Birkhäuser, 1987.
21. Genest, C., van der Goorbergh, R. and Werker, B.: Multivariate option pricing using dynamic copulas, discussion paper 122, Tilburg University, 2003.
22. Giraitis, L., Surgailis, D. *ARCH-type bilinear models with double long memory.* Stochastic Processes and their Applications, 100, 275-300, 2002.
23. Gouriéroux C., Laurent J. P. and O. Scaillet: Sensitivity Analysis of Values at Risk, *J. Empirical Finance* **7**, 225-245, 2000.
24. Markowitz H.: Portfolio Selection, *Journal of Finance* **7**, 77-91, 1952.
25. Nelsen, R. B.: *An introduction to copulas*, Lecture Notes in Statistics, **139**, Springer, 1999.
26. Patton, A.: Modelling Time-Varying Exchange Rate Dependence Using the Conditional Copula, UCSD WP 2001-09, 2001.
27. Rio, E.: *Théorie asymptotique des processus aléatoires faiblement dépendants*, Mathématiques et Applications, 31, Springer Verlag, Berlin, 2000.
28. Rosenberg, J.: Nonparametric pricing of multivariate contingent claims, WP Stern School of Business, 2001.
29. Scaillet O.: Nonparametric Estimation and Sensitivity Analysis of Expected Shortfall. *Mathematical Finance*, Vol. 14, 115-129, 2004.
30. Sklar, A.: Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris* **8**, 229-231, 1959.
31. Vaart, A. W. van der and Wellner, J.: *Weak convergence and empirical processes*, Springer, New-York, 1996.
32. Withers, C.S.: Central limit theorems for dependent variables, *Z. Wahrsch. Verw. Gebiete* **57**, 509-534, 1981 (corrigendum **63**, 555, 1983).