COPULAS OF A VECTOR-VALUED STATIONARY WEAKLY DEPENDENT PROCESS

By Paul Doukhan[a], Jean-David Fermanian[b] and Gabriel Lang[c]

July 22, 2004

[a] *ENSAE and CREST, 3 Avenue Pierre Larousse, 92245 Malakoff cedex, France. doukhan@ensae.fr*

[b] *CREST, 15 bd Gabriel Péri, 92245 Malakoff cedex, France. fermania@ensae.fr*

[c] *GRESE, ENGREF, 19 av du Maine 75732 Paris cedex 15, France. lang@engref.fr*

Address for correspondence: Jean-David Fermanian, CREST, J320, 15 bd Gabriel Péri, 92245 Malakoff cedex, 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 fidi of smoothed copula densities is proved. A theoretical analysis of conditional copulas is provided, with applications to Goodness-of-Fit tests.

Keywords: Copulas; multivariate FCLT; weak dependence.

1 Introduction

This paper is aimed to consider asymptotic results relative to copulas for weakly dependent sequences. Various notions of weak dependence have been introduced in the literature. Among them, α-mixing and β-mixing have been studied, but these notions are not fully satisfactory, as very simple processes like AR(1) processes with Bernoulli innovations may fail to satisfy any mixing condition (see Doukhan, 1994). Doukhan and Louhichi (1999) introduce a definition of weak dependence that is easier to check on various examples of stationary processes (see Doukhan, 2002). Various other applications and developments of weak dependence are addressed in Ango Nze, Bühlmann and Doukhan (2002) and Ango Nze and Doukhan (2002).

To fix the ideas, we recall the recent notion of weak dependence as defined in Doukhan and Louhichi (1999). The definition of weak dependence may be seen as a way to weaken independence. Consider two finite samples P (for past) and F (for future) of a sequence 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 keep tractable the checks of the condition in practice.

More formally, let $\theta = (\theta_r)_{r\geq 0}$ be a real positive sequence that tends to zero. 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 $||\mathbf{x}|| = ||(x_1, \ldots, x_d)||_1 = \sum_{i=1}^d |x_i|$. Define $\mathcal{L}$ as the set of functions that are bounded by 1 and have a finite Lipschitz modulus. We say that sequences of indices $i_1 \leq \cdots \leq i_u$ and $j_1 \leq \cdots \leq j_v$ such that $i_u \leq j_1$ are r-distant if $j_1 - i_u = r$. Let $\theta = (\theta_r)_{r\geq 0}$ be a real positive sequence that tends to zero. Let f and g be two functions of $\mathcal{L}$ defined on $\mathbb{R}^u$ and $\mathbb{R}^v$ respectively.

Let (ψ_1) be the functions defined on $\mathcal{L}^2 \times \mathbb{N}^2$ by

$$\psi_1(f, g, u, v) \quad = \quad (u+v)(\text{Lip }(f) \vee \text{Lip }(g)).$$

Definition 1. *We say that the d-dimensional process $(\xi_i)_{i\in\mathbb{Z}}$ is $(\theta, \mathcal{L}, \psi_1)$-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 $\mathcal{L}$ defined on $(\mathbb{R}^d)^u$ and $(\mathbb{R}^d)^v$ respectively, we have*

$$|\text{cov}(f(\xi_{i_1}, \ldots, \xi_{i_u}), g(\xi_{j_1}, \ldots, \xi_{j_v}))| \leq \psi_1(f, g, u, v)\theta_r. \tag{1.1}$$

Note that if ξ is $(\theta, \mathcal{L}, \psi_1)$-dependent and if f and g are only bounded Lipschitz functions, the previous covariance is bounded by $\|f\|_\infty \|g\|_\infty \psi_1(f, g, u, v)\theta_r$.

Considering some weakly dependent vector-valued sequences $(\mathbf{X}_i)_{i\in\mathbb{Z}}$, the main theoretical result of the paper is to prove a functional central limit theorem. It is an extension on the independent case, where the limit process is known to be a Brownian bridge B, i.e. a Gaussian process with covariance function

$$\text{cov}(B(\mathbf{x}), B(\mathbf{y})) = \mathbb{P}\left(\mathbf{X}_0 \leq \mathbf{x} \wedge \mathbf{y}\right) - \mathbb{P}(\mathbf{X}_0 \leq \mathbf{x})\mathbb{P}(\mathbf{X}_0 \leq \mathbf{y}), \tag{1.2}$$

for every vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^d$. Such a FCLT is proved in appendix A. Thus, the functional delta method applies and several applications are provided. We respectively consider empirical and smoothed copulas in this weakly dependent framework (section). Contrarily to the usual empirical process, the limiting distributions are not free of the distribution's process in this case. This is why we also perform the special case of copulas densities. They are discussed in a semi-parametric framework (section). In this case, limit laws of their finite repartitions are asymptotically gaussian and distribution-free, after a normalization. A discussion of conditional copulas and their properties is postponed in section 4. They can be applied to test the constancy of the dependence structure with respect to past observations.

2 Empirical copula processes

Copulas describe the dependence structure between some random vectors. They have been introduced a long time ago (Sklar, 1959) 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 margins are uniform and that summarizes the dependence "structure" independently of the specification of the marginal distributions.

To be specific, 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, it is defined on $[0,1]^d$, with an obvious extension to $\overline{\mathbb{R}}^d$. When F is discontinuous, there are several choices to expand C on the whole $[0,1]^d$ (see Nelsen (1999) or Joe (1997) for a complete theory).

Imagine we are faced with a sequence of random vectors $\mathbf{X}_1, \ldots, \mathbf{X}_n$, that are the realizations of some underlying process $(\mathbf{X}_i)_{i\in\mathbb{Z}}$. This process is assumed stationary. Thus, the joint law of $\mathbf{X}_i$ is independent of i and we denote by C its copula. We seek to estimate nonparametrically the copula C. Note that our results can be applied to the sequence $(\mathbf{X}_i, \mathbf{X}_{i+1})_{i\in\mathbb{Z}}$ (as in Chen and Fan, 2002), or even $(\mathbf{X}_i, \mathbf{X}_{i+1}, \mathbf{X}_{i+2})_{i\in\mathbb{Z}}$ etc. In the latter case, knowing the stationary marginal distributions, C describes fully the process $(\mathbf{X}_m)_{m\in\mathbb{Z}}$ when it is Markov.

The natural empirical counterpart of C is the so-called empirical copula, defined by

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

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

$$F_n(\mathbf{x}) = n^{-1} \sum_{i=1}^{n} \mathbf{1}\{X_{i1} \leq x_1, \ldots, X_{id} \leq x_d\}, \tag{2.1}$$

and we use the usual "generalized inverse" notations, viz, for every $j = 1, \ldots, d$, $F_j^{-}(u) = \inf\{t | F_j(t) \geq u\}$.

Empirical copulas have been introduced by Deheuvels (1979,1981a,1981b) in an i.i.d. framework. This author studied the consistency of C_n and the limiting behavior of $n^{1/2}(C_n - C)$ under the strong assumption of independence between margins. Gaensler and Stute (1987) and Fermanian *et al.* (2002) prove some functional CLT for this empirical copula process in a more general framework and provide some extensions. We will first expand these results to dependent data by applying theorem 7 in the appendix to the process $(\mathbf{Y}_i)_{i\in\mathbb{Z}}$, $\mathbf{Y} = (F_1(X_1), \ldots, F_d(X_d))$. Note that the results of Fermanian *et al.* (2002) are available under the sup-norm and outer expectations assumptions, as in Van der Vaart and Wellner (1996). Here, the natural space will be the space of càdlàg functions $D([0,1]^d)$ endowed with the Skorohod metric d_S.

Consider a centered Gaussian process $\mathbb{B}$ such that, for any vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^d$,

$$\mathrm{cov}(\mathbb{B}(\mathbf{x}), \mathbb{B}(\mathbf{y})) = \sum_{i\in\mathbb{Z}} \mathrm{cov}\left(\mathbf{1}\{\mathbf{X}_0 \leq \mathbf{x}\}, \mathbf{1}\{\mathbf{X}_i \leq \mathbf{y}\}\right). \tag{2.2}$$

Note that the previous covariance structure depends not only on the copula C (via the term associated with $i = 0$ e.g.), but also on the joint law between $\mathbf{X}_0$ and $\mathbf{X}_i$, for every i. This is different from the i.i.d. case, where $\mathbb{B}$ becomes a Brownian bridge whose covariance structure is a function of C only (equation (1.2)). Actually, the covariances of $\mathbb{B}$ depend here on every successive copulas of the random vectors $(\mathbf{X}_0, \mathbf{X}_i)$. First, we can state :

Theorem 1. *If $(\mathbf{Y}_i)_{i\in\mathbb{Z}}$ is $(\theta, \mathcal{L}, \psi_1)$-dependent, $\theta_n = O(n^{-a})$, $a > 2d+1$, if C has some continuous first partial derivatives, then the process $n^{1/2}(C_n - C)$ tends weakly towards a Gaussian process $\mathbb{G}$ in $(D([0,1]^d), d_S)$. Moreover, this process has continuous sample paths and can be written as*

$$\mathbb{G}(\mathbf{u}) = \mathbb{B}(\mathbf{u}) - \sum_{j=1}^{d} \partial_j C(\mathbf{u}) \mathbb{B}(u_j, 0_{-j}), \tag{2.3}$$

for every $\mathbf{u} \in [0,1]^d$.

See the proof in appendix B. The proof is based on a FCLT for multivariate weakly dependent sequences (theorem 7, whose proof is postponed in appendix B). Note that the covariance structure of $n^{1/2}(C_n - C)$ is involved, because of both (2.3) and (2.2).

Remarks. The same result applies in the case of sequences for whom a FCLT holds for the multivariate empirical cdf. Even if we recalled previously the problems occurring with the standard mixing conditions (see Doukhan, 1994), it seems useful that theorem 1 still holds in such mixing situations which are more widely considered and yields various other models as stressed in Doukhan (1994). Precisely:

- For stationary strongly mixing sequences, if $\alpha_n = \mathcal{O}\left(n^{-a}\right)$ for some $a > 1$. To this goal, we use Rio (2000) which states an empirical CLT for vector-valued sequences under this condition.
- In the absolutely regular case, this assumption is written $\beta_n = \mathcal{O}\left(n^{-1}\log^{-b} n\right)$ for some $b > 2$ by using Doukhan, Massart and Rio (1995)'s result.
- Many other results yielding such FCLTs for the process $B_n = \sqrt{n}(F_n - F)$ are recalled in Doukhan (2002)'s review paper.

In practice, it is often necessary to estimate smoothed copulas. From a visual point of view, the results are nicer than those obtained for the empirical copulas themselves. Since nonparametric estimation is often the first step before a parametric modelization, it is important to help intuition conveniently. Moreover, for optimization purposes, some estimates of the derivatives of underlying copulas are most of the time necessary : portfolio optimization in a mean-variance framework (Markowitz, 1952)) 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.* (2000) or Scaillet (2000)). Clearly, a smooth empirical cdf is differentiable, contrary to a usual empirical cdf. Thus, let us introduce smoothed empirical copula processes.

Consider $\hat{F}_n$ the d-dimensional smoothed empirical process

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

associated with the usual empirical process F_n (see equation 2.1), where K is the primitive function of a d-dimensional kernel k subject to the limit condition $\lim_{-\infty} K = 0$, and $h = h_n$ is a bandwidth. More precisely, $\int k = 1$, $h_n > 0$, and $h_n \to 0$ when $n \to \infty$. Similarly, for every margin, say the j-th, we can estimate nonparametrically the cdf F_j by

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

$$F_{n,j}(x_j) = n^{-1} \sum_{i=1}^{n} \mathbf{1}\{X_{i,j} \leq x_j\},$$

and K_j is the primitive function of a univariate kernel k_j. For simplicity, we have assumed the bandwidth h is the same for every margin and that $k(\mathbf{u}) = \prod_{j=1}^d k_j(u_j)$ for every $\mathbf{u}$. These two latter assumptions can be easily removed. Then, for every $\mathbf{u} \in [0,1]^d$, we can define the smoothed empirical copula process by

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

or by smoothing directly the process C_n, viz

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

As for 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 of order p, 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 2.1. *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 B. Assumption (i) is satisfied when $\mathbf{X}$ is compactly supported, invoking theorem 7. We get assumption (ii) by assuming some regularity on F, e.g. F is p-times continuously differentiable. Therefore, by mimicking exactly the proof of theorem 10 in Fermanian *et al.* (2002), we get

Theorem 2. *Assume (K) and*

- F *is* p*-times continuously differentiable,*
- $h \to 0$, $nh^{2p} \to 0$,
- $(\mathbf{X}_i)_{i\in\mathbb{Z}}$ *is* $(\theta, \mathcal{L}, \psi_1)$*-dependent,* $\theta_n = O(n^{-a})$, $a > 2d + 1$.

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

This result extends for weakly dependent processes the fidi result in Fermanian and Scaillet (2002). Moreover, we can prove lemma 2.1 replacing F_n by C_n exactly by the same ways. This provides :

Theorem 3. *Assume (K) and*

- $(\mathbf{Y}_i)_{i\in\mathbb{Z}}$ *is* $(\theta, \mathcal{L}, \psi_1)$*-dependent,* $\theta_n = O(n^{-a})$, $a > 2d + 1$,
- C *is* p *times continuously differentiable,* $p \geq 1$,
- $nh_n^{2p} \to 0$.

Then $\|\hat{C}_n^{(2)} - C_n\|_\infty = o_P(n^{-1/2})$. *Particularly, the process* $n^{1/2}(\hat{C}_n^{(2)} - C)$ *tends weakly towards the Gaussian process* $\mathbb{G}$ *in* $(D([0,1]^d), d_S)$.

Thus, the weak convergence of $\hat{C}_n^{(2)}$ is satisfied under slightly weaker assumptions than $\hat{C}_n^{(1)}$. Since formulas are nicer too, we advise to work with the former estimator rather than the latter.

3 Weak convergence of kernel copula densities

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 density by $F_j(\cdot|\theta_j^0)$ (or simpler F_j). Usually, marginal models 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 $\hat F_j(\cdot) = F_j(\cdot|\hat\theta_j)$.

The natural semiparametric copula process we consider is

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

By smoothing the 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 $\mathbb{G}$.

To be specific, set for each index i the d-dimensional vectors

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

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

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

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 (K_0)**.** 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

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). Since this speed depends on the behavior of τ, we are rather the simpler assumption (K_0).

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 (B0) can be weakened easily by assuming (K_0) with $p_K > 2$. Nonetheless, a certain additional amount of regularity is required.

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

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

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 $j = 1, \ldots, d$,

$$\hat{\theta}_j - \theta_j^0 = n^{-1} A_j(\theta_j^0)^{-1} \sum_{i=1}^n B_j(\theta_j^0, Y_{i,j}) + o_P(r_n), \tag{3.2}$$

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 Ragache (2004), 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 $(\theta, \mathcal{L}, \psi_1)$-dependent, with $\theta_n = O(n^{-a})$. 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, we prove:

Theorem 4. *Under (K_0) 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*

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

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

Remarks.

- Such a result can be used to prove some GOF tests, exactly as in Fermanian (2003).
- Replacing ψ_1-weak dependence by its causal counterpart $\psi_1'-$weak dependence where now $\psi_1'(f, g, u, v) = v\mathrm{Lip}\, g$, we derive the convergence

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

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 (2000) or Doukhan and Ragache (2004), by a different method. The corresponding result also holds for finite dimensional repartitions of this process (with independent limiting distributions).

4 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, viz copulas can be considered similarly. Patton (2001) has introduced conditional copulas, viz 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 $\mathcal{B}$. For convenience, 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 $\omega \in \Omega$, $P(X_j \leq x_j | \mathcal{A}_j)(\omega) = P(X_j \leq \tilde{x}_j | \mathcal{A}_j)(\omega)$ for every $j = 1, \ldots, d$ implies $P(\mathbf{X} \leq \mathbf{x} | \mathcal{B})(\omega) = P(\mathbf{X} \leq \tilde{\mathbf{x}} | \mathcal{B})(\omega)$.

This technical assumption is satisfied particularly when every conditional cdfs' of $X_1, \ldots, X_d$ are strictly increasing. It is satisfied too when $\mathcal{A}_1 = \ldots = \mathcal{A}_d = \mathcal{B}$. Particularly, $\mathcal{B}$ could be the σ-algebra induced by the $\mathcal{A}_i$, $i = 1, \ldots, d$, but it is not an obligation.

We introduce without proof the concept of pseudo-copula and the equivalent of Sklar's theorem for such functions (see Fermanian and Wegkamp (2004) for details):

Definition 2. *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 (1999), 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 5. *For every random vector* $\mathbf{X}$*, there exists a random variable function* $C : [0,1]^d \times \Omega \longrightarrow [0,1]$ *such that*

$$\begin{aligned} P(\mathbf{X} \leq \mathbf{u}|\mathcal{B})(\omega) &= C(P(X_1 \leq u_1|\mathcal{A}_1)(\omega), \ldots, P(X_d \leq u_d|\mathcal{A}_d)(\omega), \omega) \\ &\equiv C(P(X_1 \leq u_1|\mathcal{A}_1), \ldots, P(X_d \leq u_d|\mathcal{A}_d))(\omega), \end{aligned} \tag{4.1}$$

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* $u_i \mapsto P(X_i \leq u_i|\mathcal{A}_i)(\omega)$*,* $i = 1, \ldots, d$*.*

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 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, viz they depend on the past values. For instance, $\mathcal{A}_{i,n} = \sigma(X_{i,n-1}, X_{i,n-2}, \ldots)$ 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. Note that Chen and Fan (2002) propose to study univariate stationary Markov processes (X_n) by specifying the copula of the vector (X_n, X_{n+1}). In this very particular case, $\mathcal{A}_1 = \mathcal{A}_2 = \mathcal{B} = \{\emptyset, \Omega\}$ like in an i.i.d. case.

In this paper, we will consider two basic cases for the conditioning subsets:

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

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

Note that it would be possible theoretically to set $\mathcal{A}_{i,n} = (X_{i,n-1} = x_i)$ for some components $i = 1, \ldots, d$ only, and $\mathcal{B}_n = (\mathbf{X}_{n-1} \in [\mathbf{a}, \mathbf{b}])$ for some $\mathbf{a}, \mathbf{b} \in \bar{\mathbb{R}}^d$; or even $\mathcal{A}_{i,n} = (X_{i,n-1} \in [a_i, b_i])$, for some or every $i = 1, \ldots, d$, and $\mathcal{B}_n = (\mathbf{X}_{n-1} \in [\mathbf{x}, \mathbf{y}])$, where $[\mathbf{x}, \mathbf{y}] \subset [\mathbf{a}, \mathbf{b}]$. Nonetheless, the practical interest of these specifications is questionable. Moreover, note that our purpose does not depend on any Markov property. Even if the process does not satisfy this property, we could consider the previous σ-algebras $\mathcal{A}_{i,n}$ and $\mathcal{B}_n$. Nonetheless, it is particularly relevant to specify (i) and (ii) when the process $(\mathbf{X}_n)$ is Markov.

One key issue is often to state wether these copulas depend really on the past values, viz to test their constancy. This assumption is made most of the time in practice (Rosenberg (2001), Cherubini and Luciano (2000), 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 (2001) assumes a bivariate Gaussian conditional copula whose correlation parameter follows a GARCH-type model. Alternatively, Genest *et al.*

(2003) postulate Kendall's tau is a function of current conditional univariate variances. Now, we will 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

$$\begin{aligned}
& C_{(ii)}\left(F_{X_{1,n}}(x_1|X_{1,n-1} \in [a_1,b_1]),\dots,F_{X_{d,n}}(x_d|X_{d,n-1} \in [a_d,b_d])\right) \\
&= P\left(X_{1,n} \leq x_1,\dots,X_{d,n} \leq x_d|\mathbf{X}_{n-1} \in [\mathbf{a},\mathbf{b}]\right) \\
&= \int_{[\mathbf{a},\mathbf{b}]} \frac{P\left(X_{1,n} \leq x_1,\dots,X_{d,n} \leq x_d|\mathbf{X}_{n-1} = \mathbf{u}\right) dP_{\mathbf{X}_{n-1}}(\mathbf{u})}{P(\mathbf{X}_{n-1} \in [\mathbf{a},\mathbf{b}])} \\
&= \int_{[\mathbf{a},\mathbf{b}]} \frac{C_{(i)}\left(F_{X_{1,n}}(x_1|X_{1,n-1} = u_1),\dots,F_{X_{d,n}}(x_d|X_{d,n-1} = u_d)|\mathbf{X}_{n-1} = \mathbf{u}\right) dP_{\mathbf{X}_{n-1}}(\mathbf{u})}{P(\mathbf{X}_{n-1} \in [\mathbf{a},\mathbf{b}])}.
\end{aligned}$$

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 $\mathcal{H}_0 : C_{(ii)}(\cdot|\mathbf{X}_{n-1} \in [\mathbf{a},\mathbf{b}]) = C_0(\cdot)$ for every $\mathbf{a}$ and $\mathbf{b}$, against its opposite. Actually, a direct test of a similar zero assumption with $C_{(i)}$ is far more difficult because the marginal conditional cdfs' need to be estimated by some nonparametric techniques like kernel smoothing, nearest neighbors, or others. At the opposite, we do not need such tools with $C_{(ii)}$, because the marginal conditioning probabilities can be easily estimated empirically.

To be specific, assume we observe a weakly dependent stationary sequence $(\mathbf{X}_i)_{i=0,\dots,n}$. Then, denoting by P_n the empirical measure, we have

$$\begin{aligned}
& C_{(ii)}(\mathbf{u}|\mathbf{X}_0 \in [\mathbf{a},\mathbf{b}]) \simeq \\
& \quad P\left(X_{1,1} \leq F^{-}_{X_{1,1}}(u_1|X_{1,0} \in [a_1,b_1]),\dots,X_{d,1} \leq F^{-}_{X_{d,1}}(u_d|X_{d,0} \in [a_d,b_d])|\mathbf{X}_0 \in [\mathbf{a},\mathbf{b}]\right) \\
&\simeq \frac{P_n\left(X_{1,1} \leq \hat{F}^{-}_{X_{1,1}}(u_1|X_{1,0} \in [a_1,b_1]),\dots,X_{d,1} \leq \hat{F}^{-}_{X_{d,1}}(u_d|X_{d,0} \in [a_d,b_d]),\mathbf{X}_0 \in [\mathbf{a},\mathbf{b}]\right)}{P_n(\mathbf{X}_0 \in [\mathbf{a},\mathbf{b}])} \\
&\equiv C_{n,(ii)}(\mathbf{u}|[\mathbf{a},\mathbf{b}]),
\end{aligned}$$

where, for every $i = 1,\dots,d$, we set

$$\hat{F}_{X_{i,m}}(u_i|X_{i,m-1} \in [a_i,b_i]) = \frac{P_n\left(X_{i,m} \leq u_i, X_{i,m-1} \in [a_i,b_i]\right)}{P_n(X_{i,m-1} \in [a_i,b_i])}.$$

Note that the estimators $\hat{F}_{X_{i,m}}(u_i|[a_i,b_i])$ and $C_{n,(ii)}(\mathbf{u}|[\mathbf{a},\mathbf{b}])$ can be written as some regular functionals of the empirical cdf of $(\mathbf{X}_m,\mathbf{X}_{m-1})$. Therefore, by applying the Functional Delta method (as in theorem 1) and theorem 7, we get

Theorem 6. *For every d-vectors* $\mathbf{a}$ *and* $\mathbf{b}$*, the process* $n^{1/2}(C_{n,(ii)}(\cdot|[\mathbf{a},\mathbf{b}]) - C_{(ii)}(\cdot|[\mathbf{a},\mathbf{b}]))$ *tends weakly towards a Gaussian process in* $D([0,1]^d, d_S)$.

The proof is straightforward and it is left to the reader. Thus, a test of $\mathcal{H}_0$ can be based on the limiting behavior of $n^{1/2}(C_{n,(ii)}(\cdot|[\mathbf{a},\mathbf{b}]) - C_0(\cdot))$. Nonetheless, the limiting process and its covariance structure are particularly tedious. Thus, to state the critical values of such a test, we advocate to use some Bootstrap procedures. The convergence of the bootstrapped empirical process (see Fermanian *et al.*, 2002) justifies the method.

A Central limit theorem for the multivariate empirical process

We give a definition of dependence based on the covariance of some indicator functions. The link with the weak dependence is given in Lemma A.1.

Definition 3. *Let* f *be a function on* $\mathbb{R}^u$*,* $\mathbf{i} = (i_1, \ldots, i_u)$ *be a sequence of elements in* $\mathbb{Z}$ *and* $\mathbf{s} = (s_1, \ldots, s_u)$ *be a sequence of elements in* $\mathbb{R}^d$*. With implicit reference to a process* $\mathbf{X}$*, we define*

$$Z(f, \mathbf{i}, \mathbf{s}) = f(\mathbf{1}\{\mathbf{X}_{i_1} \leq s_1\}, \ldots, \mathbf{1}\{\mathbf{X}_{i_u} \leq s_u\}).$$

Define the normalized process $B_n = \sqrt{n}(F_n - F)$. Recalling the definition of the process $\mathbb{B}$ (cf equation 2.2), the main theoretical result of the paper is the following:

Theorem 7. *Assume that* $(\mathbf{X}_i)_{i\in\mathbb{Z}}$ *is a centered process 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* $\mathcal{L}$*, defined on* $\mathbb{R}^u$ *and* $\mathbb{R}^v$*:*

$$|\mathrm{cov}(Z(f, \mathbf{i}, \mathbf{s}), Z(g, \mathbf{j}, \mathbf{t}))| \leq \psi_1(f, g, u, v)\theta_r. \tag{A.1}$$

Assume that there exist some constants $C > 0$ *and* $a > 2d + 1$ *such that* $\theta_r \leq Cr^{-a}$*. Then* B_n *tends to* $\mathbb{B}$ *in distribution in* $D([0,1]^d, d_S)$.

See the proof in section B. The following lemma is useful to apply theorem 7.

Lemma A.1. *If* $(\mathbf{X}_i)_{i\in\mathbb{Z}}$ *is* $(\theta^1, \mathcal{L}, \psi_1)$*-dependent, then condition (A.1) is satisfied with* $\theta_r = 3(\theta_r^1)^{\frac{1}{2}}$.

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

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

Then $h_{\varepsilon,t}(x)$ is $1/\varepsilon$-Lipschitz, and $\mathbb{E}\|h_{\varepsilon,t}(\mathbf{X}_0) - \mathbf{1}\{\mathbf{X}_0 \geq t\}\|_1 \leq \varepsilon$.

JDF : il me semble qu'il y a un petit problème. Pour moi,

$$\mathbb{E}\|h_{\varepsilon,t}(\mathbf{X}_0) - \mathbf{1}\{\mathbf{X}_0 \geq t\}\|_1 \sum_k \leq P\left(X_{k0} \in [t_k^{(p)} - \varepsilon, t_k^{(p)}]\right) = O(\varepsilon),$$

si la loi de $\mathbf{X}_0$ est continue. Du coup, il me semble que la majoration n'est pas explicite et qu'on devrait écrire des $O(\cdot)$. Par exemple, que $\theta_r \sim (\theta_r^1)^{1/2}$.

Define the analogous approximation of $Z(f, \mathbf{i}, \mathbf{s})$ by

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

If $\mathbf{X}$ is $(\theta^1, \mathcal{L}, \psi_1)$-weak dependent, then, for any r-distant sequences $\mathbf{i}$ and $\mathbf{j}$,

$$|\mathrm{cov}(Z_\varepsilon(f, \mathbf{i}, \mathbf{s}), Z_\varepsilon(g, \mathbf{j}, \mathbf{t}))| \leq \|f\|_\infty \|g\|_\infty \varepsilon^{-1} \psi_1(f, g, u, v) \theta_r^1.$$

Moreover,

$$\begin{aligned}
&|\mathbb{E}\left(Z_\varepsilon(f, \mathbf{i}, \mathbf{s}) Z_\varepsilon(g, \mathbf{j}, \mathbf{t})\right) - \mathbb{E}\left(Z(f, \mathbf{i}, \mathbf{s}) Z(g, \mathbf{j}, \mathbf{t})\right)| \\
&\qquad \leq \mathbb{E}\left|Z_\varepsilon(f, \mathbf{i}, \mathbf{s})\left(Z_\varepsilon(g, \mathbf{j}, \mathbf{t}) - Z(g, \mathbf{j}, \mathbf{t})\right)\right| + \mathbb{E}\left|Z(g, \mathbf{j}, \mathbf{t})\left(Z_\varepsilon(f, \mathbf{i}, \mathbf{s}) - Z(f, \mathbf{i}, \mathbf{s})\right)\right| \\
&\qquad\qquad \leq (u\|f\|_\infty \mathrm{Lip}\ (g) + v\|g\|_\infty \mathrm{Lip}\ (f))\varepsilon \leq \|f\|_\infty \|g\|_\infty \psi_1(f, g, u, v)\varepsilon.
\end{aligned}$$

Similarly,

$$|\mathbb{E}\left(Z_\varepsilon(f, \mathbf{i}, \mathbf{s})\right) \mathbb{E}\left(Z_\varepsilon(g, \mathbf{j}, \mathbf{t})\right) - \mathbb{E}\left(Z(f, \mathbf{i}, \mathbf{s})\right) \mathbb{E}\left(Z(g, \mathbf{j}, \mathbf{t})\right)| \leq \|f\|_\infty \|g\|_\infty \psi_1(f, g, u, v)\varepsilon. \qquad (A.2)$$

Choosing ε such that $\theta_r^1 \varepsilon^{-1} = \varepsilon$, we get

$$|\mathrm{cov}(Z(f, \mathbf{i}, \mathbf{s}), Z(g, \mathbf{j}, \mathbf{t}))| \leq \psi_1(f, g, u, v) 3(\theta_r^1)^{1/2}. \square$$

B Proofs

B.1 *Proof of theorem 1*

The proof is directly adapted from Fermanian *et al.* (2002). Briefly, we can assume the law of $\mathbf{X}$ is compactly supported on $[0, 1]^d$, eventually by working with $\mathbf{Y} = (F_1(X_1), \ldots, F_d(X_d))$.

Indeed, it can be proved the empirical copulas associated with $\mathbf{Y}$ and $\mathbf{X}$ are equal on all the points $(i_1/n, \ldots, i_d/n)$, $i_1, \ldots, i_d$ in $\{0, \ldots, n\}$ (lemma 3 in Fermanian *et al.* (2002)), thus on $[0,1]^d$ as a whole.

Moreover, consider successively the mappings

$$
\begin{array}{llcl}
\phi_1 : & (D([0,1]^d), d_S) & \to & (D([0,1]^d), d_S) \times (D([0,1]), d_S)^{\otimes d} \\
& F & \mapsto & (F, F_1, \ldots, F_d) \\
\phi_2 : & (D([0,1]^d), d_S) \times (D([0,1]), d_S)^{\otimes d} & \to & (l^\infty([0,1]^d), \| \cdot \|_\infty) \times (l^\infty([0,1]), \| \cdot \|_\infty)^{\otimes d} \\
& (F, F_1, \ldots, F_d) & \mapsto & (F, F_1^-, \ldots, F_d^-) \\
\phi_3 : & (l^\infty([0,1]^d), \| \cdot \|_\infty) \times (l^\infty([0,1]), \| \cdot \|_\infty)^{\otimes d} & \to & (l^\infty([0,1]^d), \| \cdot \|_\infty) \\
& (F, G_1, \ldots, G_d) & \mapsto & F(G_1, \ldots, G_d).
\end{array}
$$

Clearly, ϕ_1 is Hadamard-differentiable because it is linear. Moreover, ϕ_2 is Hadamard-differentiable tangentially to the corresponding product of continuous functions by applying theorem 3.9.23 in Van der Vaart and Wellner (1996). Note that, for any function $h \in C([0,1])$, the convergence of a sequence h_n towards h in $(D([0,1]), d_S)$ is equivalent to the convergence in $(D([0,1]), \| \cdot \|_\infty)$. Thus, working with the Skorohod metric in not an hurdle here. At last, ϕ_3 is Hadamard-differentiable by applying theorem 3.9.27 in Van der Vaart and Wellner (1996). Thus, the chain rule applies : $\phi = \phi_3 \circ \phi_2 \circ \phi_1$ is Hadamard-differentiable tangentially to $C([0,1]^d)$. The result follows by applying theorem 3.9.4 in Van der Vaart and Wellner (1996) and our theorem 7. $\Box$

B.2 *Proof of lemma 2.1*

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

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

Since $\mathbf{v}$ belongs to a compact subset, $h\mathbf{v}$ is bounded above uniformly with respect to $\mathbf{v}$ and n. Thus, the equicontinuity of the process $n^{1/2}(F_n - F)$ and our assumptions provide the result.

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

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

which tends to zero under our assumptions. $\Box$

B.3 *Proof of theorem 4*

By a limited expansion, we get, for every $\mathbf{u} \in [0,1]^d$,

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

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

Note that τ^* is the kernel density estimator studied in Doukhan, Louhichi (2001), when applied to the weakly dependent sequence $(\mathbf{Y}_i)_{i\in\mathbb{Z}}$, which is improved in the recent paper by Doukhan, Ragache (2004).

Thus, under our assumptions, we get fidi convergence of $(nh^d)^{1/2}(\tau^*-\tau)$.

To obtain our result, it is sufficient to prove $R_1(\mathbf{u})$ and $R_2(\mathbf{u})$ are negligible. Let us first study $R_1(\mathbf{u})$. With some obvious notations,

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

where θ_j^* belongs a.e. in 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

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

and

$$
\mathbb{E}[R_1(\mathbf{u})] = O\left(\frac{1}{nh^{2+d}}\right) << \frac{1}{\sqrt{nh^d}}.
$$

Moreover, with obvious notations, the main term in the expansion of $\mathbb{E}[R_1(\mathbf{u})^2]$ is

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

To deal with the latter term, we have to 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 every cases, the definition of the weak dependence allows us to bound the expectation of $T_{i_1,j_1}\tilde{T}_{k_1,j_1}T_{i_2,j_2}\tilde{T}_{k_2,j_2}$. Dealing as in Theorem 5 of Fermanian (2003), it can be done relatively easily, under our assumptions. Therefore, $\mathbb{E}[R_1^2(\mathbf{u})] = o((nh^d)^{-1})$ and $R_1(\mathbf{u}) = o_P((nh^d)^{-1/2})$.

The second term $R_2(\mathbf{u})$ is simpler because an upper bound is straightforward

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

under our assumptions. So the result. □

B.4 *Proof of theorem 7*

B.4.1 CLT for the finite dimensional distributions of B_n

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

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

Denote also

$$\mathbf{x}_i = (x_{i,1}, \ldots, x_{i,p}) = (\mathbf{1}\{\mathbf{X}_i \leq s_1\} - P(\mathbf{X}_i \leq s_1), \ldots, \mathbf{1}\{\mathbf{X}_i \leq s_p\} - P(\mathbf{X}_i \leq s_p)), \text{ and}$$

$$\mathbf{S}_{k,m} = \frac{1}{\sqrt{n}} \sum_{k \leq i \leq m} \mathbf{x}_i.$$

Then $\mathbf{B}_n = \mathbf{S}_{1,n}$. Let γ_i be the covariance matrix of $(\mathbf{x}_0, \mathbf{x}_i)$ and Γ_m be the covariance matrix of $\mathbf{S}_{1,m}$. Note that Γ_n tends to the matrix $\sum_{i=0}^{\infty} \gamma_i$ so that this series is a positive definite matrix. Define $V_k = \Gamma_k - \Gamma_{k-1}$ and $\Gamma_0 = 0$. Then

$$V_k = \frac{2}{n} \sum_{j=1}^{k-1} Cov(\mathbf{x}_i, \mathbf{x}_k) + \frac{1}{n} Var(\mathbf{x}_k) = \frac{2}{n} \sum_{j=1}^{k-1} \gamma_i + \frac{\gamma_0}{n}. \tag{B.1}$$

For $k \geq k_0$, V_k is a positive definite matrix. Define a Gaussian independent vector process $\mathbf{y}_i = (y_{i,1}, \ldots, y_{i,p})$, $i \in \mathbb{Z}$, such that the covariance matrix of $\mathbf{y}_i$ is V_i. For $k < k_0$, define $\mathbf{T}_{k,m} = \sum_{k \leq i \leq m} \mathbf{y}_i$. Let f be a C^3 function on $\mathbb{R}^p$, and denote $C_i = \|D^i f\|_\infty$, $i = 1, 2, 3$. To obtain the fidi convergence, it is sufficient to show that $\mathbb{E}(f(\mathbf{S}_{1,n}) - f(\mathbf{T}_{k_0,n}))$ tends to zero when

n tends to the infinity. Write

$$\begin{aligned}
\mathbb{E}(f(\mathbf{S}_{1,n}) - f(\mathbf{T}_{k_0,n})) &= \mathbb{E}(f(\mathbf{S}_{1,n}) - f(\mathbf{S}_{k_0,n})) \\
&+ \mathbb{E}\left[\sum_{k=k_0}^{n} f(\mathbf{S}_{k_0,k-1} + n^{-1/2}\mathbf{x}_k + \mathbf{T}_{k+1,n}) - f(\mathbf{S}_{k_0,k-1} + \mathbf{y}_k + \mathbf{T}_{k+1,n})\right] \\
&= \mathbb{E}(f(\mathbf{S}_{1,n}) - f(\mathbf{S}_{k_0,n})) \\
&+ \mathbb{E}\left[\sum_{k=k_0}^{n} f_{n,k}(\mathbf{S}_{k_0,k-1} + n^{-1/2}\mathbf{x}_k) - f_{n,k}(\mathbf{S}_{k_0,k-1} + \mathbf{y}_k)\right],
\end{aligned}$$

where we have set

$$f_{n,k}(z) = \mathbb{E}(f(z + \mathbf{T}_{k+1,n})),$$

for every real number z. Since the random vectors $\mathbf{x}$ and $\mathbf{y}_k$ are independent from $\mathbf{T}_{k+1,n}$, the last formula is true. First, we have

$$|\mathbb{E}(f(\mathbf{S}_{1,n}) - f(\mathbf{S}_{k_0,n}))| \leq \left(\mathbb{E}(f(\mathbf{S}_{1,n}) - f(\mathbf{S}_{k_0,n}))^2\right)^{1/2} \leq C_1(k_0/n)^{1/2} \to 0.$$

The second part is equal to

$$\sum_{k=k_0}^{n} f_{n,k}(\mathbf{S}_{k_0,k-1} + n^{-1/2}\mathbf{x}_k) - f_{n,k}(\mathbf{S}_{k_0,k-1} + \mathbf{y}_k) = \sum_{k=k_0}^{n} (\Delta_{1,k,n} - \Delta_{2,k,n}),$$

where

$$\Delta_{1,k,n} = \mathbb{E}\left[f_{n,k}\left(\mathbf{S}_{k_0,k-1} + n^{-1/2}\mathbf{x}_k\right) - f_{n,k}(\mathbf{S}_{k_0,k-1}) - \frac{1}{2}\sum_{s,t=1}^{p} D^2 f_{n,k}(\mathbf{S}_{k_0,k-1})_{(s,t)} V_{k(s,t)}\right],$$

$$\Delta_{2,k,n} = \mathbb{E}\left[f_{n,k}\left(\mathbf{S}_{k_0,k-1} + \mathbf{y}_k\right)) - f_{n,k}(\mathbf{S}_{k_0,k-1}) - \frac{1}{2}\sum_{s,t=1}^{p} D^2 f_{n,k}(\mathbf{S}_{k_0,k-1})_{(s,t)} V_{k(s,t)}\right].$$

Bound of $\Delta_{1,k,n}$: By a Taylor expansion and with some obvious notations,

$$\begin{aligned}
\Delta_{1,k,n} &= \mathbb{E}\left[n^{-1/2}\sum_{s=1}^{p} Df_{n,k}(\mathbf{S}_{k_0,k-1})_s x_{k,s}\right] \\
&+ \frac{1}{2}n^{-1}\sum_{s,t=1}^{p}\left\{\mathbb{E}\left[D^2 f_{n,k}(\mathbf{S}_{k_0,k-1})_{(s,t)} x_{k,s} x_{k,t}\right] - \mathbb{E}\left[D^2 f_{n,k}(\mathbf{S}_{k_0,k-1})_{(s,t)}\right] nV_{k(s,t)}\right\} \\
&+ \frac{1}{6}n^{-3/2}\sum_{s,t,r=1}^{p} \mathbb{E}\left[D^3 f_{n,k}(R_k)_{(s,t,r)} x_{k,s} x_{k,t} x_{k,r}\right] \equiv T_1 + T_2 + T_3.
\end{aligned}$$

First, note that

$$
\begin{aligned}
T_1 &= n^{-1/2} \sum_{s=1,\dots p} \sum_{j=1}^{k-1} \mathbb{E}\left[\left(Df_{n,k}(\mathbf{S}_{k_0,j})_s - Df_{n,k}(\mathbf{S}_{k_0,j-1})_s\right) x_{k,s}\right] \\
&= n^{-1/2} \sum_{s=1,\dots p} \sum_{j=1}^{k-1} \operatorname{cov}\left(\left(Df_{n,k}(\mathbf{S}_{k_0,j})_s - Df_{n,k}(\mathbf{S}_{k_0,j-1})_s\right), x_{k,s}\right).
\end{aligned}
$$

Therefore, T_1 is bounded above by $n^{-1/2}\sum_{j=1}^{k-1} C_2\theta_{k-j} = O(n^{-1/2}k^{3-a})$.

JDF : je ne vois pas comment obtenir le resultat car

$$
n^{-1/2}\sum_{j=1}^{k-1}(k-j)^{-a} > n^{-1/2},
$$

et ensuite

$$
\sum_{k=k_0}^{n} n^{-1/2} = +\infty!
$$

Idem pour le second terme. Il faudrait peut-être supposer $\theta_r \leq a^r$ et non $\theta_r \leq r^{-a}$...et prendre k_0 assez grand.

Second, invoking (B.1), we get

$$
\begin{aligned}
T_2 &= \frac{1}{2}n^{-1} \sum_{s,t=1}^{p} \operatorname{cov}\left[D^2 f_{n,k}(\mathbf{S}_{k_0,k-1})_{(s,t)}, x_{k,s}x_{k,t}\right] \\
&- n^{-1} \sum_{s,t=1}^{p} \mathbb{E}\left[D^2 f_{n,k}(\mathbf{S}_{k_0,k-1})_{(s,t)}\right] \sum_{j=1}^{k-1} \mathbb{E}(x_{j,s}x_{k,t})
\end{aligned}
$$

that is bounded above by $n^{-1}\sum_{j=1}^{k-1}(C_2+C_3)(j+1)\theta_{k-j} = O(n^{-1}k^{3-a})$ from (A.1).

JDF : même remarque qu'au-dessus

Third, the remainder term T_3 is less than $C_3 n^{-3/2}$ because the x_k are bounded by 1. This gives

$$
\sum_{k=k_0}^{n} |\Delta_{1,k,n}| = O(n^{-1/2}).
$$

Bound of $\Delta_{2,k,n}$: By a Taylor expansion, we get similarly

$$\begin{aligned}\Delta_{2,k,n} =& \mathbb{E}\left[n^{-1/2}\sum_{s=1,\dots p} Df_{n,k}(\mathbf{S}_{k_0,k-1})_s y_{k,s}\right]\\ &+\frac{1}{2}n^{-1}\sum_{s,t=1}^{p}\mathbb{E}\left[D^2 f_{n,k}(\mathbf{S}_{k_0,k-1})_{(s,t)}y_{k,s}y_{k,t}\right]-\mathbb{E}\left[D^2 f_{n,k}(\mathbf{S}_{k_0,k-1})_{(s,t)}\right]nV_{k(s,t)}\\ &+\frac{1}{6}n^{-3/2}\sum_{s,t,r=1}^{p}\mathbb{E}\left[D^3 f_{n,k}(R_k)_{(s,t,r)}y_{k,s}y_{k,t}y_{k,r}\right].\end{aligned}$$

Using the independence between y and x, the first and second line equal zero. The last line is less than $C_3 n^{-3/2}$ because the $\mathbf{y}_k$ has finite third cross-moments. This gives

$$\sum_{k=k_0}^{n}|\Delta_{2,k,n}| = O(n^{-1/2}).$$

This proves that the finite distributions of B_n converge to the announced Gaussian distributions.

B.4.2 Tightness of B_n

As in [11], we prove a Rosenthal type inequality.

Proposition 1. *Assume that $\theta_r = Cr^{-a}$. For $l < (a+1)/2$ and (s,t) such that $|t-s| < C$:*

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

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

Proof of proposition 1: Let $s \le t$ be in $\mathbb{R}^d$, denote $x_i(s,t) = \mathbf{1}\{\mathbf{X}_i \le t\} - \mathbf{1}\{\mathbf{X}_i \le s\}$. Because process $\mathbf{X}$ has uniform margins, we get

$$\mathbb{E}(x_i(s,t)) \le \|t-s\|_1.\tag{B.3}$$

JDF : Nous n'avons pas fait l'hypothèse que les marges de $\mathbf{X}$ sont uniformes, il me semble. Peut-être faut-il le supposer dès le début de la démonstration. Mais le résultat de la proposition 1 ne sera vrai formellement que sous cette hypothèse.

For any multi-index $\mathbf{k}$ of $\mathbb{Z}$ denote $\Pi_{\mathbf{k}} = \prod_j x_{k_j}(s,t)$. For any integer $q \geq 1$, set

$$A_q(n) = \sum_{\mathbf{k} \in \{1,\dots,n\}^q} |\mathbb{E}(\Pi_{\mathbf{k}})|, \tag{B.4}$$

then

$$\mathbb{E}(B_n(s) - B_n(t))^{2l} \leq (2l)! n^{-l} A_{2l}(n). \tag{B.5}$$

JDF : Je ne suis pas d'accord. Pour moi,

$$B_n(t) - B_n(s) = n^{-1/2} \sum_{i=1}^{n} (x_i(s,t) - E[x_i(s,t)]).$$

Donc on doit se trainer des espérance de $X_i(s,t)$. Il me semble que cela change la démonstration.

For a finite sequence $\mathbf{k} = (k_1, \dots, k_q)$ of elements of $\mathbb{Z}$, let $(k_{(1)}, \dots, k_{(q)})$ 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 $k_{(i+1)} - k_{(i)}$, $j = 1, \dots, q-1$. If $k_{(j+1)} - k_{(j)} = r$, define the two non-empty subsequences $\mathbf{k}^1 = (k_{(1)}, \dots, k_{(j)})$ and $\mathbf{k}^2 = (k_{(j+1)}, \dots, k_{(q)})$. Define the set $G_r(q,n) = \{\mathbf{k} \in \{1, \dots, n\}^q \; ; \; r(\mathbf{k}) = r\}$. Sorting the sequences of indices by their gaps, we get

$$\begin{aligned} A_q(n) &\leq \sum_{k=1}^{n} \mathbb{E}|x_i(s,t)|^q + \sum_{r=1}^{n} \sum_{\mathbf{k} \in G_r(q,n)} |\mathrm{cov}(\Pi_{\mathbf{k}^1}, \Pi_{\mathbf{k}^2})| &\text{(B.6)} \\ &+ \sum_{r=1}^{n} \sum_{\mathbf{k} \in G_r(q,n)} |\mathbb{E}(\Pi_{\mathbf{k}^1}) \mathbb{E}(\Pi_{\mathbf{k}^2})|. &\text{(B.7)} \end{aligned}$$

JDF : je remarque juste que le dernier terme de droite de l'inégalité précédente est nul si on travaille sur les $x_i(s,t)$ centrés. Cela peut éventuellement simplifier.

Define $V_q(n)$ as the right hand side of (B.6). In order to prove that the expression (B.7) 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

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

To build a sequence $\mathbf{k}$ belonging to $G_r(q,n)$, we first fix one of the n points of $\{1,\dots,n\}$. 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

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

We use condition (A.1) and condition (B.3) to deduce:

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

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

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

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

$$V_q(n) \le 3^{q-1}nq!(\|t-s\|_1/C)^{1-(q-1)/a}\left(C + \frac{C2^a}{a-2l+1}\right).$$

We find that:

$$V_q(n) \le 3^q nq! k_l (\|t-s\|_1/C)^{1-(q-1)/a}, \tag{B.8}$$

and $V_q(n)$ is a function of q that satisfies condition $(\mathcal{H}_0)$ of [11]:

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

Then

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

and (B.2) 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 $s \le t \le s + m/n$. Let i be the element of $\mathbb{N}^d$ such that $s + i/n \le t < s + i^+/n$, where $i^+ = (i_1 + 1, \dots, i_d + 1)$. Then

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

By invoking the fact that B_n is the difference between two monotone functions, we get

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

JDF : il me semble qu'il faut que F soit au moins lipschitzienne pour avoir

$$|F(s+i^+/n) - F(s+i/n)| = O(1/n).$$

Thus,

$$\sup_{s \le t < s+m/n} |B_n(t) - B_n(s)| \le 3 \max_{0 \le i \le m} \left| B_n(s) - B_n\left(s + \frac{i}{n}\right)\right| + \frac{2d}{\sqrt{n}} \cdot \tag{B.9}$$

For $s \in \mathbb{R}^d$ and $m \in \mathbb{N}^d$, define the "discrete" box $U = B(m,s) = \{s + i/n, 0 \le i \le m\}$. 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

$$\max_{0 \le i \le m} \left| B_n(s) - B_n\left(s + \frac{i}{n}\right)\right| \le M(B(m,s)) + \left|B_n\left(s + \frac{m}{n}\right) - B_n(s)\right|. \tag{B.10}$$

As in [14], we use the moment inequality (B.2) to bound the distribution tail of $M(B(m,s))$. We can now prove that

$$\mathbb{P}\left(M(B(m,s)) \ge \lambda\right) \le \frac{C_p}{K_p}\left(\frac{\|m\|_1}{n}\right)^{p(1-1/a)} \lambda^{-2p}, \tag{B.11}$$

for some constant C_p provided by proposition 1, and where $K_p = \left(2^{(p(1-1/a)-d)/(2p+1)} - 1\right)^{2p+1}/2$. Note that (B.11) 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

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

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

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

and the same relation for the lower vertex yields

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

Because of the recurrence, and $\|p^>_{U(i)} - p^<_{U(i)}\|_1 \leq \|m\|_1/2n$:

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

Now we use the following result (see [3], p. 1661): If $\mathbb{P}(A \geq \lambda) \leq a\lambda^{-2p}$ and $\mathbb{P}(B \geq \lambda) \leq b\lambda^{-2p}$ then $\mathbb{P}(A + B \geq \lambda) \leq (a^{1/(2p+1)} + b^{1/(2p+1)})^{2p+1}\lambda^{-2p}$. We get

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

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

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

so that (B.11) is proved for m.

To prove the tightness of the sequence of processes B_n, we study the oscillations of B_n. Let $\varepsilon > 0$. Let n be such that $2d/\sqrt{n} < \varepsilon/4$. Let $\delta > 0$ and assume that $n\delta \geq 1$. Let $m = (2[n\delta]+1, \ldots, 2[n\delta]+1)$. Because of relation (B.9):

$$\begin{aligned}\mathbb{P}\left(\sup_{\|s-t\|_1<\delta} |B_n(t) - B_n(s)| \geq \varepsilon\right) &\leq \mathbb{P}\left(\sup_{s\leq t<s+m/n} |B_n(t) - B_n(s)| \geq \varepsilon\right) \\ &\leq 3\mathbb{P}\left(\max_{0\leq i\leq m} \left|B_n(s) - B_n\left(s + \frac{i}{n}\right)\right| \geq \varepsilon/4\right).\end{aligned}$$

Because of relation (B.10) and proposition 1, we obtain

$$\begin{aligned}&\mathbb{P}\left(\max_{0\leq i\leq m} \left|B_n(s) - B_n\left(s + \frac{i}{n}\right)\right| \geq \varepsilon/4\right) \leq \mathbb{P}\left(M(B(m,s))| \geq \varepsilon/8\right) \\ &+ \ \mathbb{P}\left(|B_n(s) - B_n(s + \frac{m}{n})| \geq \varepsilon/8\right) \\ &\leq \ \frac{C_p}{K_p}\left(\frac{\|m\|_1}{n}\right)^{p(1-1/a)} \frac{8^{2p}}{\varepsilon^{2p}} + C_p\left(\frac{\|m\|_1}{n}\right)^{p(1-1/a)} \frac{8^{2p}}{\varepsilon^{2p}} \\ &\leq \ C_p\left(1 + K_p^{-1}\right) \frac{(2d\delta + 1/n)^{p(1-1/a)}}{\varepsilon^{2p}} 8^{2p},\end{aligned}$$

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

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

proving the result. □

References

[1] Ango Nze, P. and Doukhan, P. (2002) Weak dependence: models and applications. *Dehling W. (ed.) et al., Empirical Processes Techniques for Dependent Data*, 117-137, Birkhaüser.

[2] Ango Nze, P., Bühlmann, P. and Doukhan, P. (2002) Weak dependence beyond mixing and asymptotics for non parametric regression. *Ann. Statist.* 30, 397-430.

[3] Bickel, P.J. and Wichura, M.J. (1971) Convergence criteria for multiparameter stochastic and some applications. *Ann. Math. Statist.* 42, 1656 − 1671.

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

[5] Cherubini, U. and Luciano, E. (2001) Multivariate option pricing with copulas. ICER Working Paper.

[6] Deheuvels, P. (1979) La fonction de dépendance empirique et ses propriétés. *Acad. Roy. Belg., Bull. C1 Sci. 5ième sér.* 65, 274 – 292.

[7] Deheuvels, P. (1981a) A Kolmogorov-Smirnov Type Test for independence and multivariate samples. *Rev. Roum. Math. Pures et Appl.* 26, No 2, 213 – 226.

[8] Deheuvels, P. (1981b) A nonparametric test for independence. *Pub. Inst. Stat. Univ. Paris.* 26, No 2, 29 – 50.

[9] Doukhan, P. (1994) *Mixing: Properties and Examples.* Lecture Notes in Statist. 85. Springer Verlag, New York.

[10] Doukhan, P. (2002) 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.

[11] Doukhan, P. and Louhichi, S. (1999) A new weak dependence condition and application to moment inequalities. *Stoch. Proc. Appl.* 84, 313 – 342.

[12] Doukhan, P. and Louhichi, S. (2001) Functional estimation for weakly dependent stationary time series. *Scand. J. Statist.* 28, 325-342.

[13] Doukhan, P., Massart, P. and Rio, E. (1995) Invariance principle for the empirical measure of a weakly dependent process. *Ann. I.H.P.* 31, 393 – 427.

[14] Doukhan, P. and Prieur-Coulon, C. (2000) A CLT for triangular arrays of weakly dependent sequences. *Statist. Probab. Letters* 47, 61-68.

[15] Doukhan, P. and Ragache N. (2004) Théorème de limite centrale pour des tableaux triangulaires et faiblement dépendants. To appear in *Comptes Rendus Acad. Sciences*, Paris.

[16] Fermanian, J-D., Radulovic, D. and Wegkamp, M. (2002). Weak convergence of empirical copula processes. Working paper CREST 2002-06. To appear in *Bernoulli.*

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

[18] Fermanian, J-D. (2003) Goodness-of-fit tests for copulas. Working Paper CREST 2003-34.

[19] Fermanian, J-D. and Wegkamp, M. (2004) Time-dependent copulas. Mimeo.

[20] Gaenssler, P. and Stute, W. (1987) *Seminar on Empirical Processes.* DMV Seminar, Band 9, Birkhäuser.

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

[22] Genest, C., van der Goorbergh, R. and Werker, B. (2003) Multivariate option pricing using dynamic copulas. Discussion paper 122, Tilburg University.

[23] Gouriéroux C., J.P. Laurent and O. Scaillet (2000) Sensitivity Analysis of Values at Risk. *Journal of Empirical Finance* 7, 225-245.

[24] Joe, H. (1997) *Multivariate models and dependence concepts*, Chapman and Hall.

[25] Markowitz H. (1952). Portfolio Selection. *Journal of Finance* 7, 77-91.

[26] Nelsen, R. B. (1999) *An introduction to copulas*, Lecture Notes in Statistics, **139**, Springer.

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

[28] Rio, E. (2000) *Théorie asymptotique des processus aléatoires faiblement dépendants* Mathématiques et Applications, 31, Springer Verlag, Berlin.

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

[30] Scaillet O. (2000) Nonparametric Estimation and Sensitivity Analysis of Expected Shortfall. Forthcoming in *Mathematical Finance.*

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

[32] Vaart, A. W. van der and Wellner, J. (1996) *Weak convergence and empirical processes.* Springer, New-York.