



**HAL**  
open science

# Strong approximation of empirical copula processes by Gaussian processes

Salim Bouzebda, Tarek Zari

► **To cite this version:**

Salim Bouzebda, Tarek Zari. Strong approximation of empirical copula processes by Gaussian processes. 2008. hal-00340276v1

**HAL Id: hal-00340276**

**<https://hal.science/hal-00340276v1>**

Preprint submitted on 20 Nov 2008 (v1), last revised 27 Aug 2010 (v3)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Strong approximation of empirical copula process by Kiefer process

Salim BOUZEBDA and Tarek ZARI

*L.S.T.A., Université Pierre et Marie Curie-Paris 6. 175, rue du Chevaleret, 8<sup>ème</sup> étage, bâtiment A,  
75013 PARIS FRANCE.*

---

**Abstract.** This paper investigates the problem of strong approximation of the empirical copula process for arbitrary dimension, with continuous unknown margins. The idea of the proof is based on an original idea exposed by Deheuvels *et al.* (2006), and the theorem of strong approximation for an arbitrary distribution function proved in Csörgő and Horváth (1988). Using these results, we derive the normality for smoothed empirical copula process and L.I.L. for empirical process of copulas.

*AMS Subject Classifications:* 62G30 ; 60F17

*Keywords:* Dependence function ; strong approximation ; asymptotic theory.

---

## 1 Notations and Definitions

Let  $\mathbf{X}_i = (X_{1i}, \dots, X_{di})$ ,  $i = 1, 2, \dots$ , be an independent replica of a  $d$ -dimensional random vector  $\mathbf{X} \in \mathbb{R}^d$ , defined on a suitable space  $(\Omega, \mathcal{A}, \mathbb{P})$ . We denote by  $\mathbb{F}(\cdot)$  the distribution function [df] given by  $\mathbb{F}(x_1, \dots, x_d) := \mathbb{P}(X_1 \leq x_1, \dots, X_d \leq x_d)$ . We assume that the corresponding marginal df's  $F_i(x_i) := \mathbb{P}(X_i \leq x_i)$ , for  $i = 1, \dots, d$ , are continuous. According to the famous theorem of Sklar (see Sklar (1959)), we have

$$\mathbb{F}(x_1, \dots, x_d) := \mathbb{C}(F_1(x_1), \dots, F_d(x_d)), \quad \text{for every } \mathbf{x} \in \overline{\mathbb{R}}^d.$$

So,

$$\mathbb{C}(u_1, \dots, u_d) := \mathbb{F}(F_1^-(u_1), \dots, F_d^-(u_d)), \quad \text{for every } \mathbf{u} \in [0, 1]^d, \quad (1)$$

where,  $F_i^-(u_i) = \inf\{x_i : F_i(x_i) \geq u_i\}$ ,  $1 \leq i \leq d$ , denotes the quantile function of  $F_i(\cdot)$ . Setting  $\mathbb{1}_A(\cdot)$  for the indicator function of the set  $A$ , we define, for each  $n \geq 1$ , the empirical counterparts of  $\mathbb{F}(\cdot)$ ,  $F_1(\cdot)$ ,  $\dots$ ,  $F_d(\cdot)$  and  $F_1^-(\cdot)$ ,  $\dots$ ,  $F_d^-(\cdot)$ , respectively, by setting, for  $j = 1, \dots, d$ ,

$$\mathbb{F}_n(x_1, \dots, x_d) := \frac{1}{n} \sum_{i=1}^n \prod_{j=1}^d \mathbb{1}\{X_{ji} \leq x_j\}, \quad \text{for } \mathbf{x} \in \mathbb{R}^d,$$

$$F_{jn}(x_j) := \frac{1}{n} \sum_{i=1}^n \mathbb{1}\{X_{ij} \leq x_j\} = \mathbb{F}_n(1, \dots, x_j, \dots, 1), \quad \text{for } x_j \in \mathbb{R},$$

and

$$F_{in}^-(u_i) = \inf\{t_i \in \mathbb{R} : F_{in}(t_i) \geq u_i\}, \quad \text{for } t_i \in [0, 1].$$

---

*E-mail address:* salim.bouzebda@upmc.fr

In view of the characterization (1), we define the empirical copula function of  $\mathbb{F}_n(\cdot)$  (or, equivalently, associated with  $\mathbf{X}_1, \dots, \mathbf{X}_n$ ), as any copula  $\mathbb{C}_n$ , through the following identity

$$\mathbb{C}_n(u_1, \dots, u_d) := \mathbb{F}_n(F_{1n}^-(u_1), \dots, F_{dn}^-(u_d)), \quad \text{for } \mathbf{u} \in [0, 1]^d.$$

Empirical copulas were originally introduced by Deheuvels (1979a, 1980, 1981a,b) under the name of empirical dependence functions. Deheuvels has extensively studied the empirical process of copula defined by

$$\mathbb{G}_n(\mathbf{u}) = n^{1/2}(\mathbb{C}_n(\mathbf{u}) - \mathbb{C}(\mathbf{u})), \quad \text{for } \mathbf{u} \in [0, 1]^d, \quad (2)$$

under the assumption of independence between margins. Further, set  $U_{ji} := F_j(X_{ji})$  for  $i = 1, \dots, n$ ,  $j = 1, \dots, d$ , and  $\mathbf{U}_i := (U_{1i}, \dots, U_{di})$ . Thus, the random vectors  $\mathbf{U}_i$  are distributed according to the copula  $\mathbb{C}(\cdot)$ , and the sequence  $U_{ji}$  of random variables are independent and uniformly distributed on  $(0, 1)$ . For each  $n \geq 1$  and  $0 \leq u_j \leq 1$  for  $1 \leq j \leq d$ , set

$$\tilde{\mathbb{C}}_n(u_1, \dots, u_d) := \frac{1}{n} \sum_{i=1}^n \prod_{j=1}^d \mathbb{1}\{U_{ji} \leq u_j\}, \quad \text{and} \quad G_{jn}(u_j) := \frac{1}{n} \sum_{i=1}^n \mathbb{1}\{U_{ij} \leq u_j\}.$$

Considering the empirical processes defined, respectively, for  $n \geq 1$ ,  $\mathbf{u} \in [0, 1]^d$  and  $0 \leq u_j \leq 1$ , for  $j = 1, \dots, d$ , by

$$\alpha_n(\mathbf{u}) := n^{1/2}(\tilde{\mathbb{C}}_n(\mathbf{u}) - \mathbb{C}(\mathbf{u})), \quad (3)$$

$$\alpha_n^j(u_j) := n^{1/2}\{G_{jn}(u_j) - u_j\}, \quad (4)$$

$$\beta_{jn}(u_j) := n^{1/2}\{G_{jn}^-(u_j) - u_j\}. \quad (5)$$

Then, with  $X \stackrel{d}{=} Y$  which means that  $X$  and  $Y$  have the same distribution, we have

$$\mathbb{F}_n(x_1, \dots, x_d) \stackrel{d}{=} \tilde{\mathbb{C}}_n(F_1(x_1), \dots, F_d(x_d))$$

and

$$(F_{1n}(x_1), \dots, F_{dn}(x_d)) \stackrel{d}{=} (G_{1n}(F_1(x_1)), \dots, G_{dn}(F_d(x_d))).$$

From the definition of  $F_{in}^-(\cdot)$ , we notice that

$$G_{in}^-(u_i) \stackrel{d}{=} F_i\{F_{in}^-(u_i)\},$$

and we also have

$$\mathbb{C}_n(u_1, \dots, u_d) \stackrel{d}{=} \tilde{\mathbb{C}}_n(G_{1n}^-(u_1), \dots, G_{dn}^-(u_d)).$$

Since the law of  $\mathbb{C}_n(\cdot)$  is the same for all  $\mathbb{F}(\cdot)$  whose associated copula is  $\mathbb{C}(\cdot)$ . It is more convenient to investigate the following process

$$\mathbb{G}_n(\mathbf{u}) = n^{1/2}(\tilde{\mathbb{C}}_n(G_{1n}^-(u_1), \dots, G_{dn}^-(u_d)) - \mathbb{C}(u_1, \dots, u_d)), \quad \text{for } \mathbf{u} \in [0, 1]^d. \quad (6)$$

Our aim is to provide the strong approximation of the empirical process of copulas  $\{\mathbb{G}_n(\mathbf{u}) : \mathbf{u} \in [0, 1]^d\}$ , based upon  $\mathbf{X}_1, \dots, \mathbf{X}_n$ , by a single Gaussian process. Denote by  $\{\mathbb{K}_C(\mathbf{u}, n); \mathbf{u} \in [0, 1]^d, n \geq 0\}$  a Gaussian process, with continuous sample paths, mean zero, and covariance function given by

$$\mathbb{E}(\mathbb{K}_C(\mathbf{u}, n)\mathbb{K}_C(\mathbf{v}, m)) = (n \wedge m) \{\mathbb{C}(\mathbf{u} \wedge \mathbf{v}) - \mathbb{C}(\mathbf{u})\mathbb{C}(\mathbf{v})\}, \quad \text{for } \mathbf{u}, \mathbf{v} \in [0, 1]^d.$$

Note that the process  $\{\mathbb{K}_{\mathbb{C}}(\mathbf{u}, n) : \mathbf{u} \in [0, 1]^d, n \geq 0\}$  is called Kiefer process.

The continuation of our work requires some notations, that will be introduced as follows. For each  $n > 0$  and  $u_j \in [0, 1]$ , for  $j = 1, \dots, d$ , set

$$\begin{aligned} \mathbb{K}_{\mathbb{C}}^*(\mathbf{u}, n) &= \mathbb{K}_{\mathbb{C}}(\mathbf{u}, n) - \sum_{j=1}^d \mathbb{K}_{\mathbb{C}}(1, \dots, 1, u_j, 1, \dots, 1, n) \frac{\partial \mathbb{C}(\mathbf{u})}{\partial u_j} \\ &=: \mathbb{K}_{\mathbb{C}}(\mathbf{u}, n) - \sum_{j=1}^d \mathbb{K}_{\mathbb{C}}^{(j)}(\mathbf{1}, u_j, \mathbf{1}, n) \frac{\partial \mathbb{C}(\mathbf{u})}{\partial u_j}, \quad \mathbf{u} \in [0, 1]^d. \end{aligned} \quad (7)$$

The remainder of our paper is organized as follows. In the forthcoming section we will give our main result concerning the strong approximation of empirical copula process by a Gaussian process, which is stated in Theorem 2.1 below. In section 3 we will give some applications of Theorem 2.1, more precisely we will give the limit law of smoothed empirical copula and the law of iterated logarithm for the empirical copula process. The proofs are referred to section 5.

## 2 Results

Let's formulate the main theorem in this section.

**Theorem 2.1** *Let us assume that  $\mathbb{C}(\cdot)$ , associated with  $\mathbb{F}(\cdot)$ , be twice continuously differentiable on  $(0, 1)^d$ , and the second derivative be continuous on  $[0, 1]^d$ . Then, in the same space of initial sequence of random variables, we can define a Gaussian process  $\{K_{\mathbb{C}}^*(\mathbf{u}, n) : \mathbf{u} \in [0, 1]^d; n > 0\}$ , with continuous sample paths, such as*

$$\sup_{\mathbf{u} \in [0, 1]^d} |\sqrt{n} \mathbb{G}_n(\mathbf{u}) - K_{\mathbb{C}}^*(\mathbf{u}, n)| = \mathcal{O}\left(n^{1/2-1/(4d)} (\log n)^{3/2}\right), \quad \text{a.s.}, \quad (8)$$

where  $K_{\mathbb{C}}^*(\mathbf{u}, n)$  is defined in (7).

## 3 Applications

### 3.1 Smoothed empirical copulas process

We define smoothed estimators  $\widehat{\mathbb{C}}_n$  of  $\mathbb{C}$  by

$$\widehat{\mathbb{C}}_n(\mathbf{u}) = \frac{1}{h} \int_{[0, 1]^d} k\left(\frac{\mathbf{u} - \mathbf{v}}{h^{1/d}}\right) \mathbb{C}_n(\mathbf{v}) d\mathbf{v} \quad \text{for } \mathbf{u} \in [0, 1]^d, \quad (9)$$

where  $k$  is a kernel and  $h = h(n)$  is the smoothing parameter. Similarly to previous section, we define the smoothed empirical copulas process by

$$\widehat{\mathbb{G}}_n(\mathbf{u}) := \sqrt{n}(\widehat{\mathbb{C}}_n(\mathbf{u}) - \mathbb{C}(\mathbf{u})) \quad \text{for } \mathbf{u} \in [0, 1]^d. \quad (10)$$

Assume  $\mathbb{C}$  fulfills the following condition.

(F.1).  $\mathbb{C}$  has a bounded  $s$ th derivative.

For our result we suppose that  $h$  and  $k(\cdot)$  satisfy the following regularity conditions.

(C.1).  $h = h(n) \rightarrow 0$ ,  $nh \rightarrow \infty$  and  $\sqrt{nh}h^{s/d} \rightarrow 0$  as  $n \rightarrow \infty$ ;

(C.2).  $k(\cdot)$  is a continuous density function having compact support;

(C.3).  $k(\cdot)$  is of order  $s$ .

**Corollary 3.1** *Assume that (F.1) and (C.1)-(C.3) hold. Then, in the same space of initial sequence of random variables, we can define a Gaussian process  $\{K_{\mathbb{C}}^*(\mathbf{u}, n) \mid \mathbf{u} \in [0, 1]^d; n > 0\}$ , with continuous sample paths, such as*

$$\sup_{\mathbf{u} \in [0, 1]^d} |\widehat{\mathbb{G}}_n(\mathbf{u}) - \frac{1}{\sqrt{n}} K_{\mathbb{C}}^*(\mathbf{u}, n)| = o_P(1). \quad (11)$$

*Remark 1* 1. Corollary 3.1 remains valid replacing the condition that the kernel  $k$  having compact support in (C.2) by

(C.4). *There exists a sequences of real numbers  $a_n$  such that  $a_n h$  tends to zero when  $n$  tends to infinity, and*

$$\sqrt{n} \int_{\{\|\mathbf{v}\| > a_n\}} |k(\mathbf{v})| d\mathbf{v} \rightarrow 0.$$

2. *Note that conditions of Corollary 3.1 are classical to control the deviations between classical and smoothed empirical copulas processes.*

## 3.2 L.I.L for empirical copula process

By Theorem 2.1 we have almost surely

$$\limsup_{n \rightarrow \infty} \left\{ \left( \frac{n}{2 \log \log n} \right)^{1/2} \sup_{\mathbf{u} \in [0, 1]^d} |\mathbb{C}_n(\mathbf{u}) - \mathbb{C}(\mathbf{u})| \right\} = \limsup_{n \rightarrow \infty} \frac{\sup_{\mathbf{u} \in [0, 1]^d} |K_{\mathbb{C}}^*(\mathbf{u}, n)|}{(2n \log \log n)^{1/2}}. \quad (12)$$

According to Wichura (1973), we have almost surely, as  $n \rightarrow \infty$

$$\limsup_{n \rightarrow \infty} \frac{\sup_{\mathbf{u} \in [0, 1]^d} |K_{\mathbb{C}}^*(\mathbf{u}, n)|}{(2n \log \log n)^{1/2}} \leq d + 1. \quad (13)$$

Combining (12) and (13) to give the following Corollary.

**Corollary 3.2** *under the same conditions of the Theorem 2.1, we have*

$$\limsup_{n \rightarrow \infty} \left\{ \left( \frac{n}{2 \log \log n} \right)^{1/2} \sup_{\mathbf{u} \in [0, 1]^d} |\mathbb{C}_n(\mathbf{u}) - \mathbb{C}(\mathbf{u})| \right\} \leq d + 1. \quad (14)$$

## 4 Concluding remarks

1. The approximation of empirical copula process by a Gaussian process  $B_n$ , we know not how  $B_n$  behaves as a function of  $n$ , to resolve this problem, in Theorem 2.1 we have approximated the empirical copula process by a single Gaussian process.
2. Theorem 2.1 may be applied to derive some limit laws of some statistics like Kendall's sample rank correlation coefficient and Spearman's sample rank correlation coefficient.
3. Obviously, the empirical copulas can be used in many others problems like testing independence the challenging task is to apply our results to derive the asymptotic properties of those test statistics which will be considered in our future investigations.

## 5 Proofs

### 5.1 Proof of theorem 2.1

For easy reference, we recall some facts that we will need for our demonstration.

**Fact 1.** The processes  $\mathbb{K}_{\mathbb{C}}^*(\cdot, n)$ ,  $\mathbb{K}_{\mathbb{C}}^{(j)}(\cdot, n)$  and  $\mathbb{K}_{\mathbb{C}}^{(i)}(\cdot, n)$ , for all  $1 \leq i \neq j \leq d$ , are independent.

For convenience, we will denote the sup-norm of a bounded function  $f$ , defined on  $I = [0, 1]$  or  $I = [0, 1]^d$ , by  $\|f\| = \sup_{x \in I} |f(x)|$ . The next fact, due to Kiefer (1970) and Bahadur (1966) provides a L.I.L of Bahadur-Kiefer process (see also, Deheuvels et Mason (1990)).

**Fact 2.** We have, for  $j = 1, \dots, d$ , with probability 1,

$$\limsup_{n \rightarrow \infty} n^{1/4} (\log n)^{-1/2} (\log \log n)^{-1/4} \|\alpha_n^j + \beta_{jn}\| = 2^{1/4},$$

Recall the definition (6) of  $\mathbb{G}_n$ , the empirical process of copulas can write itself, for  $\mathbf{u} \in [0, 1]^d$ , as follows

$$\begin{aligned} \mathbb{G}_n(\mathbf{u}) &= n^{1/2} (\tilde{\mathbb{C}}_n(G_{1,n}^-(u_1), \dots, G_{d,n}^-(u_d)) - \mathbb{C}(u_1, \dots, u_d)) \\ &= \alpha_n \left( G_{1,n}^-(u_1), \dots, G_{d,n}^-(u_d) \right) + n^{1/2} \left\{ \mathbb{C} \left( G_{1,n}^-(u_1), \dots, G_{d,n}^-(u_d) \right) \right. \\ &\quad \left. - \mathbb{C}(u_1, \dots, u_d) \right\} \\ &= \alpha_n \left( u_1 + n^{-1/2} \beta_{1n}(u_1), \dots, u_d + n^{-1/2} \beta_{dn}(u_d) \right) \\ &\quad + n^{1/2} \left\{ \mathbb{C} \left( u_1 + n^{-1/2} \beta_{1n}(u_1), \dots, u_d + n^{-1/2} \beta_{dn}(u_d) \right) - \mathbb{C}(u_1, \dots, u_d) \right\} \\ &= \alpha_n(\mathbf{u}) + \left\{ \alpha_n(\mathbf{u} + n^{-1/2} \beta_n(\mathbf{u})) - \alpha_n(\mathbf{u}) \right\} \\ &\quad + n^{1/2} \left\{ \mathbb{C}(\mathbf{u} + n^{-1/2} \beta_n(\mathbf{u})) - \mathbb{C}(\mathbf{u}) \right\} \\ &= \alpha_n(\mathbf{u}) + \Delta_1(\mathbf{u}, n) + \Delta_2(\mathbf{u}, n). \end{aligned} \tag{15}$$

where  $(\mathbf{u} + n^{-1/2} \beta_n(\mathbf{u})) = (u_1 + n^{-1/2} \beta_{1n}(u_1), \dots, u_d + n^{-1/2} \beta_{dn}(u_d))$ . We start with  $\Delta_2(\mathbf{u}, n)$ , by differentiability assumption on  $\mathbb{C}(\cdot)$  we make use of Taylor expansion to obtain

$$\begin{aligned} \Delta_2(\mathbf{u}, n) &= \sum_{i=1}^d \frac{\partial \mathbb{C}(\mathbf{u})}{\partial u_i} \sqrt{n} (G_{ni}^-(u_i) - u_i) \\ &\quad + \frac{\sqrt{n}}{2} \sum_{i=1}^d \frac{\partial^2 \mathbb{C}(\mathbf{u}^*)}{\partial u_i \partial u_j} (G_{ni}^-(u_i) - u_i) (G_{nj}^-(u_j) - u_j), \end{aligned}$$

where  $\mathbf{u}^*$  lies in the interior of the line segment joining  $(G_{n1}^-(u_1), \dots, G_{nd}^-(u_d))$  and  $(u_1, \dots, u_d)$ . It follows from the definition of  $\alpha_n(\cdot)$  (cf. (3)), for  $u_j \in [0, 1]$ ,  $j = 1, \dots, d$ , that

$$\begin{aligned} \sqrt{n} (G_{ni}^-(u_i) - u_i) &= -\sqrt{n} (G_{ni}(G_{ni}^-(u_i)) - G_{ni}^-(u_i)) + \sqrt{n} (G_{ni}(G_{ni}^-(u_i)) - u_i), \\ &= -\alpha_n(\mathbf{1}, G_{ni}^-(u_i), \mathbf{1}) + \sqrt{n} (G_{ni} \circ G_{ni}^-(u_i) - u_i). \end{aligned}$$

Using the fact that  $|G_{ni} \circ G_{ni}^-(u_i) - u_i| \leq 1/n$  and the Chung (1949)'s law of the iterated logarithm, one finds, almost surely,

$$\Delta_2(\mathbf{u}, n) = - \sum_{i=1}^d \frac{\partial \mathbb{C}(\mathbf{u})}{\partial u_i} \alpha_n(\mathbf{1}, G_{ni}^-(u_i), \mathbf{1}) + O(n^{-1/2} (\log \log n)^{1/2}).$$

Further more, by Stute (1982), we have

$$\sup_{u_i \in [0,1]} |\alpha_n(\mathbf{1}, G_{ni}^-(u_i), \mathbf{1}) - \alpha_n(\mathbf{1}, u_i, \mathbf{1})| = O(n^{-1/4}(\log n)^{1/2}(\log \log n)^{1/4}).$$

Then, it follows that

$$\Delta_2(\mathbf{u}, n) = - \sum_{i=1}^d \frac{\partial \mathbb{C}(\mathbf{u})}{\partial u_i} \alpha_n(\mathbf{1}, u_i, \mathbf{1}) + O(n^{-1/4}(\log n)^{1/2}(\log \log n)^{1/4}).$$

as mentioned in Stute (1984). Next we consider  $\Delta_1(\mathbf{u}, n)$  which appears as the difference of the process  $\alpha_n$  between the two points  $(\mathbf{u} + n^{-1/2}\beta_n(\mathbf{u}))$  and  $\mathbf{u}$ , what allows us to consider the modulus of continuity  $w_n^*(\cdot)$  of  $\alpha_n$ , defined by

$$w_n^*(\mathbf{a}) := \sup \left\{ \alpha_n(A) : A = \prod_{i=1}^d [u_i, v_i] \in [0, 1]^d, \text{ with } |[u_i, v_i]| = v_i - u_i \leq a_i, \forall i = 1, \dots, d \right\},$$

According to Einmahl and Ruymgaart (1987), we have

**Fact 3.** Let  $\{a_n\}_{n \geq 1}$  be a sequence in  $(0, 1)$  such as  $a_n \downarrow 0$ , as  $n \rightarrow \infty$ , and

$$i) na_n^d \uparrow \infty, \quad ii) na_n^d / \log n \rightarrow \infty, \quad iii) \log(1/a_n) / \log \log n \rightarrow \infty$$

Then, with probability 1, we have

$$\lim_{n \rightarrow \infty} \left\{ 2a_n^d \log(1/a_n^d) \right\}^{-1/2} w_n^*(a_n, \dots, a_n) = 1.$$

An application of the Chung (1949) law of the iterated logarithm shows that, for each  $j = 1, \dots, d$ , almost surely,

$$\limsup_{n \rightarrow \infty} \left\{ (\log \log n)^{-1/2} \sup_{0 \leq u \leq 1} |\beta_{jn}(u_j)| \right\} = 2^{-1/2}, \quad (16)$$

so

$$\sup_{0 \leq u_j \leq 1} |\beta_{jn}(u_j)| \leq \frac{(\log \log n)^{1/2}}{n^{1/2}} \leq \frac{(\log n)^{2/d}}{n^{1/d}} := a_n,$$

an application of fact 3 shows that, almost surely as  $n \rightarrow \infty$ , we have

$$\sup_{\mathbf{u} \in [0,1]^d} |\Delta_1(\mathbf{u}, n)| \leq w_n^*(a_n) = \mathcal{O}\left(n^{-1/2}(\log n)^{3/2}\right).$$

The next fact, due to Csörgő and Horváth (1988) provides a strong approximation result appropriate for our needs.

**Fact 4.** On a suitable probability space  $(\Omega, \mathcal{A}, \mathbb{P})$ , it is possible to define  $\{\alpha_n(\mathbf{u}) : \mathbf{u} \in [0, 1]^d\}$ , jointly with a sequence of Gaussian process  $\{\mathbb{K}_{\mathbb{C}}(\mathbf{u}, n) : \mathbf{u} \in [0, 1]^d, n \geq 0\}$ , in such a way that, almost surely, as  $n \rightarrow \infty$ ,

$$\sup_{\mathbf{u} \in [0,1]^d} |\sqrt{n}\alpha_n(\mathbf{u}) - \mathbb{K}_{\mathbb{C}}(\mathbf{u}, n)| = \mathcal{O}\left(n^{1/2-1/(4d)}(\log n)^{3/2}\right), \text{ a.s.} \quad (17)$$

By applying the triangle inequality, we have

$$\begin{aligned}
\sup_{\mathbf{u} \in [0,1]^d} |\sqrt{n}\mathbb{G}_n(\mathbf{u}) - K_{\mathbb{C}}^*(\mathbf{u}, n)| &\leq \sup_{\mathbf{u} \in [0,1]^d} |\sqrt{n}\alpha_n(\mathbf{u}) - \mathbb{K}_{\mathbb{C}}(\mathbf{u}, n)| + w_n^*(a_n) \\
&+ \sum_{j=1}^d \left| \frac{\partial \mathbb{C}(\mathbf{u})}{\partial u_j} \right| \sup_{0 \leq u_j \leq 1} |n^{1/2}\alpha_n(\mathbf{1}, u_j, \mathbf{1}) - \mathbb{K}_{\mathbb{C}}^{(j)}(\mathbf{1}, u_j, \mathbf{1}, n)| \\
&+ O(n^{-1/4}(\log n)^{1/2}(\log \log n)^{1/4}), \\
&\leq \mathcal{O}\left(n^{1/2-1/(4d)}(\log n)^{3/2}\right) + \mathcal{O}\left(n^{-1/2}(\log n)^{3/2}\right) \\
&+ \mathcal{O}\left(n^{1/2-1/(4d)}(\log n)^{3/2}\right) + \mathcal{O}\left(n^{-1/4}(\log n)^{1/2}(\log \log n)^{1/4}\right),
\end{aligned}$$

then, we have almost surely as  $n \rightarrow \infty$ ,

$$\sup_{\mathbf{u} \in [0,1]^d} |\sqrt{n}\mathbb{G}_n(\mathbf{u}) - K_{\mathbb{C}}^*(\mathbf{u}, n)| = \mathcal{O}\left(n^{1/2-1/(4d)}(\log n)^{3/2}\right).$$

The proof of Theorem 2.1 is achieved.  $\square$

## 5.2 Proof of corollary 3.1

First observe that the difference between the classical and the smoothed empirical copula processes ( $\mathbb{G}_n(\cdot)$  and  $\widehat{\mathbb{G}}_n(\cdot)$ ) can be controlled as follows. We see that

$$\widehat{\mathbb{G}}_n(\mathbf{u}) = \sqrt{n} \left( \frac{1}{h} \int_{[0,1]^d} k\left(\frac{\mathbf{u}-\mathbf{v}}{h^{1/d}}\right) \mathbb{C}_n(\mathbf{v}) d\mathbf{v} - \mathbb{C}(\mathbf{u}) \right) \quad (18)$$

$$\begin{aligned}
&= \left( \frac{1}{h} \int_{[0,1]^d} k\left(\frac{\mathbf{u}-\mathbf{v}}{h^{1/d}}\right) \sqrt{n}(\mathbb{C}_n(\mathbf{v}) - \mathbb{C}(\mathbf{v})) \right) d\mathbf{v} \\
&+ \sqrt{n} \left( \frac{1}{h} \int_{[0,1]^d} k\left(\frac{\mathbf{u}-\mathbf{v}}{h^{1/d}}\right) \mathbb{C}(\mathbf{v}) d\mathbf{v} - \mathbb{C}(\mathbf{u}) \right) \quad (19)
\end{aligned}$$

$$\begin{aligned}
&= \left( \frac{1}{h} \int_{[0,1]^d} k\left(\frac{\mathbf{u}-\mathbf{v}}{h^{1/d}}\right) \mathbb{G}_n(\mathbf{u}) \right) d\mathbf{v} \\
&+ \sqrt{n} \left( \frac{1}{h} \int_{[0,1]^d} k\left(\frac{\mathbf{u}-\mathbf{v}}{h^{1/d}}\right) \mathbb{C}(\mathbf{v}) d\mathbf{v} - \mathbb{C}(\mathbf{u}) \right). \quad (20)
\end{aligned}$$

Thus

$$\begin{aligned}
|\widehat{\mathbb{G}}_n(\mathbf{u}) - \mathbb{G}_n(\mathbf{u})| &\leq \left| \int_{\prod_{i=1}^d \left[\frac{u_i-1}{h^{1/d}}, \frac{u_i}{h^{1/d}}\right]} (\mathbb{G}_n(\mathbf{u} - h^{1/d}\mathbf{v}) - \mathbb{G}_n(\mathbf{u})) k(\mathbf{v}) d\mathbf{v} \right| \\
&+ \sup_{\mathbf{u} \in [0,1]^d} |\mathbb{G}_n(\mathbf{u})| \left| \int_{\prod_{i=1}^d \left[\frac{u_i-1}{h^{1/d}}, \frac{u_i}{h^{1/d}}\right]} k(\mathbf{v}) d\mathbf{v} - 1 \right| \\
&+ \sqrt{n} \left| \int_{\prod_{i=1}^d \left[\frac{u_i-1}{h^{1/d}}, \frac{u_i}{h^{1/d}}\right]} (\mathbb{C}(\mathbf{u} - h^{1/d}\mathbf{v}) - \mathbb{C}(\mathbf{u})) k(\mathbf{v}) d\mathbf{v} \right| \\
&+ \sqrt{n} \sup_{\mathbf{u} \in [0,1]^d} |\mathbb{C}(\mathbf{u})| \left| \int_{\prod_{i=1}^d \left[\frac{u_i-1}{h^{1/d}}, \frac{u_i}{h^{1/d}}\right]} k(\mathbf{v}) d\mathbf{v} - 1 \right| \\
&:= \xi_{1;n}(\mathbf{u}) + \xi_{2;n}(\mathbf{u}) + \xi_{3;n}(\mathbf{u}) + \xi_{4;n}(\mathbf{u}). \quad (21)
\end{aligned}$$



Under conditions (F.1), (C.1)-(C.3) and using Taylor expansion of order  $s$  we get

$$|\xi_{3;n}(\mathbf{u})| = \frac{h_n^{s/d}}{s!} \sqrt{n} \left| \int \sum_{k_1+\dots+k_d=s} u_1^{k_1} \dots u_d^{k_d} \frac{\partial^s \mathbb{C}(\mathbf{u} - h_n \theta \mathbf{v})}{\partial u_1^{k_1} \dots \partial u_d^{k_d}} K(\mathbf{v}) d\mathbf{v} \right|,$$

where  $0 < \theta < 1$ . Thus a simple application of Lebesgue dominated convergence theorem gives

$$n^{-1/2} h_n^{-(s/d)} |\xi_{3;n}(\mathbf{u})| = \frac{1}{k!} \left| \sum_{k_1+\dots+k_d=s} \frac{\partial^s \mathbb{C}(\mathbf{u})}{\partial u_1^{k_1} \dots \partial u_d^{k_d}} \int u_1^{k_1} \dots u_d^{k_d} K(\mathbf{v}) d\mathbf{v} \right|. \quad (22)$$

Then we have by condition (C.1)

$$|\xi_{3;n}(\mathbf{u})| = O(n^{1/2} h_n^{s/d}) = o(1). \quad (23)$$

By Theorem 2.1 and almost sure continuity of  $K_{\mathbb{C}}^*(\mathbf{u}, n)$  we have

$$\begin{aligned} \xi_{1;n}(\mathbf{u}) &\leq \sup_{\mathbf{u}, \mathbf{v} \in [0,1]^d} \sup_{|\mathbf{u}-\mathbf{v}| \leq h_n} |\mathbb{G}_n(\mathbf{v}) - \mathbb{G}_n(\mathbf{u})| \left| \int K(\mathbf{v}) d\mathbf{v} \right| \\ &= o_P(1) O(1) = o_P(1). \end{aligned} \quad (24)$$

We treat now  $\xi_{2;n}(\mathbf{u})$ . We have  $\sup_{\mathbf{u} \in [0,1]^d} |\mathbb{G}_n(\mathbf{u})| = O_P(1)$  and as  $n$  tends to infinity (by condition (C.2))

$$\sqrt{n} \left| \int_{\prod_{i=1}^d \left[ \frac{u_i-1}{h^{1/d}}, \frac{u_i}{h^{1/d}} \right]} k(\mathbf{v}) d\mathbf{v} - 1 \right| = o(1),$$

then  $\xi_{2;n}(\mathbf{u}) = o_P(1)$ . Using the same arguments we obtain  $\xi_{4;n}(\mathbf{u}) = o_P(1)$ . Recollecting the above results, one obtains

$$\sup_{\mathbf{u} \in [0,1]^d} |\widehat{\mathbb{G}}_n(\mathbf{u}) - \mathbb{G}_n(\mathbf{u})| = o_P(1). \quad (25)$$

Now the Corollary follows from (25) in conjunction with Theorem 2.1, and the fact that

$$\begin{aligned} \sup_{\mathbf{u} \in [0,1]^d} \left| \widehat{\mathbb{G}}_n(\mathbf{u}) - \frac{1}{\sqrt{n}} K_{\mathbb{C}}^*(\mathbf{u}, n) \right| &\leq \sup_{\mathbf{u} \in [0,1]^d} |\widehat{\mathbb{G}}_n(\mathbf{u}) - \mathbb{G}_n(\mathbf{u})| \\ &\quad + \sup_{\mathbf{u} \in [0,1]^d} \left| \mathbb{G}_n(\mathbf{u}) - \frac{1}{\sqrt{n}} K_{\mathbb{C}}^*(\mathbf{u}, n) \right|. \end{aligned}$$

and thus the proof of Corollary is complete.  $\square$

## References

- Bahadur, R. (1966). A note on quantiles in large samples. *ann. Math.Statist*, **37**, 577–580.
- Chung, K.-L. (1949). An estimate concerning the Kolmogoroff limit distribution. *Trans. Amer. Math. Soc.*, **67**, 36–50.
- Csörgő, M. and Horváth, L. (1988). A note on strong approximations of multivariate empirical processes. *Stochastic Process. Appl.*, **28**(1), 101–109.

- Csörgő, M. and Révész, P. (1975). A strong approximations of the multivariate empirical process. *Stud.Scient.Math.Hung*, **10**(10), 427–434.
- Deheuvels, P. (1979a). La fonction de dépendance empirique et ses propriétés. Un test non paramétrique d'indépendance. *Acad. Roy. Belg. Bull. Cl. Sci. (5)*, **65**(6), 274–292.
- Deheuvels, P. (1979b). Propriétés d'existence et propriétés topologiques des fonctions de dépendance avec applications à la convergence des types pour des lois multivariées. *C. R. Acad. Sci. Paris Sér. A-B*, **288**(2), A145–A148.
- Deheuvels, P. (1980). Nonparametric test of independence. In *Nonparametric asymptotic statistics (Proc. Conf., Rouen, 1979) (French)*, volume 821 of *Lecture Notes in Math.*, pages 95–107. Springer, Berlin.
- Deheuvels, P. (1981a). A Kolmogorov-Smirnov type test for independence and multivariate samples. *Rev. Roumaine Math. Pures Appl.*, **26**(2), 213–226.
- Deheuvels, P. (1981b). Multivariate tests of independence. In *Analytical methods in probability theory (Oberwolfach, 1980)*, volume 861 of *Lecture Notes in Math.*, pages 42–50. Springer, Berlin.
- Deheuvels, P. et Mason, D. M. (1990). Bahadur-Kiefer-type processes. *Ann. Probab.*, **18**, 669–697.
- Deheuvels, P. (1997). Strong laws for local quantile processes. *Ann. Probab.*, **25**(4), 2007–2054.
- Deheuvels, P., Peccati, G., and Yor, M. (2006). On quadratic functionals of the Brownian sheet and related processes. *Stochastic Process. Appl.*, **116**(3), 493–538.
- Einmahl, J. H. J. and Ruymgaart, F. H. (1987). The almost sure behavior of the oscillation modulus of the multivariate empirical process. *Statist. Probab. Lett.*, **6**(2), 87–96.
- Fermanian, J.-D., Radulović, D., and Wegkamp, M. (2004). Weak convergence of empirical copula processes. *Bernoulli*, **10**(5), 847–860.
- Kiefer, J. (1970). Deviations between the sample quantile process and the sample df. In *Nonparametric Techniques in Statistical Inference (Proc. Sympos., Indiana Univ., Bloomington, Ind., 1969)*, pages 299–319. Cambridge Univ. Press, London.
- Philipp, W. and Pinzur, L. (1980). Almost sure approximation theorems for the multivariate empirical process. *Z. Wahrsch. Verw. Gebiete*, **54**(1), 1–13.
- Révész, P. (1976). On strong approximation of the multidimensional empirical process. *Ann. Probability*, **4**(5), 729–743.
- Ruymgaart, F. H., Shorack, G. R., and van Zwet, W. R. (1972). Asymptotic normality of nonparametric tests for independence. *Ann. Math. Statist.*, **43**, 1122–1135.
- Sklar, A. (1959). Fonctions de répartition à  $n$  dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, **8**, 229–231.
- Stute, W. (1982). The oscillation behavior of empirical processes. *Ann. Probab.*, **10**(1), 86–107.

Stute, W. (1984). The oscillation behavior of empirical processes: the multivariate case. *Ann. Probab.*, **12**(2), 361–379.

Wichura, M. J. (1973). Some Strassen-type laws of the iterated logarithm for multiparameter stochastic processes with independent increments. *Ann. Probability*, **1**, 272–296.